INTERNATIONAL CONFERENCE ON TRAFFIC AND TRANSPORT ENGINEERING

# ictte

ISBN 978-86-916153-2-1

Belgrade 2014

# INTERNATIONAL CONFERENCE ON TRAFFIC AND TRANSPORT ENGINEERING

ICTTE

**November 27-28, 2014**
**Belgrade, Serbia**

**- ICTTE Belgrade 2014 -**

**Proceedings of the Second International Conference on Traffic and Transport Engineering**

ICTTE Belgrade 2014 has been jointly organized by the City Net Scientific Research Center Ltd. Belgrade, University of Belgrade, Faculty of Transport and Traffic Engineering and "Kirilo Savić" Institute. ICTTE Belgrade 2014 is co-hosted by the AIIT (Associazione Italiana per l'Ingegneria del Traffico e dei Trasporti) Research Center, Rome, Italy. The conference is supported by the EA SEA WAY project (Adriatic IPA, Cross Border Cooperation 2007-2013), and is held in Belgrade, Serbia, from 27th to 28th November 2014.

The conference covers a wide range of topics related to traffic and transport engineering, with the aim of representing the importance of all modes of traffic and transport, especially the importance of improving these industries, and their compliance to one of the most significant principles nowadays, sustainable development. ICTTE Belgrade 2014 gathers researchers, scientists and engineers whose fields of interest are traffic and transport engineering, and should provide them a good platform for discussion, interactions and exchange of information and ideas. ICTTE Proceedings have been indexed within Thomson Reuters's CPCI – Conference Proceedings Citation Index accessed via Web of Science.

For publisher: Prof. Dr Srećko Žeželj
Editor in Chief: Assist. Prof. Dr Olja Čokorilo
Publisher: City Net Scientific Research Center Ltd. Belgrade
Uzun Mirkova 10/I
11000 Belgrade, Serbia
Phone: (+381 11) 26 23 895
Cell: (+381 63) 28 12 55
Fax: (+381 11) 32 82 076
e-mail: office@city-net.rs / office@ijtte.com
www.citynetsrc.com / www.ijtte.com

**ISBN 978-86-916153-2-1**

**CIP - Каталогизација у публикацији - Народна библиотека Србије, Београд**

656.07(082)(0.034.2)
711.73(082)(0.034.2)

**INTERNATIONAL Conference on Traffic and Transport Engineering (2nd; 2014; Beograd)**
[Proceedings of the Second] International Conference on Traffic and Transport Engineering ICTTE, [November 27-28], 2014, Belgrade [Elektronski izvor] / [organized by the City Net Scientific Research Center, Belgrade ... [et al.] ; editor in chief Olja Čokorilo]. - Belgrade : City Net Scientific Research Center, 2014 (Belgrade : City Net Scientific Research Center). - 1 elektronski optički disk (CD-ROM) ; 12 cm

Sistemski zahtevi: Nisu navedeni. - Nasl. sa naslovnog ekrana. - Tiraž 200. - Napomene uz tekst. - Bibliografija uz svaki rad.

ISBN 978-86-916153-1-4
1. Čokorilo, Olja [главни уредник], 1977- 2. City Net Scientific Research Center (Beograd)
a) Саобраћај - Зборници b) Транспортни системи - Зборници c) Роба - Превоз - Зборници
COBISS.SR-ID 211435020

# PREFACE - ICTTE Belgrade 2014
# Why do we need a new transportation philosophy of history?


Ladies and gentlemen, distinguished guests and speakers, dear colleagues and readers,

I am delighted to welcome you to Belgrade and to the International Conference on Traffic and Transport Engineering (ICTTE Belgrade 2014).

ICTTE Belgrade 2014 presents state of art in the field of traffic and transport engineering. The conference is major conference in the region with the participation of researchers from more than 50 countries worldwide. Our research comprehensive network of people, research institutions and industry rapidly enlarge within ICTTE community.

The contributions to ICTTE 2014 have been high, with more than 120 papers divided into 18 sessions. Proceedings will be indexed within Thomson Reuters's CPCI – Conference Proceedings Citation Index accessed via Web of Science. After the conference, I have truly hope, that new research groups will find opportunities in some of Horizon 2020 perspectives, Danube Transnational Programme 2014-2020 strategic partnerships, Adriatic and Ionian Initiative, etc.

After the conference, a selection of papers will be edited to make a series of thematic volumes, covering broad topics of interest for the scientific community and end users. These volumes will be published by International Journal for Traffic and Transport Engineering (IJTTE) special edition by the end of 2015.

I am delighted to welcome you to Belgrade, the hart of Serbia, and I hope you will enjoy your work as much as social networking activities organized by our team. I hope old participants and new comers will join us in 2016 and fill history with their cooperation.

ICTTE 2014 Director

Olja Cokorilo

## Organizing Secretariat

**Olja Čokorilo** – *Conference Director*
Faculty of Transport and Traffic Engineering, University of Belgrade, SRB

**Ivana Čavka** – *Conference Executive Coordinator*
Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
**Srećko Žeželj**, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
**Ivan Ivković**, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB

## International Scientific Committee

**Gianluca Dell'Acqua** – *Committee President*
Department of Transportation Engineering, University of Napoli Federico II, **ITA**

Ali Payidar Akgüngör, Transportation Division, Kirikkale University, TUR
Ahmed F. Al-Kaisy, Department of Civil Engineering, Montana State University, USA
Shriniwas Shrikant Arkatkar, Birla Institute of Technology and Science, Vidya Vihar Campus Pilani, IND
Ivana D. Atanasovska, Kirilo Savic Institute, Belgrade, SRB
Nebojša Bojović, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
Chen Chao, Transportation and Management College, Dalian Maritime University, CHN
Dragan Čišić, Faculty of Maritime Studies, University of Rijeka, CRO
Olja Čokorilo, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
Tatjana Davidović, Mathematical Institute, Serbian Academy of Sciences and Arts, SRB
Vlastimir Dedović, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
Mario De Luca, Department of Transportation Engineering, University of Naples "Federico II", ITA
Momčilo Dobrodolac, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
Kasthurirangan Gopalakrishnan, College of Engineering, Iowa State University Ames, USA
Qi-Zhou Hu, School of Automation, Nanjing University of Science and Technology, CHN
Muhannad Ismeik, Department of Civil Engineering, University of Jordan, JOR
Ivan Ivković, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
Milan Janić, Delft University of Technology, NED
Rob Konings, Delft University of Technology, NED
Željko Kurtela, University of Dubrovnik, CRO
María Eugenia López Lambas, Department of Civil Engineering, Polytechnic University of Madrid, ESP
Dragana Macura, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
Vladislav Maraš, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
Goran Marković, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
Francesc Xavier Martínez de Osés, Nautical Science and Engineering Department, UPC Barcelona, ESP
Jaroslav Matuška, Jan Perner Transport Faculty, University of Pardubice, CZE
Bhimaraya Metri, Management Development Institute Gurgaon, IND
Dušan Mladenović, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
Goran Mladenović, Faculty of Civil Engineering, University of Belgrade, SRB
Snežana Mladenović, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
Nilanchal Patel, Birla Institute of Technology Mesra (A Deemed University of UGC), Ranchi, IND
Vesna M. Pavelkić, Kirilo Savic Institute, Belgrade, SRB
Laurie Pickup, Vectos Ltd. and Centre for Transport Research, University of Aberdeen, GBR
Zeng Qingcheng, School of Transportation Management, Dalian Maritime University, CHN
Valentina Radojičić, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
Kalaga Ramachandra Rao, Indian Institute of Technology Delhi, IND
Stefano Ricci, Department of Civil, Constructional & Environmental Engineeering, La Sapienza, ITA
Li Shuguang, School of Electronic and Control Engineering, Chang'an University, CHN
Arvind Kumar Shukla, Civil Engineering Department, Institute of Engineering & Technology U.P., IND
Branimir Stanić, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
Sanja Steiner, Faculty of Transport and Traffic Sciences, University of Zagreb, CRO
Milica Šelmić, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
Milan Šimko, University of Žilina, SVK
Serhan Tanyel, Department of Civil Engineering, Dokuz Eylul University, TUR
Slaven Tica, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
Ádám Török, KTI Institute for Transport Sciences Non Profit Ltd., HUN
Onder Turan, Anadolu University, Faculty of Aeronautics and Astronautics, TUR
Elen Twrdy, Dean of Faculty of Maritime Studies and Transport, University of Ljubljana, SLO
Ernestos Tzannatos, Department of Maritime Studies, University of Piraeus, GRE
Ljubiša Vasov, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB
Lidia Żakowska, Department for Civil Engineering, Cracow University of Technology, POL
Srećko Žeželj, Faculty of Transport and Traffic Engineering, University of Belgrade, SRB

# CONTENTS

## *Transport Optimization Problems*

## *Transport Modeling and Decision Making*



## *Transport Safety and Security*

## *Air Traffic Performance and Infrastructure Development*



## *Transport and Environment*

## *Transport and Economic Growth*

## Transport Infrastructure



## Land Use Development, Spatial and Transport Planning



## Transport Policy

## *Human Factors*



## *Urban Transportation Planning*

## *Transport Logistics*



## *Sustainable Transport*

*Transport Optimization Problems*

# A MICROSIMULATION CALIBRATION METHODOLOGY BASED ON SPEED-DENSITY RELATIONSHIPS FOR FREEWAYS

**Raffaele Mauro[1], Orazio Giuffrè[2], Anna Granà[3] , Sandro Chiappone[4]**
[1] *Department of Mechanical and Structural Engineering – Università degli Studi di Trento, Italy*
[2,3,4] *Department of Civil, Environmental, Aerospace, Materials Engineering – Università degli Studi di Palermo, Italy*

**Abstract:** This paper describes the calibration methodology based on speed-density relationships in the microsimulation calibration process, stated that they represent the traffic flow phenomenon in a wide range of operational conditions and well summarize all the information that may be collected on field (or following a run of the microsimulation model) on two of the three key variables of traffic flow. The matching of speed-density relationships from field and simulation was evaluated using statistical analysis as technique of pattern recognition. A test freeway segment under uncongested traffic conditions was selected as case study; based on traffic data observed at A22 Brenner Freeway, Italy, statistical regressions between the variables of traffic flow were developed. Analogous relationships were obtained using the Aimsun microscopic traffic simulator software, reproducing the on field conditions and varying some selected parameters until a good matching between field and simulation was achieved.

**Keywords**: microsimulation, freeways, speed-density relationship.

## 1. Introduction

Simulation is a sampling experiment on the real system through its model (Pidd, 1992). This means that the evolution over time of the system model should be able to imitate properly the evolution over time of the modeled system, and conclusions on the system behavior can be drawn by examining samples of the observational variables of interest through statistical analysis techniques. Thus a traffic simulation model has to represent the system behavior with sufficient accuracy so that the model can be used as a substitute for the actual system for experimental purposes. Road traffic microsimulation models, first commercially introduced in the 1990s, are one of the latest generation of available traffic models and became very popular for the development and evaluation of a broad range of road traffic management and control systems. They model the movements of individual vehicles, traveling around road networks by using car following, lane changing and gap acceptance rules; hence, traffic microsimulation models try to replicate the behavior of individual drivers. However, the "realism" sought by the representation of individual drivers introduces a level of complexity into the modeling process which must be taken into account from the stage of model calibration. Traffic microsimulation models typically include a very large number of parameters, representing various characteristics of travelers, vehicles and road system, that must be calibrated before the model is applied as a prediction tool of traffic performances (Hollander and Liu, 2008). Calibration of a traffic microsimulation model is an iterative process that consists of changing and adjusting numerous model parameters and comparing model outputs with a set of empirical data until a predefined level of agreement between the two data sets is achieved (Barcelo, 2011). Since no single model can be equally accurate for all possible traffic conditions or can include the whole universe of variables affecting real-world traffic conditions, every microsimulation software has a set of user-adjustable parameters which enable the analyst to calibrate the model to match locally observed conditions.
In order to reproduce the mechanism of a single decision made by an individual driver (e.g. the decision to change lane or to use a gap in the opposing stream to enter an intersection), every traffic microsimulation model consists of several sub-models each of which includes several parameters. Direct measurement of these parameters is complex, because many of them represent features hard to isolate, or extensive data collection is required. Thus, in the calibration process, aggregate data, which do not describe the behavior of individual drivers, are often used; however, when a model is calibrated using aggregate data, the result can limit behavioral power (see Hollander and Liu, 2008). Another question concerns which parameters have to be considered in the model calibration process. Some studies focus on the calibration of driver behavior parameters only, while assuming the others are given (see e.g. Hourdakis et al., 2003; Kim and Rilett, 2003); other studies introduce driving behavior in a broader problem, including the calibration of a route choice model and/or an o-d matrix (see e.g. Dowling et al. 2004a). There are also differences among calibration studies in terms of variation in the number of parameters that must be calibrated before the model can be used as a tool for prediction. In the case of a small number of parameters, the calibration process can be developed through a manual procedure; thus some parameters are calibrated often through multiple retries (Toledo et al., 2003). In the case of big parameter subsets, calibration process normally uses automated algorithms, which should allow a closer approach to the optimal solution; anyway automated procedure make harder to follow changes in the value of each parameter (Menneni et al., 2009). However, the selection of calibration parameters is often put in relation to the purpose of the calibration problem. The achievement of calibration targets, i.e. when the model outputs are similar to empirical data, can be influenced by the simplification of which microsimulation models are not free. This concerns some technical characteristics of micro-simulation models such us the transport system update mechanism, the representation of randomness, traffic generation, allocating driver/vehicle characteristics, vehicle interactions, etc. A further question concerning microsimulation is whether this process produces a valid model for the system in general, or the model gives

[13] Corresponding author: anna.grana@unipa.it

only a representation of the particular set of input data. In this regard it should be noted that to gain a valid model, two independent data sets are necessary: the first set of data should be used for the calibration of the model parameters; the second set should be used for the running of the calibrated model so that the resulting model output data can be compared to the second set of system output data. The comparison part is referred to as the validation of the calibrated model; it represents the process of checking to what extent the model built replicates reality (see e.g. Toledo and Koutsopoulos, 2004). The objective of this paper is to present a methodology which uses speed-density relationships in the microsimulation calibration process, stated that they represent the traffic flow phenomenon in a wide range of operational conditions and they well summarize all the information that may be collected on field (or following a run of the microsimulation model) on two of the three key variables of traffic flow. The matching of speed-density relationships from field and simulation was evaluated using statistical analysis as technique of pattern recognition. A test freeway segment under uncongested traffic conditions was selected as case study and reference will be made to it in the following. Based on traffic data observed at A22 Brenner Freeway, Italy, statistical regressions between the variables of traffic flow were developed. Analogous relationships were obtained using the Aimsun microscopic traffic simulator software, reproducing the on field conditions and varying some selected parameters until a good matching between field and simulation was achieved. The paper is organized as follows: section 2 briefly reviews some methodologies on calibration of traffic microsimulation models and issues on calibration approaches. In section 3 the path followed to construct the speed-flow-density relationships based on field observations at the A22 Brenner Freeway, Italy, is presented. Focusing on the test freeway segment selected as case study, the parameters to be calibrated within the model will be presented in section 4, whereas the calibration methodology and its implementation will be described in section 5. Discussion and conclusions are presented in section 6.

## 2. Calibration methodologies for traffic microsimulation models

Technical literature proposes various methodologies for calibration of microsimulation models. Some of the proposed methodologies use a single measure (see e.g. Kim and Rilett, 2003), whereas some others use more than one measure by performing a sequence of calibration sub-processes, each one of which uses a different traffic measure for calibrating a separate group of parameters. Among the procedures using more than one measure, we refer, for instance, to Dowling et al. (2004a) which proposed a three step methodology: first they calibrated driving behavior parameters by comparing capacities, then calibrated route choice parameters by comparing flows; in the last stage, calibration was completed by comparing travel times and queue lengths. Furthermore, in the procedure proposed by Hourdakis et al. (2003) simulated and observed flows were compared in the first stage to calibrate global parameters (e.g. maximum acceleration and other vehicle characteristics), then (simulated and observed) speeds were compared to calibrate local parameters; finally an optional calibration stage was suggested, comparing any measure chosen by the user. FHWA-HRT-04-040 Report (Dowling *et al.*, 2004b) proposed that calibration of the model to capacity be one of the steps in microsimulation calibration, having the objective to find a set of model parameters that make the model outputs as close as possible to the field-measured capacities. The capacity calibration step consists of two phases: a global calibration phase and a fine-tuning phase. Global calibration is first performed to identify the appropriate network-wide value of the capacity parameters that best reproduces conditions in the field. Link-specific capacity parameters are then adjusted to fine-tune the model so that it more precisely matches the field measurements of capacity at each bottleneck. It is also suggested that queue discharge flow rate be used to estimate a numerical value for capacity. However, loss of information can derive by defining capacity as a single numerical value, despite it clearly has to be represented by a distribution of capacity values; see Brilon *et al.* (2007) for an introduction to the stochastic nature of capacity. Menneni et al. (2009) noted that if the capacity calibration process is based on a single numerical value, matching the means of capacity distribution does not necessarily match the other important properties of a distribution, nor even other traffic parameters characterizing capacity as speed or density. In any case, it should be noted that in the calibration process, the main target should be to maximize the information suitable for replicating real system performances.

Capacity information can be derived from generalized relationships between speed and density, speed and flow, flow and density; these relationships also provide information regarding free-flow and congested regions which cannot be gained from a single numerical value or a distribution of capacities. Basing on speed-flow, speed-density, or flow-density relationships which provide information about the free-flow, congested, and queue discharge regions, a calibration procedure could replicate the whole range of traffic behavior and not just peak period. According to Menneni et al. (2009), for model calibration purposes, only a portion of one of the three graphs mentioned above instead of the entire graph could be used. Since the amount of information available in fitting empirical/simulated data is very important, more information can be obtained by using speed-flow, speed-density, or flow-density graphs; thus, a higher number of parameters can be submitted to the calibration process, resulting in a better fine-tuned simulation model. The calibration of speed-flow, speed-density, or flow-density graphs should be one step in microsimulation calibration and should be followed by route-choice calibration and system performance calibration. Despite the potential benefits in calibration process, the studies and researches which have used the fundamental relationships of traffic flow in the microsimulation calibration process are still few in number. For instance, we refer to Wiedemann (1991) which applied the concept of replicating field speed-flow relationships and used them to demonstrate closeness of field and simulated data, whereas Fellendorf and Vortisch (2001) demonstrated the ability of a simulation model to replicate speed-flow graphs from real-world freeways. An objective function based on minimizing the dissimilarity between speed-flow graphs was developed by Menneni et al (2009). They measured the dissimilarity of two graphs by calculating the amount of area that is not covered by the other. Since speed and flow measurements were represented as point sets, discretization to convert point information to area was necessary. Moreover, considering that the information derived from the field and simulation was often just partial and not a complete speed-flow graph, the comparison was only made over the space occupied by the field graph. Differently from the above mentioned approaches, in this paper the measure of the closeness between empirical data and simulation outputs is achieved through a statistical approach which included hypothesis testing using *t*-test and confidence intervals. The description of the methodology will be shown in section 4.

**3. The fundamental diagram of traffic flow for the A22 Brenner Freeway, Italy**

Experimental surveys carried out at observation sections on the A22 Brenner Freeway, Italy, have allowed to model the relationships between the fundamental variables of traffic flow (namely the speed-flow-density relationships) for a traffic flow of cars only (Mauro 2003, 2005, 2007). Data were collected over different locations and multiple days and combined to show a complete graph between the pairs of traffic flow variables. The aforesaid relationships between flow and density, speed and density, speed and flow were developed for the right lane, the passing lane and the roadway, through the treatment and the processing of traffic data measured at specific observation sections (San Michele, Rovereto and Adige) on the A22 Freeway (Mauro 2003, 2005, 2007). A procedure for the estimation of the passenger car equivalent factors was also developed and reported in (Mauro 2003, 2005, 2007). For the same reference framework, under uninterrupted flow conditions an exploratory study has aimed to propose a criterion for predicting the reliability of freeway traffic flow by observing speed stochastic processes (see Mauro et al., 2013).
First the relationship between speed and density was searched. This choice is motivated by the following: considering the real traffic flow phenomenon, the speed-density relationship is a monotonically decreasing function and implies a mathematical relation simpler than the flow-density and speed-flow relationships; furthermore, the function $V=V(D)$ represents in a direct way the interaction between vehicles in a traffic stream, where users perceive, through the mutual spacing among vehicles, the density and to it adapt their speed. The main speed-density models as proposed by literature were taken into account (e.g. Greenshield, 1935; Greenberg, 1959; Underwood, 1961; Edie, 1965; May, 1990). The single-regime models were selected; among these, May's model (May, 1990) was chosen since it appeared as the best in interpreting the available data and the traffic flow phenomena at the observed sections, particularly the maximum values of density under congested traffic conditions. According to May's model (May, 1990), the relationship between speed and density was expressed mathematically as follows:

$$V = V_{FF} \cdot \exp\left[-0.5 \cdot \left(\frac{D}{D_c}\right)^2\right] \tag{1}$$

where $V_{FF}$ is the free flow speed and $D_c$ is the critical density, namely the density to which is associated the reaching of the capacity $C$. Equation 1 can be converted into linear form by using the logarithmic transformation:

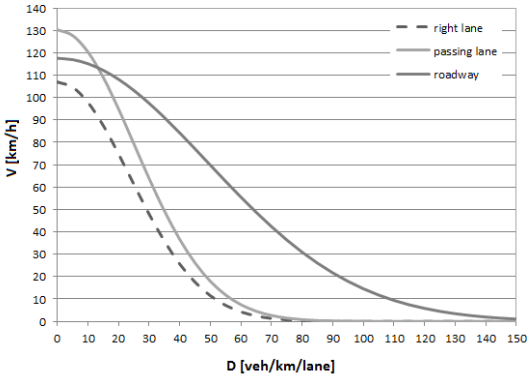
$$\ln(V) = \ln(V_{FF}) - \frac{1}{2 \cdot \mathrm{D}_c^2} \cdot D^2 \qquad \text{or else} \qquad V_1 = a + b \cdot D_1 \tag{2}$$

where $V_1$ is $ln(V)$, $a$ is $ln(V_{FF})$, $b$ is $-\frac{1}{2 \cdot \mathrm{D}_c^2}$ and $D_1$ is $D^2$, with $V_{FF}$ and $D_c$ as previously defined. Starting from the above equation, by means of the fundamental relation between flow $Q$, density $D$ and speed $V$, $Q = D \cdot V$, it was possible to obtain $Q$ as follows:

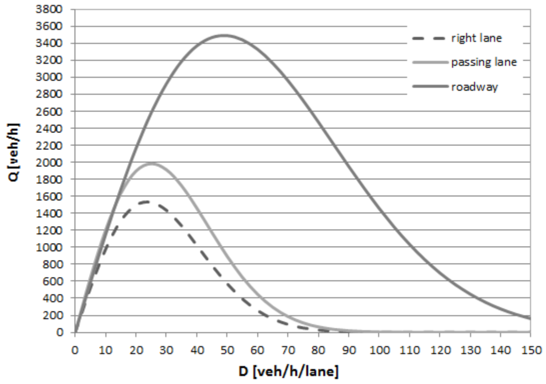
$$Q = V \cdot \sqrt{\frac{\ln\left(\frac{D}{D_c}\right)^2}{0.5 \cdot D_c^2}} \tag{3}$$

$$Q = V_{FF} \cdot D \cdot \exp\left[-0.5 \cdot \left(\frac{D}{D_c}\right)^2\right] \tag{4}$$

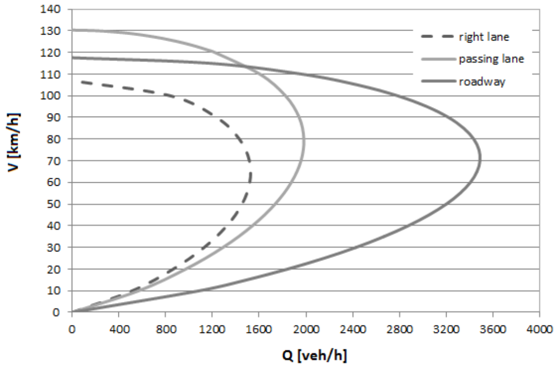
that allowed to obtain the speed-flow relationship, $V = V(Q)$, and the flow-density relationship, $Q = Q(D)$. For the complete specification of the relationships shown before the parameters $V_{FF}$ and $D_c$ were estimated. Traffic flow models were calibrated for the right lane, the passing lane and the roadway at the sections under examinations, by using the values of $Q$, veh/h, and $V$, km/h, and calculating the density $D$, veh/km/lane, from $D = Q/V$; then, for every value of speed $V$, corresponding to each lane and the roadway, the natural logarithm, $lnV$, was calculated to derive from each of the available pairs ($D$, $V$) the corresponding pair ($D^2$, $lnV$). For every observation section, basing on the corresponding scatter plot ($D^2$; $lnV$), according to equation 2, a least squares estimation was performed; then, the model calibration parameters ($V_{FF}$ and $D_c$) were calculated for all observation sections (see Mauro 2003, 2005, 2007). Thus, the relationships between the fundamental variables of traffic flow were specified for all observation sections by using equations 1, 3 and 4; estimations of capacity $C$ and speed $V_c$, corresponding to $C$, were then performed. For all cases, moreover, values of $R^2$ corresponding to ($V$; $Q$) and ($Q$; $D$) relationships are never found to be lower than 0.7. In order to calculate the speed-flow-density relationships for the right lane, the passing lane and the roadway for the A22 Freeway (Italy), the homologous determinations of $V_{FF}$ and $D_c$, corresponding to the three observation sections were averaged. Using the $V_{FF}$ and $D_c$ values the speed-flow-density relationships for the freeway under examination were obtained (see Fig. 1). In the following sections, however, empirical data which were taken as a reference in the calibration of the microsimulation model are those corresponding to S. Michele observation section (southbound), chosen as case study; for this observation section, Table 1 gives the speed-flow-density relationships.




a) Speed (v) vs. Density (D)




b) Flow (Q) vs. Density (D)




c) Speed (v) vs. Flow (Q)

| lane/lanes of travel | $V_{FF}$ | $D_c$ | C | $V_C$ |
|---|---|---|---|---|
| **right lane** | 106.95 | 23.65 | 1534.00 | 64.86 |
| **passing lane** | 130.28 | 25.09 | 1983.00 | 79.02 |
| **roadway** | 117.45 | 48.56 | 3459.00 | 71.23 |

**Fig. 1. The speed-flow-density relationships for the A22 Brenner Freeway, Italy**
*In the table the values of* $V_{FF}$ and $D_c$ *are given; the same table shows the values of C and* $V_c$ *calculated after the calibration of the speed-flow-density relationships.*
*Source: (Mauro, 2003;2005; 2007)*

**Table 1 The speed-density relationships for the S. Michele observation section**

| relationship | right lane | passing lane | roadway |
|---|---|---|---|
| speed-density | $V = 105.00 \cdot \exp\left[-0.5 \cdot \left(\frac{D}{24.36}\right)^2\right]$ | $V = 131.50 \cdot \exp\left[-0.5 \cdot \left(\frac{D}{24.67}\right)^2\right]$ | $V = 118.20 \cdot \exp\left[-0.5 \cdot \left(\frac{D}{48.35}\right)^2\right]$ |
| flow-density | $Q = 105.00 \cdot D \cdot \exp\left[-0.5 \cdot \left(\frac{D}{24.36}\right)^2\right]$ | $Q = 105.00 \cdot D \cdot \exp\left[-0.5 \cdot \left(\frac{D}{24.36}\right)^2\right]$ | $Q = 105.00 \cdot D \cdot \exp\left[-0.5 \cdot \left(\frac{D}{24.36}\right)^2\right]$ |

*Source: (Mauro, 2003;2005; 2007)*

## 4. Calibration Parameters

Traffic simulation for the freeway facility was performed with Aimsun micro-simulator. As for any other microsimulation software program, Aimsun comes with a set of user-adjustable parameters for the purpose of calibrating the model to local conditions, i.e. to minimize the difference between the empirical and the simulated values of the variables of interest. The Aimsun micro-simulator updates the vehicle position which moves along the network, basing on two driver behavior models named “car following” and “lane changing” (Barcelo, 2011). The car-following model implemented in Aimsun is an evolution of the empirical model proposed by Gipps (1981; 1986), in which the model parameters are determined by the influence of local parameters, depending on the type of driver, the road characteristics, the influence of vehicles driving in the adjacent lanes, etc. Very briefly the model consists of two components: acceleration, representing the intention of a vehicle to achieve a certain desired speed, and deceleration, reproducing the limitations imposed by the preceding vehicle when trying to drive at the desired speed. The car-following model proposed by Gipps considers only the vehicle and its leader; the implementation of this model in Aimsun also includes the influence of certain vehicles driving slower in the adjacent lane on the vehicle driving along a section. The model determines a new maximum desired speed of a vehicle in the section, considering the mean speed of vehicles driving downstream of the vehicle in the adjacent slower lane and allowing a maximum difference of speed (Barcelo, 2011). The lane-changing model can also be considered an evolution of the lane changing model proposed by Gipps (1986), according to which the lane change is modeled as a decision process analyzing the desirability of a lane change, the benefits of a lane change resulting from the attainment of the desired speed when the leading vehicle is slower, and the feasibility conditions for a lane change depending on the location of the vehicle in the road network. For the list of the car following and lane-changing model parameters for freeways the reader is referred to Barcelo (2011).

In order to find the set of parameter values for the model that best reproduces local traffic conditions at the A22 Freeway, the default values for the model parameters were used in trial simulation runs for checking any coding error; however, the outcomes of the comparison between simulation and empirical data showed that the default parameters provided simulation outputs which did not emulate properly the existing traffic flow characteristics. The fine tuning process involved the iterative changing of some parameters and simulation replications until the simulated pairs of speed and density, as closely as possible, matched the corresponding pairs observed on field. Due to unrealistic simulation results in comparison to field observations when Aimsun default parameters were used, some parameters were changed, basing on engineering knowledge and best practices. These parameters included the minimum headway, representing the time in seconds between the leader and the follower vehicle, the reaction time, or the time in seconds it takes a driver to react to speed changes in the preceding vehicle, and the minimum distance between vehicles or the distance, in meters, that a vehicle keeps between itself and the preceding vehicle when stopped. After having explored different combinations of values of the parameters (see Table 2), a value of 1.70 s was used for the minimum headway parameter instead of the default value of 2.10 s, whereas a value of 0.8 s was used for the reaction time parameter instead of the default value of 0.7 s; for the minimum distance between vehicles a values of 1 m was used instead of the default value of 1.10 m.

**Table 2 Calibration Parameters**

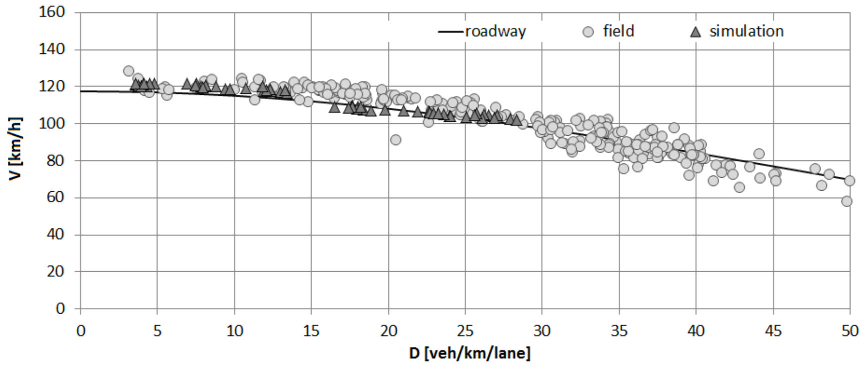
| Parameter | Default | Used | Levels | | |
|---|---|---|---|---|---|
| minimum headway - seconds | 2.10 | 1.70 | 1.70 | 1.90 | 2.10 |
| minimum distance between vehicles - meters | 1.10 | 1.00 | 1.00 | 5.00 | 10.00 |
| reaction time - seconds | 0.70 | 0.80 | 0.70 | 0.80 | 1.00 |

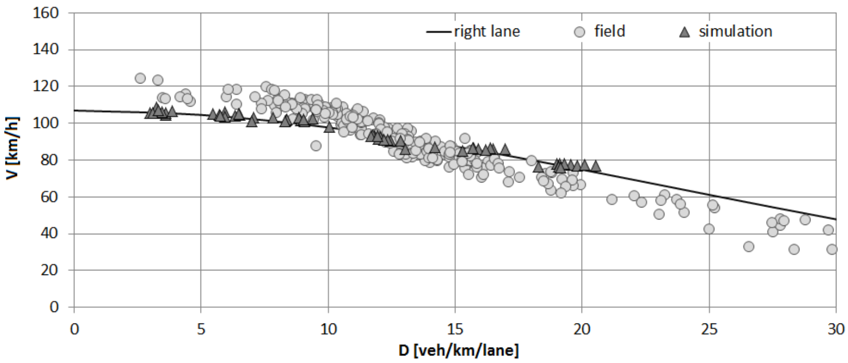
The calibration process also included the adjustments of the desired speeds, namely the maximum speed, in km/h, that a certain type of vehicle can travel at any point in the network. For example, a “car” vehicle type can be defined in Aimsun having as default values a mean desired speed of 110 km/h and a deviation of 10 km/h; desired speed for each vehicle of this type is sampled from a truncated Normal distribution (110, 10). According to observational data for A22 Freeway and what reported by Uddin and Ardekani (2002), the desired speeds on right lane were assumed lower than those in the passing lane; moreover, it was noted that the desired speed was sensitive to flow rate, tending to decrease as flow rate values became consistent (see Table 3).

**Table 3 Adjustments for the desired speed**

| flow rate [pcu/h] | desired speed, mean [km/h] | | |
|---|---|---|---|
| | right lane | passing lane | roadway |
| <1500 | 110 | 140 | 125 |
| 2000 | 100 | 140 | 115 |
| 2500 | 95 | 140 | 115 |
| >3000 | 90 | 130 | 115 |

In the simulation process, a 2 km long freeway segment centered on the S. Michele observation section (southbound) was used, having the cross section of A22 Brenner Freeway (Italy) and a grade equal to 0.09%; the aforesaid length was chosen so that all vehicles introduced into the segment exited at the end of the segment and no traffic entered and exited in the middle. For the freeway segment, 10 simulation replications were performed for 7 different values of traffic flow, increasing with step 500 from 500 veh/h to 3500 veh/h during a time interval of 4 hours; values of traffic variables generated during the first half-hour, namely the warm up, were excluded, because they were considered related to a motion condition not fully operational, and therefore unreliable. A fleet of cars only was used, choosing them within the range of cars that Aimsun allows to select. With regard to traffic generation, in the Aimsun micro-simulator different headway models may be selected as interval distributions; the exponential distribution is the default distribution among different headway models and it was chosen to model time intervals between two consecutive arrivals of vehicles. The simulation data were collected by placing detectors at exactly the same locations as detectors in the field. Furthermore, the simulated values were verified against the empirical values as indicated in the speed-density diagrams, where the plots of empirical and simulated data for S. Michele section (southbound) are shown in Figure 2; $\ln V$-$D^2$ regression lines for observed and simulated data for S. Michele section (southbound) will be shown in the next section, in which issues on implementing the methodology for calibrating the traffic microsimulation model will be introduced.



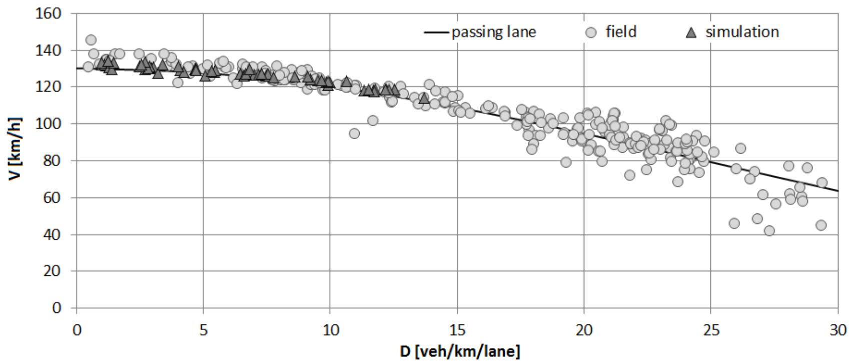




**Fig. 2. Speed-density graphs with plots of field and simulated data for S. Michele section (southbound)**
*In every graph the corresponding speed-density relationship of Table 1 is also shown .*

### 5. Using Speed-Density Relationships in the Calibration Process

The quality of the microsimulation model calibration was evaluated by comparing the $lnV$-$D^2$ linear regressions of all simulated speed/density values with the corresponding linear regressions of the empirical data. Thus, a statistical approach including hypothesis testing using *t*-test and confidence intervals was used as described briefly below. Suppose we observe, for $i = 1,...,n$, the measured variable $Y_i$ ($lnV_i$) corresponding to certain values of the input variables $x_i$ ( $D_i^2$ ) and we want to use them with the objective of estimating the regression parameters ($\alpha$ and $\beta$) in a simple linear regression model. If $A$ and $B$ are the estimators that we are searching for, then $A + Bx_i$ is the estimator of the response variable corresponding to the input variable $x_i$. In order to get the distribution of the estimators $A$ and $B$, additional assumptions necessarily have to be made. As starting point the estimators $A$ and $B$ are usually assumed to be independent, normally distributed with zero mean and constant variance $\sigma^2$. Consequently, if for $i = 1, 2, ..., n$, the measured variable $Y_i$ is the response given to the input variable $x_i$, we will assume that $Y_1, Y_2, ..., Y_n$ are independent and $Y_i \sim N(\alpha+\beta x_i, \sigma^2)$. Starting from the above proposition, a statistical test and confidence intervals for the regression parameter $\beta$ were constructed. As it is well known the hypothesis to be tested is that $\beta = 0$ (the response does not depends on the input variable, i.e. there is no correlation between the two variables). It can be easily demonstrated that the statistic for the test of interest has a *t* distribution with *n-2* degrees of freedom:

$$\sqrt{\frac{(n-2)\cdot S_{xx}}{SS_R}}\,\mathrm{B} \quad \sim t_{n-2}$$

where $S_{xx}$ is $\sum_i x_i^2 - n\bar{x}^2$ and $SS_R$ is the sum of squared residuals.

So, to test $H_0 : \beta = 0$ against $H_1 : \beta \neq 0$, at the $\gamma$ significance level, we have to:

reject $H_0$, if $\sqrt{\frac{(n-2)\cdot S_{xx}}{SS_R}}\,|\mathrm{B}| > t_{\frac{\gamma}{2},n-2}$

accept $H_0$ otherwise.

Thus an interval containing $\beta$, at the $1-\gamma$ confidence level, is given by:

$$\left(B - t_{\frac{\gamma}{2},n-2}\cdot\sqrt{\frac{SS_R}{(n-2)\cdot S_{xx}}},\; B + t_{\frac{\gamma}{2},n-2}\cdot\sqrt{\frac{SS_R}{(n-2)\cdot S_{xx}}}\right)$$

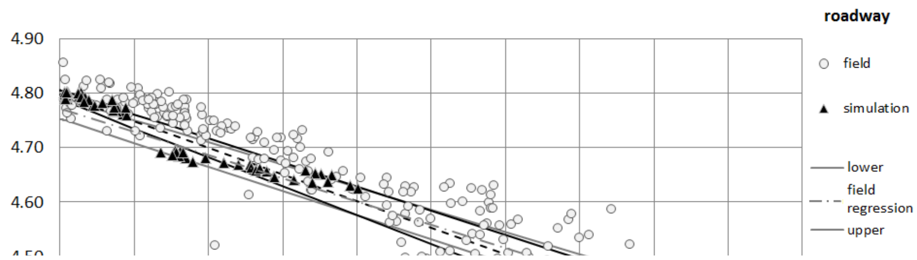
The determination of the confidence intervals and statistical tests for the regression parameter $\alpha$ was obtained as for $\beta$. So, the confidence interval at the 1- $\gamma$ level is given by:

$$A \pm t_{\frac{\gamma}{2},n-2}\cdot\sqrt{\frac{SS_R\cdot\sum_i x_i^2}{n(n-2)S_{xx}}}$$

Figure 3 shows the $lnV$-$D^2$ regression lines for observed and simulated data for S. Michele section (southbound), both for the right lane and the passing lane, as well as for the entire roadway; on each set of data, statistical inference on the regression parameters (intercept and slope) was performed by means of a *t*-test at the significance level of 5%. Comparing the two regressions from empirical and simulated data, including confidence areas, a substantial overlapping of the regression curves can be noted (see Fig. 3); however, the most important thing is that the simulated data fell almost entirely within the confidence band of the regression line fitted to the observed data. This allowed to state that the microsimulation model is able to reproduce the real phenomenon of traffic flow within a wide enough range of operations, from the free flow conditions until almost to the critical density. At the same time the methodology showed that, if only one regime of traffic flow (for instance, the congested flow conditions) had been considered, we would not have had any insurance on the ability of the model to reproduce, just as well, the real operations at different regimes of traffic flow. It should be emphasized the exploratory nature of the analysis carried out in this study in which, among all models analyzed, only the single-regime model was considered having the accuracy and consistency to interpret the experimental data which covered the three traffic regions (i.e., free-flow, congested, and queue discharge), and to represent the operating conditions for each lane and the entire roadway. Nevertheless, in order to improve the calibration process, one can hypothesize to model separately the inside lane and the outside lane and further survey should be conducted to relax the single-regime assumption.

## 6. Discussion and Conclusions

In this paper a methodology which uses speed-density relationships in the microsimulation calibration process is presented. Statistical analysis technique of pattern recognition was used to evaluate the matching of speed-density relationships from field and simulation. Traffic patterns were implemented developing relationships between the variables of traffic flow for empirical and simulated data. For the former we referred to traffic data observed at A22 Freeway (Italy); for the latter, Aimsun microscopic traffic simulator software was used with reference to a test freeway segment under uncongested traffic conditions for a fleet of cars only. Differently from the methodological approaches reported by technical literature on the topic of interest, in this paper the measure of the closeness between empirical data and simulation outputs was achieved through a statistical approach which included hypothesis testing and confidence intervals. Encouraging results were obtained from the comparison of the two sets of observed and simulated data; indeed, a substantial overlapping of the regression curves was obtained and, primarily the simulated data fell almost entirely within the confidence band of the regression line fitted to the empirical data. This allowed to state that the microsimulation model was able to reproduce the real phenomenon of traffic flow within a wide enough range of operations, from the free flow conditions until almost to the critical density. Conversely, the methodology here proposed showed that, if only one regime of traffic flow (free flow conditions or, congested flow conditions) had been considered, we would not have had any insurance on the ability of the model to reproduce, just as well, the real operations at different regimes of traffic flow.




***field***

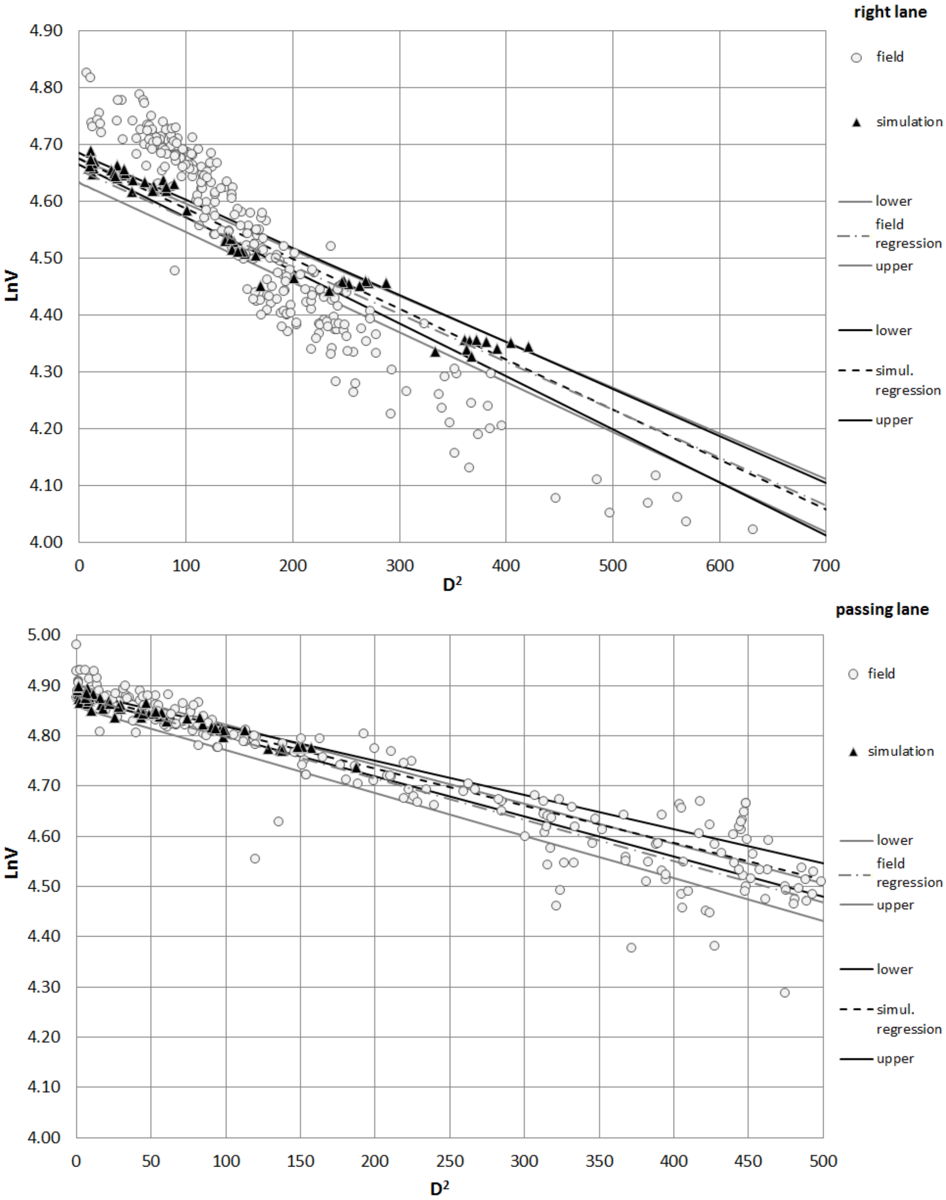
$y = -2.1390\cdot E^{-04}\,x + 4.7726$

$R^2 = 0.9067$

***simulation***

$y = -2.4417 \cdot E^{-04}\, x + 4.7972$

$R^2 = 0.9204$




***field***

$y = -8.4291 \cdot E^{-04}\, x + 4.6540$

$R^2 = 0.8762$

***simulation***

$y = -8.8134E^{-04}\, x + 4.6744$

$R^2 = 0.9655$

***field***

$y = -8.2173 \cdot E^{-04}\, x + 4.8789$

$R^2 = 0.8993$

***simulation***

$y = -7.380E^{-04}\, x + 4.8819$

$R^2 = 0.9395$

**Fig. 3. Speed-density graphs with plots of field and simulated data for S. Michele section (southbound)**

At last, the deepening of the model calibration as presented in this paper has led the authors to develop some considerations of general order that will be summarized in the following:

i. first, it should be noted that although the results of the calibration process may seem satisfactory, the analyst does not have any guarantee on his/her work: he/she may have changed (or, in the extreme, forced) some parameters, but may have neglected other parameters even more important. However, it must be said that this risk can be contained when information for the calibration process is derived from the speed-flow, speed-density, or flow-density graphs since a higher number of parameters can be submitted to the calibration process, resulting in a better fine-tuned simulation model. Moreover, the above relationships provide information about the free-flow, congested, and queue discharge regions, which cannot be gained from a single numerical value or a distribution of capacities;

ii. second, although microsimulation model gave us data that, on the whole, belong to the population of the observed data, some doubts could relate to what was developed for the right lane. One single model which fits to empirical data both for the right lane and the passing lane, as well as for the entire roadway, does not always represent the best choice. The empirical observations have gradually led to consider that the modeling of the speed-density relationship (and then the associated fundamental diagram) could be improved differentiating by each lane; for instance, with regard to the capability of the model (single regime or multi regime) to fit empirical observations reasonably well over the entire range of a traffic variable (i.e. flow, speed or density). The inability of single regime models to perform well over the entire range of density may prompt to think about fitting the data in a piecewise manner through multiple equations;

iii. third, a question that concerns the way in which vehicles are generated. At the beginning of the simulation run the system is empty; vehicles are generated at the entry nodes of the analytical network, based on the input volumes and an assumed headway distribution. The users of the microsimulation program can choice among different headway models (exponential, uniform, normal, constant, etc.), but the default distribution is often selected. It is known, however, that the choice of the distribution should depend on how much complexity is desired, as well as the behavior of the traffic. It is well known, indeed, Poisson distribution for vehicle counts and negative exponential distribution for time headways are only applicable when traffic flows are light (i.e. no interaction between vehicles, thus enabling them to move at random). As traffic becomes heavier, vehicles are restricted in their driving freedom, and interaction between vehicles increases; moreover, the exponential distribution provides nonzero probabilities even for very small values of headways. Because the poor agreement between the frequencies of headways observed in practice and the frequencies predicted by the negative exponential distribution, as well as theoretical considerations precluding very short headways, distributions different from the exponential should have to be used as a means of improving the capability of microsimulation models to replicate the real traffic phenomenon. It follows that in microsimulation the use of the usually preset headway distributions can produce inadequate choices in traffic generation, and a user-defined program is required to feed the network with vehicles; this usually requires further computational effort and time.

**Acknowledgements**
Authors wish to thank Dr. Eng. Walter Pardatscher, CEO for the A22 Brenner Freeway (Italy) for the constant availability shown during the development of this research.

**References**

Barceló, J. 2011. *Fundamentals of Traffic Simulation*. International Series in Operations Research and Management Science, vol. 145. Springer eds, 442 p.

Brilon, W.; Geistefeldt, J.; Zurlinden, H. 2007. Implementing the Concept of Reliability for Highway Capacity Analysis. . In *Proceedings of the 87th Annual Meeting of the Transportation Research Board*, Washington, D.C., USA.

Dowling, R.; Skabardonis, A.; Halkias, J.; McHale, G.; Zammit, G. 2004a. Guidelines for Calibration of Microsimulation Models: Framework and Applications. *Transportation Research Record*, 1876/2004: 1–9.

Dowling, R.; Skabardonis, A.; Vassili, A. 2004b. Traffic Analysis Toolbox Volume III: Guidelines for Applying Traffic Microsimulation Software. Report No. FHWA-HRT-04-040. Available from Internet: <http://ops.fhwa.dot.gov/trafficanalysistools/tat_vol3/foreword.htm>.

Edie, L. C. 1965. Discussion of Traffic Stream Measurements and Definitions. In *Proceedings of the 2nd International Symposium on the Theory of Traffic Flow*, OECD, Paris, France, pp. 139–154.

Fellendorf, M.; Vortisch, P. 2001. Validation of the Microscopic Traffic Flow Model VISSIM in Different Real-World Situations. In *Proceedings of the 80th Annual Meeting of the TRB*, Washington, D.C., USA.

Gipps, P.G. 1981. A behavioral car-following model for computer simulation. *Transportation Research Record: Journal of the Transportation Research Board,* 15-B: 105–111.

Gipps, P.G. 1986. A model for the structure of lane-changing decisions. *Transportation Research Record: Journal of the Transportation Research Board*, 20-B(5): 403–414

Greenberg, H. 1959. An analysis of traffic flow. *Operational Research*, 7(1): 79-85.

Greenshields, B. D. 1935. *A Study of Traffic Capacity*. Highway Research Board Proceedings, Vol. 14, pp. 448–477.
Hollander, Y.; Liu, R. 2008. The principles of calibrating traffic microsimulation models, *Transportation*, 35(3): 347-362.

Hourdakis, J.; Michalopoulos. P.G.; Kottommannil, J. 2003. Practical procedure for calibrating microscopic traffic simulation models. *Transportation Research Record,* 1852/2003: 130–139.

Kim, K.O.; Rilett, L.R. 2003. Simplex-based calibration of traffic microsimulation models with intelligent transportation systems data. *Transportation Research Record,* 1855/2003: 80–89.

Mauro, R. 2003; 2005; 2007. *Analisi di traffico, elaborazione di modelli e sistemi per la stima dell'affidabilità per l'Autostrada A22, Italia [Traffic analysis, development of models and systems for estimating reliability on the A22 Freeway, Italy]*. Technical report, Part I,II, III - Autostrada del Brennero, Trento, Italy (in Italian).

Mauro, R.; Giuffrè, O.; Granà, A. 2013. Speed Stochastic Processes and Freeway Reliability Estimation: Evidence from the A22 Freeway, Italy. *Journal of Transportation Engineering,* 139(12): 1244–1256.

May, A. D. 1990. *Traffic Flow Fundamentals*. Prentice Hall, Inc. USA. 473 p.

Menneni, S.; Sun, C.; Vortisch, P. 2009. Microsimulation Calibration Using Speed-Flow Relationships. *Transportation Research Record*, 2088/2009: 1-9.

Pidd, M. 1992. *Computer Simulation in Management Science*, John Wiley.

Toledo, T.; Koutsopoulos, H.N.; Davol, A.; Ben-Akiva, M.E.; Burghout, W.; Andreasson, I.; Johansson, T., Lundin, C. 2003. Calibration and validation of microscopic traffic simulation tools. Stockholm case study. *Transportation Research Record,* 1831/2003: 65–75.

Toledo, T.; Koutsopoulos, H. N. 2004. Statistical Validation of Traffic Simulation Models. *Transportation Research Record,* 1876/2004: 142–150.

Uddin, M.S.; Ardekani, S. 2002. An Observational Study of Lane Changing on Basic Freeway Segment. *In Proceedings of the 81st Annual Meeting of the TRB, Washington, D.C., USA*.

Underwood, R.T. 1961. *Speed, volume and density relationships*. Quality and Theory of Traffic Flow. Bureau of Highway Traffic, Yale University, New Haven, Conn., pp. 141-187.

Wiedemann, R. 1991. Modelling of RTI-Elements on Multi-Lane Roads. In *Advanced Telematics in Road Transport,* Vol. II, Elsevier, Amsterdam, Holland, 1991, pp. 1001–1019.

# GENERALIZATION OF THE ERLENKOTTER APPROACH FOR SOLVING OF THE PUBLIC SERVICE SYSTEM DESIGN

**Ján Bendík** [1]
*[1]Faculty of Management Science and Informatics, University of Zilina, Univerzitna 8215/1, 010 26 Zilina, Slovakia*

**Abstract:** This paper deals with the problem of designing the optimal structure of a public service system. Real instances of these problems are characterized by considerably big number of possible service center locations, which can take the value of several hundreds or thousands. The problem can be often formulated as forms of the weighted p-median problem. The optimal solution of smaller instances of the problem can be obtained by universal IP solvers. The universal IP solvers are very time-consuming and often fail when a large instance is solved. Our exact approach to the problem is based on the Erlenkotter procedure for solving of the uncapacitated facility location problem. We generalize the Erlenkotter dual approach to lower bounding to be able to solve the associated location problem with restricted number of the located service centers. To solve the problem we design the exact algorithm based on the branch and bound method. We verify the exact algorithm by comparison of the solutions obtained by our approach with the optimal solutions obtained by universal IP solver XPRESS-IVE in the computational time and the quality of solutions obtained for smaller instances of the problem.

**Keywords**: Erlenkotter approach, p-median location problem, public service system design

## 1. Introduction

The public service system structure is formed by deployment of a limited number of the service centers and the associated objective is to minimize total costs. The family of public service systems includes public administration system design, medical emergency system (Marianov, 2002), fire-brigade deployment and many others. The public service system design problem are often related to the p-median location problem or the uncapacitated facility location problem. The p-median location problem has become one of the most well-known and studied problems in the field of facility location. This problem is formulated as a task of determination of at most $p$ network nodes as facility locations. In real problems, the number of serviced customers takes the value of several thousands and the number of possible facility locations can take this value as well. The number of possible service center locations seriously impacts the computational time. To obtain good decision on facility location in any serviced area, a mathematical model of the problem can be formulated and some of mathematical programming methods can be applied to find the optimal solution. Solving the p-median location problem in the public service system design is the NP-hard problem (Garey 1979, Hakimi 1979). Sassano and Vasil´ev (2003) presented a branch-and-price-and-cut algorithm to solve large-scale instances of the p-median problem. Balinski (1965) provided an early integer programming formulation of the plant location problem that has historically been adapted to the p-median problem. Reese (2006) summarized the exact solution methods for the p-median problem and Mladenovic (2007) summarized the heuristic methods. Many authors have dealt with this problem. The p-median location problem is very similar with the uncapacitated facility location problem. Erlenkotter (1978) proposed one of the most effective algorithms for solving the uncapacitated facility location problem. Erlenkotter approach is based on the theory of duality and the branch and bound algorithm. The algorithm DualLoc realizes the suggested approach. These algorithms exploit the relation between the primary and dual formulation of the strong linear programming relaxation of the original problem. And when needed, the procedures, calculating the dual and induced primary solution, are followed by branch and bound method. Inspired by this approach Korkel (1989) improved the Erlenkotter approach and designed the algorithm PDLoc. Janacek and Buzna (2008) improved the Erlenkotter and Korkel approach and designed the algorithm BBDual for solving the uncapacitated facility location problem.

We generalize the Erlenkotter dual approach to lower bounding to be able to solve p-median location problem. We design the exact algorithm based on the branch and bound method and Erlenkotter approach and verify the exact algorithm by comparison of the solutions obtained by our approach with the optimal solutions obtained by the universal IP solver XPRESS-IVE in the computational time and the quality of solutions obtained for smaller instances of the problem.

## 2. Problem formulation and solution design

The p-median location problem finds the optimal location of exactly p facilities, so that the sum of the distances between the closest facilities and their costumers is minimized. The p-median location problem consists of placing facilities such as hospitals, police stations, warehouses in some sites of a given finite set $I$. The placing facilities serve the customers such as villages and cities from a given finite set $J$. The total costs of the optimal deployment of facilities in the specific network are constituted the fixed charges $f_i$ and the costs $c_{ij}$. The fixed charges $f_i$ give costs for the facility location at the location $i$. The costs $c_{ij}$ give costs for the demand satisfaction of the j-th customer from the location $i$.

[1] Corresponding author: Jan.Bendik@fri.uniza.sk

### 2.1. Mathematical model of p-median location problem

The formulated p-median location problem can be modeled using the following notation. Let the decision of the service center location at the place $i \in I$ be modeled by a zero-one variable $y_i \in \{0, 1\}$ which takes the value of 1, if a center is located at $i$, otherwise it takes the value of 0. In addition, the variables $z_{ij} \in \{0, 1\}$ for each $i \in I$ and $j \in J$ represent to assign a customer $j$ to a possible location $i$ by the value of 1. The maximal number of the facility locations represents the constant $p$. The mathematical model for the p-median location problem is defined as follows:

$$Minimize \qquad F_P = \sum_{i \in I} f_i\, y_i + \sum_{i \in I} \sum_{j \in J} c_{ij}\, z_{ij} \tag{1}$$

$$Subject\ To \qquad \sum_{i \in I} z_{ij} = 1 \qquad for\ j \in J \tag{2}$$

$$z_{ij} \le y_i \qquad for\ i \in I, \quad j \in J \tag{3}$$

$$\sum_{i \in I} y_i \le p \tag{4}$$

$$y_i \in \{0,1\} \qquad for\ i \in I \tag{5}$$

$$z_{ij} \in \{0,1\} \qquad for\ i \in I, \quad j \in J \tag{6}$$

The objective function (1) represents the minimization of the total costs of the p-median location problem which consists of the fixed charges $f_i$ and the costs $c_{ij}$. The constraints (2) ensure that each customer is assigned to the exactly one possible service center location. Binding constraints (3) enable to assign a customer to a possible location $i$, only if the service center is located at this location. The constraint (4) bounds the number of the located service centers. The obligatory conditions in the mathematical model are (5) and (6). This p-median location problem without the condition (4) gives the uncapacitated facility location problem (UFLP).

### 2.2. Erlenkotter approach for solving uncapacitated facility location problem

The basic idea of Erlenkotter approach consists in relation between linear relaxation of the original problem and the associated dual problem. After some reformulation and introduction of slack variables $u_i$ , the dual problem takes the following form (7-10):

$$Maximize \qquad F_D = \sum_{j \in J} v_j \tag{7}$$

$$Subject\ To: \qquad \sum_{j \in J} max\{0, v_j - c_{ij}\} + u_i = f_i \qquad for\ i \in I \tag{8}$$

$$v_j \ge 0 \qquad for\ j \in J \tag{9}$$

$$u_i \ge 0 \qquad for\ i \in I \tag{10}$$

The dual variables $v_j$ correspond to constraints (2). According to the weak duality theorem objective function value $F_D$ of any feasible solution is smaller or equal to any objective function value $F_p$ of any feasible solution of the linear relaxation of problem (1–3,5-6). A lower bound for optimal solution of the problem (1-3, 5–6) constitutes objective function value of arbitrary feasible solution of problem (7 – 10). Erlenkotter (1978) obtained a good lower bound from the dual problem than the combination of two procedures. The dual ascent procedure (DA) starts from an arbitrary feasible solution of the dual problem and subsequently increases the values of the $v_j$ variables as long as constraints (8-10) hold. Janacek and Buzna (2008) improved DA procedure (Fig. 1).

**Procedure DA$^*$ ($J^+$)**
Order the set $J^+$ of relevant customers into a sequence $j_1, j_2, \dots j_k \dots j_n$
ascendingly with respect to the cardinalities $K_j$.
For $k = 1$ to $|J^+|$ do:
  Set $j := j_k$.
  Set $d = min\{min\{s_i : i \in I, c_{ij} \leq v_j\}, \{c_{ij} - v_j : i \in I, c_{ij} > v_j\}\}$.
  If $d > 0$, then:
    Update $s_i$ for $i \in I$, $c_{ij} \leq v_j$ by $s_i := s_i - d$. Set $v_j := v_j + d$ and $z_D = z_D + d$.
    If the cardinality $K_j$ has increased, then reorder the sequence of customers $j_k, \dots, j_n$.
Terminate.

**Fig. 1.** *DA* procedure improved by Janacek and Buzna*
*Source: Janacek and Buzna (2008)*

The dual Adjustment (DAD) improved of the dual solution obtained by the DA$^*$ procedure. The dual adjustment procedure (DAD) searches for a configuration, in which a decrease of some variable $vj$ by value $\delta$ will create free space at least at two locations $i$ and $i^*$ from $I$ , which can be used for an ascent of at least two different variables $v_k$ and $v_l$ $(k \neq l \neq j)$. A dual solution and an induced primal feasible solution can be obtained by applying the complementary constraints (11)–(13):

$$(y_i - z_{ij})\, max\{0, v_j - c_{ij}\} = 0 \qquad for\ i \in I, \quad j \in J \tag{11}$$

$$u_i y_i = 0 \qquad for\ i \in I \tag{12}$$

$$max\{0, c_{ij} - v_j\}\, z_{ij} = 0 \qquad for\ i \in I, \quad j \in J \tag{13}$$

For construction of the associated primal solution to the dual solution Erlenkotter approach minimizes the difference between primal and dual value of objective function which it gives the following notation (14):

$$\boldsymbol{F_p} - \boldsymbol{F_D} = \sum_{i \in I} \sum_{j \in J} (y_i - z_{ij})\ max\{0, v_j - c_{ij}\} + \sum_{i \in I} u_i y_i + \sum_{i \in I} \sum_{j \in J} z_{ij}\, max\{0, c_{ij} - v_j\} \tag{14}$$

$F^*_p$ represents a value of objective function uncapacitated facility location problem (1-3, 5-6). $F_D$ represents a value of objective function of the associated dual problem.

### 2.3 Generalization of Erlenkotter approach for solving p-median location problem

Our approach is based on Erlenkotter approach for solving the uncapacitated facility location problem. We formulate the dual problem of the p-median location problem (1-6) as follows:

$$Maximize \qquad F_D^* = \sum_{j \in J} v_j + px \tag{15}$$

$$Subject\ To: \qquad \sum_{j \in J} max\{0, v_j - c_{ij}\} + x + u_i = f_i \qquad for\ i \in I \tag{16}$$

$$v_j \geq 0 \qquad for\ j \in J \tag{17}$$

$$x \leq 0 \tag{18}$$

$$u_i \geq 0 \qquad for\ i \in I \tag{19}$$

The dual variables $v_j$ correspond to constraints (2) and the new dual variable $x$ correspond to constraint (4). According to the weak duality theorem objective function value $F_D^*$ of any feasible solution is smaller or equal to any objective function value $F_p$ of any feasible solution of the linear relaxation of (1-6). A lower bound for optimal solution of the problem (1-6) constitutes objective function value of arbitrary feasible solution of (16 – 19). Our approach uses the Erlenkotter combination of procedures DA$^*$ and DAD for obtained a good lower bound from the dual problem, but with some modifications. The correct dual ascent procedure (CDA) starts from an arbitrary feasible solution of the dual problem and subsequently increases the values of the $v_j$ variables and decreases the value of the $x$ variable as long as constraints (16-19) hold. The new CDA procedure has following notation:

```
Execute Procedure DA*
While sum of dual variables u_i=0 > p and all customers are not processed
        Order the set J of relevant customer
        Choose the active customer j
        Backup value of all dual variables v_j, x, u_i
        Execute Procedure DA* with updating the dual variable x
        If lower bound does not decrease then back to the original saved values
        Find next customer j for processing
Terminate
```

**Fig. 2.** C*DA procedure*

The dual Adjustment (DAD) improved of the dual solution obtained by the CDA procedure. A dual solution and an induced primal feasible solution can be obtained by applying the complementary constraints (20)–(23):

$$(y_i - z_{ij})\, max\{0, v_j - c_{ij}\} = 0 \qquad for\ i \in I, \quad j \in J \tag{20}$$

$$(p - \sum_{i \in I} y_i)\, x = 0 \tag{21}$$

$$(f_i - x - \sum_{j \in J} max\{0, v_j - c_{ij}\})\, y_i = 0 \qquad for\ i \in I \tag{22}$$

$$max\{0, c_{ij} - v_j\}\, z_{ij} = 0 \qquad for\ i \in I, \quad j \in J \tag{23}$$

New complementary condition is condition (21). For construction of the associated primal solution to the dual solution our approach minimize the difference between primal and dual value of objective function which it gives the following notation (14):

$$F_p - F_D^* = \sum_{i \in I} \sum_{j \in J} (y_i - z_{ij})\ max\{0, v_j - c_{ij}\} + (\sum_{i \in I} y_i - p)\, x + \sum_{i \in I} u_i y_i + \sum_{i \in I} \sum_{j \in J} z_{ij}\, max\{0, c_{ij} - v_j\} \tag{24}$$

$F_p$ represents a value of objective function of the p-median location problem (1-6). $F_D$ represents a value of objective function of the associated dual problem.

### 2.4 Algorithm for solving the public service system design

Erlenkotter designed the algorithm DualLoc for solving the uncapacitated facility location problem. Korkel designed the algorithm PDLoc, Janacek and Buzna improved the Erlenkotter approach and designed the algorithm BBDual for solving same problem. The BBDual algorithm is based on the branch and bound method. Two subproblems solve by fixing the variables *yi* to ones or zeros. If a subproblem should be processed or excluded from the searching process then the obtaining the lower bound must be very effective. Lower bound in algorithm BBDual is obtained by successively performing the dual ascent and the dual adjustment algorithms for solving the uncapacitated facility location problem.

I designed the algorithm pMedBBDual for solving p-median location problem by applying the CDA procedure. This algorithm is extension of algorithm BBDual about CDA procedure. The procedure makes possible to obtain the lower bound of p-median location problem. We verify CDA procedure for obtaining the good lower bound and algorithm pMedBBDual on the experiments from Slovak road network.

### 2.5 Numerical experiments

The experiments in the Table 1 was realized on the benchmarks from the set of 100 places in the district of Banska Bystrica. Column *p* in the Table 1 gives the maximal number of the facility locations. Column $OF_{ES}$ gives the value of optimal solution obtained by univerzal IP-solver XPRESS IVE, column $OF_{UB}$ gives value of solution p-median location problem with algorithm pMedBBDual, column $OF_{LB}$ gives value of lower bound of solution obtained by algorithm pMedBBDual and columns *t(s)* gives the computational time in seconds.

**Table 1**
*Comparison of solution obtained by algorithm pMedBBDual and XPRESS IVE*

| **BB100x100** | **XPRESS-IVE** | | **pMedBBDual** | | |
|---|---|---|---|---|---|
| ***p*** | ***$OF_{ES}$*** | ***t(s)*** | ***$OF_{UB}$*** | ***$OF_{LB}$*** | ***t(s)*** |
| 2 | 3508 | 0,75 | 3508 | 3508 | 0,143 |
| 4 | 2391 | 0,86 | 2401 | 2328 | 0,257 |
| 6 | 1845 | 0,92 | 1869 | 1844 | 0,236 |
| 8 | 1507 | 0,85 | 1511 | 1500 | 0,444 |
| 10 | 1275 | 0,76 | 1331 | 1268 | 0,541 |
| 12 | 1121 | 0,65 | 1128 | 1115 | 0,402 |
| 14 | 981 | 0,66 | 983 | 977 | 0,251 |
| 16 | 879 | 0,69 | 879 | 878 | 0,374 |
| 18 | 811 | 0,88 | 821 | 804 | 0,469 |
| 20 | 745 | 0,65 | 754 | 739 | 0,565 |
| 22 | 688 | 0,99 | 694 | 688 | 1,108 |
| 24 | 639 | 1,12 | 650 | 637 | 2,144 |
| 26 | 596 | 0,81 | 599 | 595 | 0,780 |
| 28 | 558 | 1,35 | 559 | 556 | 1,088 |
| 30 | 521 | 0,75 | 521 | 520 | 2,695 |

The experiments show that we can solve the p-median location problem with applying the CDA procedure for lower bound. We obtain a lower bound in short time and with very good value in most of instances. The optimal solution was found in little bit cases, so we must design the better heuristic for obtaining the minimal set of location. The computational time is the most of instances better with the algorithm pMedBBdual, but differences are minimal.

## 3. Conclusion

Problem of designing the optimal structure of a public service system is NP-hard problem. Our approach to the p-median location problem is based on Erlenkotter approach for solving the uncapacitated facility location problem. We generalized the Erlenkotter dual approach to lower bounding to be able to solve p-median location problem. We designed the CDA procedure and algorithm pMedBBDual based on the branch and bound method and Erlenkotter approach and verified the algorithm pMedBBDual by comparison of the solutions obtained by our approach with the optimal solutions obtained by universal IP solver XPRESS-IVE in the computational time and the quality of solutions obtained for smaller instances of the problem. The experiments showed that we can solve the p-median location problem with applying the CDA procedure and obtain a lower bound in short time and with very good value in most of instances. We found the optimal solution in a few cases. The computational time is better in the most of instances with the algorithm pMedBBdual than using by the universal IP-solver XPRESS IVE, but differences are minimal. We must design better heuristic for obtaining the minimal set of location. This heuristic minimizes the relation (24). To improve the heuristic for obtaining the minimal set of location, we can obtain optimal solution and test the algorithm pMedBBDual on larger-scale problem.

## Acknowledgements

This work was supported by the research grants VEGA 1/0296/12 “Public Service Systems with Fair Access to Service” and APVV-0760-11 “Designing of Fair Service Systems on Transportation Networks”. We would also like to thank to „*Centre of excellence for systems and services of intelligent transport*“**(**ITMS 26220120050) for built up the infrastructure, which was used.

**References**

Avella, P.; Sassano, A.; Vasil´ev, I. 2007 Computational study of large-scale p-median problems. Mathematical Programming 109: 89-114

Balinski, M. 1965. Integer programming, methods, uses and computation. Management Science 12: 254–313
Erlenkotter, D. 1978. A Dual-Based procedure for uncapacitated facility location, Operations Research 26(6): 992-1009.

Garey, M. R.; Johnson, D. S. 1979. Computers and intractability: A guide to the theory of NP-completeness. W.H. Freeman and Co., San Francisco. 338 p.

Garcia, S.; Labbé, M.; Marín, A. 2011. Solving large p-median problems with a radius formulation. IN-FORMS Journal on Computing 23(4): 546-556.

Hakimi, S. L.; Kariv, C. 1979. An algorithmic approach to network location problems. SIAM Journal on Applied Mathematics 37: 539–560.

Janacek, J.; Buzna, L. 2008. An acceleration of Erlenkotter-Korkel´s algorithms for the uncapacitated facility location problem. Annals of Operations Research 164: 97-109.

Körkel, M. 1898. On the exact solution of large – scale simple plant location problem, European Journal of Operational Research 39: 157-173.

Marianov, V.; Serra, D. 2002. Location problems in the public sector. Drezner Z (ed.) et al. Facility location. Applications and theory. Berlin: Springer. 119-150 p.

Mladenovic, N.; Brimberg, J.; Hansen, P.; Moreno-Perez, J.A. 2007. The p-median problem: a survey of metaheuristic approaches. European Journal of Operational Research 179: 927–939.

Reese, J. 2006. Solution methods for the p-median problem. Networks 48(3): 125–142.

# EVALUATION OF RAILWAY LINE CAPACITY USING GENERAL PURPOSE SIMULATION TOOL

**Peter Márton**[1]
*[1] Department of Mathematical Methods and Operations Research, Faculty of Management Science and Informatics, University of Žilina*

**Abstract:** A White paper of the European Commission (Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system) states that 30% of road freight over 300 km should shift to other modes of transport, such as rail transport, by 2030, and more than 50% by 2050. It is also asserted that by 2050, the majority of medium-distance passenger transport in Europe should be on rails. In order to prepare rail transport to be able to absorb expected flows of passengers and goods, reconstruction and revitalization of main and regional railway lines is necessary. One of the goals of railway lines reconstruction is to increase their capacity. Several capacity research methods exist. They can be classified as technology-oriented types of performance evaluations and are used for determination of adequate track infrastructure and signal systems design, operating performance and quality of operation. In principle, there are two methods for capacity research – an analytical and a simulation approach. The paper will be focused on the simulation approach.

**Keywords**: simulation tool, railway line capacity, automatic block

## 1. Introduction

Reconstruction of railway lines was started in several new states of the European Union in the last decade with support of EU cohesion funds (Pélerin 2009; Mašek 2013; Mašek 2014). Railway corridors are being upgraded for speed of 160 km/h or more. In some states, regional lines are reconstructed too. An integral part of the upgraded lines are signaling and interlocking systems.

Railway line upgrade has to be planned using a holistic approach. Upgraded tracks, a new signaling and interlocking system and a newly constructed timetable should lead to customer satisfaction with railway transport services. There is a crucial interconnection (dependence) between a signaling and interlocking system and railway line capacity.

Decisions about new design (tracks, signaling and interlocking system, etc.) of an upgraded railway line have to be verified before they are realized, e.g. it is very important to verify if line capacity will be the same or higher after an upgrade. In this paper, the author deals with a simulation model of operation on a railway line. First results regarding development and building a simulation model of a railway line without a signaling and interlocking system are proposed, altogether with a description of how to build a simulation model of a railway line equipped with this system. This type of signaling and interlocking system, also called the automatic block, is used mainly on railway corridors. This system uses trackside signals that allow movement of a train along block sections. It is believed that the proposed simulation model can support some railway line simple capacity evaluation and research.

In order to make this research study and proposed methodology more understandable, brief explanations of all important terms from the railway field are included, using definitions that can be found in works of Hansen and Pachl (2008), and Theeg and Vlasenko (2009).

### 1.1. Railway line capacity

Railway line capacity can be defined as the maximum traffic flow per track, i.e. number of trains per day or peak hour. Several basic capacity types are distinguished – theoretical vs. practical and used vs. available (Gašparík 2014a; Gašparík 2014b; Jánošíková 2014; Rici 2009). This paper will refer to the first and the second capacity type. Firstly, theoretical capacity represents the number of trains that can run over a railway line, during a specific time interval, in a strictly perfect and mathematically generated environment, with trains running permanently and ideally at minimum headway. Theoretical capacity represents the upper theoretical bound. Secondly, practical capacity is a practical limit of "representative" traffic volume that can be moved on a line at a reasonable level of reliability intended in terms of punctuality. Practical capacity represents a more realistic measure, being calculated under more realistic assumptions, which are related to the level of expected operating quality and system reliability. It is capacity that can be permanently provided under normal operating conditions and it may usually attain 60-75% of theoretical capacity. A railway signaling and control system is needed for safe control of transport processes in rail traffic, i.e. for safe and reliable movements of all trains considered in a capacity calculation.

[1] Corresponding author: Peter.Marton@fri.uniza.sk

### 1.2. Signaling and interlocking systems on a railway line

Signaling systems ensure safe control of transport processes on a railway line. Elements which are monitored by those systems can be explained as "movements on a track section". Interlocking fulfils the function of "information processing". Interlocking is the central function to ensure that trains move safely (in technical terms). To achieve this, interlocking obtains information about track occupation of rail vehicles. Then, using signals, this piece of information is evaluated and movements are authorized.

In a signaled territory, tracks of an open line (i.e. the section between stations) are equipped with signaling appliances for train movement. These appliances are called trackside signals. They indicate if a movement (train) can enter a track section behind the signal. So-called "main" signals indicate if a train may enter a track section of a line. A signal that authorizes train movement requires an approach aspect at braking distance. The approach aspect is necessary for safe braking when approaching a stop signal. In a territory where distance between signals does not exceed train braking distance much (this is a case of the automatic block), the approach aspect is usually provided by a signal in rear. This signaling is called multi-section signaling since the main signal provides information for, at least, two following block sections.

The automatic block can be found on railway lines where train separation by block distances is used; whereas, a track is divided into block sections. A "block section" is a section of an open line track in which only one train is allowed at any one time (an exception is a so-called moving on sight; however, it is not dealt with in our simulation model). Block sections are limited by signals, such as entrance and exit signals (block signals). In addition, signals at stations and junctions can serve as block signals for the adjacent open line sections. A block signal together with related evaluation units for block messages is called a "block station". Block stations are unstaffed in case of the automatic block. Each block station, in cooperation with the neighboring block stations, regulates the following movements. To avoid deadlock situations, only these block stations are actively involved in opposing protection where the order of trains can be changed. This can be done in e.g. stations, loops, junctions and crossovers. In the following part, they will be called "train sequence stations". Protection of opposing movements is only regulated between the neighboring train sequence stations, whereas intermediate block stations that only forward related messages are only informed about it, or are not involved in these processes at all. A section between two train sequence stations is called "a train sequence section".

## 2. General purpose simulation tools

The aim of designing a simulation model is to support decision-making. In this particular case, a simulation model could be used to decide how to split a track to several block sections, i.e. where the block signals will be located.

There are three possibilities when choosing a simulation tool for designing a simulation model of the automatic block on a railway line. Firstly, some general purpose simulation tools that have been developed for wide use can be employed. As examples, AnyLogic (Grigoryev 2012) and ARENA (Kelton 2010) can be mentioned. Secondly, it is possible to use specific simulation software that has been developed for railway transportation modelling purposes, e.g. Opentrack, Railsys and Villon (Adamko 2009). Railway-oriented simulation software is preferable for microscopic modelling of complex railway networks and stations. Creation of a simple single-track railway line could be too complicated because of different input data needed. As a result, using simulation tools of this category could be ineffective in the case of this study (simulation of the automatic block operation on a single-track railway line). Advantages of the first and the second possibility are reduced by programming requirements, natural framework for simulation modelling, conceptual guidance, automated gathering of statistics, graphic symbolism for communication and animation. Thirdly, simulation models using some generic programming language can be created – however, developing such a model may be time consuming and sometimes too complicated, as some programming knowledge is expected.

In this paper, the use of AnyLogic, as a general purpose simulation tool, is described. Simulation theory terms and their use in a railway line simulation model will be discussed in a separate part of the paper.

AnyLogic is a simulation tool that supports the most common simulation methods in place today, i.e. discrete-event, agent-based and system dynamics methods. Flexibility of the modelling language used by this tool enables a user to capture the complexity and heterogeneity of business, economic and social systems to any desired level of detail. It is comprised of a graphical interface, tools and library objects that allow modelling diverse areas, such as manufacturing, logistics, business processes, human resources, consumer and patient behavior. It uses object-oriented model design, which allows developing modular, hierarchical and large models, which is very advantageous. AnyLogic is based on the Java programming language. Modelling is realized via model components ("objects") that are grouped into several libraries according to their function.

## 3. Simulation model

### 3.1 Model input data

In order to reach relevant outputs and results, it is important to use accurate and reliable data. To build a simulation model of railway line operation, following data types are needed:

- Railway line topology, e.g. technical data about a railway line – length, position of trackside signals, speed limit, etc.
- Operational logic and timetable related data – rules such as “one train in one block section”, running times, headways, departure times, etc.

Specific input data are needed depending on the formulation of a solved problem and required output data. Some detailed technical information (topology) about e.g. slopes, radius curvature, etc. are needed when modelling a continuous dynamic train movement. However, this is not necessary in case of models that need only running times taken from the real timetable or estimated running times based on railway line length and train average speed.

In case of this paper, firstly, a simulation model of operation on a railway line without a signaling and interlocking system equipped with an automatic block was developed. The modeled railway line is not equipped with any signaling and interlocking system at present. Secondly, a simulation model with one track and several trackside signals in one direction was developed. This was made as an intermediate step to develop the final version of the simulation model. This simulation model can be used for a double track line simulation, too. Thirdly, a simulation model with trackside signals for both directions on a single-track line was created. The simulation model was developed in order to interpret a simple case of an automatic block interlocking system, whose working principle is very important input.

### 3.2 Modelling the system

It is necessary to specify what the basic model elements are. It can be accounted for system entities (customers and resources), input variables, performance measures and functional relationships. In these simulation models, a railway line and trains are taken as system entities, arrival rate (arrival time) and running time (an analogy to service rate) are accounted as input variables, mean wait time by a specific block signal is the example of a performance measure, and average time spent on a train sequence section (an analogy to “time in system” = wait times + running times) and signal status (“clear”, “stop” or “caution”) are examples of functional relationships.

A simulation model of operation on a railway line without a signaling and interlocking system is described as following. The line should be occupied only by one train. The model layout is very simple (Fig.1.). It features one track without trackside signals. Following basic model objects are used:

- railSettings: This object offers a lower-level interface for the railway model management based on Java functions and a callback technique.
- trainSource: This object starts any railway model process flow in the AnyLogic environment. It generates trains, places them on a track and injects train entities into a train process flowchart. A new train should contain at least one wagon. In Fig.1., the train has four wagons (in the middle of the track).
- trainMoveTo: This is one and only object that controls the train movement. The train can move forwards or backwards. The train can have a target position where it should get, or just move without a target. This object has no embedded control of number of trains on the track. Several trains can move at the same time in one block section.
- trainDispose: This object removes a train from the model.

The first model is used only to introduce the Anylogic simulation tool environment. Used objects connected by connection lines create a flowchart that describes entity (train) existence in the simulation model:

1. Train entity is created in the trainSource object.
2. Train entity is sent to the next object (trainMoveTo) to receive an instruction to move to the track end. Train entity is located in this object during the movement.
3. Train entity is disposed in the last object of the model flowchart.

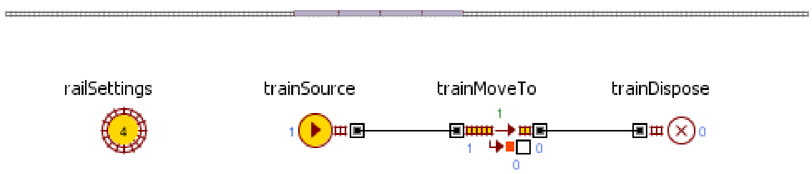



**Fig. 1**
*Simple model of railway track*
*Source: author*

The second model was built in order to develop a process flowchart for the automatic block in one direction (Fig.2.). A basic arrangement of necessary model objects was derived from the example of one train sequence section with four block sections. Each block section is bounded by two signals. This model contains only block signals. It does not contain the main signal for departures from any train sequence station, nor the main signal for arrivals to the train sequence station B. Two block sections bounded from both sides by block signals are showed in the model at full length. Two block sections bounded from one side by the main signal of the train sequence station are showed partially.
Other objects of AnyLogic Rail Library are used in the process flowchart to ensure that one entity (train) will occupy only one block section at the same time.
Following model objects are used to improve the first simulation model:

- queue: This object is used to enable creating of a queue (buffer) of entities waiting to be accepted by a next object in the process flow.
- hold: This block can hold the entity flow along a particular connection. It is used, for example, when the receiving object can accept entities, but the user does not temporarily want the entities to proceed there. In our model, we use it to hold the train on the beginning of a block section, when this one is occupied by another (previous) train.
- Function: AnyLogic enables the user to define its own functions. The function will return the value of an expression each time the user calls it from the model. In our model, function "navesti" (Signals) is called (activated, started) always by entering the hold blocks and by entering and leaving the moveTrainTo objects. This enables to update the colors showed on signals (color ovals).

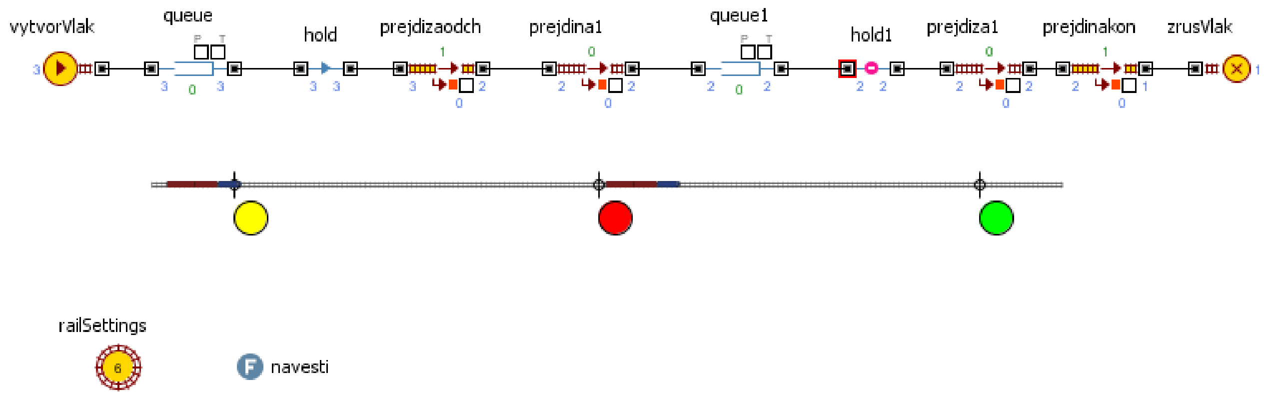



**Fig. 2**
*Simple model of automatic block with two block sections for one direction*
*Source: author*

In the second model, two trainMoveTo objects are used for each block section. The first trainMoveTo object is responsible for a train movement from the block section beginning to overlap (short distance behind the first block section signal). The second trainMoveTo object is used to move a train to the block section end (next signal).
This model corresponds to a common situation on Slovak railways. The automatic block is used on double-track lines. One track is used normally for train movements in one direction between two stations. Model blocks sequence in the process flowchart will be the same for models with more block sections on a track. The existence of entities in the second model could be described as follows:

1. On start-up, signal symbols are the places near block section borders. All signals are green. They permit the train movement.
2. A train entity is created by a trainSource object.
3. A train entity is sent from a queue object to a trainMoveTo object through a hold object in case that the hold object is unblocked. A train entity has to stay in the queue object in case that the following hold object is blocked.
4. A train entity continues to the first trainMoveTo object. This object controls the movement from the first block section until the overlap of the first block signal. Blocking of the hold object is started by a train entity entering this trainMoveTo object. Update of the colours showed by signals is activated too (using the function object).
5. A train entity continues to the second trainMoveTo object. This object controls the movement from the first block signal overlap to the second block signal.
6. A train entity is waiting in the following queue object in case the hold object is blocked. This will happen when a train entity occupies the following block section.
7. A train entity moves from a queue object through a hold object to a trainMoveTo object in case that the following block section is free (empty). First of the trainMoveTo objects controls the movement from the block signal to the signal overlap. The second trainMoveTo object controls the movement from the block signal overlap to the track end. Unblocking of the previous block section (first hold object) is activated after a

train entity enters the first trainMoveTo object. The block section behind the second block signal is blocked (the second hold object is blocked) in the same time. Update of the colors showed by signals is activated too (calling the function in function object).

8. A train entity is disposed in the last object of the model flowchart.

In the third model, the author tried to reflect a bidirectional operation on a single-track line controlled by the automatic block (Fig.3.). These model objects are used to improve the second simulation model:

- Variable: Variables are generally used to model some data units or object characteristics, changing over time. A variable has always some value assigned. A variable´s value can be changed during a simulation using the assignment operator. In the third model, two variables are used. The first variable keeps information about "direction permission". This variable is Boolean. "True" means a direction permission for train movements in one direction. "False" means a direction permission for train movements in opposite direction. The second variable informs about number of trains in the whole track. Only when this variable is equal to zero, a direction permission for train movements can be changed.
- Schedule: It is a special element that allows defining how some value changes in time, according to a defined cyclic pattern. In our model, a schedule object is used to define a timetable for train arrivals. Train arrivals are defined by a specified time value. Two schedule objects are used. Each for defining train arrivals from one train sequence station.

Some simplifications from the first and the second model were fixed in the third model. Here are the modifications, when comparing the second and the third simulation model:

- To show all the possible states on all three block signals, three hold objects are needed in the model process flowchart. The third block signal was not modelled fully in the second simulation model, as it was eliminated in the third simulation model.
- Three trainMoveTo objects are used for a train entity movement control in a block section.
- Two trainSource objects are used in the process flowchart in the third simulation model. Each is used for one train movement direction.
- Train arrivals controlled by both trainSource objects are determined by schedule objects.

Entities existence in the third model could be interpreted the following way:

1. On start-up, signal symbols are the places near block section borders. Block signals for a train movement in one direction are green and activated by calling the "inic" function defined by a function object. Block signals for the opposite direction are inactive (black).
2. A train entity is created by one of the two trainSource objects, based on the data defined in the corresponding schedule object.
3. A train entity enters a trainMoveTo object. It receives the movement instruction to the first block signal. The value of the variable for counting trains on the track is incremented.
4. A train entity is sent to the queue object and waits there when the following hold object is blocked, i.e. the block section behind the first block signal is occupied. In other case, train entity moves onwards to the following trainMoveTo object. This is the second trainMoveTo object that belongs to the first block signal. Its role is to control the train movement from the signal to its ovelap. The third trainMoveTo object controls the train movement to the next (in this case the second) block signal.
5. A train entity is sent again to another queue object, after leaving the output port of the third trainMoveTo object. This queue object and following hold object belong to the second block signal. They work the same way as the queue object and the hold object by the first block signal.
6. A train entity can approach a block section behind the third block signal when the condition for the movement is checked by the corresponding hold block. The train movement behind the third block signal is controlled by three following trainMoveTo objects.
7. A train entity is disposed in the last object of the model flowchart. The value of the variable for counting trains on the track is decremented.

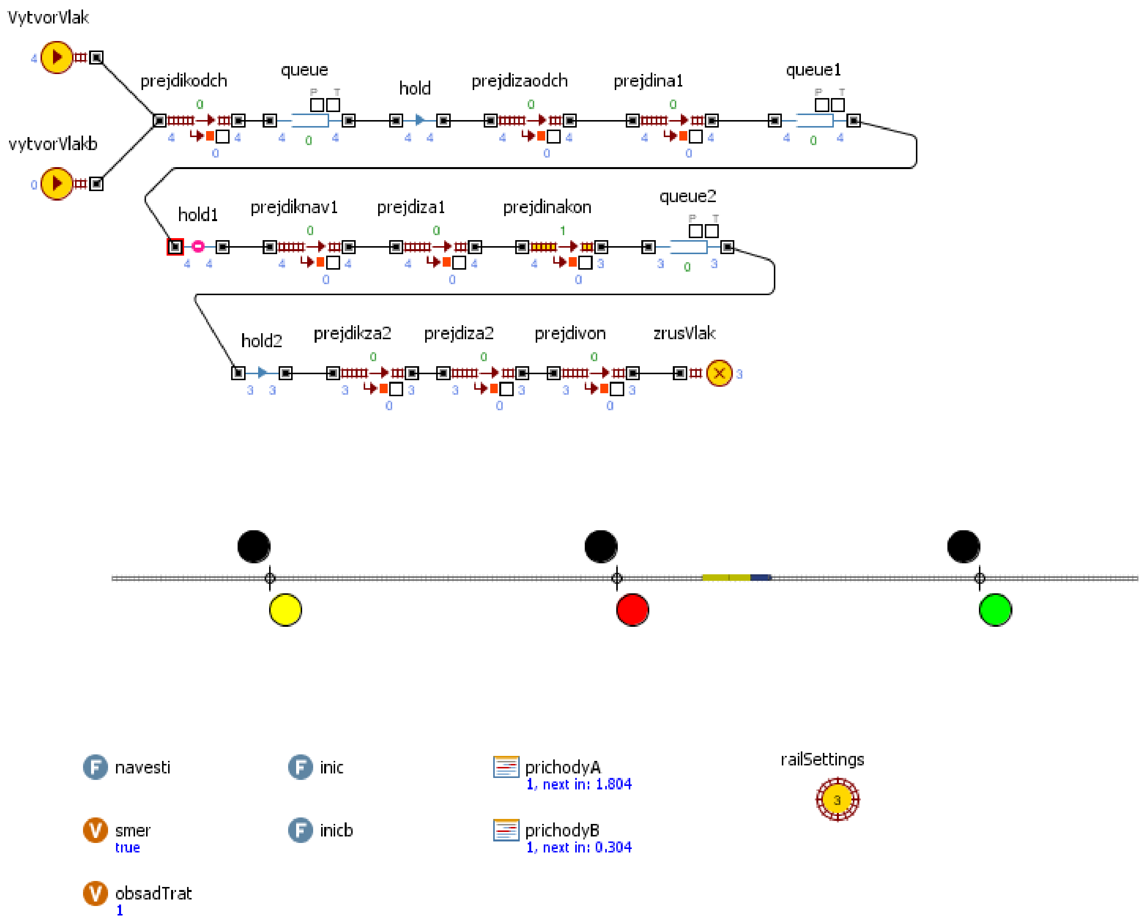



**Fig. 3**
*Simple model of bidirectional automatic block with two block sections*
*Source: author*

The third simulation model includes a direction permission handover because of a bidirectional operation on a single track. Some variable object is embedded in the model to avoid the direction permission handover in situation that the track is occupied ("obsadTrat"). The direction permission handover can be realized only when the whole track is free. This means that variable for counting trains on the track has to contain the value 0. Only when this condition is met, the value held in the variable object with direction permission information ("smer") is switched. The signal symbols in the animation are activated for the train movement in the opposite direction in the same moment.

### 3.3 Verification and validation

Verification involves checking a model to make sure that each component behaves in a manner one would expect. By verifying each section of a model, possible faults can be detected before their impact on other sections of the system. For our model a few areas were tested during verification:

- The flow of entities and the route they can take in the model flow chart.
- The parts of the model that ensure trains only proceed behind a block signal when the signal is green or yellow.

Validation involves making sure the model is an accurate representation of a real system, and as a result, that it fulfils objectives and aims of the simulation study. One step of validating the model was watching the model for a length of time and noting down how the system behaved. It was then compared to how the real automatic block behaves. Another step of validating was a model demonstration. The simulation was run and shown to a handful of people that were familiar with the automatic block principles from the railway operation. Feedback relating to the models behavior was positive.

Successful completion of verification and validation steps is important in order to receive results after a simulation is run. Result analysis enables to achieve better knowledge about a modelled system behavior, i.e. to identify the areas that need to be optimized.

## 4. Model output data and capacity research

General purpose simulation tools (as AnyLogic) offer many statistical outputs, for example, about average time spent by an entity in a system, duration of services in resources, queue length and time spent in queues and the use of resources. This type of evaluation is sufficient in non-complex and capacity-oriented research of the rail operation. The fact that people who work with different documents about the railway operation are not familiar with this type of output data is not a disadvantage.

For the model presented in this paper:

- The average time spent by an entity in the system is the running time of one train on the line.
- Duration of a service in resources could be the running time in one block section.
- Time spent in queues and queue length can be used for determining the number of trains waiting in front of the block signals (this may/might have happen/ned) because of two reasons. Either the following block section was occupied or because of the waiting time of those trains.

In AnyLogic, several objects can be used in order to save output data, namely a TimeMeasureStart object and a TimeMeasureEnd object. These objects compose a pair of objects measuring time entities (trains in our model) spend between them and remember the time when an entity goes through. Another object, a TextFile object, has the ability to write to the file. It can be used together with some model flow chart object, e.g. in a Queue object. A variable object can be used for saving output data as well. For example, it can be used to log how many trains passed the line. The objects mentioned above are not used in our model. (The objects described above are not used in the presented model). This can be perceived as the area of research in the future, as they are being elaborated already. All statistic data that are mentioned as examples for output data can be used to calculate a railway line capacity, using several techniques and methodologies. Basic input data for calculation is the amount of trains that passed the line in both directions. It is important that a train passes the line without waiting for any block signal.

## 5. Conclusion

Over the last two or three decades, simulation has found its place as the most popular operation research tool (Banks 2009). Main reasons for its popularity and its advantages when using it in the field of railway operation are:

- Ability to deal with complex of correspondingly complicated systems.
- Experiment is realized outside the examined rail line.
- It is possible to examine a projected or virtual rail line.
- Observation of impact of new equipment (interlocking system) is possible before it is installed on a rail line.
- Simulation time can be accelerated or decelerated. It is possible to receive results in a very short time. It is possible to reduce simulation speed in case that any complex processes need to be evaluated.
- Simulation experiments, comparing with the real ones (tested in real conditions) are realized on a low cost level. Experiments on the real system are often not feasible due to financial, ethic and other reasons.
- It is possible to see all the processes as often as needed, as simulations can be re-played.
- Simulation results can be used to support arguments by presentation of planned changes in a layout or operation rules in a rail line. Simulation results may serve as a very good background for managers within a decision-making process.
- Complete simulation model can be used as a training tool for railway company personnel.

However, some drawbacks regarding the use of simulations may exist. It is necessary to mention the following ones:

- No optimal solutions delivery, as this tool is only "laboratory". It is the user who has to decide about parameter settings (e.g. block section number, block signal positions), as well as about how to set the experiment.
- Iterative search of a sub-optimal solution. For example, a simulation model does not deliver the optimal count of block sections. It is necessary to increase or decrease their count step by step, each step meaning one experiment. The author of a simulation model can decide about the optimal count of block sections after the experiment is evaluated.

## Acknowledgement

This paper presents results of work supported by the Slovak Scientific Grant Agency under the project VEGA No. 1/0188/13.

**References**

Adamko, N.; Klima, V.; Márton, P. 2009. Designing railway terminals using simulation techniques, *International Journal of Civil Engineering*, Vol. 8, No. 1 (3-2010): 57-67.

Banks, J.; Carson, J.S. 2009. *Discrete-Event System Simulation (5$^{th}$ Edition).* Prentice Hall. USA. 640 p.

Gašparík, J.; Široký, J.; Pečený, L.; Halás, M. 2014a. Methodology for assessing the quality of rail connections on the network. *Communications : scientific letters of the University of Žilina*, Vol. 16, no. 2 (2014): 25-30.

Gašparík, J.; Zitrický, V. 2014b. Aspects of railway capacity and occupation time estimation. *Journal of civil engineering and architecture,* Vol. 8, No. 3 (2014): 322-331.

Grigoryev, I. 2009. *AnyLogic 6 in Three Days.* AnyLogic North America. USA. 162 p.

Hansen, I.A.; Pachl, J. 2008. *Railway timetable & traffic*. Eurailpress. EU. 228 p.

Jánošíková, Ľ.; Kavička, A.; Bažant, M. 2014. Optimal operation scheduling and platform track assignment in a passenger railway station. *Journal of rail and rapid transit : proceedings of the institution of mechanical engineers Part F,* Vol. 228, No. 3 (2014): 271-284.

Mašek, J; Kendra, M; Čamaj, J; Gašparík, J. 2013. The impact of liberalization on regional passenger transport in Slovakia. In *Proceedings of 21$^{st}$ international symposium EURO-Žel 2013,* 131-138.

Mašek, J.; Kendra, M.; Čamaj, J.; Belošević, I.; Milinković, S. 2014. The influence of transport services provided by private operator on the quality of regional railway transport of Slovakia. *Railway Transport and Logistics,* Vol. 9, No. 3 (2013): 23-29.

Kelton, W.D.; Sadowski, R.P.; Swets, N.B. 2010. *Simulation with Arena, 5/e.* McGraw-Hill. USA. 656 p.

Pelerin, J; Mircea T. 2009: *Rail Investments in Central and Eastern Europe: Towards a Competitive Railway System*. UNIFE. EU. 126 p.

Ricci S.; Kontaxi E. 2009. A technical handbook of railway capacity methods am models since 1950. In *Proceedings of the 17$^{th}$ international symposium EURO-Žel 2009,* 129-139.

Theeg, G.; Vlasenko, S. 2009. *Railway signalling & interlocking*. EU. 448 p.

# MULTI-OBJECTIVE CONTAINER TRANSPORT OPTIMIZATION ON INTERMODAL NETWORKS BASED ON MATHEMATICAL MODEL

**Radoslav RAJKOVIC[1], Nenad ZRNIC[2], Olja COKORILO[3], Snezana RAJKOVIC[4], Đorđe STAKIC[5]**

[1] *University of Belgrade, Faculty of Mechanical Engineering, Innovation center, Kraljice Marije 16, 11000 Belgrade, Serbia*

[2] *University of Belgrade, Faculty of Mechanical Engineering, Kraljice Marije 16, 11000 Belgrade, Serbia*

[3] *University of Belgrade, Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, 11000 Belgrade, Serbia*

[4] *Institute for Forestry, Kneza Viseslava 3, 11030 Belgrade, Serbia*

[5] *University of Belgrade, Faculty of Mathematics, Studentski trg 16, 11000 Belgrade, Serbia*

**Abstract:** International container shipping is one of the most dynamic economic sectors of the last few years. In terms of value, global seaborne container trade is believed to account for approximately 60 percent of all world seaborne trade, and was valued at around 5.6 trillion U.S. dollars in 2010. Reflecting the sharp growth in world trade, global container traffic was 7 times higher in 2011 than in 1990. The global container trade increased by 4.7% in 2013 and by the end of 2014 it should reach 684 million TEU. The development of container transport in recent years marks a constant intention for increased cargo flows. This paper analyzed multi-objective container transport optimization on intermodal transportation networks from Far East to Serbia through selected Mediterranean ports (Koper, Rijeka, Bar, Thessaloniki and Constanza) composed by two legs. First leg (ocean freight) observed the six world's largest container operators (Maersk Line, Mediterranean Shipping Company, CMA CGM, Evergreen Line, China Ocean Shipping Company and Hapag-Lioyd) with theirs different type of services. The second leg of the chain represents the inland component of the distribution, in which containers are routed from gateways to final destination, by road, rail or barge. The research used evolutionary multi-objective mathematical model. Our model observed multi-objective optimization, minimizing transportation cost, transit time and emissions. Inclusion of more objectives into the model we obtain accurate information to the observed objects at the same time. Cost and transit time are the two most common considerations in transport planning problems. Results provide an optimal route with shortest transit time, lowest transportation cost and minimum of emissions between Far East and Serbia and give us the possibility of observation the group of all possible solutions, which are ranked in ascending particular order.

**Keywords**: optimization, container transport, mathematical model.

## 1. Introduction

Containerization represents a revolution in the freight transport industry, facilitating both economies of scale and improvements in handling speed and throughput, with containerized traffic has surged since the 1990s. The surge in container trade is mainly attributed to the increasing penetration of Asian products in developed European and American markets. In fact, container trade flows to and from Asia present the highest growth among the three main East-West trade routes (Asia-Europe, Transpacific and Transatlantic) which make 47 percent of world maritime container trade flows. Specifically, during the period 1995-2011, container trade has increased by almost 5 times between Asia and Europe reaching about 20 million TEUs. Between Far East and Mediterranean 4.37 million TEUs was shipped using 31 different liner shipping services in 2012. (Review of Maritime Transport, 2013)

This paper analyzes the supply chain network with primary focus on import of containers from Far East (Port of Shanghai) to Serbia (city of Belgrade) through selected Mediterranean ports (Koper, Rijeka, Bar, Thessaloniki and Constanza), observing the six world's largest container operators (Maersk Line-MSK, Mediterranean Shipping Company-MSC, CMA CGM, Evergreen Line-EMC, China Ocean Shipping Company-COSCO and Hapag-Lloyd). Serbia is hinterland country and container import from Far East to Serbia needs to use different transport modes on inland to link shipping transport in the sea leg including railway, barge and truck. We considered only import containers in Serbia given the negligibly small exported quantities. (Rajkovic et al., 2014)

Transportation costs and transit time are the two most common considerations in transport planning problems. Some customers prefer lowest freight rates while some others would rather pay more for a faster delivery, but rising concerns about carbon dioxide ($CO_2$) emissions can therefore no longer be ignored when planning supply chains; on the one hand because companies have a moral obligation to operate in a sustainable fashion and on the other hand because customers are becoming more and more aware of the humongous affect that supply chain design has on $CO_2$ emissions.

In this study we developed multi-objective optimization model, which minimize transportation cost, transit time and $CO_2$ emissions of container import from Far East to Serbia. It was programmed in MATLAB and simulations were performed on an Intel Core i7-3612 QM 2.1 GHz computer. Through the minimization of time, costs and emissions for container import flow from Far East through Mediterranean nodes it was determined the most optimal route for container import to Serbia which reached a number of 42 000 TEU in 2012.

[1] Corresponding author: radoslav.rajkovic@gmail.com

The rest of this paper is organized as follows: Section 2 presents the literature review and describes the multi-objective problem which is considered in this work while the mathematical model is explained in Section 3. Section 4 reports and analyze the results of the mathematical model. Finally, Section 5 is devoted to conclusions and future developments.

**2. Background**

One of the most frequent fields of the observations in the container transport is optimization of container flows. Various multi-objective evolutionary algorithms have been developed to efficiently solve these optimization problems. Multi-objective optimization (also known as multi-objective programming, vector optimization, multicriteria optimization, multiattribute optimization or Pareto optimization) is an area of multiple criteria decision making, that is concerned with mathematical optimization problems involving more than one objective function to be optimized simultaneously. Multi-objective optimization has been applied in many fields of science, including engineering, economics and logistics where optimal decisions need to be taken in the presence of trade-offs between two or more conflicting objectives.

As there usually exist multiple Pareto optimal solutions for multi-objective optimization problems, what it means to solve such a problem is not as straightforward as it is for a conventional single-objective optimization problem. Therefore, different researchers have defined the term "solving a multi-objective optimization problem" in various ways. Many methods convert the original problem with multiple objectives into a single-objective optimization problem. This is called a scalarised problem. If scalarisation is done carefully, Pareto optimality of the solutions obtained can be guaranteed. Solving a multi-objective optimization problem is sometimes understood as approximating or computing all or a representative set of Pareto optimal solutions (Ehrgott, 2005; Carlos et al., 2007). When decision making is emphasized, the objective of solving a multi-objective optimization problem is referred to supporting a decision maker in finding the most preferred Pareto optimal solution according to his/her subjective preferences (Miettinen, 1999; Branke et al., 2008). The underlying assumption is that one solution to the problem must be identified to be implemented in practice. Here, a human decision maker (DM) plays an important role. The DM is expected to be an expert in the problem domain. The most preferred solution can be found using different philosophies. Multi-objective optimization methods can be divided into four classes (Ching-Lai, 1979). In so-called no preference methods, no DM is expected to be available, but a neutral compromise solution is identified without preference information. The other classes are so-called a priori, a posteriori and interactive methods and they all involve preference information from the DM in different ways. In a priori methods, preference information is first asked from the DM and then a solution best satisfying these preferences is found. In a posteriori methods, a representative set of Pareto optimal solutions is first found and then the DM must choose one of them. In interactive methods, the decision maker is allowed to iteratively search for the most preferred solution. In each iteration of the interactive method, the DM is shown Pareto optimal solution(s) and describes how the solution(s) could be improved. The information given by the decision maker is then taken into account while generating new Pareto optimal solution(s) for the DM to study in the next iteration. In this way, the DM learns about the feasibility of his/her wishes and can concentrate on solutions that are interesting to him/her. The DM may stop the search whenever he/she wants to.

Novikova et al., 2013, presented multi-objective optimization of container flow through Dostyk and Alashankou. The paper had aims to contribute to body of research in application of multi-objective evolutionary algorithms in the area of transport and logistics. It reviewed various optimization approaches applicable to railway border stations focusing on aspects like time-tabling, platforming, rolling stock circulation, train shunting, line planning and crew planning. Throughput and cost data for Dostyk and Alashankou were analyzed using IMEA using Matlab software.

Yang et al., 2011, developed an intermodal network optimization model to examine the competitiveness of 36 alternative routings for freight moving from China to and beyond Indian Ocean. The proposed model, which is built upon the principles of goal programming, is able to handle multiple and conflicting objective functions such as minimizing transportation cost, transit time and transit time variability while ensuring flow continuity and transit nodes compatibility among the rail, road, ocean vessel, airplane and inland waterway transports.

Winebrake et al., 2008, presented an energy and environmental network analysis model to explore trade-offs associated with freight transport. The geospatial model uses an intermodal network built by the authors to connect various modes (rail, road, water) via intermodal terminals. Routes along the network are characterized not only by temporal and distance attributes, but also by cost, energy, and emissions attributes (including emissions of carbon dioxide, particulate matter, sulfur oxides, volatile organic compounds, and oxides of nitrogen).

Tsung-Sheng, 2008, formulated the international intermodal routing as a multiobjective multimodal multicommodity flow problem (MMMFP) with time windows and concave costs. The objectives of his paper were to develop a mathematical model encompassing all three essential characteristics, and to propose an algorithm that can effectively provide answers to the model.

Hokey, 1991, developed a chance-constrained goal programming model to aid the distribution manager in choosing the most effective intermodal mix that not only minimises cost and risk, but also satisfies various on-time service reqirements.

## 3. Materials and methods

In this paper we consider an intermodal transportation chain which based on the import-way and composed by two legs. The first one represents maritime transfers from origin port-Shanghai port to gateway ports (Constanza, Thessaloniki, Bar, Rijeka, Koper). The second leg of the chain represents the inland component of the distribution, in which containers are routed from gateways to final destination-Belgrade, by road, rail and barge.

The network is composed by three categories of nodes: origin port (port of loading), gateway ports (ports of discharge) and destination (place of delivery), and two categories of links, maritime and inland.

### 2.1. Nodes

As one of the most important foreign trade partners in container imports from the Far East to Serbia is China with major port-port of Shanghai. It is the world's busiest trading port which handles a staggering 32 million containers a year, carrying 736 million tons of goods to far-flung places around the globe. The container traffic flow from 2004 to 2012 is presented in table 1.

**Table 1**
*Container traffic flow from 2004 to 2012*

| Rank | Port, Country | Volume (Million TEUs) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 2012 | 2011 | 2010 | 2009 | 2008 | 2007 | 2006 | 2005 | 2004 |
| **1** | Shanghai, China | 32,53 | 31,74 | 29,07 | 25,00 | 27,98 | 26,15 | 21,71 | 18,08 | 14,56 |
| **2** | SIngapore, Singapore | 31,65 | 29,94 | 28,43 | 25,87 | 29,92 | 27,93 | 24,79 | 23,19 | 21,33 |
| **3** | Hong Kong, China | 23,10 | 24,38 | 23,70 | 20,98 | 24,25 | 23,88 | 23,54 | 22,43 | 21,98 |
| **4** | Shenzhen, China | 22,94 | 22,57 | 22,51 | 18,25 | 21,41 | 21,10 | 18,47 | 16,20 | 13,62 |
| **5** | Busan, South Korea | 17,04 | 16,17 | 14,19 | 11,95 | 13,43 | 13,27 | 12,04 | 11,84 | 11,43 |
| **6** | Ningbo-Zhoushan, China | 16,83 | 14,72 | 13,14 | 10,50 | 11,23 | 9,35 | 7,07 | 5,21 | 4,01 |
| **7** | Guangzhou, China | 14,74 | 14,26 | 12,55 | 11,19 | 11,00 | 9,20 | 6,60 | 4,69 | 3,31 |
| **8** | Qingdao, China | 14,50 | 13,02 | 12,01 | 10,26 | 10,32 | 9,46 | 7,70 | 6,31 | 5,14 |
| **9** | Jebel Ali, Dubai, UAE | 13,30 | 13,01 | 11,60 | 11,12 | 11,83 | 10,65 | 8,92 | 7,62 | 6,43 |
| **10** | Tianjin, China | 12,30 | 11,59 | 10,08 | 8,70 | 8,50 | 7,10 | 5,95 | 4,80 | 3,81 |
| **11** | Rotterdam, Netherlands | 11,87 | 11,88 | 11,14 | 9,74 | 10,78 | 10,79 | 9,66 | 9,29 | 8,28 |
| **12** | Port Kelang, Malaysia | 10,00 | 9,60 | 8,87 | 7,31 | 7,97 | 7,12 | 6,33 | 5,54 | 5,24 |
| **13** | Kaohsiung, Taiwan | 9,78 | 9,64 | 9,18 | 8,58 | 9,68 | 10,26 | 9,78 | 9,47 | 9,71 |
| **14** | Hamburg, Germany | 8,86 | 9,04 | 7,91 | 7,01 | 9,74 | 9,89 | 8,86 | 8,09 | 7,00 |
| **15** | Antwerp, Belguim | 8,64 | 8,66 | 8,47 | 7,31 | 8,66 | 8,18 | 7,02 | 6,48 | 6,06 |

*Source: The JOC Top 50 World Container Ports*

Gateway ports are connected with origin port, but only by incoming links. From origin port it is possible to reach a gateway port but the opposite is not allowed since here we are addressing only incoming flows. The main gateways for container import to Serbia are Constanza, Thessaloniki, Bar, Rijeka and Koper port.
Destination-Serbia is hinterland country with capital city-Belgrade. This region represents the largest percentage of Serbian imports in general. It is connected to the gateway ports with a direct link, representing the shortest path to reach it from that gateway, by road, rail and barge.

### 2.2. Links

Maritime links are those between origin port and gateway ports. As for intercontinental links, there may be more than one link connecting an origin port to a gateway port, and each such link belongs to a different service with given travel time and frequency depends on different operators (MSK, MSC, CMA CGM, EMC, COSCO and Hapag-Lloyd). Top 10 container shipping companies are presented in figure 1.

| Rnk | Operator | TEU | Share |
|---|---|---|---|
| 1 | APM-Maersk | 2,828,485 | 15.2% |
| 2 | Mediterranean Shg Co | 2,520,499 | 13.5% |
| 3 | CMA CGM Group | 1,626,126 | 8.7% |
| 4 | Evergreen Line | 948,869 | 5.1% |
| 5 | COSCO Container L. | 819,202 | 4.4% |
| 6 | Hapag-Lloyd | 723,773 | 3.9% |
| 7 | CSCL | 631,750 | 3.4% |
| 8 | MOL | 600,806 | 3.2% |
| 9 | Hanjin Shipping | 598,325 | 3.2% |
| 10 | APL | 563,624 | 3.0% |

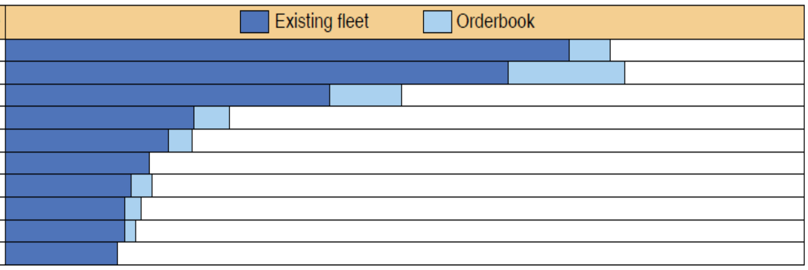



**Fig. 1.**
*Container shipping companies*
*Source: Alphaliner*

Inland links are those between gateway ports and place of delivery of containers-Belgrade. There are three available inland transportation modes which could be chosen including truck, railway and barge. Where there are available rail or barge linkages, line-haul may be done by rail or barge before last mile delivery by truck. Without such facilities, containers could also be transported from gateway ports to end-customers all the way by truck.

### 2.3. Mathematical model

Our mathematical model observed multi-objective optimization, minimizing transportation cost, transit time and $CO_2$ emission.

In this paper we analyzed transit time of container flow, from the moment of the ship departure, from the port of loading to the moment of the arrival container to the place of delivery in Belgrade, regarding the six different operators, five discharge ports and three different types of ocean services. The total time includes waiting time of containers at the port of discharge, depending on the different modes of transport from port of discharge to Belgrade.

The transport cost was considered for each of the most commonly used types of containers in the container transport, and is based on the free on board-FOB term. Total cost includes, except the cost of transport from Shanghai to Belgrade, local costs in the port of discharge, customs clearance and handling costs. During inland transport it was used different modes of transport, and because of more appropriate comparisons of the costs of rail and barge with truck, it was included also handling costs at the terminal in Belgrade and final delivery to users by truck (local delivery).

The total $CO_2$ emissions include ocean transport emissions from container ships and land emissions produced by truck, rail and barge. This paper excluded $CO_2$ emissions in ports of discharge due to their negligible share in total emissions.

Finally, our developed mathematical model provides us opportunity to observe the entire range of solutions and give us possibility of their ranking. Choosing different coefficients it was made simple selection of a group of optimal solutions, observing at the same time minimizing the transit time, transport costs and $CO_2$ emissions.

The model formulation and corresponding explanations are given as follows:

| Set | Description |
|---|---|
| N | set of nodes, let N = S ∪ E ∪ B, while S stands for origin port, E stands for gateway ports and B stands for place of delivery |
| A | set of arcs connecting an origin to a gateway (first-leg arcs) |
| C | set of arcs connecting gateways to place of delivery (second-leg arcs) |

| Decision Variable | Description |
|---|---|
| $t_{ij}$ | binary cost variable representing containers flow on first-leg arc, operator ″ i ″ to gateway ″ j ″, $t_{ij} \in \{0,1\}$ |
| $l_{jk}$ | binary cost variable representing containers flow on second-leg arc, gateway ″ j ″, mode of transport ″ k ″, $l_{jk} \in \{0,1\}$ |
| $tt_{ij}$ | binary time variable representing containers flow on first-leg arc, operator ″ i ″ to gateway, ″ j ″, $tt_{ij} \in \{0,1\}$ |
| $tl_{jk}$ | binary time variable representing containers flow on second-leg arc, gateway ″ j ″, mode of transport ″ k ″, $tl_{jk} \in \{0,1\}$ |
| $d_{ijs}$ | binary $CO_2$ emission variable representing container flow on first-leg arc, operator ″ i ″ to gateway ″ j ″ by service ″ s ″, $d_{ijs} \in \{0,1\}$ |
| $g_{jk}$ | binary $CO_2$ emission variable representing container flow on second-leg arc, gateway ″ j ″, mode of transport ″ k ″, $g_{jk} \in \{0,1\}$ |

| Parametar | Description |
|---|---|
| i | number of operator, $i \in \{1, \dots, 6\}$ |
| j | number of port, $j \in \{1, \dots, 5\}$ |
| s | type of service, $s \in \{1, \dots, 3\}$ |
| k | mode of transport, $k \in \{1, \dots, 3\}$ |
| $TSE_{ijs}$ | transit time on first-leg arcs (expressed in days) |
| $TEB_{jk}$ | transit time on second-leg arcs (expressed in days) |
| $CSE_{ij}$ | transportation cost on first-leg arcs (expressed in $) |
| EX | exchange rate (€ / $) |
| $CPC_{ij}$ | port cost (expressed in €) |
| $CEB_{jk}$ | transportation cost on second-leg arcs (expressed in €) |
| $CO_2\ SE_{ijs}$ | $CO_2$ emissions on first-leg arcs (expressed in kg/TEU) |
| $CO_2\ EB_{jk}$ | $CO_2$ emissions on second-leg arcs (expressed in kg/TEU) |

Objective functions:

Minimize Cost =

$$\sum_{(i,j)\in A} ((CSE_{\mathrm{ij}}\frac{1}{EX} + CPC_{ij}) * t_{ij}) + \sum_{(j,k)\in C} (CEB_{jk} * l_{kj}) \tag{1}$$

Minimize Time =

$$\sum_{(i,j)\in A} \sum_{s} (TSE_{ijs} * tt_{ijs}) + \sum_{(j,k)\in C} (TEB_{jk} * tl_{jk}) \tag{2}$$

Minimize $CO_2$ emission =

$$\sum_{(i,j)\in A} \sum_{s} (CO_2SE_{ijs} * d_{ijs}) + \sum_{(j,k)\in C} (CO_2EB_{jk} * g_{jk}) \tag{3}$$

Constraints:

$$\sum_{(i,j)\in A} t_{\mathrm{ij}} = 1 \tag{4}$$

$$\sum_{(j,k)\in C} l_{\mathrm{jk}} = 1 \tag{5}$$

$$\sum_{(\mathrm{i,j})\in \mathrm{A}} \mathrm{t_{ij}} = \sum_{(\mathrm{j,k})\in \mathrm{C}} \mathrm{l_{jk}},\ \forall \mathrm{j} \tag{6}$$

$$\sum_{(\mathrm{i,j})\in \mathrm{A}} \sum_{\mathrm{s}} \mathrm{tt_{ijs}} = 1 \tag{7}$$

$$\sum_{(j,k)\in C} tl_{\mathrm{jk}} = 1 \tag{8}$$

$$\sum_{\mathrm{i,s}} \mathrm{tt_{ijs}} = \sum_{\mathrm{k}} \mathrm{tl_{jk}},\ \forall \mathrm{j} \tag{9}$$

$$\sum_{(\mathrm{i,j})\in \mathrm{A}} \sum_{\mathrm{s}} \mathrm{d_{ijs}} = 1 \tag{10}$$

$$\sum_{(j,k)\in C} g_{\mathrm{jk}} = 1 \tag{11}$$

$$\sum_{\mathrm{i,s}} \mathrm{d_{ijs}} = \sum_{\mathrm{k}} \mathrm{g_{jk}},\ \forall \mathrm{j} \tag{12}$$

Corresponding Explanations:

The objective function (1) minimizes total cost of container import flow through the transport network. They include transport cost on the first leg-arc (ocean costs), port cost and transport cost on the second leg-arc (cost of using inland vehicles-truck, rail and barge). The objective function (2) minimizes total transit time of container import from Shanghai to Belgrade. The objective function (3) minimizes total $CO_2$ emissions of container import flow through the transport network. They include $CO_2$ emissions for container ships on the first leg-arc (maritime transport) and $CO_2$ emissions for inland vehicles on the second leg-arc. Constraints (4) and (5) define a single best solution for cost from a group of ordered pairs on the maritime and inland part. Constraint (6) selects the same port for the first and second leg-arc and defines one route from origin to place of delivery regarding transportation cost. Constraint (7) defines a single best solution for transit time on the ocean and depends of different type of service. Constraint (8) gives a single best solution for transit time on the second leg-arc. Constraint (9) selects the same port for the first and second leg-arc and represents one route from origin to place of delivery regarding transit time. Constraint (10) defines a single best solution for $CO_2$ emission on the ocean and depends of different type of service. Constraint (11) gives a single best solution for $CO_2$ emission on the second leg-arc. Constraint (12) selects the same port for the first and second leg-arc and represents one route from origin to place of delivery regarding $CO_2$ emission.

Calculating the "Carbon footprint" in intermodal transport:

$CO_2$ emissions = Distance x Emission Factor
g of $CO_2$ / TEU= km x [g of CO2 / (TEU x km)]

*Source: Greenhouse Gas Protocol 'Distance-based methodology' for calculating CO2 emissions (Maersk Line, 2012)*

## 4. Results

In this section we analyze the results obtained by our developed optimization model, which minimize the transportation cost, transit time and $CO_2$ emissions in intermodal transport. The model was programmed in MATLAB and simulations were performed on an Intel Core i7-3612 QM 2.1 GHz computer. We use original input data regarding period between july and december 2013. We concern multi-objective optimization and use different carbon footprint factors for different modes of transport presented in Table 2.

**Table 2**

| *Carbon footprint factors* | Ship | Rail | Barge | Truck |
|---|---|---|---|---|
| **$CO_2$ factor (kg/TEU-km)** | 0.084 | 0.205 | 0.084 | 0.072 |

*Source: Technical report (Maersk Line, 2012)*

The developed mathematical model provides us the opportunity to observe the entire range of possible solutions, and give us the possibility of their ranking, whose choice is made simple selection of the optimal solutions group, observing at the same time minimizing transportation costs, transit time and $CO_2$ emissions depending on the arbitrariness and significance of the different coefficient. Group of all possible solutions are ranked in ascending particular order and period of observation in the mathematical model does not have to be time-limited. The advantage of the model is that it can be applied to various "point to point" relations where the containers are transported by a combination of sea and land whose implementation of adequate input data can get reliable solutions of the observed problems.

Figure 2 illustrated the results of multi-objective optimization using equal value of scalars. The results show top 3 of possible 270 solutions ranking in ascending order. The optimal (rank:1) transportation cost, transit time and $CO_2$ emissions in intermodal transport between Shanghai and Belgrade are 1657 EUR/TEU, 31 days and 1515,56 kg/TEU, using see and land legs together. The operator COSCO using Far East Black Sea Express Service-ABX from Shanghai to Pireaus and Adriatic Feeder Service-AFS from Piraeus on the first-leg arc reach to the gateway port, Rijeka port, and continues on the second-leg arc with rail to the final destination Belgrade.

```
Command Window
>> modelrr15
Input: 1 - cost, 2 - time, 3 - emission, 4 - cost/time, 5 - cost/emission, 6 - time/emission, 7 - multi-objective: 7
Input scalar for cost: 1
Input scalar for time: 1
Input scalar for emission: 1
Input number for rank list (1-270): 3

* * *

Type of container: TEU, Rank: 1, Rate: 1657 EUR

Transport between: SHANGHAI - BELGRADE
Port of loading: SHANGHAI
Operator: COSCO
Number of transshipments: 1
Service: ABX / AFS
Route: SHANGHAI - Ningbo - Yantian - Hong Kong - Nansha - Singapore - Suez - PIREAUS / PIREAUS - Thessaloniki - Pireaus - RIJEKA
Port of discharge: RIJEKA
Mode of transort: RAIL
Place of delivery: BELGRADE

Transit time = 31 days
Distance (see + land): 16417 km + 666 km
Emission = 1515.56 kg/TEU
* * *

Type of container: TEU, Rank: 2, Rate: 1657 EUR

Transport between: SHANGHAI - BELGRADE
Port of loading: SHANGHAI
Operator: COSCO
Number of transshipments: 1
Service: MD1 / AFS
Route: SHANGHAI - Taipei - Hong Kong - Shekou - Suez - PIREAUS / PIREAUS - Thessaloniki - Pireaus - RIJEKA
Port of discharge: RIJEKA
Mode of transort: RAIL
Place of delivery: BELGRADE

Transit time = 33 days
Distance (see + land): 16406 km + 666 km
Emission = 1514.63 kg/TEU
* * *

Type of container: TEU, Rank: 3, Rate: 1772 EUR

Transport between: SHANGHAI - BELGRADE
Port of loading: SHANGHAI
Operator: COSCO
Number of transshipments: 1
Service: ABX / AFS
Route: SHANGHAI - Ningbo - Yantian - Hong Kong - Nansha - Singapore - Suez - PIREAUS / PIREAUS - Thessaloniki - Pireaus - RIJEKA
Port of discharge: RIJEKA
Mode of transort: TRUCK
Place of delivery: BELGRADE

Transit time = 29 days
Distance (see + land): 16417 km + 550 km
Emission = 1638.63 kg/TEU
```

**Fig. 2.**
*Multi-objective optimisation*
*Source: MATLAB Mathematical model by autor*

Mathematical model also shows a range of solutions for each of criteria separately choosing number 1, 2 and 3 from figure above (single objective optimization) and all of their possible combinations (bi-objective optimization) choosing number 4, 5 and 6. Similar research in terms of testing the bi-objective optimization minimizing cost and transit time conducted Lam and Gu, 2013, observing import and export container flow to and from inland China. The results and analysis offer managerial insights of the impact of trade-offs between cost and transit time, and the effect of different carbon footprint requirements on transport planning. Francesetti, 2005, analysed the costs of shipping containers from four Chinese ports to representative central European destinations. It is demonstrated that the sum of costs by sea and costs over land, using both truck and rail transport, clearly favours the Italian ports, above all those of Genoa and Trieste for a geographic range that does not include all the Northern countries of the European Union and Russia but does cover a considerable portion of the southernmost cities of these countries such as Milan, Munich, Vienna, Budapest, Bern, Lyon and Kiev. The purpose of his paper was to define costs in each sector (shipping costs, port costs and inland distribution costs) and to compare the relative port positions. Han et al., 2011, considered the problem of determining transportation quantity and mode in transporting international cargoes between Myanmar and her trading countries, especially focusing on the countries in South East Asia to check the extent of using short sea shipping, and inland transportation. The objective of theirs paper is to minimize transportation costs by mode between cargo origin and destination, subject to the maximum cargo volumes being handled at each seaport, in order to optimize the short sea shipping and inland transportation in Myanmar.

## 5. Conclusion and future developments

Optimization deals with the study of those kinds of problems in which one has to optimize one or more objectives that are functions of some real or integer variables. Scrupulous observation of multi-objective optimization and analysis of applied algorithms, we came to the conclusion that the development of appropriate mathematical models can be an accurate way to answer the problem making simultaneous and significant multi-criteria decision.

This paper analyzes the supply chain network with primary focus on import of containers from Shanghai to Belgrade through Mediterranean ports, observing the six world's largest container operators with theirs different type of ocean services and three different inland transport modes.

The main goal of this research is to provide an optimal route with shortest transit time, lowest transportation cost and minimal $CO_2$ emission of container import from Shanghai to Belgrade.

We propose a mathematical model, which observe the entire range of possible solutions, and give us the possibility of their ranking, whose choice is made simple selection of the optimal solutions group, observing at the same time minimizing transportation costs, transit time and $CO_2$ emissions depending on the arbitrariness and significance of the different scalars (coefficients).

The advantage of the model is that it can be applied to various “point to point” relations where the containers are transported by a combination of sea and land whose implementation of adequate input data can get reliable solutions of the observed problems. Findings of this paper show that optimization using immunized multi-objective evolutionary algorithm is a viable approach that can help managers who made a policy of the company, in order to improve their businesses following the constant changes in the market and making reliable comparisons.

Further research are needed and they can be extended in the future, importing with a lot of new initial and final nodes and comparing with some of adopted software solutions.

## Acknowledgements

The study was carried out within the Project TP - 36027, "Software development and national database for strategic management of the development of means of transport and infrastructure in road, rail, air and water transport by European transport network models", financed by the Ministry of Education Science and Technological Development, Republic of Serbia.

**References**

Review of Maritime Transport: UNCTAD/RMT/2013. United National Publication Sales no. E.13.II.D.9, ISBN 978-92-1-112872-7, e-ISBN 978-92-1-054195-4, ISSN 0566-7682. Available from Internet: <http://www. vi.unctad.org/digital-library/?task=dl_doc&doc_name...reviewma

Rajkovic, R, Zrnic, N, Stakic, Dj, Mahnic, B. 2014. The Costs of Container Transport Flow Between Far East and Serbia Using Different Liner Shipping Services, In *Proceeding of the 11th International Conference on Logistics and Sustainable Transport*, 98-102.

Rajkovic, R, Zrnic, N, Stakic, Dj. 2014. Application of Mathematical Model for Container Transport Flow of Goods: from Far East to Serbia, In *Proceeding of the International Conference on Industrial Logistics*, 159 - 166.

Ehrgott, M. 2005. Multicriteria Optimization. *Birkhäuser*, ISBN: 978-3-540-21398-7 (Print) 978-3-540-27659-3 (Online).Second edition. 290 p.

Carlos, A.; Coello, C.; Lamont, G. ; Veldhuisen, D. A. 2007. Evolutionary Algorithms for Solving Multi-Objective Problems. *Springer*. ISBN: 978-0-387-33254-3 (Print) 978-0-387-36797-2 (Online). Second edition: 800 p.

Miettinen, K. 1999. Nonlinear Multiobjective Optimization. *Springer*. 299 p.

Branke, J.; Deb, K.; Miettinen, K.;Slowinski, R. 2008. Multiobjective Optimization: Interactive and Evolutionary Approaches. *Springer*. 66 p.

Ching-Lai, H.; Abu Syed, M. 1979. Multiple objective decision making, methods and applications: a state-of-the-art survey. *Springer-Verlag*, 1-351.

Novikova, K. K. A.; Jing, L.; Kennedy, R. O. 2013. A Multi-objective Optimization of Sino-Kazakh Container Flow through Dostyk-Alashankou, *International Journal of Business, Humanities and Technology*, 3 (7): 37-47.

Yang, X.; Low, J. M. W.; Tang, L. C. 2011. Analysis of Intermodal Freight from China to Indian Ocean: A Goal Programming Approach, *Journal of Transport Geography*, 19 (4): 515-527.

Winebrake, J. J.; Corbett, J. J.; Falzarano, A.; Hawker, J. S.; Korfmacher, K.; Ketha, S.; Zilora, S. 2008. Assessing Energy, Environmental, and Economic Tradeoffs in Intermodal Freight Transportation, *Journal of the Air & Waste Management Association*, 58 (8): 1004-1013.

Tsung-Sheng, C. 2008. Best routes selection in international intermodal networks, *Computers & Operations Research*, 35 (9): 2877-2891.

Min, H. 1991. International Intermodal Choices Via Chance-Constrained Goal Programming, *Transportation Research Part A: Policy and Practice*, 25A (6): 351-362.

Lam, J. S. L; Gu, Y. 2013. Port Hinterland Intermodal Container Flow Optimisation with Green Concerns: A Literature Review and Research Agenda, In *Proceedings of International forum on Shipping, Ports and Airports: Trade, supply chain activities and transports: Contemporary logistics and Maritime issues*, 167-179.

Francesetti, D. C. 2005. Italian versus Northern Range port competitiveness: a transportation cost analysis in Chinese trade, *European Transport \ Trasporti Europei*, 30: 37-53.

Han, M. M.; Guolong L.; Bin, Y. 2011. A Linear Programming Model For Short Sea Shipping And Multimodal Inland Transportation In Myanmar, *Report and Opinion*, 3 (1): 37-43.

# A WEB BASED PLATFROM WHICH HELP INDIVIDUALS AND COMPANIES MOVE COMMODITIES WITH THE MOST ENVIROMENTAL FRIENDLY WAY, MINIMIZING EMISSIONS AND TRANSPORTATION COST

**Georgios K.D. Saharidis[1], George Kolomvos[2], George Liberopoulos[3]**
[1] *University of Thessaly, Department of Mechanical Engineering, Greece*
[2] *Kathikas Institute of Research & Technology, Cyprus*
[3] *University of Thessaly, Department of Mechanical Engineering, Greece*

**Abstract:** The objective of the proposed research is to develop a Decision Support System (DSS) for a web based platform which will help individuals and companies move commodities in the most environmental friendly way, minimizing environmental externalities (e.g. CO2 emissions) and transportation costs. The developed platform which is the final outcome of an FP7 European research project, referred to as GreenRoute project, uses existing information systems (e.g. geographical, weather, real time traffic information systems) and emission calculation models as a basis to apply two main scientific outcomes. The first scientific outcome is the development of a function that assigns a score to each arc of a transportation network referred to as the arc environmental externalities score (EESarc). The second scientific outcome is the development of a novel approach for solving the general travelling salesman problem (TSP) whose objective is to find the most environmental friendly route. In the frame of this research work, we consider the TSP in a connected graph driven by a novel cost function related to environmental factors. The cost function describes with a novel way the environmental impact of the feasible routes. This score approximates the impact of environmental externalities over the specific route. EESarc is uncorrelated to other fuel consumption factors such as the type of vehicle, the weight load and the operating conditions of the vehicle, and depends solely on the traffic conditions; the infrastructure profile; the weather conditions; and the length of each arc. We seek a path passing from all intermediate user-defined points minimising the sum of the EESarc for each arc employed in the solution. In order to accelerate solution times in the afore-mentioned context, we seek to identify novel modelling and solution approaches. We consider the Euclidean asymmetric TSP as a base model to work on. We keep the integrality constraints on the binary variables of the original problem intact and introduce additional integer variables for each node which provide the number of outgoing and ingoing arcs of two given disjoint sets. We relax the subtour elimination constraints and enforce valid separation inequalities for the start-up solution. These separation inequalities are generic and do not depend on the specific problem considered. Their number is equal to the number of nodes in the network, resulting to a few but dense rows in the simplex tableau. At each iteration of the algorithm, we inspect the solution yielded by the relaxed problem and identify subtours. For every subtour, separation cuts are added to the relaxed problem. The number of cuts is low but their density is high. We observe the solution times on a TSP testbed which includes randomly generated instances as well as problems coming from the well-known TSPLIB. We compare our results against several formulations. We observe that especially for medium to large networks of some hundreds of nodes, our approach is dominant in terms of solution time over every other one we have considered.

**Keywords**: Traveling salesman problem, green, environmental, optimization, mathematical programming.

## 1. Introduction

In this paper the application is motivated by the environmental extension of the TSP, namely, to find the environmentally friendliest tour in a directed graph, the arcs of which are weighted based on their impact on fuel consumption. At the modelling side, the vast majority of works in the TSP is focused on minimizing the total distance of the tour. Although one may claim that fuel consumption is linked to the distance travelled, this may only be true should one assume that all routes are under identical conditions e.g. same quality of tarmac, same grade, identical wind speed and direction etc. We may encounter dozens of possible daily conditions where this assumption fails dramatically. It suffices to follow a route with large alteration in grade or strong side winds and the fuel consumption may increase at such level that a longer route with environmentally friendlier conditions would have been far more economical. The TSP may be amenable to the vehicle routing problem, consequently, the results of this paper may also be used to optimize the fleet management of light- or heavy-duty vehicles.

At the solution side, the TSP literature is vast. For a review of approaches to solve the TSP, the reader is referred to the comprehensive work of Laporte [8]. A more recent review with developments and an updated set of modern areas of applications is included in [2] and [12]. The description of the TSP polytope is not yet known and its complexity relies on the subtour elimination constraints. The formulation proposed by Dantzig, Fulekrson and Johnson in [4] provides a tight description of the polytope, but the number of constraints grows exponentially with the size of the problem, rendering the formulation impractical from the computational point of view. The idea that was firstly explored in [4] was to solve the LP relaxation and subsequently add cuts which are violated by the integer solutions and not by their continuous counterparts. This was the first cutting plane algorithm proposed in the TSP literature and was made popular thereof, since it showed to perform fairly well on the case of the 49-node problem addressed by the authors. Current algorithms [1] are able to attack problems of thousands of nodes, but still parallel computing is required to obtain an optimal solution within reasonable time.

Cutting planes algorithms can be divided into generic and structured cuts algorithms by the way one follows to determine which cuts they should append at each iteration. In the former case, one can base on algebraic arguments to generate cuts such as for instance Gomory cuts, lift-and-project or mixed-integer rounding cuts [3]. More emphasis in the literature is given on structured cuts, where the underlying structure of the specific problem is exploited to generate valid inequalities at each iteration. Consequently, one obtains 2-matching inequalities [5] and comb inequalities [6]. It is not surprising that these cuts are coming or may also be applied to other problems such as knapsack and vertex packing since the TSP is linked to these problems. These are commonly called in the literature as “brunch and cut” approaches, since one solves the relaxed version of the TSP where the integer constraints are dropped; apply valid inequalities; and continues with the branch and bound method. The work of Padberg & Rinaldi in [10] is considered as a milestone in the successful application of branch and cut strategies in the TSP. Branch and cut combine cutting planes with the well-known branch and bound algorithms according to which a smart enumeration of all different combinations is performed driven by the solution to the LP relaxation.

Another perspective of viewing the TSP is as a special case of a minimum 1-spanning tree. This analogy was nicely explored by Held & Karp in [15]. The idea is to carefully create an objective function such that the result of the spanning tree which is a lower bound of the TSP closely approximates the TSP. The formulation of the minimization of 1-spanning trees by default excludes subtours, so there is no reason to enforce any subtour elimination constraints. On the other hand, in a minimum spanning tree there may be nodes with a degree greater to two, that is for instance, a node with two descendants nodes, which is prohibited in the TSP.

Christofides algorithm [14] is based on the 1-spanning view perspective and provides a tight lower bound on the original TSP. The minimum spanning tree and the perfect matching problem are the two main operations performed, based on the triangular property or Euclidean graphs. In the literature the vast majority of applications and theory is devoted to the symmetric Euclidean TSP. In this paper we are interested in the asymmetric non-Euclidean TSPs.
Heuristics algorithms are largely employed to efficiently solve the TSP of larger instances. Heuristics are divided to construction heuristics and improvement heuristics. The former aim at constructing a tour from scratch usually including Euclidean arguments (such as the greedy algorithm for the nearest or k-th nearest neighbor, see for instance). The latter include methods that are given an initial tour and are trying to improve. Such algorithms are the 2-OPT algorithm that we apply in this paper, 3-OPT and k-OPT (known as LKH) proposed by Lin & Kernighanin [9] and successfully applied by Helsgaun in [7] to provide what is known today as the best generic algorithm for solving the TSP.

## 2. Environmental Externalities Score

Let us consider two different routing options for a vehicle to travel from a point A to a point B. The need to introduce the concept of the environmental externalities score arises when one wishes to compare these two routes that will be travelled by the same vehicle in terms of their environmental impact. In both cases, we may assume that the driving attitude will not change and of course that the vehicle characteristics will remain identical.

In this work, we introduce the concept of the environmental externalities score, henceforth EES, which is a measure expressing the percentage of increase or decrease of the underlying environmental externalities compared to the nominal conditions. Each arc would have an individual EES based on the arc’s characteristics. By multiplying the EES with the values provided by any emission calculation model, we may translate the result into fuel consumption in liters per kilometer.

In this paper we will not enter into the details of how the EES is calculated. The reader is referred to [13] for a thorough study on how this measure was conceived, devised and developed. We define the instantaneous environmental externalities score function EES related to fuel consumption to be the ratio of instantaneous fuel consumption to fuel consumption at nominal conditions. The above is expressed through the following formula: $EES = \frac{FC}{\overline{FC}}$ where FC stands for fuel consumption and $\overline{FC}$ for fuel consumption at nominal conditions. The consumption at nominal conditions is provided by an emission calculation model.

## 3. Travelling Salesman Problem

Let us now turn our attention towards the solution of the resulting environmental TSP. Let G=(V,A) be a graph where V is a set of N vertices and A is a set of arcs. Let C be a cost matrix associated with A. V is the set of vertices such that V={0,1,2,…,N} and $i,j \in V$. We call the first vertex i=0. Arcs connect vertices such that edge ij connects the vertices i and j. We denote by $x_{ij} \in \{0,1\}$ the binary variable which takes the value of 1 if the arc connecting i and j is included in the Hamiltonian tour and 0 if not. x is the vector containing the values $x_{ij}$. Let $c_{ij}$ be the vector of costs associated to the edge ij. In fact $c_{ij}$ is fed into the objective function using the values EES for every arc mentioned in the previous section. The formulation of the TSP without the subtour elimination constraints (SECs) is equivalent to the assignment problem (AP) and is presented right below.

$$(Assignment\ Problem: AP)$$
$$\min_{x_{ij}} c_{ij} x_{ij}$$

*subject to:*

$$\sum_{i \in V \setminus \{j\}} x_{ij} = 1, \forall j \in V$$
$$\sum_{j \in V \setminus \{i\}} x_{ij} = 1, \forall i \in V$$
$$x_{ij} \in \{0,1\}$$

The SECs can be represented in various ways following the formulations. In this paper we follow a cutting place approach, where the subtour elimination constraints are initially dropped. The resulting problem *$(AP')$* is a relaxation of the original TSP in the sense that every feasible solution of the TSP is a feasible solution of *$(AP')$*, while the optimal value of *$(AP')$* provides a lower bound on the optimal value of the original TSP. Note that we call it *$(AP')$*, because it is the assignment problem *(AP)* augmented with the cutting planes appended at each iteration.

At each iteration of our algorithm, we solve *$(AP')$* and inspect the resulting solution $x_{ij}$. Note that we solve the *$(AP')$* using the branch and bound method. If the resulting solution satisfies the SECs then we can safely claim that we have obtained the optimal path. If it does contain at least two subtours, then we have obtained a lower bound. We identify subtours by simple inspection. That is, we inspect the solution vector $x_{ij}$ by starting from an arbitrary vertex (e.g. $i = 0$) keeping track of the path. If we reach the starting vertex without having visited all vertices then we have encountered a subtour. We enter the respective SEC and we proceed to identifying the next subtour.

We construct a feasible tour using the subtours identified earlier (we will explain later how we do this) and we apply a 2-OPT heuristic aiming to improve the incumbent solution. The role of the resulting 2-OPT feasible solution has a dual role: (a) to provide an upper bound at the current iteration and (b) to pass this information of the upper bound to the branch and bound method and accelerate branching by fathoming more nodes in the search tree. Moreover, using the subtours identified we generate a valid cut – explained below – and proceed to the next iteration.

Once all subtours identified and the respective valid inequalities entered, then we solve again *$(AP')$* to optimality and repeat the procedure. The algorithm is presented in Table 1 as a pseudocode.

**Table 1**
*Algorithm in pseudocode*

| | |
|---|---|
| Step 0: | (**Initialization**) Create the set $S = \emptyset$ that will contain the subtours identified. Set the iterator $k = 0$. Set lower bound $LB = \infty$ and upper bound $UB = -\infty$. |
| Step 1: | (**Solution of** $(\boldsymbol{AP'})$**)** Solve $(AP')$ to obtain the incumbent solution $x_{ij}^k$ at iteration $k$. Store the value of the objective function at *LB*. |
| Step 2: | (**Subtour inspection**) Let $S_l \subset S$ be the subtour $l$ which contains the vertices belonging to this subtour. Inspect the solution vector $x$ to identify subtours $S_l$ $l = 1,2$ ... until all vertices are inspected. Subtour inspection is performed starting from an arbitrary vertex and following connections. If no subtour is identified, i.e. $S = \emptyset$, then the incumbent solution $x_{ij}^k$ is optimal representing a Hamiltonian tour and the algorithm ends here. Else continue to Step 3. |
| Step 3 | (**Valid inequality**) For every subtour $S_l$, append the following cut at $(AP')$: $\sum_{i \in S_l, j \notin S_l} x_{ij} = 2y_l \ \forall S_l$ and $y_l \in \mathbb{N}^+$ *(1)* |
| Step 4: | (**Tour construction**) Consider the current subtour identified at Step 2. If this is the first subtour identified, then go to Step 2. Else, find the arc *ij* of the current tour with specific properties (see notes below) and place the identified subtour between $i$ and $j$. |
| Step 5: | (**2-OPT**) Perform a 2-OPT move. If a 2-OPT move improves the *UB*, then (a) update the $UB$ with the value of the new feasible tour and (b) set the $UB$ as cut-off value for the branch and bound method and return to Step 1. |

Let us explain a bit Steps 3, 4 and 5. In Step 3, the valid inequality (1) imposes an even number of connections between the nodes of the subtour and the remaining nodes not belonging to the subtour. Unless the incumbent solution of Step is optimal, it will include subtours. When a subtour is identified we instruct the algorithm to connect this subtour with the remaining vertices by an even number of arcs. For this reason we introduce the variable $y_l \in N$ for every identified subtour $S_l$, which denotes the number of arcs connecting the subtour with the remaining vertices. Because this number needs to be even, we actually use variable $2y_l$ and enforce constraint (1) for each subtour $S_l$ identified.

The performance of the 2-OPT heuristic in Step 4 depends on the quality of the initial tour on which improvements will be tested. The idea exploited here is to feed the 2-OPT with as many candidate initial tours as possible and let it find the best improved tour among those. We have thus implemented several strategies, such as: (a) random positions of the current subtour in the current tour; (b) placing current subtour at the end of the current tour, (c) nearest neighbor algorithm; (d) the strategy that worked best in terms of results is the following: Every time we identify a subtour we place it between the two vertices inducing the most "expensive" pass of the tour. To illustrate this and for ease of explanation, let us represent the current tour as a string with starting point the node *0* and last node say *n* (in the tour, the last node *n* is also the preceding node of *0*). Iterate among all possible pairs of nodes *i* and $j \neq i$. Figure 1 illustrates this operation. All resulting tours are sent to the 2-OPT algorithm and an improved tour, if possible is returned. We keep the tour with the lowest upper bound.

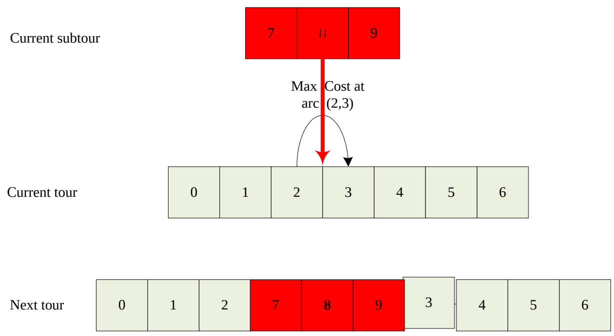



**Fig. 1.**
*Tour construction*

In Step 5, we perform a 2-OPT move. We adopt the representation of the string to illustrate this move. Name the node following $i$ as $i'$ and the node following $j$ as $j'$. Select the pair *(i,j)* for which $c_{ii'} + c_{jj'} + c_{n0} > c_{ij'} + c_{ni'} + c_{j0}$ and perform a 2-OPT move by placing the string which is contained between the nodes $i'$ and $j$, right after node $n$.

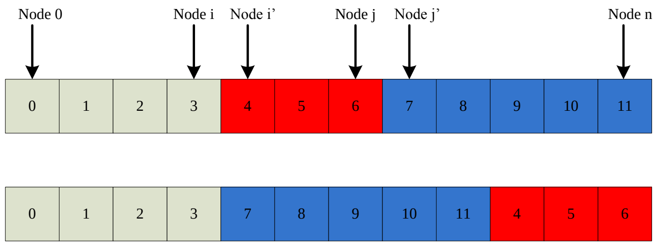



**Fig. 2.**
*2-OPT move*

### 4. Computational results

We report performance of our algorithm by testing it on the well-known TSP library. The TSP-LIB [11] is a library created by several TSP instances aimed to provide researchers with a broad set of test problems from various sources and with various properties. We have attacked problems included in the asymmetric test cases.
We performed the experiments on a dual-core 2.2GHz processor with 3GB of usable memory. The code was written on C++, the modelling of the problem was done using the IBM ILOG Concert Technology and the solution was provided by the IBM ILOG CPLEX 12.4 suite.

Table 2 presents the results of our runs. The column "Name" stores the name of the problem where the digits represent the number of nodes. The column "#iterations" stores the number of iteration until the end of the algorithm. The column "CPU time" stores the time required to reach the solution. We have divided the results in two pairs of columns: "Optimal" means that the algorithm run until the optimal solution was provided; "Gap<1%" means that the algorithm stopped once a feasible tour was obtained with a value close to optimum by at most 1%.

We have included this number, because in the frame of the real-world application on which this algorithm will be applied, a 1% solution is considered as a satisfactory approximation.

**Table 1**
*Results on asymmetric instances of the TSP-LIB*

| | Optimal | | Gap<1% | |
|---|---|---|---|---|
| **Name** | **#iterations** | **CPU time** | **#iterations** | **CPU time** |
| ftv64 | 12 | 5.679sec | 9 | 4.15sec |
| ftv33 | 5 | 1.7sec | 5 | 1.7sec |
| ftv35 | 11 | 2.543sec | 10 | 2.403sec |
| ftv38 | 11 | 2.48sec | 9 | 2.2sec |
| p43 | 16 | 7.816sec | 3 | 1.498sec |
| ftv44 | 9 | 2.698sec | 9 | 2.698sec |
| ftv47 | 10 | 3.573sec | 10 | 3.573sec |
| ry48p | 11 | 4.415sec | 11 | 4.415sec |
| ft53 | 7 | 2.168sec | 7 | 2.168sec |
| ftv55 | 12 | 4.181sec | 9 | 3.26sec |
| ftv64 | 12 | 5.679sec | 9 | 4.15sec |
| ft70 | 6 | 3.167sec | 3 | 1.95sec |
| ftv70 | 7 | 4.93sec | 7 | 4.93sec |
| kro124p | 8 | 8.627sec | 8 | 8.627sec |
| ftv170 | 11 | 46.441sec | 11 | 46.441sec |
| rbg323 | 30 | 353.92sec | 2 | 9.251sec |
| rbg358 | 70 | 1699.97sec | 1 | 6.396sec |
| rbg403 | 1 | 7.489sec | 1 | 7.489sec |
| rbg443 | 2 | 17.207sec | 1 | 9.033sec |

There may be cases where the 1%-solution is obtained relatively fast, while the optimum may take very long to reach. The most typical instance is rbg358 where the 1%-solution is obtained already from the first iteration, while the optimum is obtained at the end of the 70th iteration, almost half an hour later.

We generally observe that although LB starts with a value not very far from the optimum, it is actually UB that eventually decreases more rapidly. Figure 1 illustrates the convergence of the two bounds in the case of an instance with 55 nodes and another one with 124 nodes.

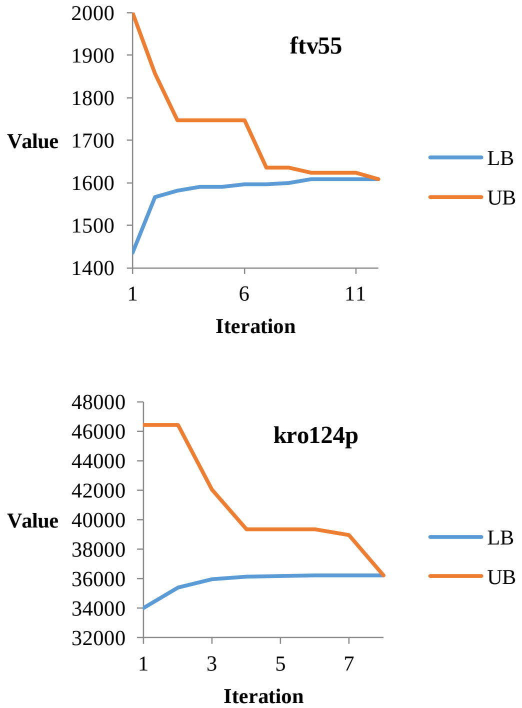



**Fig. 3.**
*Convergence of LB and UB for two instances – ftv55 and kro124p*

The cuts passed to the (AP′) are dense which results to a heavy simplex tableau. A possible future research direction will be to seek ways to lower the density of the obtained cuts in order to find a compromise between sufficient information (high-density cuts) and compactness that will allow us attack larger instances (low-density cuts). Apart from a certain instances (ftv170), the CPU time obtained for each instance does not seem to increase rapidly with the number of nodes. Figure 4 illustrates this relation.

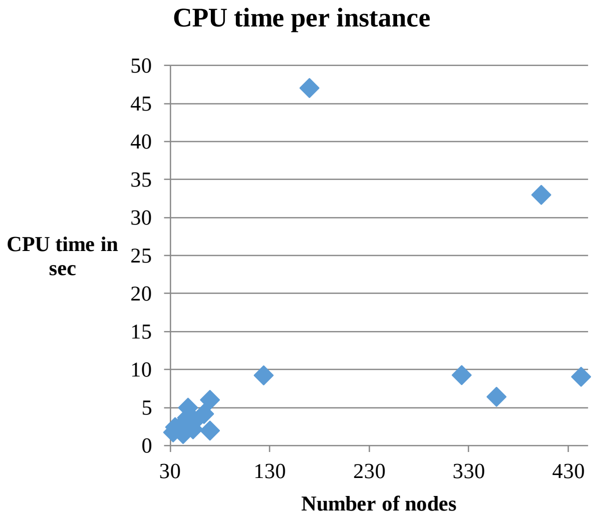



**Fig. 4.**
*Comparison of CPU time for different instances (Gap<1%)*

## 2. Conclusions

In this paper we considered the environmental TSP, the first application of the TSP in an environmental dimension where the fuel consumption of the tour is minimized. We attacked in this paper two distinct fronts. We proposed a novel cost function the aim of which is to minimize the fuel consumed throughout a route. Two major limitations of emission calculation models have motivated this work: (a) they cannot cater for real-time conditions such as weather, traffic etc, the result being that they provide a rather myopic view of the actual fuel consumption when factors such as grade, weather and use of air condition come into play; and (b) they require substantial information on the type of vehicle, engine and either characteristics which is not straightforward to input in a web platform. The data we take into account are provided through freely available web APIs. The proposed EES reflects the increase or decrease of the nominal values (provided by any calculation model) when a vehicle travels on a route segment. We based our model on available research and validated our results through comparison with an onboard diagnostic that we installed on a testing light-duty vehicle. The EES we devised feeds the objective function of the TSP.

The solution approach we followed for the TSP is based on the combination of a cutting plane strategy with 2-OPT heuristic. The cutting plane strategy appends a cut requiring an even number of connection between the nodes of each subtour and the remaining nodes outside the subtour. The resulting cuts are dense but effective. Solution of the assignment problem augmented with these valid inequalities provides an effective lower bound. The 2-OPT receives the different subtours and construct a feasible tour (not necessarily optimal however) and performs a 2-OPT move providing an upper bound. This upper bound is fed into the augmented problem when performing the branch and bound strategy to fathom nodes in the search tree. We tested the concept on the asymmetric instances of the TSP-LIB. The convergence of the proposed algorithm is reached after a few iterations, typically around 12. In the case one is interested in a solution within 1% margin of the optimum, then a tour is obtained typically within 10 seconds.

A future research direction will be to apply the environmental dimension of the TSP on heavy-duty vehicles and other transport modes. At the solution approach level, we will be investigating ways to make the appended cuts lighter (less dense) in order to benefit from economizing some computer memory.

**Acknowledgements**

The authors gratefully acknowledges financial support from the European Commission under the grant FP7-PEOPLE-2011-CIG, GreenRoute, 293753 and the grant EnvRouting SH3_(1234) of Action “Supporting Postdoctoral Researchers” of the Operational Program "Education and Lifelong Learning" (Action’s Beneficiary: General Secretariat for Research and Technology, Greece), and is co-financed by the European Social Fund (ESF) and the Greek State.

**References**

Applegate, D., Bixby R., Chvatal V, Cook W (2001). TSP Cuts Which Do Not Conform to the Template Paradigm.” In M Junger, D Naddef (eds.), *Computational Combinatorial Optimization, Lecture Notes In Computer Science*, Vol. 2241, pp. 261-304. Springer-Verlag Berlin Heidelberg.

Bektas T. (2006). The multiple traveling salesman problem: an overview of formulations and solution procedures. *Omega*, Vol. 34, No. 3, pp. 209-219.

Cornuejols, Gerard (2008). Valid Inequalities for Mixed Integer Linear Programs. *Mathematical Programming Ser. B*, 112:3-44.

Dantzig, G.B., Fulkerson, D.R. and Johnson, S.M. (1954). Solution of a large scale traveling-salesman problem. *Operations Research,* Vol. 2, pp. 393-410.

Edmonds, J. (1965). Maximum matching and a polyhedron with 0–1 vertices. *J. Res. Nat. Bur. Standards* 69B 125–130.
Grötschel M. and Padberg M.W. (1979). On the symmetric travelling salesman problem I: Inequalities, *Mathematical Programming*, Volume 16, Issue 1, pp 265-280.

Helsgaun, K., (2000). "An Effective Implementation of the Lin-Kernighan Traveling Salesman Heuristic". *European Journal of Operational Research* 126 (1): 106–130. doi:10.1016/S0377-2217(99)00284-2.

Laporte G. (1992). The Traveling Salesman Problem: An overview of exact and approximate algorithms. *European Journal of Operational Research*, 59, 1992, pp. 231-247.

Lin, S. & Kernighan, B. W. (1973). "An Effective Heuristic Algorithm for the Traveling-Salesman Problem". *Operations Research* 21 (2): 498–516. doi:10.1287/opre.21.2.498.

Padberg, M.W., Rinaldi, G., (1991). A branch-and-cut algorithm for the resolution of large-scale symmetric traveling salesman problems, *SIAM Rev*. 33 pp 60–100.

Reinelt, G., (1991). TSPLIB - A Traveling Salesman Problem Library. INFORMS Journal on Computing 01/1991; 3:376-384.

Saharidis G.K.D. (2014). Review of solution approaches for the symmetric traveling salesman problem. I*nternational Journal of Information Systems and Supply Chain Management.* 2014, unpublished.

Saharidis, G. (2012). Research report that presents the candidate TN factors for the definition of EESarc and the revision methodology following the development of new emission calculation models. FP7-PEOPLE-2011-CIG, GreenRoute: A web based platform which helps individuals and companies move commodities with the most environmental friendly way, minimizing emissions and transportation cost. Athens: Marie-Curie Grant.

Christofides, N. (1976). Worst-case analysis of a new heuristic for the travelling salesman problem. Technical Report 388, Graduate School of Industrial Administration, Carnegie Mellon University.

Held, M, and Karp, R.M. (1969). The traveling-salesman problem and minimum spanning trees. New York : IBM Systems Research Institute.

# PLANNING AND OPTIMAZATION OF A FAST-CHARGING INFRASTRUCTURE FOR ELECTRIC URBAN BUS SYSTEMS

**Alexander Kunith**[1], **Dietmar Goehlich**[2], **Roman Mendelevitch**[3]
*[1, 2] Department Methods for Product Development and Mechatronics, Technische Universität Berlin, Germany*
*[3]Workgroup for Infrastructure Policy (WIP), Technische Universität Berlin, Germany*

**Abstract:** The deployment of battery-powered electric bus systems within the public transportation sector plays an important role to increase the energy efficiency and to abate emissions. Raising attention is given to bus systems comprising fast-charging technology. This concept requires a comprehensive infrastructure to equip bus routes with charging points. The number of charging points in turn has significant influence on the financial investment needs. Therefore, an appropriate and efficient layout of the charging infrastructure is crucial to optimize the total cost of ownership of the deployed technology. The central issue of optimizing is described by a capacity set covering problem. A mixed-integer linear optimization model was developed to determine the minimum number and location of required charging stations for a bus line service while respecting operational and technology-related constraints, particularly the battery charging behavior. Within the scope of the study, different energy consumption scenarios are examined in order to reflect external factors affecting the bus energy consumption such as traffic volume and climate conditions. Furthermore, a sensitivity analysis on the battery capacity and required charging infrastructure was conducted. The findings reveal significant differences in terms of needed infrastructure based on the consumption scenarios and the daily operation time. Moreover, a trade-off between battery size and charging infrastructure has to be made. The model is an advanced tool for planning the fast-charging infrastructure of a bus system and hence supports the financial assessment of respective bus systems. The paper addresses upcoming challenges for transport authorities during the electrification process of the bus fleets and sharpens the focus on infrastructural issues related to the fast-charging concept.

**Keywords**: electric bus, charging infrastructure, fast-charging, cost optimization

## 1. Introduction

The implementation of alternative drive technologies such as battery-powered electric bus systems is of major importance to decrease the exhaust gas emissions of the public transportation. In order to evaluate the possible substitution of currently deployed conventional diesel buses, transport companies around the globe are intensifying electric bus trials (Choi et al. 2012; Erkkilä et al. 2013; Halmeaho 2014). Raising attention has been recently received by bus systems comprising fast-charging technology. This application follows the opportunity charging concept which assumes a comprehensive infrastructure of charging points to recharge the buses in operation by using high charging power (Bedell et al. 2011; Festner & Karbowski 2012; Miles & Potter 2013). Potential charging spots are the bus stops of the served bus line. Through fast-charging en route operational deployment in terms of range and operating time similiar to conventional diesel buses is feasible. Opportunity charging bus systems are characterized by smaller dimensioned battery storages compared to overnight slow-charging buses. The smaller dimensioned battery storages benefit lower vehicle weight, higher passenger capacity and lower investment needs for batteries (Goehlich et al. 2013).
The system imposes high infrastructural requirements since bus routes have to be equipped with charging points. The number of charging points in turn has significant influence on the total costs (Schwuerzinger & Puetz 2012). Therefore, an appropriate layout of the charging infrastructure is crucial to optimize the capital expenditures. The central issue of optimizing the infrastructure is described by a positioning problem of the charging points within a network and aims to determine the minimum number of required charging stations.
The emerging infrastructural challenges related to the deployment of alternative-powered vehicles such as electric vehicles (EV) are attracting widespread interest and have been recently addressed in the pertinent literature. Research has been conducted in the field of distributing charging point locations for passenger electric vehicles. Frade et al. (2010) developed a covering model to determine the optimal locations for a fixed number of public charging stations for an urban area to maximize the demand covered. Sadeghi-Barzani et al. (2014) established a model for the optimal placing and sizing of fast charging stations for electric vehicles. The approach comprises a mixed-interger non-linear (MINLP) optimization in order to minimize the total cost associated with supplying the energy demand of EV based on the number of charging stations. Liu et al. (2013) introduced a two-step screening method to identify the optimal sites of EV charging stations taking into account environmental factors and service radius, followed by the application of a mathematical algorithm to solve charging stations' optimal capacities. By employing an integer programming optimization model adapted from a Vehicle Routing Problem (VRP), Worley et al. (2012), presented an approach to perform the routing of commercial EV and distribution of charging stations simultaneously. Also based on VRP, Kameda and Mukai (2011) set up a model using the Node Insertion Algorithms for solving the location problem in context of an on-demand electrical bus system in Japan. The placement of charging stations is finally determined by means of taxi probe data. An alternative technique is proposed by Boostani et al. (2010) for an analogous location problem of compressed natural gas refueling station. The method applies a heuristic algorithm to solve the arc demand coverage problem of determining the optimal location of refueling station such that the flow demands are met and capital expenditure is minimized.

[1] Corresponding author: alexander.kunith@tu-berlin.de

However, public buses within a transport networks in general differ from the behavior of electric vehicles. In contrast to the above mentioned studies -focusing on passenger electric vehicles- the objective of this paper is to apply an optimization model for an urban public bus network. Central for bus network operations are predetermined and fixed scheduled routes. Hence, optimization models allocating a given number of charging points purposing to maximize the covered demand are not an appropriate approach to solve the infrastructural issue for fast-charging bus systems.
The application of the opportunity charging concept requires a cost related trade-off between battery capacity (battery cost and weight) and number of installed charging points. Only very limited research for electric bus operation had been carried out to broach the described issue. Chao et al. (2013) dedicated their investigation to establish a single depot vehicle scheduling model for a battery-switching station. The multi-objective model minimizes the number of vehicles to serve all trips, the number of standby batteries reserved for the continuous operation and determines the energy supply to meet daily demand. Concerning the charging schedule of fast-charging electric buses De Filippo et al. (2014) developed a simulation model to examine how electric bus charging patterns and queuing times are affected by the number of chargers deployed in the system and the charging policy employed.
A cost-optimized distribution of charging stations was not subject of these studies. To the authors' best knowledge, no comprehensive contributions are made in the field of planning a multi-charging station infrastructure for innovative electric bus systems following a cost optimization approach.
In this paper, a mixed-integer linear optimization model is presented. The model solves a capacitated set covering problem to determine the minimum required number and the respective location of charging points for a bus route while maintaining the same operational requirements which are applied for conventional diesel buses. The model simulation supports considerably the optimal infrastructure layout of fast-charging bus systems.

## 2. Problems planning the charging infrastructure for bus systems

In planning a charging infrastructure for electric fast-charging bus systems, several technology-related and operational constraints have to be respected. The following section aims to specify relevant aspects and requirements for setting up an appropriate optimization model.
The optimal distribution of charging points fundamentally results from the efficient replenishment of the bus energy consumption. Therefore, it is essential that the route specific energy consumption is represented adequately and efficiently supplied during daily operation. The energy which can be recharged at charging points mainly depends on the dwell times at bus stops and on the available charging power. The dwell times are determined by the schedule and the operational planning of the bus route and might vary according to the characteristics of the bus stops. The charging power itself is constrained by the available grid power, the battery type and the battery's state of charge (SoC) which is of particular significance for actual bus operation. The battery's charging behavior in terms of the charging power and the SoC can be described by a non-linear function. In addition, to improve the longevity of a battery the depth of discharge (DoD) is intended to be kept relatively small so that in operation the battery is attempted to perform in a predefined SoC range (Marano 2009; Millner 2010; Lam & Bauer 2013).

Several factors similar to internal-combustion vehicle such as vehicle weight, topography of operation area and energy efficiencies influence the total energy consumption. In contrast, full-electric buses feature specific consumptions characteristics, in particular in relation to climate conditions and air-conditioning and heating due to low engine waste heat. Thus, significantly higher energy consumption for extreme weather conditions has to be taken into account (Goehlich et. al 2014).

However, the energy consumption cannot be assumed constant over the entire route service. The consumption rather varies with the current circumstances the bus is exposed to. Therefore, it is crucial to determine the energy consumption for route sections. A route sectioning can be conducted in accordance with the bus stops and the section's characteristics such as traffic volume and passenger occupancy rate have to be captured by an optimization model. Since bus routes feature different driving profiles it is required to collect the relevant route section data for all bus lines considered. Furthermore, each potential bus stop needs to be evaluated individually in context of dwell times and usage but also regarding possible institutional and local structural reservations against the installation of charging points at certain bus stops. Bus stations affected by such reservations should be excluded from further investigations.
Finally, the charging infrastructure planning is ideally conducted from the network perspective in order to leverage synergies. This implies that the simulation model depicts network specific information such as bus line intersections assuming that the potential multiple use of the charging infrastructure leads to cost savings.

## 3. Methodology – optimization model

In the following section a general mixed-integer optimization model is described that is able to capture the main features of an electric bus system as described in the section above. While representing network, operational and technical constraints, it determines a cost minimizing set of charging stations and respective charging profiles for each bus in the system.

The objective function (1) in the original version of the model minimizes the number of charging stations and the resulting cost of construction at the respective stop ( $\text{ccosts}_i$ ) while serving the required daily demand. The variable $z_i$ indicates whether a charging station is built at stop $i$ .

$$\min_{z_i, e_{i,n,b}, load_{i,n,b}, u_{i,n,b,(1,2)}, load\_level_{i,n,b,l}} \quad \sum_i z_i \cdot \text{ccosts}_i \tag{1}$$

Numerous constraints are necessary to represent the network layout and the technical parameters of the charging process. A nodal energy balance constraint (2) ensures that the current energy level $e\_before_{j,n,b}$ at each stop $j$ , for each tour $n$ and for each line $b$ , is equal to the energy level at the previous station minus consumption for traveling from the previous stop $t_{i,j,b}$ [2] plus energy recharged at the current station (dwelling time of a bus at a stop $wz_j$ times a variable kW charging power $load_{j,n,b}$ ). Distinction between $e\_before_{i,n,b}$ and $e_{i,n,b}$ which generally capture the same value, namely the current energy level, become necessary due to the complex charging process and is further described later in this section. This energy balance applies to all stops except for the first stop of each tour $y_b$ . For these stops the alternative energy balance (3) applies, where $w_b$ is a stop just before the first stop.

$$e\_before_{j,n,b} = e_{i,n,b} - t_{i,j,b} + load_{j,n,b} \cdot wz_j \qquad \forall j \neq y_b, n, b \mid t_{i,j,b} > 0 \tag{2}$$

$$e\_before_{j,n,b} = e_{i,n-1,b} + load_{j,n,b} \cdot wz_j \qquad \forall j = y_b, i = w_b, n, b \tag{3}$$

Additional constraints specify the absolute lower bound and upper bound and initial charging level of the battery as well as the boundary constraints for specific cases as “depot to first stop” and “retour from final stop”.

To approximate the non-linear form of the charging power function of a battery the function is split into three different segments divided by thresholds $tr_{1,2}$ , where different charging power levels $lp_{ll}$ represent the average charging power in the respective segment. To pin down the segment of the energy level at stop $j$ before charging which depends on the battery capacity $bt_b$ , logical equations are employed. Here, $load\_level_{j,n,b,l}$ is a binary variable that flags the respective segment (c.f. Table 1**Error! Reference source not found.**).[3]

$$(e_{i,n,b} - t_{i,j,b}) / bt_b \leq 1 - (1 - tr_l) \cdot load\_level_{j,n,b,l} \qquad \forall i, j, n, b, l = 1,2 \mid t_{i,j,b} > 0 \tag{4}$$

$$tr_l \cdot load\_level_{j,n,b,l} \leq (e_{i,n,b} - v \cdot t_{i,j,b}) / bt_b \qquad \forall i, j, n, b, l = 2,3 \mid t_{i,j,b} > 0 \tag{5}$$

$$load\_level_{j,n,b,'1'} + load\_level_{j,n,b,'2'} + load\_level_{j,n,b,'3'} = 1 \qquad \forall j, n, b \tag{6}$$

**Table 1**

*Boolean values of logical conditions for different energy levels* [4]

| load_level before charging/ load_level after charging | $e \leq tr_1$ | $tr_1 \geq e \geq tr_2$ | $tr_2 \leq e \leq 1$ |
|---|---|---|---|
| l=1/4 | 0/1 | 0 | 0 |
| l=2/5 | 0 | 0/1 | 0 |
| l=3/6 | 0 | 0 | 0/1 |

Additionally, it is necessary to account for the fact that during the charging, one of the threshold levels might be passed which would result in over- or understating the battery level after charging. To correct for that the hypothetical battery level after charging is identified with the same charging power using equations (7)-(9). For the first stop of each new tour equations (4)-(9) have to be adjusted following the same logic as applied in equation (3).

$$e\_before_{j,n,b} / bt_b \leq 1 - (1 - tr_{l-3}) \cdot load\_level_{j,n,b,l} \qquad \forall j, n, b, l = 4,5 \tag{7}$$

$$tr_{l-4} \cdot load\_level_{j,n,b,l} \leq e\_bofore_{j,n,b} \cdot bt_b \qquad \forall j, n, b, l = 5,6 \tag{8}$$

$$load\_level_{j,n,b,'4'} + load\_level_{j,n,b,'5'} + load\_level_{j,n,b,'6'} = 1 \qquad \forall j, n, b \tag{9}$$

[2] The parameter $t_{i,j,b}$ also captures the routes of the lines by only allowing to travel between adjacent $i$ and $j$ for a given line $b$ .

[3] Quotation marks ‘’ refer to the section of specific entry of a set

[4] Equations (6) and (9) force the conditions to flag true whenever it is possible.

Next, the amount of energy is calculated that was incorrectly accounted for ( $u_{i,n,b,(1,2)} \cdot wz_i$ ) in equations (10) and (11). Additional constraints are required to make sure that $u_{i,n,b,x}$ is only associated with a non-zero value if a threshold is passed during the charging process. Moreover, constraints ensure that $u_{j,n,b,(1,2)}$ only has a value if one of the thresholds was passed and charging has occurred at the station under consideration.

$$e\_before_{j,n,b} / bt_b - tr_l \leq (1\text{-}load\_level_{j,n,b,l}) \cdot BIG + u_{j,n,b,l} \qquad \forall j,n,b,l=1,2 \tag{10}$$

$$e\_before_{i,n,b} / bt_b - u_{p,n,b,l-4} \geq tr_{l-4} \cdot load\_level_{p,n,b,l} + tr_{l-4} \cdot load\_level_{p,n,b,l+1} \qquad \forall j,n,b,l=5,6 \tag{11}$$

The flags from equations (4)-(6) are used to determine the maximum charging power at each stop in equation (12). Charging is only permitted if a charging station was built at stop p which is ensured by equation (13).

$$load_{j,n,b} \leq \sum_{l=1}^{3} lp_{l,b} \cdot load\_level_{p,n,b,l} \qquad \forall j,n,b \tag{12}$$

$$load_{j,n,b} \leq z_p \cdot BIG \qquad \forall j,n,b \tag{13}$$

Due to the potential overcharging the preliminary energy level at each station $e\_before_{i,n,b}$ potentially needs to be corrected to reach the final level $e_{i,n,b}$ if during charging, one of the thresholds was passed. In equation (14) (14)(14) $u_{p,n,b,(1,2)}$ is multiplied with the respective correction factor and added to or subtracted from the preliminary energy level.

$$e_{j,n,b} = e\_before_{j,n,b} - \left(u_{p,n,b,'1'} \cdot bt_b \cdot \left(1 - lp_{'2',b} / lp_{'1',b}\right)\right) - \left(u_{p,n,b,'2'} \cdot bt_b \cdot \left(1 - lp_{'3',b} / lp_{'2',b}\right)\right) \tag{14}$$

**4. Application case - model assumptions**

The application of the developed optimization model described in section 3 is conducted under certain constraints and assumptions. In the following paragraph the scope of modeled scenarios is defined and the data basis is specified.

The simulation is exemplarily applied for a reference route. Therefore, the siting of the charging points is determined for a single bus route and a 15-hour service. This bus line covers around 28.5 km (total route including return) and 30 bus stops with the turning point at the $15^{th}$ bus stop. It is assumed that the stops of the forward and return run are different and no cost benefit can be realized when placing a charging point. For each bus stop the dwell time and the distances between stops are provided as input parameters. It is assumed that the bus, when coming from the depot for the first ride, reaches the initial stop with a SoC of 90% and returning from the final stop of the bus route to the depot at the end of service with a battery capacity not below the critical level of 20%. Reflecting the battery charging behavior, it is presumed for simplification that within a range of 30-70% SoC the energy is transferred with a charging power of 200 kW, outside this range the power is reduced to 100 kW.

Furthermore, three energy consumption scenarios are defined: low corresponding to 0.8-2.4 kWh/km, medium stating 1.0-3.0 kWh/km and high 1.2-3.6 kWh/km. The different scenarios represent the factors described in section 2 such as traffic and temperature causing a varying bus energy consumption. The values within the stated ranges for each scenario are assigned to the route sections which are defined as the distance between two subsequent bus stops. Additionally, the simulation is modeled with two different bus batteries providing 90 kWh and 120 kWh.

**5. Results**

In this study, an optimization model for the layout of fast-charging infrastructure was developed and simulations on the basis of one bus line and different scenarios were carried out. The number and location of needed charging points to operate the bus route described in section 4 were obtained. The final results are presented in Fig. 1 and Fig. 2. The two figures respectively show the results for daily service with 15 round trips per bus for the low, base and high energy consumption scenario as well as for battery energy capacities of 90 kWh and 120 kWh. Since the last trip of the day determines the final number of required charging stations, the figures display the bus SoC chart (in percentage) during the $15^{th}$ trip serving the 30 bus stops. The curves feature noticeably SoC jumps which indicate that at the corresponding bus stop a charging station is placed and the bus is charged. Additionally, the charged energy in proportion of the battery capacity is given exemplarily for the base scenario.

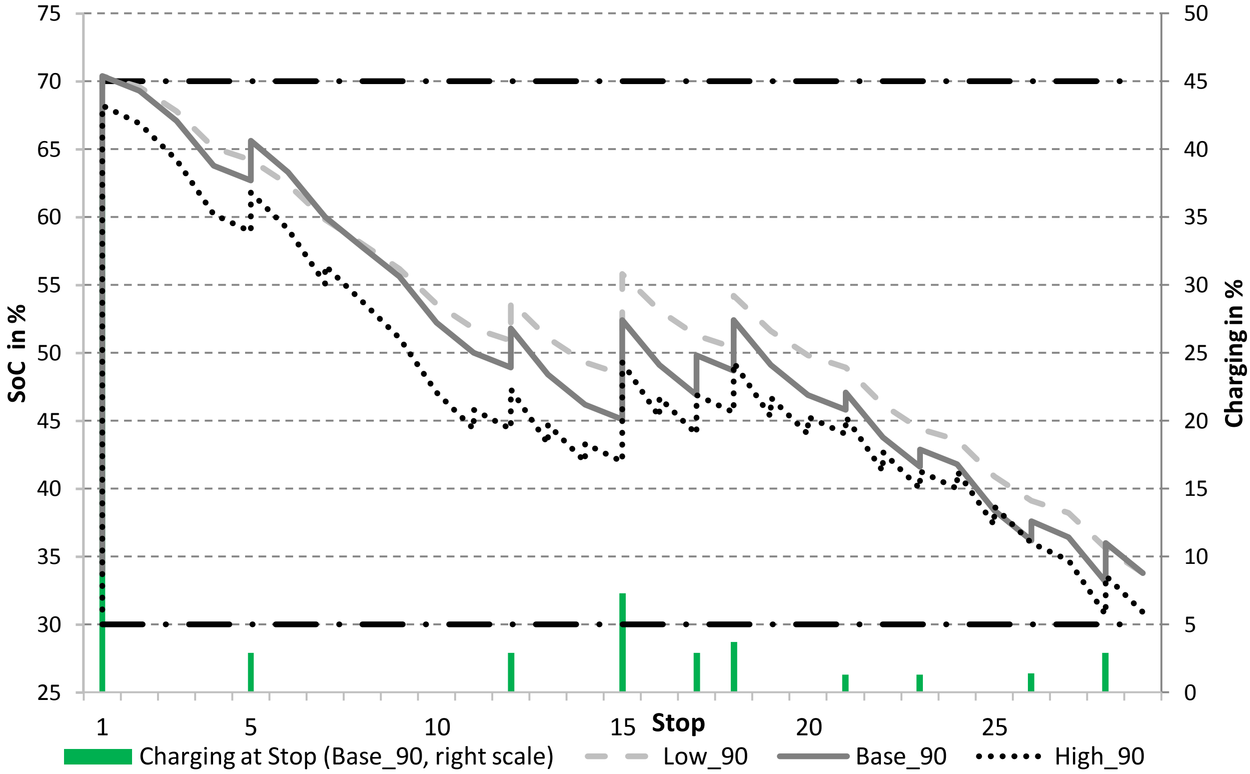



**Fig. 1.**
*Needed charging infrastructure considering three consumption scenarios and a battery energy capacity of 90kWh*

As can be seen in Fig. 1, in order to operate the bus line with a 90 kWh battery for the low, base and high scenario, 4, 10 and 19 charging points are required, respectively.

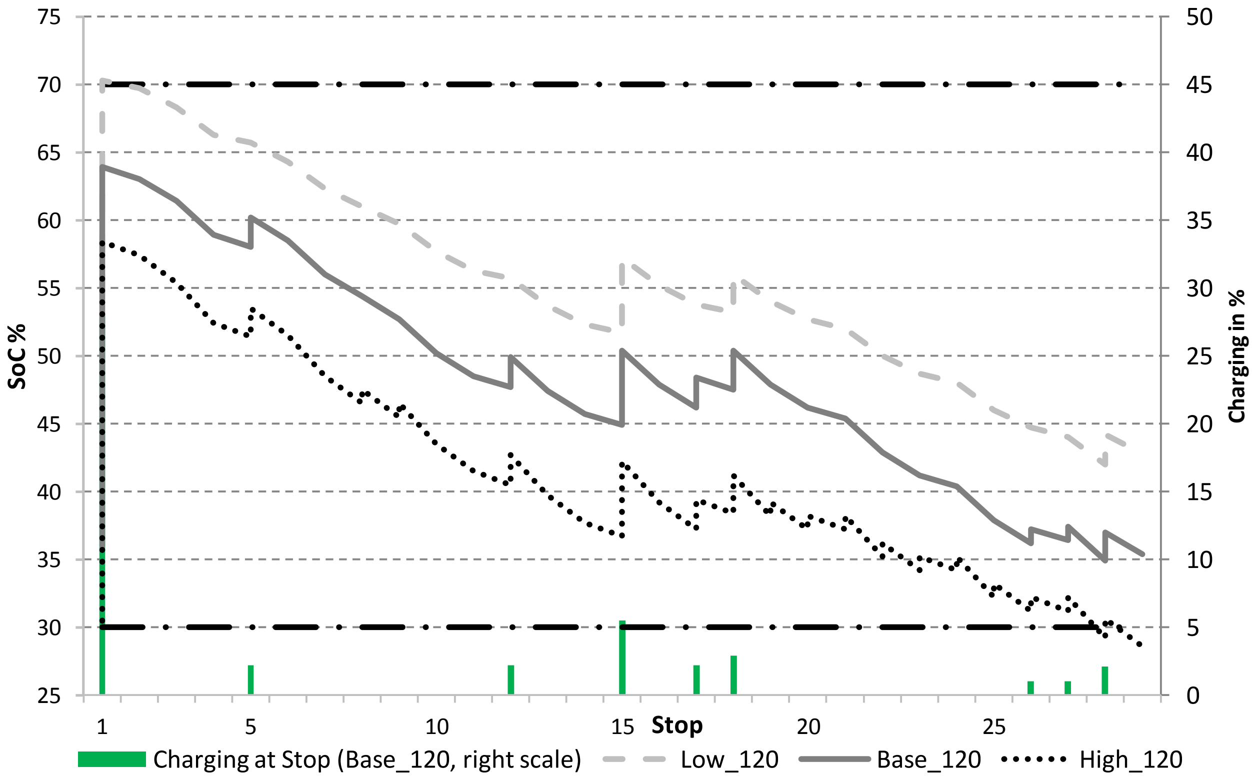



**Fig. 2.**
*Needed charging infrastructure considering three consumption scenarios and a battery energy capacity of 120kWh*

The results presented in Fig. 2 reveal that for bus operation with 120 kWh, 4, 9 and 18 charging points are needed depending on the consumption scenario. The number of charging points in Fig. 1 and Fig. 2 are slightly different comparing the equivalent scenarios. A possible significant decrease in charging points due to higher battery capacity could not be observed for the application case. Furthermore, the placement of the charging points comparing 90 kWh and 120 kWh are almost identical resulting from similar SoC curves considering the entire dataset. In addition to the described model variations, the number of round trips was reduced to fewer repetitions (2-4). In case of the 120 kWh and base scenario, for instance, the number of charging point counts 1 when simulating only 2 round trips. Therefore, the daily operation time of a bus affects considerably the infrastructure requirements.

However, these results suggest that the model is capable to provide comprehensive data on central issues of planning a fast-charging infrastructure and thus supports the evaluation of innovative bus systems.

## 6. Conclusion

The planning and optimization of fast-charging infrastructure for the deployment of electric buses was studied. A mixed-integer linear optimization model was developed to minimize the charging infrastructure costs. The applied model determines the minimum required number and the respective location of charging points for a bus route while taking operational and technology-related constraints, in particular the battery charging behavior, into account. It is demonstrated that different assumptions regarding energy consumption leads to significantly different charging infrastructure requirements. Thus, potential factors causing increased energy consumption need to be carefully analyzed. Contrarily, the battery size does not prove substantial influence on the number of charging points in the application case. Nevertheless, a possible trade-off between battery size and installed charging points has to be attended in order to optimize investment needs. Moreover, the number of scheduled trips, equivalent to the daily operation hours, can cause remarkable discrepancy of the needed infrastructure complexity in terms of installed charging spots.

The model is sophisticated tool for planning and optimizing the infrastructure for fast-charging bus systems. The optimized layout supports transit operators in their financial assessment of respective bus systems. However, some limitations are worth noting. The application case allows the limited simulation of one single bus line. Expected synergy due to the electrification of an extended bus network should be addressed in future work. Furthermore, additional constraints resulting from power grid limitations and availability of the charging spots, in particular for network operation, should be considered for further research.

## Acknowledgements

The authors gratefully acknowledge the financial support provided by the German Federal Government. This research was carried out as a part of the German Federal Government's Showcase Regions for Electromobility, the funding is issued by the Federal Ministry of Transportation and Digital Infrastructure. Furthermore, particular gratitude is expressed to Mario Kendziorski, Alexander Müller and Valentin Zinnecker, members of the research group, for their support developing the optimization model.

**References**

Bedell, R.; Westerlund, B.; Astrand, P. 2011. First Results from Field Testing of Fast Charged Hybrid Buses in Umea, Sweden. In *Proceedings of EEVC European Electric Vehicle Conference*, 1-5.

Boostani, A.; Ghodsi, R.; Miab, A. 2010. Optimal location of compressed natural gas (CNG) refueling station using the arc demand coverage. In *Proceedings of the 4th Aisia International Conference on Mathematical/Analytical Modelling and Computer Simulation*, 193-198.

Chao, Z.; Xiaohong, C. 2013. Optimizing battery electric bus transit vehicle scheduling with battery exchanging: model and case study, *Journal of Procedia - Social and Behavioral Sciences* 96: 2725-2736.
Choi, U.-D.; Jeong, H.-K.; Jeong, S.-K. 2012. Commercial operation of ultra low floor electric bus for Seoul city route. In *Proceedings of IEEE Vehicle Power and Propulsion Conference*, 1128-1133.

De Filippo, G.; Marano, V.; Sioshansi, R. 2014. Simulation of an electric transportation system at The Ohio State University, *Journal of Applied Energy* 113: 1686-1691.

Erkkilä, K.; Nylund, N.-O.; Pellikka, A.-P.; Kallio, M.; Kallonen, Kallio, M.; Ojamo, S.; Ruotsalainen, S.; Pietikäinen, S.; Lajunen, A. 2013. eBUS - Electric bus test platform in Finland. In *Proceedings of EVS27 International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium*.

Festner, L.; Karbowski, G. 2012. Foothill Transit Ecoliner Electric Bus Program. In *Proceedings of EVS26 International Battery, Hybrid, and Fuel Cell Electric Vehicle Symposium*, 1–8.

Frade I.; Ribeiro A; Antunes A.P.; Gonçalves G.A. 2011. An optimization model for locating electric vehicle charging stations in central urban area, *Journal of Transport Research Record: Journal of the Transportation Research Board* 2252: 91-98.

Goehlich, D.; Kunith, A.; Ly, T.-A. 2014. Technology assessment of an electric urban bus system for Berlin. In *Proceedings of the 20th International Conference on Urban Transport*, 137–149.

Goehlich, D.; Spangenberg, F.; Kunith, A. 2013. Stochastic Total Cost of Ownership Forecasting for innovative Urban Transport Systems. In *Proceedings of the 7th IEEE Int. Conference on Industrial Engineering and Engineering Management.*

Halmeaho, T. 2014. Electric city buses. Available from internet, viewed 25 July 2014:
<https://noppa.aalto.fi/noppa/kurssi/as-84.3210/luennot/AS-84_3210_electric_city_buses.pdf>

Lam, L.; Bauer, P. 2013. Practical Capacity Fading Model for Li-ion Battery Cells in Electric Vehicles, *Journal of IEEE Transactions on Power Electronics* 28 (12): 5910–5918.

Liu, Z.; Wen, F., Ledwich, G. 2013. Optimal planning of electric-vehicle charging stations in distribution systems, J*ournal of IEEE Transactions on Power Delivery* 28 (1): 102–110.

Marano, V.; Onori, S.; Guezennec, Y.; Rizzoni, G.; Madella, N. (2009): Lithium-ion batteries life estimation for plug-in hybrid electric vehicles. In *Proceedings of IEEE Vehicle Power and Propulsion Conference (VPPC)*, 536–543.

Miles, J.; Potter,S. 2013. Developing a viable electric bus service: the Milton Keynes demonstration project. In *Proceedings of Thredbo 13: International Conference on Competition and Ownership in Land Passenger Transport*, .

Millner, A. (2010). Modeling Lithium Ion battery degradation in electric vehicles. In *Proceedings of IEEE Conference on Innovative Technologies for an Efficient and Reliable Electricity Supply*, 349–356.

Sadeghi-Barzani, P.; Rajabi-Ghahnavieh, A.; Kazemi-Karegar, H. 2014. Optimal fast charging station placing and sizing, *Applied Energy* 125, 289-299.

Schwürzinger, J.; Pütz, R. 2012. *Report on Electrified Public Transport Bus System.* UITP - International Association of Public Transport.Belgium 67 p.

Utsumi, H.; Saito, K.; Saito, A.; Tanaka, K.; Ohmoto, H. 2013. E-Bus for Akita Prefecture, Japan. In *Proceedings of EVS27 International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium*.

Worley, O.; Klabjan, D.; Sweda, T. 2012. Simultaneous vehicle routing and charging station siting for commercial. In *Proceedings of IEEE International Electric Vehicles Conference*, 1-3.

# AN ANALYSIS OF BASES FOR DEFINING LEVEL OF THE PORT MACHINERY FLEXIBILITY IN A MULTIPURPOSE SEAPORT

**Deda Djelovic**[1]
*Port of Bar, Montenegro;*

**Abstract:** Decision on which port machinery to select in a multipurpose seaport is very complex and under influence of a broad range of factors. According to the most relevant recommended selection methodologies, basic influential factors that have to be taken into account can be classified into five major classes: port development factors, equipment costs, factors related to equipment maintenance, manning requirements and operating factors.
Level of the port machinery flexibility (its both main components: technological and positional flexibility), as an element of the class of influential factors titled "operating factors" has high importance in a multipurpose seaport due to intensive variations in volume and characters of port customers demands over the time.
After some general considerations related to system flexibility, this paper is focused on general aspects of the port machinery selection process and, than, port machinery flexibility is analyzed in details. Special attention is given to structuring a methodology for defining optimal level of the port machinery flexibility. Some of the presented concrete data are referred to the port machinery in the Port of Bar as a multipurpose seaport.

**Key words:** seaport, port machinery, flexibility.

## 1. Introduction

Many changes have occurred in the past period in the technology of ships and cargo handling and this developing trend is likely to continue. One of the key principles in planning seaport facility, therefore, is that development plan should be as flexible as possible to allow a prompt response to changing demands.

A specialized terminal (with appropriate port machinery) will normally be justified whenever a traffic forecast shown that there are sufficient customers` demands. This principle will apply to specialized terminals for cargoes such as containers, timber products, iron and steel, dry bulk commodities, etc. But, when traffic forecasts and expected values of other relevant parameters can not justify investments in specialized facility - the remaining decision is to invest in multi-purpose facility an to try, during the process of defining its elements (especially port machinery), to provide its highest possible flexibility from different aspects.

Intensive variations in nature of the customers` demands are especially characteristic for multipurpose seaports. As an illustration of the previous statement can be used data on changes of the throughput structure in the Port of Bar for the 15 years period, from 1999 to 2013 (Port of Bar documentation, 2014): share of the dry bulk cargoes in the overall throughput varied from 5% to 59%/year; share of the liquid cargoes in the overall throughput varied from 17% to 47%/year; share of general cargo in the overall throughput varied from 21% to 57%/year.

Principle objectives of this paper are to recognize port machinery flexibility as a crucially important component of the overall port flexibility, to identify key bases for optimization of the port machinery flexibility level and to create bases for further researches in this domain.

Considerations done are based on relevant literature from the domains of system flexibility and port machinery selection, as well as on results of the specific analyses performed (in the Port of Bar). Shown results and made conclusions are, dominantly, related to the port shore cranes, but with some appropriate adjustments those can be adequately connected and with other categories of the port machinery.

## 2. General considerations on a system flexibility

In this chapter is given an short overview of considerations on a system (port, etc.) flexibility, found in the relevant literature.

The stability of the competitive business environment of the past decades has been replaced by increasing uncertainty (Taneja et al., 2012). The inevitability of rapid change in the competitive environment of business is an accepted fact. This changeful environment, coupled with maturing markets, has resulted in enormous pressures of firms in their efforts to continue performing at high levels. One of the responses of firms has been to adapt to the new demands of the environment in a gradual, incremental manner. In recent times, managements of firms have had to accept the fact that rapid changes in the marketplace can not be controlled. The unpredictability of change, it is easily recognized, is a major factor limiting firm`s capability to respond effectively in a preplanned manner (Das, Elango, 1995).

[1]Corresponding author: deda.djelovic@lukabar.me;

However, in environments characterized by continuous changes the implementation of flexibility is a very important element of the firm performance. The flexibility of the organization will allow the firm to respond to various changes in the environment appropriately. These kinds of changing environments will allow the organization to gain the full benefits of its strategy compared to environments with little or no changes (Das, Elango, 1995).

*Flexibility* is defined as the ability of a system to adapt to external changes, while maintaining satisfactory system performance. System performance is characterized by parameters such as capacity, level of service, maintainability, and profitability. External changes are uncontrolled conditions that affect the system, including changes in level of demand or use, shifts in spatial traffic patterns, infrastructure loss and degradation, and changes in the price and availability of important resources such as fuel, etc. (Morlok, Chang, 2004).

There are many kinds of flexibility and indeed a sizable literature is devoted to the various kinds of flexibility. However, from the organizational point of view, all forms of flexibility present a common challenge: efficiency requires a bureaucratic form of organization with high level of standardization, formalization, specialization, hierarchy and staffs; but these features of bureaucracy impede the fluid process of mutual adjustment required for flexibility (Adler et al., 1999).

Strategic flexibility is the ability of an organization to respond to changes in the environment in a timely and appropriate manner with due regard to the competitive forces in the market place, The flexibility may be achieved by the organization in several ways, both external and internal. The significant external factors include suppliers, alliances and multinational operations. The internal factors comprise manufacturing flexibility, modular product design, employee flexibility and organizational structure (Das, Elango, 1995).

Organizations which place a premium on using flexibility as a means to attain long term viability need to realize that flexibility entails certain disadvantages. These disadvantages include higher cost, increased stress and a potential lack of focus. The managers of a firm should realize that strategic flexibility, even though desirable, is not the solution to all problems of the organization. Strategic flexibility should be part of a strategy only if the benefits far outweigh the undesirable effects of the flexibility (Das, Elango, 1995).

Ports have moved from performing nautical, cargo handling, stacking, and distribution functions to being multimodal nodes in the logistic chain. Academic research on ports has evolved from a focus on the physical infrastructure to the supply chain perspective, where a port is seen as a node in the supply chain network (Asperen, Dekker, 2009).

Modern technology is being implemented within the maritime industry. The impacts have focused on several areas: bigger vessels, new equipment configurations, new logistic concepts, new cargo handling concepts, advancements in ICT leading to development of information systems such as GIS (Geographic Information Systems), dynamic real time control of operations, efficient data collection and processing time, new camera systems, new gate processing systems, and the introduction of radiation detection monitoring systems. It can be concluded that the driving force behind the most significant trends for the areas of port planning, design, and project appraisal is an uncertain environment (Taneja et al., 2012).

The word flexibility in the context of ports can have many associations, e.g., flexible port layout, flexible infrastructures, flexible operations, flexible management, etc.

Most of the current trends in the port sector can be categorized as barriers, drivers, or enablers of flexibility (Fig. 1.).

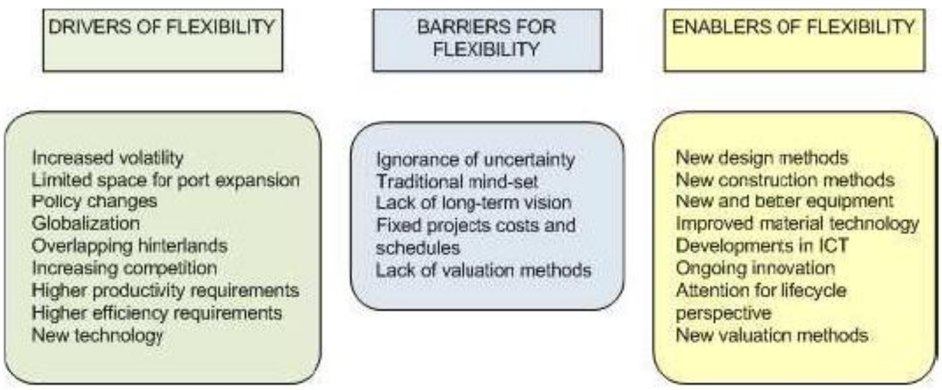



**Fig. 1.**
*Drivers, barriers and enablers of flexibility*
*Source: Taneja et al., 2012*

Respecting complex correlations of a port with elements of its changing environment and highest possible importance of the machinery in a port, further considerations are focused on some aspects of the port machinery flexibility. With achieved optimal level of the port machinery flexibility and port flexibility in general is significantly increasing.

## 3. Level of flexibility as an influential factor on port machinery selection

Since operating and other costs are not likely to be the deciding factors, the complex decision on which port machinery to select must be based on a broad range of factors. The many factors that have to be taken into account in evaluating port machinery investment decisions and in a multipurpose seaport (terminal) planning process are identified here. Results are shown by cause-effect diagram (Fig. 2.).

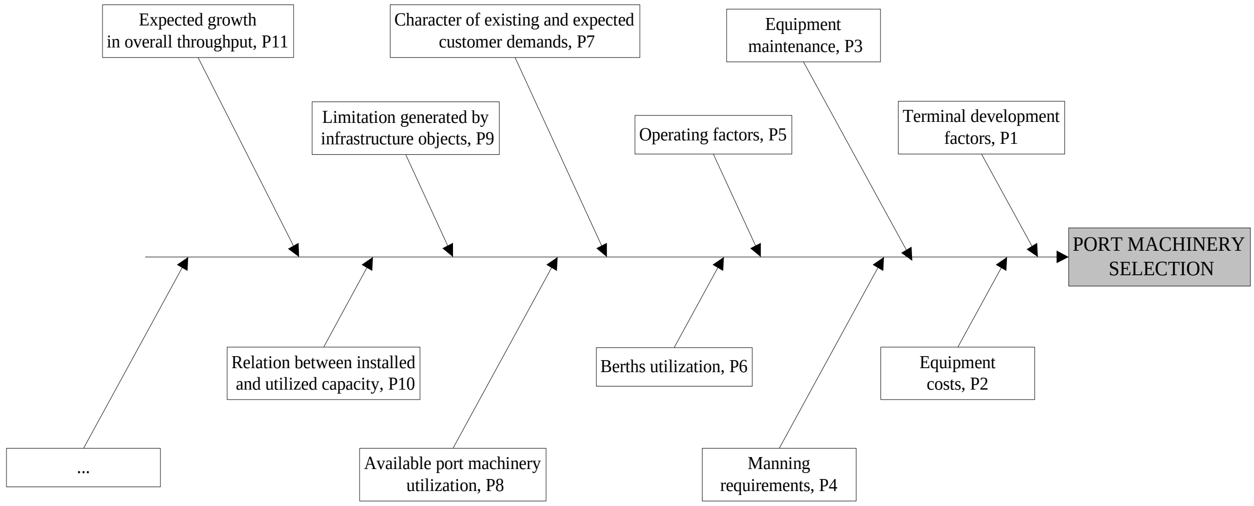



**Fig. 2.**
*Identification of influential parameters on port machinery selection*

Previously shown analyses results are based on consideration done in (Dobner, Rijsenbrij, 1999; Roach, 1987;).

All of identified groups of influential factors are complex and consist of several elements. If the group $P_5$ (operating factors of influence) is focused, than its structure can be defined:

- Group $P_5$ - operating factors of influence: **degree of flexibility, $p_{51}$;** travel speed, $p_{52}$; average hourly handling rates, $p_{53}$; lifting capacity, $p_{54}$; intermodal capability, $p_{55}$; safety records, $p_{56}$; warning devices, $p_{57}$; etc.

## 4. Port machinery flexibility

In general, it is possible to identify two main “components” of port machinery flexibility: “positional” flexibility and “technological” flexibility.

“Positional” flexibility considers that port machinery performances make possible its usage on different position within port (terminal) area and the “technological” flexibility means that exploitation characteristics of the machinery can match very wide range of technological requirements generated during the cargo handling process. Both flexibility components are determined by values of numerous factors, which are identified in the following part of this chapter. The most important factors which determine required level of port machinery “positional” flexibility are defined and shown with Fig. 3.

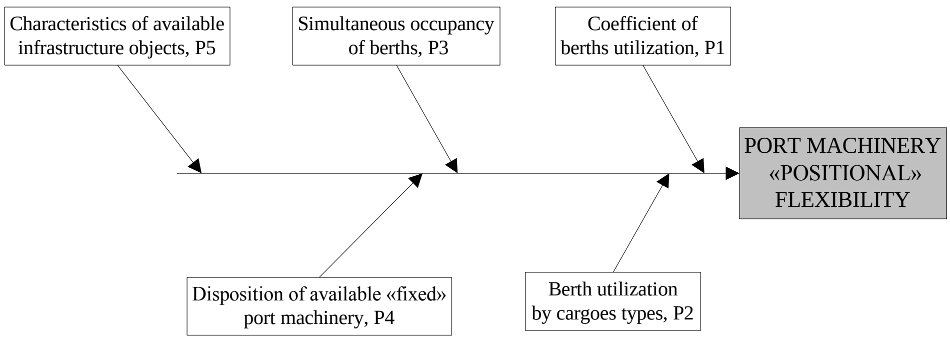



**Fig. 3.**
*Factors which determine required level of port machinery “positional” flexibility*

Factors which have dominant influence on required level of port machinery "technological" flexibility are shown by the next diagram (Fig. 4.).

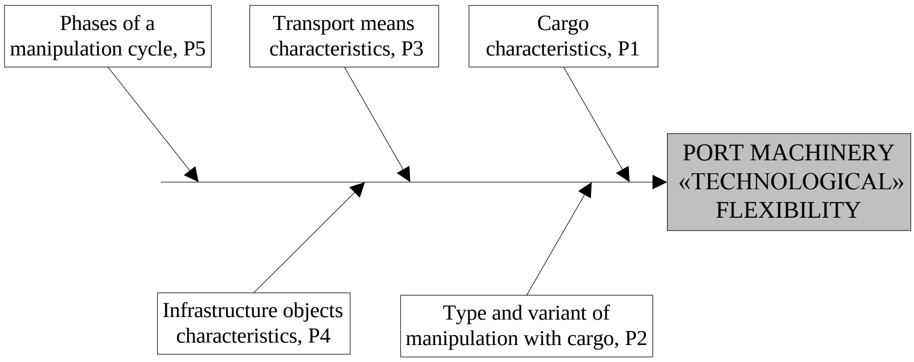



**Fig. 4.**
*Factors which determine required level of port machinery "technological" flexibility*

### 4.1 An analysis of factors which determine required level of port machinery "technological" flexibility

Port machinery "technological" flexibility has already been defined as ability of port machinery to match very wide range of technological requirements generated during the cargo handling process. Otherwise, technological requirements appeared (generated) in the cargo handling process belong to group of key bases for port machinery performances definition and, at the same time, for port machinery category selection (Djelovic, 2000). Previous statement can be illustrated with following diagram (Fig. 5).

| PORT EQUIPMENT SELECTION PROCESS | | | **Phase $PH_i$**: Analysis of technological requirements influence | | | |
|---|---|---|---|---|---|---|
| | | | | | | |
| Parameters which determine technological requirements, $P_i$ | ► | Technological requirements appeared in the cargo handling process, $TR_i$ | ► | Port machinery performances, $EP_i$ | ► | Port machinery category, $C_{PHi}$ |

**Fig. 5.**
*Influence of technological requirements in the process of port machinery selection*

Selected port machinery category is result of the final Port equipment selection process phase where categories $C_{Phi}$ are studiously compared based on defined criteria.

Technological requirements which appeared in the cargo handling process (process of port equipment exploitation) are mainly determined by following groups of parameters: cargo characteristics, $P_1$; type and variant of manipulation with cargo, $P_2$; transport means (ship, wagon, truck) characteristics, $P_3$; infrastructure objects characteristics, $P_4$; characteristic phases of a manipulation cycle, $P_5$; It can be concluded that these groups of parameters are exactly the same with groups of influential factors on port machinery technological flexibility.

Technological requiremets are the key bases for defining cargo handling technologies and are determined by numerous factors of different nature. The dominant influence on character of the technological requirements have: kind and characteristics of cargo, type of the handling operation, performances of the transport means used for transporting cargoes to the port/from the port, performances of the port infrastructure objects, etc. (Djelovic, 2004)

Basic technological requirements are generated in the following processes: unloading cargoes from ship to open storage area (and vice versa), storing cargoes, stuffing and un-stuffing containers, unloading cargoes from trucks and wagons (and vice versa), etc.

Using Cause - effects method, a partial decomposition of groups of parameters which determine technological requirements is done (Fig. 6.).

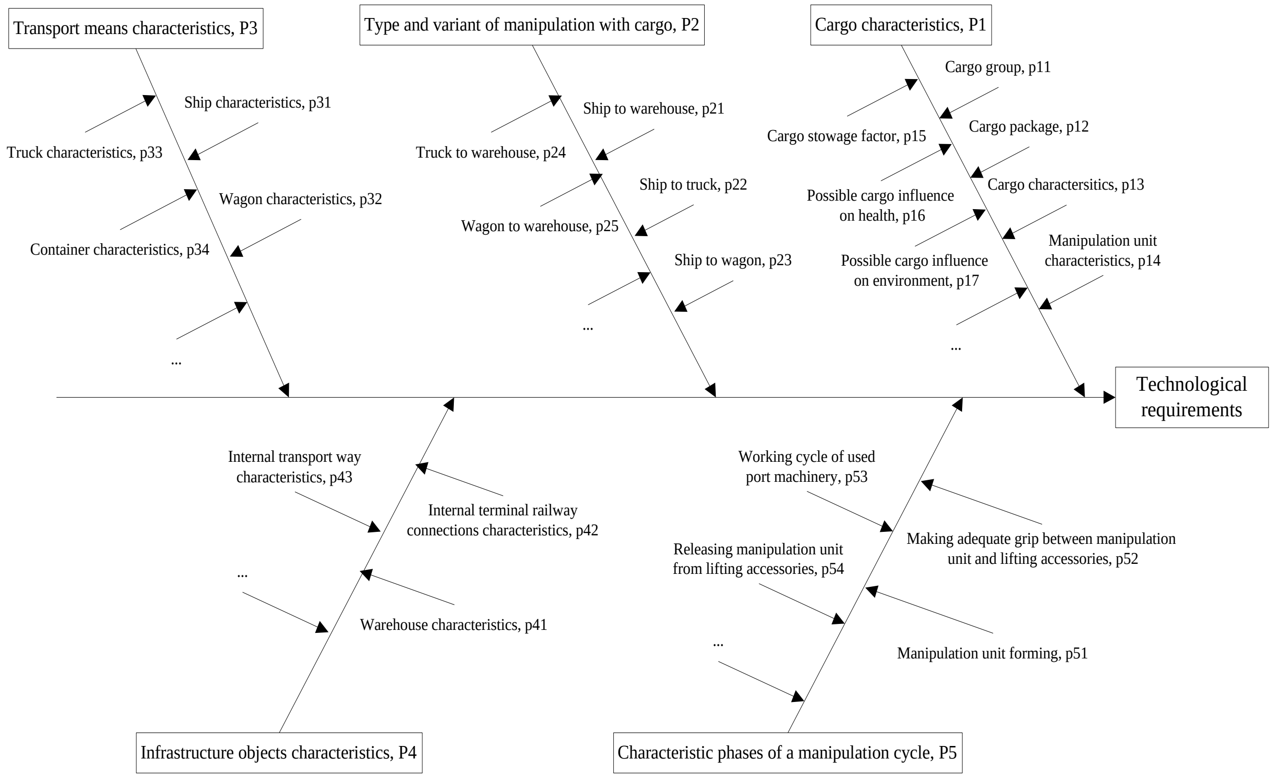



**Fig. 6.**
*Parameters which determine technological requirements in the cargo handling process*

Obviously, one from the group of the most important outputs from port equipment selection process have to be optimal relation between technological requirements generated in the cargo handling process and port equipment performances, considering all relevant influential factors. If previously mentioned is not fulfilled a lot of negative consequences on parameters which characterize port (terminal) functioning appear (cargo handling costs increasing, …).

## 5. Methodology for defining required level of port machinery flexibility

One of key consequences of the fact that multipurpose port terminals are characterized by very intensive variations in customers` requirements is necessity of having port machinery with a such performances which can match those variations. Possibility of satisfying previous demand is determined by adequate realization of port machinery selection process where defining optimal level of port machinery flexibility takes a very important place. Previously done condensed analyses strongly confirm that port machinery flexibility in a multipurpose port terminal has to be considered in an appropriate way. If not, numerous negative effects appear (increased costs of port machinery exploitation, increased costs of port machinery maintenance, etc.). Namely, an adequate methodology for defining required level of port machinery flexibility has to exist and to be implemented.

With Fig. 7. are given key elements of methodology for defining port machinery flexibility in a multipurpose port (terminal). Structure of given methodology is based on analyses done in previous paper segments. Elements of the methodology are given in a general form and all of them have complex structure and numerous components which are not elaborated in details in this paper. As per plan of the author`s further engagement in this field, further development of the methodology and completion of bases for created adequate software is planned.

Implementation of the methodology in a concrete seaport is connected with previous fulfillment of different preconditions (related on system organization, on competences of the staff, on adequacy level of the port information system, on availability of the required input data, etc.).

Key expected benefits from implementation of proposed methodology are related to increasing overall adequacy and efficiency of the port machinery selection process, to increased degree of accepted customers` demands, to optimization of costs generated in the cargo handing process, to establishing bases for improving market position of the port (contributing to wider assortments of port services offered to the customers), etc.

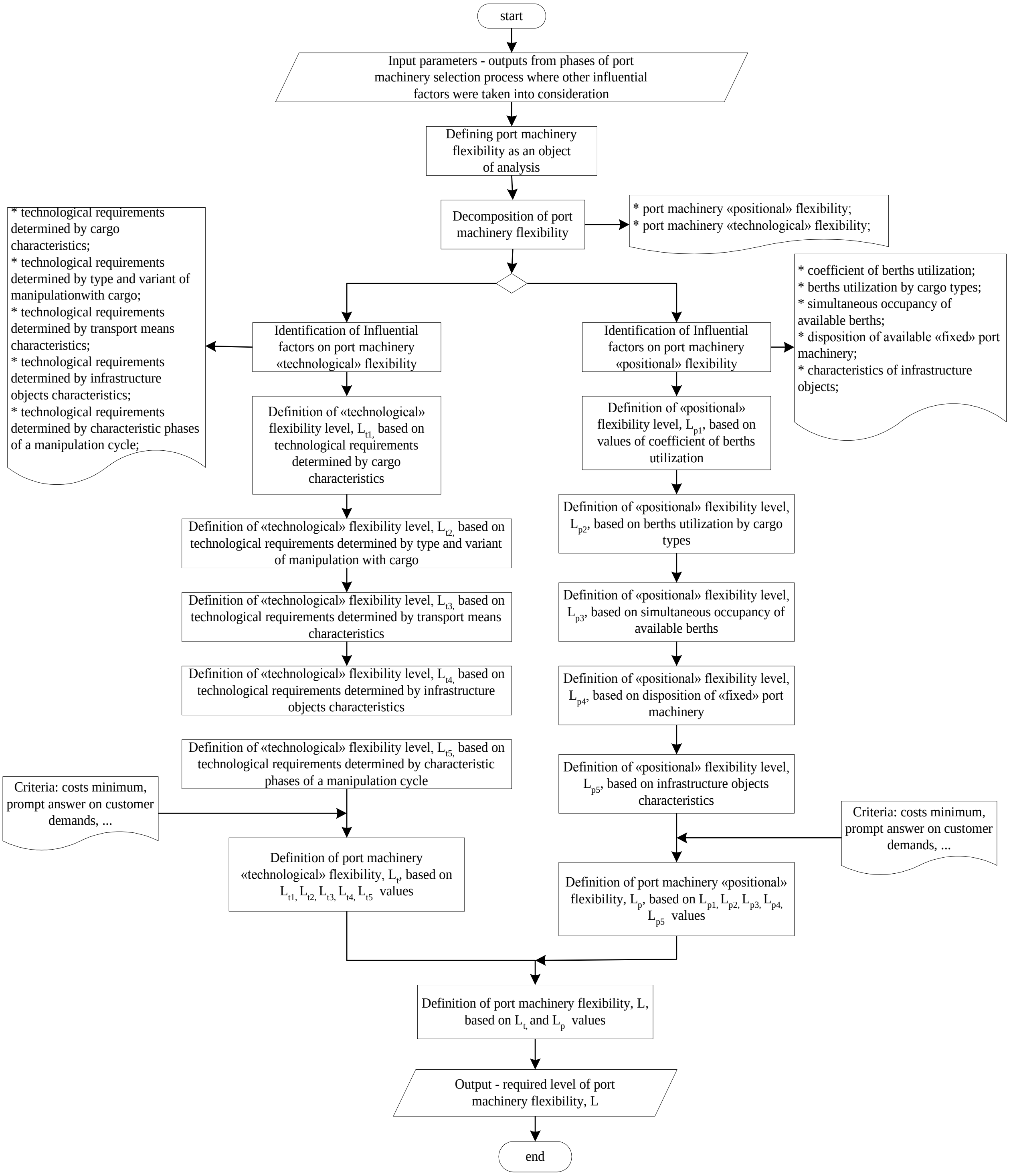



**Fig. 7.**
*Key elements of methodology of defining port machinery flexibility*

## 6. An example of flexible port machinery

A general outline of the mobile harbor crane, as an example of the flexible port machinery, is as follows (www.elit.terex.com). It consists of a traveling device with rubber-tired wheels for traveling freely on harbor areas, outrigger and a crane unit to load and unload cargo onto or from a ship. The traveling device of the crane consists mainly of a frame for traveling, an outrigger device, a steering device and an operator's cab. Rubber-tired wheels are used when the crane travels, and an outrigger is used to support the crane unit during cargo handling operations. The crane unit consists of a slewing device to swing around the crane unit, a luffing device to raise the jib and a hoisting device to lift cargo. Cargo handling operation is performed by the three different actions of the crane: slewing, luffing and hoisting.

Using mobile harbor cranes in a port is fully justified from different aspects. Its mobility (positional flexibility) and capability to be used in different handling operations (technological flexibility) are bases for achieving high utilization rates, especially in situation characterized by intensive variations in a port throughput structure (intensive variations of customers` demands). Mobile harbor cranes can be used on various berths within the port area, at different terminals etc. In other words, "a mobile harbor crane is going to the cargo" and not "waiting for cargo" (as classic shore cranes normally do) (Dundovic, 2002). In addition, possibilities to changes cranes lifting accessories (replacing a hook with a grab, etc.) in a shortest possible period of time are further advantages of these cranes.

Lifting and shifting cargoes are done in accordance with the mobile harbor crane load chart. Mobile harbor cranes are designed to work in different climate zones and there are no important specific elements related to local conditions. As well, there are no required prior civil construction (infrastructure) works (installing rails, etc.).

Mobile harbor cranes are widely used in the ports all over the world. For example, in the port of Antwerp (Belgium) (www.portofantwerp.com) are exploited more 100 mobile harbor cranes, in the port of Salerno (Italy) (www.porto.salerno.it), at its Container terminal, complete working technologies are based on mobile harbor cranes, in the regional ports (e.g. Koper – Slovenia (www.luka-kp.si), Ploce – Croatia (www.luka-ploce.hr), Durres – Albania (www.apdurres.com.al)) mobile harbor cranes are in use, as well.

Exploitation of the mobile harbor cranes is connected and with two important groups of limitations while a mobile harbor crane is under operation: limitations related on the area where a crane is positioned (port transport means, trucks and wagons can not have optimal transport route through the location where a mobile crane is positioned; difficulties in moving crane from one position to another in case of existing any kind of physical obstacles;) and limitations related to the allowed pressure on the quay construction. But, in the process of port machinery selection both limitation groups have to be taken into detailed consideration.

With the Fig. 8 is shown a mobile harbor crane positioned for work in a seaport.

**Fig. 8.**
Mobil*e harbor crane in a working position*
*Source: www.liebherr.com*

**7. Conclusion**

Based on the results of made considerations, it can be concluded that level of the port machinery flexibility is one of the key determining parameters of the overall port flexibility. Previous statement is directly connected with the high importance rank which the port machinery has in a port. Flexible port machinery is an answer on changing port environment and intensive changes in the customers demands over a period.

Two main components of the port machinery flexibility are defined: positional flexibility (capability of the machinery to work at different locations within port area) and technological flexibility (which consider that machinery exploitation performances are in line with character of wide spectrum of the technological requirements generated in the cargo handling process).

Optimization of the port machinery flexibility level considers implementation of a systematic approach and usage of an adequate methodology, which is partly (its basic elements) shown in this paper.

**References**

Adler, P.S.; Goldoftas, B.; Levine, D.I. 1999. Flexibility versus Efficiency? A Case Study of Model Changeovers in the Toyota Production System, *Journal of Organization Science*, 10(1): 43-68.

Asperen, V. E.; Dekker, R. 2009. *Flexibility in Port Selection: a Quantitative Approach using Floating Stocks*. Economic Institute of the Erasmus University Rotterdam, Report EI2009-44;

Das, T.K.; Elango, B. 1995. Managing Strategic Flexibility: Key to Effective Performance, *Journal of General Management*, 20(3): 60.

Djelovic, D. 2000. Factors which determine port machinery selection (in Serbian). *In Proceedings of the 28th Conference of the Production Engineering of Yugoslavia ( held in Serbia)*, 3.47.

Djelovic, D. 2004. Structure of a CHT Management Model and Principal Effects of Its Implementation. *In Proceedings of the IAME Annual Conference 2004 (held in Turkey)*, 618.

Dobner, M.; Rijsenbrij, J. 1999. Equipment Selection and Utilization – Key Factors for Smaller Terminals, *In Proceedings of AAPA Facilities Engineering Seminar (held in USA)*.

Dundovic, C. 2005. *Port Machinery for Cyclic Transport (in Croatian)*, University of Rijeka – Faculty od Maritime Studies.

Morlok, E.K.; Chang, D.J. 2004. Measuring capacity flexibility of a transportation system. Available from Internet: <http://www.seas.upenn.edu>.

Roach, T. 1987. *Operating and Maintenance Features of Container Handling Systems*, The World Bank.

Taneja, P.; Ligteringen, H.; Walker, H.E. 2012. Flexibility in Port Planning and Desing, *EJTIR Journal*, 12(1): 66-87.

* * * 2014. *Documentation of the Port of Bar*, Port of Bar – Development Department.

www.elit.terex.com.

www.portofantwerp.com.

www.porto.salerno.it.

www.luka-kp.si.

www.luka-ploce.hr.

www.apdurres.com.al.

www.liebherr.com.

# AN OPTIMIZATION MODEL BASED ON DATA ENVELOPMENT ANALYSIS: EMPIRICAL STUDY FROM THE POSTAL INDUSTRY

**Momčilo Dobrodolac[1], Predrag Ralević[2], Dejan Marković[1]**
*[1] University of Belgrade – Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, Belgrade 11000, Serbia*
*[2] PhD Candidate at University of Belgrade – Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, Belgrade 11000, Serbia*

**Abstract:** In this paper we proposed an optimization model for solving a combinatorial problem of limited resources. The optimization is performed based on technical efficiency measurement by using data envelopment analysis (DEA). The model provides two parallel processes. The first is related to the classification of decision making units (DMU) depending on the CCR and BCC scores. The second treats the stability of the returns to scale (RTS) classifications. The model has been tested and verified through a research carried out on the sample of 27 DMUs which were public postal operators (PPOs) from selected European states.

**Keywords**: optimization model, combinatorial problem, data envelopment analysis, public postal operators

## 1. Introduction

In the management science, an optimization problem is the problem of finding the best solution from all feasible solutions. Depending on whether the problem is described with the continuous or discrete variables, it can be solved by continuous or combinatorial optimizations. For this purpose deterministic models belonging to the field of mathematical programming are used. These models refer to the allocation of limited resources (organizational, technological, financial, etc.). They might be used in decision making, performance evaluations and a wide variety of activities in many contexts bringing a guideline about the best solutions.

There are two types of mathematical programming: linear programming and nonlinear programming. In linear programming the objective function is linear, and the constraints are presented with more linear equations and inequalities. The objective function and constraints in nonlinear programming can take a nonlinear form.
In this paper a model for solving a combinatorial optimization problem of limited resources is proposed. The optimization is performed by using data envelopment analysis (DEA) proposed by Charnes et al. (1978). It is a linear programming technique for evaluating the relative efficiency of multiple-input and multiple-output of a decision making units (DMUs) based on the production possibility set. A fundamental assumption behind this method is that if a DMU can produce a certain level of output utilizing specific input levels, other DMUs of equal scale should be capable of doing the same if they were to operate efficiently. Thus, the essence of the DEA analysis lies in finding the best DMU or DMUs which are observed as the most efficient under the given conditions, and are used to construct the efficient production frontier. The others that either make less outputs with the same inputs or make the same outputs with more inputs are inefficient. The degree of their inefficiency can be measured based on the distance from the frontier by using this method. A detailed review of theory, methodology and applications of DEA can be found in the book of Cooper et al. (2007).

The main characteristic of the model proposed in this paper is that the efficiency is considered as a comparison between observed and optimal output and input. Therefore, it is possible to compare observed output to maximum potential output obtainable by the input, or observed input to minimum potential input required to produce the output, or the combination of these two concepts. In the model the optimum is defined in terms of production possibilities and this type of efficiency is called technical. It is also possible to define the optimum in terms of financial targets where the efficiency is measured by comparing observed and optimum costs, revenue and profit. The optimum is expressed in value terms and the efficiency is economic.

The current article proceeds as follows: In Section 2, the optimization model is proposed. In Section 3, we applied the model from Section 2 on real data set of public postal operators (PPOs). The analytical results are presented and illustrated in Section 4. Finally, the conclusions are given in Section 5.

[1] Corresponding author: m.dobrodolac@sf.bg.ac.rs

## 2. The Proposed Model

The implementation of the proposed model for solving an optimization problem of limited resources should be realized through several phases as shown in Figure 1. First it needs to be defined and selected multiple-input and multiple-output of DMUs where an optimization of limited resources is required. In selecting the DMUs, we need to consider those DMUs which are comparable. In the paper by Cooper et al. (2007) there are some practical advices for selecting the DMUs suggested:

- Numerical data are available for each input and output, with the data assumed to be nonnegative for all DMUs.
- The items (inputs, outputs and choice of DMUs) should reflect an analyst's or a manager's interest in the components that will enter into the relative efficiency evaluations of the DMUs.
- In principle, smaller input amounts are preferable and larger output amounts are preferable so the efficiency scores should reflect these principles,
- The measurement units of the different inputs and outputs need not be congruent. Some may involve number of persons, or areas of floor space, money expended, etc.

The inputs and outputs are defined on the basis of experience, theory and practice in the given field and depend on the specific considered business. Also, in order to achieve the objectivity it is important that the values of the defined inputs and outputs are obtained from reliable sources and references and remain uniform for all units that are compared. A number of DMUs under consideration should be higher than the total number of inputs and outputs.

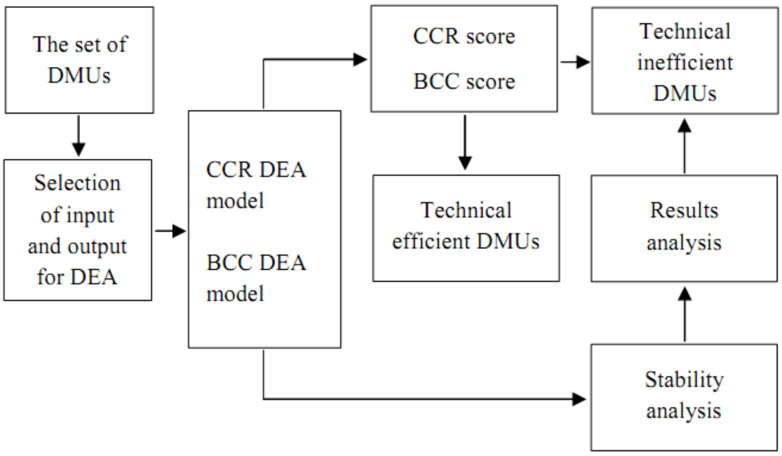



**Fig. 1.**
*The proposed model*

Further, the model provides two parallel processes. In the first process, DMUs are classified into technical efficient or technical inefficient units by using the adequate DEA models which are described in subsection 2.1. The second process of the model requires the stability analysis to be performed. This process is explained in subsections 2.2.

### 2.1 Efficiencies

The evaluation of relative efficiency for a group of DMUs is performed by using the original CCR DEA model proposed by Charnes et al. (1978). Suppose there are a set ( $j \in A$ ) uses $m$ inputs $x_{ij}$ ( $i = 1,2,3,\dots,m$ ) to produce $s$ outputs $y_{rj}$ ( $r = 1,2,3,\dots,s$ ). The CCR model evaluates the relative efficiency of a specific $DMU_o$, $o \in A$, with respect to a set of CCR frontier DMUs defined $E_o = \{ j \mid \lambda_j > 0$ for some optimal solutions for $DMU_o\}$. One formulation of a CCR model aims to minimize inputs while satisfying at least the given output levels, i.e., the CCR input-oriented model (see the M1 model). Another formulation of a CCR model aims to maximize outputs without requiring more of any of the observed input values, i.e., the CCR output-oriented model (see the M1' model). The CCR models assume the constant returns to scale production possibility set, i.e. it is postulated that the radial expansion and reduction of all observed DMUs and their nonnegative combinations are possible and hence the CCR score is called overall technical efficiency.

M1 model

$$\theta^* = \min \theta$$

Subject to: $\sum_{j \in E_o} \lambda_j x_{ij} \le \theta x_{io},\ i = 1,2,3,\ldots,m$

$\sum_{j \in E_o} \lambda_j y_{rj} \ge y_{ro},\ r = 1,2,3,\ldots,s$

$\lambda_j \ge 0,\ j \in E_o$

M1' model

$$\phi^* = \max \phi$$

Subject to: $\sum_{j \in E_o} \lambda_j x_{ij} \le x_{io},\ i = 1,2,3,\ldots,m$

$\sum_{j \in E_o} \lambda_j y_{rj} \ge \phi y_{ro},\ r = 1,2,3,\ldots,s$

$\lambda_j \ge 0,\ j \in E_o$

If $E_o = A$, then the M1 model is original form of the input-oriented CCR model. The $DMU_j$ ( $j \in E_o$ ) are called CCR efficient and form a specific CCR efficient aspect. These $DMU_j$ ( $j \in E_o$ ) appear in optimal solutions where $\lambda_j > 0$. The fact that $\lambda_j = 0$ for all $j \notin E_o$ in the M1 model when evaluating $DMU_o$ enables performing the CCR model in form of M1 model.

If we add $\sum_{j \in E_o} \lambda_j = 1$ in the M1 and M1' models, we obtain the BCC input-oriented and the BCC output-oriented models, respectively proposed by Banker et al. (1984). The BCC models assume that convex combinations of observed DMUs form the production possibility set and the BCC score is called local pure technical efficiency.

Seiford and Thrall (1990) derived the relationship between the solutions of the M1 model and M1' model. Suppose $\lambda_j^*$ ( $j \in E_o$ ) and $\theta^*$ is an optimal solution to M1 model. There exists a corresponding optimal solution $\lambda_j^{**}$ ( $j \in E_o$ ) and $\phi^*$ to the M1' model such that $\lambda_j^* = \dfrac{\lambda_j^{**}}{\phi^*}$ and $\phi^* = \dfrac{1}{\theta^*}$.

### 2.2 Stability

It is an important theme in DEA and was first examined by Seiford and Zhu (1999). The stability of RTS provides important information on the data perturbations in the DMU analysis. These information provide discussions that can be developed in performance analysis. This enables to determine the movement of inefficient DMUs on the frontier in improving directions. In this paper the authors developed several linear programming formulations for investigating the stability of RTS classification (constant, increasing or decreasing returns to scale). These authors considered data perturbations for inefficient DMUs. Charnes et al. (1985) indicated that sometimes a change in input or output or simultaneous changes in input and output are not possible. Jahanshahloo et al. (2005) developed an approach for the sensitivity analysis of efficient DMUs from the observations set.

The stability of the RTS classifications provides some stability intervals for preserving the RTS classification of a specific $DMU_o$. It enables to consider perturbations for all the inputs or outputs of $DMU_o$. Input-oriented stability of RTS classifications allows output perturbations in $DMU_o$, and output-oriented stability of RTS classifications enables input perturbations.

We can define lower and upper limit of stability intervals for inputs of DMUs based on the following theorems by Seiford and Zhu (1999):

**Theorem 1.** Suppose $DMU_o$ exhibits CRS. If $\gamma \in R^{CRS} = \left\{ \gamma : \min\left\{1, \mu_o^*\right\} \le \gamma \le \max\left\{1, \eta_o^*\right\} \right\}$. The CRS classification continues to hold, where $\gamma$ represents the proportional change of all inputs, $\hat{x}_{io} = \gamma x_{io}$ $(i = 1,2,3,\ldots,m)$, and $\eta_o^*$ and $\mu_o^*$ are defined in the M2 and M2' models, respectively.

**Theorem 2.** Suppose $DMU_o$ exhibits DRS. The DRS classification continues to hold for $\xi \in R^{DRS} = \left\{ \xi : \eta_o^* < \xi \le 1 \right\}$, where $\xi$ represents the proportional decrease of all inputs, $\hat{x}_{io} = \xi x_{io}$ $(i = 1,2,3,\ldots,m)$, and $\eta_o^*$ is defined in the M2 model.

**Theorem 3.** Suppose $DMU_o$ exhibits IRS. Then the IRS classification continues to hold for $\zeta \in R^{IRS} = \left\{ \zeta : 1 \le \zeta < \mu_o^* \right\}$, where $\zeta$ represents the proportional change of all inputs, $\hat{x}_{io} = \zeta x_{io}$ $(i = 1,2,3,\ldots,m)$, and $\mu_o^*$ is defined in the M2' model.

M2 model

$$\eta_o^* = \frac{1}{\min \sum_{j \in E_o} \lambda_j}$$

Subjec to $\sum_{j \in E_o} \lambda_j x_{ij} \le x_{io}, i = 1,2,3,\ldots,m$

$$\sum_{j \in E_o} \lambda_j y_{rj} \ge \phi^* y_{ro}, r = 1,2,3,\ldots,s$$

$$\lambda_j \ge 0, j \in E_o$$

M2' model

$$\mu_o^* = \frac{1}{\max \sum_{j \in E_o} \lambda_j}$$

Subjec to $\sum_{j \in E_o} \lambda_j x_{ij} \le x_{io}, i = 1,2,3,\ldots,m$

$$\sum_{j \in E_o} \lambda_j y_{rj} \ge \phi^* y_{ro}, r = 1,2,3,\ldots,s$$

$$\lambda_j \ge 0, j \in E_o$$

Lower and upper limit of stability intervals for outputs can be obtained by following theorems (Seiford and Zhu, 1999):

**Theorem 4.** Suppose $DMU_o$ exhibits CRS. If $\chi \in R^{CRS} = \{\chi : \min\{1, \mu_o^*\} \le \chi \le \max\{1, \eta_o^*\}\}$. The CRS classification continues to hold, where $\chi$ represents the proportional change of all outputs, $\hat{y}_{ro} = \chi y_{ro}$ $(r = 1,2,3,\ldots,s)$, and $\eta_o^*$ and $\mu_o^*$ are defined in the M3 and M3' models, respectively.

**Theorem 5.** Suppose $DMU_o$ exhibits DRS. The DRS classification continues to hold for $\beta \in R^{DRS} = \{\beta : \eta_o^* < \beta \le 1\}$, where $\beta$ represents the proportional change of all outputs, $\hat{y}_{ro} = \beta y_{ro}$ $(r = 1,2,3,\ldots,s)$, and $\eta_o^*$ is defined in the M3 model.

**Theorem 6.** Suppose $DMU_o$ exhibits IRS. Then the IRS classification continues to hold for $\alpha \in R^{IRS} = \{\alpha : 1 \le \alpha < \mu_o^*\}$, where $\alpha$ represents the proportional increase of all outputs, $\hat{y}_{ro} = \alpha y_{ro}$ $(r = 1,2,3,\ldots,s)$, and $\mu_o^*$ is defined in the M3' model.

M3 model

$$\eta_o^* = \frac{1}{\min \sum_{j \in E_o} \lambda_j}$$

Subjec to $\sum_{j \in E_o} \lambda_j x_{ij} \le \theta^* x_{io}, i = 1,2,3,\ldots,m$

$$\sum_{j \in E_o} \lambda_j y_{rj} \ge y_{ro}, r = 1,2,3,\ldots,s$$

$$\lambda_j \ge 0, j \in E_o$$

M3' model

$$\mu_o^* = \frac{1}{\max \sum_{j \in E_o} \lambda_j}$$

Subjec to $\sum_{j \in E_o} \lambda_j x_{ij} \le \theta^* x_{io}, i = 1,2,3,\ldots,m$

$$\sum_{j \in E_o} \lambda_j y_{rj} \ge y_{ro}, r = 1,2,3,\ldots,s$$

$$\lambda_j \ge 0, j \in E_o$$

### 3. Sample

In this study we implemented the proposed model on the sample of 27 DMUs. The observed DMUs are public postal operators (PPOs) in the countries of European Union and the PPO in Serbia (Table 1). The main sources of data for the year 2011 were the Universal Postal Union (UPU) and the European Commission. The PPO in Serbia as a candidate country was included in the observed production possibility set consisting of PPOs in European Union member states.

**Table 1**
*Sample*

| Universal service provider | Country |
|---|---|
| Österreichische Post AG | Austria |
| Bulgarian Posts | Bulgaria |
| Cyprus Post | Cyprus |
| Česká pošta | Czech Republic |
| Post Denmark | Denmark |
| Eesti Post | Estonia |
| Itella Posti Oy | Finland |
| La poste | France |
| Deutsche Post | Germany |
| Royal Mail Group plc | Great Britain |
| Hellenic Post ELTA | Greece |
| Magyar Posta | Hungary |
| An Post | Ireland |
| Poste Italiane | Italy |
| Latvijas Pasts | Latvia |
| Lietuvos pastas | Lithuania |
| P & T Luxembourg | Luxembourg |
| Malta Post | Malta |
| PostNL | Netherlands |
| Poczta Polska | Poland |
| CTT – Correios | Portugal |
| Posta Romana | Romania |
| Slovenská pošta | Slovakia |
| Pošta Slovenije d.o.o | Slovenia |
| Correos y Telégrafos | Spain |
| Posten Sweden Post | Sweden |

In the case of universal service provider, we employed 3 inputs and 2 outputs (see Figure 2). The number of full-time staff ( $x_1$ ), the number of part-time staff ( $x_2$ ) and total number of permanent post offices ( $x_3$ ) are the inputs while the number of letter-post items ( $y_1$ ) and parcels and post express ( $y_2$ ) are the outputs. Input and output data are listed in Table 2.

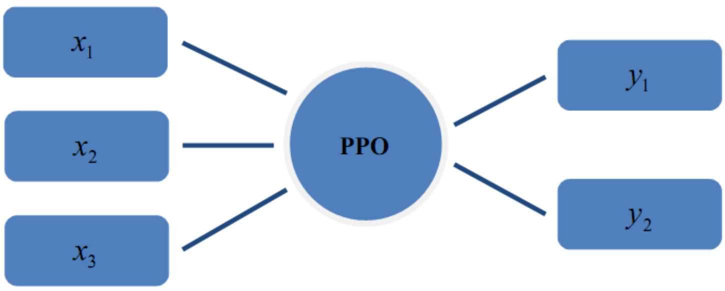



**Fig. 2.**
*The inputs and outputs of public postal operator*

**Table 2**
*Data*

| PPO Name | Full-time staff | Part-time staff | Post offices | Letters | Parcels and post express |
|---|---|---|---|---|---|
| Austria | 17233 | 3882 | 1880 | 6215000000 | 178366838 |
| Bulgaria | 8689 | 3796 | 2981 | 19159655 | 1256059 |
| Cyprus | 714 | 1034 | 1082 | 58787116 | 2398 |
| Czech Republic | 28232 | 8020 | 3408 | 2574778260 | 2338309 |
| Denmark | 12800 | 6200 | 795 | 800000000 | 30049000 |
| Estonia | 2290 | 502 | 343 | 25837400 | 1458500 |
| Finland | 20077 | 7508 | 978 | 837000000 | 28500000 |
| France | 204387 | 25900 | 17054 | 14900000000 | 266000000 |
| Germany | 512147 | 0 | 13000 | 19784000000 | 870000000 |
| Great Britain | 117206 | 38558 | 11818 | 18074291171 | 423000000 |
| Greece | 9060 | 28 | 1546 | 446505500 | 2743000 |
| Hungary | 28592 | 5368 | 2746 | 857056665 | 8890689 |
| Ireland | 7825 | 1584 | 1156 | 614320000 | 8067188 |
| Italy | 133426 | 11025 | 13923 | 4934317901 | 692761 |
| Latvia | 2438 | 2055 | 571 | 28886614 | 1370420 |
| Lithuania | 2336 | 4226 | 715 | 36599075 | 725000 |
| Luxembourg | 950 | 547 | 116 | 110800000 | 23294 |
| Malta | 490 | 123 | 63 | 35123154 | 106000000 |
| Netherlands | 13141 | 46590 | 2600 | 3777000000 | 20856767 |
| Poland | 77548 | 16534 | 8207 | 822176000 | 226900 |
| Portugal | 11608 | 315 | 2556 | 868548000 | 10552892 |
| Romania | 32630 | 1319 | 5827 | 292635204 | 3948360 |
| Slovakia | 9650 | 5081 | 1589 | 425743495 | 288002 |
| Slovenia | 6344 | 161 | 556 | 1013027273 | 661857173 |
| Spain | 65924 | 0 | 3183 | 5123200000 | 70614713 |
| Sweden | 19222 | 2918 | 1924 | 2231000000 | 3802000 |
| Serbia | 14659 | 280 | 1507 | 243130583 | 178366838 |

## 4. Results

By using the M1 or M1' model we evaluated the relative efficiency for a group of 27 PPOs. The CCR, BCC and scale characteristics of each PPO are listed in Table 3.

The results from Table 3 show that there are three PPOs which have the CCR score equal to 1. This score indicates overall technical efficiency when evaluated on the constant returns to scale assumption. These are PPOs in Austria, Slovenia and Spain. PPO in Austria is one of three best performers, and furthermore it is the PPO most frequently referenced for evaluating inefficient PPOs.

The BCC score provide efficiency evaluations using a local measure of scale, i.e. under variable returns to scale. In this empirical example four PPOs are accorded efficient status in addition to the three CCR efficient PPOs which retain their previous efficient status. These four PPOs are in Germany, Great Britain, Greece and Malta. For example, it can be concluded that PPO in Greece has the efficient operations ( $\theta^*_{BCC} = 1$ ). Additionally, it can be considered that all PPOs having the BCC score above average (0.51441) have the efficient operations. These are PPOs in Austria, Cyprus, France, Germany, Great Britain, Greece, Luxembourg, Malta, Netherlands, Slovenia and Spain.

**Table 3**
*Efficiencies*

| PPO No. | PPO Name | BCC | CCR | | Scale Score |
|---|---|---|---|---|---|
| | | Score | Score | Reference | |
| $PPO_1$ | Austria | 1.00000 | 1.00000 | | 1.00000 |
| $PPO_2$ | Bulgaria | 0.05639 | 0.00611 | $PPO_1$ | 0.10842 |
| $PPO_3$ | Cyprus | 0.77607 | 0.22830 | $PPO_1$ | 0.29417 |
| $PPO_4$ | Czech Rep. | 0.26107 | 0.25288 | $PPO_1$ | 0.96862 |
| $PPO_5$ | Denmark | 0.36212 | 0.30440 | $PPO_1$ | 0.84059 |
| $PPO_6$ | Estonia | 0.24353 | 0.03182 | $PPO_1$, $PPO_{24}$ | 0.13067 |
| $PPO_7$ | Finland | 0.30549 | 0.25888 | $PPO_1$ | 0.84744 |
| $PPO_8$ | France | 0.80392 | 0.30588 | $PPO_1$, $PPO_{24}$, $PPO_{25}$ | 0.38048 |
| $PPO_9$ | Germany | 1.00000 | 0.94551 | $PPO_{25}$ | 0.94551 |
| $PPO_{10}$ | Great Britain | 1.00000 | 0.46263 | $PPO_1$ | 0.46263 |
| $PPO_{11}$ | Greece | 1.00000 | 0.56198 | $PPO_{24}$, $PPO_{25}$ | 0.56198 |
| $PPO_{12}$ | Hungary | 0.11309 | 0.09707 | $PPO_1$, $PPO_{25}$ | 0.85834 |
| $PPO_{13}$ | Ireland | 0.28302 | 0.23249 | $PPO_1$, $PPO_{24}$ | 0.82146 |
| $PPO_{14}$ | Italy | 0.20119 | 0.17023 | $PPO_1$, $PPO_{24}$ | 0.84611 |
| $PPO_{15}$ | Latvia | 0.20098 | 0.03285 | $PPO_1$ | 0.16346 |
| $PPO_{16}$ | Lithuania | 0.21147 | 0.04344 | $PPO_1$ | 0.20543 |
| $PPO_{17}$ | Luxembourg | 0.73492 | 0.32340 | $PPO_1$ | 0.44005 |
| $PPO_{18}$ | Malta | 1.00000 | 0.19875 | $PPO_1$ | 0.19875 |
| $PPO_{19}$ | Netherlands | 0.80875 | 0.79696 | $PPO_1$ | 0.98543 |
| $PPO_{20}$ | Poland | 0.03610 | 0.03069 | $PPO_1$, $PPO_{25}$ | 0.85018 |
| $PPO_{21}$ | Portugal | 0.48475 | 0.46344 | $PPO_1$, $PPO_{24}$ | 0.95604 |
| $PPO_{22}$ | Romania | 0.08887 | 0.05130 | $PPO_1$, $PPO_{24}$ | 0.57726 |
| $PPO_{23}$ | Slovakia | 0.16045 | 0.12233 | $PPO_1$ | 0.76245 |
| $PPO_{24}$ | Slovenia | 1.00000 | 1.00000 | | 1.00000 |
| $PPO_{25}$ | Spain | 1.00000 | 1.00000 | | 1.00000 |
| $PPO_{26}$ | Sweden | 0.41553 | 0.40714 | $PPO_1$, $PPO_{24}$, $PPO_{25}$ | 0.97983 |
| $PPO_{27}$ | Serbia | 0.34138 | 0.11897 | $PPO_{24}$, $PPO_{25}$ | 0.34851 |
| **Average** | | **0.51441** | **0.34991** | | **0.64940** |

Based on the results of scale scores from Table 3 the following PPOs operate in the advantageous conditions: Austria, Czech Republic, Denmark, Finland, Germany, Hungary, Ireland, Italy, Netherlands, Poland, Portugal, Slovakia, Slovenia, Spain and Sweden. Their scale scores are higher than average value (0.64940). Some of them although working in the advantageous conditions have the inefficient operations. We can notice the examples of PPOs in Czech Republic, Poland and Portugal. There are the opposite cases where PPOs work in the disadvantageous conditions but their operations are above average, for example PPOs in Cyprus and Luxembourg. Further there are PPOs operating in the disadvantageous conditions and having the inefficient operations such as PPOs in Bulgaria, Estonia, Latvia, Lithuania, Romania, Slovakia and Serbia.

The results of the stability analysis of observed PPOs are shown in Table 4. These results indicate that PPOs in Austria, Slovenia and Spain do not need input perturbations. This could be expected because these PPOs are in the most productive scale size (MPSS). PPOs in Czech Republic, France, Germany, Great Britain, Hungary, Italy, Poland, Portugal, Romania, Sweden and Serbia should consider decreasing inputs. PPOs in Bulgaria, Cyprus, Denmark, Estonia, Finland, Greece, Ireland, Latvia, Lithuania, Luxembourg, Malta, Netherlands and Slovakia should consider increasing inputs. Considering the results from Table 4, PPOs that need to perform input perturbations can be divided in three groups.

The first group contains of PPOs with the input excess and the output deficit (the lower limit of stability interval for inputs is less than 1 while the upper limit of stability interval for outputs is higher than 1). These PPOs are in Czech Republic, Hungary, Poland, Portugal, Romania, Sweden and Serbia. For example, based on results from Table 4, PPO in Serbia has the input excess of around 45% (0.55737).. On the other hand, in this case we note the output deficit. The outputs could be increased more than four times (4.68485) from the current value. This optimization can refer to one or more observed inputs as well as one or more outputs. The future research could be related to determination of particular input or output to be changed.

In the second group there are PPOs having the input excess (the lower limit of stability interval for inputs is less than 1 while the upper limit of stability interval for outputs is equal to 1). This means they could achieve the current output level with less inputs. The examples of this kind of PPOs are in France, Germany, Great Britain and Italy.

The rest of PPOs are in the third group. The main characteristic of these PPOs is the possibility of increasing output by increased inputs (the upper limit of stability interval for inputs and the upper limit of stability interval for outputs are higher than 1). These PPOs are in Bulgaria, Cyprus, Denmark, Estonia, Finland, Greece, Ireland, Latvia, Lithuania, Luxembourg, Malta, Netherlands and Slovakia.

**Table 4**
*Stability*

| PPO No. | PPO Name | Stability interval for inputs | Stability interval for outputs |
|---|---|---|---|
| $PPO_1$ | Austria | (1.00000, 1.00000) | (1.00000, 1.00000) |
| $PPO_2$ | Bulgaria | (1.00000, 1.98331) | (1.00000, 324.37954) |
| $PPO_3$ | Cyprus | (1.00000, 24.13585) | (1.00000, 105.72044) |
| $PPO_4$ | Czech Rep. | (0.61041, 1.00000) | (1.00000, 2.41380) |
| $PPO_5$ | Denmark | (1.00000, 2.36478) | (1.00000, 7.76875) |
| $PPO_6$ | Estonia | (1.00000, 7.15360) | (1.00000, 224.80315) |
| $PPO_7$ | Finland | (1.00000, 1.92229) | (1.00000, 7.42533) |
| $PPO_8$ | France | (0.11300, 1.00000) | (0.36945, 1.00000) |
| $PPO_9$ | Germany | (0.24485, 1.00000) | (0.25896, 1.00000) |
| $PPO_{10}$ | Great Britain | (0.15908, 1.00000) | (0.34386, 1.00000) |
| $PPO_{11}$ | Greece | (1.00000, 3.39434) | (1.00000, 6.03998) |
| $PPO_{12}$ | Hungary | (0.69990, 1.00000) | (1.00000, 7.21057) |
| $PPO_{13}$ | Ireland | (1.00000, 1.84123) | (1.00000, 7.91971) |
| $PPO_{14}$ | Italy | (0.05805, 1.00000) | (0.34103, 1.00000) |
| $PPO_{15}$ | Latvia | (1.00000, 7.06850) | (1.00000, 215.15156) |
| $PPO_{16}$ | Lithuania | (1.00000, 7.37714) | (1.00000, 169.81303) |
| $PPO_{17}$ | Luxembourg | (1.00000, 18.14000) | (1.00000, 56.09206) |
| $PPO_{18}$ | Malta | (1.00000, 35.16939) | (1.00000, 176.94880) |
| $PPO_{19}$ | Netherlands | (1.00000, 1.31139) | (1.00000, 1.64549) |
| $PPO_{20}$ | Poland | (0.23138, 1.00000) | (1.00000, 7.53992) |
| $PPO_{21}$ | Portugal | (0.54957, 1.00000) | (1.00000, 1.18585) |
| $PPO_{22}$ | Romania | (0.20413, 1.00000) | (1.00000, 3.97887) |
| $PPO_{23}$ | Slovakia | (1.00000, 1.78580) | (1.00000, 14.59799) |
| $PPO_{24}$ | Slovenia | (1.00000, 1.00000) | (1.00000, 1.00000) |
| $PPO_{25}$ | Spain | (1.00000, 1.00000) | (1.00000, 1.00000) |
| $PPO_{26}$ | Sweden | (0.57863, 1.00000) | (1.00000, 1.42120) |
| $PPO_{27}$ | Serbia | (0.55737, 1.00000) | (1.00000, 4.68485) |

The obtained optimizations should be considered conditionally having in mind the public expectations about the postal systems, first of all the obligation to provide postal services on the whole territory of a state. Thus, in order to implement the proposed model further research should be performed for each specific country considering the legal limitations.

## 5. Conclusion

The optimization of limited resources is a real problem vary common in many disciplines. In this paper we proposed the model for solving one class of optimization problems which are concerned with combinatorial optimization. For this purpose we used data envelopment analysis (DEA).

The model has been applied on the sample of 27 public postal operators (PPOs) and it focused on individual characteristics of each unit. The optimization is performed through all considered PPOs. Our analytical research has shown that more than 85% observed PPOs have a possibility to improve their performance by scaling up their activities, while less than 15% PPOs operate in the most productive scale size.

The model proposed in this paper, except in the postal sector, could be implemented in other systems as well. It requires operational processes modelling in terms of performance evaluations for each particular optimization problem. These settings can not affect the basic idea and character of the model. Therefore, the model could be applied to various considerations related to limited resource allocation.

## Acknowledgement

This research was supported by Serbian Ministry of Education, Science and Technological Development with project TR 36022.

**References**

Banker, R. D.; Charnes, A.; Cooper, W.W. 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management Science*, 30(9): 1078-1092.

Charnes, A.; Cooper, W.W.; Golany, B.; Seiford, L.M.; Stutz, J. 1985. Foundation of data envelopment analysis for Pareto-Koopmans efficient empirical production functions, *Journal of Econometrics*, 30(1-2): 91-107.

Charnes, A.; Cooper, W.W.; Rhodes, E. 1978. Measuring the efficiency of decision making units, *European Journal of Operational Research*, 2(6): 429-444.

Cooper, W.W.; Seiford, L.M.; Tone, K. 2007. *Data envelopment analysis: A comprehensive text with models, applications, references and DEA-Solver software*, 2nd ed., Springer, New York, USA.

Jahanshahloo, G. R.; Hosseinzadeh Lotfi, F.; Shoja, N.; Tohidi, G.; Razavian, S. 2005. A one-model approach to classification and sensitivity analysis in DEA, *Applied Mathematics and Computation*, 169(2): 887-896.

Seiford, L.M.; Zhu, J. 1999. Sensitivity and stability of the classification of returns to scale in data envelopment analysis, *Journal of Productivity Analysis*, 12(1): 55-75.

Seiford, L.M.; Thrall, R.M. 1990. Recent developments in DEA: The mathematical programming approach to frontier analysis, *Journal of Econometrics*, 46(1-2): 7-38.

# TRAVEL CHAIN BASED URBAN MOBILITY

**Lajos KISGYÖRGY**[1], **Gergely VASVÁRI**[2]
[1] *Budapest University of Technology and Economics, Department of Highway and Railway Engineering*
[2] *Budapest University of Technology and Economics, Department of Highway and Railway Engineering*

**Abstract:** One of the basic strategies of urban transportation is the promotion of public transportation modes. However numerous travels can not be served by public transportation because there is no adequate spatial or temporal cover of service or the travel chain is to complex.
To solve this problem the design paradigm is needed: instead of vehicles the focus should be on the mobility of persons and goods, which requires attractive and competitive travel chains from the origin of the trip to the end. It means travel chains which are quick and comfortable. Travel time is the sum of the travel times on elements of the travel chain, so improving one element the decrease of travel time will appear int he total travel time as well. Comfort however is judged by the worst parts, so improving one lement could be totally ineffective from the point of view of the whole travel.
This study focuses on a strategic approach based on travel chains. It shows a method to optimize the transport system according the aspects of the user, to shift efficiently from individual vehicle usage to more sustainable transportation modes.

**Keywords**: travel chains, sustainable mobility, urban mobility.

## 1. Introduction

The urban space is composed by a complex network of systems that interact in an integrated fashion, therefore influencing one another. The traditional urban planning process, however, is quite often concentrated in just a few parts of the systems (Miranda, da Silva, 2012). This is also the case of transport planning (Litman, 1999), where conventional planning tends to assume that transport progress is linear, consisting of newer, faster modes that displace older, slower modes.

In the last years the economic and production conditions caused a suburbanization ('sprawl') process and a widespread use of private car for travelling. The dispersion of urban functions within vast areas requires a great diffusion of private vehicles capable of guaranteeing mobility inside contexts which are so dispersed that they can't be served by conventional public transport systems with economical convenience, and congestion has saturated city centres and suburbs (Migliore at al., 2011). Traffic related problems in the central areas require multimodal platforms and a modal shift from individual car usage towards more sustainable transportation modes. One of the key elements of this integrated transport system is the transfer from passenger cars to public transportation. Rebalancing relations between transportation modes is in the center of the strategy for the European Community's sustainable development (European Commision, 2003).

Sustainable development requires significant changes in existing transportation systems for increasing economic efficiency, equity and environmental security (Litman, 1999). Sustainable urban mobility is represented by the displacement of persons and goods in the urban space (Gudmundsson, 2004). Individual transport decisions must be subordinate to community's long-term strategic objectives. Planning sustainable transport requires fundamental changes in the existing transportation planning practices. Restrictive policies, however could be met with resistance or be politically costly, so measures encouraging behavioural change towards the required direction are necessary. Behavioural Economics provides a model for policy measures beyond the traditional instruments. It focuses on individual choice and the factors that motivate people to do things (Dawnay and Shah, 2005). According to this idea small interventions can be designed to help people overcome cognitive bias and make the "right" decisions automatically. These soft measures are already being applied in other sectors, but their application in transport is still rather limited. (Economides et al. 2012). The main target of behavioural interventions in transport is mode shift from private car use to other modes of transport, especially public transport (Bamberg et al., 2011).

The state of the system and the impacts of interventions can be measured precisely through indicators. According to Maclaren (1996), indicators are simplifications of complex phenomena and they often provide only an indication of the condition or state of a given element. Thus a better picture can be obtained by the combination of indicators, in order to capture the different dimensions and aspects of any particular problem (Gudmundsson, 2004). The evaluation index should reflect the main and the comprehensive information. And each index should be independent, and easy to measure (Guo, 2009).

In theory the transportation mode is chosen based on the Level of Service (LOS) of the given mode. There are many studies regarding the relationship between LOS and modal split, and some of them focus on the LOS determination mechanism and identify the effective policy measures for improvement of user's benefit. Although those studies describe the mechanism theoretically and find out the important implications in transport policy formulations, they

[1] Corresponding author: email (kisgyorgy@uvt.bme.hu)

make the actual transport operation too simplified to estimate the policy effects (Kii et al. 2005). The evaluation systems of public transportation use quantitative and qualitative indicators and generally the analytic hierarchy process is used (Guo, 2009). However integrated analysis along a travel chain was performed only for accessibility (Tsalis and Naniopoulos, 2012).

## 2. The importance of travel chains

A new bus lane was created on the Budaörsi street, one of the most loaded artery of Budapest in 2011. The intended goal was to improve the LOS and travel time of the buses on the route, making people shift to public transportation from individual passenger car usage. The project failed miserably, gigantic congestions were created, several ten thousands of people raged in their cars, and the media loved the theme. After two weeks the mayor of Budapest retreated and the bus lane was cancelled.

Clearly serious mistakes were done but the application of the bus lane was not that unprofessional as the events show. This transport corridor is mainly used by people commuting to Budapest from the western suburban area. There are several railroad lines with good service parallel to the major road direction. The Kelenföldi railway station, which is a great intermodal node with good connections to the downtown of Budapest and to the destination of the vast majority of commuters, was located just before the beginning of the new bus lane. Still people chose to wait hours in their cars and get mad. They did not even think on the possibility of modal shift.

This little story shows an example where the modal shift was not realized, and it has several lessons. From the point of you of the paper it shows that improving public transportation service in itself not enough to promote modal shift, something more is needed. The example clearly shows that existing planning practice requires fundamental changes; our current tools are not adequate.

The main goal of urban transport policies is to provide mobility with minimizing the harmful effects of transportation. However transport planning is usually focused on specific problems or on broad transport concerns at a local level. And because it tends to deal with localized problems, the solutions adopted cannot achieve the required goals in the global scale.

The bus lane project failed mainly because there was no chance for the drivers to shift to public transportation locally. There were no P+R facilities at the intermodal node. So when the drivers faced the problem they could not shift. But they did not shift on the following days either, rather spent hours in a gigantic congestion. And this shows that there are more serious underlying factors worth to investigate.

If we consider the corridor itself, it shows that there are extremely good public service providing a comfortable access to the destinations. So it is reasonable and advantageous for the drivers to shift to public transportation in a local sense. The reason that the drivers stubbornly continued to use their cars even in extremely unfavorable conditions is that the access of railway stations in the suburban areas is very uncomfortable and time consuming, due to the lack of P+R and B+R facilities and the poor bus services. This fact demonstrates the importance of thinking in travel chains. If the whole travel chain of commuters had been analyzed it would have showed that the modal shift in the given corridor could have been realized by improving the access to the railway lines (and the frequency of railway service), in addition to the bus lane.

Mobility demands can be projected by travel chains. This makes it possible to cluster the mobility demands based on fi. spatial characteristic, temporal characteristic, transportation mode or other features. Inside a cluster the optimal intermodal travel chain can be designed – optimal from the point of view of traveler and of sustainable transport system. Inside these clusters the soft measures of Behavioral Economics can be applied effectively. Travel chain based approach is very useful at assigning road surface to transportation modes as well. The available surface should be rationally divided among the clusters using it – according to their optimal travel chains and their optimal transportation modes. And that is a huge improvement because it makes possible to design locally based on global impacts.

## 3. Evaluation of travel chains

Travel chains should be evaluated from the point of view of travelers. The travelers want to get from their origin to their destination quickly and comfortably, namely they want to use travel chains which are quick and comfortable all along. This definition gives us the indicators which should be used at evaluation: speed and comfort. Theoretically these measures can be determined relatively easily. However the problem is that speed and comfort are perceived in different ways, and travelers has different expectations as well. For example a commuting industrial worker has significantly different expectation about these concepts than a mother travelling with three children to arrange daily duties or the bank director. So the evaluation of travel chains should be done always according to the expectations of a given traveler group.

The choice of an alternative is assumed to be the result of that alternative's utility being higher than its competitors in the perception space of the decision-maker. The utility of a transportation alternative itself is typically mapped to observed characteristics of the decision-maker, such as the socio-demographics of an individual, and observed characteristics of the alternatives, such as travel time and travel costs by alternative modes (Bhat and Dubey, 2014). In this paper we use travel time and travel conditions to describe the utility of alternative travel chains.

**Travel time**

Speed is well described by travel time. Travel time includes the necessary walking time to access transportation modes, the waiting time, the time of travel and transfers. But there are challenges, because the travel time is not constant, depends on disturbances and congestions. The reliability of the travel relates to the handling of these uncertainties. And the reliability of the travel time is very important to the traveler if they want to reach their destination in time. It shows the significance of reliability that very often travelers choose somewhat slower but more reliable transportation modes. According to these for travelers speed means when they have to start their trip to arrive at their destination in time.

The travel time of certain transportation modes are very much analyzed. The traditional, mode-based interventions are focusing to the improvement of these individual travel times. However the transfer time, the time needed to transfer from one transportation vehicle to another is also very important from the point of view of the traveler. The transfer time can be divided into two parts: the time needed to reach the access point of the next vehicle, and the waiting time.

The Travel Time Indicator, defined as the travel time along the travel chain, can be calculated as the sum of the expected travel times of individual transportation modes and transfers along the chain:

$$TTI = \sum_{i=1}^{2} wt_i + \sum_{i=1}^{n} tt_i + \sum_{i=1}^{n-1} trt_{i,i+1} \tag{1}$$

where

| | |
|---|---|
| *TTI*: | sum of expected travel times along the travel chain, called Travel Time Indicator, |
| $wt_1$ and $wt_2$: | the expected value of time spent by walking at the beginning and at the end of the travel, |
| $tt_i$: | expected travel time of transportation mode or service *i*, |
| $trt_{i,i+1}$: | expected transfer time between transportation mode *i* and *i+1*. |

Both travel times and transfer times have uncertainty. Congestions or queues at ticket shops can cause significant delays. If we want to handle these uncertainties than we have to express the travel time and the transfer time with a certain (f. i. 95% or 99%) confidence level, to express the relevant travel time, say the (Travel Time Reliability Indicator, TTRI):

$$TTRI^{CL} = \sum_{i=1}^{2} wt_i^{CL} + \sum_{i=1}^{n} tt_i^{CL} + \sum_{i=1}^{n-1} trt_{i,i+1}^{CL} \tag{2}$$

where

| | |
|---|---|
| *CL:* | confidence level, 95% or 99%, |
| *TTRI*: | Travel Time Reliability Indicator, sum of travel times with *CL* confidence level, |
| $wt_1^{CL}$ and $wt_2^{CL}$: | the time spent by walking at the beginning and at the end of the travel, with *CL* confidence level, |
| $tt_i^{CL}$: | travel time of transportation mode or service *i*, with *CL* confidence level |
| $trt_{i,i+1}^{CL}$: | transfer time between transportation mode *i* and *i+1*, with *CL* confidence level. |

We used a 3-point scale (with categories "Good", "Acceptable" and "Unacceptable") to categorize TTI and TTRI. Only alternatives whose TTI and TTRI values belong to the "Good" or "Acceptable" categories can be considered as viable alternatives.

It is possible to monetize the travel time indices to include travel costs. In this case the TTI or TTRI value should be multiplied by the value of time, and the travel costs are simply added to get the monetized index.

**Travel conditions**

Qualitative factors such as travel convenience, comfort and security affect the competitiveness of travel alternatives in a great extent. Unfortunately conventional planning practices tend to overlook and undervalue service quality impacts (Litman, 2008). The *individual* specific, not mode specific, latent variables of environmental preferences, safety, security, comfort, convenience and flexibility significantly influence travel mode choice. The construction of safety, security and environmental preference variables is based on behavioural indicator variables, the construction of the comfort, convenience and flexibility variables is based on attitudinal indicator variables. Thus the explanatory power of constructions based on either type can be compared (Johansson et al. 2005). In our paper we refer to these individual specific latent variables collectively as travel comfort.

Several models were developed to describe the soft psychometric measures associated with individual perceptions (Bath and Dubey, 2014), and based on the monetized value or the utility value of the model results a Comfort Index (CI) can be constructed. The CI describes the perceived quality of the transportation service and conditions. In our paper we assume that such a consistent model to determine the CI for the elements of the travel chain exists (or will be developed).

The travel comfort along the travel chain can be defined as the weighted *average* of the CI of the components in the travel chain.

In our model we weight the individual CIs with their expected travel time, respectively:

$$TCI = \frac{1}{TTI} \cdot \left( \sum_{i=1}^{2} wt_i \cdot CI_{w,i} + \sum_{i=1}^{n} tt_i \cdot CI_i + \sum_{i=1}^{n-1} trt_{ij} \cdot CI_{i,i+1} \right) \tag{3}$$

where

| | |
|---|---|
| *TCI:* | Comfort Index along the travel chain, |
| *TTI*: | sum of expected travel times along the travel chain, called Travel Time Indicator, |
| $wt_1$ and $wt_2$: | the expected value of time spent by walking at the beginning and at the end of the travel, |
| $CI_{w,i}$: | the Comfort Index of walking at the beginning and at the end of the travel, |
| $tt_i$: | expected travel time of transportation mode or service *i*, |
| $CI_i$: | the Comfort Index of transportation mode or service *i*, |
| $trt_{i,i+1}$: | expected transfer time between transportation mode *i* and *i+1*. |
| $CI_{i,\,i+1}$: | the Comfort Index of transfer between transportation mode *i* and *i+1*. The CI of transfer should reflect the discomfort of transfer compared to the direct transportation. |

TCI describes the average comfort level along the travel chain, and is very useful when the competitiveness of transport alternatives is compared. We used a 5-point scale (with categories “Dissatisfactory”, “Low”, “Medium”, “High” and “Excellent”) to categorize TCI. If two alternatives are on the same comfort level, then soft measures for behavioural changes are needed to realize modal shift. If their comfort level is different, then first the overall comfort of the promoted travel chain should be improved to make other modal shift measures effective.

However the average comfort level in itself is not enough to describe the comfort along the travel chain. The perception of comfort is heavily influenced by the worst case. The Minimal Comfort Index (MCI) shows this effect, and is defined as the lowest Comfort Index along the travel chain:

$$MCI = \min\left(CI_{w,i}, CI_i, CI_{i,i+1}\right) \tag{4}$$

where

| | |
|---|---|
| *MCI:* | the worst Comfort Index along the travel chain, |
| $CI_{w,i}$: | the Comfort Index of walking at the beginning and at the end of the travel, |
| $CI_i$: | the Comfort Index of transportation mode or service *i*, |
| $CI_{i,\,i+1}$: | the Comfort Index of transfer between transportation mode *i* and *i+1*. The CI of transfer should reflect the discomfort of transfer compared to the direct transportation. |

We classified MCI into the same 5-point scale as described above. The difference between the levels of TCI and MCI shows the reliability of comfort. If TCI and MCI belong to the same class, then the comfort along the travel chain is continuous and reliable. The higher is the difference the higher is the fluctuation in comfort and thus the lower is the attractiveness of the travel chain. This reliability is described by the Comfort Reliability Index, which is defined as the level difference between TCI and MCI subtracted from 5:

$$CRI = 5 - \left(Level(TCI) - Level(MCI)\right) \tag{5}$$

where

| | |
|---|---|
| *CRI:* | Comfort Reliability Index along the travel chain, with 5 as the maximal reliability and 1 as the worst reliability, |
| *TCI:* | Comfort Index along the travel chain, |
| *MCI:* | the worst Comfort Index along the travel chain, |
| *Level(i):* | the Level of index *i* on the 5-point scale, anchored 1 to “Dissatisfactory” and 5 to “Excellent”. |

At comparison of alternatives the MCI and CRI give very useful information about the perceived comfort and the attractiveness of alternatives.

**4. Evaluation system**

The evaluation should be done from the point of view of users. The first step of qualification is the identification of the travel group at issue: who they are and what their expectations are, how they value time, etc. There are several approaches and models for this purpose, any of them can be used.

The next step is the evaluation of the travel time indices. Both TTI and TTRI (and their monetized counterparts) should be "Good" or at least "Acceptable", else the alternative is not competitive. Our introductory case study of the bus lane clearly shows this fact. If TTI or TTRI is not acceptable along the travel chain, then no push measure to increase travel time of individual passenger car usage will be effective.

The indices show how to improve speed along the travel chain. If TTI and TTRI are on the same level than intervention for decreasing travel time are suggested. If the level of TTRI is worse than the level of TTI, then measures improving the reliability along the travel chain are needed. If the monetized indices show problem than the fare and the toll system should be adjusted.

The speed conditions along the travel chain are described by the worst category, and measures for improvement should aim those areas which belong to the worst level. Improving from "Unacceptable" to "Acceptable" means a significant improvement, while stepping from "Acceptable" to "Good" has a smaller effect.

Then the evaluation of the comfort indices comes. MCI shows the perceived lowest comfort, CRI defines the continuity of comfort and TCI gives the average comfort level. All of the components with less comfort than the average comfort level should be improved to have effect; separated local improvements will cause only marginal results. The CRI shows the extent of the necessary improvement. If the reliability is low then it is possible that by improving the worst components a higher average comfort level can be reached; so the comfort of these components should be improved onto the level above the TCI level.

If we want to encourage modal shift then first we have to determine the affected travel chains. Then the speed and comfort indices of the alternatives should be evaluated. If there are differences in level, then first these differences should be eliminated, by either push or pull measures. It is very important, that we have to think globally along the travel chains, local interventions might be very dissatisfactory. Then soft measures of behavioural changes are the most effective to promote modal shift.

**5. Optimization of road surface usage**

In urban areas one of the major challenges to divide the very limited road surface among the competing transportation modes. The old approach was to build as many lanes for passenger cars as possible, but now we try to shift the balance towards more sustainable transportation modes, such as public transport and bicycles. The environment-protecting groups continuously demands more bus lanes and bike lanes and less lanes for road vehicles.

The travel chain based analysis helps us to answer this question. Simply we have to determine the travel chains using a given corridors, and the *number of people* traveling along all travel chains. The demand for each mode is the sum of people whose travel chain contains the given mode in the given corridor. Now we have a clear picture about the demands and the patterns, and we can divide the limited road surface among the modes according to the demand for the mode. If this analysis had been made for the introductory case study of bus lane, it would have shown that the great majority of people in that corridor use passenger car. The creation of the new bus lane upturned the existing balance and this caused its failure.

It does not mean that the bus lane in corridors like this is disadvantageous and unnecessary. It only shows that we have to think a little more. If we want to give more surface for sustainable urban modes to improve their travel time characteristics, then the question of modal shift along the travel chains should be handled as well.

**6. Conclusions**

The main goal of urban transport policies is to provide mobility with minimizing the harmful effects of transportation, which requires a modal shift from individual car usage towards more sustainable modes. The usual measures of sustainable urban mobility aim to improve the attractiveness of public transportation services and discourage individual car usage. However there are cases when the application of this usual methodology is not only ineffective, but harmful.

In the current paper we have proposed a new, travel chain based approach to evaluate urban mobility. We showed that local interventions to promote modal shift can be not only unsuccessful but extremely harmful as well. We analyzed the variables influencing of attractiveness of travel alternatives, and defined indices to describe the characteristics of travel chains from the individual specific point of view of the traveler. The proposed methodology provides a tool to easily and simply identify the most effective measures to promote modal shift toward sustainable transportation modes and to evaluate the expectable effect of intended measures.

Modal shift can be achieved only if both the travel time and the reliable travel time of the sustainable alternative are at least acceptable. If not, then measures to decrease travel time or improve the reliability should be done, else the modal shift will not take place.

If the travel time indices show that the travel alternatives are competitive than the comfort factors will decide the battle. The average comfort level describes the overall travel conditions along the travel chain, while the minimal comfort level shows the perceived comfort level from the point of view of the traveler. To improve comfort the fluctuation along the travel chain should be decreased by improving the components of the lowest comfort level. So it is possible that the modal shift in a given traffic corridor can be promoted by measures outside the corridor.

Road surface should be divided among transportation modes according to their demand. If one transportation mode should be promoted by assigning more surface to it than its share, then travel chain based measures for modal shift are needed as well.

The results of our research are helpful to the evaluation the actual service level of sustainable alternatives of individual car usage and to promote modal shift towards sustainable transportation modes.

**References**

Bamberg, S., Fujii, S., Friman, M., & Gärling, T. (2011). Behaviour theory and soft transport policy measures. *Transport Policy, 18 (1), 228-235.*

Bhat, C. R., Dubey, S. K. 2014. A New Estimation Approach to Integrate Latent Psychological Constructs in Choice Modeling. Available in Internet: <http://ctr.utexas.edu/wp-content/uploads/pubs/dstop/iclv_modeling.pdf>

Dawnay, E., & Shah, H. (2005). Behavioural Economics: Seven principles for policy makers. *London, New Economics Foundation.*

Economides, S.B., Han, C.L., Orowitsch, S., Scoullos, I.M., Nuttall, W.J. 2012. Paradigm shift for future mobility: a cross country analysis of behavioural policies. *Transport Research Arena – Europe 2012.* In *Procedia - Social and Behavioral Sciences 48* ( 2012 ) 2588 – 2596

European Commission. Energy & Transport. 2003. Available at from Internet: <http://ec.europa.eu/transport/marcopolo/documents/docs_en.htm>.

Gudmundsson, H. 2004. Sustainable transport and performance indicators. In: Hester, R. E., Harrison, R. M. (Eds), Transport and the Environment-Issues in environmental Science and Technology, 20. Royal Society of Chemistry, Cambridge, UK, pp. 35-63.

Guo, X. 2009. The comprehensive Evaluation on the Service level of City Public Transport. In International Conference on Computational Intelligence and Natural computing, 2009. Available in Internet: <http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5231093>

Johansson, M. V., Heldt, T., Johansson, P. 2005. Latent Variables in a Travel Mode Choice Model: Attitudinal and Behavioural Indicator Variables. *Uppsala University, Department of Economics, Working Paper 2005:5* Available in Internet: <http://uu.diva-portal.org/smash/get/diva2:665210/FULLTEXT01.pdf>

Kii, M., Hirota, K., Minato, K. 2005. A Study On Modal Shift Potential Considering Public Transport Operation. Available from internet: < http://128.40.111.250/cupum/searchpapers/papers/paper161.pdf >

Litman, T. 1999. Exploring the paradigm shift needed to reconcile sustainability and transportation objectives. *Transportation Research Record* 1670, 8-12.

Litman, T. 2008. Valuing Transit Service Quality Improvements. *Journal of Public Transportation,* Vol. 11, No. 2. pp 43-63

Maclaren, V. W. 1996. Urban sustainability reporting. *Journal of the American Planning Association* 62 (2), 184-202.

Migliore, M., Amoroso, S., Cardaci, V., Catalano, M. 2011. The role of passenger modal shift nodes in the interaction between land use and transport system. ERSA Conference Paper, RePEc: wiw:wiwrsa:ersa10p1606, Available from internet: <http://ideas.repec.org/p/wiw/wiwrsa/ersa10p1606.html>.

Miranda, H. F., da Silva, A. N. R 2012. Benchmarking sustainable urban mobility: The case of Curitiba, Brazil. *Transport Policy* 21: 141-151.

Tsalis, P., Naniopoulos, A. 2012. *Transport Research Arena – Europe 2012*. In *Procedia - Social and Behavioral Sciences 48* ( 2012 ) 2597 – 2606

# THE DOWNS-THOMPSON PARADOX IN MULTIMODAL NETWORKS

**Rodrigo Rezende Amaral[1], El-Houssaine Aghezzaf[2]**
*[1,2]Department of Industrial Management, Faculty of Engineering and Architecture, Ghent University, Technologiepark 903, 9052 Gent-Zwijnaarde, Belgium*

**Abstract:** Users of the transportation networks generally choose their routes in an independent and uncoordinated way in order to minimize their own perceived costs. This non-cooperative behaviour can lead to a suboptimal utilization of the network and, in some situations, increasing the network capacity can make the subutilization even worse. Such phenomenon is described in literature as traffic or network paradoxes. This paper provides a review on two famous network paradoxes, and also introduces a new one.

**Keywords**: Traffic Paradoxes, Network Paradoxes, Downs-Thomson Paradox, Braess Paradox

## 1. Introduction

Transportation networks are designed to allow the movement of individuals, the distribution of goods and the execution of service operations. These are essential activities in any society and the efficient management of the transportation networks is vital for its economic development and well-being. The amount of time people spend going to and from work; the logistic costs of goods and services and the pollution level are just some examples of quality of life factors which are very closely related to the efficiency of transportation networks.

In this context, the task of traffic managers is to optimize the flows in a transportation network considering the population's needs and requirements. This means that the network should be designed, controlled and operated in a way to minimize average travel times; avoid traffic jams; reduce air, noise and visual pollution, etc.

The users of the transportation networks, however, may not have exactly the same perception and objectives as the traffic managers. For instance, individuals will choose the routes which minimize their own travel times (rather than the global average); companies will place their facilities in the most economically convenient areas (even if it causes congestion and impacts other users); service providers will schedule their activities in order to maximize their gains (even during rush hours, if the returns compensate). In general, users analyse how the network impact their businesses, not how their businesses impact the network. Not surprisingly, this selfish behaviour can lead to a subutilization of the network capacity and potentially to a net economic loss for the whole society.

Notwithstanding this fact, this lack of coordination is accepted as a natural part of the current socio-political systems of developed economies, which not unjustifiably foster freedom and concurrency as key factors for growth. Given that it cannot – and maybe should not – be easily changed, it is natural to think that the only way to improve the performance of a transportation network is by increasing its capacity. But here a more surprising conclusion has been found in research: in this non-cooperative environment, increasing network capacity (e.g. adding or expanding a road) may actually decrease the overall performance, even when the only and explicit objective of all users is to minimize their travel times.

This counterintuitive fact has been demonstrated by some paradoxes, among which the Downs-Thomson Paradox (Downs, 1962) and the Braess Paradox (Braess, 1968) can be highlighted. Real evidences of these paradoxes have been found and studied, and many experiments have been proposed based on them (Denant-Boèmont et Hammiche, 2010; A. Rapoport et al, 2009). These paradoxes have received growing attention over the recent years, as they seem to occur in other types of networks, such as computer networks and power distribution (Max Planck Institute, 2012).

This paper intends to review the above-mentioned paradoxes and provide some possibly new perspectives from which they could be studied in Operations Research. A third paradox, which is, for the best of our knowledge, still unexplored in literature, is also introduced. It is named *The Downs-Thomson Paradox in Multimodal Networks* and is described in Section 5. Before presenting the paradoxes, we first discuss some theoretical aspects on which our analysis is based.

## 2. Game Theory Applied To Transportation Networks

As stated in the previous section, the users of a transportation network will generally take decisions in a non-cooperative way in order to maximize their own benefits. Although it would be very difficult to define what each user considers as benefit, it can be assumed that almost every user intends to minimize his travel time, and this is what really matters from the network management point of view.

In the problems discussed in this document it will be considered that a number of individuals travel from one to another point of a network, and that they do it on a regular basis. Each individual wants to minimize his travel time and will therefore choose the route which he perceives as the fastest. Individuals with longer travel times tend to copy the ones with shorter travel times and, eventually, an equilibrium situation is reached, in which the travel time of all users is the same and the flows in each route remain constant. In a game theoretical point of view, such a situation can be seen as a Nash equilibrium.

To illustrate this approach, suppose that 1,000 vehicles/h travel daily from city A to city B during peak-hours using the transportation network depicted in Figure 1.

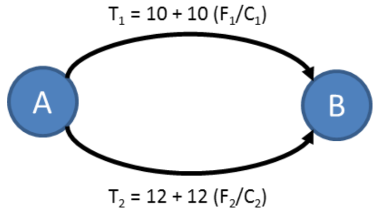



**Figure 1**
*Two separate roads connect two cities*

There are two separate roads connecting the cities. The travel time on the first road ($T_1$) is 10 minutes without traffic, and it rises linearly with the ratio of traffic flow ($F_1$) to road capacity ($C_1$). Similarly, the travel time on the second road ($T_2$) is 12 minutes without traffic, and it rises linearly with the ratio of traffic flow ($F_2$) to road capacity ($C_2$).

$$T_1 = 10 + 10\,(F_1/C_1);\ T_2 = 12 + 12\,(F_2/C_2) \qquad (1)$$

Let us consider that 600 travellers initially use the first road while the other 400 use the second road, and that the capacities $C_1$ and $C_2$ are respectively 400 and 480 vehicles/h. In this situation, the travel times are the following:

$$T_1 = 10 + 10\,(600/400) = 25 \text{ min};\ T_2 = 12 + 12\,(400/480) = 22 \text{ min} \qquad (2)$$

Once this information is spread among the travellers, some individuals using the first road will change their route on the following day. It can be expected that, after some time, the travellers will divide themselves in a way that the travel times of both routes are the same or at least very similar.

To determine the equilibrium flows and travel time, the following system of equations can be used:

$$\begin{cases} T_i = T_j \quad \forall i,j\,, i \in \{1 \dots N\}, j \in \{1 \dots N\} \\ \sum_{i=1}^{N} F_i = F_{total} \end{cases} \qquad (3)$$

where:

$T_i$ : travel time of route $i$
$F_i$ : flow of route $i$
$N$ : number of possible routes from A to B
$F_{total}$ : total flow of individuals from A to B

If the solution of the system contains a negative flow for some route $i$, it is not valid. It means that the route related to this flow will not be used in equilibrium ($F_i$ will actually be zero). In order to find a valid solution, the equations related to this route should be removed and the system should be solved again.

In the example described in this section, the system becomes:

$$\begin{cases} T_1 = T_2 \Rightarrow 10 + 10(F_1/400) = 12 + 12(F_2/480) \\ F_1 + F_2 = 1000 \end{cases} \qquad (4)$$

Solving it, we find that 540 travellers will take the first road and 460 will take the second when the equilibrium is reached. The travel time for all of them will be 23.5 minutes.

$$F_1 = 540;\ F_2 = 460;\ T_1 = 10 + 10\,(540/400) = 23.5 \text{ min};\ T_2 = 12 + 12\,(460/480) = 23.5 \text{ min} \qquad (5)$$

The solution is graphically presented in Figure 2.

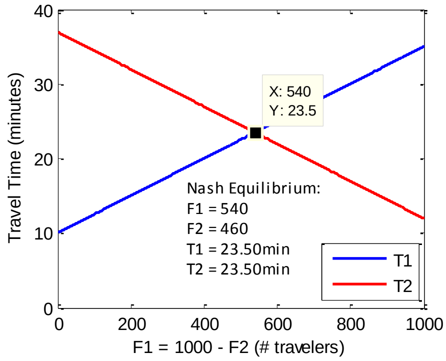



**Figure 2**
*Graphical solution for the Nash Equilibrium*

The equilibrium solution of this example is also Pareto optimal, since it is impossible to decrease the travel time of one individual without increasing the travel time of others. Furthermore, an increase of the capacities $C_1$ or $C_2$ will certainly decrease the travel time of all network users. There is no paradox in this network topology.

This game theoretical approach can be used, for instance, to minimize the equilibrium travel time in a transportation network given a limited budget and the costs of expanding the capacity of each road. Of course, in realistic problems, many other variables, constraints and objectives would be involved; multiple sources and destinations would exist; multiple types of transportation means would be available; etc. Nonetheless, the game theoretical concepts can undoubtedly provide useful insights and better solutions.

### 3. The Downs-Thomson Paradox

Suppose that in the previous example the travellers could choose between using the first road and a privately operated train line. The travel time on the road ($T_1$) is 10 minutes without traffic, and it rises linearly with the ratio of traffic flow ($F_1$) to road capacity ($C_1$). The travel time by train ($T_2$), on the other hand, depends on its frequency, which is adjusted by the train operator accordingly to the demand for the service, i.e. the number of users ($F_2$). The maximum travel time is 20 minutes, and this can be shortened in the rate of 1 minute per 300 users.

$$T_1 = 10 + 10\,(F_1/C_1);\ T_2 = 20 - (F_2/300) \quad (6)$$

This transportation network is presented in Figure 3.

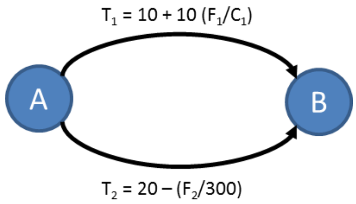



**Figure 3 – A road and a train line connect two cities**

The equilibrium solution is found by solving the following system of equations:

$$\begin{cases} T_1 = T_2 \Rightarrow 10 + 10(F_1/C_1) = 20 - (F_2/C_2) \\ F_1 + F_2 = 1000 \end{cases} \quad (7)$$

Let us consider that the road capacity $C_1$ is initially 500 vehicles/h. The equilibrium solution is then:

$$F_1 = 400;\ F_2 = 600;\ T_1 = 10 + 10\,(400/500) = 18\text{min};\ T_2 = 20 - (600/300) = 18\text{min} \quad (8)$$

In equilibrium, 400 travellers will take the road and 600 travellers will take the train. The travel time for everyone is 18 minutes. The graphical solution for this situation is presented in the graph on the left of Figure 4.

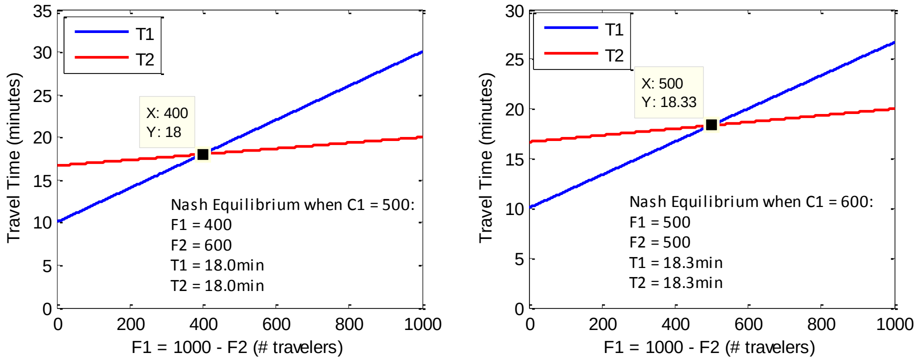



**Figure 4**
*Nash Equilibrium when $C_1$ = 500 (left) and when $C_1$ = 600 (right)*

If the road capacity $C_1$ is expanded to 600 vehicles/h, the equilibrium solution becomes:

$$F_1 = 500;\ F_2 = 500;\ T_1 = 10 + 10\ (500/600) = 18.33\text{min};\ T_2 = 20 - (500/300) = 18.33\text{min} \tag{9}$$

In this situation, 500 travellers will take the road and 500 travellers will take the train. The travel time for everyone becomes 18.33 minutes. Here is the paradox: the increase of the road capacity ($C_1$) causes an increase on the travel time for all the users. The graphical solution for this situation is presented in the graph on the right of Figure 4.

The paradox does not occur for any increase of $C_1$. For instance, if $C_1$ is increased to 2,000 vehicles/h, all travellers will take the road, since the worst travel time using it is $T_1$ = 10 + 10 (1000/2000) = 15min, which is shorter than the best travel time by train $T_2$ = 20 – (1000/300) = 16.67min.

The relation between the equilibrium travel time and the road capacity $C_1$ is shown in Figure 5. Remark that if the road capacity $C_1$ is less than 1,500 vehicles/h, the equilibrium travel time can be reduced simply by closing it (i.e., making $C_1$ = 0). In other words, the Nash Equilibrium when $0 < C_1 < 1{,}500$ is not Pareto Optimal.

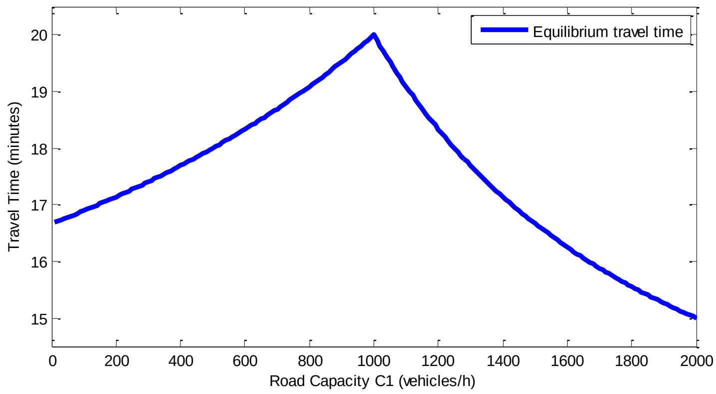



**Figure 5**
*The equilibrium travel time in function of the road capacity $C_1$*

## 4. The Braess Paradox

Suppose now that the 1,000 travellers from the previous examples use the network depicted in Figure 6.

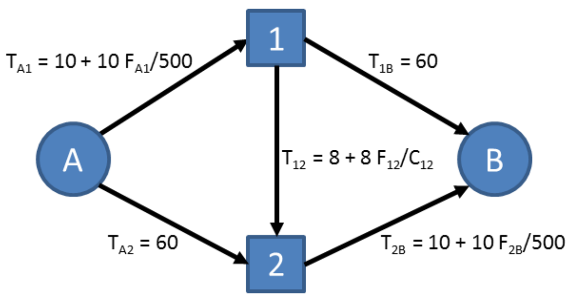



**Figure 6**
*A network with three possible paths from A to B: A-1-B, A-2-B and A-1-2-B*

The cities are separated by a river and the travellers have three options to go from city A to city B. In the first option, they can first take a narrow road from city A to point 1, in which the travel time increases with the number of vehicles, and then take a ferry from point 1 to city B. The frequency and capacity of the ferry can always suffice the demand, and the travel time is 60 minutes. The second option is similar to the first, but first they take a ferry from city A to point 2 and from there a narrow road to city B. A third option is to go from city A to point 1 using the same narrow road as in the first option, then take a narrow bridge from point 1 to point 2, and finally go from point 2 to city B using the same narrow road as in the second option. The travel time on the bridge ($T_{12}$) is 8 minutes without traffic, and it rises linearly with the ratio of traffic flow ($F_{12}$) to bridge capacity ($C_{12}$).

$$T_{A1} = 10 + 10\ F_{A1}/500 \tag{10}$$
$$T_{1B} = 60 \tag{11}$$
$$T_{A2} = 60 \tag{12}$$
$$T_{2B} = 10 + 10\ F_{2B}/500 \tag{13}$$
$$T_{12} = 8 + 8\ F_{12}/C_{12} \tag{14}$$

The travel times of the three possible paths, i.e. A-1-B, A-2-B and A-1-2-B, can be respectively written as:

$$T_{A1B} = T_{A1} + T_{1B} = 70 + 10\ F_{A1}/500 \tag{15}$$
$$T_{A2B} = T_{A2} + T_{2B} = 70 + 10\ F_{2B}/500 \tag{16}$$
$$T_{A12B} = T_{A1} + T_{12} + T_{2B} = 28 + 10\ F_{A1}/500 + 10\ F_{2B}/500 + 8\ F_{12}/C_{12} \tag{17}$$

Remark that a fourth option A-2-1-B does not have to be considered, since it would take at least 128 minutes ($T_{A2} + T_{12} + T_{1B}$), which is more than the maximum travel times for paths A-1-B and A-2-B. Since the path A-1-2-B share the arcs A-1 and 2-B respectively with the routes A-1-B and A-2-B, the flows $F_{A1B}$, $F_{A2B}$ and $F_{A12B}$ must satisfy the following equations:

$$F_{A1} = F_{A1B} + F_{A12B};\ F_{2B} = F_{A2B} + F_{A12B} \tag{18}$$

Path A-1-2-B is the only one which contains the arc 1-2. For this reason, the flow $F_{A12B}$ must equal $F_{12}$.

$$F_{A12B} = F_{12} \tag{19}$$

The equilibrium solution can then be found by solving the following system of equations:

$$\begin{cases} T_{A1B} = T_{A2B} \Rightarrow\ 70 + 10\ F_{A1}/500 =\ 70 + 10\ F_{2B}/500 \\ T_{A1B} = T_{A12B} \Rightarrow\ 70 + 10\ F_{A1}/500 =\ 28 + 10\ F_{A1}/500 + 10\ F_{2B}/500\ + 8\ F_{12}/C_{12} \\ F_{A1B} + F_{A2B} + F_{A12B} = 1000 \\ F_{A1} = F_{A1B} + F_{A12B} \\ F_{2B} = F_{A2B} + F_{A12B} \\ F_{12} = F_{A12B} \end{cases} \tag{20}$$

Considering that the bridge capacity is initially $C_{12}$ is 200 vehicles/h, the solution of the system is:

$$F_{A1B} = 180;\ F_{A2B} = 180;\ F_{A12B} = 640;\ T_{A1B} = T_{A2B} = T_{A12B} = 86.4 \text{ min} \tag{21}$$

In equilibrium, 180 travellers take the path A-1-B, 180 take the path A-2-B and 640 take the path A-1-2-B. The travel time for all of them is 86.4 minutes.

Since the paths A-1-B and A-2-B are equivalent, $F_{A1B} = F_{A2B}$ and $T_{A1B} = T_{A2B}$ in any equilibrium solution. This fact allows the problem to be solved using a two dimensional graph as demonstrated by Figure 7. The X-axis represents the sum of $F_{A1B}$ and $F_{A2B}$, which is also the same as 1000 - $F_{A12B}$. On the left, the solution with $C_{12}$ = 200 is presented.

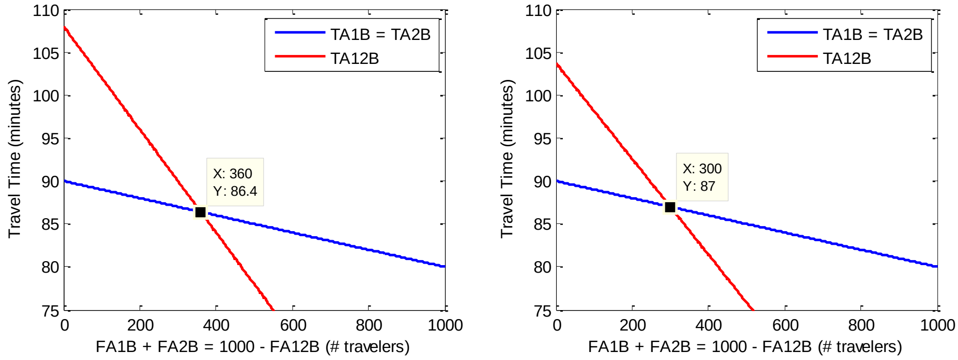



**Figure 7**
*Nash Equilibrium when $C_1$ = 200 (left) and when $C_1$ = 224 (right)*

If the bridge capacity $C_{12}$ is increased to 224 vehicles/h (right portion of Figure 7), the solution of the system is:

$$F_{A1B} = 150;\ F_{A2B} = 150;\ F_{A12B} = 700;\ T_{A1B} = T_{A2B} = T_{A12B} = 87 \text{ min} \tag{22}$$

An increased bridge capacity motivates more travellers to use it, shifting users from paths A-1-B and A-2-B to the path A-1-2-B. The travel time, however, is also increased, and this is the paradox.

Similarly to the Downs-Thomson Paradox, the Braess Paradox does not occur for any increase of $C_{12}$. If $C_{12}$ is increased to 500 vehicles/h, for instance, all travellers will take the bridge, and the travel time will be 84 minutes. This is less than the previous results found for $C_{12}$ = 200 and $C_{12}$ = 224. The relation between the equilibrium travel time and the bridge capacity $C_{12}$ can be observed in Figure 8. If the bridge capacity is less than 667 vehicles/h (approximately), the equilibrium travel time can be shortened by making it unavailable (i.e., making $C_{12}$ = 0). The Nash Equilibrium when $0 < C_{12} < 667$ is thus not Pareto Optimal.

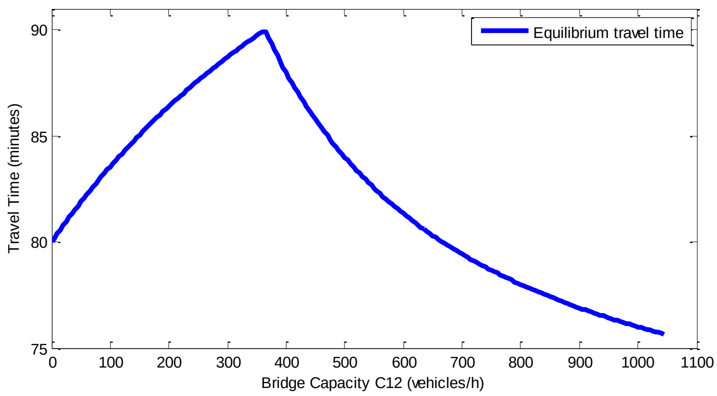



**Figure 8**
*The equilibrium travel time in function of the bridge capacity $C_{12}$*

### 5. The Downs-Thomson Paradox in Multimodal Networks

The paradox presented in this section represents a situation which can happen in multimodal transportation networks. To illustrate it, let us consider that the 1,000 travellers from the previous examples can go from city A to city B using cars or small buses. The frequency of buses is very high and users never have to wait long for them, although they are slower and the travel times without traffic are longer for them than for cars. There are dedicated lanes for them on some parts of the road; on the other parts, they share the road with cars. A network diagram of this situation is depicted in Figure 9.

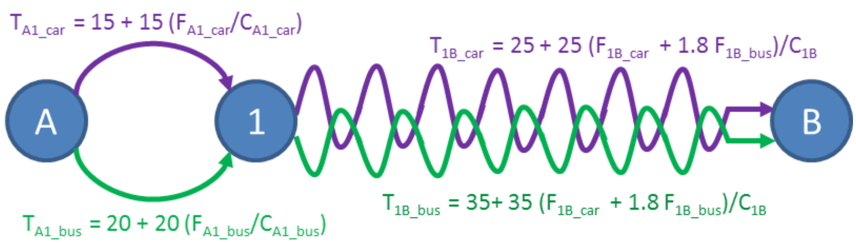



**Figure 9**
*A multimodal network*

The arcs A-1 represent the parts of the road in which there are dedicated bus lanes. On this arc, the travel time by car ($T_{A1\_car}$) depends on the flow of cars ($F_{A1\_car}$) and the capacity of the car lanes ($C_{A1\_car}$), while the travel time by bus ($T_{A1\_bus}$) depends on the flow of buses ($F_{A1\_bus}$) and the capacity of the bus lanes ($C_{A1\_bus}$).

$$T_{A1\_car} = 15 + 15\ F_{A1\_car}/C_{A1\_car};\ T_{A1\_bus} = 20 + 20\ F_{A1\_bus}/C_{A1\_bus} \tag{23}$$

The arcs 1-B represent the parts of the road shared by buses and cars. The travel time of each transportation mean depends on the flow of cars ($F_{1B\_car}$) and buses ($F_{1B\_bus}$) and the road capacity ($C_{1B}$). The congestion impact of buses is higher than the one of cars. This is taken into account in the calculation of the travel times by the inclusion of the Passenger Car Equivalent for buses ($PCE_{bus} = 1.8$) as a averaging factor for $F_{1B\_bus}$. The travel times for cars and buses are respectively given by:

$$T_{1B\_car} = 25 + 25\ (F_{1B\_car} + 1.8\ F_{1B\_bus})/C_{1B};\ T_{1B\_bus} = 35 + 35\ (F_{1B\_car} + 1.8\ F_{1B\_bus})/C_{1B} \tag{24}$$

Remark that sections A-1 and 1-B on the network diagram do not necessarily represent real traffic roads. They can be seen as an illustration of the fragments of the road where the two distinct conditions apply, i.e., parts where there are dedicated bus lanes and parts where the lanes are shared by cars and buses share. In this sense, point 1 may actually not exist and travellers are not able to change from car to buses, nor vice-versa. Mathematically, this fact implies that:

$$F_{car} = F_{A1\_car} = F_{1B\_car};\ F_{bus} = F_{1B\_bus} \tag{25}$$

Made these considerations, the travel times by car and by bus can be rewritten as:

$$T_{car} = T_{A1\_car} + T_{1B\_car} = 40 + 15\ F_{car}/C_{A1\_car} + 25\ (F_{car} + 1.8\ F_{bus})/C_{1B} \tag{26}$$
$$T_{bus} = T_{A1\_bus} + T_{1B\_bus} = 55 + 20\ F_{bus}/C_{A1\_bus} + 35\ (F_{car} + 1.8\ F_{bus})/C_{1B} \tag{27}$$

Each bus can take 10 passengers and the total flow from city A to B is again 1,000 travellers/h. The flow equation is:

$$F_{car} + 10\ F_{bus} = 1000 \tag{28}$$

The equilibrium flow and travel times for this network can then be found by solving the following system of equations:

$$\begin{cases} T_{car} = T_{bus} \\ \quad \Rightarrow 40 + 15\ F_{car}/C_{A1car} + 25\ (F_{car} + 1.8F_{bus})/C_{1B}\ =\ 55 + 20\ F_{bus}/C_{A1bus} + 35\ (F_{car} + 1.8F_{bus})/C_{1B} \\ F_{car} + 10\ F_{bus} = 1000 \end{cases} \tag{29}$$

Considering that the capacities are $C_{A1\_car} = 140$; $C_{A1\_bus} = 100$; $C_{1B} = 200$, the system yields the following results:

$$F_{car} = 510;\ F_{bus} = 49;\ T_{car} = T_{bus} = 169.6\text{min} \tag{30}$$

If the capacity of arc A-1 is increased to $C_{A1\_car} = 190$, the solution of the system becomes:

$$F_{car} = 680;\ F_{bus} = 32;\ T_{car} = T_{bus} = 190.5\text{min} \tag{31}$$

Both solutions can be graphically observed in Figure 10. As in the paradoxes previously described, the increase of the capacity resulted in an increase of the equilibrium travel time. Figure 11 exhibits the relation between the equilibrium travel time and $C_{A1\_car}$. Note that the minimum equilibrium travel time occurs for $C_{A1\_car} = 0$. This means that the shortest travel time would be achieved when all travellers take the buses. In other words, the Nash Equilibrium is never Pareto Optimum when $C_{A1\_car} > 0$.

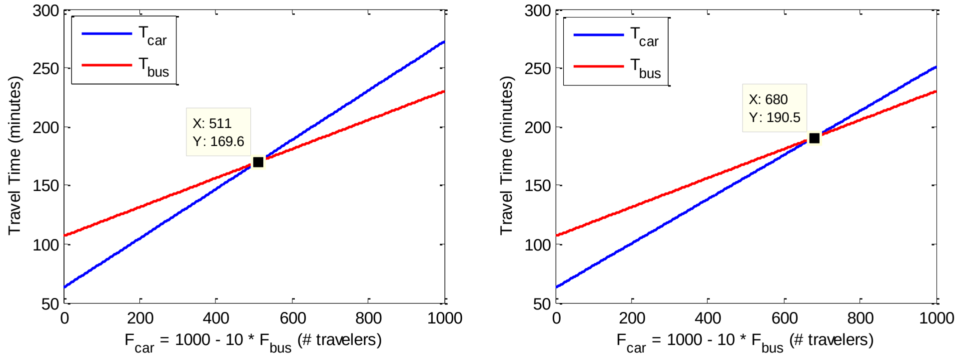



**Figure 10**
*Nash Equilibrium when $C_{A1\_car}$ = 140 (left) and when $C_{A1\_car}$ = 190 (right)*

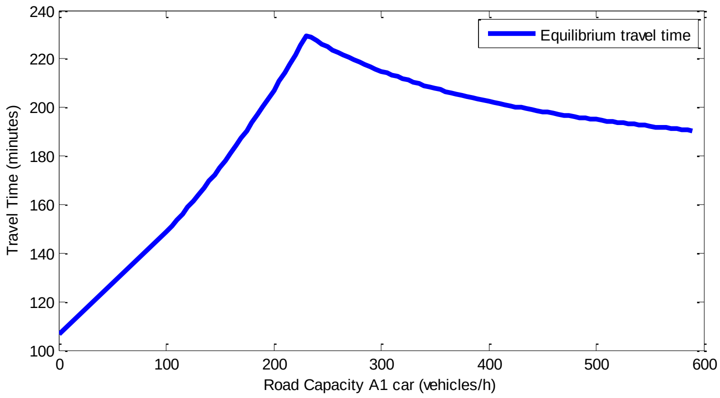



**Figure 11**
*The equilibrium travel time in function of the road capacity $C_{A1\_car}$*

The paradox presented in this section can be seen as a version of the Down-Thomson Paradox, in which equilibrium travel times decrease as users shift from public to private transportation means. An important difference is that the worsening of the travel time of public transport in the original formulation is caused by a disinvestment due to a reduced demand for the service (e.g., less frequent trains), while in this version of the paradox it is directly caused by the increase of the number of cars competing with buses on the shared roads, which is a direct result of the network users' behaviour and not an external decision.

Consequently, in the original version of the paradox, if the public transport operator would not adapt its service level to the demand, either no paradox would occur (i.e., an increase of the private transport capacity would always decrease the equilibrium travel time) or the travel time would not change (we could refer here to the Pigou-Knight-Downs Paradox, which has not been discussed in this work). This means that the equilibrium travel time could be reduced if the public transport operator would receive subsides to improve its service level (e.g., more frequent trains). In the version presented in this section, however, no action can be taken by the public transport operator (e.g., more frequent buses) in order to avoid the worsening of the equilibrium travel time.

## 6. Concluding Remarks

Arnott, Richard and Kenneth Small (1994) have cleverly pointed out two reasons for the paradoxes. The first reason is latent demand, i.e., when the capacity of a route is expanded, it is likely that the number of users of that route will increase. The second reason is that congestion is mispriced, since drivers do not pay for the losses they impose on the others. The solution proposed by these authors was congestion pricing: users should pay to use (congested) roads. They have mathematically shown how it would avoid the paradoxes, highlighted the growing technology which could make it feasible and discussed about the economic and political issues which could rise by implementing such a system. From the game theoretical point of view, this approach creates a favourable setting for the competitive strategy that the network users tend to follow, and ensures that the Nash Equilibrium is always Pareto Optimum.

Among the transport network users, however, are the logistics operators, who act in an increasingly coordinated way, using the growing knowledge and technology to organize their activities (Benjelloun and Crainic, 2008). They appear as a special class of players, governing a substantial part of the total flow in the network and having more complex objectives than simply reducing the length of a single trip. The progressive organization, rationality and importance of these players allow new optimization possibilities through the cooperation between them and the traffic managers.

A small example of how cooperation can lead to a better utilization of the transportation network is this: large logistic operators avoid overloaded routes at busy periods, while the traffic managers, in return, give them priority at traffic lights or dedicated lanes, as frequently done for public transportation. It is important to highlight, however, that to perform cooperation in an optimal way, it is necessary to integrate the systems (Traffic Management and Logistic Optimization), and this requires a very good understanding of the interactions between the variables involved and the possible trade-offs.

**Acknowledgement:** *This research work is support by CAPES Foundation (Proc. nº 9369-13-9), Ministry of Education of Brazil, Brasília - DF 70040-020, Brazil.

**References**

Arnott, R. and Small, K. (1994). The Economics of Traffic Congestion, *American Scientist*, 82: 446-455.

Benjelloun, A., and Crainic, T. G. (2008). Trends, challenges, and perspectives in city logistics. *Transportation and Land Use Interaction*, Proc. TRANSLU'08. Editura Politecnica Press, Bucharest, Romania, 269–284

Braess, D. (1968) Uber ein paradoxen der verkehrsplaning. *Unternehmensforschung* 12, 258-268.

Denant-Boèmont L. and S. Hammiche, (2010). Downs Thomson Paradox in Cities and Endogenous Choice of Transit Capacity: An Experimental Study, *Journal of Intelligent Transport Systems*: *Technology, Planning, and Operations*, 14 (3), 140-153

Downs, A., (1962). The Law of Peak-hour Expressway Congestion, *Traffic Quarterly*, 16, 393-409

Max Planck Institute (2012), Solar and wind energy may stabilize the power grid. Available from Internet: <*http://www.rdmag.com/news/2012/09/study-solar-and-wind-energy-may-stabilize-power-grid*>

Rapoport, A., Kugler, T. S., Dugar, and Gisches, E. J. (2009). Choice of routes in congested traffic networks: Experimental tests of the Braess Paradox. Games and Economic Behavior 65

# DEVELOPMENT & IMPLEMENTATION OF AN AIRPORT PAVEMENT MANAGEMENT SYSTEM

**Gianfranco Battiato**[1], **Lorenzo Cosimi**[2]

*[1,2] RODECO Group, Italy*

**Abstract:** Since the late '80s, RODECO Group has developed a Pavement Management System called RO.MA® (Road Management) PMS. The purpose of this PMS is to improve the quality of airport infrastructure in the short, medium and long term, maximizing available resources, not always adequate to the requirement.
RO.MA® PMS fulfills the requirements for an asset management tool and a PMS airport maintenance planning on airport pavements.
The survey of functional and structural pavement characteristics, done by the use of NDT (Nondestructive Deflection Testing) systems, is the basis of PMS; the most important parameters measured in the infrastructure and used by RO.MA® PMS, are the bearing capacity (HWD system) with ACN-PCN method, roughness and plano-altimetric profile, surface distresses (distress and PCI-Pavement Condition Index) and skid resistance.
RODECO Group has recently developed and introduced new technologies for automatic measurement of the PCI, IRI and skid resistance, in particular the ADE system (Automated Distress Evaluation), which allow more extensive and widespread use of high-performance systems for surveying and quality control of airport infrastructure.
The RO.MA.® PMS has been continuously updated linked to the evolution of high-performance systems, the new needs expressed by users and the international experience gained by RODECO in the field of Pavement Evaluation and Management System.
The main requirements of RO.MA.® are the following:

- provide a methodology to assess the pavement conditions with the Pavement Quality Index;
- define an optimum plan of maintenance with a priority list of interventions;
- use appropriate decay models to predict future conditions of pavements;
- propose planned pavement maintenance and rehabilitation based on optimizing the cost/benefit ratio.

The recent application of RODECO PMS in many international airports (Malta, Palermo, Olbia and Trieste) has specifically highlighted as an efficient system of PMS can optimize the maintenance on a multiannual basis, substantially reducing the overall cost of management of airport infrastructure.

**Keywords**: RO.MA.® PMS, HWD, Laser Profilometer, PCI, Skid Resistance, GPR, ADE System, Decay models, Homogeneous Sections, Priority Levels

## 1. Introduction

The purpose of a PMS is to improve the quality of airport infrastructures and plan maintenance measures in the short, medium and long term, maximizing available resources that are not always adequate to the customer needs.
RO.MA.® PMS satisfies these requirements as asset management tool, easy to use and can be customized following recommendations of airport agency.
The survey of functional and structural characteristics of the pavements, using high-performance systems, is the basis of PMS: the most important parameters, measured in the infrastructure and used by RO.MA.® PMS software, are the bearing capacity (measured with HWD) for ACN-PCN evaluation, roughness and planimetric profiles, surface distresses and PCI (Pavement Condition Index), skid resistance.
RODECO Group has recently developed and introduced new technologies for automatic measurement of the PCI, IRI and distress; in particular the ADE System (Automated Distress Evaluation) which allow more extensive and widespread use of high-performance systems for surveying and quality control of road and airport infrastructures.
The PMS RO.MA.® has been continuously updated relating to the evolution of high-performance systems, the new needs expressed by users and, in particular, the international experience gained by RODECO in the field of Pavement Evaluation and Management System.
Particular attention was engaged in software development, simplification and flexibility of data entry procedures, in order to reduce the amount of information necessary to obtain the results of the PMS.
The main requirements of RO.MA.® are the following:

- provide a methodology to assess the pavement conditions (surface and structural parameters), using the Pavement Quality Index (PQI);
- provide a device to define an optimum plan of scheduled maintenance and rehabilitation of airport pavements, with a priority list of interventions;
- estimate the optimal time to apply maintenance measures, using appropriate decay models to predict pavement conditions in the future;
- propose planned maintenance optimizing the cost/benefit ratio, with any budget constraints introduced by the user.

[1] Gianfranco Battiato - Email: gbattiato@rodeco.it

The recent RO.MA. applications, as in Malta and Palermo, have specifically highlighted as an efficient PMS system can optimize the maintenance on a multiannual basis, substantially reducing the overall cost of management of airport infrastructure.

## 2. Pavement Evaluation

RO.MA.® PMS software was designed to process data coming from the analysis of the results of the Pavement Evaluation phases (field tests). For the evaluation of structural and surface characteristics of an airport pavement network, there are different types of high-performance systems, which allow to quickly record all the parameters required for a proper evaluation of the PMS.

According to ICAO Annex 14 and ENAC "Italian Civil Aviation Authority", the typical pavement surveys, that it's possible to conduct are the following:

- definition of the pavement bearing capacity with HWD (Heavy Falling Weight Deflectometer) and calculation of ACN and PCN parameters;
- survey of longitudinal and transverse profiles using Laser Profilometer;
- measurement of pavement surface conditions and evaluation of PCI (Pavement Condition Index);
- investigations with GPR (Ground Penetrating Radar) and coring to define the pavement stratigraphy;
- survey of the skid resistance.

### 2.1 HWD (Heavy Falling Weight Deflectometer)

The structural characteristic of an airport pavement network are analyzed using the Heavy Falling Weight deflectometer (HWD), which can adequately simulate the load conditions of an aircraft.

For each HWD measuring point, through the software RO.ME. (Road Moduli Evaluation, developed by RODECO), are estimated:

- the values of the E1, E2, E3 moduli (asphalt layers, subbase and subgrade) under test conditions;
- the value of the E1 modulus (asphalt layer), referred to 20° C;
- residual fatigue life of the pavement in years;
- the critical layer;
- the calculated theoretical reinforcement necessary to support the project traffic, in mm;
- calculation of ACN/PCN in accordance with ICAO.

In addition to airport pavement analysis, bearing capacity tests of airport strips must be carried out.

The purpose of the strip survey is to evaluate the behavior of the ground when an aircraft runs off the runway and/or emergency vehicles transit.

The strips must have enough bearing capacity to support the weight of the critical airplane.

### 2.2 Laser Profilometer

The Laser Profilometer is used for the survey of the pavement roughness, longitudinal and transverse profiles and the definition of the index IRI (International Roughness Index), measured in mm/m.

The IRI is a standardized index that contains the information required to establish the regularity of a pavement surface, as defined by the World Bank Technical Paper No. 45.

The irregularities of a pavement are the result of an infinite number of wavelengths, that make up the longitudinal profile of a pavement.

The profilometer can detect the actual profile of the pavement in the XY coordinates (relative), where X is the distance measured by the odometer and Y represents the planimetric and elevation profile of the runway.

Generally, the system is capable of storing the actual average profile every 100/200 mm of each section.

The knowledge of the amplitude to short-waves (1 - 3.3 m), medium (3.3 m - 13 m) and long (13 m - 60 m) is very important to identify the cause of the irregularities.

Where the irregularity is related to the short wave, the problem can be found in the surface layers of the pavement (surface distresses), while the irregularities related to the long and medium wave may be due to problems of subsidence in the bottom layers.

The software allows, through simulation, to analyze the filtered results to obtain the values of irregularities at the wavelengths desired.

Starting from the current profile, the following parameters are estimated for sections of 25 m:

- IRI (International Roughness Index) averaged in mm/m on 25 m sections;
- irregularities filtered to short waves from 1 to 3.3 m, in mm/m, on sections of 25 m;
- irregularities filtered to medium waves from 3.3 to 13 m, in mm/m, on sections of 25 m;
- irregularities filtered to long waves from 13 to 60 m, in mm/m, on sections of 25 m;
- simulation of a 3 m straightedge for calculating the maximum deflection, as required by ICAO standards;

- the maximum deflection on sections of 45 m and the number of irregularities that in this section exceed 20-30 mm as required by ICAO standards;
- cross slope (%).

The equipment of RODECO Laser Equipment includes:
- various lasers, for a width of about 3 meters, with a sampling frequency each 5 mm;
- n. 3 accelerometers and n. 2 high-precision gyroscopes;
- software for raw data processing, to calculate cross-slope, rutting, longitudinal profiles (IRI) and irregularities for short, medium and long waves.

### 2.3 Survey Distress: PCI (Pavement Condition Index) analysis

The survey of the surface distress for rigid and flexible pavements is necessary to:
- control of surface degradations and their evolution over time;
- identify degraded areas to plan emergency actions;
- provide preventive maintenance and rehabilitation to slow or halt the process of degradation and prolong the service life of the pavements.

RODECO Group has developed advanced technologies for the automatic detection of surface distress as the ADE (Automated Distress Evaluation).
Digital images, using VIDEOCAR system, are acquired each 4 meters, on all airport pavements (runways, taxiways, aprons), and simultaneously 6 different types of surface distress (classification using 3 level of severity) are detected each 25 m section (ASTM-D5340-12).
The main surface distresses that ADE system can process are the following:
- longitudinal cracks;
- transverse cracks;
- alligator cracking;
- potholes;
- raveling;
- depressions.

Each type of surface distress should be evaluated by using 2 indices:
- quantity index: % of the area needing repair;
- quality index: L (low) = low severity - M (medium) = Moderate severity - H (high) = high severity.

The images of the pavement, as detected by the ADE system, are post-processed with a specific software called "Automated Distress Evaluation", through it's possible to recognize and codify, on a database, the different types of distresses, in a completely automatic way.
The system has a license that represents a technological innovation: Patent n. MI2008A000307 (26/02/2008).
From ADE surveys, PCI (Pavement Condition Index) are calculated for each sample unit, in accordance with the ASTM-D5340-12.
Starting from Videocar survey, global quantity of possible preventive maintenance/rehabilitation is defined, together with needs and priorities of maintenance, like crack sealing or other type of preventive treatments.
The PCI is a numerical index, ranging from 0 for a failed pavement to 100 for a pavement in perfect condition. PCI is divided into three classes:
- $70 \leq PCI < 100$ good
- $55 < PCI < 70$ fair
- $PCI \leq 55$ poor

Using this classification, a detailed map of all airport pavements is realized, that permits to identify promptly the critical sample unit (red color in the map).

### 2.4 Ground Penetrating Radar (GPR) Surveys

Through the use of GPR technology, it is possible to continuously detect the pavement stratigraphy and estimate the thickness of the different layers. The output of this survey consists of numerical tables and/or graphics that provide the thickness of the different layers of the existing airport pavements. The data obtained from the GPR surveys are used to process the HWD data, to estimate the elastic moduli of different layers, and for the PMS drawing.
RODECO Group has a radar system, that may mount up to 3 antennas operating between 200 MHz and 2 GHz, and that can investigate the pavement airport infrastructure up to 3 m of depth.

The antennas are installed on a vehicle properly equipped with the control unit, a computer process, a high-resolution color video and an encoder.
The collected data can be displayed on screen in real time and then processed to produce tables and graphs.
GPR is designed to make measurements of the thickness of the pavement in accordance with the ASTM Standard D4748-10.

### 2.5 Skid resistance

The skid data collection has a very important rule to assess the adequacy of air traffic safety level. In addition to the traditional static methods to evaluate the friction coefficient, such as the british pendulum friction tester, there are high-performance systems that can measure the friction by a moving vehicle properly equipped.
With these measurement systems, skid resistance can be detected almost continuously, every 5, 10 meter of pavement length.
The GripTester, approved by ICAO, is a very simple device, consisting of a trailer towed to a vehicle; the trailer has two side wheels and a central wheel, braked during movement, used to measure the friction coefficient.
The braked wheel is constantly sprinkled with water by a distributor during the tests; water comes from a tank installed on the driving vehicle.
The water flow is regulated with an electronic pump to ensure the desired thickness (eg. 1 mm of water) between the tire and pavement; this water thickness don't depend of the speed measurement.
During the trailer movement, two longitudinal strain gauges measure the strain "Fo" (which is opposed to the travel) and the vertical load "Fv", given by the weight of the trailer; "Fv/Fo" ratio is the Grip Number (GN).
The strain gauges are connected to an electronic system that records data value of the friction coefficient measured every 10 meters on a onboard computer. The tests are performed according to the requirements of the legislation (ICAO Annex 14 - Aerodromes Volume I - Aerodrome Design) at a speed of 65 and 95 km/h.
Measurements are performed along 4 lines, at ±3 and ± 6 meters from the runway centerline.
Processed data can be provided in charts and graphics for every test point values, and the relative average values every third of the runway length, as required by ICAO standards, are reported.

## 3. Definition and evaluation of homogeneous sections

One of the most important step in the interpretation and processing data is related to choice airport sections with similar characteristics; an homogeneous section is characterized by having virtually constant parameters throughout its considered length. The procedure to calculate the homogeneous section permits to define the pavement section having similar features: once identified, they cannot be further divided into subsections having significantly different behavior.
The subdivision of an airport network element into homogeneous sections may be based on various parameters, whose considered are the following:

- IRI;
- PCI;
- deflection basis parameters (D0, SCI "Surface Curvature Index" and BCI "Base Curvature Index");
- traffic;
- number of layers used for backanalysis.

There are various statistical techniques available for sharing data in a series of homogeneous sections. One technique is the cumulative sum method, introduced in the '50s for the industrial quality control, and resumed in the "AASTHO Design for Design of Pavement Structures guides".
This method is the best among existing in the identification of homogeneous sections, but at the same time presents a difficult to translate in an algorithm. It was therefore used another method, inside the RO.MA.® PMS software, based on dichotomization of the measures proposed by the LCPC (Laboratoire Central des Ponts et Chaussées) in 1981.
This method considers a probability of change of an $r$ – parameter at a specific significance level; $r$ - parameter is the ratio between the quadratic distance between two successive data and variance of data.
The method uses the Carré Moyen test of consecutive differences for the study of a value's randomness and Gaussian characteristics in a section. In addition, it uses a dichotomist technique to perform the segmentation. Through this, the method verifies whether data behaves under a normal probabilistic distribution before performing the segmentation.
The segmentation method searches for the point where the decision function $g(i)$ is maximized, as shown in the following equation:

$$g(i) = \frac{n}{(n-i)i} \cdot \left[ \sum_{j=1}^{i} (x_j - \bar{x}) \right]^2$$

$$g_k = \max_{1 \le j \le k} S(i)$$

*(Eq. 3.1)*

where *n* is the total number of data contained in the analyzed section, $x_j$ is the $j$-value of the measured point and $\bar{x}$ is the mean of the data in the analyzed section. The point where $g(i)$ is maximum within the section, represents the point where the mean changes. It is also the first element of the second segment of the same section. Later, the procedure is repeated in both segments, obtaining consecutive maximum points of $g(i)$. The segmentation process ends when the $g(i)$ function contains no more maximum points within either analyzed segment, or when the homogeneous section length is less than a length value fixed by the user.

## 4. Description of an Airport Pavement Management System using RO.MA.® PMS software

The RO.MA.® PMS software, developed and specialized for airports, is a tool that can provide, in the airport pavement management, the optimal maintenance and rehabilitation strategies, identified through clear and simple procedures. The software, starting from a database relating to the pavement conditions and its further division into homogeneous sections, evaluates alternative strategies in a reference period of 10 years, considering the one with the highest effectiveness/cost ratio, for each homogeneous section.
The software flow chart starts from the survey data (residual life, IRI and PCI), then proceeds to the definition of homogeneous sections, with the method following described, for all considered parameters.
As database was created, and enriched over the time by surveys described in the chapters before, the software works following four main steps, which will subsequently be detailed separately:

- database of pavement parameters referred to the current condition of the pavement, definition of the homogenous sections and priority levels set for the airport network elements.
- prediction models of the IRI, PCI, and RL (Residual Life) over the time and parameter range for specific maintenance treatments;
- maintenance measures and their characteristics;
- calculation of the optimal strategy for each homogeneous section, considering budget constraints or free budget.

### 4.1 Definition of database and priority levels

The first step is to create a database, which shall contain all the features of the surveyed pavements. First, it’s necessary to identify, inside the entire airport network, the branches which will be assigned a priority for maintenance management; this priority decreases with the importance of the branch itself. It will be necessary to report the following data for each branch:

- ID code;
- traffic flow: traffic can be divided into three categories (high, medium, low), where the lower priority level is associated with high traffic level, and so on;
- number of layers of the backcalculated model, type and thickness, and average moduli E1, E2, E3 (MPa), respectively, of the asphalt concrete layer at 20° C, subbase and subgrade;
- average IRI, PCI and RL values, calculated for the homogeneous sections;
- geometrical characteristic, such as width, length and area of the homogeneous sections, and parameters measured by the laser profilometer, such as cross slope, and straight-edge;
- average value of the PQI parameter (Pavement Quality Index), which expresses the overall pavement condition. This variable is calculated as the sum of the various analyzed parameters (IRI, PCI, and RL); each of them has an associated multiplier wi such that:

$$\sum w_i = 1$$

*(Eq. 4.1.1)*

The PQI equation is the following:

$$\boldsymbol{PQI} = \frac{100}{(IRI_{max}-IRI_{min})} \cdot (IRI_{max} - IRI) \cdot w_{IRI} + \frac{100}{(PCI_{max}-PCI_{min})} \cdot (PCI - PCI_{min}) \cdot w_{PCI} + \frac{100}{(RL_{max}-RL_{min})} \cdot (RL - RL_{min}) \cdot w_{RL}$$

*(Eq. 4.1.2)*

with maximum and minimum values of each parameter depending on the decay curves.
Once the database is given, the priorities for maintenance will be calculated for each section, needed to define the list of maintenance strategies during the PMS time period.

### 4.2 Prediction models of pavement parameters (IRI, PCI and RL) and selection of maintenance level

Prediction models for IRI, PCI and RL parameters have been provided inside the software as a function of the possible traffic class of the airport. For each parameter, basing on technical experience, a decay curve was calculated for each traffic class. These models depends on time, so the value of a particular parameter can be calculated at any time during the service life of the pavement.
It should be noted that these models have to be calibrated starting from the results of the high – performance survey that must be repeated over the time, to create an historical pavement database. For each parameter it is necessary to identify different ranges, when a type of maintenance treatment can be used, according to the three maintenance classes:

A) Preventive maintenance: required treatments that restore only the PCI parameters;
B) Partial reconstruction maintenance: required measures that restore PCI, IRI and RL parameters not in a complete way;
C) Rehabilitation maintenance: required treatments that restore PCI, IRI and RL parameters to the project conditions.

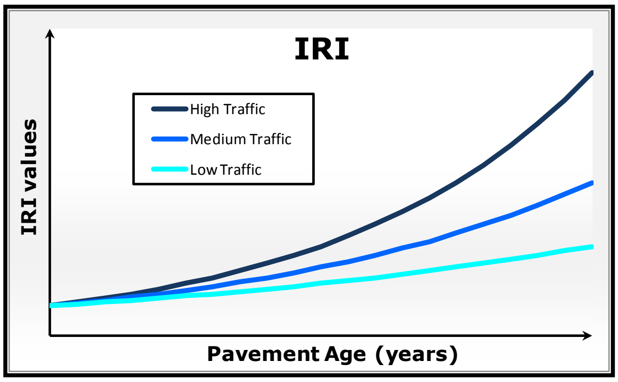



**Fig.1**
*Prediction model for IRI parameter, depending on the traffic level*

Consequently, three ranges have been defined for the parameters IRI and PCI; for the RL, 4 levels of maintenance are been considered (the “Do Nothing” condition was added).

### 4.3 Types of maintenance treatment and their characteristics

Within the RO.MA.® PMS software, it was decided to use a set of 9 maintenance treatments, divided into the three maintenance classes (see classification above). For each treatment it was necessary to define:

- the maintenance class which it is part of (preventive, partial reconstruction or rehabilitation);
- what parameters are repaired and how it restores (the recovery of each parameter is expressed as a percentage of the existing value);
- the unit cost (depending of the type of treatment);
- the expected life for each maintenance treatment (the required time of pavement to decay in the condition before applied maintenance).

Once established treatments list and their activation range, it is possible to define the treatment selection tree, the heart of the software: this tree represents the pavement situation, both before and after any maintenance, based on the calculation of PQI parameter.

### 4.4 Definition of the optimum strategy for each homogeneous section considering budget constraints or free budget

Once defined maintenance treatment and their application range, the last step of the software is to identify the best possible strategy for each homogeneous section during the PMS time.
Strategy refers to a plan of action that includes one or more specific maintenance operations designed to restore and improve the performance characteristics of the pavements.
Starting from the current value of PQI, the software calculates the best possible maintenance strategy, by combining the various types of treatments. This calculation is performed considering the possibility that the strategies are not applied to the first year of analysis, but they are deferred over the time, allowing the decay of all parameters, following their prediction model.

This approach is used because some interventions have a useful life less than the period of analysis on which the strategy is evaluated, then calculating strategies that involve the application of multiple treatments it's possible to cover the entire period analysis.
The strategies are established from the decision tree starting from the set of maintenance treatments listed in the previous chapter: a strategy will therefore consist of one or more branches of the tree, depending on the duration of the proposed maintenance period. At the end of the analysis, the results, for each strategy, will be a vector having the following features:

- ID code, related to the homogeneous section;
- maintenance treatments types included and year of application.

Each section will have different maintenance strategies, because it depends when managing authority decide to take action on the section, as shown in the Fig.2: this is linked closely to the range values of individual parameters representing the pavement conditions.

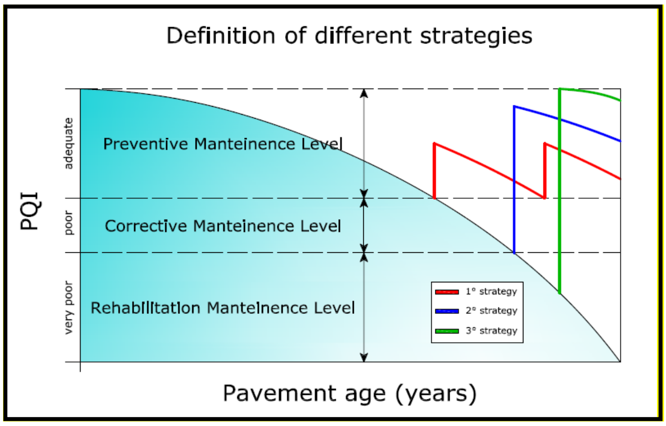



**Fig.2**
*Definition of different strategies for an homogeneous section*

The next step is to define the effectiveness of each strategy, its cost and the effectiveness/cost ratio, in order to choose which is the optimal strategy for each section. In this case it's necessary to consider two possible cases: the case where there is a free budget, and if there are economic constraints.

### 4.4.1 Free budget Analysis

For each strategy it's necessary to identify:

- his effectiveness, calculated on the PQI parameter;
- total costs of maintenance treatments considered.

About the calculation of the effectiveness of single-strategy (E), it will be identified graphically by the included area between the curve that represents the strategy after his application and the PQI curve before. About the overall costs of each strategy, they will be referred to the initial year of the analysis period (CCA), depending on the discount rate *r*.
By computing, for all possible strategies of each homogeneous section, the ratio between the effectiveness of single strategy and total cost (E/CCA), the strategy, that will have the higher value of the previous ratio, will be chosen. At the end of this process, which will be repeated for each homogeneous section, it will be possible to estimate the total budget to be invested for the implementation of the best strategies identified.

### 4.4.2 Limited budget Analysis

The method used to identify the best strategies, when a preset budget in the time of analysis (10 years) is used, consists of the following steps:

a) definition of the budget to invest;
b) definition of maintenance strategies during the period of analysis, for all homogeneous sections;
c) calculation of effectiveness "E" and total cost "CCA" of each maintenance strategy;
d) calculation of the ratio effectiveness/cost of each strategy (E/CCA);
e) selection of the strategy (referred to a *k* homogeneous section) having the higher value of E/CCA: it will be used as reference for the beginning of the iterative calculation;

f) calculation, for the *k* section, the marginal cost-effectiveness ratio MCE of all other expected strategies on the same *k* section. It is expressed in terms of effectiveness and cost of each i-th strategy identified for the section *k*, as compared with the selected strategy in step 5;

g) if MCE is negative or $E_i < E_k$, the i-th strategy is eliminated and will not be considered in further steps, otherwise the combination replaces the selected strategy in step 5, and the available budget is updated;

h) the process is repeated until no other strategy can be selected for each section in each year of analysis, and when the allocated budget is fully committed.

This method allows, in addition to the definition of the best strategies to be applied to all homogeneous sections, to perform an analysis at different budget levels: inserting a different value for the estimated budget (step 1), strategies choice will change, in accordance with the most appropriate capital to invest.

### 5. Implementation of Airport Pavement Management System: the example of Malta International Airport

Starting from 2010, a PMS study was conducted to plan maintenance strategy on all pavements of M.I.A. (Malta International Airport).
RO.MA.® PMS software was used to define the strategy of preventive maintenance and rehabilitation, once made the surveys using high-performance systems (IRI, PCI, HWD etc.). The software output was a list of maintenance measures per year and for each homogeneous section located within the airport network. In the chart obtained by the software the following parameters were shown:

- year of implementation of maintenance measure belonging to the chosen strategies for each homogeneous section;
- start and end of the homogeneous section;
- type of maintenance measure applied;
- area of maintenance treatment;
- unit cost (€/m2);
- total cost of maintenance per year (€).

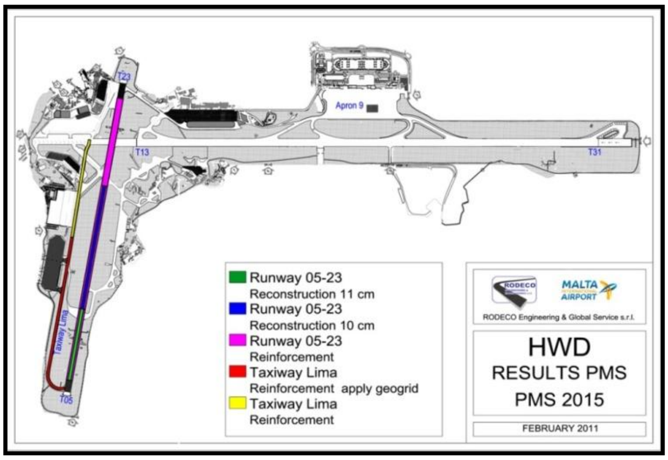



**Fig.3**
*Graphical output of RO.MA.® PMS software*

For each year, the total cost incurred by the managing authority was shown with the total budget invested at the end of the analysis time (which, in the case of limited budget, must be less than or equal to that allocated budget by the management).
Changing some parameters inside the software, it was also possible to make various analysis at different levels of budgets: the results were several tables, each with a different type of maintenance plan on the airport network examined.
The managing authority chose the maintenance plan as it deemed the best: this choice was based both on the results obtained by the program and on the budget that they were able to allocate to cover the analysis period (10 years).
The primary objective of the PMS for M.I.A. was to plan and especially "optimize" the pavement maintenance and rehabilitation strategy, i.e. to maintain the highest level of functional and structural characteristics of the pavement with minimal cost, assuring a significant savings for the managing authority.
This result has been achieved for the M.I.A. (Malta International Airport) investing in preventive maintenance and partial reconstruction to prolong the pavement service life and delay the process of deterioration of the pavements.

The first aim has been obtained through crack sealing with routing, using a special high-polymer-modified bitumen, and "Regephalt" technology, a special cold applied rejuvenating agent, that allows to stop raveling and aggregate loss from the pavement surface.
To improve the bearing capacity of flexible pavements, the use of a special fiber glass grid has been submitted by RODECO: the study, conducted in collaboration with the French company "6D Solutions" has allowed the development of design solutions that provide to scarify only 10 cm of asphalt, the installation of a special grid and the subsequent application of a high modulus binder thickness of 6 cm and a high modulus SMA thickness of 4 cm.
This pavement rehabilitation, based on new solution which saves over 60% of the cost compared to traditional reconstruction, with very low work times, was executed in April 2011 on the most critical part of the Apron 9.

**Fig.4**
*Laying of fiber glass grid on Malta International Airport*

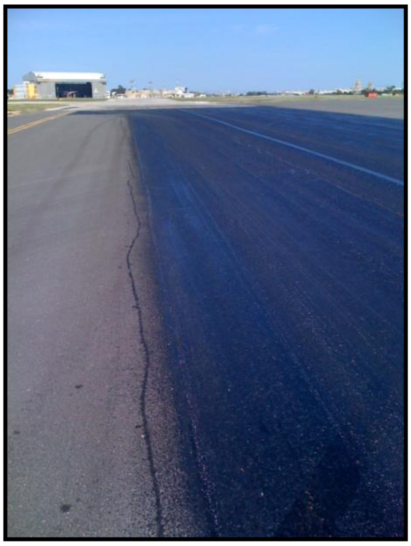


**Fig.5**
*Preventive Maintenance: Regephalt technology on Malta International Airport*

In February 2012, a full PMS study was completed of all flexible and rigid pavement of Palermo Airport, in Sicily. The functional and structural pavement condition was evaluated using the following systems:

- longitudinal pavement profile by laser profilometer;
- pavement bearing capacity by HWD;
- distress analysis with Videocar System;
- ground penetrating radar system (GPRS) to determine the thicknesses and composition of the different pavement layers.

The use of GPRS was essential to provide accurate and reliable data for use in the modulus evaluation step. A detailed care was spent for georadar survey on rigid pavement, in order to check the possible presence of voids, inside the subbase layer, under the slab corners. Indeed many corner cracks was detected by Videocar System, especially in the oldest part of the main Apron and in the taxiway thresholds.

The pavement management study produced a multiannual maintenance program, considering both free budget and fixed budget cases as suggested by the user. For both cases examined, a maintenance priority list was prepared based on cost/benefit analyses of the alternative measures.

The total cost proposed for the whole airport network was about 10 million euro.

The main maintenance measures supplied by the software are the following:

- pavement overlay with the laying of a special fiber glass geogrid, to reduce overall asphalt overlay thickness;
- injection of polyurethane resin with an high swelling pressure able, by expanding, to transmit to the surrounding soil an action of compaction with consequent filling of the voids and improvement of the mechanical and hydraulic soil features; this type of maintenance has been designed only for rigid pavements.

## 6. Conclusions

Some practical examples of the use of the RO.MA. PMS were presented for airport with a brief description of the pavement evaluation and PMS phases of the RO.MA. method. The pertinence of the proposed method has been discussed in terms of technical and economic aspects. The implementation of a database on the user's PC containing PE and PMS data allows and forces the user to prepare more realistic annual maintenance budgets based on a scientific approach to maintenance rehabilitation problems. Moreover, the use of new software for graphic representations on a PC of the airport map increases the information available for the agency to manage its network in terms of efficiency and economy, and the graphic presentation of the data allows the user to easily understand the results of the RO.MA. analysis.

The use of the system in all cases presented in this paper resulted in optimization of the maintenance budgets for all agencies. This was accomplished through the use of RO.MA. techniques to identify specific areas of the airport network that are in need of maintenance. Similarly, RO.MA. identified areas of the roadway that did not need maintenance. The users benefitted from this analysis by knowing exactly where to implement maintenance techniques, as well as what those techniques should be. By identifying the problem areas of the airport pavements, the users' maintenance budget was optimized.

**References**

Battiato, G., 1990. Evaluation of PCN (Pavement Classification Number) of airfields by means of the F.W.D.. In *Proceedings for Third International Conference on Bearing Capacity of roads and airfields, Trondheim, Norway, 3-6 July 1990*.

Battiato, G., 1998. Prove e controlli non tradizionali sulle pavimentazioni bituminose. La valutazione della qualità delle pavimentazioni stradali mediante l'impiego dei sistemi ad alto rendimento. In *Proceedings for SITEB 7° corso tecnico di base su leganti e conglomerati bituminosi, S. Donato Milanese, Italy, 25-27 November 1998.*

Battiato, G., 2000. Valutazioni e controlli di qualità delle Pavimentazioni Aeroportuali. In *Proceedings for Le pavimentazioni aeroportuali, Rome, Italy, 30 Novembre 2000,* 90-108.

Battiato, G.; Al Emadi K.M.I.; Liberati P, 2003. Description et application du systeme de gestion routier (PMS) developpe pour l'Etat du Qatar. In *Proceedings for XXII Congres Mondial Routier - C6 Comite sur la gestion routiere" Durban, South Africa, 19-25 October 2003.*

Battiato, G.; Larsen B.K., 1990. Description and Application of RO.MA.® (Road Evaluation and Pavement Management System). In *Proceedings for III International Conference on Bearing Capacity of roads and airfield, Trondheim, Norway, 3-5 July 1990.*

Battiato, G.; Amé E.; Wagner T., 1994. Description and Implementation of RO.MA.® for Urban Road and Highway Network Maintenance. In *Proceedings I for Third International Conference on Managing Pavements, S.Antonio, Texas, USA, May 22-26 1994,* 43-50.

Battiato, G.; Larsen B.K.; Ullidtz P.,1987. Verification of the analytical – empirical method of Pavement Evaluation based on FWD testing. In *Proceedings for 6th International Conference on Structural Design of Asphalt Pavements" Ann Arbor, Michigan, USA, 1987.*

# TRAFFIC FLOW SIMULATION BY NEURO-FUZZY APPROACH

**Aksel Seitllari** [1]
[1] *Department of Civil Engineering, Epoka University, Tirana, Albania*

**Abstract:** Traffic flow characteristics, which are the basic parameters of the traffic stream, play a crucial role in analyzing and modeling traffic flow. Widely, simulation software are utilized to predict the parameters such as average speed, travel time, capacity, etc. In essence, such programs are used to characterize the model input parameters for achieving an appropriate matching between the model and actual data. In this study, the actual traffic data is used for the evaluation of the susceptibility of adaptive neuro-fuzzy inference methodology to simulate the nonlinear mapping between flow and speed parameters. In the neuro-fuzzy model, Sugeno type of inference system is selected for fuzzy approximate reasoning, and the model parameters adopted by the hybrid-learning algorithm. Results denoted that neuro-fuzzy system can easily characterize such a nonlinear relationship including certain amount of uncertainty. This study is only for the demonstration of the application of the proposed methodology; nevertheless, additional parameters of traffic flow must also be considered in such a system to characterize real-world problems better.

**Keywords:** Traffic flow parameters, flow, speed, fuzzy systems, neural networks, ANFIS

## 1. Introduction

The purpose of this paper is to demonstrate the usage of mathematical tool such as fuzzy inference system (FIS) and neural network (NN) to gain model able to predict traffic flow parameters, based on the values of these parameters used as input variables.
Fuzzy inference system (FIS) and neural network (NN) is considered as a potential alternative to usage of conventional techniques. These latter restricted due to their strict assumption of linearity, normality, homogeneity and variable dependence (MA Hanna, D Ural, G Saygili, 2007).

Fuzzy inference system operates under the features of fuzzy logic which is able to deal with vague, complex and dynamic data. Using conventional mathematical tools in analyzing complex data might not effort the solving of the situation. In some cases the expressing of quantities is impossible to be done on numerical parameters, thus calling the need for linguistic approach. The if-then rules form uses common human reasoning instead of conventional mathematical equations enabling the system to deal with linguistically approached situations. The mapping of input-output space, also known as assigning of mathematical expression between input and output variables, is performed using data set (G Dell'Acqua, R Lamberti, F Abbondanti, 2002).
The concept of neural neurons is based on real neurons and their methodology of functioning. The modeling of artificial neurons consists of inputs, weights, sum function, activation function and outputs. Neural network provides a precise model based on database examples of the relationship. This due to its special ability of learning from examples, generate reasonable solutions even when the input data are incomplete, adaption ability to changing situations, rapid processing of information and clear transfers between other computing systems (I Flood, NKartam , 1994). The application of neural network in civil engineering comprise in structural optimization, structural control (HM Chen, KH Tsai, GZ Qi, CS Yang, F Amini, 1995), predicting the moment capacity of ferrocement members (MA Mashrei, N Abdulrazzaq, T Y Abdalla, MS Rahman, 2010) etc.

Fuzzy systems and neural network are capable of dealing and meliorate systems operating under imprecise, vague and noisy environments. They seem as obverse techniques. The combination of these techniques result in an integrated system known adaptive neuro-fuzzy inference system (ANFIS). This system is able to operate in the name of both techniques thus using and combining their abilities.
The adaptive neuro-fuzzy inference system was firstly introduced by Jang (Jang, 1993). It demonstrates the integration of fuzzy systems into the framework adaptive networks. ANFIS uses human knowledge in mapping input-output data and gradient decent and least squares (hybrid algorithm) on the generated data. The application of ANFIS in civil engineering consists in: modeling material properties (H Qian, B Xia, SZ Li, F Wang, 2002), permeability of granular soils ( (A Sezer, AB Goktepe and S Altun, 2009)), determine swell/shrink factor affecting earthwork optimization (A. B Goktepe, A. H Lav, S. Altun and G. Altuintas , 2008), prediction of sulfate expansion of PC mortar (G Inan, AB Goktepe, A Ramyar, 2007) etc.
The present study is aimed at demonstrating the modeling of traffic parameters, mainly flow and speed, using ANFIS. The developed models were trained, tested and validated for the Dogana-Kashar highway section, Tirana-Durres highway in Tirana. The objectives considered through this study were the developing of flow-speed ANFIS models, the validation of the models, performance evaluation and comparison of observed data with ANFIS modeled data.

[1] Corresponding author: aseitllari@epoka.edu.al

## 2. Adaptive Neuro-Fuzzy Inference System

Firstly introduced by Zadeh fuzzy set theory deals with the ambiguity accompanying the position of a variable. Fuzzy set is known to be a set that has no sharp boundaries. Contrary to crisp sets where the element is/isn't part of the set, belongingness (membership) degree is applicable in fuzzy sets.
Fuzzy set is represented by membership function which maps the membership degree of the elements within the set. The mapping process of membership functions to numerical values is set in the interval [0, 1] (Zadeh, 1965). The fuzzy inference system (FIS) is the process of establishing the mapping of given input to the output by the usage of fuzzy logic. FIS is composed of linguistic rules formulated on basis of experienced knowledge of the field (M. A. Mashrei, N. Abdulrazzaqb, T.i Y. Abdallac and M.S. Rahman, 2010).
Known as complementary technologies, fuzzy system and neural network help each other overcome their shortages. The learning capability of neural network supports in development of fuzzy systems by using the sufficient amount of data to analyze the system and generate fuzzy rules to a sufficient accuracy. Meanwhile the neuro-fuzzy system is linguistic based thus the integration of expert knowledge into the system might shorten the learning process. ANFIS is one of the popular techniques combining fuzzy inference system and back propagation algorithm (SR Jang, CT Sun and E Mizutani, 1997).
The ability of neural network to learn the fuzzy structure through the input data sets has rapidly increased the interest of researches in developing such techniques effectively. Including fuzzy associative memory, fuzzy aggregation modeling, fuzzy inference networks etc. (J Keller, R Krishnapuram and FCH Ree, 1992).
Adaptive Neuro-Fuzzy Inference System (ANFIS) arises as a practical method combining artificial neural network (ANN) and fuzzy inference system (FIS) methodologies (Jang and Sun, 1995). In this study the model is based on back-propagation learning algorithm and Sugeno type fuzzy inference system (Jang, 1993).
To have a simple view of ANFIS architecture: consider first order Sugeno fuzzy model based on two if-then rules expressed as:

Rule 1: *if x is A1 and y is B1, then* $f_1 = p_i x + q_i y + r_1$
Rule 2: *if x is A2 and y is B2, then* $f_2 = p_i x + q_i y + r_2$

where: x and y are the inputs, Ai and Bi are the linguistic labels, pi, qi and ri are output function's parameters.

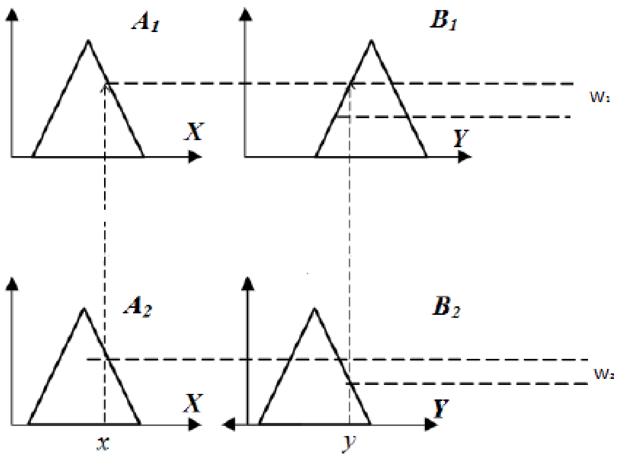



**Fig. 1.**
*Two rules Sugeno fuzzy model*

The corresponding ANFIS architecture in Fig. 2 belongs to the Fig. 1 mentioned rules, which is composed of five layers.

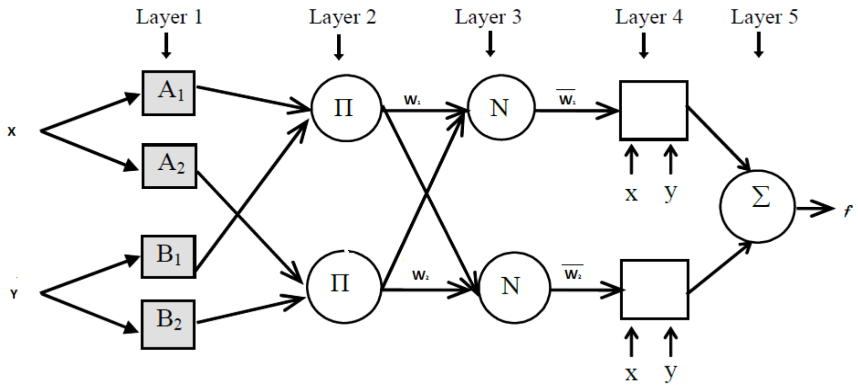



**Fig. 2.**
*ANFIS architecture*

The circular and square nodes correspond to fix and learned nodes. In general the nodes functions of the same layer stand for same functions family. Let see them below:

Layer 1. The nodes in this layer use the membership functions to assign the membership degree of crisp inputs which correspond to respective fuzzy sets. In this study Gaussian-shaped MFs are utilized:

$$\mu_A(x) = \exp\left\{-\frac{1}{2}\frac{(x-c_i)^2}{\sigma_i^2}\right\} \quad (1)$$

where $\mu_A$ is the membership function of $A_i$ fuzzy set and $c_i$ and $\sigma_i$ are the premise parameters.

Layer 2. In this layer the incoming signals from layer one are multiplied and send as outputs according to a firing strength represented as:

$$w_i = \mu_A(x) * \mu_B(y); i = 1,2 \quad (2)$$

Layer 3. The nodes in this layer determine the firing strength of $i^{th}$ rule to the sum of all rules firing strength.

$$w_i = \frac{w_i}{w_1 + w_2} \quad (3)$$

Layer 4. In this layer the effect of $i^{th}$ rules on the model function is calculated as:

$$w_i f_i = w_i\,(p_i x + q_i y + r); i = 1,2 \quad (4)$$

where $\overline{w_i}$ is the output of the $3^{rd}$ layer and $p_i$, $q_i$ and $r_i$ are known to be as consequent parameters.

Layer 5. In this layer there is a single node which computes the net output as the weighted average of all incoming signals from the $4^{th}$ layer calculated as:

$$f(x,y) = \sum \overline{w_i} f_i\,(p_i x + q_i y + r); i = 1,2\,) \quad (5)$$

Detailed explanation of ANFIS calculation is given in (Jang, 1993).

## 3. Methodology and Evaluation Criteria

In this study Matlab software is used for modelling with ANFIS to approximate real world forecasting, classifications and function approximation problems.
The study uses 1) Mean Absolute Deviation (MAD), 2) Root Mean Square Error (RMSE), 3) Correlation Coefficient (R) and 4) Coefficient of Determination ($R^2$) as evaluation criteria. Below these criteria are briefly explained:

### 3.1. Mean Absolute Deviation (MAD)

It is the average distance between the observed and estimated values. It is expressed as:

$$MAD = \frac{1}{n}\sum_{i=1}^{n} | x_i - \overline{x_i} | \qquad (1)$$

where $x$is the observed value, $\bar{x}$ is the estimated value and n is the number of observations.

### 3.2. Root Mean Square Error (RMSE)

It is also known as root mean square deviation (RMSD). It measures the difference between the predicted values and the observed values. It aggregates the residuals to a single measure. RMSE is expressed as:

$$RMSE = \sqrt{\frac{\sum_{i}^{n}(x_i - \bar{x}_i)^2}{n}} \qquad (2)$$

### 3.3. Correlation Coefficient (R)

It is the measurement of the strength and direction of the observed and predicted values. It is obtained by the division of covariance of these variables by their standard deviations product. R is expressed as:

$$R = \frac{\sum_{i=1}^{n}(x_i - \bar{x}) * (\hat{x}_i - \bar{\hat{x}})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 * \sum_{i=1}^{n}(\hat{x}_i - \bar{\hat{x}})^2}} \qquad (3)$$

### 3.4. Coefficient of Determination (R2)

The square of correlation coefficient is known as coefficient of determination. Describes the relative performance of variances of the observed and modeled data demonstrated by the linear fit. $R^2$ is expressed as:

$$R^2 = 1 - \frac{\sum_{j=1}^{n}(x_j - \hat{x}_j)^2}{\sum_{j=1}^{n}(x_j - \bar{x}_j)^2} \qquad (4)$$

## 4. Models using Adaptive Neuro-Fuzzy Inference System

The models development is based on traffic flow volume as input and traffic flow velocity as output, for Dogana-Kashar highway section.
Based on the correlation coefficient (R) and coefficient of efficiency ($R^2$) there were selected five models for the training, testing and validation. The models are trained, tested and validated using the input data of Dogana-Kashar highway section. Among the models, two of them were selected as representatives of the traffic flow velocity modeling of the highway section with the purpose of forecasting real time situations.

**Table 1**
*Flow-Speed ANFIS models*

| **Model No.** | **Best Model** | **R** | **R^2** | **RMSE** |
|---|---|---|---|---|
| 1 | ANFIS 1. | 0.942 | 0.887 | 3.856 |
| *2* | *ANFIS* 2. | *0.972* | *0.946* | *5.529* |
| 3 | ANFIS 3. | 0.865 | 0.749 | 6.303 |
| 4 | ANFIS 4. | 0.749 | 0.562 | 12.827 |
| 5 | ANFIS 5. | 0.835 | 0.698 | 7.483 |

The above table 1 lists all model's performances, which was based on Correlation Coefficient (R), Coefficient of Determination ($R^2$) and Root Mean Square Error (RMSE). Comparing the obtained results of correlation coefficient and coefficient of determination ($R^2$) for all five models, the second model ANFIS2 may be considered as the best model.

## 5. Conclusions

The present study demonstrated the usage of ANFIS methodology to model the velocity behavior of traffic flow, considering noise measurements. The efficiency of the methodology was appraised for further suggestions and application in prediction of the problem.
ANFIS modeling validation was evaluated using statistical evaluation criteria of Correlation Coefficient (R) and Coefficient of Determination ($R^2$). The results obtained are considered to be close to the observed data, thus resulting in an acceptable relationship between the volume and velocity of the traffic flow.
The application of ANFIS modeling in this study is a demonstration of a new approach to traffic flow parameters modeling techniques. Thus ANFIS may be considered to be an efficient mathematical tool to deal with the study of traffic flow parameters.

**References**

A Sezer, AB Goktepe and S Altun, n.d. Estimation of the permeability of granular soils using neuro fuzzy system.

A. B Goktepe, A. H Lav, S. Altun and G. Altuintas , 2008. Fuzzy Decision Support System to Determine Swell/Shrink Factor Affecting Earthwork Optimization of Highways. *Mathematical and Computational Applications*, 13(1), pp. 61-70.

G Dell'Acqua, R Lamberti, F Abbondanti, 2002. Adaptive neuro fuzzy inference system for highway accidents analysis, < www.iasi.cnr.it/ewgt/13conference/1_dellacqua.pdf>.

G Inan, AB Goktepe, A Ramyar, 2007. Prediction of sulfate expansion of PC mortar using adaptive neuro-fuzzy methodology. *Build Environ* , Volume 42, pp. 1264-1269.

H Qian, B Xia, SZ Li, F Wang, 2002. Fuzzy neural network modeling of material properties.. *Mater Process Tech*, Volume 122, pp. 196-200.

HM Chen, KH Tsai, GZ Qi, CS Yang, F Amini, 1995. *Neural network for structure control*, 9(2), pp. 168-176.

I Flood, NKartam , 1994. Neural networks in civil engineering, principle and understanding. *ASCE J Comput Civil Eng* , 8(2), pp. 131-148.

J Keller, R Krishnapuram and FCH Ree, 1992. Evidence aggregatin networks for fuzzy logic interface. *IEEE Trans. Neural Network* , 3(5), pp. 761-769.

Jang, J. S. R., 1993. ANFIS: Adaptive-Network-based Fuzzy Inference Systems. *IEEE Transactions on Systems, Man, and Cybernetics*, 23(3), pp. 665-685.

Jang, S., 1993. Adaptive network-based fuzzy inference system. *IEEE* , 23(3).

M. A. Mashrei, N. Abdulrazzaqb, T.i Y. Abdallac and M.S. Rahman, 2010. Neural networks model and adaptive neuro-fuzzy inference system for predicting the moment capacity of ferrocement members. *Engineering Structures*, 32(6), p. 1723–1734.

MA Hanna, D Ural, G Saygili, 2007. Evaluation of liquefaction potential of soil deposits using artificial neural networks. *Comput Aided Eng Softw* , 24(1), pp. 5-16.

MA Mashrei, N Abdulrazzaq, T Y Abdalla, MS Rahman, 2010. Neural networks and adaptive neuro-fuzzy inference system for predicting the moment capacity of ferrocement members. *Engineering Structures*, Volume 32, pp. 1723-1734.

SR Jang, CT Sun and E Mizutani, 1997. *Neuro-Fuzzy and Soft Approaching*. Prentice Hall.
Zadeh, L., 1965. Fuzzy Sets. *Information and Control*, pp. 338-353.

# ORIENTEERING PROBLEM: BEE COLONY OPTIMIZATION APPROACH

**Dragana Drenovac**[1], **Predrag Jovanović**[2], **Ranko Nedeljković**[3]
[1,2,3]*University of Belgrade Faculty of Transport and Traffic Engineering*

**Abstract:** In this paper the orienteering problem is solved by the metaheuristic technique Bee colony optimization (BCO). Literature review with applications and applied techniques are given. The proposed algorithm is applied to a problem in which a travelling salesman has a number of locations which he could visit, but he does not have time to visit all of them.

**Keywords**: Orienteering problem, Bee colony optimization, Travelling salesman.

## 1. Introduction

Orienteering is a sport which is a mixture of cross-country running and navigation through a forest, using a map and compass. A number of control points each with an associated score are placed in the forest and their locations are marked on the competitors' maps. Competitors start at intervals of, say, a minute and are required to visit a subset of the control points from the start point (node 1) so as to maximize their total score and return to the end point (node n) within a prescribed amount of time (Golden et al. 1987, Tsiligirides 1984).
In the literature (Laporte and Martello 1990, Gendreau et al. 1998, Thomadsen and Stidsen 2003) orienteering problem is also known as selective traveling salesman problem. In addition, the same problem in the literature (Butt and Cavalier 1994) can be found as the problem of maximum collection. Orienteering problem can be viewed as a variant of the traveling salesman problem with profits in which cost is referred to constraint and collected profit to objective function (Fillet et al. 2005).
Orienteering problem is solved by applying the Bee colony optimization. The proposed algorithm is used to solve a problem in which a travelling salesman has a number of locations which he could visit, but he does not have to visit all of them (Tsiligirides 1984). The problem was successfully solved by the developed metaheuristic algorithm.

## 2. Orienteering problem

Orienteering problem (OP) can be defined in the following way (Vansteenwegen at al. 2011): Given a graph $G=(V, A)$, where $V=\{v_1, \ldots, v_N\}$ is the vertex set and A is the arc set. The nonnegative score $S_i$ is associated with each vertex $v_i \in V$ and the travel time $t_{ij}$ is associated with each arc $a_{ij} \in A$. The OP consists of determining a Hamiltonian path $G' \in G$ over a subset of $V$, including preset start ($v_1$) and end ($v_N$) vertex, and having a length not exceeding $T_{max}$, in order to maximise the total collected score.
The OP can be formulated as an integer programming problem.

$$\max \sum_{i=2}^{N-1} \sum_{j=2}^{N} S_i x_{ij} \tag{1}$$

$$\sum_{j=2}^{N} x_{1j} = \sum_{i=1}^{N-1} x_{iN} = 1, \tag{2}$$

$$\sum_{i=1}^{N-1} x_{ik} = \sum_{j=2}^{N} x_{kj} \le 1; \ \forall \mathrm{k} = 2,\ldots,\mathrm{N-1}, \tag{3}$$

$$\sum_{i=1}^{N-1} \sum_{j=2}^{N} t_{ij} x_{ij} \le T_{\max}, \tag{4}$$

$$2 \le u_i \le N; \ \forall \mathrm{i} = 2,\ldots,\mathrm{N}, \tag{5}$$

$$u_i - u_j + 1 \le (N-1)(1 - x_{ij}); \ \forall \mathrm{i, j} = 2,\ldots,\mathrm{N}, \tag{6}$$

$$x_{ij} \in \{0,1\}; \ \forall \mathrm{i, j} = 1,\ldots,\mathrm{N}, \tag{7}$$

The following decision variables are used: $x_{ij} = 1$ if a visit to vertex $i$ is followed by a visit to vertex $j$ – 0 otherwise; $u_i$ is the position of vertex $i$ in the path.
The objective function (1) is to maximise the total collected score. Constraints (2) guarantee that the path starts in vertex 1 and ends in vertex $N$. Constraints (3) ensure the connectivity of the path and guarantee that every vertex is visited at most once. Constraint (4) ensures the limited time budget. Constraints (5) and (6) are necessary to prevent sub-tours.

---
[1] Corresponding author: drenovac@sf.bg.ac.rs

Golden et al. (1987) prove that the OP is NP-hard. This implies that exact solution algorithms are very time consuming and for practical applications heuristics will be necessary.
In literature (Vansteenwegen at al. 2011) the orienteering problem, the team orienteering problem (goal is to determine $P$ paths, each limited by $T_{max}$, that maximizes the total collected score), the orienteering problem with time windows and the team orienteering problem with time windows (each vertex is assigned a time window and a visit to a vertex can only start during this time window) could be found.
There are some interesting applications of the orienteering problem in the literature. The first application (Tsiligirides 1984) refers to the case when a salesman does not have enough time to visit all the cities. Based on expected sales in the nodes he tries to achieve as higher sale at available time.
The following problem (Golden et al. 1984) is a combination of inventory management problem and routing problem. The truck fleet delivers fuel to consumers on daily basis. Forecasted inventory levels can be seen as a measure of urgency for replenishment. First, the subset of users who urgently need fuel delivery on a certain day is chosen. That part of the problem is modeled as the OP. Score of the control points is a measure of urgency for restocking. Then, assignment of the trucks to consumers, as well as the effective paths for each truck are determined.
Orienteering problem found its application in tourism (Souffriau et al. 2008). Visiting a city or region, tourists usually do not have time for all of its attractions. Therefore, it is necessary to specify some of the locations that will be visited. Locations are characterized by attractiveness. The most attractive locations are selected in a given period of time.
Orienteering problem may have military applications (Wang et al. 2008). The submarine or unmanned aircraft can be sent in a reconnaissance action on a set of locations. At each location it can collect some information. The importance of information is the score associated with the location. It is necessary to visit a number of locations and collect data, taking into account the fuel or the available time.
In the literature, orienteering problem was solved by different techniques, from exact algorithms (Laporte and Martello 1990, Ramesh et al. 1992), through heuristics (Tsiligirides 1984, Golden et al. 1987) to metaheuristics - tabu search, genetic algorithms, ant colony optimization (Gendreau et al. 1998 Tasgetiren 2001, Liang et al. 2002), as well as artificial neural networks (Wang et al. 1995).

## 3. Bee colony optimization

The bee colony optimization metaheuristic (Teodorović 2008) was developed and used by Lučić and Teodorović (2001, 2003). Artificial bees represent agents, which collaboratively solve complex combinatorial optimization problems. Each artificial bee is located in the hive at the beginning of the search process, and makes a series of local moves, thus creating a partial solution. Bees incrementally add solution components to the current partial solution and communicate directly to generate feasible solutions. The best discovered solution of the first iteration is saved and the process of incremental construction of solutions by the bees continues through subsequent iterations.
Artificial bees perform two types of moves while flying through the solution space: forward pass or backward pass. Forward pass assumes a combination of individual exploration and collective past experiences to create various partial solutions, while backward pass represens return to the hive, where collective decision-making process takes place. It is assumed that bees exchange information and compare the quality of the partial solutions created, based on which every bee decides whether to abandon the created partial solution and become again uncommitted follower, continue to expand the same partial solution without recruiting the nestmates, or dance and thus recruit the nestmates before returning to the created partial solution. During the second forward pass, bees expand previously created partial solutions, after which they return to the hive in a backward pass and engage in the decision-making process as before. Series of forward and backward passes continue until feasible solutions are created and the iteration ends.

## 4. Solving the orienteering problem by bee colony optimization

Hive with artificial bees is located in node 1. Bees bounce from the hive and begin search process of area of allowable solutions. Let $V_i$ be the bee's benefit of choosing the $i^{th}$ node to serve. It is accepted that the bee's benefit when selecting a node is even greater if the score is greater and the travel time is shorter. Namely, it could be presented in the following way:

$$V_i = \frac{S_i}{{t_{li}}^{a}} \tag{8}$$

where $S_i$ is score in node $i$, $t_{li}$ is time required to reach node $i$, from the last selected node $l$ , while $a$ is parameter.
Let $p_i$ be the probability that a bee will choose the $i^{th}$ node. The route has to start at node 1 and end at node $n$. Before the node selection, a subset of those nodes that satisfy the maximal length of the route (their inclusion does not exceed the maximal length) has to be determined. Later, nodes are being chosen randomly from the subset.
Logit model was adopted as a choice model and the probability of selection is:

$$p_i = \frac{e^{V_i}}{e^{V_1} + e^{V_2} + \ldots + e^{V_m}} \tag{9}$$

where $m$ is the cardinal number of the subset. The formula (9) shows the greater the bee's benefit from the choice of a node the higher the probability of selection of that node.

During the first forward pass each bee chooses a predefined number of nodes. Having returned to the hive, bees start to communicate. They calculate values of the objective function as the ratio of the total score and the total time required to visit chosen nodes and compare their partial solutions. In the proposed algorithm, the objective function of each bee comprises time because the greater route efficiency the greater number of included nodes and the greater achieved score. Then, bees make decision about their loyalty to the solution (whether to keep the solution or to abandon it).

Let $\Pi_j$ be the objective function value generated by the $j^{th}$ bee ($j=\overline{1,b}$, where $b$ is the number of bees). Let $\Pi_{norm_j}$ be the normalized value of the objective function value. It is calculated as follows:

$$\Pi_{norm_j} = \frac{\Pi_j - \Pi_{\min}}{\Pi_{\max} - \Pi_{\min}}, \ \Pi_{norm_j} \in [0,1], \ j=\overline{1,b} \tag{10}$$

Where $\Pi_{\max}$ and $\Pi_{\min}$ are the minimum and the maximum value of the objective function.

The probability that the $j^{th}$ bee will be loyal to its partial solution at the beginning of the next forward pass is calculated as follows:

$$p_j^{u+1} = e^{-\frac{\Pi_{norm_{\max}} - \Pi_j}{u}}, \ j=\overline{1,b} \tag{11}$$

where $u$ is ordinal number of the forward pass and $\Pi_{norm_{\max}}$ is maximal normalized value of the objective function.

Relation (11) shows that the bee, which generated the best partial solution, in the next forward step will stay loyal to its solution, since $p_j^{u+1} = 1$. The ordinal number of forward pass also affects this possibility. The more forward passes bees made the less freedom to abandon partial solutions already generated and to stay uncommitted.

After making decisions about the loyalty to their solutions, bees gather in the dance floor area. The bees that decide to keep their partial solutions start dancing and thus recruit uncommitted bees. Uncommitted followers choose which of the loyal bees to follow in the next forward pass.

During the search process time is updated and when bees can no longer expand the solution, the search ends. Having coming back to the hive after an arbitrary forward pass, the bees can create partial as well as final solutions. Despite of that, all of them participate in information exchange, evaluation of the solutions and decision-making concerning the loyalty of the solutions.

If the number of loyal bees is equal to $r$ before the next forward pass, the probability that the uncommitted bee will join the $k^{th}$ loyal bee in the next forward pass is equal to:

$$p_k = \frac{e^{\Pi_{norm_k}}}{e^{\Pi_{norm_1}} + e^{\Pi_{norm_2}} + ... + e^{\Pi_{norm_r}}}, \ k=1,\dots,r \tag{12}$$

Based on these probabilities, uncommitted bees are coupled with committed bees. The bees then start a new forward pass flying together to the last node of the partial solutions generated by the committed bees. After that, each bee individually extends its partial solution.

Each iteration gives a particular solution. The best solution obtained during the pre-specified number of iterations is selected.

## 5. Numerical example

In order to demonstrate the proposed metaheuristic approach, the following problem was considered. It occurs when a travelling salesman (Tsiligirides 1984) has a number of locations which he could visit, but he does not have time to visit each and every one. He knows the number of sales he can expect to make in each location and therefore he wishes to plan his route so as to maximize his total number of sales, whilst keeping the total length of his route within the distance he can travel in one day (or week). This formulation could be referred to as a selective travelling salesman problem in which the locations take the role of the control points and the numbers of sales take the place of the scores. The objective is stated in terms of maximizing the sales (or scores), so that the time needed to visit the corresponding nodes is the smallest one which does not exceed the maximum time available. The above problem could be considered, alternatively, by placing bounds upon the length of the distance of each route. In this case, the objective is to find the shortest route when the length of the distance which will be covered does not exceed a predetermined value $T_{max}$.

The benchmark problems, 33-node network and 2 out of 20 examples ($T_{max}$=60 and $T_{max}$=80), were taken from the following internet address: www.mech.kuleuven.be/cib/op, and modified by dividing $T_{max}$ and distances by 10, without loosing the original example. Then, it is accepted that the measure of time is [h], and consequently $T_{max}$=6 h and $T_{max}$=8 h. It means that a travelling salesman has 6 h (the first numerical example) or 8 h (the second numerical example) in one day to visit a subset of 33 locations (nodes), each with assigned score representing the number of sales. The solutions obtained by the proposed algorithm are given in Table 1 and Table 2.

**Table 1**
*The solution of the first numerical example*

| | |
|---|---|
| **$T_{max}$** | 6 h |
| **Route length** | 5.99 h |
| **Score** | 580 |
| **Route** | 1, 24, 22, 7, 5, 14, 4, 20, 17, 16, 21, 15, 13, 3, 6, 2, 8, 31, 12, 29, 30, 26, 33 |

**Table 2**
*The solution of the second numerical example*

| | |
|---|---|
| **$T_{max}$** | 8 h |
| **Route length** | 7.97 h |
| **Score** | 710 |
| **Route** | 1, 24, 22, 7, 5, 28, 14, 4, 20, 17, 16, 15, 13, 3, 6, 2, 8, 31, 12, 29, 30, 11, 19, 18, 10, 9, 25, 33 |

The solutions can be compared to those obtained by the metaheuristic Ant colony optimization (Liang and Smith 2006). It could be noticed that both metaheuristics give the same results in scores.
Routes obtained by applying the BCO are given on Figure 1 and Figure 2.

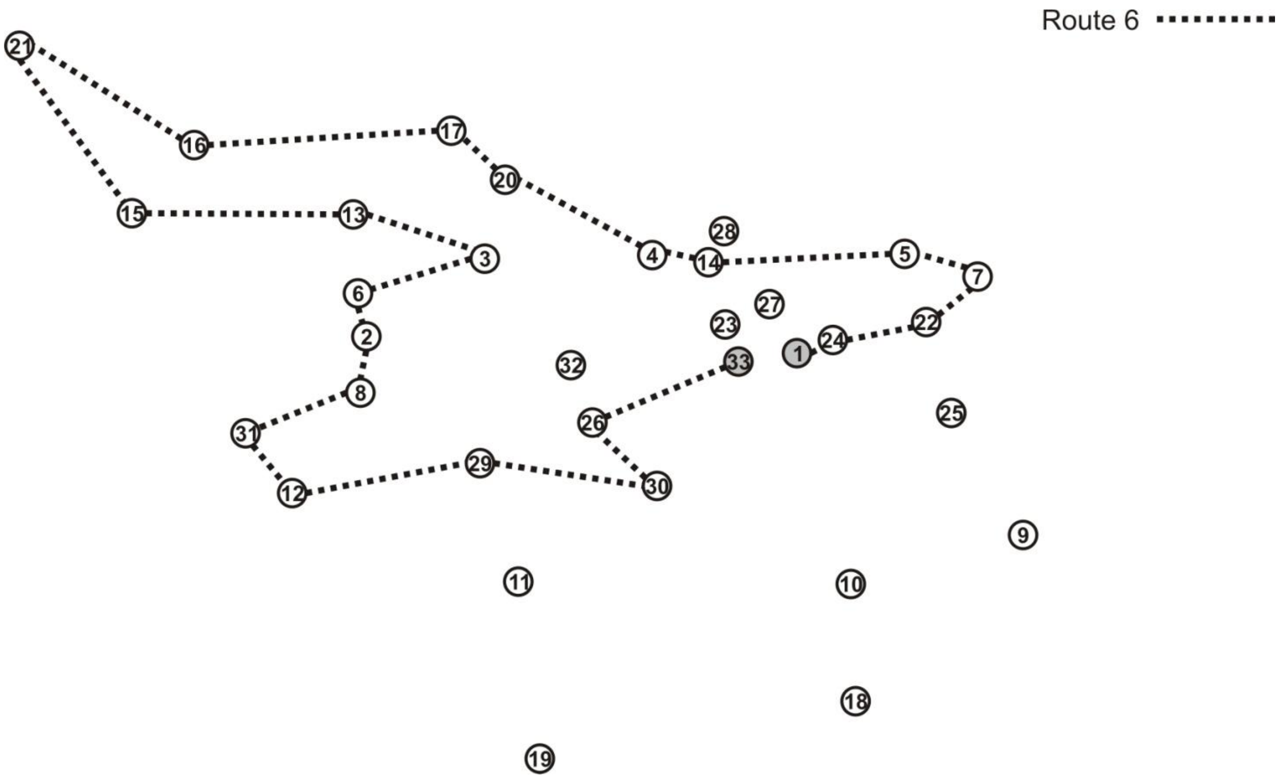



**Fig. 1.**
*Obtained route ($T_{max}$=6 h)*

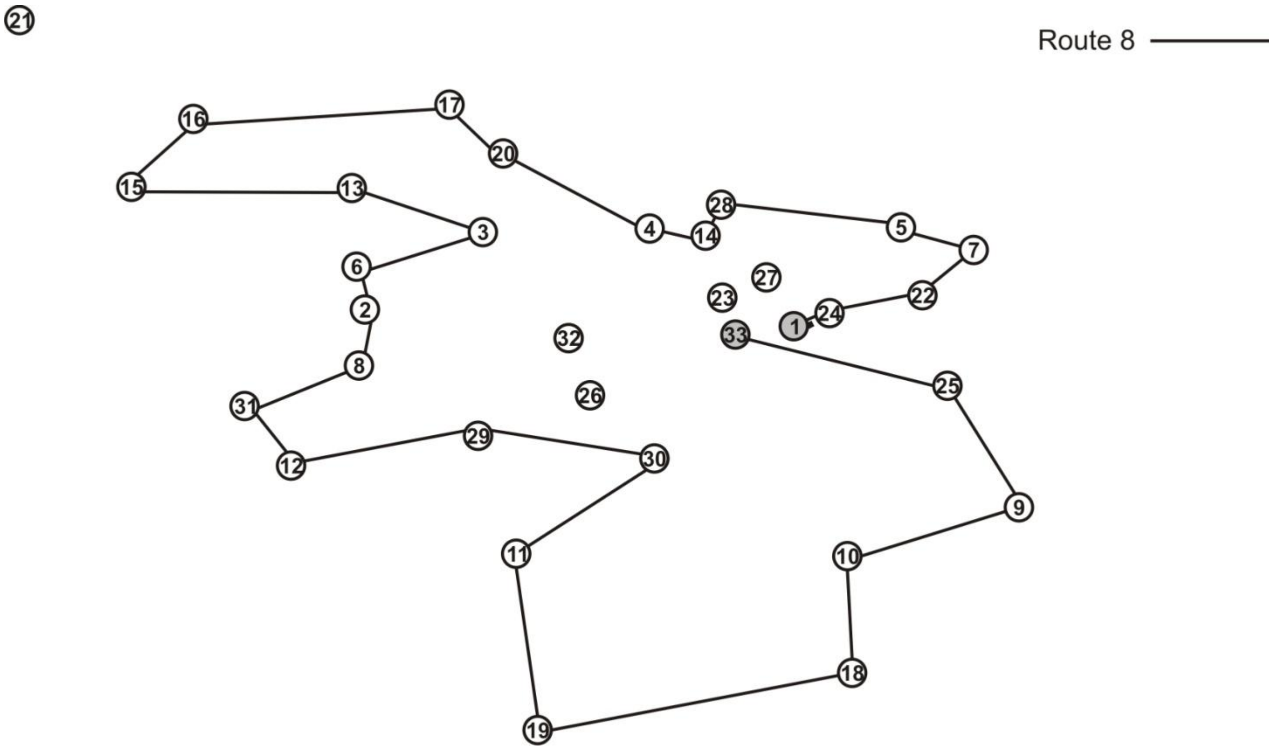



**Fig. 2.**
*Obtained route ($T_{max}$=8 h)*

## 6. Conclusion

In this paper orienteering problem is solved by applying the Bee colony optimization. The problem when travelling salesman chooses to visit a subset of given nodes is solved by the proposed algorithm. The developed algorithm, used for solving orienteering problem for the first time, yields acceptable results. It also might be applied to all those problems which are possible to be modelled by orienteering problem.

**References**

Butt, S.; Cavalier, T. 1994. A heuristic for the multiple tour maximum collection problem, *Computers and Operations Research*, 21: 101–111.

Fillet, D.; Dejax, P.; Gendreau, M. 2005. Travelling salesman problem with profits, *Transportation science*, 39(2): 188-205.

Gendreau, M.; Laporte, G.; Semet, F. 1995. *A branch-and-cut algorithm for the undirected selective travelling salesman problem*. Publication CRT-95-80. Centre de recherche sur les transports. Montreal. Canada.

Gendreau, M.; Laporte, G.; Semet, F. 1998. A tabu search heuristic for the undirected selective travelling salesman problem, *European Journal of Operational Research*, 106: 539–545.

Golden, B. L.; Assad A.; Dahl R. 1984. Analysis of a large scale vehicle routing problem with an inventory component, *Large Scale Systems*, 7: 181–190.

Golden, B. L.; Levy L.; Vohra R. 1987. The Orienteering Problem, *Naval Research Logistics*, 34: 307-345.

Laporte, G.; Martello, S. 1990. The selective travelling salesman problem. *Discrete Applied Mathematics*, 26: 193–207.

Liang, Y.; Kulturel-Konak, S.; Smith, A. 2002. Meta heuristics for the orienteering problem. In *Congress on Evolutionary Computation*, Hawaii, Honolulu, 384–389.

Liang, Y., S.; Smith, A. 2006. An Ant Colony Approach to the Orienteering Problem, *Journal of the Chinese Institute of Industrial Engineers*, 23(5): 403-414.

Lučić, P.; Teodorović, D. 2001. Bee system: modeling combinatorial optimization transportation engineering problems by swarm intelligence. In: *Preprints of the TRISTAN IV Triennial Symposium on Transportation Analysis*, Sao Miguel, Azores Islands, Portugal, 441–445.

Lučić, P.; Teodorović, D. 2003. Computing with Bees: Attacking complex Transportation Engineering Problemsl *International Journal on Artificial Intelligence Tools*, 12(3): 375-394.

Ramesh, R.; Yoon, Y.; Karwan, M. 1992. An optimal algorithm for the orienteering tour problem, *ORSA Journal on Computing*, 4: 155–165.

Souffriau, W.; Vansteenwegen, P.; Vertommen, J.; Vanden Berghe, G.; Van Oudheusden, D. 2008. A personalised tourist trip design algorithm for mobile tourist guides, *Applied Artificial Intelligence*, 22 (10): 964–985.

Tasgetiren, M. 2001. A genetic algorithm with an adaptive penalty function for the orienteering problem, *Journal of Economic and Social Research*, 4 (2): 1–26.

Teodorović, D. 2008. Swarm intelligence systems for transportation engineering: Principles and applications, *Transportation Research Part C*, 16(6): 651-667.

Thomadsen, T.; Stidsen, T. 2003. The quadratic selective travelling salesman problem. Informatics and Mathematical Modelling, *Technical Report*, 17, Technical University of Denmark.

Tsiligirides, T. 1984. Heuristic methods applied to orienteering, *Journal of the Operational Research Society*, 35(9): 797-809.

Vansteenwegen, P.; Souffriau, W.; Van Oudheusden, D. 2011. The orienteering problem: A survey. *European Journal of Operational Research*, 209: 1–10.

Wang, Q.; Sun, X.; Golden, B.; Jia, J. 1995. Using artificial neural networks to solve the orienteering problem, *Annals of Operations Research*, 61: 111–120.

Wang, X.; Golden, B.; Wasil, E. 2008. Using a genetic algorithm to solve the generalized orienteering problem. In: B. Golden, S. Raghavan, E. Wasil (Eds.), *The Vehicle Routing Problem: Latest Advances and New Challenges*. New York: Springer. 263–274.

# 3D DESIGN SOLUTION AND SIMULATION OF A WAREHOUSE, CASE OF FREE ZONE CITY OF NIŠ

[1]Vojislav Tomić, Zoran Marinković, Boban Nikolić, Miloš Milošević
[1] *Faculty of Mechanical Engineering, Center for engines and motor vehicles, Niš*

**Abstract:** The choice of warehouse technology and variant design of warehouse system are complex tasks demanding special attention. The paper puts focus onto the procedure of developing a technical project 3D design and simulation solution of a small railway-highway Container terminal warehouse in the free zone - city of Nis. An analytical approach to the issue implies the application of the method of planning, modeling, simulation and analysis of warehouse systems. The methodology gains on significance when used with developed software packages in order to perform a number of computer simulations. Software package Flexim has been used in the work to address the variant warehouse design with selective pallet racking. The methodology enables rather inexpensive variant design of warehouse systems and a relatively simple choice of optimal solution.

**Keywords**: *Container terminal, design, railway, road, logistics.*

## 1. Introduction

Container Terminal (CT) is a facility where cargo containers are transshipped between different transport vehicles, for onward transportation (Arnold, 1998). CT and ther warehauses have one of the most important roles in the logistic supply chains, i.e. in transport networks. Various types of transport meet each other at CT warehauses, where also transformation of material flows is accomplished (Lippolt, 2005), (Zečević, 2006), (Marinković et al., 2012), (Aiying et al., 2009). One of the crucial factors for a highly efficient distribution network is a suitable choice of CT warehouse design. The main aim of this paper is to find the most suitable design for the CT warehouse.
The paper gives the procedure of developing a technical project design solution of a public CT warehouse in the free zone - city of Nis. One of the approaches to the issue implies the application of the method of planning, modeling, simulation and analysis. The methodology gains on significance when a number of computer-based simulations are performed, i.e. when a variant design of CT warehouse system is performed with the aim of determining the optimal solution. An example of a CT warehouse with selective pallet racking demonstrates the procedure of variant warehouse design. The standard procedure of warehouse evaluation, a PC and software package *Flexim* are used. In this paper structural and parametric solutions of all necessary elements of the CT warehouse are given, including the design of a public warehouse. Such a system, offers a great synergic effect as well as a great influence onto the industrial development in the city of Nis and southern part of Serbia. (Marinković et al., 2011), (Marković et al., 2011), (Vasić, 2012).

## 2. Literature review

The warehouse layout depends on many factors e.g: the items stored, space available, height, the layout of road, and rail tracks around the warehouse, etc. Because of that design planning and location of warehouse has been formally studied and researched as a discipline since the mid 1950s. In literature, Simon (Simon, 1975) suggests that the layout problem is a design problem but the location problem is an optimization problem. On the other hand, according to Francis et al (Francis et al., 1992) location problem should be treated as a design problem not an optimization problem. In our opinion, the layout and location problems have elements of both design and optimization problems. A confirmation of our thinking we found in the source (Sunderesh, 2008). In any case, all authors agree in one, warehousing is a time-consuming activity that does not add value. However, the need to provide better service to customers and be responsive to their needs appears to be the primary reason to have warehouse (Kulwiec, 1980). Based on all stated above, it can be concluded that designing a warehouse is a complex problem. Because of that, Rouwenhorst et al. (Rouwenhorst, et al., 2000) classify the warehouse design and planning problems into three levels of decisions – strategic, tactical and operational. In this paper, we used the strategic decision as a relevant decision.

## 3. Container terminal design

There are several reasons for building and operating warehouses for CT. In many cases, the need to provide better services to customers and to responsive their needs appears to be the primary reason. Although it may seem that the only function of a warehouse is warehousing, that is, temporary storage of goods but in reality many other functions are performed. Because of that, characteristics of flows, which pass through of the CT warehouse, are necessary to analyze before design process. The first step in the design of a CT warehouse is to determine the general flow pattern for material, parts and work in process through the system. Based on the aforementioned, the material flows model of CT warehouse in Niš is analyzed in this paper and it is shown in fig 1. Flow pattern refers to the overall pattern in which the product flows from beginning to end. Beginning started from raw material at the receiving stage (goods are unloaded from wagons), through storage of goods in racks and finally distributed goods in trucks.

[1]Corresponding author: Vojislav@masfak.ni.ac.rs

According to the presented flows model of CT warehouse in Niš, the load is delivered to the CT warehouse in transport logistic units (TLU). The load is delivered using both road and rail sides of the terminal. The load arriving in TLU can be directly transformed into the final output TLU using the fast line. The goods that arrive in inadequate TLU, or which are not for any other reason ready to be shipped directly, are sent to the section where TLU are built or broken (work stations).

Such a great amount of load is consolidated into TLU in the construction section using workforce at work stations. After building of an adequate TLU, the goods are directly stored or shipped using the fast line to the output transport.

If the input transport is in inadequate transport logistic units, it is necessary to break it at work stations. At these stations, load is separated, sorted, scanned, and wrapped in thermal shrinking foil, building a compact warehouse TLU. Such new TLU is then stored, and the goods a wait for the commissioner to remove them.

When the commissioner receives an order for removing a certain quantity of goods, TLU are sent to the TLU construction zone, where new TLU are built in accordance with the picking list. As mentioned above, certain TLU do not need to be sent to the construction section, but they are directly moved to the storing zone for shipment preparation. In the same way, certain transloaded transports can be moved directly to the next connection transport over the fast line without the need to break and build TLU again. This happens only in the case when the connection route is compatible with the previous one, that is, when it requires the same TLU containing the exact same goods.

The possible goods at the work stations in this model are:

- output goods that need to be built,
- input goods that need to be broken,
- and transloaded goods that need both to be broken and built.

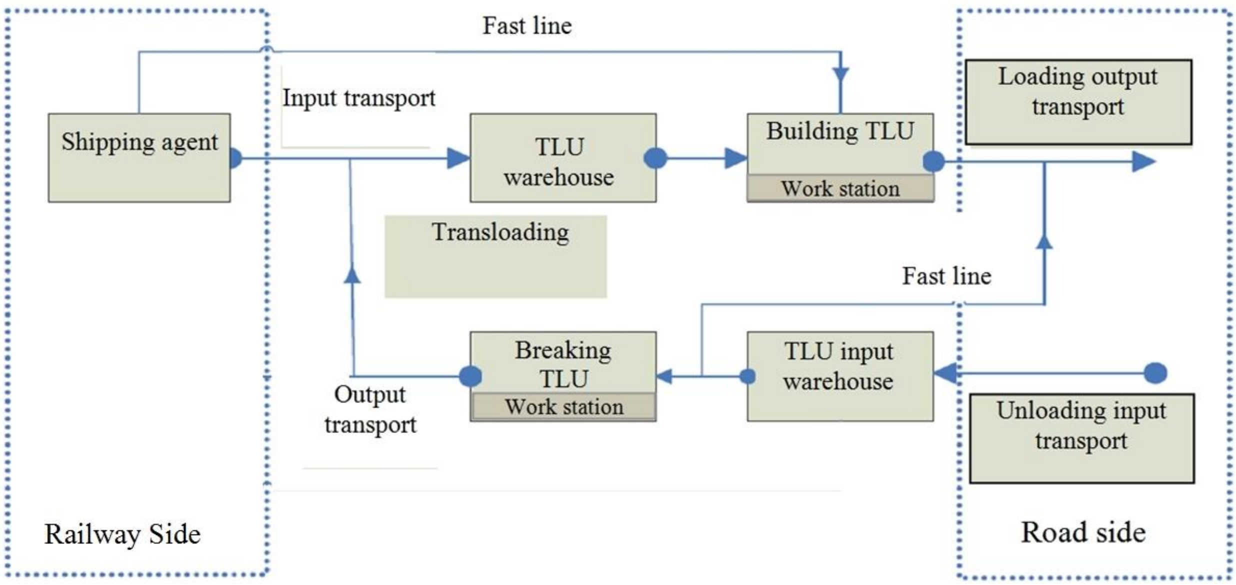



**Fig 1.**
*Material flows model of CT warehouse in Free zone Niš*

**4. Variant design and simulation of storage processes**

On the basis of general flow in fig 1 we present a design solution of a zone covered by selective pallet regales in fig 2. For that zone 3D simulations are done by *Flexim* software tools. The modern approach to storage design, aiming at efficient, economic and safe systems, demands the application of the method of planning, modeling, simulation and analysis. The task of planning a system implies defining procedures for obtaining the solution starting from the proposed objective and up to its realization. Modeling and simulation imply description, emulation and analysis of real systems by means of their equivalent and mathematical models. This methodology requires development of variant solutions, which implies that for a given set of criteria, i.e. objective function and constraints, most suitable solutions are extracted. Nowadays this problem is successfully solved by developed software packages and a PC.

In this paper, an example is used to demonstrate the procedure of variant storage design, i.e. the design of the warehouse zone with selective pallet racking. For approximate dimensions of the storage zone and exactly defined standard storage unit, under the constraint of using specific transportation machinery for storage processes, four variants have been modeled. The objective is to choose the optimal solution based on the analysis of storage capacity (*UBSJ*), surface-($\eta_A$) and volume- ($\eta_V$) usage coefficients. The problem is solved by software package *Flexim* and a PC.

The properties of the storage zone are:
length $Lsz = 20\ m$ and
width $Bsz = 12\ m$,
while the height is a variable and it depends on the rack height and the type of transport machine. The standard storage unit has:
dimensions $800 \times 1200 \times 1500$ mm ,
surface $A_{SJ} = 1.2 \cdot 0.8 = 0.96\ m^2$ and
volume $V_{SJ} = 1.2 \cdot 0.8 \cdot 1.5 = 1.44\ m^3$ .
For such two storage units, the dimensions of storage cell are: length *2 m*, height *1.6 m* , and depth *1.2 m*.

### 4.1 The variant with low-height rack and a head-on forklift

With this storage variant, a head-on forklift is used as a transport machine, with the lift height of *5.0 m*. In this case the bay width is $b_k = 3.4\ m$, so that 4 selective racks can be used, with the height of $H^* = 6.4\ m$. According to the dimensions of the storage cell, these racks have:
$B_k = 20/2.0 = 10$ columns
and $B_r = 6.4/1,6 = 4$ rows,
i.e. the overall number of storage cells is
$B_k \cdot B_r = 10 \cdot 4 = 40$.
Within one rack $Bk \cdot Br \cdot nsj = 10 \cdot 4 \cdot 2 = 80$ storage units can be stored, so that the capacity of this variant is $UBSJ = NR \cdot Bk \cdot Br \cdot nsj = 4 \cdot 10 \cdot 4 \cdot 2 = 320$. Fig. 3 shows the layout of storage bay, while Fig. 3 gives the 3D-model of a storage zone for this variant, modeled by means of *Flexim*.

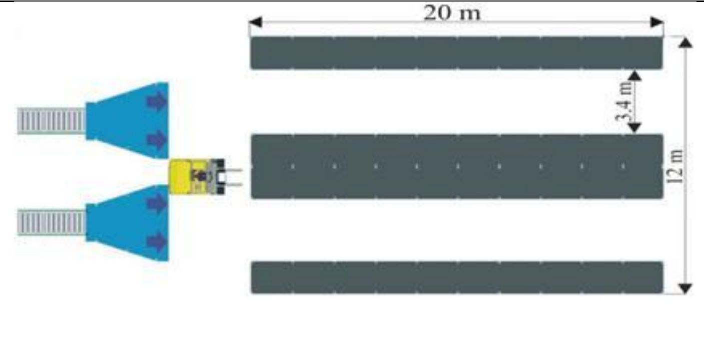



**Fig. 2**
*Modeled layout of the rack bays – 1st variant*

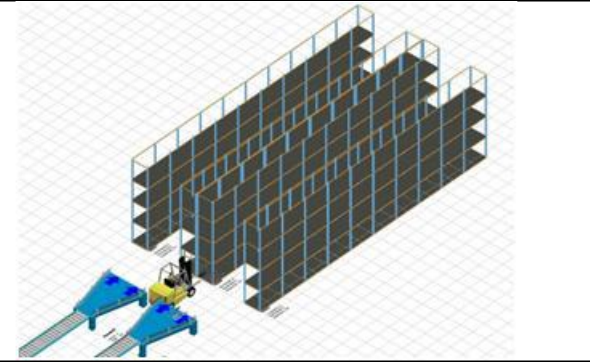


**Fig. 3**
*3D-model of the storage zone – 1st variant*

### 4.2 The variant with low-height rack and a sideways forklift

In order to reduce the bay width and to enlarge the rack height, a forklift with side lift, the height of which is *8.0 m*, is chosen in the second variant. The bay width is now $bk = 2.0\ m$, so that the width of the storage zone for 4 racks ($NR = 4$) is less compared to the 1st variant, and it is $Bsz = 9.2\ m$.
Due to larger lift height, the rack height is $H^* = 9.6\ m$, thus the number of rows of storage cells increased to $Br = 9.6/1,6 = 6$, while the number of columns remained the same $Bk = 20/2.0 = 10$. The capacity is now calculated to be $UBSJ = NR \cdot Bk \cdot Br \cdot nsj = 4 \cdot 10 \cdot 6 \cdot 2 = 480$ storage units. Fig. 4 depicts the modeled layout of rack bays for the 2nd variant.

### 4.3 The modified variant with low height rack and a forklift with side lift

This variant is distinguished from the 2nd variant by the fact that is uses one more bay, i.e. the number of racks is 6, and the rest is just the same. In this case the width of the storage zone has increased to $Bsz = 13.8\ m$. Since the number of racks with the same properties as in the 2nd variant has increased, the capacity of this variant has enlarged to $UBSJ = NR \cdot Bk \cdot Br \cdot nsj = 6 \cdot 10 \cdot 6 \cdot 2 = 720$ storage units. Fig. 5 shows the modeled layout of rack bays for the 3rd variant.

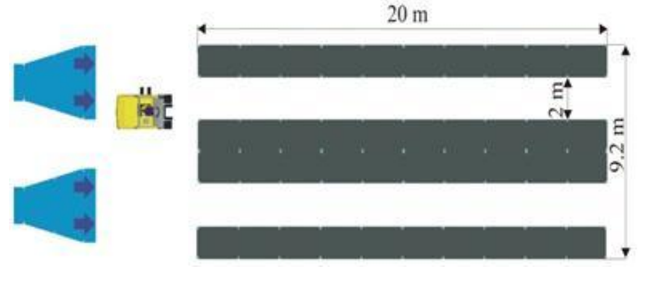



**Fig. 4**
Modeled layout of rack bays – 2nd variant

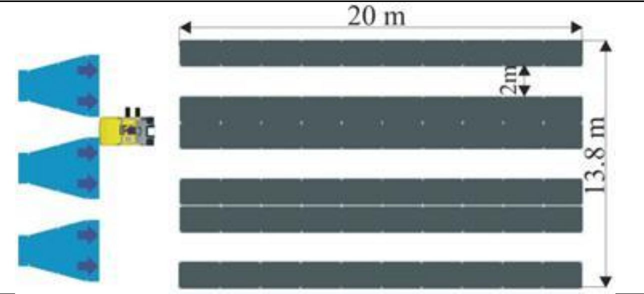



***Fig. 5***
*Modeled layout of rack bays – 3rd variant*

### 4.4 The variant with high racks and rack-cranes

With this variant, the dimensions are kept the same as in the 1st variant, i.e. $Lsz = 20$ *m*, $Bsz = 12$ *m*. Six high racks ($NR = 6$), served by the rack-cranes, can be placed in the warehouse. Racks have the height of $H^{*}=16$ *m*, while the bay width is $bk = 1.4$ *m*. The number of columns in the rack is $Bk = 20/2.0 = 10$ and the number of rows $Br = 16/1{,}6 = 10$, hence the capacity of $UBSJ = NR{\cdot}Bk{\cdot}Br{\cdot}nsj = 6{\cdot}10{\cdot}10{\cdot}2 = 1200$ storage units. Fig. 6 depicts the modeled layout of rack bays for the 4th variant.

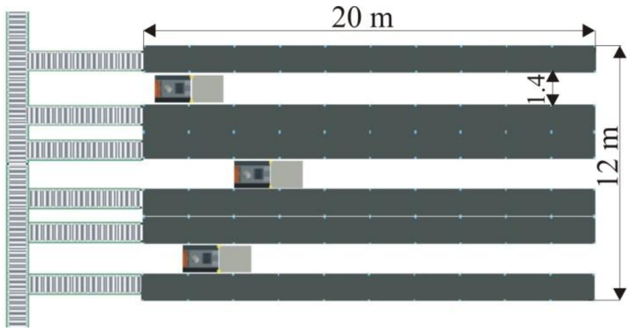



**Fig. 6**
*Modeled layout of rack bays – 4th variant*

Table 1 gives the relevant properties of the four variants with selective pallet racking. It would be interesting to compare the capacities, given as the overall number of storage units (Eq. (1)), surface-usage coefficient (Eq. (2)) and volume-usage coefficient (Eq. (3)) of the storage zone. The two coefficients are determined for the storage zone with the surface of $A_{SJ} = 0.96\ m^2$ and the volume of $V_{SJ} = 1.44\ m^3$. Obviously, the 4th variant yields the greatest capacity of UBSJ = 1200 storage units and the greatest values for the surface- usage, $\eta_A = 4.80$, and the volume-usage, $\eta_V = 0.45$, coefficients.

***Table 1.***
*Evaluation properties of the four designed storage variants*

| Variant | Storage zone | | | Corridor | Rack | | Capacity | Coefficient of use | |
|---|---|---|---|---|---|---|---|---|---|
| | Lsz (m)<br>Bsz(m) | Hsz =<br>H*(m) | Asz(m²)<br>Vsz(m³) | NK<br>Bk (m) | NR | Br<br>Bk | UBSJ<br>(nsj = 2) | surface<br>$\eta_A$ | volume<br>$\eta_V$ |
| **1.** Short racks and head-on fork lift | 20<br>12 | 6.4 | 240<br>1536 | 2<br>3.4 | 4 | 10<br>4 | 320 | 1.30 | 0.30 |
| **2.** Short racks and sideways fork lift | 20<br>9.2 | 9.6 | 184<br>1766.4 | 2<br>2.0 | 4 | 10<br>6 | 480 | 2.50 | 0.39 |
| **3.** Short racks and sideways fork lift | 20<br>13.8 | 9.6 | 276<br>2649.6 | 3<br>2.0 | 6 | 10<br>6 | 720 | 2.50 | 0.39 |
| **4.** High racks and storage and retrieval | 20<br>12 | 16 | 240<br>3840 | 3<br>1.4 | 6 | 10<br>10 | 1200 | 4.80 | 0.45 |

The choice of the optimal variant based only on capacity, surface- and volume-usage coefficients would be inadequate und could lead towards a wrong decision. The final decision has to include the economic analysis of expenses for each solution, which includes the costs for transport machinery, racks and buildings. In this case, the difference in price for high racks and low-height racks does not play a major role. However, there is an obvious difference in the price of transport machinery. Most inexpensive is the universal head-on forklift, sideways forklift is somewhat more expensive and the high-rack crane is most expensive. Universal head-on forklift offer the advantage of their additional employment in other areas, whereas the rack-crane cannot leave the storage zone. Warehouses with low-height racks demand additional objects (buildings), while high-racks are actually the carrying structure of the warehouse and are used to carry side panels as walls as well as the roof structure. The solution based on high-rack storages, which can be used in automatic mode, is justified in the case of great number of articles that need to be handled and if the inventory is rapidly depleted and restocked.
As can be noticed, *Flexim* enables efficient modeling of various variants of storage systems and their convenient 3D representation, which is important for the choice of necessary equipment. The software package also provides the possibility of simulating the work of the equipment, i.e. simulating the storage processes, and their convenient presentation in a video format. By varying the parameters of transportation machinery, such as transportation velocity, accelerations, decelerations, lengths of transport patterns, etc., it is possible to analyze how they influence the choice of optimal variant.

## 5. Conclusion

The following conclusions can be drawn from the preceding analysis:

- warehouse is an important subsystem of the logistic concept of industry which provides a link between acquisition, production, distribution and sale of material and goods,
- design of warehouse systems is a preveiling topic in logistics that needs to be addressed from several aspects with the aim of achieving efficient, inexpensive and safe warehousing,
- a significant number of typical warehouse technologies has been developed and the appropriate choice is affected by a number of influencing factors, which need to be analyzed,
- technology of storing goods in selective pallet racks, especially in high-rack automatized warehouses, is nowadays quite frequent due to a number of advantages it offers in the case of a quick turnover and a large asortiment,
- high-rack warehouses, with a height of up to 40 m and length of up to 200 m, offer large capacities and surface-usage coefficients ($\eta_A \approx 10$),, while the volume is less utilized ($\eta_V \approx 0.5$)
- software package *Flexim* demonstrates its full potential if a variant solutions are investigated, whereby the solution parameters are varied with the aim of obtaining the solution with optimal technical parameters.

**References**

Arnold, D. 1998. *Materialflusslehre*, Vieweg-Verlag, Springer. Translation on Serbian TEMPUS 17019. 2004.

Aiying, R.; Martin, G. 2009. Shift designs for freight handling personnel at air cargo terminals, *Journal of transportation research*. DOI: http://www.sciencedirect.com/science/article/pii/S1366554509000180, 45(5): 725-739.

Francis, R. L.; McGinnis, L. F.; White, J. A. 1992. *Facility layout and location*, An analytical approach (2nd ed.). Englewood Cliffs, NJ, US: Prentice-Hall.

Kulwiec, R.A. 1980. *Advanced Material Handling*, The Material Handling Institute, Charlotte, NC.

Lippolt, C. 2005, *Lager-und Distributations systeme*, Vorlesung, IFL, Universitaet Karlsruhe, Karlsruhe.

Marinković, Z.; Marinković, D., Tomić, V. 2012. Concept solution of logistic center in the free zone Pirot, *International Conference MHCL 2012*, 227-230.

Marinković, Z.; Marinković, D.; Tomić, V.; Marković, G. 2011. Importance of establishing a logistic (Cargo) center in Nis and his influence on regional economy development, *Serbian fourth symposium with international participation, TIL 2011*, DOI: http://ttl.masfak.ni.ac.rs/til2011/ZR/TIL2011_P6.pdf, 31-38.

Marković, G.; Gašić, M.; Marinković, Z.; Tomić, V.; 2011. Essence and importantance of establishing a regional logistic concept: strategic directions of development, *Serbian fourth symposium with international participation TIL 2011*, DOI: http://ttl.masfak.ni.ac.rs/til2011/ZR/TIL2011_P2.pdf, 18 (3), 7-14.

Rouwenhorst, B.; Reuter B.; Stockrahm V.; Houtum, G.J; Mantel, R.J.; Zijm, W.H.M. 2000. Warehouse design and control: Framework and literatures review, *European Journal of Operational Research*, DOI: http://www.sciencedirect.com/science/article/pii/S037722179900020X, 515-533.

Simon, H.A. 1975. *Style in design in C.M.*, Spatial Synthesis in Computer-Aided Building Design, Wiley, New York.
Sunderesh, S. 2008. *Facilities Design third edition*, Taylor & Francis Group, New York.

Vasić, G. 2012. Inženjerski pristup dizajniranju distributivnog centra u okviru slobodne zone Pirot *(The engineering approach to design a distribution center in the Free Zone Pirot)*, Master diploma thesis, Faculty of Mechanical Engineering Nis.

Zečević, S. 2006. Robni terminali i robno-transportni centri *(Freight terminals and material-transport centers)*, Transport and Traffic Engineering, University of Belgrade, Belgrade.

*Transport Modeling and Decision Making*

# MOBILITY IMPACTS OF LINE 1 OF THE METRO OF NAPLES: THE CASE STUDY OF GARIBALDI STATION

**Francesca Pagliara[1], Claudio Troisi[2], Luca Pelliccia[3]**
*[1,3]Department of Civil, Architectural and Environmental Engineering –University of Naples Federico II*
*[2]Department of Architecture –University of Naples Federico II*

**Abstract:** In this paper the focus will be on the area surrounding Garibaldi square in Naples in Italy following the inauguration of the station Garibaldi of Line 1 of the metro in January 2014. Indeed, many interventions of different types have been carried out, with a subsequent reconfiguration of the whole transport system connecting this node.
Private and public transport needed to be reorganized after the inauguration of the new station with the objective of making mobility efficient as a function of the transport demand on the network.
In order to evaluate both transport demand and supply, it has been necessary to undertake two surveys: the first one aimed at identifying the present circulation scheme and the one induced by the inauguration of the station. The second one aimed at quantifying the entity of traffic flows in “critical points”, by counting the vehicles during peak periods. From the surveys, it has come out that in the peak morning period (7am -10am) 6,000 vehicles have been counted towards the rail node Garibaldi. This value decreases of almost 15% in the afternoon peak period (12 am - 3 pm) as well as in the evening one (5pm -8pm). From the simulation of the future scenario, which changes radically the access to Garibaldi Square and to the surrounding areas (where the CBD and the historical centre are placed), a new distribution of the vehicles is registered with a decrease of flows coming from the motorway A3 of more than 20% and of more than 30% approaching Garibaldi square. The implemented model is also able to assess future scenarios promoted by the local administration.

**Keywords**: Metro station; Direct Impacts, Accessibility, Flows distribution.

## 1. Introduction

Impacts of a transportation system project can be defined as the consequences of the project relevant for some of the actors involved. Thus the definition of the relevant impacts is the main indicator of the approach followed and the breadth of the evaluation activity. The spectrum of the effects considered has widened with the passing of time in concert with improvements in models and computing power and with the expansion and classification of the different and often contrasting objectives and goals of actors and decision-makers (Cascetta, 2009). The impacts are considered for all users, both present and project-induced, calculating the variations in generalized costs, perceived and not perceived, for the different transportation modes (Pagliara and Papa, 2011).
In this paper the focus will be on the Garibaldi strategic node in the city of Naples in Italy, following the inauguration of the station Garibaldi of Line 1 of the metro.
Campania region is one of the twenty administrative regions making up Italy. It has about 5.7 million inhabitants within a 13,590 km$^2$ area and is the second largest region in Italy. It is divided into five administrative provinces and is characterised by a twofold distribution of population and activities. There is a very large central metropolitan area centred around Naples with 3.5 million inhabitants. Naples is a Large Urban Zone (LUZ), i.e. an area with a significant share of residents travelling to work within the city, with one of the highest residential densities of the whole of Europe (an average value of 2305 inh/km$^2$ and peaks of 13,323 of inh/km$^2$). During the second half of the last century, very limited investments were made to expand and/or upgrade the existing railway network. Furthermore, these efforts followed an un-coordinated process in which decisions were taken by individual transport companies, thereby limiting an integrated vision of the regional railway system. Moreover, expansion projects were not coordinated with land-use decisions; in fact, the latter were often made independently of, or even in open contrast with, the rail system. The new planning approach, starter in 1996 in Naples and extended in 2001 to the whole region with the Regional Metro System (RMS) project, is based on the idea that only a highly integrated and extended railway system can provide sustainable mobility in an area with densities such as those of the central area of Campania (Cascetta and Pagliara, 2008).
Between 2001 and 2007 several interventions have been carried out on the urban rail network with the opening of new rail lines with the upgrading and opening of many stations both in central and peripheral areas of the city of Naples (see Table 1).

[1] Corresponding author: fpagliar@unina.it

**Table 1.**
*The RMS project infrastructures*

| | *Base scenario (2001)* | *Completed (2011)* | *Under construction* | *Planned* | *Total* | *% Change* |
|---|---|---|---|---|---|---|
| Total extension of the railway network (km) | 1,179 | 54 | 50 | 66 | 13,49 | +14% |
| Number of stations | 340 | 39 | 30 | 33 | 442 | +30% |
| Investments (M€) | - | 3,115 | 2,525 | 3,500 | 9,140 | |

| | *New* | *Revamped* | ***Total new and revamped*** |
|---|---|---|---|
| Number of trains | 94 | 75 | **169** |
| Number of buses | 1,250 | - | **1,250** |
| **Total number of trains and buses** | **1,344** | **75** | **1,419** |

*Source: Cascetta et al. (2014)*

The objective of this contribution is to provide a reorganization of the Garibaldi node, today totally congested and to analyse the effects post inauguration of the Garibaldi metro station.
This paper is organised as follows. In section 2 the description of Line 1 of the metro of Naples is reported. In section 3 the node of Garibaldi is described, while in section 4 conclusions and further perspectives are reported.

## 2. The metro system in Naples

Metro Line 1, in the Municipal Plan of transport, serves the metropolitan area of Naples, also thanks to the many interconnections with the rail system and park & ride nodes placed in the north of the city, allowing to leave the car in favour of public transport.
Line 1 connects the north of the city with the stop of Piscinola/Scampia to the west with Garibaldi station, which in turn connects the national railways with the High Speed Rail system (Cascetta et al., 2011), passing through the hilly area of the borough of Vomero and the old city center. It serves a total of 17 stations over a distance of 18 km.
The stations have a total of 115 escalators, moving walkways, stair lifts and platform lifts and 52 elevators. Almost the entire route is developed in the gallery, except the stretch Colli Aminei/Piscinola, which runs mostly along the viaduct. The stations served are: Garibaldi, Università, Toledo, Dante, Museo, Materdei, Salvator Rosa, Quattro Giornate, Vanvitelli, Medaglie d'Oro, Montedonzelli, Rione Alto, Policlinico, Colli Aminei, Frullone, Chiaiano, Piscinola.
Line 1 operates on weekdays about 242 trips per day from 6.00 a.m. to 11.00 p.m., with a frequency during peak hours of 8 minutes. On average about 135,000 passengers use the service on weekdays and 50,000 on weekends (www.comunedinapoli.it).
The design and construction of Line 1 are assigned to MN Metropolitana di Napoli SpA, company of the Municipality of Naples engaged in building the infrastructures. The completion of the line involves the construction of a rail link that will end in Piscinola passing the CBD and Capodichino airport, greatly extending the current route length (from 18 km to 25). In Figure 1, the metro stations of Line 1 are reported, while in Figure 2 the future scenario is represented when all stations will be inaugurated.

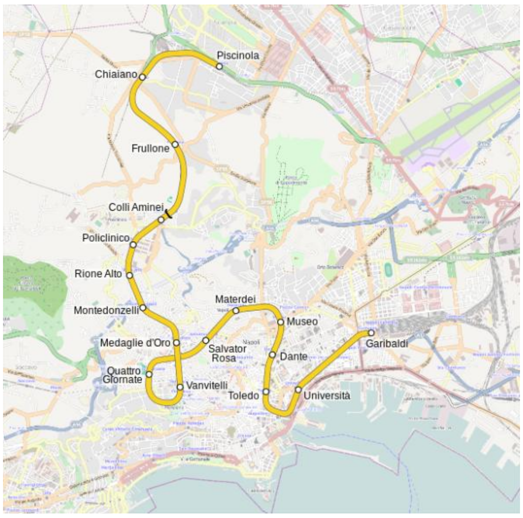



**Fig. 1.**
*Metro Stations of Line 1 in Naples*
*Source:www.metro.na.it*
*Source: Authors' elaboration*

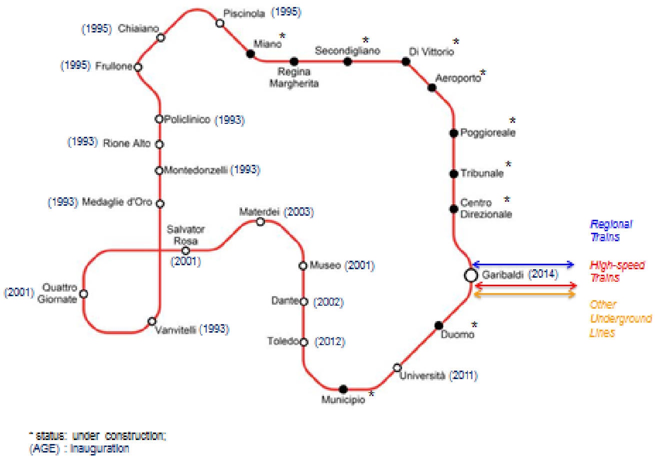



**Fig. 2.**
*The future scenario of Line 1*

Garibaldi square represents a fundamental node in the actual urban transportation system (See Fig. 3). From the public transport perspective, it represents the most important interchange node between rail and road. The reasons are to be found in the number of facilities serving not only the metropolitan area, but also the regional and national demand incoming to the node. In particular, the station Napoli Centrale, Garibaldi Square and Circumvesuviana (from 1980 till 2012, this was the company managing the urban and sub-urban rail systems in the metropolitan area of Naples, since 2012 this service has been managed by Ente Autonomo Volturno (in the common language people still call it "Circumvesuviana"), define an interchange node both internal and external to the transportation system of the city. Internal, since it links the line 2 of Ferrovie dello Stato (The Italian National Railway) with the lines Naples-Sorrento, Naples-Sarno and Naples-Nola-Baiano of the Circumvesuviana.
The presence of a bus terminal both at urban and extra-urban level makes this node more attractive. From the private transport perspective, Garibaldi Square represents the only crossing node for flows coming from the coast and directed to north, but also this node receives the flows from the motorways A1 and A3.

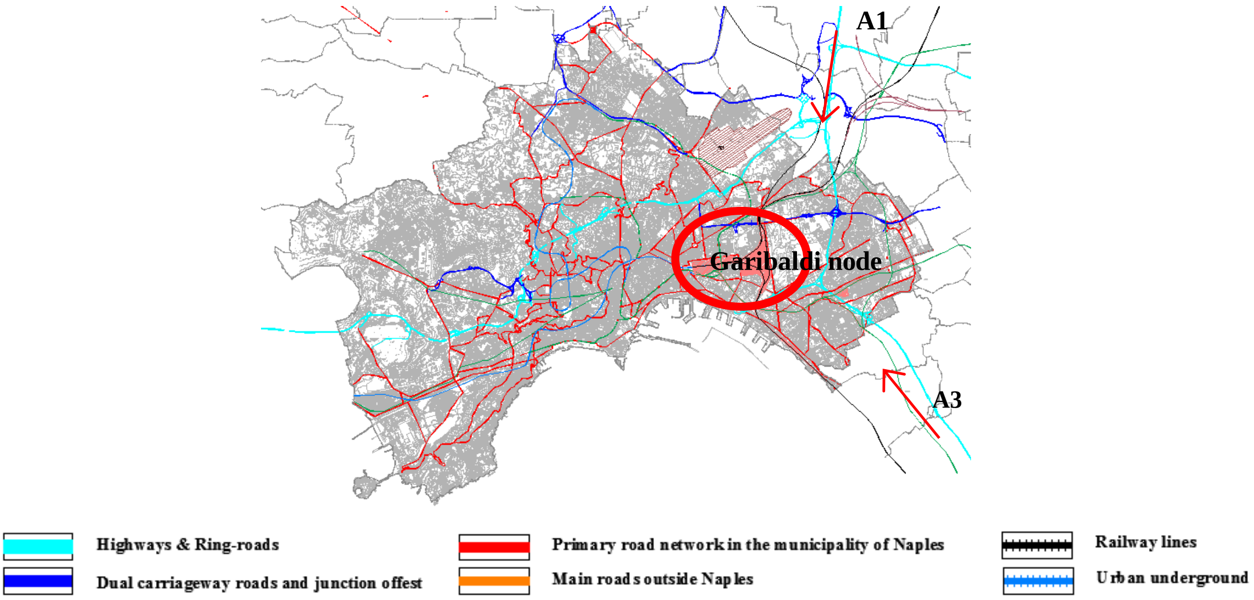



**Fig. 3.**
*Garibaldi node*
*Source: Authors' elaboration*

### 3. The node of Garibaldi within the urban mobility

Private and public transport needed to be re-organized after the inauguration of the new station with the objective of making mobility efficient as a function of the transport demand on the network.
In order to evaluate both transport demand and supply, it has been necessary to undertake two surveys: the first one aimed at identifying the present circulation scheme and the one induced by the inauguration of the station, and the other one aimed at quantifying the entity of traffic flows in "critical points", by counting the vehicles during peak periods.
Subsequently, a GIS software has been used to modify some nodes and links, by intervening on a detailed graph. A macrosimulation software has been considered for modifying nodes and links in order to understand the distribution of flows and the corresponding level of congestion in the area interested by the intervention. Finally, a micro-simulation model has been built. A Deterministic User Equilibrium assignment model has been considered for interpreting the users' behaviour on the road.
A survey has been carried out during the month of February 2014 (on the 5th and the 6th) and during the morning and afternoon peak periods (i.e. 7.00-10.00 a.m.; 12.00-15.00 p.m. and from 17.00 to 20.00 p.m.) with the objective of analyzing the actual scenario of the vehicles crossing Garibaldi Square in order to highlight the critical points w.r.t. the data of 2012. This information has been used to calibrate and therefore simulate traffic flows through the software MT-Model (version: V4.1.010).
The activities which have been carried out are:

− Identification of the actual circulation scheme of both private and public means of transport.
− Flows of cars, buses, motorbikes crossing the square.

The main outcome has been that in the peak morning period (7.00 am -10.00 am) 6,000 vehicles have been counted to the direction of the rail node of Garibaldi. This value decreases of almost 15% in the afternoon peak period (12.00 am - 3-00 pm) as well as in the evening one (5.00 pm -8.00 pm).

In Fig. 4 the distribution of flows by typology is reported, referring to all the periods during which counts have been carried out. From the simulation of the future scenario which changes radically, both from an infrastructural and functional point of view, the access to Garibaldi Square and to the surrounding areas (where the CBD and the historical centre are placed), a new distribution of the vehicles has been registered with a decrease of the flows incoming from the motorway A3 of more than 20% and incoming to Garibaldi square of more than 30%.

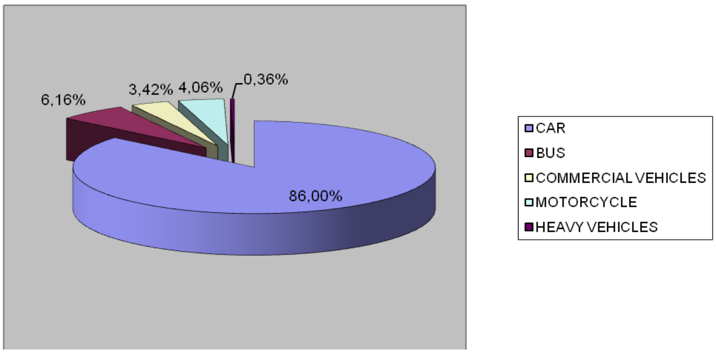



**Fig. 4.**
*Percentage of counted vehicles*
*Source: Authors' elaboration*

Figure 5 refers to September 2012, before the inauguration of Garibaldi station, from there it is clear that it exists a high level of congestion incoming to the square due to the high pedestrian and car flows.

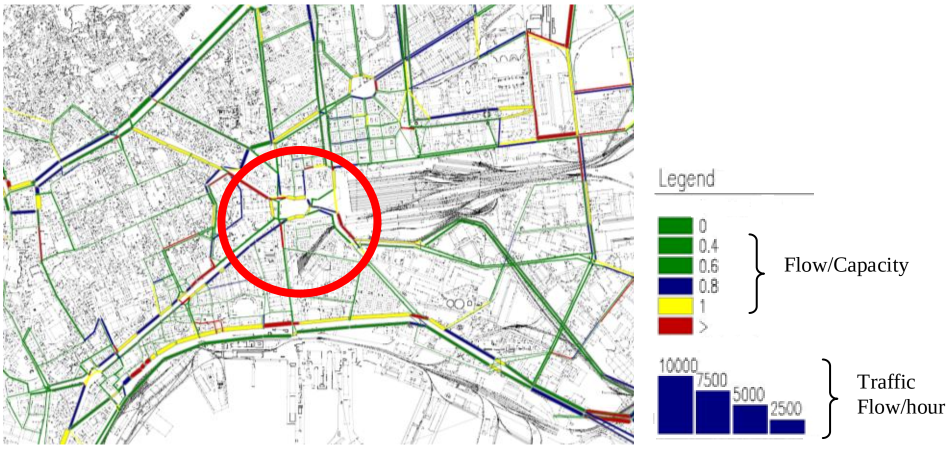



**Fig. 5.**
*Flows in Garibaldi square – September 2012*

In Figure 6 the same situation is reported but referring to December 2013, where some changes have been carried out in order to make the circulation in this node more sustainable. Figure 6 represents the actual scenario w.r.t Figure 5, cars are not allowed to cross the square from north to south with the excpetion of buses, taxis and vehicles for loading and unloading freight. Finally, in Figure 7 the future scenario has been implemented and simulated. The area becomes a traffic limited zone where only public transport is allowed to cross and two underpasses (circled in red) have been introduced as well.

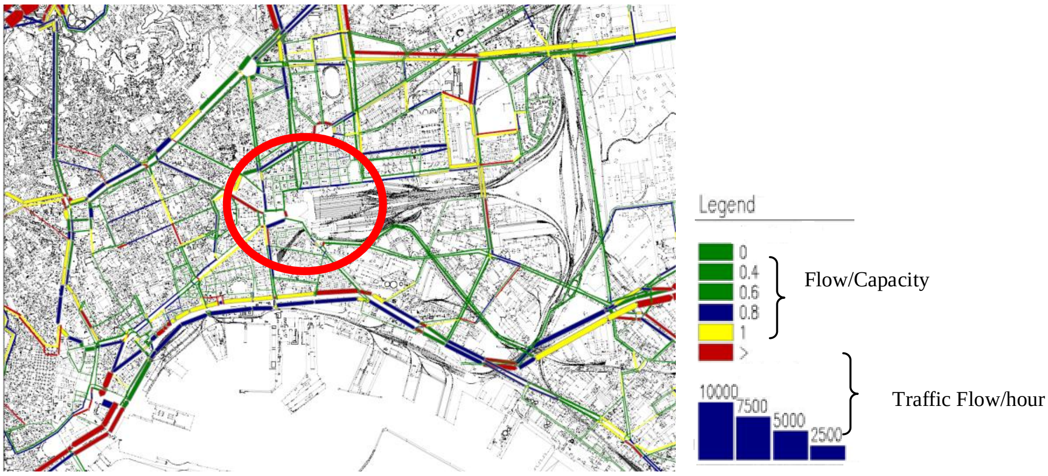



**Fig. 6.**
*Flows in Garibaldi square in December 2013 with traffic's device*
*Source: T-Model results*

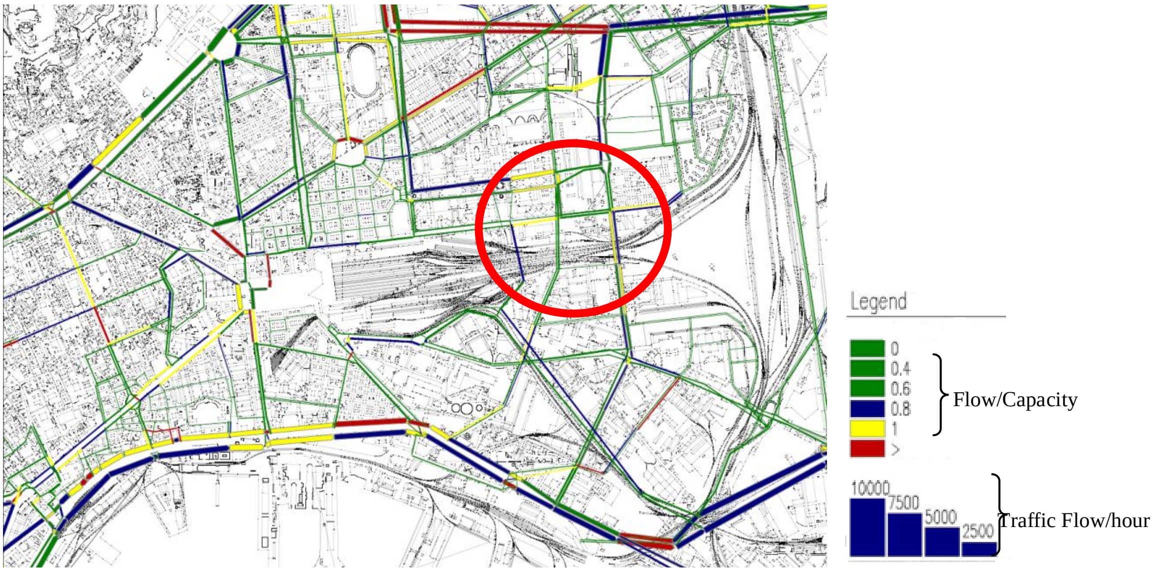



**Fig. 7.**
*Flows in Garibaldi square in the future scenario with a Limited Traffic Zone and two underpasses circled in red*
*Source: T-Model results*

## 4. Conclusions and further perspectives

In order to evaluate the impacts of the different interventions the three different scenarios have been compared, i.e. before inauguration, after the inauguration with infrastructural interventions (reference scenario) and the intervention scenario.

The before inauguration scenario refers to the year 2012; the reference scenario refers to February 2014 after the inauguration of the metro station Garibaldi and the intervention limiting cars crossing the square from the north to the south. The intervention scenario will consider two underpasses.
The indicators computed and referred to the system simulated in the peak morning period are:

1. Total km on the network (flow * link length)
2. Total time spent (flow * link travel time)
3. Average speed (total km/total time spent)

The comparison has been considered for different roads typology.

From the analysis of the results reported in Tables 2 and 3 referring to the absolute value and to the absolute and percentage change respectively, it results, by comparing the indicators of the before inauguration scenario (2012) and those of the reference scenario (2014), an increase of 2% of km travelled on the motorway network and an increase of almost 6% on other roads. An increase of 4% is registered in the total of km travelled. The greatest increase is registered in the km of other types of roads following the interventions of limiting the cars. Indeed users are pushed to choose other roads for reaching the different destinations close the Garibaldi node.
It is also true that the year 2012 was an year of crisis and that the flows on the network have been resulted to be extremely less.
This is confirmed by the increase of 43% of the total time spent on other roads, linked to the high traffic flow on these roads. As for as the time spent on urban roads it is possible to notice a reduction of almost 5%. Finally, by comparing the average speed it is possible to highlight a reduction on the entire network equal to 19%.
In conclusions the analysis of the indicators of the two scenarios confirms the speed decrease, with still “places” of local congestion, which determines an increase of the total time spent on the network. There are weaknesses within the road network and therefore interventions are needed in order to reduce the load and making the general state of the network working better.

By comparing the future scenario with the reference one (2014), it is possible to observe that for the total km on the motorways the difference is less than 1%. An increase of 3% is registered on the total km on the other roads. The percentage decrease is due to the opening of the two underpasses and the limitation of the cars. Also for the total time spent it is possible to notice a reduction equal to 10%. On the other roads the decrease is equal to 13%.
Through the comparison of the average speeds, the change is really insignificant both on motorways and on other roads.
In conclusions the future infrastructural interventions, i.e. the opening of the two underpasses will allow the redistribution of traffic flows making a satisfactory level of functioning of the entire network.

**Table 2.**
*Comparison among indicators (absolute values)*

| | Before Garibaldi station inaguration | After Garibaldi station inaguration | Intervention scenario |
|---|---|---|---|
| **Performance indicators** | **2012** | **2014** | **future** |
| **Tot Km** | **782016** | **812940** | **797420** |
| Motorways | 420015 | 429548 | 427212 |
| Other roads | 362001 | 383392 | 370208 |
| **Total time spent** | **36716** | **47212** | **42218** |
| Motorways | 11242 | 10699 | 10631 |
| Other roads | 25474 | 36513 | 31587 |
| **Average speed (Km/h)** | **21,30** | **17,22** | **18,89** |
| Motorways | 37,36 | 40,15 | 40,19 |
| Other roads | 14,21 | 10,50 | 11,72 |

**Table 3.**
*Comparison among indicators (absolute values and % values)*

| | **Absolute variation** | | **% variation** | |
|---|---|---|---|---|
| **Performance indicators** | **2014-2012** | **Intervention scenario-2014** | **2014-2012** | **Intervention scenario-2014** |
| **Tot Km** | **30924** | **-15520** | **3,95%** | **-1,91%** |
| Urban motorways | 9533 | -2336 | **2,27%** | **-0,54%** |
| Other roads | 21391 | -13184 | **5,91%** | **-3,44%** |
| **Total time spent** | **10496** | **-4994** | **28,59%** | **-10,58%** |
| Urban motorways | -543 | -68 | **-4,83%** | **-0,64%** |
| Other roads | 11039 | -4926 | **43,33%** | **-13,49%** |
| **Average speed (Km/h)** | | | **-19,16%** | **9,69%** |
| Urban motorways | | | **7,46%** | **0,09%** |
| Other roads | | | **-4,81%** | **0,47%** |

This paper is a clear example of a wicked problem. A “wicked problem” is a phrase used in social planning to describe a problem that is difficult or impossible to solve because of incomplete, contradictory, and changing requirements that are often difficult to recognize. Moreover, because of complex interdependencies, the effort to solve one aspect of a wicked problem may reveal or create other problems (Rittel and Weber, 1973). Congestion can be considered a wicked problem according to the six critical characteristics raised by Conklin (2009), i.e.:

1. You don’t understand the problem until you have developed a solution.
2. Wicked problems have no stopping rule.
3. Solutions to wicked problems are not right or wrong.
4. Every wicked problem is essentially unique and novel.
5. Every solution to a wicked problem is a “one shot operation”.
6. Wicked problems have no given alternative solutions.

Further research will reconsider these points and analyse them in depth for the case study analysed here.

**References**

Cascetta, E. 2009. *Transportation Systems Analysis. Models and Applications*. Springer, New York.

Cascetta, E. and Pagliara, F. 2008. The Regional Metro System project of Naples and Campania, *Transport Policy*, Vol. 15, N.2, pp. 81-93.

Cascetta, E.; Papola, A.; Pagliara, F. and Marzano, V. 2011. Analysis of mobility impacts of the high speed Rome–Naples rail link using withinday dynamic mode service choice models, *Journal of Transport Geography*, 19: 635-643.
Conklin, J. (2009) Building Shared Understanding of Wicked Problems, Rotman Magazine, the alumni magazine of Rotman School of Management.

Cascetta, E.; Cartenì A., Pagliara, F. and Montanino, M. 2014. A new look at planning and designing transportation systems. Part II: A decision-making model based on cognitive rationality, stakeholders engagement and quantitative methods, submitted to Transport Policy, forthcoming.

Pagliara, F. and Papa, E. 2011. Urban rail systems investments: an analysis of the impacts on property values and residents'location, *Journal of Transport Geography*, 19: 200-2011.

Rittel, H.; Webber, M. 1973. Dilemmas in a General Theory of Planning, *Policy Sciences*, Vol. 4, pp. 155–169.

# CALIBRATION AND VALIDATION PROCEDURE OF MICROSCOPIC TRAFFIC SIMULATION MODEL: A CASE STUDY

**Daniela Koltovska[1], Kristi Bombol[2], Dragan Ilievski[3]**
[1,2,3]*Sv. Kliment Ohridski University, Bitola, Faculty of Technical Sciences, Bitola, Republic of Macedonia*

**Abstract:** Increased application of microscopic traffic simulation models imposes the need for high accuracy and precision. However, in research literature an exact and precise calibration and validation procedure definition of microscopic traffic simulation models is lacking. Informal practical experiences with barely full description procedure are present. Very often validation has not been carried out at all or if it has, it has been done ad hoc. As input model parameters, default parameters are usually used without being adjusted to local urban network features and its users. In most cases this has been regarded as delimitation of simulation models to be able to reflect the real on site conditions for appropriate decisions to be taken upon. Hence, the adjustment of input parameters according to the local urban network features and its users is becoming a calling issue.
In this paper a system procedure for calibration and validation of parameters in a designed microscopic traffic simulation model for a real small urban network is presented. The calibration and validation is done in VISSIM. The traffic simulation results were compared with detailed recorded field data. The input model parameters calibration and validation proved both accuracy and precision of the designed microscopic traffic simulation model.

**Key words:** microscopic modelling, simulation, calibration, validation, urban network

## 1. Introduction

Microscopic traffic simulation models contain standard values for each of the parameters which cannot always genuinely reflect the real local traffic conditions. Most often, such lack of a genuine picture of the said conditions poses limitations to the simulation models set to precisely “simulate” the field conditions on which corresponding decisions would be grounded. The problems identified impose themselves as an immediate research requesting field.

In order to achieve the reliability of a microscopic model, the most specific and significant phase in its development is the one related to the calibration and validation. The research presented in literature most frequently is concerned with (1) the general needs of the procedure for calibrating simulation models (Hourdakis et al., 2003; Chu et al., 2006; FHWA., 2004; Hellinga., 1998) or (2) to a detailed description of the model calibration process with an accent on wide traffic networks and on traffic congestion conditions (Fellendorf., 2001; Park et al., 2005; Park et al., 2006).

VISSIM software calibration parameters can be divided into several categories, primarily in accordance with the research objectives and the defined effectiveness measures. It has to be pointed out that the selection of data to be calibrated requires that the following points are needed to be taken into consideration:

1. **Difficulties when collecting data.** When collecting data, it is very important to check if data is available to access
2. **Sensitivity of selected parameters of traffic conditions.** Parameters should be “sensitive” to field traffic conditions because calibration becomes difficult if the parameter is not sensitive to traffic conditions
3. **Consistency and reliability of model calibration.** Once the calibration has been completed, it is necessary to approach analytically when calculating the accuracy with which the simulation model reflects the reality (which is done by means of statistical tests)

In the Republic of Macedonia, so far, there has been no procedure developed for calibration and validation of microscopic simulation models, regardless of the size of the various urban networks. Thus, it is in this paper that a procedure for calibration and validation of a microscopic simulation model has been developed for the first time in VISSIM and for a small town. The procedure suggested below was applied in a case study of an urban network of eight intersections in the centre of the small town of Vinica.

The paper consists of the following sections: (1) development of the procedure for calibration and validation, (2) application of the procedure and results regarding the center of the town of Vinica (3) discussion and conclusions.

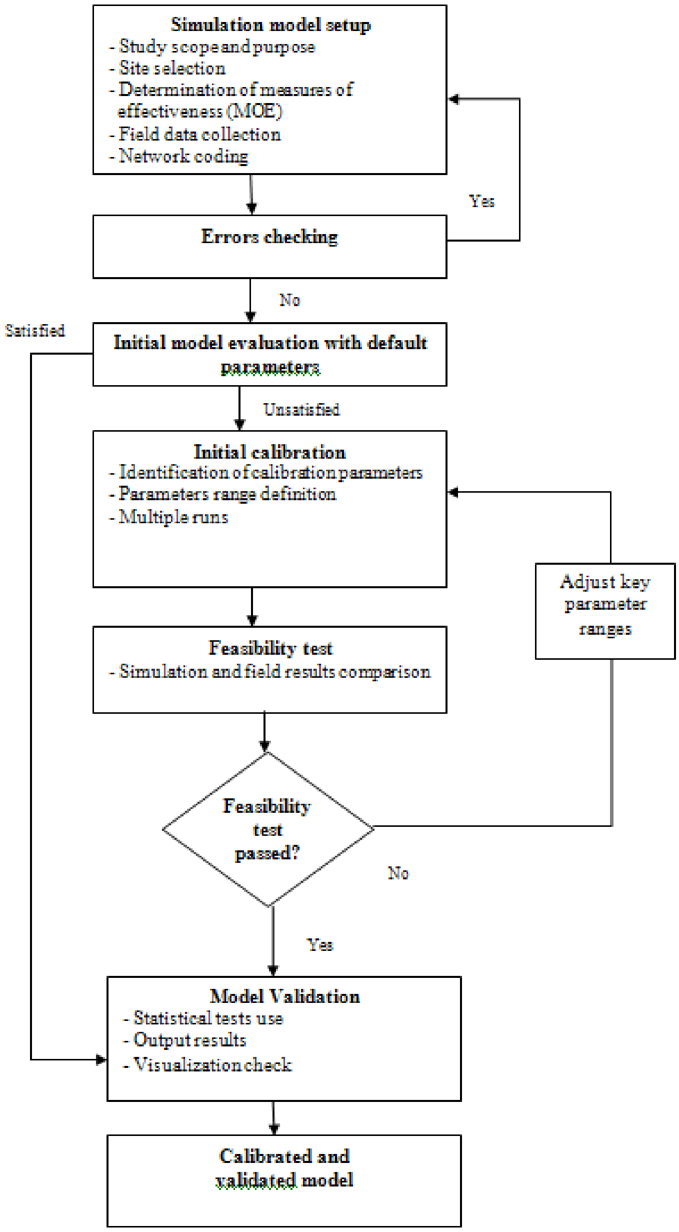
## 2. Calibration and Validation Procedure Steps

Having conducted basic theoretical examination of the model and sub models of VISSIM, the actual parameters for adapting the model were determined in accordance with the characteristics of the urban traffic network of Vinica and its users.

[1] Corresponding author: daniela.koltovska@tfb.uklo.edu.mk

The designed procedure for calibration and validation of microsimulation model for a small town is presented in Fig.1.

**Fig. 1.**
*Calibration and validation procedure*
*Source: Authors*




**Step 1: Design and development of a simulation model.** Comprises defining the study scope and objectives, site selection, determination of measures of effectiveness, field data collection and network coding
**Step 2: Errors checking.** Refers to data input errors, errors of software tools and errors when checking the visual animation
**Step 3: Initial model evaluation with default parameters.** The output results obtained with the model with default parameters are compared to empirical data. If the output results are satisfied the model developed is considered appropriate and validation begins. If they are not satisfied, steps of initial calibration as well as the feasibility test should be performed
**Step 4: Initial calibration**. Identification of calibration parameters is performed, acceptable parameters range is defined, and multiple runs are performed
**Step 5: Feasibility test.** It aims at determining whether empirical data correspond to model output data. If deviations occur, the ranges of parameter values are adjusted, and initial calibration is performed again

**Step 6: Model validation.** Most frequently a variety of statistical tests are applied: *t*-test, hypotheses test, descriptive statistics, $\chi^2$ test, ANOVA, MANOVA, etc. The account of output results is given in tables (with numerical values) and linear graph or histogram (for comparison). After the visualisation testing, the model can be used for further research or for corresponding engineering steps

## 3. Application of the procedure for the centre of the town of Vinica

The procedure developed has been implemented in the centre of the town of Vinica. The text below gives a detailed account of the procedure.

1. *Design and development of a micro-simulation model for the town of Vinica*

Eight intersections in the narrow centre of the town were analyzed with the research. Travel time and travel speed on the road section were determined as effectiveness measures. Traffic counts were realized by means of camcorders.

Field measurements of travel time and travel speed on the referential section of the "Marshal Tito" Blvd between the intersection of the boulevard with "Partizanska" st. and the one with "Braka Miladinovci" st. (Fig. 1), according to the type of vehicles are:

- Passenger cars (PC): 28.8 seconds, and 41 km/h
- Heavy good vehicles (HGV): 33.6 seconds, and 35 km/h
- Buses: 39.5 seconds, and 31 km/h

VISSIM network coding consisted of road network design, data on traffic demand, and data on traffic management.

In Fig. 2. the referential road section and the measurement spots for traffic counts are presented.

**Fig. 2**. *Referential road section*
*Source: Authors*


2. *Checking for errors*

Two types of checking for the VISSIM model are conducted. The first one runs automatically along with the initialization of the simulation. The second one is visualization checking of the simulation. In this research case noncompliance of priority was spotted among pedestrians and motor vehicles. These errors were corrected.

3. *Initial evaluation of the model with default parameters*

Output results of the initial evaluation are presented in Table 1.

**Table 1:**
*Travel time on the road section*

| Type of vehicles | Empirical measurements (s) | Model values (s) | Difference (%) |
|---|---|---|---|
| **PC** | 28.8 | 29.8 | 3.5% |
| **HGV** | 33.6 | 30.4 | 9.5% |
| **BUS** | 39.5 | 30.5 | 22.8% |
| **Average travel time (s)** | 29.3 | 29.8 | 1.7% |

*Source: Authors*

Total average travel time measured on the selected section is 29.3 seconds. Total average travel time of the model is 29.8 seconds. The difference of 0.5 seconds only (1.7%) is insignificant. However, according to vehicle type, travel time is as follows:

1. Passenger cars (PC): 28.8 seconds on the selected section, and 29.8 seconds of the model. The difference in percentages amounts 3.5%
2. Heavy good vehicles (HGV): 33.6 seconds on the selected section, and 30.4 seconds of the model. The difference in percentages amounts 9.5%
3. Buses: 39.5 seconds on the selected section, and 30.5 seconds of the model. The difference in percentages amounts 22.8%.

In accordance with the defined value of maximum difference of 10% with the output model parameters and the empirical measurements [8], it was necessary to perform travel speed calibration of HGV and buses.

4. *Initial calibration*

*a) Identification of calibration parameters*

On the grounds of the insight into the experiences (Kim., 2006; Otković et al., 2013), the possible list for selection of input parameters in VISSIM and their limitations in the calibration process is presented in Table 2. The final selection of input parameters to be entered for the calibration process is grounded on the availability of real data and the impact of the selected parameter on the simulation outcome. In this research, the output results for travel time were influenced by the desired travel speed. Empirical measurements of the other parameters given in Table 2 were not performed.

**Table 2**
*Selection of input parameters for model calibration in VISSIM*

| P | Input parameters | Range | Step | Default |
|---|---|---|---|---|
| P1 | Simulation resolution | 1÷10 | 1 | 5 |
| P2 | Number of observed proceedings vehicles | 1÷4 | 1 | 2 |
| P3 | Max look ahead distance (m) | 100÷300 | 1 | 250 |
| P4 | Min look ahead distance (m) | 0÷20 | 1 | 0 |
| P5 | Average standstill distance (m) | 1÷3 | 0.1 | 2 |
| P6 | Additive part of desired safety distance (m) | 1÷5 | 0.1 | 2 |
| P7 | Multiplicative part of desired safety distance (m) | 1÷6 | 0.1 | 3 |
| P8 | Desired speed (km/h) | 25÷50 | 10 | 40 |

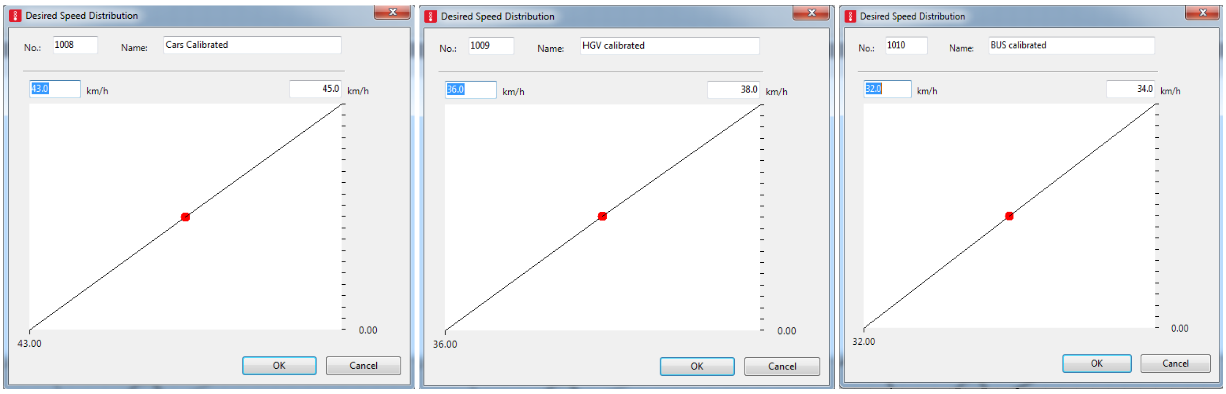
*Source: Otković I.I. et al., Analysis of the influence of car-following input parameters on the modelled travelling time. Tehnički vjesnik 20, 5(2013), 919-925*

*b) Determination of acceptable ranges for parameter values*

At this stage, due to the prevalence of passenger cars (80%) testing the differences in means of travel time on the selected section of the model and of the empirical measurements of all vehicle types was performed with *t*-test. The test results showed that passenger cars are also to be included in the calibration process.

The diagrams of desired travel speed of PC, HGV, and buses are presented in Fig. 3.

**Fig. 3.** *Desired speed of vehicles for calibrated model*
*Source: Authors*

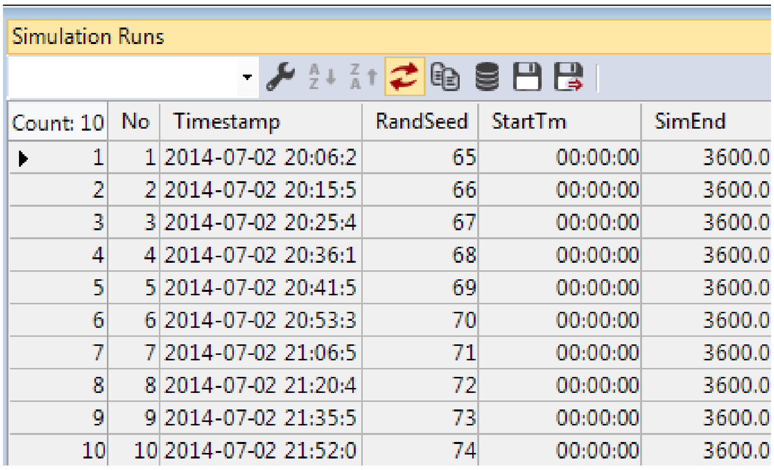


A slight range shift was made with PC by increasing the minimal value from 43 km/h to 45 km/h. HGV travel speed was decreased, and decided upon to be within the speed range between 36 km/h and 38 km/h. The buses travel speed was decided upon to be within the speed range between 32 km/h to 34 km/h in relation to the initial range between 40 km/h and 45 km/h.

*c) Multiple simulation*

Details of the ten simulations initialized with VISSIM are presented in Fig. 4.

**Fig. 4.** *Multiple runs in VISSIM*
*Source: Authors*



Simulation Runs

| Count: 10 | No | Timestamp | RandSeed | StartTm | SimEnd |
|---|---|---|---|---|---|
| 1 | 1 | 2014-07-02 20:06:2 | 65 | 00:00:00 | 3600.0 |
| 2 | 2 | 2014-07-02 20:15:5 | 66 | 00:00:00 | 3600.0 |
| 3 | 3 | 2014-07-02 20:25:4 | 67 | 00:00:00 | 3600.0 |
| 4 | 4 | 2014-07-02 20:36:1 | 68 | 00:00:00 | 3600.0 |
| 5 | 5 | 2014-07-02 20:41:5 | 69 | 00:00:00 | 3600.0 |
| 6 | 6 | 2014-07-02 20:53:3 | 70 | 00:00:00 | 3600.0 |
| 7 | 7 | 2014-07-02 21:06:5 | 71 | 00:00:00 | 3600.0 |
| 8 | 8 | 2014-07-02 21:20:4 | 72 | 00:00:00 | 3600.0 |
| 9 | 9 | 2014-07-02 21:35:5 | 73 | 00:00:00 | 3600.0 |
| 10 | 10 | 2014-07-02 21:52:0 | 74 | 00:00:00 | 3600.0 |

*5. Feasibility test*

*a) Comparison of simulation outcome with empirical measurements*

In Table 3 measurements of travel time on the selected road section and on the calibrated model are presented.

**Table 3**
*Travel time from empirical (field) measurements and calibrated model*

| Type of vehicles | Empirical measurements (s) | Model values (s) | Difference (%) |
|---|---|---|---|
| **PC** | 28.8 | 29.1 | 1.0 % |
| **HGV** | 33.6 | 34.2 | 1.8 % |
| **BUS** | 39.5 | 37.9 | 4.0 % |
| **Average travel time (s)** | 29.3 | 29.4 | 0.1 % |

*Source: Authors*

The average travel time for all vehicle types with VISSIM model and travel time on the field differs for 0.1 second. The values of passenger cars travel time differ for 0.3 second (1%), of HGV for 0.6 seconds (1.8%), and of buses for 1.6 seconds (4.0%). Thus, it indicates that high level of simulation model confidence has been achieved.

## *6. Model validation*

### *a) t-test application*

*t*-test was applied for testing the significance of difference between the means of travel time obtained by the model and of measured on the field for all the vehicle types. Travel time was tested with significance level of 95%. *t*-test results are presented in Table 4.

**Table 4**
*t-test results*

| | **Empirical Data** | | | **Calibrated Model** | | | ***t* - test** | | |
|---|---|---|---|---|---|---|---|---|---|
| | **No of Vehicles** | **Travel time (s)** | **SD** | **No of vehicles** | **Travel time (s)** | **SD** | **SD** | **Determined *t*-value** | ***t*-value from the table of values from Student's t-distribution** |
| **PC** | 200 | 28.8 | 4.3 | 223 | 29.1 | 0.1 | 0.30 | -0.99 | 1.96 |
| **HGV** | 11 | 33.6 | 6.7 | 7 | 34.2 | 1.2 | 2.07 | -0.29 | 2.12 |
| **Bus** | 4 | 39.5 | 2.1 | 4 | 37.9 | 0.9 | 1.14 | 1.40 | 2.45 |
| **Total** | 215 | 29.3 | 4.7 | 234 | 29.4 | 0.2 | 0.32 | -0.31 | 1.96 |

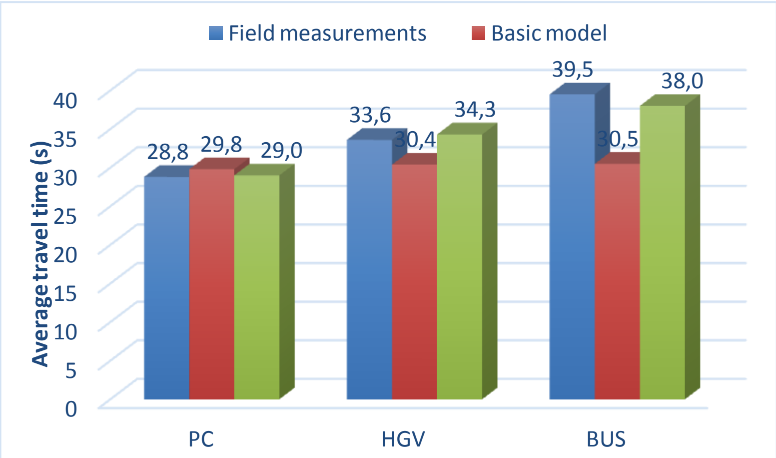
*Source: Authors*

As absolute values for *t*-variables for all vehicles categories are lower than the limit values from the table of values from Student's *t*-distribution with 95% significance level, we conclude that differences between the model values and of those from the field are insignificant. It means that the model realistically reflects the conditions on the field with risk of 5%.

### *b) An account of the output results*

Comparison of travel time from field measurements, from the basic model and from the calibrated model of different vehicles types is presented in Fig. 5.

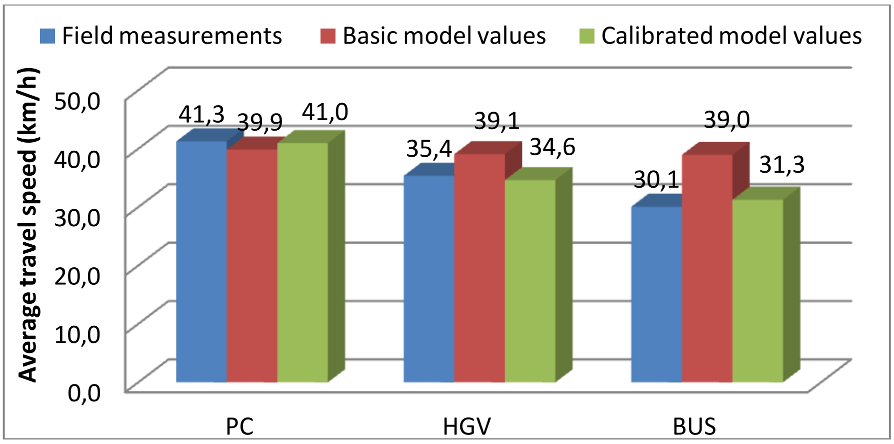
**Fig. 5.**
*Comparison of travel time*
Source: *Authors*




Travel time values from field measurements and from the calibrated model are nearly identical. Thus, the calibration model validity is confirmed with regards to travel time.

Fig.6 presents the travel speed of vehicles types measured on the field, on the basic model, and on the calibrated model.

**Fig. 6.** *Comparison of travel speed*
*Source: Authors*




Travel speed values from field measurements and from the calibrated model are nearly identical. Thus, the calibration model validity is confirmed with regards to travel speed.

*c) Visualization check of the model*

Visualization check showed the priorities regular set, regular pedestrian flow as well as compliance to traffic signals.

The advantages of the performed model calibration and validation are:

- Obtaining/Developing a model which would give an authentic picture of the on-field conditions
- Lessening the possibility of obtaining unreal output results
- Quality and significance of the model
- Possibility to use this model to develop and analyze draft traffic solutions, free of the concern that their implementation on the field may produce unsatisfactory results

However, the calibration procedure has its disadvantages:

- Additional time and funds for developing the model
- When geometric features and/or traffic control devices are changed, the previously calibrated model will have to be calibrated again

## 4. Conclusion

This research has established the foundations of the procedure for calibration and validation of traffic micro-simulation models designed for small towns. In the Republic of Macedonia this is the first time that a general procedure for calibration and validation of traffic micro-simulation models for small towns has been proposed and evaluated.

The model calibrated and validated as above can be used for draft traffic solutions, new traffic control measures and other engineering measures for the purpose of improvement of traffic system in small towns.

Research should proceed in this course for the purpose of expanding the parameters to be measured on the field (e.g.: delays, queue length, saturation flow). It would mean enlarging the number of input parameters of simulation models relevant for analyses.

**References**

Chu L.; Liu H.; Oh J.; Recker W. 2006. A Calibration Procedure for Microscopic Traffic Simulation, *83rd Annual Meeting CD-ROM, Transportation Research Board*.

Federal Highway Administration. 2004. *Analysis Toolbox Volume III: Guidelines for Microsimulation Modeling*. FHWA Publication Number FHWA-HRT-04-040. 450 p.

Fellendorf, M.; Vortisch, P. 2001. Validation of the Microscopic Traffic Flow Model Vissim in Different Real World Situations. *In 80th Annual Meeting Preprint CD-ROM, Transportation Research Board*.

Hellinga, B. R. 1998. Requirement for the Calibration of Traffic Simulation Model. Available from Internet: <http://www.civil.uwaterloo.ca/bhellinga/publications/publications/csce-1998-calibration.pdf>.

Hourdakis J.; Michalopoulos P.; Kottommannil J. 2003. A Practical Procedure for Calibration Microscopic Traffic Simulation Models, *Journal of the Transportation Research Board*, 1852: 130 -139.

Kim, S. J. 2006. Simultaneous calibration of a microscopic traffic simulation model and OD matrix, Ph dissertation. Available from Internet: <https://repository.tamu.edu/handle/1969.1/4409>.

Otković Ištoka I. et al., 2013. Analysis of the influence of car-following input parameters on the modelled travelling time, *Tehnički vjesnik* 20 (5): 919-925.

Park, B.; Jongsun, W.; Ilsoo, Y. 2006. Application of Microscopic Simulation Model Calibration and Validation Procedure: A Case Study of Coordinated Actuated Signal System. *Journal of the Transportation Research Board*, 1978: 113–122.

Park, B.; Qi, H. 2005. Development and Evaluation of Simulation Model Calibration Procedure. *84th Annual Meeting.*

Park, B.; Won, J. 2006. *Microscopic Simulation Model Calibration and Validation Handbook.* Richmond, Virginia, 32 p.

Park, B; Qi, H. 2005. Development and Evaluation of a Procedure for the Calibration of Simulation Models. *Journal of the Transportation Research Board*, 1934: 208–217.

# THE CALIBRATION OF MACROSCOPIC TRANSPORT SUB-MODELS THROUGH TRAFFIC COUNTING

**Gabriela MITRAN[1], Sorin ILIE[2], Viorel NICOLAE[3]**
*[1,2,3]University of Pitesti, Faculty of Mechanics and Technology, Automotive and Transport Department*

**Abstract:** The estimation of transport demand at a future time moment represents a complex process in which are integrated data and information from multiple and diverse sources, as demography, economics, industry, land use, taking into consideration their implications on transport activity. One of the key elements that contribute to obtain a high degree of prediction confidence, besides the quality of data previously mentioned and the skill of expert that harmonizes these, is the knowledge of current situation of transport demand and behavior of transport systems users. Knowing the current situation of transport demand can be achieved by traffic surveys, but when we are dealing with a transportation network, metering traffic flows on each sector requires significant resources, both for collecting and for processing and interpretation of results. Thus, the best method for highlighting the transport demand on each network element is transport modeling. The classic model used in this respect is called "the four-step model" and contains inter-related mathematical models concerning trip generation, their distribution on destinations, mode choice and trip assignment. Representation with a fidelity as high as possible of the modeled reality requires calibration and validation of each sub-model of the transport model. In this paper is highlighted the way in which the traffic data, collected with the help of a pneumatic tubes system, are used to calibrate the inter-related mathematical models. The case study is applied for the transport model of Pitesti City, Romania.

**Keywords**: traffic survey, pneumatic tubes, transport model calibration, transport sub-models.

## 1. Introduction

Knowing the transport demand and traffic flows for the base year are essential elements which significantly influences the operation of transport systems, both at the time of analysis, and in future time moments. The results of traffic forecast process are inputs for transport planning and related sectors (air quality management, land planning). The values of predicted traffic flows are directly dependent on a range of socio-economic, demographic and land use factors forecasted for the study area, having as calculation basis the values of traffic flows specific for base year *(Yao and Sun, 2013; Lippi et al., 2010)*. The current situation of the transport demand can be assessed through traffic surveys *(Potocnik and Govekar, 2011)*, but when we are dealing with a transportation network, metering of traffic flows on each sector requires significant resources, both for the collection and for the processing and interpretation of results. Thus, the best method for highlighting the transport demand on each network element is transport modeling. The classical model used in this respect is called "four-step model" and contains inter-related mathematical models regarding *trip generation, their distribution on destinations*, *modal choice* and *trip assignment on itineraries (Mitsakis et al., 2014; Ortuzar and Willumsen, 2011; Henser and Button, 2007)*. The representation with as higher fidelity of the modeled reality requires the calibration and validation of each sub-model within the transport model *(Federal Highway Administration, 2010).* In this paperwork is highlighted the way in which the traffic data, collected using a system with pneumatic tubes, are used to calibrate the inter-related mathematical models. The case study is applied for the transport model of the city of Pitesti, Romania.

## 2. Methodology

As a basis in various applications in transport domain and in the fields interacting with it, the result of transport models shows a major importance, therefore it requires a very high accuracy. In this respect, in the frame of modeling process, should not miss *calibration* and *validation* components. The implications of these phases can be observed in Figure 1.

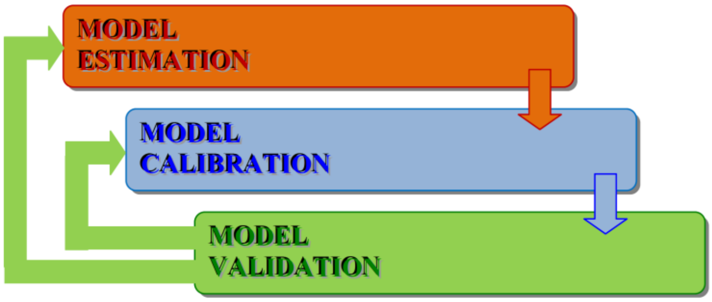



**Fig. 1.**
*The elaboration process of transport model*
*Source: adaptation from: Federal Highway Administration, 2010*

[1] Corresponding author: sorin.ilie@upit.ro; sorin_ilie_dis@yahoo.com

Mathematical models applied for estimating the transport demand and traffic flows are working interdependently. Thus, in order to reduce the phenomenon of errors propagation from one stage to another, is necessary the calibration at the level of each modeling step. This leads to a calibrated final model, but requires additional resources to carry out surveys specific for the calibration of each individual stage *(Azad and Boushehri, 2014).* Figure 2 presents the scheme of transport model in which is proposed the calibration of each constituent sub-model.

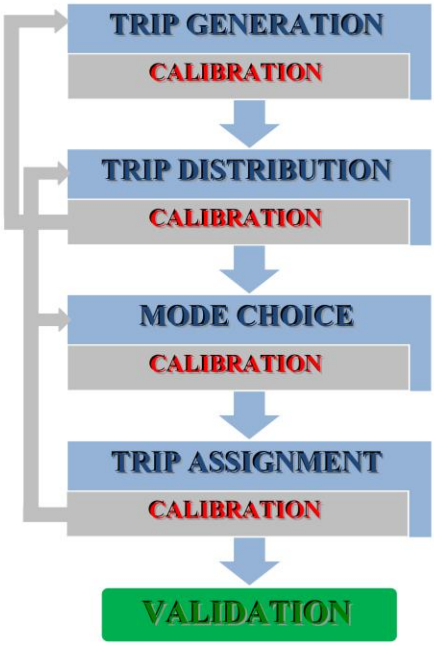



**Fig. 2.**
*The scheme of the macroscopic transport model*

Creating a transport model based on the scheme in Figure 2 leads to increased accuracy of the model, but involves costs associated to activities of collecting and processing data needed for the calibration of inter-related models. The categories of surveys through which are collected the data used to calibrate every sub-model are specified in Table 1.

**Table 1**
*Survey types applied for transport model calibration*

| Macroscopic sub-model | Survey type |
|---|---|
| Trip generation | Household survey |
| Trip distribution | Origin - Destination survey |
| Mode choice | Stated preference travel surveys / Household survey / On-board transit survey |
| Trip assignment | Traffic count |

The surveys through which are collected specific data for calibrating each sub-model are different and expensive (especially Household and Origin-Destination surveys). Therefore, in developing a transport model, a balance must be ensured between the accuracy of the model, established according to the project objectives, and the resources allocated for collecting and processing the data needed for calibration and validation. In this respect, in this paper is proposed the calibration of each sub-model integrated in frame of the transport model, based on data collected within traffic surveys.

### 2.1. Traffic counting methods

There is a wide range of methods for collecting traffic data in order to estimate the *ex-post* transport demand. Depending on the observer's placement related to road surface, these can be divided into two main categories:

- *intrusive methods* – involve the placement of sensing device in contact with road surface;
- *non-intrusive methods* – entail the use of observation techniques from distance.

The most commonly applied intrusive methods consist in using the following means *(Rodrigue et al., 2009)*:

- *inductive loop:* a wire incorporated into road surface as a rectangle and which creates a magnetic field through which relates information with a counting device located outside the roadway. The device shows a low viability, because it may be damaged by heavy vehicles and is predisposed to installation errors;

- *piezo-electric transducer:* a device located in a slot formed in the roadbed of the lane to be counted. This electronic counter is used for measuring the mass and speed of the vehicles on the monitored lane. The mounting operations can affect the integrity of the embankment and reduce the pavement lifetime;
- *pneumatic tubes:* a set of tubes made of rubber, which is placed perpendicular on the road axis, and uses the pressure variations to record the crossing of each axle, through a counting device located on the side of the road;
- *bending plate:* a weight pad attached to a steel plate which is incorporated into the road surface in order measure the mass and the speed of each axle of the vehicle. The use of this device is costly and requires interventions at the level of wearing layer of the road.

Among non-intrusive methods and means, the most used are:

- *manual counting:* traditional method that involves the placement of human observers in certain recording points to count the number of vehicles transiting in front of the observer. In classical manner the observers use registration forms in which they note the number and type of vehicles. There are also electronic devices that are operated by observers by pressing a button corresponding to category of which it belongs the vehicle that passed through the front of registration point. Through this method there can be achieved a detailed traffic monitoring by type of vehicles and travel directions. As downside, manual counting generates traffic safety issues;
- *video recording:* video cameras are used to record vehicles by category and their instantaneous speed. With the help of various software systems are analyzed the video files. Adverse weather conditions may affect the accuracy of counting;
- *Doppler/Radar microwave sensing:* a device that counts the vehicles and records their travelling speed. Except for Radar equipment, they have difficulty in detecting vehicles closely spaced and cannot detect stationary vehicles. The counting devices with microwaves are not affected by meteorological conditions;
- *passive magnetic sensing:* a magnetic sensor that counts the vehicles, recording the speed and type of thereof. In the operation of these counting devices occur difficulties in classifying the vehicles circulating at small distances one from each other;
- *passive and active infrared sensing:* a sensor that detects the presence, speed and type of vehicle by measuring the infrared energy radiating in the detection zone. The device is mounted above the flow of vehicles, on top of a pole or on a bridge. This method shows a reduction in performance during unfavorable weather conditions and a limited coverage of the road lane;
- *ultrasonic and passive acoustic sensing:* devices that use sound waves or energy for detecting vehicles. The ones based on ultrasounds are located above the road infrastructure to record the presence of the vehicle. Their operation can be affected by temperature and turbulences. The acoustic devices are located along the road infrastructure and can detect vehicles by categories.

## 3. Case study

### 3.1. Study area

The study area is represented by the administrative territory of the city of Pitesti, Romania. Pitesti Municipality has a population of 164664 inhabitants and an area of 40.7 km$^2$, being the administrative center of Arges County. The transport network of the city has a longitudinal shape, along the Arges River (Figure 3).

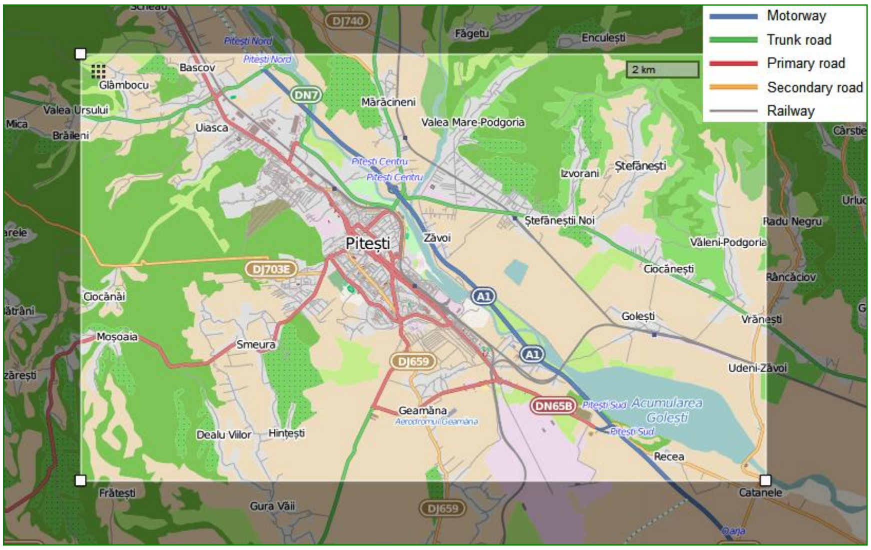



**Fig. 3.**
*The map of the city of Pitesti*
*Source: http://www.openstreetmap.org*

During authors' studies regarding the planning of transport network, was developed a transport model for the mentioned study area for the base year 2012. The influence territory was divided into 92 traffic zones, to which were added another nine, representing external areas with which take place traffic exchange through national and county roads that are in the extension of the street network. The transport network was formalized through a graph with arcs and nodes *(Elefteriadou, 2014)* (229 nodes and 594 arcs).

### 3.2. Transport model calibration

The calibration of transport model was performed at all stages of modeling. The transport network was calibrated in terms of length and average speed on road sections. In framework of model, the transport demand was calibrated at each of the four steps. The option was for calibration based on traffic data recorded at the level of street network. This was possible through the procedures available within *VISUM* software (updating demand matrix with *TFlowFuzzy* procedure, projecting path volumes, calibrating a matrix), with the help of it was implemented the transport model. In total were used the data from 20 survey posts (Figure 4).

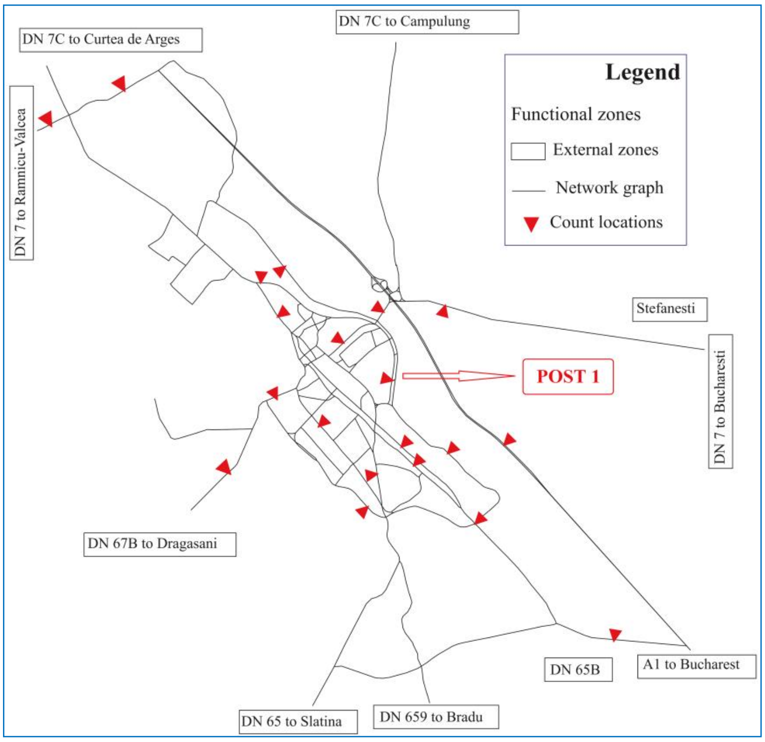



**Fig. 4.**
*Traffic counting points*

#### 3.2.1. Traffic count system

The method for automatic collection of traffic data which has been applied in this case study is the intrusive one which consists in counting and classifying the vehicles with the help of two devices with pneumatic tubes – *MetroCount 5600 Vehicle Classifier System* (Figure 5). One such device has two main components *(MetroCount, 2008; MetroCount, 2007)*:

- *hardware component* – an assembly consisting of a central unit and two axial sensors (pneumatic tubes) installed parallel to a predetermined distance (1 meter) across the roadway; these emit a signal (impulse given by increasing the pressure in tubes) when a vehicle cross over that tube. Signals are sent to the central unit (which can retrieve them with a frequency of less than 1 $ms^{-1}$), where are recorded, processed and saved. By using this assembly, it is possible to get traffic information, such as vehicle speed and vehicle type (by deducting the number of axles, the distance between them and measuring the time elapsed between the events of crossing the axles over tubes).
- *software component* – installed on a portable computer, working together with hardware component, allows the setting of system parameters, retrieving basic information from central unit, their processing and delivering them in the form of reports, such as:
  - the number of vehicles in each category within the range of specified time;
  - the effective speeds of each of them;
  - grouping counted vehicles in classes depending on speed;
  - tracking interval of vehicles;
  - real-time monitoring of recorded vehicles flows;
  - the direction of travel.

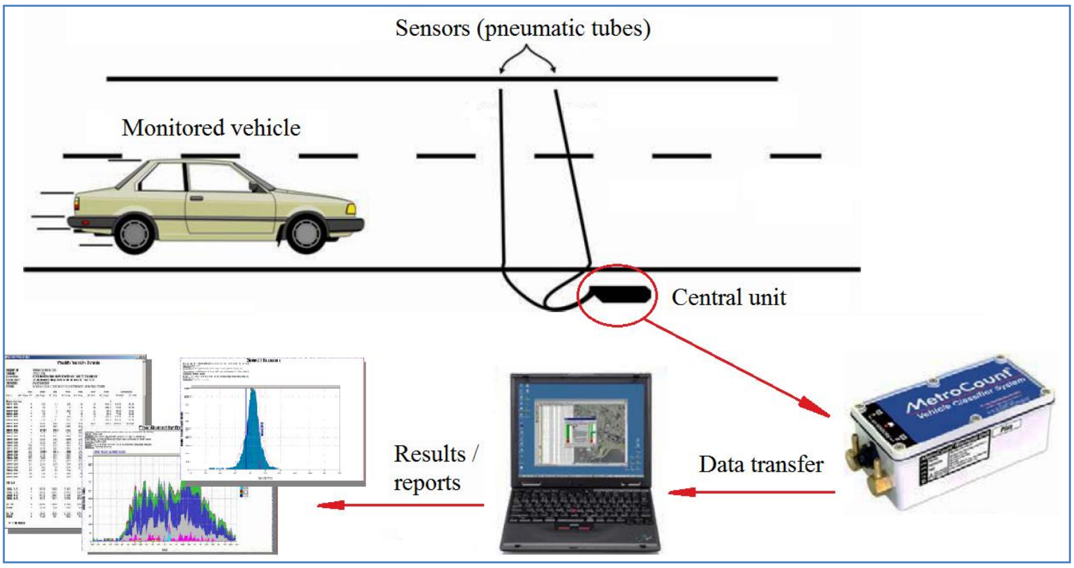



**Fig. 5.**
*Automated counting and classifying of vehicles with MetroCount 5600 Vehicle Classifier System*

The use of this automated system for traffic data collection has the advantage of continuously monitoring traffic flows, can be revealed the travel behavior in terms of time moment at which the movement is performed, the category of vehicle used, the average moving speed, the routes chosen by users to perform travels. These aspects are easily obtained, being output elements of the data processing software.
Further are presented data collected in the survey post number 1 in the way that has been processed with this tool.

### 3.2.2. Obtained results

To exemplify, there have been considered the most important reports generated from data recorded in Post 1 on the direction of traffic North - South. The software of the system allows displaying the results in several ways: as tables or graphs, in raw or processed form, etc. at different time intervals and on different vehicle considered classes. Also, hereinafter is briefly described the way these data are used to calibrate the transport sub-models.

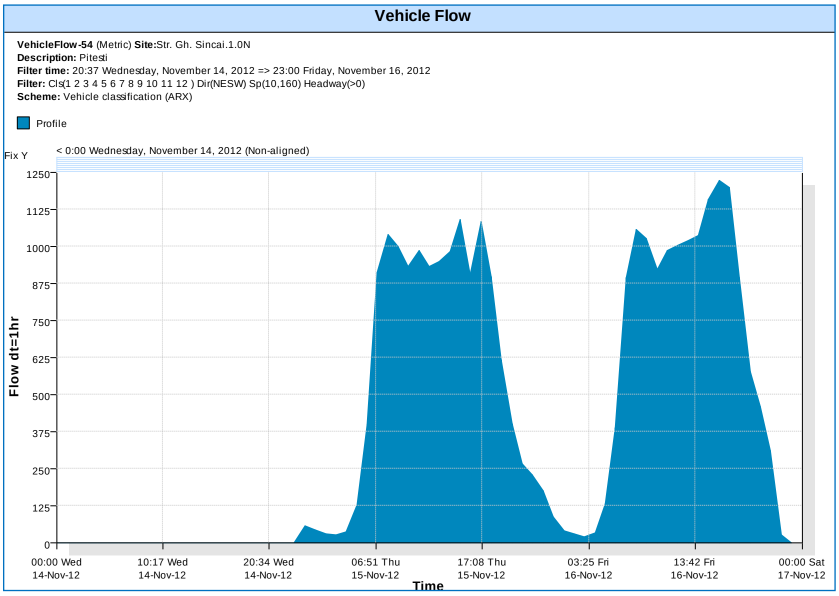



**Fig. 6.**
*Vehicle flow chart*

In Figure 6 is presented *the vehicle flow chart*, showing the total vehicles volume (of all 12 classes specific to ARX classification scheme *(MetroCount, 2008)*) in a time-based graph. The vehicle flow report is a measure of vehicles per time period. With an integration time of one hour, each point on the graph represents total vehicles per hour. The hourly distribution of total traffic volume emphasizes the travel behavior of users in the study area, based on it being possible to identify the peak traffic periods. Knowing the total number of vehicles transiting the key points of transport network is very important for calibration and validation of overall traffic model.

Moreover, the system software allows to generate *the flow stacked by class* report (Figure 7), which is similar to *flow report*, but with either vehicle class bin differentiation. Total flow per integration period is represented by a stack of bars for each included class bin. Fewer classes bins may be required to give sufficient detail. Aggregating a class scheme (usually light, medium and heavy vehicles) often gives extremely useful results, which can be used for calibrating the sub-model for trip assignment; these correct O-D matrices, calibrating the trip distribution sub-model and consequently the trip generation one.

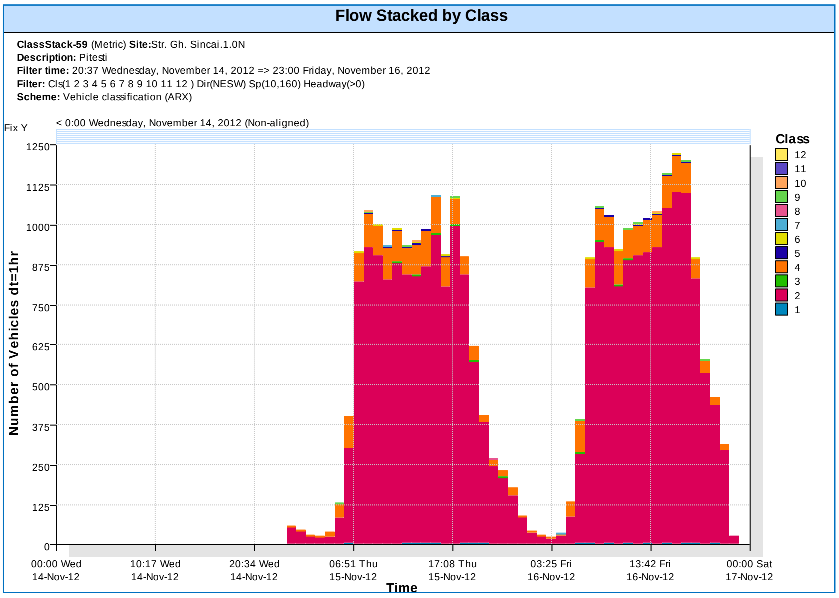



**Fig. 7.**
*Flow stacked by class*

In calibration and validation of transport models, besides traffic flows, a very important role has the knowing the moving speed of vehicles. The hourly representation of average travel speed highlights the intervals with difficult circulation. For each survey point is possible to represent a *speed histogram*. Built for data collected in the point taken for example, it has the allure presented in Figure 8. The histogram of speed shows the speed profile at a counting site. A normal curve, with the same mean and standard deviation, is plotted to help gauge the skew of the speed distribution. The vertical markers show the speed percentile, speed pace and posted speed limit.

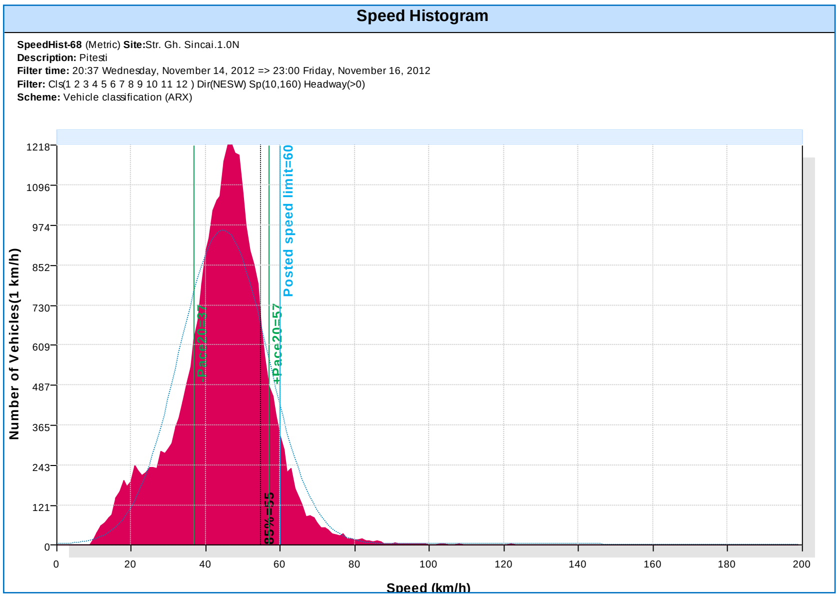



**Fig. 8.**
*Speed histogram*

In Figure 9 is presented *the speed report*, a time-based plot of average vehicle speed per selected integration period. The report also shows a plot of maximum vehicle speed, and a horizontal line for the posted speed limit (PSL = 60 km/h).

The green line shows the average speed of 85% of vehicles which crossed the monitored road section in the considered period. These data are fundamental for calibration of modeled transport network and of sub-model for transport demand assignment. The data regarding speed variation, corroborated with *velocity dispersion* and *flow stacked by speed bins*, are the base of calibration of volume-delay functions, specific to each time interval and each link type.

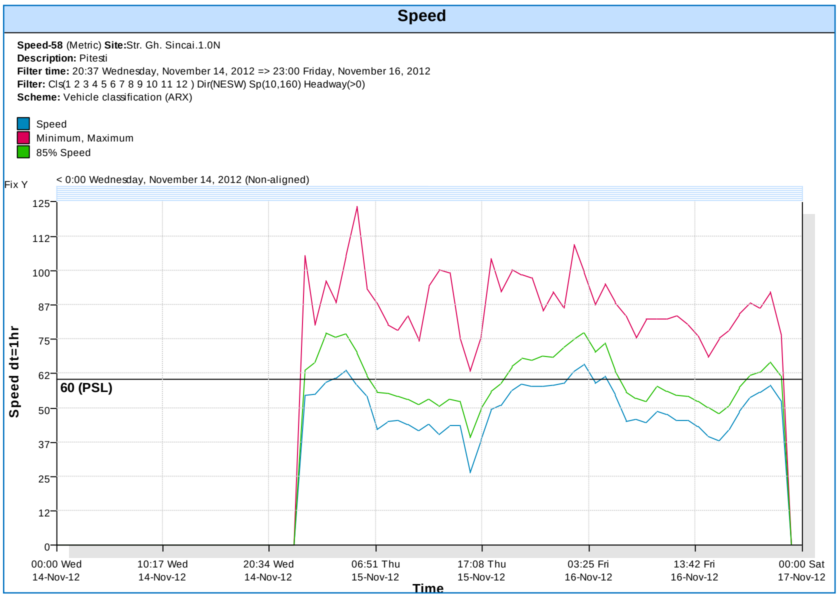



**Fig. 9.**
*Speed report*

The report issued as *velocity dispersion* is a time-based graph indicating relative speed densities (Figure 10). This report is useful for establishing the relationship between speed and traffic density in the measuring location. On this kind of graph is easy to identify the congestion periods, particularly.

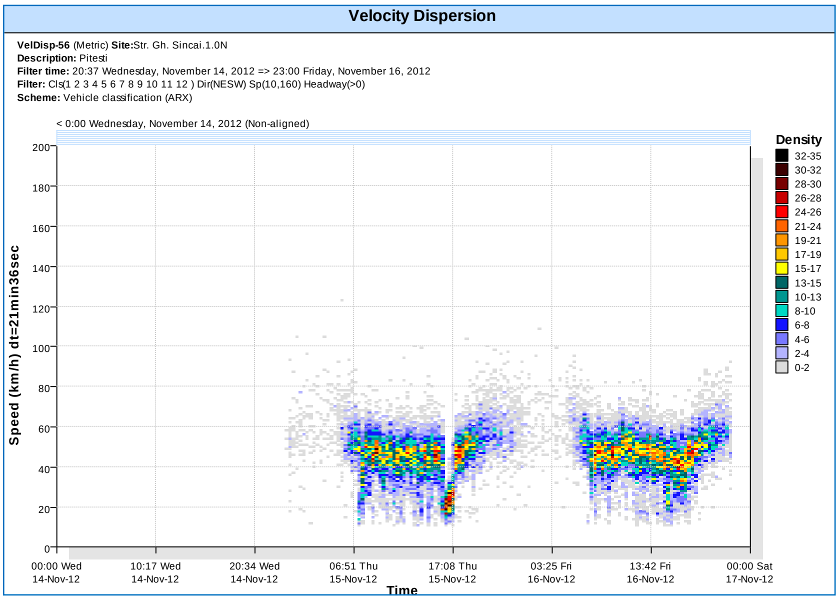



**Fig. 10.**
*Velocity dispersion*

In Figure 11 is shown the diagram of *flow stacked by speed bins* in which the total number of vehicles is differentiated in speed classes, from 10 to 10 km/h.
There are many other types of reports (predefined or customizable) that system offers and which are very useful in transport and traffic studies, not only in the calibration and validation stages of transport model / sub-models.

## 4. Conclusions

Emphasizing once again the importance of calibrating each modeling stage in order to obtain a valid transport model, in this paper were indicated calibration procedures (related to those offered by VISUM software) which, based on traffic volumes and information about average traveling speeds of vehicles, lead to overall transport model calibration, as a result of calibrating Origin - Destination relations and, consequently, of trips generation.

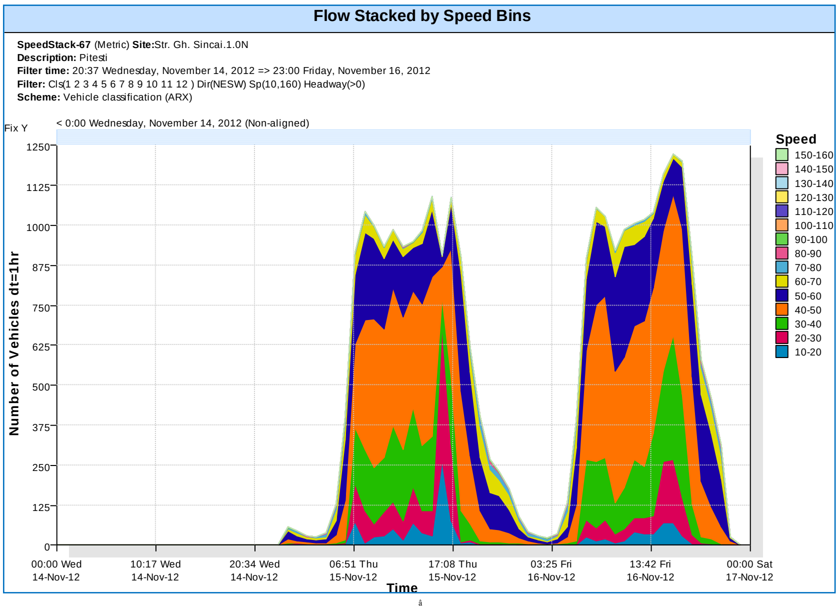



**Fig. 11.**
*Flow stacked by speed bins*

There were detailed the categories of data provided by the software of counting system with pneumatic tubes which were used for calibration and validation of traffic model for Pitesti Municipality. The most important types of reports and how they are used in these procedures were also explained. Bringing a wide range of benefits, including the one related to how it processes and displays counting results, the pneumatic tube system has proved a very useful tool in the stages of collecting data on traffic flows (traffic volumes on vehicle categories, vehicles classification, average speeds), but also in the procedures of calibrating and validating the transport sub-models

## Acknowledgements

This work was funded through the project *PERFORM - Sustainable performance in doctoral and postdoctoral research*, ID: *POSDRU/159/1.5/S/138963*, co-financed by the European Social Fund – Investing in People, within the Sectoral Operational Programme Human Resources Development 2007-2013.

**References**

Azad, H.; Boushehri, N. 2014. Billboard advertising modeling by using network count location problem, *International Journal for Traffic and Transport Engineering,* 4(2): 146 – 160.

Elefteriadou, L. 2014. *An Introduction to Traffic Flow Theory*. ISBN: 978-1-4614-8434-9 (Print) 978-1-4614-8435-6 (Online), Springer Science+Business Media, New York, 251 p.

Federal Highway Administration, 2010. *Travel Model Validation and Reasonableness Checking Manual* Second Edition, Cambridge Systematics, Inc.

Henser, D.; Button, K. 2007. *Handbook of Transport Modelling*. Elsevier Science, Pergamon.

Lippi, O.; Bertini, M.; Frasconi, P. 2010. Collective Traffic Forecasting. In *Proceedings of European Conference, ECML PKDD 2010, Barcelona, Spain, September 20-24, 2010, Part II - Machine Learning and Knowledge Discovery in Databases*, Springer Berlin Heidelberg, 259-273.

MetroCount, The Traffic Data Specialists 2007. Classifier Instalation MC5600 Series RSUs - Operator Guide 3.2. Available from Internet: <http://www.metrocount.com>.

MetroCount, The Traffic Data Specialists 2008. MTE v3.2 Training Manual UK Edition. Available from Internet: <http://www.metrocount.com>.

Mitsakis, E.; Aifadopoulou, G.; Grau, J.; Chrysohoou, E.; Morfoulaki, M. 2014. Combination of macroscopic and microscopic transport simulation models: use case in Cyprus, *International Journal for Traffic and Transport Engineering,* 4(2): 220 – 233.

Ortuzar, J.D.; Willumsen, L. 2011. *Modelling Transport*. 4$^{th}$ Edition, John Wiley and Sons, Ltd., Publication, ISBN: 978-0-470-76039-0, 608 p.

Potočnik, P.; Govekar, E. 2011. Analysis and Short-Term Forecasting of Highway Traffic Flow in Slovenia. In *Proceedings of the 10th International Conference, ICANNGA 2011, Ljubljana, Slovenia, Part I - Adaptive and Natural Computing Algorithms,* Springer Berlin Heidelberg, 270-279.

Rodrigue, J.; Comtois, C.; Slack, B. 2009. *The geography of transport systems.* Routledge Taylor & Francis Group, ISBN13: 978-0-203-88415-7, UK.

Yao, L; Sun, L. 2013. Practical Methods in Traffic Demand Forecasting Model, *Computational Intelligence for Traffic and Mobility, Atlantis Computational Intelligence Systems*, Volume 8, pp 297-319.

# TOWARD A UNIFIED MANAGEMENT OF A COMPLEX ROAD NETWORK

**Fabio Torta [1], Francesca Fermi[2]**
*[1, 2] TRT TRASPORTI E TERRITORIO srl, Via Rutilia 10/8, 20141 Milano (Italy)*

**Abstract:** Road infrastructure charging plays a key role in the EU framework aiming at an effective and fair use of transport policies to promote the 'user pays' and 'polluter pays' principle. The framework contributes to the internalisation of the external costs related to environmental and social impacts generated by the use of the road infrastructure. Furthermore, the application of common rules on road tolls is foreseen to apply harmonised and transparent charging mechanisms across the EU.
The paper aims at investigating and evaluating the application of more efficient integrated motorways toll schemes in the Lombardia Region (Italy), testing the potential of road charging taking into account not only economic and financial aspects, but also environmental and transport impacts, and fairness towards road users.
The paper analyses first the issues related to legal, regulatory and economic aspects for the implementation of an integrated management approach, taking into account current toll differentiation on network segments managed by different companies. Then it presents the results of a transport model which has been implemented to test different pricing policies, including differentiation by emission standard, truck vehicle dimension, time period of the day, road type.
The model results show significant impacts particularly in terms of traffic distribution on the network, while congestion and emissions are less affected.
Impacts are also relevant in terms of revenues for private operators, the most significant ones on newly planned motorway, highlighting how the adoption of different schemes could affect private investment sustainability and revenues distribution among the operators.

**Keywords**: road charging, motorway, road management.

## 1. Introduction

The EU is encouraging the use of transport infrastructure charging in the most effective and fair manner in order to promote the 'user pays' and the 'polluter pays' principle (European Commission, 2011). The framework contributes to the internalisation of the external costs related to environmental and social impacts generated by the use of the road infrastructure. Furthermore, the application of common rules on road tolls is foreseen to apply harmonised and transparent charging mechanisms across the EU. Moreover, road charging can also be a useful instrument to generate new sources of revenue to help the development of transport network and support also cleaner and less energy consuming modes of transport.
In this light, during recent decades road infrastructure charging for heavy goods vehicles and cars have been implemented in several Countries applying various approaches and for different purposes.
As an example, with reference to heavy goods vehicles above 12 tonnes, Belgium, Denmark, Luxembourg, the Netherlands and Sweden have implemented a common system of user charges called the 'Eurovignette' system (www.eurovignettes.eu ). This system allows hauliers after the payment of a specified amount (related to the Emission standard and number of axles of vehicles) to use motorways for a given period (i.e. a day, a week, a month or a year).
In Germany, toll for heavy goods vehicles have been introduced for travelling on motorway network, because of the steady growth of heavy goods traffic (German Federal Trunk Road, 2011). The toll system combines modern mobile communications with satellite-based positioning technology, calculating the charge on the basis of toll route segment, number of axles and emissions class of the vehicles (www.toll-collect.de ).
Nowadays, the technological progress in the field of toll collection system and ITS gives the opportunity to explore the effects of road charges related to the real-time use of the infrastructure and modulated spatially and temporally (B. OEHRY,2010). In this sense, road charging systems play currently a key role to achieve broader objectives, e.g. in the field of internalisation of the external costs related to environmental and social impacts generated by the use of the road infrastructure. Moreover, road tolls are the leverages to provide incentives and disincentives for traffic management, with the purpose of reducing the negative impacts (congestion, pollution) and improve the level of service of infrastructures.
This paper aims at investigating and evaluating the application of more efficient integrated motorways toll schemes in the Lombardia Region (Italy), testing the potential of road charging taking into account not only economic and financial aspects, but also environmental and transport impacts, and fairness towards road users.
The paper is organized as follows. In the first section the issues related to legal, regulatory and economic aspects are analysed, taking into account current toll differentiation on network segments of the Lombardia Region. The following section presents the results of a modelling application where different pricing policies have been tested, including differentiation by emission standard, truck vehicle dimension, time period of the day, road type. In the last section the main results and findings are discussed.

[1] Corresponding author: Fabio Torta: torta@trt.it

## 2. Integrated management approach in road tolls in Italy: legal, regulatory and economic aspects

Road tolls are currently largely differentiated on the Italian network, since the segments of motorways are managed by different companies. In the Lombardia Region there are nine motorways which are managed by six companies with different management contracts and level of road toll. Seven more segments of the motorway network are planned and the management is foreseen in charge of different companies. The resulting picture of the road charges in the region is fragmented and extremely variable on the network also within the same vehicle category.

**Table 1**
*Road tolls on motorway network in the Lombardia Region*

| Management company | Segment of motorway network | Collection system | Toll for cars (€/km) | Toll for trucks (range) (€/km) |
|---|---|---|---|---|
| Autostrade per l'Italia Spa | Autostrada del Sole (A1) | closed (with entry/exit tolls) | 0.069 | 0.071 - 0.143 |
| | Autostrada Dei Laghi (A8 E A9) | open (with mainline barrier toll plazas) | 0.069 | 0.071 - 0.143 |
| | | | 0.082 | 0.084 - 0.167 |
| | Milano - Brescia (A4) | Closed | 0.069 | 0.071 - 0.143 |
| Autostrada Brescia-Verona-Vicenza-Padova Spa | Brescia - Padova (A4) | Closed | 0.065 | 0.066 - 0.134 |
| SATAP Spa (To-Mi) | Milano – Torino (A4) | Closed | 0.106 | 0.108 - 0.214 |
| | | | 0.109 | 0.111 - 0.219 |
| SATAP Spa (To-Pc) | Torino – Piacenza (A21) | Closed | 0.084 | 0.086 - 0.171 |
| Autostrade Centro Padane Spa | Piacenza - Brescia (A21) | Closed | 0.063 | 0.064 - 0.130 |
| Autostrada del Brennero Spa | Autostrada del Brennero (A22) | Closed | 0.062 | 0.064 - 0.129 |
| Milano Serravalle - Milano Tangenziali Spa | Milano Serravalle (A7) | Closed | 0.057 | 0.059 - 0.120 |
| | Tangenziale Est di Milano (A51) | Open | 0.129 | 0.132 - 0.259 |
| | Tangenziale Nord di Milano (A52) | Open | 0.126 | 0.130 - 0.254 |
| | Tangenziale Ovest di Milano (A50): Terrazzano | Open | 0.098 | 0.100 - 0.198 |
| | Tangenziale Ovest di Milano (A50): To-Mi Rho, Arluno e Boffalora | Open | 0.044 | 0.045 - 0.093 |
| | Tangenziale Ovest di Milano (A50): To-Mi-tutte le direzioni | Open | 0.118 | 0.121 - 0.238 |
| | Tangenziale Ovest di Milano (A50): Mi-Ge | Open | 0.060 | 0.061 - 0.125 |
| | Tangenziale Ovest di Milano (A50):Mi-Na e Melegnano | Open | 0.059 | 0.060 - 0.123 |

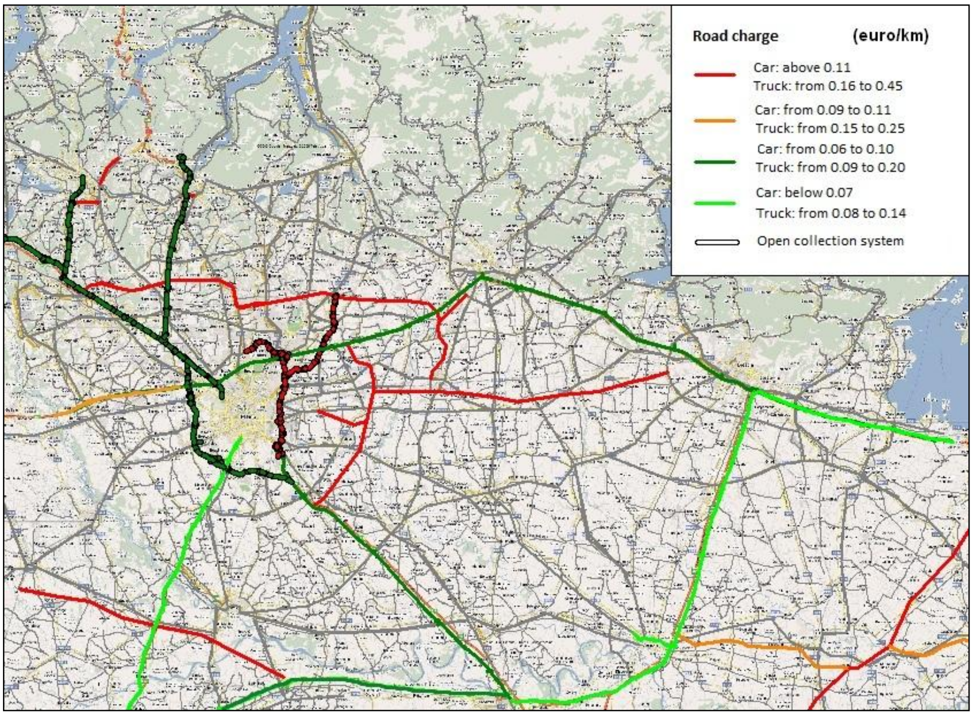



**Fig. 1.**
*Road tolls on existing and planned motorways in the Lombardia region*
*Source: TRT elaboration*

In the Lombardia Region, as throughout the country, the tolled network management has been often characterized by the persistent extension of existing contracts with the motorway companies, formally justified by the need of new investments and/or to solve disagreements arisen over time.
Although a national directive sought to limit this phenomenon, the tendency to grant the motorway companies with extensions of contracts hasn't been curbed (Directive Costa-Ciampi, 1997). This situation is in fact an obstacle to the introduction of tariff systems with different approach with respect to the existing configuration.
Furthermore, the grantors of tolled motorway concessions are state-owned, regional or mixed: therefore a number of competing interests need to be balanced to apply an integrated approach in road tolls.

In Italy, a national Authority for transport regulation has been established by law in 2011 and it is operational since September 2013 (Italian Transport national authority, 2011). It is responsible for the regulation of the transport sector and the access to infrastructure and related ancillary services. Its tasks include the definition of the levels of quality of transport services and the minimum content of the user rights. The Authority is an independent administrative authority and reports annually to the Parliament.
In the field of the road motorway charging, the Authority establishes the tolls tariff systems for the new concessions and defines the grant schemes to be respected with reference to the management or construction of motorway infrastructures.

Having in mind a process of standardization of road charging schemes, the Authority represents the subject through which the negotiation with stakeholders should be developed.
In fact, in order to allow the acceptability of a new scheme, the methods of estimation and evaluation of variations of revenues and traffic levels over time should be agreed with the stakeholders. Furthermore, compensation mechanisms and incentives should be defined depending on the configuration resulting e.g. from the application of a new integrated approach.
An option could be to remove the direct link between the tolls paid by users and the profits of the companies: the grantors would be required to pay back into a common fund the differences between the user charges collected and the level of revenues preliminarily agreed. The fund would be used to support new investments and to compensate the companies with an amount of charges collected lower than the revenues agreed with the Authority.
Another hypothesis could involve the use of the "unbundling": once current contracts are expired, the activities related to toll collection, maintenance of infrastructure and ancillary services would be assigned to other private companies through the use of public procurement procedures. The difference between the revenues from user fees and the cost of managing the network would be allocated to a common fund for financing new infrastructures.
Finally, there is a growing interest in the possibility of applying a road charge for the use of the primary road network, currently free of charge.

This option might be considered only with a "fiscally neutral" approach, with consequent cuts of other transport taxes (i.e. circulation and registration taxes), in order to avoid to be perceived as an additional increase in the tax burden at local level. In addition, for reasons of equity, the revenues from this new road charge should be allocated to the regional road network and the transport sector as a whole, with the aim of improving the overall performances of the regional transportation system.

In any cases, the Authority should play a key role in the process, allowing the application of generalized and sophisticated toll collection systems, which might operate in a flexible way in space and time with the aim of optimizing the traffic flows on a regional scale.
In this perspective, the mechanisms of motorway concession would require a major revision: if tariffs and revenues are centrally managed at regional level both the need and the opportunity for long lasting concessions would be undermined. Building and maintenance activities could be assigned with separate contracts. For maintenance and toll collection activities (in case it is not completely centralized with the ITS systems) the duration of contracts could be in the range of 5-7 years, which is the time required to implement management strategies and amortize the related investments.
The role of private capital investment (in some cases more formal than substantial) would be replaced by lines of credit negotiated with banks on the basis of forecasts on the revenues from traffic of the centralized collection system.

The following table summarises two options of an integrated toll system, applied on the motorway network or extended to the primary network. The table is prepared with the approach of a SWOT analysis, which allows to highlight strengths and weaknesses as well as opportunities and threats.

**Table 2**
*SWOT analysis of the road charging options for the Lombardia Region*

| | **Integrated management approach with road charge on motorway network** | **Integrated management approach with road charge on both motorway and primary network** |
|---|---|---|
| **Strengths** | Harmonised management approach on the regional network<br>Fairness towards road users<br>Road toll related to usage of infrastructure | Harmonised management approach on the regional network (core and primary)<br>Fairness towards road users<br>Road toll related to usage of infrastructure |
| **Weakness** | Long time for the application due to multiplicity of subjects involved<br>Problems with motorways in Project Financing<br>Legislative issues<br>Differences with respect to other Italian regions | Cost and time for the implementation of charging infrastructures, control systems and vehicle equipment<br>Issues of consensus<br>Management of the secondary road network (congestion, safety, etc.) |
| **Opportunities** | Revenues to be invested in the road management sector / the transport sector (i.e. also PT and railways) | Revenues to be invested in the road management sector / the transport sector (i.e. also PT and railways) |
| **Threats** | Further changes of national legislation<br>Appeal | Further changes of national legislation<br>Appeal |

*Source: TRT analysis*

Both options of integrated management toll system show similar strengths: the application of an harmonised management approach on the regional network (core or extended to the primary roads) , the fairness guaranteed to all users, and the implementation of a charge related to the actual use of the road network. The quantification of the benefits in terms of travel time, emissions, revenues, but also safety (to the extent that it is related to the amount of traffic) has been estimated with a transport model, as described in the following paragraph.
In terms of weakness, the implementation of an integrated toll system on the motorway network involves a multiplicity of subjects and the revision of current legislation would take a long time. In addition, problems might emerge with motorways in Project Financing since their tolls are significantly higher than those of existing infrastructures: of course the toll applied with an integrated and unified approach could not be settled at the same value and the financial sustainability of the new motorways would be difficult without some form of compensation. Finally, although the Lombardia region might benefit by the application of an integrated approach, the surrounding regions might have different strategies and the overall system might still be fragmented (e.g. the Veneto region and its approach of management of the Mestre ring-road).
With reference to the application of a road charge on the primary network, the main weakness point is related to the investment costs required for the implementation in terms of charging infrastructures, control systems and vehicle equipment, in order to avoid that the toll collection slows down the traffic flow and affect transport performances. Nevertheless, these investments represent at the same time a great potential in the near future where sustainable mobility is foreseen as a driving aspect of European policies and the Galileo system is fully operational.

Another critical issue is related to the consensus of individuals, used to travel on the ordinary roads for free and assuming they are financed by general taxation: from this point of view the tariff system should be carefully defined, encouraging freight vehicles toward the use of motorways and travelling with less polluting vehicles and during off-peak time period. Finally, this policy would require a careful monitoring and management of the secondary road network (not involved in the charging system), which is likely to face an increase of traffic which might result in congestion and safety problems.

The main opportunity of both options is the possibility of collecting additional resources, which might be used to support maintenance and management of the existing network as well as to contribute to the new investments. Moreover, investments could be related to the transport sector as a whole (e.g. TPL, cycling networks, etc..): the improvement, adaptation and development of the regional transport network and its services can adequately compensate the expectation of the users, with improvements of performances and safety over time. Nevertheless, the timing is a key variable to be taken into account, in order to avoid that the benefits associated with the new investments are enjoyed at times excessively delayed with respect to the application of the toll.

Threats include in both cases mainly administrative and legal aspects related to the possibility of further changes in national legislation on road pricing, as well as appeals which may come from different institutional actors with respect to a regional legislation introducing new concepts, procedures and methods of toll collection, in comparison to the traditional approaches applied on the rest of national territory. In particular, in Italy the application of road charging on the ordinary network has been currently experienced only in urban areas (although in complex situations, e.g. the city of Milan) (Martino, 2013); nevertheless, the legal feasibility analyzed in the next paragraph seems to provide enough peace of mind in this sense.

In general, the current constitutional settlement of the division of legislative and administrative powers between State, Regions and local authorities seems to provide a consolidated authority of the Region in the field of management of roads of regional interest (Reform of Title V of the Italian Constitution, 2001). Nevertheless, a revision of this aspect is currently under discussion at Parliament level and the division is expected to be different in the future.

With reference to the application of road charging on the primary road network, an explicit law is missing at national and/or regional level. Therefore, in order to avoid possible appeals, the Lombardia region should apply its right to intervene in legislative and administrative terms on the field of road charging of the regional road network.

Concerning the implementation of an integrated toll system on the motorway network, the legislative and administrative revision of the authority in terms of road tolls schemes have necessarily to take into account the existing concession contracts and related agreements, including the financial plan and the user tariff adjustment. Although this aspect does not affect the authority of the Region in general terms, it constitutes a limitation for the complete application.

In fact, the Region needs to comply with existing agreements until the contracts are expired, i.e. within a time period ranging up to 2038 for the motorway in the Lombardia territory. The national Authority for transport regulation could play a key role in this sense, driving the negotiation with stakeholders to find an agreement for the implementation of an integrated toll system. Nevertheless, in case the consensus with stakeholders is not reached, another option would be the application of the integrated toll system from the user point of view, combined with some form of compensation to comply with current concession contracts.

Given the defined level of road user charge, the Authority would assume the liability to compensate the companies (either directly or through other forms of facilitation) to cover the difference between the amount of charges collected and the revenues agreed with current concession contracts. For such a solution there aren’t evident obstacles in legal terms. The situation should be easier with reference to new motorway infrastructures for which the national Authority is already in charge of the definition of the tolls tariff systems.

## 3. Modelling pricing policies in the Lombardia Region

Starting from the analysis of the issues related to legal, regulatory and economic aspects mentioned above, a transport model has been implemented to test different pricing policies on the road network of the Lombardia region.

The quantitative evaluation of alternative motorways toll schemes has been developed with the aim of identifying an integrated approach increasing the fairness toward road transport users, with explicit parameters for toll differentiation.

### 3.1. The modelling approach

The methodological approach is based on a modelling application. The transport model implemented for the analysis represents the road network in the Lombardia region at a future stage (the year 2020), including the new motorway infrastructures planned at regional level.

The model, built as an application of the software MEPLAN, is a composition of several solid and validated local network models developed by TRT for the evaluation of traffic demand on new motorway infrastructures.

The model simulates the route choice by drivers of cars, light and heavy duty vehicles travelling in the Lombardia region. Transport demand is exogenously given in terms of origin-destination matrices with a high level of segmentation (by emission standard, truck vehicle dimension, time period of the day), in order to allow the simulation of complex road toll systems.

The main purpose of the modelling application is to provide strategic information on the impacts of pricing policies in terms of relative differences of transport, environmental and economic indicators in comparison with a reference situation (where nothing changes). Results are provided at aggregated level (i.e. by types of roads, types of area, etc.) through the following indicators:

- The level of service of the network, measured in terms of average speed and total user costs;
- The revenues collected from passenger and freight traffic;
- The environmental impacts, measured in terms of pollutant and GHG emissions (CO, NOx, VOC, PM, CO2).

Relative differences of these indicators are estimated taking into account short and medium term reactions of users, i.e. mainly in terms of choice of alternative routes and, in a simplified form , modal shift. The model does not deal with the possible long-term changes, such as the renewal of vehicle fleet or relocation of activities.

The model is applied to test the application of integrated road charging policies, involving in some cases both motorways and roads of the primary network.

### 3.2. Main impacts of integrated road charging policies

The model has been applied to test different integrated road charging policies, to be compared with a reference scenario where current and planned charges have been applied to the correspondent existing and planned tolled motorway network.

At first, a “flat” toll has been implemented to analyze the effects of an homogenous approach on existing and planned motorways. A set of values for the charge has been defined on the basis of the analysis of current and planned tolls, applying charge differentiation by truck vehicle dimension only (as in the current system). Since in the reference scenario planned motorways are supposed to apply higher tolls in comparison to existing motorways, the main impacts of the tests are observed in terms of traffic distribution on the network. Depending on the value applied, traffic flows move from routes on existing motorway to planned motorways or, in case of high values of toll, to primary roads free of charge. Despite this re-routing, minor effects could be observed in terms of congestion and emissions. From an economic point of view, the tests of integrated tolls’ schemes have shown to be not neutral in terms of revenues, especially for planned motorways. Although in most of the tests the total revenue from road charging is unchanged or increased with respect to the reference scenario, a loss of earnings is often observed when looking at new infrastructures only. Financial sustainability of private investments and distribution of incomes among different companies is therefore a crucial issue to be taken into account.

Depending on the value applied, user costs are unchanged or slightly increased with respect to the reference scenario.

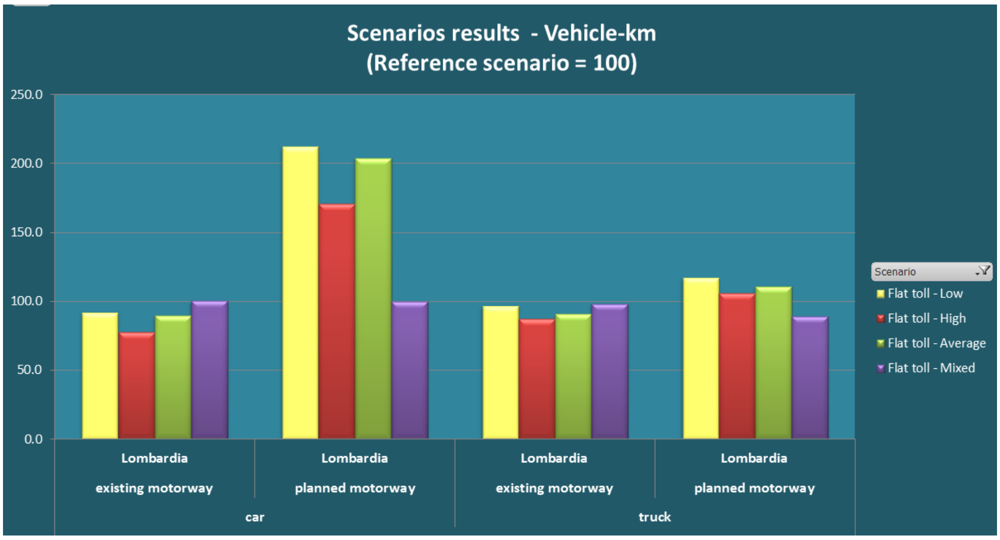



**Fig. 2.**
*Variation of vehicle-km of cars and trucks on the motorway network*
*Source: TRT elaboration*

A further step of the analysis has been done toward a “sustainable” approach for the definition of integrated road charges, with explicit aim of reducing pollutant emissions and improving the level of services. Therefore, tolls has been differentiated by emission standard, truck vehicle dimension, time period of the day and road type, assuming that the road network subject to the integrated tariff includes also the primary network (partially or as a whole). The impacts of this set of tests are extremely variable and strictly related to the characteristic of user segments of the transport demand estimated; therefore, results should be carefully evaluated since different assumptions on several aspects (e.g. vehicle fleet composition) might give rise to different reactions and configuration.

In general, the tests lead to a different traffic distribution with flows moving from existing motorway and ordinary network (when tolled) to the new motorway infrastructure: as a result of a better distribution on the network, congestion is reduced with benefits in terms of travel time. Positive impacts are especially observed during peak time, in case toll is differentiated by time period with incentives to travel during off-peak time period.
The integrated charging system gives rise to an increase of the total level of revenues; nevertheless, this result is depending on the application of a charge to a part of the ordinary road network, since the revenues from tolls collected on the motorway network are always reduced with respect to the reference scenario.
From the users point of view, the total costs are basically unchanged with respect to the reference scenario; nevertheless, the distribution of costs is unequal between trucks and cars, with a slight benefit for the former and a slight detriment for the latter. This result is however influenced by the setup of the tests (motorway tolls are slightly reduced for trucks in order to attract freight traffic on the core network).

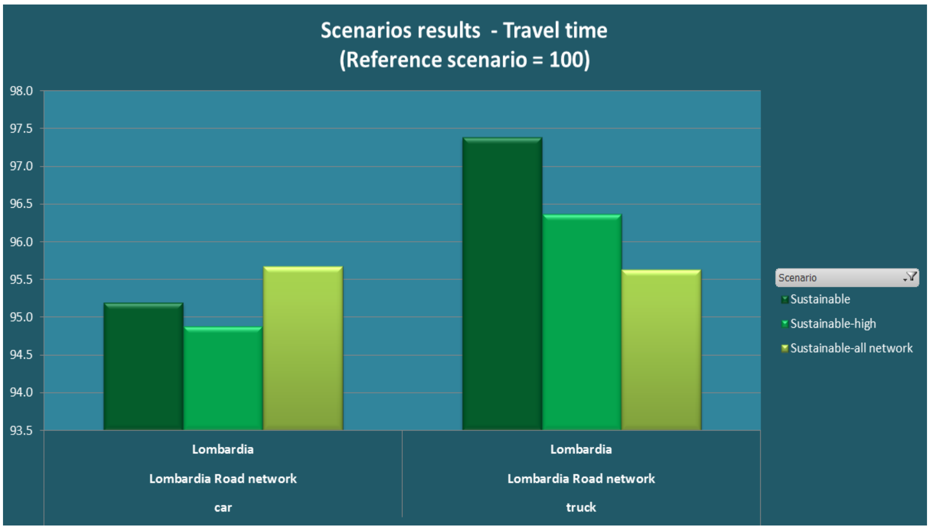



**Fig. 3.**
*Variation of travel time of cars and trucks on the road network*
*Source: TRT elaboration*

Finally, another approach for toll differentiation has been tested strictly related to the wear and tear of infrastructures, i.e. taking into account that bigger vehicles produce a higher infrastructure's deterioration. Therefore, the charge has been differentiated by truck vehicle dimension and applied also on the primary network (assuming higher tolls for truck vehicles).
As a result, freight traffic moves from tolled primary network toward planned motorways and non-tolled roads. Nevertheless, negative impacts might arise in terms of congestion on the ordinary network: in order to compensate this effect, a test has been simulated implementing bans on the ordinary roads for truck vehicles. In this case the heavy duty vehicles basically travel (as expected and desired) on the motorway network, provided that the motorway toll is not too high. The resulting traffic distribution produces positive effects in terms of travel time and it is neutral in terms of pollutant emissions. Otherwise (without bans) a slight increase of pollutant emissions is observed, strictly connected to the impacts on travel time.
From the perspective of network operators, in these tests revenues are generally increased; however, a significant contribution comes from the application of a charge to a part of the ordinary road network.
In terms of user costs, trucks are subject to a higher tariff as defined by definition of the tests. The introduction of bans further accentuates the disparity in the distribution of costs between truck and car drivers.

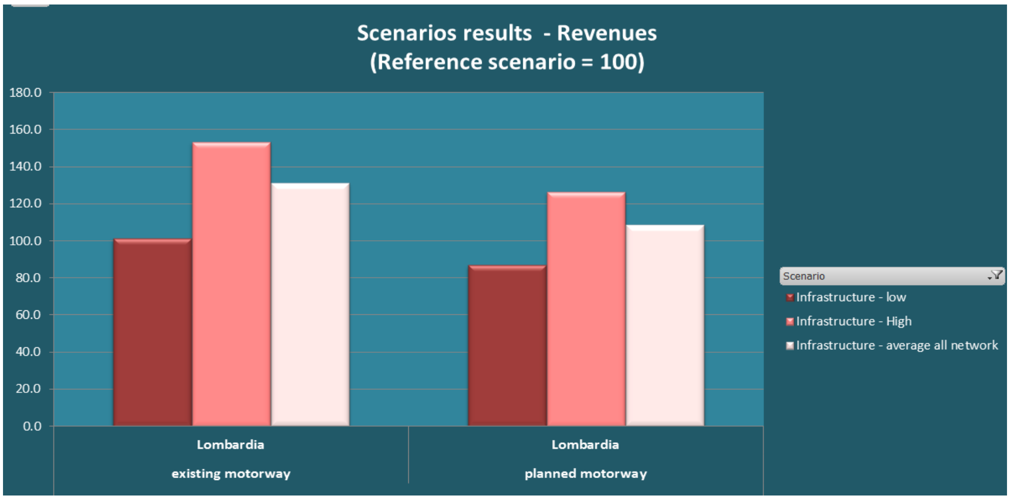



**Fig. 4.**
*Variation of revenues by motorway network*
*Source: TRT elaboration*

In summary, the quantitative analysis shows that significant impacts could be observed in terms of traffic distribution on the network (motorway or primary network), while congestion and emissions are less affected.
Impacts are also relevant in terms of revenues for private operators, mainly on newly planned motorway, highlighting how the adoption of different schemes could affect private investment sustainability and revenues distribution among the operators. Moreover, in many tests the total level of revenues is unchanged or increased with respect to the reference scenario thanks to the application of a charge on part of the ordinary road network.
A scenario of partial integration could be a first step in the direction of an integrated road toll system: the toll applied on existing motorways could be equal to the average of the current ones, and the toll applied on planned motorways equal to the average of their tolls. This road pricing scheme would be substantially neutral from all points of view (level of service, revenues, etc.) while introducing significant elements of simplification and greater equity.

## 4. Conclusion

This paper presents the results of the analysis of the road charging system in the Lombardia Region (Italy) from a legal regulatory and economic point of view on one hand and considering quantitative aspects (in modelling terms) on the other hand.
The first part of the analysis underlined the key role of the national Authority for transport regulation, which represents the main subject to develop and coordinate the negotiation between stakeholders in the light of a process of standardization of road charging schemes. Furthermore, it would be crucial to support the concept of reallocation of revenues considering the transport sector as a whole, with the aim of improving the overall performances of the regional transportation system.
From the quantitative point of view, the analysis of the Lombardia case study suggests that road pricing is a useful instrument to generate new sources of revenue, but seems to be less effective to address issues related to congestion and the environment. Nevertheless, it would be interesting to investigate the impacts with long-term analysis, taking into account the feedback on demand generation.

## Acknowledgements

The paper is based on the study developed by TRT for IreR –Regional Research Institute of Lombardy (Istituto Regionale di Ricerca della Lombardia).

**References**

European Commission, 2011. *Directive 2011/76/EU of the European Parliament and of the Council of 27/09/2011*

Eurovignette road user charge system. Available from Internet: <http:// www.eurovignettes.eu >.

German Federal Trunk Road, 2011. *Toll Act (BFStrMG), 19 July 2011*

Truck toll system in Germany. Available from Internet: <http://www.toll-collect.de>

B. OEHRY, 2010. *Critical Success Factors for Implementing Road Charging Systems*. Discussion Paper No. 2010-3 ITF/OECD Round Table 2010

Italian Transport national authority, 2011. *Art. 37 del decreto-legge 6 dicembre 2011, n. 201*

Directive Costa-Ciampi, 1997. *Decreto Ministeriale del 20 Ottobre 1998, n. 283*

Reform of Title V of the Italian Constitution, 2001. *Gazzetta Ufficiale n. 248 del 24 ottobre 2001*

Martino, et al. 2013. *From pollution charge to congestion charge in Milano, Italy*. WCTR 2013, Rio (Brazil).

# MODELING OF THE INERTIAL TORQUE TRANSFORMER WITH METHODS OF INTERVAL ANALYSIS

**Semen L. Blyumin[1], Alexander V. Galkin[2], Pavel V. Saraev[3], Anton S. Sysoev[4]**
[1,2,3,4] *Lipetsk State Technical University, Lipetsk, Russia*

**Abstract:** Modern technical devices used in the production of vehicles have a complex structure, which creates difficulties for their mathematical modeling. The main problem is the fact that approximate methods of such calculating contain a computational error. In designing such devices the value of computational error modeling can greatly influence the determination of their parameters. It may lead to the deterioration of reliability and durability of these devices. The inertial torque transformer (ITT) can be taken as an example of such device. ITT is an infinitely variable mechanical transmission which has the ability of automatical changing the gear ratio, depending on the angular velocity of the output shaft and the magnitude of the external load resistance. The advantages of this device are high coefficient of efficiency, compact construction, opportunity to work on the mode of direct transmission. The disadvantage is the fragility of ITT free-wheel mechanism due to its constant work with high frequency and load. Certain methods for selecting optimal parameters of the device should be developed in order to improve the durability of the ITT. The mathematical model of ITT is a set of composite systems of nonlinear differential equations. The exact analytical solution of these systems cannot be obtained. But usually Runge Kutta methods are used in order to solve these equations. These methods are approximate. The improvement of the accuracy of these methods is achieved by reducing the step. It may lead to an increase in the computational error. As a result, the solution is not accurate. But it also possible to estimate the exact solution, indicating the lower and the upper limit values which guarantee the assessment solution. The paper contains an interval problem of modeling the workflow of ITT, the modification of interval Runge Kutta method and the interpretation of the obtained solution.

**Keywords**: inertial torque transformer (ITT), mathematic modeling, interval analysis.

## 1. Introduction

In the present article we analyze the modeling of ITT. Modeling is used to solve the differential equations describing the operation of the device. Modeling is an important step to determine the optimum process parameters. Due to the fact that the applied method of solution is numerical, and not all the values of parameters of the device are determined reliably the obtained modeling values are not reliable. The lack of information about the possible flow of the process may strongly influence the choice of non-optimal parameters. In this regard, the article is to investigate application of interval analysis to the solution of systems of differential equations describing the workflow of ITT. This method allows determining reliable estimates of the possible values obtained in modeling.

## 2. Mathematical model of inertial torque transformer

The physical model of inertial torque transformer with a generalized impulse mechanism and two fixed free-wheel mechanisms (FWM) is shown in Fig. 1. The model is described in [Leonov (1978)], [Bazhenov (2003)].

The work of a transformer is performed in the following way. The master flywheel 1 sets in motion the freight links 10 of the impulse mechanism (as any impulse mechanism may be used in the scheme the freight links are shown conventionally). Inertial forces of freight links arising from the displacement of the center of mass relative to the rotational axis link, create a sign-variable torque on the transmission shaft 2 (reactor). The reactor is not rigidly linked to the master and slave flywheels. Direct impulse tends to rotate the reactor in the rotational direction of the shaft of the engine, and is transmitted through the outgoing FWM 7 to the outgoing flywheel 4. The back impulse tends to rotate reactor in the opposite direction and is carried by FWM 8, which is rigidly connected with housing 5 of the inertial transformer.

The work of the inertial transformer is cyclical. During the cycle, there are four segments in which the differential equations of motion are not changing:

– reactor acceleration to the angular velocity of the slave flywheel;
– joint motion of the reactor and the slave flywheel;
– braking of the reactor to a stop;
– motion dwell of the reactor.

[4]Corresponding author: anton_syssoyev@mail.ru

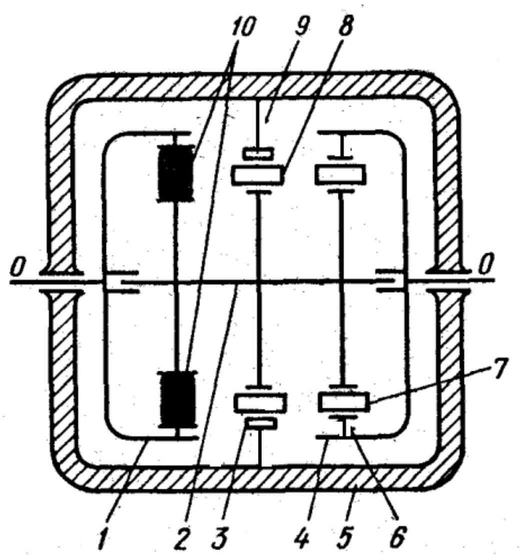



**Fig. 1.**
*General layout of an inertial transformer*

The first segment begins simultaneously with the commencement of the direct impulse of the inertial forces acting on the intermediate shaft, which causes it to rotate in the direction of rotation of the master flywheel. Slave and master flywheels work independently of each other. Reactor acceleration ends when the reactor runs up the angular velocity of the slave flywheel.

Acceleration of the slave flywheel under the influence of the direct impulse transmitted through the reactor occurs at the area of joint motion of the reactor and the slave flywheel. Joint movement ends with the expiration of the direct impulse.

At the braking segment of the reactor the speed of the reactor decreases under the influence of the back impulse. The segment of braking ends when the speed of the reactor comes to zero.

At the moment of the motion dwell of the reactor the back impulse by means of housings of FWM is transferred to the housing of inertial transformer.

The mathematical model of ITT is the systems of differential equations describing the motion of the master flywheel, the reactor and the slave flywheel at each part of the cycle.

The system of differential equations describing the workflow of ITT at acceleration segment of the reactor looks like (1).

$$(1)\quad \begin{cases} B_1(\psi)\ddot{\varphi}_{21} + B_2(\psi)\ddot{\varphi}_{22} - B_4(\psi)(\dot{\varphi}_{21} - \dot{\varphi}_{22})^2 + B_6(\psi)\dot{\varphi}_{22}^2 = M_Д; \\ B_2(\psi)\ddot{\varphi}_{21} + B_3\ddot{\varphi}_{22} - B_6\dot{\varphi}_{21}^2 = 0; \\ J_П\ddot{\varphi}_1 = -M_C. \end{cases}$$

The transition to the second segment occurs when the angular velocity of the reactor reaches the angular velocity of the slave link, i.e. $\dot{\varphi}_1(t_1) = \dot{\varphi}_{22}(t_1)$ .

At the stage of joint motion the reactor and the master flywheel move as one unit and therefore $\dot{\varphi}_{21} = \dot{\varphi}_1$. The system of differential equations is (2).

$$(2)\quad \begin{cases} B_1(\psi)\ddot{\varphi}_{21} + B_2(\psi)\ddot{\varphi}_{22} - B_4(\psi)(\dot{\varphi}_{21} - \dot{\varphi}_{22})^2 + B_6(\psi)\dot{\varphi}_{22}^2 = M_Д; \\ B_2(\psi)\ddot{\varphi}_{21} + B_5\ddot{\varphi}_{22} - B_6(\psi)\dot{\varphi}_{21}^2 = -M_C. \end{cases}$$

Condition of transition is a twist of the satellite in the relative motion at π radians, i.e. $\varphi_{21}(t_2) - \varphi_{22}(t_2) = \frac{\pi}{a}$. .

The third section has the same equations as the first one (1). Upon reaching the reactor the angular velocity equal to zero the system goes to the fourth part of the work, i.e. $\dot{\varphi}_{22}(t_3) = 0.$

The fourth segment is described by the following equations

$$(3)\quad \begin{cases} B_1(\psi)\ddot{\varphi}_{21} - B_4\dot{\varphi}_{21}{}^2 = M_Д; \\ J_П\ddot{\varphi}_1 = -M_C, \end{cases}$$

where $\varphi_i, \dot{\varphi}_i$ – generalized coordinates and generalized velocities,

$B_1(\psi) = J_{21} + nme^2 + 2nmed(1+a)\cos\psi + nJ_Г(1+a)^2,\ \ B_2(\psi) = -anJ_Г(1+a) - nmaed\cos\psi,\ \ B_3 = J_{22} + nJ_Г a^2,$

$B_4(\psi) = nmaed(1+a)\sin\psi,\ \ B_5 = B_3 + J_П,\ \ B_6(\psi) = nmaed\sin\psi,\ \ \psi(t) = a(\varphi_{21} - \varphi_{22}),$

$a$ – inner gearing ratio; $n$ – number of freight links; $m$ – mass of freight link; $d$ – distance from the axis of rotation of a freight link to the center of its gravity; $e$ – distance from the axis of rotation of ITT to the axis of rotation of a freight link; $J_{21}$ – reduced torque of inertia of the leading elements; $J_{22}$ – reduced torque of inertia of the leading portion of the reactor; $J_{П}$ – reduced torque of inertia of slave elements; $J_{Г}$ – reduced torque of inertia of the freight link, $M_{Д}$ – engine torque, $M_C$ – resistive torque. Initial conditions for the first section $\varphi_{21}(0)=\varphi_{210}$, $\dot{\varphi}_{21}(0)=\dot{\varphi}_{210}$, $\varphi_{22}(0)=\varphi_{220}$, $\dot{\varphi}_{22}(0)=\dot{\varphi}_{220}$, $\varphi_{1}(0)=\varphi_{10}$, $\dot{\varphi}_{1}(0)=\dot{\varphi}_{10}$. For the initial values for the subsequent sections the final values of the previous sections are used, which is the result of the continuity of the process.
An example of the solutions of (1) - (3) obtained by the Runge-Kutta fourth-order method of accuracy is shown in Figure 2.

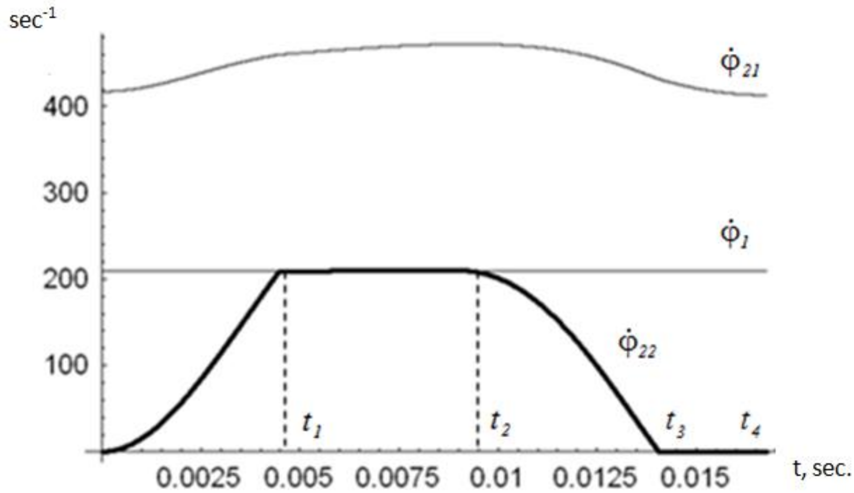



**Fig.2.**
*Graphs of angular velocities variation*

Number of model parameters is measured with errors i.e. their true values are within the ranges determined by the errors of measuring instruments. In particular, the evaluation of the true value of engine torque is the interval $[M_{Д}-\varepsilon_1, M_{Д}+\varepsilon_1]$. Substituting this interval to the model (1) - (3), we obtain interval workflow model of ITT. Intervals will also be a solution of such a model $[\underline{\varphi}_i, \overline{\varphi}_i]$.
The problem is formulated as follows: – to solve a system of differential equations (1) - (3) with interval parameters under the given initial conditions.

### 3. Algorithm for solving the span task

To solve this task, the methods of interval analysis may be used. Elements of interval analysis are described in [Shary (2012)], [Kalmykov, Shokin, Yuldashev (1986)]. Interval is any simply connected subset of the set of real numbers
$[x]=[\underline{x},\overline{x}]=\{x \mid \underline{x}\le x\le \overline{x}\}$.
The following basic binary operations are defined over intervals:
- multiplication by a scalar

$$a[x]=\begin{cases}[a\underline{x},a\overline{x}], a\ge 0,\\ [a\overline{x},a\underline{x}], a<0,\end{cases}$$

- addition of intervals

$[x]+[y]=[\underline{x}+\underline{y},\overline{x}+\overline{y}]$,

- difference of the of intervals

$[x]-[y]=[\underline{x}-\overline{y},\overline{x}-\underline{y}]$,

- product of intervals

$[x][y]=[\min\{\underline{x}\underline{y},\underline{x}\overline{y},\overline{x}\underline{y},\overline{x}\overline{y}\},\max\{\underline{x}\underline{y},\underline{x}\overline{y},\overline{x}\underline{y},\overline{x}\overline{y}\}]$,

- division of intervals

$$\frac{[x]}{[y]}=[x]\left[\frac{1}{\overline{y}},\frac{1}{\underline{y}}\right],\ 0\notin[\underline{y},\overline{y}].$$

The method for solving this task is examined on the example of the system of equations (1). At present time various algorithms for solving interval differential equations are developed. In [Shary (2012)] interval Runge-Kutta methods are presented. To solve the given task it is necessary to take into account its features.

The function argument is the time, which is not the value of the interval. Only part of the parameters is interval. An algorithm based on posteriori error estimation is considered [Dobronets (2004)]. The main steps of this algorithm as applied to the system (1):

1. The system is solved (1) by the Runge-Kutta fourth-order method for the specific values of the parameters $M_Д$ и $M_C$ on the equally spaced grid $\omega_h$. For the solution the system (1) is converted to normal view

(4)
$$\begin{cases} \dot{\varphi}_{21} = \chi; \\ \dot{\chi} = \dfrac{B_6(\psi)(B_3\xi^2 + B_2(\psi)\chi^2) - B_3B_4(\psi)(\chi-\xi)^2 - B_3M_Д}{B_2^2(\psi) - B_1(\psi)B_3}; \\ \dot{\varphi}_{22} = \xi; \\ \dot{\xi} = \dfrac{B_2(\psi)B_4(\psi)(\chi-\xi)^2 - B_6(\psi)(B_2(\psi)\xi^2 + B_1(\psi)\chi^2) + B_2(\psi)M_Д}{B_2^2(\psi) - B_1(\psi)B_3}; \\ \dot{\varphi}_1 = \varsigma; \\ \dot{\varsigma} = -\dfrac{M_C}{J_П}. \end{cases}$$

2. Using the obtained values of the functions $\varphi_{21}$, $\varphi_{22}$, $\varphi_1$, $\chi$, $\xi$, $\varsigma$ in grid nodes $\omega_h$ and their derivatives calculated approximately by the system of equations (4) and cubic Hermite splines are set up $s_i$ [Stechkin, Subbotin (1976)]. Functions are defined

(5) $$\phi_i(t,s) = f_i(t,s) - \frac{ds_i}{dt},$$

where $f_i(t,s)$ – functions of the right hand side of the equation (4). Approximation of solutions by using cubic Hermite splines is caused by the following aspects: for Hermite splines approximating the numerical solution of differential equations, as shown in [Dobronets (2004)], the following relation for the evaluation of deviations from the exact solution is correct

$$\left|\left|\frac{d^v(y_i - s_i)}{dt^v}\right|\right| \le c(h^{4-v}\,||y_i|| + h^k),$$

where $v$=0,1; $k$ is determined by the order of accuracy of the numerical methods, in particular by using the Runge-Kutta fourth-order method $k$=4; $y_i$ – values of the system solutions (1) in $i$-grid nodes $\omega_h$.

The solution error $(y_i - s_i)$ consists of the error in the numerical method and the approximation error. And, in its turn, the size of the error determines the interval for the true values of the system of differential equations. As a result, in step 2, we construct the function of the deviation of approximation of the numerical solution from the true.

3. Two systems of differential equations are set up and solved numerically $\frac{du}{dt} = Wu + w$, $u(0) = 0$; $\frac{dv}{dt} = Wv$, $v(0) = 0$,

where $w$ – unit vector, and the elements of matrix $W$ are defined as follows:: $W_{ii} = \frac{\partial f_i}{\partial y_i}(t,s)$ и $W_{ij} = \left|\frac{\partial f_i}{\partial y_j}(t,s)\right|$, $i \ne j$.

For the system of equations (4), the matrix $W$ will be as follows

$$W = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ W_{21} & W_{22} & W_{23} & W_{24} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ W_{41} & W_{42} & W_{43} & W_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

The elements of the matrix W are as follows

$$W_{21} = \frac{(ademn(a(a^2n^2(J_g + aJ_g + dem\cos\psi)^2 - (J_{22} + a^2J_gn)(J_{21} + (1+a)^2J_gn + e^2mn + 2(1+a)emn\cos\psi))\cdot}{(a^2n^2(J_g + aJ_g + demn\cos\psi)^2 - (J_{22} + a^2J_gn)(J_{21} + (1+a)^2J_gn + e^2mn + 2(1+a)demn\cos\psi))^2}$$

$$\frac{(-(aJ_gn(\chi + a\chi - a\xi)^2 + J_{22}((1+a)\chi^2 - 2(1+a)\chi\xi + a\xi^2))\cos\psi - ademn\chi^2\cos 2\psi) + 2(-(1+a)J_{22} + a^2demn\cos\psi)\cdot}{(a^2n^2(J_g + aJ_g + demn\cos\psi)^2 - (J_{22} + a^2J_gn)(J_{21} + (1+a)^2J_gn + e^2mn + 2(1+a)demn\cos\psi))^2}$$

$$\frac{\sin\psi(-M_Д(J_{22} + a^2J_gn) - ademn(aJ_gn(\chi + a\chi - a\xi)^2 + J_{22}(1+a)\chi^2 - 2(1+a)\chi\xi + a\xi^2) + ademn\chi^2\cos\psi)\sin\psi)))}{(a^2n^2(J_g + aJ_g + demn\cos\psi)^2 - (J_{22} + a^2J_gn)(J_{21} + (1+a)^2J_gn + e^2mn + 2(1+a)demn\cos\psi))^2},$$

$$W_{22} = \frac{(2ademn(-(1+a)(J_{22} + a^2J_gn)(\chi - \xi) - an\chi(J_g + aJ_g + dem\cos\psi))\sin\psi)}{(a^2n^2(J_g + aJ_g + demn\cos\psi)^2 - (J_{22} + a^2J_gn)(J_{21} + (1+a)^2J_gn + e^2mn + 2(1+a)demn\cos\psi))},$$

$$W_{23}=\frac{(ademn(a(a^2n^2(J_g+aJ_g+dem\cos\psi)^2-(J_{22}+a^2J_gn)(J_{21}+(1+a)^2J_gn+e^2mn+2(1+a)emn\cos\psi))\cdot}{(a^2n^2(J_g+aJ_g+demn\cos\psi)^2-(J_{22}+a^2J_gn)(J_{21}+(1+a)^2J_gn+e^2mn+2(1+a)demn\cos\psi))^2}$$

$$\frac{((aJ_gn(\chi+a\chi-a\xi)^2+J_{22}((1+a)\chi^2-2(1+a)\chi\xi+a\xi^2))\cos\psi-ademn\chi^2\cos2\psi)-2(-(1+a)J_{22}+a^2demn\cos\psi)\cdot}{(a^2n^2(J_g+aJ_g+demn\cos\psi)^2-(J_{22}+a^2J_gn)(J_{21}+(1+a)^2J_gn+e^2mn+2(1+a)demn\cos\psi))^2}$$

$$\frac{\sin\psi(-M_{\text{Д}}(J_{22}+a^2J_gn)-ademn(aJ_gn(\chi+a\chi-a\xi)^2+J_{22}(1+a)\chi^2-2(1+a)\chi\xi+a\xi^2)+ademn\chi^2\cos\psi)\sin\psi)))}{(a^2n^2(J_g+aJ_g+demn\cos\psi)^2-(J_{22}+a^2J_gn)(J_{21}+(1+a)^2J_gn+e^2mn+2(1+a)demn\cos\psi))^2},$$

$$W_{24}=\frac{(2ademn(\chi+a\chi-a\xi)\sin\psi)}{(-J_{21}(J_{22}+a^2J_gn)-n((1+a)^2J_gJ_{22}+e^2J_{22}m+a^2e^2J_gmn)+demn\cos\psi(-2(1+a)J_{22}+a^2demn\cos\psi))},$$

$$W_{41}=\frac{(a^2demn((a^2n^2(J_g+aJ_g+dem\cos\psi)^2-(J_{22}+a^2J_gn)(J_{21}+(1+a)^2J_gn+e^2mn+2(1+a)emn\cos\psi))\cdot}{(a^2n^2(J_g+aJ_g+demn\cos\psi)^2-(J_{22}+a^2J_gn)(J_{21}+(1+a)^2J_gn+e^2mn+2(1+a)demn\cos\psi))^2}$$

$$\frac{(-(J_{21}\chi^2+n(e^2m\chi^2+(1+a)J_g(\chi+a\chi-a\xi)^2))\cos\psi-demn((1+a)(2+a)\chi^2-2a(1+a)\chi\xi+a^2\xi^2)\cos2\psi+M_{\text{Д}}\sin\psi+}{(a^2n^2(J_g+aJ_g+demn\cos\psi)^2-(J_{22}+a^2J_gn)(J_{21}+(1+a)^2J_gn+e^2mn+2(1+a)demn\cos\psi))^2}$$

$$\frac{2n(-(1+a)J_{22}+a^2demn\cos\psi)\sin\psi(-(1+a)J_gM_{\text{Д}}-dem(J_{21}\chi^2+n(e^2m\chi^2+(1+a)J_g(\chi+a\chi-a\xi)^2\sin\psi+dem\cos\psi}{(a^2n^2(J_g+aJ_g+demn\cos\psi)^2-(J_{22}+a^2J_gn)(J_{21}+(1+a)^2J_gn+e^2mn+2(1+a)demn\cos\psi))^2}$$

$$\frac{(-M_d-demn((1+a)(2+a)\chi^2-2a(1+a)\chi\xi+a^2\xi^2)\sin\psi))))}{(a^2n^2(J_g+aJ_g+demn\cos\psi)^2-(J_{22}+a^2J_gn)(J_{21}+(1+a)^2J_gn+e^2mn+2(1+a)demn\cos\psi))^2},$$

$$W_{42}=\frac{(2ademn(-J_{21}+(1+a)^3J_g+e^2m)n)\chi+a(1+a)^2J_gn\xi-(1+a)demn((2+a)\chi-a\xi)\cos\psi)\sin\psi)}{(a^2n^2(J_g+aJ_g+demn\cos\psi)^2-(J_{22}+a^2J_gn)(J_{21}+(1+a)^2J_gn+e^2mn+2(1+a)demn\cos\psi))},$$

$$W_{43}=\frac{(a^2demn((a^2n^2(J_g+aJ_g+dem\cos\psi)^2-(J_{22}+a^2J_gn)(J_{21}+(1+a)^2J_gn+e^2mn+2(1+a)emn\cos\psi))\cdot}{(a^2n^2(J_g+aJ_g+demn\cos\psi)^2-(J_{22}+a^2J_gn)(J_{21}+(1+a)^2J_gn+e^2mn+2(1+a)demn\cos\psi))^2}$$

$$\frac{((J_{21}\chi^2+n(e^2m\chi^2+(1+a)J_g(\chi+a\chi-a\xi)^2))\cos\psi+demn((1+a)(2+a)\chi^2-2a(1+a)\chi\xi+a^2\xi^2)\cos2\psi-M_{\text{Д}}\sin\psi-}{(a^2n^2(J_g+aJ_g+demn\cos\psi)^2-(J_{22}+a^2J_gn)(J_{21}+(1+a)^2J_gn+e^2mn+2(1+a)demn\cos\psi))^2}$$

$$\frac{2n(-(1+a)J_{22}+a^2demn\cos\psi)\sin\psi(-(1+a)J_gM_{\text{Д}}-dem(J_{21}\chi^2+n(e^2m\chi^2+(1+a)J_g(\chi+a\chi-a\xi)^2\sin\psi+dem\cos\psi}{(a^2n^2(J_g+aJ_g+demn\cos\psi)^2-(J_{22}+a^2J_gn)(J_{21}+(1+a)^2J_gn+e^2mn+2(1+a)demn\cos\psi))^2}$$

$$\frac{(-M_d-demn((1+a)(2+a)\chi^2-2a(1+a)\chi\xi+a^2\xi^2)\sin\psi))))}{(a^2n^2(J_g+aJ_g+demn\cos\psi)^2-(J_{22}+a^2J_gn)(J_{21}+(1+a)^2J_gn+e^2mn+2(1+a)demn\cos\psi))^2},$$

$$W_{44}=\frac{(2a^2demn^2(\chi+a\chi-a\xi)(J_g+aJ_g+dem\cos\psi)\sin\psi}{(-J_{21}(J_{22}+a^2J_gn)-n((1+a)^2J_gJ_{22}+e^2J_{22}m+a^2e^2J_gmn)+demn\cos\psi(-2(1+a)J_{22}+a^2demn\cos\psi))}.$$

The use of these auxiliary systems of differential equations is explained as follows. The equation $\frac{ds_i}{dt}=f_i(t,s)-\phi_i(t,s)$, obtained in accordance with (5), and the initial system of differential equations leads to the following equations for estimating the error in the numerical solution: $\frac{d(y_i-s_i)}{dt}=f_i(t,y)-f_i(t,s)+\phi_i(t,s)$, and using the mean-value theorem [Dobronets (2004)], we obtain the following system $\frac{d(y_i-s_i)}{dt}=\sum_{j=1}^{n}\frac{\partial f_i}{\partial y_j}(t,r_i)(y_j-s_j)+\phi_i$.

In a general view it can be rewritten as $\frac{d\varepsilon}{dt}=V\varepsilon+\zeta$. In [Dobronets (2004)] it is mentioned that under the following conditions $V_{ii}\geq\frac{\partial f_i}{\partial y_i}(t,r_i)$, $V_{ij}\geq\left|\frac{\partial f_i}{\partial y_j}(t,r_i)\right|$, $\zeta_i(t)\geq|\phi_i(t,s)|$ the following estimate is valid for solution $\varepsilon_i(t)\geq|y_i(t)-s_i(t)|$.

Thus, the approximation of the numerical solutions of the auxiliary systems of differential equations may be used to estimate the interval width containing the true solution of the system (1).

4. Hermite splines are set up $s_i^{(1)}$ и $s_i^{(2)}$ for solutions obtained for the third step *u* and *v* on grid $\omega_h$. Solution of the system (1) with interval parameters may be written as $y_i=s_i+[-1;1]s_i^{(2)}+as_i^{(1)}$, where *a* – some interval constant.

Thus, the approximation of the numerical solution of the initial system of differential equations serves as an estimate of the center of true solutions, and approximation of numerical solutions of the auxiliary systems for error estimates, i.e. to evaluate the variation of the true solution relative to the center.
5. A constant $a$ value is calibrated at the last step. At this stage to estimate the interval width of the exact solution of the system (1) the nature of the interval functions is taken into account $f_i(t,y)$, caused by interval values of a set of parameters. For this purpose we introduce the interval functions $\frac{\partial f_i}{\partial y_i}(t,\theta) \in f_{y,i,j}(t,[\underline{\theta};\bar{\theta}]) = [\underline{f}_{y,i,j};\bar{f}_{y,i,j}]$ by means of interval extension. The centered differential expansion is used [Shary (2012)]. Centered differential expansion of functions is calculated by the formula $f_{mv}(X) = f(c) + f'(X)(X-c)$, where $X$ for the described equations is the interval for the parameters $[M_Д - \varepsilon_1, M_Д + \varepsilon_1]$. Respectively, the derivative $f'(X)$ is calculated with the same parameters. The interval function is assigned $[\underline{\phi}_i;\bar{\phi}_i] = f_i(t,s) - \frac{ds_i}{dt}$. Here for the evaluation of the function $f_i(t,y)$ the centered differential expansion is used.
By using constants $\delta_i$ the values are defined $r_i = s_i + [-1;1]s_i^{(2)} + [-\delta_i;\delta_i]s_i^{(1)}$. Next the following values are defined

$$\tilde{f}_{y,i,j}(t) = \max(|\underline{f}_{y,i,j}(t,r)|, |\bar{f}_{y,i,j}(t,r)|), \quad \Phi_i(t) = \max(|\underline{\phi}_i(t,s)|, |\bar{\phi}_{ii}(t,s)|) - ds_i^{(2)}(t)/dt + \sum_{j=1}^{m} \tilde{f}_{y,i,j}(t)s_i^{(2)}(t),$$

$$\Psi_i(t) = ds_i^{(1)}(t)/dt - \sum_{j=1}^{m} \tilde{f}_{y,i,j}(t)s_i^{(1)}(t).$$

The value of $a$ is calculated by the formula $a = \max_{i=1,\dots,m}(\Phi_i(t)/\Psi_i(t), 0)$.

The result is the interval solution of the system of differential equations (1). This algorithm includes the following main points: the definition of the numerical solution without the inclusion of interval parameters; error estimation of the numerical solution and its approximation; definition of intervals containing the exact solution, taking into account error estimation of the method and interval extensions of functions, including the interval parameters.
The results of applying this algorithm to the system (1) are shown in Figure 3.
Values obtained by the Runge-Kutta method and upper and lower bounds of the angular velocities for some nodes are presented in Table 1.

**Table 1**
*Values of angular velocities*

| **t, sec** | $\dot{\varphi}_{21}$, sec$^{-1}$ | $\bar{\dot{\varphi}}_{21}$, sec$^{-1}$ | $\underline{\dot{\varphi}}_{21}$, sec$^{-1}$ | $\dot{\varphi}_{22}$, sec$^{-1}$ | $\bar{\dot{\varphi}}_{22}$, sec$^{-1}$ | $\underline{\dot{\varphi}}_{22}$, sec$^{-1}$ |
|---|---|---|---|---|---|---|
| 0,0005 | 418.946 | 418.981 | 418.91 | 2,95969 | 2.96674 | 2.95263 |
| 0,001 | 421.744 | 421.883 | 421.605 | 12.8021 | 12.8291 | 12.7751 |
| 0,0015 | 426.101 | 426.416 | 425.786 | 28.9009 | 28.9602 | 28.8415 |
| 0,002 | 431.569 | 432.137 | 431.001 | 50.2718 | 50.3754 | 50.1682 |
| 0,0025 | 437.626 | 438.533 | 436.719 | 75.708 | 75.8666 | 75.5495 |
| 0,003 | 443.757 | 445.097 | 442.417 | 103.923 | 104.146 | 103.7 |
| 0,0035 | 449.528 | 451.405 | 447.651 | 133.677 | 133.973 | 133.381 |
| 0,004 | 454.625 | 457.156 | 452.094 | 163.869 | 164.243 | 163.494 |
| 0,0045 | 458.873 | 462.187 | 455.558 | 193.592 | 194.05 | 193.134 |

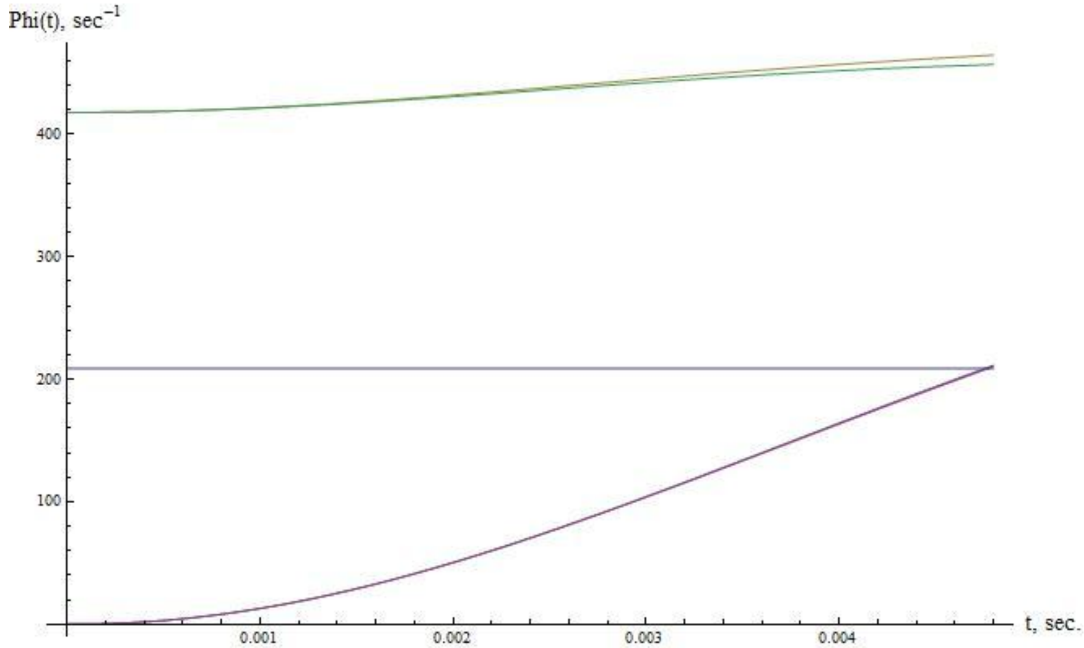



**Fig.3.**
*Graphs of the variation of the angular velocities at the stage of reactor acceleration*

## 3. Conclusion

In the present article the mathematical model of ITT workflow was described. A method for accurate estimation of the solution of differential equations describing the workflow of ITT was also presented. The results show the application of the method to determine the range within which workflow of ITT runs. The knowledge of all possible trajectories of the process may allow designers to be more precise in selecting the optimal parameters for this device.

## Acknowledgements

This research was supported by the Ministry of education and science of the Russian Federation in the bounds of the list of scientific and research work of main part of the government job, project number 970.

## References

Bazhenov, S.P. 2003. *Infinitely variable transmission of traction and transport vehicles*. Lipetsk State Technical University, Lipetsk. Russia. 81 p.

Dobronets, B.S. 2004. *Interval mathematics*. Krasnoyarsk State University, Krasnoyarsk. Russia. 216 p.

Galkin, A.V., Dyatchina, D.V. 2011. The numerical solution of mathematical models of objects specified by component systems of differential equations. *Contemporary Problems of Science and Education,* 6. <http://www.science-education.ru/100-5196>.

Kalmykov, S.A.; Shokin, Yu.I.; Yuldashev, Z.H. 1986. *Methods of interval analysis.*, Nauka Science Publ., Novosibirsk. Russia. 222 p.

Leonov, A.I. 1978. *Inertial automatic torque transformers*. Mashinostroenie Science Publ., Moscow. Russia. 224 p.

Shary, S.P. 2012. *Finite-dimensional interval analysis.* <http://www.nsc.ru/interval>.

Stechkin, S.B.; Subbotin, Yu.N. 1976. *Splines in numerical mathematics*. Nauka Science Publ., Moscow. Russia. 248 p.

# INTERVAL CYCLIC HYPERGRAPHS FOR MODELING OF TRANSPORTATION SYSTEMS

**Semen L. Blyumin[1], Alexander V. Galkin[2], Denis I. Prikhodko[3], Pavel V. Saraev[4], Anton S. Sysoev [5]**
[1,2,3,4,5] *Lipetsk State Technical University, Lipetsk, Russia*

**Abstract:** the paper presents mathematical foundations of matrix description of interval cyclic hypergraphs relevant for the modeling of widespread ring transport routes such as different urban transport routes, taxi fixed-routes, airplane and rail routes etc. Transport routes are spatially distributed systems. Thus grapho-structural modeling is a relevant mathematical tool for the simulation of such systems. The vertex set of interval cyclic hypergraph is defined as a cyclic group. Intervals are defined as consecutive sequences of vertices in cyclic groups i.e. sequences of vertices without gaps in correspondence with transport routes. The matrix description includes incidence, valence, adjacency matrices, Laplacians and relations which interconnect them. Namely Laplacian is a product of incidence matrix by its conjugation; simultaneously it is a sum or difference of valence and adjacency matrices. This relation plays a controlling role for matrix manipulations in grapho-structural modeling. Both indirected and directed (oriented) complete interval cyclic hypergraphs are considered. The orientation serves for the fixation of the order of vertices in intervals. This is essential for ring transport routes. Incidence matrices of complete interval cyclic hypergraphs are circulants in contrast with incidence matrices of general complete hypergraphs. Consequently Laplacians of complete interval cyclic hypergraphs are circulants just as Laplacians of general complete hypegraphs. The circulants can be presented as polynomials from cyclic permutation matrices. The eigenvalues of circulants are efficiently calculated as the same polynomials from primitive unity roots of suitable degree. Circulantness is very useful for mathematical analysis of ring transport routes. Some examples of applications of interval cyclic hypergraphs for modeling of urban ring transport routes are presented.

**Keywords**: interval cyclic hypergraphs, circulants, ring transport routes.

## 1. Introduction

Technical systems consisted a composition of units, which form different configurations, could be referred to a class of distributed systems. An appropriate mathematical basis for describing and analysis such systems is grapho-structural modeling, which wide methods from the graph theory to the theory of hypergraphs, metagraphs, iterated hypergraphs and other hierarchical arrangements.

There are technical systems, which have pipeline ring structure. In the field of the transportation science it could be ring transport routes, such as different urban transport routes, taxi fixed-routes, airplane and rail routes, etc.

The aim of the paper is to construct the basis of matrix description for interval cyclic hypergraphs, which could be used for modeling elementary pipeline ring systems.

## 2. Hypergraph and its characteristics

### 2.1. The definition of hypergraph

The study (Bollobas, 1986) holds in fact (but not in full) the concept of interval cyclic hypergraph. *Hypergraph* HG is defined as a pair (V, HE) where V is a finite set of vertices, $|V|=m$, $HE \subseteq B(V)$ is a set of hyperedges (subsets of the set of vertices), $B(V)=2^V$ is the Boolean for the set V (the set of all subsets of V), $|2^v|=2^m$. The *complete hypergraph* is defined as a pair $(V,2^V)$. *k-uniform hypergraph*, k=0,1,...,m, is defined the fact that all its hyperedges consist k vertices ($HE \subseteq C_V^k$, where $C_V^k$ is a set of all k-elements subsets of the set of vertices, $|C_V^k| = C_m^k$). *Complete k-uniform hypergrap* is defined as a pair $(V, C_V^k)$. At the same time $2^V=\cup_{k=0}^m C_V^k$.

Bolobas (1986) proposes to represent the set of vertices of a hypergraph as a cyclic group of integer numbers with respect to base m, $V \cong \mathbf{Z}_m=\{0,1,\ldots,m-1\}$, and this defines the cyclic order of vertices. An interval of length k is defined as a consistent set of vertices, in other words this is a hyperedge $\{a,a+1,\ldots,b=a+(k-1)\}$ (here + should be understood as the modulo m addition, while (k-1) - in the ordinary arithmetic sense). Thus, an interval does not consist gaps of vertices in their fixed cyclic order, while when $b \geq m$ the end of an interval returns in the beginning of the group $\mathbf{Z}_m$. The set of all intervals of length k is denoted $I_V^k \subseteq C_V^k$, $|I_V^k|=m \leq C_m^k$, $k=1,\ldots,m-1$, $|I_V^0|=|I_V^m|=1$. The *complete k-uniform interval cyclic hypergraph* is defined as a pair $(V, I_V^k)$.

[5]Corresponding author: anton_syssoyev@mail.ru

### 2.2. Matrix characteristics for interval hypergraph

As an illustration it is convenient to use an incidence matrix of a hypergraph, which has rows marked with vertices and columns marked with hyperedges. Until a graph is undirected, this matrix has no sign (it is denoted as $I^{sl}$, comes from signless).

Thus, for the complete 2-uniform hypergraph (in other words for graph) with 5 vertices this matrix is

$$I^{sl}(5,2)=\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix},$$

and for the complete 3-uniform hypergraph with 5 vertices this is

$$I^{sl}(5,3)=\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

Intervals (as mentioned above) are presented in the matrix $I^{sl}(5,2)$ with the 1st, 5th, 8th, 10th and 4th columns and in the matrix $I^{sl}(5,3)$ - with the 1th, 7th, 10th, 6th and 3rd columns. These columns (in that order) form incidence matrices $II^{sl}$ (the symbol comes from “incidence” and “interval”) for complete interval cyclic graph

$$II^{sl}(5,2)=\begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

and for complete 3-uniform interval cyclic hypergraph

$$II^{sl}(5,3)=\begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}.$$

These examples demonstrate the specificity of $II^{sl}(m,k)$ incidence matrices for complete k-uniform interval cyclic hypergraphs with m vertices - these matrices are circulants (in terms of Voevodin (1987)). A transposed circulant, sum and product, so a degree of cicrulants and also a circulant multiplied by a number, keep circulants too. The curculant C is a polynom from a permutation matrix

$$P_m=\begin{bmatrix} 0 & 0 & . & 0 & 0 & 1 \\ 1 & 0 & . & 0 & 0 & 0 \\ 0 & 1 & . & 0 & 0 & 0 \\ . & . & . & . & . & . \\ 0 & 0 & . & 1 & 0 & 0 \\ 0 & 0 & . & 0 & 1 & 0 \end{bmatrix},$$

in other words

$$C = c_0 P_m^0 + c_1 P_m + c_2 P_m^2 + \ldots + c_{m-1} P_m^{m-1},$$

where $P_m^0 = I_m$ – unity matrix of order m.

In examples, given before, $II^{sl}(5,2)=I_5 + P_5$ , $II^{sl}(5,3)=I_5 + P_5 + P_5^2$ . In the general case $II^{sl}(m,k)=I_m + P_m + \ldots + P_m^{k-1}$.

Together with signless matrices $I^{sl}$ it is also possible to describe undirected graphs, using valence matrices D, adjacency matrices A and signless Laplacians $L^{sl}$ (Cvetkovic, 1986). The following expression connect these characteristics

$$L^{sl} = D + A = I^{sl} \cdot (I^{sl})^T .$$

The same more general expressions could be founded out for hypergraphs. It come from papers connected to directed hypergraphs and complete hypergraph by Bluymin (2010). In case of undirected complete k-uniform interval cyclic hypergraps with m vertices signless the Laplacians (as a product of circulants) are circulants too and could be presented as

$$LI^{sl}(m,k)=II^{sl}(m,k)\cdot II^{sl}(m,k)^T =(I_m + P_m +\ldots+ P_m^{k-1})\cdot( I_m + P_m +\ldots+ P_m^{k-1})^T =$$

$$=( I_m + P_m + \ldots + P_m^{k-1})\cdot( I_m + P_m^{m-1} + \ldots + P_m ) =$$

$$= p_0 P_m^0 + p_1 P_m + p_2 P_m^2 + \ldots + p_{m-1} P_m^{m-1},$$

where $p_i$, i=0,1,…,m are integer numbers (matrices $P_m^0$ , $P_m$ , $P_m^2$ ,…, $P_m^{m-1}$ form cyclic multiplicative group, $(P_m^i)^T=(P_m^i)^{-1}=P_m^{-i} =P_m^{m-i}$). Thus,

$$LI^{sl}(5,2)=( I_5 + P_5 )( I_5 + P_5^4 )=2I_5+P_5+P_5^4,$$

$$LI^{sl}(5,3)=( I_5 + P_5 + P_5^2 )( I_5 + P_5^4 + P_5^3 )=3I_5+2P_5+P_5^2+P_5^3+2P_5^4=$$

$$=\begin{bmatrix} 3 & 2 & 1 & 1 & 2 \\ 2 & 3 & 2 & 1 & 1 \\ 1 & 2 & 3 & 2 & 1 \\ 1 & 1 & 2 & 3 & 2 \\ 2 & 1 & 1 & 2 & 3 \end{bmatrix},$$

and, for example,

$$LI^{sl}(6,4)=4I_6+3P_6+2P_6^2+2P_6^3+2P_6^4+3P_6^5.$$

One of the lack of the presentation of interval hypergraphs with signless incidence matrices is the fact that all vertices, included in a some interval, are offered with the only number 1. This fact does not allow to show the order of vertices within the interval (and it is connected with the orientation of hypergraph) which can be a very essential for some applications.

In case of a graph this problem is solved in the following way: when orientating the edges (transforming undirected graph into directed) the sign "-" shows that the edge leaves a vertex and the sign "+" shows that the edge enters a vertex. Thus, this matrix becomes signless and the connections between incidence, valence, adjacency matrices and Laplacians (further incidence matrices and Laplacians have sign) could be presented with the following standard expression

$$L = D – A = I \cdot I^T .$$

So, now we have

$$I(5,2)=\begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & -1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 & -1 & -1 \end{bmatrix}$$

and, correspondingly,

$$\mathrm{II}(5,2)=\begin{bmatrix} 1 & 0 & 0 & 1 & -1 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \end{bmatrix}=I_5 - P_5 \,,$$

$$\mathrm{LI}(5,2)=2I_5 - P_5 - P_5^4.$$

In case of a hypergraph for the orientation its hyperedges (transforming them into directed hyperedges and a graph into a directed hypergraph) it is proposed (Blyumin, 2010) to use (as a generalisation of squared roots of a unit ±1 in case of a graph) complex roots of degree k of a unit for the presentation of vertices for a hyperedge of size k

$$\varepsilon_k^j = \exp\left(\frac{2\pi i}{k} j\right), \; j=0,1,\ldots,k-1.$$

Now we have

$$\mathrm{I}(5,3)=\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ \varepsilon & \varepsilon & \varepsilon & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ \bar\varepsilon & 0 & 0 & \varepsilon & \varepsilon & 0 & \varepsilon & \varepsilon & 0 & 1 \\ 0 & \bar\varepsilon & 0 & \bar\varepsilon & 0 & \varepsilon & \bar\varepsilon & 0 & \varepsilon & \varepsilon \\ 0 & 0 & \bar\varepsilon & 0 & \bar\varepsilon & \bar\varepsilon & 0 & \bar\varepsilon & \bar\varepsilon & \bar\varepsilon \end{bmatrix},$$

where $\varepsilon = \varepsilon_3 = \exp\left(\frac{2\pi i}{3}\right) = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$, $\bar\varepsilon = \varepsilon_3^2 = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$, and, correspondingly,

$$\mathrm{II}(5,3)=\begin{bmatrix} 1 & 0 & 0 & \bar\varepsilon & \varepsilon \\ \varepsilon & 1 & 0 & 0 & \bar\varepsilon \\ \bar\varepsilon & \varepsilon & 1 & 0 & 0 \\ 0 & \bar\varepsilon & \varepsilon & 1 & 0 \\ 0 & 0 & \bar\varepsilon & \varepsilon & 1 \end{bmatrix}=I_5 + \varepsilon P_5 + \varepsilon^2 P_5^2$$

(it should be noted, that there are cyclic permutations of the vertices for two last intervals).

In the general case

$$\mathrm{II}(m,k)=I_m + \varepsilon_k P_m + \ldots + \varepsilon_k^{k-1} P_m^{k-1}.$$

These matrices and, as consequence, respective Laplacians (as described above) are circulants, and now the incidence matrix is multiplied not by transposed matrix but (because of using complex numbers) by conjugate to it

$$\mathrm{LI}(m,k)=\mathrm{II}(m,k)\cdot \mathrm{II}(m,k)^* =$$

$$=( I_m + \varepsilon_k P_m + \ldots + \varepsilon_k^{k-1} P_m^{k-1})( I_m + \varepsilon_k P_m + \ldots + \varepsilon_k^{k-1} P_m^{k-1})^*=$$

$$=( I_m + \varepsilon_k P_m + \ldots + \varepsilon_k^{k-1} P_m^{k-1})( I_m + \varepsilon_k^{k-1} P_m + \ldots + \varepsilon_k P_m^{k-1})=$$

$$= p_0 (\varepsilon_k )P_m^0 + p_1 (\varepsilon_k )P_m + p_2 (\varepsilon_k )P_m^2 + \ldots + p_{m-1} (\varepsilon_k )P_m^{m-1},$$

where coefficients $p_j(\varepsilon_k)$, j=0,1,…,k-1 are dependent on $\varepsilon_k$.

So, now we have

$$\mathrm{LI}(5,3)=( I_5 +\varepsilon P_5 +\varepsilon^2 P_5^2 )( I_5 + \varepsilon^2 P_5^4 + \varepsilon P_5^3 )=$$
$$=3I_5+2\varepsilon P_5+\varepsilon^2 P_5^2+\varepsilon P_5^3+2\varepsilon^2 P_5^4=$$

$$=\begin{bmatrix} 3 & 2\bar{\varepsilon} & \varepsilon & \bar{\varepsilon} & 2\varepsilon \\ 2\varepsilon & 3 & 2\bar{\varepsilon} & \varepsilon & \bar{\varepsilon} \\ \bar{\varepsilon} & 2\varepsilon & 3 & 2\bar{\varepsilon} & \varepsilon \\ \varepsilon & \bar{\varepsilon} & 2\varepsilon & 3 & 2\bar{\varepsilon} \\ 2\bar{\varepsilon} & \varepsilon & \bar{\varepsilon} & 2\varepsilon & 3 \end{bmatrix},$$

and, for example,

$$LI(4,3)=( I_4 +\varepsilon P_4 +\varepsilon^2 P_4^2 )( I_4 + \varepsilon^2 P_4^3 + \varepsilon P_4^2 )=3I_4+2\varepsilon P_4 -P_4^2+2\varepsilon^2 P_4^3.$$

Applications of hypergraphs described above to the problems of multi-agent systems are presented in papers by Blyumin (2010, 2011, 2012). Blyumin (2012) investigates possibilities of more flexible E-ean (the opposite to Boolean one) approach when the Boolean $2^V$ is replaced by the E-ean $E^V$, where E is the assessed set which is more general than the Boolean {0,1}. Thus, the choice as E the interval [0,1] leads to fuzzy hypergraphs. Using interval cyclic hypergraps in this case is an open problem which is need to be solved.

### 3. Examples of practical using hypergraphs in modeling transportation systems

To demonstrate the concept of usage hypergraphs in modeling of transportation systems let us have a look on the following two examples.

#### 3.1. Hypergraphs in modeling urban traffic regulated intersection

This example shows the possibility of the usage a hypergraph for a description (and its incidence matrix for modeling) of a variable structure system. In the context of transportation science such systems could be used for modeling of an urban regulated traffic intersection.

Sysoev (2014) presents the following situation. Let us we have a regulated intersection, presented on the fig. 1. The fig. 1 also shows the hypergraph HG associated with this intersection: the directions for vehicles could be understood as the set of hypervertices $V=\{V_2,V_3,V_4,V_5,V_6,V_7,V_8,V_9,V_{12}\}$ and phases of regulation could be presented as a set of hyperedges $HE=\{\{V_2,V_3,V_8,V_9\},\{V_7,V_{12}\},\{V_6,V_9,V_{12}\},\{V_3,V_4,V_5\}\}$ .

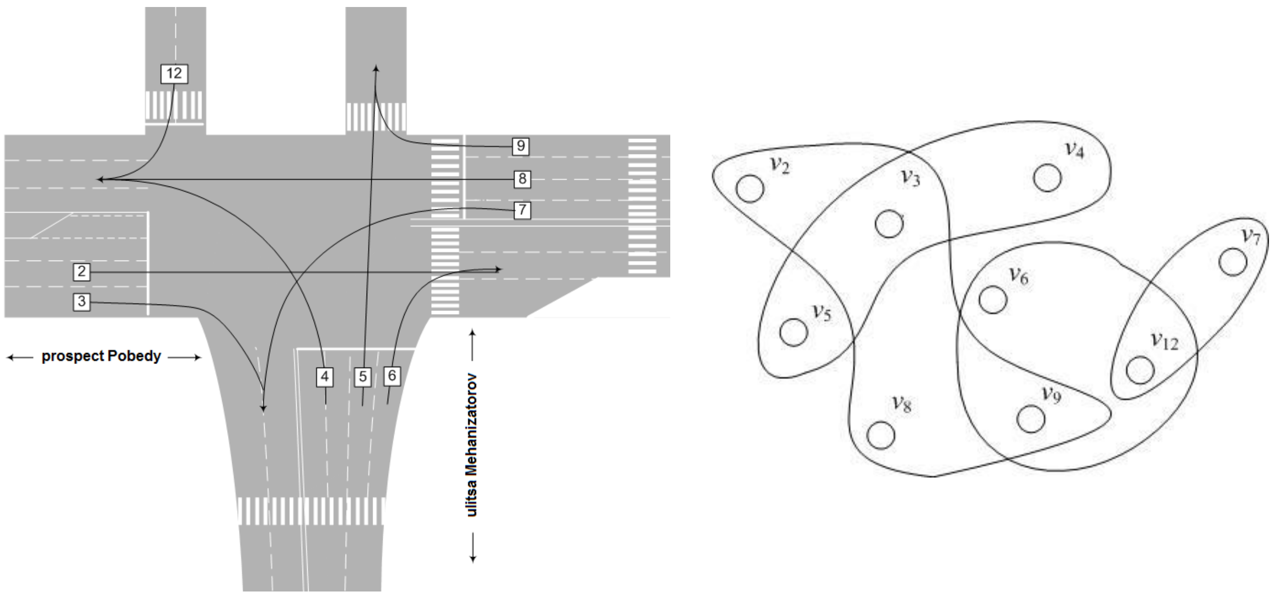



**Fig. 1.**
*Regulated intersection and its hypergraph*
*Source: Sysoev, A.S. PhD thesis, 2014.*

The incidence matrix of the hypergraph HG in this example is

$$C = \begin{bmatrix} 1 & 1 & \xi & \xi & \xi & \xi & 1 & 1 & \xi \\ \xi & \xi & \xi & \xi & \xi & 1 & \xi & \xi & 1 \\ \xi & \xi & \xi & \xi & 1 & \xi & \xi & 1 & 1 \\ \xi & 1 & 1 & 1 & \xi & \xi & \xi & \xi & \xi \end{bmatrix},$$

where $\xi$ is a neutral element in terms of idempotent (max-plus) algebra.

Thus, also in terms of max-plus algebra, the following system could be used for modeling of an urban regulated traffic intersection

$$D(k) = A(k) \oplus W(k) \oplus (D^T(k-1) \otimes \Xi(k))^T,$$

where D(k) is a vector of departure times for vehicles with the number k from all approaches to the intersection, A(k) is s vector of arrival times for vehicles with the number k from all approaches to the intersection, $\Xi(k)$ is a vector of random times for vehicles with the number k from all approaches to the intersection and a vector of nonrandom waiting times W(k) contains the incidence matrix C in the next way: its elements $w_l(k) = \bigotimes_{i=cur}^{c_{ij}\neq\xi} \bar{c}_{ij}^{g_i}$, $l, j = \{2,3,4,5,6,7,8,9,12\}$ (here $c_{ij}$ is an element of the incidence matrix C, cur is a counter of phases from the beginning of modeling, $g_i$ are durations of phases and $\oplus$ and $\otimes$ are max and plus operators correspondingly).

### 3.2. Hypergraphs in modeling urban ring transport routes

Cyclic interval hypergpahs (in the general case) also could be used for modeling of different complex cyclic ring transportation systems. For example, these could be ring route highways with local towns and villages nearby, which in this case are control points (hypervertices); route lines of municipal transport (in this case the proposed approach could consider gaps of stops, which is essential in a case of road or equipment fixing). Furthermore, in the general case every transport route could be presented as a ring, in which for the last stop point follows the first again. And intervals mentioned before are nothing but distances between them. The fig. 2 demonstrates the "Kolcevaya" line of Moscow underground system as an example of such kind of urban route traffic system.

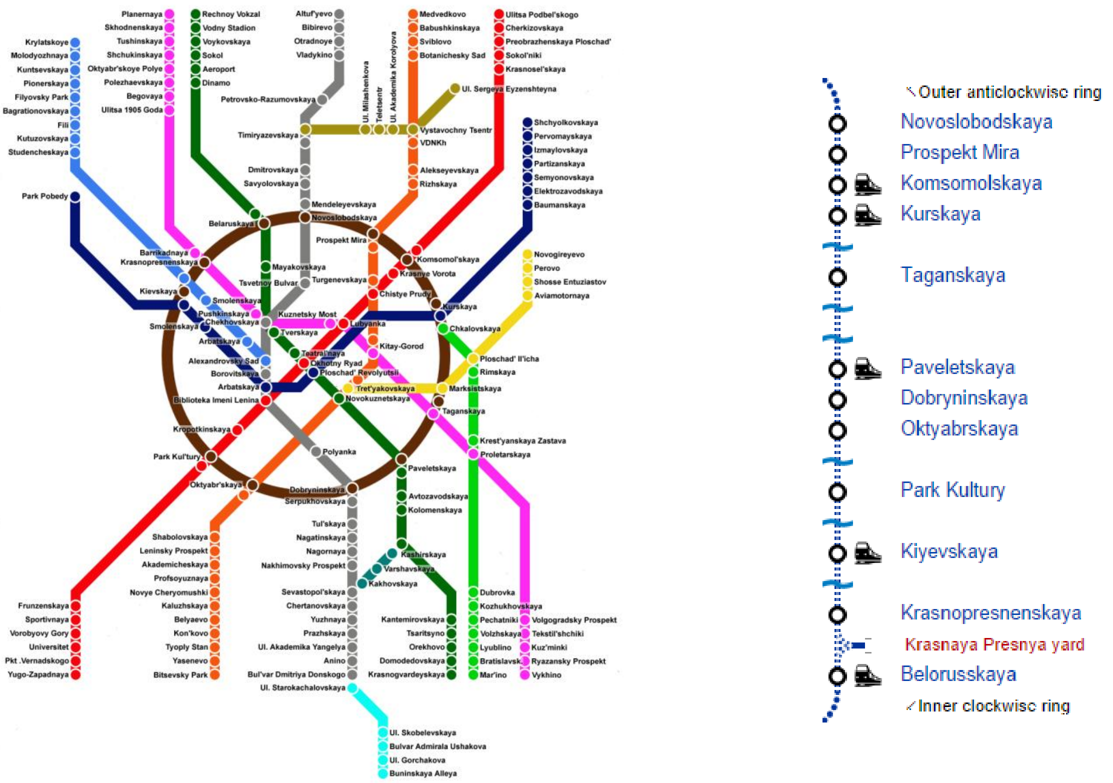



**Fig. 2.**
*Moscow Underground System and Kolcevaya Line*
*Source: en.wikipwdia.org, "Moscow Metro", 2014.*

The main advantage of the proposed approach is the usage of special structure matrix characteristics of graphs. This point makes it easy to process and to keep the information in the computer memory.

**Acknowledgements**

This research was supported by the Ministry of education and science of the Russian Federation in the bounds of the list of scientific and research work of main part of the government job, project number 970.

**References**

Bollobas, B. 1986. *Combinatorics. Set Systems, Hypergraphs, Families of Vectors and Combinatorial Probability*. Cambridge: CUP. 187 p.

Blyumin, S.L. 2010. Directed hypergraphs: matrix characteristics, *Large-scale systems control*. 30.1 Network models in control science: 22-39.

Blyumin, S.L. 2010. Complete hypergraphs. Spectrums of Laplacians. Multi-agent systems, *Large-scale systems control*. 30: 5-23.

Blyumin, S.L. 2011. Multi-agent systems: group interactions, weighted directed, *Control systems and information technology*. 2(44): 49-53.

Blyumin, S.L. 2012. E-nary hypergraphs, Laplacians and multi-agent systems, *Control systems and information technology.* 3(49): 6-10.

Cvetkovic, D.; Simic, S. 2010. Towards a Spectral Theory of Graphs Based on the Signless Laplacian, *Appl. Anal. Discrete Math*. 4: 156-166.

Sysoev, A.S. 2014. Modeling and optimization of variable structure systems using idempotent mathematics and analysis of finite fluctuations (PhD thesis), Lipetsk State Technical University, Lipetsk. Russia.

Voevodin, V.V.; Tyrtyshnikov, E.E. 1987. Computation processes using Toplitz matrixes. Moskva: Nauka. 230 p.

# HIGHWAY TRUCK PARKING PREDICTION SYSTEM AND STATISTICAL MODELING UNDERLYING ITS DEVELOPMENT

**Marek Brabec[1,2] , Ondrej Konar[1], Ivan Kasanicky[1], Emil Pelikan[1,2], Marek Maly[1]**
[1] *Department of Nonlinear Modeling, Institute of Computer Science, Academy of Sciences of the Czech Republic, 182 08, Praha 8, Pod Vodaraenskou vezi 2, Czech Republic*
[2] *Department of Transport Telematics, Faculty of Transportation, Czech Technical University, 110 00, Praha 1, Konviktska 20, Czech Republic*

**Abstract:** In this paper, we will describe a system for on-line prediction of truck parking demand along highway system in the Czech Republic. We will describe structure of the system developed during the TACR TA02031411 project and mention some of its specific functionalities. Further, we will explain in detail statistical modeling methodology which underlies the forecasting model in the core of the prediction procedure. Whole system relies on the use of indirect but very precise and relatively cheap to obtain toll transaction data (accessible through a cooperation with Kapsch Telematic Services, Inc.). Our statistical modeling starts with a recognition of the fact that the number of trucks parking at a given lot and given time is a latent variable to be estimated from the observable toll transaction data (which are available in the form of times when individual truck pass toll gates). After constructing appropriate proxy variable, we formulate a flexible class of statistical semi-parametric models constructed in a Markovian fashion. In fact, our model can be viewed as a non-homogeneous Markov chain, whose Poissonian transition probabilities change with several external covariates (describing e.g. weekly and daily periodicity of parking intensities) as well as spatially. Once the model is estimated (its parametric and nonparametric parts are estimated simultaneously), it is used for real time prediction for several short to medium horizons, using Monte Carlo simulations to obtain efficient and robust software implementation. We will demonstrate practical performance of the prediction system under routine conditions, based on evaluation against manual parking lot counting.

**Keywords**: Highway Truck parking, Prediction System, Dynamical Statistical Modeling, Generalized Additive Model

## 1. Introduction

For longer truck trips, it is necessary for the driver to have parking places of sufficient capacity and suitable distribution along the route in order to adhere to pause schedules required by law and to keep obvious safety precautions. This problem is obviously accentuated on highways – both due to the intense truck traffic and necessarily limited parking lot capacity allotted to trucks along a given motorway. Extension of the present parking places is costly and can occur only in longer time horizon. This, in turn stresses the need for effective use of the existing capacities.

One possible strategy in this context is to provide truck drivers with real-time predictions of parking lots availability ahead of them. Naturally, only the parking lots situated along the same highway, downstream in the direction of the truck movement and lying in a reachable distance are considered. Such information enables the truck drivers to plan their pauses effectively. In order to achieve this, the predictions should be available with a multitude of prediction horizons to cover different speeds and different decision priorities of individual drivers. With the provided information drivers can make more informed decisions about whether to park immediately at a current parking lot, or whether there is a good chance to find a place at a parking lot(s) in a reasonable distance ahead. Not only that such a prediction system can enhance drivers' convenience and safety, it also contributes to a more uniform utilization of existing parking capacities and hence to alleviate, to some extent, the need for building new costly structures.

Obviously, building the prediction system and delivering its outputs to the end-users (truck drivers) is a complex task and consists of several related, but very different steps:

1) real time data acquisition for characterizing current parking situation (plus their formalized pre-processing, if necessary)
2) development of a statistical model providing the forecasts
3) implementation of the resulting forecasting algorithm into a fast, online running software
4) delivering the forecasts to the end-users (drivers) in an easy-to-grasp form (e.g. via mobile application with red/yellow/green colors denoting full/almost full/almost empty parking lot statuses based on suitable discretization of the predicted number of parking trucks)

Such systems are being developed in several countries on different principles in order to accommodate to locally-specific conditions and regulations and infrastructures (Austrian Unterweg and Netherland's ParckR can serve as successful examples). Authors of this paper are members of a team (consortium of research institution, university, data provider and commercial company) which develops a pilot version of online parking prediction system for the Czech Republic within a larger project sponsored by the Czech Technological Agency. Different specialists there cover all items 1), 2), 3), 4) in order to deliver running product capable of online deployment.

[1]Corresponding author: Marek Brabec, mbrabec@cs.cas.cz

In this paper, we will deal with the part 2) which constitutes the core of the prediction system, determining its quality. Nevertheless, since in the Czech Republic (as in many other countries), there are not direct measurements of individual parking lots availability in an on-line regime, we have to deal also with statistical part of 1). Namely, we have to derive the free capacity of parking lots from external data in the form of a proxy variable upon which the predictions from 2) and 3) are based. In situations where direct occupancy data are available (e.g. when parking lots are equipped with counting devices delivering the data fast enough to a central database), one would start directly with statistical modeling right in 2.

## 2. Statistical development of the truck parking prediction system

In order to predict the number of trucks parking on a given parking lot (or equivalently, for a fixed parking capacity, the number of free parking places), one needs to have information about the current and possibly also past occupancies. If a direct measurements of the occupancies are available on-line and with sufficient precision (e.g. through automatic processing of video surveillance images, via underground loops or other techniques), one can built the forecasting system on a statistical model of the observable counts.

On the other hand, if no such direct measurement of parking lot occupancies exists, but suitable indirect data are available, one can built a proxy variable for the unobservable (latent) counts from the available (observable) measurements. Then the prediction model can be built upon the proxy. In principle, such a framework corresponds to a two level, state-space formulation, Harvey (1989), Durbin and Koopman (2001).

The situation just descried is common in various countries. In the Czech Republic, for instance, there are no long-term and spatially distributed on-line measurements of highway parking lot occupancies. But there are very accurate, cheaply obtainable, indirect massive on-line data from the highway truck toll collection system on all Czech highways. They are in the form of times when individual trucks passed individual toll transaction gates along the highway. Obviously, for non-billing purposes, the data have to be anonymized, but that does not matter as long as one is able to work with individual sequences of times of gate passing moments. The toll gates are located essentially between every two highway inlets/outlets so that individual use of every road segment can be billed separately. Along the D5 motorway, where we conduct a pilot study, the average distance between the gates is about 8 km which is providing rather accurate information from which a proxy for the parking count can be built.

The full prediction procedure then consists of two steps:

i) construction of a proxy
ii) building and implementing a statistical model for producing the actual forecasts, based on the proxy from the previous step

In the next section, we describe the first step, i.e. the construction of a proxy from available toll transaction data.

### 2.1. Construction of a proxy for (latent) parking occupancy from observable toll transaction data

From the toll transaction system we can get times $T_{i,1}, T_{i2}, \ldots, T_{i,n_i}$ when individual truck passes different gates $G_{i1}, G_{i2}, \ldots, G_{in_i}$, for an *i*-th individual truck. The times are recorded with very fine precision (seconds). Note that these individual time series are highly irregular – in general, different trucks can visit different number of gates located at different places along the highway. The gate times are transferred to the central database with a negligible time delay (unlike, e.g. in the German satellite-based toll-collection system).

It is clear that these data contain substantial information about whether a particular truck is parking on a given lot or not, as a long passage time between two consecutive gates containing parking lot between them suggests that the truck has been parking with high probability. It is, however, not entirely trivial to construct a proxy in a formalized way. Fig. 1 illustrates that while it is clear that the left part of the apparent segment average velocity histogram is relevant for the parking proxy classification, it is not a priori obvious even where to place the cutoff.

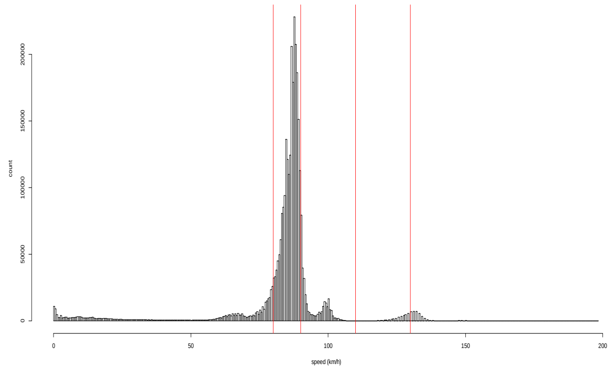



**Fig. 1.**
*Histogram of apparent segment average velocities*

After experimenting with several alternatives and comparing their results with respect to various criteria used for assessing classification quality, like total misclassification, false positive, false negative rate, ROC (receiver operating characteristic) Hastie et al. (2009), we reached the following proxy showing good results on verification data. The verification was done with manual parking truck counting of selected lots throughout an interval of several days in the role of the “gold standard”. In addition, an approximation to the false positive rate can be obtained also from inter-gate segments not containing any parking lot (as long as the road characteristics of the parking-lot-containing and parking-lot-free segments are sufficiently similar).

Individual truck is classified as parking in a given parking-lot-containing segment when satisfying the following three criteria simultaneously:

1) apparent segment-average velocity for the segment is smaller than for a previous segment (parking lots are not located in consecutive segments)
2) when considering typical moving-truck speed of 85 km/h and the time elapsed between passing entry and exit gate of the segment, the time left for (potential) parking is larger than 20 minutes (in order to exclude very short stops that often do not contribute to a real parking lot load)
3) parking time is smaller than 3 days (in order to exclude data errors, inconsistencies, etc.)

In real operating system the proxy is needed at a time before a given truck passed exit gate, especially in case of actually parking trucks. Consequently, there is a need for working with “incomplete” segment timing which brings additional complications along the lines of right censored data, Meeker, Escobar (1998).

Once we have proxy classification for individual trucks – say in the form $C_{g,i,t}$ (assuming 1 if the *i*-th truck was classified as parking on parking *g* at time *t* and 0 otherwise), we can easily obtain the proxy $N_{g,t}$ for truck parking occupancy at a given parking and time as $N_{g,t} = \sum_i C_{g,i,t}$. Time for $N_{g,t}$ computations has the minute-resolution. As an alternative, we also derived expected number of parking trucks for each time using a few basic probability theory tools. Nevertheless, for the practical predictions, we stick to the proxy described above because of its simplicity and slightly better performance under the test we did so far. Fig. 2 shows a good agreement between manually counted parking truck numbers and proxy presented above.

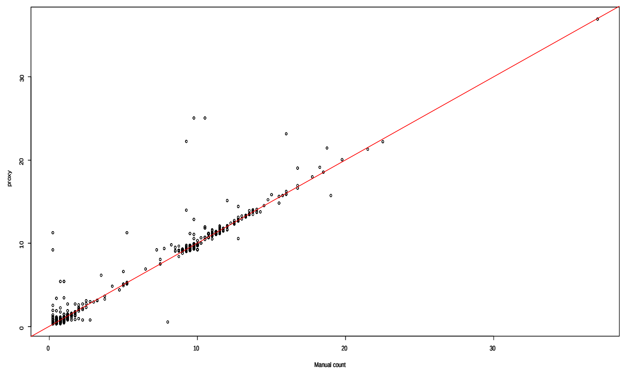



**Fig. 2.**
*Comparison of the proxy and manual count for a control campaign run on one of the parking lots for about 3 days.*

$N_{g,t}$ is the variable that will be used in the next step for statistical modeling of the next section. Resulting dynamical model can then be used to get predictions of the parking lot occupancies for various prediction horizons $h > 0$, say $\hat{N}_{g,t+h}$. As long as the parking lot capacity is known and fixed (not changing over time), this is enough to compute predictions of the number of parking places still available.

### 2.2. Statistical model for prediction of future occupancy

Once we have a proxy (or a direct occupancy measurement), we need to build a (dynamical) statistical model which will serve as a formalized basis for practical implementation of the prediction system. In order to be able to run the forecasting-related computations fast enough, one needs a highly parsimonious model utilizing immediate past of the occupancy process in an effective way. To this end, we use a Markovian approach, Resnick (1992), Cox, Miller (1965).

In order to take into consideration both long-term (or typical) and local (in time) behavior of the occupancy process, we build our statistical prediction model with the two major components utilizing fundamentally different information sources, whose weight is set automatically during the identification of the model (i.e. estimation of its parameters and non-parametric parts):

i) long-term average shape of the occupancy proxy trajectory over a week

ii) immediate occupancy proxy history revealing possible local (in time) departures from the typical behavior

It is our philosophy that a statistical model should come to the data and not vice versa. This means that the model should respect salient features of the reality rather than that a reality is squeezed into a technically convenient model. Therefore, we work with a flexible, semiparametric class of regression models: generalized additive models (GAM), Wood (2006), Hastie, Tibshirani (1990), McCullagh, Nelder (1989). The main reason is that models of this class can easily accommodate not only the simultaneous contribution of the two information sources listed just above (in the style analogous to multiple regression), but they can also take into account important features of the occupancy distribution. For instance, the (relatively low) truck numbers on a given parking lot are necessarily non-negative integers, with relatively large variability, behaving roughly like Poisson random variables. For such a data, ad hoc approximations based on linear regression are very inappropriate.

Therefore, we use a GAM Poisson model with the canonical log link. Namely, our model for a given parking lot (for notational simplicity, we skip the index *g*) is:

$$
\begin{aligned}
&N_t \mid N_{t-1} \sim Poi(\lambda_t) \\
&\log(\lambda_t) = f_t + \beta \log(\tilde{N}_{t-1}) \\
&\tilde{N}_t = \max(N_t, c) \text{ for a given } c > 0
\end{aligned}
\tag{1}
$$

We can easily see that the model (1) indeed specifies a Markovian type of model of first order (conditionally on the current *N*, the past does not matter). In particular, we can also recognize (1) as an inhomogeneous Markov chain, whose first-order transition probabilities depend both on the position within a week and on the occupancy from the immediately preceding time.

Here, $f_t$ is the long-term, or typical behavior part. It is modeled as a cyclic cubic regression spline, Eubank (1988), Wood (2006). It is periodic in order to reflect typical periodic behavior within a week. It depends on actual *t* only through the relative time within a week given in minutes since Sunday's midnight. It is important that $f_t$ is *not* given as a sum of daily time and day of week – different days of the week have very much different parking profiles. This amounts to what would be called an interaction between day of week and time within day effects. Nevertheless, the interaction is not saturated (like in the standard ANOVA models), but it is modeled in a restricted, parsimonious way which leads to much more effective estimation and more precise predictions.

On the other hand, the $\beta \log(\tilde{N}_{t-1})$ term (with unknown coefficient $\beta$ to be estimated from data) models the influence of the immediate past and its departures from the long-term-typical behavior. In order to account for (very rare, but possible) situations with totally empty parking lot, we have to switch from $N_{t-1}$ to $\tilde{N}_{t-1}$, under the presence of the logarithm in the predictor. Concrete choice of the constant *c* did not matter in practical computations.

As an alternative to (1), we also tested a negative-binomial-distribution-based-model. This is based on the well-known negative binomial generalization of the Poisson, Venables, Ripley (2002). Such a more complicated (over-dispersed) model did not lead to an improvement in the prediction ability.

Further, we experimented with the Zeger, Quaqish (1988) formulation in order to improve separation of the two parts of (1), but without substantial improvement. Since (1) is a standard GLM whose estimation is computationally *much* easier than the non-GLM Zeger, Quaqish (1988) formulation, we stick to the model (1), with obvious implementational advantages. Specifically, it is easy to estimate all model parameters in a standard statistical software, e.g. within the R environment, R core team (2014).

Model (1) has unknown quantities ($\beta$ and coefficients implicitly contained in $f_t$). In order to fully identify the model (1), they have to be estimated from available data. The approach that we used and tested successfully in our implementation relies on off-line estimation of the model parameters from past (training) data. It is vital that they are of sufficient size (e.g. from a full calendar year).

Once the model is identified, it is used for practical forecasting computations without continuous re-estimation. This approach brings a lot of speed and computational efficiency since the estimation procedure (based on the very effective iteratively reweighted least squares, IRLS, Wood, 2006) is much more demanding than the practical forecasting computations. On the other hand, the prediction quality does not suffer from batch-like parameter re-estimation (e.g. once-a-year) since we observed that their changes over time are generally very small and slow.

On the other hand, since there are substantial spatial differences in the parking pattern (both in terms of long-term within-week behavior $f_t$ and in the short-time memory part measured by $\beta$), it is very important to estimate (identify) model (1) for different parking lots separately. Trying to impose a common shape of $f_t$ and/or a common $\beta$ for different parking lots would lead to disastrous results. This is because not only size of different parking lots differs substantially (by approximately an order of magnitude), but also dynamics of their utilization is very much different, depending on various local characteristics (e.g. distance to a larger city, proximity of major logistic centers or topology and quality of local road network). We have found that even parking lots lying on the opposite sides of the same highway kilometer can have very much different $f_t$ shapes. This is connected with the well-known within-week periodicity of truck traffic direction in the Czech Republic (west-prevailing traffic at the beginning of a workweek and east-prevailing traffic before the weekend). For these and other reasons, we effectively stratify on individual parking lots (imposing no a priori similarity among the parameters of different lots) when identifying the statistical model to be used for forecasting.

Now, once the model is identified, the forecasting computations are very straightforward in principle, due to its Markovian character. In fact, (1) gives us one-step transitional (time inhomogeneous) probability in a countable state Markov chain. Numerical computations can be done via Chapman-Kolmogorov equations, Resnick (1992), Cox, Miller (1965), Norris (1998), deriving the *h*-step-ahead distribution for the horizon of *h* time units ahead from the appropriate (time-varying 1-step-ahead) transitional probabilities and then computing the expected value for each desired time *t* and *h* combination. In practical online implementation, the time *t* is running in 5 minute intervals and horizons up to 2.5 hours ahead are required. This means that $N_t$ is known and $\hat{N}_{t+h}$ for $h$ =1, 2, …, 30 are needed. In practical terms, the numerical precision and speed deteriorate quickly to render the above mentioned "definitional" style of computations impractical. Instead, we utilize a Monte Carlo approach, Robert, Casella (2004). Specifically, in order to compute predictions for all 30 horizons and a given *t*, we generate a large number *K* (say 1000) of independent Markov chains of length 30 starting from the value of proxy at *t* as of the *0*-th state. These are averaged over *K* for each time-ahead (*h*) to get the ensemble average as the predictor for each of the 30 horizons. Such computations are easy and straightforward, requiring only Poisson random variable generator and a few appropriate aggregations. This can be easily done on-line e.g. by calling a small R (R core team, 2014) script, exporting only the forecasts for each horizon separately.

Fig. 3 shows performance of the prediction procedure for a given parking lot over about three days, summarized by the plot of RMSE (root mean square error in the units of vehicles) criterion against the horizon, *h*. We can see easily that the performance is not bad, considering the nominal capacity of this particular parking lot of about 85 trucks. Moreover, the criterion is computed by comparing proxy count predictions to the manual counts of paring trucks as to a "gold standard". The fact that the manual count is not entirely perfect leads to a slight RMSE inflation. Generally, the prediction quality is worse for very small parking lots. RMSE deteriorates with increasing horizon *h* in a concave way, as expected. Inspecting the MSE decomposition into bias and variance shows that absolute majority of the RMSE comes from the random variability. The bias is very small. In the case illustrated, it is less than 0.5 vehicle.

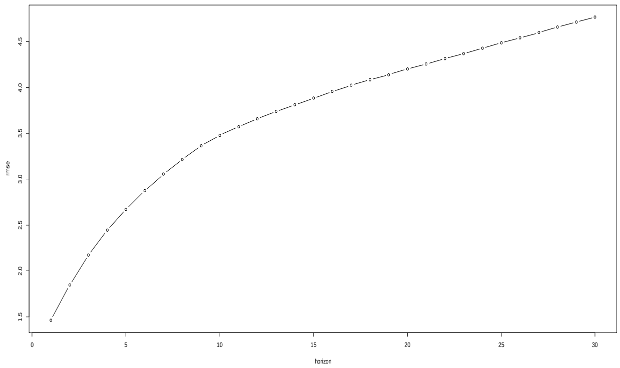



**Fig. 3.**
*Comparison of the proxy and manual count for a control campaign run on one of the parking lots for about 3 days.*

It is obviously not necessary to restrict attention to the first-order Markov models when formulating the statistical model to underlie the prediction system implementation. For instance, second-order model:

$$
\begin{aligned}
&N_t \mid N_{t-1} \sim Poi(\lambda_t) \\
&\log(\lambda_t) = f_t + \beta_1 \log(\tilde{N}_{t-1}) + \beta_2 \log(\tilde{N}_{t-2}) \\
&\tilde{N}_t = \max(N_t, a) \text{ for a given } a > 0
\end{aligned}
\tag{2}
$$

might be expected to improve the prediction performance. The likelihood ratio test, Schervish (1995) shows (2) as a significant improvement over (1). Interestingly enough, this does not matter too much with respect to the practical performance! This is because the improvement is very small and is restricted only to $h$=1 (too short prediction horizon of 5 minutes to be interesting practically). For longer horizons, the RMSE of (2) is uniformly worse than that of (1). Upon a little bit of reflection, this interesting morale is not entirely unexpected. It arises because the current short-term dynamics (capturing the subtle interplay of $t$-1 and $t$-2 lags) does not propagate too far to the future. More complex and hence less effectively estimated model (2) does not win for longer horizons of more practical interest.

## 3. Conclusion

In this paper, we described a statistical framework for prediction of the truck parking along a highway. A fully formalized model suitable for on-line prediction is presented and its structure is discussed. For the case that direct measurements of parking trucks are not available (or that they are not available in an on-line fashion), we describe how to construct a proxy variable from observable measurements, e.g. those obtainable from routine toll transaction system. We show how to compute the prediction practically, using Markov property of our model and simple Monte Carlo simulations. Performances of the predictions are illustrated on comparisons against manual counting of selected parking lots.

## Acknowledgements

The work on this project was supported by the Czech Technological Agency grant TA02031411 (Increasing the usage of parking capacity on highways using prediction models).

**References**

Cox D.R.; Miller H.D. 1965. *The theory of stochastic processes*. Chapman and Hall. London.

Durbin, J.; Koopman, S.J. 2001. *Time series analysis by state space methods*. Oxford University Press. Oxford.

Eubank, R.L. 1988. *Spline smoothing and nonparametric regression*. Marcel Dekker. New York.

Hastie, T.; Tibshirani, R.; Friedman J.H. 2009. *The elements of statistical learning: data mining, inference, and prediction*. Springer. New York.

Hastie, T.; Tibshirani, R. 1990. *Generalized additive models*. Chapman and Hall. London.

Harvey, A.C. 1989. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press. Cambridge.

McCullagh, P.; Nelder, J.A. 1989. *Generalized linear models*. Chapman and Hall. London.

Meeker, W.Q.; Escobar, L.A. 1998. *Statistical methods for reliability data*. Wiley. New York.

Norris, J.R.. 1998 *Markov chains*. Cambridge University Press. Cambridge.

R Core Team 2014. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, available from Internet: <http://www.R-project.org>

Resnick, S.I. 1992. *Adventures in stochastic processes*. Birkhauser. Boston.

Robert, C.P; Casella, G. 2004. *Monte Carlo statistical methods*. Springer. New York.

Schervish M.J. 1995. *Theory of statistics*. Springer. New York.

Venables, W.N.; Ripley, B.D. 2002. *Modern applied statistics with S*. Springer. New York.

Wood, S.N. 2006. *Generalized additive models. An introduction with R*. Chapman and Hall. London.

Zeger, S.L.; Quaqish, B. 1988. Markov regression models for time series: a quasi-likelihood approach. *Biometrics* 44, 1019-1031.

# IMPROVED GRAVITY MODELS OF COMMUTING CONDITIONS: A CZECH CASE STUDY

**Jiří Horák[1], Igor Ivan[2], Jan Tesla[3]**
[1,2,3] *VŠB-Technical University of Ostrava, Institute of Geoinformatics, 17.listopadu 15, Ostrava-Poruba, Czechia*

**Abstract:** The aim of this paper is to present the development of improved gravity models of commuting conditions based on alternatively-assessed "mass" of origins and destinations, as well as on improved distance evaluation. For the case study the Ostrava region (CZ) was selected and delimited as a catchment with significant everyday commuting flows to Ostrava, observed in census 2011. The distribution of employers was constructed using the Albertina - Company Monitor. The number of employees in all companies within each municipality and each economical sector represents the attractiveness of municipalities as destinations in the gravity model. The origins in the gravity model are represented by municipalities. Their power to generate commuting flows is directed by the population of employed inhabitants. The distances between municipalities were enumerated using road network analysis and substituted by railway distances if they were shorter. For selected economic sectors (i.e. processing industries, internal trade) the specific gravity models of regional daily commuting flows were calibrated. The calibration was based on observed commuting flows between municipalities for each economic sector adjusted for missing census' answers. Differences among sector gravity models are discussed. Their Root Mean Square Errors and Mean Errors show the overall fitness of models. The variability of models' parameters generally corresponds to expected differences among economic sectors. Finally, extreme differences of observed and expected flows for each gravity model were analysed. The results prove the usefulness of applied improved gravity models to understand commuting conditions, discover drawbacks of existing public transport services and challenges to improve the current conditions towards better efficiency of transport organisation. The parameters of gravity models can be utilised in stochastic simulations of random commuting flows in the region.

**Keywords**: commuting, gravity model, regional mobility, transportation.

## 1. Introduction

The popularity of gravity models is given by the fact they express basic characteristics of interaction data – the strength of relationships between two objects directly depends on mass of both objects and indirectly on the distance between them. The theoretical basis of gravity models for commuting purposes are entropy maximization (Wilson, 1967, Bailey, Gatrell, 1995) and random utility theory (Anas, 1983).
The family of gravity models differs by their mathematical formulations, constrains, data sources, methods of calibration, convergence criteria and other parameters. The selection and modification should be fitted to the required purpose.
The form of mathematical specification of gravity models depends mainly on the formulation of deterrence function between two zones or objects. These are exponential function $e^{-\alpha(cij)}$; power function $c_{ij}^{\beta}$ ; Tanner (or Gamma) function a $* e^{-\alpha(cij)} * c_{ij}^{\beta}$ (Celik, 2010), log-normal function (Shrewsbury, 2012), Box-Cox, Pareto function and its modifications (Halás et al, 2014). Most applications of gravity models are based on the exponential deterrence function (Uboe, 2004). The appropriateness of the functional form for the deterrence function should be always critically examined (Fik, Mulligan, 1998).
The traditional form of trip distribution problem is solved using a doubly constrained model formulation. It refers constrains where the marginal totals of the trip matrix are considered to be given (Uboe, 2004). The calibration of gravity models requires a good representation of flows between zones (objects) in trip matrix implying a large sample size. Celik (2010) approves a sample size around 1000 for each trip purpose would produce approximately the same parameter estimate as models with a very large sample size. The calibration methods may be based on iterative procedures (i.e. BPR approach see Viton 1995), maximum likelihood estimations (or other regression based models), Hyman's calibration algorithm (Williams, 1976), three-point rational function interpolation method (Williams 1977), stochastic heuristic optimization procedures (Tsekeris, Stathopoulos 2006) and other approaches.
For regression based calibration the equations of gravity models are linearized using appropriate transformations. Viton (1995) rewrite the usual specification of gravity model into a following form:

$$V_{ij} = \frac{T_{ij}}{O_i * A_j} \qquad V_{ij} = \theta * d_{ij}^{-\beta} * \varepsilon_{ij}$$

where θ and β are unknown parameters, ε is a random variable, $d_{ij}$ distance, $T_{ij}$ number of trips, $O_i$ originations, $A_j$ attractions. The transformation using $V_{ij}$ leads to a single regression solution.
More often the gravity model is directly linearized into multiple linear regression form (i.e. Khadaroo, Seetanah, 2008, Grange et al., 2009). Advantages are possibilities to avoid a hidden multicollinearity between variables and evaluation of influences of all independent variables suitable for interpretation of relationships in the gravity model.

---

[1] Corresponding author: jiri.horak@vsb.cz

The formulation and performance of gravity models are dependent on many other factors, namely (Celik, 2010): choice of spatial separation measure, choice of travel mode, choice of matrix type (i.e. production–attraction or origin–destination matrix), choice of time interval, and choice of travel mode.

The gravity models represent the appropriate theoretical tool for evaluation of spatial interactions, features and behaviours of origins and destinations, enable to predict and simulate future or conditional behaviour. Nevertheless, their capabilities are constrained and influenced by deteriorating factors including:

- Time heterogeneity (changes of gravity models' parameters over time – Mikkonen, Luoma, 1999; temporal evolution of the trip distribution within study time period – Tsekeris, Stathopoulos, 2006)
- Space heterogeneity (gravity based estimates of parameters vary systematically across space – Uboe 2004)
- Real geographical/transport constrains (i.e. limited accessibility of some potential destinations)
- Limited spatial resolution (distance between the city centers do not coincide with the average distance between internal and external commuting – Uboe 2004)
- Workers heterogeneity (neither qualifications in the labour market and their response to distance are homogeneous - Uboe 2004)
- Jobs heterogeneity (not all destinations provides a proportional distribution of all relevant job offers).
- hierarchical order of potential destinations and their competition (Fotheringen, O'Kelly, 1989, proposed competing destination models),

The gravity models and their behaviour and limitations are subjects of continuous research including partially constrained and unconstrained models (Bailey, Gatrell, 1995), models for dynamic transport planning (reflecting the temporal development of travel behaviour - Tsekeris, Stathopoulos, 2006), two-dimensionally constrained choice model (Vrtic et al., 2007). Gravity models are applied not only for trip distribution studies, but frequently also for economic applications like evaluation of mutual business exchange between countries or regions, studies of tourism interactions, for geographical applications (Halás, Klapka 2010 for estimations of spheres of influence for Czech cities).
The preferred double constrained gravity models may be inappropriate in condition when not all trips are known. One of the reasons is a weak data source. In the Czech Republic, the main data source for evaluation of journey-to-work condition is census. Nevertheless, it is provided one time per ten years (with delaying 2-3 years) and furthermore, the quality is deteriorated by still a large share of missing or curious answers.

## 2. Description of study area and data

The Ostrava region represents an industrial region in the NE corner of the Czech Republic, close to borders with Poland and Slovakia. Main cities are Ostrava (296224 inhabitants, census 2011), Havířov (76694), Opava (58351), Karviná (56897) and Frýdek-Místek (56356). The role of traditional mining, metallurgy, and heavy machinery industries are increasingly substituted in the last decade by automobile industry, IT and other high-tech companies. These changes are interlinked with a transformation of labour market and changes in labour force demand.
The Ostrava region is delimited as a catchment with significant daily commuting flows to Ostrava, observed in census 2011, and it contains 230 municipalities (fig. 1), with total population 1.1 million.

### 2.1 Flows and characterisation of origins and destinations

Main data was extracted from results of **census 2011** (relevant date 25.3.2011) including recorded flows between municipalities (all, daily, journey-to-work all and classified according to main economic sectors), aggregated demographic, economic and commuting characteristics of individual municipalities (namely employed and commuting, both characteristics expressed as all and classified according to main economic sectors).
Unfortunately, census 2011 suffers from the large portion of missing answers which reflects low and still decreasing willingness of respondents to answer questions especially those sensitive from the personal point of view.
The main problem was recognised for recorded commuting flows between municipalities. The sum of flows from the each municipality was compared with the number of commuting workers. The last one is usually about 3-4% higher, but in some towns the difference is more than 40% (Budišov) and in the case of Ostrava (the largest city) the difference is 30% (!) which cannot be neglected. The original data was adjusted for missing values by coefficient representing a portion of missing answers for the starting municipality.
The distribution of employers was constructed using the **Albertina - Company Monitor** (ACM). Based on the name of resident municipality, records of companies from Ostrava region has been exported including classifications of company's activities according to basic economic sectors, size category of employees, and the institutional form.
First, companies located in the region were checked for indication of liquidation or deletion from the official registers; such companies were excluded from the list (3405 from 231336 which is only 1.5%). This reduction was discovered as not fully sufficient. The register contains more companies than those really acting in the labour market; some of them are "sleeping" (especially small) or transformed. Unfortunately, there are limited possibilities to verify the status of all companies. Data is considered excessive and appropriate corrections are applied to reduce this surplus (see later).

The important step is the estimation of the number of employees for each size category. The mean for each size category has been estimated using a sample of companies with known number of employees (provided by a labour office). In case of missing evaluation of the company by the size category, their classification according to an institutional form (altogether 87 different forms) was used. The population of these companies was estimated using average value for each institutional form. Finally, the total number of employees (as a number of jobs) was aggregated in each municipality.

Nevertheless, due to the previously mentioned excessive number of companies in municipalities, data was adjusted using the total sum of employment in each municipality provided by the Czech Ministry of Finance.

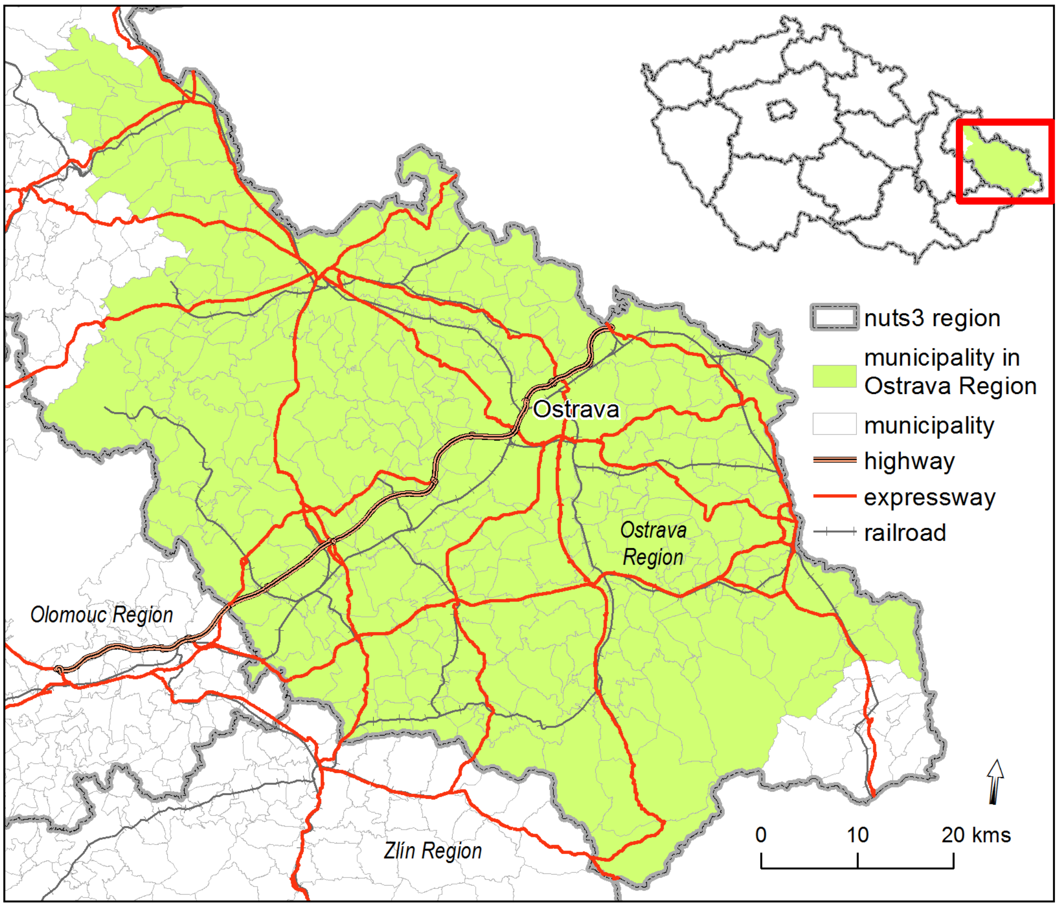



**Fig. 1. Study area**

*Source: administrative borders (RSO CZSO), transport infrastructure (AcrCR500, RMD), own definition of the region*

Following economic sectors were investigated: processing industries, internal trade, and education. Other important sectors (i.e. civil engineering, transport and logistics) were not selected due to the fact that in these branches the residence of the company frequently does not correspond to the job place.

### 2.2 Distances

The road network valid in 2011 was obtained from the Road and Motorway Directorate of the Czech Republic (RMD). The road distances between municipalities were generated using network analysis in ArcGIS 10.1 between municipality's centres. The second data source for inter-municipal distance estimations was a database of public transport connections (Horák et al., 2014). In this database the public transport connectivity and its attribute for each pair of municipalities within 100 km are stored using both bus and train transport modes. The database is regularly updated using a special SW application TRAM. The process of public transport connection searching, processing and storage, SW architecture and design are described in Fojtík et al. (2009). Distances between municipalities based on the public transport mode were selected from the database (updated in March 2011) and compared with distances derived by the network analysis. On the assumption that the shorter distances of public transportation are relevant more to the real commuting conditions the longer road distances were replaced by those of public transport. Such replacements occur in 8.5%.

## 3. Methodology and findings, discussion

Based on data availability and its modifications following data has been used in gravity models:

- **Trips** – journey-to-work in the given sector between municipalities (adjusted data from census 2011) (TRIPS).
- **Origins** – sum of adjusted trips (in the given sector) from the starting municipality (ORIGINS, O),
- **Number of employees** in the given sector from census 2011 (EMPLOYEES, E).
- **Destinations** (attractiveness) - sum of adjusted trips (in the given sector) to the ending municipality (DESTINATIONS, D),
- **Number of jobs** in the given sector from ACM adjusted by data of the Ministry of finance (JOBS, J).

Following type of models have been selected for testing: power (PW), exponential (EXP), log-normal (LN). The universal distance-decay function derived by Halás et al. (2014) for Czech regional centres did not provide satisfactory results in our case ($R^2$ is 0.184). Original forms of equations have been linearized using logarithm transformation. Obtained multiple linear regression forms were optimised in SPSS v.18. All models were developed in two variants – double constrained models use variables ORIGINS and DESTINATIONS (hereafter OD models) and unconstrained models utilize variables EMPLOYEES and JOBS (hereafter ED models). Tested mixed variants with both OD and EJ parameters showed extremely high multicollinearity (due to a high correlation between O and E, D and J respectively) and they were excluded from further analysis. Finally six types of gravity models for each economic sector were analysed: EJPW $Y_{ij} = \beta^0 * E_i^{\beta^1} * J_j^{\beta^2} * d_{ij}^{\beta^3}$, ODPW $Y_{ij} = \beta^0 * O_i^{\beta^1} * D_j^{\beta^2} * d_{ij}^{\beta^3}$, EJLN $Y_{ij} = \beta^0 * E_i^{\beta^1} * J_j^{\beta^2} * e^{(\beta^3 * \ln^2(d_{ij}))}$, ODLN $Y_{ij} = \beta^0 * O_i^{\beta^1} * D_j^{\beta^2} * e^{(\beta^3 * \ln^2(d_{ij}))}$, EJEXP $Y_{ij} = \beta^0 * E_i^{\beta^1} * J_j^{\beta^2} * e^{\beta^3 * d_{ij}}$, and ODEXP $Y_{ij} = \beta^0 * O_i^{\beta^1} * D_j^{\beta^2} * e^{\beta^3 * d_{ij}}$.

First, exploratory data analysis for all variables and their logarithm transformation has been provided. The data distribution was checked and outliers have been excluded to improve the behaviour of linear regression. The results of model optimisation (power and log-linear models, both OD and EJ variants) for educational sector approved the positive result of outliers' elimination – index of determination was increased by 3-6% and RMSE by 9-11%. All gravity models were optimised for data without outliers.

As a second step, results of flow data modification (section 2) were verified on power and exponential models for both OD and EJ variants in the educational sector. The change of RMSE is very small (0-2%), but the positive change of the mean error is strong (table 1).

**Table 1**

*Comparison of models with original and adjusted data*

| Parameter | Original flows, EJPW model | Adjusted flows, EJPW model | Difference in % | Original flows, EJEXP model | Adjusted flows, EJEXP model | Difference in % |
|---|---|---|---|---|---|---|
| RMSE | 8,52 | 8,54 | -0,19 | 8,36 | 8,20 | 1,82 |
| Mean error | -2,09 | -1,05 | 49,58 | -1,95 | -0,96 | 50,97 |

**Table 2**
*Results of optimisation of multiple linear regression models for industrial (INDU), educational (EDU) and trade (RET) sectors*

| Model | $R^2$ | RMSE | Mean error | Part correlation $\beta^1$ | Part correlation $\beta^2$ | Part correlation $\beta^3$ |
|---|---|---|---|---|---|---|
| INDU-EJPW | 0,466 | 43,09 | -4,05 | 0,519 | 0,599 | -0,453 |
| INDU-ODPW | 0,535 | 42,00 | -4,98 | 0,538 | 0,648 | -0,504 |
| INDU-EJLN | 0,474 | 42,80 | -3,89 | 0,524 | 0,604 | -0,462 |
| INDU-ODLN | 0,545 | 41,72 | -4,98 | 0,544 | 0,654 | -0,513 |
| INDU-EJEXP | 0,482 | 42,34 | -4,09 | 0,533 | 0,604 | -0,547 |
| INDU-ODEXP | 0,542 | 40,98 | -4,84 | 0,547 | 0,645 | -0,510 |
| EDU-EJPW | 0,520 | 8,54 | -1,05 | 0,558 | 0,649 | -0,454 |
| EDU-ODPW | 0,550 | 8,48 | -1,40 | 0,580 | 0,656 | -0,451 |
| EDU-EJLN | 0,530 | 8,55 | -1,18 | 0,563 | 0,655 | -0,464 |
| EDU-ODLN | 0,559 | 8,40 | -1,36 | 0,584 | 0,661 | -0,461 |
| EDU-EJEXP | 0,542 | 8,26 | -0,96 | 0,565 | 0,660 | -0,477 |
| EDU-ODEXP | 0,567 | 8,20 | -1,31 | 0,585 | 0,663 | -0,470 |
| RET-EJPW | 0,501 | 18,11 | -2,08 | 0,496 | 0,636 | -0,488 |
| RET-ODPW | 0,520 | 17,96 | -2,50 | 0,499 | 0,645 | -0,495 |
| RET-EJLN | 0,510 | 18,07 | -2,14 | 0,501 | 0,641 | -0,498 |
| RET-ODLN | 0,528 | 18,01 | -2,68 | 0,503 | 0,650 | -0,503 |
| RET-EJEXP | 0,512 | 17,75 | -1,92 | 0,501 | 0,638 | -0,499 |
| RET-ODEXP | 0,522 | 17,58 | -2,31 | 0,501 | 0,643 | -0,499 |

All β coefficients in regression equations and F statistics for ANOVA in all cases are significant ($p<0.001$).
Comparison between OD and EJ models shows quite similar results. One of possible explanation is the labour market in the region is relatively closed (less influenced by external commuting to work) and EJ model represents a suitable substitution for OD models.
Shapes of curves for different gravity models in education sector and hypothetical value of E=2000 and J=20000 (which simulate conditions of commuting to Ostrava from other larger towns) are depicted in fig. 2 (left). The elimination of outliers increases the estimated values of flows, mainly for short distances. The significant differences in shapes are between exponential models and the group of power and log-normal models – the deterrence function is moderate for exponential models while it is quite steep for the other evaluated models. As expected, other sectors demonstrate similar shape of regression (see i.e. modelled flows for E=2000, J=20000 for processing industry sector in fig. 2 (right).
Finally, the exponential type of models has been selected because of:

- Low errors (RMSE, ME)
- Relatively high index of determination (relatively to other models)
- Higher part (semipartial) correlation
- Satisfactory distributed regression residuals (i.e. Fig. 3)
- Better behaviour of the regression curve

The potential weakness of exponential models lies in higher deviation from normal distribution of the distance variable which is highly skewed.

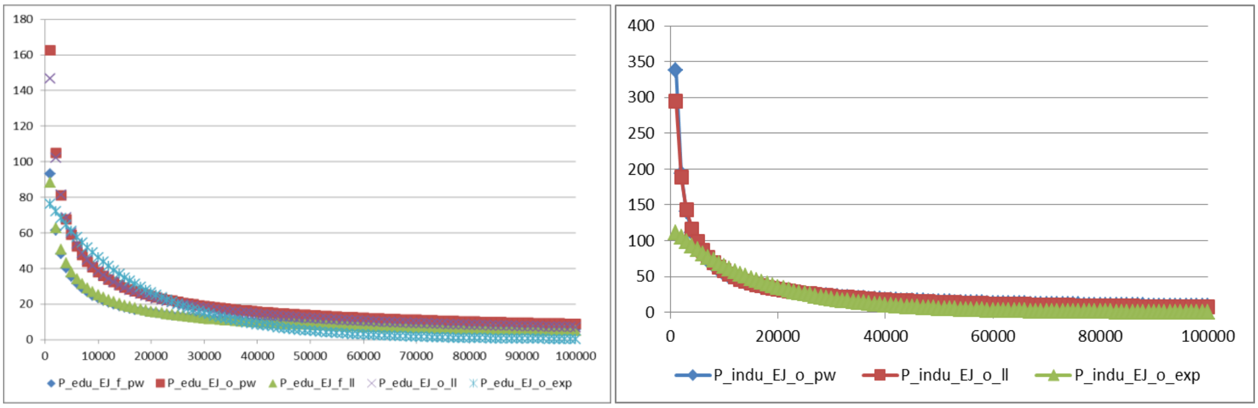



**Fig. 2.**
*Regression EJ models (E=2000, J=20000) for education (left) and processing industry (right)*
*Source: own calculation*

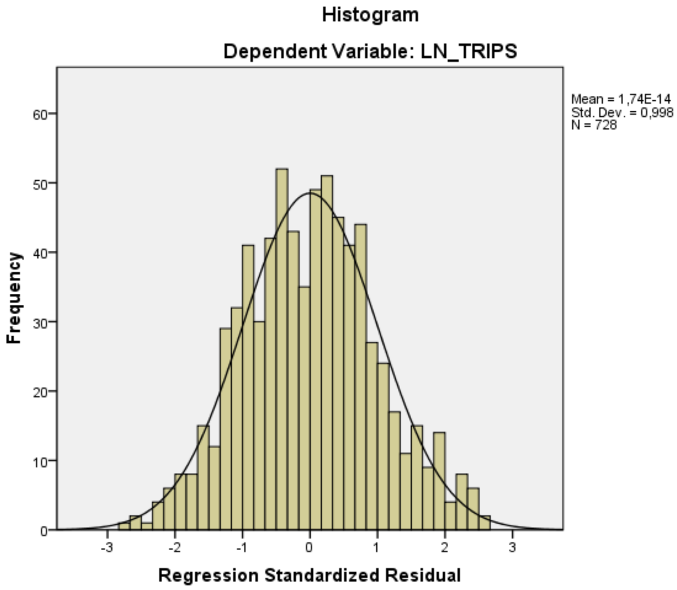



**Fig. 3.**
*Histogram of standardized regression residuals for exponential model EDU-EJEXP*
*Source: own calculation*

The optimised exponential gravity models (unconstrained, based on E-J variables, adjusted, outliers eliminated) has been applied for 3 explored economic sectors. The example with E=2000 and J=20000 was used again to enable comparison of behaviour of these models. As it can be seen in Fig. 4, the behaviour of these gravity models for education and trade sectors is almost the same which indicates no significant differences in attractiveness and the power of opportunities for both sectors. The industrial model shows clearly different behaviour – it attracts flows much more than in the case of education or trade (the shape is steeper and the values are usually by 50% higher for short distances). It indicates higher mobility of processing industry labour force. This result approves the real general conditions (it seems to reflect the differences among sectors) – the processing industry in our region is characteristic by numerous large employers located partly outside cities (industrial landscape with large industrial plants between towns) which obviously have to generate large flows of employees. Another factor enhancing flows is the higher level of salaries.
The variability of models' parameters generally corresponds to expected differences among economic sectors.
Finally, extreme differences of observed and expected flows for each gravity model were analysed.
Commuting trips to processing industry show largest differences (fig. 5) for Havířov-Ostrava. Havířov was built after the 2nd World War as a new settlement with dominant resident function and almost missing employment function. Main flows (both relative and absolute) from Havířov are to Ostrava and to Nošovice. Relatively large flows point to dominant industrial centres in the Ostrava region. Nošovice with large automotive companies (Hyundai 3500 workplaces, Mobis Automotive 850, Dymos CR 420) attracts main flows from Ostrava, Havířov and Frýdek-Místek. Traditional metallurgy and machinery centre Třinec is also frequently recorded in relatively high flows, but these flows comes mainly from close neighbourhood. Large flows attracts Frenštát pod Radhoštěm (important electrotechnic companies Continental Automotive Czech Republic, Siemens (former Siemens Electromotors), information technologies Logis etc.). Unexpected high flow of industrial employees was recognised also from Opava to the small town Hradec.

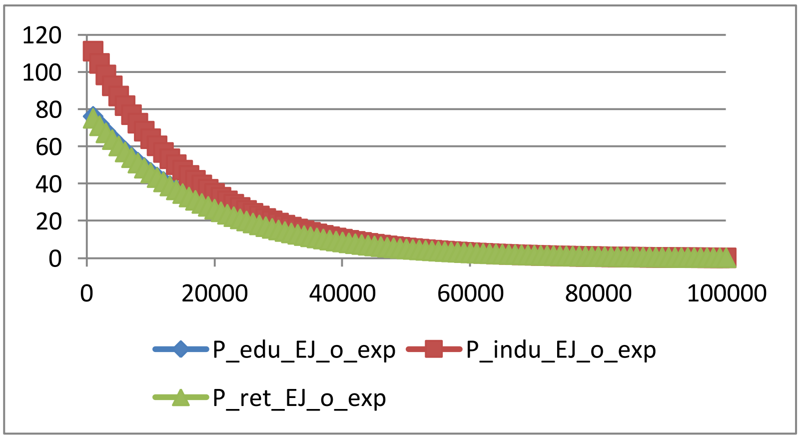



**Fig. 4.**
*Exponential regression EJ models (E=2000, J=20000) for education, processing industry and trade*
*Source: own calculation*

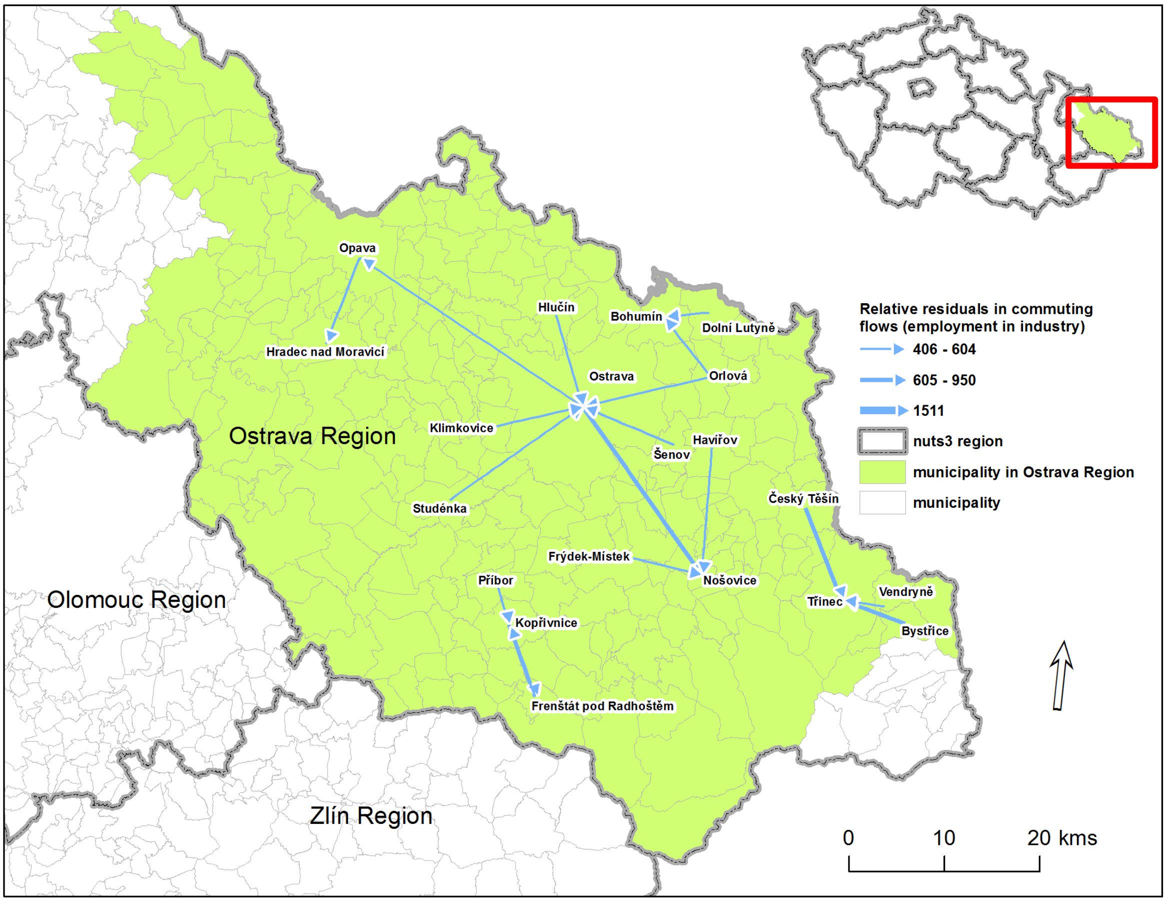



**Fig. 5.**
*Large unexpected flows (commuting in processing industry) by high relative residuals (INDU-EJEXP)*
*Source: own calculation*

Commuting trips to education jobs indicate large differences for several small municipalities and towns around larger cities (mainly Ostrava and Frýdek-Místek) which can be easily explained by suburbanisation. This suburbanisation is geographically asymmetric in both cases. Extreme high mutual flows are documented for Ostrava-Opava connection. Both Ostrava and Opava host universities and numerous secondary schools and these flows declare the strength of exchange of teachers (and other workers in educational sector). Very high outflows are depicted again for Havířov to more distant places (Ostrava, Karviná, Č.Těšín).
Extremely low flows according to this gravity model were discovered mainly for journeys coming from central cities to smaller towns in the surroundings. This finding is expected – the educational opportunities are concentrated in larger towns and more in regional centres, thus return flows should be quite weak.
The results are used also for evaluation of existing public transport services with the aim to improve the current conditions towards better efficiency of transport organization. It may decrease a growing stress of traffic due to rapid suburbanization processes (i.e. Burian et al., 2014).

## 4. Conclusion

The set of gravity models for journey-to-work in three economic sectors (processing industry, education, trade) was developed in Ostrava region (CZ). The data from census 2011 was combined with estimations from Albertina - Company register and statistical data from the Czech Ministry of Finance. The flows registered in census 2011 have been adjusted and outputs of modelling partly approved usefulness of such modifications.

The double constrained models using sum of flows as origins and destinations have provided similar statistical characteristics and function parameters as for unconstrained gravity models with estimations of number of employees (trip production) and estimations of number of jobs (trip attraction). Nevertheless, indexes of determination are not satisfactory (0.46-0.56) and further development and improvement of gravity models are envisaged. The exponential unconstrained models have been selected to represent commuting conditions for investigated economic sectors. The behaviour of gravity models for education and trade sectors is almost the same which indicates no significant differences in attractiveness and the power of opportunities for both sectors. The industrial model shows the industrial employers attracts flows much more than in cases of education and trade.

The analysis of relative regression residuals has provided valuable results for interpretation due to discovering unexpected flows in the region for industrial and educational sectors. The parameters of optimised gravity models can be utilised in stochastic simulations of random commuting flows in the region to discover local constrains (limited accessibility) in the region.

**Acknowledgements**

The research is supported by the Czech Science Foundation, project Spatial simulation modelling of accessibility, No. 14-26831S.

**References**


Anas, A. 1983. Discrete choice theory, information theory and the multinomial logit and gravity models. *Transportation Research B*, 17: 13–23.

Bailey, T.C.; Gatrell, A.C. 1995. *Interactive spatial data analysis*. Longman Scientific & Technical, Essex. 413 p.

Burian, J.; Zajickova, L.; Tucek, P.; Vozenilek, V.; Langrova, B.; Boori, M. 2014. Traffic intensity changes and their influence on spatial distribution of suburbanization. In *14th International Multidisciplinary Scientific GeoConference SGEM 2014*.

Celik, H.M. 2010. Sample size needed for calibrating trip distribution and behavior of the gravity model. *Journal of Transport Geography,* 18(2010): 183–190.

Fik, T. J.; Mulligan, G.F. 1990. Spatial flows and competing central places: towards a general theory of hierarchical interaction. *Environment and Planning A.* 22: 527–549.

Fojtík, D.; Horák, J; Šeděnková, M.; Ivan, I. 2009. Software support for automatic creating database of public transport connections. In *Proceedings of the International Carpathian Control Conference ICCC 2009, Zakopane, Poland, ISBN 8389772-51-5,* pp. 517-520.

Fotheringen, S.; O'Kelly, M.E. 1989. *Spatial Interaction Models: Formulations and Applications*. Kluwer Academic Publishers, Dordrecht.

Grange, L.; Troncoso, R.; Ibeas, A.; González, F. 2009. Gravity model estimation with proxy variables and the impact of endogeneity on transportation planning. *Transportation Research Part A.* 43 (2009), 105–116.

Halás, M.; Klapka, P.; Kladivo, P. 2014. Distance-decay functions for daily travel-to-work flows. *Journal of Transport Geography,* 35(2014): 107–119.

Horák, J.; Ivan, I.; Fojtík, D.; Burian, J. 2014. Large scale monitoring of public transport accessibility in the Czech Republic. In *Proceedings of ICCC 2014, Velké Karlovice, 28-30.5.2014,* pp. 1-7.

Khadaroo, J.; Seetanah, B. 2008. The role of transport infrastructure in international tourism development: A gravity model approach. *Tourism Management,* 29(2008): 831–840.

Mikkonen, M.; Luoma, M. 1999. The parameters of the gravity model are changing - how and why? *Journal of Transport Geography*, 7(1999): 277-283.

Shrewsbury, J. 2012. Calibration of trip distribution by generalised linear models. Available from Internet:.<http://www.nzta.govt.nz/resources/research/reports/473/docs/473.pdf>

Tsekeris, T.; Stathopoulos, A. 2006. Gravity models for dynamic transport planning: Development and implementation in urban networks. *Journal of Transport Geography,* 14(2006): 152–160.

Uboe, J. 2004. Aggregation of gravity models for journeys to work. *Environ. Plan. A.* 36 (4), 715–729.
Viton, P.A. 1995. Calibrating the Gravity Model. Available from Internet: <http://facweb.knowlton.ohio-state.edu/pviton/courses2/crp5700/calibrating.pdf>

Vrtic, M.; Frohlich, P.; Schussler, N.; Axhausen, K.W.; Lohse, D.; Schiller, C.; Teichert, H. 2007. Two-dimensionally constrained disaggregate trip generation, distribution and mode choice model: Theory and application for a Swiss national model. *Transportation Research Part A.* 41(2007): 857–873.

Williams, I. 1976. A comparison of some calibration techniques for doubly constrained models with an exponential cost *function. Transportation Research 10*, 91–104.

Williams, I. N. 1977. Algorithm 1: Three-point rational function interpolation for calibrating gravity models. *Environment and Planning 9A*, 215–221.

Wilson, A.G. 1970. *Entropy in Urban and Regional Modeling*. Pion, London.

# WAVE MODELING WITH DATA ASSIMILATION TO SUPPORT THE NAVIGATION IN THE BLACK SEA CLOSE TO THE ROMANIAN PORTS

**Dorin Butunoiu[1], Eugen Rusu[2]**
[1,2] *"Dunarea de Jos" University of Galati, Department of Mechanical Engineering, Galati, Romania*

**Abstract:** The objective of the present work is to develop a reliable wave prediction system, based on numerical models, in the Black Sea in order to support the navigation and the harbor operations in the Romanian nearshore. This system focuses, in a multilevel computational scheme, on the Romanian coastal environment. The SWAN model was implemented and calibrated for the entire sea basin against satellite data. Nevertheless, in some particular cases, as for example the extreme storm conditions, the numerical predictions might be less accurate and such behavior is sometimes even more accentuated in the sea environment than in the ocean. For these reasons, data assimilation techniques are considered to combine the measured wave parameters with those predicted by the numerical models in order to derive better estimations of the sea states. In the present work, the wave measurements from the Gloria drilling unit have been considered for the assimilation procedure. A scheme based on the successive corrections method has been implemented to perform the assimilation in both the geographical and spectral spaces. In this way, the model predictions in the Romanian nearshore are improved and the wave modeling system developed can provide a better real time support to coastal navigation and harbor operations. Finally, the extension of this scheme to the forecast products is also considered.

**Keywords**: numerical wave models, Black Sea, data assimilation, navigation risks, Romanian harbors.

## 1. Introduction

A better prediction of the wave conditions in the coastal environment of the Black Sea became in the last decades an issue of increasing importance. The recent economic developments, including the last years offshore activities related to the oil and gas exploitations, enhanced substantially the marine activities, in general and the maritime traffic, in special (Gasparotti and Rusu, 2012). The western side of the sea has been subjected to the highest traffic since on this side the mouths of the Danube River and the entrance in the navigation channel that links the Black Sea with the Danube River are located. Moreover, via the seventh Pan European transportation corridor (the inland navigation system Danube-Rhine-Maine), the Black Sea is connected to the North Sea, and for this reason the coastal navigation is significantly higher in this area (Ivan et al, 2012). Since the Romanian harbors at the Black Sea represent an important link in this logistic chain, the increased maritime traffic induces also a higher risk of maritime accidents with unwished and sometimes even catastrophic consequences on the sea and coastal environments.

On the other hand, the risks of marine and coastal hazards are also enhanced due to the major climate transformations that are more and more evident in the last years in the marine areas. Some relevant insights concerning the climate dynamics in the Mediterranean and Black seas are given in Ardhuin et al (2007), Onea and Rusu (2014), Mateescu and Coman (2008), and Maslova et al (2010).

An appropriate way to reduce the risks of marine and coastal hazards would be to provide in a more realistic way predictions of the wave conditions in the coastal areas that are subjected in general to high navigation traffic. This can be done by using spectral phase averaged wave models. Such models present the advantage that can cover large geographical spaces and also can provide nowcast and forecast products (Rusu et al, 2006).

In order to calibrate the wave models in specific areas, extended hindcast studies are usually carried out, performing validations against remotely sensed data and in situ measurements. A method to increase the reliability of the wave models is to use data assimilation (DA) techniques. This means in general to combine the model results with the measurements (either remotely sensed or in situ) in order to increase the reliability of the model predictions. Such techniques are even more valuable in the enclosed seas (as the Black Sea is), which are more difficult environments from the point of view of the wave modeling processes (Rusu, 2010 and 2011).

From this perspective, the objective of the present work is to validate a wave prediction system based on numerical spectral models in the basin of the Black Sea and to implement a DA scheme for improving the reliability of the wave predictions close to the major Romanian harbors.

## 2. Reliability of spectral wave models in the Black Sea

A wave prediction system SWAN based was implemented and validated in the basin of the Black Sea. SWAN (Booj et al, 1999, acronym from Simulating Waves Nearshore) is a third generation spectral wave model that integrates the spectral action balance equation (1) in time, geographical and spectral spaces.

[2] Corresponding author: erusu@ugal.ro

The spectrum that is considered in most of the actual wave models is the action density spectrum *(N)*, rather than the energy density spectrum, since in the presence of currents, action density is conserved whereas energy density is not. The action density is equal to the energy density *(E)* divided by the relative frequency ( $\sigma$ ). Thus the action balance equation is:

$$\frac{DN}{Dt} = \frac{S}{\sigma}. \tag{1}$$

*S,* from the right hand side of the action balance equation, represents the source terms. In deep water, three components are significant in the expression of the total source term. They correspond to the atmospheric input, whitecapping dissipation and nonlinear quadruplet interactions, respectively. Various parameterizations for these source terms are alternatively available in SWAN and tunable coefficients are in general defined for each case. Besides these three terms, in shallow water additional source terms corresponding to phenomena like bottom friction, depth induced wave breaking and triad nonlinear wave-wave interactions may play an important role, and the total source becomes:

$$S = S_{in} + S_{dis} + S_{nl} + \underbrace{S_{bf} + S_{br} + S_{tri} + \ldots}_{finite\ depth\ processes} \tag{2}$$

Although the model is not considered appropriate for ocean scale simulations, the recent developments implemented in SWAN in relationship with the propagation scheme and with some physical processes (SWAN team, 2013) make the model quite suitable for simulations at sub oceanic scales and definitively it represents the best choice for the enclosed seas, as the Black Sea is. Moreover, in such environments SWAN presents the advantage that one single model can cover the full scale of the wave modeling process in a multi-level wave modeling system that can be focused on the most sensitive coastal areas, as for example the harbors. Some relevant results related to the SWAN predictions in the Black Sea basin are provided in Rusu and Ivan (2010) and Butunoiu and Rusu (2012).

Additional information concerning the model settings and the physical parameterizations considered for the SWAN simulations carried out in the basin of the Black Sea is given in Table 1. Thus, the implementation of the SWAN model was made for 36 directions and 30 frequencies logarithmically spaced from 0.12 Hz to 1.2 Hz. The model was executed without the influence of currents. The computations were performed in the non-stationary mode with a 20 minute time step. The number of iterations was set to 4, so the numerical accuracy would be increased between iterations. These specifications are presented in Table 1, along with the characteristics of the computational domain and indicating also the physical processes activated. The quantities in Table 1 have the following signification: $\Delta x$ and $\Delta y$ represent the resolution in the geographical space, $\Delta\theta$ – resolution in directional space, $\Delta t$ – time resolution, *nf* – number of frequencies in spectral space, $n\theta$ – number of directions in spectral space, *ngx* – number of grid points in *x* direction, *ngy* – number of grid points in *y* direction and *np* – total number of grid points. The significations of the input fields and of the physical processes activated are: wave – wave forcing, tide – tide forcing, wind – wind forcing, curr – current field input, gen – generation by wind, wcap - whitecapping process, quad – quadruplet nonlinear interactions, triad – triad nonlinear interactions, diff – diffraction, bfric – bottom friction, set up – wave induced set up, br – depth induced wave breaking.

**Table 1**
*SWAN model configuration for the computational domain corresponding to the entire Black Sea level*

| **Sea level SWAN configuration** | **Coordinates** | | **Δx × Δy (°)** | | **Δθ x Δt (°)-(min)** | **Mode** | **nf** | **nθ** | **ngx × ngy = np** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Spherical | | 0.08 x 0.08 | | 10 x 20 | non-stat | 30 | 36 | 176 x 76= 13376 | | | |
| **Input / Process** | **wave** | **wind** | **tide** | **crt** | **gen** | **wcap** | **quad** | **tri** | **diff** | **bf** | **set up** | **br** |
| | X | X | 0 | 0 | X | X | X | 0 | 0 | X | 0 | X |

Considering the above mentioned settings, the present work presents some results of the simulations performed in the five-year period 2000-2004. Validations have been performed both against satellite data and in situ measurements. Thus, Fig. 1 illustrates the $H_s$ scatter diagrams, SWAN against the satellite data along the Black Sea, for the entire 5-year period (a), and only in the winter time periods (b). In these analyses the winter time represents all the periods between September to March (inclusively). At the same time, Fig. 2 illustrates the $H_s$ scatter diagrams, SWAN against measurements at the Gloria drilling unit for the entire 5-year period (a) and only for the winter time periods (b). The Gloria drilling platform operates in the western sector of the Black Sea (44$^0$31'N, 29$^0$34'E) at a location where the water depth is about 50m.

The measurements were performed daily during the 5-year period considered at 6-h intervals, the percentage of valid data being about 92% and the technique developed by Makarinskyy et al (2005) was considered to fill the gaps in the measurements.

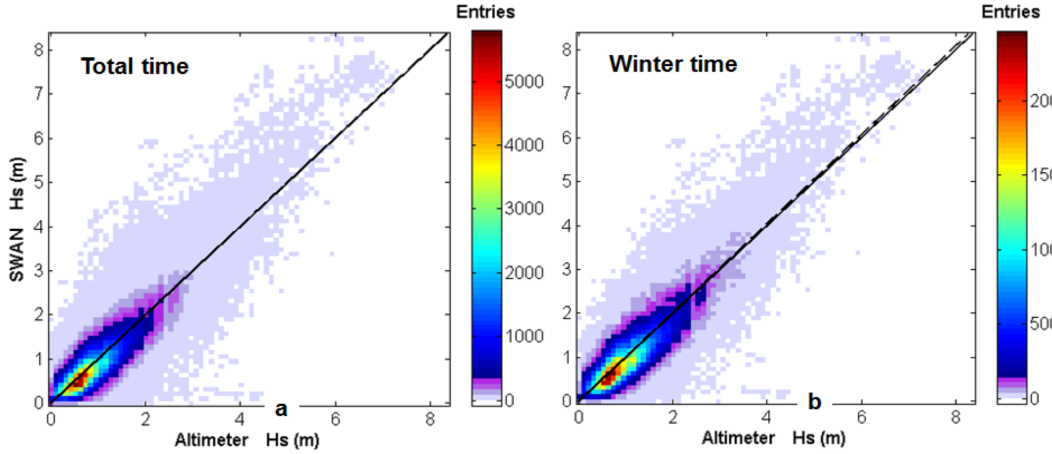



**Fig. 1.**
*$H_s$ scatter diagrams, SWAN against satellite data in the Black Sea, corresponding to the time interval 2000-2004; a - the entire 5-year period and b - winter time periods*

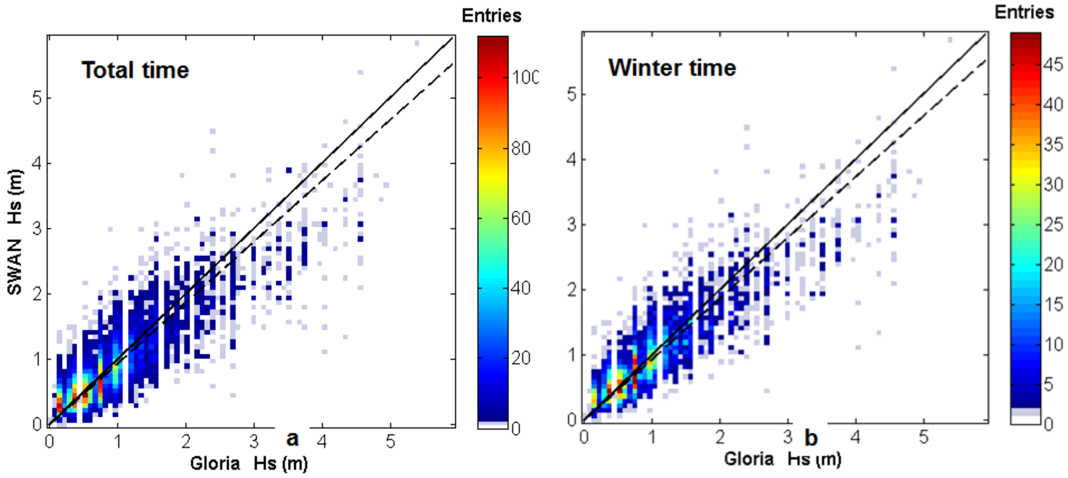



**Fig. 2.**
*$H_s$ scatter diagrams, SWAN against measurements at the Gloria drilling unit, corresponding to the time interval 2000-2004; a - the entire 5-year period and b - winter time periods*

Table 2 from bellow illustrates the $H_s$ statistics, SWAN model simulations against remotely sensed data in the Black Sea and SWAN results against in situ measurements at the Gloria drilling unit corresponding to the total time and only to the winter time, respectively. The statistical parameters considered are those commonly used to evaluate the wave model output, such as the mean error (*Bias*) computed as the difference between the average values of the simulated (*MedSim*) and observed (*MedObs*) data, the mean absolute error (*MAE*), the root mean square error (*RMSE*), the scatter index (*SI*) defined as the ratio of standard deviation of error to the mean observed values, the linear correlation coefficient (*R*) and the symmetric slope (*S* - for *S* >1 the model overestimates the observations).

**Table 2**
*$H_s$ statistics, SWAN model simulations against remotely sensed data in the Black Sea and SWAN results against in situ measurements at the Gloria drilling unit corresponding to the total time and only in the winter time, respectively.*

| | MeanM | MeanS | Bias | MAE | RMSE | SI | R | S | N |
|---|---|---|---|---|---|---|---|---|---|
| **Satellite-Total** | 1.07 | 1.04 | -0.03 | 0.28 | 0.38 | 0.36 | 0.87 | 1.00 | 323709 |
| **Satellite-Winter** | 1.26 | 1.23 | -0.03 | 0.30 | 0.41 | 0.33 | 0.88 | 1.01 | 172117 |
| **Gloria-Total** | 0.97 | 0.94 | -0.03 | 0.27 | 0.38 | 0.39 | 0.85 | 0.93 | 6965 |
| **Gloria-Winter** | 1.14 | 1.11 | -0.03 | 0.29 | 0.41 | 0.36 | 0.87 | 0.93 | 3577 |

The results presented in Figs. 1 and 2 and in Table 2 show that the computational framework implemented provides in general at a global scale reliable wave predictions.

However, as regards the Romanian nearshore, it has to be highlighted that the comparisons against the in situ measurements carried out at Gloria platform, although still acceptable, appear to be less accurate. An example of direct comparison, SWAN simulations against Gloria measurements, is presented in Fig. 3 from bellow. This figure shows comparison results in terms of maximum wave height ($H_{max}$), corresponding to the time intervals: 1999/12/26-2000/01/05 (a) and 2003/01/29-2003/03/07 (b). According to the results presented in this figure, in the case of the energetic peaks there might be significant differences between the model predictions and the measurements (in general underestimations by the model of the energy peaks) higher than one meter (or some times even higher than two meters) in term of maximum wave height.

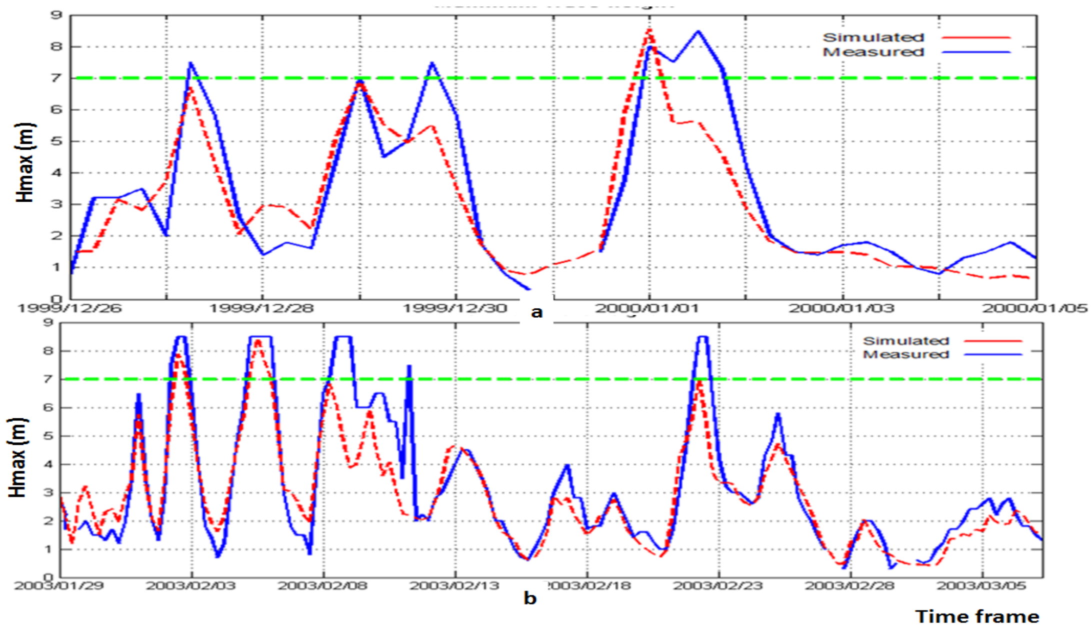



**Fig. 3.**

*$H_{max}$ direct comparisons, SWAN against measurements at the Gloria drilling unit, corresponding to the time intervals: a 1999/12/26-2000/01/05 and b - 2003/01/29-2003/03/07*

From this perspective, it appears very important to design some DA techniques in order to increase the reliability of the wave predictions close to the most sensitive areas, as the Romanian harbors, where the maritime traffic is increased inducing in this way a higher risk of accidents and where a better capacity to predict accurately the wave conditions might by fundamental in avoiding the accidents.

## 3. Data assimilation to improve the wave predictions close to the Romania ports

In order to implement the DA procedure for improving the wave predictions in the Romanian nearshore, a second computational domain (denoted as the coastal level domain) was nested inside the sea level computational domain above described. The characteristics of this new computational domain, as well as the physical processes activated, are presented in Table 3 while the geographical extent of this area together with the location of the data source (the Gloria drilling platform) is illustrated in Fig. 4.

**Table 3**

*SWAN model configuration for the computational domain corresponding to the Romanian nearshore*

<table>
<tr><td rowspan="2">Coastal level SWAN configuration</td><td colspan="2">Coordinates</td><td colspan="2">Δx × Δy (º)</td><td>Δθ x Δt (º)-(min)</td><td>Mode</td><td>nf</td><td>nθ</td><td colspan="4">ngx × ngy = np</td></tr>
<tr><td colspan="2">Spherical</td><td colspan="2">0.005 x 0.005</td><td>10 x 20</td><td>non-stat</td><td>30</td><td>36</td><td colspan="4">221 x 221= 48841</td></tr>
<tr><td rowspan="2">Input / Process</td><td>wave</td><td>wind</td><td>tide</td><td>crt</td><td>gen</td><td>wcap</td><td>quad</td><td>tri</td><td>diff</td><td>bf</td><td>set up</td><td>br</td></tr>
<tr><td>X</td><td>X</td><td>0</td><td>0</td><td>X</td><td>X</td><td>X</td><td>X</td><td>X</td><td>X</td><td>0</td><td>X</td></tr>
</table>

As reflected by Table 3, in comparison with the sea level computational domain, some additional processes, specific to shallow water, were included in the coastal level domain, as triad wave-wave interactions and diffraction.

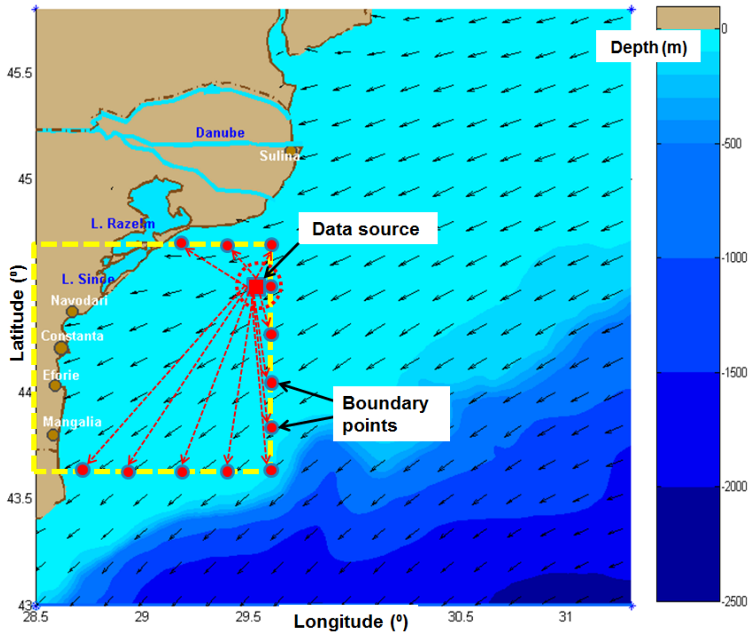



**Fig. 4.**
*Coastal level computational domain and the DA scheme considered for improving the wave predictions in the Romanian nearshore, in the background the bathymetric map is illustrated together with the most common pattern for the wave propagation.*

The DA procedure assumed to be implemented is also suggested in Fig. 4 and considers the use of the in situ measurements performed at the Gloria drilling unit in order to correct the wave predictions in the coastal level computational domain. That is to propagate the correction between the measurements at Gloria and the results of the SWAN simulations corresponding to the same location and time frame in the geographical space and particularly in the boundary points of the higher resolution computational domain that focuses on the Romanian harbors. Thus, in a point *B* (from the boundary) the assimilated value of the significant wave height corresponding to the simulation performed for the time frame *Tf* $\left(H_{sBTf}^{Assim}\right)$ is computed with the following relationship:

$$H_{sBTf}^{Assim} = H_{sBTf}^{Sim} + C_1(B,Tf)\left(\Delta H_{sGTf}\right) \tag{3}$$

$H_{sBTf}^{Sim}$ represents the model predicted value of the significant wave height in the point *B*, $\Delta H_{sGTf}$ is computed as the difference between the measured and the simulated values of $H_s$ at the buoy location and the correction $C_1(B,Tf)$ was defined as:

$$C_1(B,Tf) = \frac{H_{sBTf}^{Sim}}{H_{sGTf}^{Sim}}. \tag{4}$$

After performing corrections of the significant wave heights in the points *(B)*, that control the variable boundary conditions of the coastal level domain, the next step is to transfer this correction in the spectral space. In the SWAN model, the 2D wave energy density spectrum corresponding to a certain simulation ( $E_{BTf}^{Sim}$ ) is discretized in a matrix form as follows:

$$E_{BTf}^{Sim} = \left[E_{BTfnm}^{Sim}\right], \tag{5}$$

the index *n* (*n* = *1*, *nf*, with *nf* the number of frequencies) defines the distribution of the wave energy along the frequencies and the index *m* (*m* = 1,*nd*, with *nd* the number of directions) defines the distribution of the wave energy along the wave directions. Thus, $E_{BTfnm}^{Sim}$ represents the wave energy density (or variance density in m$^2$/Hz/degr) corresponding to the frequency number *n* and to the direction number *m* (as defined in the model settings) for the point and to the simulation performed for the time frame *Tf*. The spectral matrix can be written in a normalized form as follows:

$$E_{BTf}^{Sim} = E_{BTf\,\max}^{Sim}\left[E_{BTfNnm}^{Sim}\right] \tag{6}$$

In order to propagate the $H_s$ correction in the spectral space the following approach is considered:

$$E_{BTf}^{Assim} = C_2(B,Tf)\, E_{BTf\,\max}^{Sim} \left\lfloor E_{BTfNnm}^{Sim} \right\rfloor \tag{7}$$

The spectral correction $C_2(B,Tf)$ is defined as:

$$C_2(B,Tf) = \left( \frac{H_{sBTf}^{Assim}}{H_{sBTf}^{Sim}} \right)^2 , \tag{8}$$

Following this approach, the correction of the significant wave height is propagated in the spectral space keeping as invariant the normalized spectral matrix before defined. This means that while the significant wave height is corrected, the shape of the spectrum is kept unchanged. Some results in relationship with the changes induced in the wave predictions by this DA scheme are discussed in the next section.

## 4. Results and discussions

Several situations have been considered and analyzed and two relevant case studies will be presented and discussed next. The first reflects the patterns of a regular storm for the Black Sea basin and the situation when the model underestimates the measurements at the Gloria drilling platform. The second case study that will be presented is related to some average wintertime non storm conditions, and in this particular case the model results over estimate the measurements at the Gloria drilling platform.

Thus, the first case study is illustrated in Fig. 5. This figure presents the SWAN results for the simulation corresponding to the time frame 2002/02/18 that reflects regular storm conditions. In the upper panel the significant wave height fields and wave vectors for the entire Black Sea level are presented. For this time frame, the results of the simulation at the location of the Gloria drilling platform indicate a *Hs* value of 5.64m while the measurements indicate for the same parameter a value of 6.72m, which means a model under evaluation with 1.08m. The lower panels present the results of the SWAN simulations corresponding to the coastal level, to the left side the model results without DA, while to the right side the model results with DA.

The second case study is illustrated in Fig. 6. This figure presents the SWAN results for the simulation corresponding to the time frame 2002/03/24 that reflects regular winter average non storm conditions. In the upper panel the significant wave height fields and wave vectors for the entire Black Sea level are presented. For this time frame, the results of the simulation at the location of the Gloria drilling platform indicate a *Hs* value of 3.63m while the measurements indicate for the same parameter a value of 2.61m, which means in this case a model over evaluation with 1.02. The lower panels of the figure present the results of the SWAN simulations corresponding to the coastal level, to the left side the model results without DA, while to the right side the model results with DA.

From the analysis of the significant wave height fields presented in Figs. 5 and 6 some conclusions also arise. Thus, although in both cases presented the absolute value of the difference between simulation and measurement is around the same value (a little bit higher than one meter), the propagation of the DA scheme in the geographical space is considerably higher in the second case when the direction from where the waves are coming at the Gloria drilling platform is 65 degrees (in Nautical convention, which means that the waves are measured from the North) in comparison with the first case study when the wave direction at Gloria is about 20 degrees.

Since in most of the cases the dominant wave direction in the Gloria sector is from East-North East, it results that the correction of the DA scheme proposed herewith may have a relevant impact in the computational domain defined for the coastal level simulations.

## 5. Conclusions

The wave measurements from the Gloria drilling platform have been considered in the present work in order to implement a DA procedure. A scheme based on the successive corrections method has been implemented to perform the assimilation in the geographical and spectral spaces. In this way, the model predictions in the Romanian nearshore are improved and the wave modeling system developed can provide better nowcast products concerning the wave conditions in the vicinity of the major Romanian harbors.

Moreover, such DA scheme can be also considered for providing forecast products. Rusu (2014) developed another DA procedure, based on the linear regressions, applied to perform corrections in time to the model predictions at the location of the data source (the Gloria drilling unit) in order to fit better the observations. The results presented show that an improvement in the model prediction is obtained for both the two wave parameters considered ($H_s$ and $T_m$).

The statistical analysis shows that better results were obtained for a training period higher than 20 days, especially for $H_s$. By combining the two procedures, a complete DA scheme can be accomplished to perform the assimilation at three levels: time, geographical and spectral spaces, which means an improvement for both nowcast and forecast products.

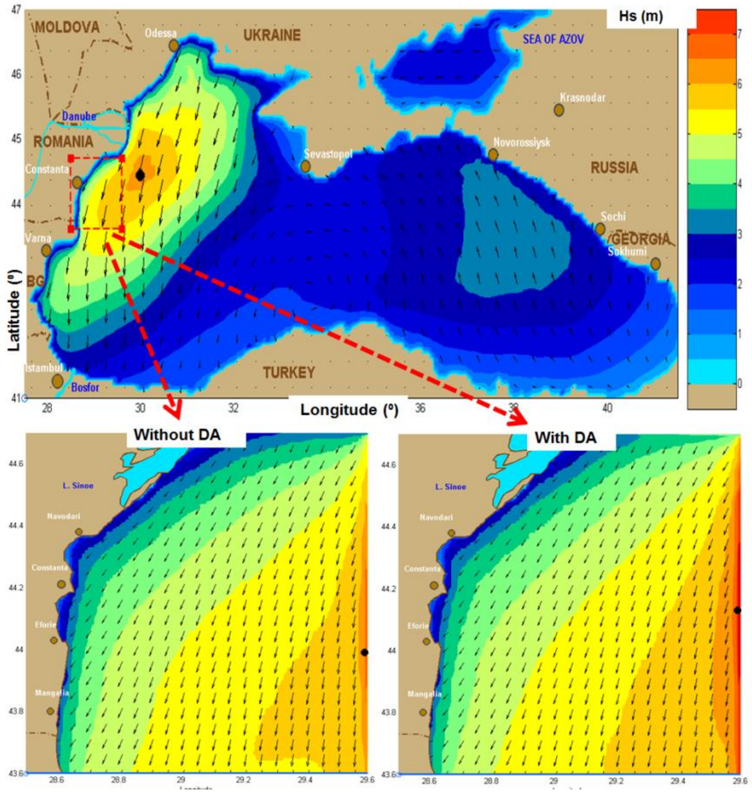



**Fig. 5.**
*Results for the SWAN simulation corresponding to the time frame 2002/02/18 – storm conditions, significant wave height fields and wave vectors. Upper panel, the Black Sea level, lower panels coastal level, left side model results without DA, right side model results with DA.*

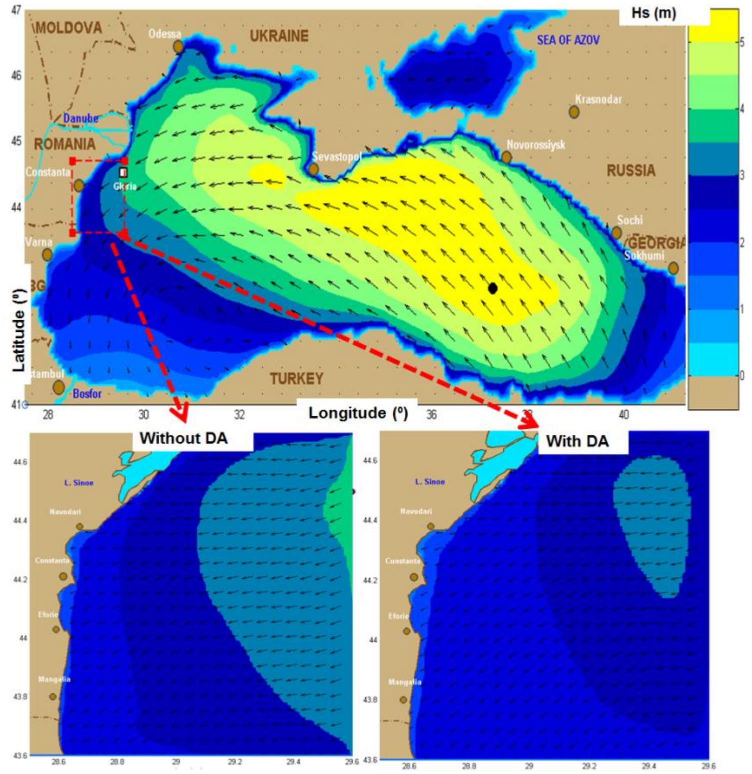



**Fig. 6.**
*Results for the SWAN simulation corresponding to the time frame 2002/03/24 – non storm conditions, significant wave height fields and wave vectors. Upper panel, the Black Sea level, lower panels coastal level, left side model results without DA, right side model results with DA.*

Finally, the comparisons performed between the SWAN results and the in situ measurements carried out at the Gloria drilling unit show that the greatest differences occur in general in high storm conditions, and exactly in such situations it is more important to provide reliable wave predictions, because they usually represent the main source of navigation risks and coastal hazards. From this perspective, the DA scheme implemented herewith can provide a computational framework that could increase the safety of both navigation and harbor operations in the Romanian coastal environment of the Black Sea.

**Acknowledgements**

This work was supported by a grant from the Ministry of National Education, CNCS – UEFISCDI, project number PN-II-ID-PCE-2012-4-0089 (project DAMWAVE). The first author also acknowledges the receipt of a postdoctoral fellowship in the framework of the SOP HRD project PERFORM (ID 138963).

**References**

Ardhuin, F.; Bertotti, L.; Bidlot, J.R; Cavaleri, L.; Filipetto, V.; Lefevre, J. M.; Wittmann P. 2007. Comparison of Wind and Wave Measurements and Models in the Western Mediterranean Sea. *Ocean Eng*, **34** (3–4): 526-539.

Booij, N.; Ris, R. C.; Holthuijsen, L. H. 1999. A Third Generation Wave Model for Coastal Regions. Part 1 - Model Description and Validation. *J Geophys Res*, 104, 7649-7666.

Butunoiu, D.; Rusu, E. 2012. Sensitivity Tests with two Coastal Models. J Environ Prot Ecol, 13 (3), 1332-1349.

Gasparotti, C.; Rusu, E. 2012. Methods for the Risk Assessment in Maritime Transportation in the Black Sea Basin, *J Environ Prot Ecol*, 13 (3A): 1751-1759.

Ivan, A.; Gasparotti, C.; Rusu E. 2012. Influence of the Interactions between Waves and Currents on the Navigation at the Entrance of the Danube Delta. *J Environ Prot Ecol*, **13** (3A), 1673-1682 (2012).

Makarynskyy, O.; Makarynska, D.; Rusu, E.; Gavrilov, A. 2005. Filling Gaps in Wave Records With Artificial Neural Networks, *Maritime Transportation and Exploitation of Ocean and Coastal Resources*, Guedes Soares, C., Garbatov, Y. and Fonseca, N. (Eds.),Taylor & Francis Group, London, 1085-1091.

Maslova, V.; Voskresenskaya, E.; Bardin, M. 2010. Variability of the Cyclone Activity in the Mediterranean-Black Sea Region. *J Environ Prot Ecol*, 11 (4): 1366-1372 (2010).

Mateescu, R.; Coman, C. 2008. Evolution of the Romanian Black Sea Shore under the Actual Anthropogenic Impact. *J Environ Prot Ecol*, 9 (2): 357-365.

Onea, F.; Rusu, E. 2014. Evaluation of the Wind Energy in the North-west of the Black Sea. *Int J Green Energy*, 11 (5): 465-487.

Rusu, E.; Ventura Soares, C.; Rusu L. 2006. Computational Strategies and Visualization Techniques for the Waves Modelling in the Portuguese Nearshore, *Maritime Transportation and Exploitation of Ocean and Coastal Resources* (Eds C. G. Soares, Y. Garbatov, N. Fonseca). Taylor & Francis Group, London, Vol II, 2006, 1129-1136.

Rusu E. 2011. Strategies in Using Numerical Wave Models in Ocean/Coastal Applications. *J Mar Sci Technol* - Taiwan, 19 (1), 58-73.

Rusu, L. 2010. Application of numerical models to evaluate oil spills propagation in the coastal environment of the Black Sea. *Journal of Environmental Engineering and Landscape Management* 18 (4), 288-295.

Rusu, L., Ivan, A., 2010. Modelling Wind Waves in the Romanian Coastal Environment. *Environmental Engineering and Management Journal* 9(4), 547-552.

Rusu, L. 2014. A data assimilation scheme to improve the wave predictions in the western side of the Black Sea, 14$^{th}$ International Multidisciplinary Scientific GeoConference SGEM, Section Marine and Ocean Ecosystems, 16-20 June 2014, Albena Bulgaria.

SWAN team, 2013. *Scientific and technical documentation*, SWAN Cycle III version 40.91. Delft University of Technology, Department of Civil Engineering, the Netherlands

# TEMPORAL DATA WEREHOUSE APPROACH FOR ACCIDENT MONITORING IN CROSSROADS

**Georgia Garani[1], George K. Adam[2], Dimitrios Ventzas[3]**
*[1,2,3]Department of Computer Science and Engineering, Technological Educational Institute of Thessaly, Larissa, Greece*

**Abstract:** Nowadays, methodologies and techniques used for road traffic data processing and analysis are continuously evolving with specific statistical analysis methods being proposed for this purpose. An example of this is crossroad monitoring, where the increasing complexity of human factors, as well as the diversity of technological aspects, affect the road control system and require suitable tools for its management. Furthermore, increased traffic is observed at crossroads where accidents constitute a major problem. Conclusively, the recording and processing of a vast amount of data relevant to accident occurrences in crossroads is of high priority, as it could assist in minimizing such occurrences by improving our understanding of and ability to confront their causes.

Data warehouses are collections of data specifically designed for the support of decision making by data querying, reporting and analysis. A data warehouse supporting historical data is called a temporal data warehouse.

In this paper a temporal data warehouse model is proposed and evaluated for accident monitoring in crossroads. The crossroads' accident data warehouse is logically modeled using a temporal starnest schema which forms an integration of the star schema and the snowflake schema. The expressive power of the model is validated by a number of temporal SQL queries in Oracle 11g. Evaluation results have shown that the proposed approach is promising compared to the star schema and the snowflake schema.

**Keywords:** Crossroad, Road Traffic Data, Data Warehouse

## 1. Introduction

The increased traffic at crossroads causes a number of problems such as traffic delays, transport chaos, high levels of stress, and inevitable, frequent accidents. A number of different researchers, such as engineers, data analysts, traffic controllers and traffic police officers, deal with this kind of problems. Hence, there is a strong need to gather transport related information for processing and statistical analysis confronting the problems caused at crossroads. Data is collected and then processed for modeling, retrieving and querying. Traffic data consists of large and complex collection of data sets and therefore belongs to the 'big data' category. Traditional relational database management systems cannot efficiently manage and process big data.

Data warehouses (DWs) are used for decision support by querying, reporting and analyzing huge amount of data. They store current and historical data. Stored data cannot change. Therefore, insertions, deletions and updates do not take place in a DW. Data is used only for querying and consequently, querying performance is essential to be as high as possible.

DWs are thus the perfect solution for processing traffic data and extracting valuable information regarding the causes of the various problems traffic generates. Particularly for monitoring accidents taking places in crossroads, the need for data analysis is even more essential, as reducing the number or severity of such incidents would decrease the number of injuries and casualties. DWs designed for monitoring accidents in crossroads also need to support time dependent data, i.e., historical data. Hence, temporal DWs are used for this purpose.

Temporal DWs have been developed the last decade for describing information changing over time. Temporal DWs are concerned mainly with aggregation in connection with time-varying data and temporal queries, which has been improved through knowledge gained from thorough research on temporal databases. Temporal databases represent and manage information varying over time (Malinowski & Zimányi, 2008). Two time dimensions are considered in temporal databases, valid-time and transaction-time. Valid-time expresses the time when a fact is true in the modeled reality. Transaction-time represents the time when a fact is registered in the database. Temporal databases are thus divided in three categories, valid time databases, transaction time databases and bitemporal databases. A bitemporal database supports both time dimensions. Depending on whether or not the first normal form property is satisfied, databases are also divided into two different categories, normalized databases and unnormalized or nested databases. Normalized databases are the databases where all relations contained are at least in first normal form. On the contrary, nested databases contain at least one relation not in first normal form, i.e., containing subrelations.

Temporal and nested databases can also be combined and applied in a similar way in the field of DWs. In traditional DWs a DW schema consists of two types of objects, fact tables and dimensions tables. A fact table stores business measurements. It contains measures which are quantitative descriptions of the fact and foreign keys referencing to dimension tables. A dimension table contains a primary key and dimensional attributes, i.e. a set of descriptive attributes. Dimensional attributes are expressed by hierarchies. A hierarchy is described by a set of attributes linked by many-to-one associations (Golfarelli & Rizzi, 2009a). For example, dimensional attributes City, State, Country express a hierarchy represented by the functional dependencies City → State and State→ Country.

The logical modeling of a DW is mainly based on two different types of relational schemata, the star schema and the snowflake schema. A star schema consists of a fact table and a set of dimension tables.

Each dimension table corresponds to one dimension. In snowflake schema some dimension tables are split up into smaller normalized tables.
The starnest schema is a new schema type proposed in Garani & Helmer (2012). The starnest schema is based on the nested approach, where hierarchies are represented as nested tables. The starnest schema is extended in Garani, Adam & Ventzas (2014) to support time. Time in the temporal starnest schema is not treated as another dimension but as time attributes in every temporal dimension, i.e., dimension tables dependent on time.
This paper deals with the implementation of a temporal DW for accident monitoring in crossroads. The implementation is carried out using Oracle 11g. Two different approaches are used, the star schema approach and the starnest schema approach for evaluation purposes. Preliminary results of query execution were promising for the starnest approach because the number of tuples accessed for query retrieval is reduced due to temporal nested dimension tables.
This paper is organized as follows. Section 2 presents related work that has been undertaken by other researchers. In Section 3 DW modeling is discussed. Section 4 presents the temporal DW for crossroad accidents. Implementation is discussed in Section 5. Section 6 concludes the paper.

## 2. Related work

A number of researchers have made proposals in the field of temporal DWs. A brief presentation of related research work is given below in chronological order.
Kimball (1996) introduced the term 'slowly changing dimensions' for dimensions used for storing slowly changing historical data. He proposed three different techniques for dealing with attributes changing over time, either by overwriting the value, adding a dimension row, or adding a dimension column and a number of hybrid methods. However, schema evolution and dimension updates are not generally considered.
A number of requirements that a multidimensional model must fulfill are presented in Pedersen, Jensen & Dyreson (2001) where an extended multidimensional data model is also developed. The proposed model supports all features implied by these requirements. The temporal support of the extended multidimensional data model can manage changes over time. Two time operators are defined in the corresponding algebra, valid-timeslice operator and transaction-timeslice operator. The two operators return the part of the multidimensional object that satisfies the corresponding operator and transform the temporal type of the object to transaction-time or valid-time respectively.
A bitemporal DW model is proposed in Koncilia (2003). The model is an extended version of the COMET metamodel (Eder, Koncilia & Morzy, 2002) which supports valid time and transaction time both at instance and schema levels. The model allows all possible changes of schema and structure of a DW with the introduction of suitable transformation functions.
In Rechy-Ramírez & Edgard (2006) a bitemporal multidimensional schema evolution DW model is presented based on multidimensional database versions with the same valid time but different transaction times. Sixteen schema evolution operators for dimensions and cubes are associated with the proposed model. A SQL-like language for the model is suggested. It is capable of creating and modifying multidimensional versions.
The conceptual multidimensional MultiDimER model is temporally extended in Malinowski & Zimanyi (2006) for supporting valid time and transaction time. DW loading time is also supported. The model is suitable for representing time-varying levels, attributes and hierarchies. It distinguishes time variant elements from time invariant elements and treats them separately.
An enhanced solution is introduced in Nguyen, et al (2007) namely event-fed comprehensive slowly changing dimension approach, which combines the two of the three techniques proposed in Kimball (1996). This is performed by creating a new table, the TRANS table, consisting of a new record for each dimensional attribute change. The authors claim that the proposed approach reduces the number of records to be processed compared to the snapshot-based approach.
The modeling and querying of a graph based temporal semi-structured DW is proposed in Combi, Oliboni & Pozzi (2009), together with a suitable query language which stores data in a similar way as an XML document does.
A survey of issues associated to temporal data warehousing is presented in Golfarelli and Rizzi (2009b) distinguishing three different topics, handling changes in the DW, handling data changes in the data mart and handling schema changes in the data mart.
A semi-structured bitemporal model is designed in Aslan (2013). The model supports nested tables and nested types. Multiple attributes (a value and the corresponding temporal values) are stored in one column as a complex structure. This includes for instance valid time lower and upper bound values as well as transaction time lower and upper bound values. Several queries were tested in Oracle. It is shown that less disk space is used and performance is improved using this approach compared to the equivalent relational database in SQL Server.

## 3. Data warehouse modeling

### 3.1 Overview of the temporal starnest schema

The temporal starnest schema (Garani, Adam & Ventzas, 2014) consists of a temporal fact table and a number of dimension tables. Only valid time is supported.

The fact table consists of the composition of the foreign keys referencing the dimension tables, quantitative attributes as measures and one or two valid time attributes depending on the duration of the fact it is described, one valid time attribute for a time point and two valid time attributes for a time interval. The time dimension is not included in the temporal starnest schema; accordingly, join operations between the fact and the time dimension table are avoided. Dimension tables dependent on time are called temporal dimension tables. Temporal dimension tables are also timestamped. Two valid time attributes, the start and stop time points of the corresponding time interval, are included in each temporal dimension. For temporal dimension tables referring to objects with a time point duration, only one valid time attribute is required. Hierarchies in dimension tables are presented as nested tables and thus, data are clustered naturally.

A temporal nested dimension table consists of hierarchical attributes, dimensional attributes and time attributes which can also be nested inside less detailed attributes.

Therefore, the scheme of a (valid time) temporal nested dimension table is defined as follows:

**Definition 1 (*Scheme of a valid time temporal nested dimension, $D_{tm}$*)**

The scheme of a (valid time) temporal nested dimension, $D_{tm}$, with $h_m$, $h_{m-1}$, …, $h_m$, …, $h_1$ hierarchical attributes ($h_m$ being the most aggregated level and $h_1$ the most detailed one, i.e. $h_1 \rightarrow h_2 \rightarrow \dots h_m$), $d_{tm}^{1}$, $d_{tm}^{2}$, …, $d_{tm}^{c(m)}$ temporal dimensional attributes at level m (where c(m) the cardinality of $d_{tm}^{c(m)}$) and $d_{t1START}$, $d_{t1STOP}$, …, $d_{tmSTART}$, $d_{tmSTOP}$ valid time start and valid time stop time attributes of $D_{t1}$, …, $D_{tm}$ respectively.

The tree representation of a (valid time) temporal nested dimension is shown in Figure 1.

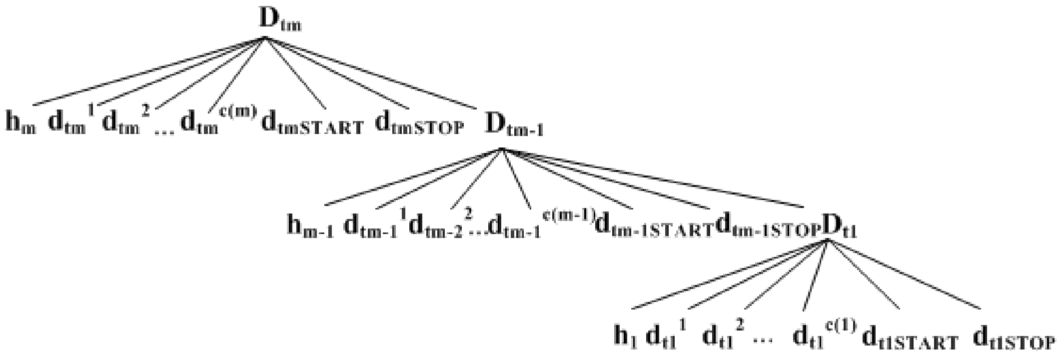



**Fig. 1.** *Tree representation of a (valid time) temporal nested dimension*

Transitive functional dependencies are not present in the temporal starnest schema. 1:N relationships link the fact table and the nested dimension tables. Each relationship is materialized by a foreign key attribute of the fact table associated to one of the most detailed hierarchical attribute of a nested dimension.

## 4. The crossroads' accident temporal data warehouse

### 4.1 Temporal data warehouse starnest schema

The crossroads' accident DW is logically modeled using a temporal star schema and subsequently, the same DW is converted to a temporal starnest schema. The schemas presented below are those implemented in Oracle 11g.

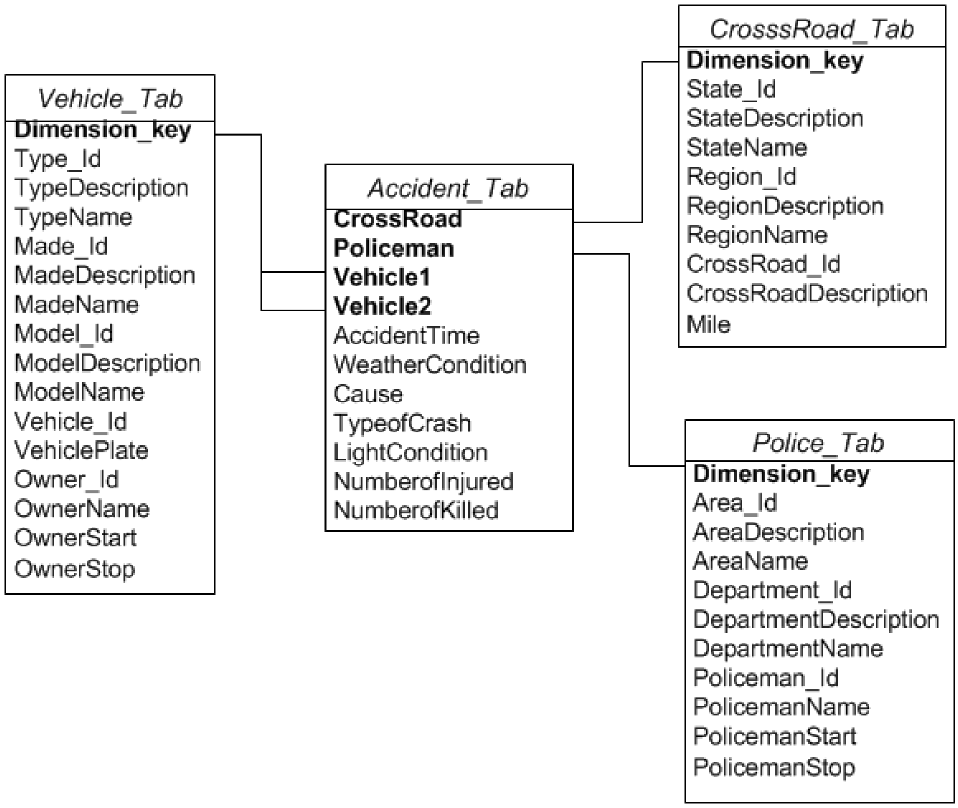



**Fig. 2.** *Crossroads' accident temporal star schema*

The Crossroads' accident temporal star schema consists of four flat tables, the Accident fact table and three dimension tables, CrossRoad, Police and Vehicle. The schema is presented in Figure 2. The Accident fact table consists of the composition of the foreign keys referencing the dimension tables, one time attribute, AccidentTime which is a time point representing valid time, i.e., the time when the accident took place, two quantitative descriptions of an accident, NumberofInjured and NumberofKilled as measures and WeatherCondition, Cause, TypeofCrash, LightCondition attributes. In Oracle, the relation scheme of the Accident fact table is Accident_Tab(CrossRoad, Policeman, Vehicle1, Vehicle2, AccidentTime, NumberofInjured, NumberofKilled, WeatherCondition, Cause, TypeofCrash, LightCondition). The Vehicle dimension exists twice in the Accident fact table, Vehicle1 and Vehicle2, as foreign keys referencing the one dimension table that models two different vehicles involved in an accident. Thus, the Crossroads' accident star schema contains a shared hierarchy since a hierarchy is replicated twice in the fact schema (Golfarelli & Rizzi, 2009a). The existence of a shared hierarchy is a good design solution because it omits the duplication of the same dimension table at the logical level.

The dimension tables CrossRoad, Police and Vehicle are not in 3NF as transitive functional dependencies exist given that attributes of a hierarchy occur in the same relation. Dimension tables Police and Vehicle are temporal dimensions as they contain two valid time attributes, i.e. representing the start and end time points of the corresponding time interval. In Figure 2 the attributes of the dimension tables are also given.

The crossroads' accident temporal starnest schema shown in Figure 3 consists of a flat fact table and three nested dimension tables, the fact table Accident and the nested dimension tables CrossRoad, Police and Vehicle since all dimension tables have hierarchies. Police and Vehicle are modeled as temporal nested tables where valid time is expressed as two temporal attributes representing the start and end time points of the corresponding time interval, similar to the star schema approach. CrossRoad and Police tables have two nesting levels and Vehicle table three nesting levels. The nesting levels represent the hierarchies of each dimension.

The representation of dimension table Police as a temporal nested dimension table is Police_Nest(Dimension_key, Area_Id, AreaName, *Department(Department_Id, DepartmentName, *Policeman(Policeman_Id, PolicemanName, PolicemanStart, PolicemanStop))).

The representation of dimension table Vehicle as a temporal nested dimension table is Vehicle_Nest(Dimension_key, Type_Id, TypeName, *Made(Made_Id, MadeName, *Model(Model_Id, ModelName, *Vehicle(Vehicle_Id, VehiclePlate, Owner_Id, OwnerName, OwnerStart, OwnerStop)))).

The representation of dimension table CrossRoad as a nested dimension table is CrossRoad_Nest(Dimension_key, State_Id, StateName, *Region(Region_Id, RegionName, *CrossRoad(CrossRoad_Id, Mile))).

A name of an attribute with the prefix '*' denotes that this is a nested attribute, i.e. a subrelation.

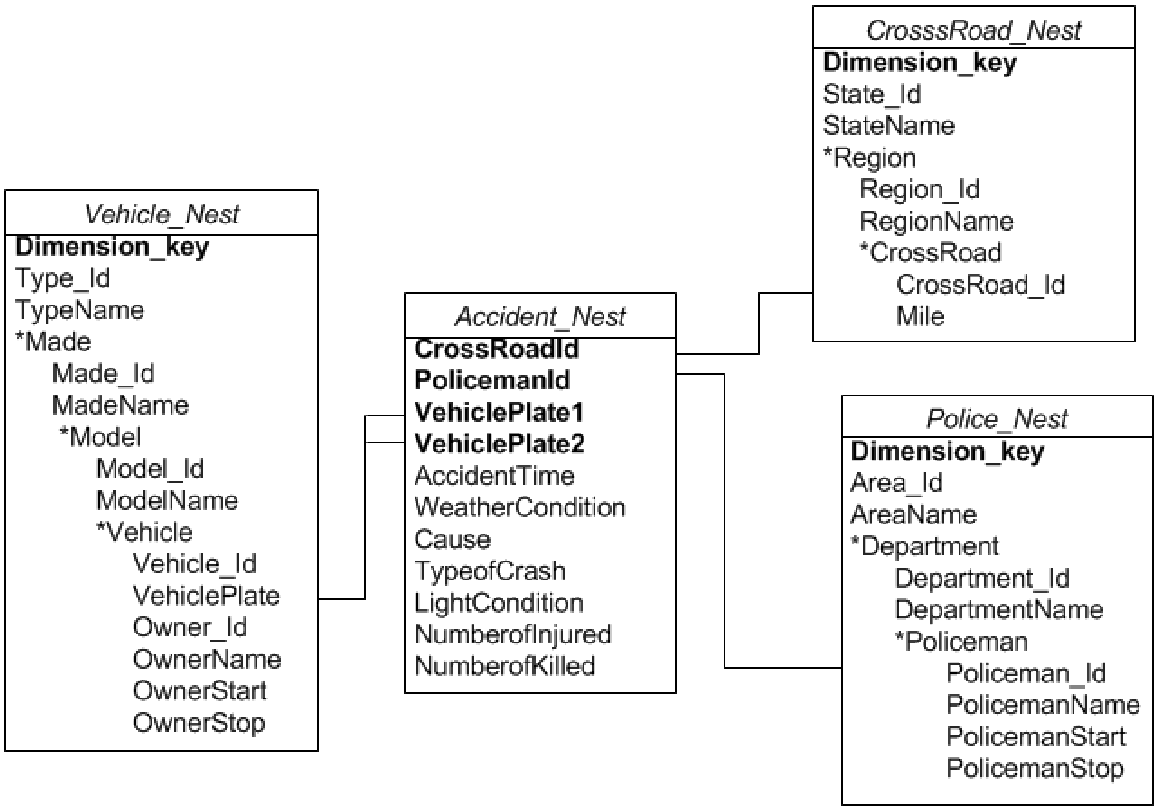



**Fig. 3.** *Crossroads' accident temporal starnest schema*

## 5. Implementation

Both approaches, the star schema approach and the starnest schema approach, have been implemented for the same data and queries. Results and run time are compared and evaluated. All queries have been executed on an Intel(R) Core(TM) i5 processor, running at 2.3 GHz, with 2.5 GB ram memory, under Windows 7 (32bit). The DW was built in Oracle Data Warehouse builder 11.2.0.1 and Oracle SQL Developer 4.0.3 was used.

The Accident cube table contains 20000 rows in both approaches. Although Police and CrossRoad dimension tables contain 20000 rows in the star schema approach and the Vehicle dimension table 40000 rows, in the starnest schema approach Police dimension contains 10 rows, CrossRoad dimension contains 14 rows and Vehicle dimension 5 rows. This is because in the starnest schema approach subrelations are included in each dimension and thus, in a more nesting level more rows are included. When the dimension is unnested, it must contain the same number of rows as the corresponding dimension in the star schema approach. Therefore, while the star schema approach occupies 16.3 MB disk space, the starnest schema approach uses only 3.2 MB.

In what follows six queries are written in SQL for both approaches. Each query is executed ten times. For each query a chart demonstrates the run times in milliseconds and the last measure shows the average run time.

**Query 1**

Who are the owners of the vehicles involved to accidents where more than two people were killed?

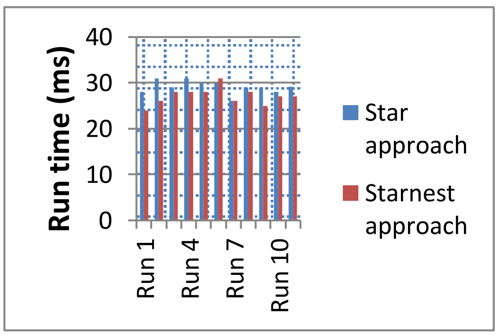
**Star schema approach**

```
SELECT v.OwnerName
FROM  Accident_Tab a, Vehicle_Tab v
WHERE a.NumberofKilled > 2
AND (v.Dimension_key = a.Vehicle1 OR v.Dimension_key = a.Vehicle2)
AND a.AccidentTime BETWEEN  v.OwnerStart AND v.OwnerStop;
```

**Starnest schema approach**

```
SELECT s.OwnerName
FROM Accident_Nest a, Vehicle_Nest v, table(Made) u, table(u.Model)  f,
table(f.Vehicle) s
WHERE a.NumberofKilled > 2
AND (s.VehiclePlate = a.VehiclePlate1 OR s.VehiclePlate = a.VehiclePlate2)
AND a.AccidentTime BETWEEN  s.OwnerStart AND s.OwnerStop;
```

**Table 1** *Query 1 run comparison chart*



**Query 2**

How many accidents for the vehicle type Van happened at the 46th mile of region Texas at December 2010?

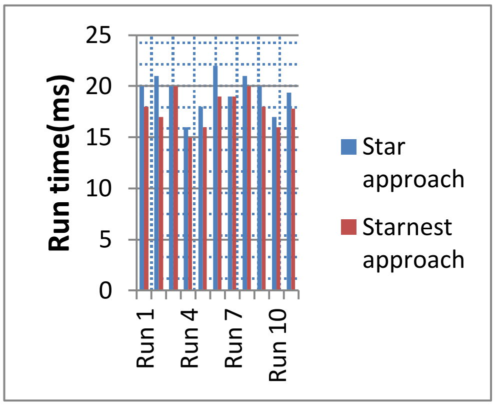
**Star schema approach**

```
SELECT COUNT(a.CrossRoad)
FROM   Vehicle_Tab v, Accident_Tab a, CrossRoad_Tab cr
WHERE cr.Mile = 46
AND cr.RegionName = 'Texas'
AND a.CrossRoad = cr.Dimension_key
AND (v.Dimension_key = a.Vehicle1 OR v.Dimension_key = a.Vehicle2)
AND a.AccidentTime BETWEEN  v.OwnerStart and v.OwnerStop
AND v.TypeName='Van'
AND a.AccidentTime BETWEEN  '01/12/10' AND  '31/12/10';
```

**Starnest schema approach**

```
SELECT COUNT(a.CrossRoadId)
FROM Accident_Nest a, CrossRoad_Nest cr, table(Region) reg, table(reg.CrossRoad)
creg, Vehicle_Nest v, table(Made) u, table(u.Model) f, table(f.Vehicle) s
WHERE creg.Mile = 46
AND reg.RegionName = 'Texas'
AND a.CrossRoadId = creg.CrossRoad_Id
AND (s.VehiclePlate = a.VehiclePlate1 OR s.VehiclePlate = a.VehiclePlate2)
AND a.AccidentTime BETWEEN  s.OwnerStart and s.OwnerStop
AND  v.TypeName ='Van'
AND a.AccidentTime BETWEEN  '01/12/10' AND  '31/12/10';
```

**Table 2** *Query 2 run comparison chart*




**Query 3**

Who is the policeman who recorded Stewart's accident and when this accident happened?

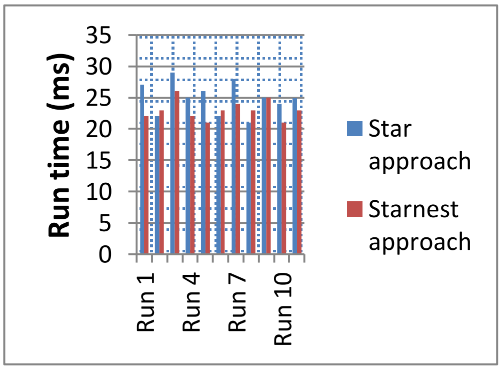
**Star schema approach**

```
SELECT p.PolicemanName, a.AccidentTime
FROM Vehicle_Tab v, Accident_Tab a, Police_Tab p
WHERE v.OwnerName = 'Stewart'
AND (v.Dimension_key = a.Vehicle1 OR v.Dimension_key = a.Vehicle2)
AND a.AccidentTime BETWEEN  v.OwnerStart and v.OwnerStop
AND p.Dimension_key = a.Policeman
AND a.AccidentTime BETWEEN p.PolicemanStart AND p.PolicemanStop;
```

**Starnest schema approach**

```
SELECT q.PolicemanName, a.AccidentTime
FROM Vehicle_Nest v, table(Made) u, table(u.Model)  f, table(f.Vehicle) s,
Accident_Nest a, Police_Nest p, table(Department) d, table(d.Policeman) q
WHERE s.OwnerName = 'Stewart'
AND (s.VehiclePlate = a.VehiclePlate1 OR s.VehiclePlate = a.VehiclePlate2)
AND a.AccidentTime BETWEEN  s.OwnerStart and s.OwnerStop
AND a.PolicemanId = q.Policeman_Id
AND a.AccidentTime BETWEEN  q.PolicemanStart AND q.PolicemanStop;
```

**Table 3** *Query 3 run comparison chart*




**Query 4**

In which region an accident happened and the policeman who recorded the accident was John Smith on 7/7/2012?

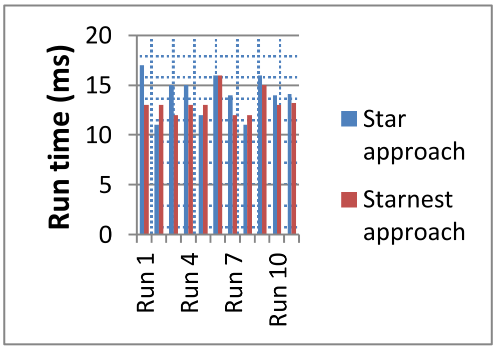
**Star schema approach**

```
SELECT cr.RegionName
FROM Accident_Tab a,  Police_Tab p, CrossRoad_Tab cr
WHERE a.AccidentTime = '07/07/12'
AND p.Dimension_key = a.Policeman
AND  p.PolicemanName = 'John Smith'
AND a.AccidentTime BETWEEN  p.PolicemanStart and p.PolicemanStop
AND cr.Dimension_key = a.CrossRoad;
```

**Starnest schema approach**

```
SELECT reg.RegionName
FROM Accident_Nest a, CrossRoad_Nest cr, table(Region) reg, table(reg.CrossRoad)
creg, Police_Nest p, table(Department) d, table(d.Policeman) q
WHERE a.AccidentTime = '07/07/12'
AND q.PolicemanName = 'John Smith'
AND a.PolicemanId = q.Policeman_Id
AND a.AccidentTime BETWEEN  q.PolicemanStart and q.PolicemanStart
AND a.CrossRoadId = creg.CrossRoad_Id;
```

**Table 4** *Query 4 run comparison chart*



**Query 5**
Find the names of vehicle's owners who were involved in accidents that happened when it was raining.

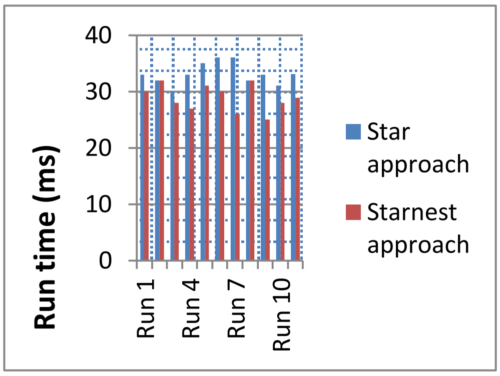
**Star schema approach**

```
SELECT v.OwnerName
FROM Vehicle_Tab v, Accident_Tab a
WHERE a.WeatherCondition = 'Rain'
AND (v.Dimension_key = a.Vehicle1 OR v.Dimension_key = a.Vehicle2)
AND a.AccidentTime BETWEEN  v.OwnerStart and v.OwnerStop;
```

**Starnest schema approach**

```
SELECT s.OwnerName
FROM Vehicle_Nest v, table(Made) u, table(u.Model)  f, table(f.Vehicle) s,
Accident_Nest a
WHERE a.WeatherCondition = 'Rain'
AND (s.VehiclePlate = a.VehiclePlate1 OR s.VehiclePlate = a.VehiclePlate2)
AND a.AccidentTime BETWEEN  s.OwnerStart and s.OwnerStop
```

**Table 5** *Query 5 run comparison chart*




**Query 6**
What is the number of accidents per year which have been recorded from LA DC department?

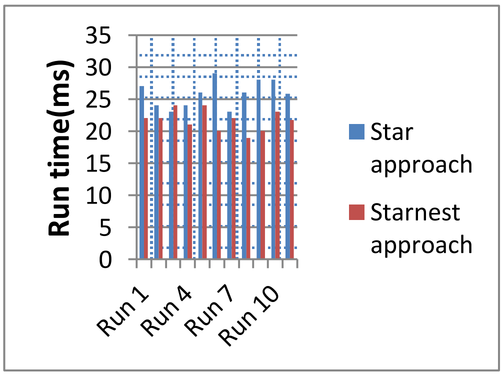
**Star schema approach**

```
SELECT COUNT(year), year
FROM
        (SELECT
        To_CHAR(a.AccidentTime,'YYYY') as year
        FROM  Accident_Tab a, Police_Tab p
        WHERE p.Dimension_key = a.Policeman
        AND p.DepartmentName = 'LA DC'
        AND a.AccidentTime BETWEEN  p.PolicemanStart AND p.PolicemanStop)
group by year
order by year;
```

**Starnest schema approach**

```
SELECT COUNT(year), year
FROM
        (SELECT To_CHAR(a.AccidentTime,'YYYY') as year
        FROM  Accident_Nest a, Police_Nest p,
        table(Department) d, table(d.Policeman) q
        WHERE d.DepartmentName = 'LA DC'
        AND a.PolicemanId = q.Policeman_Id
        AND a.AccidentTime BETWEEN  q.PolicemanStart and q.PolicemanStop)
group by year
order by year;
```

**Table 6** *Query 6 run comparison chart*




The results obtained from the six executed queries lead to the following conclusions: The DW implemented using the starnest schema approach occupies significantly less disc space than the corresponding DW where the star schema approach is used. Generally, queries are faster with the starnest schema approach. Specifically, when the selection condition involves exclusively one or more attributes at the top level of the dimension hierarchy and only a one dimension table (without joining it with the fact table), the execution time of the starnest schema approach queries is reduced remarkably.

## 6. Conclusion

An accident temporal DW is developed in this paper for helping the traffic monitoring in crossroads. The model is implemented in Oracle 11g. Two different approaches have been evaluated and compared, the star schema approach and the starnest schema approach. Results have shown that the starnest schema approach is quite promising compared to the star schema approach, as it occupies significantly less disk space and runs faster.
Future work includes the implementation of optimization techniques for an efficient evaluation of complex temporal nested queries. Also, transaction time can be added to the implemented model. The extension to spatial dimension to support spatial data is another research challenge.

**Acknowledgements**

The research is implemented through the Operational Program "Education and Lifelong Learning" and is co-financed by the European Union (European Social Fund) and Greek national funds, under the Research Project Archimedes III with title "Intelligent Monitoring of Traffic Load and Flow of Vehicular Crossroads with Image and Video Processing Techniques".
The authors would like to thank George Tsaprazlis, undergraduate student from the Department of Computer Science and Engineering of the Technological Educational Institute of Thessaly, who helped with the implementation of this research work during his diploma thesis.

**References**

Aslan, G. 2013. *Modeling and Querying Bitemporal Semistructured Data Warehouses*. Dokuz Eylül University. Graduate school of natural and applied sciences. MSc Thesis. 85p.

Combi, C.; Oliboni, B.; Pozzi, G. 2009. Modeling and Querying Temporal Semistructured Data Warehouses. In Kozielski, S. and Wrembel, R. (Eds.): New Trends in Data Warehousing and Data Analysis, Annals of Information Systems, 3: 299–324, Springer, New York, USA.

Eder, J.; Koncilia, C.; Morzy, T. 2002. The COMET Metamodel for Temporal Data Warehouses. In *Proceedings of the 14$^{th}$ International Conference on Advanced Information Systems Engineering (CAiSE)*, Toronto, Canada, 83-99.

Garani, G.; Adam, G.K.; Ventzas, D. 2014. Temporal Data Warehouse Logical Modelling, *International Journal of Data Mining, Modelling and Management*, accepted for publication.

Garani, G.; Helmer S. 2012. Integrating Star and Snowflake Schemas in Data Warehouses; *International Journal of Data Warehousing and Mining* 8(4): 22-40.

Golfarelli, M.; Rizzi, S. 2009a. *Data Warehouse Design. Modern Principles and Methodologies*. McGraw-Hill Osborne, Italy. 458p.

Golfarelli, M.; Rizzi, S. 2009b. A Survey on Temporal Data Warehousing, *International Journal of Data Warehousing & Mining* 5(1): 1-17.

Kimball, R. 1996. *The Data Warehouse ToolKit*. Wiley Computer Publishing, USA. 374p.

Koncilia, C. 2003. A Bi-Temporal Data Warehouse Model. In *Proceedings of the 15$^{th}$ International Conference on Advanced Information Systems Engineering (CAiSE)*, Klagenfurt, Austria, 77-80.

Malinowski, E.; Zimányi, E. 2006. A Conceptual Solution for Representing Time in Data Warehouse Dimensions. In *Proceedings of the 3rd Asia-Pacific Conference on Conceptual Modeling (APCCM),* Hobbart, Australia, 45-54.

Malinowski, E.; Zimányi, E. 2008. *Advanced Data Warehouse Design. From Conventional to Spatial and Temporal Applications*. Springer, Germany. 444p.

Nguyen, T.M.; Tjoa, A.M.; Nemec, J.; Windisch M. 2007. An approach towards an event-fed solution for slowly changing dimensions in data warehouses with a detailed case study, *Data & Knowledge Engineering* 63(1): 26–43.

Pedersen, T.B.; Jensen, C.S.; Dyreson C.E. 2001. A foundation for capturing and querying complex multidimensional data, *Information Systems* 26(5): 383–423.

Rechy-Ramírez, E.-J.; Edgard, B.-G. 2006. A Model and Language for Bitemporal Schema Versioning in Data Warehouses. In *Proceedings of the 15$^{th}$ International Conference on Computing (CIC),* Mexico City, Mexico, 309-314.

# EFFECTS OF FACILITY LOCATION ON URBAN ROAD TRAFFIC

**Şerban RAICU[1], Vasile DRAGU[2] , Ştefan BURCIU[3], Dorinela COSTESCU[4], Florin RUSCĂ[5]**
[1]*Transport, Traffic and Logistics Department, Transport Faculty, University Politehnica of Bucharest, ROMANIA*

**Abstract:** The aim of the paper was to determine the effects of locating a new commercial/logistic facility on the urban traffic. The location decision together with the one of goods' supply/distribution support the general framework of taking transport decisions: weather to realise the transport or not, choosing the destination, mode, time and frequency of the transport to be realised. Necessity of locating and/or relocating different facilities must foresee the transport flows that they would generate and that would eventually become traffic flows on an already congested urban network. The paper also tries to establish a best practice example in the field of intersection design and traffic management for further developments of new logistic activities. Considering the actual O-D matrix with traffic flows measured on the field we have determined a solution for creating access to a new commercial/logistic centre without creating congestion or important irregularities on a major route of Bucharest road network. The proposed solutions took into consideration traffic parameters like: flow, speed, queue length, delay, density and stop time.

**Keywords**: facility location, traffic flow, planning, simulation, control strategies.

## 1. Introduction

Sustainable development issues, even though often without the aimed results, are frequently brought up as necessities, the solution for satisfying beneficiaries' demands being or having to be adopted in concordance with multiple criterion that integrate divergent goals of those directly and/or indirectly involved. Ignoring the correlation between mobility (transport) and land use brings up major risks in social-economic and natural environment evolution, investment attractivity and, finally, everyone's life quality (Popa, 2004, Macket, 1992).
Freight movement needs (for supply, distribution, construction, etc.) and people's need for mobility (work, education, fun and shopping), determining the so called transport demand, characterized by relation (origin-destination), size and structure are directly connected to land use decisions. These mobility needs of the entire social-economic activities, transport demands, lead successively to transport flows and traffic flows on the transport infrastructure networks (Raicu, 2007). To be able to distinguish the necessity of actions for location/ relocation different activities one must know from the beginning the flows that they would generate, the so called "in" and "out" movements determined by the new locations (transport and traffic flows). The essential role of transportation is to reduce distance and to allow people and economic agents to ignore space. Theoretically, all routes are possible in a continuous space. In reality, defining the transport system characteristics is subject to different restriction:

- time - transport might be seen as a time-space change (long time trips or trades limit the transport demand),
- transport costs, as infrastructures need important financing,
- space, itself, as infrastructure needs space,

that might be surmounted through certain conditions in favour of spatial continuity (cheap and split able infrastructures, high service quality, etc.). Some transport means, generally the individual ones, answer easily spatial continuity condition, while the others need important infrastructures, pushing away from spatial continuity condition (Waller, 1997, Burciu, 2013). Another limit of spatial continuity of the transport system resides in its connection with human activities, as individuals move to activity zones and goods from production to commerce and consumption. Spatial location economy proves that even within a homogenous and isotropic space human activities regroup in hierarchy spaces and the appeared discrepancies lead to city or central zones creation (Allen et al., 2012). Transport flows have origin and/or destination in these human activity concentrations, mainly in the high level ones, making the relation between transport and human activities a direct one.
Transport infrastructures guide the habitat and the facility location. This power of attraction of the different modes of transport infrastructures is defined by the influence zone (for networks) and by the service zone (for economic activities) (Popa, 2004).
Accessibility, as a consequence of land use strategic decisions, leading to attractivity, is determined for the transport demand. Its transformation in transport flow, meaning resources consumption (for the transport means and loading units), is marked by the choosing of the transport operators.
Road traffic congestion has become one of the most irritating problems of the modern world which comes as a result of the increasing in motorization rate, extension of city outskirts, increase of urban population and its mobility needs. While the development of the urban road network is extremely limited both from physical and financial reasons, methods to mitigate congestion in urban agglomerations are mainly aimed to reduce social mobility (in terms of citizen travel needs and freight transfer) and to take actions for the better use of the existing road infrastructure (by reducing spatial and temporal variability of the traffic flows and by improving the traffic flows fluency as a result of the traffic systematization in conflict areas - intersections) (Raicu, 2007).
Too often the link between spatial planning/ urbanism and transports/ traffic is ignored. The socio-economical objective and residential areas location is established by designated authorities without being preceded by a sufficiently detailed study on the accessibility assured by the existent transport infrastructure (Roşca et al., 2013).

[3]Corresponding author: stefanburciu@yahoo.com

These skills (fig. 1) is responsible for the traffic characteristics meaning also for resources consumption for transporters (vehicles, energy, human resources, tolls, insurances), infrastructures' administrators but with negative external effects on the riverans and not only (pollution, stress, etc.).

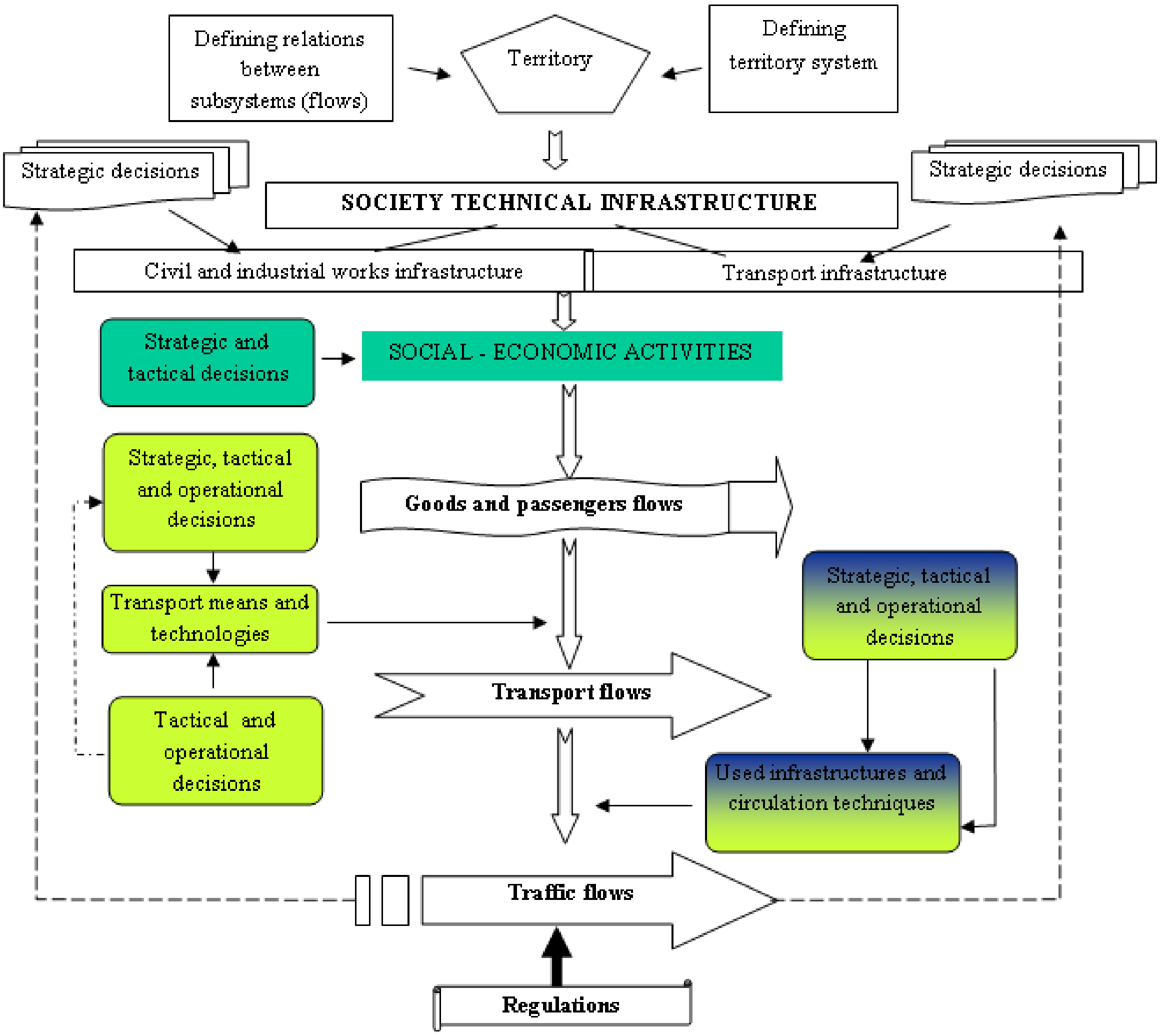



**Fig. 1.**
*Land use and transportation strategic, tactical and operational*
*Source: (adapted from Raicu, 2007)*

Finding and selecting the most appropriate locations for production/ commercial units is a challenge to all the specialists in the field as these solutions might be generated by the so called location-allocation models (Sabater, 2004). These models select optimal locations from a set of possible locations in the most efficient manner and links demand for a certain service to these locations based on the demand distribution. The essential objective is to assure covering of all the demands within an area for a certain trip length or in a standard response time, without taking into consideration the effects on road traffic (Burciu, 2013). The study case within the paper tried to determine these effects and limit their action on the urban road network.

## 2. Study hypothesis

The present paper tries to determine the effect on traffic in a situation where we had to find a proper "ex-post" solution for assuring the access in a large commercial complex. Situated on an access way to A2 highway with seasonal, daily and hourly traffic flows marked by significant irregularities, the mentioned objective is distinguished by superposing the mobility specific to commercial and agreement areas with the one afferent to daily activities of a few hundreds of employees from the office building.
In the study we tried to assure such accessibility to the commercial complex so that its attractiveness would be amplified. In the same time, we tried not to significantly affect the transit traffic from and to the highway. Furthermore, we ameliorated the access conditions for RATB depot, situated in area. That's why, the studied area was extended to other intersections for which traffic control systems was correlated (STAS, 1991).

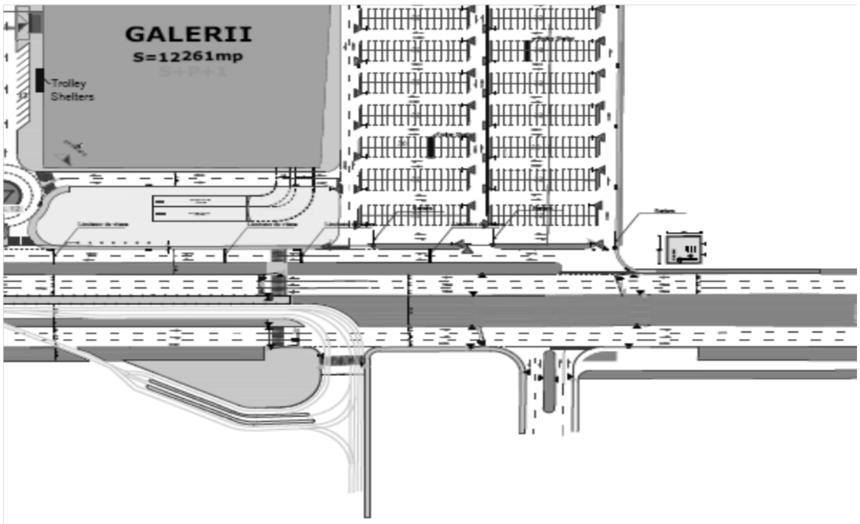



**Fig. 2.**
*Present road network configuration*

The traffic engineer is responsible for assuring the flow's quality parameters (speed, stops' frequency and duration, traffic safety, external effects limitations, energetic consumption and specific costs) and for founding the infrastructure development strategies and traffic control systems. For this purpose, equivalent traffic entities were determined starting from values obtained by traffic surveys.

The area estimated traffic (in equivalent traffic entities) is being presented but we also took into consideration the further traffic demand predicted for the year 2015. The performances of the six configurations proposed for the access were estimated with computer simulation (using AIMSUN Transport Simulation Systems) in five calculation hypotheses as AIMSUN is considered as software which does not need much calibration (Fang et al., 2005). Traffic simulation models might be used in transport planning as they are flexible and feasible in testing a multitude of situations that might happen in the real-world (Lieberman et al., 2004, Manstetten et al., 2005, Nourinejad et al., 2014).

In congestion conditions, no matter what the infrastructure's construction standards are, the average speed won't exceed and won't even approach 50 km/h. In other words, the traffic is self-regulating according to the speed imposed by the safety conditions. All the drivers are constrained to move in column, with speeds under 50 km/h. In this way, congestion makes the maximum allowed speed impossible to overcome, especially within the city.

This was actually the case in our study and so we had to design an intersection starting from a prospective of their classifications. Depending on the nature of the flows that intersect, modifying the character of the traffic on every link of involved infrastructures there are (Popa et al., 1999, Salter, 1989, Wagner, 2010):

- homogenous intersections, when the flows belong to one transport mode;
- non-homogenous intersections, when the flows belong to distinct transport modes.

In relation with the number of infrastructure ways that intersect or merge, there are:

- simple intersections, in which only two ways are involved;
- complex intersections, when three or more circulation ways are intersecting.

Taking into consideration the order in which the transport means (traffic entities) occupy the intersection area, there are:

- no priority intersections, when the first arrived occupies the intersection;
- priority intersections, when traffic entities from a certain way cross the intersection without being retained, meaning that those traffic entities have priority in relation to the ones from the other ways.

The priority can be established in relation to the character of traffic entities. We can emphasize absolute priorities - the priority traffic entities can't be retained under any circumstance and relative priorities – the non-priority demands, already arrived, that occupy the intersection in the moment of priority demand arrivals, can produce delays.

According to the traffic character on the areas that converge into the intersection, there are:

- rigid schedule intersections, the order of occupying the intersection is determined in advanced for a certain time horizon;
- non-rigid schedule intersections, the order of occupying the intersection is established with the arrival of the traffic entities.

When traffic flows intersect or merge, there are retentions (additional stationery or speed reduction) of the traffic entities. On rail and road infrastructures, when the access ways are not very solicited, retaining the traffic entities leads to stop or speed reduction in the intersection area. At higher solicitations, the waiting queues determine the traffic character on large areas of the infrastructure. That's why, the constructive and organizational solutions for solving the traffic flows intersection problem must take into consideration the consequences of retaining the traffic entities, including the influence on the maximum debit (capacity) of the convergent infrastructures.

So, we came up to the solution of designing a non-homogenous, complex and priority intersection.

Given the circumstances in the field we also had to create a non-rigid schedule intersection that would have a signalling cycle different in the five time periods considered, correlated to the other intersections in the area and that would also ease the access to the city for RATB vehicles. Pedestrians were also taken into consideration.

### 3. Junction configurations and characteristics

The configurations that were taken into account:

- dedicated lanes intersection with traffic lights;
- “roundabout” with give way signs and two direct lanes to Constanta;
- roundabout and traffic signalised ”roundabout”;
- island junction with traffic lights (one lane for the commercial centre turn);
- two turn left lanes with traffic lights (two lanes for the commercial centre turn).

These configurations have been studied taking into consideration some hypothesis regarding the traffic during the day: peaks during working day and also on weekend.
The traffic flow, in equivalent traffic entities, is presented in table 1. The numbers from 422 to 426 represent zone centroids, the ones responsible for the traffic (422 – Bucharest, 423, 424 – the commercial centre, 425 – RATB depot and 426 – Constanta exit).

**Table 1.**
*O-D matrix (evening peak, working day)*

| Centroid | 422 | 423 | 424 | 425 | 426 |
|---|---|---|---|---|---|
| **422** | 0 | 0 | 600 | 20 | 480 |
| **425** | 0 | 0 | 0 | 0 | 20 |
| **426** | 680 | 10 | 360 | 0 | 0 |

The figures from 3 to 7 present the six configurations of the proposed junction:

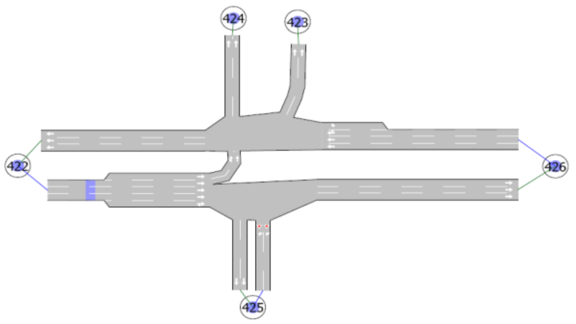



**Fig. 3.**
*Dedicated lanes intersection with traffic lights*

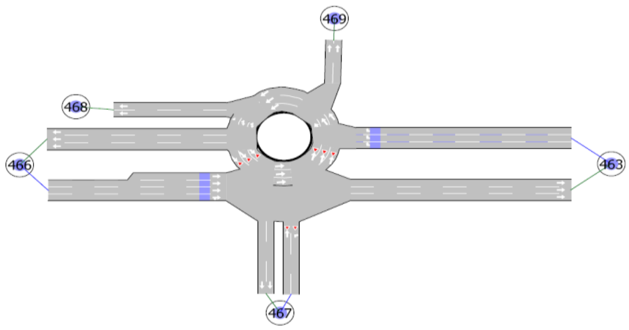



**Fig. 4.**
*“Roundabout” with give way signs and two direct lanes to Constanta*

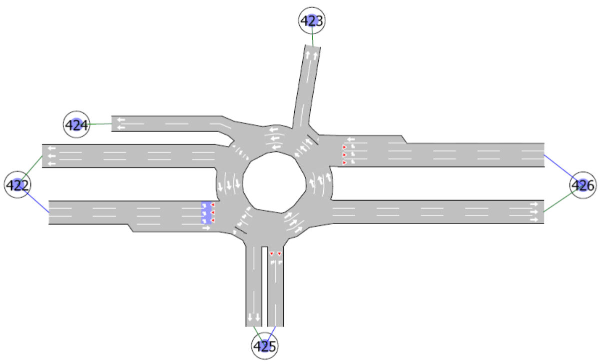



**Fig. 5.**
*Roundabout and traffic signalised ”roundabout”*

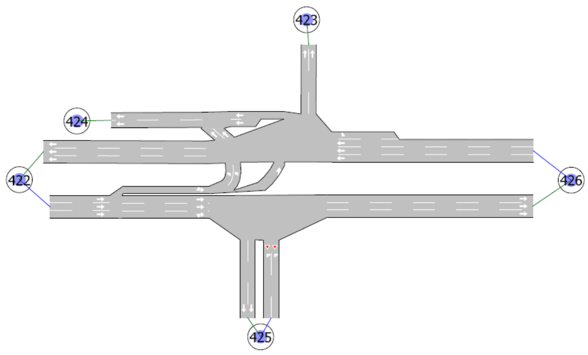



**Fig. 6.**
*Island junction with traffic lights (one lane for the commercial centre turn)*

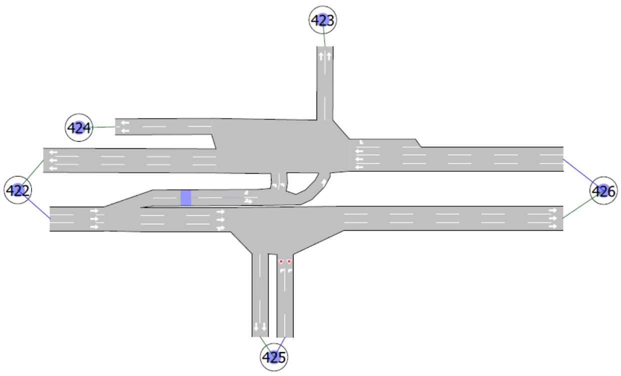



**Fig. 7.**
*Two turn left lanes with traffic lights (two lines for the commercial centre turn)*

## 4. Comparison of the proposed configurations and effects on traffic

The comparison achieved took into account more traffic parameters like: traffic flow, speed, queue length, delay, density and stop time. The following charts represent only an example (fig. 8-11) for one configuration of the most significant parameters in only one traffic peak (TSS, 2009).

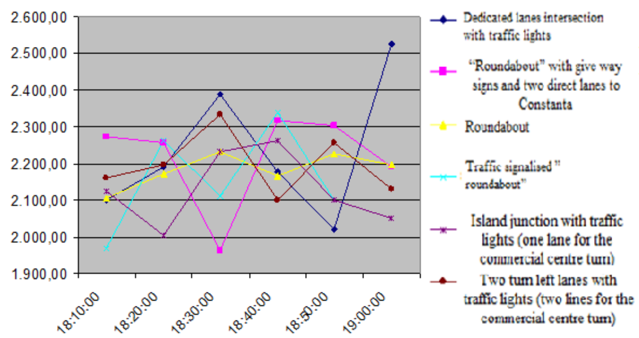



**Fig. 8.**
*Traffic flow (vehicles/h)*

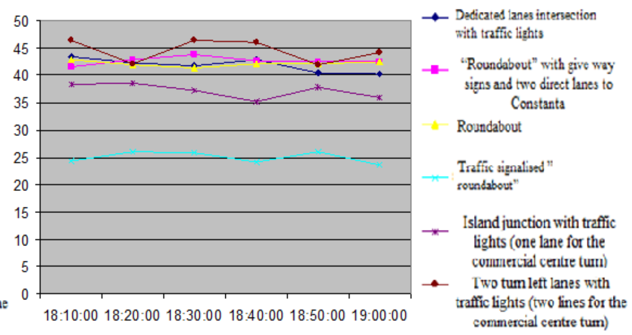



**Fig. 9.**
*Speed (km/h)*

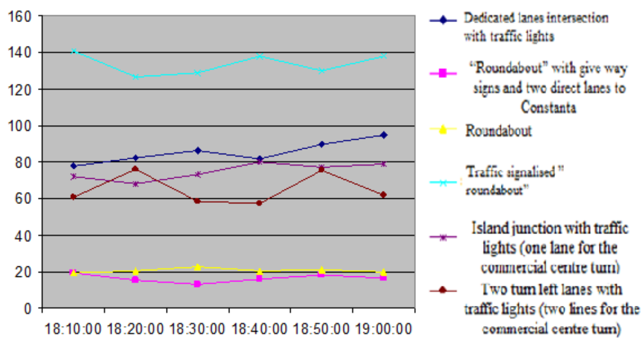



**Fig. 10.**
*Delays (seconds/km)*

After realising the study and in concordance with the attributes revealed by the simulation the two turn left lanes with traffic lights (two lines for the commercial centre turn, fig. 11) solution was selected as an appropriate systematization and recommended for implementation.

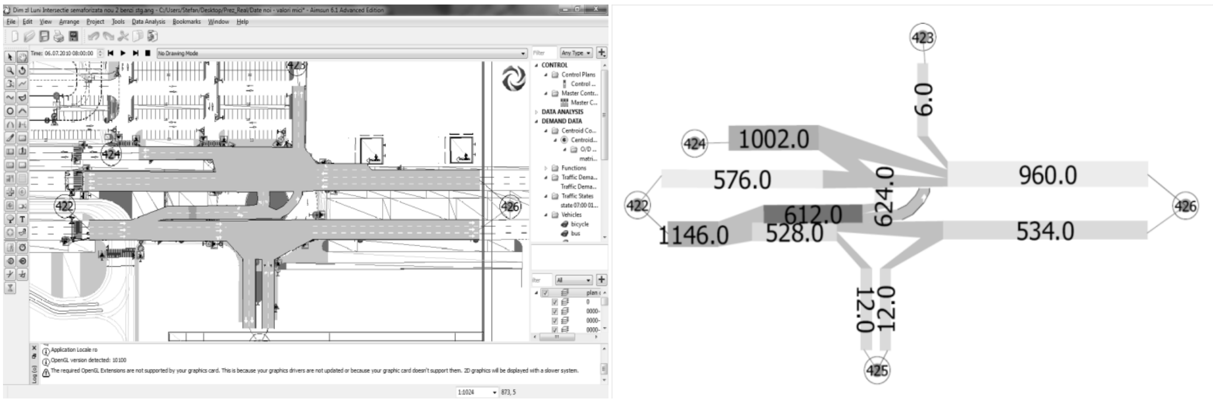



**Fig. 11.**
*Design and traffic flows on the proposed configuration*

The capacity of the proposed configuration is appropriate for the present and also for the future traffic volumes (fig. 11), so, with a proper design and traffic management (based on a signal light cycle ) the negative effects of locating a facility close to a major urban boulevard can be overcome (Wagner, 2010).

## 5. Conclusions

As location decisions are strategic and long term, logistic activities locating models must consider future uncertainties. From an economic point of view of producers and/or suppliers whose actions are determined by market evolution or authorities' policy these models were developed, by including spatial variables – spatial distribution of resources and distances, within microeconomic theory.
Locating a new commercial facility leads to reducing the total distance covered by the beneficiaries and also determines cost of transport, energy and pollution reduction with positive social effects (new work places, improved infrastructures). But, as it is strategic, clients and demand might vary daily leading to daily route changes, with effects on traffic, even though facilities keep their positions for a long time. With an efficient configuration and traffic management technique these influences on the road network can be reduced.
Traffic flow managing techniques are adapted to roads' configuration. Traffic flow control actions' objectives aim for:

a) maximum existing infrastructures usage;
b) traffic fluidization for congestion and external effects reduction;
c) some network area restrictions to avoid or limit congestion.

Urban traffic irregularity during the day determines an alternation of the above mentioned objectives, a step back in flexibility.
Traffic simulation models have proven to be helpful in analysing complex traffic situations that exist beyond the scope of the traditional analytical methods. Traffic simulation models could also be applied in transportation planning process, due to their flexibility and feasibility in testing different alternatives that do not currently exist in the real-world (Roşca et al., 2013).
Control strategies were also used in establishing the framework of the proposed objective. If traffic congestion is sufficient to determine primer congestion on a link then the control strategy must focus on avoiding secondary congestion development.
Linked to the loading level of the infrastructure's elements the control strategies must focus on:

1. prevention of reaching the saturation degree;
2. saturation handling, when not avoided;
3. immediate intervention on the causes determining saturation congestion.

Control and planning strategies must be extended to correlated intersections, as we have done, in order to determine the effects on the whole network/area and not only in some isolated points.
The aim of the paper was not only to determine a configuration of a future junction that appeared necessary after a new facility was located but also to prove that with the right measures negative effects could be diminished.

**Acknowledgement**
The work has been funded by the Sectorial Operational Programme Human Resources Development 2007-2013 of the Ministry of European Funds through the Financial Agreement POSDRU/159/1.5/S/132397.

**References**

Allen, J., Browne, M., Cherrett, T., 2012. Investigating relationships between road freight transport, facility location, logistics management and urban form, *Journal of Transport Geography 24*, 45–57.

Burciu, Ş., 2013. *Locating industrial facilities that generate major transport flows,* PhD Thesis, University Politehnica of Bucharest.

Fang, F. C., Elefteriadou, L., 2005. Some guidelines for selecting microsimulation models for interchange traffic operational analysis, *Journal of Transportation Engineering,* Vol. 131, 7, 535–543.

Lieberman, E., Rathi, A.K., 2005. Traffic Simulation, *Revised Monograph on Traffic Flow Theory* (N. H. Gartner, C.J. Messer, and A. K. Rathi, ed.), United States Transportation Research Board, 1–23.

Mackett, R.L., 1992. Structure of linkages between transport and land use, *Transportation Research Board* 27B, 189–206.

Manstetten, D., Krautter, W., 2005. Schwab, T., Traffic Simulation Supporting Urban Control System Development, *Proceedings of the 4th World conference on ITS,* Seoul.

Nourinejad, M., Wenneman, A., Habib, K.N., Roorda, M., 2014. Truck parking in urban areas: Application of choice modelling within traffic microsimulation, *Transportation Research Part A,* 64, 54–64.

Popa Gabriela, Beloiu, D. M., 1999. *Trafic şi securitate rutieră.* Îndrumar de laborator, Editura Datavision, Bucureşti.

Popa Mihaela, 2004. *Elemente de economia transporturilor*, Ed.BREN, Bucuresti, ISBN-973-648-311-8.

Raicu, Ș., 2007. *Sisteme de transport,* Ed. AGIR, ISBN 978-973-720-152-2.

Raicu, Ş., Dragu, V., Burciu, Ş., Ştefănică Cristina, 2011. Junction design and planning for the penetration of a major route determined by a new centroid, *Proceedings of the 5th International Conference on Manufacturing Science and Education*, 479–483.

Roșca, M.A., Roșca, E., Ruscă, F., Cârlan,V., 2013. Computer simulation for operational traffic improvement in urban intersections, *Advances in Automatic Control, Modelling& Simulation, ACMOS-15, Brasov,* 2013.

Sabater, G., 2004. *Productive and Logistics Systems Design*, Universitat Politècnica de València.

Salter, R. J., 1989. *Highway traffic analysis and design,* Macmillan Education LTD, London.

STAS 10144_1-6, 1991.

TSS-Transport Simulation Systems, 2009. *Aimsun User's Manual*, Draft Version – July 2009.

Wagner Tina, 2010. Regional traffic impacts of logistics-related land use, *Transport Policy,* 17, 224–229.

Waller, P.F., 1997. Transportation and Society, *Traffic & Transport Psychology*, Rothengater T. & Carbonell V.E. (Eds.), Pergamon Press, Oxford, 435–444.

# AN IMPROVED ENSEMBLE NEURAL NETWORKS MODEL FOR PREDICTION OF PUSHBOAT SHAFT POWER

**Aleksandar RADONJIC[1], Zlatko HRLE[2], Vladeta COLIC[3]**
[,1,2,3] *Faculty of Transport and Traffic Engineering, University of Belgrade, Vojvode Stepe 305, Belgrade, Serbia*

**Abstract**:This research intends to introduce new neural networks in the existing Ensemble Neural Networks (ENN) model and intends to build new ENN model for predicting pushboat shaft powers. As a continuation of the previous papers, research sets prerequisites for proper determination of shaft powers at a given navigating conditions. Authors made proposals that should contribute to improvement of the propulsive characteristics of the pushboats. Application of the ENN model may indirectly reduce fuel consumption and gas emissions.
It is proposed that speed-power curves obtained by ENN model are compared with curves obtained by 3rd order polynomial. It is also expected that new ENN model finds new functional relationship between shaft powers and convoy speeds.
Better generalization capability of the new ENN model is achieved through accuracy of additional component networks. It is shown that 4 additional component networks in the model can significantly improve the output results. All proposed conditions are satisfied while parameters like mean absolute error (MAE), root-mean square error (RMSE) and mean relative errors are notably reduced.
Research finds out that new ENN model provides more accurate shaft powers than the previous ENN model. However, there are still incorrect parts of the speed-power curves which can be solved by further implementation of the component networks in the model and introduction of the new methodologies for the calculation of output results.

**Keywords**: Full-scale trials, Towboat shaft power, Ensemble Neural Networks, 3rd order polynomial

## 1. Introduction

Full-scale speed-power trials are carried out to establish relationships between the speed, shaft power and propeller revolutions per minute under specified conditions of displacement, draft and trim. They are run for various purposes that include collection of experimental data for use in the future designs of ships, for improving the prediction of ship powers and for use by the ship as an aid to navigation (Lewis, 1988). In this paper they are used for the navigation planning of pushed convoys in order to minimize time travel and fuel consumption.
From the above points of view, the authors made functional relationship between ship speed and ship shaft power using the data collected from the full-scale speed-power trials published in various papers (Colic, 2006). Ensembles of neural networks (ENN) have been used to make functional relationship between speed and shaft power. Results obtained from this research could be used by the crews and captains as an aid to navigation of pushed convoys. This paper is a continuation of the work described in (Radonjic, 2014) and (Radonjic and Vuakdinovic, 2014).
The research subjects in this paper are pushboats and pushed convoys from full-scale trials. It was considered that all trials were conducted in deep water and with the same propulsive efficiency. Pushboat shaft power prediction was performed on the basis of the following input variables: pushed convoy slenderness ratio (length-displacement ratio), pushed convoy length-to-beam ratio, pushed convoy drought-to-beam ratio and Froude number. The main aim of the paper is to predict pushboats shaft powers on the basis of different combinations of pushboats and barge convoys from dataset presented in (Radonjic, 2014) and (Radonjic and Vuakdinovic, 2014). Research includes further investigation of the existing AIC based ENN model and its improvement and enhancement on the different data values.
Although the results of the previous papers are remarkable in terms of mean absolute errors (MAE), root-mean square errors (RMSE) and relative errors, the authors will investigate further possibility of correct prediction of pushboat shaft powers. Investigation will be conducted on the intermediate pushed convoy speeds from minimum to maximum values within the data sets per full-scale trial runs. It will be assumed that all speed-power curves per full-scale trial runs correspond to 3rd order polynomial (Ricketts and Hundley, 1997).
Most of the calculated pushboat shaft powers by AIC based ENN model agree well with the corresponding speed-power curves. However, in a few cases it is found that there are significant errors between data obtained by AIC based ENN model and corresponding speed-power curves obtained by using 3rd order polynomial. It reveals new insight into the accuracy of the intermediate pushboat shaft powers. Therefore, authors recommend introduction of the new component networks in AIC based ENN model.
The idea of the paper is to show that there is no need for the additional component networks in AIC based ENN if the so called extreme data points can be replaced under conditions which will be explained in detail. On the other side this paper gives clear understanding of actual propulsion characteristics to the crew and captains. It sets prerequisites for speed-power functional relationship at a given navigating condition. On the basis of the final results fuel consumption and gas emissions can be reduced. Also, shipping companies will be able to form various pushed convoys knowing that they (convoys) will navigate with the lower fuel consumption and higher speed at the given conditions.

[1] Corresponding author: a.radonjic@sf.bg.ac.rs

## 2. Literature review

Over the last two decades there has been extensive research in the field of ANNs for predicting relationships between various data related to the ship. ENN methodology appears in those kinds of papers only recently.

Ray et al. (1996) predicted a container capacity and the estimation of added mass coefficients for asymmetric bodies of revolution on the basis of the main dimensions of the ship. The authors concluded that ANNs represent a potential solution to wide variety of problems in the field of naval architecture and marine engineering.

Reich and Berai (2000) created a neural network model to predict the propeller thrust coefficient, the propeller torque coefficient and the propeller efficiency. Input variables were propeller pitch-propeller diameter ratio, expended area ratio, number of blades, advance coefficient and cavitation number. They concluded that ENN model building should be done carefully; starting from data collection, model quality estimation, to solution deployment. Further improvement of the model quality is possible through advanced methods.

Petersen and Jacobsen (2012) presented two statistical models of fuel efficiency in ship propulsion. They investigated and compared ANN and Gaussian processes (GP). The authors concluded that comparison with similar data is almost inapplicable as the trial data differs from test to test.

Radonjic and Vukadinovic (2014) proposed an AIC based ensemble neural network (ENN) and the single neural network (ANN) with two hidden layers to predict pushboat shaft power. These two models were compared on the basis of their calculated MAE values, RMSE values and relative errors. Computational results from this numerical example showed that ENN definitely outperformed single ANN with two hidden layers.

## 3. Methodology

In this paper Ensemble Neural Networks (ENN) have been used to predict towboats shaft powers. ENN model was made by using Akaike Information Criterion (AIC) (Sugiura, 1978). AIC based ENN methodology is known to have high generalization capability through the accuracy of every single component network (Ren and Zhao, 2002). Constrained truncated Newton algorithm (Nash, 2000) was selected to train component networks. Data set was separated into the training and testing data sets. Validation data set was used to detect overfitting and stop training process. Applied training stopping technique is “early stopping” (Priddy and Keller, 2005).

AIC based ENN model is composed of feed-forward neural networks (component networks) with one hidden layer. Component networks were chosen to form ENN on the basis of the following formula (Ren and Zhao, 2002):

$$0 < N_h < (N_{tr} - 1)/(N_i + 2) \quad (1)$$

where $N_i$ is the number of the input neurons; $N_h$ is the number of the hidden neurons; and $N_{tr}$ is the number of training data. Component networks are ranked according to their AIC Values (Zhao et al., 2008) and then selected to make ENN model (Radonjic, 2014). All data sets stayed unchanged during any training and testing processes. Prediction of pushboats shaft powers was made by the following expression:

$$P_i^{(i_{pb},i_{bf})} = \left\{ \left(P_i^{(i_{pb},i_{bf})}\right)_I, \text{ if } \min\left\{ \left| \left(P_i^{(i_{pb},i_{bf})}\right)_1 - P_{pi}^{(i_{pb},i_{bf})} \right|, \ldots, \left| \left(P_i^{(i_{pb},i_{bf})}\right)_M - P_{pi}^{(i_{pb},i_{bf})} \right| \right\} = \left| \left(P_i^{(i_{pb},i_{bf})}\right)_I - P_{pi}^{(i_{pb},i_{bf})} \right| \right. \quad (2)$$
$$I = 1,2,\ldots,M,\ i = 1,2,\ldots,m^{(i_{pb},i_{bf})}$$

where:

$P_{pi}^{(i_{pb},i_{bf})}$ - pushboat shaft power calculated by 3rd order polynomial and expressed by the following formula[2]:

$$P_{pi}^{(i_{pb},i_{bf})} = P_{pi}^{(i_{pb},i_{bf})}\left(v_i^{(i_{pb},i_{bf})}\right) = a_i^{(i_{pb},i_{bf})} \cdot \left(v_i^{(i_{pb},i_{bf})}\right)^3 + b_i^{(i_{pb},i_{bf})} \cdot \left(v_i^{(i_{pb},i_{bf})}\right)^2 + c_i^{(i_{pb},i_{bf})} \cdot v_i^{(i_{pb},i_{bf})} + d_i^{(i_{pb},i_{bf})} \quad (3)$$
$$a_i^{(i_{pb},i_{bf})} \in R, b_i^{(i_{pb},i_{bf})} \in R, c_i^{(i_{pb},i_{bf})} \in R, d_i^{(i_{pb},i_{bf})} \in R, i = 1,2,\ldots,m^{(i_{pb},i_{bf})}$$

$v_i^{(i_{pb},i_{bf})}$ - intermediate pushed convoy speed

M - number of component networks

$m^{(i_{pb},i_{bf})}$ - number of intermediate pushed convoy speeds per full-scale trial run[3] which is given by:

$$m^{(i_{pb},i_{bf})} = \max\left\{ m \in Z \,\middle|\, m \le \left(v_{\max}^{(i_{pb},i_{bf})} - v_{\min}^{(i_{pb},i_{bf})}\right) \cdot 10 \right\} \quad (4)$$

$v_{\min}^{(i_{pb},i_{bf})}$ and $v_{\max}^{(i_{pb},i_{bf})}$ are minimal and maximal convoy speeds recorded during the individual full-scale trial runs.

$i_{pb}=i_{tow}$ – pushboat number (Radonjic and Vukadinovic, 2014) ($i_{pb}$ = 1,2,…,15)

$i_{bf}$ – barge formation number (Radonjic and Vukadinovic, 2014) ($i_{bf}$ = 1,2,…,48)

p – 3rd order polynomial

R – the set of real numbers

Z – the set of integers

[2] Full-scale trial run is determined by the combination of $i_{pb}$ and $i_{bf}$.

[3] Regression was used to model pushboats shaft powers as a response to pushed convoy speed.

Authors adopted AIC based ENN methodology with 4 component networks to analyze MAE, RMSE and relative errors and now applied existing AIC based ENN model to the new intermediate convoy speeds per full-scale trials. It revealed new insights into both predicted shaft powers and concavity of the newly obtained speed-power curves.

Therefore, new AIC based ENN model is proposed if the predicted shaft powers and speed-power curves does not satisfy the following conditions:

1. $P_m^{(i_{pb},i_{bf})} > P_{m-1}^{(i_{pb},i_{bf})} > \ldots > P_2^{(i_{pb},i_{bf})} > P_1^{(i_{pb},i_{bf})}$ for $v_m^{(i_{pb},i_{bf})} > v_{m-1}^{(i_{pb},i_{bf})} > \ldots > v_2^{(i_{pb},i_{bf})} > v_1^{(i_{pb},i_{bf})}$ is not true for all pairs $\left(v_i^{(i_{pb},i_{bf})}, P_i^{(i_{pb},i_{bf})}\right), i = 1,2,\ldots,m^{(i_{pb},i_{bf})}$;
2. Two thirds of the speed-power curve obtained by the AIC based ENN model is concave and mean relative error is higher than 5%;
3. Mean relative error is higher than 10%;
4. More than 50% of the speed-power curve obtained by AIC based ENN model is concave;
5. More than one relative error of the predicted pushboat shaft power is higher than 10%;
6. Entire 3rd polynomial speed-power curve is convex and AIC based ENN speed-power curve is concave.

The condition number 1 is satisfied if there are enough values from component networks that can replace shaft powers that satisfy the following condition:

$$P_{i+1}^{(i_{pb},i_{bf})} < P_i^{(i_{pb},i_{bf})} > P_{i-1}^{(i_{pb},i_{bf})} \text{ for } v_{i+1}^{(i_{pb},i_{bf})} > v_i^{(i_{pb},i_{bf})} > v_{i-1}^{(i_{pb},i_{bf})} \quad (5)$$

The next expression is applied if there is at least one shaft power which does not meet the condition number 2:

$$P_i^{(i_{pb},i_{bf})} = \left\{\left(P_i^{(i_{pb},i_{bf})}\right)_{I'}, \text{if } P_{i+1}^{(i_{pb},i_{bf})} > \left(P_i^{(i_{pb},i_{bf})}\right)_{I'} > P_{i-1}^{(i_{pb},i_{bf})}\right. \\ I' = 1,2,..I-1, I+1,..M, i = 1,2,\ldots,m^{(i_{pb},i_{bf})} \quad (6)$$

$P_{i-1}^{(i_{pb},i_{bf})}$ and $P_{i+1}^{(i_{pb},i_{bf})}$ are predicted on the basis of expression (2). After the application of the expression (6) it is necessary to verify if there is another shaft power that does not meet the condition 1. If there is such a shaft power and more than one shaft power that do not meet the condition 1, the expression (6) will be constantly applied until the all predicted shaft powers satisfy the condition 1. It is clear that values of MAE, RMSE and relative errors per full-scale trial run are increasing by applying expression (6).
Mean relative error is counted by the following formula (Pedersen and Larsen, 2009):

$$\hat{\mu}_G = \frac{1}{m^{(i_{pb},i_{bf})}} \sum_{i=1}^{m^{(i_{pb},i_{bf})}} \left| \frac{P_{pi}^{(i_{pb},i_{bf})} - P_i^{(i_{pb},i_{bf})}}{P_{pi}^{(i_{pb},i_{bf})}} \right| \quad (7)$$

Condition which will ensure the convexity of the speed-power curve obtained by AIC based ENN model is expressed according to the following formula (Boyd, 2004):

$$P_{pi}^{(i_{pb},i_{bf})}\left(t_i \cdot v_1^{(i_{pb},i_{bf})} + (1-t_i) \cdot v_m^{(i_{pb},i_{bf})}\right) \le t_i \cdot P_{pi}^{(i_{pb},i_{bf})}\left(v_1^{(i_{pb},i_{bf})}\right) + (1-t_i) P_{pi}^{(i_{pb},i_{bf})}\left(v_m^{(i_{pb},i_{bf})}\right) \quad (8)$$

where:

$$t_i = \left(v_m^{(i_{pb},i_{bf})} - v_i^{(i_{pb},i_{bf})}\right) / \left(v_m^{(i_{pb},i_{bf})} - v_1^{(i_{pb},i_{bf})}\right), i = 1,2,\ldots,m^{(i_{pb},i_{bf})} \quad (9)$$

If there are shaft powers data points that are not satisfy proposed conditions, new component networks will be introduced in AIC based ENN model. The training process will be performed for all of them. New component networks will be introduced until the all conditions are satisfied

**4. Numerical example**

In order to predict new pushboat shaft powers, AIC based ENN model is used. New shaft powers are estimated at intermediate speeds that increased to 0.1 km/h from minimum to maximum values within the data set per full-scale trial run. Shaft powers predicted by AIC based ENN model are compared with shaft powers calculated by 3rd order polynomials. In this paper, only speed-power curves (per full-scale trial runs) that do not satisfy one or more conditions (presented in chapter 3) will be presented. Remaining speed-power curves will not be presented as their shaft powers satisfy all previous conditions. They will be presented with the values of MAE, RMSE and mean relative errors of the differences between values predicted by an AIC based ENN model and the values obtained by 3rd order polynomial.

### 4.1. Speed-power curves (3rd order polynomial and initial AIC based ENN)

In order to investigate the AIC based ENN model, predicted pushboats shaft powers are compared with the powers predicted by using 3rd order polynomials as shown in the following speed-power curves in Fig. 1.

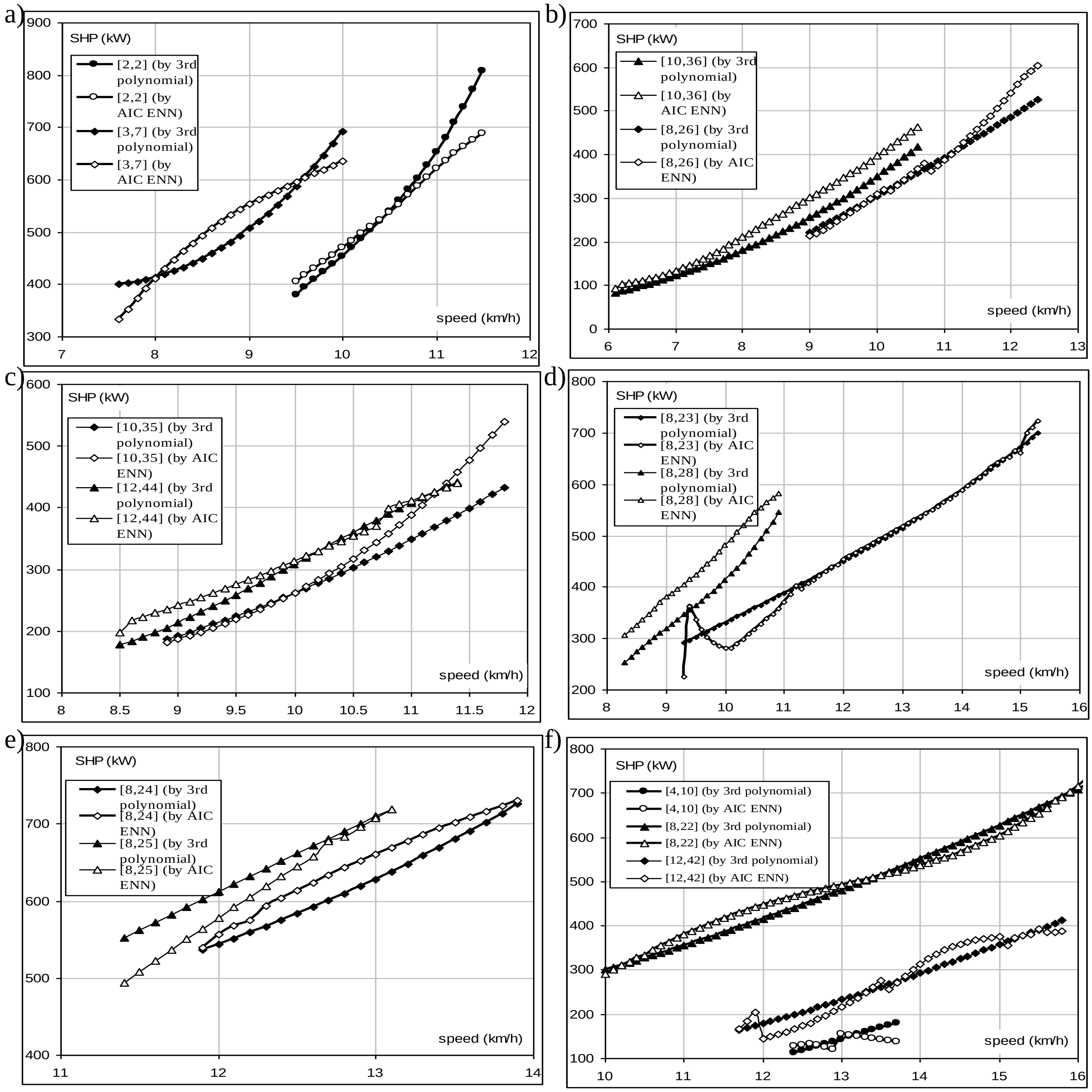



**Fig 1.**
*Comparison of shaft powers predicted by 3rd order polynomial and by AIC based ENN model*

In the Fig. 1 pair [2,2] denotes the values of $i_{pb}$ and $i_{bf}$. Pushboat shaft powers predicted by 3rd order polynomials are presented as markers filled with black color while those powers predicted by AIC based ENN model are shown as markers with no fill. All speed-power curves from Fig. 1 satisfy condition 1 except the full-scale trial run with $i_{pb}$ = 4 and $i_{bf}$ = 10 (f)). Predicted shaft powers from this speed-power curve do not meet any of the conditions while the values of MAE, RMSE and mean relative errors are the biggest of all.
Values of MAE, RMSE, mean relative errors and unsatisfied conditions are shown in the table 1.
Overall value of MAE is 16.67 kW which is between values of MAE calculated for the training data set and values of MAE calculated for the testing data set (Radonjic, 2014). Overall value of RMSE is 20.60 kW and is lower than values of RMSE calculated for the training data set and values of RMSE calculated for the testing data set (Radonjic, 2014). Overall mean relative error is 3.8% and is almost identical to mean relative error calculated for training data set in (Radonjic, 2014). Reasons why some speed-power curves are not acceptable are shown in column 8.

The conclusion is that there has to be new component networks on the basis of the results presented in table 1 and speed-power curves shown in Fig. 1. In this paper one by one new component network is added to the initial AIC based ENN model until the all conditions are satisfied for all speed-power curves.

**Table 1.**
*MAE, RMSE and relative errors of pushboat shaft powers predicted by previous AIC based ENN model*

| Pushboat | $i_{pb}$ | $i_{bft}$ | $\left[v_1^{(i_{pb},i_{bf})}; v_m^{(i_{pb},i_{bf})}\right]$ (km/h) | MAE (kW) | RMSE (kW) | Mean relative error | Unsatisfied conditions (number) |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| "Bukovik" | 1 | 1 | [7.6;10.3] | 6.040 | 7.470 | 0,020 | – |
| "Deligrad" | 2 | 2 | [9.5;11.5] | 30.179 | 43.255 | 0,049 | 5 |
| | | 3 | [8;10] | 5.003 | 5.825 | 0,009 | – |
| | | 4 | [8.5;10] | 9.287 | 11.464 | 0,016 | – |
| "Vinodol" | 3 | 7 | [7.6;10] | 33.439 | 37.699 | 0,069 | 6 |
| "Kablar" | 4 | 10 | [12.4;13.7] | 16.420 | 20.132 | 0,109 | 1,2,3,4,5,6 |
| | | 11 | [9.8;11.4] | 2.381 | 3.582 | 0,016 | – |
| "Sloga" | 5 | 12 | [8.8;13.3] | 10.433 | 14.082 | 0,021 | – |
| | | 13 | [9.2;15.8] | 3.916 | 4.555 | 0,007 | – |
| | | 14 | [11.2;16.7] | 27.911 | 31.758 | 0,038 | – |
| | | 16 | [11.6;18.5] | 5.441 | 6.745 | 0,011 | – |
| "Romanija" | 6 | 17 | [8.7;12.1] | 3.190 | 4.068 | 0,005 | – |
| | | 18 | [7.5;10.9] | 4.492 | 6.019 | 0,009 | – |
| | | 19 | [6.4;10] | 7.579 | 9.785 | 0,014 | – |
| "Sumadija" | 7 | 20 | [8.7;10.7] | 3.518 | 4.369 | 0,007 | – |
| | | 21 | [7.6;9.4] | 5.368 | 8.357 | 0,013 | – |
| "Pinki" | 8 | 22 | [10;16.2] | 17.429 | 19.993 | 0,038 | 2 |
| | | 23 | [9.3;15.3] | 14.322 | 23.152 | 0,039 | 1 |
| | | 24 | [11.9;13.9] | 22.928 | 24.922 | 0,037 | 4 |
| | | 25 | [11.4;13.1] | 26.363 | 32.385 | 0,044 | 4 |
| | | 26 | [9;12.4] | 19.101 | 29.728 | 0,044 | 1 |
| | | 27 | [7.2;11.6] | 18.174 | 20.158 | 0,045 | – |
| | | 28 | [8.3;10.9] | 59.488 | 59.928 | 0,162 | 3 |
| "Panonija" | 9 | 29 | [7.9;11.3] | 4.380 | 5.281 | 0,010 | – |
| | | 30 | [10.1;11.4] | 7.660 | 8.949 | 0,016 | – |
| | | 31 | [7.2;10.3] | 9.260 | 11.300 | 0,029 | – |
| | | 32 | [8.42;11.6] | 10.294 | 13.558 | 0,025 | – |
| | | 33 | [8.8;12.3] | 6.554 | 11.359 | 0,014 | – |
| "Banat II" | 10 | 34 | [10.4;13.9] | 10.241 | 11.761 | 0,036 | – |
| | | 35 | [8.9;11.8] | 27.965 | 41.950 | 0,078 | 2 |
| | | 36 | [6.1;10.6] | 31.679 | 34.840 | 0,142 | 5 |
| "Kadinjaca" | 11 | 37 | [9.8;13.5] | 82.640 | 96.518 | 0,094 | – |
| | | 38 | [8.7;10.6] | 19.903 | 25.163 | 0,024 | – |
| | | 39 | [10.5;14.8] | 7.573 | 10.393 | 0,009 | – |
| | | 40 | [9.9;12.3] | 24.698 | 30.005 | 0,028 | – |
| | | 41 | [12.3;16.2] | 14.316 | 18.983 | 0,010 | – |
| "Srem" | 12 | 42 | [11.6;15.8] | 19.487 | 22.608 | 0,078 | 1,2,3,4,5,6 |
| | | 43 | [10.1;13.5] | 12.239 | 14.239 | 0,044 | – |
| | | 44 | [8.5;11.4] | 12.970 | 16.780 | 0,056 | 5 |
| "Kladovo" | 13 | 45 | [8.8;14.2] | 16.991 | 26.097 | 0,016 | – |
| "Truzenik-1" | 14 | 46 | [9.2;11.3] | 16.897 | 18.646 | 0,070 | – |
| | | 47 | [11.3;14.1] | 6.672 | 7.665 | 0,028 | – |
| "Bor" | 15 | 48 | [10.9;15.5] | 25.855 | 30.159 | 0,038 | – |

*Source for coulmns 1, 2 and 3: (Radonjic and Vukadinovic (2014) and Radonjic (2014))*

### 4.2. Computational results obtained by additional component networks

The process of adopting new component networks is completed when all additional networks are incorporated in AIC based ENN model. New shaft powers are predicted on the basis of the training of new networks and on the basis of 4 previously trained networks in the ENN model. All previous speed-power curves that satisfied conditions were not tested in the process of adding component networks. Introduction of new component networks was completed when all speed-power curves satisfied all proposed conditions. During the process of introducing new networks in AIC based ENN model it was assumed that all previously adopted speed-power curves would remain unchanged or would become more accurate. Final results showed that only condition 1 had to be applied on the previously accepted speed-power curves in order to correct newly emerged errors.

Networks were sequentially added based on the AIC values presented in (Radonjic and Vukadinovic, 2014). The first network that was added in AIC based ENN was network with 11 neurons in the hidden layer. Since the predicted shaft powers with this network in the model did not meet proposed conditions, new network with 12 hidden neurons was introduced. Process of introducing component networks was continued until the 4 new networks along with 4 previous networks made new AIC based ENN model. It occurred when all of the speed-power curves (see Fig. 1) satisfied the previous conditions. Therefore, on the basis of the ranked networks in the Table 2, AIC based ENN model with 4 component networks was extended with 11, 12, 13 and 9 one hidden layer neural networks. Thus, the performance of the ENN model was enhanced. The new AIC based ENN model had 8 component networks with the following number of neurons in hidden layer: 8, 9, 10, 11, 12, 13, 14 and 15. Upon the completion of training process of the new networks new speed-power curves were obtained and shown in Fig. 2.

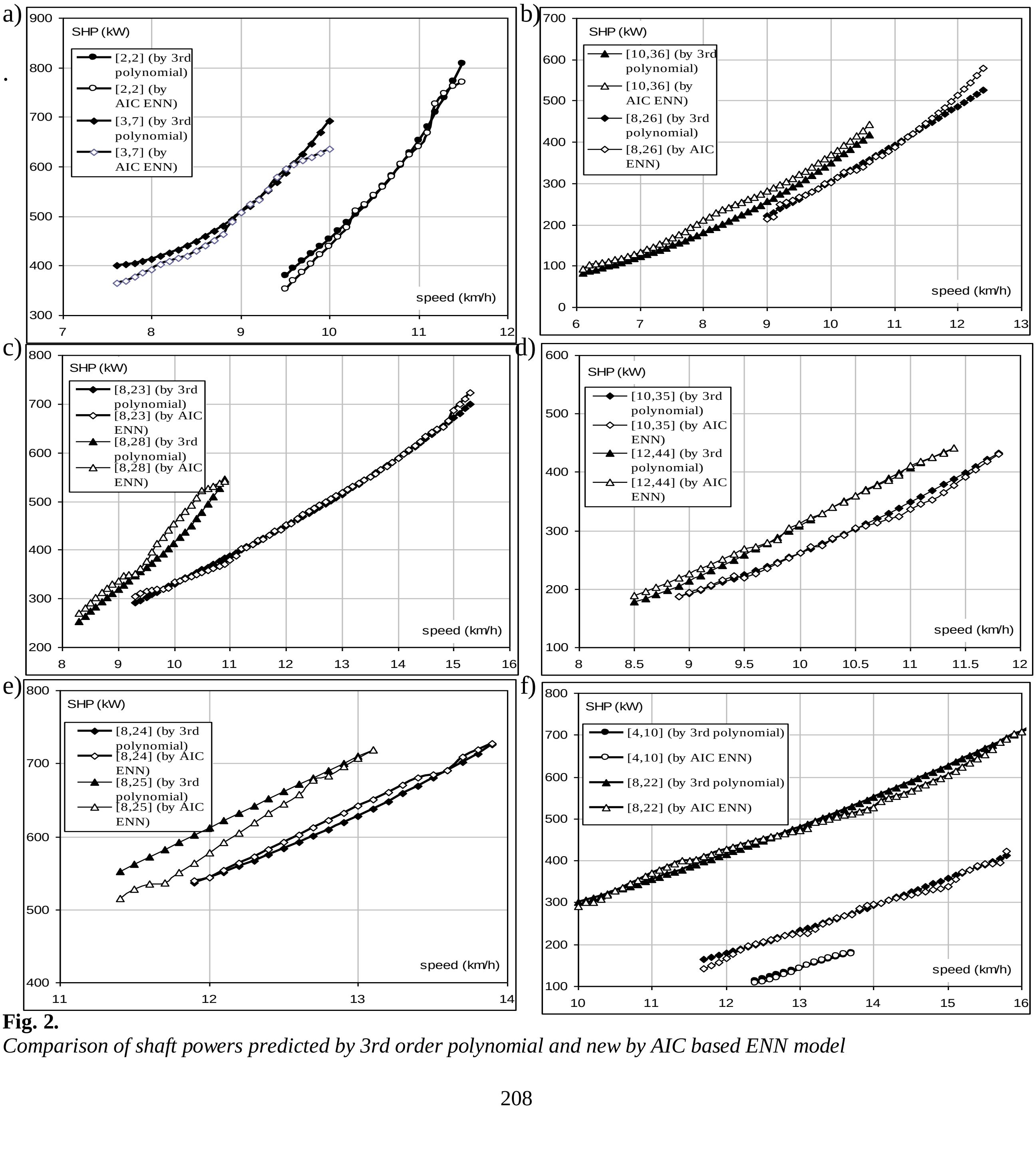



**Fig. 2.**
*Comparison of shaft powers predicted by 3rd order polynomial and new by AIC based ENN model*

In the Fig. 2 speed-power curves obtained by new AIC based ENN model correspond much better to speed-power curves obtained by 3$^{rd}$ order polynomial. It is clearly visible that errors between predicted shaft powers are lower than errors obtained in the Fig. 1.

Speed-power curves from the Fig. 2 satisfy proposed conditions and are accepted as the final solution in this paper. In order to verify the validity of results, MAE, RMSE and mean relative errors of the predicted shaft powers by the new AIC based ENN are shown in Table 2. Results from the two ENN models show that significant accuracy gains can be made using the later ENN model. The value of MAE is 8.78 kW which is half of the previous value recorded with former AIC based ENN model. The RMSE value is 16.09 kW and is lower than previous value too. Calculated mean relative error is 1.9% which is also half of the previous percent of 3.8. Following the possible errors that could appear in the already accepted speed-power curves it was tested whether new speed-power curves matched the previous conditions. It was found that few of them didn't match the first condition and expression (6) was applied. All previous speed-power curves (presented in Fig. 1) satisfy all the rules and it is not required to add new component networks.

**Table 2.**
*MAE, RMSE and mean relative errors of pushboat shaft powers predicted by new AIC based ENN model*

| Pushboat | $i_{pb}$ | $i_{bft}$ | $\left[v_1^{(i_{pb},i_{bf})}; v_m^{(i_{pb},i_{bf})}\right]$ (km/h) | MAE (kW) | RMSE (kW) | Mean relative error |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| “Bukovik” | 1 | 1 | [7.6;10.3] | 1.440 | 1.820 | 0.005 |
| “Deligrad” | 2 | 2 | [9.5;11.5] | 12.821 | 16.278 | 0.025 |
| | | 3 | [8;10] | 2.962 | 3.431 | 0.005 |
| | | 4 | [8.5;10] | 5.611 | 7.305 | 0.010 |
| “Vinodol” | 3 | 7 | [7.6;10] | 18.688 | 22.877 | 0.038 |
| “Kablar” | 4 | 10 | [12.4;13.7] | 3.633 | 4.117 | 0.027 |
| | | 11 | [9.8;11.4] | 1.519 | 1.851 | 0.011 |
| “Sloga” | 5 | 12 | [8.8;13.3] | 6.582 | 9.558 | 0.013 |
| | | 13 | [9.2;15.8] | 3.318 | 4.068 | 0.006 |
| | | 14 | [11.2;16.7] | 2.597 | 3.471 | 0.004 |
| | | 16 | [11.6;18.5] | 2.390 | 3.030 | 0.004 |
| “Romanija” | 6 | 17 | [8.7;12.1] | 2.396 | 3.232 | 0.004 |
| | | 18 | [7.5;10.9] | 2.809 | 3.441 | 0.005 |
| | | 19 | [6.4;10] | 5.028 | 6.269 | 0.009 |
| “Sumadija” | 7 | 20 | [8.7;10.7] | 3.398 | 4.316 | 0.007 |
| | | 21 | [7.6;9.4] | 1.923 | 2.492 | 0.004 |
| “Pinki” | 8 | 22 | [10;16.2] | 11.937 | 13.992 | 0.025 |
| | | 23 | [9.3;15.3] | 4.825 | 7.130 | 0.011 |
| | | 24 | [11.9;13.9] | 7.547 | 8.734 | 0.012 |
| | | 25 | [11.4;13.1] | 23.213 | 27.536 | 0.038 |
| | | 26 | [9;12.4] | 10.505 | 16.554 | 0.026 |
| | | 27 | [7.2;11.6] | 16.015 | 18.756 | 0.040 |
| | | 28 | [8.3;10.9] | 23.825 | 26.793 | 0.063 |
| “Panonija” | 9 | 29 | [7.9;11.3] | 3.882 | 4.851 | 0.009 |
| | | 30 | [10.1;11.4] | 6.152 | 7.657 | 0.013 |
| | | 31 | [7.2;10.3] | 4.283 | 6.279 | 0.012 |
| | | 32 | [8.42;11.6] | 6.696 | 7.980 | 0.017 |
| | | 33 | [8.8;12.3] | 2.690 | 3.743 | 0.007 |
| “Banat II” | 10 | 34 | [10.4;13.9] | 5.235 | 6.560 | 0.020 |
| | | 35 | [8.9;11.8] | 5.382 | 6.981 | 0.017 |
| | | 36 | [6.1;10.6] | 20.886 | 22.032 | 0.107 |
| “Kadinjaca” | 11 | 37 | [9.8;13.5] | 47.989 | 66.057 | 0.058 |
| | | 38 | [8.7;10.6] | 13.621 | 16.082 | 0.017 |
| | | 39 | [10.5;14.8] | 6.329 | 9.711 | 0.008 |
| | | 40 | [9.9;12.3] | 20.649 | 26.021 | 0.023 |
| | | 41 | [12.3;16.2] | 5.302 | 9.012 | 0.004 |
| “Srem” | 12 | 42 | [11.6;15.8] | 6.602 | 9.111 | 0.026 |
| | | 43 | [10.1;13.5] | 5.852 | 7.757 | 0.021 |
| | | 44 | [8.5;11.4] | 5.733 | 7.511 | 0.025 |
| “Kladovo” | 13 | 45 | [8.8;14.2] | 14.995 | 23.919 | 0.014 |
| “Truzenik-1” | 14 | 46 | [9.2;11.3] | 3.188 | 3.841 | 0.013 |
| | | 47 | [11.3;14.1] | 3.938 | 5.336 | 0.018 |
| “Bor” | 15 | 48 | [10.9;15.5] | 13.460 | 17.177 | 0.021 |

The new AIC based ENN model performs significantly better than the previous AIC based ENN model. These results show that the new ENN model can be used to predict shaft powers and that all speed-power curves satisfy proposed conditions. The relationships between shaft power and convoy speeds were found to be in good order with that of 3rd order polynomials. The number of combinations of pushboats and barge formations is significantly improved. With the new model, every pushboat is now a candidate to perform transportation of barge formations that were the subjects of full-scale trial runs. However, only combinations with the pushboats of the same principal dimensions and power are proposed to make an extended set of combinations of $i_{pb}$ and $i_{bf}$ in this paper. Thus, the number of ($i_{pb}$, $i_{bf}$) pair combinations are increased from 48 (Tables 1 and 2) to 114.

**5. Conclusion**

In this paper new AIC based ENN model is proposed to predict pushboat shaft power and make a better quality results. The study is the continued work described in (Radonjic and Vukadinovic, 2014) and (Radonjic, 2014). Authors tried to improve already obtained results by analyzing the modification of the ENN model with additional component networks. They also developed a list of conditions that should contribute to prediction of accurate shaft powers.

The subject of analysis was data obtained from full-scale speed-power trials. The main goal was to demonstrate that additional neural networks in ENN model could significantly improve the generalization ability of the ENN model. The other objective was to indicate the importance of the predicted shaft power to the shipping companies. By using the new ENN model it is possible to increase number of combinations of ($i_{pb}$, $i_{bf}$) pairs. Also, the crews and the captains will get valuable information about the pushboat powering characteristics for the given barge formation by using AIC based ENN model. Computational results showed that ENN model with 8 neural networks performed significantly better than ENN model with 4 neural networks.

However, there are still concave parts of the speed-power curves. Concave parts can be solved in two ways: by adding more component networks to ENN model and by introducing ranges of different methodologies for the calculation of shaft powers including usage of best combination weights of the ENN model (Zhao et al., 2008 and Yao and Liu, 1998). As a next step, the authors are studying the impact of the changes of the main convoy dimensions. As another next step, the authors will investigate further improvement of the results by diversity of the members of ENN model (Brown et al., 2005). This research and AIC based ENN model can be applied to various maritime fields that include stability of ships, ship propulsion (propellers and ship resistance), ship hydrodynamics, gas emissions, loading capacity etc.

**References**


Boyd, S.P.; Vandenberghe, L. 2004. *Convex Optimization*. Cambridge University Press, 716p.

Brown, G.; Wyatt, J.; Harris, R.; Yao, X. 2005. *Diversity creation methods: a survey and categorization*, Information Fusion, 6(1):5–20.

Colic, V. 2006. *Research of navigational, technical, energetic and propulsive characteristics of Danube towboats*. University of Belgrade, Faculty of Transport and Traffic Engineering, Belgrade, 350p.

Lewis V, E. (Ed.) 1988. *Principles of Naval Architecture.* Volume II, Resistance, Propulsion and Vibration, Section 11, *Ship Standardization Trials*. The Society of Naval Architects and Marine Engineeres, 328p.

Nash, S.G., (2000) „A survey of truncated-Newton methods", Journal of Computational and Applied Mathematics 124(1-2), 45-49.

Pedersen, B.; Larsen, J. 2009. *Prediction of full-scale propulsion power using artificial neural networks* In: Proceedings of the 8th international conference on computer and IT applications in the maritime industries (COMPIT'09), Budapest, Hungary, May, 10–12, 537–550. Available from the internet: <http://www2.imm.dtu.dk/pubdb/p.php?5840>.

Petersen, J.P.; Jacobsen, D.J. 2012. *Statistical modelling for ship propulsion efficiency*, Journal of Marine Science and Technology 17(1), 30-39.

Priddy, K.L.; Keller, P.E. 2005. *Artificial neural networks: An introduction*. SPIE Publications, 180p.

Radonjic, A. 2014. *Analytical Methods for Calculation of Danubian Vessels' Performance Characteristics in order to Improve Navigation*. University of Belgrade, Faculty of Transport and Traffic Engineering, Belgrade, 236p.

Radonjić, A.; Vukadinović, K. 2014. *Application of Ensemble Neural Networks to Prediction of Towboat Shaft Power.* Journal of Marine Science and Technology, DOI: 10.1007/s00773-014-0273-2.

Ray, T.; Gokarn, R.P.; Sha, O.P. 1996. *Neural network applications in naval architecture and marine engineering*, Artificial Intelligence in Engineering 10(3): 213–226.

Reich, Y.; Barai, S.V. 2000. *A methodology for building neural networks model from empirical engineering data,* Engineering Applications of Artificial Intelligence 13(6): 685-694.

Ren, L.Q.; Zhao, Z.Y. 2002. *An optimal neural network and concrete strength modeling*. Advances in Engineering Software 33(3), 117–130.

Ricketts, J. B.; Hundley, L. L. 1997. *Naval Ship Self-Assessment of Hull Powering Performance Using Propulsion Shaft Torsionmeters and GPS*. Naval Engineers Journal, 109(3): 181–192.

Sugiura, N. 1978. *Further analysis of data by Akaike's information criterion and the finite corrections*, Communications in Statistics - Theory and Methods 7(1), 12–26.

Zhao, Z.; Zhang, Y.; Liao, H. 2008. *Design of ensemble neural network using the Akaike information criterion*. Engineering Application of Artificial Intelligence 21(8), 1182-1188.

Yao, X., Liu; Y. 1998. *Making use of population information in evolutionary artificial neural networks*, IEEE Transactionson Systems, Man, and Cybernetics, PartB, 28(3):417–425


.

# MODELING CONTAINERS THROUGHPUT IN PORT OF KOPER

**Elen Twrdy[1], Milan Batista[2]**
[1,2]*University of Ljubljana, Faculty of Maritime Studies and Transport, Pot pomorščakov 4, SI – 6320 Portorož, Slovenia*

**Abstract:** The global container transport increase of 8 to10% on a yearly basis and reach over 601,8 million TEU in 2012 (1). The growth was stopped in 2008 when the global financial and economic recession started to have full effect in ports and maritime industry. The ports North Adriatic did not followed the average of increase of container transport in the last 20 years (they had an increase of 7% on yearly base) and the decline in container transport in this ports in the years 2008 to 2009 was minimal. In this condition the fastest growth of container throughput was recorded at the Port of Koper, at an average of 14% per year, and reached 600.441 TEU in year 2013. The growth of container transport in the Port of Koper as well as the beginning of construction on the new container terminal has made the reconstruction and extension of the current container terminal an absolute priority. The extension is in line with the estimated growth of traffic as well as with the exploitation of present and future terminal capacities.
In paper we present dynamic models that help us to forecast a container throughput in port of Koper. The models are prepared in the base of available data and include model of growth, model of share and model of growth-share. As the port of Koper is limited with the capacity it is important to know the situation and to prepare the resources.

**Keywords**: North Adriatic ports, port of Koper, container throughput

## 1. Introduction

The Port of Koper is some 2000 nautical miles closer to destinations east of Suez than the ports of Northern Europe and it's located at the northern tip of Adriatic Sea near to the market of Central and Eastern Europe. From Koper there are regular and reliable shipping container lines to all major world ports. More than thirty container lines use the Port of Koper. Land transport from Koper by road and by railway to the main industrial centres in Central Europe is approximately 500 km shorter then from North European ports. Some two-thirds of cargo is transported by rail, which means that more than 500 wagons arrive and leave the port on a daily basis.
According to Notteboom Port of Koper is a part of Multi – port gateway region of North Adriatic (NA), presented by ports of Ravenna, Venice, Triest, Rijeka and Koper and container throughput dynamics can best be analyses by using "multi-port gateway regions" as unit of analysis (2). These ports have made in year 2010 association in the field of promotion called NAPA (North Adriatic Ports Association), however today Port of Ravenna is not a partner any more (3). Various goods are throughput in these ports but in this paper we will primarily deal with the container traffic and the fastest growth of container traffic was recorded at the port of Koper. Despite the global financial and economic crises in the near past the decline in overall container transport in NA ports was insignificant. The reason for that is probably in the fact that in the last years some of manufacturing companies, especially major vehicle and vehicle parts producers, have moved their production to Central Europe, partly to follow their main customers and partly to take advantage of the qualified and cheap labor force for export production. This production movement leads to larger bi-directional East-West flow within the European Union of raw materials and consumer products in containers thought the NA ports. A great increase in containers transport that is observed recently is thus in our opinion a consequence of the introduction of the direct transport lines between Asia and the North Adriatic and this was probably also one of the main reasons that in the first period of economic crises NA ports did not lost a lot of container traffic.
The Port of Koper is striving to become the main container terminal in the northern Adriatic. All NA ports operate in a relatively closed system in which the market and customers are limited and therefore the ports are forced to co-operate while at the same time they compete with each other. They are located in three countries, with different transport policies and plans of development. The Port of Koper is Slovenia's only port, and therefore extremely important, contributing significantly to the Slovenian GDP. The state of Slovenia is the largest shareholder and the future development of the port depends on decisions made by the Ministry of Infrastructure. The growth of container traffic in the Port of Koper as well as the beginning of construction of the new container terminal has made the reconstruction and extension of the current container terminal an absolute priority. The extension is in line with the estimated growth of traffic as well as with the exploitation of present and future terminal capacities.
In the article we will analyze the throughput in NA ports – Koper, Trieste, Venice, Ravenna and Rijeka in the last twenty years and based on this numbers we find a simple but efficient model to forecast the possibility of increasing or decreasing of traffic in next year. We used the same methodology to forecast the traffic volume for the next year for port of Koper. The dynamical model includes two characteristics: relative growth and market share.

## 2. Container throughput in North Adriatic ports

In the analysed period of twenty two years 1991-2013 the total container traffic in the NA ports has increased almost exponentially, on average 7% per year (this was lower than average of all European ports) but the rate has varied among ports.

[1] Corresponding author: Elen.Twrdy@fpp.uni-lj.si

In the Figure 1 we can see the fastest growth of container traffic was recorded at the Port of Koper, at an average of 14% per year; in Venice the growth was constant; while at Ravenna the traffic barely increased at all. The minimum throughput was and remains at the Port of Rijeka, which lost a great deal of traffic due to the state of war in Croatia; since about 2003 the increase in Rijeka's container throughput has been more in line with that of Koper, Trieste, and Venice.

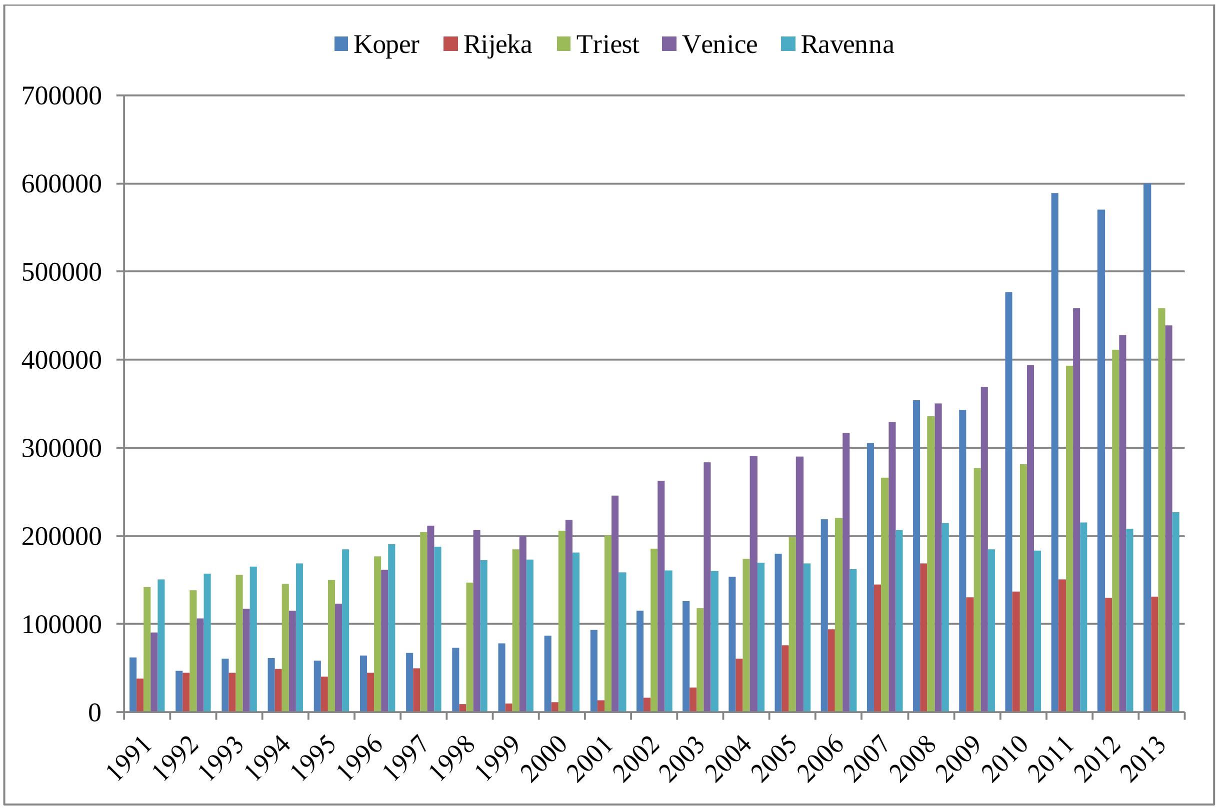



**Fig. 1:**
*Container throughput in NAPA ports during 1991-2013 (1000 TEUs)*

In the year 2008 and 2009 when the financial and economic crises has started the throughput in the North Adriatic ports was decreased for about 15 % but in this situation only the Port of Venice still had a slow growth (5%). The largest drop in traffic was recorded in Trieste, a decrease of more than 58,000 TEUs (17.5 %), though by percentage Rijeka fared worse, declining at the rate of 22.5 % (38,000 TEUs fewer).
The overall throughput for the period 1992-2013 is shown on graph on Figure 2. In accordance with this graph we prepared the simple two state systems. We take in consideration the state of the market - it could be in the phase of growth or fall. From the graph we can see that in analyzed period the market was 9 times in the state of fall and 11 times in the state of growth. Now we can count that when market is in the state of growth there is four times case that market next year will fall, and five times case that market insist growing. If we have the state when market has fall down than we have six cases when the market has growth following year and in only two cases when market has fall down following year.

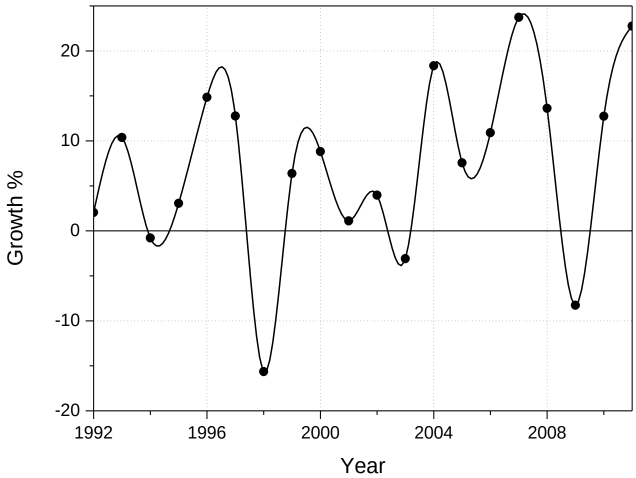



**Fig. 2.**
*Relative growth of throughput in NA ports*

**The dynamics model of total container traffic in the NA ports**
The dynamical model of this behavior is shown on Figure 3and 4 where together with twenty year observation the ten year observation is shown. It is seen from the model that ten and twenty year prediction for the transition from state of market fall is same, while probability of transition from market growth to market fall state decrease a bit. This indicates that the system has possible relatively stable behavior.

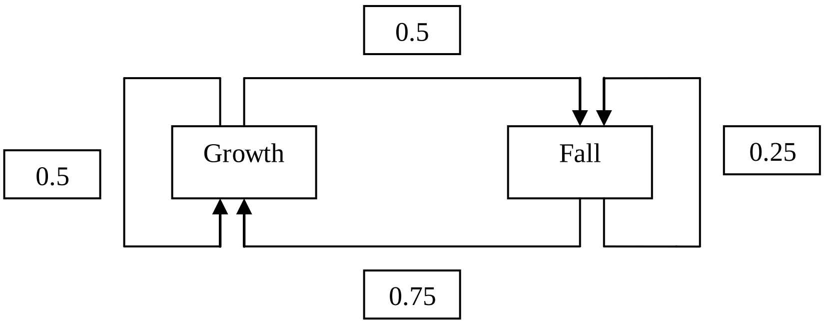



**Fig. 3.**
*Dynamics of relative growth of total market (1992-2002)*

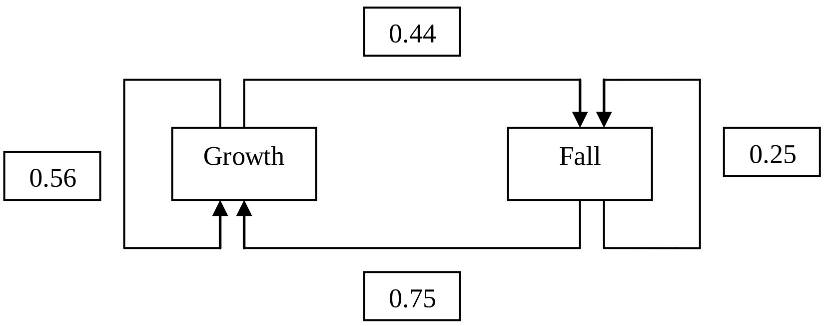



**Fig. 4.**
*Dynamics of relative growth of total market (1992-2011)*

From the model we can see that if we are in the state of market growth then there are practically equal chances the next year the market will grow or fall, however if the system is in the state of fall there are 75% chances that in next year market will grow again. Also the chance that market fall two successive years is about 20% and that fall successively three years the chance is about 5%. The chance that market grows for two successive years is about 30%, and that grows for three successive years is about 18%.

As the market potential for the NA ports in the container market in 2030 appears to be ambitious in terms of the absolute growth it implies +348% traffic growth from 2010 compared to 73% growth in the market as a whole and in terms of market share growing from the current 4.3% to reach 11.3% in 2030 (3).

**3. Container throughput for port of Koper (PK)**

In this section we will develop the dynamical model for port of Koper by similar procedure as for the overall throughput of NA ports. The biggest share of throughput among the NA ports was reached by the port of Koper probably due to the expansion of terminal capacities and good market orientation. According to the presented date we prepare models of growth and models of share.
The Figure 3 and 4 shows the graph of relative growth of containers throughputs in NA ports. By setting two states of share growth (or better not fall) and share fall we arrive to the model displayed on Figure 5.

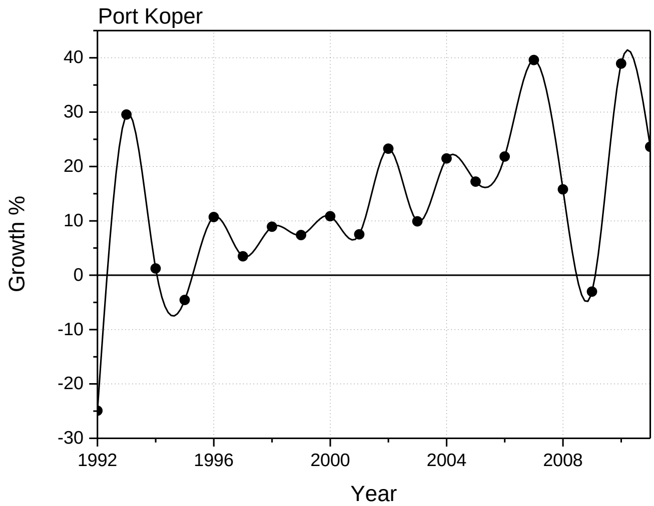



**Fig. 5.**
*Relative growth of throughput in port of Koper*

Again we observe relative similar behavior of system for ten and twenty year of observation. In particular the transitions from fall state are practically the same. From the model we see that most probable behavior is that port state alternate between yearly grow and fall. For two successive years the most unlikely is successive growth which is about 1%, while transition between states is about 70% which is most likely.

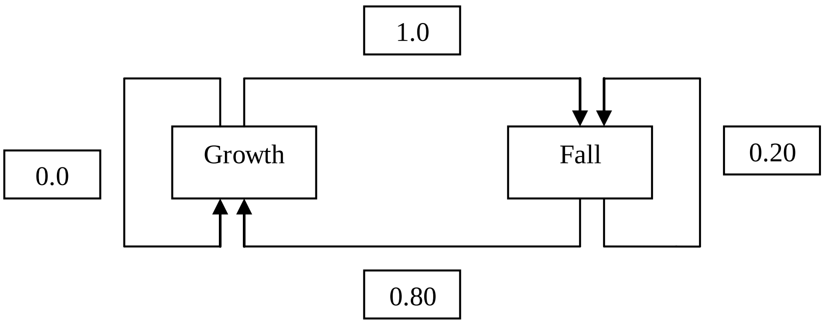



**Fig. 6.**
*Dynamics of relative growth of market for port of Koper (1992-2002)*

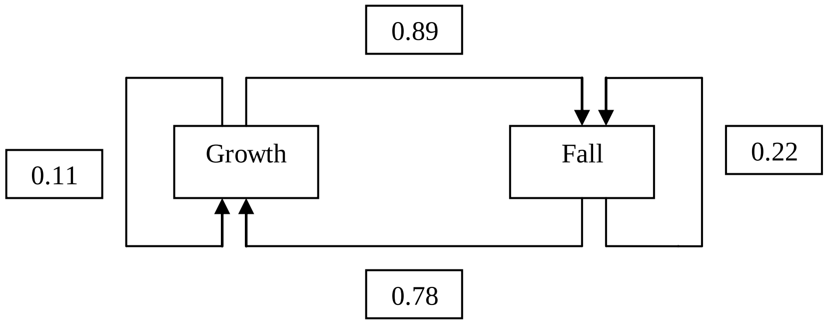



**Fig. 7.**
*Dynamics of relative growth of market for port of Koper (1992-2011)*

It is also interesting to compare total throughput growth with the growth obtained in port of Koper. We see by counting the cases that at eight cases when total throughputs growth this is the same in the port of Koper - throughputs growth, in seven cases when both throughput fall, three cases when throughput of Koper decrease and total relative throughput increase and only there cases when throughput of Koper increase while total throughout decrease. From the graph on Figure 1 we see that in twenty years the total market was in a state of growth for nine years. On the base of these observations we can construct probability tree shown on Figure 8. From the data on the figure follows a bit surprising conclusion: the total probability that throughput in port of Koper (PK) will increase is about 44%, and that will decrease is about 56%. (If we count the realized numbers over the period we will get 47% and 53% which is in very good agreement with these predicted total probabilities).

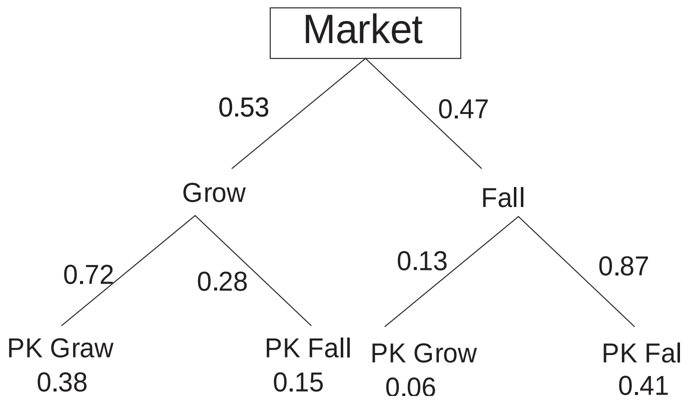



**Fig. 8.**
*Probability tree for port Koper (PK) throughput grow*

The similar model is now established for port of Koper marked share. The dynamics of market share is displayed on graph on Figure 9. Counting the numbers of share decrease and increase yield the model shown on Figure 10. Now it is seen from the Figure 9 that market share practically grows all over the observed years.

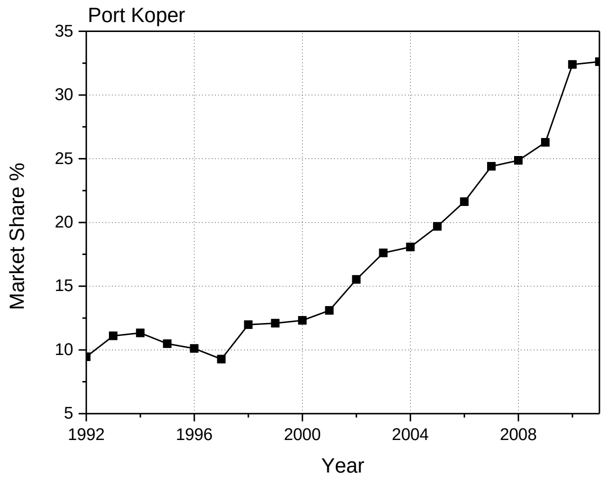



**Fig. 9.**
*Market share of port of Koper*

This is also reflected in the model shown on Figure 10 and 11. There are about 6% chances that market – share once port is in the state of share growth – will fall. However the chance that growth is again established is about 33 %. This means that only after three successive years of shares falling the chance to transit to share growth is again more probable.

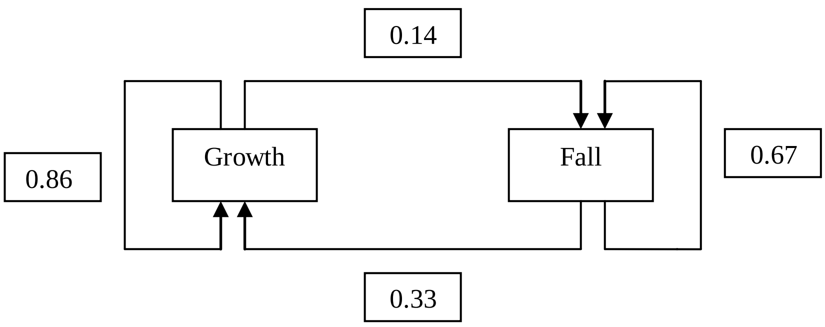



**Fig. 10.**
*Dynamics of market share of port of Koper (1992 – 2002)*

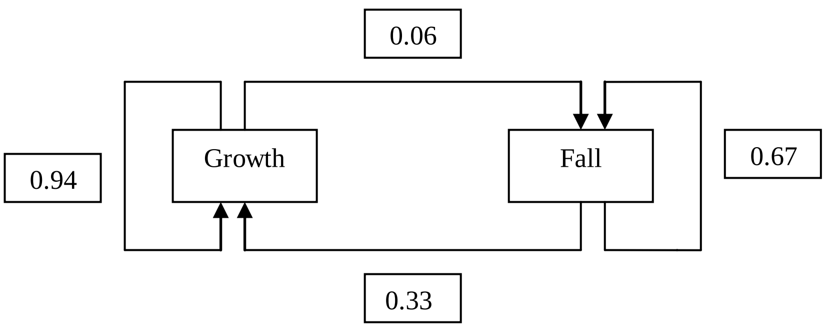



**Fig. 11.**
*Dynamics of market share of port of Koper (1992 – 2011)*

Now we build the probability model that will answer to the question if market share increase with growth of throughput. Again if we count the number of case we get the probability tree shown on Figure 12. From that tree we conclude that total probability that marked share of port of Koper will increase is about 89%, so there is only 11% chance that it will fall down.

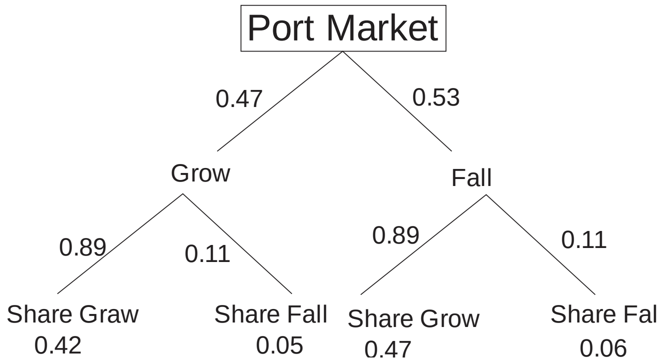



**Fig. 12.**
*Probability tree for port of Koper (PK) share*

## 4. Conclusions

Container throughput in Europe has change in last year's partly because of economic crises, market orientation of consumers and partly because of different dynamic structural changes in logistics. Although the total container traffic in the northern Adriatic ports increased in recent years it still represents a negligible proportion in total throughput of the European ports.

The data indicate that container traffic in NA ports in the European Common throughput shows a slight increase – in 2008 it was 1.6 percent and it amounted to almost 2 percent in 2011. In the proportion - the throughput of all NA ports present just 15.2 percent of the throughput, which has created Europe's largest port Port of Rotterdam in 2011, and 16 percent of the throughput in 2013. New trends in maritime transport favour the use of big container ships (economy of scale) and the ports on NA will have to join the forces to attract shipping lines to this part of Mediterranean. Consequently, collaboration and competition between NA ports will remain, and even more, it will probably be intensified in the future. Even if the NA multiport region is one of the smallest Europe it lost just a small part of throughput of container and in year 2010 they obtain the highest volume of throughput than in year 2008 – before the recession started. In this condition port of Koper had a biggest growth of throughput.

In the article we conduct several simple models that help us to forecast a state of port Koper. Among others it is shown that most probable state of growth of throughput is annually alternating between grow and fall, while with relatively high probability port Koper will not lose its market share. The authors are aware that more data should be available to build more reliable models.

**References**

United Nations Conference on Trade and Development: Review of maritime transport 2013, UNTCAD/RMT/2013, ISSN 0566-7682, United Nations, New York, Geneva, 2013. 204 p.

NAPA: Market study on the potential cargo capacity of the North Adriatic ports system in the container sector, Draft report, MDS Transmodal Limited, 2011. 66 p.

NOTTEBOOM, Teo. Concentration and the formation of multi-poort gateway regions in the European container port system: an update, Journal of Transport Geography 18 (2010), 567-583

TWRDY, Elen, BATISTA, Milan. A dynamical model of container throughput of the North Adiatic multiport gateway region. *Pomorstvo*, ISSN 1332-0718, 2013, vol. 27, no. 2,. 361-367.

TWRDY, Elen, TRUPAC, Igor, KOLENC, Jurij. Container boom in the Port of Koper. *Promet*, ISSN 0353-5320, 2012, vol. 24, no. 2, 169-175

GI-TEA, Yea et al, Evaluating the competitiveness of container ports in Korea and China, Transport Research Part A, Vol. 42 (2008), 910-921

WEI, Yim Yap et al, Developments in Container Port Competition in East Asia, Transport Electronic materials
North Adriatic Ports Association, Availabile at http://www.portsofnapa.com

*Transport Safety and Security*

# ROAD SAFETY EVALUATION USING TRAFFIC CONFLICTS: PILOT COMPARISON OF MICRO-SIMULATION AND OBSERVATION

**Jiří Ambros[1], Richard Turek[2], Jiří Paukrt[3]**
[1, 2] *Centrum dopravního výzkumu, v.v.i., Líšeňská 33a, 636 00 Brno, Czech Republic*
[3] *SIAS Limited, Opletalova 55, 110 00 Prague, Czech Republic*

**Abstract:** For local road safety evaluations traffic accident frequency has been traditionally used; however it has been also known that accident occurrence is statistically rare and thus their data collection is time consuming. To this end various other means have been investigated, including observation of traffic conflicts (near-accidents). Traffic conflict happens when two road users approach each other to such extent that, if no action is taken, a collision occurs. These surrogate measures may enable collection of larger samples and quicker safety assessment.
There have been various traffic conflict techniques developed around the world, using different approaches to assessing the conflict types and severity levels. For example Czech traffic conflict technique uses assessment based on the severity of evasive manoeuvre assessed by observers on the site.
However using evaluations by human observers has been criticized as subjective and potentially biased. Among others the use of microscopic traffic simulation was thus attempted. Micro-simulations are usually used to estimate the operational performance of road networks; their traffic safety applications have appeared more recently.
The paper presents the pilot study whose objective was to compare the results of simulated traffic conflict data with the observation using Czech traffic conflict technique. A busy signalized junction in Brno was chosen for this purpose. Software S-Paramics was used for the micro-simulation, employing the severity indicator of time-to-conflict, in which two vehicles will collide if no actions are taken. To test the results reliability, traffic conflicts were both observed and simulated in two terms with different traffic volumes (summer and autumn 2013). The paper presents the study approach and results, as well as discussion of pros and cons and practical conclusions.

**Keywords:** road safety, traffic conflict, micro-simulation.

## 1. Introduction

Traffic has been necessary part of everyday life, however its negative outcomes include traffic accidents. According to European Road Safety Observatory, each year over 1 million people are killed and 50 million injured on roads around the world. In the European Union, road accidents comprise over 90% of all transport deaths and accident costs and are the leading cause of death and hospital admission for people younger than 50 years (ERSO, 2009).

This toll includes huge socio-economical losses as well. The annual number of people dying on the European roads equals the number of inhabitants of a medium town (EC, 2010). The first step in managing the mentioned issue is the evaluation of current situation. In this paper the focus is on local road safety evaluations. As a measure traffic accident frequency has been traditionally used; however it has been also known that accident occurrence is statistically rare and thus their data collection is time consuming. To this end various other means have been investigated, including observation of traffic conflicts. Traffic conflict is internationally defined as ‘an observable situation in which two or more road users approach each other in space and time to such an extent that a collision is imminent if their movements remain unchanged’ (Amundsen and Hydén, 1977). The frequency of conflicts, considering their severity and types, may serve as an indirect safety performance indicator (surrogate safety measure). Compared to traditional indicators based on traffic accidents, conflicts are more frequent and thus enable collection of larger samples and quicker safety assessment. There have been various traffic conflict techniques (TCTs) developed around the world, using different approaches to assessing the conflict types and severity levels. Some of them use qualitative definitions, some are more quantitative; for example Older and Shippey (1980) presented more than 10 different techniques.

Nevertheless TCTs have been usually dependent on human observers; they have thus been criticized as subjective and potentially biased, with unproven or insufficient reliability and validity (Williams, 1981; Hauer and Gårder, 1986; Chin and Quek, 1997). One way of circumventing this issue has been using computer microscopic traffic simulations (micro-simulations). Traffic simulation models utilise stochastic sampling of the distributions of driver behaviour to replicate the interactions between vehicles in a traffic stream to determine the consequences of their actions (Young et al., 2014). Micro-simulations have thus been used to estimate the operational performance of road networks – they can provide estimates of traffic system capacity, delay and general flow conditions.
Their traffic safety applications – modelling the stochastic process involving driver behaviour and vehicle movement on transport infrastructure – have appeared more recently. Notwithstanding the progress in software capabilities and computing power, according to recent state-of-the-art review, road safety simulation models are still in an early stage of development. Vehicle behaviour is based on the family of car-following, lane-changing and gap acceptance models (Huguenin et al., 2005). It is also important that simulations have still focused primarily on vehicular traffic (Young et al., 2014).

[1] Corresponding author: jiri.ambros@cdv.cz

In the context of safety the idea has been to use the micro-simulation to estimate the number and the location of conflicts predicted by the model. These estimates should then be validated against observed conflicts and/or recorded accidents. One of first example applications was reported by McDowell et al. (1983): their study focused on rural T-junctions in UK. Using gap acceptance as safety indicator proved to yield the results comparable to 5-year injury accident history. Sayed et al. (1994) tested their simulation on a sample of 4-arm urban junctions in Canada and also found correlation between observed and simulated conflicts. Further studies were conducted for example by Archer (2005) on Swedish urban T-junctions, by Cunto and Saccomanno (2008) in California or by Astarita et al. (2012) in Italy.

These authors used various safety indicators to define potential conflicts: e.g. gap acceptance, time-to-conflict (TTC), post-encroachment time (PET) or deceleration rate (DR). The most common measures (TTC and PET) indicate the likelihood of accidents, while the other class of measures is more indicative of the severity of accidents (Gettman and Head, 2003).

Important step in progress of safety micro-simulations was development of Surrogate Safety Assessment Model (SSAM). This software is able to derive surrogate measures from data output (vehicle trajectories) through several compatible traffic simulation models (Gettman et al., 2008). One of them is S-Paramics software from SIAS Limited, UK. It has been used for example by Pirdavani et al. (2010) or Dijkstra (2013).

To sum up there have been several applications of using simulated conflicts to assess local safety conditions. However they used different conflict indicators and also conflicts observed with different techniques. Specifically with Czech traffic conflict technique, which uses assessment based on the severity of evasive manoeuvre assessed by observers on the site, no micro-simulation has been attempted. The pilot study reported in the paper aimed to fill this gap. Its objective was to compare the results of micro-simulation traffic conflict data with the observation using Czech traffic conflict technique. A busy signalized junction in Brno was chosen for this purpose. Software S-Paramics was used for the micro-simulation, employing the severity indicator of time-to-conflict. To test the results reliability, traffic conflicts were both observed and simulated in two terms with different traffic volumes (summer and autumn 2013). The paper presents the study approach and results, as well as discussion of pros and cons and practical conclusions.

## 2. Methodology

The study aim was to compare observed and simulated conflict frequency. A busy signalized 4-arm junction in Brno (local name ‘Nové Sady’) was selected. In order to test the results reliability two terms were chosen – summer and autumn week-day (July and September 2013). The methodology behind the two indicators is described in following sections.

### 2.1 Observed conflicts

Traffic conflicts were observed according to the Czech technique (Ambros and Kocourek, 2013). It is based on physical observation on-site or video observation in the office. Observers detect conflicts and assign them conflict types (turning, rear, front, etc.) and severity grades. The severity grades (0, 1, 2, 3, 4) are based on conflict evolution scheme – see Fig. 1.

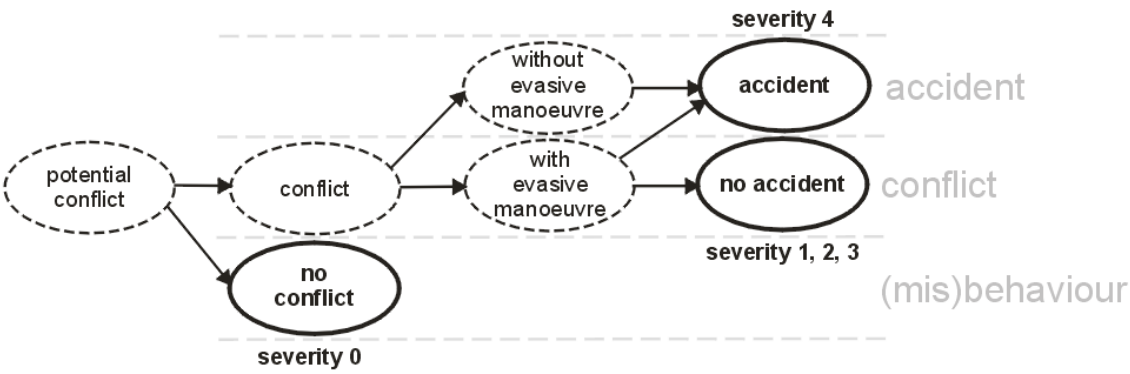



**Fig. 1.**
*Definition of severity grades in the Czech TCT within the conflict evolution scheme*

Table 1 shows the characteristics of severity grades which are assigned to observed conflict situations based on severity of an evasive manoeuvre. Situations of specific behaviour or misbehaviour have severity grade 0, since they are situations of one user only and thus do not conform to a conflict definition.

Severity grades 1, 2, 3 (highlighted in the Table 1) are assigned to conflict according to the observed evasive manoeuvre severity, together with physical reactions and other characteristics. Obstruction and endangerment, used to distinguish between 2nd and 3rd severity grade, is defined according to the Czech law (Road Act No. 361/2000 Coll.). Severity grade 4 belongs to a traffic accident with property-damage-only and/or injury consequences.

**Table 1**
*Characteristics of severity grades according to the Czech TCT (traffic conflicts are highlighted)*

| Severity grade and description | Severity | Physical reactions | Events | |
|---|---|---|---|---|
| | | | Related to vehicles | Related to pedestrians |
| 0 – (mis)behaviour | none | none | breaking the rules without consequences, misbehaviour of road users | breaking the rules, e.g. crossing outside of crossing |
| 1 – slight conflict | low | common | fluent, controlled, predictable manoeuvres | change of walking course, e.g. overtaking |
| 2 – medium conflict | obstruction | sudden | pronounced, sudden, unpredictable manoeuvres | change of walking speed, sudden entering the crossing |
| 3 – severe conflict | endan-germent | sharp | critical, emergency manoeuvres | shocking manoeuvres |
| 4 – accident | various levels (property damage only or injury consequences) | | | |

Video camera was used, recording time was between 2 and 4 hours. In the office video record was observed in order to detect conflicts and assess their severity using the above mentioned definitions.

Traffic volume data (in terms of total number of entering vehicles) were also detected from video records. Since the first term occurred in summer vacation time, traffic volumes were lower, compared to the autumn term (see Table 2).

**Table 2**
*Overview of observation and simulation parameters*

| Term | Date | Approximate hourly number of entering vehicles | Observation period | Simulation period |
|---|---|---|---|---|
| summer | Tuesday (July 9, 2013) | 1800 | 4 hours (7:00 – 11:00) | 1 hour (8:00 – 9:00) |
| autumn | Wednesday (September 18, 2013) | 2100 | 2 hours (7:30 – 9:30) | |

### 2.2 Simulated conflicts

Micro-simulation in S-Paramics was based on following input data: junction geometry, categorized directional traffic count, signal plans and queue lengths. This data was obtained during above mentioned conflict observation (on-site and from video). The model was calibrated using the observed and modelled queue lengths according to recommendations in UK Design Manual for Roads and Bridges (Highways Agency, 1996).

S-Paramics generated trajectory files for each simulated vehicle, including its position, speed, acceleration, direction and other information such as category or dimensions (defaultly twice per second). Some indicators (position, speed and direction) were used to calculate time-to-conflict (TTC). The process used was inspired by SSAM approach (Gettman et al., 2008) and Dijkstra (2010) and developed into internal S-Paramics module (Paukrt, 2010). Its functioning may be shortly described as follows:

- All combinations of modelled vehicle pairs were created. Their distances were calculated.
- Calculated distances were compared with threshold value (multiplied by 1.5), which was set as potential distance travelled in 2 seconds (critical TTC), based on speed and acceleration of involved vehicles. Distant vehicles (i.e. above threshold value) were discarded.
- Potential collision course of selected vehicle pair trajectories was determined, based on speed and size of collision area (given by vehicle dimensions). Should the vehicles trajectories intersect with TTC below 2 seconds, situation was considered a conflict.
- Conflict positions were stored and their counts summed up.

### 2.3 Comparison

Results of traffic volume survey showed the most pronounced streams. These are the two direct streams depicted in Fig. 2: they involve 16% and 20% of all entering vehicles. Since left turns are typically the most critical, the combinations A and B were chosen for the comparison.

Given the fact that micro-simulations focus on vehicles, observed conflicts including pedestrians were not used in comparison.

Only conflicts of severity grade 1, 2, 3 were used in comparison.

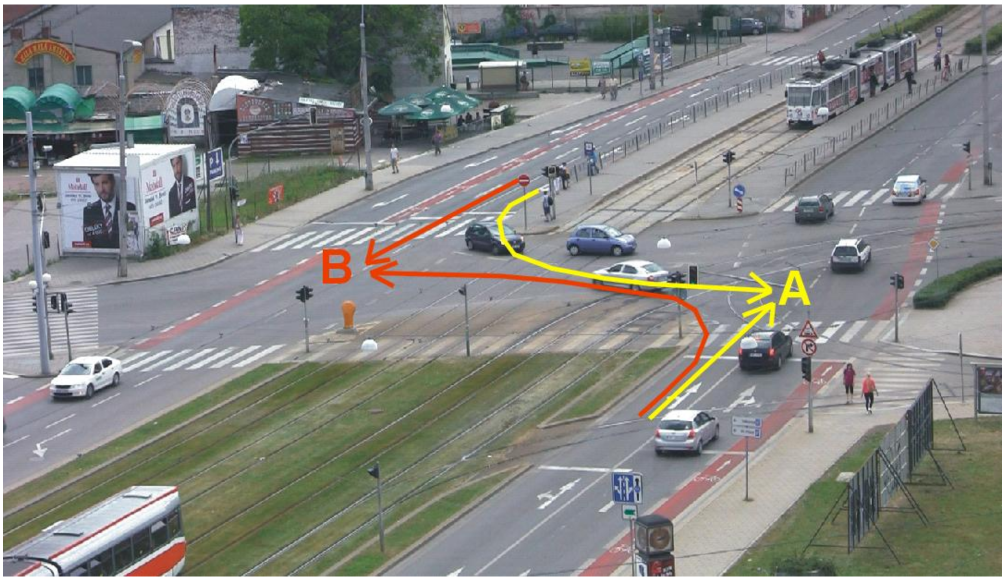



**Fig. 2.**
*View of the junction, with studied combinations of traffic streams A and B*

## 3. Results and discussion

Table 3 reports the results for summer and autumn term:

- Total observed conflict frequency and hourly value (divided by number of hours) (*O*)
- Hourly simulated conflict frequency (*S*)
- Ratio of hourly simulated conflict frequency to hourly observed conflict frequency (*S/O*)

**Table 3**
*Results of observed and simulated conflicts*

| Left turns | Summer term | | | | Autumn term | | | |
|---|---|---|---|---|---|---|---|---|
| | Observed conflicts (*O*) | | Simulated conflicts (*S*) | Ratio *S/O* | Observed conflicts (*O*) | | Simulated conflicts (*S*) | Ratio *S/O* |
| | 4 hours | 1 hour | 1 hour | | 2 hours | 1 hour | 1 hour | |
| A | 8 | 2.00 | 7 | 3.5 | 4 | 2.00 | 8 | 4.0 |
| B | 3 | 0.75 | 3 | 4.0 | 2 | 1.00 | 4 | 4.0 |

All but one ratio have value 4, which means there were 4 times more simulated conflicts compared to observed conflicts.

In addition to investigate the relation between observed/simulated conflicts and accidents, Police accident records from studied junction were retrieved. In total 20 accidents happened in junction in available time period 2009 – 2012 (4 years). However accidents and conflicts should be compared in comparable conditions (Older and Spicer, 1976), thus accidents which happened on weekends, at night, in wet conditions or with pedestrians, were excluded. In further step only accidents from time period 8 – 9 AM and studied left turns were selected. After this filtering the sample contained only 1 accident: it was an accident with left turn B.

It would be ideal to prove the validity by quantifying the relation between conflicts (both observed and simulated) and accidents (Fig. 3) in so called ‘two-stage validation process’, where firstly modelled conflicts are related to observed conflicts and secondly observed conflicts are related to accidents (Darzentas et al., 1980).

However it is obvious that accidents, being filtered according to specific conditions, are relatively infrequent. Using longer time period is a solution, however the conditions may not be stable for a longer time. This only stresses the need for surrogate safety measures using more frequent events in shorter time periods.

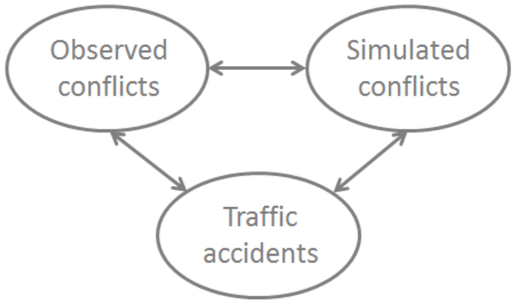



**Fig. 3.**
*Scheme of ideal validity triangulation between conflicts and accidents*

Nevertheless even the relation between observed and simulated conflicts is not definite. Some relation was found in the study (see Table 3), however probably only due to narrowing the focus to most critical left turns and excluding pedestrian conflicts.

## 4. Summary and conclusions

Traffic conflict studies provide surrogate safety measures which may complement or even substitute traditional safety studies based solely on traffic accidents. In order to overcome some of weaknesses of conflict studies, conflict micro-simulation applications have appeared recently. In this respect a link between Czech traffic conflict technique observations and micro-simulation results was sought. With this objective the reported pilot study compared observed and modelled conflict counts in signalized junction. To test the results reliability, two terms were modelled (summer and autumn 2013).

In results ratio between simulated and observed conflicts was around value of 4, i.e. there were 4 times more simulated conflicts compared to observed conflicts. The one exception (ratio 3.5) appeared in summer term, probably due to lower traffic volumes.

Although it was only a pilot study, the results are potentially promising and should be validated in other sites and conditions. Also other limitations should be considered:

- Calibration of micro-simulation model was done only by comparison of queue lengths. While it may be sufficient for capacity studies, models for safety purposes should be probably calibrated more thoroughly. However Yang and Ozbay (2011) noted a limited simulation calibration experience with an emphasis on safety evaluation – no recommended procedure exists.
- The method used (currently an in-house S-Paramics module) is using only several parameters of modelled vehicles; set of data parameters may be enlarged in order to better describe real driving behaviour.
- Presented micro-simulation model was only 1-hour long; an extension is necessary. The same hold for study limitation to the most critical turning movements and vehicle conflicts only.

These points are consistent with general objections towards micro-simulations: insufficient calibration and reliability of driving behaviour and dynamics (Wood, 2012). Simulation models have also been seen as ‘black boxes’, producing the conflicts different from the actual observed ones (Dijkstra, 2013). Nevertheless there has been large progress in this field in recent years and it was concluded that simulations will become a useful tool in analysing the safety of the traffic system (Young et al., 2014). In this context the authors of the paper are willing to continue in effort in order to enhance the quality of described tools with objective of practical proactive road safety assessment.

## Acknowledgements

The study was conducted with support of the Technology Agency of the Czech Republic project TA01030096 “Czech Traffic Conflict Technique Methodology“ (KONFLIKT). Help of following colleagues with observations is also acknowledged: Soňa Fenclová, Ondřej Gogolín, Martin Kovář, Jan Křenek.

**References**

Ambros, J.; Kocourek, J. 2013. *Metodika sledování a vyhodnocování dopravních konfliktů*. Transport Research Centre, Brno and Czech Technical University in Prague. Czech Republic.

Amundsen, F.H.; Hydén, C. (Eds.). 1977. *Proceedings: First workshop on traffic conflicts*. Institute of Transport Economics, Oslo. Norway.

Archer, J. 2005. *Indicators for traffic safety assessment and prediction and their application in micro-simulation modelling: A study of urban and suburban intersections*. PhD thesis. Royal Institute of Technology, Stockholm. Sweden.

Astarita, V.; Guido, G.; Vitale, A.; Giofré, V. 2012. A new microsimulation model for the evaluation of traffic safety performances, *European Transport*, (51): 1-16.

Chin, H.-C.; Quek, S.-T. 1997. Measurement of traffic conflicts, *Safety Science*, 3(26): 169-185.

Cunto, F.; Saccomanno, F.F. 2008. Calibration and validation of simulated vehicle safety performance at signalized intersections, *Accident Analysis and Prevention*, 3(40): 1171-1179.

Darzentas, J.; Cooper, D.F.; Storr, P.A.; McDowell, M.R.C. 1980. Simulation of road traffic conflicts at T-junctions, *Simulation*, 5(34): 155-164.

Dijkstra, A. 2010. Do calculated conflicts relate to observed crashes? Retrieving conflicts from S-Paramics. In *The 10th Annual S-Paramics User Group Meeting & Conference*.

Dijkstra, A. 2013. Assessing the safety of routes in a regional network, *Transportation Research Part C*, (32): 103-115.

European Comission (EC) 2010. Towards a European road safety area: policy orientations on road safety 2011-2020. Available from Internet: <http://ec.europa.eu/transport/road_safety/pdf/com_20072010_en.pdf>.

European Road Safety Observatory (ERSO) 2009. Road Safety Management. Available from Internet: <http://ec.europa.eu/transport/road_safety/specialist/knowledge/pdf/road_safety_management.pdf>.

Gettman, D.; Head, L. 2003. *Surrogate Safety Measures from Traffic Simulation Models: Final Report*. Publication No. FHWA-RD-03-050. Federal Highway Administration, McLean. USA.

Gettman, D.; Pu, L.; Sayed, T.; Shelby, S. 2008. *Surrogate safety assessment model and validation: Final report.* Report No. FHWA-HRT-08-051. Federal Highway Administration, McLean. USA.

Hauer, E.; Gårder, P. 1986. Research into the validity of the traffic conflicts technique, *Accident Analysis and Prevention*, 6(18): 471-481.

Highways Agency 1996. Design Manual for Roads and Bridges. Volume 12: Traffic Appraisal of Roads Schemes. Part 1: Traffic Appraisal in Urban Areas. Available from Internet: <http://www.dft.gov.uk/ha/standards/dmrb/vol12/section2/12s2p1.pdf>.

Huguenin, F.; Torday, A.; Dumont, A.-G. 2005. Evaluation of traffic safety using microsimulation. In *Proceedings of the 5th Swiss Transport Research Conference*.

McDowell, M.R.C.; Wennell, J.; Storr, P.A.; Darzentas, J. 1983. *Gap acceptance and traffic conflict simulation as a measure of risk.* Supplementary Report 776. Transport and Road Research Laboratory, Crowthorne. UK.

Older, S.J.; Shippey, J. (Eds.) 1980. *Proceedings of the Second International Traffic Conflicts Technique Workshop*. Supplementary Report 557. Transport and Road Research Laboratory, Crowthorne. UK.

Older, S.J.; Spicer, B.R. 1976. Traffic Conflicts – A Development in Accident Research, *Human Factors*, 4(18): 335-350.

Paukrt, J. 2010. *Road Safety Assessment*. SIAS Technical Note.

Pirdavani, A.; Brijs, T.; Bellemans, T.; Wets, G. 2010. A simulation-based traffic safety evaluation of signalized and un-signalized intersections. In *Road Safety on Four Continents*.

Sayed, T.; Brown, G.; Navin, F. 1994. Simulation of traffic conflicts at unsignalized intersections with TSC-Sim, *Accident Analysis and Prevention,* 5(26): 593-607.

Williams, M.J. 1981. Validity of the traffic conflicts technique, *Accident Analysis and Prevention*, 2(13): 133-145.

Wood, S. 2012. Traffic microsimulation – dispelling the myths, *TEC Magazine,* October 2012: 339-344.

Yang, H.; Ozbay, K. 2011. Calibration of Micro-Simulation Models to Account for Safety and Operation Factors for Traffic Conflict Risk Analysis. In *3rd International Conference on Road Safety and Simulation (RSS2011).*

Young, W.; Sobhani, A.; Lenné, M.G.; Sarvi, M. 2014. Simulation of safety: A review of the state of the art in road safety simulation modelling, *Accident Analysis and Prevention*, (66): 89-103.

# REAL-TIME CRASH RISK PREDICTION MODELS USING LOOP DETECTOR DATA FOR DYNAMIC SAFETY MANAGEMENT SYSTEM APPLICATIONS

**Ali Pirdavani[1], Maarten Magis[2], Ellen De Pauw[3], Stijn Daniels[4], Tom Bellemans[5], Tom Brijs[6]; Geert Wets[7]**

*[1]Postdoctoral Researcher, Research Foundation-Flanders, Egmontstraat 5, B-1000 Brussels, Belgium; and Transportation Research Institute, Hasselt University, Wetenschapspark 5–Bus 6, B-3590 Diepenbeek, Belgium*
*[2]Master student, Hasselt University, Campus Diepenbeek, Agoralaan gebouw D, B-3590 Diepenbeek, Belgium.*
*[3]PhD candidate, Transportation Research Institute, Hasselt University, Wetenschapspark 5–Bus 6, B-3590 Diepenbeek, Belgium.*
*[4, 5, 6, 7] Professor, Transportation Research Institute, Hasselt University, Wetenschapspark 5–Bus 6, B-3590 Diepenbeek, Belgium.*

**Abstract:** There is a growing trend in development and application of real-time crash risk prediction models within dynamic safety management systems. These real-time crash risk prediction models are constructed by associating crash data with the real-time traffic surveillance data, collected by loop detectors. The objective of this paper is to develop a real-time prediction model that will potentially be utilized within safety management systems. This model aims to predict the traffic safety condition of a motorway. Given that the dependent variable (i.e. traffic safety condition) is considered dichotomous ("no-crash" or "crash"), the binary logistic regression technique is selected for model development. The crash and traffic data used in this study were collected between June 2009 and December 2011 on a part of the European route E313 in Belgium between Geel-East and Antwerp-East exits, on the direction towards Antwerp. The results of analysis show that several traffic flow characteristics such as standard deviation of speed and occupancy at the upstream loop detector, and the difference in average speed on upstream and downstream loop detectors are significantly contributing to the crash occurrence prediction. The final chosen model is able to predict more than 60% of crash occasions while it predicts more than 90% of no-crash instances correctly. The findings of this study can be used to predict the likelihood of crashes on motorways within dynamic safety management systems.

**Keywords:** Binary Logistic Regression Model; Real-Time Crash Risk Prediction; Dynamic Safety Management Systems; Traffic Surveillance Data

## 1. Introduction

In the recent years, proactive traffic management systems have increasingly attracted researchers and policy makers' attention. These systems, which are mainly implemented on motorways, are meant to improve traffic safety by smoothening the traffic flow. In such dynamic safety management systems, real-time crash risk prediction models are major elements. These models estimate the likelihood of crash occurrence by using real-time traffic flow characteristics that are collected by traffic surveillance systems such as loop detectors. These models can dynamically evaluate the traffic safety condition of motorways and identify crash conditions that would potentially lead to crash occurrence. When such crash condition is identified, proactive safety countermeasures can be implemented to alleviate crash occurrence risk. Among others, variable speed limits (Lee et al., 2004, 2006c; Abdel-Aty et al., 2006; Jo et al., 2012), ramp metering (Abdel-Aty et al., n.d.; Lee et al., 2006b) and intelligent speed adaptation (Chen et al., 2002; Carsten and Tate, 2005; Servin et al., 2008; Lai et al., 2012) are effective measures that are known to improve the traffic safety. These measures are intended to smoothen the traffic flow, increase average time headways, reduce speed variation and subsequently improve the traffic safety. For instance, safety benefits will be gained by simultaneously lowering down the speed upstream and increasing the speed downstream of the location where a crash condition is identified by the real-time crash risk prediction models.

In recent years, several studies were conducted where real-time crash risk prediction models were developed by associating real-time traffic flow data with crash data (Lee et al., 2002, 2003; Chang and Chen, 2005; Oh et al., 2005; Abdel-Aty et al., 2006; Lee et al., 2006c; Oh et al., 2006; Pande and Abdel-Aty, 2006a; Abdel-Aty et al., 2007; Zheng et al., 2010; Pande et al., 2011; Xu et al., 2012; Abdel-Aty et al., 2012; Xu et al., 2013; Ahmed and Abdel-Aty, 2013).

The most commonly used modeling technique in developing real-time crash prediction model is logistic regression. Lee et al. (2002, 2003) investigated a number of traffic flow characteristics that would be linked with crash occurrence on the Gardiner Expressway in Toronto. They developed an aggregate log-linear model by associating the crash precursor variables with crash data observed over a period of 13 months.

The results of their analysis revealed that the variation of speed, speed difference between upstream and downstream loop detectors and traffic density are significant predictors of crash occurrence. In a study conducted by Abdel-Aty et al. (2004) matched case-control logistic regression models were developed to associate real-time traffic flow characteristics with crash likelihood.

[1] Corresponding author: ali.pirdavani@uhasselt.be

The results of this study showed that the likelihood of crash occurrence would be predicted by speed variation on downstream loop detector station and the average occupancy on the upstream loop detector station. Abdel-Aty et al. (2005) further extended their study by developing crash prediction models under low speed and high speed traffic regimes. Zheng et al. (2010) also used a matched case-control logistic regression model to assess the impacts of speed variance on crash occurrence, whereas Ahmed and Abdel-Aty (2012) used this technique to develop a real-time crash risk prediction model. Hourdos et al. (2006) developed a binary-response logistic model where crash-prone conditions were identified by using real-time traffic characteristics. The results of analysis showed that the crash likelihood increases when the difference in speed variability increases. In another study conducted by Lee et al. (2006a) the real-time traffic factors associated with sideswipe crashes were investigated by means of logistic regression models. The results revealed that the variation in speed and the variation in flow were among the most significant predictors of sideswipe crashes. Golob et al. (2008) investigated the level of safety on freeways, based on data collected from single loop detectors and by means of multinomial logistic regression models. The results showed significant relationships between traffic flow data and crash likelihood. Recently, Xu et al. (2013) utilized binary logistic models to predict crash likelihood and severity. In their study, several traffic characteristics such as the upstream occupancy, the upstream speed variance, the downstream speed variance and etc. were found as significant predictors of crash likelihood.

Besides application of logistic regression technique, other approaches were followed to associate crash likelihood and real-time traffic flow characteristics. Oh et al. (2005) developed a Bayesian model in which the standard deviation of speed was found to be an appropriate predictor of hazardous traffic conditions. Abdel-Aty and Pande (2005) applied the probabilistic neural network model to predict crash occurrences on freeways. In another study conducted by Chang et al. (2005) classification and regression tree technique was employed to analyze freeway accident frequency. The results of this study revealed that the average daily traffic and precipitation were the two major determinants of crash occurrence. Pande et al. (2006b) also developed a crash risk prediction model based on the classification tree and neural network while random forests technique was used by Pande et al. (2011). In the context of real-time crash prediction modeling, there are also other employed techniques such as Bayesian semi-parametric Cox (Ahmed et al., 2012) and Stochastic Gradient Boosting (Ahmed and Abdel-Aty, 2013).

As can be seen from the literature, there exist different traffic flow characteristics that are associated with crash occurrence, such as speed, speed variation, traffic density, occupancy and etc. The primary objective of this research is set to investigate the possibility of developing real-time crash prediction models based on the traffic flow characteristics that are collected by loop detector stations in Flanders, Belgium. To this end, the binary logistic regression technique will be employed for model development. This is considered as the first step in realization of a proactive highway safety management system. When the developed models predict the crash condition appropriately, the transportation authorities are enabled to implement crash prediction countermeasures within available dynamic traffic management systems in order to improve the traffic safety conditions of motorways.

## 2. Data preparation

### 2.1 Study area

The study area in this research is a part of the European route E313 in Belgium between Geel-East and Antwerp-East exits, on the direction towards Antwerp. The total length of the studied road segment is about 42.5 km. However, because of problems with the loop detectors, a 10 km length segment was excluded from the study. From the starting point of the study area up to 10 km before the ending point, the E313 has two lanes (+ hard shoulder) and the speed limit is 120 km/h. In the last 10 km, the motorway has three lanes (+ hard shoulder/bus lane) and the speed limit is 100 km/h.

All crashes that are included in this study occurred between June 2009 and December 2011. In total 78 crashes occurred during this period at the study area. Due to the necessity of having precise crash occurrence time (in the order of 1 minute) and since the crash data (gathered by the police) were obtained from a different authority than the one which provides traffic flow data (Ministry of Mobility and Public Works, Flemish Traffic Center), the accuracy of crash occurrence time was double checked by matching these two datasets. To this end, for each crash record the traffic flow data derived from adjacent loop detectors were collected for a period of one hour around the crash occurrence time. Subsequently all crash records were verified whether their corresponding traffic flow data show any speed-drop event or not.

This was carried out to ensure that each crash record perfectly matches with its linked traffic flow data. If a speed-drop event is observed for a crash record, then this record is considered as a valid record and, therefore, is added to the final dataset. An example of this matching task is shown in Figs. 1, 2 and 3 where changes in traffic flow characteristics such as average speed, traffic volume and occupancy (i.e. percentage of time that a loop detector is occupied by a vehicle) are observed for a crash, which occurred at 17:11. These figures are the outputs of the program “Mindat” which are provided by the Flemish Traffic Center. This program enables users to derive different traffic flow characteristics for any specific place and time. Traffic flow data were collected from four consecutive loop detectors spaced at approximately 750 meters; one loop detector downstream (DS1) and three loop detectors upstream (US1-3). The first upstream loop detector station is named US1; the subsequent stations in the upstream direction were labeled US2 and US3.

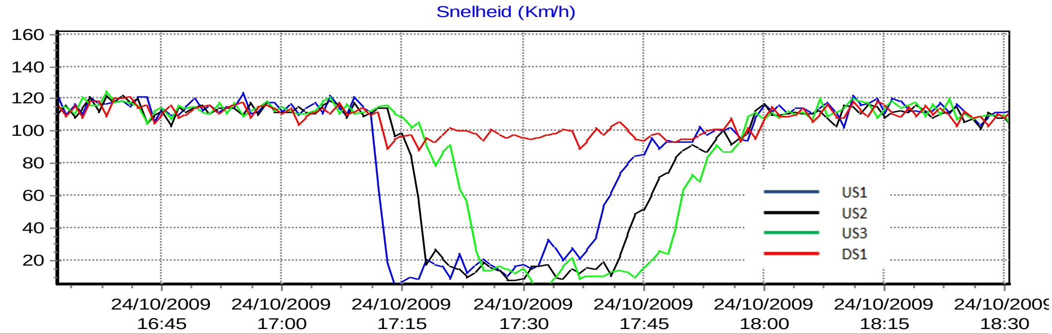



**Fig.1.** *"Mindat" program output: Speed (km/h)*

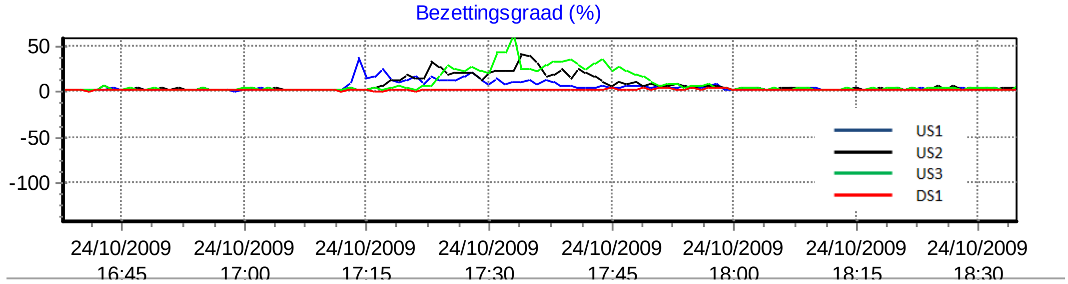



**Fig. 2.** *"Mindat" program output: Occupancy (percentage)*

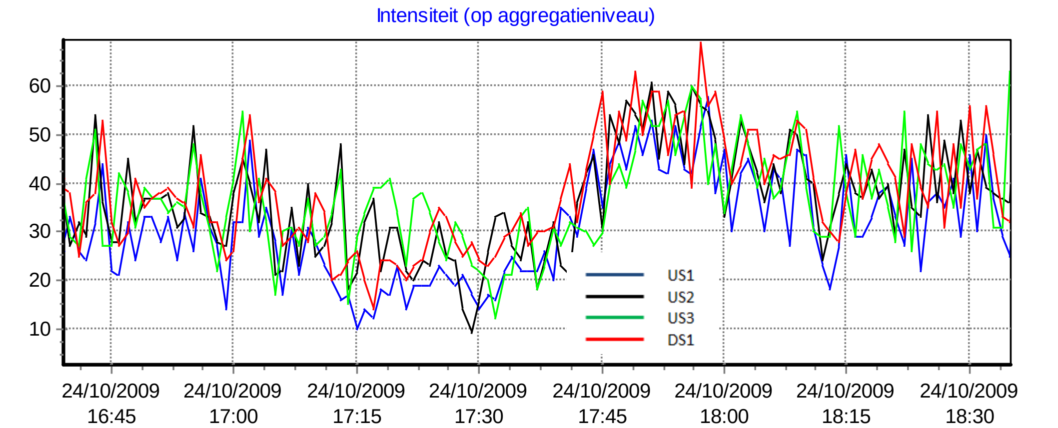



**Fig. 3.** *"Mindat" program output: Traffic volume (veh/min)*

### 2.2 Aggregation levels

The next step in data preparation is the data aggregation. The 1-minute raw data seemed to have random noise and, therefore, the primary raw data should be combined into 5-min level (Abdel-Aty et al., 2012). The extracted raw data were then aggregated to three different aggregation levels, namely 5-minute, 10-minute and 15-minute intervals prior to crash occurrence time. All these three aggregation levels will be investigated to identify the best level that will result in better crash prediction.

In the next step and for preparing the complete dataset which will be used for the modeling task, four non-crash cases were also taken from the same location, the same day of the week and the same time given the condition that no crash had occurred within one hour time period around the targeted time.

To eliminate the seasonal effects and to avoid possible bias resulting from dissimilar traffic patterns on different days of the week, non-crash cases were extracted from exactly one and two weeks before and after the crash occurrence time. All non-crash cases matched the condition that no crash had occurred within a one hour time period around the targeted time. This results in utilizing traffic flow data for the following records and for each location:

- Exactly two weeks before crash occurrence
- Exactly one week before crash occurrence
- Crash occurrence
- Exactly one week after crash occurrence
- Exactly two weeks after crash occurrence

To summarize, the final dataset consists of the traffic flow data corresponding to each crash record and four matched non-crash records. This dataset includes 390 observations (i.e. 78 crashes and 312 non-crash records).

### 2.3 Traffic flow characteristics

There are several traffic flow variables that are collected by loop detectors and might be relevant to this study. However, in order to save time and effort in the model development stage, a pre-analysis is performed to minimize the number of potential explanatory variables. To this end, firstly the non-parametric Spearman's correlation test was performed to investigate which variables do have a significant correlation with the dependent variable (i.e. safety condition). After removing all uncorrelated variables (e.g. occupancy and average speed on the first downstream loop detector station), due to the existence of inter-relationship among remained variables, the variance inflation factor test needs to be performed to ensure that the collinearity issue do not exist among explanatory variables (Kutner et al., 2004). Final variables that will be considered for model development are listed in Tables 1-3 together with their descriptive statistics. These variables are prepared for different time intervals before crash occurrence time (i.e. 5-minute, 10-minute and 15-minute interval). Here is a short definition of each variable:

TV_US1: Traffic volume on first upstream loop detector station

STDEV_SP_US1: Standard deviation of speed on first upstream loop detector station

OC_US1: Occupancy (% time that a loop detector is occupied) on first upstream loop detector station

Diff_STDEV_SP_US1-DS1: Difference between standard deviation of speed on first upstream and downstream loop detector stations

Diff_SP_US1-DS1: Difference between average speed on first upstream and downstream loop detector stations

TV_DS1: Traffic volume on first downstream loop detector station

STDEV_SP_DS1: Standard deviation of speed on first downstream loop detector station

**Table 1**
*Descriptive statistics of final variables for 5-minute interval*

| | TV_US1 | STDEV_SP_US1 | OC_US1 | Diff_STDEV_SP_US1-DS1 | Diff_SP_US1-DS1 | TV_DS1 | STDEV_SP_DS1 |
|---|---|---|---|---|---|---|---|
| Traffic safety condition: Crash (dependent variable = 1) | | | | | | | |
| Min | 0.1 | 1 | 0.3 | 0.1 | 0.2 | 0 | 0 |
| Max | 34.1 | 46.5 | 46.8 | 40.4 | 101 | 24.2 | 53.7 |
| Mean | 13.65 | 17.71 | 20.13 | 12.71 | 35.97 | 11.99 | 9.73 |
| St. Deviation | 6.66 | 11.97 | 11.47 | 10.99 | 28.21 | 6.087 | 9.31 |
| Traffic safety condition: No-crash (dependent variable = 0) | | | | | | | |
| Min | 1 | 1 | 0.3 | 0 | 0 | 0.8 | 0.9 |
| Max | 33.9 | 57.5 | 48.3 | 52.3 | 120.5 | 33.9 | 36.2 |
| Mean | 18.27 | 7.72 | 10.45 | 4.65 | 17.82 | 16.53 | 8.06 |
| St. Deviation | 6.35 | 7.38 | 7.83 | 6.53 | 27.42 | 6.99 | 6.85 |

**Table 2**
*Descriptive statistics of final variables for 10-minute interval*

| | TV_US1 | STDEV_SP_US1 | OC_US1 | Diff_STDEV_SP_US1-DS1 | Diff_SP_US1-DS1 | TV_DS1 | STDEV_SP_DS1 |
|---|---|---|---|---|---|---|---|
| Traffic safety condition: Crash (dependent variable = 1) | | | | | | | |
| Min | 0.1 | 0.8 | 0.2 | 0.2 | 0 | 1.2 | 3 |
| Max | 34.5 | 46.3 | 37.1 | 42.4 | 83.2 | 27.3 | 49.6 |
| Mean | 15.76 | 22.74 | 16.43 | 11.85 | 27.99 | 13.91 | 15.51 |
| St. Deviation | 6.72 | 11.03 | 8.88 | 9.97 | 24.43 | 6.44 | 10.71 |
| Traffic safety condition: No-crash (dependent variable = 0) | | | | | | | |
| Min | 1.2 | 1.2 | 0.4 | 0 | 0 | 1.2 | 1.5 |
| Max | 33.6 | 38.7 | 43.4 | 30.4 | 119.9 | 33.2 | 38.8 |
| Mean | 18.27 | 9.19 | 10.26 | 3.91 | 16.39 | 16.52 | 9.64 |
| St. Deviation | 6.15 | 7.55 | 7.08 | 5.21 | 26.87 | 6.81 | 7.58 |

**Table 3**
*Descriptive statistics of final variables for 15-minute interval*

| | TV_US1 | STDEV_SP_US1 | OC_US1 | Diff_STDEV_SP_US1-DS1 | Diff_SP_US1-DS1 | TV_DS1 | STDEV_SP_DS1 |
|---|---|---|---|---|---|---|---|
| Traffic safety condition: Crash (dependent variable = 1) | | | | | | | |
| Min | 1.2 | 8.9 | 0.6 | 0.1 | 0 | 0 | 0 |
| Max | 33.7 | 52.5 | 30.5 | 42.2 | 88.7 | 29 | 47.7 |
| Mean | 16.73 | 26.04 | 14.85 | 11.92 | 21.46 | 14.88 | 17.05 |
| St. Deviation | 6.01 | 10.02 | 7.25 | 10.23 | 20.08 | 6.62 | 10.57 |
| Traffic safety condition: No-crash (dependent variable = 0) | | | | | | | |
| Min | 0 | 0 | 0 | 0 | 0 | 1.3 | 1.7 |
| Max | 33 | 35.6 | 36 | 90.7 | 118.4 | 31.4 | 90.7 |
| Mean | 18.15 | 10.16 | 10.19 | 4.61 | 16.45 | 16.58 | 11.48 |
| St. Deviation | 5.94 | 7.96 | 6.74 | 10.52 | 28.06 | 6.68 | 11.70 |

## 3. Model development

### 3.1 Model structure

This study aims to predict the traffic safety condition of motorways by associating crash data with traffic flow characteristics that are collected by traffic loop detectors. Due to the dichotomous nature of the dependent variable Y (i.e. dependent variable can only take two values; Y=1 for crash condition and Y=0 for no-crash condition), application of binary logistic regression technique is appropriate. This type of model facilitates the probability estimate of being involved in a crash or no-crash condition based on the independent variables incorporated into the regression model. These independent variables can be categorical or continuous. The binary logistic regression model assumes a binomial distribution for the dependent variable. The probability of experiencing a crash or no-crash condition is modeled as the following:

$$\pi(x) = \frac{e^{g(x)}}{1+e^{g(x)}} \qquad (1)$$

The Logit transformation of the $\pi(x)$ logistic function is shown in Eq. 2:

$$g(x) = ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n \qquad (2)$$

Where $\pi(x)$ describes the probability of experiencing crash or no-crash condition. This probability falls between 0 and 1 (i.e. $0 \leq \pi(x) \leq 1$); values close to 1 signify crash conditions while values close to 0 denote no-crash conditions. $x_i$'s are the independent variables and $\beta_i$'s are the regression coefficients for each variable. These coefficient estimates determine the odds ratio of crash occurrence. The odds of an event are defined as the probability of the outcome event occurring divided by the probability of the event not occurring (Hosmer, Jr. et al., 2013). The odds ratio that is equal to the exponential of the $\beta_i$ indicates the relative amount by which the odds of the outcome increase (ratios greater than 1.0) or decrease (ratios less than 1.0) when the value of the independent variable increases by 1.0 unit.

### 3.2 Model validation technique

In this study the n-fold cross-validation technique is employed to validate the accuracy and robustness of the prediction models (Olson and Delen, 2008). The n-fold cross-validation technique minimizes the possible bias caused by the random sampling of the training and testing datasets. In the n-fold cross-validation technique, the complete dataset is equally divided into n subsets. In each step of the model development one subset is considered as for validation dataset while n-1 subsets are used as for training dataset. The cross-validation process is then repeated n times, when each of the subsamples will be used only once as the validation data. Subsequently the n results from the n developed models can be averaged or combined to deliver one single estimation. In this study, a 10-fold cross-validation approach is followed.

### 3.3 Model development

For developing the real-time prediction models, the final variables (see Tables 1, 2 and 3) were considered separately for each time interval. In other words, three binary logistic regression models were developed using explanatory variables of each time interval (i.e. 5-minute, 10-minute and 15-minute intervals). The results of analysis showed that the model developed based on 5-minute data outperforms the other two models and, therefore, is selected as the final prediction model. All models' classification results are reported in Tables 4, 5 and 6.

### 4. Model performance evaluation

The classification of results are shown in Tables 4, 5 and 6; commonly referred to as contingency table or confusion matrix. In a binary prediction problem, the outcomes are labeled either as positive or negative. In the context of this study and since the ultimate objective is to predict crash conditions, a positive outcome is set to be a crash condition while predicting a no-crash condition is considered as a negative outcome. Hence, there will be four possible outcomes by which the prediction accuracy of the model can be evaluated. If the outcome of a prediction is positive and the observed value is also positive, then this condition is considered as true positive (TP) while if the observed value is negative then it is stated to be a false positive (FP). Similarly, a true negative (TN) will occur when both the prediction outcome and the observed values are negative, and false negative (FN) is when the prediction outcome is negative while the observed value is positive. Model outcomes that are labeled with this convention are shown in Tables 4, 5 and 6.

**Table 4**.
*The contingency table of the 5-minute prediction model*

| Contingency Table[a] | | | | |
|---|---|---|---|---|
| Observed | | Predicted | | ROC = 0.853 |
| | | Traffic safety condition | | Percentage Correct |
| | | Crash | No-crash | |
| Traffic safety condition | Crash | 47 (TP) | 31 (FN) | 60.26 |
| | No-crash | 29 (FP) | 283 (TN) | 90.71 |
| Overall Percentage | | | | 84.62 |

*a. The classification threshold value is 0.4*

**Table 5**.
*The contingency table of the 10-minute prediction model*

| Contingency Table[a] | | | | |
|---|---|---|---|---|
| Observed | | Predicted | | ROC = 0.806 |
| | | Traffic safety condition | | Percentage Correct |
| | | Crash | No-crash | |
| Traffic safety condition | Crash | 38 (TP) | 40 (FN) | 48.72 |
| | No-crash | 22 (FP) | 290 (TN) | 92.95 |
| Overall Percentage | | | | 84.1 |

*a. The classification threshold value is 0.4*

**Table 6**
*The contingency table of the 15-minute prediction model*

| Contingency Table[a] | | | | |
|---|---|---|---|---|
| Observed | | Predicted | | ROC = 0.819 |
| | | Traffic safety condition | | Percentage Correct |
| | | Crash | No-crash | |
| Traffic safety condition | Crash | 41 (TP) | 37 (FN) | 52.56 |
| | No-crash | 24 (FP) | 288 (TN) | 92.31 |
| Overall Percentage | | | | 84.36 |

*a. The classification threshold value is 0.4*

The prediction performance of a binary model of outcome can be illustrated by means of a graphical plot which is called the receiver operating characteristic (ROC) curve (Olson and Delen, 2008). This graph is constructed by plotting the true positive rate (i.e. TP divided by total observed positives) against the false positive rate (i.e. FP divided by total observed negatives). Fig. 4 illustrates the ROC curve for the final prediction model. Each instance or prediction outcome of the contingency table represents one point in the ROC space. The best possible prediction would yield a point in coordinate (0,1) of the ROC space (i.e. the upper left corner) which implies no FN and no FP. Hence, the larger the area under the ROC curve, the better the prediction accuracy. The area under the ROC curve for the 5-minute model is 0.853, greater in comparison with other model results, indicating an appropriate predictive performance of this model. Moreover, it is important to distinguish between different costs imposed by different false or true predictions. In many cases optimizing the classification rate without considering the cost of the errors often leads to misleading results. A very well-known example of this would be applicable in loan decisions. As it is the case for the current study, the cost of lending to a defaulter is far greater than the lost-business cost of refusing a loan to a non-defaulter (Witten, Frank, & Hall, 2011). The same rule is applicable in the context of this study where the costs of having more false negatives is greater than false positives. The results reported in Tables 4-6 showed that the 5-minute model outperform other models by predicting less false negatives.

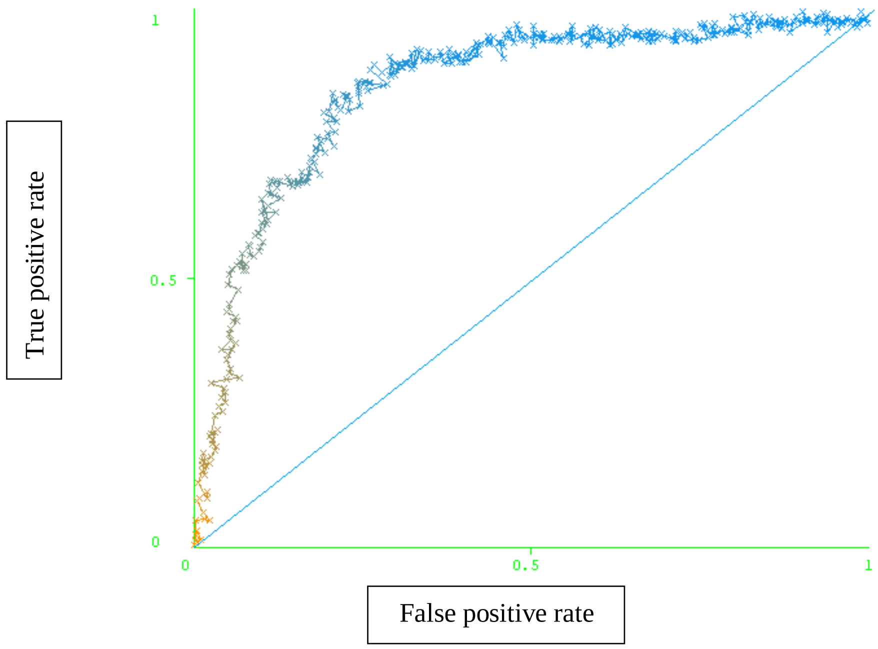



**Fig. 4.** *The ROC curve of the 5-minute prediction model*

Expectedly the prediction accuracy of the model increases if the acceptable FP rate (also referred to as false alarm rate) also increases. However, the trade-off between the prediction accuracy of the model and the false alarm rate needs to be considered and an appropriate threshold should be set by traffic authorities. To this end, the classification threshold (indicated in Tables 4, 5 and 6) can be adjusted in order to deliver appropriate classification accuracy for both traffic safety conditions (i.e. no-crash and crash conditions). Depending on the application and the definition of dependent variable, this classification threshold can be increased or decreased to provide less false positive and false negatives. In the context of this research, it would be beneficial to decrease this threshold and conservatively predict more crash occasions (even if they are not observed as crash occasions) in order to stay on the safe side. In practice and in the case of predicting a crash condition, different countermeasures can be implemented (e.g. intelligent speed adaptation or variable speed limits) to avoid a potential crash occurrence.

Although this crash condition might not be correctly predicted (i.e. it might not eventually lead to crash occurrence), implementation of safety countermeasures would be anyhow beneficial since it reduces traffic flow disturbance. With the classification threshold of 0.4, the false alarm rate of the 5-minute model is less than 10% which is significantly lower in comparison to the results of previous studies reported by Xu et al. (2013).

## 5. Model results

Based on the discussion of the previous section, the 5-minute model is selected as the final model. As can be seen from the model output in Table 7, three predictor variables have significant association with traffic safety condition.

**Table 7**
*Coefficient estimates and odds ratios for 5-minute model*

| | Coefficient estimate | p-value | Odds ratio |
|---|---|---|---|
| Intercept | -4.3078 | 0.000 | - |
| STDEV_SP_US1 | 0.1044 | 0.000 | 1.1101 |
| OC_US1 | 0.0832 | 0.000 | 1.0868 |
| Diff_SP_US1-DS1 | 0.0149 | 0.000 | 1.0151 |

The final chosen model's formulation is shown in Eq.3.

$$ln\left[\frac{\pi(x)}{1-\pi(x)}\right] = -4.3078 + 0.1044 \times \text{STDEV\_SP\_US1} + 0.0832 \times \text{OC\_US1} + 0.0149 \times \text{Diff\_SP\_US1} - \text{DS1} \quad (3)$$

Where $\pi(x)$ denotes the probability of experiencing crash or no-crash condition. All estimate signs are also in line with intuitive expectations. As can be seen from the results, standard deviation of speed upstream the crash location has a positive association with crash occurrence. This implies that higher standard deviation of speed at a location will potentially increase the risk of crash occurrence at a downstream location. Occupancy at the upstream loop detector has also a positive sign. This indicates that higher occupancy will also increase the likelihood of crash occurrence. Another significant variable is the difference in average speed of upstream and downstream loop detectors. This interesting result reveals the importance of speed and its derivative variables in crash likelihood prediction. This signifies that if the difference in average travel speed of two consecutive locations is becoming greater, there will be a higher probability of crash occurrence in between those two locations. Figs. 5, 6 and 7 illustrate the distribution of traffic safety condition in association with the final chosen explanatory variables.

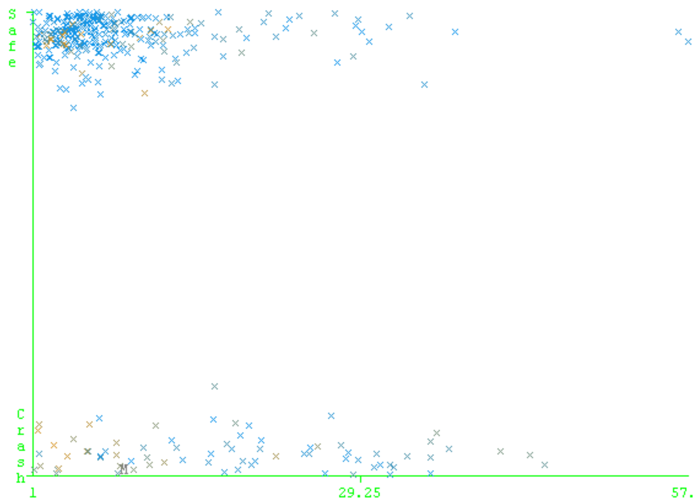



**Fig. 5.** *The distribution of traffic safety condition in association with variable STDEV_SP_US1*

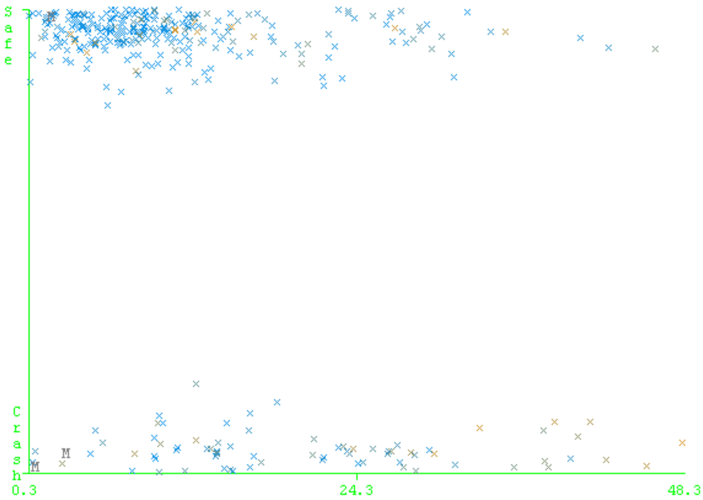



**Fig. 6.** *The distribution of traffic safety condition in association with variable OC_US1*

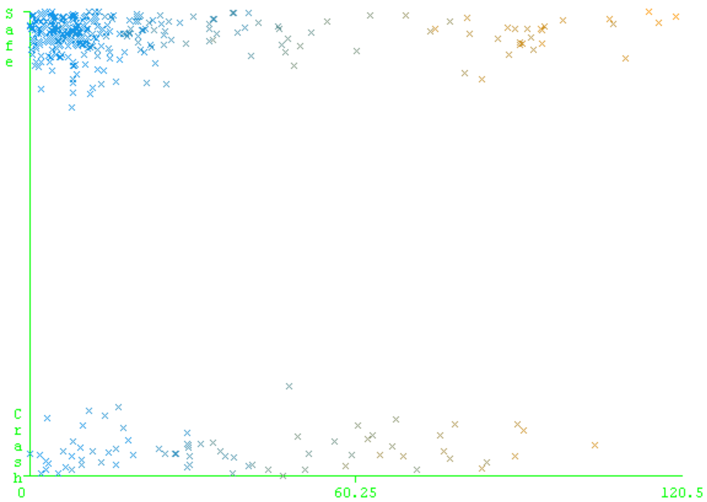



**Fig. 7.** *The distribution of traffic safety condition in association with variable Diff_SP_US1-DS1*

### 6. Conclusions and Discussion

The main objective of this study was to explore the possibility of predicting traffic safety conditions on motorways by means of traffic flow characteristics collected by loop detector stations. Various variables such as traffic volume, occupancy, average speed, standard deviation of speed, difference between average speeds on two consecutive loop detector stations were among the potential predictor variables that were considered for model development. The raw data were at 1-minute level of aggregation, which would potentially bias the results due to their random noise. To avoid this problem, the primary data were aggregated into three different levels, namely 5-minute, 10-minute and 15-minute intervals prior to crash occurrence time. This also enables us to identify the best level of aggregation that will result in better crash prediction accuracy. All of these three aggregation level data were used to develop individual prediction models by means of the binary logistic regression technique.

The results of analysis showed that the 5-minute model outperforms the other two models by means of more correctly predicted crash traffic conditions. The results showed that the 5-minute model was able to correctly predict more than 60% and 90% of crash and no-crash instances respectively. The false alarm rate (i.e. false positive in this study) resulted from the 5-minute model is less than 10%, significantly lower than false alarm rates reported in the literature. This low percentage of false alarm rate allows system users to decrease the classification threshold (i.e. currently set to be 0.4). By doing so, the number of true positive predictions (i.e. number of crash instances that are predicted correctly) will increase, meaning that the predictability of crash conditions will be improved. In return, decreasing the classification threshold yields to more no-crash occasions being incorrectly predicted as crash occasions. Depending on the application of safety management systems, this increase in false alarm rate does not impose any critical problem on road users. In application of some dynamic safety management systems (e.g. ramp metering or variable speed limits), road users are not aware of the reason behind a change in ramp metering rate or a temporary reduced speed limit. Therefore, a controlled decrease of the classification threshold will improve crash condition prediction accuracy. This trade-off between the prediction accuracy and the false alarm rate must be determined cautiously by traffic authorities based on their own specific preferences.

The first requirement in realization of a proactive highway safety management system is having accurate real-time prediction models. The performance of the developed prediction model in this study (i.e. the 5-minute model) appropriately fulfills this condition by predicting an acceptable rate of crash conditions. Having said that, there is always room for improving the accuracy level of developed model by enriching the crash and traffic data and by employing other modeling techniques. This would improve model accuracy and robustness and subsequently would increase the acceptability of the prediction model by traffic authorities who are going to utilize this model in their dynamic safety management system. Another extension for future research would be the transferability check of the developed prediction model. To this end, the model should be validated against crash and traffic data collected from various motorways. This should be carried out to ensure that the final model is able to correctly predict traffic safety conditions on any motorway under the same jurisdiction and with the same infrastructural basis (e.g. speed limit, traffic volume order or geometric conditions).

**Acknowledgements**

The authors thank the Flemish Traffic Centre for their support during the data gathering and for their collaborated efforts in conducting the survey. This research was partly supported by a grant from the Research Foundation Flanders (FWO). The content of this paper is the sole responsibility of the authors.

**References**

Abdel-Aty, M.A., Hassan, H.M., Ahmed, M., Al-Ghamdi, A.S., 2012. Real-time prediction of visibility related crashes. Transp. Res. Part C Emerg. Technol. 24, 288–298. doi:10.1016/j.trc.2012.04.001

Abdel-Aty, M., Dilmore, J., Dhindsa, A., 2006. Evaluation of variable speed limits for real-time freeway safety improvement. Accid. Anal. Prev. 38, 335–345. doi:10.1016/j.aap.2005.10.010

Abdel-Aty, M., Haleem, K., Cunningham, R., Gayah, V., n.d. Application of variable speed limits and ramp metering to improve safety and efficiency of freeways. Department of Civil, Env. & Construction Engineering University of Central Florida.

Abdel-Aty, M., Pande, A., 2005. Identifying crash propensity using specific traffic speed conditions. J. Safety Res. 36, 97–108. doi:10.1016/j.jsr.2004.11.002

Abdel-Aty, M., Pande, A., Lee, C., Gayah, V., Santos, C.D., 2007. Crash Risk Assessment Using Intelligent Transportation Systems Data and Real-Time Intervention Strategies to Improve Safety on Freeways. J. Intell. Transp. Syst. 11, 107–120. doi:10.1080/15472450701410395

Abdel-Aty, M., Uddin, N., Pande, A., 2005. Split Models for Predicting Multivehicle Crashes During High-Speed and Low-Speed Operating Conditions on Freeways. Transp. Res. Rec. J. Transp. Res. Board 1908, 51–58. doi:10.3141/1908-07

Abdel-Aty, M., Uddin, N., Pande, A., Abdalla, F., Hsia, L., 2004. Predicting Freeway Crashes from Loop Detector Data by Matched Case-Control Logistic Regression. Transp. Res. Rec. J. Transp. Res. Board 1897, 88–95. doi:10.3141/1897-12

Ahmed, M., Abdel-Aty, M., 2012. The viability of using automatic vehicle identification data for real-time crash prediction. IEEE Trans. Intell. Transp. Syst. 13, 459–468.

Ahmed, M., Abdel-Aty, M., 2013. A data fusion framework for real-time risk assessment on freeways. Transp. Res. Part C Emerg. Technol. 26, 203–213. doi:10.1016/j.trc.2012.09.002

Ahmed, M., Abdel-Aty, M., yu, 2012. A Bayesian updating approach for real-time safety evaluation using AVI data, in: The 91st Annual Meeting of the Transportation Research Boar. Washington, DC., USA.

Carsten, O.M.J., Tate, F.N., 2005. Intelligent speed adaptation: accident savings and cost–benefit analysis. Accid. Anal. Prev. 37, 407–416. doi:10.1016/j.aap.2004.02.007

Chang, L.-Y., Chen, W.-C., 2005. Data mining of tree-based models to analyze freeway accident frequency. J. Safety Res. 36, 365–375. doi:10.1016/j.jsr.2005.06.013

Chen, G., Meckle, W., Wilson, J., 2002. Speed and safety effect of photo radar enforcement on a highway corridor in British Columbia. Accid. Anal. Prev. 34, 129–138. doi:10.1016/S0001-4575(01)00006-9

Golob, T.F., Recker, W., Pavlis, Y., 2008. Probabilistic models of freeway safety performance using traffic flow data as predictors. Saf. Sci. 46, 1306–1333. doi:10.1016/j.ssci.2007.08.007

Hosmer, Jr., D.W., Lemeshow, S., Sturdivant, R.X., 2013. Applied Logistic Regression. John Wiley & Sons.

Hourdos, J., Garg, V., Michalopoulos, P., Davis, G., 2006. Real-Time Detection of Crash-Prone Conditions at Freeway High-Crash Locations. Transp. Res. Rec. J. Transp. Res. Board 1968, 83–91. doi:10.3141/1968-10

Jo, Y., Yoon, K., Jung, I., 2012. Variable Speed Limit to Improve Safety near Traffic Congestion on Urban Freeways. Int. J. Fuzzy Syst. 14, 278–288.

Kutner, M.H., Nachtsheim, C.J., Neter, J., 2004. Applied Linear Regression Models, 4th ed. McGraw-Hill, Boston; New York.

Lai, F., Carsten, O., Tate, F., 2012. How much benefit does Intelligent Speed Adaptation deliver: An analysis of its potential contribution to safety and environment. Accid. Anal. Prev. 48, 63–72. doi:10.1016/j.aap.2011.04.011

Lee, C., Abdel-Aty, M., Hsia, L., 2006a. Potential Real-Time Indicators of Sideswipe Crashes on Freeways. Transp. Res. Rec. J. Transp. Res. Board 1953, 41–49. doi:10.3141/1953-05

Lee, C., Hellinga, B., Ozbay, K., 2006b. Quantifying effects of ramp metering on freeway safety. Accid. Anal. Prev. 38, 279–288. doi:10.1016/j.aap.2005.09.011

Lee, C., Hellinga, B., Saccomanno, F., 2003. Real-Time Crash Prediction Model for Application to Crash Prevention in Freeway Traffic. Transp. Res. Rec. J. Transp. Res. Board 1840, 67–77. doi:10.3141/1840-08

Lee, C., Hellinga, B., Saccomanno, F., 2004. Assessing Safety Benefits of Variable Speed Limits. Transp. Res. Rec. J. Transp. Res. Board 1897, 183–190. doi:10.3141/1897-24

Lee, C., Hellinga, B., Saccomanno, F., 2006c. Evaluation of variable speed limits to improve traffic safety. Transp. Res. Part C Emerg. Technol. 14, 213–228. doi:10.1016/j.trc.2006.06.002

Lee, C., Saccomanno, F., Hellinga, B., 2002. Analysis of Crash Precursors on Instrumented Freeways. Transp. Res. Rec. J. Transp. Res. Board 1784, 1–8. doi:10.3141/1784-01

Oh, C., Park, S., Ritchie, S.G., 2006. A method for identifying rear-end collision risks using inductive loop detectors. Accid. Anal. Prev. 38, 295–301. doi:10.1016/j.aap.2005.09.009

Oh, J.-S., Oh, C., Ritchie, S.G., Chang, M., 2005. Real-Time Estimation of Accident Likelihood for Safety Enhancement. J. Transp. Eng. 131, 358–363. doi:10.1061/(ASCE)0733-947X(2005)131:5(358)

Olson, D.L., Delen, D., 2008. Advanced data mining techniques. Springer, Berlin.

Pande, A., Abdel-Aty, M., 2006a. Comprehensive Analysis of the Relationship Between Real-Time Traffic Surveillance Data and Rear-End Crashes on Freeways. Transp. Res. Rec. J. Transp. Res. Board 1953, 31–40. doi:10.3141/1953-04

Pande, A., Abdel-Aty, M., 2006b. Assessment of freeway traffic parameters leading to lane-change related collisions. Accid. Anal. Prev. 38, 936–948. doi:10.1016/j.aap.2006.03.004

Pande, A., Das, A., Abdel-Aty, M., Hassan, H., 2011. Estimation of Real-Time Crash Risk. Transp. Res. Rec. J. Transp. Res. Board 2237, 60–66. doi:10.3141/2237-07

Servin, O., Boriboonsomsin, K., Barth, M.J., 2008. Preliminary Design of Speed Control Strategies in Dynamic Intelligent Speed Adaptation System for Freeways. Presented at the Transportation Research Board 87th Annual Meeting.

Witten, I.H., Frank, E., Hall, M.A., 2011. Data mining practical machine learning tools and techniques, third edition. Morgan Kaufmann Publishers, Burlington, Mass.

Xu, C., Liu, P., Wang, W., Li, Z., 2012. Evaluation of the impacts of traffic states on crash risks on freeways. Accid. Anal. Prev. 47, 162–171. doi:10.1016/j.aap.2012.01.020

Xu, C., Tarko, A.P., Wang, W., Liu, P., 2013. Predicting crash likelihood and severity on freeways with real-time loop detector data. Accid. Anal. Prev. 57, 30–39. doi:10.1016/j.aap.2013.03.035

Zheng, Z., Ahn, S., Monsere, C.M., 2010. Impact of traffic oscillations on freeway crash occurrences. Accid. Anal. Prev. 42, 626–636. doi:10.1016/j.aap.2009.10.009

# ASSESSMENT AND OPINION ON THE CAUSE OF AN ACCIDENT AND DAMAGE TO A PRIVATE VESSEL THROUGH THE ANALYSIS OF A REAL-LIFE EVENT

**Tatjana Stanivuk[1], Neven Kralj[2], Frane Prolić[3]**
[1] *University of Split, Faculty of Maritime Studies, Zrinsko-Frankopanska 38, 21000 Split, Croatia*
[2] *Standing court expert, T. P. Marovica 8a, 21209 Mravince, Croatia*
[3] *Student, University of Split, Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, Ruđera Boškovića 32, 21000 Split, Croatia*

**Abstract:** A detailed examination of a completely sunken private vessel was carried out after an insurance company requested an expert opinion on the cause of the accident and the extent of the damage. The examined private vessel was a fibreglass planing boat powered by two petrol engines whose owner used her for sports and recreation. The vessel was examined on the fenced dry dock on the premises of a licensed service centre. The purpose of the examination of the vessel and the material evidence was to find out what actually happened. This paper describes in detail the entire procedure, involving the inspection of the hull, machinery space and propulsion system. In addition to presenting the assessment and conclusions, the paper features a number of facts, figures and photographs related to this real-life event.

**Keywords:** vessel, damage, accident, material evidence.

## 1. Introduction

The person who handled the herein discussed vessel at the time of the maritime accident (the skipper) made a statement to the insurance company about the circumstances and the cause of the accident. The statement explained that the vessel was under way and lost speed after 6 to 7 nautical miles. The dials at the control desk in the cockpit indicated that one of the two engines shut down so the skipper went to the stern of the vessel to open the engine room cover to see what went wrong. When he lifted the cover he found out that the machinery space was filled with smoke. The skipper also noticed that the after part of the vessel was flooding rapidly. He tried to stop the penetration of the sea water into the engine room. He tried to locate the very spot where the fracture in he hull shell was, but his right forearm got burned so he did not manage to find the spot where the water was breaking into the vessel. As the herein discussed vessel was fitted with two engines, the skipper promptly tried to start the other engine and approach the shore as close as possible. However, after a short while, the second engine shut down and the vessel started sinking. Having no choice, the skipper abandoned the vessel jumping into the sea and started to call for help, hoping that the vessels sailing nearby would notice him. According to his statement, it was only after 4 hours that he was rescued by a British sailboat while his vessel remained floating with her bow above the sea surface.

After a few days, local fishermen came across the damaged vessel by chance, fixed the spot where the water penetrated into the damaged vessel, pumped out the water from the hull and towed the vessel into the nearest harbour.

The vessel's owner and skipper lodged a claim for indemnity. Their claim was followed by the inspection of the damaged vessel (investigation into the extent of the damage).

In order to make the analysis that follows easier to understand, it should be noted that the discussed vessel was a private planing boat used for sports and recreation, built by Progression company in 1990. It was a 2.65 GT fibreglass monohull planing boat powered by two inboard Mercruiser 454 engines having the output of 380kW each. The length of this high performance boat was 9.57 m, the beam 2.10 m, and the draft 55 cm. According to all these features, the vessel belonged to the sailing category III a – inshore waters.

## 2. Inspection of the damage to the discussed vessel – investigation into the damage

As it was already mentioned, owing to local fishermen the vessel was towed to a safe harbour where the inspection of the damage (see Fig. 1) took place. It was found out that the hull platitude at the starboard bow was damaged (a crack of 6 cm in diameter as well as a 4 x 13 cm crack at the fore deck). These cracks could not be the cause of flooding as they were above the waterline (see details in paragraph 2.1.). They were neither related to the observed accident nor to the sequence of events as described by the skipper because the vessel did not have any of the damaged after the accident took place. The cracks rather looked like an attempt to entirely sink the partly foundered vessel by flooding the air pocket in the bows. However, due to the multi-layer ("sandwich") hull design, this type of vessel is practically unsinkable (Kurtela, 2000).

**Figure 1.**
*The vessel under discussion (high performance planing power boat)*
*Source: Authors*

The investigation of the deck revealed a considerable fracture of the windshield glass and a broken rail at the stern. The stern platform was well bended which might be the result of the salvage operation. A bow cleat and a navigation light were missing as well as the stern light which had been fitted in the cockpit just before the machinery space.
The inner space at the bow was entirely flooded with the interior components (upholstery, carpets and pillows) completely soaked and destroyed by sea water action (see Fig. 3.). The investigation also revealed damaged engine control instruments with visible sea water residues and the destroyed VHF station ICOM ISM-45 (see Figure 2.).

**Figure 2.**
*Control stand of the damaged vessel*
*Source: Authors*

**Figure 3.**
*Inner (flooded) space of the vessel in question*
*Source: Authors*

As expected, the examination of the engine room revealed that the propulsion units Mercruiser 454 were entirely flooded with sea water. During the actions of salvaging, delayed disassembling and immersing into oil, the steel and iron engine components were destroyed by corrosion. Engine electrical installation, batteries and other devices (starter, alternator, pumps) were entirely destroyed by the action of the sea water. By checking the oil dipsticks in both engines it was noticed that both engines were fully filled with water and suffered considerable inner damage (more details in section 2.2.).
One of the fishermen who towed the damaged vessel and pulled her out of water found out that the water broke into the vessel through the cooling pipeline in the port Z-drive propulsion pod. He fixed the fault by connecting and sealing the pipeline in order to make the vessel floating again and suitable for towing.

In order to ascertain the state of the underwater part of the hull and the propulsion pods, an additional examination of the hull, propulsion unit and equipment was carried out upon bringing the vessel to the service centre (Stanivuk et al., 2014). As for the hull, the following parameters were determined for reaching final conclusions regarding the cause of the sinking:

### 2.1. Examination of the hull

The examination of the vessel's hull ascertained that the underwater part of the hull had not suffered any damage that could eventually result in flooding the bilge and the bottom of the vessel. Only minor scratches on the gel coat surface, caused probably during salvage operation, were detected. It should be pointed out that serious damages to the hull are placed above the waterline at the bow. These damages were the result of pounding the hull with a sharp object (see Fig. 4. and Fig. 5.).

**Fig. 4.**
*Damage to the starboard side*
*Source: Authors*

**Fig. 5.**
*Damage to the starboard bow*
*Source: Authors*

The propulsion pods did not experience any major damage (Martinović, 2005). Sea water did not break into them and they can be repaired (Morton, 1974). As the stern platform is heavily bended, it should be dismounted, repaired and mounted again with additional strengthening. Except for the above mentioned items, there was not any other visible damage.

### 2.2. Inspection of the machinery space and the propulsion system

Through the examination of the machinery space it was ascertained that the vessel was fitted with two inboard Mercruiser petrol V 8 engines having the swept volume of 7.4 l and the torque transmission to the Bravo Z-drive. The basic characteristic of the propulsion plant cooling system is the direct cooling featuring the supply centrifugal impeller sea water pump placed below the engines (see Fig. 6. and Fig. 7.).


**Fig. 6.**
*Machinery space of the vessel under discussion*
*Source: Authors*

**Fig. 7.**
*Z-drive fitted with sea water inlet*
*Source: Authors*

The pump is belt driven by the crankshaft directly, the transmission ratio being 1:1, and it draws sea water through the intake strum box and aluminium braided rubber hose fitted with decorated terminals where stainless steel clamps secure the rubber hose to the strum box outlet which is designed as a barbed 19 mm tube attached to the intake strum box by two screws (Inst. Mar. Eng., 1981).
The clamp tightening test was performed at the service centre and it was ascertained that the clamp was faultless and that it could not get loose during engine operation (see Fig. 8. and Fig. 9.).

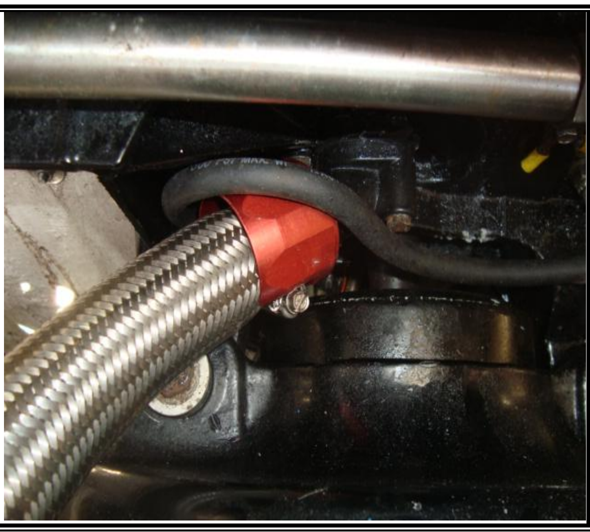


**Fig. 8.**
*Position of the cooling hose at the intake*

*Source: Authors*

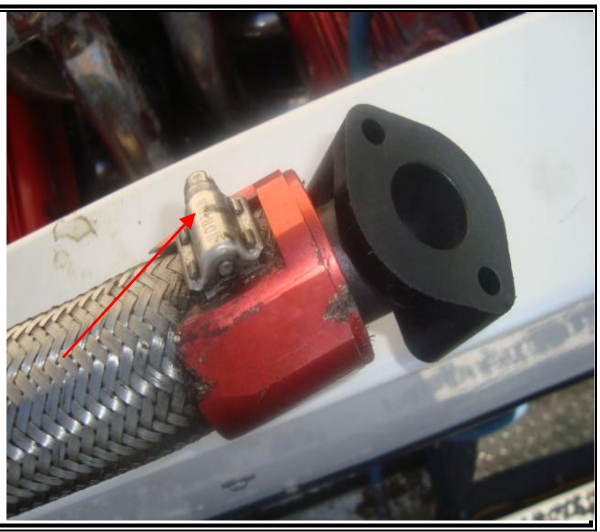


**Fig. 9.**
*The cooling hose secured to the outlet of the seawater intake by a clamp*
*Source: Authors*

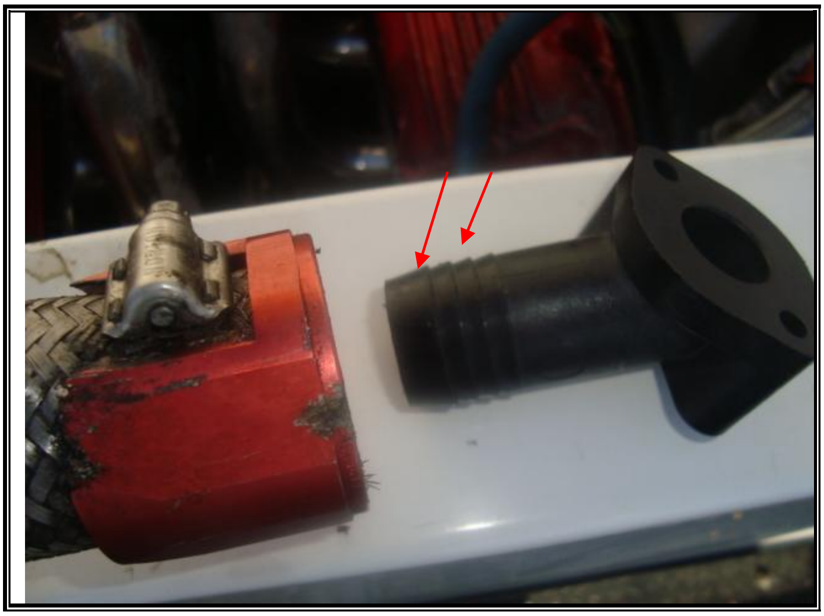


**Fig. 10.**
*Design of the outlet terminal of the seawater intake – barbed fitting to the rubber hose secured by a stainless steel clamp*
*Source: Authors*

The capacity of the seawater supply centrifugal pump at minimal speed (600–650 rpm) is about 28 l/min, whereas at full speed the engine speed amounts to about 3600 rpm and the capacity of the seawater supply centrifugal pump is around 150 l/min. The pump design is rather simple and features a Bakelite box housing a rubber impeller rotating at 1500 rpm at cruising speed (see Fig. 10).

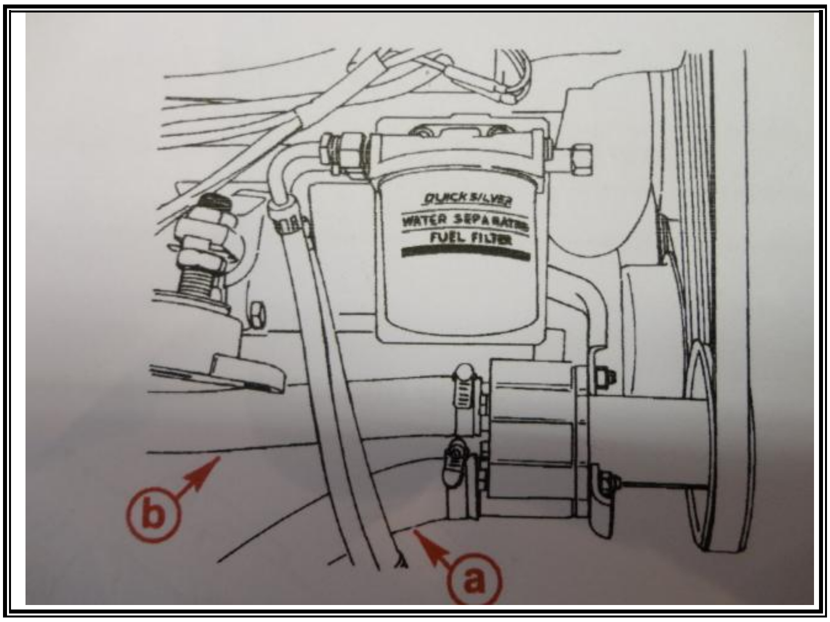



**Fig. 11.**
*Diagram of the seawater supply pump position with the intake and pressure hoses*
*Source: Authors*

The inspection of the seawater pump was performed and it was ascertained that there were no damages to the rubber impeller due to the potential lack of sea water during operation (see Fig. 12. and Fig. 13.).

**Fig. 12.**
*Disassembled sea water supply centrifugal pump*

*Source: Authors*

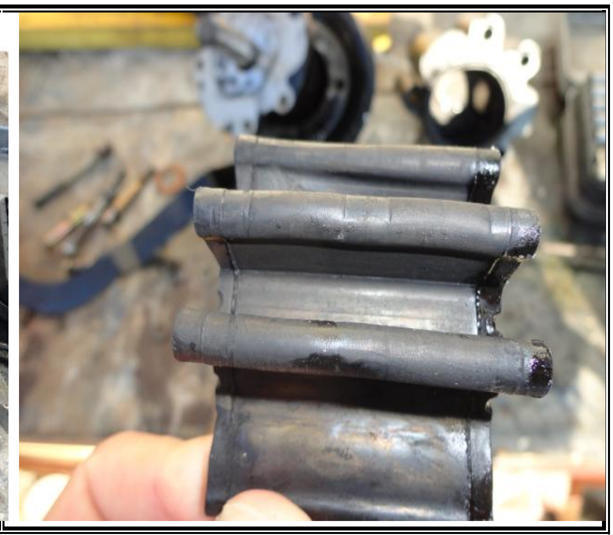


**Fig. 13.**
*The supply pump rubber impeller blades having no visible damage due to potential burnout*
*Source: Authors*

The rubber hose taking the sea water from the Z-drive pod into the hull has its terminal placed around 10 cm above the waterline (see Fig. 14.).

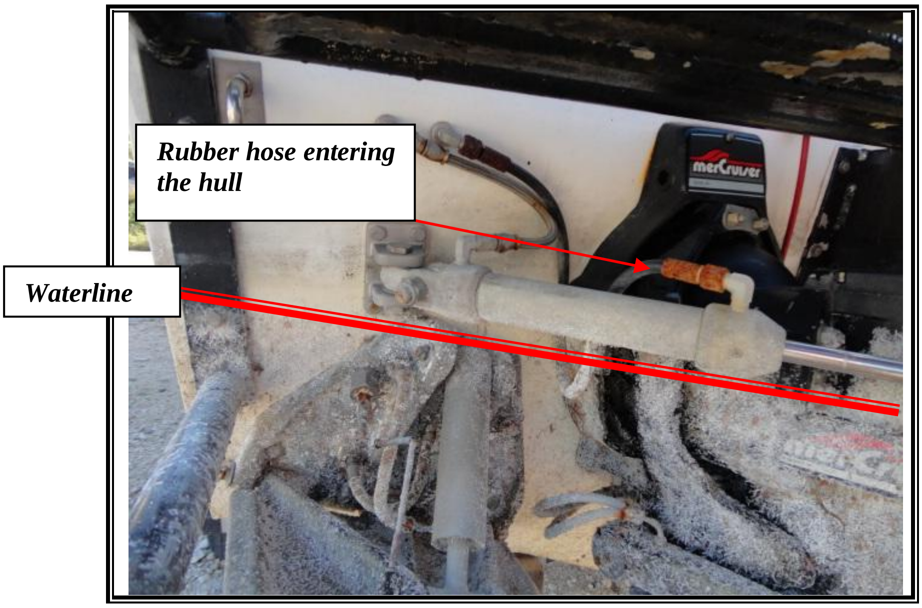



**Fig. 14.**
*Position of the waterline with regard to the vessel's hull – the transom and the place where the rubber hose enters the hull*
*Source: Authors*

Subsequently, the inspection of the bilge pump having the minimum capacity of 150 l/min was performed. It was ascertained that the pump remained in good condition without any visible faults.
As the Mercruiser 454 engine is not manufactured anymore, it is possible to replace it by Mercruiser racing 525 efi engine which has similar performances but is somewhat stronger and features new technologies so that, in case of replacement, it is necessary to replace the propulsion pods as well (http://www.mercuryracing.com/sterndrives/hp525efi.php).

## 3. Opinion

The inspection of the above described items and the skipper's statement has led to the following conclusions:

- It has been ascertained that below the waterline the hull of the vessel in question had no damages that may have resulted in flooding the cockpit and that may have eventually caused the sinking of the vessel.
- The skipper's statement affirms that the port engine shut down and that the water broke into the machinery space after approximately 6 hours of sailing and it was subsequently ascertained that the flooding was caused by the detachment of the intake hose of the port engine off the intake strum box.
- The examination of the cooling system proved that the rubber hose could not be detached by itself for the following reasons:
    - The above discussed terminal of the rubber hose is placed at the intake part where the seawater flows at 150 l/min, i.e. there is a considerable underpressure during operation, tending to press the hose towards and not to detach it from the strum box connector as the engine speed increases. Moreover, the connector has a barbed design in order to fit perfectly.
    - Therefore, in cases like this, the stainless steel clamp serves merely as a safety device that is not subjected to any load during sailing; nevertheless the testing of the clamp was performed at the licensed service centre and it was proved that the clamp is in order. Even when it is tightened to a minimum degree, it is impossible to detach the rubber hose from the barbed fitting of the intake strum box (Mrakovčić, 2005).
    - Any detachment of the rubber hose during sailing would completely destroy the rubber impeller within around 5 seconds as the impeller would be left without sea water and would rotate in a dried Bakelite pump housing. Clearly this is not the case in this accident because the inspection confirmed that the impeller was not damaged at all.
    - If there was no sea water at the pump intake, the supply of water into the machinery space would not be 150 l/min – matching the pump capacity – but only about 10 litres per minute, as a result of the pressure arising from the sailing action. Given the engine room volume, it would take a long time for the skipper to become aware of the problem; meanwhile, the entire port engine and its cooling system would have burned due to cooling failure, which was not the case in the accident under discussion (Šretner, 1970).
- Furthermore, the skipper stated that he noticed that the water broke into the machinery space when he stopped the vessel and opened the engine room cover. The measuring performed during the inspection ascertained that, when the vessel is not under way, the sea water intake is placed about 10 cm above the sea level so the hull could not be flood under these circumstances.
- Finally, through the inspection of the bilge pump having the minimum capacity of 150 l/min it has been found out that the pump was in order at the time of accident and was capable to pump out an amount of water much larger than allegedly breaking into the vessel, as reported by the skipper.

## 4. Extent of the damage

### 4.1. Partial damage

The partial damage, i.e. the partial average, refers to a complete loss of parts of the insured property (e.g. loss of anchor, propeller, etc.) or to a damage to the insured property. During the procedure of average settlement the lost items are replaced by new ones and the damaged parts are repaired and brought into their original condition.

If it results from the insured risks, the average involving a partial loss or damage to the insured property are compensated through the marine insurance policy. The lost or damaged property is compensated according to the extent of the damage, the indemnity being sufficient to replace the lost items or to repair the damaged property and bring it to the original condition. Pursuant to Article 712, alinea 1 of the Maritime Code, the indemnity does not extend above the insured value. On the other hand, the insurer shall compensate for a number of consecutive losses and damages covered by the same policy, even if the overall amount of compensation exceeds the insured value (Article 717, alinea 1 of the Maritime Code).

When establishing the extent of the average occurred due to the lost or damaged property, it is important to know whether the property is under-insured or over-insured. As a rule, the extent of compensation is proportional to the insured value. In case of under-insurance, an item is insured for less than it is worth and the average will apply to reduce the amount payable; and if the value of the insured item is not set, the compensation is proportional to the insured and real value. When an item is over-insured, the claim adjustment will take into consideration only the real value of the insured item (Pavić, 2003.).

If he damaged and insured item is not repaired or the lost parts are not replaced, which subsequently causes the total loss of the item within the period of time covered by the insurance policy, the insurer is liable to compensate for the total loss only and is not liable for the unrepaired damage. In such cases, in addition to the total loss, the insurer is liable to compensate for the costs that the insured had with regard to the partial loss or damage, such as the costs of the inspection (survey, calculation etc.).

In the real-life event that is discussed in this paper, it would be necessary to carry out a number of repairs and replacements of parts in order to bring the damaged vessel in the condition prior to the sea accident, which would eventually amount to 121,249.90 EUR (see Table 1.).

**Table 1:**
*Extent of the costs of salvage and costs related to the necessary repairs to the insured vessel in question*

| Service / product | Price (EUR) |
|---|---|
| Replacement of the upholstery and cushions in the cabin | 3,212.85 |
| Replacement of Mercruiser 454 propulsion engines (2 pieces) | 97,724.23 |
| Services of the mechanic (2 x 40 hours x 37.48 EUR) | 2,998.40 |
| Propulsion engine instruments and installation | 5,354.75 |
| VHF set ICOM | 495.31 |
| Replacement of the electric installation (2 x 24 h x 37.48 EUR) | 1,799.04 |
| Installations, batteries and other materials | 2,008.03 |
| Replacement of the radio and loudspeakers | 93.71 |
| Repairs of the fibreglass and painting the vessel | 6,291.83 |
| Replacement of the bow cleat and the navigation light | 374.83 |
| Replacement of the stern rail | 896.92 |
| **Total:** | **121,249.90** |

*Source: Authors*

The Progression high performance planing boat that is discussed in this paper was in service for 21 years before the accident. Through the detailed examination of the vessel (state of the shell plating, engines, interior spaces etc.), and taking into consideration the supply and demand for such vessels on the market (the market demand is rather low due to extraordinary output of her petrol engines) it has been ascertained that the actual market value of the vessel amounts to 34,136.55 EUR (http://uk.yachtworld.com).
Hence, it would be necessary to pay 121,249.90 EUR to bring the discussed vessel to the original condition, as it was before the accident, which exceeds her market value by more than 3 times, her market value on the day of the accident being 34,136.55 EUR, which leads to the conclusion that the repair of the vessel is not economically justified and that the vessel suffered complete or total damage.

### 4.2. Complete (total) damage

The extent of the salvage costs and costs related to the necessary repairs to the insured vessel in question is larger than the agreed or real value of the insured vessel and the insured is entitled to claim the compensation as if the complete (total) loss of the insured vessel has taken place (Maritime Code, Article 710).
The total loss may be actual or constructive total loss. The insurer and the insured can agree to settle potential damage as an actual total loss although the insured property has not experienced a real total loss. This type of the total loss is called the agreed (value) total loss (Pavić, 2000).
If the value of the insured item is agreed on the total loss policy, the compensation for the damage extends to the amount of the agreed value, regardless of the total loss being actual or constructive. If the agreed value is insured only partially, the compensation for the damage extends only up to the amount of the agreed value. If the value of the insured item is not specified on the total loss policy, the compensation matches the current real value of the insured property (Article 709, alinea 2 of the Maritime Code). The real value of the insured property is estimated in the event of damage and according to the value of the property at the beginning of the insurance policy period, unless otherwise specified and agreed (Article 695 of the Maritime Code).
By examining the maritime insurance agreement (full insurance policy) it was found out that there was no agreed value clause on the policy of the vessel under discussion so that her actual market value on the day of the sea accident had to be estimated.
The extent of the total damage can be calculated by deducting the depreciation (amortisation) from the vessel's purchase price in order to define the market value of the vessel on the day of the accident which is then reduced by the real value of the rescued parts (Ćurković, 2002).
In our case, the market value of the vessel amounted to 34,136.55 EUR on the day of the accident (http://uk.yachtworld.com). The value of the rescued parts of the vessel (real value of the vessel's remnants on the market after the accident) was 12% of the vessel's value on the day of the accident, i.e. 4,096.39 EUR, so that the real extent of the total loss amounted to **30,040.16 EUR** (34,136.55 EUR – 4,096.39 EUR = 30,040.16 EUR).

## 5. Recommendation

The vessel in question was built in 1990. At the time of her accident the vessel was in service for 21 years. In the accident that is discussed in this paper the electrical installation, equipment, interior and both propulsion engines were completely destroyed whereas the vessel's hull was partly damaged. The overall costs of repairs and replacements of the engines, equipment, fittings and interior would amount to 121,249.90 EUR and it can be concluded that the vessel in question suffered a complete damage or a total loss. Taking into consideration all her features, the vessel's real market value in the Republic of Croatia is about 34,000.00 EUR.

The facts referring to the damage that have been discussed in detail in this paper, as well as the technical data supported by real-life photographs of the vessel in question, clearly indicate that the vessel was not flooded due to the detachment of the intake hose of the cooling sea water supply pump while the vessel was under way. In other words, the course of the sea accident did not take place as described in the skipper's statement.

As the vessel's hull below the waterline did not suffer any damage that may have resulted in flooding the vessel and its sinking, it is obvious that breaking of water into the hull and the vessel's partial sinking were due to the intentional detachment of the sea intake hose of the port engine when the engine was not running. This conclusion is supported by the state of the rubber impeller in the sea water supply pump. After detaching the intake hose on purpose, the after part of the vessel was loaded with weight in a way that remains unknown, in order to submerge the strum box terminal below the sea level (it is normally above the sea level when the vessel is not under way), which allowed sea water to break into the machinery space, flood the hull and to partially sink the vessel eventually.

The damages to the starboard bow and deck were caused by an unknown sharp object, probably with the purpose of breaking the air pocket formed in the bow section and sinking the vessel completely.

Hence the actual damage could not be covered by the insurance policy as the damage resulted from the intention to sink the vessel on purpose.

*The results presented in the paper have been obtained in the scientific research project No. 250-2502209-2364 and the international research Project "The possibilities of reducing pollutant emissions from ships in the Montenegrin and Croatian Adriatic implementing Annex VI of the MARPOL Convention" supported by the Ministry of Science, Education and Sport of the Republic of Croatia.*

**References**

Ćurković, M. 2002. *Zbirka propisa iz osiguranja (Collection of regulations on insurance)*, Inženjerski biro / Engineering Bureau, Zagreb.

Inst. Mar. Eng. 1981. *Marine Engineering Practice* (Parts 1 to 10), London.

Kurtela, Ž. 2000. *Osnove brodostrojarstva (Basics of marine engineering)*, Polytechnics in Dubrovnik.

Martinović, D. 2005. *Brodski strojni sustavi (Marine engine systems)*. Faculty of Maritime Studies, University of Rijeka.

Mercuryracing. 2014. Available at: http://www.mercuryracing.com/ [accessed 2014-08-28]

Morton, T. 1974. *Steam Engineering Knowledge for Marine Engineers*, London.

Mrakovčić, T. 2005. *Brodski sustavi (Ship's systems)*, Faculty of Engineering, University of Rijeka.

Pavić, D. 2000. *Pomorsko pravo* - *Knjiga treća* - *Pomorske nezgode, pomorsko osiguranje (Maritime Law – Book Three – Sea accidents, marine insurance)*. Nautical High School, Split.

Pavić, D. 2003. *Pomorske havarije i osiguranje (Maritime average and insurance)*. 2nd edition, Faculty of Maritime Studies, University of Split.

Pomorski zakonik (Maritime Code), 2014. Available at: http://www.zakon.hr/z/310/Pomorski-zakonik [accessed 2014-08-28]

Stanivuk, T.; Kralj, N.; Slipčević, J. 2014. "Case study of causes of the ship's main propulsion failure", in: *Suvremeni promet: časopis za pitanja teorije i prakse prometa* 34, 3-4; pp. 236-239.

Šretner, J. 1970. *Brodski motori s unutarnjim izgaranjem (Marine internal combustion engines)*. University of Zagreb.
World's Largest Selection of Brokerage Boats for Sale. 2014. Available at: http://www.yachtworld.co.uk/ [accessed 2014-08-28]

525efi. 2014. Available at: http://www.mercuryracing.com/sterndrives/hp525efi.php [accessed 2014-08-28]

# ANALYSIS OF MOTOR VEHICLE COLLISION AND EFFECT OF MORE SAFETY REQUIREMENTS USING POLICE RECORD INFORMATION FOR IRAN ROADS IN SPRING HOLIDAYS

**Mousa Amiri[1], Fereidoon Moghadas Nejad[2], Javad Tanzadeh[3], Alireza Noory[4], Rashid Tanzadeh [5]**
[1]*PhD student of Shahid Beheshti, University of Medical Sciences, Tehran-Iran*
[2]*Associate Professor & Head of Transportation Group, Dept. of Civil & Environmental Engineering, Amirkabir University of Technology*
[3]*Assistant Professor, Islamic Azad University Bandar Anzaly Branch, Gilan –Iran*
[4]*PhD student of road and Transportation, Amirkabir University of technology*
[5]*PhD Candidate, Dept. of Civil Engineering, Semnan University, Semnan-Iran*

**Abstract:** Iran Roads are one of the most hazardous ways in the world that in this regard annual on-going cost of traffic accidents entered on the country's economy. High rate of injuries, fatality and property damage caused by a conflict with the police, search and rescue institution, hospitals and families affected by the accident is remarkable. New Year holiday about 20 days in early spring due to ever-Iranian tradition, is one of the most crowded times, it is notable that the high rate of traffic is considerable. In recent years with increasing road safety, harder regulation and oversight role of the police, the number of traffic accidents has declined. In this study, police role in significant reduction of traffic accidents recorded by the police for the years of 2012 and 2013 been investigated. Also, research has shown that greater oversight road with speed cameras, find black spots, binding rules, heavier fines, child training as police helpers were effective in reducing traffic accident rate.

**Keywords**: Motor vehicle collision, safety requirements, Iran roads, police role

## 1. Introduction

Any person who suffers from damage or injury in a crash either becomes disabled or loses his/her life. Since no person is independent from others and is living in a communication network with others (including family, friends, or any person who is affected by this accident), crash consequences are not limited to the person by him/herself and all the people who are involved in his/her life are also influenced (Qorbani.,2006). High crash rate among motorcyclists in Iran has been always one of the main concerns of transportation safety authorities such as Iran's Traffic Police. In 2012, high crash rate made the police apply more accurate and stricter policies in terms of observing safety rules by motorcyclists including the rule about lack of helmet use in order to control mortality rate caused by crashes of motorcyclists in the country. This paper first investigated the situation of accidents and global macro policies along with some of the functions leading to the safety improvement of motorcyclists. Then, situation of the dead was examined during three definite weeks in Iran in Nowruz holidays. Finally, measures undertaken by the police were evaluated.

## 2. Mortality rate of crashes in the world

In general, millions of people all over the world face death, impairment, or disability of their family members due to crash injuries. Since accurate evaluation of human life and the resulting pain and injury of losing a relative is not possible, accurate estimation of costs of a crash seems to be an impossible issue.
Damage and injury caused by road traffic impose a heavy load not only on the global and national economy but also on the family economy. Most families whose breadwinners have either died or disabled in crashes will suffer from severe poverty (Qorbani.,2006). Table 1 shows main reasons of mortality in the world in 2002. The interesting point is that, at the ages of 15 to 44 years old which are the main years of production and employment, one of the three main reasons of mortality in the world is traffic accident (Peden et al., 2004). This issue has been estimated to continue up to 2020 and crash would be the third mortality factor in the world (Qorbani.,2006).Moreover, car accidents have been predicted to be the fifth mortality factor in the world by 2030 (Ramli.,2014).Unfortunately, despite all these injuries and damage, few measures have been undertaken in terms of preventing the emergence of car accidents and small amounts of money have been invested in this regard (Qorbani.,2006). According to the report by World Health Organization (WHO), without investment increase and attempts for preventing the occurrence of road accidents, transportation mortality rate would increase by about 65% between 2000 and 2012 in the world. In low- and middle-income countries, this increase would be about 80% (Peden et al., 2004).

[5] - Corresponding author: Rashidtanzadeh@yahoo.com

**Table 1:**
*Main reasons of mortality in the world for separate age groups (Peden et al., 2004)*

| *Leading causes of deaths by age group, world, 2002* | | | | | | | |
|---|---|---|---|---|---|---|---|
| Rank | 0–4 years | 5–14 years | 15–29 years | 30–44 years | 45–59 years | ≥60 years | All ages |
| 1 | Lower respiratory infections 1 890 008 | Childhood cluster diseases 219 434 | HIV/AIDS 707 277 | HIV/AIDS 1 178 856 | Ischaemic heart disease 1 043 978 | Ischaemic heart disease 5 812 863 | Ischaemic heart disease 7 153 056 |
| 2 | Diarrhoeal diseases 1 577 891 | Road traffic injuries 130 835 | Road traffic injuries 302 208 | Tuberculosis 390 004 | Cerebrovascular disease 623 099 | Cerebrovascular disease 4 685 722 | Cerebrovascular disease 5 489 591 |
| 3 | Low birth weight 1 149 168 | Lower respiratory infections 127 782 | Self-inflicted injuries 251 806 | Road traffic injuries 285 457 | Tuberculosis 400 704 | Chronic obstructive pulmonary diseases 2 396 739 | Lower respiratory infections 3 764 415 |
| 4 | Malaria 1 098 446 | HIV/AIDS 108 090 | Tuberculosis 245 818 | Ischaemic heart disease 231 340 | HIV/AIDS 390 267 | Lower respiratory infections 1 395 611 | HIV/AIDS 2 818 762 |

The United Nations proposed a decade called "Decade of Action for Road Safety" in 2010; the objective of this decade which started from 2011 and will end in 2020 is to maintain and reduce the growing trend of road mortality. According to the presented approximations which can be found in Fig. 1, correct implementation of this plan is expected to lead to saving the lives of about 5 million people in road crashes until 2020 (Global status report on road safety., 2013).

*Goal of the Decade of Action for Road Safety 2011-2020*

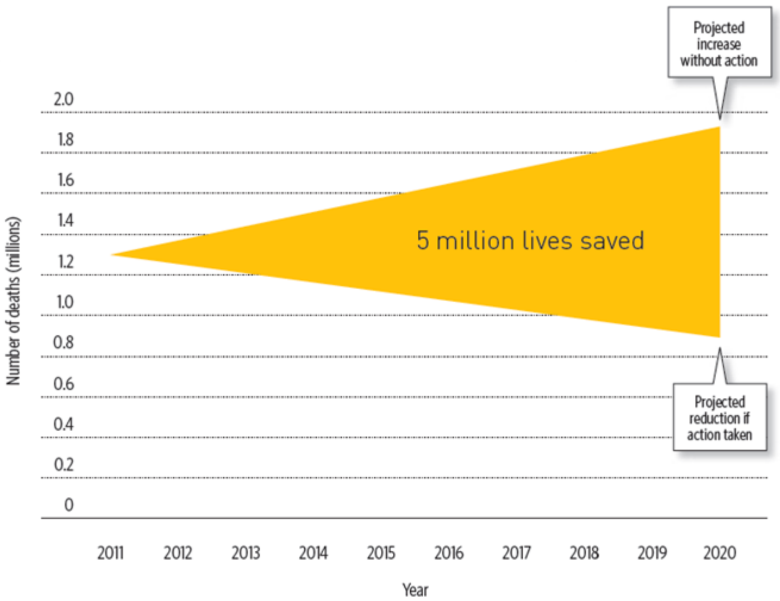



**Fig. 1:**
*Objectives of Decade of Action for Road Safety (Global status report on road safety., 2013).*

**3. Situation of crash mortality of motorcyclists in the world and its reduction strategies**

Motorcyclists constitute a considerable share of death in traffic crashes due to lack of stability and protection of pillions against collisions. Considerable growth in the number of two-wheeled motor vehicles in most countries of the world has a direct relationship with the growth of injury and mortality rate of their users (Toroyan.,2006). Motorcyclists constitute about one-third of all traffic crash death in the south-east of Asia and the Pacific region. Moreover, this statistic has demonstrated a growing rate in African and American countries (Liu., 2005).Head injuries are the most important causes of injury, impairment, and mortality among motorcyclists. In European counties, head injury has been the cause of about 75% of deaths among motorcyclists. In some low- or middle-income countries, head injury has been the cause of 88% of death; this figure has been estimated, since there is no accurate data record and collection system in these countries (Liu., 2005).Around 30 years ago in the USA, William Haddon proposed traffic and transportation system as a system without any design and plan (human-machine) and declared that this system needs structures and systematic behavioral understanding and perception. He introduced a matrix (currently known as Haddon matrix) in which three components of human, vehicle, and environment (road) are in interaction. Also, he presented three crash phases as follows:

1. Pre-crash phase
2. Crash phase
3. Post-crash phase

In three by three interaction of the above cases, a 9-component matrix is generated which is demonstrated in Table 2 (Peden et al., 2004). Accordingly, studies showed that factors such as training can be effective for reducing mortality of motorcyclists and emergence of severe injuries in a long run. Also, they mentioned that speed control, especially in highways, has a considerable role in reducing losses in pre-crash phase (Ramli.,2014).Ensuring health of motorcycles and proper functioning of their different parts is among the effective factors for pre-crash phase. During crashes, one of the most important effective and controllable factors is health and standard helmet use by motorcyclists.

**Table 2:**
*Haddon matrix (Peden et al., 2004)*

| | | FACTORS | | |
|---|---|---|---|---|
| **PHASE** | | **HUMAN** | **VEHICLES AND EQUIPMENT** | **ENVIRONMENT** |
| Pre-Crash | Crash prevention | Information<br>Attitudes<br>Impairment<br>Police enforcement | Roadworthiness<br>Lighting<br>Braking<br>Handling<br>Speed management | Road design and road Layout<br>Speed limits<br>Pedestrian facilities |
| Crash | Injury prevention during the crash | Use of restrains<br>Impairment | Occupant restrains<br>Other safety devices<br>Crash-protective design | Crash-protective roadside objects |
| Post-Crash | Life sustaining | First-aid skill<br>Access to medics | Ease of access<br>Fire risk | Rescue facilities<br>Congestion |

Research has demonstrated that using standard and high-quality helmet can reduce mortality risk by 40% and emergence risk of severe injuries by 70% (Liu.,2005).Further, studies have indicated that law enforcement for motorcyclists has a major role in increasing helmet use and reducing head injuries among motorcycle pillions (Kraus.,1995).Many activities in different countries of the world for approving and enforcing law for mandatory helmet use for both motorcyclists and other pillions have had a major role in this regard. Thus, the number of countries which have this rule has increased from 131 countries in 2008 to 155 in 2011. In other words, 88% of the world population is under the provision of this law. To effectively reduce the number of injured people related to the use of motorcycles, it is necessary to implement mandatory use of helmets and provide motorcyclists with different types of standard helmets, the use of which leads to decreased head injury and losses among pillions. The interesting point is that, based on the conducted studies by Chiang et al. (2014), there is no statistically considerable difference in rate and pattern of injury for motorcyclists and pillions (Valerie.,2014). Additionally, according to Haddon matrix, another important step for mortality reduction in motorcycle crashes is related to post-crash phase.

One of the most important post-crash strategies is management of crash scene and proper and fast relief assistance; also, relief agents must be completely familiar with different types of accidents in motorcycle crashes and injury management.

**4. Investigating statistics of the deceased in motorcycle crashes in Iran during Nowruz 2012**

Iranians celebrate turn of the year in spring, which is called Nowruz. During this period, all schools and organizations are closed; thus, it is a good opportunity for having fun and meeting relatives. During this period, beside motorcyclists, a more number of women and children travel as pillions compared with other days of the year. In 2012, this issue along with lack of helmet use and common belief about discarding violations of motorcyclists led to lack of proper observance of traffic rules and increased accidental violations among the motorcyclists. Thus, in 2012, 239 people were killed in motorcycle crashes during 3-week Nowruz holiday (according to police reports), among whom 22 people (i.e. 9.2%) belonged to Tehran province (capital city of Tehran is located in this province).225 and 14 of the mentioned 239 people were men and women, respectively. In Iran, women are not allowed to ride motorcycles and these 14 people were the pillions. Also, in Tehran, 21 and 1 deaths belonged to men and women, respectively. Education of most of these people was at secondary school level, as can be observed in Fig. 2. Also, minimum number of them had Bachelor's degree. This point can be justified as follows: in Iran, higher level of education mainly leads to increased income; also, using motorcycles has potential risks for people and considered a low-class action. Thus, higher education leads to considerable reduction in the use of this vehicle and high-income people usually use cars.

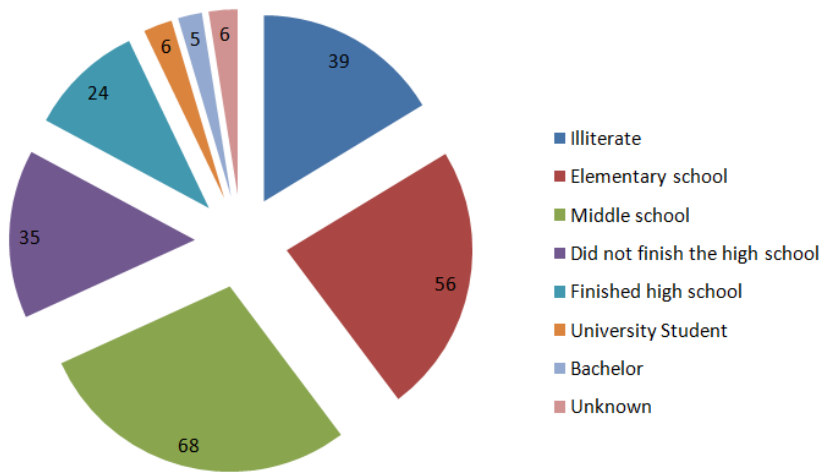



**Fig. 2:**
*Education of the deceased in motorcycle crashes during Nowruz 2012*

Out of these crashes, 129 cases occurred during daytime and 77 cases were at night. Also, 26 crashes occurred during the sunset or sunrise. Seven crashes have been reported as "occurring at other times".

**5. Changing mortality situation of motorcycle crashes**

In Iran, cases like moving at the sidewalk, moving in the opposite direction, making noise, transporting unconventional cargo, doing dramatic and risky movements such as single pirouette, using three-wheeled motorcycles, moving on special lines, lack of helmet use, being smoky, lack of third party insurance, not having driving license, simultaneously committing two accidental violations, and driving without having driving license are among violations from traffic regulations and lead to seizing of motorcycles. In spite of the approval of these regulations, little rigidity is applied in their enforcement. Since 2012, police has followed emergency plans and seriously punished the motorcyclists who violate rules in major cities in Iran, such as the capital city. This issue led to about 75% increase in helmet use within a three-week interval since the beginning of one of these plans.

Moreover, violations were reduced by 30% during this period and death statistics of motorcyclists reached zero in the capital city of Tehran. In big cities, traffic of motorcycles on special bus lines was considerably reduced (Iranian Student News Agency).

**Fig. 3:**
*Lack of helmet use, dramatic movements, and not considering safety points of motorcycles were among the increasing factors of losses in motorcycle crashes in Iran.*

**Fig. 4:**
*Decisive and precise encounter of Iran's police with violators*

In addition to precise and strict encounter of Iran's police with violating motorcyclists, plans such as free distribution of helmet among motorcyclists and encouraging them via the banners reminding about death or impairment as the consequences for lack of helmet use and the one showing their dear ones are looking forward to them are among the measures undertaken by Iran's police for building culture among drivers. An example of cultural affairs performed by Iran's police is presented in Fig. 5.


**Fig. 5:**
*Some of the cultural activities conducted by Iran's police for increasing helmet use*

These measures led to 18% reduction in the number of deceased in motorcycle crashes in Nowruz 2013 in Iran. During three weeks of Nowruz 2013, the number of the deceased in motorcycle crashes reached 195 people, which had 44 people decrease compared to the previous year. This statistic (15 deaths in Tehran) showed about 32% reduction compared with the previous year. As was mentioned, in 2012, gender distribution of crash death was 94% men and 6% women.

In 2013, this gender distribution did not show any considerable change and 97% and 3% of deaths belonged to men and women, respectively, which indicated one-third reduction in women's death compared with the previous year. Moreover, distribution of death based on education demonstrated a change in 2013, in which the deceased having a Bachelor's degree had maximum decrease in number. Maximum reduction in terms of the number of death in 2013 was related to the group with secondary school education who were 22 people.

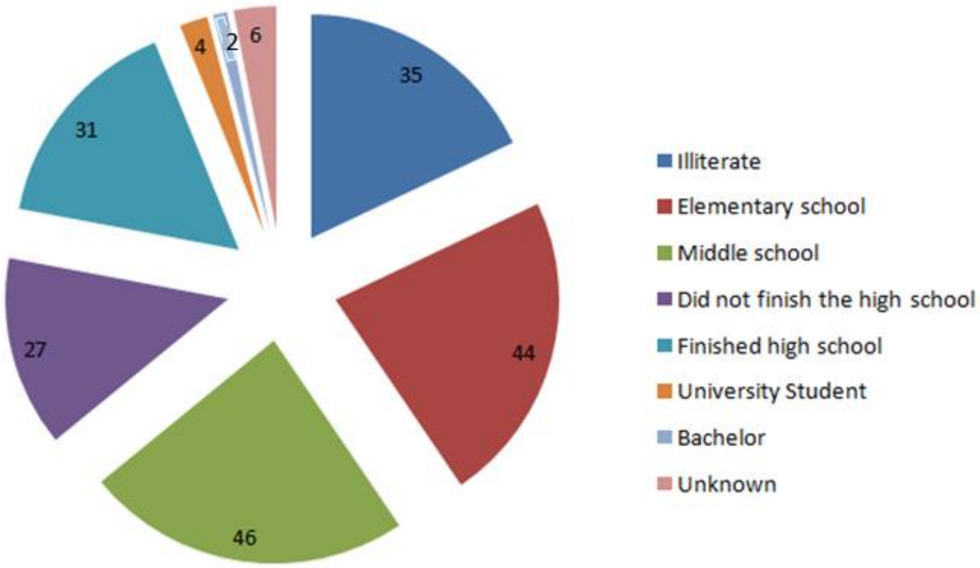



**Fig. 6:**
*Education of the deceased in motorcycle crashes in 2013*

Lighting and pedestrian conditions along with fatigue/lack of fatigue of motorcyclists which has a considerable effect on the distribution of mortality percent based on daytime did not show any change in 2013. As mentioned in the previous section, about 53.9% of crashes in 2013 occurred during daytime; in 2013, this percent reached 54.3% and did not show any considerable change. Percentage of the number of crashes during night and sunset was 32.2% and 10.87% in Nowruz 2012, respectively; in Nowruz 2013, these figures were 33.3% and 9.2% and did not show any considerable changes.

## 6. Conclusions

Applying controlling and strict strategies to motorcyclists, especially in two cases, could lead to reduced number of deceased. The first case was helmet use, the effect of which has been previously considered. The second case was to encourage motorcyclists to observe traffic rules via culture building. Based on the mentioned cases and high losing rate of motorcycle crashes during Nowruz in Iran, precise and intensive operations have been undertaken by traffic police in Iran during the last two years which have been based on these two cases. The result was considerable reduction in death statistic despite the increased number of numbered motorcycles per year. Strategies applied by the police in this regard which were previously mentioned can be classified according to Haddon matrix as follows:

**Table 3:**
*Haddon matrix*

| | Factors | | |
|---|---|---|---|
| **Phase** | **Human** | **Vehicles and Equipment** | **Environment** |
| Pre-crash | Enhancing knowledge of motorcyclists about accidents<br>Culture building about traffic safety of motorcycles<br>Enforcing law to violating motorcyclists | Being strict about technical situation of motorcycles | Installing speed control equipment and reducing traffic speed |
| Crash | Increasing use of standard helmets | - | - |
| Post- crash | - | - | - |

Using these strategies could reduce the number of death in motorcycle crashes by 18% during one year all over the country; this issue led to 32% reduction in the number of death caused by motorcycle crashes in Tehran province in Nowruz 2013.

**References**

Global status report on road safety 2013: supporting a decade of action .Organization, World Health .20 Avenue Appia, 1211 Geneva 27, Switzerland : World Health Organization .ISBN 978 92 4 156456 4.

Head of Tehran Road Police announced: Seizing cars and motorcycles which move on special lines since Saturday, Iranian Student News Agency (ISNA), Friday, 20 December 2013, News code: 92092919989.

Kraus J, Peek C, Williams A , Compliance with the 1992 California motorcycle helmet-use law .American Journal of Public Health, 1995 .American Journal of Public Health,1995, 85:96–98.

Liu B et al,2005.Helmets for preventing injury in motorcycle riders .al., The Cochrane Database of Systematic Reviews, issue 5

Margie Peden, Richard Scurfield,David Sleet, Dinesh Mohan,Adnan A. Hyder, Eva Jarawan and Colin Mathers 2004.World report on road traffic injury prevention .Geneva : World Health Organization. ISBN 92 4 156260 9.

Qorbani. 2006. World report on preventing injuries caused by road crashes, World Health Organization (WHO), Translated by Farhad Mehryari Leylami, Mehran, Tehran: Ministry of Road and Transportation. 978-964-6299-73-3.

Roszalina Ramli, Jennifer Oxley, Faridah Mohd Noor , Nurul Kharmila Abdullah .Fatal injuries among motorcyclists in Klang Valley, Malaysia Kuala Lumpur, 2014. Malaysia : Journal of Forensic and Legal Medicine. 2014 .Journal of Forensic and Legal Medicine 26 39-45.

Toroyan T et al.2006. Helmets: a road safety manual for decision-makers and practitioners .al, Geneva : ,World Health Organization.(www.who.int/roadsafety/projects/manuals/helmet_manual/en/index.html).

Valerie X.Y. Chiang, Jolene Y.X. Cheng , Z.C. Zhang , Li-Tserng Teo, 2014 . Comparison of severity and pattern of injuries between motorcycle riders and their pillions: A matched study ..Singapore : Injury . Injury, Int. J. Care Injured 45 333–337.

# COMPARISON OF THE ACCURACY OF SEVERAL BRIDGE/RAIL COUPLING MODELS WITH EXPERIMENTAL AND NUMERICAL RESULTS

**Vojtěch Stančík[1], Pavel Ryjáček[2]**
[1, 2]*Czech Technical University in Prague, Faculty of Civil Engineering, Dep. of Steel and Timber Structures Thakurova 7, 166 29 Prague 6, Czech Republic*

**Abstract:** Current demand of faster, comfortable and competitive railroad infrastructure results in the removal of older jointed rails in order to place the continuously welded rails (CWR) instead. These are used for their ability to create a comfortable joint – free track supporting the minimization of dynamical impact on both the substructure and the vehicle itself. Unfortunately the placement of CWR is limited for bridges due to its interaction with the supporting structure that is coupled with the rail. The scientific data, needed for simple and reliable analysis, aren't currently sufficient, because the values vary immensely. There was therefore done an extensive experimental measurement on the railway bridge near the town Děčín with the intention of improving current state of knowledge of the CWR/bridge interaction. The main goal of this paper rests on performing numerical analysis using various numerical models of the coupling interface and the subsequent comparison with the experimental and prescribed values. The complex overview with the results of the experimental evaluation, numerical model verification and data comparison is offered in this paper including some interesting conclusions as well. The conclusions of the paper provide with advices for better numerical modelling of the bridge/track interaction.

**Key words:** continuously welded rail, steel bridges, interaction, nonlinear joints

## 1. Introduction

### 1.1 Motivation

The continuously welded rail (CWR) is nowadays being placed on almost all modern railway tracks including the bridges. The replacement of the older rail type - the jointed rail - allows even more comfortable passenger transport, so as it allows increasing the train speed. It also ensures a positive influence on the bridge substructure due to the reduction of the dynamical impact arising from crossing of expansion gaps by the train vehicle. However this rail type has one significant disadvantage. The removal of the expansion joints means limitation of the rail expansion movement, which results in additional rail stress increase. When considering the mechanical behavior of the CWR in the open track, the stress change depends more or less only on the temperature change. When the CWR reaches bridge structure the problem suddenly becomes more complicated, since the bridge/rail interaction is being affected also by other properties - for example the bridges span arrangement and its structural parameters or the track arrangement and its structural parameters. Loads needed to be considered by performing structural analysis of the bridge/rail combined response are the classified vertical loads, horizontal loads coming from train vehicles braking or acceleration and the aforementioned temperature change. Reaching the limit state by combining these loads may cause either the rail break or its buckling. Rail interaction with the bridge also influences the dimensions of the substructure and bearings, which may greatly affect the costs of the whole structure. These consequences describe the most adequate way the importance of this phenomenon proper numerical analysis.

**Fig. 1.**
*Rail buckling*

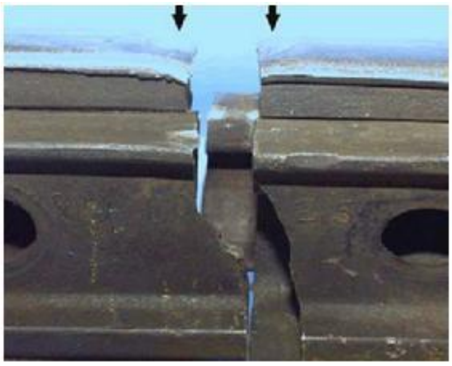


**Fig. 2.**
*Rail break*

[1] Corresponding author: vojtech.stancik@fsv.cvut.cz

### 1.2 Bridge/rail interaction description

The interaction between the rail and the structure is being generally described using the nonlinear stiffness law, formed by longitudinal restoring force $r_x$ and the relative bridge to rail displacement $u_x$. The restoring force can be determined by the ratio of the rail normal force change related to the corresponding rail section length in which this force change occurs. The longitudinal restoring force value changes nonlinearly accordingly to the relative bridge/rail displacement till it reaches its limit value of $r_0$, which is called the plastic resistance. The European codes simplify this behavior using the bilinear dependence of $r_x$ on $u_x$. In this case the plastic resistance is achieved with the limit value of bridge/rail displacement $u_0$. The slope of the chart's elastic part defines the longitudinal stiffness of the coupling - $k_x$. Unfortunately the longitudinal stiffness and the corresponding plastic value of restoring force depend on the tracks state of load and on its location in the structure, respectively these values vary for loaded and unloaded track on the bridge and also for open railway track. The railway track of course deforms in the vertical direction as well. There has been therefore determined the vertical stiffness $k_z$ described as the ratio of the vertical load related to the bridge/track vertical relative displacement. The latter stated nonlinear stiffness law and the vertical stiffness together define the coupling interface between the rail and the bridge.

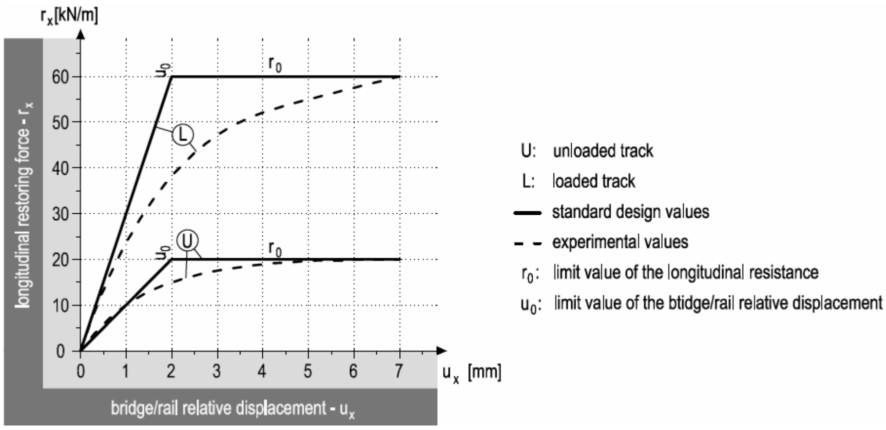



**Fig. 3.**
*Nonlinear stiffness law*

It should be also emphasized that the functions describing coupling interface differ accordingly to the various possible performances of bridge deck types and the appropriate type of rail fastening. This means various coupling parameters for directly fastened rails, ballasted tracks and embedded rails are used.

### 1.3 Research goals and methods

The bridge/rail interaction phenomenon is being lately a widely discussed structural problem because of the diverse and inconsistent coupling parameters values offered by the European scientific literature and experimental measurements. The main goal of this contribution is therefore to verify standardly used coupling parameters, so as to find possible weak spots and uncertainties in the numerical bridge/rail coupling analysis. Simultaneously the performed numerical analysis offers some advices for relatively simple and fast modelling of this problem. The research aims at bridges with ballasted track, which is the most frequently used bridge deck type on latest railway bridges. The whole study is based on the data offered by the experimental measurement carried out on the steel truss bridge near the town Děčín. These data are subsequently verified using the numerical model of the bridge. In the case there have been registered a good concordance between the experimentally measured data and the corresponding numerical values the analysis process also simultaneously provides with the validation of the numerical model. The last step of this study is to carry out numerical analysis using the coupling parameters taken from European codes and to determine their accuracy compared with the verified experimental data.

## 2. Experiment

A brief familiarization with the experiment is offered in the following paragraph. As said before, the in-situ testing was performed on the steel truss bridge near the town Děčín. The bridge consists of three simple spans supported by massive brick substructure followed by arch bridges on both sides of the abutments. The ballasted railway track is being transferred via the orthotropic bridge deck across the Elbe River.

**Fig. 4.**
*In-situ testing on the steel truss bridge in Kolín*

**Fig. 5.**
*Bridge/rail relative displacement gauge*

Whole measurement was carried out during the static load test which allowed simultaneous and therefore economical utilization of the static load burdens. Burdens used to induce prescribed level of stress involved two EDK 750 railway cranes and one locomotive HV 749. Thus the experiment focused on coupling behavior testing for both the unloaded and loaded track. For fulfilling the announced goals there were measured following quantities:

- Temperature of the rail and the bridge (main truss girder, bridge deck)
- Rail stress distribution
- Relative horizontal and vertical bridge/rail displacement
- Bearing displacement and rotation

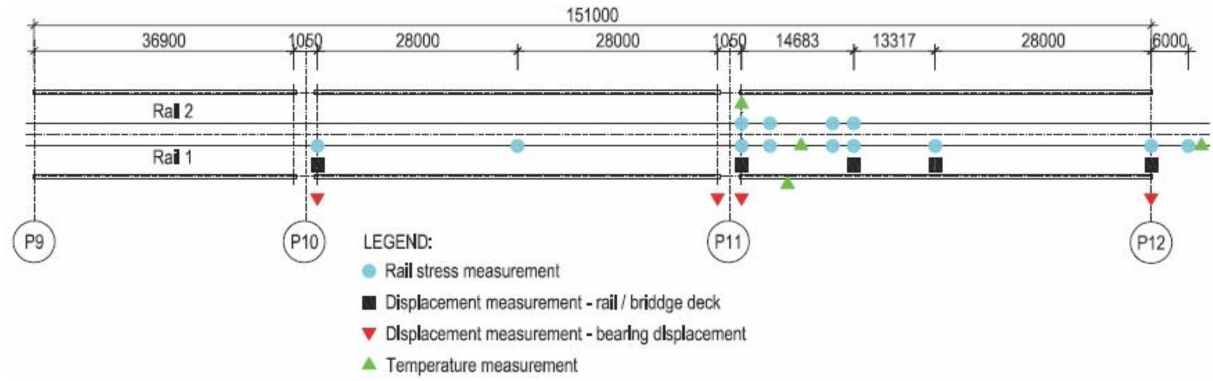



**Fig. 6.**
*Gauge placement scheme*

Relating the normal force change to the corresponding section length allowed determining the longitudinal restoring force distribution, which together with the relative displacement described the nonlinear horizontal coupling behavior under various loading levels of the bridge/rail combined system. The vertical stiffness had to be evaluated by iterative procedure of numerical calculations. This involved iterative setting of vertical stiffness in the coupling interface model in order to achieve experimentally detected vertical bridge/rail relative displacement in the ballast. Extensive description of the whole in-situ testing process may be found in the recently published paper (Stančík et al., 2014).

**3. Numerical Analysis**

For the case of this study, the most convenient way of finite element method (FEM) modelling is to create a 2D numerical beam model in some commonly used FEM solving software, which grants the applicability of nonlinear joints and analysis, then the analysis also provides with helpful information about the accuracy of this frequently used type of modelling. Accordingly to latter facts the bridge model was created in the commonly used software SCIA Engineer.

**3.1 Structural 2D beam model**

As said before the interaction depends significantly on the combined bridge/rail systems structural and geometrical arrangement, which greatly influence the demand of precise structural modelling. There has been therefore applied identical bearing arrangement and profiles (including the effective part of the orthotropic bridge deck) like on the real bridge.

CWR profile being placed on the bridge model is the rail type 60 E1 which changes into rail type 49 E1 in the transition zones near both abutments. In order to ensure trustworthy results, CWR has been extended by the value of 50 meters behind both abutments, which has to be considered for there is not being placed any expansion device. Interesting thing was finding out, that the experimental results measured on both individual rails slightly vary. The analysis was therefore carried out separately for both main truss girders.

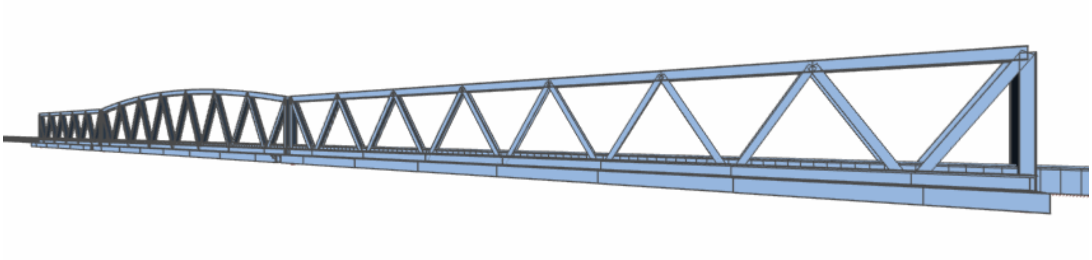


**Fig. 7.**
*Structural 2D beam model*

Considerable attention was given to the ballast and fasteners modelling. These structural elements are replaced with the fictive beams, rigid in both directions and embedded with their bottom face in the bridge deck. Of course the spacing of latter described fictive elements was modeled in accordance with the real sleeper spacing. On the upper face there is situated the connection with the CWR. This joint represents the coupling interface with both the horizontal and vertical nonlinear functions, thus the nonlinear behavior of the longitudinal restoring force and the vertical stiffness of the ballast. Fictive beams' height has been adjusted accordingly to the designed distance of the rail and the truss bottom chord neutral axis. Of course the bearings rotation point distance from the bottom truss chord has been modified as well. The latter two improvements ensure appropriate bridge/rail relative displacement.

### 3.2 Coupling parameters settings

$$r_x = r_0 \cdot \left(1 - e^{-\frac{k_x}{r_0} \cdot u_x}\right) \tag{1}$$

As the numerical analysis was divided into two procedures involving experimental measurement verification and subsequent comparative analysis of the European codes' accuracy, the coupling interface nonlinear model had to be adapted appropriately to both this tasks. The first setting was therefore applied with the approximated coupling functions, which fitted the best to the gathered experimental data (see Fig.8). By the approximation process it was looked for the most convenient parameters of $k_x$ and $r_0$ (defined in 1.2) by using the equation (1) and the method of least squares. Because of the variable character of the measured values there were applied different coupling functions in particular sections lengthwise the bridge (see Fig.9). The authors are aware that the extrapolation doesn't offer precise values of the plastic resistance $r_0$ as the regression analysis isn't supported with sufficient amount of experimental data. Unfortunately the observed bridge doesn't allow achieving of bigger relative bridge/rail displacement values and the corresponding longitudinal forces thanks to its structural character, even despite the fact that there is applied the most extreme loading caused by the EDK 750 railway cranes. Nevertheless this extrapolation is meaningful for the general comparison with the parameters prescribed by the European codes.

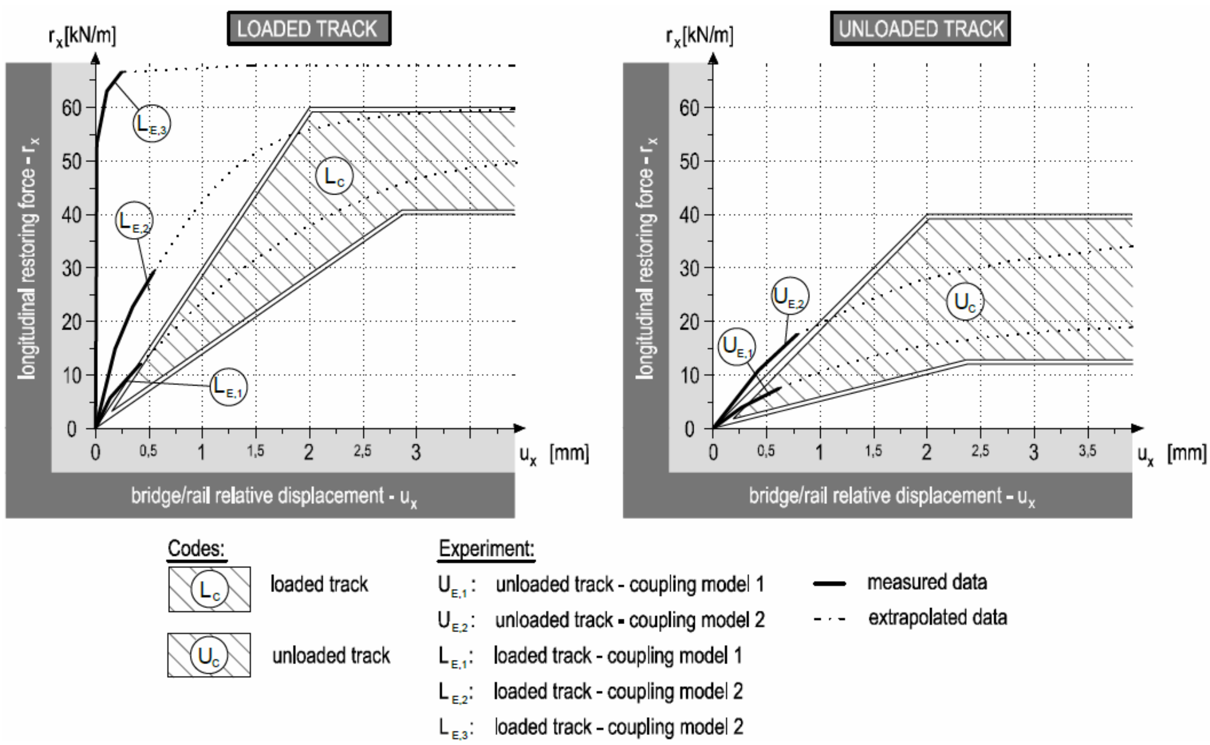



**Fig. 8.**
*Settings of coupling interface functions*

Next setting considered the coupling parameters recommended by the European codes. In this case the longitudinal arrangement of the taken nonlinear functions was slightly different compared to the previous one (see Fig.9). There shall be assumed different values for loaded (L) and unloaded (U) track on the bridge and also different values for the interaction description in open track. For our case the open track functions are designated as "behind abutment" (BA) and they are presumed to be similar to the ones valid for unloaded track on the bridge, since the steel bridge is being followed behind both abutments by arch bridges. The coupling interface model used for both latter analyses involved the beforehand evaluated function of vertical stiffness as well.

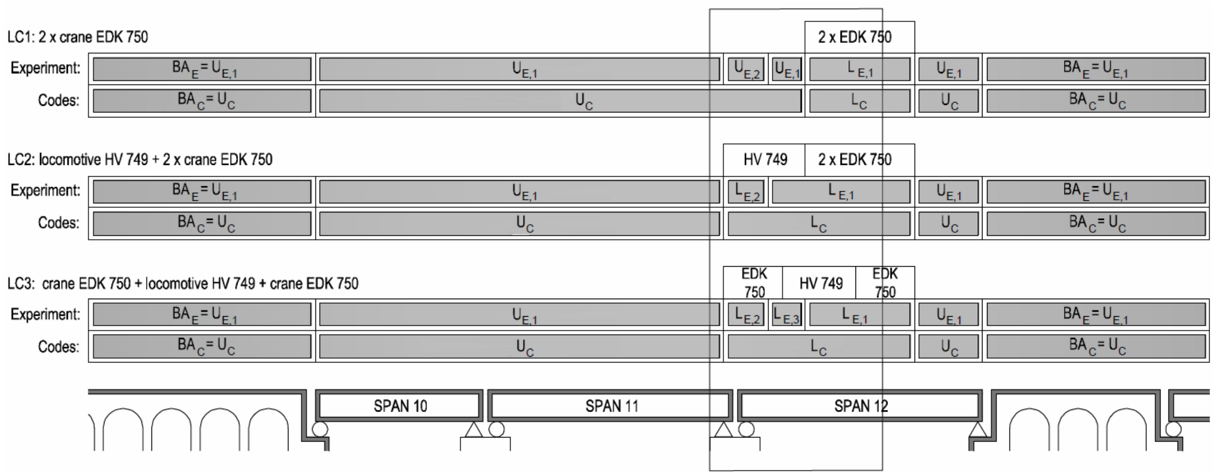
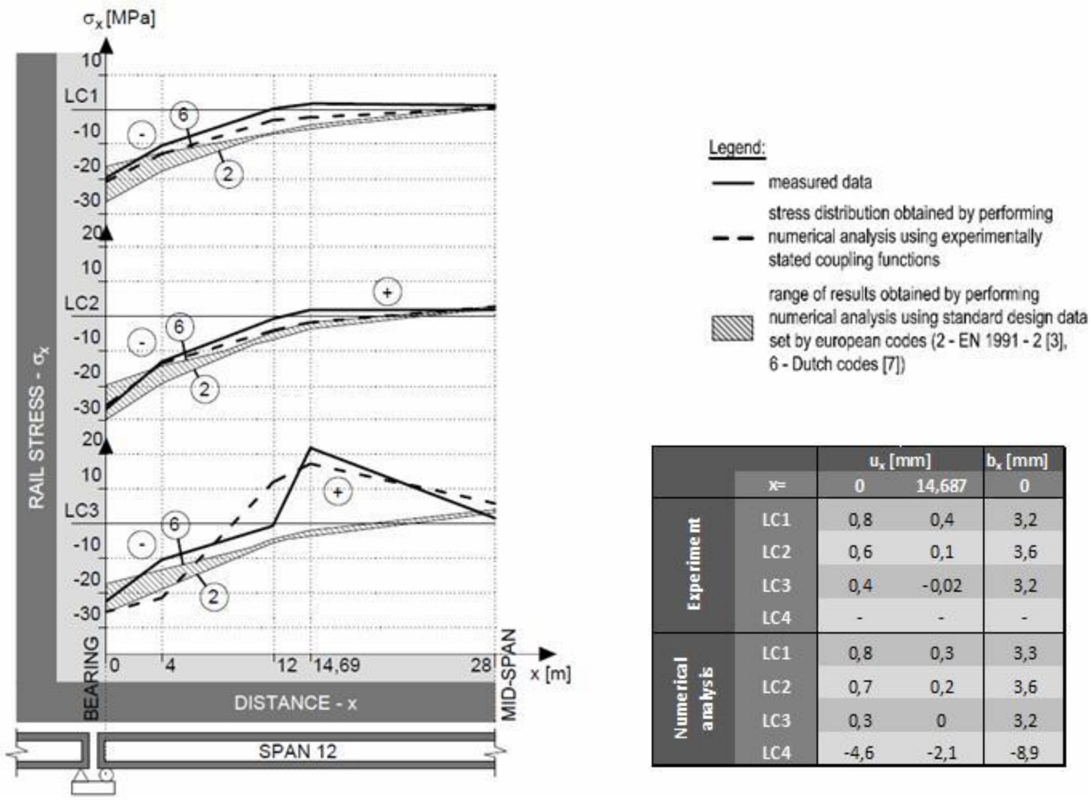



**Fig. 9.**
*Coupling interface arrangement for various LCs*

For the purpose of carrying out this study, the bridge structure was tested and numerically analyzed in three different load cases under different loading level above the sliding bearing in span 12. The arrangement of burdens is shown on figure 9. The calculation also involved the fourth load case providing with illustrative information about the temperature change influence, which may result in different nonlinear interaction behavior, since the structure straining may be affected by the load history as stated by Ruge and Birk (2007). The temperature decrease was assumed variable for the bridge deck, rail and main truss girder with the approximate value of $\Delta T$ = 16,5 °C, which was determined as difference of the measured temperature and the minimal allowed limit for rail neutral temperature. Nevertheless, the possible impact of load history has to be examined using a complex numerical model in an advanced FEM analyzing tool, which is a good motivation for further research.




| | | $u_x$ [mm] | | $b_x$ [mm] |
|---|---|---|---|---|
| | x= | 0 | 14,687 | 0 |
| Experiment | LC1 | 0,8 | 0,4 | 3,2 |
| | LC2 | 0,6 | 0,1 | 3,6 |
| | LC3 | 0,4 | -0,02 | 3,2 |
| | LC4 | - | - | - |
| Numerical analysis | LC1 | 0,8 | 0,3 | 3,3 |
| | LC2 | 0,7 | 0,2 | 3,6 |
| | LC3 | 0,3 | 0 | 3,2 |
| | LC4 | -4,6 | -2,1 | -8,9 |

**Fig. 10.**
*Numerical analysis results overview*

**3.3 Results**
Performed numerical analysis has shown good concordance between the experimental and computed data. In the above attached figure there is offered the stress distribution comparison according to each load case. Obtained values of relative bridge/rail displacement $u_x$ and relative displacement in sliding bearing $b_x$ are sufficiently accurate as well. Therefore the assessed experimental data may be considered verified. The approximate values describing the horizontal coupling functions are evaluated to:

- $U_{E,1}$: $r_0$ = 20 kN/m, $k_x = 15.10^3$ kN/m$^2$ (expected value of $k_x = 10.10^3$ - $20.10^3$ kN/m$^2$)
- $U_{E,2}$: $r_0$ = 35 kN/m, $k_x = 25.10^3$ kN/m$^2$ (expected value of $k_x = 10.10^3$ - $20.10^3$ kN/m$^2$)
- $L_{E,1}$: $r_0$ = 55 kN/m, $k_x = 30.10^3$ kN/m$^2$ (expected value of $k_x = 30.10^3$ kN/m$^2$)
- $L_{E,2}$: $r_0$ = 55 kN/m, $k_x = 75.10^3$ kN/m$^2$ (expected value of $k_x = 30.10^3$ kN/m$^2$)
- $L_{E,3}$: $r_0$ = 65 kN/m, rigid-plastic behavior (expected value of $k_x = 30.10^3$ kN/m$^2$)

The detected vertical stiffness is also a bit bigger than the generally expected value of $40.10^3$ kN/m$^2$. There were also found a slightly nonlinear relation between the measured vertical bridge/rail relative displacement - $u_z$ and the mentioned vertical stiffness - $k_z$. The approximate values are given below:

- $u_z$ = 0,55 mm relates to $k_z = 72.10^3$ kN/m$^2$
- $u_z$ = 1,82 mm relates to $k_z = 55.10^3$ kN/m$^2$

Fig.9 also introduces an overview of the normative analysis accuracy and the possible scope of achieved results.

## 4. Conclusions

As the verification procedure confirmed, numerical analysis using the rather simple 2D beam model offers sufficiently accurate results. Main problem of the bridge/rail interaction behavior examination therefore rests in proper understanding and subsequent modelling of the coupling interface and the related geometrical and structural input parameters of the combined bridge/rail system.

The experiment indicated a variable horizontal stiffness of the ballast lengthwise the bridge. Mainly the area above the sliding bearing has shown higher values of $k_x$ than usually recommended. Vertical stiffness evaluation also provided with some interesting results, referring to its nonlinear character and slightly higher value than the normative one - $k_z$ = 40 kN/m$^2$. In general, the vertical stiffness doesn't affect the rail stress distribution very much. It can't be on the other hand by performing the analysis neglected, thus an appropriate value has always to be taken into account. Furthermore the experiment failed to achieve the plastic resistance; however the extrapolation indicated a relative good concordance with the recommended values.

In spite of the fact that there were evaluated higher values of the horizontal stiffness, the bilinear idealization recommended by the codes demonstrated especially during the LC1 and LC2 a relatively good accuracy. In LC3 there was indicated a concordance above the sliding bearing but there was found an unexpected local maximum of the rail stress at the approximate distance of 14,5 m from the sliding bearing in span 12 (see Fig.9), which relates to the interval with the evaluated rigid-plastic coupling function $L_3$. This phenomenon probably occurred because of the previous temperature change, since there was indicated a possible depletion of the ballast elastic capacity in LC4 (see Fig.9 - table) Missing this local maximum may lead to a dangerous design when considering possible loading combinations.

Some general remarks about the normative bridge/rail analysis are emphasized in the following paragraph. The absence of horizontal stiffness or any limit value of plastic relative bridge/rail displacement in the EN 1991 – 2 means it is almost useless without the knowledge of other scientific literature. The UIC 774-3, Norway and German recommendations provided with relatively reliable results with the aforementioned exception of the LC3 stress distribution. On the other hand the Dutch codes introduce the lower level of the ballasted track horizontal stiffness which for our case resulted in lower values of rail stress amplitudes that had been indicated within the experiment.

## Acknowledgements

Research reported in this paper was supported by Competence Centres program of Technology Agency of the Czech Republic (TA CR), project Centre for Effective and Sustainable Transport Infrastructure (no. TE01020168)

**References**

EN 1991-2. 2003.. *Eurocode 1: Actions on structures – Part 2: Traffic loads on bridges.* Brussels: European Committee for Standardization, 2003. 164 p.

Freystein, H., Geißler, K., 2013. (Rail/Bridge Interaction on Steel Bridges with Examples. *Steel Structures* 82 (2)): 78 – 86. [in German]

Freystein, H., Geißler, K., 2010. (Rail/Bridge Interaction – State of Art and Examples. *Steel Structures* 79(3)): 220-231. [in German]

(Regulations for Structures. 2003. *Appendix 6. Part 1: Loads.*) [in Dutch]

Ruge, P.; Birk, C., 2007. Longitudinal forces in continuously welded rails on bridgedecks due to nonlinear track–bridge interaction. *Computers and Structures* 85: 458–475.

Ryjáček, P.; Vokáč, M., 2014. Long-term monitoring of steel railway bridge interaction with continuous welded rail. *Journal of Constructional Steel Research* 99: 176–186.

Stančík, V., Ryjáček, P., Očadlík, P., Vokáč, M., 2014. In-situ testing of railway bridges with CWR. In *Proceedings of the 12th International Conference on Steel, Space & Composite structures,* 327-335.

(Technical Regulations for Railway Structures: Design of Structures. 2007. *Part 5: Loads on Railway Bridges.*) [in Norwegian]

UIC code 774-3 R, 2nd edition, 11/2001, *Track/bridge interaction. Recommendations for calculations.* UIC

# ANALYSIS OF CHARACTERISTICS AND APPLICATION OF WATERBORNE ROAD MARKING PAINT

**Darko Babić[1], Tomasz Burghardt[2], Petr Prusa[3]**

[1]*University of Pardubice, Jan Perner Transport Faculty, Department of Transport Management, Marketing and Logistics in Transport, and Logistics, Studentská 95, 53210 Pardubice, Czech Republic;*
*University of Zagreb, Faculty of Transport and Traffic Sciences, Vukelićeva 4, 10000 Zagreb, Croatia*
[2]*SWARCO Limburger Lackfabrik GmbH, Robert-Bosch-Str. 17, 65582 Diez, Germany*
[3]*University of Pardubice, Jan Perner Transport Faculty, Department of Transport Management, Marketing and Logistics in Transport, and Logistics, Studentská 95, 53210 Pardubice, Czech Republic*

**Abstract:** Road markings are one of the essential safety features of modern roadways. All of the horizontal road markings systems consist of a pigmented coating containing partially embedded retroreflective elements such as glass beads. In addition to durability and functionality of the road marking, ease of application and their effect on human health and environment are also primary considerations.
Road markings can be divided into plural component systems that cure due to chemical reaction that occurs at the site of application, thermoplastics that require heat for application, and paints. The first two systems shall not be extensively discussed herein.
Traffic paints can be separated into solventborne and waterborne materials. Over 100 years old solventborne technology includes paints based on alkyd, alkyd-acrylic, styrene acrylics, and lately also 100% acrylic resins — all of these paints afford consistent application properties, are easy to dilute to the desired application viscosity, and provide quite reasonable "no track" time under a variety of conditions such as lower temperatures and high humidity. Their environmental and human health impact is significant.
Modern waterborne paints are based on a 100% acrylic technology and incorporate developed in the 1990s quick-set chemical mechanism for drying. In most cases, properly formulated and applied waterborne paints provide advantages as compared to other road marking materials — ease of application without the need to dilute or stir, unparalleled safety during manufacturing, transport, and use, and exceptional durability together with very good adhesion of glass beads. Their drawback is slow development of washout resistance, particularly during application at high humidity and low temperature. Waterborne road marking paints are extensively and successfully used in North America and in Europe in Scandinavia and also in France, Italy, Spain, and some areas in Germany. Impact of waterborne traffic paints on human health and our environment is very significantly reduced as compared to solvent-based materials.
Analysis of characteristics of waterborne road marking paints and their application shall be presented in comparison with solventborne paints. Based on the comparison and results from test application on the road, intelligent decisions regarding future use and function of waterborne road marking paints shall be possible.

**Keywords**: waterborne road marking paint, road markings, traffic paint, solvent borne paint, retroreflection

## 1. Introduction

Road safety is a focus of the European Commission policies: In 2009, more than 35,000 people died on the roads of the European Union and no fewer than 1,500,000 persons were injured. The corresponding social cost was approximately 130 billion euros (European Commission, 2010). The importance of horizontal road markings and their influence on driving safety, especially at night and in adverse weather conditions (fog, rain, etc.) is well known (Migletz et al., 1994).
Horizontal road markings can be defined as a set of longitudinal and transversal lines, signs and symbols on the surface of transportation infrastructure. They represent significant fraction of overall traffic signalisation and cannot be at present replaced by other markings or regulations. Efficiency and durability of road markings are required by both road users and the road authorities. From the users' point of view, road markings provide an optical path by means of contrast of colour and luminance with the road surface and it would appear that these properties should be maximised (Horberry et al., 2006). Cost and performance influence the roads administrators' choices - they invest considerable amounts of money to keep the markings at adequate performance level. Even though the overall long-term performance and cost should be balanced, unfortunately, quite frequently short-term financial considerations plays more significant role.
Assessment of effects of various road markings has not produced definitive results as to which road markings maximise safety benefits (Asdrubali et al., 2013). Studies showed that with high visibility markings on outer roads the drivers tend to increase speed, thus nullifying the benefits on safety (Retting et al., 2000). At urban level, things get even more complicated because of factors like pedestrian crossings, turnabouts, stop signals and also a different way of driving (Retting et al., 2000).
This article presents analysis of characteristics and application of waterborne road marking paints and show comparison with solventborne material. Based on these comparisons it will be possible to define the future application and use of waterborne road marking paint. First results from trial application in Croatia are provided.

[1] Corresponding author: dbabic@fpz.hr

## 2. The role of road markings

Road markings are one of the most important components of traffic signalisation because of their position in the central area of driver attention. Their function is to warn drivers about conditions of the road and its construction characteristics and to help in determining lateral or transverse position of their and other vehicles. Field of driver attention is changing depending on the construction characteristics of the road, the traffic conditions, and the vehicle speed (decreasing with increase if the speed and increasing when reducing speed), (Thurston, 2009). Important feature of road markings is their continuity along the entire length of the road, which is a significant fact in the process of orientation. Horizontal road markings are important factor for road safety: The presence of only the centre and edge lines were reported to reduce all accidents by 20% (Miller, 1992).

The main tasks of road markings are:

- Drawing attention to the situation around and in the area in front of the vehicle which requires special attention and caution.
- Ensuring traffic management, particularly at approaches to intersections.
- Defining the road in its course and layout.
- Giving drivers clear orientation and safe guidance by day and night.
- Informing drivers about certain legal restrictions.
- Helping to regulate traffic flow in an optimal way.
- Helping drivers to safely reach their destinations.

It can be said that road markings are telling the driver, with a specific language, what to do and how to behave in certain situations in traffic, which is especially beneficial in poor visibility conditions. All drivers interact with environmental clues during driving; indeed, driving is a series of decisions based 90% on visual clues (Thurston, 2009). As the overall society age increases, it becomes increasingly important that road systems incorporate sufficient tolerances that cater for deteriorating light perception and also for the longer time it takes for the elderly to react to all of these visual clues. Clear and visible horizontal road markings improve recognition of those clues.

Two important factors describe connection between driver and road markings:

- A driver must be able to see road markings at a certain distance to perceive, process, and react to the information that the pavement marking presents in order to receive adequate information to safely guide the vehicle. Since the required distance increases as the speed of the vehicle increases, it is often described as constant preview time.
- "It has been established that for night time low-beam conditions, a driver requires a minimum recommended preview time (comprising both eye fixation time and driver reaction time) of 3.65 seconds at 80 km/h, of oncoming road geometry to enable safe negotiation without the driver requiring to shift attention away from the road, to look for other clues." (Zwahlen et al., 1998)

Two important components must be evaluated when deciding which road marking material to use. The first component is whether the line or the marking on the pavement is visible during the day. The second component is the retroreflectivity, i.e. visibility at night when headlights reflect off the line. Both components are necessary for the marking to be useful to drivers. Typically, glass beads are dropped on top of the material that is used to give the line marking its retroreflectivity, as shown in Fig. 1. Proper bead embedment is critical.

Retroreflectivity is the portion of incident light from a vehicle's headlights reflected back toward the eye of the driver of the vehicle. Retroreflectivity is provided by road marking materials by glass or ceramic beads that are partially embedded in the surface of the material (Migletz et al., 1998). Retroreflectivity is achieved as long as the reflective elements remain properly embedded — upon their loss or damage, only daytime reflectivity is maintained, as long as the paint is not discoloured and adheres to the surface.

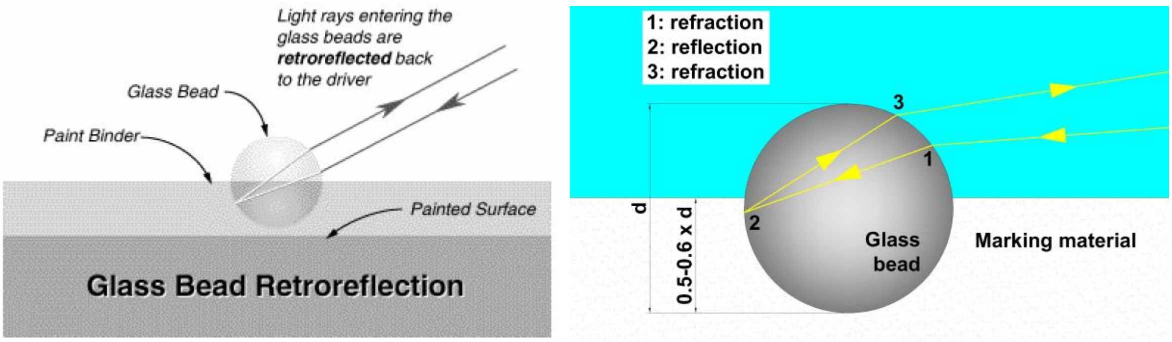



**Fig. 1:**
*Retroreflectivity of glass beads in road marking*
*Source: Asdrubali, F., Buratti, C., Moretti, E., D'Alessandro, F., Schiavoni, S.: Assessment of the Performance of Road Markings in Urban Areas: The Outcomes of the CIVITAS RENAISSANCE Project, The Open Transportation Journal, 2013, 7, 7-19.*

### 3. Materials for road markings

Pavement marking materials can be categorised in several ways. Assigning to categories based on the material type – paint, thermoplastic, plural component systems, and tape is rather frequent. Another method is based on the solvent: Solventborne, waterborne, and solvent-free. Yet another categorising is based on their durability, into conventional, durable, and temporary marking products. Dividing road markings types based on their chemistry is also possible and frequently employed.

Road authorities could divide the marking systems based on the initial or overall expense of application and maintenance. Indeed, many data sources list durability and relative prices of various road markings (Montebello et al., 2000). Even though such information can serve as general guideline, one must remember that only side-by-side comparison could provide reliable scientifically valid information. (Simultaneously, such comparison would not necessarily be fair due to different typically applied film builds and design of different materials for particular layer thickness.) No references of such comparisons could be found.[2] Hence, such data provided by various road authorities has to be taken with a grain of salt. Comprehensive cost-benefit analysis that would include all of the cradle-to-grave factors, including also financial and society costs, is still to be done.

Therefore, none of the categorising is definite. The issue in proper categorising of some waterborne paints shall be discussed below. Table shows the summary of the road marking materials and their assignment to different categories. Impossibility of clear assignment should be obvious.

Road marking materials other than paint are not a focus of this paper and shall not be described in detail, even though some of their strengths and weaknesses are mentioned in Table. Glass beads and other materials to provide retroreflectivity are a topic in themselves and shall not be discussed, either.

[2] One could use as the reference results from the German BASt, but in their carefully controlled laboratory test one again fails into the issue of dissimilar film builds.

**Table 1.**
*Road marking systems*

| Type | Price(a) | Durability | Resin | Solvent | VOC | Risks — Health | Risks — Environmental | Risks — Transport and storage | Typical applied thickness | Advantages | Disadvantages |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Paint / 1K(b) | Moderate | High | Acrylic | Water | <50 g/l (1-3%) | Minimal | Minimal | None | wet 300-1200 µm, dry 200-900 µm | - High durability.<br>- Low environmental footprint.<br>- Safety.<br>- Quick tack-free time of some paints.<br>- Fast drying at proper conditions.<br>- Adhesion to various surfaces. | - Cost.<br>- Slow development of water resistance (risk of washout).<br>- Application in warm dry weather preferred.<br>- Requires the use of stainless equipment. |
| Paint | Very low | Very low | Alkyd or styrenic | Organic | 400-600 g/l (20-35%) | High (solvents) | High (solvents) | Flammable. | wet 300-400 µm, dry 200-250 µm | - Initially inexpensive.<br>- Easy to apply.<br>- Easy clean-up.<br>- May use inexpensive equipment. | • - Very low durability.<br>- Poor UV resistance.<br>- Rather slow drying.<br>- High environmental impact.<br>- Alkyds require organometallic driers.<br>- Solvent harmfulness and odour. |
| Paint | Low | Low to moderate | Acrylic | Organic | 400-600 g/l (20-35%) | High (solvents) | High (solvents) | Flammable. | wet 300-400 µm, dry 200-250 µm | - Initially inexpensive.<br>- Easy to apply.<br>- Dry in adverse conditions.<br>- Easy clean-up.<br>- May use inexpensive equipment. | - Low durability.<br>- High environmental impact.<br>- Solvent harmfulness and odour. |
| Paint / 2K(c) | High | High | Epoxy or urethane or urea (also modified) | Organic | 400-600 g/l (25-35%) | High (solvent, epoxy toxicity) | High (solvent, toxicity to aquatic life) | Flammable. | wet 300-400 µm, dry 200-250 µm | - High durability.<br>- Reasonable cost-performance balance. | - Quite expensive.<br>- Risk of human allergic reaction.<br>- High environmental impact.<br>- Solvent odour.<br>- Poor ultraviolet stability.<br>- Poor flexibility.<br>- Slow drying. |
| 2K | Very high | Very high | Epoxy or urethane or urea (also modified) | — | <50 (0-1%) | High (epoxy toxicity, amine corrosive) | High (toxicity to aquatic life) | Flammable. | 500-1500 µm | - High durability.<br>- No solvents.<br>- Super-quick drying. | - Very costly.<br>- Special heated application equipment needed.<br>- Corrosive amine for curing.<br>- Rather poor UV stability. |

| Type | Price(a) | Durability | Resin | Solvent | VOC | Risks: Health | Risks: Environmental | Risks: Transport and storage | Typical applied thickness | Advantages | Disadvantages |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Coldplastic(d) | Very high | Very high | Acrylic or polyester | — | <50 (0-1%) | Moderate | Moderate | Flammable, uncontrolled polymerisation. Organic peroxide transport and storage. | 1000-3000 µm | - Very high durability.<br>- No solvents.<br>- Quick drying.<br>- Excellent adhesion to some surfaces. | - Costly.<br>- Special application equipment needed.<br>- Sensitive to moisture and temperature (location-specific formulations needed).<br>- Risk of uncontrolled polymerisation. |
| Tape(d) | Moderate to extremely high(d) | Moderate to very high | Acrylic and other | — | <50 (0-1%) | None known | None known | None | 1000-3000 µm | - Superior durability.<br>- Tailoring durability with specific requirements.<br>- Ease of application.<br>- Very high retroreflectivity.<br>- Ease of removal of temporary markings. | - Extremely high cost.<br>- Application surface preparation (poor adhesion can lead to complete peel-off). |
| Thermo | Very low | High | Alkyd or hydrocarbon | — | <50 (0-1%) | Moderate (heat during application, unpleasant odour) | Minimal | None. | 2000-5000 µm | - Durability with quite constant retroreflectivity.<br>- Low cost.<br>- Ease of application. | - Requires heat to apply.<br>- High environmental cost.<br>- Discolouration from asphalt oils possible.<br>- Surface preparation flaws may lead to failures.<br>- Sensitive to air temperature (location-specific formulations needed). |

(a) Relative price estimate of materials and application per unit weight; application film build or durability not taken into consideration. (b) Due to irreversibility of the solidification reaction, quick-set waterborne single component material is somewhat different as compared to the typical paint (*vide infra*). (c) Two-component solventborne materials are applied like paint. (d) Tape comes in many varieties, including designed for temporary marking; hence, its price depends on durability.

*Source: By Authors*

**Paints**

Paint remains the most widely used road marking material in the world since it was first applied as a centreline in 1911 in Michigan, U.S.A. Paint application is done by simple spraying using high- or low-pressure equipment; it does not require heating or special technologically advanced application machines. Paints generally have good affinity for glass beads, which sometimes are pre-mixed or more frequently dropped-on or injected to the formed film.

All paints are composed of the resin (historically, chlorinated rubber or alkyd were used; more recently styrene-acrylic, acrylic-alkyd blends or 100% acrylic are preferred), pigments and fillers, solvents (organic or water), and numerous additives that assure stability, grinding, absence of foam, proper film formation, etc. Paints can be divided into solventborne and waterborne, depending on the solvent used, but one must remember that physicochemical process of drying/curing of the currently used high-end waterborne paints is simultaneously the major difference.

Partially as a part of the worldwide movement to limit emissions of Volatile Organic Compounds (VOC) and to minimise the effects on human health and partially to seek improved performance, solventborne paints are slowly being replaced by waterborne paints or solvent-less road marking materials.

****Solventborne paints**

Modern solventborne paints are based on acrylic resins (rarely styrenic- or alkyd-acrylic blends) that are dissolved in organic solvents like esters or ketones. Some countries still permit aromatic solvents, like toluene, despite their harmfulness for the environment and human health hazards[3] (McMichael, 1988). Heavy metals like chromium or lead are not used anymore in developed countries any more[4]. VOC content of a typical high-solids solventborne road marking paint is 400-500 g/l (~25%).

In solventborne paints, solid polymer (the resin) is dissolved in an organic solvent. After application, evaporation of the solvent causes the polymer to solidify and thus the film forms. Penetration of the surface cracks by these low surface tension materials assures good adhesion. Since there is no chemical reaction during drying, the resulting material can be re-dissolved. In-can stability of solventborne paints is most of the time rather poor – settling and thickening occur, but they can be easily alleviated by the users by mixing and addition of solvent.

Solventborne paints can be conveniently applied at wet film builds reaching only 400 μm – higher film builds lead to exponential increase in drying time. For their application, air temperatures should be 5-40 °C and above dew point; the road surface must be dry and clean and should not be hotter than 50 °C. Their drying is strongly influenced by air and surface temperature, air movement, applied film build, and the utilised solvent blend, but is generally not affected by humidity.

Typically, solventborne paints, either alkyd or acrylic, are not suitable for high-volume roadways due to their poor durability: The expected service life on a side-line marking is only 6-12 months.

****Waterborne paints**

Before waterborne paints are discussed, a property unique to them must be explained: washout resistance. The applied waterborne paint achieves washout when it no longer is affected by water, which is later and in some cases much later than dry-through. There is no standard protocol of testing washout. Procedure using exposing the paint to running water to imitate heavy rain is frequently used. A method preferred by the authors is a water droplet – finger push-and-twist method: A drop of water is applied to paint surface for 1 minute and then finger with pressure roughly 3 kg (subjective, but quite reproducible with experienced user – as a more scientific laboratory alternative, one could utilise an instrument such as Erichsen *Dry Time Tester*) is applied together with twisting motion that is to imitate car tyres moving on the surface – washout is achieved when the paint is no longer affected by the action. Such test can be conveniently performed at the application site, too. Washout time very strongly depends on the utilised resin and additives and is affected by weather conditions, in particular humidity. Under normal dry conditions, there is no need to wait for washout time to open the road to traffic – this property is absolutely critical only when there is a risk of rain.

First waterborne road marking paints were commercialised in the 1980s. Those paints were quite similar to solventborne coatings in terms of solvent evaporation causing drying, but their main solvent was water. They suffer from sluggish drying and very slow development washout resistance; hence, their use is quite limited. On the other hand, good durability and in-can stability could be achieved with these systems.

Quick-set binders were developed and patented by Morton International and in the early 1990s (Clinnin et al., 1991). Generally, the technology requires the polymers to be modified by acidic moieties. The presence of carefully selected amine and high pH assure that the resin does not precipitate. Upon application, drop of pH occurs and the polymer solidifies, which leads to film formation. Irreversibility of the reaction and significantly enhanced durability may suggest that the quick-set waterborne paint is something in-between a one-component and two-component system. Hence, dried water-based paint cannot be re-dissolved in water or even in most common organic solvents. Surface tension of waterborne paints is higher as compared to solventborne materials, yet their adhesion to road surface and old markings is generally really good.

From this technology comes another unique property of waterborne road marking paints – they stop being tacky very quickly. Clear distinction between tack-free, dry-through, and washout resistance times can be measured. For the applicator, quick tack-free time means lesser unsightly marks caused by careless or impatient drivers encroaching on freshly painted markings.

Application at film builds reaching 1 mm is possible without very significant increase in drying times – unlike solventborne paints, increase in dry-through and washout times appears linear.

[3] (a) In Poland up to 2% of aromatic solvents are permitted per "Warunki Techniczne. Poziome znakowanie dróg. POD-2006. Seria „I" - Informacje, Instrukcje." IBDiM, Warszawa, 2007; (b) In the United States: Federal Specification TT-P-115F (5.12.1985): "Paint, traffic (highway, white and yellow)" does not limit the use of solvents and indeed, even xylenes are occasionally used (*Source: Authors*).

[4] (a) United States Federal Specification TT-P-1952E (6.08.2007): Paint, traffic and airfield marking, waterborne. (b) EN 1436 norm.

VOC content of a typical waterborne road marking paint is less than 50 g/l (~2%), coming from the required additives. Properly designed waterborne paints have shelf life of about one year – no settling or skin formation should occur during storage and the paint is ready to apply as delivered. Minor viscosity increase is normal and the paints can be thinned by the applicators with a tiny amount of water. Paint has to be protected from frost and extreme heat[5]. The same equipment and procedures as are utilised for application of solventborne paints can be followed, with two caveats: only stainless steel couplings are allowed and the paint cannot be permitted to dry in the equipment (which forces the applicators to maintain clean work environment).

Waterborne road marking paints are compatible with a variety of surfaces, including bituminous and concrete materials. As with all other materials, they show sensitivity to freshly applied concrete or asphalt. Waterborne paints adhere well to the existing markings independently on their type, even to thermoplastic masses.[6]

## 4. Characteristics and application of waterborne road marking paint

Waterborne road marking paint is commonly used in the United States – not necessarily because of its environmental friendliness, but due to excellent cost-performance balance. Recently, due to huge volume, cost of binders for waterborne paints is lower as compared to Europe – hence, additional advantage on that market is the price. In Europe, waterborne paints are used in Scandinavia, where solventborne paints are banned, and also in the France, Italy, Spain, and some areas of Germany. Other European countries use mostly, if not solely, paints based on organic solvents.

A cradle-to-grave Life Cycle Analysis (LCA) that measured environmental impact in numerous categories was prepared and presented simultaneously by Dow, major manufacturer of the binders for waterborne paints (Kheradmand, 2012), and by Evonik, major manufacturer of monomers for coldplastic (Klein, 2012). Both analyses indicate advantages of products made by their companies. The authors herein do not assess correctness or validity of either analysis.

For the road marking applicator, the advantage of using waterborne paint is the absence of hazardous and flammable ingredients, which directly translates into lower transportation and storage costs. However, a disadvantage is that the paint must be protected from frost and extreme temperature. At the application time, advantages and disadvantages of waterborne road marking materials can be somewhat schizophrenic. Drying at favourable conditions can be much faster as compared to other materials (see Table below), but simultaneously at less favourable humid conditions or in case of sudden rain there is a risk of washout. The applicators must be properly trained and must understand the limitations of the systems that they are applying – in Scandinavia, where rain is frequent, there are no problems with the paint. There, waterborne paint is frequently applied at speeds over 25 km/h, leading to improved throughput.

For the road authorities, advantage of waterborne paints are numerous: faster drying at proper conditions means lesser traffic congestion, higher durability means lower overall road maintenance cost, and the environmental friendliness means lower impact on our and our planet's health. The disadvantage is the immediate price tag that must be paid – even though long-term performance seems to be balancing the initial cost.

One must remember that there are several qualities of waterborne paints and it would be unfair to compare the low-end material with, for example, coldplastic – especially with different applied film builds. That unfair comparison would apply particularly to waterborne paints based on slowly-drying low-end waterborne binders (Montebello et al., 2000). Similarly, some documents lists waterborne paints as having no pick-up times of up to 15 minutes, dry-through of up to 1 hour – that is all true, but does not apply to fast-dry paints.

Some waterborne paints can be accelerated to furnish drying and washout within a few minutes without affecting the paint performance, by the means of a special catalyst that is intermixed with glass beads. Hence, application season can be meaningfully extended and jobs in areas with highly congested traffic can be done very quickly. During our field evaluation, one such system was included – drying times were outstanding, indeed, as shown in Table and the initial retroreflectivity results are matching the standard system.

Durability in the field is typically measured by retroreflectivity. This is affected by the utilised glass beads, weather conditions at the application site, traffic, road surface quality, winter conditions and maintenance, etc.

Under laboratory conditions, durability of the binder itself can be conveniently measured using abrasion, according to Taber procedure per ASTM C1353. Fig. 2 illustrates binder durability achieved with the road marking materials that were applied during our evaluation. Poor performance of solventborne paints was surprising, but repeat experiments confirmed its low durability. The result of this on the road is visible after loss of glass beads, which in addition to providing retroreflectivity serve to protect the paint – solventborne paints very quickly deteriorate whereas waterborne paints are retained. Indeed, in a study done by Dow at an application site in Pennsylvania, U.S.A., a high-end waterborne paint applied at 600 μm (900 g/m$^2$) wet film build (about 400 μm dry film) outperformed in long-term testing a thermoplastic road marking material applied at 3000 μm (Randazzo, 2013).

[5] United States specifications requires freeze-thaw and heat stability (Cf. TT-P-1952E.). In Europe, there is no freeze-thaw requirement and the heat stability demands are less strict (Cf. specification EN 1871).

[6] www.nzrf.co.nz/techdocs/NZRF-Materials-Guide.pdfthe New Zealand data.

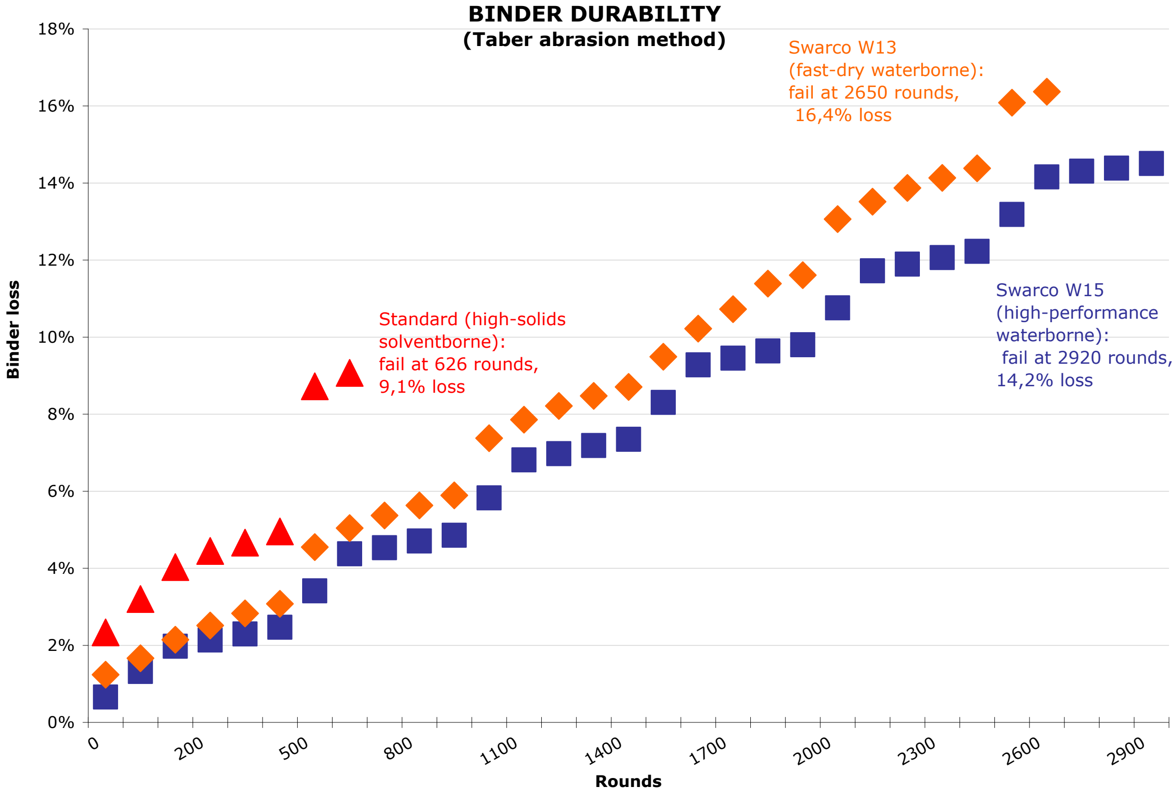



**Fig. 2.**
*Taber abrasion testing of binder durability (600-µm drawdown).*
*Source: By Authors*

### Evaluation

Recently, a field application was done to (1) assess feasibility and ease of applying waterborne road marking paints in Croatia, (2) evaluate the paints' performance in comparison with other road marking materials, and (3) test various glass beads for retroreflectivity and their retention under normal use conditions on a heavily travelled road.
The selected road stretch near Zagreb is used by approximately 8000 vehicles per day, has two-years old asphalt in good condition, and was previously marked with the standard solventborne paint commonly utilised in Croatia (Croatian Roads Ltd., 2013). The application work was done by a crew from Chemosignal d.o.o (Zagreb, Croatia), under the direction of Faculty of Transport and Traffic Science, Department for Traffic Signalization, University of Zagreb, using paints from SWARCO Limburger Lackfabrik GmbH (Diez, Germany) and drop-on reflective materials from M. Swarovski GmbH (Amstetten, Austria). For comparison, a standard high-solids solventborne was included.
The crew has never worked with waterborne paints, but was experienced with the application of solventborne paints and coldplastic and thermoplastic masses. They were instructed in regards to the required cleaning procedures and cautions. After a few adjustments of the machinery (pressure, nozzle, etc.) the waterborne paints were applied at target film builds. Change of paint and change of film builds were accomplished without meaningful difficulties, but adjustments were necessary each time. Lesser required coning and absence of the tyre marks from vehicles driven by careless and impatient drivers who were encroaching on the freshly marked lines were noted. Accidentally, the crew run out of paint and that lead to initiation of paint drying in the drum – immediate cleaning of the filters solved the problem. There were no noted issues with glass beads flow and embedment. (A little curiosity must be noted here: The utilised small American-made application machine (Graco) has a container for glass beads holding only 49 kg of beads, so two full bags did not fit!)
Systems were applied on side and centre lines, each at the road length of 500 m, at 15-cm line width, at two film builds and with several classes and sizes of glass beads. Amongst the systems, one must note high-performance paint W15 that was applied with dropped-on glass beads that included drying accelerator. That system, applied at 600 µm wet film dried-through in less than 5 minutes and achieved washout in less than 10 minutes! The applicators immediately noticed exceptionally quick drying.
Drying and washout times shown in Table were achieved in the field. The weather conditions were perfect: mostly sunny with mild wind, roadway temperature 30-40 °C and air temperature 23-28 °C at 40-50% relative humidity.

**Table 2.**
*Drying and washout resistance of various road marking materials during application near Zagreb.*

| Road marking material | Layer thickness | Dry time | Washout resistance |
|---|---|---|---|
| Swarco W13, fast-dry waterborne paint | 400 μm | 4 min. | 8 min. |
| Swarco W15, high-performance waterborne paint | 600 μm | 15 min. | 36 min. |
| Swarco W15, high-performance waterborne paint, with drying accelerator | 600 μm | 4 min. | 8 min. |
| Standard solventborne high-solids paint | 400 μm | 17 min. | — |

*Source: By Authors*

Initial retroreflectivity depended on the film build and glass beads type and size. It is much too early for any conclusions. Preliminary observations indicate that care should be taken with initial measurements as it takes some time of exposing the marking to traffic before the maximum retroreflectivity is achieved, that film build and bead size must be matched, and that type of glass beads and their size profoundly influence retroreflectivity and the rate of its diminishment. Selected data is provided in Table.

**Table 3.**
*Initial measurements of retroreflectivity.*

| Paint | Applied film build | Glass beads type and size (μm) | Initial retroreflectivity (mcd/m²/lx)(a) |
|---|---|---|---|
| W15 high-performance waterborne, accelerated | 600 μm | Megalux-Beads® 600-2000 | 439 |
| W15 high-performance waterborne | 600 μm | Megalux-Beads® 600-2000 | 487 |
| W15 high-performance waterborne | 600 μm | Swarcolux50 212-1400 | 389 |
| W15 high-performance waterborne | 400 μm | Swarcolux50 212-1400 | 494 |
| W13 fast-dry waterborne | 400 μm | SolidPlus100 212-850 | 702 |
| W13 fast-dry waterborne | 400 μm | SolidPlus30 300-1000 | 555 |
| W13 fast-dry waterborne | 400 μm | Swarcolux50 212-1400 | 267 |
| Standard solventborne high-solids paint | 400 μm | Swarcolux50 200-800 | 374 |

(a) Readings obtained using dynamic testing equipment, average of readings taken every 25 m throughout the applied line length.

*Source: By Authors*

**5. Conclusions**

Waterborne road marking paints seem a viable alternative to the currently utilised solventborne paints and even have a chance to expand into the application areas of plural component materials. Durability of these paints, as measured by Taber abrasion, is outstanding and significantly outperformed the compared standard solventborne materials. Drying and washout resistance times measured in the field during perfect application conditions were excellent and faster as compared to simultaneously applied solventborne paints.

From the point of view of application, there are no problems as long as the personnel is appropriately trained, assures that there is no risk of oncoming rain, and maintains clean work environment. Advantage for the applicator companies includes savings in transport costs and absence of hazardous flammable and toxic goods. Lower emissions and lack of harmful chemicals benefit human health. For the road authorities, the use of waterborne paints would mean higher initial cost that would be offset by longer usable paint life. In addition, quick drying under good conditions would permit lesser road blocking and easing traffic congestion.

The biggest disadvantages of waterborne road marking paints remain sensitivity to rain until washout is achieved and sluggish drying under high humidity conditions. Price of these materials is also not low, but one must remember that the technology was not designed for minimising cost, but for maximising the overall benefit.

During the test application work near Zagreb, the crew was satisfied with applying waterborne paints. Education of the application crews appears critical to assure acceptance of the technology where it is currently not utilised. With minor training, the limitations of waterborne paints can be explained, which should assure high quality of the finished work and avoidance of problems. Testing of retroreflectivity shall be done periodically — the initial results are meeting our expectations.

**Acknowledgement**

The Ministry of Education, Youth and Sports of the Czech Republic, Project POSTDOK, CZ.1.07/2.3.00/30.0021 “Enhancement of R&D Pools of Excellence at the University of Pardubice“, financially supported this work.

**References**

Asdrubali, F., Buratti, C., Moretti, E., D'Alessandro, F., Schiavoni, S.: Assessment of the Performance of Road Markings in Urban Areas: The Outcomes of the CIVITAS RENAISSANCE Project, The Open Transportation Journal, 2013, 7, 7-19.

Clinnin, D. D., Heiber, W. G., Lewarchik, R. J.: Fast dry waterborne traffic marking paint. Patent US5340870 (28.06.1991).

Croatian Roads Ltd.: Traffic Counting on the State Croatian Roads, 2013.

European Commission, Communication from the Commission to the European Parliament, the Council, the European economic and social committee and the Committee of the regions. Towards a European road safety area: policy orientations on road safety 2011- 2020, Brussels, Belgium, July 2010. Available from: http://ec.europa.eu/transport/road_safety/pdf/com_20072010_en.pdf

Horberry, T., Anderson, J., Regan, M.A.: "The possible safety benefits of enhanced road markings: a driving simulator evaluation", Transportation Research, Part F: Traffic Psychology and Behaviour, vol. 9, pp. 77-87, 2006.
http://pubchem.ncbi.nlm.nih.gov/rest/chemical/toluene

Kheradmand, H.: Life Cycle Assessment. Road Marking Technologies Eco-Profile. Intertraffic, Amsterdam, 2012.

Klein, A.: Lifecycle under the lens, Intertraffic World, Annual Showcase 2012, Infrastructure, p 210.

McMichael A. J.: Carcinogenicity of benzene, toluene and xylene: epidemiological and experimental evidence. IARC Sci. Publ. 1988; (85): 3-18 and a plethora of other references available through

Migletz, J., Fish, J., Miles, J.: "Roadway Delineation Practices Handbook", U.S. Dept. of Transportation, Federal Highway Administration, Washington, D.C., U.S.A., Tech. Rep. FHWA-SA- 93-001 August 1994.

Migletz, J., J. Graham, K. Bauer, and D. Harwood. Field Surveys of Pavement Marking Retroreflectivity. Transportation Research Record 1657. Transportation Research Board, National Research Council, Washington, D.C., 1999, pp. 71-78.

Miller, T. R. (1992). Benefit–cost analysis of lane marking. Transportation Research Record, 1334, 38–45.

Montebello, D., Schroeder, J.: Cost of Pavement Marking Materials, Final Report, Minnesota Local Road Research Board, March 2000.

Randazzo, C.: Waterborne Traffic Markings Technology and Testing, Dow Coating Materials, 2013.

Retting, R.A., McGee, H.W., Farmer, C.M.: "Influence of Experimental Pavement Markings on Urban Freeway Exit-Ramp Traffic Speeds", Transportation Research Record, vol. 1705, pp. 116-121, 2000

Thurston, P.: Pavement Markings Role in Enhancing Road Safety Strategies, Roadmarking Industry Association of Australia, 2009.

Zwahlen, H. T., Schnell, T.: "Visibility of Road Markings as a Function of Age, Retro-reflectivity Under Low-beam and High-beam Illumination at Night," Paper No. 980285, presented at the 77th Annual Meeting of the Transportation Research Board, Washington, DC, January 11-15, 1998

.

# TRAFFIC SIGN ANALYSIS AS A FUNCTION OF TRAFFIC SAFETY ON CROATIAN STATE ROAD D3

**Dario Babić[1], Anđelko Ščukanec[2], Mario Fiolić[3]**
[1,2,3]*Faculty of Transport and Traffic Science, Vukelićeva 4, 10000 Zagreb, Croatia*

**Abstract:** Traffic systems are dynamic systems where many conflicting situations may result in traffic accidents. The analysis of possible causes of accidents showed that there are four main factors contributing to their occurrence: human, vehicle, road and environment. Road as a safety factor is characterized by a number of elements such as: road route, technical elements, pavement condition, equipment and road lightning, intersections, etc. An important element of traffic safety is road equipment, including traffic signs. Traffic signs are elements defined as a set of specially coded labels intended for traffic participants, located in the vertical plane with respect to the traffic area and they represent the basic means of communication between the authorities and road users.
Improper installation and maintenance of traffic signs can affect the process of signs perception and quality of transmitted messages, thus affecting the drivers' reaction and general traffic safety. For these reasons, traffic sign inspection should be conducted at least once a year to test their retroreflection and other technical characteristics.
In this article an analysis of retroreflective and technical characteristics of road signs on state road D3 in Zagrebačka County will be presented. The main objective of the detailed analysis is to gain insight into the condition and quality of traffic signs on the road.

**Keywords:** road safety, traffic signs, retroreflection, traffic accidents

## 1. Introduction

Traffic signs are elements defined as a set of specially coded labels intended for traffic participants, located in the vertical plane with respect to the traffic area created from the need for traffic management and which is their most important function. Their other functions are related to the: traffic regulation, information and orientation/guidance. With respect to traffic regulation, their significance lies in the fact that they provide information about speed limits, possible dangers etc. to road users. In addition, one of the functions of traffic signs is also orientation or guidance to specific locations.
Classification of traffic signs depends on observed characteristics. The most common classification is by function, shape, size, their meaning, retroreflection material (class), degree of standardization, method of production etc.
Classification of traffic signs by their function and retroreflection material is the major classification. In Croatian regulations traffic signs are, by their function, classified into[2]:

- danger signs
- mandatory signs
- information signs
- guide signs
- additional panels
- traffic equipment
- variable traffic signs.

From the above it can be concluded that traffic signs are primary means of communication between the road authorities and traffic participants and must guarantee a safe and comfortable journey in all conditions.
Knowing that there are four main factors contributing to traffic accidents (human, vehicle, road and environment), it can be concluded that the traffic engineers' main focus is on the road factor. Road as a safety factor is characterized by a number of elements, such as: road route, technical elements, pavement condition, equipment and road lightning, intersections, etc. An important element of traffic safety is road equipment, including traffic signs.
As traffic signs are primary means of communication between the road authorities and traffic participants, their improper installation and maintenance can affect the perception of road signs, the quality of the transmitted message that they carry, thus affecting drivers' reaction time and general road safety.
For quality and safe traffic flow, all users must receive clear, readable, continuous and timely information in order to comply with them and adjust their behavior, thereby avoiding non desirable situation. To achieve these demands, retroreflection materials are used for traffic signs. These retroreflective materials contain spherical reflectors (glass beads) or micro prisms which ensure visibility of traffic signs in low visibility conditions when the risk of traffic accidents is the largest.

[1] Corresponding author: dario.babic@fpz.hr
[2] Ordinance on traffic signs, signals and equipment on roads (OG 33/05, 64/05, 155/05, 14/11)

Glass beads or micro prisms are molded in durable, transparent material from 0.14 to 0.22 mm thick and are produced in all the colors that are used for traffic signs (Stanić et al., 2003.). Today, there are three types of retroreflective materials for traffic signs:

a) Class I
b) Class II
c) Class III

Retroreflection of traffic signs can be measured using dynamic or static retroreflectometer. Dynamic measurement is a system that uses retroreflectometer mounted on the vehicle which conducts measurement while driving. Static measurement uses handheld retroreflectometer containing an internal light source and a photoreceptor, and when testing, it is put on the surface of the sign in order to exclude the impact of daylight.
In this article retroreflection and technical characteristics of road signs on Croatian state road D3 in Zagrebačka County were analyzed. The main objective of the analysis is to gain insight into the condition and quality of traffic signs. Based on the stated analysis and analysis of traffic accidents, a solution related to the increase of the road signs quality in order to reduce traffic accidents on the analyzed road will be proposed.

## 2. Methodology for measuring retroreflection of traffic signs

Traffic signs should be inspected at least once a year to verify their retroreflection and other technical characteristics. The coefficient of retroreflection is defined as the coefficient of luminous intensity, $R_A$ of a plane retroreflecting surface to its area or as a ratio of returned intensity to incident illumination divided by the area of the retroreflector. The metric unit for retroreflection coefficient is $cd/lx1/m^2$ (candelas per lux per square meter) (Austin et al., 2009.).
As said earlier, retroreflection of traffic signs can be measured using dynamic or static retroreflectometer. The geometry of these handheld instruments should match the values of the European Standard (EN 12899-1: Fixed, vertical road traffic signs - Part 1: Permanent signs) which implies an observation angle of $0.33^0$ and an entrance angle of $5^0$. The entrance angle is primarily determined by the position of the sign on the side of the road and the geometry of an oncoming vehicle position and represents the angle that is formed between the light rays falling on the surface of the sign and the line that goes vertically from the surface. The observation angle is the angle between the incoming light ray and the reflected ray[3].
Entrance and observation angle are shown in Figure 1.

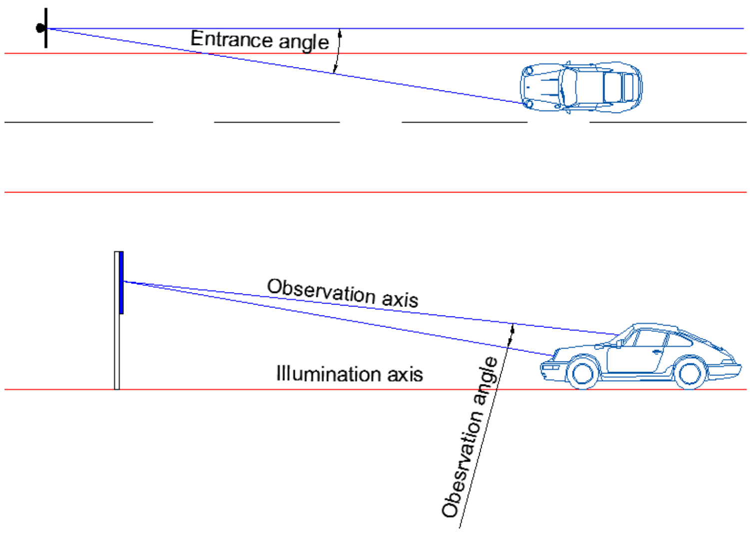



**Figure 1.**
*Entrance and observation angles for traffic sign*
*Source: Ščukanec, A., Babić, D., Sokol, H.: Methodology for measuring traffic signs retroreflection, 1st International Scientific Forum, Tirana 2013., ISBN: 978-608-4642-15-2*

Since it is assumed that the retroreflective material reflects light directly back to the source, the optimum observation angle should be zero. However, in reality it is not so because drivers eyes are higher than the front headlights.
The minimum initial coefficient of retroreflection $R_A$ ($cd/lx1/m^2$) retroreflective signs must match the values shown in Table 1, Table 2 and Table 3. The coefficient of retroreflection ($R_A$) of all printed colours, except white, shall be not less than 70 % of the values in tables[4].

[3] EN 12899-1: Fixed, vertical road traffic signs - Part 1: Permanent signs
[4] EN 12899-1: Fixed, vertical road traffic signs - Part 1: Permanent signs

**Table 1.**
*Retroreflection coefficient $R_A$: Class I ($cd/lx^1/m^2$)*

| Geometry of measurements | | Colour | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| α | β1 ($\beta_2$=0) | white | yellow | red | green | blue | brown | orange | gray |
| 12´ | +5° | 70 | 50 | 14.5 | 9 | 4 | 1 | 25 | 42 |
| | +30° | 30 | 22 | 6 | 3.5 | 1.7 | 0.3 | 10 | 18 |
| | +40° | 10 | 7 | 2 | 1.5 | 0.5 | # | 2.2 | 6 |
| 20´ | +5° | 50 | 35 | 10 | 7 | 2 | 0.6 | 20 | 30 |
| | +30° | 24 | 16 | 4 | 3 | 1 | 0.2 | 8 | 14.4 |
| | +40° | 9 | 6 | 1.8 | 1.2 | # | # | 2.2 | 5.4 |
| 2° | +5° | 5 | 3 | 1 | 0.5 | # | # | 1.2 | 3 |
| | +30° | 2.5 | 1.5 | 0.5 | 0.3 | # | # | 0.5 | 1.5 |
| | +40° | 1.5 | 1.0 | 0.5 | 0.2 | # | # | # | 0.9 |
| # Indicates "Value greater than zero but not significant or applicable". | | | | | | | | | |

*Source: EN 12899-1: Fixed, vertical road traffic signs - Part 1: Permanent signs*

**Table 2.**
Retroreflection coefficient RA: Class II ($cd/lx^1/m^2$)

| Geometry of measurements | | Colour | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| α | β1 ($\beta_2$=0) | white | yellow | red | green | dark green | blue | brown | orange | gray |
| 12´ | +5° | 250 | 170 | 45 | 45 | 20 | 20 | 12 | 100 | 125 |
| | +30° | 150 | 100 | 25 | 25 | 15 | 11 | 8,5 | 60 | 75 |
| | +40° | 110 | 70 | 15 | 12 | 6 | 8 | 5,0 | 29 | 55 |
| 20´ | +5° | 180 | 120 | 25 | 21 | 14 | 14 | 8 | 65 | 90 |
| | +30° | 100 | 70 | 14 | 12 | 11 | 8 | 5 | 40 | 50 |
| | +40° | 95 | 60 | 13 | 11 | 5 | 7 | 3 | 20 | 47 |
| 2° | +5° | 5 | 3 | 1 | 0,5 | 0,5 | 0,2 | 0,2 | 1,5 | 2,5 |
| | +30° | 2,5 | 1,5 | 0,4 | 0,3 | 0,3 | # | # | 1 | 1,2 |
| | +40° | 1,5 | 1,0 | 0,3 | 0,2 | 0,2 | # | # | # | 0,7 |
| # Indicates "Value greater than zero but not significant or applicable". | | | | | | | | | | |

*Source: EN 12899-1: Fixed, vertical road traffic signs - Part 1: Permanent signs*

**Table 3.**
*Retroreflection coefficient $R_A$: Class III ($cd/lx^1/m^2$)*

| Geometry of measurements | | Colour | | | | | |
|---|---|---|---|---|---|---|---|
| α | β1 ($\beta_2$=0) | white | yellow | red | green | blue | orange |
| 0.1° | +5°<br>+20°<br>+30° | 850<br>600<br>425 | 550<br>390<br>275 | 170<br>120<br>85 | 85<br>60<br>40 | 55<br>40<br>28 | 260<br>130<br>95 |
| 0.2° | +5°<br>+20°<br>+30° | 625<br>450<br>325 | 400<br>290<br>210 | 125<br>90<br>65 | 60<br>45<br>30 | 40<br>30<br>20 | 140<br>100<br>70 |
| 0.33° | +5°<br>+20°<br>+30° | 425<br>300<br>225 | 275<br>195<br>145 | 85<br>60<br>45 | 40<br>30<br>20 | 28<br>20<br>15 | 95<br>65<br>49 |

*Source: Consiglionazionaledellericerche: Common understanding of assessment procedure (CUAP): Microprismatic retro-reflective sheetings, 2002.*

On state road D3, measurements of traffic signs retroreflection were conducted by the Department for Traffic Signalization at the Faculty of Transport and Traffic Sciences using handheld retroreflectometer Zehntner ZRS 6060 in accordance with the European and National standards and specifications. When measuring retroreflection, each sign was measured four times: up, down, left and right. The relevant value of retroreflection represents the average values of all four measurements. Except the retroreflection value, several other elements were analyzed: sign name and code, graphic display (sign picture), dimension and height and distance from the edge of sign, colours of the surface, edge and symbols, the way the sign is installed and fixed, information about the sign producer, retroreflective material, etc (Ščukanec et al., 2013.).

## 3. Analysis of the current state of road signs and traffic accidents on the state road D3

As mentioned earlier, road signs retroreflection was measured using handheld retroreflectometer Zehntner ZRS 6060 in accordance with the European and National standards and specifications. Except the retroreflection, technical characteristics such as height, distance from the edge, compliance of signs with the Regulations, etc. were analyzed.
In this section, an analysis of the current state of road signs and traffic accidents on the state road D3 will be presented.

### 3.1. Analysis of retroreflection measurements and technical characteristics of traffic signs on the state road D3 in Zagrebačka County

On the state road D3 in Zagrebačka County, in the length of 17.8 km, on the section Dubovec (g.ž.) - Lužan (g.ž.), a total of 414 traffic signs were measured. Signs are according to Croatian Regulations classified by their meaning, as shown in Table 4.

**Table 4.**
*Total number of analysed traffic signs classified by meaning*

| Road | D3 | |
|---|---|---|
| Section | Dubovec (g.ž.) - Lužan (g.ž.) | |
| **Sign classification** | | **Total** |
| Danger signs - A | | 110 |
| Mandatory signs - B | | 74 |
| Information signs – C | | 174 |
| Guide signs - D | | 30 |
| Additional panels - E | | 15 |
| Traffic equipment - K | | 11 |
| **Total** | | **414** |

From a total of 414 traffic signs which are located on the state road D3, 110 traffic signs are warning signs, 74 are mandatory signs, the 174 information signs, 30 signs are guide signs, 15 are additional panels and 11 are signs that are classified as traffic equipment (Figure 2).

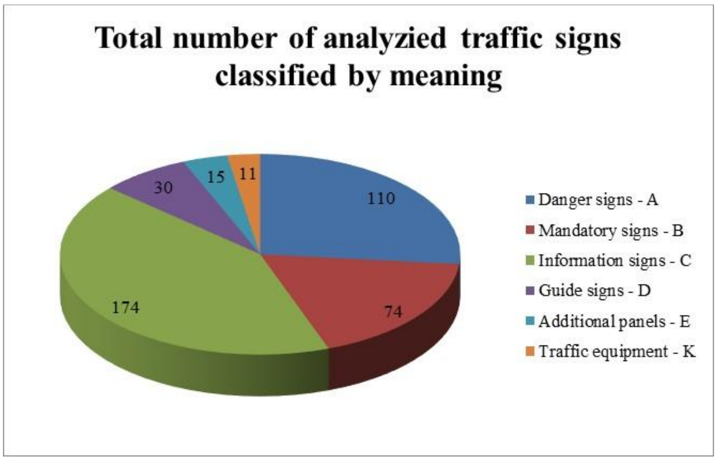



**Figure 2.**
*Total number of analysed traffic signs classified by meaning*

The analysis of the retroreflective material from which traffic signs are made showed that 314 signs are made of class I, 61 from class II and 39 from class III material, as shown in Figure 3.

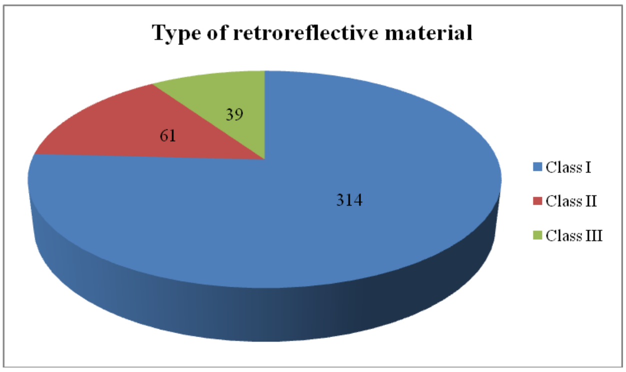



**Figure 3.**
*Total number of analysed traffic signs classified by class of retroreflective material*

Retroreflection measurements showed that 284 or 68.3% of traffic signs meet minimum prescribed value while 132 or 31.7% of them do not. Figure 4 shows the number of traffic signs classified by meaning that meet or do not meet the minimum value of retroreflection.

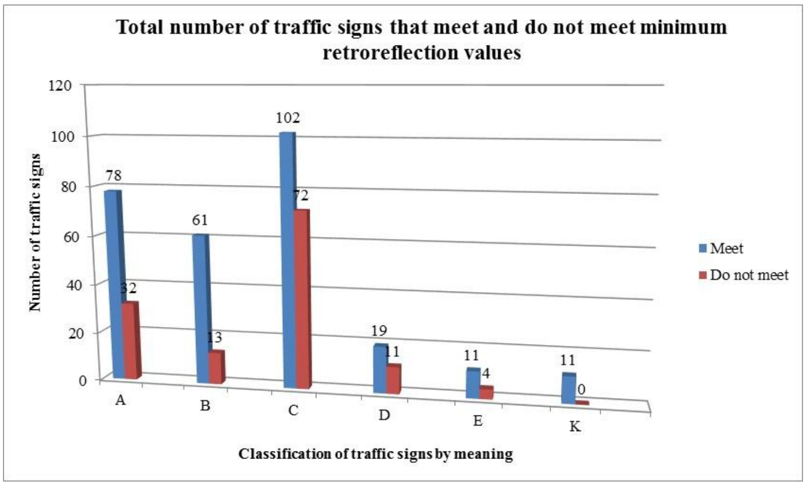



**Figure 4.**
*Total number of traffic signs that meet and do not meet minimum retroreflection values*

Figure 5 shows the total number of traffic signs classified by the class of retroreflective material that meet and do not meet the minimum value of retroreflection. From the Figure it can be seen that out of 314 signs made of class I, 210 of them (66.8%) meet the minimum retroreflection values, and 104 (33.2%) do not. From 61 sign made of class II material, 49 (80.3%) of them meet minimum values and 12 (19.7%) do not. Also, 23 (59%) traffic signs made of class III meet, while 16 of them (41%) do not meet the prescribed minimum values.

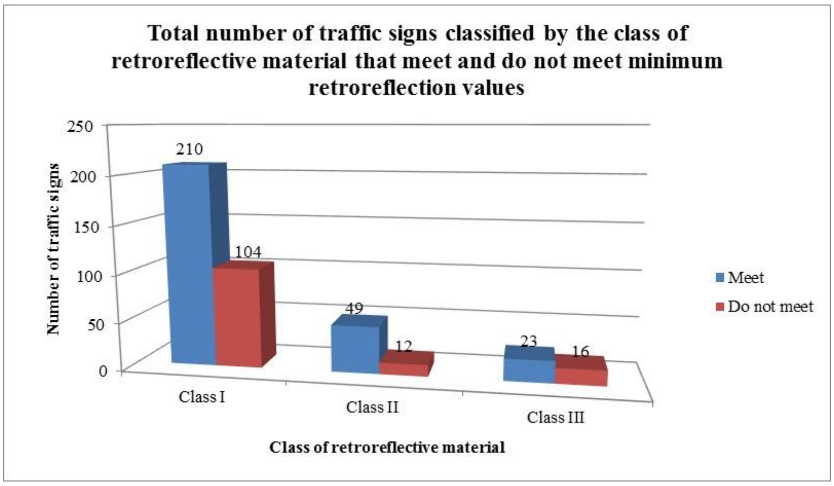



**Figure 5.**
*Total number of traffic signs classified by the class of retroreflective material that meet and do not meet minimum retroreflection values*

From a technical point of view, the analysis showed that 127 signs do not meet the technical specifications (signs are damaged, bent, blocked, not in accordance with the Regulations), of which 24 are danger signs, 16 mandatory signs, 76 information signs, 5 guide signs, 5 additional panels and 1 sign that falls into the category of traffic equipment. As the state road D3 has total of 414 traffic signs, from all the above it can be concluded that 30.6% of traffic signs do not meet the technical requirements.

From all the above, and based on the analysis of technical and retroreflection requirements of traffic signs it can be concluded that on the state road D3 in Zagrebačka County, of 414 signs, there are 213 or 51.4% which do not meet prescribed requirements (some signs do not meet prescribed technical and retroreflection requirements at the same time).

### 3.2. Traffic accidents analysis on the state road D3 in Zagrebačka County

According to data received from the Police department in the period from 2010 to 2013, on the state road D3 in Zagrebačka County 205 accidents were reported, of which 4 with fatalities. Of 205 accidents, 76 were including injuries and 125 material damage. Of 76 injuries, four ended with fatalities, 11 with serious injuries and 118 with minor injuries (Table 5)

**Table 5.**
*The number and consequences of traffic accidents on the state road D3 in Zagrebačka County*

| Number of traffic accidents | | The consequences of accidents | |
|---|---|---|---|
| 2010-2013 (data collected up to October 2013) | | | |
| With fatalities | 4 | Fatalities | 4 |
| With injuries | 76 | Serious injuries | 11 |
| With material damage | 125 | Minor injuries | 118 |
| **Total** | **205** | **Total** | **133** |

*Source: Adapted by authors*

From a total of 205 accidents, 118 of them happened during the day, 79 during the night, six during sunrise and two during dusk, as shown in Table 6. From the above it is evident that 38.5% of traffic accidents on the analyzed road happened during the night. If one adds to this number accidents that occurred at the time of dusk and dawn (3.9%), when visibility is reduced, then it can be concluded that the total of 42.4% traffic accidents happened in conditions with reduced visibility.

**Table 6.**
*Number of traffic accidents on the state road D3 classified by the time of day when they happened*

| Number of traffic accidents on the state road D3 classified by | |
|---|---|
| Day | 118 |
| Night | 79 |
| Dawn | 6 |
| Dusk | 2 |
| **Total** | **205** |

*Source: Adapted by authors*

## 4. Conclusion

Losses resulting from traffic accidents (fatalities, injuries, material and environmental damage) present a significant problem from social, economic and health point of view. The amount of these losses is estimated as 0.35% up to 4.0% of GDP, and even more in undeveloped countries. The overall damage from accidents includes not only medical costs, property damage (directly on the vehicle and indirectly by reducing the value of the vehicle), the loss due to absence from work, but also the damage that cannot be directly expressed: reduction of economic activity, temporary or permanent inability of earning, loss of work ability, inability to perform daily activities, the direct medical costs of reproduction or vocational rehabilitation etc. According to the World Health Organization, the costs for one person killed in a car accident has been estimated at one million euros while other hardly assessable costs of traffic accidents in the EU grow to around 166 billion euros per year[5].

In the period from 2003 to 2012 there were 574 000 traffic accidents on Croatian roads. In these accidents 223 000 people were injured: 5 590 people were killed, 40 000 seriously injured and 177 000 people suffered minor injuries. In 2012 there were 4 915 traffic accidents on state roads, which represented 13.3% of total traffic accidents on all the roads in the Republic of Croatia that year. Out of these 4 915 traffic accidents that happened on state roads, 108 accidents had fatal consequences, while in 1 893 accidents people were injured[6].

As seen from the analysis, in the period from 2010 to 2013 on state road D3 (Zagrebačka County) 205 accidents happened, of which there were four fatalities, 11 serious injuries, 118 minor injuries and 125 accidents resulting in material damage. In addition, the analysis showed that 42.4% of accidents, which is almost half of all the traffic accidents on the analyzed road, had happened under the conditions of reduced visibility.

Since in conditions of reduced visibility the amount of visual information that a driver receives is significantly reduced due to the limited capability to use peripheral vision, drivers' decision-making process is affected. In addition to reduced peripheral vision, the possibility of depth perception and color vision are also reduced in conditions of low visibility.

Further to all the above, the analysis of traffic signs on state road D3 showed that 314 or 75.8% signs are made of class I which has the weakest retroreflective properties. On the other hand, signs made of class II or III share less than 30% in the total number of traffic signs. Also, 31.8% of traffic signs does not meet the minimum retroreflective values and 30.6% of traffic signs does not meet the technical requirements (signs are damaged, bent, blocked, not in accordance with the Regulations, etc.).

[5]http://www.sigurno-voziti.net/razmisljanja/razmisljanja09.html (06.09.2014.)
[6]Ministry of Interior Croatian: Bulletin of Road Safety 2012., Zagreb 2013., ISSN 1331-2863 ()

Based on the analysis it can be concluded that of 414 traffic signs, 213 or 51.4% do not meet the minimum prescribed conditions. From the above it is evident that more than half of the traffic signs on the analyzed road is not technically correct, and as such they are not able to perform their function of conveying information to drivers.

As the amount of information that the driver receives from the environment, in conditions of reduced visibility, is relatively small, technical accuracy of traffic signs (retroreflection values, proper installation of the sign, compliance of the sign with the Regulations, etc.) is one of the key elements to road safety.

As stated earlier, nearly half of traffic accidents (42.4%) on state road D3 occurred in conditions of reduced visibility (night, dusk and dawn). One way of preventive action for the reduction of the number of accidents on road D3 should be a replacement of traffic signs which are not technically correct with new and better signs made of class II or III material. The necessity of replacing signs is also evident in the fact that the average age of the analyzed signs on the state road D3 is eight years.

Installing new signs with better retroreflective properties (class II or III) would significantly increase their visibility in conditions of low visibility, and thus should reduce the number of road accidents and increase general safety on the road. As class II and III materials have a guarantee for 10 years, replacement of incorrect signs with new and better would ensure relatively long-term retroreflection quality.

**References**

Austin, L. R., Schultz, J. R.: Guide to retroreflection safety principles and retroreflective measurements, RoadVista, 2009., ISBN 0-9710215-0-3

Consiglionazionaledellericerche: Common understanding of assessment procedure (CUAP): Microprismatic retro-reflective sheetings, 2002.

EN 12899-1: Fixed, vertical road traffic signs - Part 1: Permanent signs

Ministry of Interior Croatian: Bulletin of Road Safety 2012., Zagreb 2013., ISSN 1331-2863

Ordinance on traffic signs, signals and equipment on roads (OG 33/05, 64/05, 155/05, 14/11)

Stanić, B., Zdravković, P.S., Vukanović, S., Milosavljević, S.: *Elementi saobraćajnog projektovanja – vertikalna signalizacija*, Univerzitet u Beogradu, Saobraćajni fakultet, 2003.

Ščukanec, A., Babić, D., Sokol, H.: Methodology for measuring traffic signs retroreflection, In Proceeding of the 1st International Scientific Forum, Tirana 2013., ISBN: 978-608-4642-15-2

Zuber, N. 2009. Available from the Internet: http://www.sigurno-voziti.net/razmisljanja/razmisljanja09.html (06.09.2014.)

# PRAGMATIC ASPECT OF MARITIME ENGLISH LANGUAGE

**Sanela Kovačević Pejaković**[1]
[1]*Faculty of Maritime Studies, Kotor*

**Abstract:** Maritime English is a specific register. Therefore, it requires a specific treatment. This paper implies the restricted aspect of a complex system such as English for specific purpose in the nautical education and profession. The sophisticated system of the maritime English makes it a plain instrument of communication. We have tried to look into certain subregisters within Standard Marine Communication Phrases (SMCP) by pointing to the main characteristics. Seaspeak is a vital instrument of communication on board a ship. The importance of English language has always been emphasized, especially the importance of the standardized maritime phrases used in vessel-to-vessel vessel-to-port communication.

**Keywords**: li*nguistic, restricted language, maritime terms, pragmatics, seaspeak, specific purpose, ship, sail, rudder, derrick, VHF-equipment*.

## 1. Introduction

The Standard Marine Communication Phrases (SMCP) is a set of key phrases in the English language (which is the internationally recognized language of the sea), supported by the international community for use at sea and developed by the IMO. They aim to explain:1) external communication phrases – ship to ship & ship to shore communication; 2) on-board communication phrases – communication within the ship. The SMCP were adopted by the 22nd Assembly of the IMO in November 2001 in a resolution which also promoted the wide circulation of the SMCP to all prospective users and all maritime education authorities. The SMCP includes phrases which have been developed to cover the most important safety-related fields of verbal shore-to-ship (and vice-versa), ship-to-ship and on-board communications. The aim is to reduce the problem of language barriers at sea and avoid misunderstandings which can cause accidents.

As navigational and safety communications from ship to shore and vice versa, ship to ship , and on board ships must be precise, simple and unambiguous, so as to avoid confusion and error, there is a need to standardize the language used. This is of particular importance in the light of the increasing number of internationally trading vessels with crews speaking many different languages since problems of communication may cause misunderstandings leading to dangers to the vessel,the people on board and the environment. In 1973 the IMO Maritime Safety Committee agreed at its twenty-seventh session that, where language difficulties arise, a common language should be used for navigational purposes and that language should be English. In consequence the SWNV was developed, adopted in 1977 and amended in 1985. In 1992 the IMO Maritime Safety Committee at its sixtieth session instructed the IMO Sub-Committee on Safety of Navigation to develop a more comprehensive standardized safety language than the SMNV, 1985, taking into account the changing conditions in modern seafaring and covering all major safety-related verbal communications. At its sixty-eighth session in 1997 the IMO Maritime Safety Committee adopted the Draft Standard Marine Communication Phrases (SMCP) developed by the IMO Sub Committee on Safety of Navigation. The Draft SMCP, following international trials, was amended at the forty-sixth session of this Sub-Committee and final consideration given at the IMO Maritime Safety Committee. Under the International Convention on Standards of Training, Certification and Watch-keeping for Seafarers, 1978, as revised 1995, the ability to understand and use the SMCP is required for the certification of officers in charge of a navigational watch on ships of 500 gross tonnage or more.

## 2. Position of the SMCP in maritime practice

The Standard Marine Communication Phrases (SMCP) has been compiled:
- to assist in the greater safety of navigation and of the conduct of the ship
- to standardize the language used in communication for navigation at sea, in port-approaches, in and on board vessels with multilingual crews, and
- to assist maritime training institutions in meeting the objectives mentioned above.

These phrases are not intended to supplant or contradict the International Regulations for Preventing Collisions at Sea, 1972 or special local rules or recommendations made by IMO concerning ships' routeing, neither are they intended to supersede the International Code of Signals and when applied in ship's external communication this has to be done in strict compliance with the relevant radiotelephone procedures as set out in the ITU Radio Regulations. Furthermore, the SMCP, as a collection of individual phrases, should not be regarded as any kind of technical manual providing operational instructions.
The SMCP meets the requirements of the STCW Convention, 1978, as revised, and of the SOLAS Convention, 1974, as revised, regarding verbal communications; moreover, the phrases cover the relevant communication safety aspects laid down in these Conventions.

1Corresponding author: sanelap@ac.me

Use of the SMCP should be made as often as possible in preference to other wording of similar meaning; as a minimum requirement users should adhere as closely as possible to their wording in relevant situations. In this way they are intended to become an acceptable safety language, using English for the verbal interchange of intelligence among individuals of all maritime nations on the many and varied occasions when precise meanings and translations are in doubt, increasingly evident under modern conditions at sea. The accompanying CD/Cassette is designed to familiarize users with the pronunciation of the phrases.

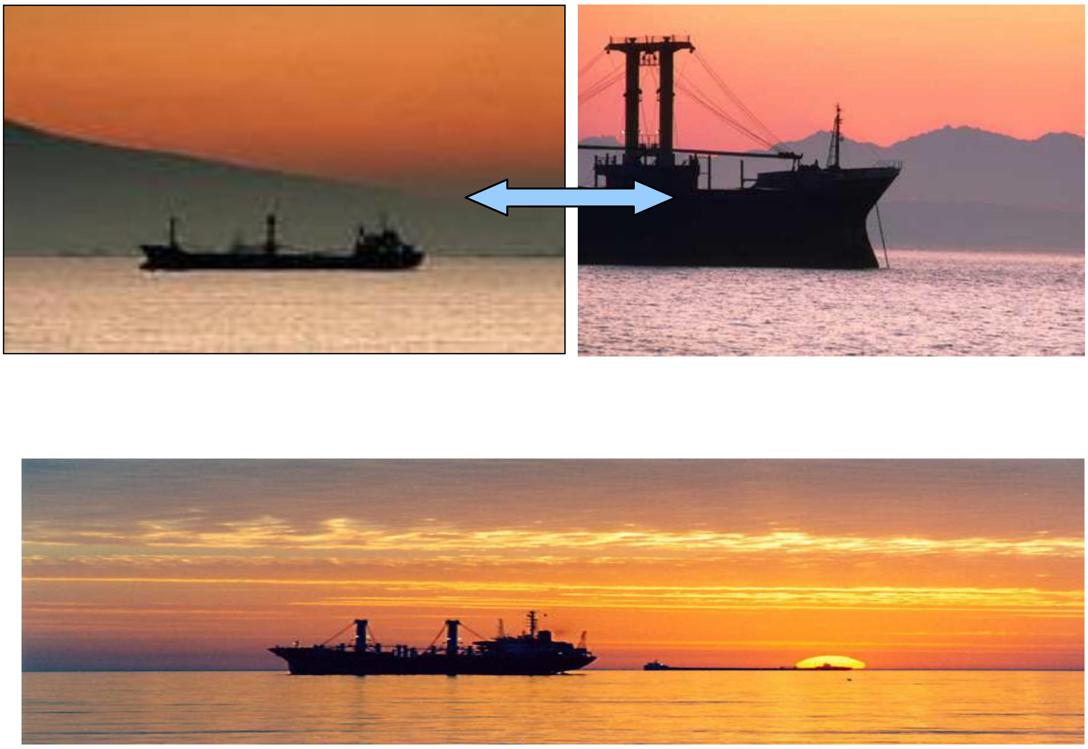


**Fig. 1.**
*Ships communicate using standardized maritime language in a confined context. However, when we compare two persons to ships that pass in the night, we imply that they can meet once and never again.*

**2.1.Organization of the SMCP**

The SMCP is divided into External Communication Phrases and On-board Communication Phrases as far as its application is concerned, and into PART A and PART B as to its status within the framework of the STCW, 1978, as revised. PART A covers phrases applicable in external communications and which may thus be regarded as the replacement of the Standard Marine Navigational Vocabulary, 1985, which is requested to be used and understood by the STCW Code, 1995, Table A-II/I. This part was enriched by essential phrases concerning ship handling and safety of navigation to be used in on-board communications, particularly when the Pilot is on the bridge, as required by Regulation 14(4), Chapter V, SOLAS 1974,as revised.

PART B calls attention to other on-board standard safety-related phrases which, supplementary to PART A, may also be regarded useful for Maritime English instruction.

**2.1.1. Position of the SMCP in Maritime Education and Training**

The SMCP does not intend to provide a comprehensive Maritime English syllabus which is expected to cover a far wider range of language skills to be achieved in the fields of vocabulary, grammar, discourse abilities, etc., than the SMCP could ever manage… However, PART A in particular, should be an indispensable part of any curriculum which is designed to meet the corresponding requirements of the STCW Convention, 1978, as revised. In addition, PART B offers a rich choice of situations covered by phrases well suited to meet the communication requirements of the STCW Convention, 1978, as revised, which are implicitly expected to be satisfied by mariners.

The SMCP should be taught and learnt selectively, according to the users` specific needs rather than completely. The respective instruction should be based on practice in the maritime environment and be implemented through appropriate modern language teaching methods.

The SMCP builds on a basic knowledge of the English language. It was drafted on purpose in a simplified version of Maritime English to reduce grammatical, lexical and idiomatic varieties to a tolerable minimum, using standardized structures for the sake of its function aspects, i.e. diminishing misunderstanding in safety related verbal communications, thereby endeavouring to reflect present Maritime English language usage on board vessels and in ship-to-shore/ship-to-ship communications.

This means, in phrases offered for use in emergency and other situations developing under considerable pressure of time or psychological stress as wells as in navigational warnings, a block language was applied which sparingly uses, or frequently omits, the function words the, a/an, is/are as done in seafaring practice. Users, however, may be flexible in this respect. Further communicative features may be summarized as follows:

- avoiding synonyms
- avoiding contracted forms
- providing fully worded answers to "yes/no"-questions and basic alternative answers to sentence questions providing one phrase for one event, and
- structuring the corresponding phrases after the principle: identical invariable plus variable.

Ship to ship and ship to shore communication is regulated speech in action. It is organized within a confined maritime affairs.

### 2.1.2. Typographical conventions

**( )** brackets indicate that the part of the message enclosed within the brackets may be added where relevant; text, text,
**...** dots indicate that the relevant information is to be filled in where the dots occur;
**~** tides stand for the invariable part of an aforementioned standard phrase which is followed by a variable addendum

## 3. Linguistic restrictedness and accuracy in safety-related communication

When it is necessary to indicate that the SMCP are to be used, the following message may be sent :

- "Please use Standard Marine Communication Phrases."
- "I will use Standard Marine Communication Phrases".

### 3.1.Spelling of letters

When spelling is necessary, only the following spelling table should be used:

| Letter | Code | Letter | Code |
|---|---|---|---|
| **A** | Alfa | **N** | November |
| **B** | Bravo | **O** | Oscar |
| **C** | Charlie | **P** | Papa |
| **D** | Delta | **Q** | Quebec |
| **E** | Echo | **R** | Romeo |
| **F** | Foxtrot | **S** | Sierra |
| **G** | Golf | **T** | Tango |
| **H** | Hotel | **U** | Uniform |
| **I** | India | **V** | Victor |
| **J** | Juliet | **W** | Whisky |
| **K** | Kilo | **X** | X-ray |
| **L** | Lima | Y | Yankee |
| **M** | Mike | **Z** | Zulu |

### 3.2. Responses

When the answer to a question is in the affirmative, say:"Yes, .... " - followed by the appropriate phrase in full.
When the answer to a question is in the negative, say:"No, ..." - followed by the appropriate phrase in full.
When the information requested is not immediately available, say: "Stand by" - followed by the time interval within which the information will be available.
When an INSTRUCTION (e.g. by a VTS-Station, Naval vessel or other fully authorized personnel ) or an ADVICE is given, respond if in the affirmative:
"I will/can ... " - followed by the instruction or advice in full; and, if in the negative, respond:"I will not/cannot ... " - followed by the instruction or advice in full. Example: "ADVICE. Do not overtake the vessel North of you."Respond: "I will not overtake the vessel North of me."Responses to orders and answers to questions of special importance both in external and onboard communication are given in wording in the phrases concerned.

### 3.3. Distress, urgency and safety signals

MAYDAY to be used to announce a distress message
PAN - PAN to be used to announce an urgency message
SÈCURITÈ to be used to announce a safety message

### 3.4. Standard organizational phrases

The phrase "How do you read (me)?" is followed by:
"I read you ...bad/one with signal strength one (i.e. barely perceptible) poor/two with signal strength two (i.e. weak)
fair/three with signal strength three (i.e. fairly good)
good/four with signal strength four (i.e. good)
excellent/five with signal strength five (i.e. very good)
When it is advisable to remain on a VHF Channel / frequency say:"Stand by on VHF Channel ... / frequency ... "
When it is accepted to remain on the VHF channel / frequency indicated, say: "Standing by on VHF Channel ... / frequency ... "
When it is advisable to change to another VHF Channel / frequency, say:"Advise (you) change to VHF Channel ... / frequency ... ." "Advise(you) try VHF Channel .. / frequency... ."
When the changing of a VHF Channel / frequency is accepted, say: "Changing to VHF Channel ... / frequency ... ."
When the information requested cannot be obtained, say:"No information."

### 3.5. Corrections and repetition

When a mistake is made in a message, say: "Mistake ..." - followed by the word: "Correction ... " plus the corrected part of the message Example: "My present speed is 14 knots – mistake. Correction, my present speed is 12, one-two, knots
If any part of the message are considered sufficiently important to need safeguarding, say:"Repeat ... " - followed by the corresponding part of the message. Example: "My draft is 12.6 repeat one-two decimal 6 metres.""Do not overtake - repeat - do not overtake."

## 4. Numbers in Standardized Marine Communicative phrases

A few digits and numbers have a modified pronunciation compared to general English: All numbers except whole thousands should be transmitted by pronouncing each digit separately: "Wun-fife-zero" for 150

| number | spelling | pronunciation |
|---|---|---|
| **0** | zero | ZEERO |
| **1** | one | WUN |
| **2** | two | TOO |
| **3** | three | TREE |
| **4** | four | FOWER |
| **5** | five | FIFE |
| **6** | six | SIX |
| **7** | seven | SEVEN |
| **8** | eight | AIT |
| **9** | nine | NINER |
| **1000** | thousand | TOUSAND |

All numbers except whole thousands should be transmitted by pronouncing each digit separately.
A conversation in Seaspeak follows to illustrate the above mentioned communicative features.

Western Sky (WS) is approaching Singapore (SPO).
WS: Singapore Port Operations. This is Western Sky. Information: My ETA position: East Johore pilot station is time: one-three-four-five UTC. Over.
SPO: Western Sky. This is Singapore Port Operations. Mistake. Time is : One-four-three-zero UTC now. Stay on. Over.
WS: Singapore Port Operations. This is Western Sky. Correction. My ETA is one-five-four-five UTC. Over.
SPO: Western Sky. This is Singapore Port Operations. Information-received: Your ETA position: East Johore pilot station is time : one-five-four-five UTC. Instruction: anchor in General Purpose Anchorage, reason: your berth is occupied. Over.
WS: Singapore Port operations. This is Western Sky. Instruction-received: anchor in the General Purpose Anchorage. Nothing more. Over.
SPO: Western Sky. This is Singapore Port Operations. Out

**Table 1.**
*The most frequent verbs and collocations in the Standard Marine Communicative Phrases*

| Glagol | Broj linija konkordance | Karakteristične kolokacije |
|---|---|---|
| *report* | 175 | *Check the lifebuoys and report; Report the distance; Report the distress position; Report seaworthiness of the holds.* |
| *have* | 110 | *You have the watch now; We have minor damage; We have grounded.* |
| *do* | 66 | *Do you have a bow thruster?;Do you need help?;Do you carry any dangerous goods?; Do you need assistance?* |
| *check* | 51 | *Check the electrical lighting; Check the watertight door; Check the operation of the hoses.* |
| *require* | 44 | *MV requires tug; Do you require assistance?; We require tug service; We require a strecther; I do not require a pilot.* |
| *must* | 39 | *We must restow cargo; We must transfer fuel / ballast; Yes, you must take a pilot - pilotage; Convoy ... must wait / moor at; All passengers must attend this drill.* |
| *can* | 38 | *Can you make rendezvous in position? Yes, I can proceed to distress position; Can you pick up survivors?Yes, I can pick up survivors.* |
| *go* | 36 | *Let go port /starboard /both anchors; Let go the tug(s); Stand by for letting go towing line; Let go lifeboat(s); Go through engine room; Go to your lifeboat stations.* |
| *keep* | 30 | *Keep clear of this area;Keep calm;Keep the wind on the port side; Keep your eyes on the person.* |
| *inform* | 22 | *Inform pollution control; Inform on radio; Inform coast guards;Iinform the Master/Chief engineer;* |
| *make* | 19 | *Make fast tug; Make fast forward and aft; Make a lee; Course to make good; Make a boarding speed; Make announcement.* |
| *give* | 11 | *Give "Vessel aground" signal; Give assistance when entering port; We need to give way; Give notice of readiness.* |
| *get* | 10 | *We are not ready to get underway; Are you underway? We will be ready to get underway in ... minutes.* |
| *may* | 8 | *During the voyage you may hear some other sound signal; You may leave your children under; You / MV ... may stop search.* |
| *need* | 4 | *Assist those who need help; No, you need not take a pilot; No, you need not take tug(s);We need not give way.* |
| *advise* | 3 | *Give sound signals; Give way to the vessel abeam of you; Give assistance; Give notice of readiness.* |

## 5. Conclusion

We can conclude that maritime English is a *restricted language* which is characterized by a great many specific features on the phonological, morphological, lexical and syntaxical level. These particularities are the most obvious on the lexical level, for it implies maritime terms. also under the greatest influence of the real world, that is the world outside language sphere.
Since the maritime language is further subdivided into registers and subregisters, we can make a distinction between maritime language and maritime languages. The complexity of the nautical register makes learning, that is acquisition very difficult. One should firstly learn maritime terms in his own mother tongue and then move into the maritime English world.

**References**

Lakić, I., 1997. “ ESP at the end of the century: Genre analysis”, symposium, Language and Literature at the end of XX Century, Collection of Papers, Podgorica.

Kluijven, P.C. Van, 2003. *The International Maritime Language Programme. An English Course for students at Maritime Colleges and for on-board training*, Alk & Heijnen Publishers, AlkmaarThe Netherlands.

Thompson, S. .1994. “Frameworks and contexts: a genre-based approach to analysing lecture introductions”, English for Specific Purposes, Vol.13.

Thornbury, S., 2002. *How to Teach Vocabulary*, Essex, Pearson Education Limited

Widdowson, H.G., 1998. "Communication and Community the Pragmatics of ESP", English for Specific Purposes, Vol. 17, No 1.

# RAILROAD GEOMETRIC DEFECTS: A METHODOLOGY TO DEVELOP DEGRADATION MODELS FROM EXPERIMENTAL DATA

**Mauro D'Apuzzo[1], Bruna Festa[2], Elisabetta Miraglia[3]**
[1]*University of Cassino and Southern Lazio*
[2]*University of Naples "Federico II"*
[3]*University of Naples "Federico II"*

**Abstract:** In recent years, almost all railway management companies are equipping themselves with more complex high performance devices in order to monitor track condition. Periodic control of railway characteristics is strongly advised to prevent and reduce transportation risk. Wear defects, as well as the erroneous design of railway structures, are among the main causes of derailment. Various approaches have been considered to analyze, manage and reduce risk.
Local transportation companies should have available and reliable predictive degradation models, based on empirical and mechanistic considerations obtained from experimentally collected rail defect data. These would be useful to predict the evolution of track conditions to be used in decision support systems for Railway Track Maintenance. In this paper an analysis approach following a large experimental measurement campaign carried out on a suburban narrow-gauge railway network operating in the Naples metropolitan area is proposed. The data have been obtained by the Plasser device (POS/TG System), directly mounted on a vehicle Fe ETR 220's wheelset during its regular operation. Track cant, level, gauge and twist have been collected almost every month in a year since 2007 to 2013 with a 25 cm sampling interval on track sections up to 40 Km long. Preliminary results seem to indicate that by coupling track geometric data analyzed so far with alignment and traffic, a more effective evaluation of track defect can be obtained.

**Keywords**: railway superstructure maintenance, rail defects, automatic detection algorithms.

## 1. Introduction

Railways infrastructures have been interested in a continuous increase of traffic load and length, weight and speed of trains. This corresponds to an increase in stress and in rate of evolution of rolling stock and infrastructures degradation phenomena.
Monitoring this kind of degradation phenomena is necessary to plan maintenance operations aimed at restoring original railway characteristics at minimum cost to ensure an adequate safety level and passenger comfort.
To assure the highest levels in the security of these systems, rail infrastructure managers increase inspective activities, which leads to increased costs and decreased availability of the transport systems.
In last decades automatic diagnostic systems installed on specialized rolling stock, were introduced in the European market. These solutions require high initial investments and human resource to implement and to analyze output data, especially for local Railway company. It can be necessary to have monitoring solutions to achieve high levels of track geometric standards for safe and efficient train operations while minimizing long-term track maintenance costs and extending track component service lives. In this connection, mathematical models able to forecast rail defect evolution may play a critical role in the railway maintenance process.
This paper is organized as follows. In next sub section the literature background on rail defects monitoring and maintenance will be briefly described . In particular, the new devices used to collect the required track geometrical data are described. Then the experimental analysis carried out with data collected, and preliminary results provided by the statistical methodologies used to derive stochastic degradation models are presented.

## 2. Background

The measurement and improvement of track quality are key issues in determining both the restoration time and cost of railway maintenance in order to guarantee minimum safety level against derailment, on one hand, and to mitigate environmental impact as far as traffic induced vibrations are concerned (D'Apuzzo et al., 2008) (D'Apuzzo et al. 2012), on the other. According to the Federal Railway Administration database most of train accident are caused by Track geometry defect and wide gauge (Liu X. et al. 2011), so it is necessary to keep under control this parameter and their evolution during the time in absence of maintenance operations.
However, executing early maintenance works can save the railway track owner from a more costly failure, so it is important to use predictive methods to plan maintenance and on the other hand to avoid service interruptions that represents other additional costs (Wireman, T. 2008). There is a growing interest in the technical literature on railway maintenance planning and operations. . Some studies are related to search the optimization equilibrium between costs and performance of different kinds of maintenance works (Grimes, et al. 2006). In (Lovett et al. 2013) an interesting framework for the planning and scheduling of track maintenance works composed by a track degradation module a maintenance activity selection and evaluation module and a schedule optimization module is presented, whereas in (Higgins et al., 1998) (Peng et al., 2011) heuristic approaches are introduced to look for the optimum railway track maintenance schedules. In (Macchi et al., 2012) a reliability based maintenance programming framework that is going to be implemented for the Italian Railway Administration is describe whereas Zio et al. in 2007 proposed a an approach for improving the service level of the railway network, while maintaining high standards of safety (Zio et al. 2007).

As far as the degradation of a railway superstructure is concerned, according to International Union of Railways (UIC 2002) railway defect can be classified into different classes : track defects; rail defects; structural defects; and geometric defect. Ripage and twist are track defects and they are two most important problem that cause derailment. The first one exists when all track expands and or moving on support surface (Panagin, 1990). Twist is the algebraic difference between two cross levels measured at a predefined distance, usually it is expressed as inclination in percentage and it can be measured directly with a step equal to the axis of the vehicle, or the proceeds from the measures of cross level (UNI-EN 13848-1). Rail defect can be classified in micro and macro defect or superficial imperfections (eg breaking, corrosion, cracking buckling, squats, weld, corrugation). This kind of defect are identified from UIC with a 4 digit code to identify the longitudinal and lateral position of defect, the shape and additional characteristics and differentiations.
Structural defect interested each rail track components, like superstructure, ballast, tie, rail, switches, crossing and other civil marginal works. Latest geometric defects (eg gauge, cant) became defects when they was poorly designed or degraded during the time.
As far as rail geometrical defects are concerned, International Union of Railways (ORE; 1988) propose to evaluate track degradation indexes in terms of standard deviation of the longitudinal level, cross level and alignment in track segments of 200 m of length. The standard deviation is also one of the performance indicators specified by the European Standard too (UNI EN13848). For this geometric parameter and for Gauge, Twist and Cant Current European standards and Italian Railway Administration can be reffered to (RFI 2004) where threshold limit values at which the track is defined deteriorated are reported. It has to be highlighted that gauge threshold values, proposed by UNI EN 13848-5 have been evaluated with reference to the standard gauge value (1435 mm) ;. for narrow-gauge railway network (gauge value: 950 mm) threshold values must be conveniently converted (D'Apuzzo, et al. 2014).).
In this paper an analysis of geometrical rail defects (gauge, twist and cant) collected by a high performance devices was carried out in order to derive reliable mathematical prediction tools aimed at improve maintenance intervention programs.

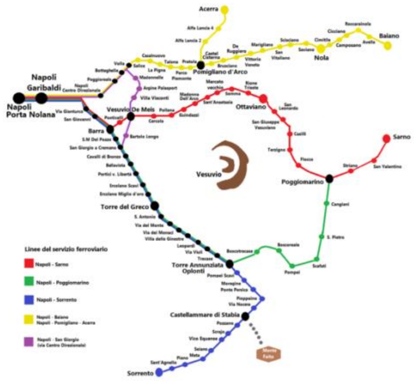
## 3. Data Collection

Nowadays different measurement equipments and methods are used by Italian railway companies to monitor track quality in terms of rail and track defects within maintenance programs. However static or low-speed measurement devices such as profiling trolleys have been gradually abandoned in favour of high-speed and high performance measurement equipments. These latter can be mainly split into two kind: instrumented conventional rolling stocks and diagnostic rolling stocks. In the former case, measuring devices are mounted on a normally operating railway cars whereas in the latter specific rolling stocks have been designed to host several measurement equipments and are solely used for the monitoring of the quality of the railway track Circumvesuviana, now embedded in Ente Autonomo Volturno (EAV), is a local narrow-gauge railway company operating in the Naples' area whose rail network is extended around Vesuvius for an overall amount of 142 km distributed among 6 lines and 96 stations (see Fig. 1a). The measurement system used by EAV is essentially composed of an inertial platform, an odometer, a differential Global Positioning System (GPS), and two pairs of optical detectors mounted close to the axles of the wheel sets of a normally operating rolling stock (see fig. 1b). The back optical detectors verify the accuracy of the measurements of the front ones and are an element of alert for any sampling errors. Data are usually collected in terms of: gauge, twist, cant, curvature, alignment, longitudinal level, latitude, longitude and altitude, GPS coordinates and train speed with 25 cm sampling lenght. The output are graphical and numerical. Data are collected by EAV maintenance staff on a routine basis and geometrical defects so far evaluated are checked against threshold level in order to detect critical points that need to be further investigated. However no in-depth analysis seem to be undertaken in order to evaluate the evolution of the geometrical track defects with time.
In a previous paper (D'Apuzzo et al. 2014) a preliminary analysis on cant twist a gauge data has been presented over a three year long period (2009-2011) for a specific Circumvesuviana railway line. Additional data have been recently integrated and a more extended analysis has been carried out. Results so far obtained are reported in the followings.

## 4 Data Analysis and discussion of results

The advantage of high performance measuring devices for track data collection consist of their quick set up and the possibility to collect a large number of data. However it has to be underlined that the huge amount of data collected is the one of the most critical issue for the analysis of the evolution of the geometrical track properties. If a 25 cm long sample length is assumed, a single sweep of the measuring device on the overall extension of the railway network examined can yield to more than one million data values for each of track defects previously reported. Additional issues are related to the accurate identification of invalid readings and to the calibration of distance sensors (odometer) in order to correctly locate the track defects among the different railway lines. Preliminary efforts have been therefore devoted to the analysis of a specific railway line, (namely the Naples-S.Giorgio line) where more accurate and complementary information on alignment are available. The analysis of the other Circumvesuviana railway lines is being undertaken in order to eventually confirm results so far obtained. The Naples-S.Giorgio railway line is about 12 Km long and comprises 11 stations. Railway track is a traditional narrow gauge (design value: 900 mm) ballast type railway superstructure apart from underground sections where a concrete slab ballastless track is employed.

Since the aim of the study was to evaluate the evolution of geometrical track defects with time and the degradation of railway superstructure is primarily affected by the type of railway superstructure (Esveld 2001), data collected on ballastless sections were conveniently expunged from the overall analysis.



(a)

(b)

**Fig. 1.**
*Circumvesuviana rail network layout (a), rail quality measuring device employed (b).*

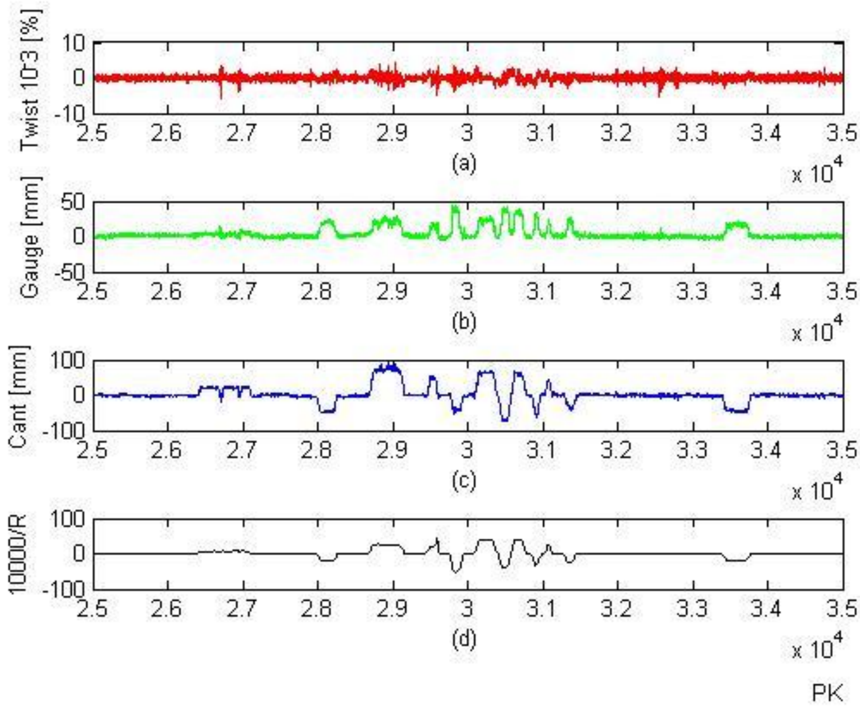



**Fig. 2.**
*Typical pattern of a geometric rail properties as a function of progressive kilometric in a part of the track.*

Furthermore it has to be acknowledged that wheel-rail dynamic interaction forces are the main responsible for track degradation, therefore it seems reasonable to separately evaluate the evolution of degradation for each alignment section (i.e. tangents, transition curves, and circular curves). Basing on these premises, track geometrical defects have been individually analyzed on each specific alignment segment.
It was therefore necessary to apply a sectioning procedure in order to analyze data for specific homogeneous sections. Curvature values were derived by post-processing GPS data. Data were subsequently grouped according to radius ranges for curves and length ranges for tangents detailed in the following Table (Table 1):
Synthetic statistical descriptors, such as mean and standard deviations have been conveniently derived. It has emerged that regardless of date of survey, for a specific geometrical track defect, Coefficient Of Variation (COV) i.e. the ratio between the standard deviation and the mean value, are very low (see fig. 2). Therefore mean values of track geometrical properties have been evaluated with time. Main results are detailed in the followings.

**Table 1**
*Different groups of curves and radius analyzed.*

| | CURVES - R (m) | STRAIGHTS - L (m) |
|---|---|---|
| Group 1 | 0 ÷300 | 0 ÷100 |
| Group 2 | 300 ÷600 | 100 ÷200 |
| Group 3 | 600 ÷900 | 200 |
| Group 4 | > 900 m | - |

### 4.1 Gauge data,

As it is known, conventional gauge has to be conveniently increased in bends in order to accommodate the curve negotiation for the wheelsets. According most of railway design standards, gauge design increment value is evaluated basing on the circular curves geometrical properties and therefore each bend is characterized by a specific augmented gauge value. To develop a more general model for the stud of the evolution of this track geometric property, gauge value recorded by the devices on a specific circular curve has been therefore divided by gauge value measured in the first survey day in order to derive a normalized gauge value on which a variation with time has been evaluated (see Figure 3). As it can be observed by figure, a cyclic decrease of the normalized gauge increment can be detected on annual basis. It has been assumed that this behavior may be affected by climatic factors. A history of temperature data were therefore conveniently retrieved by national database in Naples' district in order to make a comparison with normalized gauge increments. As it can be highlighted, in the summer period a fairly decrease of the normalized gauge increment can be identified that may be ascribed to thermal expansion effects involving the rails. Therefore apart from seasonal variation, no clear increasing or decreasing trend of gauge with time can be detected among the period investigated (from October 2009 to September 2013).

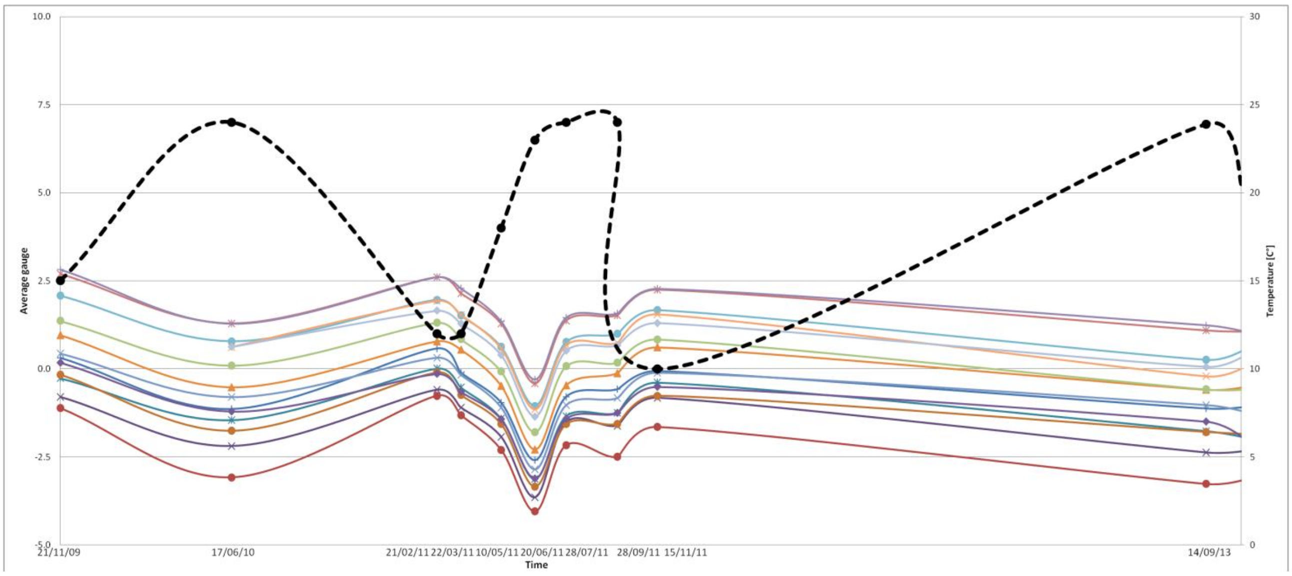



**Fig.3**
*Normalized average gauge increment values for curves vs temperature data.*

### 4.2 Cant and Twist data

According to the same approach employed in the analysis of gauge data, standard deviations and mean values have been evaluated on circular and transition curves. In detail, twist data have been investigated for spiral curves whereas cant values have been examined for bends. Even for these geometrical track properties, resulting COV values were very low. It has to be reminded that, according to railway alignment standards, design cant values are dependent on circular curve geometric properties; specific twist design values are similarly imposed for transition curves. Therefore in order to obtain more general degradation model normalization is needed.

Cant and twist normalized values obtained by performing the ratio between the mean value estimated at a certain survey day and the value evaluated in the first survey day for each specific transition and circular curve have been derived and analyzed as a function of time. Examined data showed an increasing or decreasing trend of normalized values depending on circular and transition curve properties and therefore it was assumed that this different behavior may be ascribed to the vertical forces exerted by the rail vehicle to the inner and outer rail during normal operation. Therefore the following general empirical degradation model describing the evolution of normalized *increment* of track cant and twist was proposed as a function of independent variables X1 and X2 respectively:

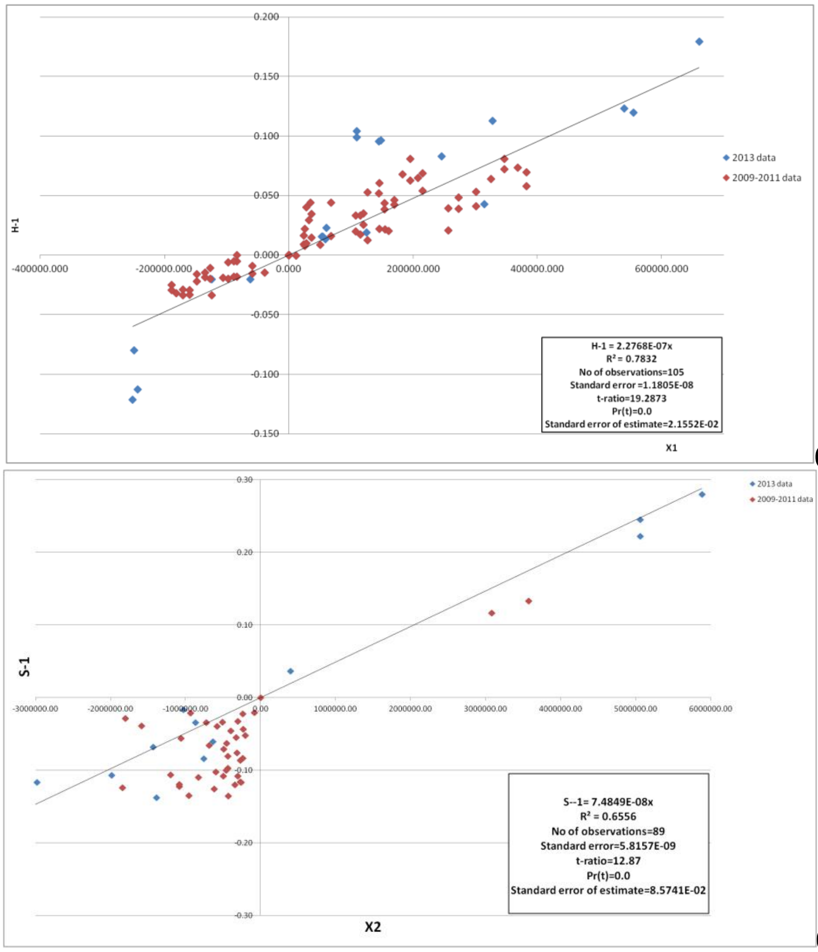
$$H_{cant} - 1 = K_{cant} \cdot X1 = K_{cant} \cdot \frac{V_{op}^2 - V_c^2}{R} \cdot CT \tag{1}$$

$$H_{twist} - 1 = K_{twist} \cdot X2 = K_{twist} \cdot \frac{V_{max}^2 - V_c^2}{L} \cdot CT \tag{2}$$

where:
$H_{cant}$ and $H_{twist}$ are the previously defined normalized cant and twist values respectively;
$K_{cant}$ and $K_{twist}$ are empirically calibrated factors for cant and twist degradation models respectively;
$V_{op}$ is the train operating speed on the specific circular curve examined derived from measuring device ;
$V_{max}$ is the maximum train operating speed on the specific transition curve examined;
*CT* is the *Cumulative Traffic* i.e. the number of train passes on the specific circular or transition curve examined since the initial survey day;
*R* is the radius of curvature for the specific bend;
*L* is the length of specific transition curve examined;
$V_c$ is the train *Critical Velocity* i.e. the velocity value when centrifugal forces are counterbalanced by the track cant and that can be evaluated by means of the following expression:

$$V_c = \sqrt{\frac{H \cdot g \cdot R}{S}} \tag{3}$$

where *H*, *R* and S are the cant, radius and gauge design values for the specific bend examined respectively and *g* is the acceleration of gravity.
In the following figure a linear regressions of data with relevant statistics is presented. As it can be seen, the resulting correlation coefficient appears to be fairly good and higher than that provided by previous analysis (D'Apuzzo et al. 2014) and recent collected data seem to confirm the overall trend previously detected.



(a)
(b)

**Fig.4**
*Calibrated empirical degradation models for normalized cant increment (a) and normalized twist increment (b).*

Markov Chain probabilistic degradation models can been derived from the empirical degradation model so far proposed. Basing on a six month period, cant and twist increment range have been selected in order to define specific degradation levels and corresponding Transition Probability Matrix (TPM) for each circular curve, have been calibrated by solving the following optimization problem (Otiz-Garcia et al., 2006):

$$Z = \min \sum_t [y(t) - y_m(t)]^2 \quad (4)$$

where y(t) denotes the value provided by the empirical degradation model at time t, and ym(t) is the average value estimated by the TPM. A typical TPM for a circular curve with radius equal to 178 meters is also conveniently reported below (Tab. 3).

**Table 2**
*Typical TPM obtained for a circular curve with radius equal to 178 m.*

P

| | | | | |
|---|---|---|---|---|
| 0.973416 | 0.024252 | 0.002332 | 0.000000 | 0.000000 |
| 0.000000 | 0.681825 | 0.253545 | 0.064630 | 0.000000 |
| 0.000000 | 0.000000 | 0.395825 | 0.362602 | 0.241573 |
| 0.000000 | 0.000000 | 0.000000 | 0.121525 | 0.878475 |
| 0.000000 | 0.000000 | 0.000000 | 0.000000 | 1.000000 |

## 5. Conclusion

This paper introduces a novel modeling methodology aimed at supporting maintenance management of railway infrastructures based on a large experimental measurement campaign carried out on a suburban narrow-gauge railway network operating in Naples metropolitan area.
In particular, track cant, gauge and twist have been collected with a 25 cm sampling interval by using an high-performance measuring device (directly mounted on conventional rail vehicles) several times a year since 2009 to 2011 on a specific Circumvesuviana line. In this paper additional analysis following further experimental surveys carried out in 2013 has been presented. The results so far obtained showed a good agreement with previous analysis and seem to confirm that:

- gauge values on bends are mainly affected by climatic factors;
- cant and twist values for circular and transition curve respectively are affected by vertical dynamic forces exerted by the rolling stock and by cumulative traffic.

Basing on these premises, general empirical degradation models for cant a twist have been developed and Markov Chain Stochastic models have been calibrated in order to provide a useful prediction tool for Railway Managers for maintenance purpose. Although further investigations are needed in order to validate the methodology proposed on other railway lines, preliminary results seem to be promising.

## Acknowledgements

The authors are gratefully acknowledged to the EAV maintenance team for providing the track geometrical defect data.

## References

D'Apuzzo M., Aiello V., Boiero D., Socco L.V. & Silvestri F. 2008.Experimental and numerical analysis of vibrations induced by underground trains in urban environment. Structural Control & Health Monitoring, vol. 15, p. 315–348, ISSN: 1545-2255, doi: 10.1002/stc.247.

D'Apuzzo M., Nicolosi V. & Bogazzi E. 2012. A Unified Approach for the Prediction of Vibration Induced by Underground Metro. Procedia: Social & Behavioral Sciences, vol. 53, p. 62–71, ISSN: 1877-0428, doi: 10.1016/j.sbspro.2012.09.860.

D'Apuzzo, M., , Festa B . and Miraglia E .; 2014. Evaluation of rail defects provided by high performance measuring devices Sustainability, Eco-efficiency, and Conservation in Transportation Infrastructure Asset Management Edited by Massimo Losa and Tom Papagiannakis CRC Press pp. 649–Print ISBN: 978-1-138-00147-eBook ISBN: 978-1-315-75712-4 doi: 10.1201/b16730-95.

Esveld, C. 2001. Modern railway Track. MRT Production, Zalt Bommel, The Nerthelands.

Grimes, G. A. and Barkan, C. P. L. 2006. Cost-Effectiveness of Railway Infrastructure Renewal Maintenance, Journal of transportation engineering asce.

Higgins, A. 1998. Scheduling of railway Track Maintenance Activities and Crews. Journal of Operational Research Society, Vol. 49, No 10, 1026-1033.

Liu, X., Barkan et alii. 2011. Analysis of Derailments by Accident Cause: Evaluating Railroad Track Upgrades to Reduce Transportation Risk; Transportation Research Record 2261.

Lovett, A H et al. 2013. Development of an integrated model for the evaluation and planning of railroad track maintenance. In proceedings of the 2013 Joint Rail Conference April 15-18.

Macchi, M. et al. 2012. Maintenance management of railway infrastructures based on reliability analysis, Reliability Engineering and System Safety 104.

Ortiz-Garcia J., Costello S., Snaith M., 2006, Derivation of Transition Probability Matrices for Pavement Deterioration Modelling, Journal of Transportation Engineering, Vol. 132, No. 2, ASCE. doi: 10.1061/(ASCE) 0733-947X(2006)132:2(141)

Panagin, R.. 1990. La dinamica del veicolo ferroviario. Libreria editrice universitaria Levrotto&Bella.
Peng, F.et al 2014. Optimal clustering of railroad track maintenance jobs. Computer-Aided Civil and Infrastructure Engineering Volume 29, Issue 4, 235–247.

RFI 2004, Track geometric quality standards for speed up to 300 Km/h. Part1. Part 2, Part 3. Italian Railway Company (in Italian).

UIC 2002, Handbook of Rail Defects. UIC Code 712R (4th edn), Paris.

UNI EN 13848 Track geometry quality, Part 1: Characterization of track geometry; Part 2: Measuring systems, Track recording vehicles; Part 3: Measuring systems, Track construction and maintenance machines; Part 5: Geometric Quality Levels.

Wireman, T. 2008. Preventative Maintenance, Industrial Press, Inc., New York.

Zio, E. et al. 2007 Importance measures-based prioritization for improving the performance of multi-state systems: application to the railway industry. Reliability Engineering & System Safety; 92:1303–14.

# PRELIMINARY THEORETICAL INVESTIGATION OF INFLUENCE OF TRACK GEOMETRICAL DEFECTS ON WHEEL-RAIL FORCES

**Mauro D'Apuzzo[1], Bruna Festa[2], Laura Mancini[1]**
[1]*University of Cassino and Southern Lazio*
[2]*University of Naples "Federico II"*
[3]*University of Naples "Federico II"*

**Abstract:** Rail and track defects are one of the most important sources of dynamic loads generated by a moving train. The oscillations and vibrations decrease the ride comfort, increasing the track deterioration and noise emission. In the spectrum of the vibrations of railway vehicles, the track geometrical irregularities generate excitations at low frequencies, usually below 30Hz, that are responsible for the dynamic behavior of the vehicle. A comprehensive multi-body simulation model of a four-axle passenger rail vehicle is implemented in the multi-body simulation software (MBS) and it is used for analyzing the potentially dangerous running condition. On basis of these multi-body model simulations, the influence of the track irregularities on the vehicle–track dynamic interaction, mainly in terms of lateral (Y) and vertical (Q) wheel-rail forces, their ratio (Y/Q) and contact angle (δ), is investigated. Four kinds of the track related irregularities, according to the ORE B176 standard, are statistically determined: lateral, vertical, roll and gauge. Each one of these, is examined in two different levels of severity (low, high) by means of excitation signals in distance domain on the basis of a Power Spectral Density (PSD) function. In order to individuate the worst run condition varying the train speed within of this range: 70 - 200 km/h, the two different combinations of irregularity levels are implemented. Finally, the relations between the main dynamic parameters (Y, Q, Y/Q, δ) and the track irregularities severity level are analyzed, under certain geometrical conditions. These studies are to be considered in preparation to future investigations aimed to obtain the best simplified model to describe the railway alignment defects occurrence and their evolution.

**Keywords**: track geometrical defects, dynamic behavior, wheel-rail forces, multi-body simulation.

## 1. Introduction

The deterioration of the track geometry is manifested as development of track geometry irregularities. If maintenance is neglected, not performed properly, or performed too late the consequences could range from more extensive and expensive maintenance actions to fatal accidents. As one knows, the fundamental safety requirement for railway transportation is that no derailment accidents take place, otherwise, it may bring about severe losses of lives and materials. Severe track lateral and vertical irregularities and gauge widening are significant causes of derailments that usually occur from a combination of unfavorable vehicle and track conditions (Iwnicki S., 2006) . The understanding of dynamic responses at the wheel-rail interface due to track geometry irregularities and/or wheel and rail defects is necessary for the railway infrastructure engineers to work out the track maintenance strategies. Therefore, measuring and monitoring of the track irregularities represent one of key aspect in order to facilitate maintenance management (D'Apuzzo M. et al 2014). Experience in the rail field underlined that the track geometry deviations (or geometrical defects) act as a major source of vehicle movements and therefore wheel-rail forces, accelerations, damage and safety issues as a cause for derailment. Also, the dynamic component of vertical wheel–rail contact forces, as induced by irregularities in track geometry is an important source to ground-borne vibration and ground-borne noise (D'Apuzzo M., 2012) (Aiello et al., 2008). UIC518 and following EN 13363: 2005 state that the following track geometric quality parameters should be taken into account because they have an impact on vehicle dynamic behavior: vertical alignment, lateral alignment, twist, track gauge. Also, the European standard EN 14363: 2005 specifies the test conditions that must be met to ensure acceptance of the driving conditions of the rolling stock.
Among the various critical conditions that may present, some are related to the stability loss or the crisis of the structural components, while others are related to the loss of functionality or ride comfort of the system. In general, the crisis scenarios can be grouped into various categories, depending how the crisis occurs and the causes that precede it: climbing the wheel to the outside of the track, derailing for structural crisis; the fall of the wheel between the rails , collision or impact with other rolling stock or any other object. The climbing of the flange on the rail is the only of derailment type that occurs as a result of specific running dynamic conditions of the vehicle. Even though many works have been done in the past, the flange climb derailment mechanism is still not well known because of the complexity of the wheel/rail system. Recent researches have demonstrated that the limit value Y/Q, above which derailment occurs, depend on different parameters related to both track and vehicle typology, and it can be defined by only experiments or by means of numerical simulation (Diana G., 2012), whereas in (Pombo J, 2011) an interesting numerical analysis on rail-wheel forces taking into account wheel-rail wear mechanism and track irregularity is presented. In this paper, the emphasis on the interaction between the vehicle and the rail track, with particular attention to the contact problem between wheel and rail, is stressed. The sensitivity of the dynamic responses in terms of lateral forces (Y), vertical forces(Q), contact angle (δ) and ratio Y/Q, to the variation of irregularities level and run speed are investigated.

[1] Corresponding author: laura.mancini@unina.it

## 2. Numerical modeling of the rail track – vehicle system

### 2.1 Simulation tool

In this work, the commercial multi-body software used to study the dynamic behavior of railway vehicles is SIMPACK. A multi-body approach allows to simulate the dynamic performance of integrated railway systems, that are essentially composed of a vehicle, a track and a wheel-rail contact interaction, including the masses and inertias of the structural elements and the characteristics of suspensions. Quasi-static, modal, frequency-response and time-domain analyses can be performed (SIMPACK AG, 2014). In this study mainly time-domain analyses were performed. The simulation package also includes supporting add-in, e.g. for generation of track irregularity input files from measured irregularities; however in these analysis the track irregularities were generated according to ERRI B176 rules.

### 2.2 Model of rail track geometry and irregularities

A railway track defines the three-dimensional run of a railroad in the space (track line) and its role is to guide the trains in a safe e economic manner. This work aimed at solely analyzing the dynamic influence of track geometrical defects on parameters that govern the wheel-rail contact and in particular on the wheel-rail interface forces, that is the rails are mounted on inertia-fixed track and follow exactly the track and its irregularities. In order to investigate the vehicles response in the worst running conditions, EN 14363: 2005 standards defines the tortuosity as one of the key features that a test track must possess, understood as a sequence of different geometric elements, both longitudinal and transverse direction. The track geometry can be described by the nominal geometry and irregularities. As far as the nominal track geometry concerns, it consists of vertical curvature, horizontal curvature, gradient (i.e. slope of track), track gauge and cross level (cant) (Fig. 1(a)). The track considered for dynamic studies conducted in this work is 1344 m long and is composed of straight segment followed by one left curve opposite to one right curve, both with same radius of 275 m; then the track continues with a straight segment. Between all these elements, the transition curves with continuously changed radii are interposed. Also, the cant is changed gradually over the transition length leading to so-called super-elevation ramp; it represents the cant variation along the transition, ensuring a smooth cant evolution from a zero value (at the straight track) to the nominal cant (160 mm) of the circular curve.

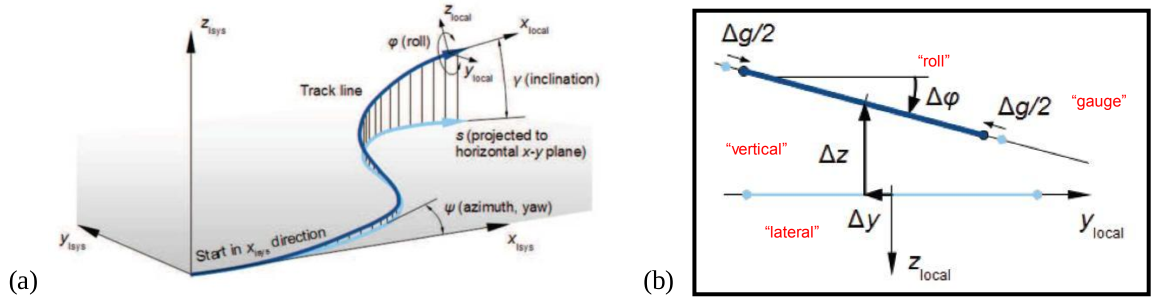
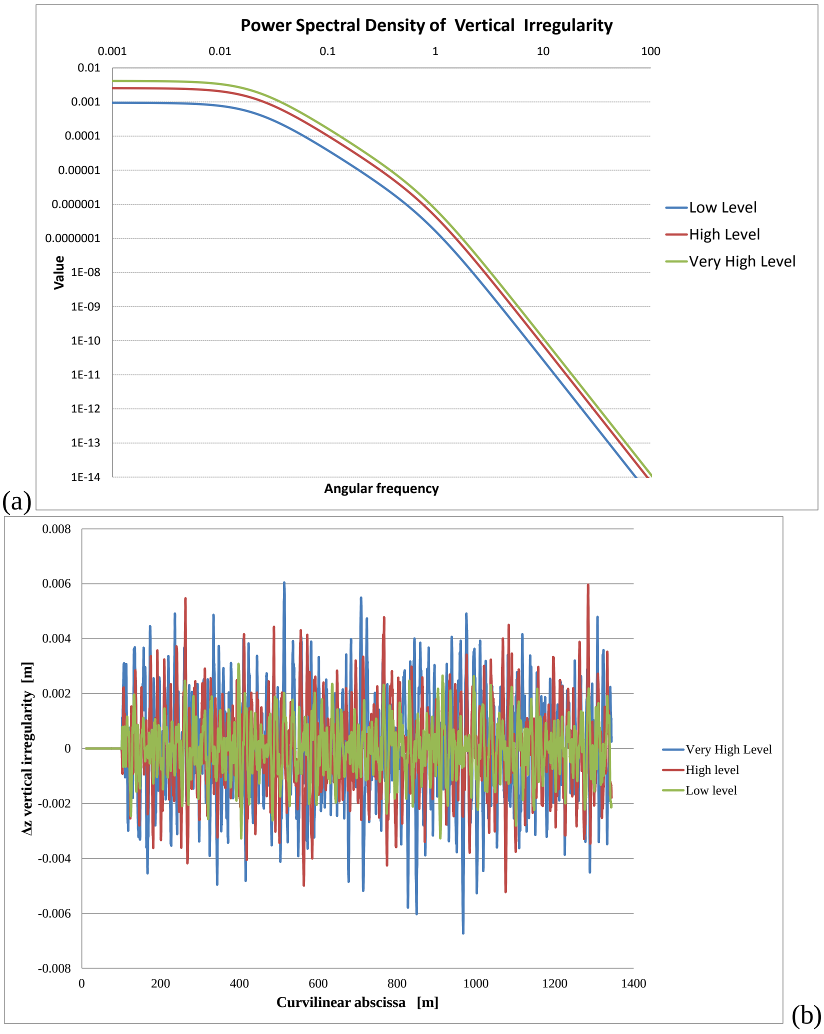



**Fig. 1.**
*Configuration of the (a) track in the spatial domain and (b) track related irregularities, in SIMPACK environment.*

Track irregularities are defined as the deviation of the actual track geometry from the nominal (designed) track geometry (EN 13848-1, 2008). The relation between wheel/track irregularity wavelengths and excitation frequencies at given vehicle speeds is $f = v/\lambda$, i.e. a periodic wheel/track irregularity with wavelength $\lambda$ will generate a dynamic excitation at frequency $f$. In according to ERRI B-176, the track irregularities are defined through the power spectral density (PSD) in distance domain; these PSDs are defined in lateral (horizontal), vertical, roll and gauge direction (Fig.1(b)). Three polynomial PSD functions, which represent gradually increasing amplitude signals, corresponding to the three levels of irregularity, are regarded: low, high and very high. In Fig. 2 (a), the three PSDs of vertical irregularity are reported, corresponding to aforementioned three levels of irregularities and in (b), three typical signals of vertical irregularity assigned to the rail track are represented, and originated from the PSDs above mentioned.

(a)




(b)

**Fig. 2.**
*(a) Typical PSDs according to ERRI B176 and (b) signals of vertical irregularity (Δz) reproduced by PSDs.*

## 2.2 Rail Vehicle Model

The 3D vehicle model used in all the analysis of this paper presents some typical characteristics of railway vehicles in circulation, and is used in the international field as a reference vehicle for the railway simulations. It is called the Manchester Wagon, and consists of a total of seven rigid bodies: a wagon, two bogies and four wheel-sets, whose inertial properties are listed in Table 1. As in each rail vehicle, the involved bodies are variously connected between them, by means of force elements (e.g. suspensions) or by means of auxiliary bodies of mass and inertia null (dummy bodies). To properly simulate the ride of actual vehicles, it is necessary that each link is provided with a certain stiffness and a certain damping, features that are well highlighted in the force elements of the SIMPACK interface. In the present study, it is assumed that the loading is symmetrically distributed on the two rails, at start conditions and the following hypotheses are made: carbody, bogie frames and wheelsets are considered to be rigid; all of the dampers (primary and secondary) are of viscous damping; for each vehicle component three displacements (longitudinal, lateral and vertical) and three rotations (pitching, rolling and yawing) are considered. In the multi-body model, the nominal wheel profiles S1002 and the nominal rail profiles UIC 60 with an 1/20 inclination are employed; also, the rails and the wheels with new profiles are considered.

**Table 1**
*Inertial properties of the principal vehicle components*

| Component | Mass [kg] | Inertia Jxx [kgm$^2$] | Inertia Jyy [kgm$^2$] | Inertia Jzz [kgm$^2$] |
|---|---|---|---|---|
| Carbody | 32000 | 56800 | 1970000 | 1970000 |
| Bogie | 2615 | 1722 | 1476 | 3076 |
| Wheelset | 1813 | 1120 | 112 | 1120 |

## 3. Analysis of results

By means of a multi-body model able to simulate vertical and lateral vehicle-track dynamics and to evaluate corresponding interaction forces at wheel-rail level a sensitivity analysis has been performed. . The purpose is to analyze several indicators and parameters in order to assess how the geometrical track defects influence the vehicle dynamic behavior. In all simulations, the track irregularities (Fig. 1(b)) are considered active simultaneously. The vehicle speed ranges between 70 km/h and 200 km/h, corresponding to real service conditions of a typical conventional rail vehicle; furthermore, a value of coefficient of friction equal to 0.4 was adopted in the performed analysis. Most significant results of simulations will be discussed below with reference to the dynamics of the right wheel of the vehicle leading wheelset where most critical conditions are generally observed.

### 3.1 Vertical (Q) and Lateral (Y) Forces

The resultant force acting at the wheel-rail contact area can be decomposed into longitudinal, lateral and vertical components. In particular, the lateral force Y lays in the plane of the upper surface of the rails, while the vertical component Q is orthogonal to the upper surface of the rails  and acts on the flanging wheel (Fig. 4. (a)), usually during a curve entry or exit transient.

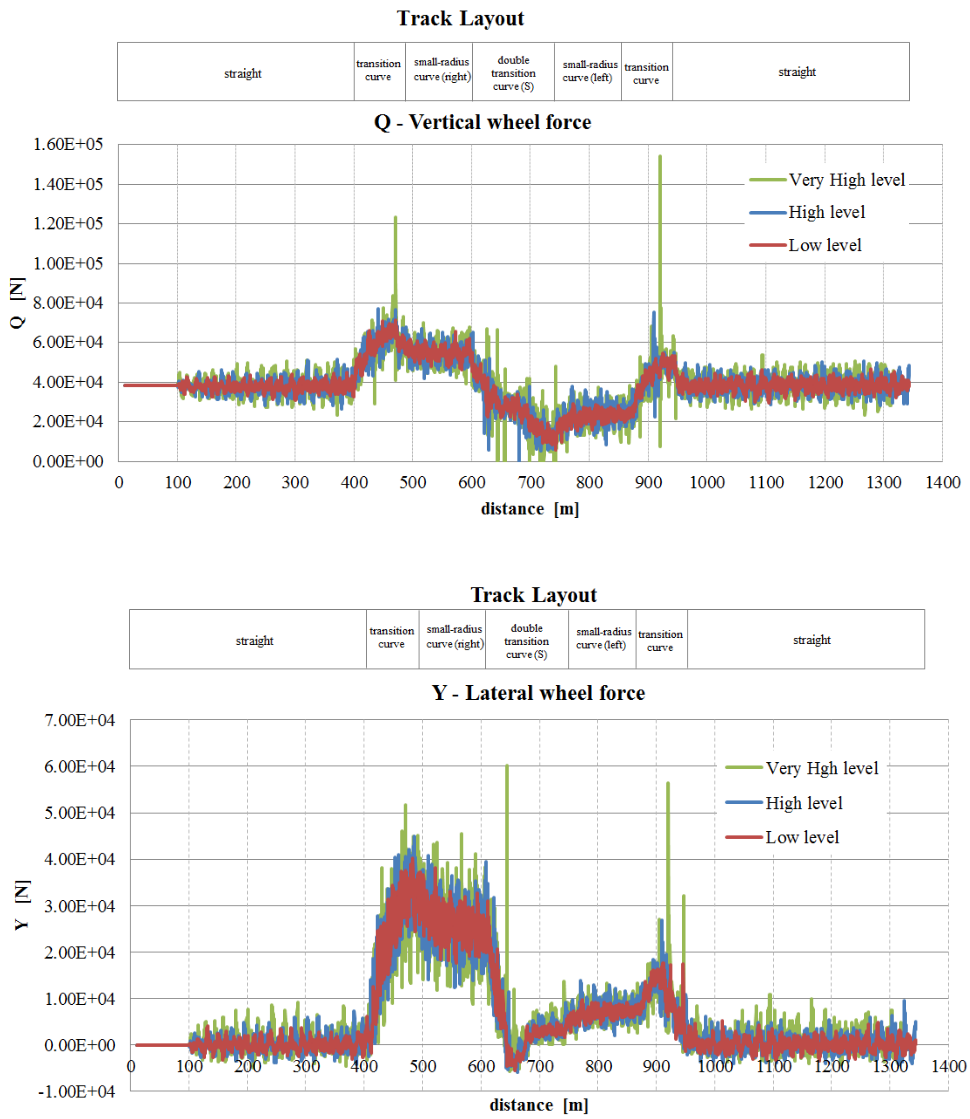



**Fig. 3.**
*Vertical force Q and Lateral force ( speed of run = 90 km/h).*

Since the wheel is assumed to be constantly following the rail, mechanically connected, it is easy to imagine that vertical irregularities would be well correlated to the vehicle reaction in the vertical direction. In the lateral direction, the wheels are connected to the rails via both friction and mechanical contact, unless the wheel flange is in contact with the rail. Analysing Fig. 3, it is possible to see as the maximum values of Q and Y are recorded at track elements with variable curvature; in particular the maximum peaks of the two contact forces are reached at the transition curves' entry and/or exit. Obviously, it can be observed that the dynamic components are larger increasing the irregularities severity, but the dependence of the two forces from them are different. A lateral irregularity might therefore not is expected to have a direct effect on the lateral vehicle forces in the same way as for the vertical direction, in particular with more severe irregularities. However, the combination of larger lateral force and higher vertical load tends to keep the Y/Q ratio below critical levels; at same time, the sensitivity of the vehicle to the longitudinal irregularities, such as the twist, is very important for the derailment conditions. In fact, the unloading of the external wheel of the first wheelset along vertical direction, result to be very dangerous, especially when the vehicle negotiates narrow curves. The combination between the small value assumed by the vertical force Q and the large value of the lateral force Y, being the wheel the flanging one, produces very large values of ratio Y/Q (Fig. 5 (a)).

### 3.2 Derailment coefficient Y/Q

Most of scientific papers published in the past agreed that the fundamental parameter influencing the derailment phenomenon is the Y/Q ratio, (namely the *Derailment Coefficient*); since it is a critical factor in predicting the risk of flange climbing and of rail turnover. Flange climbing, which is one common reason for derailment, means that the wheel flange climbs onto the top of the rail and continues rolling over the rail head. Usually, it occurs in combination with flange steering in curve negotiation but it can also occur when the vehicle is running faster than its critical speed (Iwnicki S., 2006; Panagin R., 1990). The criterion most well-known and historically used for the risk assessment of derailment is the Nadal's criterion (1908), which provides a limit value of the Y/Q ratio, in specific running condition, denoting the Y/Q value at which derailment is supposed to take place. In practice, an Y/Q ratio value in the range of 0.8 - 1.2 is generally considered as the minimum for wheel climb to be likely to occur (Andersson E., 1998). This limit Y/Q value only depends upon friction coefficient, μ, and the maximum contact angle, $\delta_{max}$. This latter parameter depends on itself wheel profile type; for a standardized S1002 profile is near to 70°. In Fig. 4(b), at 70° and 0.4 of friction coefficient, corresponds a limiting Y/Q value equal to 1.11.

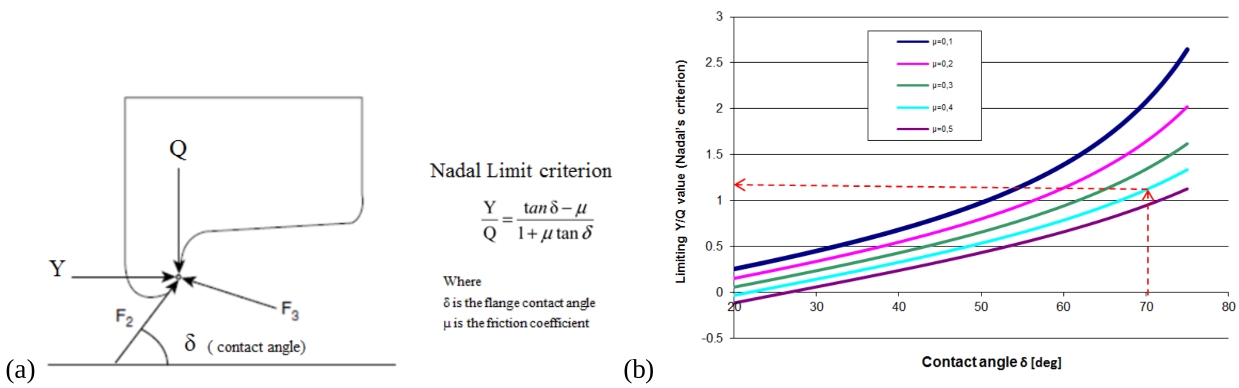



**Fig. 4**
*(a) Nadal's Limit Criterion: forces and contact angle ; (b)Y/Q as function of δ and μ.*

Actually, research works have proved by means of both experimental tests in line and on roller rig and numerical simulation that the Nadal's criterion results too restrictive in practical applications, since underestimates the real behavior of the wheel-rail, and the limiting wheel Y/Q ratio is affected by other parameters, such as the angle of attack of the flanging wheel over the rail and the longitudinal forces magnitude acting at wheel contact (Diana G., 2012); indeed, other simplified formulations set this limiting ratio of Y/Q at 1.2 and indicate that this boundary condition should be maintained during the run for a minimum distance of 2 meters. Therefore, the duration of the occurrence is also an important parameter when studying this issue in addition to knowing which are the track line and running conditions leading to high Y/Q values that may exceed the limit. In Fig. 5 (b), the Y/Q values resulting from a performed simulation are represented at varying irregularity levels (low, high and very high) for 90 km/h of speed, In Fig. 5 (c), the trend of the observed $Y/Q_{max}$ values obtained by 15 vehicle run simulations, are reported; at the highest levels of irregularity correspond more severe values of the ratio, at varying speed too. In these analysis, the vehicle derailment phenomenon was occurred, for all the irregularity levels considered, at both 150 and 200 km/h speed, whose corresponding values of Y/Q are far greater to the limit 1.11. In the others case, although the Y/Q value is greater than the limit value, determined using Nadal's criterion, the wheel did not show any tendency to override the rail during the simulation test.

### 3.3 Contact angle δ

The contact angle δ is the angle of the tangent line in the contact point between the wheel and rail (Fig. 4). It is one of the parameters that influences the mechanism of contact between wheel and rail. In the simulation of the full vehicle, it is observed large variation of the contact angle value mainly in curve, indicating that the contact point of the wheel varies continuously with the track curvature. In fact, when the right wheel negotiates the first curve, is subjected to a centrifugal force that moves the contact point from the wheel plane towards the flange. Also, a more severe track irregularity contributes to increase such oscillations, as represented in Fig. 6. In addition, from the theoretical point of view, the contact point on the profile of the wheel, and the contact angle consequently, moves as that increases the running speed of the curve.

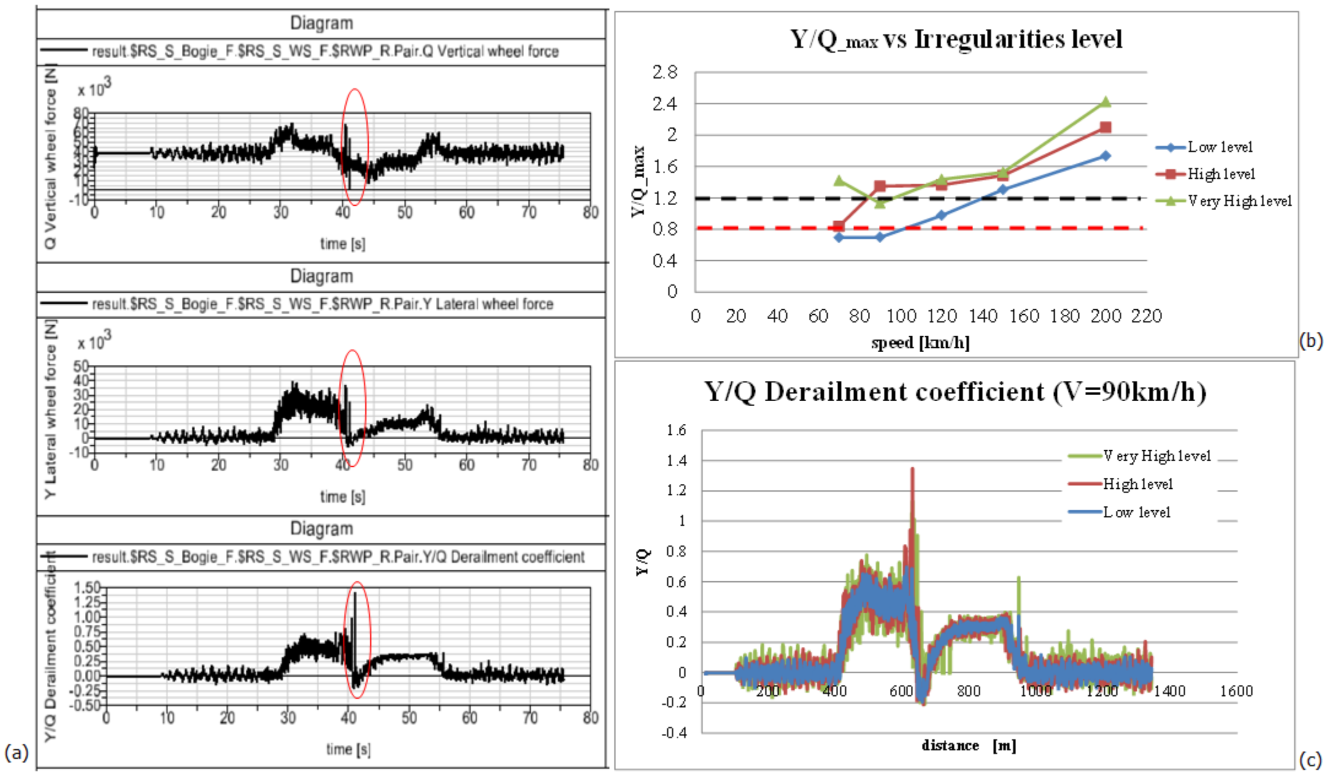



**Fig. 5**
*(a)Q, Y and Y/Q value (with V=90 km/h and very high level of irregularity); variability of maximum Y/Q_max value with speed (b)and Y/Q value with irregularities(c).*

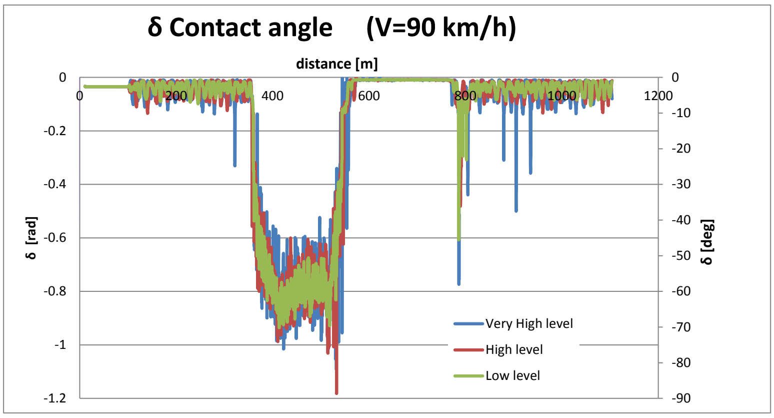



**Fig. 6**
*δ value (for 90 km/h) at varying irregularity level.*

In order to compare the dynamic response of the vehicle at varying speed and irregularity level, in terms of Y/Q and δ, further detailed analysis are performed focusing on the track section from 400 to 430 m, where the simulations have shown the worst dynamic conditions of the vehicle. In Fig. 7 is possible to observe that the trends of Y/Q ratio and δ values, averaged over moving window of 2 m, are inclined to increase with increasing speed and irregularity level, i.e. tending towards the conditions of incipient derailment.

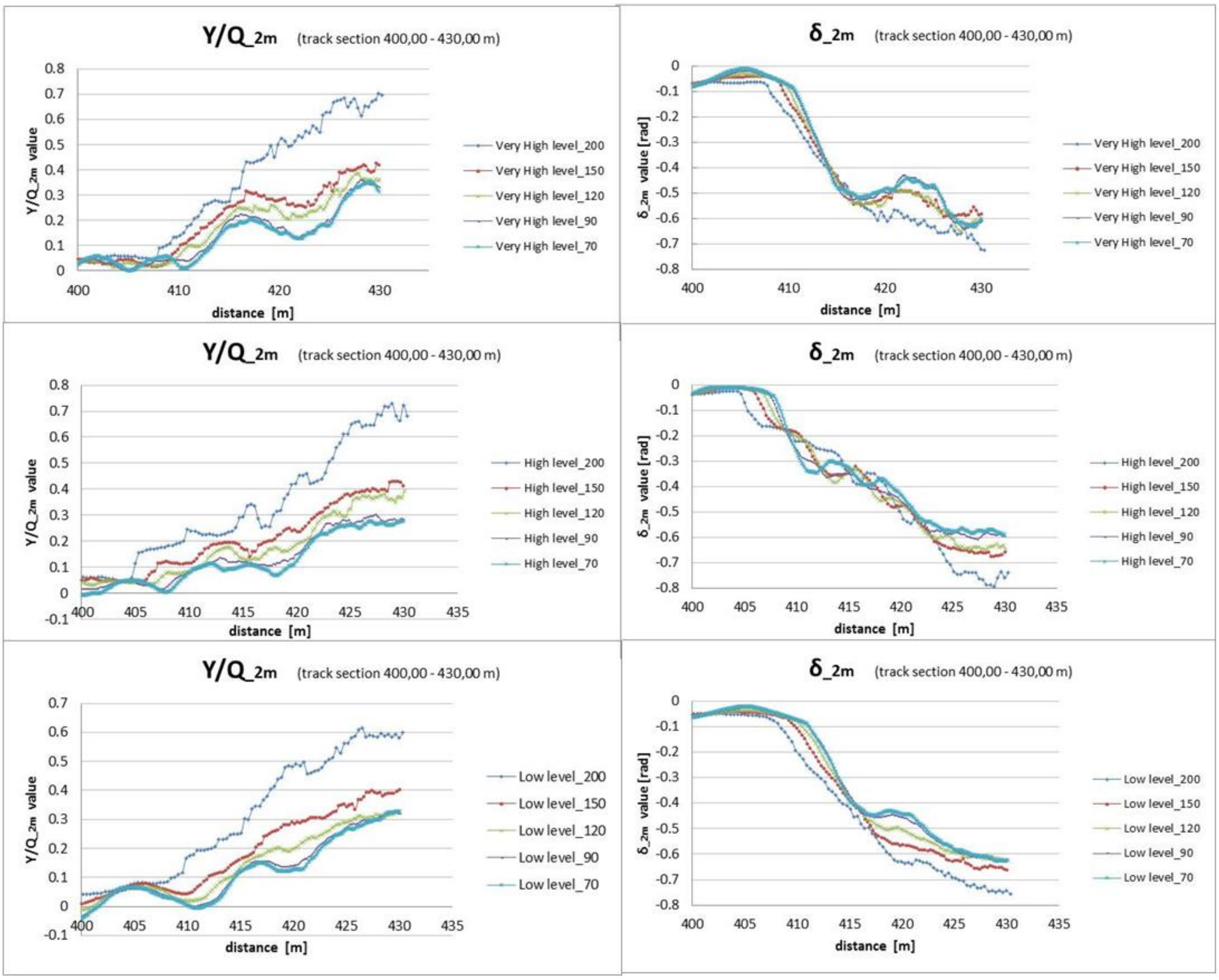



**Fig. 7**
*δ mean value and Y/Q mean ratio for the track section from 400 m to 430 m.*

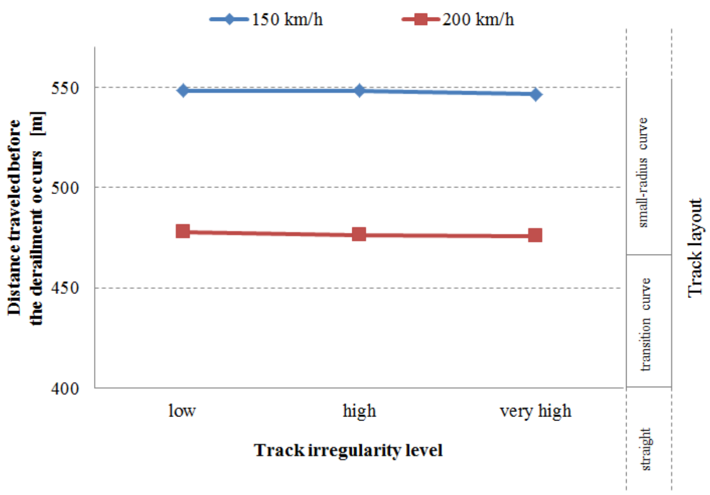



**Fig. 7**
*Distance travelled before the derailment occurs.*

In Fig. 7 is plotted the *Travelled Distance Before* the *Derailment* is occurred (or TDBD) at varying irregularity level and at 150 and 200 km/h speed. In particular, the TDBD increases with decreasing speed irregularity level, maintaining a fixed irregularity level; on the other hand, the TDBD decreases with increasing level of irregularity, although this reduction appears to be not so significant. Note that, all derailment events are occurred in closeness of first transition curve of the rail track.

## 4. Conclusions

In this paper, a 3D multibody rail-vehicle model, developed in SIMPACK simulation environment, has been used to examine the wheel rail interaction response due to track irregularities. For a better understanding and a wider assessment of the derailment phenomenon, it is important to tackle two issues:1) study the operation conditions that promote the loss of the lateral guidance at the wheel and rail interface, and; 2) evaluate the consequences on the dynamic behavior of the railway vehicles. Both problem are addressed in this work with regard to wheel climb derailments and simultaneously presence of all track geometrical defects. The dynamic interactions of the wheel-rail interface have been predicted for four types of track irregularities, namely, lateral, vertical, roll and gauge variations, in terms of Y/Q derailment coefficient, Y lateral force, Q vertical force and δ contact angle. Three combinations of irregularity levels are considered and numerous simulations at varying run speed (70, 90, 120, 150 and 200 km/h) have been performed. Following points can be highlighted to summarize the main results achieved in the present work:

1) the presence of track irregularities worsens the dynamic behavior of railway vehicle and the dynamic components are larger increasing the irregularities severity, assuming the hypothesis of rigid track;
2) the lateral and vertical forces at the wheel-rail interface are influenced by track defects even though in different way;
3) the criticalness of a given Y/Q ratio value is a function of the prevailing dynamic conditions;
4) the derailment conditions are defined not only by the Y/Q ratio, also the duration of the occurrence and flange contact angle are important parameters;
5) the transition curves would seem dangerous points of the track since very pronounced oscillations of the Y, Q and Y / Q values are observed in them.

It has to be underlined that these results so far obtained should be considered as very preliminary and future developments of this work should concern the extension of the range of significant variation of parameters shown here, and the investigation of additional parameters. In particular, further investigations will need to take into account the flexibility of the track, the influence of the singular irregularities related to wheel (flats) or rail (welded joints) and the influence of ballast' settlement.. As far as the vehicle model is concerned, it is intended to develop a simplified wheelset model able to capture the significant dynamics of a whole rolling stock model in order to speed up simulations and provide a prediction tool to investigate on track degradation evolution due to railway traffic.

**References**

Aiello V., Boiero D., D'Apuzzo M, Socco L.V. & Silvestri F. 2008. Experimental and numerical analysis of vibrations induced by underground trains in urban environment. *Structural Control & Health Monitoring*, vol. 15, p. 315–348, ISSN: 1545-2255, doi: 10.1002/stc.247.

Andersson E.; Berg M.; Stichel S. 1998. *Rail Vehicle Dynamics, Fundamentals and Guidelines*. Royal Institute of Technology (KTH), Stockholm.

D'Apuzzo, M., , Festa B . and Miraglia E .; 2014. Evaluation of rail defects provided by high performance measuring devices *Sustainability, Eco-efficiency, and Conservation in Transportation Infrastructure Asset Management* Edited by Massimo Losa and Tom Papagiannakis CRC Press pp. 649–Print ISBN: 978-1-138-00147-eBook ISBN: 978-1-315-75712-4 doi: 10.1201/b16730-95.

D'Apuzzo, M.; Nicolosi V.; Bogazzi, E.. 2012. A Unified Approach for the Prediction of Vibration Inducedby Underground Metro. *PROCEDIA: SOCIAL & BEHAVIORAL SCIENCES*, vol. 53: 62-71, ISSN:1877-0428, doi: 10.1016/j.sbspro.2012.09.860

Diana, G; Bruni, S.; Corradi, R.; Di Gialleonardo, E. .2012. On the derailmentof a railway vehicle. Influence of different parameters, *Scienza e Tecnica Ingegneria ferroviaria*, 2(12): 109-135.

EN 13848-1:2003+A1:2008 (E). 2008. *Railway applications – Track – Track geometry quality – Part 1:Characterisation of track geometry*. European committee for standardization, 25 pp.

ERRI B176/RP1. 1989. *Volume 1: Preliminary studies and specifications. Volume 2: Specifications for bogie with improved curving characteristics.Volume 3: Specifications for a bogie with improved curving characteristics designed forbody tilt*. ETF - Edition, Paris.

Esveld, C.. 2001. *Modern Railway Track*. TU Delft, Duisburg.

Iwnicki, S.. 2006. *Handbook of Railway Vehicle Dynamics*.Taylor & Francis.

Panagin, R.. 1990. *La dinamica del Veicolo Ferroviario*. Leprotto & Bella, Torino.

Pombo, J.; Ambrosio J.; Pereira M.; Verardi R.; Ariaudo C.; Kuka N. 2011. Influence of track conditions and wheel

SIMPACK AG, 2014. *SIMPACK Release 9.6, Simpack Documentation*. Copyright by SIMPACK AG© 2001- 2014.

UIC 518. 2007. *Testing and Approval of Railway Vehicles from the Point of View of their Dynamic Behaviour- Safety-Track Fatigue-Running Behaviour*. UIC, Paris, France.

wear state on the loads imposed on the infrastructure by railway vehicles, *International Journal Computers & Structures*, 89(21-22):1882-1894.

# ASSESSMENT OF RELATIVE ACCIDENTS' RISK DURING DARKNESS

**Stanislaw Gaca, Mariusz Kiec**[1]

[1]*Institute of Road and Railway Engineering, Cracow University of Technology*

**Abstract:** The paper presents a study of relative risk of accidents on Polish roads in darkness. In recent years, about 39% of all accidents took place during insufficient conditions of roads lighting, i.e.: night, dusk and dawn, while the share of traffic, which is exposure to risk during that period, was approximately 29%. Winter season is particularly unfavorable, when due to daylight saving time (DST) changes, the period of adverse lighting conditions is extended.

The paper presents methods identification and assessment various factors on changes risk of accident during darkness. The analysis covers rural sections and roads passing through built-up areas excluding cities, which mostly have adequate artificial lighting.The study distinguished two groups of roads:

- national roads, which mainly carry traffic between major cities with a high share of through traffic, higher geometric standard and better infrastructure
- regional roads with less traffic, which serve local and regional traffic, and also are characterized by lower geometric parameters.

Assessment of accident risk during the darkness based on accident rates taking into account the impact of traffic changes during the day and year. Also, the results of analyses with application of odds ratio are presented. The comparisons included the accidents which took place during the lack of natural light and during daylight. The risk of fatalities on national roads at night-time is about 4-times higher on rural roads and about 2.5-times higher on roads in built-up areas comparing to daytime.

**Keywords:** road safety, accident risk, odds ratio, darkness, type of accident

## 1. Introduction

One of many environmental factors which considerably affect road safety are lighting conditions. During insufficient lighting, i.e.: darkness, dusk or dawn, about 39% of all accidents on Polish national and regional roads occur, fatalities of which account for 54% of all road fatalities. A significant increase in the proportion of fatalities is associated with pedestrian traffic, because more than 40% of accidents on roads in built-up areas during insufficient lighting involved pedestrians, despite their potentially low share in the traffic during the night-time (darkness). Traffic volume, which is the basic variable of risk exposure (Elvik, Erke and Christiansen, 2009; Hermans, Wets andVan den Bossche, 2006) in darkness accounts for only 29% of Annual Average Daily Traffic – *AADT*. The proportions presented are similar to those in other countries (Elvik et al, 2009). An increase in road safety risk during darkness is caused mainly due to the worsening of road perception and worse motoric skills, and drivers' concentration at night-time (Fors and Lundkvist, 2009; Schlag et al, 2009*)* because of visibility restrictions.

Road lighting conditions greatly affect the accident risk in darkness (Elvik et al, 2009; Wanvik, 2009; CEDR, 2010). They improve road area perception and are the most frequently analysed influence on road accidents in darkness. Among other analysed in study factors affecting the accident risk in darkness we can distinguish: weather conditions, road surface conditions, drivers' fatigue, location (urban and rural roads, intersections, section of roads), which were compiled by Johansson et al. (2009). In darkness it can also be noticed that a greater share of drinking-and-driving as well as young drivers, who have less experience and are more prone to risk taking while driving. In studies relative changes are often estimated, not only referring to the number of accidents but also the severity of accidents, types of accidents or their participants, especially considering pedestrians. Statistical analysis of road accidents data clearly confirms an increased risk of road accidents in darkness. The highest ratio of darkness to daylight accidents risk refers to pedestrian accidents – even 7.4-times higher risk, whereas in other accident group it is 4.6. In darkness, the severity of accidents increases significantly, too (Johansson et al., 2009).

There is relatively shallow knowledge of infrastructural road solutions including road lighting conditions affecting road incidents. Mainly, studies focused on the influence of various geometric and infrastructural factors like road and shoulders widths, cross-section type, type of intersection, horizontal parameters (radius, deflection angle), curvature in relation to road safety, hardly ever include their influence in different daylight and darkness periods or lighting conditions (Elvik et al, 2009; Lamm et al., 1999; Lippold, 1997; TRB, 2001; TRB 2012; Vieten et al., 2010).

Speed prediction models for night-time are not frequent in the literature. It is caused by the lack of evidence for statistical significance of influence of road and traffic factors on accidents occurrence. Also the authors attempted to approach the issue of Safety Performance Function development during the night-time as well as to identify the factors affecting road safety based on regression analyses, however, the obtained results were unsatisfactory. Therefore, the assessment of the night-time effect on road safety was carried out using relative measures taking into account traffic volume. For different periods of time (day, night) a number of relative measures of risk were proposed and the results were compared.

[1] Corresponding author: mkiec@pk.edu.pl

Such approach results from the belief that with improving technical standards of roads, it is human factors that mainly affect road accidents. Therefore, having reached a certain level of road infrastructure standards, further improvement in road safety can only be achieved by drivers' education and raising their awareness and affecting their behaviour with preventative and repressive actions. Obviously, road and driver's behaviour interactions are also evaluated. With those interactions in mind, it can be said that infrastructural solutions can also significantly improve the road perception in darkness. Identifying those most crucial road elements that affect accidents risk in darkness comes as a basis for mastering road infrastructure designing. Such a study was carried out in Poland by the authors.

Road sections of single roadway in rural and small town areas were evaluated. City streets, which are mostly sufficiently artificially lit, were excluded. Two groups of roads were identified, i.e. national roads with through traffic and regional ones, with lower traffic volume and a different function (with a high share of local traffic). National roads are mostly of good technical standards and infrastructure, and regional ones are usually of lower technical standards. Winter season is an especially adverse time of the year, when with Daylight Saving Time (DST) transitions the periods of unfavourable lighting conditions and traffic volume increase.

## 2. Methods of analyses

The study on influence of darkness on road safety due to visibility restrictions included two stages:

- assessment measures description of relative road safety risk during darkness for different times of day (24h period)
- relative accident risk assessment , i.e. by comparing types, causes and locations of accidents at different times of day and night,
- comparison of relative road safety measures for roads with different technical standards

The national and regional roads were differentiated for evaluation. Thus, the estimation of influence of technical standards of design and road infrastructure on road safety in darkness was facilitated.

The study analysed sections of national and regional roads with a total length of about 47,000 km, and approximately 90,000 accidents in the years 2005-2009. In this period of time on rural national and regional (local) roads, 48,585 accidents occurred, out of which 18,686 were in conditions of limited visibility (dusk, dawn, night); 40,257 took place in built-up area (through roads), including 14,523 in conditions of limited visibility. Accidents in which the victim or the perpetrator were intoxicated, were removed from the database - 8.9%.

### 2.1. Assessing the risk of accident associated with darkness

In order to assess the risk of accident associated with darkness the methods, based on an odds ratio and an accident rate ratio, were used. The ways, in which analyses were conducted, are described in detail in (Elvik et al, 2009). The analyses were made in correspondence with assessment of road safety at night-time, which were suggested in previous studies (Sullivan, Flannagan, 1999; 2001; 2002; 2003), and later developed by (Johansson et al., 2009).
In the analysis the method of an odds ratio was used to eliminate the influence of traffic volume variety by comparing the number of accidents in pairs, at specified time periods (hours), regardless of traffic volume. Apart from such an approach, results of accident rate ratios were also estimated based on the share of traffic volume in specified hours of the day and night. In order to do so, an extensive database of 488 test sites all over Polish national and regional roads network was used (Gaca et al., 2003-2008).
Using the data of road accidents and traffic volume shares in specified hours of the day and night, the authors estimated a relative accident risk in darkness and night-time (accident rate ratios) for different time periods of daylight and darkness. It is crucial to select for the comparison suitable periods which are determined by the accidents data availability. For further analyses a constant and an average period of daytime and night-time or daylight and darkness was defined.

### 2.2. The selection of time intervals for road safety risk analyses

Based on Sullivan and Flannagan's (1999), to select the time intervals for road safety risk analyses, seasonal variations of natural lighting as well as Daylight Saving Time (DST) transitions in March and October were taken into consideration. The hour and month periods for the analysis are based on sunrise and sunset times for 52°15' North latitude and 21°00' East longitude (Warsaw). In the first stage, the time periods with night-time or daytime only throughout the whole year were selected. Figure 1 shows the optimal, for the analysis, time intervals selection with night-time or daytime only, regardless of the season of the year. For night-time period the time between 10 PM and 4 AM was selected, whereas for daytime it was 8 AM – 3 PM. Such approach ensures a suitable correlation between accidents and the time of the day or night. However, it restricts numbers of accidents in the analysis, so that is why attempts were also made to estimate an average daytime and night-time, including the real sunrise and sunset times, which significantly increases the number of the accidents in the analysis (especially in darkness).

To define the times of "day" (daytime only) and "night" (night-time only) a 15-minute accuracy is used as well as summer and winter times. Table 1 illustrates periods of specified lighting conditions depending on the season of the year, according to the assumptions above.

**Table 1**
*Periods of Specified Lighting Conditions Depending on the Season of the Year*

| | **Summer time (April-October)** | **Winter time (November – March)** |
|---|---|---|
| Day | 5:45AM – 7:30PM | 7:15AM - 4:00PM |
| Dusk | 7:30PM – 8:30PM | 4:00PM - 5:00PM |
| Night | 8:30PM – 4:45AM | 5:00PM - 6:15AM |
| Dawn | 4:45AM – 5:45AM | 6:15AM - 7:15AM |

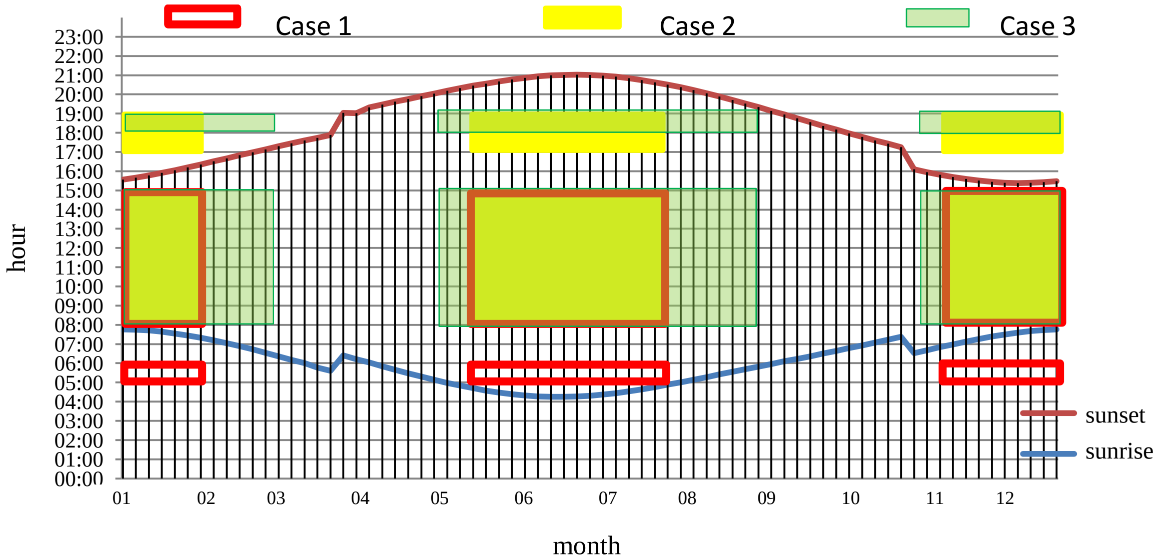



**Fig. 1.**
*Selection of an hour periods daytime and night-time and for relative analysis including an odds ratio with respect to sunrise and sunset time variations throughout the year*

In order to make a relative estimation of road safety risks applying an odds ratio, the methodology described by (Johansson et al., 2009) was used, which was modified by including the geographical location of Poland, hence day and night light conditions changes. The modification of the method was designed to achieve the highest possible accidents sample size by extending the analysed and compared time interval, and to eliminate the transition light periods of dusk and dawn. This way 3-time periods, illustrated in Figure 1 and Table 2, were selected for the analysis.

**Table 2**
*Selection of Risk Estimation Cases with an Odds Ratio Method*

| | **Analysis hours** | **Comparison hours** | **Summer time** | **Winter time** |
|---|---|---|---|---|
| Case 1 | 5.00-6.00 | 8.00-15.00 | 11 MAY – 31 JUL | 12 NOV – 31 JAN |
| Case 2 | 17.00-19.00 | 8.00-15.00 | 11 MAY – 31 JUL | 12 NOV – 31 JAN |
| Case 3 | 18.00-19.00 | 8.00-15.00 | 1 MAY – 31 AUG | 1 NOV – 28 FEB |

As Figure 1 illustrates, the cases differ one from another in terms of analysed times both in daylight and darkness, and season of the year aspects. Comparison hours in all three cases were the same during the day, but in the third case the analysed time during the year was longer. The second case comprises a two-hour analysed period during the day, whereas the first and the third case comprise only one hour.

### 2.3. Estimates of relative risk in road safety during darkness

All measures used for calculating relative risks associated with darkness are presented below. In the analysis of the effects of restricted visibility associated to darkness on road accidents the following accident rate ratios were defined:

- an hour relative risk ratio $D_h$ obtained according to the following equation:

$$D_h = \frac{Acc_i/STV_i}{\Sigma Acc_i/\Sigma STV_i} \; [-] \tag{1}$$

where:
$Acc_i$ – number of accidents in the analysed hour in the analysed period,
$STV_i$ – share of traffic volume of the specified hour during the day [%],
$i = 1,2,...,24$.

This ratio defines the accident risk in the specified hours of day and night. As the exposure risk variable the percentage share of traffic volume in the specified hour with respect to the whole day and night period is used;

- "night-time" relative risk ratio $D_{n/d}$ obtained according to the following expression:

$$D_{n/d} = \frac{Acc_{22-4}/STV_{22-4}}{Acc_{8-15}/STV_{8-15}} \, [-] \tag{2}$$

where:
$Acc_{22-4}$ – number of accidents in the night-time period, hours 10:00PM-4:00AM,
$Acc_{8-15}$ – number of accidents in the daylight period, hours 8:00AM-3:00PM,
$STV_{22-4}$ – share of traffic volume in the period of night-time, hours 10:00PM-4:00AM [%],
$STV_{8-15}$ – share of traffic volume in the period of daylight, hours 8:00AM-3:00PM [%].

This ratio defines the accident risk in the night-time period with respect to daylight period. As the exposure risk variable the percentage share of traffic volume in night-time and daylight is used;

- "accident" relative risk ratio $D_{acc\_night/acc\_day}$ obtained according to the following expression:

$$D_{acc\_night/acc\_day} = \frac{Acc_{night}/STV_{night}}{Acc_{day}/STV_{day}} \, [-] \tag{3}$$

where:
$ACC_{night}$ – number of accidents classified as night ones according to “traffic accident report”,
$Acc_{day}$ – number of accidents classified as daylight ones according to “traffic accident report”,
$STV_{night}$ – mean traffic share in the darkness [%],
$STV_{day}$ – mean traffic share in the daylight [%].

This ratio allows implementing a different approach and it was obtained in order to compare the increase values in risks in darkness periods, obtained according to the "night-time" ratio, as shown in equation 2. It was obtained using all accidents classified as “night” or “day” ones as described in the police database with respect to lighting conditions (traffic accident report). In order to simplify the calculation, an average traffic share in “night” and “day” – estimated by the authors basing on cyclical traffic measurements (3-4 times per year) in years 2003-2008 in Poland (Gaca et al., 2003-2008) – when accidents occurred, were used;

- "seasonal" relative risk ratio $D_{Si}$ in darkness with respect to season of the year:

$$D_{Si} = \frac{Acc_{darkness}/STV_{darkness}}{Acc_{daylight}/STV_{daylight}} \, [-] \tag{4}$$

where:
$Acc_{darkness}$ – number of accidents in darkness in the analysed year period,
$Acc_{daylight}$ – number of accidents in daylight in the analysed year period,
$STVN_{darkness}$ – average traffic share in darkness in the analysed period [%],
$STVN_{daylight}$ – average traffic share in daylight in the analysed period [%],
$i$ – season of the year: summer, winter.

The seasonal ratio is suggested to analyse risk of accidents in darkness in comparison to the daylight ones, with respect to the season of the year. It was obtained by applying average traffic share values in darkness or daylight. Such an approach allows to assess the potentially worse influence of weather conditions on road safety during the winter season in comparable lighting conditions. The total number of accidents in periods and hours were used according to the time periods shown in Table 1;

- odds ratio method

In the case of estimates of risk in the darkness period developed by the odds ratio method (Wanvik, 2009; Johansson et al., 2009), it is possible to estimate relative risk changes excluding the traffic share risk exposure factor. Seasonal and day variations are identified by applying an odds ratio with respect to specified analysis periods and comparison period (see Figure 1 and Table 2). The odds ratio estimate was then obtained as follows:

$$\text{Odds ratio} = \frac{\text{Accidents in analysed hours winter/Accidents in analysed hours summer}}{\text{Accidents in comparison hours winter/Accidents in comparison hours summer}} \, [-] \tag{5}$$

For all estimates of ratios cases their confidence intervals were obtained (95%) including variability of the phenomenon described, according to (Elvik et al., 2009).

Risk assessment method based on odds ratio may lead to overestimation of the relative change in the number of accidents. The odds ratio value strongly depends on the selection of the compared hour which is related to time-dependent changes in risk exposure. The most comparable results based on this method are available only in countries which use Daylight Saving Time (DST) transitions.

All relative measures shown here allow to assess the change in accident risk during conditions of changing lighting conditions for various entry data. Further part of this paper shows their applications, differences and recommendations for use.

## 3. Results

Estimated results of the ratios defined in methods of analyses are presented in this section. The estimate value of $D_h$ is illustrated in Figure 2. It shows that the hazardous period associated with darkness on rural roads falls on 5:00PM-5:00AM. On road sections through towns and villages it is 4:00PM-3:00AM, and on regional roads 2:00PM-11:00PM. The most hazardous hours are: 9:00PM-2:00AM on rural roads and 4:00PM-2:00AM on road sections through towns and villages. The change in trend of the $D_h$ ratio in the last 4 hours in case of regional roads can be attributed to a significant decrease of traffic intensity on these roads, especially when compared with national roads after 8:00PM.

While assessing the estimate values of $D_{n/d}$ it was stated that :

- the risk of an accident occurrence in night-time in comparison to daytime on national roads increases by 60%, and on regional roads by 15%. This difference can be a result of a greater night/day split of traffic volume on the regional (lower standard) roads than on the higher standard "national" roads
- the number of fatalities in night-time in comparison to daytime on national roads increases 3.7-times ($D_{n/d}$=3.77), and on regional roads over 2.5-times ($D_{n/d}$=2.50).

The ratio $D_{acc\_night/acc\_day}$ obtained according to the equation (3), where there is no constant hour interval for "darkness" and "daylight" times was 1.52 for national roads and 1.28 for regional roads. The values do not differ significantly from the values obtained by applying the method with night-time and daylight hours (the difference is of 5% for national roads and of 10% for regional roads). Due to seasonal variations of daylight and darkness times in Poland, the obtained average traffic share in darkness or daylight is only a rough estimation. Simplifications applied when calculating this ratio were more significant than when calculating the ratios according to the equations (1) and (2). Nevertheless, the obtained value of the increase ratio of road accident risk associated with darkness is similar to the one estimated with more accurate data estimation used for its calculation.

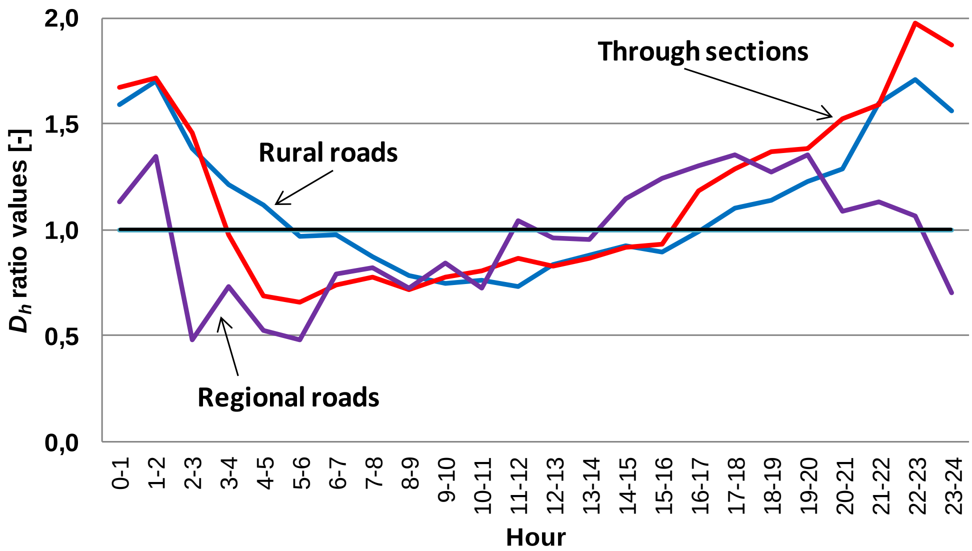



**Fig. 2.**
*$D_h$ ratio values on specified road sections*

In the last stage of analysing a relative increase of road safety risks in darkness applying an accident ratio methodology the data from Daylight Saving Time (DST) transitions periods, both in spring and autumn, was used. Such approach allows separation of adverse weather conditions in winter. The relative accident risk ratios in darkness obtained according to equation (4) are presented in table 3.

**Table 3**
*Relative Accident Risk Ratios in Darkness Obtained According to Equation (4)*

| | **Summer Time (April-October)** | | **Winter Time (November-March)** | |
|---|---|---|---|---|
| | **National roads** | **Regional roads** | **National roads** | **Regional roads** |
| Seasonal risk ratio in darkness | 1.63 | 1.23 | 1.55 | 1.31 |

The ratio $D_{Si}$ (equation 4), showing the risk of road safety in darkness to daylight, indicates that the risk in darkness is greater throughout the year, being slightly higher in summer time on national roads and in winter time on regional roads. Weather conditions may affect these findings as well as lower road standards and their winter maintenance.

The odds ratio results, obtained according to equation (5) for various cases of hour periods analysis, were shown in table 4. The justification for the selected periods of analysis are given in methods and analysis. The morning period and evening periods shown in table 2 were included. Greater attention should be paid to the evening period since it was represented by a larger sample size, and the results of the study so far indicate a greater decrease in safety in the evening rather than morning periods.

As it was expected, the influence of darkness on road safety (5:00PM – 7:00PM Case 2 and 6:00PM – 7:00PM Case 3) is far greater. During the morning hours (Case 1 – 5:00AM – 6:00AM) the lack of natural light resulted in only a 6% increase on national roads. Taking this estimation into account, attention must be paid to the choice of the compared hour. Selecting times 5:00PM-6:00PM from case 2 leads to significant overestimation, especially for local roads. Greater values of ratios for early hours are a result of greater risk exposure in this period (greater traffic intensity, high pedestrian activity). Results far more comparable with those obtained using methods shown here, were obtained when the comparison period was extended to 5:00PM-7:00PM (case 2). This clearly indicates the importance of comparison period selection and its influence on the representativeness of the results in the scale of the whole 24h period.

**Table 4**
*Relative Road Safety Risk Ratios in Darkness Estimated Applying the Odds Ratio Methodology According to Equation (5) for All Users*

| | **Case 1** | | | **Case 2** | | | **Case 3** | | |
|---|---|---|---|---|---|---|---|---|---|
| | period 11.05-31.07 12.11-31.01 | odds ratio national roads | odds ratio regional roads | period 11.05-31.07 12.11-31.01 | odds ratio national roads | odds ratio regional roads | period 01.05-31.08 01.11-28.02 | odds ratio national roads | odds ratio regional roads |
| analysed hours | 5.00AM-6.00AM | **1.06** | **1.31** | 5.00PM-7.00PM | **1.65 (1.72)*** | **1.26 (1.48)*** | 6.00PM-7.00PM | **1.55** | **1.15** |
| comparison hours | 8.00AM-3.00PM | | | 8.00AM-3.00PM | | | 8.00AM-3.00PM | | |
| 95 % confidence interval | | (1.04 ; 1.09) | (1.28 ; 1.35) | | (1.64 ; 1.66) (1.71 ; 1.75)* | (1.22 ;1.29) (1.47 ; 1.50)* | | (1.54 ; 1.56) | (1.14 ; 1.17) |

**(...) – results for 5:00PM-6:00PM period*

Table 5 shows results of ratios calculated for different road accidents types on both national and regional roads with using "night-time" relative risk ratio $D_{n/d}$ (2) and odds ratio (5). Such a comparison was considered as justified since it allows assessing ratios and additional influence of a road technical standard on accidents in darkness for various periods (evening (5) and night-time (2)). For that purpose, the values of risk changes ratios, for accidents in darkness for various analysis hours, were estimated. The ratio calculated according to equation (2) includes the hours when an additional influence of a driver's fatigue an accidents may arise.

**Table 5**
*Different types of accidents risk changes estimation results for accidents in darkness and at night-time*

| | Type of accident | National roads | | Regional roads | |
|---|---|---|---|---|---|
| | | Best estimate | Confidence interval (95%) | Best estimate | Confidence interval (95%) |
| "night-time" relative risk ratio $D_{n/d}$ (equation 2) | head-on collision | 1.36 | 1.32 - 1.40 | 0.57 | 0.55 - 0.60 |
| | side collision | 0.86 | 0.83 - 0.88 | 0.42 | 0.41 - 0.44 |
| | rear-end collision | 0.86 | 0.83 - 0.89 | 0.46 | 0.44 - 0.48 |
| | pedestrian | 2.79 | 2.70 - 2.88 | 1.54 | 1.49 - 1.59 |
| | hitting other vehicle. | 1.66 | 1.50 - 1.84 | 0.97 | 0.85 - 1.11 |
| | hitting the animal | 8.29 | 6.45 - 11.34 | 3.39 | 2.75 - 4.30 |
| | rollover | 2.50 | 2.39 - 2.61 | 1.95 | 1.87 - 2.03 |
| | hitting the obstacles | 3.22 | 3.09 - 3.35 | 2.80 | 2.72 - 2.89 |
| | others | 2.57 | 2.41 - 2.75 | 1.87 | 1.75 - 1.99 |
| odds ratio (equation 5) | head-on collision | 0.89 | 0.87 - 0.92 | 0.68 | 0.64 - 0.71 |
| | side collision | 0.59 | 0.57 - 0.60 | 0.39 | 0.37 - 0.39 |
| | rear-end collision | 0.80 | 0.77 - 0.83 | 0.34 | 0.30 - 0.36 |
| | pedestrian | 4.90 | 4.88 - 5.16 | 4.54 | 4.50 - 4.88 |
| | hitting other vehicle. | 0.58 | 0.31 - 0.84 | 1.33 | 0.76 - 2.84 |
| | hitting the animal | 2.00 | - | 1.33 | - |
| | rollover | 0.58 | 0.53 - 0.62 | 0.24 | 0.19 - 0.25 |
| | hitting the obstacles | 0.58 | 0.53 - 0.62 | 0.56 | 0.51 - 0.59 |
| | others | 0.74 | 0.61 - 0.86 | 0.52 | 0.39 - 0.59 |

The data compared in table 5 indicates that:

- the selection of analysis time periods, in order to estimate the accidents in darkness risk changes ratio, affects the results of the estimation. For instance, an increase in pedestrians-involving accidents in darkness calculated according to equation (5) is significantly higher than, when calculated according to equation (2). The huge difference appears also in the estimation of risk for accidents at night-time when hitting obstacles in the road surrounding. Referring to the ratio calculated using equation (2), it can be observed a significant increase in the accidents risk at night-time mentioned above, and using equation (5) a decrease in the accidents risk in darkness. Such a difference can result from a role of drivers' fatigue at night-time, which is not the case in the evening hours (included in eq. 5).
- the increase in the accidents risk at night-time on regional roads is smaller than on national roads with reference to almost all accident types. Only risk of the hitting an obstacle in the road surrounding at night-time is higher on regional roads rather than on national, which can be caused lower technical standards for regional roads.

**Discussions**

Accident risk on national roads estimated with relative ratios increases by about 60% at "night", compared to "day", and on regional roads by about 15-20% (depending on ratios).
The number of fatalities at night-time on national roads may almost quadruple and may be 2.5-times higher, compared to daytime, on regional roads. That night increase corresponds with the type of accident. The risk is especially high in terms of pedestrian-vehicle accident or driving into obstacles like trees, pillars, road barriers or animals. The likelihood of head-on collisions on national roads is also slightly higher at night-time or darkness whereas side-on or rear-end collisions risks in both day and night conditions are similar.
Pedestrians are also the group especially prone to an increased accident risk at night-time. The relative accident risk in that group at night-time is 1.34 on built-up areas roads and a staggering 8.67 on rural sections of national roads. This value, however, does not reflect a full phenomenon of an increased pedestrian risk in darkness since no pedestrian volume was involved in the analyses. Artificial road lighting decreases the severity of accidents by decreasing the accident severity ratio. On through sections of national roads in towns and villages a huge increase in pedestrians share was noticed outside daylight periods. On a road unlit at night-time the driving into an obstacle accident share also increased.
Regional roads, in comparison to national roads, are generally of lower technical standards and they carry both through and local traffic. This affects the distribution of road accidents both in darkness and daylight. In darkness risk of driving into an obstacle accident increases significantly, but less so in case of pedestrian–vehicle type. Pedestrian–vehicle accident share is also large with artificial road lighting, which may result from the fact that it is located near pedestrian crossings.
Among the most predominant causes for darkness accidents, apart from speeding, causes related to pedestrian–vehicle accidents can be noted (careless pedestrian or driver road crossing, and other pedestrians' wrongdoings).

The comparison of accident ratios and an odds ratio estimation methodologies provides similar findings in a longer-term analysis period (2 hours). Conducting analysis for the most unfavourable hour (5:00PM-6:00PM) leads to overestimation of accident risks associated with darkness (Johansson et al., 2009). This method does not allow to assess the variability of accident risk in times of day (24h period) when drivers' fatigue may significantly influence the number of accidents (night). Neither does it allow to assess the accident risk change between periods with significantly different traffic characteristics.
The presented measures indicate comparable analyses results in respect to general changes of accident risk. In cases of more detailed analyses (types of accidents, causes, participants), the selection of analysis period is significant. In such cases, use of the odds ratio method is limited.
Use of relative road safety measures for accidents during night-time is particularly important, as statistical significance of independent variables was not achieved when attempting to develop Crash Modification Factor separately for night and day. Therefore, it is reasonable to try to use relative measures for assessment of road safety improvement methods (for various times of the 24h period) instead of CMF. However, such an approach is applicable only when assessing certain types of accidents. For example the effectiveness of pedestrian crossing with flasher (permanently flashing) is 0.842 when compared with a zebra crossing (accidents involving pedestrians) throughout the whole 24h period and 0.699 during the night-time. This was calculated basing on a ratio of values obtained using equation 3 for pedestrian crossing with flasher and a zebra crossing. It confirms that the effectiveness depends on the time of the day (24h), which is not taken into account in the HSM approach. The next step of the research can be assessment of possible use of CMF with a correction factor taken into account, which would compensate for changing traffic conditions. Relative measures, described above, can be used in such a correction factor.

**Conclusions**

The conducted analyses prove an increase in road safety risks in darkness. The values of a risk ratio of road safety increase in darkness $D_h$ were 1.3–2.2, depending on a road group. Much higher values of that ratio, compared to daylight, was obtained in reference to selected accident groups: pedestrian-vehicle accidents – 4.90 (national roads) and 4.54 (regional roads) in darkness period, driving into an obstacle at the roadside accidents - 3.22 (national roads) and 2.88 (regional roads) at night-time.
Taking into consideration all the presented methods of relative road safety risks estimations in darkness compared to daylight, it can be stated that the accident risk in darkness on national roads is about 60% higher, compared to daytime. On regional roads that risk is much smaller (1.15-1.30). The accident risk increase depends on the season of the year and time of day or night (the risk is higher in winter time and singificant at dawn, compared to national roads). Additionally, results support that in case of regional roads, the role of infrastuctural factors in the accidents determinants group is bigger than on national roads. The analysis described in the paper, which used different ratios for evaluation of changes in accident risks in darkness with reference to daylight, proved a significant influence of choosing the analysis period on the estimation results.

**Acknowledgements**

Research was conducted as a part of a research project N509 254437 titled "*Identification of road safety determinants in conditions of night-time visibility restrictions*" financed by National Science Centre in Poland.

**References**

CEDR. 2010. *Road lighting & safety*. Publication Conference of European Directors of Roads (CEDR), Cedr's Secretariat General.

Elvik R., Erke A., Christensen P. 2009. Elementary units of exposure. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2103: 25–31.

Elvik R., Hoye A., Vaa T., Sorensen M. 2009. *The Handbook of Road Safety Measures*. Emerald Group Publishing, The second edition,.

Fors C., Lundkvist S.-O. 2009. *Night-time traffic in urban areas*. Publication VTI rapport 650A, Linköping Sweden.

Gaca S., Jamroz K., Ząbczyk K. et al. 2003-2008. *The nationwide Study of the Measurements od Vehicles' Speed and Safety Belt Use*. SIGNALCO Kraków – Trafik Gdańsk – BIT Poznań. Available from Internet: <www.krbrd.org.pl>.

Hermans E., Wets G., Van den Bossche F. 2006. Describing the Evolution in the Number of Highway Deaths by Decomposition in Exposure, Accident Risk, and Fatality Risk. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1950: 1–8.

HSM. 2010. *The Highway Safety Manual*. American Association of State Highway Transportation Professionals, Washington, D.C.

Johansson Ö., Wanvik P.O., Elvik R. 2009. A new method for assessing the risk of accident associated with darkness. *Accident Analysis & Prevention*, Vol. 41: 809-815.

Lamm R., Psarinos B., Mailander T. 1999. *Highway Design and Traffic Safety Engineering Handbook*. New York, Mc Graw-Hill.

Lippold, Ch. 1997. *Weiterentwicklung ausgewählter Entwurfsgrundlagen von Landstrassen*, Dissertation Technische Hochschule Darmstadt.

Schlag B., Petermann I., Weller G., Schulze Ch. 2009. *Mehr Licht – mehr Sicht – mehr Sicherheit?* VS Verlag für Sozialwissenschaften GWV Fachverlage GmbH, Wiesbaden.

Sullivan J. M., Flannagan M. J. 1999. *Assessing the potential benefit of adaptive head lighting using crash databases*. Report No. UMTRI-99-21, The University of Michigan Transportation Research Institute.

Sullivan J. M., Flannagan M. J. 2001. *Characteristics of Pedestrian Risk in Darkness*. Report No. UMTRI-2001-33, The University of Michigan Transportation Research Institute.

Sullivan J. M., Flannagan M. J. 2002. The role of ambient light level in fatal crashes inferences from daylight saving time transitions. *Accident Analysis & Prevention*, Vol. 34: 487-498.

Sullivan J. M., Flannagan M. J. 2003. Risk of fatal rear-end collisions is there more to it than attention? *Proceedings of the Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, Park City, Utah, 2003.

TRB Executive Committee. 2001. *Recent Geometric Design Research for Improved Safety and Operations*, *NCHRP Synthesis 299*. Transportation Research Board of the National Academies, Washington, D.C.

TRB Executive Committee. 2012. *Recent Roadway Geometric Design Research for Improved Safety and Operations. A Synthesis of Highway Practice*, *NCHRP Synthesis 432.* Transportation Research Board of the National Academies, Washington, D.C.

Vieten M., Dohmen R., Dürhager U., Legge K. 2010. *Quantifizierung der Sicherheitswirkungen verschiedener Bau-, Gestaltungs- und Betriebsformen auf Landstraßen*. Berichte der Bundesanstalt fur Strassenwesen, Verkehrstechnik, Heft V 201, Bergisch Gladbach.

Wanvik P. O. 2009. Effects of road lighting - An analysis based on Dutch accident statistics 1987–2006. *Accident Analysis & Prevention*, Vol. 41: 123-128.

# ADRIATIC SEA WEATHER HAZARD AND ITS INFLUENCE ON SEA TRANSPORT

**Ružica Popović[1], Captain Željko Šore[2], Dario Medić[3]**
[1]*Ph.D. Student, Meteorological and Hydrological Service, Marine Meteorological Centre, Glagoljaška ll, 21000 Split, Croatia*
[2]*Meteorological and Hydrological Service, Marine Meteorological Centre, Glagoljaška ll, 21000 Split, Croatia*
[3]*Ph.D. Student, University of Split – Faculty of Maritime Studies Split, 21000 Split, Croatia*

**Abstract:** The Adriatic Sea is the northernmost gulf of the Mediterranean, which penetrates deeply towards the Middle Europe. It extends for 783 km between the Dinaric-Balcanic mountains on the north-east and the Apennine peninsula on the south-west, having the mean width of 248 km. Its area during the mean sea level is 138595 km$^2$. The maximum depth is 1330 m. Although the Adriatic Sea is just a small closed gulf of the Mediterranean Sea it has often had extreme and dangerous maritime meteorological phenomena. Main Adriatic weather hazard are storm winds (wind speed ≥ 17.2 m/sec) and hurricanes (≥ 32.7 m/sec), high waves (≥ 2.5, 3, 4, 5, 6, 7, 8 and 10 m) and limited visibility – fog (visibility <1000m). The above mentioned phenomena will be observed separately or in linked interactive groups which increase the danger and which should be announced in time as a part of the system of preventive meteorological warnings aiming to protect human lives, nature and goods in the entire region of the Adriatic sea. Maritime meteorology participates materially in safety, regularity and efficiency of maritime transport by studying meteorological elements and phenomena from a viewpoint of influence on maritime technique and conditions of navigation. Negligence weather conditions leads to failure of maritime tasks, and sometimes even accidents.

**Keywords**: Adriatic sea, hurricane, high waves, fog, maritime meteorology, safety

## INTODUCTION

Sea transport is one of rare economic activities which is in constant increase. Not even the world economic crisis, wich has hit many economies in the last years, has left difficult or lasting consequences on sea transport. Increasing number of nations that do not have access to the sea use sea transport as the main way of transport, mostly because of relatively low prices (in comparison to other ways of transport) and the possibility of easy connection to any part of the world.
Ship transport in the Adriatic Sea is bigger every day, as well as the danger of possible accidents. Along with the danger of big freight and passenger ships transport in the Adriatic Sea, smaller ship transport also presents significant danger, especially during tourist season. The consequences of smaller ship accidents are generally smaller, however the frequency of these accidents in the end is not negligible, especially concerning the loss of human lives.
Also, smaller vessels with their numerosity significantly endanger the environment, and eventually by disruption of bigger ship transport and causing their accidents can also cause catastrophic consequences with great human and material losses and environmental pollution. In the last few years certain measures have been taken in the sense of navigation safety increase in the Adriatic Sea, for instance the establishment of traffic separation system on some parts, mandatory reporting for certain types of ships, surveillance through automatic identification system, etc.
Considering that ship transport in the Croatian part of the Adriatic is constantly increasing, it is necessary to increase the capacity and type of preventive action.
One of the significant segments of sea transport security is also the meteorologic safety of navigation with all its elements, which will be discussed in further text.

## 1. TRANSPORT IN THE ADRIATIC

According to all relevant indicators transport in the Adriatic is in increase, regardless of economic, social and political situation in our area.
Passenger and vehicle transport on the lines in public coastal maritime transport in the Adriatic is generally mildly increasing with respect to year 2012. In 2013 the total of 11.350.111 passengers and 2.785.395 vehicles were transported which is increase of 1,7 per cent of passengers and 0,8 per cent of vehicles. Ferries transported 9.338.359 passengers and 2.785.359 vehicles in 26 lines which is increase of 2,1 per cent of passengers and 0,8 per cent of vehicles.

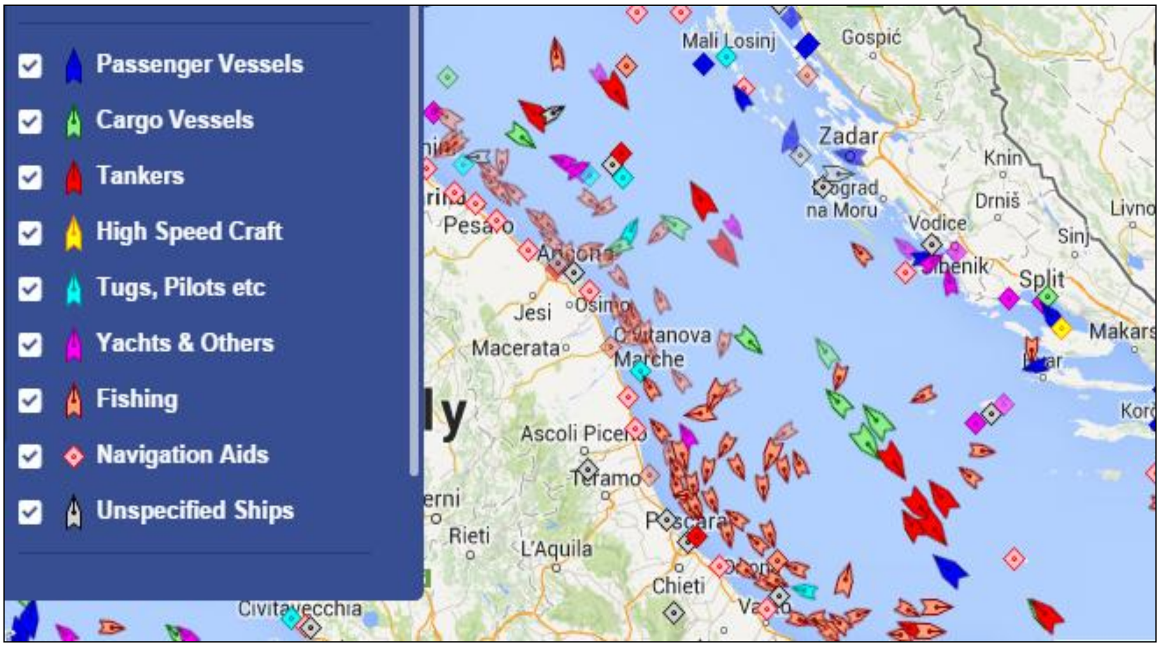



**Fig. 1.**
*Maritime transport and ship positions in the Adriatic sea*[1]

Ships that sail on the lines between Croatian and Italian ports, whose timetables are approved by the Agency for coastal maritime liner transport, in 2013 transported 507.078 passengers and 68.364 vehicles.

Motorways of the sea, a well branched network of waterways with regular and safe service of transport of goods is a well known project of the European Union with the scope of turning part of freight transport from roads to the sea, as part of an effort to establish sustainable transport system in Europe in the next thirty years. Jams on the European network of highways in the last three decades have acquired such proportions that they even threaten the competitiveness of the European economy. The answer was found in seamanship and in one of its oldest branches – short sea shipping, which the EU has upgraded into the program of the Motorways of the sea as a concept of traffic routes which consists of transport and logistic services and infrastructure, and as a mutual consequence they have a unique and complete transport door-to-door service system, with great flow of goods and passengers.

The Adriatic has been under scrutiny of European economy and institutions in the last years, firstly because of its strategic position. The growth of world trade, regardless of the crisis which has temporarily slowed it down, has especially been reflected in the Adriatic, resulting in increased need for transport services.

Short sea shipping has grown in the last decades in a very similar pace as road transport, but unlike the roads, the sea has not been congested and sea transport in general is still far away from maximum capacity of the European Union and surrounding countries. Additional advantage is the fact that the sea, unlike any other way, does not recquire special investment nor time for construction. Along with a very rugged coast and numerous ports, Europe has an additional advantage of branched network of canals and river flows that connect almost all parts of the continent.

This planning of sea transport development in the Adriatic will obviously rise all aspects navigation safety, as well as the meteorological navigation safety.

[1] *Source: https://www.marinetraffic.com/hr/*

## 2. MARITIME METEOROLOGY AS A FACTOR OF METEOROLOGICAL SAFETY OF NAVIGATION

Maritime meteorology is one of the important factors of navigation safety, as well as safe and economic earning in seamanship. Commercial seamanship has significantly decreased the dependancy on weather conditions, but the cost-effectiveness and regularity of navigation still depend on it, hence it is very important to quickly and efficiently deliver meteorological notifications to all interested users. At the same time, general (recreational) navigation is also increasing together with many other maritime activities which demand development of maritime meteorology in the sense of navigation safety. Maritime meteorology participates materially in safety, regularity and efficiency of marine transport by studying meteorological elements and phenomena from a viewpoint of influence on maritime technique and conditions of navigation. It elaborates and perfects the ways and shapes of meteorological safety and aid to navigation. The fact that bad weather is an important factor to the safety of marine transport is confirmed by the analysis of causes of maritime accidents, which in the area of the Adriatic Sea amounts to 20% with bigger and 40% with smaller ships and boats. Although the Adriatic Sea has a small closed gulf of the Mediterranean Sea, extraordinary and extremely dangerous maritime meteorological phenomena take place there.

Even though navigation along the eastern coast of the Adriatic, thanks to numerous islands and bays, has always been seen as relatively safe navigation compared to external mediterranean navigations, the remains of a number of ships from shipwrecks along our littoral area show that the safety on sea is a dubious term.

There is a certain number of meteorological phenomena which directly limit and endanger maritime transport thus earning our attention. Along with the wind, there are also sea conditions (waves height as an oceanographic element) and reduced visibility as a limitation factor in maritime navigation.

Maritime meteorology elaborates and perfects the ways and shapes of meteorological safety and aid to navigation. Weather conditions sometimes complicate and in other cases, in regular and timely estimation, help the navigation. Negligence weather conditions leads to failure of maritime tasks, and sometimes even accidents. Any type and size of vessel has its own navigational possibilities. Marine transport safety is endangered by phenomena and processes of great (sinoptical) – macro, and especially smaller mezo and micro proportions. Dangerous phenomena are mostly small proportion phenomena. Most phenomena and processes relevant to navigation are in areas of big gradients (discontinuities): fronts, various inversions and similar. In those relatively narrow areas of various proportions develop and strenghten gradients of most meteorological and oceanographic elements (temperature, pressure, air humidity, wind etc.). These processes manifest in energy accumulation which acts in different termodynamic processes – development of cloudy systems, rainfall, sea currents, waves. Phenomena which are not a consequence of gradients increase and manifestation of energy but precisely reverse processes – they are a consequence of smaller gradients (continuities) and energy as well as high stability; also influence maritime safety. These include fog and bad visibility.

Navigation safety, anchoring and mooring are endangered when certain critical meteorological conditions are exceeded, such as meteorological minimum for ships, ports and facilities, phases of navigation and participants in marine transport. Many of those minimums are not determined by direct meteorological measures nor obtained by special analysis, but are determined by estimation and experience. They are usually within the jurisdiction of port authorities or deck officers. Weather conditions (meteorological parameters) influence maritime transport by influencing:

- ships and phases of navigation,
- objects in maritime transport system and
- participants in maritime transport (active and passive).

Maritime meteorology has a significant economic role in navigation planning, which is reflected in the choice of seaway – routes and times of navigation. It is evident that the effects of meteorological navigation are most significant during winter months, when limitations occur or cancelling of navigation in certain areas.

## 3. MARITIME METEOROLOGICAL SERVICE

World Meteorological Organization - WMO and International Maritime Organization - IMO under United nations, according to International convention for the Safety of Life at Sea, have set and developed standards and recommendations for work for this branch of meteorology, and gathered them in several documents. WMO demands from their members to satisfy the requirements set in WMO document *Techincal rules*, which primarily means establishment and maintenance of observing stations on land and sea of certain type and structure, marine meteorology service, different communication systems, education of staff, research application and development of technique and technology. WMO and IMO act within the system World Weather Watch - WWW. Assignement of Marine Meteorology Service - MMS gathers necessary climatological data, analysis, forecasts and warnings for atmosphere and world waters.

Meteorology service for marine transport is organized by National Meteorological Centre - NMC. It is National Meteorological Institute in Croatia, consisting of Marine meteorological center in Split, meteorological office in Rijeka and meteorological stations. Direct coordination with ships, which are included in system of voluntery observation of meteorological elements, is held by appropriate serveces (and porth Authority) respectively Port meteorological Officer - PMO.

National meteorological headquarters receives meteorological data (through satellites, computer networks, etc.) and monitors and forecasts meteorological conditions significant for marine transport. If it is necessary, they issue special warnings of dangerous meteorological phenomena over great area for which they are in charge, prepare analysis and forecasts of ground and high altitude meteorological maps and forward them to other meteorological services or offices and coastal radio stations (Rijeka - radio, Split - radio and Dubrovnik - radio.)

## 4. METEOROLOGICAL WARNINGS

Warnings for dangerous meteorological phenomena are important for safety of all types of boats. Warnings are part of regular meteorological newsletter which is issued three times a day in Split by Marine meteorological center. Regular warnings are issued in accordance with dangerous meteorological phenomena in regular periods, but emergency warnings can also be issued when dangerous meteorological phenomena are noticed.
Warnings are issued for these meteorological elements:

- strong and stormy winds,
- state of the sea,
- reduced visibility (fog) and
- possible thunderstorms.

Flow system and in this regard winds on Adriatic depend, primarily, of distribution of air pressure on European land, Adriatic and Mediterranean sea. Wind speed depends on gradient of air pressure. Gradient is significantly increasing on Adriatic, and in this regard wind speed, because of the orographic effect, insular and coastal canals and slash on mountain (Dinaridi) line along the coast which divides land and the Adriatic in two very different and opposite types. Most of the year there are dominant winds of south and north directions in the Adriatic (jugo and bora).
Wind strength, according to its effects, is determined by Beaufort scale (table 1.), while sea state is by Douglas scale (table 2.).

**Table 1.**
*Beaufort scale*[2]

| Beafort number (Bf) | WMO | Wind Speed (kn) |
|---|---|---|
| **0** | Calm | < 1 |
| **1** | Light Air | 1-3 |
| **2** | Light Breeze | 4-6 |
| **3** | Gentle Breeze | 7-10 |
| **4** | Moderate Brezee | 11-16 |
| **5** | Fresh Breeze | 17-21 |
| **6** | Strong Breeze | 22-27 |
| **7** | Near Breeze | 28-33 |
| **8** | Gale | 34-40 |
| **9** | Strong Gale | 41-47 |
| **10** | Storm | 48-55 |
| **11** | Violent Storm | 56-63 |
| **12** | Hurricane | >64 |

**Table 2**.
*Douglas scale*[3]

| Sea state | WMO | Waves height (m) |
|---|---|---|
| **0** | Calm (glass) | 0 |
| **1** | Calm (rippled) | 0-0.1 |
| **2** | Smooth | 0.1-0.5 |
| **3** | Slight | 0.5-1.25 |
| **4** | Moderate | 1.25-2.5 |
| **5** | Rough | 2.5-4 |
| **6** | Very rough | 4-6 |
| **7** | High | 6-9 |
| **8** | Very high | 9-14 |
| **9** | Phenomenal | > 14 |

For ships that sail through open sea and large boats, in accordance with the WMO (*World Meteorological Organization*) criteria, warnings are issued for the sea state of 5 (more wavy) and more. These warnings refer to meteorological reports that are broadcasted through coast radio stations Rijeka Radio, Split Radio i Dubrovnik Radio.

## 5. METEOROLOGICAL ELEMENTS AND PHENOMENA

### 5.1 Wind

Airflow is foundation of weather processes, and wind as a horizontal component of airflow is significant meteorological and climatic element. Cause of every circulation are termical differences and differences in air pressure of neighboring areas, or natural aspiration of atmosphere for balancing air pressure. Established circulation is followed by characteristic changes of other meteorological values.
Wind speed depends on gradient of air pressure. Gradient is considerably increasing in the Adriatic, and thus wind speed because of orographic effect, island and coastal canals and cuts on mountain (Dinarides) line along the coast, which divides land and the Adriatic in two very different and opposite types. Most of the year winds of south and north direction dominate on the Adriatic. Southeast and south winds dominate in cold period of the year. South wind usually blows on the front side of mediterranean or adriatic cyclone and belongs among most frequent directions in the Adriatic, both on open seas and along the coast. Mostly moderate and strong, but sometimes stormy and often longlasting wind with storm impact (≥40 m/s), which can be seen in table 3.

**Table 3**.

[2] *Source: World Meteorological Organization*
[3] *Source: World Meteorological Organization*

*Maximum wind impacts from some meteorological stations along the Adriatic in the period from the year 2001 to 2013* [4]

| MAXIMUM WIND GUST (m) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **station** | **Rijeka** | **Bridge Krk** | **Maslenica bridge** | **Šibenik** | **Split** | **Lastovo** | **Makarska** | **Dubrovnik** |
| **direction** | S | NE | NNE | NNE | NNE | NNW | ENE | NE |
| **speed (m/s)** | 42,0 | 58,9 | 69,0 | 32,3 | 48,5 | 43,0 | 59,0 | 34,4 |

### 5.2. Visibility

Among most important elements of meteorological report for boats and seamen on sea, along with the wind and waves data, there is information and warning about decreased visibility as a critical factor of navigation safety. That particularly comes to the fore when visibility is decreased because of the haze (visibility from 1 to 10 km), fog (visibility < 1 km) and heavy rain with rain curtain. Horizontal (meteorological) visibility is defined as the smallest visibility in circle of 360˚ at the altitude of ≈ 10 m above ground.
Navigation in fog conditions and at limited visibility is eased by turning on appropriate lights, in the sense of „being seen“ and giving sound signals. Radar uses also facilitate navigation, which is particularly true for coastal areas and areas with high intense traffic. Because of that lighthouses, weather on certain islands, rocks or as port lights, facilitate navigation during low visibility. Significance of low visibility can be presented also with knowledge that ships stopping is up to ½ nautical mile, and sometimes more. This is also true for navigation direction changing.
It should be pointed out that fog is the most often and lasts longest on the west coast of Istria, so the area from Kopar to Pula has significantly larger number of days with fog than the rest of the Adriatic.

**Table 4**.
*Average annual number of days with fog occurence from some meteorological station along the Adriatic from the year 2000 to 2013*[5]

| NUMBER OF DAYS WITH FOG | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **station** | **Pula** | **Rijeka** | **Senj** | **Zadar** | **Šibenik** | **Split** | **Palagruža** | **Dubrovnik** |
| **number of days** | 4,6 | 4,7 | 1,9 | 0,8 | 1,3 | 0,5 | 3,5 | 0,1 |

### 5.3. Sea state

The main and most frequent cause of wave movement on sea surface is definitely the wind: its direction, speed and duration, then size of wind path over which blows and topography of sea bottom (sea depth). In the wind zone sea surface has very complex and always new appearance that is constantly changing. Short or long wind duration and its small or large intensity causes certain wave movement, depending on wind path on sea. Regarding that, every small and closed sea can have its own special standard on wave height, depending of many other conditional factors of that sea.
Surface waves generated by wind cause numerous hazards for people, ships and navigation safety, and many other activities related to sea. It can be expected in the Adriatic when stormy south and north wind is blowing in areas that have large wind path.
The Adriatic is a relatively small and closed sea without long enough wind path, where waves of ocean height can't appear (H > 30 m). However, with longlasting stormy south wind on the Adriatic there can be high waves (H > 10 m) which represent grave danger for marine traffic. South (SE, SSE i S directions) and north (NNW, N i NE directions) make the highest waves on the Adriatic. Waves that appear from south wind are highest on the open sea of north and middle Adriatic. Certain waves with stormy south wind reach up to 8-10 and more meters, and with north wind wave height are smaller and are 5-7 m. Although north wind reaches stormy strength, because of the smaller wind path waves are lower, shorter and steeper. The highest measured wave in the Adriatic was 10.8m high and is registered on south wind.

[4] *Source: Meteorological and Hydrological Institute of Croatia*
[5] *Source:Meteorological and Hidrological Institute of Croatia*

## 6. CONCLUSION AND RECOMMENDATIONS

In modern marine traffic it is very important to bring the boat from one port to another, with necessary concern for people, for travel time and for loaded cargo. Among other factors that make that journey safe, good meteorological preparation of the trip is necessary, and so is tracking of development of weather condition during navigation. Very well organized meteorological service in domestic scale and worlwide helps with that.
From descriptions of marine and meteorological phenomena it can be concluded that Adriatic, although small and closed sea, is sometimes full of surprises and hazards for people, boats and navigation generally. Hazard is brought and caused by stormy winds generating high and dangerous waves on this area up to even 10.8 m. Although the fog on Adriatic is rare phenomenon with average annual appearance from 1 to 5 days, depending on area presents grave marine danger.
Marine meteorological center in Split is responsible for meteorological navigation safety in Croatia, which is a part of State hidrometeorological institute of Croatia. Marine meteorological center Split transmits marine meteorological reports to boats by Croatian coastal stations (CRS Rijeka, CRS Split, CRS Dubrovnik).
The fact that expert services of Marine meteorological center Split, together with related services from company for maintaining of waterways „Plovput“ Split, are included in constant work of Commission for navigation safety by the Department of sea, tourism, transportation and development of Republic of Croatia says a lot about signifiance and perspectivity of Marine meteorological center Split, and all to increase navigation safety on the Adriatic east coast. Also, teams of experts from the field of marine meteorology are permanent members of commission for investigation of marine accidents by the same Department.
We should continue the practice of involving marine meteorological service and marine meteorologists in VTS – Vessel Traffic Service system, in sense of increasing meteorological navigation safety.

**REFERENCES**

Gelo B.,: Meteorološko osiguranje plovidbe morem i unutrašnjim vodama, Rijeka, 1992, Zbornik FPS

Gelo B.,: Opća i prometna meteorologija, Zagreb, 2000.

Gelo, B.: Mezomodeli atmosfere i sigurnost prometa, Suvremeni promet 10, 1988.

GMDSS Manual, 2009 Edition, IMO, 2009

Hodžić, M.: Ekstremne pomorsko meteorološke opasnosti na hrvatskom Jadranu, Jadranska meteorologija, LIV – 13

Hodžić, M.: Modro i plavo, zbornik, Matica hrvatska, Kaštela, 2007.

http://www.agencija-zolpp.hr/tabid/1534/articleType/ArticleView/articleId/1157/PROMET-PUTNIKA-I-VOZILA-U-2013-GODINI.aspx

http://www.zadarskilist.hr/clanci/26012011/jadran---morska-autocesta

Klimatološki atlas Jadranskog mora (1960.- 1990.), Državni hidrometeorološki zavod Hrvatske, Split, 2000.

Kulenović, Z.: Metodologija istraživačkog rada, Pomorski fakultet Sveučilišta u Splitu, Split, 2005.

Peljar za male brodove, HHI, 2000.

Penzar, B. And all.: Meteorologija za korisnike, Školska knjiga, Zagreb,1996.

Pomorski zakonik, NN, 118/2004.

Popović, R., Kulović, M., Šore.: Meteorological warnings in the safety of navigation system, The 4th International Maritime Scinence Conference – IMSC 2012, Split

Popović, R., Kulović, M.: Meteorological safety of navigation, 15th International Conference on Transport Science, 28th May 2012, Portorož, Slovenia

Popović, R.: *Model studije vjetrovne klima za pomorske gradnje*, The 3rd International Maritime Science Cobference – IMSC 2011, May, 21th, 2011

Simović, A., *Pomorska meteorologija,* ŠK, 1978.

SOLAS, Consolidated edition 2009, IMO, 2009

Štefančić, G.: Metode znanstvenog i istraživačkog rada, Fakultet prometnih znanosti u Zagrebu, Zagreb, 2004.

Vučetić, M., Vučetić, V.: *Meteorologija za nautičare,* Fabra, Zagreb, 2002.

WMO, 1981. *Manual on Marine Meteorological Services* (MMS-558)

WMO, 1981. *Manual on Marine Meteorological Services* (MMS-558)

Zelenika, R.: Metodologija i tehnologija izrade znanstvenog i stručnog dijela, Ekonomski fakultet u Rijeci, Rijeka, 2000
.

# ROAD SAFETY MANAGEMENT AT THE NATIONAL LEVEL: THE CASE STUDY OF THE REPUBLIC OF SRPSKA AND SERBIA

**Krsto LIPOVAC[1], Alan ROSS[2], Milan TESIC[3]**
[1] *The Academy of Criminal and Police Studies, Cara Dusana 196, Zemun, Serbia,*
[2] *PhD Road Safety Management Adviser, Newcastle upon Tyne, United Kingdom,*
[3] *Road Traffic Safety Agency, Zmaj Jovina 18, Banja Luka, Republic of Srpska,*

**Abstract:** Public sector represents the stakeholders at the national level. Taking into account the importance of their position, it is essential that each individual is competent in its jurisdiction. Public sector need to provide efficient and safe mobility of citizens. In order to achieve impressive results, it is necessary to strengthen the capacity. One of the steps of the road safety management is strengthening the profession capacity, i.e. strengthening road safety by raising the level of professionalism, creating the preconditions for development of strategies, action plans, etc.. Thus the first and concrete results can be expected only after the profession has risen above personal and political interests. This will allow for the understanding and expertise of the public sector in solving the problem of unsafe roads. This paper presents the way in which road safety is managed at the national level in Serbia and the Republic of Srpska. This encourages a further evolution of the concept of road safety based on science and scientific research, breaking away at the same time from the traditional unprofessional and ad hoc road safety management.

**Keywords**: professionalism, profession, capacity strengthening

## 1. Introduction

The issue of road safety has been changing throughout the history, while theories concerning the causes of road safety problems and the ways in which to solve these issues have also been evolving. First phases of the road safety problem evolvement are characterized by confusion and negligence of the issue, which did not require a particular profession or road safety professionals (Lipovac et al. 2014). However, the tables have turned in terms of road safety globally. The last phases of the road safety development (the fourth stage of the road safety management and the fifth phase of the global governance) have been based on the road safety profession, professionalism and professionals.

### 1.1. Literature Review

The importance of the road safety management (hereinafter referred to as: RS) at the national and local level is emphasized in the Global Plan for the Decade of Action for Road Safety 2010-2020 (WHO, 2010). All activities are classified into five pillars. The first pillar - organization and management of RS, includes strengthening of the RS profession and the professionalism, which is a sufficient evidence of the importance of this systemic solution aimed at improving RS.

Lipovac et al. (2010a, 2010b) have worked out the duties of local communities in the process of implementation of the Road Safety Law (RSL). However, in carrying out these responsibilities, in all units of local governments, there is a problem of a low level activism. The situation is even worse in those local government units where activism is present among unprofessional RS practitioners (Lipovac et al. 2014). The consequence of such a situation is unprofessional, ad hoc "handling" of RS, the impossibility of defining efficient, high quality measures, the lack of monitoring of effects of implemented measures, spending of modest financial resources and other resources on countermeasures providing no effects or desired effects, etc..

Various concepts of road safety management have been implemented world-wide. The results have been almost identical: only coordinated activities of stakeholders can contribute to a higher level of road safety (Bekefi, T. 2006; Bourgondien, M. 2012 i CARRS, 2011). FHWA (2006) developed its first program titled *"Strategic Highway Safety Plans: A Champion's Guide to Saving Lives"* which gives a precise definition of the multidisciplinary dependence among stakeholders, the importance of their cooperation and communication in all segments, the role of leadership in RS, with due priority given to the organizational structure in charge of road safety management. Also, DaCoTa 2012 is of opinion that RS is a shared responsibility in a complex multisectoral system and that organizational management requires the implementation of ISO standards in this field.

Great Britain has a long history of RS education development. Programs have been first intended for school children, but later on, they were expanded to the whole community focusing in the first place on the public sector. Bradbury, et al. (2008) are of opinion that it is necessary to implement such programs in low and middle development countries, as these countries account for 85% of fatalities. Klein, R. (2009) also indicates the importance of the multisectoral approach to road safety.

[1] Corresponding author: milan.te.sicm@gmail.com

Numerous projects concerning institutional capacity building have been carried out world-wide, for example in Sweden, Norway, Ghana, Kenya, China, Australia, Mali, etc. SIDA, (2005) presented a project of institutional sustainability and capacity development in Sweden. The project was implemented in several countries with various government structures, level of socio-economic development, culture, mentality, organizational structure, etc. The effects of the project have been well noted and very positive.

Resolving weaknesses in RS management must be given the highest priority, since the current initiatives are not sufficient for the achievement of sustainable and systemic changes (Bliss et al. 2009). In order to achieve better results, a big challenge lies in creating the political will, as well as global, regional and national leadership and distribution of resources necessary for the successful implementation of recommendations from the World Report.

Thomas, R. (2007) points out the responsibility of the state in all RS aspects. The state is the one that plans, designs, builds, leads and maintains a large part of road infrastructure, adopts and applies all RS laws, prescribes procedures for driving training and obtaining of driving licenses and leads state programs aimed at encouraging safe behavior in traffic. Having in mind the scope of these responsibilities, the state employs thousands of road safety professionals, and therefore it must have the central role in the efforts intended for development of RS profession and professionalism.
Elvik R. at al. 2008 are of opinion that a successful road safety management, from the point of view of its goals and objectives, will be facilitated if the following conditions have been met:

1) The lead authority should strongly support the objectives and be highly committed to their implementation;
2) The objectives should represent a challenge, but must be attainable in principle;
3) There should not be many objectives set, having in mind the political instruments available for their implementation;
4) The Agency whose task will be to choose the best way for the implementation of objectives should have the authority for determining priorities using all available political instruments;
5) The authorized Agency should have sufficient funds for the implementation of all cost-effective road safety measures;
6) There should be a system for monitoring progress in the implementation of objectives and giving information on their performances to the ministries in charge (benchmarking) and
7) There should be incentives and secure sources of funds in order to guarantee the involvement of all stakeholders in charge of implementation of the objectives set.

### 1.2. Subject and objective of the paper

This paper gives an overview of two national activities aimed at improving the profession, knowledge, methods and procedures for the advancement of RS situations in the Republic of Srpska and Serbia.

The project whose aim is to help improve the profession and professionalism of employees, at the local level, titled „Strengthening capacities of the local self-government units in the field of road safety in the Republic of Srpska“, as from 2012, was launched in the Republic of Srpska. The project is of continuous character, and is intended for implementation in this area on a yearly basis.

The original form of implementation of the project in Serbia, with the task of improving the profession and professionalism in the field of RS, at the national level, has been missing. Other projects are implemented instead, having an indirect impact on the increase of the level of safety on the roads. One of the most important activities relating to the improvement of the profession among employees, who are thinking about road safety, eight hours a day, are the scientific and professional conferences, seminars and workshops whose goal is to improve the awareness of employees. The organization of the scientific conference named "Road safety in the local community” is among the most important national activities aimed at improving the knowledge in the field of RS. This activity is also of continuous character, and has been organized since 2007.

The primary goal is to promote good practice in the road safety management, and to raise awareness of the profession and professionals in this field. This model of professionalism development among the staff at the national and local level can be also applied in other scientific fields, which will contribute to a more efficient use of resources and better results.

## 2. The basis for a successful road safety management

The RS system is very complex, involving the participation of a large number of stakeholders. This is what makes the RS management so complicated. Moreover, there are skeptics who still dominate the world, and who also think that this is impossible Lipovac (2008), DaCoTa (2012). Global efforts have been made over the last decades in order to spread the awareness of the possibilities of RS management.

Today, the promotion of the possibilities offered by RS management has been facilitated, as developed countries accepted this idea and achieved significant effects and practical results.

**Managing a certain system or process means undertaking managing measures in order to bring closer the current state of the system (process) to the targeted (desired) state** (Lipovac, K. (2008).

According to Lipovac, 2008, a system is successfully managed when:
- The existing state is well known;
- Targeted-desired state has been defined and
- Best managing measures have been selected, designed and implemented which will make the existing state closer to the targeted-desired state.

Apart from the mentioned concept of RS management, it is very important to provide a permanent model of RS funding. However, it can be noticed that the funding system and the possibility of allocating funds depend exclusively on the political will and on professionals at the positions intended for transport, at the national and local level. The conlcusion that follows therefrom is that funding of RS is directly proportional to the level of profession and professionalism among the activists.

## 3. Strengthening the profession and professionalism in the Republic of Srpska

The Republic of Srpska has a good basis for RS improvement. The Road Safety Council and the Road Safety Agency have been formed. "*The Road Safety Strategy of the Republic of Srpska 2013-2023"* and *"The Road Safety Action Plan of the Republic of Srpska"* have been also adopted. There are around 160 fatalities annually on the roads in the Republic of Srpska, while over 3.200 people get injured. Due to road crashes, the economy of the Republic of Srpska annually bleeds away over 174 million of KM (around 90 million EUR, over 20% of the national gross domestic income), (The results of the research of the Faculty of Economic studies in Banja Luka). It is therefore necessary to undertake urgent activities aimed at reducing the loss of lives and mentioned economic costs.

The Road Safety Agency of the Republic of Srpska has recognized the importance of RS professionalism at the local level, identified the lack of professional staff, insufficient commitment to RS, which were eventually the result of the low level of RS knowledge and awareness thereof. Based on the priorities set by the Strategy for training of staff in local self-government units in the Republic of Srpska, the Ministry of the Government and Local Self-Government of the Republic of Srpska and the Road Safety Agency of the Republic of Srpska organized training courses on the scope and type of RS problems, global trends, importance of strategic approach to RS management, etc.

### 3.1. Methodology

The training was carried out in 4 centers: Banja Luka, Bijeljina, Istocno Sarajevo and Trebinje, encompassing the municipalities that gravitate towards these centers. The first cycle of the training was conducted from May till September 2013. The seminars lasted for 6 hours (from 10.00 till 16.00 hours), according to the specially designed program. The subject of the first cycle of the training was: "The strategy and program of road safety in local self-government units". The aim of the training was to present the guidelines for the making of local RS strategies and RS programs, as a basis for a serious RS monitoring and management. Complete presentations were of interactive character and aimed at providing assistance to local communities when making the database necessary for the improvement of RS (road accident database, traffic offense database, etc.).

The second part of the training was conducted in September 2014. The training included the same area (4 centers) and the duration was limited to 8 hours per day. The subject of the second cycle of the training was: „Contemporary methods for road safety monitoring in local communities". Thus the participants have been made familiar with the importance and types of RS monitoring. The focus here was put on the making of the action plan with the financial framework for each local community, as well as on the making of the databases of significance for RS improvement (road accident database, traffic offense database, etc.).

The training consisted of two parts: theory and practical-interactive approach.

Both during the first cycle and the second cycle of the workshops, the speakers included RS professors and professionals, experts from local police units and local self-government units (hereinafter referred to as: LSU). The trainees of the seminar were heads of departments whose scope of work included RS tasks and assignments, and also officers of the Republic of Srpska LSU dealing with RS issues. Participants of the seminar have improved their practical RS knowledge and RS management at the local level.

### 3.2. Results of the trainings

The organization of seminars in LSU is of great importance for the RS management system building in the Republic of Srpska. The following objectives have been achieved through the organization of the seminars in question (Lipovac et al. 2014):

- Participants have improved their theory and practical RS related knowledge;
- LSU representatives have become familiar with the method of assessing the situation of the RS system in the local community (benchmarking, SWOT analysis, creation of databases);
- LSU representatives have acquired initial knowledge on the importance and possibilities of RS management;
- Participants have recognized great local capacities for the RS improvement;
- Activities related to the preparation of the local RS Strategy and Program in all LSU have been initiated and certain LSU have launched the activities concerning the creation of road accident database, and
- Exchange of experiences among LSU and identification of common problems have been also enabled.

The forthcoming period should see the continuation of training activities in LSU, in cooperation with the Ministry of Government and Local Self-Government and the Association of municipalities and cities. The number of training sessions should be increased and longer trainings (lasting several days) should be also planned. This is the only way to make basic preconditions for active inclusion of LSU into RS management process, reinforce the profession and recruit professionals in this field, and also achieve a permanent improvement of RS in local communities in the Republic of Srpska.

The main and essential goal of these workshops is the promotion of profession as a key to the easier and more efficient implementation of activities aimed at improving RS.

## 4. Strengthening the profession and professionalism in Serbia

The platform for the promotion of RS at the national level includes the following elements: management, strategy documents, horizontal coordination with all stakeholders from the RS systems and vertical coordination with local communities. RS management in Serbia got its "healthy" management with the creation of the Road Safety Agency. Furthermore, the National Assembly, as the highest legislative body, adopted the Bases of the national road safety strategy 2014. Other stages required for the systemic RS monitoring and improvement evolved into periodic, but continual activities in the form of international conferences, seminars, workshops, etc.

Strengthening RS professionalism and profession in Serbia through conferences, seminars, workshops, etc. represents the most important and effective way of mutual communication of local communities, as well as of other stakeholders from the system (police, automobile and motorcycle clubs, health system, education system, etc.). It is of particular importance to present best practices of certain local communities, from various RS segments (for example, repair of hazardous locations, improvement of children’ safety in school zones, presentation of results of various studies related to RS indicators, methods of raising awareness in local communities, creation of local databases important for RS management, etc.).

One of the national initiatives to improve RS at the local level is the International conference "Road safety in the local community," which is supported by the Ministry of Education and Science of the Republic of Serbia. This conference has been organized since 2007, solely for the purpose of establishing communication with local communities, sharing knowledge and reconsidering RS problems at the local level. In addition, this conference provides an exchange of experiences of various stakeholders that contribute to improving road safety, with due consideration of their positions within the system.

The exchange of contacts, knowledge, experience, problems among the participants of the conference help develop the profession and the professionalism through a diverse range of employees at the national and local level (representatives of ministries being directly and indirectly related to the RS system, professors, heads of police departments, heads of departments of spatial and urban planning, senior associates, students, and others). All the above mentioned participants give their contribution to the development of the RS profession, as their knowledge and awareness enable a facilitated implementation of activities intended for the road safety improvement.

## 5. Main recommendations for the strengthtening of road safety profession at the local level

The worst thing that can happen to a country, region or local community is the lack of understanding for resolving problems in RS (or in any other scientific discipline), for implementing best practices (experiences of other countries that led to outstanding results), etc. These problems appear exclusively in cases where experts, the science and experience are missing. Unfortunately, Serbia and the countries in the region have this problem which can be overcome by reinforcing RS institutional capacities and profession.

Trainings, seminars and workshops intended for the strengthening of capacities and profession make a very important and indispensable factor of advancement of any science. If the professionalism among the stakeholders who are implementing activities is missing, then it will be very difficult to expect positive results. Therefore, the academic and professional public (associations of citizens, alliances, etc.) should put efforts in providing „the right professionals at the right positions“.

Institutional development and strengthening of capacities help employ a significant number of people who perform their RS work every day and are committed to the improvement of RS on a *"full time"* basis. Thus the level of professionalism is being reinforced in important institutions in LSU, and RS related attitudes of the general public improved. This will contribute to a better dedication, better use of existing capacities, reinforcement of capacities and integrity of institutions and individuals, increase of the willingness to pay for RS, etc.

The cooperation of all RS stakeholders is of crucial importance for developing a multidisciplinary RS profession (DaCoTa 2012). Therefore, the Road Safety Agency should support the organizations, associations, alliances dealing with RS on a broad basis, whose central aim will be the institutional development, improvement of road safety profession and professionals (Thomas, R. (2007)). The participation of local authorities is of particular significance because of the roles the municipalities have in the organization of RS related tasks.
National and local authorities should also seek support from the private sector organizations, universities and professional associations whose common interest is to develop RS professionals. It is very desirable to have the participation of professional societies, educational and research institutions.

Organizations should lead the activities related to the RS profession, in various directions, (Thomas, R. (2007)), such as:

- **Promoting multidisciplinary road safety professionals** who understand and apply scientific RS principles,
- **Offering support (by lead associations**) in engaging and developing RS professionals,
- **Encouraging continual development and wider implementation of basic competencies in education, training and improvement of road safety professionals,** who have mastered the scientific methods and are able to find solutions in the field of road safety, on a systemic level,
- **Promoting the work in the RS field,** as a special profession and an opportunity for a good career,
- **Persuading universities and other scientific-educational institutions of the value of creating partnerships for RS education**, their inclusion into organized transfer of knowledge, creation of own knowledge, traffic education and upbringing, professional training and education of numerous RS professionals, and
- **Supporting scientific RS studies aiming at developing RS theory and practice and attracting the best students and pupils to dedicate themselves to RS** (scholarships, donations, covering training costs, engagement of university professors and research centers in the implementation of RS studies and projects, etc.).

The abovementioned and other actions should help the representatives from the public and private sector understand that profession-oriented and committed RS professionals are a must. This is of essential importance for sustaining a long-term support for the development of such professionals, but also for the implementation of RS improvement plans.

The possibility of creating one or more specialized centers for professional training and education of RS professionals should be also explored. There is a significant number of RS professionals needed at the national level, which requires a sustainable development. Therefore, one should also explore the needs for and ways in which to organize various courses, seminars and other types of education, professional training and upgrading through existing programs or through the introduction of new programs dedicated to the improvement of RS profession. Given the similar working conditions and a simple understanding (there are no language barriers), it would be an optimal solution to create a regional **ROAD SAFETY CENTER** that would fulfill the needs of several countries in the region.

**References**

Bekefi, T. (2006). „The Global Road Safety Partnership and Leasons in Multisectoral Collaboration“, Coorporate Social Responsability Initiative Report No.6, Cambridge, MA: John, F. Kennedy School of goverment, Hardward University.

Bliss, T. and Breen, J. (2009). Country Guidelines for the Conduct of Road Safety Management Capacity Reviews and the Specifi cation of Lead Agency Reforms Investment Strategies and Safe System Projects. The World Bank Global Road Safety Facility. Washington DC.

Bradbury, A. and Quimby, A. (2008). Community road safety education: an international perspective. Proceedings of the Institution of Civil Engineers. Volume 161, 2, 137-143.

Bourgondien, M. (2012). Institutional Capacity Building to the Transport/Road Sector, Kenya.

Centre for Accident Research & Road Safety – Queensland (2011). Building Road Safety Capacity. Queensland University of Technology, Kelvin Grove Campus.

DaCoTA (2012) Road Safety Management, Deliverable 4.8p of the EC FP7 project DaCoTA.

Rune, E. (2008). Road safety management by objectives: A critical analysis of the Norwegian approach. Accident Analysis and Prevention 40, 1115–1122.

Klein, R. (2009). Building capacity for road safety data systems. 4th IRTAD CONFERENCE, 16-17 September, 2009, Seoul, Korea, 33-40.

Lipovac, K., Jovanović, D. i Nešić, M. (2010a). Obaveze jedinica lokalne samouprave u sprovođenju odredbi ZoBSa, V stručni seminar: Uloga lokalne zajednice u bezbednosti saobraćaja, (Zbornik radova) str. 43-66, Kovačica.

Lipovac, K. (2009). Obaveze beogradskih opština u sprovođenju ZoBSa, Seminar: Novi Zakon o bezbednosti saobraćaja: velika šansa i izazov za lokalne zajednice - grad Beograd, (Zbornik radova) str.19-42, Beograd.

Lipovac, K., Nešić, M., Jovanović, D. (2010b): Unapređenje bezbednosti saobraćaja na putevima gradova i opština prema novom zakonu o bezbednosti saobraćaja, Bezbednost, Vol. 52, br. 1, 2010, str. 9-40.

Lipovac, K., Vujanić, M. i Jovanović, D. (2014). Osnove bezbednosti saobraćaja, Kriminalističko-policijska akademija, Beograd.

Norwegian Agency for Development Cooperation (1991). “Technical Assistance to Roads Department Botswana, Project Review.

Norwegian Agency for Development Cooperation (2007). “The experiences gained from 30 years of development cooperation in Botswana.

Norwegian Centre for Transport Research (2002). “Review of the Tanzania Road Sector Program”, Working Paper TR/1111/2002

PIARC Commission on Technological Exchanges and Development (2010). The Concept of Institutional Capacity Building and Review of Road Sector Projects.

Strategic Highway Safety Plans: A Champion’s Guide to Saving Lives. Federal Highway Administration, Washington, D.C., 2006.

Swedish International Development Cooperation Agency (2005). “Institutional Sustainability and Capacity Development within SIDA financed Road safety Projects”, SIDA Evaluation 05/29.

The Royal Society for the Prevention of Accident (2003). Strategic Guidance for Road Safety Professionals.

Thomas R. Menzies, Jr. (2007). Building the Road Safety Profession in the Public Sector. Special Report 289. *Transportation research board, Washington DC.*

World Health Organisations (2013). Global Status Report on Road Safety: Supporting a decade of action.

World Health Organisations (2009). Capacity building for preventing injuries and violence: Strategic Plan 2009-2013.

Nova Larousse enciklopedija (1999). JRJ, Zemun.

www.who.int

www.bslz.org

# THE IMPACT OF DIFFERENT LEGISLATIONS REGARDING DRIVING LICENCE DURATION ON TRAFFIC SAFETY IN SELECTED EUROPEAN COUNTRIES

**Marina Zanne**
*University of Ljubljana, Faculty of Maritime Studies and Transport, Pot pomorščakov 4, SI – 6320 Portorož, Slovenia*
*marina.zanne@fpp.uni-lj.si*

**Abstract:** Driving license entitles the holder to operate the vehicle of relevant category on any road. The way of obtaining driving permit is rather consistent throughout Europe, however each Member State can decide upon duration of the driving license and upon its renewal process.
In the present paper, the author first checks the applicable legislation and regulations in selected European countries in regards to driving license duration, then she examines participation of elderly drivers in traffic accidents in Slovenia as well as in other European countries and finally gives an overview of the impact of different approaches to driving license duration on road safety.

**Keywords**: driving license, driving license legislation, non-professional drivers, senior citizens, elderly drivers, comparative analysis, traffic safety

## 1. Introduction

Most individuals travel because they wish to benefit from the social, recreational, educational, employment and other opportunities which become accessible with movement (Button, 1996). Travelling needs and the ways of meeting these needs change during person's life, but in European Union as well as in Slovenia, they are very often satisfied by the use of a personal car as can be seen in Figure 1. This happens for several reasons, for example the feeling of independence, more flexibility in regards to public passenger transport or the lack of public transport.

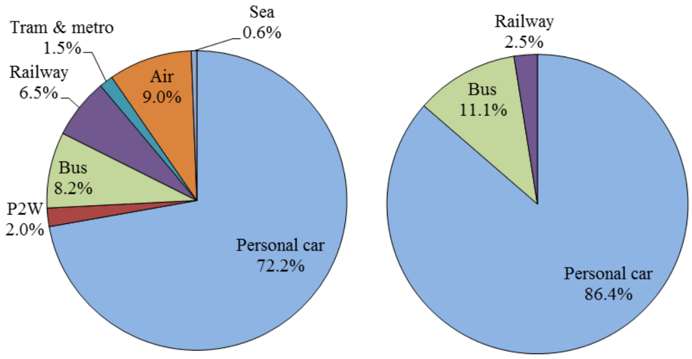



**Fig. 1.**
*Modal split of passenger transport in EU-27 and Slovenia in 2012*
*Source:* (EC, 2014b)

Citizens need to obtain a valid driving license in order to be able to use the personal car. The procedure of obtaining the driving license is nowadays pretty much harmonized throughout the European countries, but the validity of the licence differs. As the population ages, the percentage of senior citizens and elderly drivers increases. In addition the lifestyle of senior citizens is changing, thus it seems reasonable to expect that their share in traffic flows will increase.
Studies show, that driving licence rules have an impact on road safety and in this paper the author examines how different validity periods and renewal processes across European Union influence the involvement of senior citizens in road traffic accidents.

## 2. European legislation on driving licenses

European legislation on driving licences has a direct impact on nearly all Europeans as an estimated 60% of the Union's population holds a valid driving licence, that's around 300 million citizens (EC, 2013a).
On 29th of July 1991, the Council of Transport Ministers adopted Directive 91/439/EEC on driving licences. The directive entered into force on 1st of July 1996. With several amendments it still represents a major step towards harmonisation and facilitation of driver licensing procedures in EU. Another big step towards harmonisation of drivers' licences was done in 2006 with the Directive 2006/126/EC as there were more than 110 different driving licence models with different entitlements and validity periods circulating in the Member States. Provisions from this Directive came into force on 19th of January 2013.
In general, a candidate has to pass a theoretical and a practical test in order to obtain a driving licence. The minimum age to obtain a driving licence of category B is 18 years in majority of European Member States.

After having obtained a driving licence a novice driver can be a subject to a probationary licensing system, which can last for several years or there are no restrictions or special requirements. The procedure of obtaining driving licence is thus pretty much harmonised throughout Member States, however, there are four different systems regarding the renewal of driving licence in the European countries, namely (Siren A. et al., 2013):

- Licenses of unlimited validity in Austria, Belgium, Bulgaria, France, Germany, and Poland;
- Licenses that generally do not require a medical examination (Sweden) to be renewed but only an administrative procedure and/or a self-report of medical conditions (UK);
- Licenses that require a medical check by (at least) a physician to be renewed with the age of first assessment at a certain age:
  - at 50 (Italy, Portugal), 60 (Czech Republic, Luxembourg), 65 (Greece, Slovakia) or 70 years (Cyprus, Denmark, Finland, Ireland, Malta, The Netherlands, Slovenia;
- Licenses that have to be renewed every ten years including a medical examination for all age groups (Romania) with increasing frequency for increasing age at 40 (Hungary), 55 (Lithuania), 60 (Latvia), 65 (Estonia, Spain).

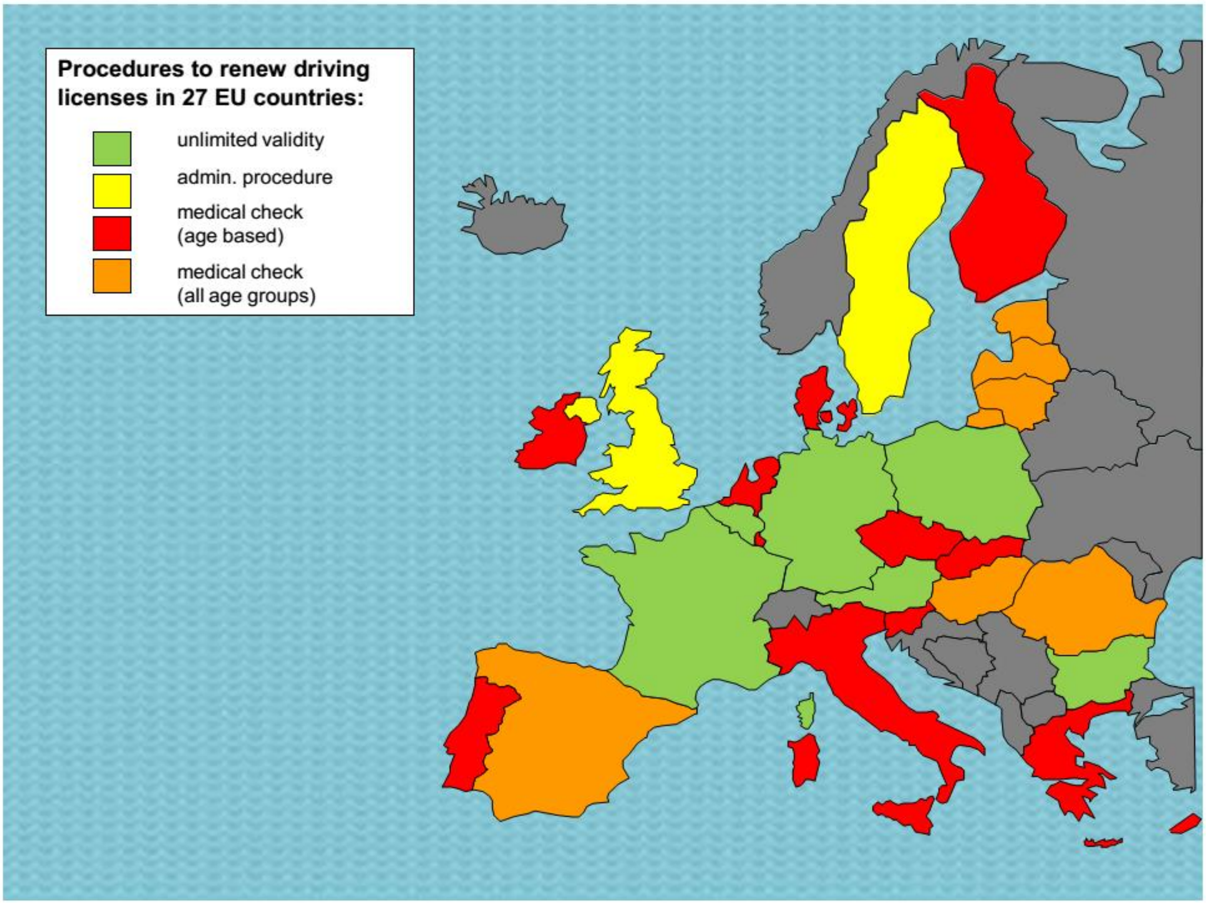



**Fig. 2.**
*Types of driving licenses in EU*
*Source: Siren A. et al., 2013*

**3. The analysis**

The “older driver problem” was recognised as a scientific and social issue the late 1960s and early 1970s (OECD, 2001), however there are not many available studies on the influence of driving licence renewal systems on road safety, but from those available it is possible to tell, that age-based mandatory assessment programmes targeting older drivers are unlikely to produce safety benefits and may have counter-productive results (see eg. Mitchell, 2008; Langford and Koppel, 2006; Langford et al., 2004; Hakamies-Blomqvist et al., 1996). However, a large number of studies have found increased accidents rates among drivers with several types of medical conditions (Elvik et al, 2009).
Approximately 1.24 million people die (fatalities within 30 days) every year on world’s roads, and another 20 to 50 million sustain non-fatal injuries as a result of road traffic crashes. Road traffic injuries are estimated to be the eight leading cause of death globally and they are the leading cause of death if young people aged 15–29 years (WHO, 2013). In European Union 28,126 persons were killed in road accidents in 2012, which was 8.3 % fewer than in 2011 (EC, 2014b).

Elderly people (> 65 years old) are vulnerable road users (Broughton et al., 2012a; Broughton et al., 2012b; Arrive Alive, 2014):

- In 2010, more than 6,500 elderly people died in road traffic accidents in 24 European countries;
- The number of elderly people who died in the EU-19 countries fell by 29,6% between 2001 and 2010 while the number of all fatalities on European roads fell by 42,3% in the same period;
- Every fifth person killed on roads in Europe is aged 65 or over (more precisely 21.8%) - it is estimated that by 2050 one death out of three will be an elderly person, if their safety level does not improve.
- Among the larger countries, the proportion of elderly fatalities ranged between 17% in Poland and 29% in the Netherlands;
- Among the larger countries, the percentage of elderly fatalities who were pedestrians is greatest in Romania (62%) and least in the Netherlands (14%);
- The proportion of elderly fatalities who were car drivers ranged between 6% in Romania and Slovakia and 50% in Ireland.

The percentage of older citizens in Europe's population is constantly increasing. In fact, the number of people over 65 was 17.5% in 2010 and 18.2% in 2013 (Eurostat, 2014) and is expected to be 24% in 2030. This means that the share of citizens over 65 years will range between 10.4% and 37.3%, while back in 2008, this range was between 9.1 % and 26.8%. The median age of European population in 2030 is projected to be between 34.2 years and 57.0 years, while in 2008 the range was between 32.9 years and 47.8 years (Giannakouris, 2010).

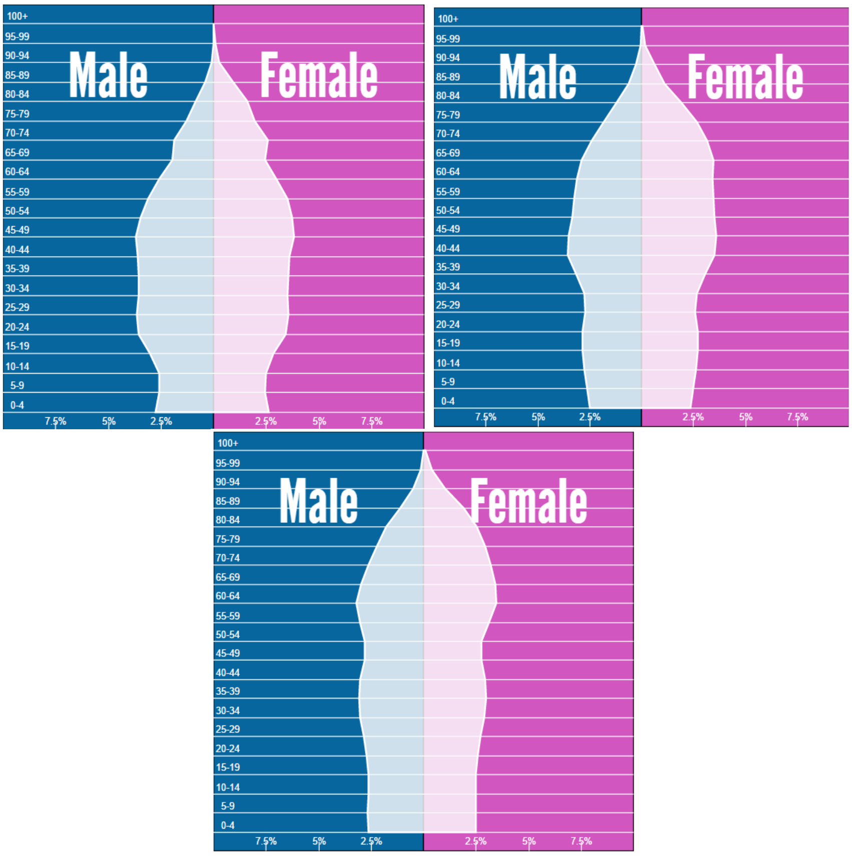



**Fig. 3.**
**Population structure in Europe in 2010, 2030 and 2050**
**Source: Europe, 2012**

Although many senior citizens are in good health, ageing often brings limitations and difficulties that can affect traffic safety**.** In fact, the road safety of older road users is to a large extent determined by two factors: functional limitations and their physical vulnerability.

**Table 1.**
*Relative risk of accidents for drivers with diseases in relation to relative risk of drivers without disease*

| **Disease** | **Relative risk** |
|---|---|
| *Diabetes mellitius* | *1.56* |
| - Epilepsy/other seizures | 1.84 |
| *Mental disorders* | *1.72* |
| - Serious mental disorders | 2.01 |
| - Serious disorders related to ageing (eg. Alzheimer, dementia) | 1.45 |
| *Pharmaceuticals and psychoactive substances* | *1.58* |
| - Misuse of pharmaceuticals/drugs | 1.96 |
| - The use of cyclic antidepressants | 1.42 |
| - Prescribed use of pharmaceuticals/drugs | 1.49 |
| *Coronary diseases* | *1.23* |
| - Angina pectoris | 1.52 |
| - Cardiac arrhythmia | 1.27 |
| *Visual impairments (all types)* | *1.09* |
| - Reduced visual acuity (<25%) | 1.24 |
| *Hearing impairments* | *1.19* |
| *Arthritis/Motor disorders* | *1.17* |

*Source: Elvik et al., 2009*

Elderly citizens are more likely to hurt themselves than to put others at risk as they are weaker and more fragile and thus have less chances to survive and injury than young people. In fact, the fatality rate for drivers over 75 is more than five times higher than the average (while the fatality rate of the 65-74 year olds is about twice that of the 30-64 year olds), and their injury rate is twice as high (EC, 2014a). People who drive less, and older people do drive less, have more accidents per kilometres driven. This can partly be explained by lack of routine but also by their driving patterns where they gain their exposure. People, who drive long distances usually drive many of their kilometres on motorways while people, who drive little, tend to drive more in urban areas and thus more complicated traffic situations where the probability for accidents is higher (EFPA, 2011). In fact, a comparison with the middle-agaed fatalities, reveals that there are fewer elderly fatalities on motorways and on rural roads as can be seen from Table 1.

**Table 2.**
*Characteristics of urban/rural fatal accidents in EU*

| **Urban area fatal accidents** | **Rural area fatal accidents** | **Motorway fatal accidents** |
|---|---|---|
| 37% are pedestrians | 9% are pedestrians | 8% are pedestrians |
| 10% are cyclists | 6% are cyclists | 0% are cyclists |
| 30% are elderly | 24% are elderly | 12% are elderly |
| 27% are women | 22% are women | 22% are women |

*Source: EC, 2013b*

The largest share of elderly drivers among all fatalities is in the countries that have administrative procedure of renewal of driving licence, which is a little unexpected in regards to the group of countries that have unlimited driving licence validity. But as can be seen from Table 4, the countries that have administrative procedure of renewal of the driving licence are the countries that have significantly better safety of elderly citizens than any other European country.

**Table 3.**
*The share of senior citizens' (>65 years old) fatalities in all fatalities in 2010*

| Licence renewal system | Country | Elderly fatalities | Elderly drivers as share of all fatalities | Elderly pedestrians as share of all fatalities | Elderly cyclists as share of all fatalities |
|---|---|---|---|---|---|
| Unlimited validity | Austria | 25.4% | 9.1% | 9.4% | 2.8% |
| | Belgium | 18.2% | 6.0% | 4.7% | 4.2% |
| | France | 19.2% | 7.3% | 6.1% | 1.3% |
| | Germany | 24.9% | 7.5% | 6.2% | 5.5% |
| | Poland | 17.2% | 2.2% | 9.5% | 2.8% |
| | | **20.4%** | **5.8%** | **7.2%** | **3.2%** |
| Administrative procedure | Sweden* | 25.7% | 11.6% | 5.4% | 2.8% |
| | UK | 19.9% | 6.6% | 7.4% | 1.0% |
| | | **20.8%** | **7.3%** | **7.1%** | **1.3%** |
| Age-based medical check | Czech Rep. | 21.4% | 4.9% | 10.1% | 2.8% |
| | Denmark | 26.3% | 5.5% | 7.4% | 5.0% |
| | Finland | 23.5% | 9.6% | 5.2% | 4.5% |
| | Greece | 21.3% | 4.7% | 7.9% | 0.4% |
| | Italy | 25.9% | 8.0% | 9.1% | 3.6% |
| | Ireland | 14.2% | 7.1% | 4.7% | 1.4% |
| | Netherlands* | 29.0% | 6.4% | 4.1% | 10.7% |
| | Portugal | 29.7% | 5.0% | 11.6% | 1.5% |
| | Slovakia | 12.9% | 0.8% | 7.8% | 2.2% |
| | Slovenia | 21.7% | 5.9% | 8.0% | 3.7% |
| | | **24.5%** | **6.5%** | **8.6%** | **3.3%** |
| Medical check for all age groups | Estonia | 18.4% | 4.0% | 8.1% | 2.0% |
| | Hungary | 20.1% | 3.0% | 9.3% | 4.2% |
| | Latvia | 20.2% | 2.8% | 10.5% | 2.2% |
| | Romania | 20.8% | 1.2% | 12.9% | 1.9% |
| | Spain | 21.3% | 6.2% | 8.5% | 0.6% |
| | | **20.8%** | **3.6%** | **10.4%** | **1.7%** |
| | **EU-24** | **21.8%** | **5.7%** | **8.2%** | **2.8%** |

*Note:* * *- data for 2009*
*Bulgaria, Lithuania, Malty and Cyprus are omitted*
*Source: Author, based on Broughton et al., 2012a; Broughton et al., 2012b; Pace et al., 2012; ERSO, 2014*

The elderly pedestrians and elderly cyclists are overrepresented in the structure of road victims. While the share of elderly cyclists in the total number of fatalities has a lot to do with the country's tradition of using bicycles (eg. the Netherlands or Denmark), the share of pedestrian fatalities can be a consequence of purchasing power of individuals, availability of public transportation or driving licence renewal regime. The countries that have a mandatory medical check for all age groups have the lowest share of elderly drivers' fatalities in the entire number of road fatalities, but the share of elderly pedestrian fatalities is by far the largest in this group of countries. The share of elderly pedestrian fatalities is especially high in the countries with low BDP, like Romania (5,800 EUR pc[1]/ 11,400 EUR pps[2]), Poland (9,200 EUR pc / 15,300 EUR pps) or Slovakia (12,100 EUR pc / 17,900 EUR pps).
According to the theory, all analysed countries have better safety records for entire population than for the senior citizens; in average 78 elderly citizens die in road accidents per million senior inhabitants in comparison to 62 victims per million inhabitants in the entire population. In all but two countries (Sweden and Ireland), the number of elderly pedestrian and cyclist fatalities is higher than the fatalities of elderly drivers; in fact, the number of dead pedestrians and cyclist is in average twice the number of deaths of elderly drivers.

[1] pc – GDP per capita (Eurostat, 2013)
[2] PPS - purchasing power standards (Eurostat, 2013)

**Table 4.**
*The structure of fatalities of senior citizens (>65 years old) and the structure of fatalities in entire population in 2010*

| Licence renewal system | Country | Fatalities per mio people | Elderly fatalities per mio | Driver | | Pedestrian | | Cyclists | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Elderly | Entire population | Elderly | Entire population | Elderly | Entire population |
| Unlimited validity | Austria | 66 | 95 | 36.0% | 41.3% | 37.0% | 17.8% | 11.0% | 5.8% |
| | Belgium | 77 | 82 | 33.0% | 38.3% | 26.0% | 12.6% | 23.0% | 8.3% |
| | France | 64 | 73 | 38.0% | 39.0% | 32.0% | 12.1% | 7.0% | 3.7% |
| | Germany | 45 | 54 | 30.0% | 36.8% | 25.0% | 13.0% | 22.0% | 10.4% |
| | Poland | 103 | 131 | 13.0% | 28.8% | 55.0% | 31.6% | 16.0% | 7.2% |
| | | **71** | **87** | **28.5%** | **35.4%** | **35.4%** | **18.6%** | **15.6%** | **7.0%** |
| Administrative procedure | Sweden* | 38 | 54 | 45.0% | 45.8% | 21.0% | 12.3% | 11.0% | 5.6% |
| | UK | 32 | 39 | 33.0% | 29.1% | 37.0% | 21.8% | 5.0% | 5.8% |
| | | **35** | **46.5** | **35.3%** | **31.7%** | **34.0%** | **20.4%** | **6.1%** | **5.7%** |
| Age-based medical check | Czech Rep. | 76 | 108 | 23.0% | 35.7% | 47.0% | 20.9% | 13.0% | 10.0% |
| | Denmark | 46 | 74 | 21.0% | 36.5% | 28.0% | 17.3% | 19.0% | 10.2% |
| | Finland | 51 | 70 | 41.0% | 43.0% | 22.0% | 12.9% | 19.0% | 9.6% |
| | Greece | 111 | 125 | 22.0% | 29.1% | 37.0% | 14.2% | 2.0% | 1.8% |
| | Italy | 68 | 87 | 31.0% | 31.9% | 35.0% | 15.0% | 14.0% | 6.4% |
| | Ireland | 47 | 59 | 50.0% | 37.3% | 33.0% | 20.8% | 10.0% | 2.4% |
| | Netherlands * | 39 | 74 | 22.0% | 32.5% | 14.0% | 9.8% | 37.0% | 5.0% |
| | Portugal | 88 | 145 | 17.0% | 23.8% | 39.0% | 20.8% | 5.0% | 3.5% |
| | Slovakia | 68 | 72 | 6.0% | 27.8% | 60.0% | 34.0% | 17.0% | 7.3% |
| | Slovenia | 67 | 89 | 27.0% | 31.9% | 37.0% | 18.8% | 17.0% | 12.3% |
| | | **66.1** | **90.3** | **26.4%** | **31.5%** | **34.9%** | **16.6%** | **13.6%** | **5.9%** |
| Medical check for all age groups | Estonia | 73 | 79 | 22.0% | 34.7% | 44.0% | 23.5% | 11.0% | 7.1% |
| | Hungary | 74 | 86 | 15.0% | 26.2% | 46.0% | 25.9% | 21.0% | 12.4% |
| | Latvia | 97 | 113 | 14.0% | 23.4% | 52.0% | 36.2% | 11.0% | 6.0% |
| | Romania | 111 | 154 | 6.0% | 20.4% | 62.0% | 36.5% | 9.0% | 7.7% |
| | Spain | 54 | 68 | 29.0% | 31.6% | 40.0% | 19.0% | 3.0% | 2.7% |
| | | **81.8** | **100** | **17.4%** | **26.2%** | **50.0%** | **27.6%** | **8.0%** | **6.1%** |
| | **EU-24** | **62** | **78** | **26.2%** | **32.1%** | **37.8%** | **19.9%** | **12.8%** | **6.4%** |

*Note: * - data for 2009*
*Bulgaria, Lithuania, Malty and Cyprus are omitted*
*Source: Author, based on Broughton et al., 2012a; Broughton et al., 2012b; Pace et al., 2012; ERSO, 2014*

## 5. Conclusion

In the European Union the personal car is still the dominant mode of personal transportation and with the increasing number of elderly population and the changes in the dynamics of life, the problem of senior citizens transportation is an important issue for the European Commission and national governments. Although the number of senior fatalities is decreasing across European Union, regardless of the renewal regime of driving licence, it is decreasing at lower pace than the number of fatalities in general, making the share of senior fatalities higher and thus making the problem of elderly drivers' safety a high priority issue.

The analysis yields inconclusive results in regards to the effect of different types of driving licence renewal systems on the safety of senior citizens, however if we remove the outliers, than it is possible to state, that mandatory medical checks, both age based and for all age groups, indicate a bit better safety performance in regards to senior drivers than unlimited validity of driving licence or administrative renewal procedure do. However, in the countries that have a mandatory medical check for a driving licence renewal the ratio between killed elderly pedestrians and cyclist on the one hand and killed senior drivers on the other is much higher than in the countries that do not limit driving licence validity or require just administrative procedure to obtain a new driving licence.

The available consolidated data does not allow further analysis of causers of road traffic accidents, so with such a data it is impossible to tell how different driving licence renewal systems affect the causation of accidents by senior drivers or senior citizens. This is an important subject to analyse, although senior citizens are more dangerous to themselves than to others. The further work will thus consist of identification of adequate countries and the thorough analysis of national data on traffic accidents in these countries. Only in this way it will be possible to make firmer statements about the influence of different renewal regimes for senior drivers on traffic safety of entire population.

**References**

Arrive Alive. (2014). *Road Safety and the Elderly / Older Road Users*. Retrieved September 10, 2014, from Arrive Alive: https://www.arrivealive.co.za/Road-Safety-and-the-Elderly-Older-Road-Users

Broughton J. et al. (2012a). *Basic Fact Sheet "Main Figures".* Brussels: European Road Safety Observatory. 20 p.

Broughton J. et al. (2012b). *Basic Fact Sheet "The elderly".* Brussels: European Road Safety Observatory. 20 p.

Button, K. J. (1996). *Transport economics* (2nd ed.). Cheltenham: Edward Elgar. 269 p.

EC. (2013a). New European driving licence for more security, safety and free movement. Brussels: European Commission

EC. (2013b). *Road safty vademecum - Road safety trends, statistics and challenges in the EU 2011-2012.* Brussels: European Commission – DG for Mobility and Transport.

EC. (2014a). *Elderly drivers*. Retrieved August 30, 2014, from Mobility and Transport - Road safety: http://ec.europa.eu/transport/road_safety/users/eldery-drivers/index_en.htm

EC. (2014b). *Statistical pocketbook 2014.* Brussels: European Commission.

EFPA. (2011). *Age based population screening for fitness to drive does not produce any safety benefits.* Brussels: European Federation of Psychologists. 6 p.

Elvik, R., Høye, A., Vaa, T., & Sørensen, M. (2009). *Handbook of Road Safety Measures.* Bingley: Emerald Group Publishing Limited. 1124 p.

ERSO. (2014). *Road safety - Statistics*. Retrieved September 9, 2014, from Mobility and transport: http://ec.europa.eu/transport/road_safety/specialist/statistics/index_en.htm

Europe. (2012). Retrieved September 8, 2014, from Population Pyramids of the World from 1950 to 2100: http://populationpyramid.net/

Eurostat. (2013). Regional GDP per capita in the EU in 2010. (2013, March 21). *Eurostat Newsrelease*(46), 7 p.

Eurostat. (2014). *Demography*. Retrieved September 2, 2014, from Eurostat: http://epp.eurostat.ec.europa.eu/portal/page/portal/population/data/main_tables

Giannakouris K. (2010). Population and social conditions. *Statistics in focus 1/2010*. 19 p.

Hakamies-Blomqvist, L., Johansson, K. & Lundberg, C. (1996). Medical Screening of Older Drivers as a Traffic Safety Measure – A Comparative Finnish-Swedish Evaluation Study. *Journal of the American Geriatrics Society*, 44: 650 – 653.

Langford, J., & Koppel, S. (2006). The case for and against mandatory age-based assessment of older drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 9*(5), 353-362.

Langford, J., Fitzharris, M., Newstead, S., & Koppel, S. (2004). Some consequences of different older driver licensing procedures in Australia. *Accident Analysis & Prevention, 36*(6), 993-1001.

Mitchell, C. G. (2008). The licensing of older drivers in Europe– A case study. *Traffic Injury Prevention, 9*(4), 360-366.
OECD. (2001). *Ageing and transport: Mobility needs and safety issues.* Paris: OECD. 11 p.

Pace J. F. et al. (2012). *Basic Fact Sheet "Pedestrians".* Brussels: European Road Safety Observatory. 22 p.

Siren A. et al. (2013). *Driver licensing legislation.* Brussels: CONSOL. 48 p.

WHO. (2004). *World report on road traffic injury prevention.* Geneva: World Health Organization. 219 p.

WHO. (2013). *Global status report on road safety 2013: Supporting a decade of action.* Geneva: World Health Organization. 304 p.

# GEOGRAPHICAL INFORMATION SYSTEMS AIDED TRAFFIC ACCIDENT ANALYSIS CASE STUDY: CITY OF ERZURUM NORTH RING ROAD

**Muhammed Yasin ÇODUR[1], Ahmet ATALAY[2], Ahmet TORTUM[3], Ahmet Özgür DOĞRU[4]**

[1]*Erzurum Technical University, Engineering and Architecture Faculty, Department of Civil Engineering, Erzurum/Türkiye*

[2]*Ataturk University, Narman Vocational High School, Department of Construction, Narman/Erzurum/Türkiye*

[3]*Ataturk University, Engineering Faculty, Department of Civil Engineering, Erzurum/Türkiye*

[4]*İstanbul Technical University, Civil Engineering Faculty, Geomatic Engineering Department, İstanbul/Türkiye*

**Abstract:** The population growth in developing World, technological development and urbanization is directly connected with the systems of transportation. Although transportation sector offers pretty much alternatives nowadays, people heavily prefer to highway transportation mode in the east of Turkey. The studies which have been done for providing road safety decrease traffic accidents. Geographical Information System (GIS) technology has been a popular tool for accident data and analysis of black spots in highways. Many traffic agencies have been using GIS for accident analysis. Accident analysis studies aim at the identification of high rate accident locations and safety deficient areas on the highways. A case study, using GIS aided traffic accident analysis for The North Ring Road of Erzurum/Turkey were developed by using the historical data, between 2005 and 2012. In conclusion, this study focused on the practicability of the GIS with the maps for traffic safety analysis.

**Keywords:** Traffic safety, Geographical Information System, Traffic Accidents

## 1. Introduction

Technological improvements have increased and diversified production based on people's needs. The most important needs include traffic safety and speed. An increasing number of motor vehicles and drivers have brought focus on the importance of traffic safety. Turkey is a large country which is a junction point between Europe and Asia in terms of its economic and social location. The population is 76 667 864, and the number of total vehicle and accident number is 17 939 447 and 1 207 354 respectively, according to the report of Road Motor Vehicle Statistics 2013.

In Turkey, one of the factors affecting traffic safety is highways. Highway transport has a great importance in Turkey. On the highways traffic accidents have been recognized as one of the major causes of human and economic losses, both in developed and developing countries. This problem is of great concern in developing countries, because of its seriousness and the limited resources available to develop feasible countermeasures for reducing this ever-growing challenge (Berhanu, 2004). The number of accidents on a given highway section during a certain period of time is probabilistic in nature and is a non-negative integer. Despite the fact that accidents are random and unpredictable at micro-level, statistical models can predict reliable estimates of expected accidents by relating aggregates of accidents to various explanatory measures of flow, site characteristics, and road geometry at macro-level. Numerous empirical relationships between vehicle accidents and these explanatory variables have been established in several previous studies (Maher and Summersgill, 1996; Miaou, 1994, Miaou and Lum, 1993).

Traffic accident analysis is an extremely complex topic due to the existence of numerous factors affecting accidents spatially. Therefore, traffic accident reports need to be more detailed and formatted properly for spatial and statistical analysis (Erdoğan et al., 2008). Most of these accidents result from human error, mainly carelessness of the drivers or pedestrians. Hence, the probability of accident occurrence, and its severity, can often be reduced by the systematic analysis of the incident scenario and by resorting to appropriate solutions involving the application of proper traffic control devices, suitable roadway design practices and effective traffic police activities. However, the task of devising effective solutions warrants analysis of spatial and temporal patterns in the zonation of traffic accidents, which can be achieved through the application of geospatial technology (Levine et al., 1995; Harkey, 1999; Li et al., 2007; Cheng and Washington, 2008). The non- random distribution of accidents, both in time and space, often raises questions about the location and the reasons for that location (Maher and Mountain, 1988; Liang et al., 2005; Randall et al., 2005; Schurman, 2009). Unlike the conventional methods, spatial thinking helps to identify the patterns and suggest reasons for the pattern characteristics. GIS technology has been a popular tool for visualization of accident data and analysis of traffic accidents and hence is being used by many traffic agencies (Miller, 1999; Cook et al., 2001; Ghosh et al., 2004; Mitchell, 2005; Anderson, 2006; Deepthi and Ganeshkumar 2010).

[1] Corresponding Author: mycodur@erzurum.edu.tr

The vehicle density in the city of Erzurum has shown drastic increase in the past few years, but the road conditions remain same resulting in increased incidents of traffic accidents and the situation warrants a detailed scientific study on the spatial and temporal probability of accident occurrence. As conventional techniques of accident analysis had failed in reducing the occurrence of traffic related accidents, the present study aims to evaluate and represent the traffic accident locations using Geographical Information System (GIS) technology.

## 2. Case Study Area and Data

Erzurum is the biggest city in the East of Turkey, where drastic changes in land use and land cover have taken place. The region, in general, has hilly topography with high mountains. The city is situated 1757 meters above sea level. Erzurum has a humid continental climate. Summer is brief, but summer days are warm, though with cool nights. Winters are cold, with an average minimum during January of around -15 °C. The city is a popular destination in Turkey for winter sports at the nearby Palandöken Mountain and hosted the 2011 Winter Universiade Olympic Games.

The city is categorized as fast growing, it is not properly planned and most of the roads are narrow causing heavy traffic congestion. The fast and speedy development, in turn, has increased the number of vehicles without proportional augmentation in road infrastructure. Since the roads in the area are always under excessive pressure and the likelihood of accidents is more, the analysis of accidents are very important to Directory of Erzurum Traffic Region as well as transportation planners and traffic engineers. Highway traffic accidents continue to be a major problem also in Erzurum, both from the public health and socio-economic perspectives. On the Erzurum Highways in the six-year period 2005-2012 and 8000 traffic accidents were recorded (Çodur, 2012).

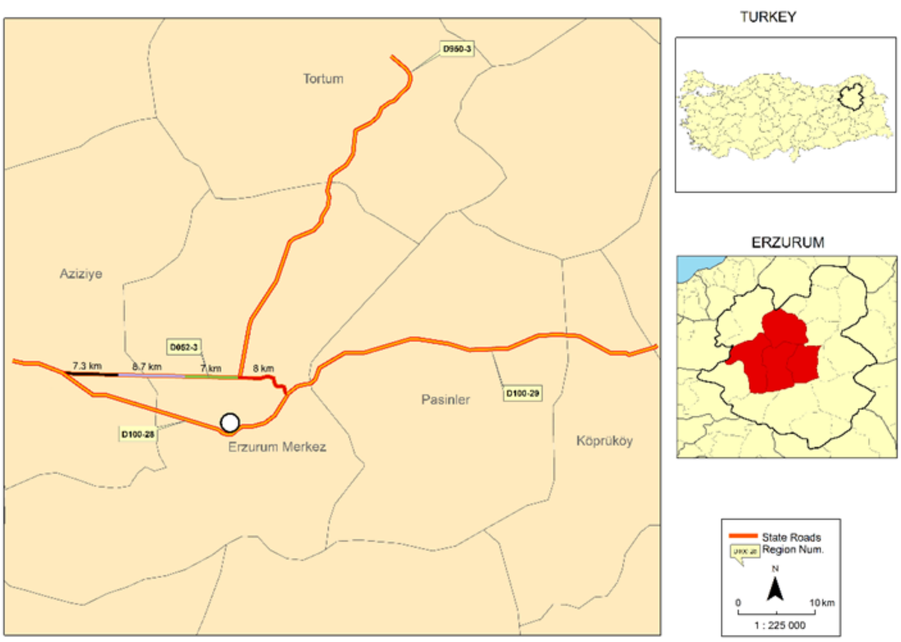



**Fig.1.**
*Erzurum's highway and sections*

Traffic accidents often can be reduced by the application of proper traffic precautionary measures (Hauer, 1996). The success of these precautionary measures depends on the analysis of traffic accident records. So it is important to collect the accurate, precise and reliable data with the traffic accident reports. Traffic accident reports are collected from the Directory of Erzurum Traffic Region. These accident reports include collected accident parameters such as the date, hour/minutes, kilometer of crash, code of highway, age, sex and alcohol consumption of driver, weather conditions, lighting conditions, vehicle type, number of persons injured/killed etc. The goal of this data collection exercise is to divide these highways into segments with homogenous characteristics. After reviewing several highways around Erzurum, it was decided that D 052-03 a four-lane median-divided Erzurum's highway is most appropriate for this task. D 052-03 were divided into four homogenous sections which are 8km, 7km, 8.7km and 7.3km respectively. The international airport of Erzurum is also on this highway (Figure 2) (Çodur, 2012).

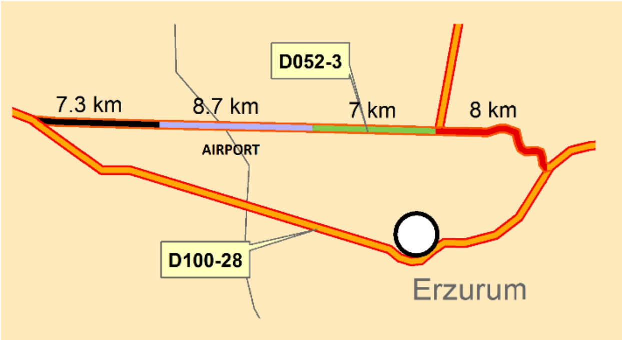



**Fig.2.**
*D 052-03 Erzurum north ring road*

## 3. Methodology

A Geographic Information System (GIS) is a computer system designed to capture, store, manipulate, analyze, manage, and present all types of spatial or geographical data. Maps are used to communicate the geographical information. They enable to observe both the attribute and geographical relationships at once. The map is an efficient tool to explore the geographical relationships between earth observations. This is an easy way to rapidly browse relationships between many variables. Data can be presented on the maps by different methods: dot-density maps, proportional circles, proportional spheres, grey-scale (choropleth) maps, contour (isopleths) maps, cartograms and 3D surface plots (Kraak and Ormeling, 1996; Terry, 1999; Uluğtekin et al., 2006; Çodur, 2012).

GIS permits users to display database information geographically. It can also provide a common link between two or more previously unrelated databases. The most useful aspect of GIS as a management tool is its ability to associate spatial objects (street names, milepost, route number, etc.) with attribute information (accidents, cause, etc.). Most of the reviewed documents consider the use of GIS in transportation under either for general data maintenance (primarily inventory of transportation-related incidents) or for simple data analysis (Liang et al., 2005).

GIS has been proven to work well in addressing transportation problems related to safety. Affum and Taylor (1997) described the development of a Safety Evaluation Method for a Local Area Traffic Management, which is a GIS-based program for analyzing accident patterns over time and the evaluation of the safety benefits. GIS can also be implemented in determining roadway and surface conditions. This was proven by Gharaibeh et. al. (1994) when they proposed to use GIS to obtain statistical and spatial analyses of roadway characteristics such as safety, congestion level and pavement conditions.

There also have been studied that aimed at showing how GIS can be applied in accident management systems. Faghri and Raman (1992) developed a GIS-based traffic accident information system for Kent County, Delaware. Their system included knowledge about the occurrence of crashes, such as conditions of incident site, and frequency of incidents at any given location (mile-point) on a roadway. Since the early stages of GIS, it was noticed that a vision of information technology outside the traditional transportation data analysis and even outside GIS was needed to implement this technology (Lewis, 1990).

## 4. Conclusion

Within the context of this study, Geographic Information System technology were used for analyzing the accidents occurred on D 052-3 highway. Under these circumstances a Geographic Information System was developed for the purpose of reducing number of the accidents. This system consists of following steps. First, accident records from 2005 to 2012 were collected from the Directory of Erzurum Traffic Region and input into an MS Access database. Meanwhile D 052-3 highway in the field of Erzurum district were digitized by using Arc GIS 9.0 software. The location of the accidents positioned with the “kilometer of crash” data on the route of highways using the “linear referencing” toolset in Arc GIS 9.0. Network Analyst tool were used to visualize the accidents based on their location. The main characteristics of the D 052-3 highway such as number of the vertical and horizontal curves were also visualized together with the number of the accidents in order to introduce the relations between road characteristics and accidents. Additionally accidents were also analyzed based on seasons by using GIS.

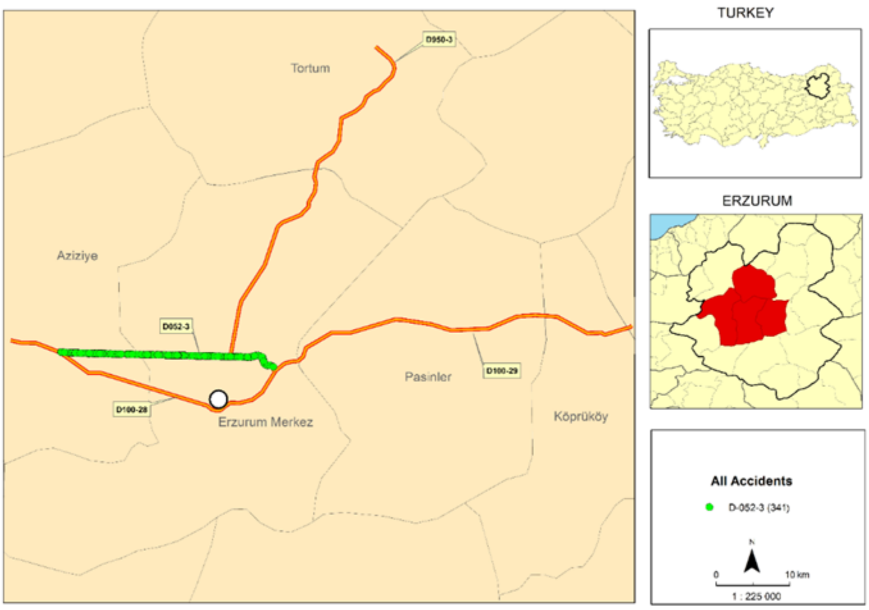



**Fig. 3.**
*All traffic accidents on the D 052-3*

Based on the accident reports totally 341 accidents occurred on the D 052-3 highway during the study period. Locations of these accidents were visualized in Figure 3. As the second step of the GIS analysis, accidents were examined depending on the official road segments and their characteristics. As it is seen in Figure 1, D 052-3 highway is divided in to 4 segments. For a detailed examination of accidents, road segments were classified based on the number of horizontal and vertical curves. As it is seen in Figure 4a, as the number of the horizontal curves increases the number of the accidents increases. 94 accidents occurred on 4$^{th}$ road segment of D 052-3 Highway with highest number of horizontal curves. Figure 4b displays similar relation with number of vertical curves and number of accidents.

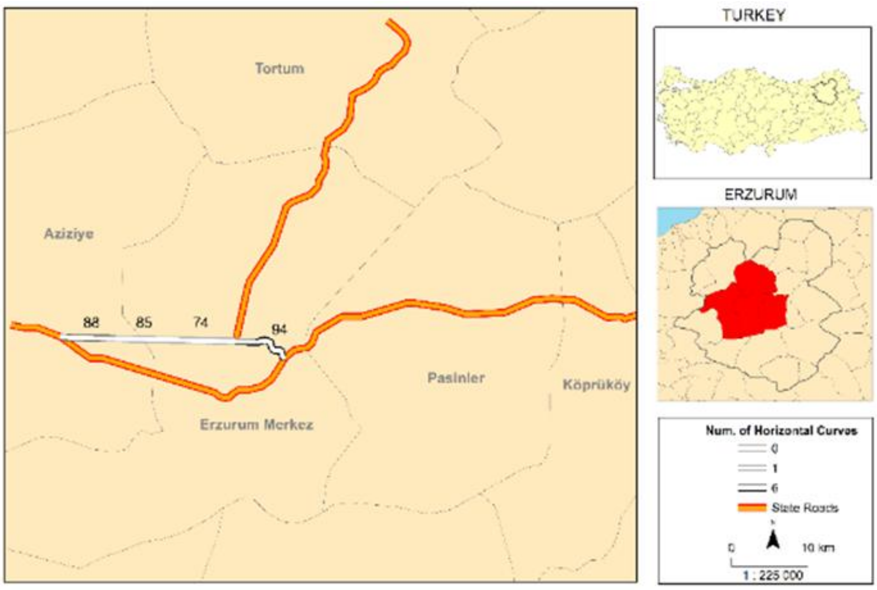



**Fig. 4. a.**
*The number of accidents on the horizontal curves*

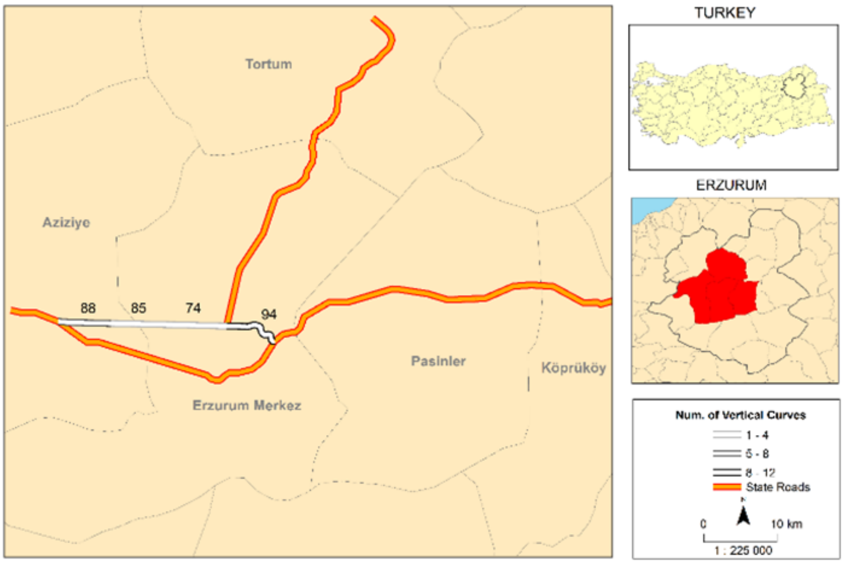



**Fig. 4. b.**
*The number of accidents on the vertical curves*

Additionally, seasonal distribution of the accidents were also analyzed within the context of the study. Figure 5 displays accidents by season. It is clearly understood that spring term can be considered as the least risky season it terms of accidents.

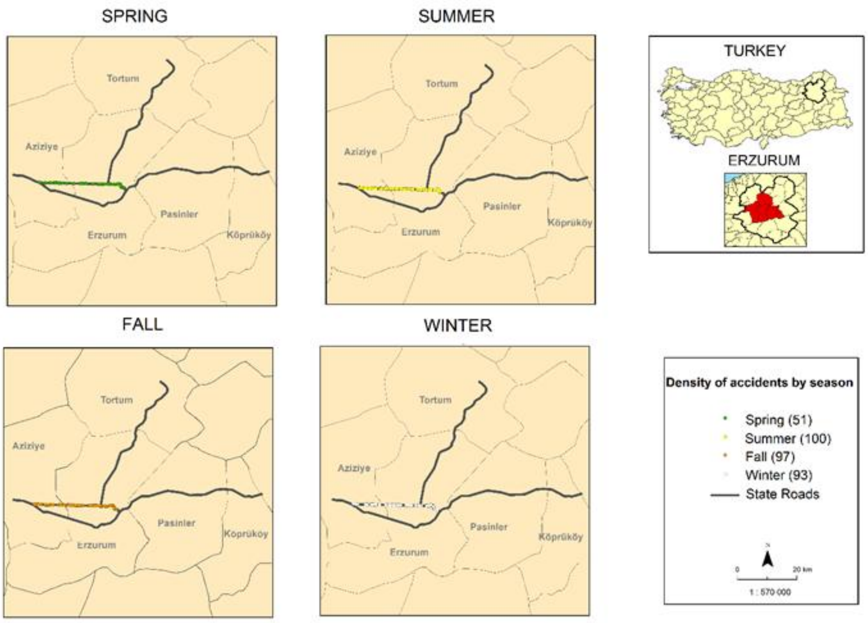



**Fig. 5.**
*Traffic accidents depends on seasons*

These results have implications for policy makers, transportation system designers, and researchers. Transportation safety designers cannot easily identify factors, make recommendations for incremental changes in the factor, and hope to achieve major differences in accident levels. The problems have to be analyzed and attacked from a multidimensional perspective: a wide variety of geometric and traffic characteristics. In this study it is showed that GIS is able to handle traffic accident analysis easily. In addition to this GIS system has a lot of advantages such as quick access for obtaining information, data storage, output and integrity.

**Acknowledgements**

The authors are grateful to reviewers to accept earlier version of the paper. We would also like to thank people who are working in the Directory of Erzurum Traffic Region and 12th Highway Regional Directorate of Erzurum. The authors also would like to thank Prof. Dr. Nesibe Necla ULUĞTEKİN who is an academician in İstanbul Technical University, Civil Engineering Faculty, Geomatic Engineering Department.

**References**

Affum, J.K., Taylor, M.A., 1997. SELATM – A GIS based program for evaluating the safety benefits of local area traffic management schemes. Transportation Planning and Technology 21, 93–119.

Anderson T., 2006. Comparison of spatial methods for measuring road accident hotspots: a case study of London. J Maps; 55- 63.

Berhanu, G., 2004. Models relating traffic safety with road environment and traffic flows on arterial roads in Addis Ababa. Accident Analysis and Prevention; 36 (5), pages 697–704.

Cheng W, Washington S, 2008. New criteria for evaluating methods of identifying hotspots. Transportation Research Record: Journal of the Transportation Research Board, No. 2083, Transportation Research Board of the National Academies, Washington, D.C, 2008, pages 76–85.

Cook WD, Kazakov A, Persaud BN. 2001. Prioritizing highway accident sites: a data envelopment analysis model. J Oper Res Soc 2001; 52: 303-309.

Çodur M.Y., 2012. Traffic Accident Prediction Models: Applications for Surrounding Highways of Erzurum. PhD Thesis, Graduate School of Natural and Applied Sciences, Department of Civil Engineering, Ataturk University Turkey.

Deepthi JK, Ganeshkumar B. 2010. Identification of accident hot spots: A GIS based implementation for Kannur District, Kerala. Inter J Geomatics and Geosc 2010;1 (1): 51-59.

Erdoğan S., Yılmaz I., Baybura T., Gullu M., 2007. Geographical information systems aided traffic accident analysis system case study: city of Afyonkarahisar, Accident Analysis and Prevention 40 (2008) 174–181.

Erdogan S., Yilmaz, I., Baybura, T., Gulu, M., 2008. Geographical information systems aided traffic accident analysis system case study: City of Afyonkarahisar, Accident Analysis and Prevention 40, 174–181.

Faghri A., Raman N., 1992. Evaluation of geographical information system applications in Delaware. Report 92-DTC-6 to DELTOT, Delaware Transportation Center, and Department of civil engineering, University of Delaware, Newark DE USA.

Ghosh SK, Parida M, Uraon JK., 2004. Traffic accident analysis for Dehradun city using GIS. ITPI Journal;1(3): 40-54.

Harkey DL., 2007. Evaluation of truck crashes using a GIS-based crash referencing and analysis system transportation research record. J Trans Res Board; 1686, 13-21.

Hauer, E., 1996. Identification of Sites with Promise, Transportation Research Record, Issue: 1542, 74th Annual Meeting, Washington, USA, pp. 54-60.

Kraak, M.J.; Ormeling, F.J. Cartography., 1996. Visualization of Spatial Data; ISBN 0-582-25953-3. Longman Limited: England; 222 pp.

Levine N, Kim KE, Nitz LH., 1995. Spatial analysis of Honolulu motor vehicle crashes: I. spatial patterns. Acc Anal and Prev; 27(5): 663-674.

Lewis, S., 1990. Use of geographical information systems in transportation modeling. ITE Journal, March, 34-38.

Li L, Zhu L, Sui DZ., 2007. A GIS-based bayesian approach for analyzing spatial–temporal patterns of intra-city motor vehicle crashes. J Trans Geog;15 (4): 274-285.

Liang LY, Hua LT, Ma'some DM 2005. Traffic accident application using Geographic Information System. J Eastern Asia Soc Trans Studies 2005; 6: 3574 – 3589.

Maher MJ, Mountain LJ., 1988. The identification of accident black spots: a comparison of current methods. Acc Anal and Prev; 20(2): 143–151.

Maher, M.J., Summersgill, J., 1996. A comprehensive methodology for the fitting of predictive accident models. Accident Analysis and Prevention; 28 (3), pages 281–296.

Miaou, S.P., Lum, H., 1993. Modeling vehicle accidents and highway geometric design relationships. Accident Analysis and Prevention; 25 (6), pages 689–709.

Miaou, S.P., 1994. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. Accident Analysis and Prevention; 26 (4), pages 471–482.

Miller HJ., 1999. Potential contribution of spatial analysis to geographic information systems for transportation (GIS-T). Geog Anal; 31(4): 373–399.

Mitchell A., 2005. The ESRI guide to GIS analysis. Volume 2: Spatial measurement and statistics. ESRI Press, Redlands, California.

Randall TA, Churchill CJ, Baetz BW. 2005. Geographic information system (GIS) based decision support for neighborhood traffic calming. Can J Civil Eng; 32: 86-98.

Schurman N., Cinnamon J, Crooks VA, Hameed SM., 2009. Pedestrian injury and the built environment: an environmental scan of hotspots. BMC Public Health; 9(233).

Terry, A.S., 1999. Thematic Cartography and Visualization; ISBN 0-13-209776-1, Prentice Hall: Englewood Cliffs, New Jersey; 293 pp.

Uluğtekin N, Alkoy S., Seker D, Goksel Ç., 2006. Use of GIS in Epidemiology: A Case Study in Istanbul. Journal of Environmental Science and Health Part A, 41:2013–2026.

# WAYFINDING SYSTEM IN PUBLIC TRANSPORT FOR PEOPLE WITH VISUAL IMPAIRMENT

**Pavol Korček[1], Lea Rollova[2]**
*[1,2]Center of Design for All, Faculty of Architecture, Slovak University of Technology in Bratislava*

**Abstract:** The quality of wayfinding system significantly affects the quality of the environment of transport buildings. In design it is necessary to take into consideration that the environment for many passengers entering the transport building for the first time is new and unknown. Most wayfinding information in transport building is provided in a visual format. However, for the majority of visually impaired people, this format is inappropriate. It follows that the wayfinding information should be provided in different formats which are perceivable by various senses.
Existing wayfinding systems in the transport buildings for persons with visual impairments are partly defined in standards, but one is still looking for solutions which would improve the system or supplement the system by new knowledge.
In the first part, the paper describes the existing wayfinding systems for people with visual impairments. The most commonly used are tactile, acoustic and visual wayfinding systems**.** The main part of the paper is focused on the research and presentation of partial results in the research. Center of Design for All, Faculty of Architecture STU in Bratislava in cooperation with Slovak Blind and Partially Sighted Union created the basic principles of wayfinding system in Slovakia which were incorporated into legislation of Slovakia. The system should continue to be improved and for this reason an additional survey and evaluation of the existing system has been executed. The main objective of this research is to investigate the current condition of the Slovak transport system with regards to wayfinding for visually impaired or blind passengers inside the selected transport buildings. The results of the project could be used to make particular recommendations to include the specific needs of travellers with visual impairment in the design of a public transport infrastructure and as a theoretical basis for renewal of the Slovak legislation.

**Keywords**: transport building, universal design, access audit, visual impairment, wayfinding system

## 1. Introduction

The information is an important part of travelling based on which passengers decide and navigate in space and time. Many passengers enter a transport building for the first time; the environment is new and unknown for them so they could not find their way without properly designed wayfinding system. The existing wayfinding systems in transport buildings are mostly based on visual form. These visual systems are not suitable especially for passengers with visual impairments. Therefore it is necessary to create a multisensory wayfinding system which provides information in different kinds of sensory perception, accessible to a wide range of users, including users with sensory impairments.

## 2. Statement of the problem

For persons with visual impairment it is difficult to not only gather information in a built environment, but also to find the way independently, especially in an unfamiliar environment. Research by Passini and Proulx (Passini, Proulx, 1988) points out that the visually impaired and the blind usually rely on their previous training and experience to move around inside the built environment. Visually impaired people should be able to use transport environments with confidence and without stress. For their safe movement it is therefore essential to have enough clear information and orientation aids by using a clear spatial *navigation* strategies.
„*Wayfinding is typically divided into two categories: orientation and mobility. Orientation concerns the ability for one to monitor his or her position in relationship to the environment; and mobility refers to one's ability to travel safely, detecting and avoiding obstacles and other potential hazards*" (National Institute on Disability and Rehabilitation Research, 2001). The ability to navigate in space depends on the type and severity of visual impairment. For example, some blind and severely visual impaired people perceive only light and outlines. Partially sighted people and people with disorders of binocular vision have a serious problem with spatial perception. Visually impaired people must therefore replace eyesight with other sensory channels in the cognitive process. Blind persons use touch, hearing and smell for the movement and orientation. Partially sighted persons use mainly remnants of their vision and, depending on the degree and type of visual defects also touch, hearing and smell (Slovak Blind and Partially Sighted Union, 2013).
Existing wayfinding systems in the transport buildings for persons with visual impairments are partly defined in international and national standards, but one is still looking for solutions which would improve the system or supplement the system by new knowledge.
If the environment of transport building meets the needs of passengers with visual impairments, it must be designed according to the principles of Universal Design. Following these principles leads to a non-discriminatory design approach and provides increased usability for everyone.

[1] Corresponding Author: korcek@fa.stuba.sk

### 3. Brief Review the Current Available Solutions

As mentioned above, wayfinding systems for people with visual impairments are based on a set of measures, which will improve the overall perception of space. We call it multisensory perception which uses multiple (at least two) modes of sensoric perception in receiving information or provide communication and orientation in space. This method is based on the principles of universal design. It is the principle 4 and guideline 4a:

- the design communicates necessary information effectively to the user, regardless of ambient conditions or the user's sensory abilities
- use different modes (pictorial, verbal, tactile) for redundant presentation of essential information.

The most commonly used combination is tactile and visual perception or auditory and visual perception. Besides the mentioned methods, persons with visual impairments use a guide dog or personal assistance in movement and orientation. In the recent time, the wayfinding assistance for visual impaired persons has been receiving much attention focused on navigation and orientation assistive systems based on information and communications technology.

#### 3.1. Tactile wayfinding systems

For spatial navigation, visual impaired persons used different tactile techniques, especially long cane technique, touch (foot contact) and trailing (gliding finger technique). But in order to effectively use these techniques, they must be supported by a coherent system of guide lines and other navigation elements which must be a necessary part of the architectural design of passageways in transport building.

The basic elements of tactile wayfinding systems include:

- tactile ground indicators
- embossed maps and 3D touch models
- embossed labels.

The purpose of **tactile ground indicators** is to give travellers who are blind, or who have vision impairment, a warning of hazards and directional information to aid navigation of the urban or interior environment.

**Embossed maps and 3D models** can provide information about the layout of space of transport building and a system of guide lines. Using the touch contact with embossed map or 3D model, visually impaired persons can get familiar with the space layout of individual room spaces or localize the search area (Samova, Ceresnova, 1997).

Using the **embossed labels**, important areas or elements must be marked. For example, the embossed label provides information on the handrails of stairs or on the walls next to elevator doors about a floor number, respectively platform number. The embossed label next to the door (or above door handle) provides information about the purpose of the room. In the elevator the embossed labels are placed on the control panel and indicate the number of floors or functional parts of the building (entrance, exit, platform, departure, arrival etc.).

#### 3.2. Acoustic wayfinding systems

With acoustic perception, an option is used to obtain information through natural and targeted sounds from the environment, as well as voice information. There are various systems for transmitting information using sound, the most common being the following:

- basic acoustic information
- acoustic signalling
- acoustic beacons
- audio recordings
- demand (command) sets.

**The basic acoustic information** is used in almost all transport buildings. It is important to ensure that all spaces of transport buildings are sound distributed, intended for public use, including platforms. **Acoustic signalling** is mainly used in urban environment at pedestrian crossings and at railway crossings.

**Acoustic beacons** provide sound information about places which are significant in terms of spatial navigation, while they determine and help keep the direction of visually impaired persons. The basic acoustic signal may be supplemented by voice information in the area, the deployment of services in large buildings, hall spaces of transport buildings and etc. (Matuška, 2009).

**Audio recordings** can provide more detailed information to visually impaired passengers, e.g. about the space. These can be solved in different ways; the most common include integration into typhlographic aids or integration into the handrail (activated by pressing or photo sensor).

For better orientation visually impaired persons used **demand sets** in public transport. To activate them a remote control is used (which is universal and can also control the acoustic signalling and acoustic beacons). The demand sets, if supported by other facilities and vehicles, may serve multiple purposes, aiming to get more information.

### 3.3. Visual wayfinding systems

Blind people are only a part of the visually impaired population. Most people with visual impairments can use the rest of the sight and can therefore use appropriately designed optical elements simplifying their spatial navigation and movement. The main principle of the spatial navigation is based on:

- creation of optimal size of objects
- significant colour contrast of important elements
- quality lighting.

Significant contrast in colour is an important element of the design of transport buildings. A main route in a terminal is marked with colour differentiation in a form of contrasting colour solutions. The colour contrast is used for floors and walls, stairs, ramps, but also interior elements such as entrance doors, information desks and ticket counters, etc.
Quality lighting of all spaces is important for persons with visual disabilities to be able to effectively use the rest of the sight. Lighting is then dispersed and will not dazzle. Lighting can also illuminate effectively elements important for orientation in the interior.

### 3.4. Prevention of injuries

A travel hazard is an architectural or environmental obstruction in the path of travel that cannot be readily detected and negotiated with standard long cane techniques, e.g. hanging objects (fire extinguishers, shelves) on the walls or stairs with open backs. Hazardous are also glass walls which are not perceptible enough for visually impaired travellers, moveable obstacles include both semi-permanent features, such as seating, tables, and temporary obstacles left by users. All these features need to be appropriately located away from passageways so that they avoid causing a hazard, or if they cannot be removed an indent in the floor level shall be designated so that they are identifiable by the long cane technique.

## 4. Research and methodology

Wayfinding systems for passengers with visual disabilities have different forms in each country, a unified system does not yet exist, which makes it significantly difficult to travel across borders for people with visual disabilities. Currently, several ongoing research projects aimed at improving wayfinding systems in transport buildings are under way. Also, the international effort to unify elements of wayfinding systems is being desired. Center of Design for All, Faculty of Architecture STU in Bratislava in cooperation with Slovak Blind and Partially Sighted Union created the basic principles of wayfinding system in the Slovak Republic which was incorporated into legislation of the Slovak Republic. The system should continue to be improved and for this reason an additional survey and an evaluation of the existing system have been executed. The main objective of this research is to investigate the current condition of Slovak transport system with regards to wayfinding for visually impaired or blind passengers inside the selected transport buildings. The results of the project could be used to make concrete recommendations to include the specific needs of travelers with visual impairment in the design of a public transport infrastructure and as a theoretical basis for renewal of Slovak legislation.

*To address these overall objectives this project shall:*

- provide access audit of selected transport buildings
- determine problems and barriers encountered by visually impaired on current transportation systems through on-site analysis
- identify the main difficulties which public transportation users with visual impairment must manage
- verify the effects which colour design and lighting have on the environment and safety
- look for ways how spatial design, orientation and safety features, colour, contrast and lighting can improve the perception of the built environment and safety of movement,
- prepare design guidance for architects and designers for wayfinding in a inclusive transport infrastructure.

## 5. Partial Results

In the initial phase of the project the access audit of selected transport buildings was carried out. Based on the results of the audits, we can conclude that the wayfinding system for persons with visual disabilities exists only in those transport buildings which have been adapted. The existing systems in the assessed transport buildings are often executed imperfectly and have a number of shortcomings. The shortcomings are in all areas of the system, especially tactile, acoustic and visual elements of the wayfinding system.

After executing status-quo analysis we found out the following conclusions:

- there is a lack of control mechanisms to verify the implementation of barrier-free accessibility of buildings in Slovak legislation
- there are no authorized persons in Slovakia who would provide consultancy and assess compliance of legislation with the execution in the approval phase and in the phase of final inspection of buildings
- there is low awareness of planners which is linked to insufficient public awareness of barrier-free design and Universal design.

The most important findings resulting from the surveys undertaken were related to inappropriate solutions, which made spatial navigation for people with visual impairments difficult and ideally they should be included in the Slovak legislation:

- hardly identifiable entrance to the building - problem often occurs with glazed facades
- poor orientation in the entrance hall due to poor layout - especially inappropriately placed staircase and guide line pointing to the escalator
- improper selection of surface materials which makes it very difficult for spatial perception
- missing marking of the first and last stair step in the stairwell
- absence of acoustic navigation systems.

Persons with visual disabilities often failed to identify position of the main entrance on the glazed facades on the transport buildings. Although the legislation stipulates that each public building must have a lead guide line to the main entrance, it is often not there. The major shortcoming of the legislation is, however, that the glazed wall as well as the entrance glazed door must be marked with a line (symbol) in the eye level. The identical graphic marking on both the glazed wall and the entrance door is commonly used. If it is not, it may cause that the entrance door may not be fully visually perceptible.


**Fig. 1.**
*Hardly identifiable entrance to the railway station*
*Source: Pavol Korček*

**Recommendations:** The door can be highlighted in different ways, e.g. design solution of the facade, roofing the entrance, colour distinction of the entrance, change of graphic marking on the entrance doors.

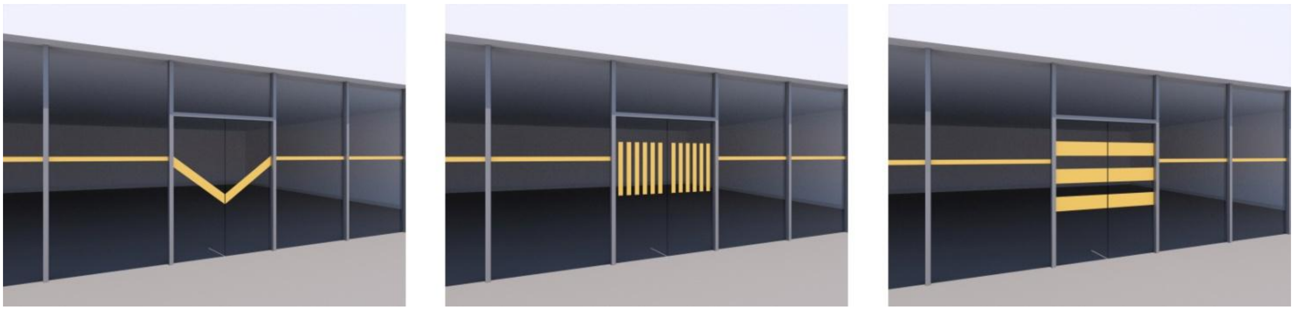


**Fig. 2.**
*Examples of suitable design of graphic marking on the entrance doors*
*Source: Pavol Korček, Lea Rollova*

The majority of visually impaired persons preferred to use a staircase rather than an escalator or an elevator.
The visually impaired persons using the elevator often lose orientation at a random stop on a floor (if there is not acoustic floor information announcement device available). The escalator is dangerous for persons with visual impairment.

**Recommendations:** The preferred staircase should be located on the easily identifiable place on the main path of travel. The tactile ground indicators must navigate to staircase rather than to escalators.

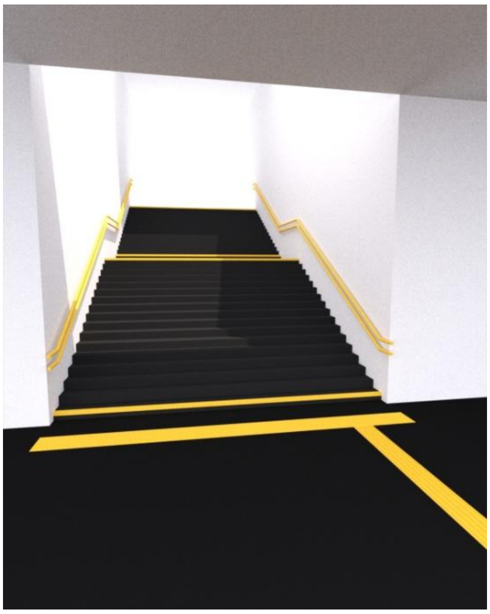


**Fig. 3.**
*Tactile guideline navigating to handrail of the staircase*
*Source: Pavol Korček, Lea Rollova*

The elevator must have the acoustic information announcement device installed. It is recommended to attach the embossed signs with a floor number on the wall next to the elevator door. It is necessary to make sure that all important spaces are lined back to back along the main path of travel (information desk, ticket sale, baggage deposit, toilets etc.).
Glaring flooring is often used in transport buildings (airports, railway stations). The floor reflections distort perception of space, especially for visually impaired persons. If both the flooring colour and wall colour are identical (non-contrastive) it will also make the orientation for them difficult.
**Recommendations:** The floor must be non-glaring, also dark floor is desired. The dark floor can improve better perception of space, and other interior elements are visible better, too. The colour solution of interior design plays a crucial role in visual orientation of partly sighted persons in large space.
In the transport buildings are often missing marking of the first and last stair step in the stairwells. In other cases, marking is just colored without tactile indication.

**Fig. 4.**
*Missing tactile marking of the first and last stair step in the stairwell*
*Source: Pavol Korček*

**Recommendations:** First and last stair step in the stairwells should be marked by colour and also by tactile ground indicator.

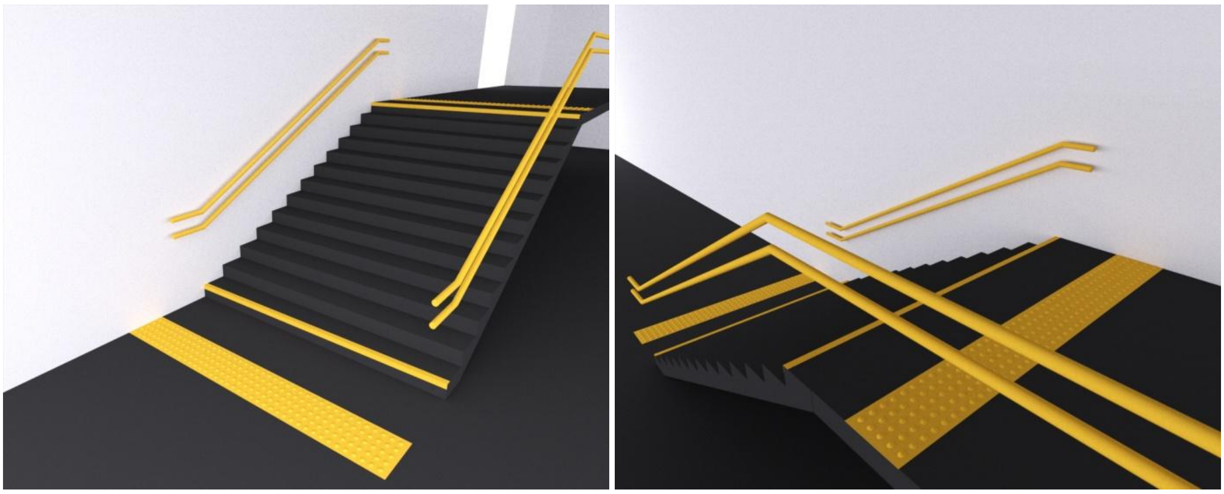


**Fig. 5.**
*Suitable marking of the first and last stair step in the stairwell*
*Source: Pavol Korček, Lea Rollova*

The acoustic navigation systems which can significantly improve orientation of visually impaired persons are very often absent in Slovak transport buildings.
**Recommendations**: A positive example of use of acoustic navigation systems is e.g. Prague Integrated Transport: „*Orientation in selected stations is made easier by special acoustic beacons informing the blind by means of an acoustic signal (tone tune) in underpasses and vestibules about the entrance to the passenger processing area of the station. At some stations acoustic beacons also provide voice information. Blind passengers can activate acoustic beacons by means of a transmitter as required*“ (Barta, 2009).

## Conclusion

It can be said that several wayfinding systems suitable for visually impaired people are used in the transport buildings at present. The problem is that architects and designers do not know them well and cannot properly design or use them. It is important that architects plan ahead using wayfinding systems suitable for people with visual impairments from the earliest stages of transport building enviroment design.
A good building design is one that is easily understood by people with impaired vision. Architects, planners and designers must focus on distinctiveness and simplicity with sufficient accessible information regarding the built environment. They can also make a significant improvement in spatial learning, wayfinding and navigation by implementing various measures in buildings.

## Acknowledgements

This work was financially supported by the Scientific Grant Agency of The Ministry of Education, Science, Research and Sport of the Slovak Republic.

**References:**

Barta, F. 2009. Information for persons with impaired mobility. Available from Internet: <http://www.eltis.org/docs/studies/CR-Prague.pdf>.

Matuška, J. 2009. *Barrier-free transport*. Institut Jana Pernera. Czech republic. 200 p.

National Institute on Disability and Rehabilitation Research. 2001. Notice of proposed funding priorities for fiscal years 2001–2003 for three disability and rehabilitation research projects. Available from Internet: <http://www2.ed.gov/legislation/FedRegister/finrule/2001-3/070601b.html>.

Passini, R.; Proulx, G. 1988. Wayfinding without vision: an experiment with congenitally totally blind people, *Environment and Behavior*, 20(2): 227-252.

Samova, M.; Ceresnova, Z. 1997. *Barrier-free design for visually impaired people – methodology guide*. Slovak Blind and Partially Sighted Union. Slovak republic. 32 p.

Slovak Blind and Partially Sighted Union. 2013. Visual impairment and its impact on mobility. Available from Internet: <http://architektonickebariery.sk/node/3>.

# THE RELATIONSHIP BETWEEN ROAD ACCIDENTS AND INFRASTRUCTURE CHARACTERISTICS OF LOW-VOLUME ROADS IN ISRAEL

**Victoria Gitelman**[1], **Etti Doveh**[2], **Roby Carmel**[3], **Fany Pesahov**[3]
[1,3]*Ran Naor Road Safety Research Center, Technion - Israel Institute of Technology, Haifa, Israel*
[2]*Technion Statistical Laboratory, Haifa, Israel*

**Abstract:** This study aimed at the identification of relationships between infrastructure characteristics and accident occurrences, on low-volume non-urban roads in Israel. The low-volume roads were defined as single-carriageway rural roads under the National Transport Infrastructure Company's (NTIC) responsibility, with daily traffic volume below 3,000 vehicles. The study's database included 1358 road sections, with a total length of 1228 km. For each section, accident data and traffic volumes were extracted from the Central Bureau of Statistics' files, while road infrastructure characteristics were produced using the 2010 NTIC road survey's data. Statistical explanatory models were fitted to five accident types including multiple-vehicle collisions, single-vehicle accidents, severe accidents, injury and all accidents. Summarizing the models, it was found that, as expected, an increase in traffic volume leads to an increase in accidents; an initial increase in shoulder width, up to 2-2.5 m, is associated with an increase in accidents, whereas further shoulder widening leads to accident reduction; an initial increase in lane width, up to 3-3.25 m, leads to a reduction in most accident types, where further extension of lane width is associated with an increase in single-vehicle and all accidents. In addition, the increase of minimum horizontal radius is associated with injury accident reduction; the improvement in roadside conditions contributes to a reduction in single-vehicle accidents; vertical grade affects all accidents only, where higher grade is associated with accident increase. Using the models, accident modification factors related to changes in the road infrastructure characteristics were produced. The models developed by the study can serve as a basis for selecting low-volume road section parameters, both during the design and road maintenance processes or black-spot treatment.

**Keywords**: Low-volume roads, road accidents, infrastructure characteristics, explanatory model, accident modification factor.

## 1. Introduction

Single-carriageway roads are typically known for their worse safety records compared to other road types (e.g. Mallschutzke et al, 2006). Among them, two-lane low-volume roads present one of the problematic groups. Many of such roads were built in the far past using obsolete design standards. Others serve as access roads to small villages, and thus, according to the design guidelines, should satisfy a low-service level which means low requirements for most road infrastructure characteristics. As a result, such roads are frequently characterized by poor geometry, non-forgiving roadsides, old road arrangements, and so on. Due to bad infrastructure conditions, such roads may be associated with higher-than-usual accident frequencies and characterized by high values of accidents/injury related to exposure. For example, Fitzpatrick and Brewer (2004) examined infrastructure and accident characteristics of low-volume roads in Texas, and found higher accident rates on roads with more vertical and horizontal curves, more narrow lanes and shoulders, and higher frequency of rigid obstacles on roadsides. Hall et al (2002) surveyed the experience of maintenance and operation of low-volume roads by various road agencies in the US and reported that higher accident rates were observed on road sections with infrastructure deficiencies such as: lack of signing; sharp curves; narrow lanes; lack of shoulders; steep slopes or rocks on roadsides, without safety barriers; etc.

This study aimed at the identification of relationships between the geometric and other characteristics of single-carriageway low-volume roads and accident frequencies and types on such roads, in Israel. The study was commissioned by the National Transport Infrastructure Company (NTIC), which is responsible for the non-urban (rural) road network of the country. The study included the following steps: a survey of international literature on the topic; study's database preparation; data analysis and development of explanatory models. Furthermore, using the models, quantitative tools for engineers were developed to be applied in the processes of new roads' design and/or improving existing roads. In addition, recommendations were provided for implementation of the study's findings in the process of design of low-volume roads in Israel.

## 2. Literature survey

A survey of professional and research literature was undertaken focusing on the issues of definition of low-volume roads; design guidelines on such roads' characteristics, in different countries; findings from the development of explanatory models on the relationships between traffic volumes, road characteristics and accident frequencies, on single-carriageway roads, in the international experience; and local research findings on the safety of low-volume roads. The literature survey demonstrated that, in all countries, a low-volume road is a rural single-carriageway two-way road (OECD, 1986; TRB, 2003). However, among the countries, differences exist as to the upper values of daily traffic volume for the definition of low-volume roads.

[1] Corresponding author: trivica@technion.ac.il

The spectrum of values of traffic volumes on such roads ranges from 100 to 6,000 vehicles per day, where in the majority of countries, this range lies between 400 and 2,000 vehicles per day. At the same time, low-volume roads present a substantial share of the rural road network, e.g. about 50% in the Netherlands, Denmark, Norway and 70%-90% in Australia, Belgium, Canada, Finland, Spain and other countries (OECD, 1986).

As to the design guidelines, due to minor traffic volumes, the budget for building and maintenance of low-volume roads is typically low and, thus, the bottom design values for such roads are lower compared to similar values for single-carriageway roads with higher traffic volumes (OECD, 1986; AASHTO, 2001). The design parameters commonly reduced in this context are: design speed, lane width and shoulder width.

Numerous examples of models for predicting accidents on rural single-carriageway roads can be found in the literature (Reurings et al, 2005). Using the models, a quantitative expression can be suggested for the effect of certain characteristics on accident numbers or rates (accidents per vehicle-kilometers traveled). Design components which were found as influential in such models are: lane width, shoulder width, a combination of lane and shoulder width, a ratio between paved shoulder and unpaved shoulder, shoulder type, super-elevation deficiency (in relation to the design guidelines), curvature, design consistency and grade. For example, based on findings of Zegeer et al (1981), Harwood et al (2000), the impact of shoulder width and lane width was estimated for low-volume roads, where for wider lanes or shoulders lower accident numbers are expected.

An Israeli study of local single-carriageway roads with four-digit numbers (Hakkert and Shokir-Hadad, 2009) demonstrated that accident rates on such roads are high compared to other rural road types. The study also found a relation between the number of injury accidents on such roads and the geometric complexity of the road. Another Israeli study found a relation between the increase in the design consistency of two-lane roads and accident rate reduction (Mattar-Habib et al, 2008). Polus and Cohen (2010) considered the safety level of low-volume roads in Israel defining a low-volume road as a road with traffic volumes of up to 3,000 vehicles per day. Based on accident data on such roads, a model was created for predicting the number of severe accidents according to traffic volumes.

In total, extensive experience of developing models for assessing accident frequencies on two-lane roads, based on accident data, traffic volumes and road geometric characteristics, can be found in the literature. However, the composition of the explanatory variables remaining in the model is not fixed but varies depending on the definition of road sections examined, data availability, analysis objectives, etc.

## 3. The study's database

As mentioned above, the literature survey revealed that among various countries, differences exist as to the upper value of daily traffic volume defining low-volume roads. In the majority of countries, this range lies between 400 and 2,000 vehicles per day, where road sections with such low traffic volumes are rare in Israel. Following a discussion with the NTIC, low-volume roads on the rural Israeli network and the study's road sections, respectively, were defined as single-carriageway roads under the NTIC responsibility, with daily traffic volumes below 3,000 vehicles (as a rounded figure). Having matched the information from the Central Bureau of Statistics' (CBS) files and the NTIC Safety Management System (SMS), over 160 pre-defined road sections of this kind were identified on the Israeli rural road network, with a total length of 1451 km.

For the study's road sections, detailed data were collected. The information on geometric and other road infrastructure characteristics came from a 2010 road survey, which was carried out by the NTIC and supplied the data to the NTIC SMS. Accident data and traffic volumes were extracted from the CBS files. Within the SMS, the data on road characteristics were collected for segments of about 100 meters in length. For the study's purpose, the SMS segments were combined into short road sections of about 1 km in length. For each short section, a range of infrastructure characteristics was produced, including lane width, shoulder width, speed limit, horizontal radius, side slope, convex vertical radius, concave vertical radius, vertical grade, and "safety level" of roadsides, in accordance with safety barriers' characteristics and recovery zone width. For some characteristics, e.g. lane and shoulder width, an average value, across the segments composing the road section, was applied; for each type of radius - a minimum value was estimated (as an average of the three lowest values across the segments), whereas for side slope and vertical grade - a maximum value was produced (as an average of the three highest values across the segments). In addition, for each type of radius, five categories were defined based on the existing design guidelines, with category "1" corresponding to the lowest and category "5" corresponding to the best design standard, where for each road section a certain category of each radius was assigned, respectively.

Finally, for roadside conditions, four categories of "safety level" were defined for the situation with and without safety barrier's presence. In the case with a barrier, barrier's position was examined, i.e. the distance from the right traffic lane and the free space behind the barrier and in front of rigid obstacles; in the case without barriers on roadsides, the recovery zone width was considered. In both cases, category "1" corresponded to the lowest and category "4" - to the best design standard, in line with the existing guidelines.

For each short section, an average category of "safety level" of the roadside conditions was estimated using the values of segments composing the section. When within the section there were segments with and without barrier's presence, a combined value of the "safety level" of roadside conditions was produced, as well, accounting for the length of barriers' presence.

The data on accident numbers, on each road section, were collected for five years, 2006-2010, according to several accident types, which are: (1) multiple-vehicle collisions, (2) single-vehicle accidents, (3) severe accidents (fatal and serious together), (4) total injury accidents, (5) all accidents (section accidents only, excluding junctions). The last group includes total injury accidents plus "general with casualties" accidents. In Israel, two types of accident files are collected by the police: an injury accident file, with cases investigated by the police examiners, which serves as a basis for the official accident figures, and a "general with casualties" file, with cases reported to the police but not investigated. The first file includes all injury severity levels: fatal, serious and slight, where slight injury accidents should satisfy certain selection criteria concerning the time passed since the occurrence and accident participants. The second file includes cases with slight injuries only (not hospitalized), which did not satisfy the selection criteria of the "injury" file. As the amount of records in the second file versus the first presents a 80% to 20% relation, it is reasonable to conduct analyses involving both accident files.

Having excluded road sections with irregular infrastructure characteristics and/or accident numbers, the study's database included 1358 short road sections with a total length of 1228 km. Table 1 provides descriptive statistics of the road section characteristics considered by the study. Table 2 shows the number of accidents observed on the study's sections, including the share of those where accidents occurred. It can be seen that injury accidents took place on a quarter of the study's sections, where any accident occurred on half of them. As expected, collisions, single-vehicle and severe accidents were rarer and reported on only 10%-16% of the sections.

**Table 1**
*Descriptive statistics of the road section characteristics considered by the study*

| Section characteristic | Variable name | N | Average | St.dev. |
|---|---|---|---|---|
| Traffic volume, thousand of vehicles | AADT | 1358 | 1.86 | 0.84 |
| Length of section, km | length | 1358 | 0.94 | 0.16 |
| Lane width, m | lane_w | 1358 | 3.19 | 0.35 |
| Speed limit, kph | sp_limit | 1358 | 79.4 | 15.0 |
| Shoulder width, m | sh_to_w | 1358 | 1.61 | 0.74 |
| Roadside conditions with safety barrier's | bar_est | 1080 | 2.24 | 0.67 |
| Roadside conditions without barriers | no_bar_est | 1321 | 1.23 | 0.30 |
| Roadside conditions, a combined estimate | rside_est | 1358 | 1.56 | 0.49 |
| Minimum horizontal radius, m | rh_min | 1338 | 1543 | 594 |
| Category of horizontal radius | rh_cat | 1338 | 3.88 | 1.47 |
| Minimum convex vertical radius, m | rvp_min | 1336 | 810 | 209 |
| Category of convex vertical radius | rvp_cat | 1336 | 1.22 | 0.63 |
| Minimum concave vertical radius, m | rvn_min | 1336 | -806 | 1225 |
| Category of concave vertical radius | rvn_cat | 1336 | 1.30 | 0.86 |
| Maximum side slope, % | ss_max | 1338 | 4.18 | 1.60 |
| Maximum vertical grade, % | gr_max | 1338 | 5.25 | 3.04 |

## 4. Development of explanatory models

### 4.1. Method

Various univariate and multivariate analyses were conducted on the data collected in order to develop explanatory models for the relationship between low-volume road characteristics and accident occurrences. In an early step of the analysis, the data were explored to understand the correlations among road characteristics in order to select characteristics with low correlation which are best predictors for accident occurrences. For this purpose such techniques as Pearson correlations and common factor analysis, were applied. Furthermore, using regression trees, by means of the *rpart* function of R software, and Generalized Boosted Regression Modeling (GBM), for each result variable (accident type), from the whole list of road section characteristics, a list of most important variables was selected.

**Table 2**
*Frequency of accidents observed on the study's sections, in 2006-2010*

| Accident type | No of accidents | No of sections where accidents occurred (% of all study sections) |
|---|---|---|
| Multiple-vehicle collisions | 345 | 217 (16%) |
| Single-vehicle accidents | 225 | 181 (13%) |
| Severe accidents | 179 | 141 (10%) |
| Total injury accidents | 598 | 341 (25%) |
| All accidents | 1900 | 680 (50%) |

Based on the literature, a strong relation is expected between the exposure variable and accident occurrences, which typically has an exponential form. Thus, an accident-exposure relationship was explored, using the logarithm of traffic volume and the logarithm of expected accident numbers (with section length as an *offset*). Generalized additive models were fitted to the data, using the *R mgcv* library of Wood (2006). A negative binomial regression was fitted for each dependent variable (accident type), where the explanatory variable was log(100*AADT). In order to explore the shape of the relationship between the exposure variable and each of the responses visually, the fitted smooth functions were plotted using the *plot.gam* function of the *mgcv* library. The plots showed a piecewise linear pattern, with a breaking point at 3.9 that is equal to AADT of 0.495. Further data examination revealed that below the breaking point there were 17 observations only; hence, it was decided to omit these cases and to fit explanatory models for sections with AADT over 500 vehicles.

For each accident type, firstly, a smooth (non-parametric) model was fitted, given the exposure, to explore a relationship between each infrastructure characteristic and accidents. Variables demonstrating no relationship were omitted from further consideration. On the other hand, for some variables, a broken (piecewise) relationship was recognized with the accident numbers and, thus, the range of values of such explanatory variable was split into two areas, allowing for a possibly different relation in each part. Secondly, a linear negative binomial regression model was fitted, for accident counts in five years, with log of section length as an *offset* and all the explanatory variables suggested by the GBM (excepting those that were omitted). The model parameters were estimated by means of the *glm.nb* function of the MASS R library (Venables and Ripley, 2002). Thirdly, more parsimonious models were fitted, in which explanatory variables were selected from the variables used at the previous stage, according to the minimum value of Akaike information criterion, using *stepAIC* function of MASS. Most of the explanatory variables remained in the final models were significant (with $p<0.05$), where some non-significant road infrastructure variables were also kept in the models if those were judged as essential from an engineering viewpoint.

The models were fitted to each of the five accident types that were introduced above. Each model estimates the expected accident number, on a road section, in five years. The general form of the model is:

$$E\{ACC\} = \exp \sum [\beta_i x_i] \qquad [1]$$

where *E{ACC}* is the expectancy of accident numbers on the road section; $x_i$ - section characteristics, $\beta_i$ - model coefficients.

### 4.2. Results

Table 3 provides an example of the model fitted to total injury accidents. The explanatory variables in the model are: traffic volume, lane width, shoulder width, horizontal radius (minimum value), side slope (maximum value) and section length.

The impact of shoulder width on accidents is different in two areas: below 2.5 m and over 2.5 m. In addition, a different relationship is observed between the side slope and accidents, depending on the value of horizontal radius: below 300 m (I_LT_300=1) or over 300 m (I_LT_300=0). The variables of traffic volume and shoulder width are significant, whereas those of lane width, horizontal radius and side slope are not (yet, impact of side slope is close to significant when the horizontal radius is below 300 m).
The fitted model parameters demonstrate that the number of accidents increases with an increase in traffic volume (impact coefficient, IC=0.32, $p<0.05$) and decreases with an increase in the lane width (IC= -0.09, p=0.61) and in the minimum horizontal radius (IC= -0.12, p=0.14), yet in the latter cases the impact is relatively low. As to the impact of shoulder width, the number of accidents increases when shoulder width increases up to 2.5 m (IC=0.41, $p<0.01$) and decreases when the shoulder width grows further, over 2.5 m (IC= -0.70, p=0.06). The number of accidents decreases with an increase in the side slope when the horizontal radius is below 300 m (IC= -0.60, p=0.05) but increases with an increase in the side slope when the horizontal radius is over 300 m (IC=0.31, p=0.26). The last relationship may reflect a positive impact of super-elevation under the conditions of sharp horizontal curves.

**Table 3**
*An explanatory model fitted to total injury accidents on low-volume roads*

| Parameter | Estimate | Std. Error | z value | Pr(>\|z\|) |
|---|---|---|---|---|
| (Intercept) | -2.670 | 1.205 | -2.22 | 0.027 |
| log(1000*AADT) | 0.321 | 0.125 | 2.56 | 0.010 |
| lane_w | -0.094 | 0.185 | -0.51 | 0.612 |
| sh_to_w | 0.413 | 0.126 | 3.26 | 0.001 |
| pmax(sh_to_w-2.5, 0) | -1.108 | 0.438 | -2.53 | 0.011 |
| log(rh_min) | -0.120 | 0.082 | -1.46 | 0.145 |
| sqrt(ss_max) | 0.314 | 0.281 | 1.12 | 0.264 |
| I_LT_300 | 2.043 | 0.936 | 2.18 | 0.029 |
| sqrt(ss_max): I_LT_300 | -0.915 | 0.418 | -2.19 | 0.028 |

*Notes to Table 3: variable names are given in Table 1. log - logarithm; sqrt - square root; pmax - maximum of the two values; I_LT_300 - indicator of the value of horizontal radius, I_LT_300=1 when horizontal radius is below 300.*
- *Model statistics: Dispersion parameter 0.482 (0.062); Residual deviance: 829.9 on 1232 degrees of freedom;*
- *2 x log-likelihood: -2069.458; AIC: 2089.5*

Table 4 summarizes the findings across the five models fitted to the five accident types. It can be seen that:
- The effect of various infrastructure characteristics is consistent across the models, though, the number of explanatory characteristics varies between the models;
- In all the models, as expected, an increase in traffic volume brings to an increase in accident numbers;
- An initial extension of shoulder width, up to 2-2.5 m, is associated with increase in accidents, while further shoulder widening, above 2-2.5 m, brings to an accident reduction;
- An initial increase of lane width, up to 3-3.25 m, brings to a reduction in most accident types, where further increase of lane width, above 3-3.25 m, is associated with an increase in single-vehicle and all accidents;
- The increase of horizontal radius is associated with injury accident reduction, including multiple-vehicle collisions and single-vehicle accidents;
- The value of side slope affects multiple-vehicle collisions, injury accidents and all accidents, where, in all cases, accident numbers increase with an increase in this characteristic's value, under the conditions of large horizontal radius (over 300 m), but decrease under the conditions of small horizontal radius, below 300 m. In other words, a positive effect of super-elevation under the conditions of sharp horizontal curves was observed;
- The improvement in roadside conditions contributes to a reduction in single-vehicle accidents, yet, the extent of the effect is relatively low;
- The vertical grade affects all accidents only, where higher positive grade is associated with accident increase.

The convex and concave vertical radiuses did not remain among the explanatory variables in any model. The impact of these road characteristics on accident occurrence on low-volume road sections was not found in the current study.

## 5. Quantitative tools for engineers

Based on the models developed, quantitative tools were built enabling to demonstrate the relationship between the infrastructure characteristics of low-volume roads and expected accident numbers. Those tools are: (a) a graphical presentation of the quantitative relations between each one of the road infrastructure characteristics and accident rates; (b) evaluation of accident modification factors associated with changes in the characteristics of road sections.

In the graphical presentation, the shape of the relation between the values of road characteristics and accident frequencies is given using a "multiplier" method, where for each change in the characteristic a relative change in accident numbers is shown compared to the value of the characteristic which was chosen as a basis for the comparison and for which the "multiplier" equals to "1". For example, using the "multiplier" method, Fig.1 demonstrates a relative change in total injury accidents following the change in lane width and shoulder width, related to basic values of the infrastructure characteristics: lane width 2.8 m; shoulder width 2 m. In line with the model introduced in Table 3, Fig.1,a indicates a monotonic decrease in accident numbers, following an increase in the lane width, where Fig.1,b shows an increase in accidents with an increase in shoulder width up to 2.5 m and a consequent decrease in accidents for higher shoulder width.

**Table 4**
*Summary of the effects of infrastructure characteristics on accident occurrence on low-volume roads, by accident type*

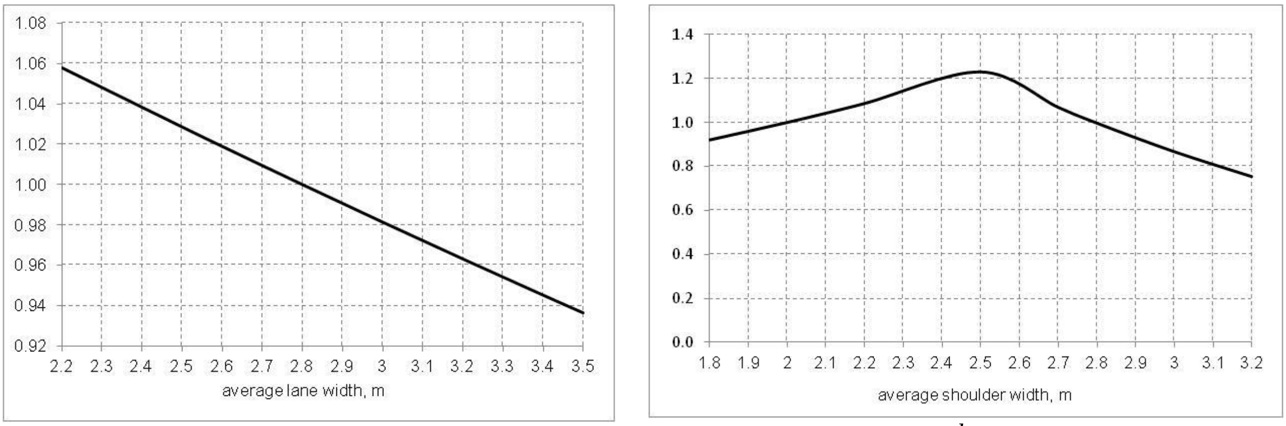
| Explanatory variable | Accident change following an increase in variable value | | | | |
|---|---|---|---|---|---|
| | **Multiple-vehicle collisions** | **Single-vehicle accidents** | **Severe accidents** | **Total injury accidents** | **All accidents** |
| Traffic volume, thousand vehicles | Increase | Increase | Increase | Increase | Increase |
| Average shoulder width, m | Increase (shoulder width up to 2.75 m) | Increase (shoulder width up to 2 m) | Increase (shoulder width up to 2.5 m) | Increase (shoulder width up to 2.5 m) | Increase (shoulder width up to 2 m) |
| Average shoulder width, m | Decrease (shoulder width above 2.75 m) | Decrease (shoulder width above 2 m) | Decrease (shoulder width above 2.5 m) | Decrease (shoulder width above 2.5 m) | Decrease (shoulder width above 2 m) |
| Average lane width, m | -- | Decrease (lane width up to 3.25 m) | Decrease | Decrease | Decrease (lane width up to 3 m) |
| Average lane width, m | -- | Increase (lane width above 3.25 m) | -- | -- | Increase (lane width above 3 m) |
| Minimum horizontal radius, m | Decrease | Decrease | -- | Decrease | -- |
| Maximum side slope, % | Increase (1) | -- | -- | Increase (1) | Increase (1) |
| Maximum side slope, % | Decrease (2) | -- | -- | Decrease (2) | Decrease (2) |
| Roadside conditions, a combined value | -- | Decrease | -- | -- | -- |
| Maximum vertical grade, % | -- | -- | -- | -- | Increase |

*Notes to Table 4: (1) where minimum horizontal radius is 300 m or larger; (2) where minimum horizontal radius is less than 300 m.*

Accident modification factor (AMF) estimates the change in safety following a change in a road design element, where it presents a relation between the expected number of accidents "with (infrastructure) improvement" to the expected number of accidents "without improvement" (Lord and Bonneson, 2006). Using the models developed by the study, AMF can be estimated as:

$$AMF=N_{new}/N_{old} = \exp [IC * (value_{new}-value_{old})] \qquad [2]$$

where $N_{new}$ and $N_{old}$ are expected numbers of accidents with *new* and *old* value of the explanatory variable (infrastructure characteristic); IC - impact coefficient of the variable estimated by the model; $value_{new}$ and $value_{old}$ are *new* and *old* values of the infrastructure characteristic, respectively.




*a* *b*

**Fig. 1**
*Multipliers of the total injury accidents on a low-volume road section related to change in: (a) lane width (with basic value 2.8 m), (b) shoulder width (with basic value 2 m)*

Applying this approach, accident modification factors associated with changes in the infrastructure characteristics of low-volume roads were estimated. For example, Table 5 provides AMFs for total injury accidents, associated with lane width changes. Each value in Table 5 shows an expected accident change where the lane width increases from an old value (given in row) to a new value (given in column). One can note that any increase in the lane width is associated with an accident reduction, where a higher accident reduction is expected following a larger increase in the lane width: in each row of Table 5 lower AMF values can be seen in the right end, with higher new values of the lane width.

**Table 5**
*AMFs for total injury accidents on a low-volume road section, related to a change in the lane width*

| **Lane width, m** | | **$value_{new}$** | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | **2.2** | **2.4** | **2.6** | **2.8** | **3.0** | **3.2** | **3.4** | **3.5** |
| **$value_{old}$** | **2.0** | 0.98 | 0.96 | 0.95 | 0.93 | 0.91 | 0.89 | 0.88 | 0.87 |
| | **2.2** | | 0.98 | 0.96 | 0.95 | 0.93 | 0.91 | 0.89 | 0.89 |
| | **2.4** | | | 0.98 | 0.96 | 0.95 | 0.93 | 0.91 | 0.90 |
| | **2.6** | | | | 0.98 | 0.96 | 0.95 | 0.93 | 0.92 |
| | **2.8** | | | | | 0.98 | 0.96 | 0.95 | 0.94 |
| | **3.0** | | | | | | 0.98 | 0.96 | 0.95 |
| | **3.2** | | | | | | | 0.98 | 0.97 |
| | **3.4** | | | | | | | | 0.99 |

## 6. Conclusions

The models developed in the study for measuring the relationships between the infrastructure characteristics and accident numbers on low-volume roads may assist in decision-making, both during the design and road maintenance processes or black-spot treatment. Concerning the safety impact of most infrastructure characteristics, the study's findings were in line with the international experience, except for the breaking points observed, in the local results, in the relationship between the lane and shoulder widths and accident occurrences.

Having considered the models' results, literature findings and existing rural roads' design guidelines (NTIC, 2011), recommendations on preferable infrastructure characteristics of low-volume roads, under Israeli conditions, were provided, such as: increasing shoulder width to 3.0 m; increasing lane width to 3.2 m; reducing the use of horizontal curves with radius below 300 m; eliminating gaps between the existing road conditions and the guidelines' demands related to the super-elevation values on curves; providing forgiving roadsides by means of widening recovery zone or installing safety barriers; adding warning measures on sections with high vertical grades. The aforementioned values of shoulder and lane width and of the horizontal curve radius are higher than the bottom thresholds permitted by the guidelines. However, they are suggested for application, in the local practice, due to related safety benefits.

In general, while reconstructing a low-volume road section, it is recommended to carry out a detailed examination, using the tools developed by the study, in order to compare the safety outcomes expected from various design alternatives.

**Acknowledgements**

Appreciation is extended to the National Transport Infrastructure Company of Israel that commissioned this study.

**References**

AASHTO, 2001. *Guidelines for Geometric Design of Very Low Volume Local Road*. American Association of State Highway and Transportation Officials.

Fitzpatrick, K., Brewer, M. 2004. *Summary of Treatments for Crashes on Rural Two-Lane Highways in Texas*. Report FHWA/TX-05/0-4048-S, Texas Transportation Institute.

Hakkert, S., Shokir-Hadad, L. 2009. *The impact of infrastructure of local rural roads on road accidents*. Research Report 318-2009, Transportation Research Institute, Technion, Israel.

Hall, J.W., Rutman, E.W., Brogan, J.D. 2002. *Highway Safety Challenges on Low-Volume Rural Roads*. Final Report, Federal Highway Administration, New Mexico Division.

Harwood, D. W., Council, F. M., Hauer, E., Hughes, W. E., and Vogt, A. 2000. *Prediction of the Expected Safety Performance of Rural Two-lane Highways*. Publication No FHWA-RD-99-207, Federal Highway Administration.

Lord, D. and Bonneson, J.A. 2006. Role and Application of Accident Modification Factors within Highway Design Process. *Transportation Research Record* 1961: 65-73.

Mallschutzke, K., Galvez, I., Gatti, G., Van de Leur, M., Dell, O.L. 2006. *Secondary road safety: Actual knowledge and experiences*. RIPCORD-ISEREST research project.

Mattar-Habib, C., Polus, A. and Farah, H. 2008. Further evaluation of the relationship between enhanced consistency model and safety of two-lane rural roads in Israel and Germany. *EJTIR* 8(4): 320-332.

NTIC, 2011. *Guidelines for geometric design of roads*. Ministry of Transport and National Transport Infrastructure Company, Israel.

OECD, 1986. *Economic design of low-traffic roads*. Organization for Economic Co-operation and Development, Paris.

Polus, A. and Cohen, M. 2010. Models of Accidents on Low-Volume Rural Roads. *TRB 89th Annual Meeting Compendium of Papers*, Transportation Research Board.

Reurings, M., Janssen, T., Eenink, R., Elvik, R., Cardoso, J., Stefan, C. 2005. *Accident prediction models and road safety impact assessment: a state-of-the-art*. RIPCORD-ISEREST research project.

TRB, 2003. Eight International Conference on Low-Volume Roads 2003. *Transportation Research Record* 1819, Transportation Research Board.

Zegeer, C. V., Deen, R. C. and Mayes, J. G. 1981. Effect of Lane Width and Shoulders Width on Accident Reduction on Rural Two-lane Roads. *Transportation Research Record* 806: 33-43.

Venables, W. N. and Ripley, B. D. 2002. *Modern Applied Statistics with S.* 4th edition. Springer.

Wood, S.N. 2006. *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC.

# A TRUCK PARKING FACILITIES ADEQUACY STUDY IN BRAZILIAN FEDERAL HIGHWAYS

**Amir Mattar Valente[1], Caroline Mazzuco Furlan[2], Eduardo Lobo[3], Jorge Destri Jr[4]**
[1,2,3,4] *Universidade Federal de Santa Catarina, Brazil*

**Abstract:** According to Brazilian reports on accidents occurred on its Federal highway network involving large trucks, accidents are mainly caused by drivers fatigue and excess of speed while trying to meet short delivery deadlines for long distances. A new federal law was approved to establish limits on truck drivers working times, including maximum continuous driving time and maximum daily and weekly on duty time. Due to this law and since the Brazilian Ministry of Transportation did not perform studies to verify its applicability, we present a research that aims to analyze which existing rest areas provide enough room for the actual and future demands of trucks needing to stop. We establish the following method: in the first step we review and study different models found in the literature and evaluate their potential application to Brazilian conditions. In the next step we develop our own model and finally, we validated our model by applying it to the BR-101 Federal highway in the Santa Catarina State. Therefore, this study describes the models evaluated in the first step, presents our new model and the results obtained on its application.
**Keywords**: maximum driving time, minimum rest period, truck, rest area, parking space

## 1. Introduction

Considering that many highway accidents in Brazil involve professional truck drivers, regulating the professional's maximum driving time and working hours became essential (DNIT Partial Report – 2R05). Law 12.619/2012 (BRASIL, 2012), enacted by the Brazilian government, regulates the working hours of freight and passenger vehicles drivers, aiming to decrease the amount of accidents caused by the fatigue of excessive driving time.
A partnership between the National Department of Transport Infrastructure (DNIT), Ministry of Transports and the Transports and Logistics Laboratory of the Federal University of Santa Catarina (LabTrans/UFSC) has the intent of create a calculation methodology that allows the evaluation of the adequacy of Brazilian highways to attend the demands of the new law. To stablish the new method, a three-step process was adopted: analysis, implementation and validation, which are described respectively in sections 2, 3 and 4. In the first stage, the existing literature models are revised, using as reference the United States (US Hours of Service. 2011), Canada (CCMTA, 2007) and the European Union (EUROPEAN UNION, 2006). In the implementation stage, an appropriate model to Brazilian reality is elaborated, and in the validation stage, the new method is applied to a stretch of a Brazilian federal highway as a pilot project. When the method is properly developed, a computational tool will be available to allow its application by associated government agencies. In this study are presented the results of the first two stages and part of the method evaluation.

## 2. Theoretical Foundation

Law 12.619/2012 provides for the exercise of the profession of driver, regulating and disciplining work hours and driving time of the professional driver. According to this law, a professional driver is prohibited to drive continuously for a period longer than four hours; after this time, the driver must respect a resting time of 30 minutes. However, there is a grace period of up to one hour to allow the driver to find an adequate and secure resting site. In addition, the driver must obey a resting time of 11 hours in a 24 hours period, which may be divided in two periods, one of nine hours, other of two hours, as long as they occur in the same day.
The law implies in a relation between supply and demand concerning the amount of drivers that must stop and the availability of parking spaces, since the drivers are obligated to observe the resting periods, being them short or long, using a limited amount of parking spaces. In this context, among the analyzed references, three models stand out, which represent the state of the art.
The model created by Lüttmerding (2009) calculates the parking demand of trucks throughout a stretch of highway to each of its traffic directions; the demand for truck parking spaces (TPS) by kilometer, is based on the parameters of average daily traffic volume (DTV) in a determined direction; average daily driving distance (ØDD) that he drives in the segment; average parking duration (ØPD); night-time or the period commonly used by the drivers to rest ($t_N$); the proportion of the cumulated pausing time within night-time ($P_N$) and the proportion of long-distance traffic within ($P_{LD}$). The resulting values are usually applicable only to the surveyed stretch or region where they were obtained. As a limitation or restriction to the model is the fact that it does not involve parameters related to the preferences of the drivers. The demand is expressed in number of parking spaces per kilometer and determined by the following calculation:

$$\frac{TPS}{km} = \left(\frac{DTV}{\phi DD}\right) \times \left(\frac{\phi PD}{24}\right) \times \left(\frac{24}{t_N}\right) \times P_N \times P_{LD}$$

[1] Corresponding author: amir.valente@ufsc.br

In the second analyzed model, originally developed by the Minnesota Department of Transports/USA (MnDOT) and presented by Garber and Wang (2004), the demand is expressed in number of parking spaces (NTSPACES). According to this model, determination of the number of parking spaces follows the parameters: average daily traffic (ADT); total percentage of mainline traffic stopping at rest area/truck parking (P); peak factor comparing the mean values of the months with higher traffic with the annual daily average (FP); design hour usage compares the design hourly volume to the AADT (DH); percentage of truck parking spaces ($D_t$) and number of vehicles parked per hour per space (VHS). As time goes by, other institutions as the Virginia Department of Transportation (VDOT) and the Apogee Research, developed researches to verify, or even adequate, these parameters to new scenarios/case studies. Such as the model proposed by Lüttmerding (2009), this one does not consider the factors related to drivers preferences and to legal questions relative to the HOS (Hours of Service of USA, 2011) regulation. The demand in this model was calculated as follows:

$$NTSPACES = \frac{ADT \times P \times DH \times D_t \times PF}{VHS}$$

The third model is based in the fact that in the USA there are two kinds of places where freight vehicles may stop: *rest areas*: places operated by state public service, with minimum services (water, toilets, self-service machines), which usually presents restrictions about the maximum standing time; and *park areas*: private places which present an improved infrastructure to attend the drivers necessities in short or long duration stops (Georgia Department of Transportation, 2011). The model adopted by the Federal Highway Administration – FHWA (2002), in its full version, have different provisions for vehicles that cover long distances (need more stops) and for those that cover short distances, as well as for the kind of preferred stopping places. Considering only the vehicles that cover long distances and that do not have a preferred stopping place, the average truck-hours of travel per day (THT) is estimated with the parameters: percent of vehicles that consists of commercial trucks ($P_t$); the annual average daily traffic (AADT); the length of the roadway segment (L) and the average speed of the trucks (S) (COLEMAN and TRENTACOSTE, 2002).

$$THT = P_t \times AADT \times \frac{L}{S}$$

Next, the following parameters are estimated: the average stop time for each driving hour for long-haul drivers ($T_{DRIVING}$); the time at home for long-haul drivers ($T_{HOME}$); the loading and unloading for long-haul drivers ($T_{LOAD/UNLOAD}$); the time at shipper/receiver for long-haul drivers ($T_{SHIPPER/RECEIVER}$) – the last term refers to the average stop time per hour throughout a daytime, with other purpose than rest.

$$PR_{LH} = \frac{\left(8days \times \frac{24hr}{day}\right) - T_{DRIVING} - T_{HOME} - T_{LOAD/UNLOAD} - T_{SHIPPER/RECEIVER}}{T_{DRIVING}} + \frac{5\min}{60\min} = 0.7833$$

The simplified demand model (D) is based on the total truck-hours of travel per day (THT) and the average parking time per truck-hour of travel ($P_{avg}$).

$$D = THT \times P_{avg}$$

Between the three presented models, the one adopted by FHWA (2002) is considered the most similar to the one proposed in this study, since its parameters consider short and long stopping time. However, similar to the other ones, it does not consider attractiveness factors.

### 3. Proposed Methodology

The methodology elaborated in this study follows a basic assumption, which was determined by the entities involved in the project, i.e. the division of the Brazilian highway network in segments delimitated by Federal Highway Police (PRF) stations, as illustrated in Figure 1. Analyzes of parking spaces supply and demand are made for each segment individually, and data are calculated separately by traffic direction, consolidated in cases where there are no restriction to lane transposition. Both short and long stopping times provided in law are considered.

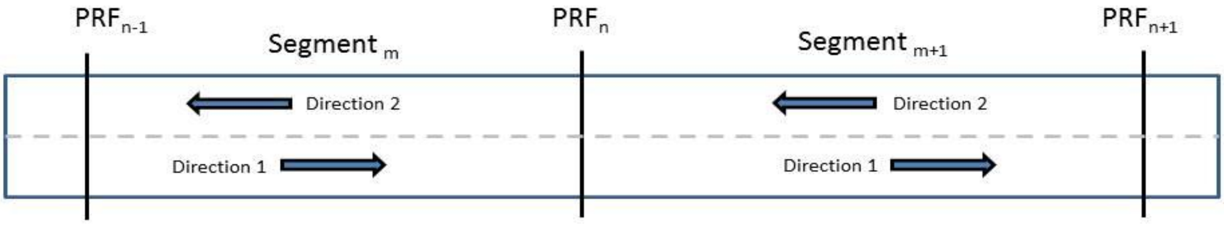



**Fig. 1.**
*Representation of the limits of the segments and the respective directions.*
*Source: LabTrans*

Initially, the demand for vacancies is calculate, i.e. how many drivers need to stop, and for how long, in a determined highway segment; then, the vacancy availability in the same segment is evaluated. Once in possession of these two values, the excess or shortage of parking spaces for the segment may be assessed, so it may support the pertinent decision making for each case.

1. The methodology consists in the following steps:
2. Obtaining legal parameters relative to the restriction to work hours, in other words, minimum rest period, maximum time between rest periods, maximum time of continuous driving, minimum stopping time between driving hours;
3. Determine coefficients that will be used as safety margin to the calculations;
4. Determine the highway segments to be assessed;
5. Obtain, for each segment, its operation parameters, i.e. daily average traffic volume (VMD), percentage of the VMD per traffic direction, VMD's frequency distribution throughout the day, travel time probability distribution for each hour of the day (due to high cost and difficulty in the achievement of this parameter, one hour may be used for the whole day, considering the peak hours, or one for each representative period, as morning, afternoon and night), percentage of drivers who need to stop for rest at each hour of the day;
6. Identify, for each segment, the facilities that may be used by freight vehicle drivers to stop and rest in each direction of the highway, registering the number of available parking spaces in each one;
7. Calculate the expected traffic volume per hour for each segment and traffic direction, based on the VMD, the division of the traffic by direction and the frequency of the VMD throughout the day;
8. Calculate the expected number of short length stops per hour for each segment and traffic direction, based on the expected traffic volume, travel time probability distribution, legal parameters and safety values;
9. Calculate the expected long length stops (rest) per hour for each segment and traffic direction, based on the expected traffic volume, expected percentage time for rest, legal parameters and safety values;
10. Calculate the availability of parking spaces through the sum of the existing parking places;
11. Verify the excess or shortage of parking spaces in each segment, compared to the expected number of short length stopping time plus the expected number of long length stopping time with the calculated offer, grouped by traffic direction, if reasonable.

**4. Evaluation of the Proposed Methodology**

With the aim of validate the elaborated method, a pilot project was executed in a stretch of the highway BR-101, which crosses the State of Santa Catarina, located in South Brazil. This stretch is 465 km long and counts with seven PRF stations as shown in Figure 2.


**Fig. 2.**
*Federal highway BR-101, Santa Catarina State, and Federal Police Stations.*
*Source: LabTrans*

Operational data were obtained from four different sources:

- Government agencies;
- Highway toll operators;
- Quantitative research for the gathering of data relative to the times of driving, stop and rest made in the PRF stations. 3,343 freight vehicle drivers were interviewed, an average of 238 interviews per station, per traffic direction. In two segments the parked vehicles were also counted;
- Traffic volume classification counts.

Data on facilities with parking spaces were obtained from a research made by the Ministry of Transports. Figure 3 shows the distribution curve of the average travel time (since the last stop) for a sample of 212 trucks, in the segment between the Joinville and the Barra Velha PRF stations, in the south-north direction.

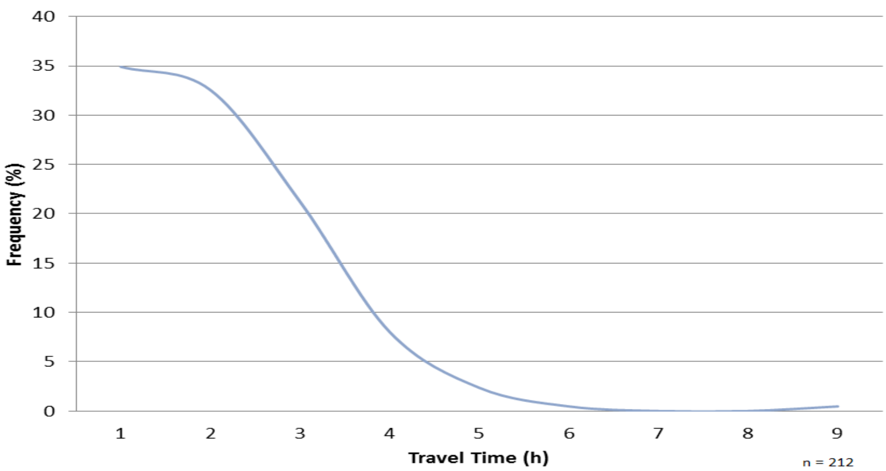



**Fig. 3.**
*Frequency Distribution Curve by travel time*
*Source: LabTrans*

For the same segment and traffic direction, Figure 4 shows, in percentages, how the volume of heavy vehicles is distributed throughout 24 hours, considering a typical day.
At the present, data relative to traffic volume for a segment are still being collected, while some data originated from the quantitative research are receiving statistical treatment. For this treatment, a mixed linear model is being used, were several observations on the same experimental unit throughout the time are being gathered (PINHEIRO and BATES, 2000.), with the aid of the software R (BATES, 2010). Longitudinal planning is frequently used in several research areas, since they allow evaluating global or individual changes in a time scale (SINGER; NOBRE; ROCHA, 2009), in way to eliminate outlier values.
Despite this situation, a prototype computer program was created to receive data and perform the calculations predicted in the methodology. Available data are already stored in the program and the missing ones were extrapolated from existing data in non-official sources. Initial tests were performed to verify the consistency and the integrity of the proposed calculations. From the initial tests two facts caused alterations in the methodology initially proposed: first, the demand generated by the stopping time (long haul) should reflect in the demand in the following hours from the beginning; second, there was no need to represent the travel time since the last stop with a probability distribution, the percentage time necessary to stop would be enough, at each hour of the day. After the adjustments in the model and in the program, new tests were performed and the results were consistent with the proposal, even using some fictional data.

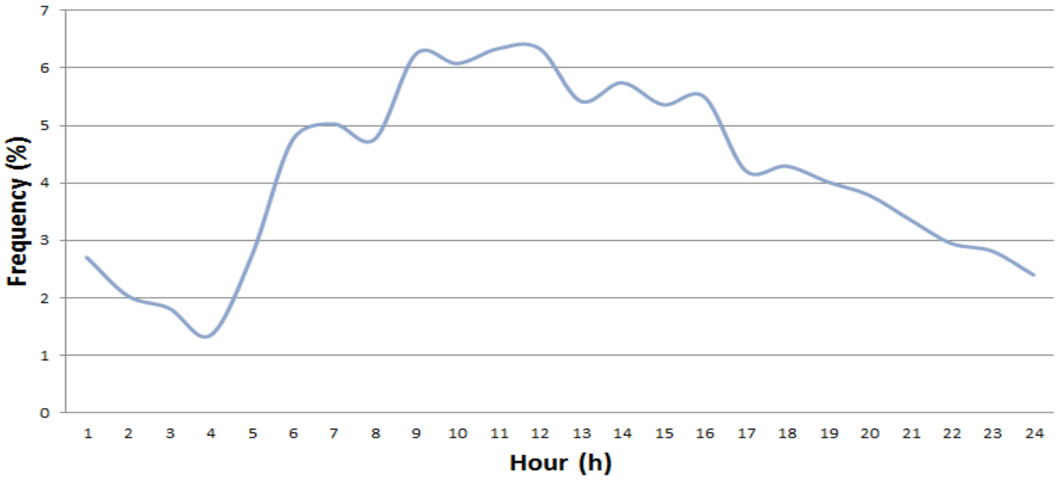



**Fig. 4.**
*Volume Frequency Distribution Curve*
*Source: Autopista Litoral Sul/LabTrans*

## 3. Conclusion

The results of the tests performed so far allow us to affirm that the method is adequate to its purpose. However, tests using all data from the pilot project are still to be performed to verify the sensitivity of the method when using real data, and if the results reflect the reality of the situations. Modifications in these indicators may suggest the necessity to adjust the safety values or to refine data on travel time and rest time. In a next stage, the method should be applied in other places with different characteristics from those found the pilot project: few available parking sites, extended segments – typical situation in the North, Northeast and Center-West Brazilian regions.

## Acknowledgements

The authors would like to thank the National Department of Transport Infrastructure (DNIT) for sponsoring this research.

**References**

Bates, D. M. (2010) – lme4: Mixed-effects modeling, with R. Springer. http://lme4.r-forge.r-project.org/book/

Canadian Council of Motor Transport Administrators - CCMTA. Commercial Vehicle Drivers Hours of Service Regulations of 25 October 2005 - Application Guide. 2007. Available from Internet: <http://laws-lois.justice.gc.ca/eng/regulations/SOR-2005-313/page-1.html>.

Coleman, J.A., Trentacoste, M.F. 2002. Model Development for National Assessment of Commercial Vehicle Parking. FHWA-RD-01-159. Federal Highway Administration, U.S. Department of Transportation

Department of Transportation, US. Commercial Driver Rest and Parking Requirements: Making Space for Safety. 1996. FHWA-MC-96-0010, Federal Highway Administration, U.S.

Laboratório de Transportes e Logística da UFSC – Universidade Federal de Santa Catarina. Estudos, pesquisas e programas de capacitação para desenvolvimento e consolidação de métodos e processos para suporte à gestão de competências da CGPERT vinculadas às áreas de segurança viária e operações rodoviárias. Objeto 2 – Segurança Viária. 2R05 – Relatório Parcial. 2013. DNIT - Departamento Nacional de Infraestrutura de Transportes.

Garber, N. J., Wang, H. 2004. Estimation is the Demand for Commercial Truck Parking on Interstate Highways in Virginia – Final Report, Virginia Transportation Research Council.

Georgia Department of Transportation.Georgia. Statewide Freight and Logistics Plan - Truck Modal Profile. 2011.

Hours of Service of Drivers. 2011. FMCSA – Federal Motor Carrier Safety Administration. U.S. Department of Transportation. Available from Internet: <http://www.gpo.gov/fdsys/pkg/FR-2011-12-27/pdf/2011-32696.pdf>.

Brasil. Lei Nº 12.619/2012 de 30 de abril de 2012. Dispõe sobre o exercício da profissão de motorista; altera a Consolidação das Leis do Trabalho – CLT. Available from Internet: <http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2012/lei/l12619.htm>.

Lüttmerding, A. 2009. Calculation model for estimating the demand of truck parking areas along motorways. In *Proceedings of the 12th International Conference on Transport Science, Portorož, Slovenija.*

Pinheiro J.C. e Bates, D. Mixed effects model in S and S=PLUS, New York: Springer – Verlag, 528p., 2000

European Union. Regulation (Ec) Nº 561/2006 of the European Parliament and of the Council of 15 March 2006. Available from Internet: < http://eur-lex.europa.eu/resource.html?uri=cellar:5cf5ebde-d494-40eb-86a7-2131294ccbd9.0005.02/DOC_1&format=PDF >

Singer, J.; Nobre, J.S.; Rocha, F.M.M. Análise de Dados Longitudinais. Apostila, 158p., 2009.

Study of Adequacy of Commercial Truck Parking Facilities –Technical Report. 2002, FHWA RD-01-158. Federal Highway Administration, U.S. Department of Transportation.

Trucking Research Institute, Apogee Research, Inc., and Wilbur Smith and Associates. 1996. Commercial Drivers Rest Area Requirements: No Room at the Inn. Federal Highway Administration.

# MAINTENANCES OF URBAN ROADS INFRASTRUCTURE IN KOGI STATE, NIGERIA

**Musilimu Adeyinka Adetunji**[1]

[1]*Department of Geography, Faculty of Arts & Social Sciences, Federal University Lokoja, Kogi State, Nigeria*

**Abstract:** Poor designed and maintenance of urban roads in Nigeria is great obstacles to the free movement of people, freights and information. Many urban commuters are made to waste valuable period of time in their intra city movement. It is on this premise that this study examines the urban road infrastructure and how they are maintained in Lokoja metropolis, a city located in the north central part of Nigeria. In the ancient city of Lokoja, three major roads were randomly selected from three tiers of governments in Nigeria to assess the maintenance of the urban transport routes in the city. A systematic random sampling technique was employed to select an average of fifteen buildings along the road segments on the principle of one in every 20th building. In each of the residential building sampled, only one head of household is selected and interviewed on the quality of road networks, provision and their maintenance. Descriptive statistics was used to analyze the data. The finding reveals that many of the urban routes in the city do not either pedestrian walkways or roads media which compounded mobility crises in the city. Pedestrians, motorcyclists and motorists and other road users compete on urban routes for their day to day transactions. Further analysis shows that more than 60% of urban routes in Lokoja have non drainage system. Those roads that have drainage facility have been filled up by sand from the erosion activity. Many urban routes have no traffic light. The erratic power supply has made it difficult for urban commuters to travel in the night. The study therefore concludes that further research needs to be conducted on the designing of the urban road networks in Nigeria so as to achieve sustainable transport development.

**Keyword:** urban roads, drainage, pedestrian walk way and traffic light

## 1. Introduction

Transportation plays a significant role in socio-economic and political development of any country of the world. It enhances the evacuation of goods from the producers to the ultimate consumers, overcome the friction of distances, promote business transactions and improve the standard of living of the people of the country who have well developed functioning and efficient transport system. These include road, rails, water and air transport. Of all the modes, road transportation has been the most utilized in Nigeria's urban transportation system (Oni, 2004). It is pertinent to note that as at 1995, it was estimated that the road network in Nigeria had an asset nominal replacement value of almost 20billion US Dollars. The annual loss due to bad roads is valued at N80 billion, while additional vehicle operating cost resulting from bad roads is valued at N53.8 billion, bringing the total loss per annum to N133.3 billion (Oni, 2004; Arizona-Ogwu, 2008). This figure excludes the hours lost on roads due to traffic congestion, physical and psychological trauma people pass through during their trips (Arizona-Ogwu, 2008). Also, Arizona-Ogwu (2008) affirms that only 10% of the fund requested by the Federal Ministry of Work and Housing was released by the federal government between 1999 and 2002. This is an indication of nonchalant attitude towards road maintenance in the country. To worsen the mobility crises in Nigeria, railway system which was designed in 1908 in Nigeria to enhance the movement of bulky goods has collapsed in the 80's due to corruption, politics and poor management.

An overview of intra-city movement in Nigeria reveals that more than 80% of urban trips are made on road and commuters depend on foot for short distances, while public transport modes (para-transit) are patronized for long distances ( Adetunji, 2013; Ogunsanya,2004). Incidentally, the designing of urban transport routes in Nigeria did not consider the pedestrian walkway, and other transport facilities such as street lights, fly overs, drainage and many others which are germane to easing the movement of people, goods and information during intra-city movement. It is important to note that the provision of these road transport facilities enhances the free flow of vehicular transport and guarantees the safety of pedestrians and other road users. For instance, the design of pedestrian walkways help to provide access for all types of pedestrian travel to schools as well as work, parks, shopping areas, transit stops and other destinations. (http://guide.saferoutesinfo.org/engineering/sidewalks.cfm). A study of the California SRTS programme has shown that providing sidewalks is one of the most effective engineering measures in encouraging children to walk to school (Boarnet et al, 2005).

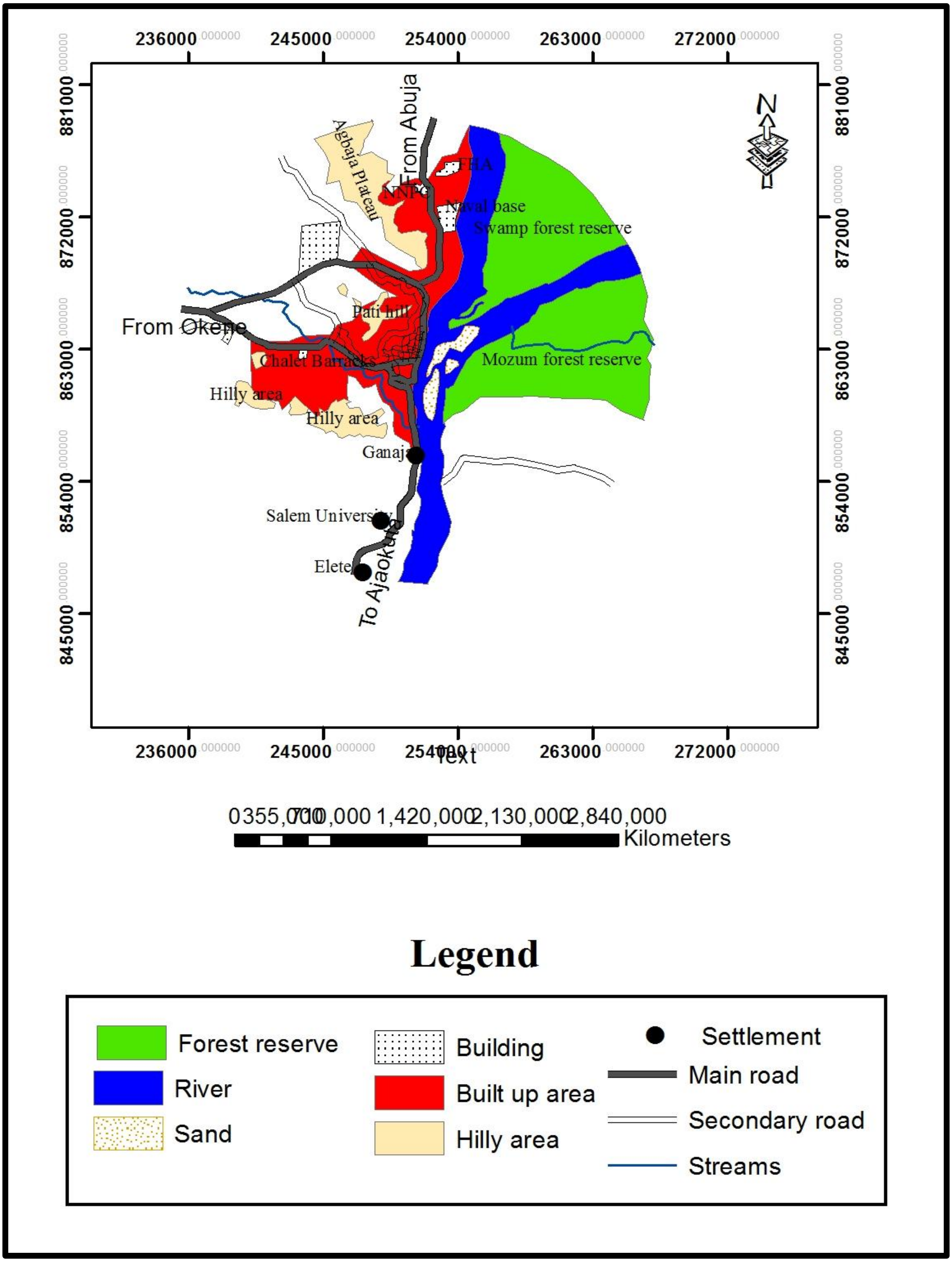
Similarly, the provision of street light is equally important as this will enable commuters to see each other clearly and ensure safety to make their trip easier during the night journey ( Pegrum, 1972; Freedman et al, 1975) . It is disheartening that in some of the African countries, most especially in Nigeria, some engineers and architects who copied the road network design in developed countries into urban centres in Nigeria do not consider these specifications during road construction. This invariably poses some treat to pedestrians and motorists on highways. It is on this premise that this study is designed to examine the existence and maintenance of transport infrastructure in Lokoja, the first administrative headquarters of the Northern Nigeria during the colonial era.

[1] Corresponding author: maadetunji@yahoo.com

## 2. Study Area

Lokoja metropolis is the study area and it is located on the confluence of River Niger and River Benue. The built up area of Lokoja is located in the eastern parts of River Niger, while the surrounding rural communities in the city are located at the western part of River Benue. These two rivers serve as mode of transport to Lokoja by rural communities living at the riverine areas of the state. According to Olawepo (2009), Lokoja became the headquarters of Kogi Local Government Area as far back as 1976 and was later made Kogi State capital in 1991. Since then, there have been massive changes in all activities of Lakoja, including its size, structures, population and other socio-economic status. Within the last three decades, Lokoja metropolis like many other urban centres in Nigeria has witnessed tremendous population increase. The phenomenon increase in the population over the years has put more pressure on road transport system, which is the principal mode of transportation in Lokoja and many other cities in Nigeria.

Three types of road networks are found in Lokoja. These are Trunk A- Federal Roads (highways), Trunk B- State Roads, and Trunk C- Local Government Roads. The Trunk A Roads comprise the federal highway that connects Lokoja to other towns and states of the federation. The Trunk B- State Roads are those that link the built –up areas in Lokoja where different activity centres are taking place on daily basis. The last category of urban route in Lokoja metropolis is Trunk C- Local Government Roads, which connect residential neighborhoods and low density area of the city. Some of these roads are earth surface roads, with no drainage and other route transport facilities. It is against this background that this study was designed to examine the provision, utilization and maintenance of transport infrastructural facilities in Lokoja, a former capital territory of Northern Nigeria during the colonial administration. This will enable the urban planners, policy formulators, government and other stakeholders to develop an appropriate transport policy that will cater for the designing and maintenance of transport infrastructure for Lokoja and other similar cities in Nigeria.


236000
245000
254000
263000
272000
881000
872000
863000
854000
845000
From Abuja
Agbaja Plateau
FHA
NNPC
Naval base
Swamp forest reserve
Pati hill
From Okene
Chalet Barracks
Mozum forest reserve
Hilly area
Hilly area
Ganaja
Salem University
Elete
To Ajaokuta
Text
0355,000710,000 1,420,000 2,130,000 2,840,000
Kilometers
Legend
Forest reserve
River
Sand
Building
Built up area
Hilly area
Settlement
Main road
Secondary road
Streams

## 3. Materials and Methodology

Primary and secondary sources of data were used for this research. Two main sources of primary data were employed, namely, household survey and measurement of pedestrian walkways. Also, field observations were equally adopted to check the functionality of drainage system as well as the condition of road surface on the sampled routes in the study area. The household survey was carried out in the first week of May, 2014, using structured questionnaire to solicit information on the households' perception on the provision, utilization and maintenance culture of transport infrastructure connecting their residential units. In an ancient city of Lokoja, where the research was conducted, three roads each were randomly selected from Trunk A- Federal Roads (highways), Trunk B-State Roads, and Trunk C- Local Government Roads to assess the maintenance of the urban transport corridors in the city. This implies that information was collected on nine roads selected from all areas of the city. In each of the sampled road, a systematic random sampling technique was employed to select an average of fifteen buildings each along the selected road segments on the principle of one in every 20$^{th}$ building. In each of the residential building sampled, only one head of household is selected and interviewed. Once the initial household has been selected for questionnaire administration, others follow in sequential order. Descriptive statistics such as tables of percentages was used to analyze the data.

## 4. Data Analysis and Discussion

### 4.1: Characteristics of Roads transport Infrastructure in Lokoja

The findings reveal that three categories of road networks are found in Lokoja metropolis. These include Trunk A- Federal Roads (highways), Trunk B-State Roads, and Trunk C- Local Government roads (Table 1). The Trunk A routes is federal highways that connect Lokoja with other towns or states of the federation. These include Abuja- Lokoja route, Lokoja- Ajaokuta / Ganaja road and Muritala Muhammed way. These routes are dual carriage way, well tarred with bitumen and motorable throughout the year but with numerous potholes on them in nearly every two electric poles along the road segments. The pedestrian walkways are completely absent along these major routes except from Ganaja junction- Bank of the North where the pedestrian walkway of less than 3metre is designed. The resultant effect of non-consideration of pedestrian's walkway, Zebra crossing and fly overs by Nigerian engineers and architects during road constructions constitute conflict between pedestrians and vehicular traffic. The utility of provision of sidewalks on road transport networks are documented elsewhere. For instance, research commissioned for the Florida Department of Transportation in 2005, reveal that in Florida, the Crash Reduction Factor resulting from installation of sidewalks average 74% (Gan Albert; Joan Shen, Adriana Rodriquez (2005) In a similar study, the U.S. Department of Transportation, University of Carolina, affirmed that presence or absence of sidewalks and the speed limit are significant factors in the likelihood of a vehicle/pedestrian crash. Sidewalk presence had a risk ratio of 0.118, which means that the likelihood of a crash on a road with a paved sidewalk was 88.2 percent lower than one without a sidewalk (McMahon, et al, 2002). Further analysis reveal that Trunk B- State Roads are those routes that link different urban land use activities centres such as residential, commercial, public and recreational land uses. Notable numbers of these routes are Taiwo Road- Post Office, Stadium Road, Ado Ibrahim Road, Beach Road, Adankolo Roads, Aliu Attah Roads, Jane Ekundayo Road, Hassan Katisina Way/ Zone 8 - Crusher and Sule Oyidi Street. Some of these routes are tarred with one lane except, Hassan Katisina Way- Crusher and Taiwo Road that have dual carriage ways. Also, some parts of the length of Janet Ekundayo route are dualized. Initially, some of these roads had good drainage system but today more than half of them have been blocked by sand from erosional activities. Also, some of these roads drainage have been filled up by waste generated by urban residents in Lokoja metropolis. The Trunk C –Local Government Roads are those that link residential units as well as low density part of the city. The last category of this route has earth surface roads, with little or no drainage system, no pedestrian walkway and others transport infrastructure.

**Table 1:**
*Characteristics of Roadway in Lokoja Metropolis*

| Roadway | Number of Lane/ Width of Road | Road Surface and Drainage |
|---|---|---|
| I Trunk A- Federal Roads(highways) | | |
| (i) Abuja- Lokoja route, | Dual carriage way | Well tarred with smooth surface in some parts of the length, but with conspicuous numerous potholes, erosional sand has encroached the road surface particularly at Kogi State Polytechnic, No drainage |
| (ii) (ii) Lokoja- Ajaokuta / Ganaja road | Dual carriage way | Tarred with smooth surface, Non-pedestrian walkway, partial drainage |
| (iii) Muritala Muhammed way. | Dual Carriage | Partially smooth surface, numerous potholes, partial pedestrian walkways |
| II. Trunk B- State Roads | | |
| (i) Adankolo Road | One single lane road | Fairly smooth surface, Partial drainage facility filled up with sand and household waste generated |
| (ii) Cemetery Road | One single lane road | Tarred, smooth surface, Partially blocked drainage facility, no pedestrian walkway |
| (iii) Karaworo Road/ Palace Road | One lane | Earth surface road, rough surface, no drainage , no pedestrian walkway |
| III. Trunk C – Local Government Roads | | |
| (i) Sango Daaji Palace Road | One single lane | Earth surface road, rough surface, no drainage facility, No pedestrian walkway |
| .(ii) First Gate- Phase Two Housing Estate- Arigbede Secondary School road | One Single Lane | Earth surface road, rough, no drainage facility, no pedestrian walkway |
| (iii) Lokongoma Phase One Gate route- Chavid Hotel | One lane | Fairly smooth surface, partial drainage, many potholes |

*Source: Author's Field Survey, 2014*

### 4.2 Road Maintenance Culture

Generally, three tiers of governments, namely, federal, state and local governments are responsible for road construction, rehabilitation and maintenance in Nigeria. The divided responsibility in road maintenance among the three tiers of government has adverse effects on the quality of road transport infrastructure. On several occasions, many of the intra-city roads which have reached deplorable conditions are left for federal government to maintain. Sometimes it is very difficult to differentiate between, Truck A- Federal, from Truck B- State and Truck C-Local Government Roads in urban centres in Nigeria and these tiers of government often shift their responsibility to one another for maintenance of road transport infrastructure. Lokoja Metropolis is not exempted from this problem. Table 2 reveals that 10. 7% of the road transport infrastructure in Lokoja are constructed and maintained by the Federal government, 55% of these roads are maintained by the State government, while, 34.3% are shared by both the Local government and community who live at the city suburb in Lokoja metropolis.

**Table 2:**
*Agency responsible for Road Construction and Maintenance in Lokoja*

| • Road Construction & Maintenance | • Percentage of Contribution |
|---|---|
| • Federal Government | • 10.7% |
| • State Government | • 55 %. |
| • Local Government/ Community | • 34.3% |
| • Total | • 100.0% |

*Source: Author's Field Survey, 2014*

### 4.3 Road Drainage Facility in Lokoja

Efficient and well-functioning drainage facility is the most important factor to be considered in road design and construction. The Truck A- federal roads in Lokoja have partial drainage; some of these drainage facilities are blocked by sand from erosional deposits and waste generated by urban dwellers. Also, the Truck B- State Roads in Lokoja metropolis have drainage facility but they are poorly maintained both by the state as well as the inhabitants of the city. Table 3 reveals that 63.7% of the respondents in the city claims that they do not have functioning roads drainage facility connecting their residential unit, while the remaining 36.3% of the inhabitants indicates that they have good drainage facility.

**Table 3:**
*Designing of Road Drainage Facility in Lokoja*

| • Drainage Facility | • Percentage |
|---|---|
| • Yes | • 63.7% |
| • None | • 36.3% |
| • Total | • 100% |

*Source: Author's Field Survey, 2014*

In term of the maintenance of road drainage facility in Lokoja, Table 4 indicates that 86% of the urban residents who has road drainage facility in their neighborhoods claims that they are responsible for the maintenance of the drainage in the city, while 10.4 % of the drainage facility is maintained by the State government. Only 1.8% each of the drainage facility in Lokoja metropolis is maintained both by Local and the Federal governments respectively.

**Table 4:**
*Agencies responsible for Maintenance of Road Drainage Facility in Lokoja.*

| • Agencies | • Percentage |
|---|---|
| • Individual member of the household | • 86 % |
| • Truck C-Local Government | • 1.8 |
| • Truck B-State Government | • 10.4 |
| • Truck A-Federal Government | • 1.8 |
| • Total | • 100.0 |

*Source: Field Survey, 2014.*

Respondents were asked to indicate their frequency of cleaning and maintaining the road drainage facilities connecting their neighborhoods, Table 5 indicates that 34% of the respondent indicates that they clean their drainage on weekly basis. Another 23.4% of the respondents interacted with claims that they do the cleaning of the drainage on monthly basis; while 6.4% indicates that they do cleaning of the road drainage on every quarter of the year. Approximately, 36.2% of the habitants of the city interviewed claimed that they have never participated in cleaning of the road drainage connecting their neighborhoods. This could largely be responsible for the reason while, virtually every part of the drainage in the city is partially or totally blocked which frequently results in flooding shortly after each torrential down pour in the city.

**Table 5 :**
*Period of Cleaning of Road Drainage in Lokoja*

| • Time of Cleaning of Road Drainage | • Percentage |
|---|---|
| • Weekly | • 34.0 % |
| • Monthly | • 23.4 % |
| • Quarterly | • 6.4 % |
| • Never | • 36.2 % |
| • Total | • 100.0 % |

*Source: Author's Field Survey, 2014*

### 4.4 Pedestrian walkway

Absence of pedestrian walkway is a major characteristic of urban roads in Nigeria. The situation in Lokoja metropolis is not different from other cities in Nigeria. An assessment of pedestrian walkways in Lokoja reveals that one Truck A - Federal road- Muritala Muhammed way has pedestrian walkway on some parts of their length ,while the remaining two major federal roads in the city have no pedestrian walkway and drainage facility. Few of the Truck B- State Roads which have pedestrian walkways are less than four metres wide, two pedestrians cannot move freely without body contact. In some places, the pedestrian walkways have been encroached by residential buildings, while it serves as walk-shops for artisans most especially road side vulgarizers. Pedestrian walkway is completely absent at the Truck C route. Some of these roads are untarred and rough, they are mostly found at the outskirts of the city.

In terms of the accessibility of urban residents to pedestrian walkway in Lokoja metropolis, Table 6 reveals that only 29.5 % of the households claim that they have access to pedestrian walkway connecting the road network in their neighborhoods or their work places. As high as 70.5% of urban residents indicate that the pedestrian walkway is not install in the road network connecting their residential neighbourhoods. Some of the inhabitants interviewed have the opinions that if the pedestrian walkway is provided in the road networks connecting their residential area, they would prefer their wards to commute on foot to schools rather than allow them rely on motorcycles for their journey to schools. This observation is similar to the study of Boarnet (et al, 2005) carried out in California SRTS programme which shows that providing sidewalks is one of the most effective engineering measures in encouraging children walk to school.

**Table 6:**
*Household Accessibility to Pedestrian Walkways in Lokoja*

| • Accessibility to Pedestrian Walkway | • Percentage |
|---|---|
| • Yes | • 29.5% |
| • No | • 70.5% |
| • Total | • 100.0 % |

*Source: Author's Field Survey, 2014.*

With respect to the utilization of available pedestrian walkway in the city, Table 7 shows that 58.2% of inhabitants of the city indicate that they frequently make use of pedestrian walkway installed in Lokoja metropolis, while 16.4% of the residents in the city indicates that they make use of pedestrian walkway occasionally when demand arises. Almost, 25.4% of the household interviewed claims that they do not make use of pedestrian walkway provided in the city. The last category of this respondent competes with motorists and motorcyclists on intra -city movement and consequently exposes them to traffic collusion.

**Table 7:**
*Utilization of Pedestrians Walkway installed in the study area.*

| Utilization | Percentage |
|---|---|
| Occasionally | 16.4% |
| Frequently | 58.2% |
| Not at all | 25.4% |
| Total | 100.0% |

Source: Author's Field Survey, 2014

Many of the residents who claim that they do not make use of pedestrian walkway in Lokoja express their dissatisfaction about the condition of pedestrian walkway installed in the city. Table 8 indicates that approximately 40.9% of the respondents claims that the pedestrian walkway in the city have been occupied by roadside traders or artisans, while 22.8% indicates that waste generated in the city have been deposited on pedestrian walkway. Another 36.3% of the respondents claim that pedestrian walkway has been encroached by the fence of buildings and other structural development in the city.

**Table 8:**
*Reasons for non- utilization of pedestrian walkways in Lokoja*

| | |
|---|---|
| Occupancy by roadside traders/ artisans | 40.9% |
| Deposition of waste | 22.8% |
| Encroachment by building fences | 36.3% |
| Total | 100.0% |

*Source: Author's Field Survey, 2014*

**4.5 Street light**

In some of the urban centres in developed countries of the world, street lights are provided along the transport corridor to illuminate areas of pedestrian activity and reduce the incidence of night traffic crashes by pedestrians and motorists and increase the vision of commuters (Pegrum, 1972; Freedman et al, 1975). However, the converse is the case in many of the medium sized cities in Nigeria where Lokoja metropolis belongs to. For instance, Table 9 reveals that 88.7% of the households interviewed claims that the street lights were not installed on the road networks connecting their neighborhoods , majority of these people in this category live in karaworo area, Adankolo street, Sango Daaji Area, Phase Two and many other parts of low density area of Lokoja metropolis. Approximately, 11.3% of the respondent indicates that street lights are provided along the roads connecting their residential unit. People in this category reside at Taiwo Road- Post Office, Stadium Road, Ado Ibrahim Road, Aliu Attah Road and Jane Ekundayo Road to mention a few. To worsen the situation, most of the street lights are not Solar powered and electric power supply is very irregular. This has made it difficult for pedestrians, motorists as well as other road users to make any meaningful night trips in Lokoja metropolis.
Apart from inadequate provision of pedestrian walkways and street light in Lokoja metropolis, there is only one fly -over and non- pedestrian crossing in the whole city. This shows that the road network in the city is poorly designed and little effort is made on the installation of road transport infrastructure that can enhance smooth trips during intra-city movement. The provision of street lights, pedestrian walkways and many other transport facilities would enhance the free flow of vehicular and non-vehicular movement in Lokoja and similar other cities in Nigeria at large.

**Table 9:**
*Installation of Street Light*

| Possession of Street Light | Percentage |
|---|---|
| Yes | 88.7% |
| No | 11.3% |
| Total | 100.0% |

*Source: Author's Field Survey, 2014*

## 5. Conclusions and Recommendations

Poor design and maintenance of road transport infrastructure are major characteristics of Nigerian roads. These impede the free movement of pedestrians and vehicular traffic on intra-city roads. This circumstance is becoming pronounced in Lokoja metropolis as a result of increase in its size, structure and pattern of development since its creation as the capital of Kogi state in 1991. It is against this background that this study was designed to assess the provision of road transport infrastructure and its maintenance status in Lokoja metropolis, Nigeria. Both primary and secondary sources of data were used for the study. Three major roads each were randomly selected from each of the three tiers of governments in Nigeria to assess the level of maintenance of the urban transport routes in the city. The findings revels that many of the urban routes in the city do not either pedestrian walkways or drainage facilities. Inadequate installation of pedestrian walkways has forced the pedestrians, motorcyclists, motorists and other road users to compete on urban routes for their day to day transactions. Those roads that have drainage facilities have been filled up by sand from the erosion activity. Many of the urban routes have no traffic lights. Also, the erratic power supply has made it difficult for urban commuters to travel during the nighttime. The study therefore concludes that further research needs to be conducted on the designing of the urban road networks in Nigeria so as to achieve sustainable transport development. Also, the study recommends that the inhabitants of Lokoja metropolis should be enlightened that maintenance of road transport infrastructure is a collective responsible of all, both by the three tiers of governments and inhabitants of the city so as to achieve a sustainable city development.

**References**


Adetunji, M.A (2013): Assessment of the Quality of Urban Transport Services in Nigeria. Academic Journal of Interdisciplinary Studies,(AJIS) Sapienza University of Rome, Italy. Volume 2, No. 1, 2013 page 49 -58

Arizona-Ogwu , I . Chinedu (2008): Road Maintenance . Government Task or People Tax Also Available on: http://www.nigeriansinamerica.com/articles/2339/1/Road-Maintenance-Government-Task-Or-Citizens-Tax/Page1.html

Boarnet, M., Day, K., Anderson, C., McMillan, T., Alfonzo, M. (2005): California's Safe Routes to School Program: Impacts on Walking, Bicycling, and Pedestrian Safety. Journal of the American Planning Association, 71(3), 301-317.

Federal Highway Administration (FHWA). (2009). The manual on uniform traffic control devices (MUTCD); Part 7 traffic control for school areas. Washington DC. Retrieved from http://mutcd.fhwa.dot.gov/pdfs/2009r1r2/part7.pdf

Freedman, M., Janoff, M., Koth, B., & McCunney, W. (1975): Fixed Illumination for Pedestrian Protection. (FHWA-RD-76-8). Washington DC, Federal Highway Administration.

Gan, Albert; Joan Shen, Adriana Rodriquez (2005): Update of Florida Crash Reduction Factors and Countermeasures to Improve the Development of District Safety Improvement Projects. State of Florida DOT. BD015-04.

McMahon, Patrick J.; Charles V. Zegeer, Chandler Duncan, Richard L. Knoblauch, J. Richard Stewart, Asad J. Khattak (2002). An Analysis of Factors Contributing to "Walking along Roadway" Crashes, Research Study and Guidelines for Sidewalks and Walkways. Federal Highway Administration. FHWA-RD-01-101. Cite uses deprecated parameters

Olawepo, R. A. (2009): Evaluating Housing Problems through Participatory Rural Appraisal in Lokoja, Nigeria. African Research Review: An International Multi-Disciplinary Journal Vol.3 (1). Pp 77-96.

Ogunsanya, A. A: Vandu-Chikolo , I and Sumaila, A.G. 2004. Perspectives on Urban Transportation in Nigeria. Nigeria Institute of Transport Technology (NITT), Zaria. M O D Press, Kaduna pp1-26

Oni , I (2004): Development of Urban Transportation in Nigeria in perspectives on Urban Transportation in Nigeria. Nigerian Institute of Transport Technology NITT, Zaria. Editedby I Vandu-Chikolo, A Ade Ogunsanya and A.G Sumaila. Pp 55.

Pegrum, B. (1972): The application of certain traffic management techniques and their effect on road safety. Washington DC, National Road Safety Symposium.

http://guide.saferoutesinfo.org/engineering/sidewalks.cfm

http://en.wikipedia.org/wiki/Pedestrian_walkway.acessed on 24--05-201

# ROAD SAFETY – PERFOMANCE AND PERCEPTION MONTENEGRO CASE STUDY

**Vladimir Pajkovic[1], Mirjana Grdinic [2]**
*[1,2]University of Montenegro, Faculty of Mechanical Engineering, Podgorica, Montenegro*

**Abstract:** Over the last decade nearly all EU countries record a decrease in number of traffic casualties year after year. In the same period Montenegro's figures generally point in the opposite direction. The fatality rate on Montenegro's roads, measured as deaths per capita, is 50% higher than that of EU average. The most of road deaths are among drivers. Efforts have been made in recent years to address this issue. Different users have different perception about road safety problems. Drivers, both frequent and occasional, and non-drivers don't have same perspective and have different general attitudes towards road safety. Case study launched in Montenegro confirms this claim.

**Keywords**: road safety, perception of road safety problems, drivers, non-drivers

## 1. Introduction

More than a million people die each year on the world's roads. According to World Health Organization (WHO, 2013), approximately 1.24 million people die every year in car accidents and more than 50 million are injured. It has been estimated that if everybody fastened their seatbelt, respected speed limits and did not drive under the influence of alcohol, more than 12,000 lives could be saved each year on European roads (EC, 2014). Also, road traffic injuries take an enormous toll on individuals and communities as well as on national economies. They result in considerable financial costs, particularly to developing economies. It is estimated to cost countries up to 4% of their gross national product. In high-income countries, number of road deaths is decreasing in last few years, but low-income countries, such as Montenegro, saw increases over the same period.

**Table 1**
*Road network density*

| | Road density | |
|---|---|---|
| | km of roads per 1000 km$^2$ | km of roads per 1000 inhabitants |
| **Montenegro** | 500 | 11.1 |
| Southeast Europe, average | 555 | 5.9 |
| Albania | 657 | 3.5 |
| Bosnia and Herzegovina | 427 | 5.6 |
| FYR Macedonia | 513 | 6.4 |
| Serbia | 500 | 5.2 |
| New EU member states | 1427 | 19.9 |
| Czech Republic | 1646 | 12.5 |
| Estonia | 1320 | 41.2 |
| Hungary | 1733 | 15.7 |
| Slovenia | 1007 | 10.2 |
| Croatia | 506 | 6.4 |

Montenegro has a road network totaling approximately 7,000 km – with approx. 900 km of main and primary roads, 950 km of regional and secondary roads, and around 5,000 km of local roads. This is equivalent to a road density of 500 km per 1,000 km$^2$. This figure is broadly consistent with the density of some new EU member states (see Table 1).

**Table 2**
*Number of vehicles in Montenegro*

| | 2010 | 2011 | 2012 | 2013 |
|---|---|---|---|---|
| Registered cars | 164653 | 171973 | 173865 | 178662 |
| per 1000 inhabitants | 266 | 277 | 279 | 286 |

Regarding road infrastructure quality, Montenegro ranked only 107$^{th}$, of 131 countries surveyed, for the quality of its road infrastructure. 90% of Montenegro's road network is high risk, including those segments with a high traffic volume (COWIS, 2008). Also, 47% of the entire road network is in poor or very poor condition, reflecting inadequate maintenance. To improve the road safety situation on the main road network, costs an estimated €105–138 million.

[2]Corresponding author: grdinicm@yahoo.com

Other factors, such as the distribution and density of population and the country's geography, play a considerable role in determining country's road network. These effects become evident when making the comparison on a different measure of road density – viz. road kilometers per 1,000 people. With this measure, with more than 11 km of road per 1,000 inhabitants, Montenegro is ahead of most of its regional comparators and comparable to those of the new EU countries, (EC, 2012b). During the last few years, the number of registered vehicles is around 200,000 counting all vehicles category and about 85% are passenger cars. In 2011 this rate was approx. 277 and in 2013 approx. 286 passenger cars per 1,000 populations, see Table 2 and Fig. 1 (Pajkovic and Grdinic, 2014a).

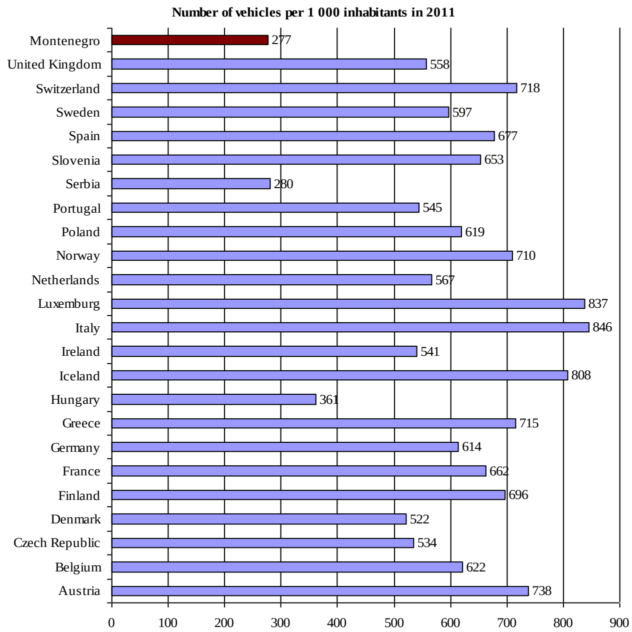



**Fig. 1**
*Numbers of vehicles per 1000 inhabitants in 2011*

## 2. Road Safety Performance

The safety performance of the road traffic system is the result of the (right) combination of the functionality, homogeneity, and predictability of the road network, the road environment, and the traffic involved. Assessment of road safety performance is based on statistical data. Statistical indicators are proposed as supportive tools in comparing the level of road safety between various countries and are useful in terms of monitoring changes over time.

In all European countries, the official accident statistics that usually rely on police reports, play an important role in road safety policies. Governments use the number of fatalities or injured casualties as a method to analyze the current road safety situation and also to define prospective goals for road safety in their countries. In order to make progress in road safety, comprehensive data collection and in-depth data analysis are essential in terms of designing effective safety strategies, setting challenging targets, determining intervention priorities and monitoring program effectiveness. Number of road accidents, deaths and injuries in EU countries is decreasing, despite larger number of vehicles. Number of accidents was reduced by 15% in period 1990-2010, fatalities by 25%, while number of injuries remained the same, Fig. 2, (EC, 2012b). Published in 2010, the European Commission’s new road safety policy orientation aims to halve the number of road deaths in the European Union by 2020 and reduce injuries. Countries of West Balkans as developing countries are trying to comply with the rules and regulations of the EU. Despite that, all indicators of road safety in Montenegro have been increased over the previous ten years. Provisional accident data collected for this paper reveals that road safety has become a serious and increasing economic and social problem in Montenegro. Compared to 2000, in 2010 the number of road accidents went up 63.2%, fatalities by 17.3% and number of injured by 8.5% (Fig. 3).

Over the last decade more than 1.000 people died on Montenegrin roads and more than 23,300 were injured. Mortality rate was the highest in period 2007-2009, over 10.000 accidents in 2008 and 122 killed in 2007 (Pajkovic and Grdinic, 2014). Years 2007 and 2008 were particularly evil, which was the “price” of uncontrolled growth of motorization (in precedent years) and inadequate road infrastructure that couldn’t “bare” an increasing traffic volume. The highest density of fatalities is found on main north-south route Border with Serbia − Bijelo Polje − Podgorica − Cetinje − Budva and Podgorica city zone.

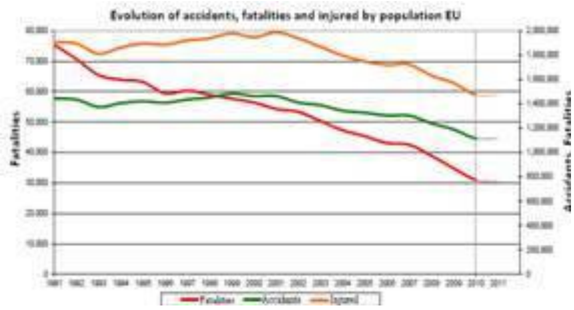



**Fig. 2**
*Accidents indicators in EU*

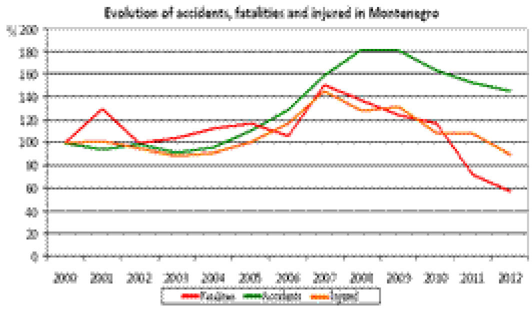



**Fig. 3**
*Accidents indicators in Montenegro (2000=100)*

During 2011 on Montenegro's roads 87 people lost its lives in traffic accidents (Fig. 4). It is not a large number comparing to other EU countries. However, comparison of road safety performance depends somewhat on which indicator is used as a measure of exposure to risk; population, number of registered vehicles or distance travelled by motorized vehicles.

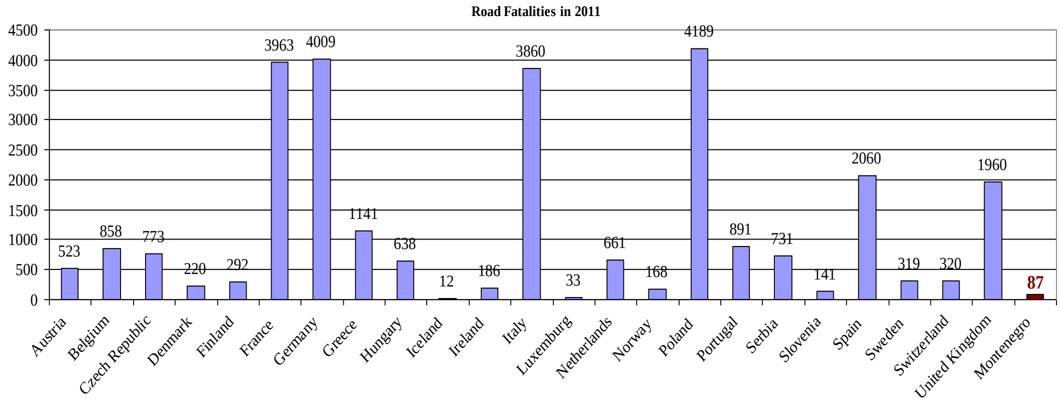



**Fig. 4**
*Road fatalities in 2011*

*Fatalities per 100.000 head of population* expresses the mortality rate, or an overall risk of being killed in traffic, for the average citizen. This is a particularly useful indicator to compare risk in countries with the similar level of motorization. According to EU statistics (ITF/OECD, 2013), the rate in terms of fatalities per 100.000 populations in Montenegro in 2011 was 14.3. Comparing to EU countries, the risk of dying as a result of a road traffic injury on Montenegro roads is significantly higher than anywhere else (Fig. 5).

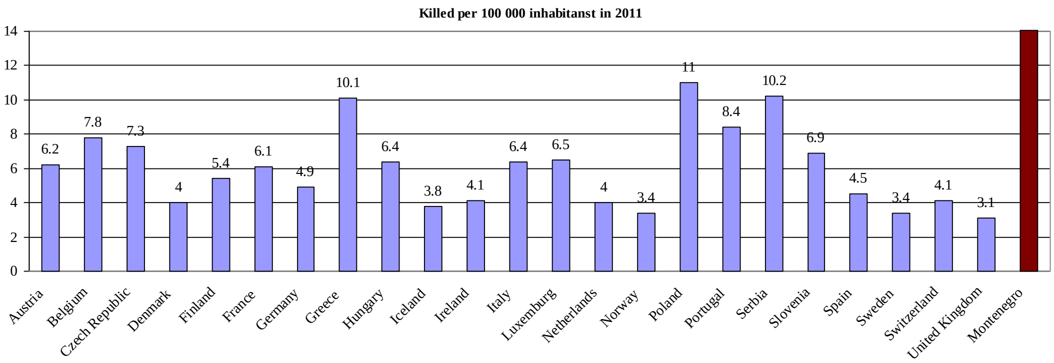



**Fig. 5**
*Fatalities per 100 000 inhabitants in 2011*

*Fatalities per 10.000 registered vehicles* is a more objective indicator of situation on the roads. However, it's useful only when comparing the safety performance between countries with similar traffic and car-use characteristics. Number of fatalities per 10.000 registered vehicles in Montenegro is also higher compared to other countries. However, there is a matter of vehicles from foreign countries in accidents on Montenegro's roads during the tourist season which are not included in the number of 10.000. Considering that fact, this rate (in terms of fatalities per 10.000 registered vehicles) is at same high as neighbor countries (Fig. 6).

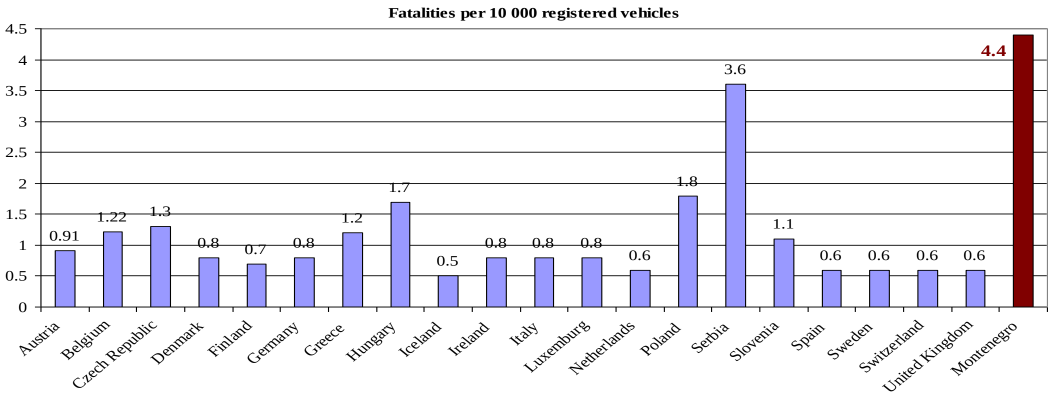



**Fig. 6**
*Fatalities per 10 000 registered vehicles*

### 3. Countermeasures

Many road crashes are preventable and history provides evidence that the right intervention can make a significant positive impact. In May 2011, the United Nations launched a Decade of Action for Road Safety. The goal of the Decade (2011–2020) is to stabilize and reduce the increasing trend in road traffic fatalities, saving an estimated 5 million lives over the period. This report builds on the 2009 report, and provides additional data in a number of important areas. It is supported by a Global Plan for Road Safety (EC, 2012b).

Speeding, drink driving and failure to wear a seat belt are the three main risk factors on the road. It is estimated that if average driving speeds dropped by only 1 km/h on all roads across the EU, more than 2.200 road deaths could be prevented each year. At least 3.500 deaths could have been prevented in 2009 if drivers concerned had not drunk before taking the wheel. Across the EU, an estimated 12.400 occupants of light vehicles survived serious crashes in 2009 because they wore a seat belt. Another 2.500 deaths could have been prevented if 99% of occupants had been wearing a seat belt. Experience from fast progressing countries shows that progress in fighting speeding and drink driving and increasing seat belt use can be fast and save thousands of lives, (ETSC, 2010). The European Union has passed stricter legislation making the use of seat belts mandatory throughout the EU. The European Commission also published two relevant Recommendations, the 2001 Recommendation on maximum permitted blood alcohol content of 0.2g/l for novice and professional drivers and 0.5g/l for other drivers, and the 2004 Recommendation on enforcement in the field of road safety.

As a response at Decade of Action for Road Safety, several countries released or updated their national road safety strategies in 2011. In Austria, the National Council amended road traffic regulations which, from 31 May 2011 (when Decade was established), mandate that children under the age of 12 years wear helmets while cycling or while being transported on cycles. In China, a new legislation entered into force on 1 May 2011. From that date, drinking and driving has been criminalized and penalties for transgressions have increased exponentially. In France, the Interministerial Committee for Road Safety took the decision with immediate effect to strengthen enforcement of legislation on speeding by, for example, criminalizing speeding in excess of 50 kilometers per hour over the speed limit, and drinking and driving by, for example, increasing the number of points lost on one's driving permit for drivers with a blood alcohol content level above 0.08 g/dl, among others. France is the only country where speed reductions have been achieved on all types of roads between 2001 and 2009. The average speed of light vehicles on all road types taken together has decreased steadily by 10 km/h (or 12%) over the last eight years. Great Britain and Austria also recorded reductions in mean speeds on both urban roads and motorways. Drivers have slowed down markedly in cities in the Czech Republic and Ireland.

In 2010 National Coordination Board was established in Montenegro to monitor road safety parameters. In addition, government adopted Strategy for improvement of road traffic safety. With this strategy, governments make commitments and take responsibility for critical issues in the system. However, the government can only do so much and it need the support of organizations, industry, businesses, community groups and individuals. The main goal of this strategy is to reduce number of fatalities by 30% and injuries by 20% before 2014 (compared to 2007). The long-term goal is to reduce these figures by 50% and 30% respectively by 2019, compared to the same year. However, the level of trauma reduction that can actually be achieved by 2019 will depend on the costs and policy changes that the community is prepared to accept in return for a safer road transport system. Also, Montenegro has started a Black spot management with risk mapping in 2007 under the responsibility of Road Directorate and through the regular maintenance and there are proposals under discussion to add a special department for road safety issues (road safety unit) in the level of the Ministry of Transport.

Sadly, the growing interest in reducing road traffic risk is not accompanied with active road safety policies. Proposed measures are inadequate and insufficient – partly because addressing the situation properly requires considerable investments.

One of the latest road safety campaigns in Montenegro is related to drinking and driving. In May 2014, government realized campaingn “When I drink, I don’t drive” (*Kad pijem, ne vozim!*). Brochures that were distributed within this campaign describe effects of alcohol on drivers depending on amount of alcohol consumed, showing penalties for driving under the influence of alcohol as well. Also, a few campaigns were realized by several non-government organizations.


**Fig. 7**
*Road safety campaign in Montenegro*

**4. Perception**

Although the number of road fatalities and injuries are the ultimate indicators of road safety, the other factors such as road users’ attitudes and behavior can help quantify the road safety status of a country. Opinions, beliefs and behavioral intentions of different groups of users regarding matters of road safety are a source of relevant information, since they reveal characteristics of their attitudes in relation to mobility and safety, their relations with other users, their experiences and their concerns. Any differences which may be found form the basis on which to implement specific measures to try to reduce road hazards to a minimum and improve mobility conditions.

Many surveys have been conducted with broad purpose to monitor public opinion and behavior on a range of important road safety issues. Specific objectives are to observe perceptions about the seriousness of road safety problems, to identify community attitudes to major road safety issues, to determine road safety problems that should receive more attention from national governments, to identify any significant changes in attitudes and perceptions compared to previous years or differences in such attitudes and perceptions between countries, also to determine measures that national governments should focus on to improve road safety.

“*Road safety – Analytical report*” is survey conducted in June 2010 by The Gallup Organization, Hungary, upon the request of Directorate-General Mobility and Transport. The objective of the survey was to derive greater insights into which road safety problems are perceived as the most serious by EU citizens, to improve the understanding of the areas of road safety where EU citizens would like national governments to do more and those where they feel that governments are already doing enough and also to identify those areas of road safety policy that EU citizens would like their national governments to prioritize. Similar survey was conducted during June and July by University of Montenegro, Faculty of Mechanical Engineering, Dept. of Road Traffic. The target population of those surveys was the general adult population. Sampling variables were: sex (male, female), age and driving experience. The surveys were carried out by face-to-face interviews.

When asked if they drove a car, about 7 in 10 EU citizens (69%) responded positively: about half of respondents said that they drove a car *most days* of the week (49%), 16% drove a car *1-3 times per week*, 3% drove *1-3 times per month* and 1% drove *less than once per month*. About 3 in 10 EU citizens (31%) said that they did not drive a car. In Montenegro, about 8 in 10 citizens (76%) responded positively: about half of respondents said that they drove a car *most days* of the week (51%), 11% drove a car *1-3 times per week*, 7% drove *1-3 times per month* and 15% drove *less than once per month*. About 24% of citizens said that they did not drive a car (Fig. 8).

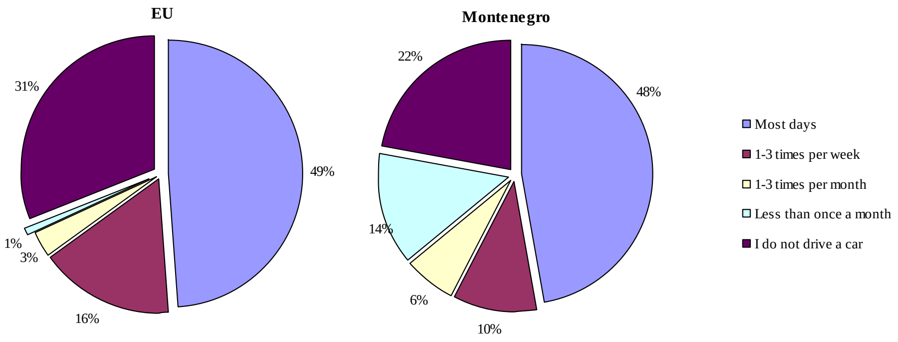



**Fig. 8**
*Frequency that respondents drive a car*

In order to evaluate general attitudes to road safety in Montenegro all interviewees were asked to give opinions on if roads in Montenegro are safe, fairly safe, fairly unsafe or unsafe. 13% of Montenegro's described road safety as *safe*. A further 40% described it as *fairly safe*, 19% described it as *fairly unsafe* and 28% as *unsafe*. Overall, 53% described the roads as *safe* or *fairly safe*, slightly higher than from those who describe as *fairly unsafe* or *unsafe* 47% (Fig. 9).

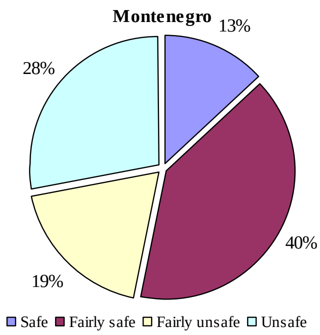



**Fig. 9**
*General attitudes about road safety in Montenegro*

Citizens were asked about five potential road safety problems: for each one, they were requested to say if it was perceived as a problem in their country and if so, whether it constituted a *major* or *minor* safety issue. In Montenegro survey couples of questions were added. Both EU and Montenegro citizens clearly perceived the issue of people driving under the influence of alcohol to be the *major* safety problem (94% in EU and 80% in Montenegro), with just 1 in 20 (both EU and Montenegro) respondents seeing this as a *minor* issue. Virtually none of the respondents said it was *not a safety problem* in their country. Drivers exceeding speed limits were considered to be a *major* safety problem by 78% of respondents in EU and 72% in Montenegro. 81% of Montenegro's citizens think that driving without respecting legal regulations is a *major* safety problem. In EU talking on a mobile phone without a hands-free kit (76%) and drivers/passengers not wearing seatbelts (74%) are listed as *major* safety problems while in Montenegro only 16% thinks that not wearing seatbelts is a *major* safety problem. Large number of Montenegro's citizen (35%) thought that people driving while talking on a mobile phone without a hands-free kit was *not a problem at all*.

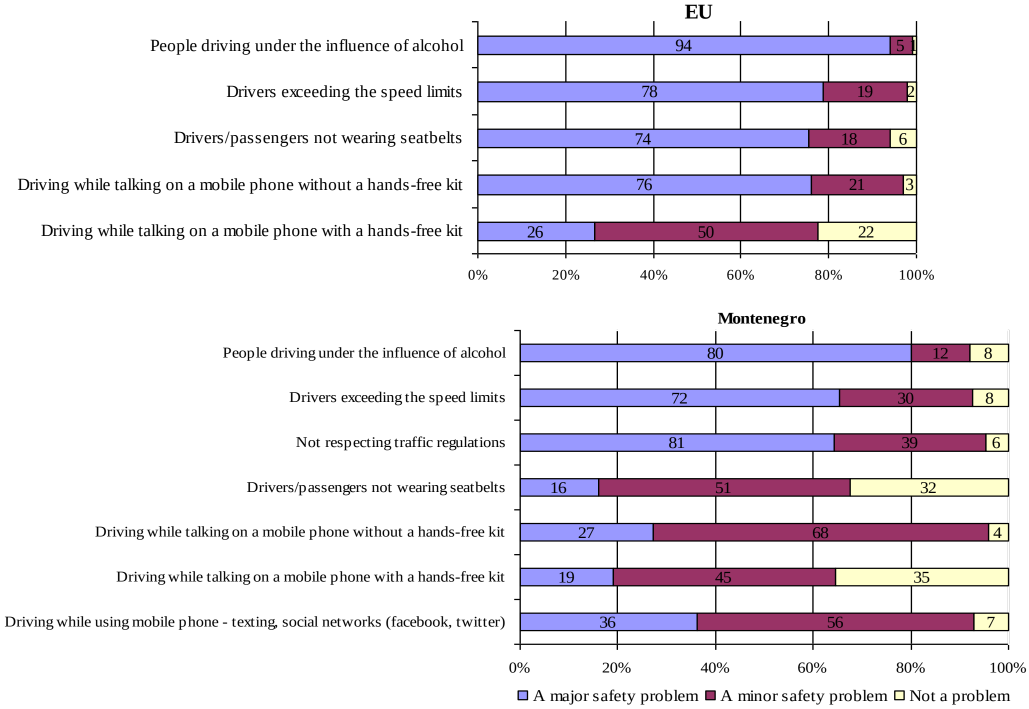



**Fig. 10**
*Perceptions about the seriousness of road safety problems in EU and Montenegro*

For the entire listed problem areas, EU citizen were asked whether it should receive more attention from their national government or whether enough was already being done. In Montenegro's questionnaire was offered to choose one of the following answers which describe what government should do: increase penalties, better education or done enough.

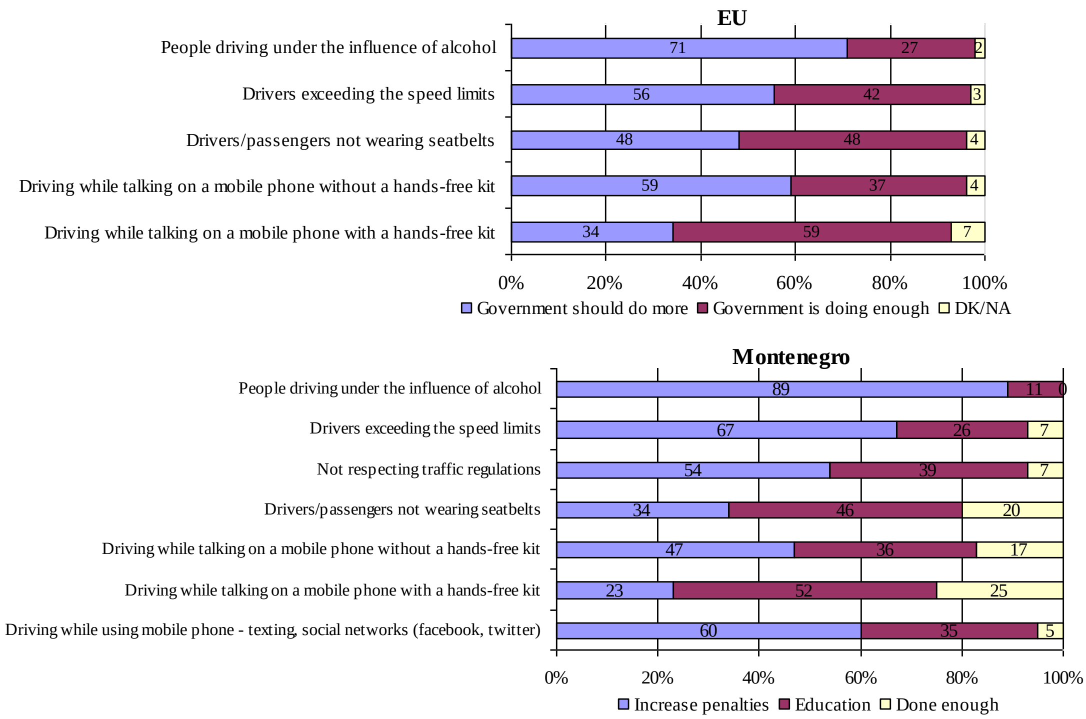



**Fig. 11**
*What to do?*

EU citizens were clearly most likely to say that their national government should do more to reduce the problem of people driving under the influence of alcohol: 7 in 10 respondents felt this way (71%) and about a quarter (27%) said that enough was being done to fight the problem of drunk-driving. In Montenegro, widely accepted opinion is that government should increase penalties for such behavior (89%). Montenegro is a low-income country and people think that penalties are the most effective way for increasing the level of road safety awareness. All respondents were asked which measures national governments should focus on to improve road safety. A slim majority of EU and Montenegro citizens (31% and 51%, respectively) said that road infrastructure safety should be improved as a first priority. Improving the enforcement of traffic laws was the second most frequently selected measure in EU (42%) and increasing traffic monitoring and improvement of pre-licensing in Montenegro (11%), followed by improving the enforcement of traffic laws (10%). 8% of interviewed Montenegrin citizens think that strict penalties for not respecting traffic regulations are a measure that the government should concentrate on in order to enhance road safety. About a quarter of EU citizens (26%) and 6% of Montenegrin citizens held the view that their government ought to assign priority, for all drivers, to the introduction of periodic driver re-training schemes in order to improve road safety. 13% respondents in EU and 3% in Montenegro were of the opinion that their national government should initiate more road safety awareness campaigns as a priority action in order to improve road safety.

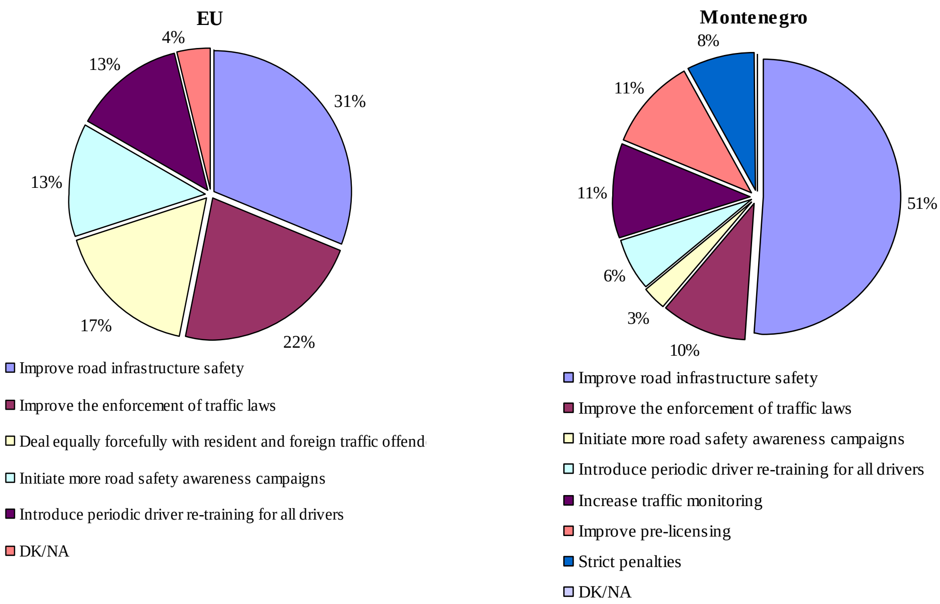



**Fig. 12**
*Measures that national governments should focus on to improve road safety*

## 3. Conclusion

According to the data provided by World Health Organization, traffic accidents claim approximately 1.2 million victims annually, with 50 million being hurt. Whether this growing trend of traffic accidents continues in the coming period, it is estimated that they would take a third position on a ‘list of ten leading causes of death’. Fortunately, the number of people killed in car crashes decreased significantly over ten years in all European countries. Meanwhile, Montenegro’s figures generally point in the opposite direction. Some improvements were made (government adopted Strategy for improvement of road traffic safety), but the overall risk indicators in Montenegro remain above the EU average, not showing a decrease in their trend and a lot of work is yet to be done. Many surveys have been conducted with broad purpose to monitor public opinion and behavior on a range of important road safety issues because attitudes and behaviors can help to quantify the road safety status of a country. The objective of surveys conducted in EU and Montenegro was to derive greater insights into which road safety problems are perceived as the most serious by citizens, to improve the understanding of the areas of road safety where citizens would like governments to do more and those where they feel that governments are already doing enough and also to identify those areas of road safety policy that EU citizens would like their national governments to prioritize.

**References**

COWIS A/S, 2008. *World Bank – Road Safety Survey Montenegro, Final Report*

EC-European Commission, 2012. *EU Transport in figures – Statistical pocketbook*

EC-European Commission, 2012. *Road safety evolution in EU*

EC-European Commission, 2014. Mobility and transport-Road safety-Behaviour. Available from: <http://ec.europa.eu/transport/road_safety/topics/behaviour/index_en.htm>

ETSC-European Transport Safety Council, 2010. *Tackling the three main killers on the roads – PIN Flash n.16*

ITF/OECD, 2013. *IRTAD – Road Safety Annual Report 2012*.

MONSTAT, Podgorica, 2011. *Traffic in Montenegro 2005–2010.*

Pajkovic, V.; Grdinic, M. 2014. Causes and effects of road traffic accidents in Montenegro. In *Proceedings of the 8$^{th}$ Road accidents Conference,* 650-660 (in Serbian)

Pajkovic, V.; Grdinic, M. 2014. Road traffic safety perfomance in Montenegro. *International virtual journal Machines, Technologies, Materials (MTM),* 2/2014: 26-28.

WHO-World Health Organization, 2013. *Global status report on road safety*

# THE RISKS OF TRANSPORT AND POSSIBILITY OF THEIR ASSESSMENT

**Barbora Schüllerová[1], Vladimír Adamec[2], Karol Balog[3]**
[1,2] *Institute of Forensic Engineering, Brno University of Technology*
[3] *Faculty of Materials Science and Technology in Trnava, Institute of Safety, Environment and Quality, Slovak university of Technology in Bratislava*

**Abstract:** Transportation of the hazardous materials is currently an essential part of the functionality and operability of the national and international economy. Transported hazardous materials include for example petroleum products such as diesel or gasoline, technical gases and other chemicals. The post focuses primarily on the form of the road transport (transportation according to the international agreement of the international transport ADR) and the risks associated with it (accidents with spills of hazardous materials).. Besides risks arising from the carriage, there may appear threats during the transportation of the hazardous materials that can have serious consequences for human life and health and have a potential to cause serious damage to the environment, property etc.. It is necessary to take into account the loading work (filling transport containers, tanks, pumping, etc.) with the risk of leaking hazardous materials. The assessment of the transport risks is a needed pre-planned part of the safe transportation. Although hidden threats may occur during every transport, these risks have not been usually analyzed. The post focuses on the methods of the assessment of the transport risks including hidden risks that may cause serious accidents during the transport with all its negative impacts on human health and the environment.

**Keywords**: transport, assessment, risk, hazardous materials, hidden processes.

## 1. Introduction

The function of industry and its production has in all developed countries great importance for their economies and its correct functioning must be secured and supported by many other activities, such as transport of hazardous materials. With increase of the industry, grow the need of increase the frequency and amount of the transported volumes in various ways, such as road transport, rail, ship and air transport or the pipeline transport. In comparison, for example, the road transport is considered riskier than the transport on rail. Although according to the statistics available the probability of accident of vehicles carrying hazardous substance is several times lower than the normal vehicular traffic, its significance is much higher. Especially with regard to the damage to human health, the environment and property, particularly in cases with the release of hazardous substances, its initiation and manifestations of the negative attributes. However, the leaks of hazardous substances can occur not only in the crash and in damage of the transport unit. There are the so-called hidden processes that may happen during the transport and may have a negative influence on the transport unit causing the leakage of hazardous substances. It is important to have these hidden processes in mind when planning the safe transport and risk assessment, which accompanies the transport of hazardous substances.

## 2. Current status of hazardous materials transport in the Czech Republic

Currently, the issue of transporting hazardous substances is embedded in the form of regulations, directives, laws and methods tending to have an international force. They are implemented into national legislation and carriers must strictly obey them. One of the fundamental European legislation is the European Agreement concerning the International Carriage of Dangerous Goods by Road (ADR), that was implemented into the Czech legislation and the legislation of other states, was agreed in Geneva on September 30$^{th}$, 1957 and entered in to force on January 29$^{th}$ 1968 (ADR, 2013). This agreement is regularly updated and currently is valid for 48 states. Similar legislation can be found elsewhere in the world, such as the United States of America (DIAMANT etc.).
Under this agreement, only an authorized person is enabled to participate to plan the transport of hazardous substances by road. This person has to be an eligible safety consultant with a valid test for a safety consultancy. The international ADR agreement also specifies further conditions. Transportation has to be carried out by specialized firms employing trained drivers that have their own special means of transport and packaging for the transport of hazardous substances. There are high demands both on drivers, who must be specially trained, so on means of transportation carrying hazardous substances.
Failure to comply with these requirements thus increases the risk of traffic accidents, spills, and deteriorates the consequences when accidents occur (environmental remediation, hazardous substances violent reaction hazards to human health and life, etc.).

[1] Corresponding author: barbora.schullerova@usi.vutbr.cz

During the risk assessment, is necessary for the expert, in addition to another, monitor the updates to existing legislation and changes such as the transition from the R and S sentences on H and P phrases, when these change of classification can lead to erroneous results in the calculations.

### 2.1. Statistical data of transport of hazardous materials

Statistics concerning traffic accidents that involve leaking of hazardous substances during the carriage, pursuant to ADR in the European Union (EU) and Czech Republic (CZ) are reported until 2012 so far. Data from 2012 also take into account the data of first half of 2013. In the last years (2009 - 2012), there have been 394 traffic accidents in which participated vehicles according to ADR mode. In the years 2009 - 2011 totally 151 accidents were directly caused by their drivers, (from a total of nine deaths was caused by the ADR driver only one death (Police, 2014; Dopravni noviny, 2012).
Totally 22 of 394 accidents occurred to leakage of hazardous substances (Police, 2014). Statistical data for 2003 – 2011 (Dopravni noviny, 2012) indicate that there have been total of 1,425 accidents with the participation of ADR vehicles, with consequence killing of 39 people. Police for this period also recorded 10 people seriously injured and 59 slightly wounded (Dopravni noviny, 2012). The thirty of them were caused by drivers carrying hazardous substances (Dopravni noviny, 2012). Czech Police also maintains statistical records of accidents including transportation of hazardous substances.
Although the number of accidents involving leaking of hazardous substances during the transport is lower than in previous years, their importance cannot be neglect due to the growth of transport volumes and new types of hazardous substances and mixtures that are provided by science in the development of industry and population.
Following chapters describe risk factors and risk areas that need to be included in the risk assessment in every analysis of risk transportation.

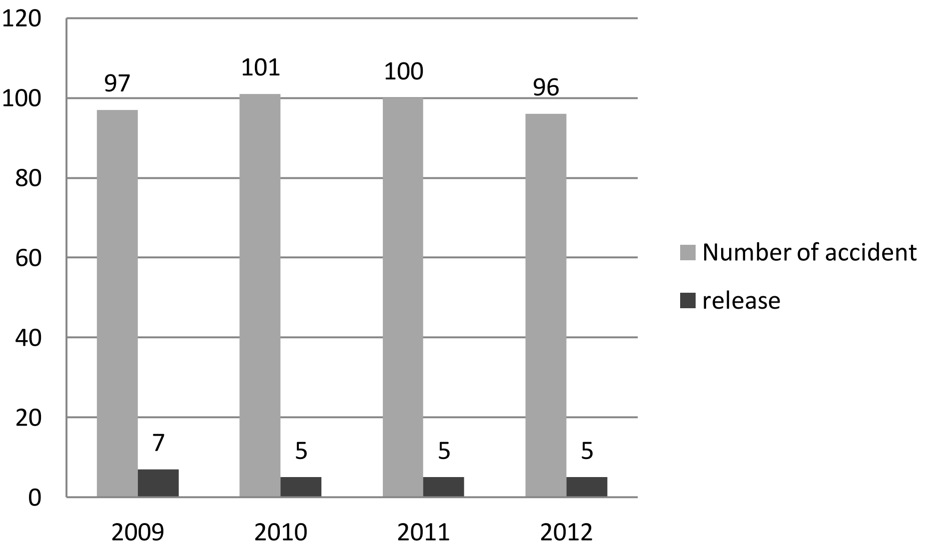



**Fig. 1.**
*Statistics of traffic accidents and spills of hazardous substances during transportation*
*Source: Police CZ, author*

### 3. Risks of transport of hazardous materials

The risks accompanying the transport of hazardous substances by road are part of the process of transportation. Their origin can be:
- Anthropogenic,
- Technical,
- Natural.

The transport of hazardous substances does not include only the moment when the transport unit with a hazardous substance is moving on the road. It is necessary to include into the transport process the situation when cargo carrier is unloading or loading the cargo or is waiting for example at the border, cleaning the transport unit etc. These activities are called loading works. Among the most often accidents is considered the leakage of hazardous substances. The leakage is mostly caused by rough handling, by damage of packaging or by intentional overload (overload axle vehicle) of higher volume of cargo that may have serious impact on transport itself, especially driving the vehicle (Brozova, 2008), a longer stopping distance, unwanted movement of cargo etc.

The other risks of leakage of hazardous substances during loading can include situations such as poor load placement, poor attachment and securing of loads, inadequate decontamination of vehicles, the choice of unsuitable packaging or vehicles for transportation, loading banned substances for joint loading, violation of safety rules such as no smoking, opening packages of hazardous substances etc. (Brozova, 2008).

According to the data from the police of the CR and professional firms as ÚAMK, accidents and leakage during the transport are most commonly caused by violations of the road rules such as insufficient distance between vehicles, traffic rules or by obstacles on the road, failure to comply with safety breaks, failure of a vehicle to adapt traffic situation (density of traffic, time of year) etc. (Vomacka, 2014).

There are also other risks like a low quality of driver´s training, failure or deliberate violation of the rule regulations and of OSH (driver overloading, fatigue, physical and mental fatigue, pain caused by sedentary, stress etc.). (European Agency OSHA, 2014). Example of risk analysis in the form of tree diagram for the conditions of release of hazardous substances during the transport (vehicle in motion) is shown in Fig. 1

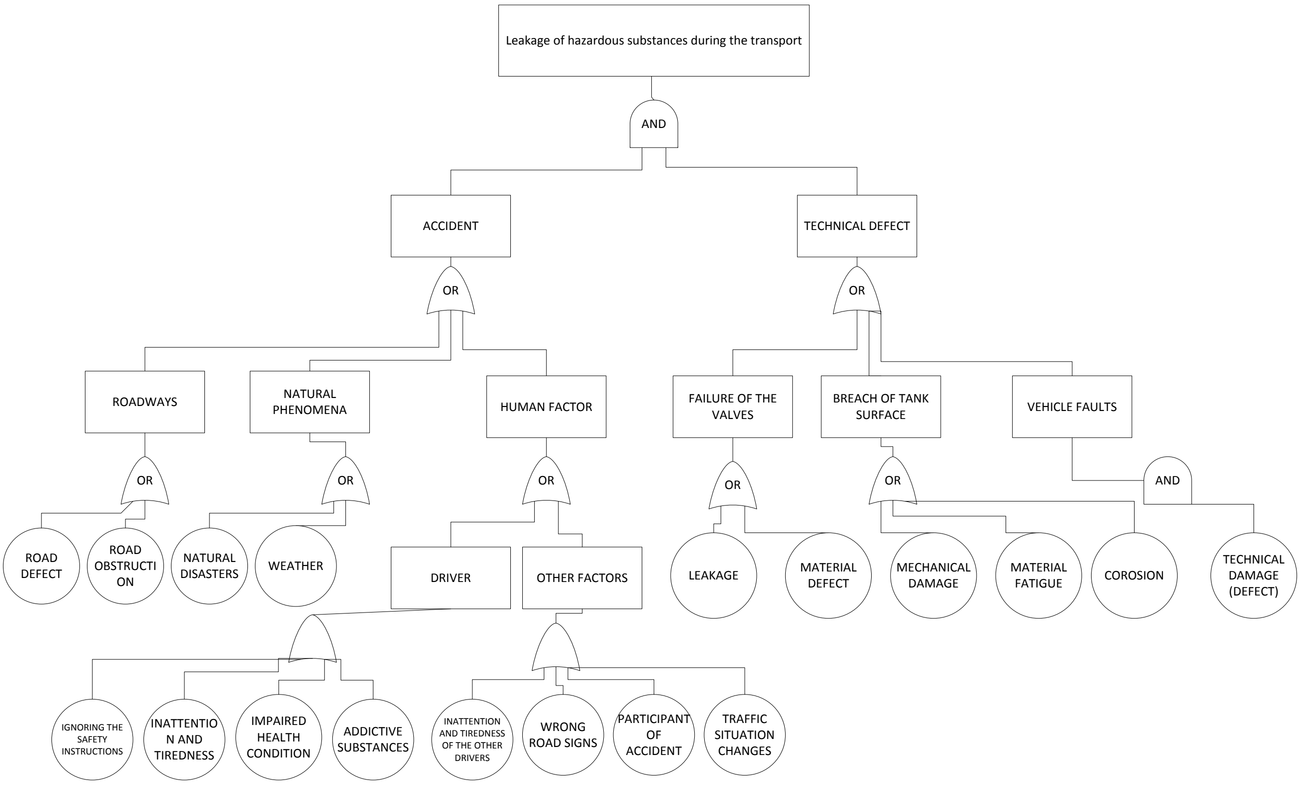



**Fig. 2.**
*FTA diagram of release of hazardous materials during the road transport*
*Source: Author*

FTA Diagram in Figure 1 analyzes causes of traffic accidents from the basic view of current traffic accidents. In more thorough analysis, it would be necessary to include additional factors that could be added into the branch "Accident - The Human Factor". Those factors involve potential risks and terrorist attack (transported unit has a little scope to safety measures), an intentional discharge of substances, in the technical failure then deliberate damage to vehicles or transport units. Although this is often a very unlikely event, it is necessary to calculate with their possible occurrence.

Data for the calculation of the probability of different scenarios mentioned in the analysis may be obtained from available statistics (most frequently police records or transport companies). Additional procedures arise from the calculations of probability and rules of probability of conjunction and disjunction (Bernatik, 2006).

## 4. Risk Assessment

Every step of the hazardous substances transportation may lead to leakage of these substances in a greater or lesser extent, such as spills from the fuel vehicle. In the case of a fuel spill, there is a space where the substance is distributed and that area is classified as hazardous. The area of danger zone depends on several factors such as the concentration, the quantity of hazardous substances, the ability of the dispersion in the environment and other physical properties of the substance or mixture. The spillages in this area can spread for example a dangerous form of cloud formation or puddles hazardous substances etc. Some other factors also influence the extent of the danger zone, especially the outside temperature, speed and wind direction, terrain topography, buildings, and the properties of the rock environment (Kocianova, 2012).

In the danger zone where hazardous substances have been released there may be the signs of the dangerous properties of the substances such as toxicity, flammability and explosiveness. The worst-case scenario connected with the release of hazardous substances is the combination of these properties, namely fire and subsequent explosion substance, with a subsequent release of the toxic vapors and gases, which under unfavorable climatic conditions can spread very quickly. The planning of transport routes has strict criteria, which must be taken into account by the carrier. It concerns especially the protection of landscape area and drinking water sources with their reservoirs, protection zones, size and population density in the place and close to transport, route changes (new roads, tunnels, underpasses, bridges etc.). Similarly, it must be taken into account predictable risks during transportation, which must not be neglected.
In the context of transport the risk maps are created, which include pre-selected risk routes. However, due to the increase of traffic and changes of routes the updates are often missing. Another factor that contributes to increasing the risk during transportation of hazardous substances is also the fact that current science has placed the new chemical substances and mixtures on the markets that are not used only in the industry. That also requires a special mode of transport and safety measures and the knowledge of their dangerous properties. Currently, it is a difficult task to find and use an appropriate and unified methodology for screening risk assessment in connection with the carriage of hazardous substances that can cause a serious damage and threaten population and the environment. It is important to maintain the uniformity of access following basic steps of risk assessment methodology that can be applied generally to the issue of leakage of the hazardous substances, including transportation. The application of methodological approaches provides the possibility to appropriately solve the technical problems (EU, 1996; Roosberg et al., 2002).

**Table 1**
*Methodology of Risk assessment*

| | |
|---|---|
| **Methodology of Risk Assessment** | The initiation and problem formulation<br>Data collection and evaluation<br>Identification of the hazards and dangerous scenarios<br>Formal measures of values to be protected<br>Assessment of the frequency and probability<br>Assessment of the consequence<br>Exposure assessment<br>Risk characterization |

*Source: European Commission*

### 4.1. Risk Assessment – Road maps of risk

Assessment and Risk Analysis, which is used for the transport of hazardous substances, involves a combination of several methodologies and procedures that are most commonly applied to the sections of the route through the hazardous substance will be transported with regard to the parts where there is a pre-determined level of risk. Different actors from various European countries deal with determination of the risk level and risk sections of roads. For example EuroRAP, non-profit organization of European car clubs, focuses on the systematic monitoring and assessment of the road traffic safety (EuroRAP, 2014).
The resulting risk maps are based on the official statistics and analysis of accident causes (EuroRAP, 2014). The most recent map of risks data for the Czech Republic created for years 2010 – 2011. The map shows how is on the stretch of risk when a road user can become a participant of an accident with serious or fatal consequences. Maps can be compared with other European countries that participate in society and especially because of the uniform methodology company that creates the maps (EuroRAP, 2014).
The maps of risk for the individual sections of roads are very effective basis for the risk assessment on and along of transport routes. In the Czech Republic, the mapping is at this moment performed on approximately 6,500 km network of highways and main roads (EuroRAP, 2014; EuroRAP, 2014b). Although most of the risk analysis methods are designed for the stationary sources of hazardous substances, these methods may also be used with a certain modifications for the sources of risks that are not stationary. These functions include the freight vehicles with the transport units containing a hazardous substance or already empty units, which were not subjected to cleaning. Part of the level of risk is then transported depends on the amount of hazardous substances.

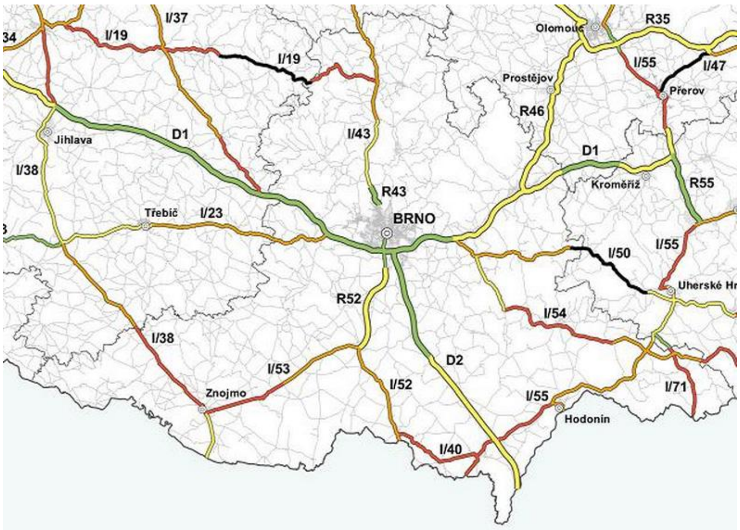



**Fig. 2.**
*Map of Road Risk in South Moravian Region*
*Source: EuroRAP*

### 4.2 Risk Assessment – Volume of Hazardous Substances

Another important factor to include to the risk assessment of transported hazardous substance is its amount. If we take the example of chlorine transport by road in the Czech Republic the ADR transport of this material in tanks can be negotiated, but in the Czech Republic, it can be transported by road only in cylinders and drums. Higher amounts of chlorine are transported only by trains, which is statistically and by the perspective of risk assessment safer. The chlorine dividing into smaller transport units reduces the risk of greater spillage of the substances into the environment and endangering not only the habitants, but also the environment. Another factor in the risk assessment is divided according by the amount of transported material (lower and upper limits of the volume of hazardous substances), which also may include the case of fuel tank with the fuel, that for trucks is usually approximately 200 l (lower limit). In the case of its leak, there is a significant risk of a dangerous area formation. The upper limit is the maximum possible volume of transported substances. Limiting the risk assessment for the transport of hazardous substances becoming methodologies that have pre-defined amount for the modeling and calculation of the level of risk and it cannot be applied especially in cases of under-limit amount. However, this sublimit amount, may involve not only the risk of acute, but also risk of chronic, which may manifest itself only after time. The input of hazardous substances to the environment at the same time can get into of such the important systems such as the food chain. Contamination of its individual components, threatens all consumers in this system.

### 4.3. Risk Assessment – Risk analysis and methods

As mentioned above, the transport risk assessment based on specific risk analysis and methods, has limited possibilities of choice and combination methods by reason the mobility of sources of risk and thus difficult to predictions of the severity of the risk. The best approach is the study of the characteristics and the application of preventive and safety measures according to the worst-case scenario, based on the analysis of the highest level of risk. For example, the transport of hazardous substances through densely populated areas, areas with the vicinity of source of drinking water, protected landscape area, etc. All of these factors increase the level of risk and may occur simultaneously in combination.

Among the analyzes that are used most often are included in the Quantitative Risk Analysis (QRA), the recommended method Purple Book, index methods, such as Fire&Exposure Index (F&EI) and H&V index (environment, Czech Republic), IAEA TECDOC 727 etc. These also includes the software tools such as freely available ALOHA, software QRAM (particularly the issue of transport of dangerous goods by road tunnels and determining the amount of social risk) and many others. Very effective for analyzing the causes and consequence are the tree diagrams (Fault Tree analysis – FTA, Event Tree analysis - ETA), and additional methods such as Failure Mode and Effect analysis (FMEA) etc. Any of these methods can be applied in the context of transport of hazardous substances (Bernatik, 2006). Although the analysis IAEA TECDOC 727, is considered by some experts as outdated and chosen instead of it other methods. However, this method is still used, e.g. Fire Brigade. Most of these methods are focused on the risk analysis of social or individual, the consequences for the inhabitants near the accident, their vulnerability. In the case of risk assessment issues for the environment, it is necessary to use other methods.

*Environment*

In the risk analysis is necessary to choose a combination of methods, which is important in terms of results and their complexity since so far there is no such method, which analyze all vulnerable fields and risks at a time. Examples of specifically focused methods in the field of environment, can be a methodology developed at the University of Ostrava in 2003, "Methodology for analyzing the impact of accidents with the participation of hazardous substances in the environment - H&V index"(Ministry of the Environment of the Czech Republic, 2003), and is recognized and recommended by the Ministry of the Environment. The methodology was developed for monitoring environmental vulnerability. It also includes the overall level of damage to individual objects and therefore also of the environment. The methodology deals with the ecotoxicity of the substances to soil and aquatic environments. The method also used methodology of other methods of risk analysis, as just mentioned methodology IAEA TECDOC 727.
A similar methodology is "Methodology for vulnerability analysis environment - ENVITech03." Methodology focuses on the vulnerability of the biotope of the environment - surface water - groundwater. The vulnerability of the environment is expressed by the relationship between the probability that it will release (accident) and simultaneously reducing capabilities and scope of the accident. Methodology ENVITech03 does not include factors that can cause damage. The methodology is based on the physical properties of objects (Ministry of the Environment of the Czech Republic, 2003). With usage of these techniques, it is possible to analyze the vulnerability of the environment along the transport routes and not just on the road. The methods can also be applied in terms of railway transport. These methods can be included in the risk assessment for the environment, including its non-production functions within each of biotopes. Currently, the Institute of Forensic Engineering BUT researches the usability of these methodologies as a support tools for determining of the amount of environmental damage, arising on non-production environment ingredients. Their application, could be not only the issue of valuation, but and risk assessment and cooperation with safety advisors in the planning of transport routes.

Application of the above methods in Chap. 4.3 is an essential part of the overall assessment of the risks associated with the transport of hazardous substances. As already mentioned, it is important to not forget the hidden processes - possibilities of small-scale damage and transport units, which are not routinely and immediately apparent. Their origin can be technical and prove with on the vehicle equipment and its functionality. It may be, for example the drips of hazardous substances (mostly dripping fuel), when is a vehicle in safety mode breaks can occur during this time in the form of dripping leak substances into the environment and causing damage that requires intervention and remediation measures. In some cases, where there is a minor release, nature is able regeneration. The example can be a leakage of fuel into the environment, when in the release of 10 l/ $m^2$ gasoline soil regeneration takes approximately one year. Recovery time then rises at higher fractions of soil contamination (Rabl, 1991). Precisely these technical malfunctions of the vehicle are very often detected by police checks, which are conducted regularly throughout the year. In most of the faults are found on ADR vehicles transporting substances within the national transport. It is important to mention that during the inspection of vehicles carrying hazardous substances control authorities must choose a place where the level of the risk is minimal. Control would not take place in the village or place where it can be dangerous to a greater numbers of healths or life of people and the environment.
Another factor, that can be hidden could imply a hazard to the transport and occur at any time during the transport, is the human influence. It could be a normal vehicle driver fatigue in not only ADR mode, but also surrounding drivers who lose their attention. The sudden changes in behavior caused by the external influences, such as aggression against another driver or the contrary condition of apathy, which can be caused by the monotonous activities. Despite this undergo the course of obtaining the professional driving license, drivers of cars psychological tests, do not include any special conditions for drivers of vehicles under the ADR mode and awareness on the hazardous properties or risks associated with the transport of that mental condition.
It is important for the calculation of probabilities and determining the economic aspects of measures, and keeps a reserve for these undesirable effects, which, even with a small probability can occur and may cause more serious accidents and other negative manifestation.

**Conclusion**

The comprehensiveness of the risk assessment of transport is due to non-stationarity of a very challenging task. Many factors can have a negative affect during the transport on the transport unit and can result in the release of hazardous substances and the full manifestation of its physical properties. In the assessing and risk analysis it is important to include all its phases, as well as the loading works, safety breaks during the transport, where the hidden processes can take place simultaneously and may lead to serious accidents. In the context of the transport planning and transport routes, it is necessary to be mindful of the trail character, the character of landscape protection areas around which the transportation of certain substances is not prohibited.
In the terms of risk assessment a risk engineer (expert) should always include factors such as transport volume, hazardous properties of the substance being carried (the full manifestation of the hazardous properties of the substance being carried), the possibility of initiation, the characterization of transport and transport routes (road conditions, safety regulations, traffic intensity, risk segments, etc..), protected landscape areas and biotopes, population density along the transport routes, sources of drinking water, the risk of accidents caused by human factors (ADR mode fault of the

driver, other road users, the human factor within the loading work - technicians, etc.). Assigning the values to various risks is then dependent on the probability of their occurrence and selected methods. Using a combination of several methods is then also the verification of the correctness of the results and possible application of security measures. For risk assessment and risk analysis, it is always advisable to determine the research team, which would at least consult specialist areas with known professionals or experts who are working in the environment and known, e.g. technical equipment well.

The actual results of the risk assessment should be comprehensive and include not only a threat for population, but also for the environment and property. It is good to always check the results with one of the methods of risk analysis such as the FMEA, where it should be established whether the proposed safety measures for reducing risks fulfils its function are sufficient or on the contrary are excessive, costly, or the contrary, in some cases, mean the inception of new risk factors.

The importance of risk assessment for the transport of dangerous goods by road shows itself as an important source of information and supporting data in the forensic sciences, where in the case of an accident the causes of it are retrospectively investigated as well with the seriousness of the consequences. Particularly concerning the pre-selected security measures, choice of routes ensuring the adequate technical condition of vehicles, etc. Further step in the case of importance evaluation of the risks of transporting of hazardous materials, is also linked to the forensic monitoring of the environment and surroundings after the accident, where the leakage of hazardous substances has been. In that case, the evaluation and findings of the severity of the risks involved with the substance after the leakage when the evidence of the Environmental Forensic can be applied. Based on the determination of the extent of environmental contamination, the acute and chronic risks can be analyzed, which may occur even after several years.

The risk assessment, whether the risks are obvious or hidden, is a very important field that can also be applied in multiple disciplines and domains such as the mentioned forensic science. With an appropriate choice of risk analysis methods, it is possible to prevent accidents with larger and smaller leakage of the hazardous substances or at least ensure such measures to reduce the consequences of the accident to a minimum.

Finally, if we compare the accidental release of the hazardous substances on the road and on the rail, the frequency of events is greater during the road transport. However, it is important not just for the conditions in the Europe to not forgetting about the rail transport, where there are almost nearly the same risks. Although the units for the transport by rail are considered stable, their traffic volumes are higher and there is a risk of leak and other manifestations in all phases of transport with serious consequences. Another serious risk is the availability of wagons waiting for the transport which could become relatively easy target e.g., for a terrorist attack. The issue of risk assessment of railway transport will be the subject of further research.

**References**


Bernatik, A. 2006. Prevention of major accident II. SPBI in Ostrava, Czech Republic. 86 p. (language: Czech)

Brozova, K. 2008. The risks in connection with road transport of dangerous matters. In. *Perners Contacts*, 3(3): 3 – 7. Available from internet: <http://pernerscontacts.upce.cz/10_2008/Brozova.pdf>. (language: Czech)

European Agency for Safety and Health at Work. *Available from internet:* <*https://osha.europa.eu/en/sector/road_transport/introduction-to-managing-risks-to-road-transport-drivers*>. (language: Czech)

European Agreement concerning the International Carriage of Dangerous Goods by Road, ADR, 1 January 2013

European Commission. 1996. *Guidance on risk assessment at work*. Luxembourg: Office for Official Publications of the European Communities. Inc. Germany. 64p.

EuroRAP – Czech Republic. Available form internet: <http://www.eurorap.org/media/178715/EuroRAP_CZ_2009-2011.pdf >.

EuroRAP – Czech Republic, 2014. *Risk Mapping and Performance Tracking.* Available from internet: < http://eurorap.org/partner-countries/czech-republic/>.

Kocianova, S. 2012. *Dangerous chemicals:* The spread of hazardous substances at spill. HZS CR. Available from internet: <http://www.hzscr.cz/clanek/prostredky-individualni-ochrany-nebezpecne-chemicke-latky.aspx>. (language: Czech)

Ministry of the Environment of the Czech Republic. Methodology of Department of Environmental Risks to determine the vulnerability of the environment by ENVITech03 and analysis of the impact of accidents with the participation of hazardous substances to the environment by H&V index. Available from internet: http://www.mzp.cz/C1257458002F0DC7/cz/metodicke_pokyny_odboru_enviro_rizik/$FILE/oer-MP%20ZP-2003.pdf. (language: Czech)

Rabl, V. 1991, *Technology of oil.* Institute of Chemical Technology Prague. Czech Republic, Prague. 256 p. (language: Czech)

Roosberg. J. Thorsteinsson. D. 2002. Environmental and Health Risk Management for Road Transport of Hazardous Material. University of Lund, Lund Institute of Technology Sweden. 29 p.

Statistical data of transport of dangerous goods (2009 – 2012). Police of the Czech Republic. May 2014. (language: Czech).

The number of traffic accidents ADR drivers is very high. 2012, *Dopravni noviny* 20/2012. Available from internet.: <http://www.dnoviny.cz/silnicni-doprava/pocet-dopravnich-nehod-ridicu-adr-je-velmi-vysoky>. (language: Czech)

Vomacka. P. Ten of the most common causes of the accidents. Available from internet: <http://www.uamk.cz/item/1974-deset-nej%C4%8Dast%C4%9Bj%C5%A1%C3%ADch-p%C5%99%C3%AD%C4%8Din-nehod>. (language: Czech)

# ELECTRIC ENERGY CONTROL SYSTEM SIMULATOR FOR METRO RAIL

**N R Mohanan**[1]
*Tata Consultancy Services, Bldg 7, Commerzone, Off Airport Road, Yerwada, Pune, Maharashtra, India 411006.*

**Abstract**: Urban population in developing countries has been growing phenomenally in the last two decades. Accordingly, the transportation needs are also increasing significantly. In order to meet this requirement, many cities are opting for new modes of transport systems like Monorail, Light Rail Transit (LRT), Mass Rapid Transit (MRT) or expanding their existing transport network to new locations.
Safety, reliability and speed are the critical requirements of these transport systems, which are nearly 100% automated in their operation. To achieve this, sophisticated control systems are required and they must be tested extensively under all operating conditions. This calls for a highly automated testing of the control system and this requires a simulator. Though hardware-based simulators can meet some of the testing requirements, they are limited in features, time consuming and more expensive as compared to software-based simulators.
This paper describes a software-based simulator for testing the electric energy control system of a metro rail. The electric network with substation and field devices, remote terminal units and the control system constitute the electric energy system of the metro rail. Each of the field devices has different functions and is controlled centrally as well as locally. Monitoring the field devices and knowing their status is essential for the control system to function. The simulator helps in the validation of the electric energy control system by simulating the remote terminal units, the field devices and the communication between them, in a software lab, thus minimizing the time, effort and cost required during commissioning. Using the simulator various 'what if' scenario studies can be carried out in the lab, which helps in the operational decision at critical junctures and in analyzing operational eventualities. The simulator can be easily customized for new projects and offer many more features and advantages.

**Keywords:** Metro Rail, Electric Energy Control System, Simulator, Testing

## 1. Introduction

Railways can be broadly classified into two groups namely, mainline and transit or metro. While the mainline deals with long distance cross-country rail transport, the metro deals with rail transport within city limits. With the exponential increase of urban population in developing countries in the recent past and with the increasing nature of this trend, it is quite clear that metro rail is becoming an integral part of such cities.

Metro rail offers faster and less expensive mass-transport as compared to other modes. The demand for such transport services at high frequency, reliability and safety, is a characteristic of any large city, and is one of the main requirements to be satisfied by metro rail. Its continuous service to the population cannot be compromised under any circumstances. Hence, such rail system should be robust in design and highly automated in operation. Extensive end-to-end testing of such system is essential to ensure a fail safe and reliable operation.

This paper describes a software-based simulator for testing the electric energy control system of a large metro rail. This control system is part of the Centralized Traffic Control (CTC) of the metro rail. A brief introduction to the CTC, electric energy system and design and development challenges of the electric energy control system is provided in the initial sections of the paper and subsequently the key aspects of the simulator namely its functions, its architecture, its components and testing using simulator are described.

## 2. Metro Rail

From an operational point of view, metro rail has two key elements under it, namely, traffic system and electric energy system. The traffic system deals with the operation and movement of rail cars along the rail network while the electric energy system deals with the distribution of power to third rail (for traction), equipment, passenger stations and auxiliary devices for the safe and reliable operation of the rail cars at the required speed and frequency. In any modern metro rail, the monitoring and control of its traffic and electric energy systems are centralized under the CTC and can be automated to nearly 100%.

### 2.1. Centralized Traffic Control (CTC)

The electric energy control system of a metro rail is part of the CTC, which consists of a number of high-end servers, consoles, mimic panel and communication interfaces. It performs monitoring and control activities in real time and ensures safe and reliable operation of the metro rail. Fig. 1 shows the layout of a CTC that is employed in large metro rails.

[1]Corresponding author: nr.mohanan@tcs.com

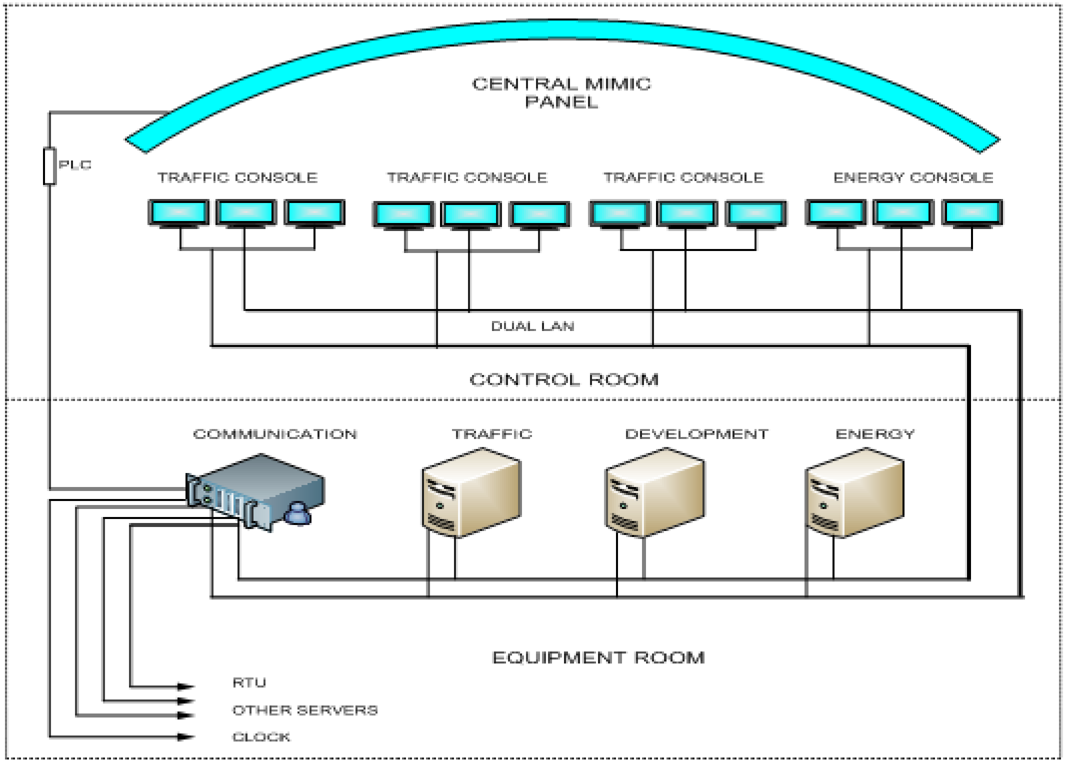



**Fig. 1.**
*Layout of Centralized Traffic Control*

The CTC has a few operational consoles for monitoring and controlling the traffic and electric energy systems. These consoles typically have multiple screens displaying the traffic movement and electric energy system status in a seamless manner, across two or three stations. The operator can fetch in any station of the traffic or electric energy system for his view by pressing few keys on the console.

The back-end servers are the nerve centre of CTC, performing sophisticated tasks such as monitoring and controlling both the traffic and electric energy systems in real time. These servers are connected to stations, Remote Terminal Units (RTU) and field devices through communication interfaces. The CTC of a network may also be connected to the control centres or computer systems of other networks, for an overall integrated operation of the city rail traffic.

### 2.2. Electric Energy System

Fig. 2 shows the schematic diagram of the electric energy network of a metro rail for one of its lines. The electric energy network has primary stations, rectifier stations and auxiliary stations. The primary stations receive 88 KV power supply from the feeder grid and convert it into 22 KV for further use. The rectifier stations convert the 22 KV power supply to 750 V DC and feed it to third rail for traction. The auxiliary stations convert the 22 KV power supply to 480 V to meet the requirements at the passenger stations and electric equipment.

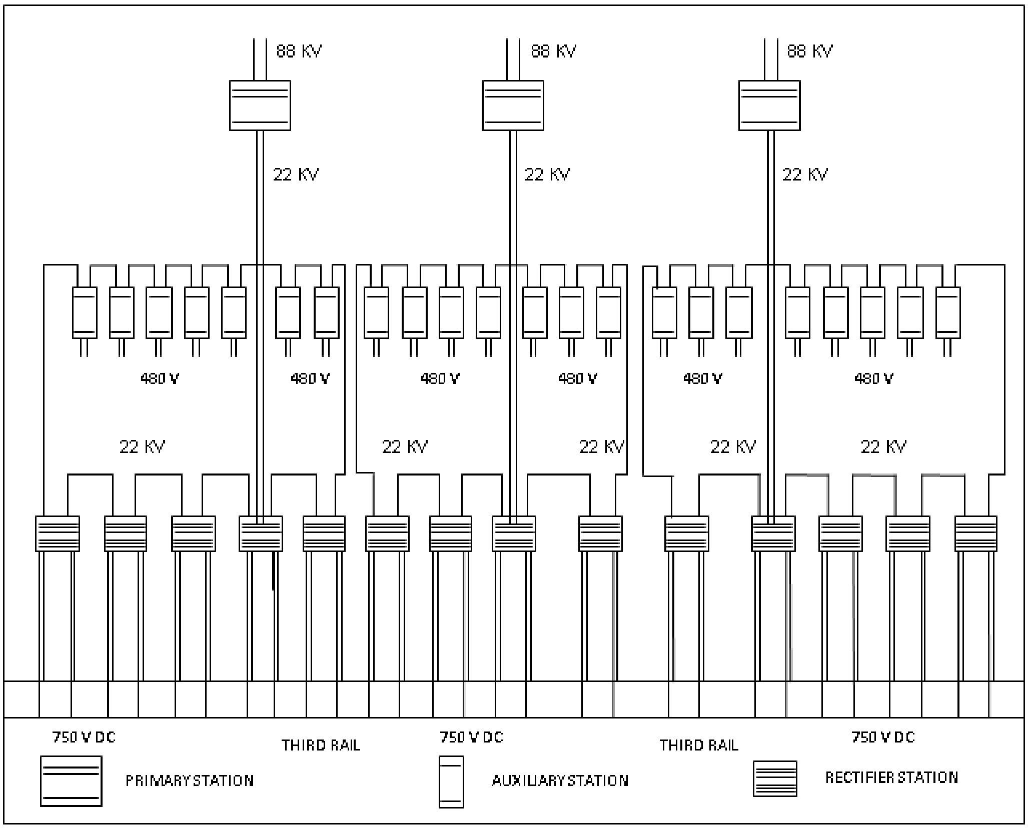



**Fig. 2.**
*Schematic of Electric Energy Network*

The primary stations are interconnected through switches and each of them is capable of meeting the power requirements of the entire line. Each alternate passenger station has a rectifier station that ensures an even distribution of electric energy to the third rail. Each passenger station has an auxiliary station that supplies 480V electric energy to meet the demands of elevators, air-conditioning system, interlocking system, signalling system and so on.

Each of the primary, rectifier and auxiliary stations also has remote/local switches to switch the operation to remote mode when the control is to be done from CTC. Each of these stations also has auto/manual switches to make the transformer switch-over automatic, whenever this is trouble in the bus feeding the transformer. In this way, the entire electric energy network is interconnected and automated through various switches and circuit breakers, for an uninterrupted power supply.

All the field devices and control switches at stations are wired to a RTU and all such RTUs are in turn linked to the CTC through separate communication channels thus establishing connection between the CTC and field devices for its monitoring and control. The status of the field device is detected by the RTU and this information is communicated to the CTC. Similarly, CTC sends commands the field devices through RTU, which are executed by the field device.

### 2.3. Design and Development Challenges of Electric Energy Control System

The design and development of the electric energy control system go almost in parallel with the engineering, design and development of the infrastructure and equipment of the project. Once the details of the track layout and stations, electric network and stations, RTU and field devices are available, there is fair amount of information to start the design and development process of the electric energy control system.

The activities involved in this design and development can be broadly classified into:

- Consolidation of the substation and field device data
- Design and development of the Human Machine Interface (HMI)
- Customization and modification of the control system
- Configuration of the control system
- Design and development of the communication interface between the control centre and RTU
- Testing and validation along with the RTU and field devices
- The electric energy control system with communication to RTU, RTU configuration and communication to field devices is to be tested for the following scenarios and cases:
- Correct display of all the field device status at control centre HMI
- Correct function of the commands sent from control centre at the field devices
- Remote and local operations from the primary, rectifier and auxiliary stations
- Automatic and manual operations at the primary, rectifier and auxiliary stations
- Automatic sequence of operation at primary, rectifier and auxiliary stations
- De-energisation of third rail

This testing, verification and validation is a lengthy and complex process for the following reasons:

- Combination of stations, RTUs and field devices to be tested
- Repetition of testing in a consistent manner during the development life cycle
- Changes in the system till the commissioning date

The testing and validation is to be 100 % to ensure safety and reliability. The testing activity is to be completed fast in each stage of the development life cycle, so that the overall time from design to implementation is minimal. Systematic completion of all the testing activities is very critical so that there are no surprises during commissioning of the system, thus saving precious time and cost during commissioning. Quick changes in the RTU configuration in line with the changes happening in the design should be possible during testing .Typically when a hardware-based simulator is used for testing, there are many limitations and is not a desired solution.

## 3. Simulator for Testing Electric Energy Control System

The complexity and amount of testing required justify the development and implementation of a software-based simulator for testing the electric energy control system. The simulator should provide a menu option, displaying the functionalities provided by it. It should simulate the functioning of all the primary, rectifier and auxiliary stations, the functioning of the RTUs, the communication between the RTU and field devices, and the communication between the RTU and control centre to enable end-to-end testing. It should facilitate stopping and resuming testing as per the convenience of the tester. It should log all the events and facilitate printouts. Most importantly, it should give a real-time response of two seconds and should be easily configurable. The following sections describe its features, and how they are achieved.

### 3.1. Simulator Functions

The simulator should have the functions described in Table 1 to facilitate the testing in a user-friendly manner. The user should be able to carry out the testing activities by clicking the respective button from the simulator screen.

**Table 1**
*Simulator Functions*

| S. No. | Function Name | Function Description |
|---|---|---|
| 1 | INITIALISE | Simulator initialises communication and loads a default station and field devices |
| 2 | LOAD | Simulator loads a previously stored snap and updates the display to facilitate the continuation of testing activities |
| 3 | SAVE | Simulator saves a snap of the status of all the devices |
| 4 | LOAD SCRIPT | Simulator loads a script file with testing steps. Testing steps consist of actions on devices with a time stamp |
| 5 | RUN SIMULATION | Simulator in execution mode; accepts commands , changes the device status, communicates with CTC etc. |
| 6 | RUN VERIFICATION | Simulator helps to verify the actions specified in the script file. No communication with CTC during this process |
| 7 | RUN CONFIGURATION | Simulator facilitates status changing of device before testing begins. No communication with CTC during this process |
| 8 | PAUSE | Simulator pauses the execution |
| 9 | PRINT SCREEN | Simulator prints the current screen |
| 10 | PRINT LOG | Simulator prints the log file |
| 11 | OPEN | Simulator opens a field device |
| 12 | CLOSE | Simulator closes a field device |
| 13 | DEENERGIZE | Simulator de-energises third rail |
| 14 | EXIT | Simulator exits |

### 3.2. Simulator Architecture

A high-level architecture of the simulator and its interfacing with the CTC is shown in Fig. 3. The simulator consists of essentially three processes; (a) Simulation Controller (b) HMI and Field Device Simulation and (c) Communication Interface.

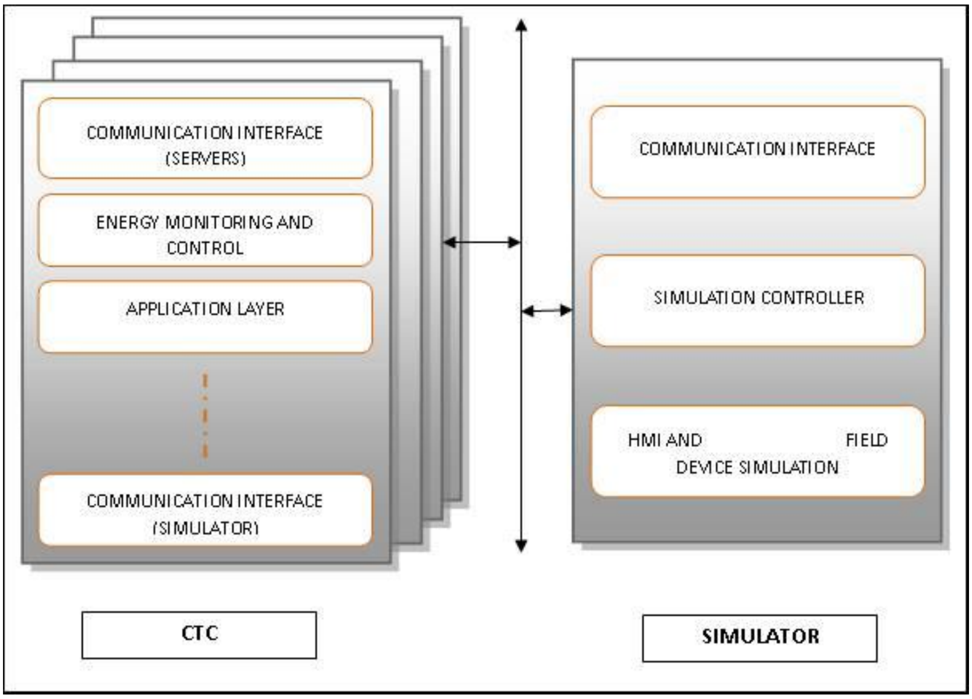



**Fig. 3.**
*Organization of Electric Energy Control System Simulator*

While the Simulation Controller is the main process, the HMI and Field Device Simulation Process manage the display and user interaction and the Communication Interface Process manages the communication with CTC. In real life scenario the communication between the CTC and the RTU is through serial line, which is simulated through the LAN by making the required changes in the CTC.

#### 3.2.1. Simulation Controller

This is the supervisory process controlling all the simulation activities. It facilitates all the simulator functions as described in Table 1, by providing a menu bar with selection options. The transition between the simulator functions is explained in the state transition diagram shown in Fig. 4.

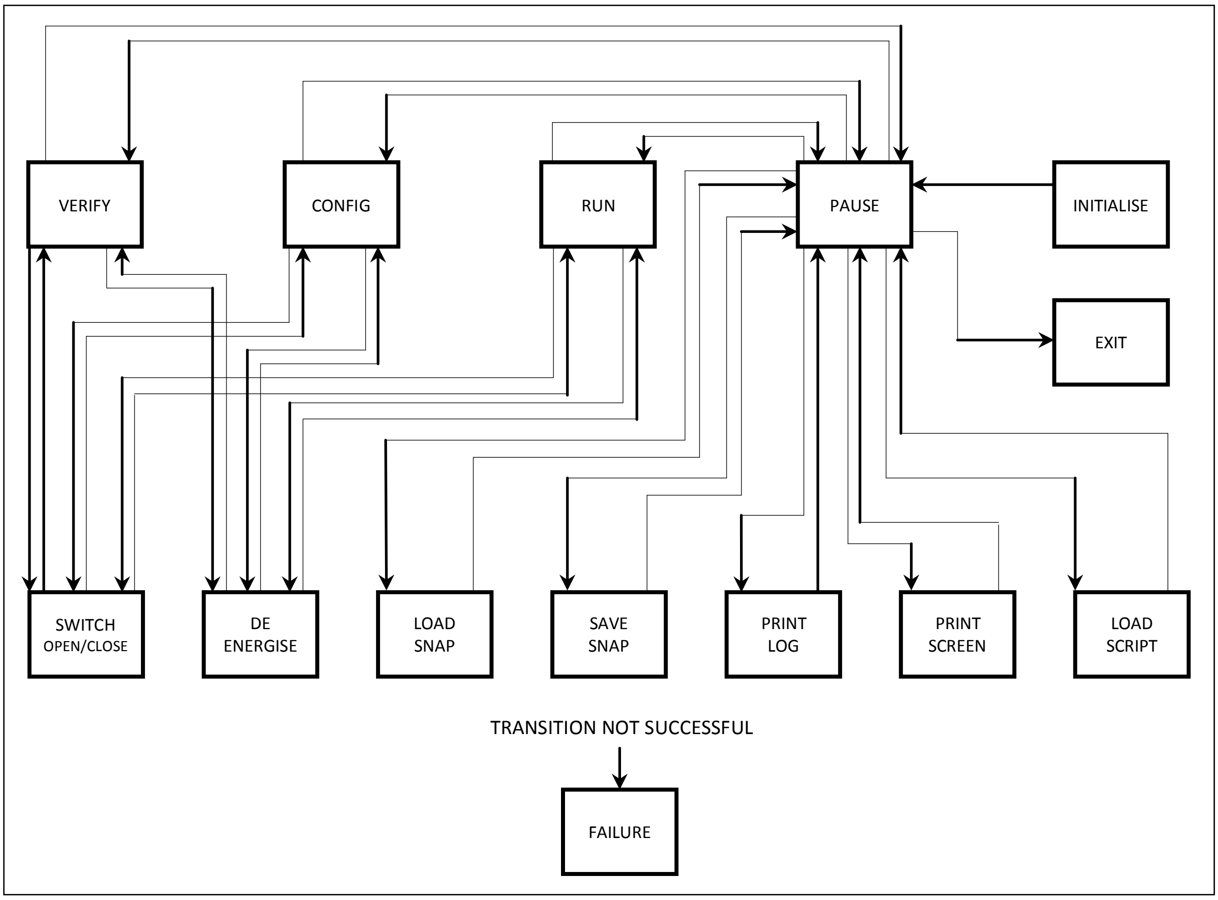



**Fig. 4.**
*Electric Energy Control System Simulator State Transition Diagram*

From the INITIALISE state, the system navigates to PAUSE state. From the PAUSE state the tester can navigate to RUN, CONFIG, VERIFY, LOAD SCRIPT, PRINT SCREEN, SAVE SNAP and LOAD SNAP states. From the RUN state, the tester can navigate to SWITCH OPEN/CLOSE and DE ENERGISE states. From the CONFIG and VERIFY states, the tester can navigate to SWITCH OPEN/CLOSE and DE ENERGISE states. From the PAUSE state the tester can navigate to EXIT state.

The mapping between the electric energy control system tags, RTU tags and field device tags are done in this module. One of the criteria while deciding the architecture and communication protocol of this module is the real time performance of the simulator. The statuses of all the field devices stored in the HMI and field device module are transmitted to the electric energy control system in less than two seconds and similarly any command coming from the electric energy control system is passed on to the HMI and field device module and it is executed in less than two seconds. The effect of this is again communicated back to the electric energy control system, operating in a cyclic manner.

This module also facilitates the internal communication between all the processes running in the simulator computer.

### 3.2.2. HMI and Field Device Simulation

The HMI consists of the simulation menu bar with buttons for all the functions described in Table 1and screens for all the primary, rectifier and auxiliary stations. The simulation menu bar is displayed at the top of the HMI and below which the screen for the electric network is displayed for the selected station. While the simulation menu bar remains constant across all the screens, tester can bring in the screen for any of the primary, auxiliary or rectifier stations by keying in the respective code for that station. This module has its own tags for the simulator menu items and the field devices. As seen in the simulator organization diagram this module is a separate process and it communicates with the simulation controller process through shared memory and message Qs. Sufficient attention is given in the design and development of this module to get a real time response of two seconds.

### 3.2.3. Communication

The communication interface module facilitates the communication between the simulator and CTC through LAN. Both at the simulator and CTC ends, this module carry out the protocol conversions for communication through LAN. At the simulator end, message Qs are employed for the inter process communication between this module and the simulation controller module. Again sufficient attention is given in the design and development of this module to get a real time response of two seconds.

## 3.3. Testing using Simulator

The tester needs to prepare a test plan for all the tests to be carried out, namely, validation of the field device indications at control centre, validation of the commands from the control centre at the field devices, validation of the remote and local operation at substations, validation of the auto/manual operation at substations, and validation of the automatic sequence operation at substation and so on.

The electric energy control system tags, RTU tags, field device tags and the mapping between these are to be completed and configured in the system to commence the testing activities.

After initializing the simulator, the tester either starts from the very beginning or loads a snap file from where he wants to continue the testing activities. The tester can perform any of the operations as if doing from the field or substation and see its effect both at the control centre HMI and the simulator HMI. The tester can also load a script file, where the operational steps are specified with a time stamp, which will be executed by the simulator.

When the simulator is in pause mode, the tester can print any of the screens, print the log of all the events happened in the simulator, take a snap of the simulator status or load a snap to continue the testing from a different scenario.

By following these guidelines, the tester carries out all the tests as per the test plan and validates the system.

# 4. Conclusion

The simulator developed as per this architecture has helped to test and validate the functioning of the electric energy control system by simulating the RTUs, field devices and communication between them in a lab replacing hardware-based simulator. The simulator was able to provide a real time performance of less than two seconds, meeting the project requirement.

One of the challenges faced during the development of the simulator is in RTU mapping, which kept on changing until the last stage. In the simulator, these changes were quickly accomplished by loading the new mapping into the system and continuing with the test. This has saved time and has helped in the overall development cycle of the project.

Testing by running the test scripts, automates the testing process and became very helpful when tests were to be consistently repeated. This has helped to save time, cost, and in achieving consistency.

Overall, the simulator has helped to complete the end-to-end testing of the electric energy control system in a systematic and consistent manner. It has helped to document all the results along with test plans. All the critical stages during the testing could be saved as snap files, which could be used for future reference or testing activities. It has helped to minimize the issues during commissioning.

The simulator is a very handy tool to quickly analyse any future events and study ‘what if’ scenarios, which could occur at any point of time in the life of the metro rail. This is a useful feature to take operational decisions at critical junctures, as well as in analysing operational eventualities.

The simulator could be easily modified to include other lines and to accommodate any future expansions in the existing line. The simulator also helps in timely decision-making and design optimization, which could be very vital from a cost perspective.

**Acknowledgement**

The author wishes to express his sincere thanks to the management of Tata Consultancy Services’ Engineering and Industrial Services unit for the encouragement and support provided.

**References**

Bach, Maurice J. 1986. *The Design of Unix Operating System*, Prentice Hall of India Pvt Ltd.

Page-Jones, Meilir Page 1981. *Practical Guide to Structured System Design*, Yourdon Press New York.

Stevens, W. Richard 1990. *Unix Network Programming*, Prentice Hall of India Pvt Ltd.

# TRAFFIC CHARACTERISTICS AND ACCIDENT ANALYSIS ON THE UNDIVIDED ROADS WITH INCREASED NUMBER OF LANES

**Michal Radimský[1], Radka Matuszková[2], Michal Kosňovský[3], Luboš Kabeš[4], Jiří Čepil[5], Martin Smělý[6]**
*[1,2,3,4,5,6] Brno University of Technology, Faculty of Civil Engineering, Institute of Road Structures, Veveří 331/95, Brno, Czech republic*

**Abstract:** Increasing the number of lanes in steep ascent and descent is designed due to the decreasing speed of slow and very slow-moving vehicles affects the smoothness of traffic flow and safety. But even in such arrangement lanes there are risks that can cause traffic accidents, especially at the point of lane reduction. A common problem is improper placement of the point of lane reduction on the horizon. This article analyzes the traffic accidents in this type of arrangement, and assesses the main causes of accidents, or the accumulation of accidents in particular place. Further offers information about the difference between the speed of the vehicle categories and the number of overtaken vehicles on two-lane sections and sections with increased number of lanes.

**Keywords**: accident analysis, undivided road, slow vehicle, lane reduction, overtaken, relative accident rate.

## 1. Introduction

To maintain level of service, increased number of lanes is proposed on route sections where longitudinal tilt decreases the speed of slow and very slow vehicles. Increased number of lanes can be also proposed on long downhill route sections, but this is not usually used in Czech Republic (Čepil et al., 2012). Another usage of the increased number of lanes is the 2+1 road lane arrangement. This is a special arrangement of two-lane road that has been introduced in the standard for rural roads in 2009. Principle of the 2+1 road lane arrangement is alternate allocation of middle lane for each direction. It is appropriate to coordinate lane adding with the longitudinal tilt and use this extra lane as an uphill lane. The 2+1 arrangement is the main reason for accident analyzes in areas with increased number of lanes (Marvan, 2013). This issue is addressed in the context of a Technology Agency grant project.

## 2. Accident analysis

72 sections with additional uphill slow lane were chosen for the analysis. All these sections are located on 1st category roads. For each analysed section were noted and compared main design and traffic characteristics as length and type of reduction. Also the intensity and the share of trucks in the traffic flow was noted from National traffic census 2010. Geographic information system Unified transport vector map provided information about traffic accidents in 2007 – 2013 (Ministerstvo dopravy et al., 2006). In the final accident analysis are also included following characteristics: number of accidents in the area, the number of accidents at lane connection section, number of accidents resulting in death or injury, number of accidents with animals and a major cause of traffic accidents. The number of traffic accidents is considered as a sum in both directions except the accidents at the lane connection section. Accidents related to usage of alcohol were removed from the list.

### 2.1. Basic accident indicators

For the analysis of accidents, we used several standard indicators, which take into account not only the value of a number of accidents, but also the intensity, time period or degree of injury. Relative accident rate indicator, density of accidents indicator and the indicator of societal losses was quantified for each section.
Relative accident rate indicator is the most commonly used characteristic and indicates the likelihood of accidents in the area in relation to the transport performance.

$$R = \frac{N_O}{365 * I * L * t} 10^6 \tag{1}$$

In formula we work with these variables: total number of accidents in the period ($N_O$), average daily traffic intensity (I; vehicles/24hrs), length (L; km) and period (t; years). Unit of the relative accident rate indicator (R) is the sum of accidents/million vehicle-km and year).
Indicators values are relative and usually range between 0.1 and 0.9. Higher value of the indicators is pointing out at a security gap and values higher than 1.6 mean essential road safety shortcoming s (VŠB-TU Ostrava, 2014).

[1] Corresponding author: radimsky.m@fce.vutbr.cz

### 2.2. Type of lane reduction

In Czech Republic there are two ways to reduce the number of lanes. The first one, often used and based on earlier legislation, is based on cars traveling in the slow uphill lane joining the fast lane. In the following text we will call this type of lane reduction as an EXTERNAL (Fig.1.). The external type can lead to dangerous situations as the slowly moving vehicle joins into a fast moving continuous lane and then an unexpected deceleration in the continuous lane can lead to an accident (Matuszková et al., 2014)
The second method, which is recommended by current applicable regulations, the continuous lane is the right slow lane and the added lane is the left fast lane so faster vehicles join the slower continuous lane. In the following text we will call this type of lane reduction as an INTERNAL (Fig.2.). This sorting method is also used in road layout 2+1. Placement of the lane connection section has also impact on accidents or dangerous arrangement.

**Fig. 1.**
*Type of lane reduction - external*
*Source: (Authors)*

**Fig. 2.**
*Type of lane reduction - internal*
*Source: (Authors)*

Change of the number of lanes should be designed in locations with adequate stopping sight distance and should not be located at the horizon and also should be designed at the location where the standardized slow vehicle already reach sufficient speed. If these requirements are not met, significant deceleration in traffic flow can cause dangerous situations on the road network (Derr, 2003).

### 2.3. Statistical evaluation

In all 72 sections with an increased number of lanes relative accident rate indicator was quantified. The average value of the relative accident was 0.8 accident / million vehicle-km and year. The average indicator value thus falls within the safe range, however, as shown in the histogram (Fig.3.), there are also sections that exceed the value of 0.9, and nine sections exceeds 1.6, which indicates major deficiencies in safety.


**Fig. 3.**
*Histogram of relative accident rate*
*Source: (Authors)*

During the accident evaluation a large share of accidents with animals was found. In almost half of the sections it is greater than 20%, including 16 sections where the share is greater than 50% (Fig.4.).

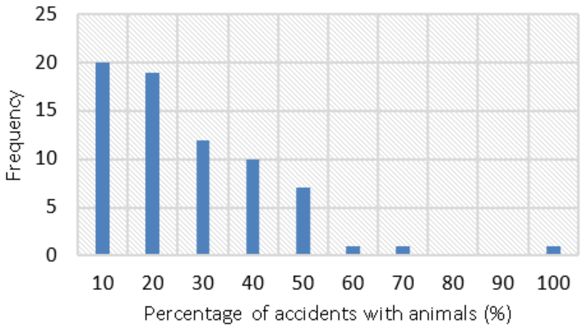
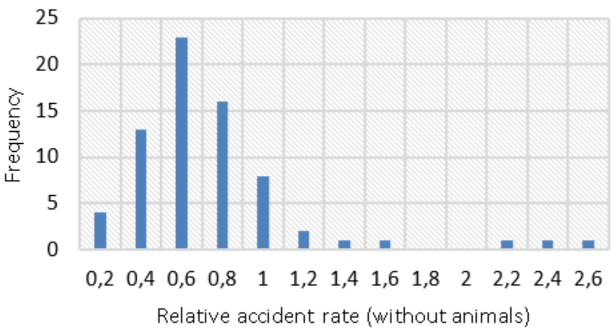



**Fig. 4.**
*Histogram of percentage of accident with animals*
*Source: (Authors)*

This finding has several possible reasons. One is the environment, increasing number of lanes are mostly designed in hilly, rugged unpopulated areas where fields and forests became typical. Another factor is the insufficient outlook for overtaking, where the overtaking vehicle cannot see moving animals over the slower car and the width of the road, which extends the route of passing animals. When disregarding animal related traffic accidents for calculating the relative accident rate indicator the resulting histogram changes (Fig.5.) and values are logically reduced.

**Fig. 5.**




*Histogram of relative accident rate without animals accidents*
*Source: (Authors)*

The occurrence of a large share of animal related accidents was also confirmed by the analysis of the accidents causes on various road sections. On 33% of the sections the most frequent cause was "not caused by driver" (Fig.6.). This cause indicates either animal related accidents or pedestrian related accidents. Due to the nature of the environment pedestrian fault was rejected. The second most frequent cause was the failure to adapt the speed to road conditions. Not on one of the sections the most common cause of accidents evaluated had a direct relationship with overtaking. However, it can be concluded that the "failure to comply with safe distance behind the vehicle" and "failing to control vehicle” refers to vehicle joining the continuous traffic flow.

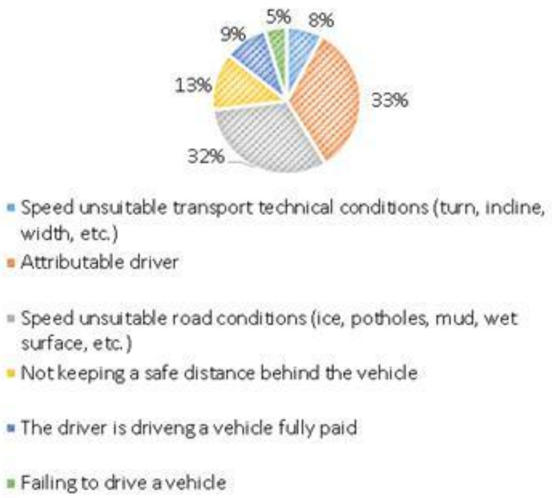



**Fig. 6.**
*Major cause of traffic accidents*
*Source: (Authors)*

During the analysis, we focused on lane connection section, and therefore, we quantified the proportion of traffic accidents on lane connection section and the total number of accidents. On more than half of the sections the proportion of accidents varies up to 20%. The rest of the section has a larger share of accidents. In one case 100% of the accidents happened on the lane connection section. The results are shown in the histogram (Fig.7.).

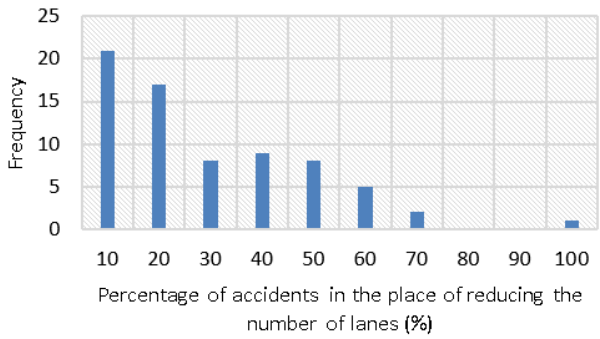



**Fig. 7.**
*Percentage of accident in the place of reducing the number of lanes*
*Source: (Authors)*

As described above, there are two ways of designing lane connection sections. Only in 10 of the 72 1st category sections were designed with the internal connection. The proportion of the internal and external lane reduction is 14% to 86%. Due to the small sample we cannot clearly determine which type is better from a security perspective. When quantifying the average value of accident share at the critical point for all sections we receive number 22.5%. When dividing these two types we get 23% for external and 18% for internal.
In the next part of the analysis relationship between various indicators was investigated. For some characteristics the assumptions were confirmed, for other the dependency was not found.
One of the dependencies that have not been found is the relationship between the length of the section and relative accident rate indicator. The graph (Fig.8.) was plotted with sections sorted by ascending length (bar chart) and curve with values of relative accident rate indicator in red with animal related accidents and in green without animal related accidents. The result can be seen in the following chart.

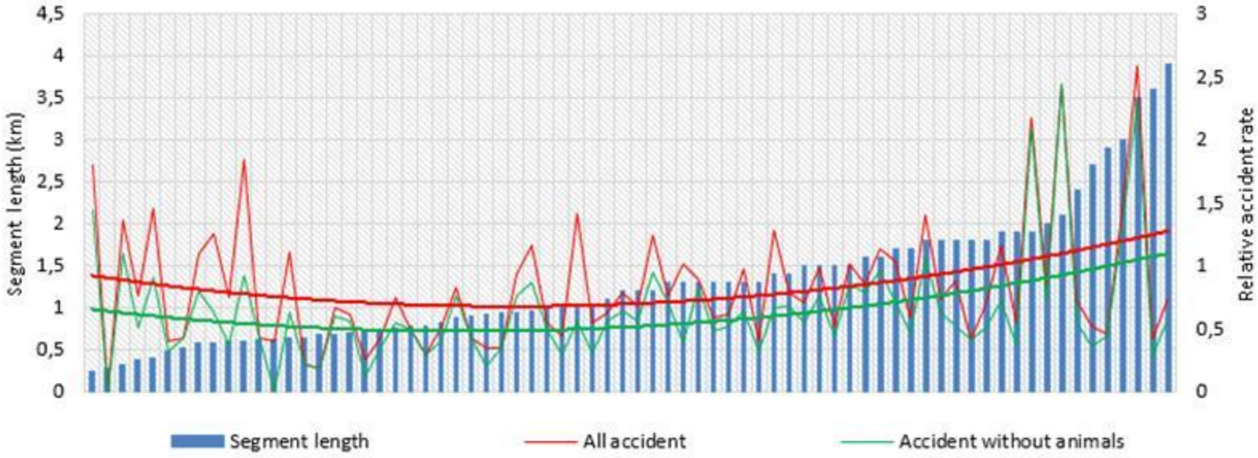



**Fig. 8.**
*Relationship segment length and relative accident rate*
*Source: (Authors)*

One of the results of indirect dependency is a relationship between the section length and percentage of accidents on the lane reduction section. The graph (Fig.9.) shows that with increasing section length, the proportion of accidents drops at the lane reduction section compared to the total number of traffic accidents. On shorter section is more problematic to join the continuous flow and more traffic accidents occur at the lane reduction section.

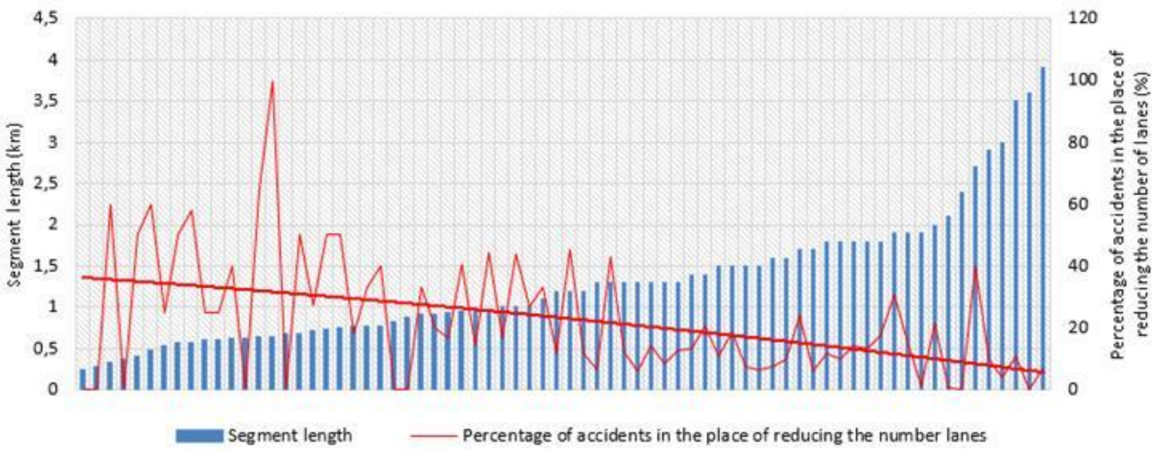



**Fig. 9.**
*Relationship segment length and relative accident rate*
*Source: (Authors)*

Another comparison was made with the relative accident rate indicator. Sections sorted in ascending order according to their indicator values and added curve of their other parameters. The graphs (Fig.10., Fig.11.) can be seen that intensity of the traffic flow and the traffic flow share of trucks in percentages has no direct effect on the relative rate of accidents in the area.

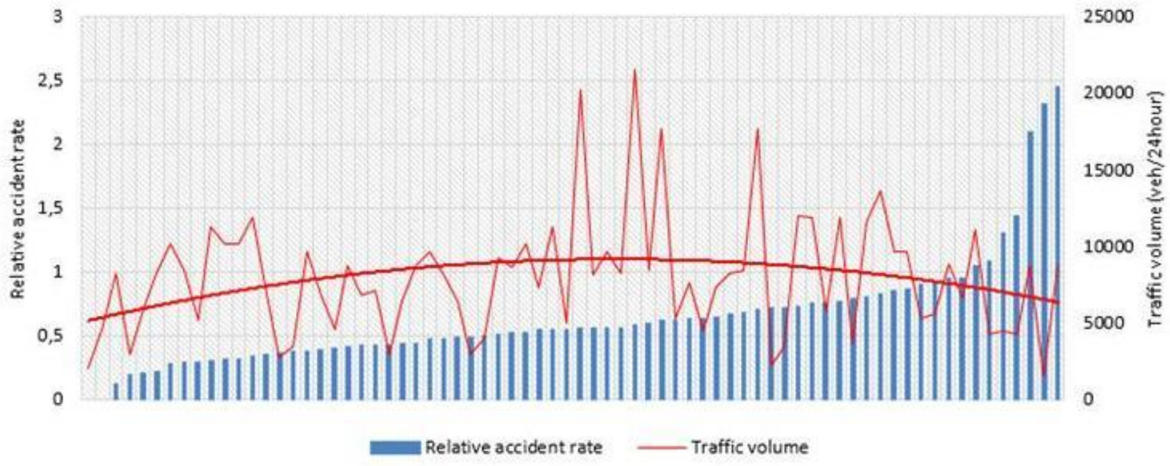



**Fig. 10.**
*Relationship segment length and relative accident rate*
*Source: (Authors)*

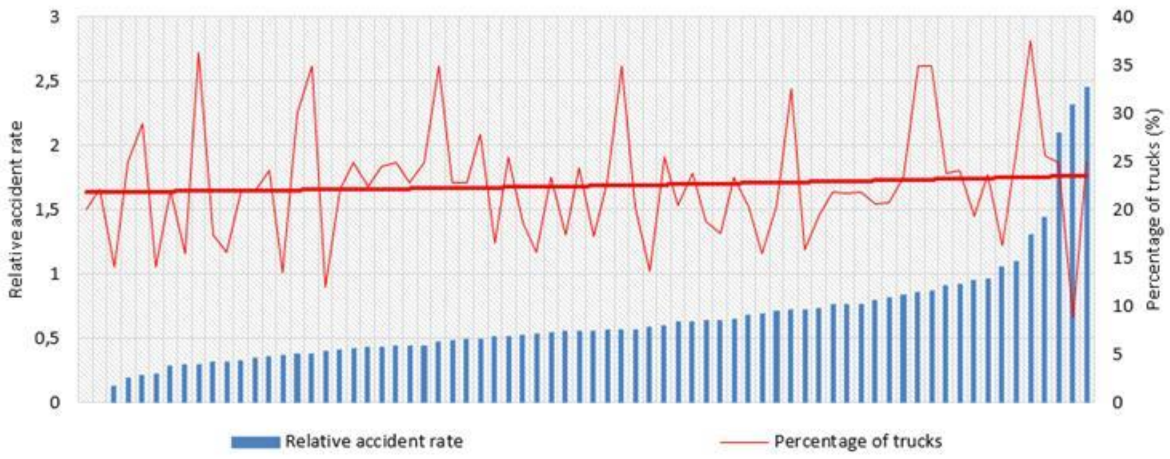



**Fig. 11.**
*Relationship segment length and relative accident rate*
*Source: (Authors)*

**3. Measurement transport characteristics**

The measuring principle was to find out change the traffic flow section speed, when uphill lane to the two-lane configuration and how it will affect overtaking vehicles. To measure transport characteristics was selected 7 sections consisting of 2 segments – two-lane (1+1) and three-lane arrangement (2+1). All sections have the same speed limit. Average traffic flow section speed was obtained from captured videos with software to analyze car registration number. The cameras were placed at three posts. The first was located at the beginning of the three-lane arrangement, the second station was in place, where three-lane turns to two-lane arrangement, and the last camera was positioned on the two-lane arrangement. Distances between cameras were the same (Čepil et al., 2013).

The cameras record always one running direction. Time in the cameras was synchronized (accuracy 0,02 s), to register exact time when a vehicle passes by the camera's post.

Speeds were analyzed for four categories: passenger car, trucks, buses and trailer sets. Traffic flow section speed was counted from the time and length of the segments (between cameras). By changing the sequence of registration marks on camera's posts the number of overtaking was determined. The following table (Table 1) and graphs (Fig.12., Fig.13.) show that the speed of cars on all measured three-lane segments was higher than on two-lane segments. The average increasing of the segments passenger car speed is 17%. The minimum increase is 6 %, maximum value is 31 %. On all three-lane segments was higher longitudinal slope than on two-lane. For small slopes could be even greater improvement (Kabeš et al., 2014).

**Table 1**
*Section car speed*

| Speed (km/h) | Section 1 | | Section 2 | | Section 3 | | Section 4 | | Section 5 | | Section 6 | | Section 7 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Two - lane | Three - lane | Two - lane | Three - lane | Two - lane | Three - lane | Two - lane | Three - lane | Two - lane | Three - lane | Two - lane | Three - lane | Two - lane | Three - lane |
| **Passenger cars** | 85 | 97 | 83 | 97 | 80 | 87 | 76 | 94 | 96 | 101 | 61 | 79 | 68 | 86 |
| **Trucks** | 80 | 80 | 78 | 85 | 78 | 71 | 74 | 70 | 85 | 80 | 60 | 70 | 62 | 72 |
| **Buses** | 86 | 81 | 81 | 79 | 78 | 68 | 71 | 77 | 0 | 0 | 57 | 70 | 62 | 71 |
| **Trailer sets** | 83 | 74 | 71 | 71 | 74 | 65 | 71 | 74 | 87 | 78 | 59 | 69 | 63 | 72 |

*Source: (Authors)*

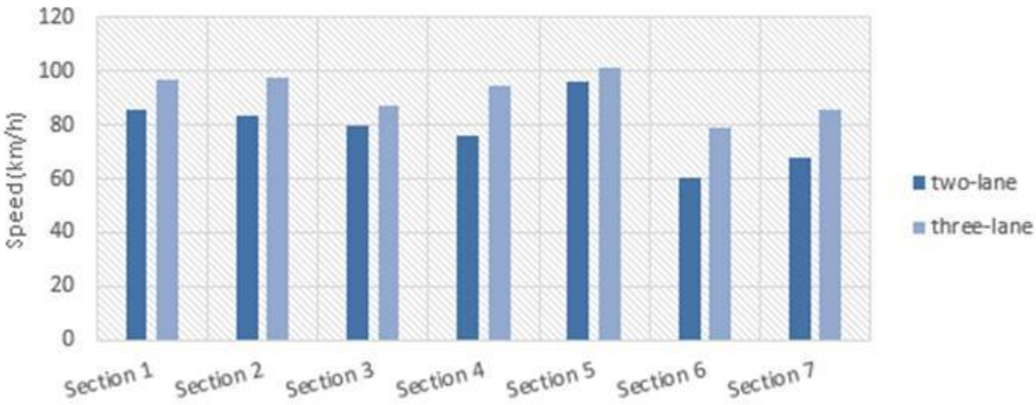



**Fig. 12.**
*Section passenger car speed comparison*
*Source: (Authors)*

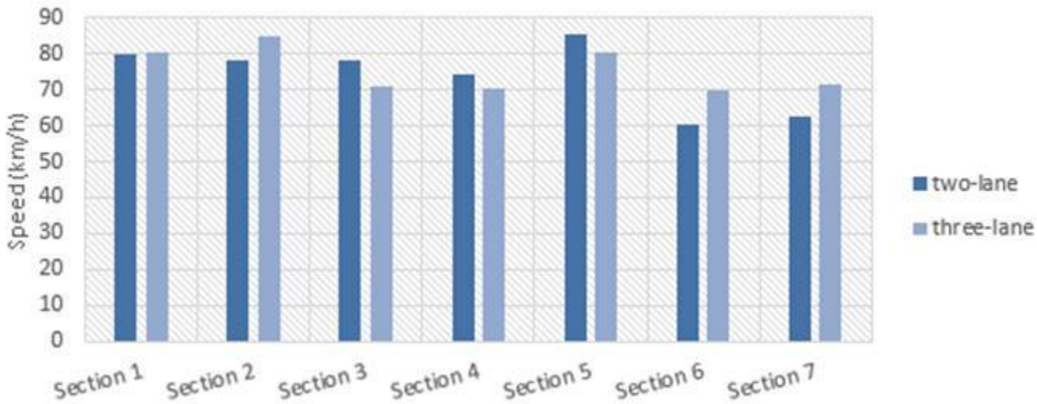



**Fig. 13.**
*Section truck speed comparison*
*Source: (Authors)*

Situation for trucks is different. Only on five three-lane segments truck speed was higher than on the two-lane. The reason for these differences is probably higher longitudinal slope.

When compare number of overtaking (all vehicles), results show significant improvement, as shown in the following table (Table 2) and graph (Fig.14.). Minimal improvement in the number of overtaking was 26%, the maximum 900 %.

**Table 2**
*Number of overtaking for all vehicles*

| | Section 1 | | Section 2 | | Section 3 | | Section 4 | | Section 5 | | Section 6 | | Section 7 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Number of overtaking** | **Two-lane** | **Three-lane** | **Two-lane** | **Three-lane** | **Two-lane** | **Three-lane** | **Two-lane** | **Three-lane** | **Two-lane** | **Three-lane** | **Two-lane** | **Three-lane** | **Two-lane** | **Three-lane** |
| | 79 | 310 | 23 | 45 | 82 | 103 | 151 | 261 | 105 | 161 | 81 | 308 | 32 | 311 |

*Source: (Authors)*

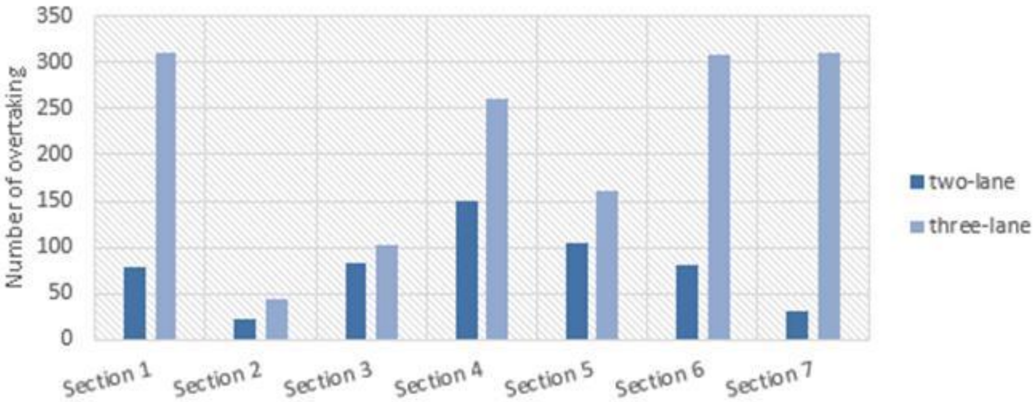



**Fig. 14.**
*Comparison the Number of overtaking*
*Source: (Authors)*

## 4. Conclusion

The article was intended to summarize the main conclusions of the accident analysis on 72 1st category road sections with increased number of lanes. Most of the sections were lanes for slow vehicles, or in 2+1 arrangement.
The main outcome of accident analysis is that with the increasing section length the rate of accidents on the lane reduction section decreases. Selection of the additional lane optimal length will be examined. It was also verified that these sections do not show a greater number of traffic accidents, as shown by the relative accident rate indicator. Another of the findings is that the share of trucks has no effect on the traffic accidents quantity.
Potentially critical point on these sections is the lane reduction section, where it is necessary to ensure the proper location with sufficient sighting distance and sufficient speed of slow vehicles. In the analysis, we focused on the type of lane reduction that can be either internal or external. Given the small sample of the internal reduction is not possible to clearly determine what the impact is on safety, but with the available data in sections with internal type lane reduction there are less traffic accidents.
In the examination of the accident was also reported large proportion of traffic accidents with animals, which can be attributed primarily to the environment in which these sections are proposed.
By evaluating traffic characteristics has been positively proven that although three-lane roads are mostly designed at upgrades, the traffic speed is higher than on two-lane roads with the equivalent traffic flow and intensity. An interesting fact is also that a truck speed increases at some sections which was likely caused by a gentler upgrade that did not affect a travel speed so much. As expected, a volume of passing vehicles also shows, that despite an upgrade the volume of passing vehicles on three-lane roads is significantly greater than on two-lane roads. So, roadways with 2+1 cross section can significantly increase level of service and do not even display worse statistics of accidents which is proven by accident rate.


## Acknowledgements

The research was supported by the project of Technology Agency of the Czech Republic TA02030548 – Update of the extra region road design elements. Publication was supported by the project OP VK CZ.1.07 / 2.3.00 / 20.0227 New Approaches to Transport Modelling.

**References**

Čepil, J.; Smělý, M.; Hornoch O.; Kosňovský M. 2012. *TA02030548 Odborná zpráva o postupu prací a dosažených výsledcích za rok 2012: Uspořádání 2+1.* Technology Agency of the Czech Republic. Czech Republic. 28 p.

Čepil, J.; Smělý, M.; Kosňovský M. 2013. *TA02030548 Odborná zpráva o postupu prací a dosažených výsledcích za rok 2013: Uspořádání 2+1.* Technology Agency of the Czech Republic. Czech Republic. 32 p.

Derr Ray B. 2003. *Application of European 2+1 Roadway Designs.* National Cooperative Highway Research Program. 31 p.

Kabeš, L.; Matuszková, R.; Kosňovský M. 2014. *TA02030548 Odborná zpráva o postupu prací a dosažených výsledcích za rok 2014: Průměrné úsekové rychlosti a předjíždění na komunikacích s uspořádáním 2+1.* Technology Agency of the Czech Republic. Czech Republic. 46 p.

Marvan, V. 2013. *Možnosti aplikace uspořádání 2+1 na silnici I. třídy v ČR.* Sudop Praha a.s. Czech Republic. 59 p.

Matuszková, R.; Radimský; Kosňovský M. 2014. *TA02030548 Odborná zpráva o postupu prací a dosažených výsledcích za rok 2014: Analýza nehodovosti na úsecích se zvětšeným počtem jízdních pruhů.* Technology Agency of the Czech Republic. Czech Republic. 22 p.

Ministerstvo dopravy, Centrum dopravního výzkumu. 2006. Jednotná dopravní vektorová mapa. Available from Internet: <http://www.jdvm.cz/>

VŠB-TU Ostrava. 2014. Dopravní inženýrství. Available from Internet: <http://projekt150.ha-vel.cz/node/97>

# INFORMATION AND COMMUNICATION INFRASTRUCTURE FOR THE ORGANISATION OF RAILWAY PASSENGER TRANSPORT

**Igor MATULIĆ[1], Mario Musa[2], Dragan Peraković [3]**
[1]*Railway track constructions Ltd., Međimurska 4, 10000 Zagreb, Croatia*
[2]*Croatian lottery Ltd, Ulica Grada Vukovara 72, 10000 Zagreb, Croatia*
[3]*Faculty of Transport and Traffic Engineering, Vukelićeva 4, 10000 Zagreb, Croatia*

**Abstract.** Information and communication infrastructure (ICI) enable transmission and application of all kinds of information and represent the most assertive generic technology of today. Railway on the other hand, is a very mature sector of traffic and transportation industry but has strategically importance for the European Union (EU) where in the large extent projects of railway network modernisation are taking place which primarily includes control systems and traffic management solutions based on ICI. The task of ICI for the organisation of railway passenger transportation is to increase security and reliability, compatibility among railway networks, stimulate and develop the awareness of ecological way of transportation, increase development and investment of new technologies in the railway network, which will correspond better to the challenges of the future economic development of Croatia. This paper examines the importance of ICI for control and traffic management, actual situation on Croatian railway network, and modern ICI solutions that are implemented and used across European Union.

**Keywords.** Railway Passenger Transportation, Information and Communication Infrastructure (ICI), Traffic Management System (TMS), European Railway Traffic Management System (ERTMS).

## 1. Introduction

Railway information and communication infrastructure (ICI) systems for passenger transport purposes can be divided into three (3) subsystems: a. subsystem provides management with timely information on the number, structure and flows of passengers as well as revenue respectively information required for making business-related decisions; b. subsystem provides passengers with essential information on travelling possibilities and quality journey; c. subsystem is used for the purposes of transport organization and safe traffic. This paper describes the current status of railway ICI systems for the purposes of passenger transport organization, further ICI system development with regard to particular examples and the current status of railway ICI systems in developed European countries, as well as the advantages provided for by those systems.

ICI represents a connectivity of microelectronics with computer technology and communications, and enables reception, storage, transmission and simple use of all information. Timely and correct information and knowledge based on such information are the basic resource of the companies today, particularly in railway transportation, where such information has an important impact on the security and reliability of the traffic flow. Development and presence of ICI depends on a number of administrative and regulatory, as well as economy and market conditions, where the railway companies, i.e. their management are playing the key role.

Successful companies have to be exceptionally adaptable, react strategically and develop a system that would facilitate flexible innovations in order to respond to faster and unpredictable changes of the environment. For the companies the introduction and implementation of new technologies implies:

- cost saving;
- increased safety;
- increased flexibility and interoperability;
- improvement of productivity;
- improvement of services rendered;
- engaging and keeping highly qualified employees;
- increased motivation and satisfaction of employees;
- increased quality of work.

When researching and collecting data, the inductive method was used, whereas desk research includes development and use of these systems in Europe and Croatia.
The classification of railway lines shown in Table 1 is based on clearly defined classification criteria in each of the three railway line groups. Through application of the evaluation according to those criteria, particular railway lines passed over to a higher or lower class, in comparison to the previous status. The dispatch and acceptance of passengers during 2014 is performed on 516 official places, i.e. railway stations and stops (Croatian Railway Passenger Transportation, 2013).

[3] Corresponding author: dragan.perakovic@fpz.hr

There are 980 km railway lines were electrified, i.e. (Croatian Railway Infrastructure, 2014):

- 977 km via alternate electric traction system 25 kV, 50 Hz
- 3 km via direct electric traction system 3 kV

**Table 1**
*Classification of railway lines in the Republic of Croatia*

| Railway line class | Network length (m) (in 2014) | | |
|---|---|---|---|
| | **Single-track railway line** | **Double-track railway line** | **Total length of railway lines** |
| Railway lines of importance for international traffic (30 railway lines) | 1.208,377 | 251,108 | 1.459,485 (56,0%) |
| Railway lines of importance for regional traffic (8 railway lines) | 626,373 | - | 626,373 (24,1%) |
| Railway lines of importance for local traffic (16 railway lines) | 518,474 | - | 518,474 (19,9%) |
| Total HZ (54 railway lines) | 2.353,224 | 251,108 | 2.604,332 |

*Source: Government of the Republic of Croatia, 2014*

**2. Information and communication infrastructure for the organization of transportation**

Train transportation, including signalization, regulation, acceptance and dispatching of trains, communication regarding train transportation on railway line network is operated by signalling and safety units and telecommunication devices. In view of traffic regulation on the network of Croatian Railways, except for permissions and notices of train departures, an announcement is to be given to the neighbouring station for each train departing from the station as well as to all official and workplaces between them.

**2.1. Signalling and Safety Systems**

The signal signs are signalized and given through signals. By means of those signals the railway staffs communicates to each other in a quick and reliable way with regard to train running, shunting, forbidden and permitted running across a certain site, track status and the need to reduce speed (Malinović, S., 2004). The rail lines of international importance are secured by relay signalling and safety devices of different manufacturers. The devices are between 25 and 40 years old, whereas the signalling device at Zagreb Main railway station (Zagreb GK) has been equipped with signalling and interlocking system in function since 1939. The long period of operation of the system, the impossibility of obtaining spare parts and some functions that are required today are reasons for the replacement of existing interlocking with a new electronic interlocking, installation of new digital telecommunication system as well as station truck junctions (Malinović, S., 2004). New equipment allows better train traffic flow throughout Zagreb node, and reduces the impact of bottlenecks in the railway transportation in X. and V.b Pan-European Corridor.

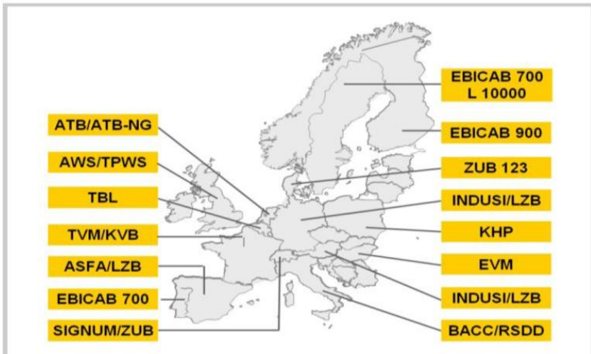



**Fig. 1.**
*Diversity of implemented railway national signalling and communication systems*
*Source: European Railway Agency*

The basic safety device for ATP (Automatic Train Protection) in the Republic of Croatia is an inductive automatic train stop device (AS) of INDUSI type (I 60), which is used in Germany, Austria, Slovenia, Bosnia and Herzegovina, Serbia, Macedonia, Montenegro and Romania. The system is partially applied in Croatia, i.e. on electrified tracks only. Traction vehicles (locomotives and motor trains) which are used for train traction, with passenger transportation are equipped with AS device.

AS device is used for the purposes of monitoring train movement on the track. According to the mode of operation, it belongs to the group of devices for monitoring train movement in one point, the so-called point-bases system unlike uninterrupted systems, where transmission of information from the track to traction vehicle is carried out continuously along the whole line. The purpose of AS device is to increase safety in rail transport in cases where a train driver does not notice or does not take any adequate measures to reduce speed ahead of “Speed Limit” sign respectively measures to stop the train ahead of “Stop” sign.

**2.2. Communication Systems**

The communication between rail regulators who regulate train traffic and train regulators, and rail regulators and train drivers is carried out by means of telephone, teleprinter, radiophone and IT devices. Notifications concerning train movement control are given through devices for evidence-based communication. The communication concerning traffic management between rail regulator and train driver is done in Croatian.

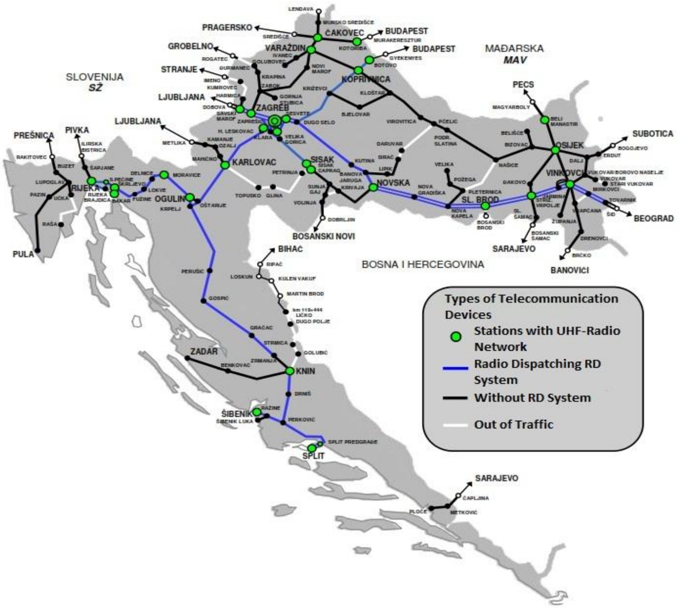



**Fig. 2.**
*Types of Telecommunication Devices – Overview*
*Source: Croatian Railway*

The status of telecommunications on particular track sections is similar to that of signalling devices. The built-in cables are between 25 and 70 years old. The telecommunication devices on stations are between 18 and 40 years old on average. A modernization respectively conversion from analogue to digital platform took place on most major routes and stations. A radio and dispatcher device, which is between 5 and 35 years old, operates along corridor tracks. The railway set up telecommunication cables and its own telephone line system along the entire network (the so-called ŽAT - Railway automatic), which is actually a closed communication system for the purpose of communication between railway employees.

The implementation of digital ŽAT switchboards as well as their cross-linkage with a new digital ŽAT network was carried out in almost all stations. Hubs (stations) are interconnected with the existing digital ŽAT network, as well as via ISDN PRA links (Integrated Services Digital Network Primary Rate Access) for ŽAT switchboard link with fixed public telephone network. The ŽAT switchboard system was set up with telephony switch digital business communication system and is fully IP (Internet Protocol) eligible communication system with up-to-date mobile functions. The users are free to choose between one or more system terminals; IP, digital or analogue fixed telephone, DECT (Digital Enhanced Cordless Telecommunications) or WLAN (Wireless Local Area Network) wireless phone or even mobile GSM (Global System for Mobile Communications) telephone. Each type of public terminal can entirely be integrated into the system. Due to its distributed architecture, it can support from 50 to 50,000 users with more than 500 users per independent unit in one system. Additional systems can be interconnected via IP network or traditional digital links enabling virtually unlimited standards in big networks.

Local Area Networks (LAN) - a structured cabling for local IT networks was performed in all designated railway stations. Likewise, an active equipment was implemented. Thus those LANs were put in local circulation.

SDH (Synchronous Digital Hierarchy) transmission infrastructure based on optical backbone as such was implemented on the Corridor X. of Corridors V.b and V.c, as well as partially on other corridors (all major railway stations). The implementation is still underway. Along with SDH transmission system upgrade, a telecommunications network management system upgrade to TNMS Core version is envisaged, whereupon hardware and application system upgrade to the level is planned which shall satisfy the network expansion on the corridors, as well as SDH network upgrade on any part thereof in future.

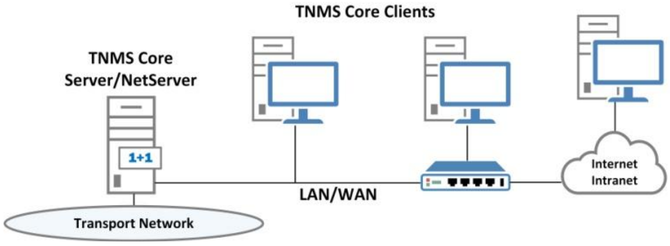



**Fig. 3.**
*Scheme of upgraded Telecommunications Network Management System TNMS Core*
*Source: Authors*

The SDH backbone set-up was a prerequisite for ŽAT switchboards linkage via optical infrastructure. The coupling of SDH device and ŽAT switchboard is carried out via E1 interface (2 Mbit/s, G.703). IP Network backbone, which is designed for transmission of data traffic, is a part of modernization project in terms of telecommunication and IT equipment on railway corridors. The solution is based on IP/Ethernet network points, which enables the construction of transport IP infrastructure for central network video surveillance system, remote control of IT screens and displays, as well as other applications which set high requirements on transmission capacity and quality of IP infrastructure. As a basic network for data transmission, IP network, based on Gigabit Ethernet and 10G Ethernet transport technology was implemented. The network is built by connecting Gigabit Ethernet switch through optical transmission application. Initial backbone capacity is 10Gbps for connecting major railway stations, whereas a network of capacity 1 Gbps is built between those hubs, which connects minor railway stations. Through creating such zones, a protection is provided in case of device failure or junction between two locations for the purpose of avoiding the interruption of communication between locations and undisturbed operation of applications, even those operating in real-time, used (or to be used) by Croatian Railways.

### 2.3. Optical network monitoring system

Within the modernization project, optical network monitoring system (ONMS), i.e. optical cables was installed. ONMS represents a system for continuous (24/7/365) monitoring of optical fibres. By applying the aforementioned system, Croatian Railways gained enormous safety in terms of all communication services via optical cables. In addition, the system can be expanded by other diverse application in future by using optical effects on cable. One of such applications is entry remote control, transit control regarding forbidden zones, theft protection, aerial cable sag monitoring for ice or heat, etc.

Likewise, significant application with increased safety is maintaining the agreed service level agreement (SLA) between users and quality of service (QoS) if a portion of network capacity is being hired. Through system, precise information on the status of a portion of hired network can be obtained more easily and quickly. Furthermore, the outside user is enabled to access directly via Internet solely to a portion of the network used by him. The system is very susceptible to all changes within the network and creates data base. Thus a defect, proactively based on tendency within the network, can be predicted beforehand and prevented before it actually occurred. ONMS facilitates the efficient property management because it guarantees that the optical cable is prepared for usage as expected of it and that the level of service is as planned. The system can create a link with other systems in monitoring centre, thus providing a high level of quality and safety, as expected in networks used in terms of railways and communications.

The backbone of the system is the central server with Oracle data base, where ONMS application is installed which collects and processes information from distant test units (OUT – Optical Test Unit) established on strategic points within the network. The users or clients are linked with the system via IP network, as well as commuted telephone network POTS (Plain old telephone service) or ISDN (Integrated Services Digital Network). In case of communication failure, test results on distant test units are stored in local memory and sent to central server, as soon as the communication is being established. The system is modular and scalable. Thus it can grow along with the network.

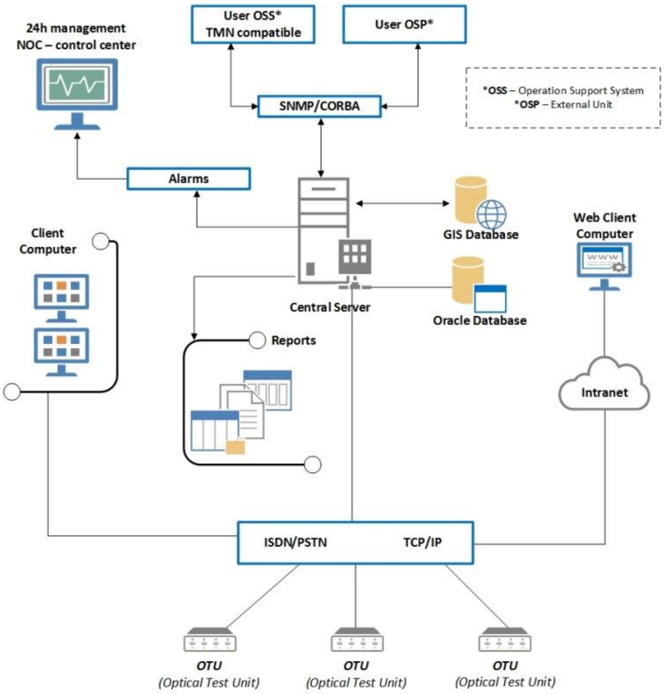



**Fig. 4.**
*ONMS – system architecture*
*Source: Authors*

### 3. Traffic Management System

The traffic of reverse and successive trains is regulated by rail regulators from occupied railway stations, by granting permission respectively consent and giving notice of train departure. On tracks where rail regulators have a visual control over the status of railway station spatial sections, permissions (consents) and notices of departure are given via signalling and safety devices operated by rail regulators, whereas on tracks where there is no visual control over the status of railway station spatial sections, the rail regulators give permissions and notices of departure via telecommunication devices in proven and tested way. On track portions, where the traffic of successive trains is regulated in spatial sections of notice of departure, notices of departure are given, except by rail regulators, by persons in charge of announcing notice of train departure as well. Exempt from the previous section, the traffic of reverse and successive trains on Vinkovci – Tovarnik route section, track M104 Novska - Tovarnik – state border, which is equipped with remote control devices, is regulated by remote control dispatcher from remote control centre located in Vinkovci station. The basic condition for regulating train traffic is that one train only can be situated in one spatial section, on the same track and at the same time. Except for permission and notice of departure, the announcement is to be given to the neighbouring station as well as all official places and workplaces between them for each train departing the station. In particular cases, it is required that the running of the train be pre-announced.

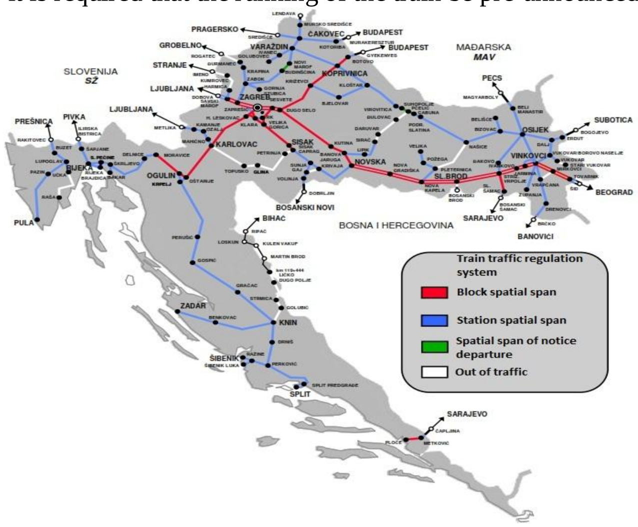



**Fig. 5.**
*Train traffic regulation system – Overview*
*Source: Croatian Railway*

Succession of trains in spatial span is regulated in three ways:

- Station spatial section – when the succession of trains is regulated by two neighbouring stations in a station span (MO)
- Spatial section of notice of departure – when the succession of trains is regulated by two neighbouring notice of departure or station and neighbouring notice of departure in a span of notice of departure (OP)
- Block spatial section – when traffic of succession trains is regulated by automatic setting of spatial signals into the position of permitted or forbidden running of the train using Automatic Track Protection (ATP).

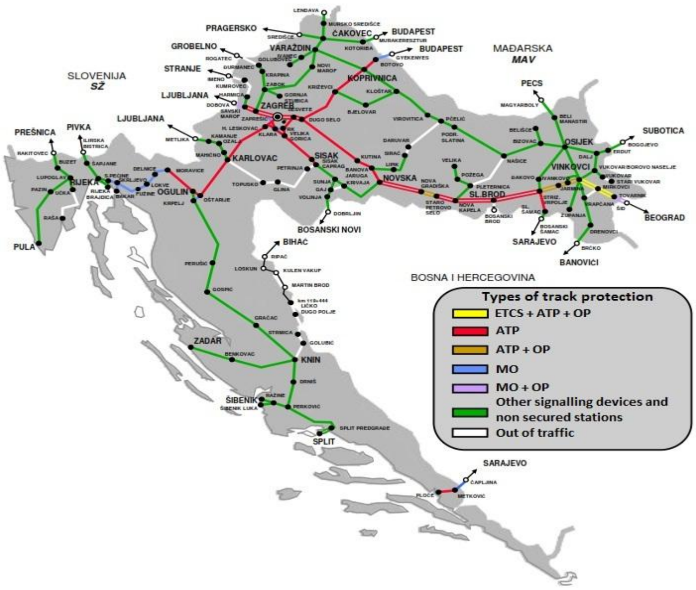



**Fig. 6.**
*Types of track protection – Overview*
*Source: Croatian Railway*

The tracks, where traffic in both directions is performed, are equipped with signalling and safety devices, which facilitate requesting and granting consent in both directions for each track and the control of track availability, as well as other signalling and safety devices which provide for safe train traffic in both directions and telecommunication links required for safe realization of such traffic. The following sections of double-track railway lines are equipped with devices, which facilitate the realization of traffic in both directions on both tracks:

- M101 DG (State Border) – S. Marof – Zagreb GK
  track section Savski Marof – DG - (Dobova) SŽ
  track section Zagreb ZK (Zagreb West station) – Zagreb GK
- M102 Zagreb Gk – Dugo Selo
- M104 Novska –Tovarnik – DG
  track section Nova Gradiška - Staro Petrovo Selo and
  track section Strizivojna Vrpolje – Tovarnik

At present, there is no automatic traffic management system available on rail network. Train traffic, including signalization, regulation, acceptance and dispatch of trains, communication regarding train traffic on rail network is regulated by signalling and safety devices and means of telecommunication.

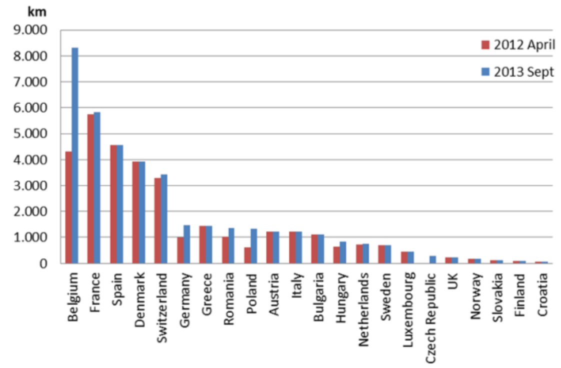
**4. ERTMS – common signalling and communication system**

Railways act in a competitive environment and have to provide operating security, international interoperability, increase of transport capacities as well as reduction of travel times and reduction of equipment life cycle costs. Nowadays the European railways have to deal with at least six (6) different rail electrification types and approx. 20 different traffic control and management systems. Each of them is extremely expensive in terms of maintenance and management. It occupies space on command board of the locomotive – it is complex and presents an additional cost in terms of cross-border traffic. This limits the competition and hinders the competitiveness of the European railway sector vis-à-vis road traffic by creating technical obstacles to international journeys (Spajić, I., 2012).

At present, ERTMS (European Railway Traffic Management System) has two basic components:

- ETCS (European Train Control System) is an automatic rail management and control system, which shall substitute the existing national ATP systems
- GSM-R (Global System for Mobile Communications - Railway), radio system for providing voice and data communication between track and train, is based on GSM standards by applying separate frequencies for rail applications with specific features and advances functions.

The purpose of ERTMS is to replace various national rail traffic control and management systems in Europe. As a unique European rail traffic control and management system, ERTMS is designed for the purpose of gradually substituting the existing incompatible systems across Europe. This shall bring considerable advantages to rail sector, as well as boost international passenger transports. ERTMS is the rail traffic control and management system which brings considerable advantages in terms of maintenance cost saving, safety, reliability, punctuality and increase of traffic capacities. The purpose of ERTMS is to make rail sector, in comparison to road traffic, more competitive, which ultimately shall provide considerable gains in terms of ecology.



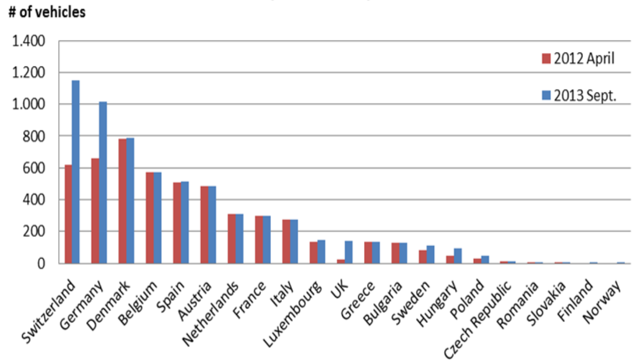




**Fig. 7.**
*ERTMS investment in Europe (trackside km and number of vehicles)*
*Source: UNIFE*

ETCS is a traffic control and management system defined by the UIC (Union Internationale des Chemins de fer). The system is elaborated in detail for the purpose of rail traffic liberalization in order to provide rail competitiveness by means of the same system in all European countries. The system shall provide cost reduction as a result of economy of scale, reduction of costs of development as well as quality cross-border services as a result of broad compatibility between on-board control and command equipment and corresponding railway track fittings. ETCS provides a unique signalization system and paves the way for boosting cross-border operations for high-speed trains and regular trains for passenger transportation trains (European Commission, 2012). From ETCS, the track sends data to the train, thus enabling that its highest speed limit is continuously calculated. On routes, where there is a signalization available by the track (lights and traffic signs make it possible for the driver to know the speed limit), this information can be sent by means of standard signalling equipment (Eurobalise), situated by the track. This is known as ETCS - level 1. For ETCS - level 2, the information can also be forwarded via radio (GSM-R). In this particular case, it is not necessary to keep track signals. This provides for considerable cost savings in terms of investments and maintenance. The position of trains at this level is still being detected by means of track signals. Ultimately, for ETCS - level 3, the train alone sends its current position, which enables the optimization of railway line capacities and additional track equipment reduction. At all levels, train-computer (the so-called Eurocab) compares the speed of the train with maximum speed and automatically reduces the speed of the train if the speed has been exceeded. ETCS system - level 1 is built in on refurbished Vinkovci – Tovarnik and Okucani – Novska track sections. However, it is not used because not a single traction vehicle of Croatian Railways has any equipment for ETCS.

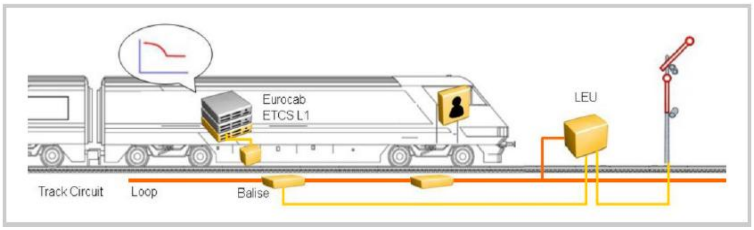



**Fig. 8.**
*ETC system - level 1*
*Source: UIC, 2008*

This system is useful if the entire section is equipped with Eurobalise, traction vehicles which run on that section by means of locomotive ETCS devices. Otherwise, „dual mode“ traction vehicle equipment is indispensable for conventional ATP and ETCS, which considerably enhances the price of the vehicle performance alone. Deployment of the speed control component (ETCS) is slower. However, 6,000 kilometres are already equipped or in the process of being equipped. Projects are underway in almost all European countries (ERTMS Net, 2013).

GSM-R is a part of ERTMS standard and conveys information on signalization directly to train signalization unit, providing for higher speed of train and traffic density at high safety level. Technical specifications and standards are drawn up with UIC within EIRENE project (European Integrated Railway Radio Enhanced Network). In order to achieve interoperability across Europe by applying one communication platform, GSM-R standard combines all key functions and experiences from 35 analogue systems used across Europe. GSM-R has not been implemented anywhere on the network of Croatian Railways yet. In Europe, GSM-R can now be deployed very rapidly, and according to data from the UIC, the GSM-R deployment in Europe in 2012 was as follows: out of the 154 284 km of track planned to be equipped with GSM-R in Europe, 85 332 km of track are with GSM-R installed (55.30% of the planned network) and 70 211 km of track are in operation (45.5% of the planned network), 40130 Cab Radios and 1900 ETCS Data Only Radios are activated (ERTMS Net, 2013).

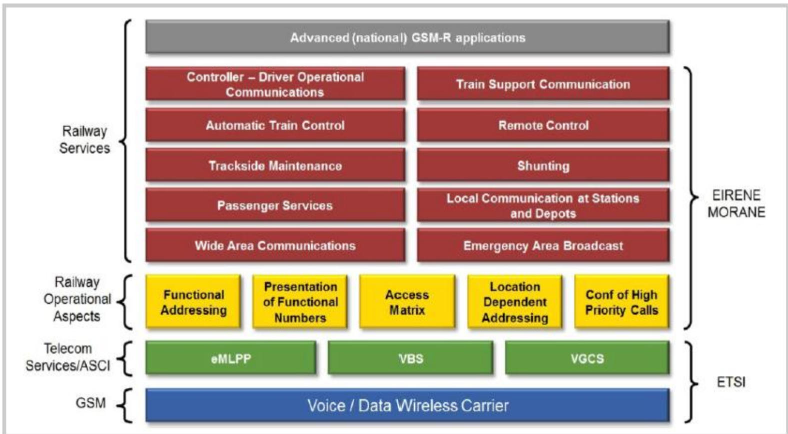



**Fig. 9.**
*Scope of services to be provided by GSM-R network*
*Source: ERTMS Net, 2013*

GSM-R provides a safe platform for voice and data communication between operating staff of railway companies, including train driver, link centre, members of switch staff, railway station and engineers and train controllers. The system provides advanced possibilities such as group calls, voice broadcasting, location-based linkage as well as emergency calls, which considerably improves communication, co-operation and management safety via operating staff members. Through GSM-R service standardization, the operators shall be able to considerably reduce management and maintenance costs, whereas travel times should be reduced, particularly in terms of cross-border traffic.

## 5. Conclusion

Railway companies that are eager to be successful have to sell their tailor-made services to potential customers. Likewise, they have to provide high-level quality service at eligible price. The fact that passengers are provided with information on services provided by a certain railway company affects both the image and the performance of the company. First of all, the modernization of signalling and safety systems, including the modernization of communication systems and network is a prerequisite. This is followed by the implementation of traffic control and management system in accordance with ERTMS standards. All of the above should result in higher operating safety, international operability, increasing of transport capacities, punctuality and reliability, reduction of management and maintenance costs as well as reduced travel times. Below ICI projects are shown, which are to be implemented for the purpose of boosting business operation of Croatian Railways, as well as integration into a unique European rail system to:

- exploit up-to-date infrastructure where available – network backbone can be expanded in future by other various applications by using optical effects on a cable. One of such applications is entry remote control, transit control regarding forbidden zones, theft protection, aerial cable sag control for ice or heat, etc.,
- upgrade TNMS to the level which shall satisfy the expansion of network on corridors, as well as to upgrade SDH network on any part of it in future,
- hire a portion of network capacity along with maintainance of the agreed SLA between users and QoS,
- implement GSM-R on tracks of importance for international and regional traffics in order to provide a broad range of services provided for by this technology. Through standardization of GSM-R services, the operators

shall be able to considerably reduce management and maintenance costs. Travel times should be reduced, particularly in terms of cross-border traffic and
- to implement automatic traffic management system in compliance with ERTMS standards – ETCS 1 on tracks of importance for regional traffic and ETCS 2 for international traffic.

Analogue with the aforementioned, traction vehicles should also be equipped with ECTS devices and the entire section with Eurobalise. Due to the fact that it is impossible to implement this on the entire network for economic and financial reasons, „dual mode“ equipment for traction vehicles is indispensable for regular ATP i ETCS equipment.

The purpose of all of this is to boost passenger traffic by rail and to make it more competitive, which ultimately shall provide considerable gains in terms of ecology, as well as social and economic gains. Since Croatian Railways are of strategic importance to the Republic of Croatia and since rail transport is one of strategic interests of the EU, by providing expert and quality preparation of technical, financial and operating feasibility studies, it is possible to implement these projects with the support of the Government of the Republic of Croatia as well as EU funds.

**6. References**

Akillilar, M. and Koppe, C. 2008. Increase of effectiveness and efficiency in the interlocking system ESTW L90 5, Signal und Draht, vol.100, no.4, pp.26-30.

Croatian Railway Passenger Transportation. 2013. Business plan for the year 2014. HŽ Putnički prijevoz, Zagreb

Croatian Railway Infrastructure. 2014. Network report 2015. HŽ Infrastruktura, Zagreb. 24 p.

ERTMS Net,2013. ERTMS. Available from Internet <http://www.ertms.net/ertms/>

European Commission. 2012. Rail: ERTMS - European Rail Traffic Management System. Available from Internet: <http://ec.europa.eu/transport/modes/rail/interoperability/ertms/>.

Government of the Republic of Croatia. 2014. Decision on Railway Line Classification (NN (OG) 3/14). Official Gazette of the Republic of Croatia, Zagreb

Letia, T. Hulea, M. and Miron, R. 2006. Distributed scheduling for real-time railway traffic control. International Multiconference on Computer Science and Information Technology, pp.679-685.

Malinović, S. 2004. Prijevoz putnika u željezničkom prometu (Passenger transportation in railway traffic), ŽTŠ, Zagreb.

Spajić, I. 2012. Replacement of the signalling and interlocking system in Zagreb Main railway stations. In Proceedings of the 32nd Korema Conference on Transportation Systems – Automation in transportation, 232-235 p.

UIC. 2008. ETCS Implementation Handbook. UIC, Paris

# DYNAMIC FRICTION IN THE BRAKING, TIRE – ROAD CONTACT

**Ivan Ivković[1], Željko Janjoš[2], Gradimir Danon[3], Srećko Žeželj[4]**
[1,4] *University of Belgrade, Faculty of Transport and Traffic Engineering, Belgrade, Serbia*
[2] *SR Center City Net Ltd., Belgrade, Serbia*
[3] *University of Belgrade, Faculty of Forestry, Belgrade, Serbia*

**Abstract:** The differential equation of the tire-road contact sliding speed is presented in the paper. It is based on the equations of: normal pressure distribution in contact, available friction and tangential stress within the tire-road contact. Solving this equation makes it possible to calculate the sliding speed and available friction within the contact path. Some numerical examples are presented in the paper. Indoor testing of radial tires was performed and the results of the measurement were used for these calculations. The formulated equation may be incorporated in the friction model of tires, for the purpose of making a more exact prediction about the friction behaviour of the braking tire on the roadway of known characteristics.

**Keywords:** tire-road contact, adhesion, braking, tire sliding.

## Introduction

The forces required to support and control a vehicle all appear in the tire contact area. In contact, an element of a tire tread is acted upon by a force vector which can be expressed in terms of two components. One is perpendicular to the contact surface and one tangential to the contact surface. The transfer of forces in the tire and road contact is limited by available friction in contact, and it is inevitably followed by a slip. The slip phenomenon is manifested as the vector difference of the translator tire speed and longitudinal speed of the tread part in contact. The compensation of this difference is possible by tread deformation or by sliding of the tread part in the plane of the contact path.

The contact path of a braking tire is separated in two regions. In the first, named “adherence” region, the dominant is a deformation component of the slip and there is no move between the tire and road surface. Available friction ($\mu_o$), in this region, is nearly constant, because the sliding of the tread pattern against road surface is negligible. Sliding begins when the value of shear stress, which is proportional to the strain of a small piece of tire tread in contact ($\tau_x=K_x \cdot S_{slip}$), reaches available unit friction in contact ($\mu_o \cdot p$). From this moment on, sliding begins between a piece of the tire tread and surface. Sliding speed value duly depends on exterior efforts and contact pressure distribution, and it has a great effect on available friction. By formulating the differential equation of sliding speed and solving it, available and obtained friction within the contact path may be calculated. The equation of sliding speed, available and obtained friction, which is presented, may be incorporated as useful in the tire friction models.

This paper outlines the equation of the sliding speed in the tire and road contact and calculations of the sliding speed distribution that are made for the straight line motion of the tire under braking effort. The equation for the sliding speed in contact is formulated on the basis of corresponding assumptions of simplification, as the outcome of the investigation of the results (Danon, 1988; Clark, 1981; Pillai and Fielding-Russell, 1986; Janicijević and Danon, 1987; Danon and Markun, n.p.; Zeng-Xin et al., 2001; Pottinger, 2005; Burke and Olatunbosun, 1997; Grigolyuk, 2004). The equations for calculating tire dimensions in the function of air pressure, radial deformation, pressure in contact and dimensions of contact are stated in the paper.

## Dimensions of an inflated pneumatic tire

On the basis of research results presented in (Clark, 1981; Pillai and Fielding-Russell, 1986), it is assumed that the circumference, diameter and width of an unloaded tire are not significantly modified regardless of compressed air it is filled with. The assumption is tested by means of corresponding laboratory measurements (Danon, 1988). The measurements are carried out on four 145 R 13 Best tires and mean results are shown in Table 1.

**Table 1**
*145 R 13 Best radial tire dimensions measurements (Danon, 1988)**

| Tire dimensions | Unit of Measure | Tire pressure [bar] | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | 0.5 | 1.3 | 1.5 | 1.7 | 1.9 | 2.1 | 2.3 |
| Tire free diameter | mm | 565.60 | 565.45 | 565.52 | 565.52 | 565.52 | 565.60 | 565.66 |
| Total tire width | mm | 144.3 | 145.5 | 145.5 | 145.6 | 145.6 | 145.7 | 145.9 |

* Tire was fitted around a 4Jx13 wheel rim

[1] Corresponding author: i.ivkovic@sf.bg.ac.rs

The results of the measurement have corroborated the stated assumption. A possible explanation for this kind of tire behaviour is belt, which all radial tires have and which is practically inelastic, i.e. its circumference, regardless of the internal air pressure variation, does not alter.

**Dimensions of tire contact**

The vertical load acting on the tire makes it deformed and changed in shape. The deformation and change in shape are mostly evident in contact with the road surface. The contact of a stationary wheel with a rigid road surface is most frequently of a rounded rectangular shape (Ivanov, 2010). The radial deformation, area and dimensions of contact of the stationary wheel and the surface greatly depend on air pressure in the tire and vertical load. Here the following rule applies-the larger the load on the tire, the larger the contact patch, and the larger the inflation pressure, the smaller the contact patch (Taylor et al., 2010).

In Fig. 1, the results of the measurement of change in radial deformation of the tire are given in the function of the vertical load for a 145R13 Best tire (Danon, 1988). Different vertical forces acted on the examined tires and in addition radial deformation, area and contact dimensions were measured.

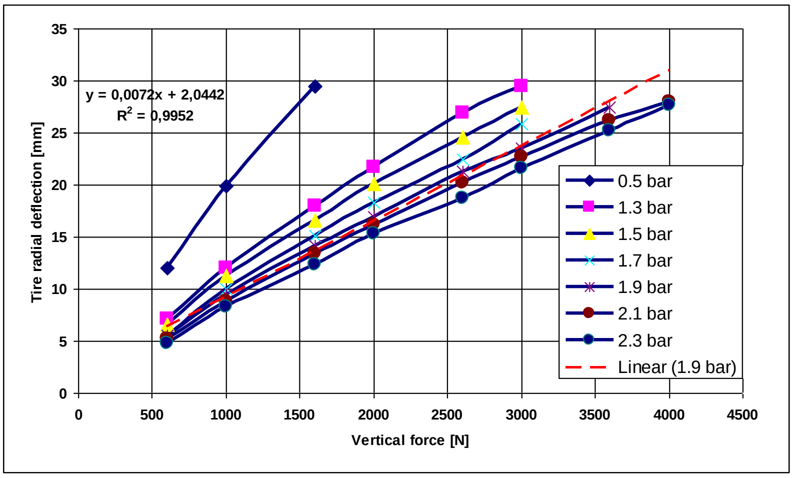



**Fig. 1.**
*Influence of air pressure and normal load on radial deformation of the 145R13 Best tire*

It is established that the dependence between vertical force and radial deformation, for all examined pressure pumping, is approximately linear. In the mentioned measurements, a linear trend line illustrates that the radial deformation dR= $f(Fz)_{p=1,9\text{ bar}}$ has consistently risen over the observed load range. Notice that the $R^2$ value is 0.9952 (see Fig. 1), which is a good fit of the line to the data.

The radial tire deflection dR consists of two components. One of them is tread deflection and the other is carcass deflection, Eq. (1):

$$dR = dR_t + dR_c \tag{1}$$

The value of the tread deflection ($dR_t$) depends on the vertical load (Z) and total tire deflection (dR). The tire changes its shape and volume on deformation. Deformation energy is partially consumed by the deformation of the tire structure and partially by interior air compressing. The carcass deformation ($dR_c$) is proportional to the vertical load and tire inflation pressure.

In Fig. 2, the schematic representation of the tire road contact is given, where the tire tread, carcass and air pressure are symbolically represented by springs.

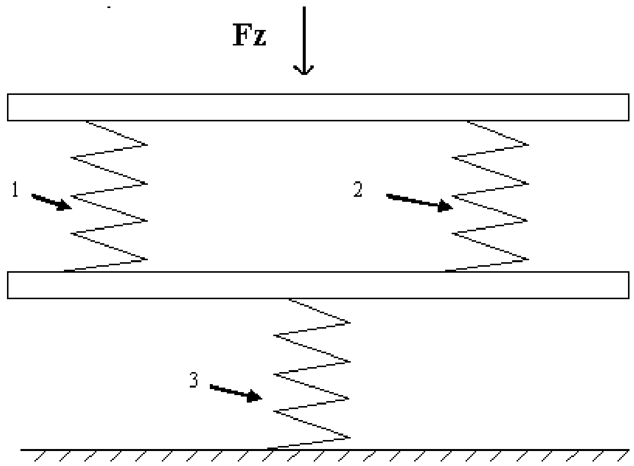



**Fig. 2.**
*Scheme of the tire tread and tire carcass behaviour undergoing a vertical force; Springs of equivalent stiffness: 1-tire carcass; 2-tire air pressure; 3-tire tread.*

If it is assumed that we are really dealing with springs, it is possible, with an appropriate simplification, instead of a bunch of curves shown in Fig. 1, to formulate the universal dependence of the vertical load and radial deformation for any tire, Eq. (2) (Janićijević and Danon, 1987; Danon and Markun, n.p., Grečenko, 1995; Knoroz, 1963; Biderman, 1963):

$$dR = \frac{C_2 \times F_Z}{2 \times (p_a + p_o)} + \sqrt{\left(\frac{C_2 \times F_Z}{2 \times (p_a + p_o)}\right)^2 + C_1 \times F_Z}, \quad (2)$$

where: $F_Z$ is the vertical load; $p_a$ is the tire air pressure; $p_o$ is the measure of structural rigidity of the tire carcass;

According to (Danon and Markun, n.p) $p_o$ is usually between 0 and 0.3 bar. Higher values are characteristic of truck and off-road tires. For radial tires for passenger cars $p_o$ can be neglected; the parameter $C_1$ depends on the tire radius, tread section radius, applied material characteristics, net/gross tread pattern ratio and its thickness; the parameter $C_2$ depends on the tire radius, tread section radius and tire construction; the $C_1$ and $C_2$ parameters are determined experimentally.

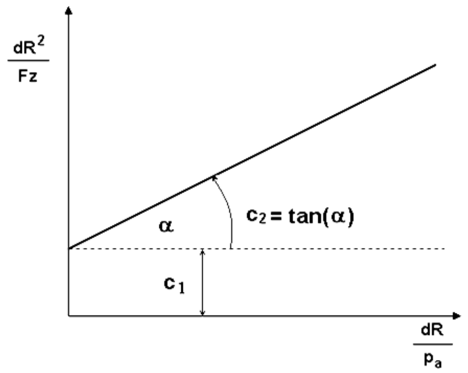



**Fig. 3.**
*Universal characteristic of tires*

The calculated constants C1 and C2 for examined tires were:
Tire 145R13 Best C1= 0.027, C2= 0.011, Tire 155R13 Best C1=.0.0495, C2= 0.0119.

At load variation, apart from radial deformation, the length of contact and the shape of contact are partially changed while the width of contact ($w_c$) remains approximately constant. The results of the measurement for 145 R 13 Best tires are shown in Table 2.

**Table 2**
*Effect of air pressure and vertical load on the contact area and size of 145 R 13 Best tire deflection*

| Pressure | Vertical load | Tire dimensions [mm] | | |
|---|---|---|---|---|
| bar | N | Width ($w_c$) | Length ($l_c$) | Deflection (dR) |
| 1.3 | 2,400 | 92 | 166 | 24,74 |
| 1.5 | 2,700 | 92 | 167 | 25,52 |
| 1.7 | 3,000 | 92 | 168 | 26,21 |
| 1.9 | 3,300 | 92 | 170 | 25,8 |
| 2.1 | 3,600 | 93 | 175 | 26,32 |
| 2.3 | 3,900 | 94 | 174 | 27,95 |

The same data but for the 155 R 13 tire are given in Table 3.

**Table 3**
*Effect of air pressure and vertical load on the dimensions of contact and size of 155 R 13 Best tire deflection*

| Pressure | Vertical load | Tire dimensions [mm] | | |
|---|---|---|---|---|
| bar | N | Width ($w_c$) | Length ($l_c$) | Deflection (dR) |
| 1.4 | 3,300 | 100 | 207 | 32.8 |
| 1.7 | 3,000 | 99 | 171 | 26.6 |
| 2.1 | 3,600 | 100 | 175 | 27.0 |
| 2.4 | 3,300 | 97 | 150 | 23.5 |
| 1.9 | 2,300 | 96 | 144 | 20.1 |
| 1.9 | 2,800 | 99 | 150 | 23.5 |
| 1.9 | 3,300 | 98 | 170 | 26.5 |
| 1.9 | 3,800 | 99 | 184 | 30.5 |
| 1.9 | 4,300 | 102 | 207 | 33.2 |
| 1.9 | 4,800 | 102 | 210 | 36.5 |

The data shown in Table 3 indicate that the variation in the contact length for a large range of pressures and load is marginal. This can be seen in Fig. 4 where the variation in the contact dimensions of the 155R13 Best radial tire (inflated at a pressure of 1.9 bar and subjected to different vertical loads) is given.

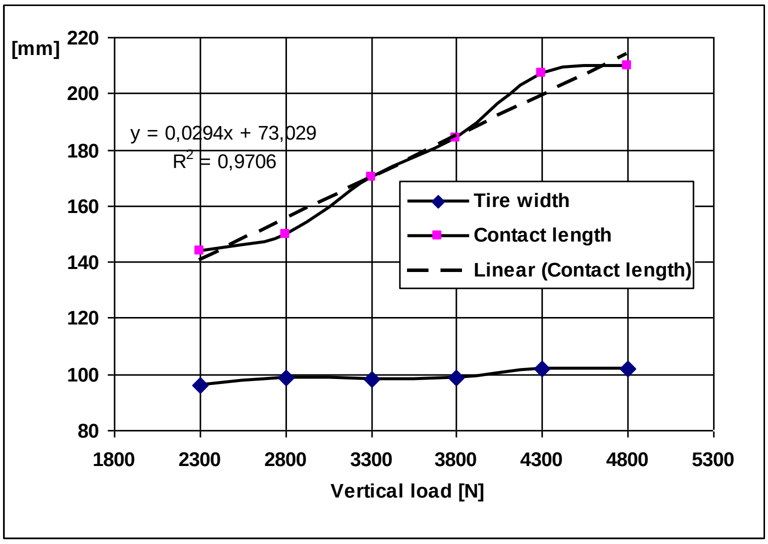



**Fig. 4.**
*Influence of vertical load on 155 R 13 Best tire footprint dimensions tire pressure was 1.9 bar*

The variation in the contact length is approximately linear and directly proportional to the variation in the vertical load. This holds if air pressure does not change. The high $R^2$ value which amounts to 0.9706 confirms it.

The equation which suits the data (the contact length and the respective tire deflection) is Eq. (3):

$$l_c = 2 \cdot \sqrt{(k \cdot R_o)^2 - (R_o - dR)^2} \quad (3)$$

The coefficient k depends on tire deflection. The increase of the radial tire deflection entails the decrease of the coefficient k. The linear dependence is presumed Eq. (4):

$$k = A - B \cdot dR \quad (4)$$

The coefficients of the Eq. (4) have been determined for all values given in Table 3 by means of the method of least squares Eq. (5) (Danon and Markun, n.p):

$$k = 0,9766 - 0,00078 \cdot dR \quad (5)$$

The measured and calculated values of the contact length as well as the determined difference are given abreast in Table 4.

**Table 4**
*Calculated versus measured static footprint lengths of the 155 R 13 Best tire*

| Pressure | Tire length [mm] | | The difference |
|---|---|---|---|
| bar | Measured | Calculated | % |
| 1.4 | 207 | 139 | 3.59% |
| 1.7 | 171 | 158 | -5.03% |
| 2.1 | 175 | 158 | -5.03% |
| 2.4 | 150 | 172 | -1.28% |
| 1.9 | 144 | 173 | -0.96% |
| 1.9 | 150 | 174 | 0.29% |
| 1.9 | 170 | 190 | -3.08% |
| 1.9 | 184 | 199 | 3.92% |
| 1.9 | 207 | 200 | 3.17% |
| 1.9 | 210 | 213 | -1.28% |

The determined differences are between -4% and +5%, which might be considered satisfactory. The influence of the vertical load on the variation in the contact width is irrelevant. According to the data set, shown in Table 3, it has been stated that the contact width increased by 6 mm in length change by 50 %. Therefore, it is safe to assume that the width of contact remains constant regardless of the level of vertical load.

During the wheel's rotation, the variation in radial deflection occurs. In order to determine the effect of the circumferential speed on radial deformation, the corresponding investigations on a tire test drum have been conducted (Danon, 1988). The tests for different inflation pressures and different vertical loads have been carried out. Table 5 shows the measured results for the 155R13 Best tire for air pressure to be equal to 1.9 bar and vertical load of 3,300 N. The values of static measurements are given in the column marked with V = 0.

**Table 5**
*Measurements of radial tire deflection of the 155R13 Best tire on test drum. Tire air pressure was 1.9 bar and vertical load 3,300 N*

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Drum speed | km/h | 0 | 40 | 60 | 80 | 100 | 120 | 140 |
| Tire deflection | mm | 29.5 | 29.5 | 28.5 | 28.5 | 28.5 | 28.5 | 27.5 |

The variation of radial deformation is very low around 2 mm within the speed range of up to 140 km/h. The tire deflection of the stationary tire (29.5 mm) mounted on the test drum is somewhat higher in relation to the result presented in Table 3 (26.5 mm). The difference is 3 mm and it is the result of the variation in the contact surface. The measurements presented in Table 3 are carried out on a flat surface, and the surface of the test drum is curved. It can be assumed that the variation in radial deformation could have been smaller if the dynamic investigations of tires had been carried out on a flat surface.

For calculating the dimensions of contact and speeds in contact it has been accepted that the statistic and dynamic deformation are equal ($dR \approx dR_d$) (Springer and Weisz, 1975).

**Speeds in tire contact**

Let us assume that a vertically loaded tire moves rectilinearly at a constant speed (V), i.e. it is rotated at an angular speed (ω). Let us also assume that the tire deformation is limited to the contact area and that it is very small, i.e. negligible outside it. Then we can say that the circumferential speed of a point on the surface of the tire tread not subjected to contact is likewise constant and equal to $V_c = \omega \cdot R_o$ and that it is approximately equal to the longitudinal speed of the wheel (V). The circumferential speed remains constant until the beginning of contact. From the beginning of contact the situation keeps changing. The longitudinal speed ($V_1$) in contact is not constant but varied, due to radial and tangential deformations of the tire tread, along the contact (Uljanov, 1982). The deformation speed ($V_2$) and the circumferential speed ($V_3$) are components of the speed in the tire and road contact, Eqs. (6), (7) and (8):

$$V_1 = \frac{\omega \cdot h}{\cos^2(\alpha)} \tag{6}$$

$$V_2 = \frac{\omega \cdot h \cdot \sin(\alpha)}{\cos^2(\alpha)} \tag{7}$$

$$V_3 = \frac{\omega \cdot h}{\cos(\alpha)} \tag{8}$$

where: h is the distance between the wheel axle and road surface of the rotating tire, Eq. (9):

$$h = R_o - dR_d \tag{9}$$

Where: $R_o$ is the free radius of the wheel and $dR_d$ is the tire radial deformation, and it is given in the equation stated in (Janjićijević and Danon, 1987):

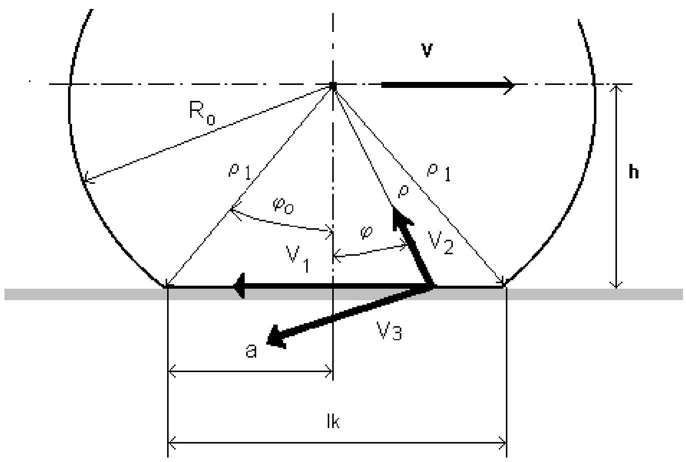



**Fig. 5.**
*Tire and road contact*

With a braking wheel the situation is somewhat different. The slipping speed of the tire and road contact, in the case of straight line braking, can be expressed as an ordinary difference of the translator speed of the tire (V) and longitudinal speed in contact ($V_1$) , Eq. (10):

$$V_{slip} = V - V_1 \tag{10}$$

Slipping speed can be divided in two components:

- deformation speed ($V_{def}$), which is dominant in the adhesion part of the tire and road contact and
- sliding speed ($V_{slide}$), which is characteristic of the sliding part of contact, Eq. (11):

$$V_{slip} = \lambda \cdot V = V_{def} + V_{slide}, \tag{11}$$

Where the slip coefficient is, Eq. (12)

$$\lambda = \frac{V - V_1}{V}, \tag{12}$$

Since there is a variation in speed ($V_1$) along the trajectory the slip coefficient ($\lambda$) is variable even when the longitudinal speed of the wheel is constant (V=const).
In the first part of contact (A-C) the slip path is approximately equal to the tangential deformation of the tread part, Eq. (13):

$$s_{slip} = \left\{ (a - x) - \frac{h}{1-\lambda_o} \cdot \left[ atn\left(\frac{a}{h}\right) - atn\left(\frac{x}{h}\right) \right] \right\}, \tag{13}$$

where: *a* is half of the contact length, *x* is a coordinate in the contact path and $\lambda_0$ is the slip coefficient in the centre of contact for *x*=0.

Due to this, the deformation speed can be expressed as the first derivative of ($\tau$) tangential stress (within contact path) per time, Eq. (14):

$$V_{def} = \frac{dS_{slip}}{dt} = \frac{1}{K_x} \cdot \frac{d\tau}{dt} \tag{14}$$

In the second part (C-E), after the friction limit has been reached, the tangential deformation of the tread part in contact changes only due to normal pressure and obtained friction. The difference between a growing slip path and decreasing tangential deformation is compensated by sliding of the tread part over pavement (Fig. 6).

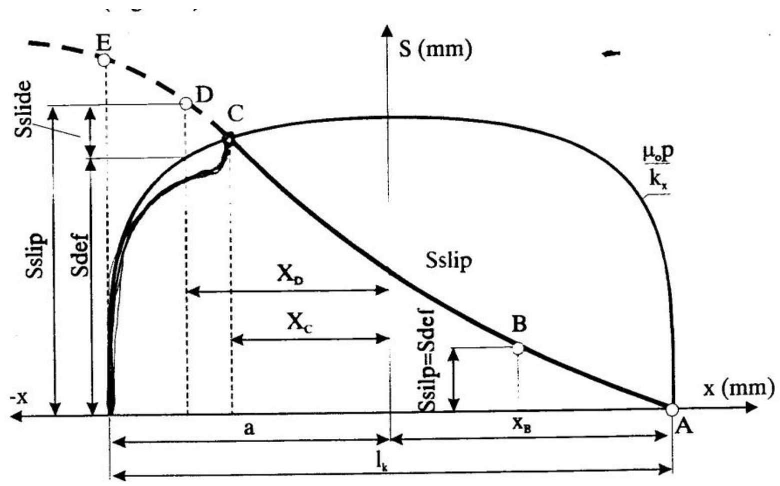



**Fig. 6.**
*Slip path in the tire-road contact*

Sliding speed can be expressed in the equation of available friction in the tire and road contact ($\mu$) (Moore, 1975). If the linear dependence between friction coefficient and sliding speed is assumed, the equation for sliding friction is, Eq. (15):

$$\mu = \mu_o - \beta \cdot V_{slide} \tag{15}$$

where: $\mu_0$ is the “static” coefficient of friction in contact, while $V_{slide}$= 0; $\beta$ is the coefficient of proportionality.

A linear function is assumed for two reasons:
It is much easier to solve final differential equation and
Linear function is not too far from reality, especially for the speeds among 10 and 100 km per hour.
The realized coefficient of friction (μ) varied along the contact path in the function of the value of sliding speed (refer to Eq. (15)). Therefore, we can say that obtained friction is equal to Eq. (16):

$$\mu = \frac{\tau}{p(x,y)} \tag{16}$$

where: $p_{(x,y)}$ is the normal contact pressure in the point with coordinates x and y. In our example, we have assumed that the pressure distribution along the contact length of tire is parabolic, and along the contact width uniform, i.e. the dependency applies Eq. (17):

$$p = p_{max} \cdot \left(1 - \frac{x^2}{a^2}\right) \tag{17}$$

Where the maximum value of pressure is calculated as, Eq. (18):

$$p_{max} = \frac{3 \cdot F_z}{2 \cdot l_c \cdot w_c} \tag{18}$$

If by increasing the vertical force ($F_z$), pressure reaches the limit value ($p_{lim}$), in some segment of contact, the form of distribution is changed into trapezoidal and the contact length is increased. With the tire that rolls, the pressure distribution in contact varies due to hysteresis and becomes asymmetric against vertical axis. Part of energy that the tire absorbs in the load area remains captured within the tire's structure and it turns into heat.

For the assumed friction-$V_{slide}$ relation Eq. (15) it is easy to solve the equation Eq. (16) as follows, and this equation is valid from x= $x_a$ to the end of contact, Eq. (19):

$$V_{slide} = V_{slip} \cdot \left[1 - e^{\left(\frac{-K_x \cdot t}{\beta \cdot p}\right)}\right] \tag{19}$$

In other cases, if direct solving is not possible, the Runge-Kutta method can be applied.

The calculated variations of slip and sliding speeds and the friction coefficient (μ) are shown in Fig. 7. The shown distributions of velocities within the tire and road contact have been calculated for the input values given in Table 6.

**Table 6**
*Input values for contact speed calculations*

| Input values | Symbol | Measure | Value |
|---|---|---|---|
| Inflation pressure | $p_v$ | bar | 2 |
| Vertical load | Z | N | 3000 |
| Radial deformation coefficients | $C_1$<br>$C_2$ | $m^2/N$<br>$m^{-1}$ | 0.0447<br>0.0116 |
| Slip coefficient | $\lambda_0$ | % | 15 |
| Longitudinal stiffness | $K_x$ | $N/mm^2$ | 400 |
| Translator speed | V | m/s | 40 |
| Static friction coefficient | $\mu_0$ | - | 1 |
| Friction correction coefficient | β | $s/m^2$ | 0.005 |

By looking at the Eq. (19) we can note that the magnitude of sliding speed depends primarily on the value of the slip speed and it depends on the slip coefficient ($\lambda$) and on the translation speed (V). Fig. 7 shows the variation of the friction coefficient ($\mu$), the slipping and sliding speed in the tire-road contact.

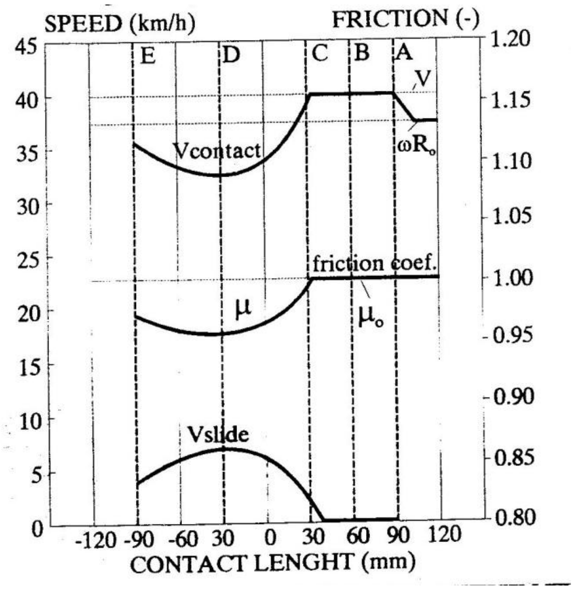



**Fig. 7.**
*Speed and friction distribution within the contact*

The form of the sliding speed is not usual and differs from those usually found in the literature (Clark, 1981; Ivanov, 2010). In the cited references, it is assumed that the longitudinal speed in contact is constant which makes the change of sliding speed in contact to increase monotonously. In our case, we have taken the speed variability ($V_1$) in contact into consideration. At the beginning of contact, the speed ($V_1$) is as follows, Eq. (20):

$$V_1 = \sqrt{\left(h^2 + a^2\right)} \cdot \omega = \rho_1 \cdot \omega \qquad \text{for } x = a \tag{20}$$

where: $\rho_1$ – the shortest distance from point (A) at the beginning of contact to the axis of rotation, usually taken to be equal to $R_o$.

After the beginning of contact this speed decreases and falls onto minimum in the middle of contact, Eq. (21):

$$V_1 = h \cdot \omega \qquad \text{for } x = 0 \tag{21}$$

After completing half of contact ($V_1$) the speed increases and at the end of contact it has nearly the same value as at the beginning of contact. The circumferential speed of the wheel at the beginning of contact is different from the longitudinal speed of the wheel (V). With the braking wheel, the circumferential speed is lower than the longitudinal speed of the wheel. The amount of variation depends on the slip coefficient ($\lambda$). From the beginning of contact (point A) the difference between these two speeds is compensated by the tangential deformation of the tire tread. Sliding begins at point (C) and it lasts until the end of contact.

## Conclusions

The presented procedure for the calculation of adherence is simple and has quite a few advantages compared to the procedures which have been applied so far.

The first advantage appears at the calculation of the tire radial deformation. Instead of a bunch of curves for the dependence of radial deformation and vertical load, the so-called carpet diagram, we only applied the Eq. (2) for which we only had to know the coefficients $C_1$ and $C_2$. The tests have shown that the deviations, for one dimension of the tire of the same construction, are not considerable (Pottinger, 2005; Taylor et al., 2010).

If the data on the inflation pressure and vertical load of the tire are available, it is possible to apply the presented formulas to accurately calculate radial deformation, footprint area and distance between the wheel axle and road surface. By applying these data, it is possible to calculate the rolling radius of a free wheel.

This paper assumes the parabolic distribution of normal pressure in contact. The formula can include some other distribution if it is more convenient for the actual conditions.

The longitudinal tangential stresses ($\tau$) acting in the plane of contact of the braking wheel were determined by a sum of stresses generated by the rolling of the free wheel ($\tau_0$) and additional stresses ($\tau_\beta$) obtained by the application of a braking torque. Using the Eqs. (16) and (17) it is possible to calculate the obtained force in contact and achieved adherence.

The formula for the calculation of achieved adherence includes the influence of sliding speed. The equation was formulated for this purpose, for the sliding speed in the tire and road contact. Sliding speed, shown in Fig. 5, includes the assumption that the speed in contact ($V_1$) is variable. The shown variation differs from those shown in (Clark, 1981; Ivanov, 2010) where sliding speed is presented as an exponential curve.

The friction in contact depends on the instantaneous sliding speed. Although we applied linear dependence here, this procedure does not exclude an application of some other dependence. The presented equations may be incorporated in the frictional model of the tire. The purpose is to improve the prediction of the friction behaviour of the braking tire. For the application of the presented procedure it is necessary to know a number of input data on the tire and exterior wheel load.

**Acknowledgements**

This paper is based on the project TR36027: "Software development and national database for strategic management and development of transportation means and infrastructure in road, rail, air and inland waterways transport using the European transport network models" which is supported by the Ministry of science and technological development of Republic of Serbia (2011-2014).

**References**

Biderman, V. 1963. *Avtomobilnije šini*. Moskva :GHI, 378 p.

Burke, A., Olatunbosun, O. 1997. Static Tyre/Road Interaction Modelling, *Meccanica* 32(5), 473-479..

Clark, S.(1981). *Mechanics of Pneumatic Tires*. USA: National Bureau of Standard, 931 p.

Danon, G. 1988. *Dynamic adhesion during straight line car braking*. Doctoral Thesis: Univeristy of Belgrade, Faculty of Mechanical Engineering, 232 p.

Danon, G. 1990. Matematički model elastičnog točka putničkog vozila, *u Zborniku VI Međunarodni simpozijum "Motorna vozila i motori 90"*, 7 p.

Grečenko, A. 1995. Tyre Footprint Area on Hard Ground Computed from Catalogue Values, *Journal of Terramechanics* 32(6),325-333.

Grigolyuk, E., Kulikov, G., Plotnikova, S. 2004. Contact Problem For a Pneumatic Tire Interacting with a Rigid Foundation", *Mechanics of Composite Materials*, 40(5), 427-436.

Ivanov, V. 2010. Analysis of Tire Contact Parameters Using Visual Processing, *Advances in Tribology* 2010, 11 p.

Janjićijević, N., Danon, G. 1987. Comparison of Experimental and Calculated Dynamic and Steady-state Tire Adhesion during Braking, *in Proceeding of EAEC Conference*, 135-142.

Knoroz, V., Klenikov, E. 1975. *Šini i kolesa*. Moskva: Mašinostroenie, 220 p.

Lenasi, J., Žeželj, S., Danon, G. 1995. *Motor Vehicles*. Belgrade: Faculty of Transport and Traffic Engineering, 433 p.

Moore, H. 1975. *Friction of the Pneumatic Tires*. Amsterdam: Elsevier Scientific Publishing Company, 220 p.

Pillai, P.,bFielding-Russell, G. 1986. Empirical Equations for Tire Footprint Area, *Rubber Chemistry and Technology* 59(1),155-159.

Pottinger, M. 2005.*Contact Patch (Footprint) Phenomena in The Pneumatic Tire*, Edited by Gent, A. and Walter, J. USA: National Highway Traffic Safety Administration, 48 p.

Springer, H., Weisz, G. 1977. Theoretical Investigation on Steady-State Rolling of Steel Belted Tire, *Automobile-Industry* 3, 87-94.

Taylor, R., Bashford, L., Schrock, M. 2000. Methods for Measuring Vertical Tire Stiffness, *Transactions of the ASAE* 43(6),1415-1419.

Uljanov, N. 1982. *Kolesnie dvigitelji*. Moskva: Masinostroenie, 325 p.

Zeng-Xin, Y., Hui-Feng, T., Xing-Wen, D., Li, S. 2001. A Simple Analysis Method for Contact Deformation of Rolling Tire, *Vehicle System Dynamics*, 36(6), 435-443.

*Air Traffic Performance and Infrastructure Development*

# ANALYSIS OF AIRSPACE COMPLEXITY IN FAB CE

**Tomislav Mihetec[1], Ivan Miljak[2], Sanja Steiner[3]**
[1] *INTEGRA A/S, Brussels, Belgium*
[2] *Croatia Control ltd., Zagreb, Croatia*
[3] *Air Transport Department, Faculty of Transport and Traffic Sciences, University of Zagreb, Croatia*

**Abstract** – With continuous increase of air traffic in Europe the term of complexity appears more frequent than ever. In air traffic control complexity is closely related to airspace and sector capacities describing the level of difficulty of operations and it can be used as additional parameter for understanding aircraft en-route delays, as well as the workload of air traffic controllers. Thus, new model for complexity assessment was introduced by EUROCONTROL. Based on few main indicators, such as adjusted density, vertical, horizontal and speed interactions, the air traffic complexity model presents only a part of airspace complexity, which is, as a rule always of higher value. Air traffic complexity indicators are presented and compared on the basis of air navigation services providers. Later, EUROCONTROL's software tool – SAAM was used for detailed analysis of air traffic complexity in the area of Functional Airspace Block Central Europe (FAB CE) with emphasis on Croatian sector groups. This study shows differences in complexity between existing sector groups in FAB CE according to complexity indicators which are presented as relative values in respect of the whole Functional Airspace Block. The traffic sample for analysis is based on radar data instead of flight plan data. Two scenarios are analysed – for lower and upper airspace including analysis of conflict types and shares in FAB CE.

**Keywords** – air transport, air traffic management, functional airspace block, airspace complexity, complexity indicators

## Introduction

Complexity can be generally defined as a condition, which is difficult to resolve and analyse, i.e. as a state of difficulty of an operation or a task. In air traffic, complexity represents the term that is usually related to the tasks of the pilot and air traffic controller, who are exposed to higher psychophysical working conditions. According to Majumdar interaction between sector and traffic features in the determination of controller workload is referred as ATC complexity (*Majumdar, 2003*). Although complexity as an indicator is very often used in order to describe the difficulties of certain activities or a series of connected activities, there is still no generally accepted definition of complexity as part of air traffic management.

It is a known fact that complexity has significant influence on the air traffic controller workload. Grossberg defined complexity as a construct that has dynamic and static characteristics that affect the rate at which controller workload increases (*Grossberg, 1989*). The concept of ATC complexity is subjectively defined by the controller, and is developed from its own perception and interaction with the sector and traffic within it (*Reber, 1985*). Schmidt related workload or control difficulty to the frequency of occurrence of events, which require decisions and actions to be taken by the controller team (*Schmidt, 1976*). According to the FAA, controller workload is influenced by four factors; i.e. constellation of ATC complexity factors, cognitive strategies used by the controllers, quality of the equipment and individual differences between the controllers (*FAA, 1995*). The assumption is that the controller's workload will long remain the highest operative restriction of the ATM system capacity. Therefore, regarding the forecast traffic increase, as well as the development of technology and procedures in air traffic control, it is very important to understand the cognitive capacities of air traffic controller and identification of safety limits of their workload. Workload can also be presented as the controller physical and mental processes to effectively control the traffic (*Pawlak et al., 1996*). One of the factors that restricts the most the safety limits of the physical and mental workload is the air traffic complexity.

Apart from the characteristics of air traffic there is an entire number of factors that affect the usage of airspace, which leads to the conclusion that the air traffic complexity accounts for only a part of the entire complexity of airspace.

The paper includes the research of airspace complexity in Croatia regarding the Functional Airspace Block Central Europe (FAB CE) which includes also Zagreb Area Control Centre Area of Responsibility. The methodology used for complexity analysis in this paper is defined by the EUROCONTROL. The airspace complexity is analysed by the use of EUROCONTROL software tools SAAM[2].

## Air Traffic Complexity Indicators – EUROCONTROL Methodology

There are many elements that are commonly accepted as indicators of ATC complexity: traffic pattern, density of flights in the sector, size and shape of sector, mix of aircraft types and performance characteristics, interface and interactions with adjacent sectors, separation standards, route crossing or convergence points, technical system limitations, restricted airspace, flow entropy , inadequate procedures, etc. (*Manning et al., 2000; Flaker et al., 2000; Histon et al., 2001; Delahaye et al., 2000; Majumdar et al., 2000; Mehadhebi, 1996*)

The model of air traffic complexity presented by EUROCONTROL defines the basic complexity factors, which include the procedures in air traffic control, traffic characteristics and external complexity (*Eurocontrol, 2006*).

[1] Corresponding author: tom@integra.dk

[2] *SAAM - System for Traffic Assignment and Analysis at Macroscopic Level*

The procedures understand parameters such as organization of airspace, route structure, sector configuration and finally the application of operative procedures and concepts for aircraft separation. The previously mentioned parameters are classified as internal factors.

The external factors include the traffic characteristics and external constraints. The characteristics of traffic include the issues of traffic concentration and the type of interactions between the aircraft, and external restrictions include the influence of military actions, interactions of the adjacent air traffic centres and some specific phenomena and events (*Eurocontrol, 2006*).

The internal factors such as organization of airspace or technological achievements can significantly reduce the complexity of the air traffic controller's tasks and improve the productivity and the cost-effectiveness. However, they are not considered as means that reduce complexity but rather as the means that only improve the air traffic management. Therefore, it is the external factors that influence the controller's workload and on whose characteristics this model is based (*Eurocontrol, 2006*).

The traffic density as a factor of complexity describes the distribution of aircraft in the airspace. The uniformity of their distribution regarding the space volume is one of the factors that determine the complexity of air traffic, and thus also of airspace. The potential indicators for the assessment of traffic density are (*Eurocontrol, 2006*):

- raw density,
- adjusted density,
- concentration within specific time intervals,
- spatial and temporal density,
- sector transit time, and
- number of conflicts.

The climbing and/or descending traffic is usually entitled as traffic in evolution. This factor of complexity results from the vertical traffic movements. The lower the number of climbing and descending aircraft, the smaller volume of airspace occupied for the flying of other aircraft, and the lower the number of potential conflicts. A larger number of evolutions lead to greater complexity. Possible indicators for this measure of complexity are (*Eurocontrol, 2006*):

- number of request for level changes,
- level changes,
- number of airports in the vicinity, and
- number of conflicts.

The situation that is described with a certain number of aircraft that fly in the same directions and at the same speeds can be defined as a very simple flow structure. Naturally, this is not the case in practice. The traffic flow structure is the consequence of horizontal movement of aircraft. The greatest complexity consists of the traffic flows that cross and intersect at certain crossing points. Some of the proposed indicators of this complexity factor are the number of conflicts and the number of crossing points (*Eurocontrol, 2006*).

The next parameter, which describes the complexity factor is the traffic mix. This term is introduced in order to define the aircraft performance differences that are in the airspace, usually characterized by different cruising speeds and climbing or descending speeds.

Every airspace that belongs to an Area Control Centre is divided into an arbitrary number of sectors. The distribution of sectors depends also on the structure of routes of the respective airspace. The optimal configuration of sectors affects significantly the complexity of the airspace. The number of operative sectors at some point of time depends on the traffic demand, and can be limited by the available number of air traffic controllers.

The changing of the route structure including the conditions and restrictions of the use, such as the letters of agreements, flight level restrictions, uni/bi-directional routes, affect the reduction of the controller's workload, increase the efficiency and airspace capacity. As is the case with the selection of sectors and sector configurations, military flying can greatly restrict the structure of routes as well. The proposed indicators used for the airspace design structure are:

- number of routes within a given airspace volume,
- number of crossing and merging points within airspace volume,
- direction of flows – uni/bi-directional routes.

Military flying has a significant influence on airspace capacities, airspace complexity and the controller's workload. Among other factors that can affect the complexity there are the influence of military exercises, potential adverse weather conditions and unexpected additional traffic due to certain specific events.

In order to distinguish the difference between the mental and physical workload it is necessary to understand the cognitive complexity. The term cognitive complexity implies the complexity of the air traffic controller performing the planned tasks for the safe separation of aircraft. It can also be defined as the time required to perform the tasks of the executive controller within a certain time interval (usually 1h) (*Eurocontrol, 2004*).

The model and the selection of load values is based on the data from Table 1. The load threshold is equal to the total load in the recorded minutes within the hour divided by one hour of work. Thus, the hour load of 30 minutes corresponds to the 50% load threshold. Such a definition of workload is applied only for the executive controller position, not the planning one. The work overload is defined at the 70% of the hour load threshold (*Eurocontrol, 2005*).

Table 1 – Values and descriptions of air traffic controller's workload threshold

| *Load threshold [%]* | *Load description* | *Recorded working time within 1h [min]* |
|---|---|---|
| 70 & more | Overload | 42 + |
| 54-69 | High | 32-41 |
| 30-53 | Medium | 18-31 |
| 18-29 | Light | 11-17 |
| 0-17 | Very light | 0-10 |

*Source:Eurocontrol, 2005*

Based on the basic tasks of the executive ACC air traffic controller, the model for the assessment of workload has been defined. The model aggregates three main tasks of the controllers. The first is called the routine load and it is based on the presence of the aircraft in the airspace and performing tasks such as radio contact and the like. The next is the task of giving clearance for climbing or descending to the required flight level and monitoring of these actions. The situation in which aircraft use the same airspace resources (overflight points, routes, etc.) at the same moment leads to the obligation of monitoring and resolving potential conflicts. The data required for the calculation of the workload include the duration of tasks and their occurrence within a certain time interval (*Eurocontrol, 2005*).

The equation can be written as follows:

$$WL = t_{AC} \cdot O_{AC} + t_{Cnf} \cdot O_{Cnf} + t_{Cl} \cdot O_{Cl}$$

where:

$O_{AC}$, $O_{Cnf}$ te $O_{Cl}$ – the number of routine tasks, tasks of climbing/descending and resolution of conflicts within the time interval (usually 1h);

$t_{AC}$, $t_{Cnf}$ *te* $t_{Cl}$ – duration of routine tasks, the task of climbing/descending and resolution of conflicts.

In order to connect the impact of complexity on the workload values it is necessary to collect valid data about the duration of tasks for sectors of different complexities and their occurrence. The data on the duration of air traffic controller's tasks are defined and weighted regarding the different sector complexities. The complexity indicators have been selected on the basis of which the differences between the sectors are found. They are divided into the complexity of flight and the complexity of interactions, and they do not differ much from the previously proposed indicators. The complexity of flight includes the number and type of aircraft within a certain interval (number of entries and overflights), data for three peak hours, volume of climbing and descending traffic, and the vicinity of borders of the area centres. The complexity of interactions includes the number and type of conflicts, flow crossing points, routes that converge at low angles, mix of aircraft performances, standards of separation, traffic densities and the time between detection and resolution of conflicts. Taking into consideration all the parameters, the necessary times (durations) to perform the defined tasks are determined (*Eurocontrol, 2005*).

**Complexity Indicators**

In previous sections the basic terms have been presented, which encompass, describe and define the air traffic complexity in a wider sense. The airspace complexity is a wider concept than the air traffic complexity. The air traffic complexity accounts only for a part of the airspace complexity and as a rule it is always of lower value.

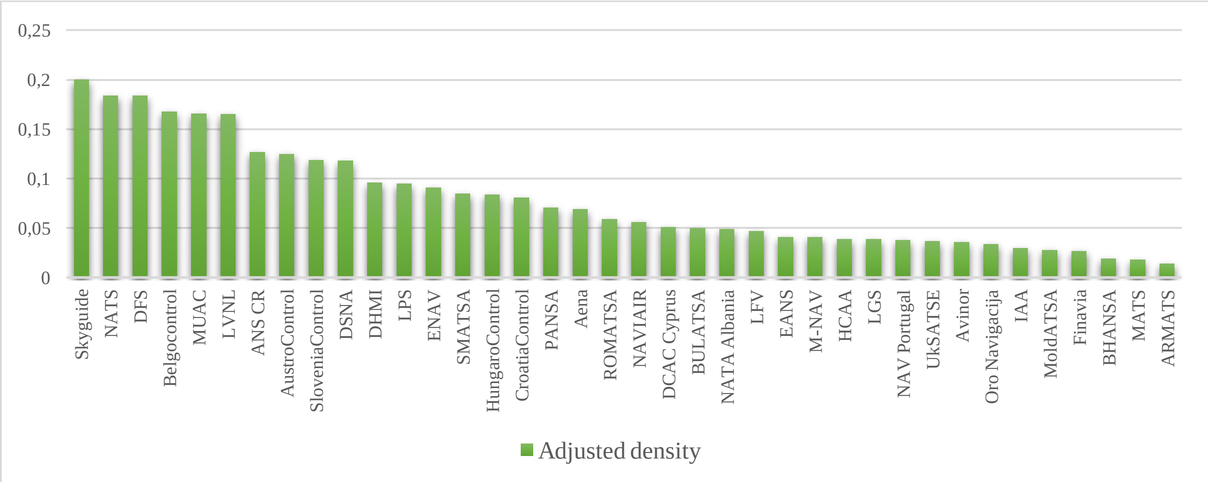
Four basic indicators of complexity have been defined: adjusted density (*AD*), vertical interactions or vertical complexity indicator (*VDIF – Vertical Different Interacting Flows*), horizontal interactions or horizontal complexity indicators (*HDIF – Horizontal Different Interacting Flows*), and speed interactions or speed complexity indicator (*SDIF – Speed Different Interacting Flows*). Although applicable to the terminal airspace, they refer mainly to *en-route* traffic, i.e. traffic which is under the jurisdiction of the Area Control Centre *(ACC)*. The drawback of the model of analysis lies certainly in the impossibility to encompass the internal factors of complexity with the emphasis on the structure of routes, which represents one of the inseparable properties of the traffic flow patterns. The problem of route network has been somewhat approximated by more detailed splitting of airspace into the network of a certain number of cells. It reduces, but does not eliminate, the influence of the route structure (*Eurocontrol, 2013*).

*Adjusted density*

Adjusted density represents complexity indicator selected in relation to the traffic density as the dimension of complexity. Traffic density is defined as the volume of traffic, i.e. the number of aircraft within a certain volume of airspace and within a certain time period. Although for the description of the dimension of complexity, raw density is often used as the ratio of the number of aircraft (number of hours of flight) and the studied volume of airspace, it does not take into consideration the uniform distribution or possibly expressed concentration of aircraft in the busiest parts of airspace or the peak hour. Adjusted density has proven as a better complexity indicator (*Eurocontrol, 2013*).
Adjusted density is a dimensionless parameter. It is defined as the ratio of the interaction hours ($D_k$), and hours of flight ($T_k$), and is expressed by the following equation (*Eurocontrol, 2006*):

$$AD_{ACC} = \frac{\sum_{days}\sum_{cells} D_k}{\sum_{days}\sum_{cells} T_k}$$

The advantage of applying the density defined in this manner lies in the fact that the cells in which there are no aircraft are not included in the calculation since there are no interactions in them nor are there any flight hour values. If there is only a single aircraft present in the cell, there are no interactions either, but the flight hours are taken into consideration. In this way the irregular traffic flow distribution in airspace is simulated. The results can be interpreted as the number of interactions per flight where this value represents an average number of aircraft in interaction (*Eurocontrol, 2006*).
Graph 1 shows the value of adjusted densities for the European air service providers in 2013. *Skyguide* leads with 0.2, whereas the value for CCL is 0.081.




***Graph 1 –** Adjusted density of air service providers in Europe (2013)*
*Source: Eurocontrol, 2013*

*Vertical complexity indicator*

The vertical complexity indicator results from the vertical movement of aircraft in various stages of flight. These movements are usually caused by pre-defined requests of the pilots to change the flight level, recorded in the flight plans. Besides, significant influence is also from bilateral agreements between countries of border ACCs, the so-called Letters of Agreement (LoA), which define the requests for the transfer of aircraft control at pre-arranged flight levels. These flight levels are contained in the form of Flight Level Allocation Schemes (FLAS).
Two flights are in the vertical interaction if they are at the same time in the same cell and have different flight levels. This includes the following positions of two aircraft: cruising/climbing, climbing/descending, cruising/descending. The situations in which both aircraft climb, both aircraft descend or both are in the cruising phase are not considered as vertically interdependent. The aircraft is considered to be cruising if at the moment of entering the cell it reaches the climbing speed lower than 500 feet/minute. The following equations presents the method for determining the vertical complexity indicator. It is defined as the ratio of the summed expected hours of potential vertical interaction of all the cells within the ACC and the summed flight hours of all cells that belong to the respective ACC (*Eurocontrol, 2006*).:

$$V_k = \sum_i \left( \sum_j t_i \cdot t_j \right),$$

$$VDIF_{ACC} = \frac{\sum_{days}\sum_{cells} V_k}{\sum_{days}\sum_{cells} T_k},$$

Where:
$V_k$ – hours of vertical interactions [h],
$t_i$ – time spent by aircraft *i* in cell *k* [h],
$t_j$ – time spent by aircraft *j* in cell *k* [h],
$T_k$ – hours of flight [h],
$VDIF_{ACC}$ –indicator of vertical complexity,
$i,j \in$ cell *k*.

The European system value for VDIF is around 0.03 hours of vertical interactions per flight hour.

*Horizontal complexity indicator*

Similar as in case of vertical complexity indicator, horizontal interactions are the results of different headings. Therefore, two aircraft are in the state of horizontal interaction if they are at the same time in the same cell, and their headings differ by more than 20° (*Pawlak et al., 1996*).
In practice a different terminology is used for lateral interdependence of aircraft. The terms of same tracks, reciprocal tracks and crossing tracks are used.
The indicator of horizontal complexity is the result of the ratio of potential hours of horizontal interactions and hours of flight. The numerator of this ratio is equal to the product of the number of horizontal interactions in the cell and the average time of interactions between two aircraft headings different by more than 20°. The expression is written in the following manner (*Eurocontrol, 2006*):

$$H_k = \sum_i \left( \sum_j t_i \cdot t_j \right), \text{ [h]}$$

where $i \in$ cell *k; j* $\in$ cell *k; i* and *j* have different flight headings. The indicator of horizontal complexity is written by expression:

$$HDIF_{ACC} = \frac{\sum_{days} \sum_{cells} H_k}{\sum_{days} \sum_{cells} T_k}, \text{ [h].}$$

Where:
$H_k$ means indicator of horizontal complexity, and
$T_k$ means hours of flight.

The expected duration of all potential horizontal interactions in all the cells associated with an ANSP/ACC are added together. This is then divided by the total flight hours within the ANSP/ACC to give the value of the HDIF indicator. The European system value for HDIF is around 0.05 hours of horizontal interactions per flight hour (*Eurocontrol, 2006*).

*Speed complexity indicator*

Airspace complexity can be expressed also by the value of speed complexity indicator which is direct consequence of different speeds of aircraft in the cell. Minimal difference in the speeds between aircraft, at the same time being the precondition of the speed interaction amounts to 35 knots (about 65 km/h) and more. This indicator is most similar to the horizontal complexity indicator, naturally, with the difference in the reference value (more than 20°, i.e. more than 35 knots) (*Eurocontrol, 2006*).
The value of speed complexity indicator is obtained as the ratio of the potential hours of speed interactions and hours of flight, and is described by following equations (*Eurocontrol, 2006*):

$$S_k = \sum_i \left( \sum_j t_i \cdot t_j \right),$$

$$SDIF_{ACC} = \frac{\sum_{days} \sum_{cells} S_k}{\sum_{days} \sum_{cells} T_k},$$

Where:
$S_K$ – potential hours of speed interactions [h],
$t_i$ – time spent by aircraft *i* in cell *k* [h],
$t_j$ – time spent by aircraft *j* in cell *k* [h],
$SDIF_{ACC}$ – indicator of speed complexity.

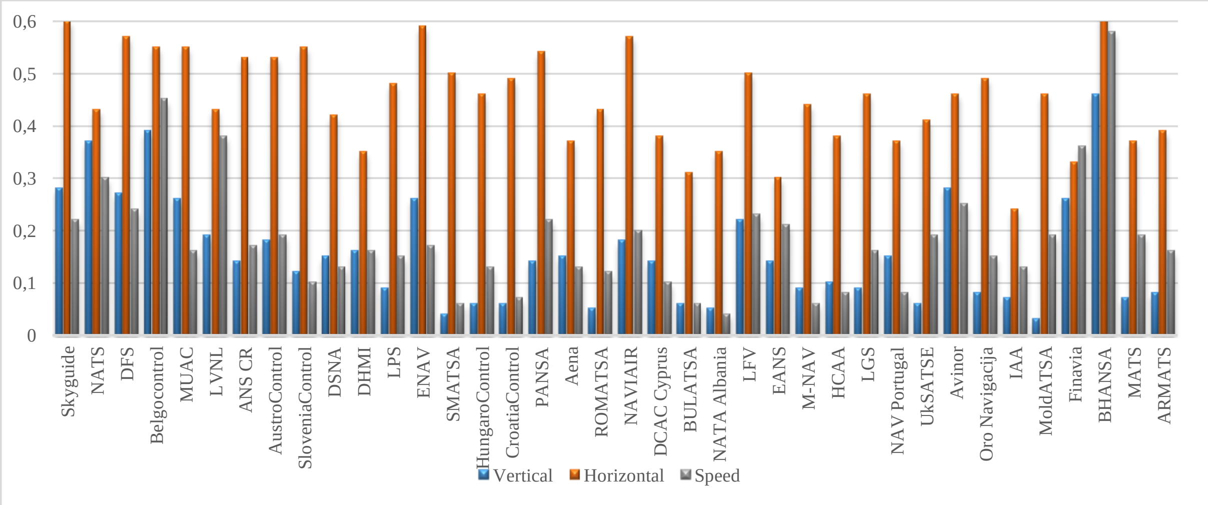
*Complexity result*

For a wider picture of complexity additional complexity indicators were introduced. They consist of the structural index and complexity score. The results of complexity indicators (vertical, horizontal and speed) are mainly under the influence of the traffic flow structure, whereas the adjusted density was conditioned by the traffic volume. In order to separate the traffic flow structure and the traffic volume, relative indicators are introduced (*r_DIF*). They are calculated by dividing the three basic ones with the adjusted density. Thus, adjusted density reflects the traffic volume (AD), and the structural index reflects the structure of traffic flows ((DIF indicators). The structural index is obtained by summing the relative indicators of vertical, horizontal and speed complexity. The complexity score at the level of the air traffic control centre combines these two values and can be presented by equations (*Eurocontrol, 2006*):

$$Score_{ACC} = VDIF_{ACC} + HDIF_{ACC} + SDIF_{ACC}$$

$$Score_{ACC} = AD_{ACC} \cdot (r\_VDIF_{ACC} + r\_HDIF_{ACC} + r\_SDIF_{ACC}).$$

The next graph sums three complexity indicators for service providers in Europe (members of ECAC[3] area) according to the 2013 data.




***Graph 2***
*Basic complexity indicators for air service providers (2013)*
*Source: Eurocontrol, 2013*

The European system value of the complexity score is around 0.10.

**Analysis of Airspace Complexity in Functional Airspace Block Central Europe**

Along with continuous increase in air traffic there is tendency to reduce the costs of air navigation service provision as well as increase of the overall efficiency of the air traffic control system. With this goal, the idea about the establishment of the Functional Airspace Blocks (FABs) has been developed. Such structures of airspace represent airspace blocks whose purpose is the improvement of the cooperation between the air service providers. They have been established regardless of the state boundaries, and every FAB Member State ensures service provision in FAB in its applicable i.e. allocated airspace. Currently, at the level of Europe, nine functional airspace blocks have been established, and Croatia belongs to FAB Central Europe (FAB CE) together with Bosnia and Herzegovina, Slovenia, Austria, Hungary, Slovakia and the Czech Republic (F*AB CE*, *2012*).
Before adopting the implementation plan and airspace distribution plan in 2012, the FAB CE task groups had carried out the static analyses of the airspace including the route structures, conflict points, etc. Based on this analysis, FAB CE is divided into sector groups and sector families. The division and delegation of FAB CE airspace is presented in the following table (*FAB CE, 2012*).

[3] *ECAC – European Civil Aviation Conference*

**Table 2**
*Division and delegation of FAB CE airspace*

| Sector group | FIR | Vertical limits |
|---|---|---|
| *SG1, SG3 and SG4* | Zagreb | **FL205 - UNL**[4] |
| *SG2 and SG4** | Sarajevo | **FL165 – UNL** |
| *SG5* | Ljubljana | **FL175 – UNL** |
| *SG6, SG7 and SG8* | Vienna | **GND**[5] **- UNL** |
| *SG9 and SG10* | Budapest | **GND – UNL** |
| *SG11* | Bratislava | **FL195 – UNL** |
| *SG12 and SG13* | Prague | **GND – UNL** |

*Source: FAB CE Master Plan, 2013*

Within groups, also the sector families are defined and their number can vary, with the aim of managing unit sector configurations. The following figure shows the distribution of sector groups from GND to FL 285 (lower airspace), i.e. from FL 285 to FL 660 (upper airspace).

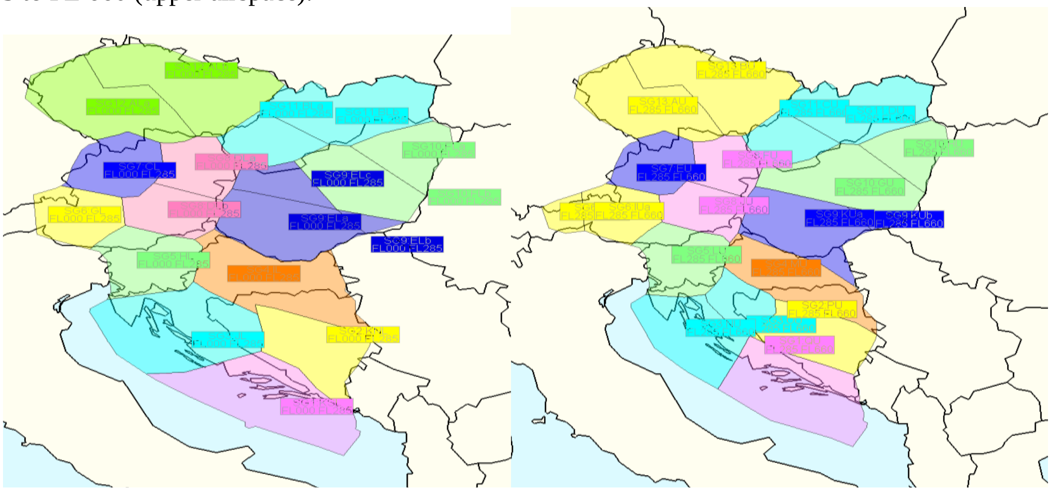


**Figure 1**
*Distribution of sector groups in FAB CE from GND to FL 285 (left) and from FL 285 to FL 660 (right)*

The results obtained by the analysis of airspace complexity which correspond to FAB CE are presented further in this paper in the form of tables and graphs. The already mentioned characteristics of airspace complexity such as sector group loads, sector group load indicators, basic complexity indicators, indicators and results of complexity, conflict loads including 3D density of conflicts have been analysed.

The analysis is based on the data of the radar image of the busiest workday of the year 2012. This was Saturday, 18 August 2012 when 27,536 flights over Europe were recorded (ECAC area). Regarding all the characteristics of airspace complexity the FAB CE airspace has been compared for the heights from the earth surface to FL 285 and from FL 285 to FL 660.

*Scenario for lower airspace according to radar image data*

The next part of analysis has been performed on the same space according to the radar image data. The difference in the data about the flight plans and radar image data is obtained when one takes into consideration the flight delays, cancellations, ATFCM regulation implementation at the level of service providers, vectoring of aircraft in order to prevent conflicts, etc. The data obtained by the analysis are presented in Table 3.
In this case Sector group *SG8* is the most loaded one, with 890 aircraft a day, peak minute load of 25 aircraft and about 70 entries of aircraft in a 20-minute period of the peak hour. Sector group *SG1* recorded 440 flights a day, peak minute load of 15 aircraft and maximal number of entries of aircraft of about 40 in a 20-minute period of the peak hour. Sector group *SG1* best describes FAB CE airspace, since the values of all the parameters of this group are closest to the average values for all the sector groups.

[4] *UNL – Unlimited*
[5] *GND – Ground*

**Table 3**
*Indicators of traffic load of FAB CE airspace from GND to FL 285 according to radar image data*

| GND - 285 *Radar data (.m3)* | | | | |
|---|---|---|---|---|
| *Sector group* | *Number of flights* | *Average distance (NM)* | *Average time (min)* | *Peak minute traffic load* |
| ***SG1*** | **440** | **55.1** | **12.4** | **15** |
| *SG2* | 143 | 56.8 | 10.7 | 5 |
| ***SG3*** | **442** | **65.1** | **12.9** | **13** |
| ***SG4*** | **251** | **44.6** | **9.1** | **8** |
| *SG5* | 311 | 57.9 | 14.1 | 15 |
| *SG6* | 396 | 47 | 8.2 | 10 |
| *SG7* | 496 | 39.8 | 7.2 | 11 |
| ***SG8*** | **888** | **56.9** | **13.1** | **25** |
| *SG9* | 415 | 67.3 | 11.8 | 14 |
| *SG10* | 313 | 45.2 | 9.5 | 9 |
| *SG11* | 317 | 49.7 | 9.4 | 9 |
| *SG12* | 772 | 68.8 | 13.2 | 21 |

Sector group *SG2* represents the airspace of Bosnia and Herzegovina, i.e. its south-eastern part. The results show very low values of airspace load and this airspace volume can be used to balance the loads of sector groups.
*Scenario for the upper airspace according to radar data*
The analysis of the radar image data for the upper airspace gives approximately similar results like for the lower airspace. Here again, regarding the number of flights, the busiest sector in Croatia is Sector group *SG3* – sector *West* with 1,358 flights, i.e. about 50 flights less than foreseen by flight plans. Here, contrary to the results for the lower airspace, *SG3* is the most loaded also regarding peak minute traffic with 31 aircraft. Sector group *SG1* – sector *South* is in the upper airspace also the least busy, and *SG4* – sector *North* represents the average of results of *SG1* and *SG3*. Sector group *SG13* is here again, along with the Hungarian *SG10* the busiest regarding the number of flights (1,751 per day), i.e. per peak minute load which amounts to 51. The data is presented in Table 4.

**Table 4**
*Indicators of FAB CE airspace traffic load from FL 285 to FL 660 according to radar image data*

| 285 - 660 *Radar data (.m3)* | | | | |
|---|---|---|---|---|
| *Sector group* | *Number of flights* | *Average distance (NM)* | *Average time (min)* | *Peak minute traffic load* |
| ***SG1*** | **814** | **126.6** | **16.8** | **23** |
| *SG2* | 817 | 82.8 | 11.2 | 17 |
| ***SG3*** | **1358** | **90.7** | **12.1** | **31** |
| ***SG4*** | **1097** | **102** | **13.7** | **25** |
| *SG5* | 1078 | 89.2 | 11.7 | 26 |
| *SG6* | 1548 | 71.2 | 9.5 | 33 |
| *SG7* | 890 | 67.9 | 9.2 | 17 |
| *SG8* | 1529 | 84.2 | 11.4 | 31 |
| *SG9* | 1188 | 73.8 | 10 | 25 |
| *SG10* | 1717 | 97.6 | 13.1 | 37 |
| *SG11* | 1317 | 87.1 | 11.6 | 29 |
| ***SG13*** | **1751** | **136.7** | **18.1** | **51** |

*Analysis of sector group complexity indicators*

The analysis of the complexity indicators is performed on the basis of the basic complexity indicators and the end results in the form of horizontal, vertical, speed complexity indicators and the results of complexity score. The total results are presented in the form of relative values in relation to overall airspace which is defined by sector groups. Isolated are just some of the basic complexity indicators that serve for the calculation of *AD*, *Hdif*, *Vdif*, *SDif* and the complexity score.

*Scenario for lower airspace according to radar data*

The results for the lower FAB CE airspace according to the radar image data are presented in Table 5. The most complex Sector group is *SG7*, which belongs to the Austrian airspace with relative value of 7,377. The results show that generally the Austrian airspace is the most complex one, whereas the results of Sector group *SG4* represent the Croatian airspace as the least complex and demanding one. The complexity result for *SG4* is the result of very low number of flights via this sector group on the observed day. Significantly, busier and more complex spaces are Sectors *West* and *South*, although still below the average of FAB CE airspace complexity.

**Table 5**
*Values of complexity indicators of FAB CE lower airspace according to radar image data*

| GND/285 (*.m3*) | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| ***Sector group*** | ***Nb Cells*** | ***Sum FL_i*** | ***Sum HC_i*** | ***Sum VC_i*** | ***AD*** | ***HDif*** | ***VDif*** | ***SDIF*** | ***Score*** |
| ***SG1*** | **787** | **144 448** | **501** | **250** | **5,147** | **0,049** | **0,025** | **29,288** | **-1,499** |
| ***SG2*** | 528 | 103 712 | 139 | 76 | 3,961 | 0,022 | 0,012 | 26,575 | -4,361 |
| ***SG3*** | **1241** | **233 824** | **1071** | **670** | **5,024** | **0,059** | **0,037** | **28,306** | **-0,798** |
| ***SG4*** | **801** | **160 304** | **183** | **179** | **4,001** | **0,017** | **0,017** | **27,179** | **-4,183** |
| ***SG5*** | 892 | 168 288 | 705 | 644 | 5,898 | 0,051 | 0,047 | 37,269 | 1,022 |
| ***SG6*** | 450 | 91 640 | 774 | 532 | 5,768 | 0,095 | 0,065 | 46,034 | 4,279 |
| ***SG7*** | 855 | 170 680 | 1675 | 1404 | 8,022 | 0,096 | 0,080 | 50,538 | 7,337 |
| ***SG8*** | 1742 | 325 808 | 2856 | 2068 | 7,555 | 0,085 | 0,061 | 39,419 | 4,527 |
| ***SG9*** | 833 | 161 872 | 429 | 441 | 4,919 | 0,036 | 0,037 | 21,435 | -2,352 |
| ***SG10*** | 878 | 168 712 | 275 | 203 | 4,718 | 0,022 | 0,016 | 15,887 | -4,604 |
| ***SG11*** | 1349 | 263 496 | 928 | 625 | 4,733 | 0,049 | 0,033 | 22,214 | -2,169 |
| ***SG12*** | 2687 | 523 208 | 4069 | 2822 | 5,790 | 0,087 | 0,060 | 36,681 | 2,800 |

*Scenario for upper airspace according to radar data*

Like in the case of lower airspace, the results of complexity show the Croatian airspace as relatively less complex in relation to the overall FAB CE upper airspace. The values of the complexity score range from -2 to -3, but are still not the lowest. The lowest complexity score value is accounted by the Hungarian Sector group *SG9* (about -5). The highest value is recorded by the part of airspace that corresponds approximately to the Slovenian airspace. This is Sector group *SG5*. This is the result of a large number of flights that form the south-eastern flow and in this area of the upper airspace the tops of descent for flights arriving in Austria or Germany are situated.
The values of the complexity indicators obtained for the upper airspace according to the radar image data are presented in Table 6. Generally, lower are the values of adjusted density, indicators of vertical, horizontal and speed complexity. This fact can be explained by the influence of control and separation of aircraft by the air traffic controller. Their role is to identify and regulate the traffic flows anticipated by the flight plans that do not take into account the increased airspace complexity or possible conflict situations.

**Table 6**
*Values of complexity indicators of FAB CE upper airspace according to radar image data*

| 285/660 *(.m3)* | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| ***Sector group*** | ***Nb Cells*** | ***Sum FL_i*** | ***Sum HC_i*** | ***Sum VC_i*** | ***AD*** | ***HDif*** | ***VDif*** | ***SDIF*** | ***Score*** |
| ***SG1*** | **982** | **346 928** | **1819** | **384** | **10,3** | **0,05** | **0,01** | **12,3** | **-3,68** |
| ***SG2*** | 719 | 255 576 | 2036 | 380 | 11,0 | 0,06 | 0,01 | 14,2 | -2,79 |
| ***SG3*** | **1095** | **384 080** | **2570** | **1386** | **10,5** | **0,06** | **0,03** | **18,4** | **-0,74** |
| ***SG4*** | **1009** | **358 856** | **5005** | **751** | **12,6** | **0,10** | **0,01** | **17,5** | **-0,008** |
| ***SG5*** | 891 | 313 144 | 9690 | 1921 | 15,7 | 0,18 | 0,03 | 24,0 | **6,21** |
| ***SG6*** | 968 | 336 872 | 7362 | 2769 | 11,9 | 0,14 | 0,05 | 27,7 | 4,91 |
| ***SG7*** | 713 | 248 912 | 1667 | 520 | 8,61 | 0,06 | 0,02 | 23,6 | -1,92 |
| ***SG8*** | 1420 | 495 920 | 9550 | 1912 | 11,7 | 0,14 | 0,03 | 22,9 | 2,15 |
| ***SG9*** | 971 | 343 624 | 1762 | 320 | 8,90 | 0,05 | 0,02 | 17,6 | -3,76 |
| ***SG10*** | 1734 | 615 416 | 4869 | 422 | 10,6 | 0,06 | 0,005 | 20,1 | -2,46 |
| ***SG11*** | 1839 | 651 616 | 6875 | 558 | 10,9 | 0,08 | 0,007 | 24,3 | -1,008 |
| ***SG13*** | 2785 | 972 960 | 13739 | 3461 | 10,4 | 0,12 | 0,03 | 35,7 | 3,11 |

Sector group *SG4 - North* features a significantly lower complexity score value and best describes the complexity of the overall FAB CE airspace (-0.008). Sector group *SG1* features the lowest value of all sector groups, and the Slovenian Sector group *SG5* features the highest value.

*Analysis of conflicts per sector group*

The last part of the analysis of the airspace complexity refers to the analysis of conflicts per FAB CE airspace sector groups. Although conflicts, according to EUROCONTROL complexity model, do not represent the primary complexity indicator, they are very important for understanding the airspace complexity. Besides, they represent very well the air traffic controller's workload, which is in turn in closest relation to the airspace complexity.
The results have been classified in two ways. The first represents the vertical position of aircraft in conflict, and the other the lateral position and interaction between aircraft in conflict. According to the vertical aircraft position, conflicts can be observed between the evolving aircraft (*evolving/evolving*), between evolving aircraft and cruising aircraft (*evolving/cruising*) and between cruising aircraft (c*ruising/cruising*). Another way of presenting the results is in relation to lateral position of aircraft in conflict. Thus they are divided into parallel, reciprocal and crossing ones.

*Scenario of conflicts for lower airspace according to radar data*

Table 7 shows the results of the analysis of conflicts in the FAB CE lower airspace according to the radar image data. According to these results, there is large difference in the planned and real number of conflicts for the observed day. The recorded number of conflicts amounts to 326, which is a significantly lower number in relation to the assumed 750. This is explained by the influence of the air traffic controller on the timely identification of aircraft in potential conflict and their separation.

**Table 7**

*Results of conflict analysis in FAB CE lower airspace according to the radar image data*

| **GND/285 (*.m3)*** | | | |
|---|---|---|---|
| Number of conflicts | ***326*** | | |
| Number of conflicts per type (%) | | | |
| *Evolving / Evolving* | ***53.37*** | | |
| *Evolving / Cruising* | ***41.1*** | | |
| *Cruising / Cruising* | ***5.52*** | | |
| | | *Evolving / Evolving* | *14.11* |
| *Parallel* | ***24.54*** | *Evolving / Cruising* | *7.98* |
| | | *Cruising/ Cruising* | *2.45* |
| | | *Evolving / Evolving* | *8.9* |
| *Reciprocal* | ***18.71*** | *Evolving / Cruising* | *9.2* |
| | | *Cruising / Cruising* | *0.61* |
| | | *Evolving / Evolving* | *30.37* |
| *Crossing* | ***56.75*** | *Evolving / Cruising* | *23.93* |
| | | *Cruising / Cruising* | *2.45* |

The highest amount of conflicts according to the vertical position of aircraft in conflict occurs between pairs of evolving aircraft (about 51%). Somewhat lower share is accounted for pairs of *Cruising/ Evolving* aircraft (41%). High values of both combinations of aircraft positions result from the fact that flight levels up to 285 are used for cruising of the turbofan aircraft and climbing of turbojet aircraft in order to reach the desired cruising levels. In lateral sense, the highest amount of conflicts has been recorded between the crossing aircraft (57%). Out of the total number of conflicts, with the combination of the analysis of vertical and lateral position of aircraft, the largest part, about 30%, are accounted for pairs of aircraft that are crossing and evolving.

*Scenario of conflicts for upper airspace according to radar data*

The results of the analysis of conflicts in the upper airspace according to radar data are presented in Table 8.
According to the radar image data in the upper airspace there is an even more pronounced share of conflicts in the sense of vertical relation *evolving / cruising*. It amounts here to about 65%. In lateral dependence of aircraft position, unlike the results for the lower airspace, the largest share in the structure of conflicts in the upper airspace are accounted for crossing flights with about 48%. Out of this in the total structure, about 30% of conflicts belong to *evolving / cruising* ones.

**Table 8**
*Results of analysis of conflicts in FAB CE upper airspace according to radar image data*

| **285/660 (*.m3)*** | | | |
|---|---|---|---|
| Number of conflicts | **661** | | |
| Number of conflicts per type (%) | | | |
| *Evolving / Evolving* | **22.39** | | |
| *Evolving / Cruising* | **65.51** | | |
| *Cruising/Cruising* | **12.1** | | |
| | | *Evolving / Evolving* | **9.83** |
| *Parallel* | **28.29** | *Evolving / Cruising* | **14.98** |
| | | *Cruising / Cruising* | **3.48** |
| | | *Evolving / Evolving* | **2.87** |
| *Reciprocal* | **23.75** | *Evolving / Cruising* | **20.27** |
| | | *Cruising / Cruising* | **0.61** |
| | | *Evolving / Evolving* | **9.68** |
| *Crossing* | **47.96** | *Evolving / Cruising* | **30.26** |
| | | *Cruising / Cruising* | **8.02** |

## Conclusion

By implementing the complexity model for FAB CE airspace, the paper has shown that the Croatian airspace, regarding the traffic load values, complexity indicators and share of conflicts, belongs to the average complex airspace (within the frames of the defined FAB CE sector limits) regarding the overall space volume encompassed by the functional block of Central Europe.

At the level of Europe the efficiency of air service is monitored according to four key performance areas. They include safety, capacity, environment and cost-efficiency. Apart from these key areas, there has been an idea for some time now about introducing the airspace complexity as an indicator of efficiency. By analysing the literature, much diversity in defining the complexity has been determined at the level of air service provider, particularly in defining the complexity indicators. Thus, the biggest barrier in the development of the idea of complexity is the absence of a generally accepted model, and the key complexity indicators, and as consequence, the absence of legal regulations.

Regardless of the absence of standards and obligations of applying the airspace complexity in monitoring the air traffic service performance, it is increasingly used today in comparisons of interdependent airspace volumes. Some generally accepted complexity indicators are used, and these are used also in this paper, in order to compare the sector groups of the Croatian airspace with other sector groups of the functional airspace block of Central Europe (FAB CE).

FAB CE airspace complexity has been analysed for the busiest day in the year 2012, 18 August (Saturday), regarding the data obtained by radar image correlation. It has been found that the Croatian sector groups, from the aspect of lower and upper airspace, belong to a relatively less complex FAB CE airspace volume. This conclusion is based on the data obtained for the number of overflights and peak traffic loads. Also, the results of complexity indicators indicate lower complexity of the Croatian sector groups, that as a rule record negative results, both in the lower and upper airspace. The negative complexity score value of the sector group indicates that the complexity of this airspace volume is lower than the average complexity of all sector groups together. The complexity score values of sector groups *SG1*, *SG3* and *SG4* are mainly in the range from 0 to -4. There are certain differences in the values of complexity indicators per sector groups in the lower and upper airspace. Thus Sector group *SG4* (sector North) has been estimated as the least complex in the lower airspace (complexity score of about -5), and in case of the upper airspace it describes best the overall FAB CE airspace (-0.008).

Generally, the least complex airspace volume represents the eastern part of Bosnia and Herzegovina to which Sector group *SG2* has been allocated, and the Hungarian airspace. Possible balancing of the values of complexity and load of sector groups at FAB CE level by restructuring the flows is certainly recommended in the area of these sector groups. The most complex sector groups according to the performed analysis are the Slovenian upper airspace and generally the Czech airspace. The analysis of conflicts shows significantly lower values regarding radar image data, i.e. real conflicts in relation to the forecast ones. This confirms the positive influence of air traffic controllers on the reduction of the forecast airspace complexity, which are in continuous influencing interrelation.

The main drawback of this method of comparing the complexity is expressed in the form of restriction of the SAAM software to present the absolute values of airspace complexity. It is therefore very difficult to confirm the thesis about the higher complexity of the lower airspace in relation to the upper airspace. Possible upgrades of this analysis should include a larger number of representative days and data about the traffic through a year, in order to obtain some average values for every sector group. Also, at the beginning of FAB CE operative action, re-sectorization of airspace within all members has been planned. Thus, the sectors will no longer be based on state borders, but rather on sector groups themselves. The analysis of airspace by applying new sector configurations would provide an even better insight into the lack of uniformity of FAB CE complexity. Additional upgrade of SAAM software, including this analysis, could be achieved by the development of the model applicability, and on terminal areas, below flight level 085.

**References**

Delahaye, D., Puechmorel, S., 2000, *Air traffic complexity : towards intrinsic metrics*, 3rd ATM 2000, ENAC

EUROCONTROL, 2004. EC Experimental Centre: *Cognitive Complexity in Air Traffic Control,* a literature review, March 2004.

EUROCONTROL, 2005. EC experimental Centre: *Adoption of Workload Model by Optimization Algorithms and Sector Capacity Assessment*, Project COCA, March 2005.

EUROCONTROL, 2006. Performance Review Commission: *Complexity Metrics for ANSP Benchmarking Analysis.* Report by the ACE Working Group on complexity, 2006.

EUROCONTROL, 2013. Performance Review Commission: Performance Review Report, *An Assessment of Air Traffic Management in Europe*, May 2013.

EUROCONTROL, 2013. *Seven-Year Forecast: Flight Movements and Service Units*, February 2013.

FAB CE Airspace Planning Design Group, 2012. *FAB Central European Implementation Phase: Airspace Plan 2012,* February 2012

Federal Aviation Administration, 1995. *The Complexity Construct in Air Traffic Control: A Review and Synthesis of the Literature*, Virgina, 1995

Flaker, J.M. et al. 2000, *An analytical model of the impacts of airspace restriction*, 4th ATM 2001

Grossberg, M. 1989. *Relation of sector complexity to operational errors*. In Quarterly Report of the FAA Office of Air Traffic Evaluations and Analysis. Washington, DC: Federal Aviation Administration, 1998

Histon, J. et al., 2001, *Introducing structural considerations into complexity metrics*, 4th ATM 2001, MIT LOG/ CENA

Majumdar, A.2003. *A Framework For Modelling The Capacity Of Europe's Airspace Using A Model Of Air Traffic Controller Workload,* PhD Thesis, University of London, 2003

Majumdar et al., 2000, *The factors affecting air traffic controller workload: a multivariate analysis based upon simulation modelling of controller workload,* London, Centre for Transport Studies

Manning et al., 2000, *Investigating the validity of Performance and Objective Workload Evaluations Research* (POWER), FAA.

Mehadhebi, K., 1996, *Comment representer la charge de travail du controleur*, CENA.

Pawlak,W.S., Briton, C.R., Crouch, K., Lancaster,K., 1996. *A Framework For The Evaluation of the Air Traffic Control Complexity,* American Institute of Aeronautics and Astronautics,1996

Reber, A. S. 1985. *The penguin diction psychology*. New York, NY: Viking Penguin Inc, 1985

Schmidt, D. K. 1976. *On modelling A TC workload and sector capacity*. Journal of Aircraft.

# LEARNING FROM ERRORS – CASE STUDY OF AN AIRCRAFT ACCIDENT

**Ivana Čavka[1] , Michelangelo Mariani[2], Petar Mirosavljević[3], Francesco Abbondati[4], Olja Čokorilo[5]**
[2, 4] *University of Napoli Federico II, Department of Transportation Engineering, Via Claudio 21, 80125 Naples, Italy*
[1, 3, 5] *University of Belgrade, Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, 11000 Belgrade, Serbia*

**Abstract:** The process of aircraft accident and serious incident investigation is the basis for understanding the underlying causes of these events. The sole objective of the investigation is making safety recommendations intended to prevent recurrences. Following the international standards and recommended practices as described in ICAO Annex 13 - Aircraft accident and incident investigation, it is possible to obtain and understand factual information about the accident/incident and to develop measures which could mitigate and/or reduce future risks. This paper deals with the results of an investigation of the accident of Flight 1153 from Bari (Italy) to Djerba (Tunisia) that was simulated in order to get insight into the possible errors and their potential avoidance. Obtained results could serve as an important factor in identifying hazards that led to the accident, one of the primary purposes of the aviation safety management system.

**Keywords:** aircraft, Safety Management System (SMS), accident investigation, error.

## 1. Introduction

The concept of safety in the aviation sector can imply the absence of any hazard or risk that might lead to the accident (Waikar and Nichols, 1997), but also the avoidance of attacks of various types (Watson et al., 2011). Therefore, it is important to understand and identify the factors that cause or might cause damage to the aviation system as a whole. Contemporary research on hazards in aviation widely consider the risk assessment as part of the safety management system (Čokorilo et al., 2011; Hurst et al., 1996; Lee, 2006; Shyur, 2008; Lawrence and Gill, 2007; Oztekin and Luxhøj, 2010).

The experience of the last decade on a range of accidents and disasters in aviation, where the so-called “management failures” have proved to be the main causes of such events, has indicated the need to establish a policy for the identification of potential risks arising from weaknesses in the organization, through the adoption of a Safety Management System (SMS) for organizations that operate within civil aviation.

The SMS is a systematic process for managing the risks associated with flight operations, airport operations (De Luca and Dell'Acqua, 2014) and air traffic control in order to obtain high levels of safety. This objective can be achieved through identification, assessment and elimination or mitigation of potential risks to acceptable levels.

ICAO international rules contained in Annex 6 (Operation of Aircraft), Annex 11 (Air Traffic Services), Annex 14 (Aerodromes), and Annex 19 (Safety Management), propose that the member states have to adopt a Safety Program that ensures an acceptable level of safety for their operations. With a new special document named “Safety Management Manual” (ICAO Doc 9859-AN/474, 2013), ICAO developed a complete unified strategy for the implementation of SMS systems in relation to flight operations, air traffic services and airport operations.

Despite the knowledge, experience and competence of the people involved in the aviation sector and even though they are not intentionally procured, the mistakes are frequent. In other words, the possibility of human error in aviation is enormous. Fortunately few errors lead to negative consequences and usually these errors are timely identified and controlled without undesirable outcomes. It would be impossible to eliminate all human errors which would also be an unrealistic goal.

The Error Theory by James T. Reason classifies human errors in four groups as follows:

- Slips – errors due to non-completion of the procedures;
- Lapses – errors due to omissions, negligence, etc.;
- Mistakes – errors due to loss-making organizations, conflicting objectives, or management decisions;
- Violation – errors due to deliberate violations of established procedures.

Reason (1990) developed a model for the analysis of human errors (ICAO Doc 9683-AN950, 1998). This model explains the organizational accident concept: accidents require the combination of a number of factors, each necessary but not sufficient alone to breach the defenses of the system.

*Latent conditions* are conditions present in the system before a bad outcome has experienced and is made evident by local triggers. They can remain inactive for a long time, and become evident once the system's defenses are breached. They are generally created by people distant in time and space from the event.

*Active failures* are actions or missing actions, including errors and violations, which have an immediate negative effect. They are generally associated with the front-line personnel (pilots, air traffic controllers), and take place in an operational context which includes latent conditions.

[1] Corresponding author: i.cavka@sf.bg.ac.rs

A recent study conducted by Boeing (2013) showed that the situations that generate the highest number of fatalities on board are: the loss of control of the aircraft (during the both landing and taking-off from runway) and misunderstandings (sometimes due to an imperfect knowledge of the English language by one of the parties) between the control tower and the pilot. The same study also highlighted the relationship between fatal accident and accident with hull loss (with complete damage to the aircraft).
Brady and Hillestad (1995) performed a study of risk outside Schiphol airport in Amsterdam. This study was motivated by the crash in October 1992 of an El Al cargo into an apartment building near Amsterdam, killing more than 40 people. The study assessed possible measures enhancing safety to be implemented. This model, called "Safety Assessment of the Ground Environment of the Airport" (SAGE-A) was developed as a general tool for the assessment of risks outside the airport.
Eisenhawer et al. (2009) presented the use of Logic Gate Models (LGMs), the hierarchical models that includes event, fault and decision trees, to perform a risk-benefit analysis for an advanced ATS technology in the Next Generation Air Transportation System (NextGen). To demonstrate the application of LGMs to systems analysis of advanced technologies for NextGen, a Logic Evolved Decision Analysis (LED) model was developed for Airborne Precision Spacing (APS). APS is a Flight Deck Merging and Spacing (FDMS) tool developed by NASA to allow aircraft to maintain consistent inter-aircraft spacing intervals at the runway threshold. The risk model combines accident scenarios with historical data for operational incidents and estimates of the rates of human errors. The benefit model uses the time delay of arrival and the increase in the rate of arrival at the airport to assess the changes in the system performance. Risk-benefit calculations are performed for each scenario. The analysis showed that the APS-based operations of merging and spacing show significantly lower risk and greater benefit than the current practice in a range of demand states and changes in the system design.
Čokorilo et al. (2014) conducted a study on accidents in aviation on a global scale. The study covers aircraft accidents in the period 1985-2010; in particular, a cluster analysis technique was used to develop a predictive model of aircraft accidents in terms of their annual average.
The study carried out in this paper analysis the air accident that occurred on TUI 1153 flight. The study was divided into two phases: in the first phase a list of human errors was identified through the use of the Reason's model; in the second phase a laboratory simulation of the TUI 1153 flight was performed to assess the behavior of aircraft pilots during the emergency procedure.

**2. The Aircraft Accident**

The accident occurred on August 6, 2005 to an ATR 72-202 originating from Bari to Djerba. One hour after leaving Bari with 39 people on board, including the crew cabin, both aircraft engines shut down. After several attempts by the pilots to turn the engines on and after 16 minutes of gliding, the aircraft made an emergency landing in the Bay of Capo Gallo near the airport of Punta Raisi, Palermo.
The contact with the sea surface caused the breakdown of the aircraft into three parts (Fig. 1). The accident resulted in fatal injuries to 16 people and various serious injuries for the rest of people on board.

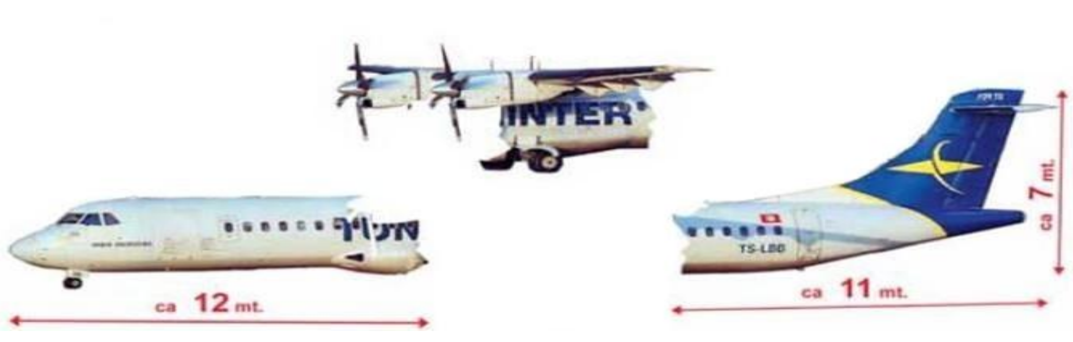



**Fig. 1.**
*Three Sections the Aircraft Broke Into Following Impact with the Sea Surface*
*Source: ANSV Report*

### 3. The Accident Official Report

The accident investigation was conducted by the National Agency for the Safety of Flight (ANSV) in accordance with the provisions of the ICAO Annex 13. It was determined that the main cause of the accident was the replacement of the FQI (Fuel Quantity Indicator); the technician erroneously installed on the ATR 72 an FQI made for an ATR 42. These two aircraft have wing tanks with different shape and capacity. After the replacement took place, the amount of fuel on board was 790 kg (395 left tank - 395 right tank), but the quantity shown by the instruments was 3050 kg (1525 left tank - 1525 right tank), as depicted in Fig. 2.
This accident, like most aircraft accidents, was imposed by a series of linked and successive events that led to the accident (Čavka and Čokorilo, 2012).

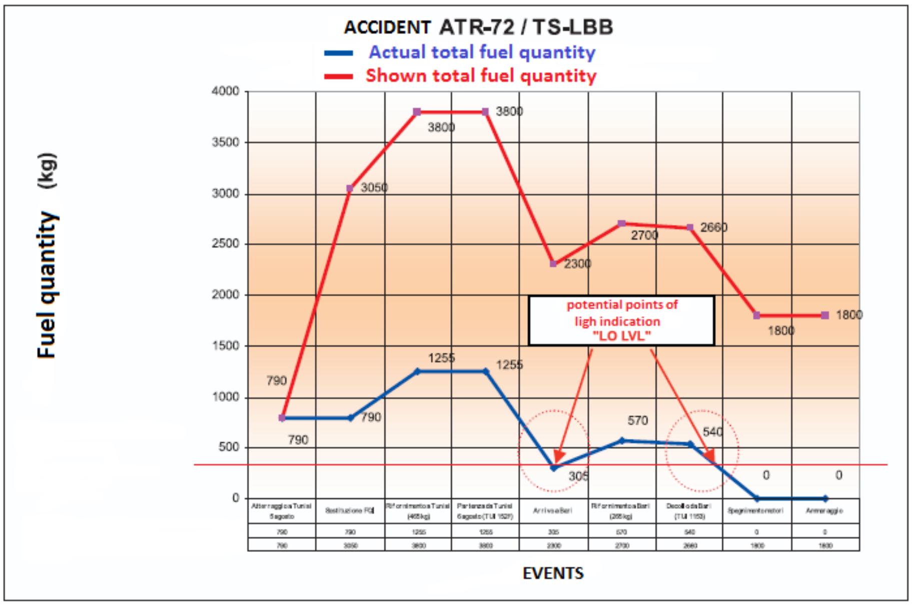



**Fig. 2.**
*The Quantity of Fuel Actually on Board (Blue Line) and the Quantity Indicated by the FQI (Red Line)*
*Source: ANSV Report*

### 4. Analysis of the Errors that Caused the Accident

For the investigation of errors in preparation and management of the flight, the Reason's model was used. In particular, beyond the main cause of the accident due to the wrong replacement of the FQI, a series of errors have been identified as contributing causes of the mishap. The errors shown below are in the order that have contributed to the accident in question:

- Lack of accuracy of data entered into the spare parts management system;
- Lack of training in using the spare parts management system;
- Lack of maintenance and control technical management;
- Lack of completion of flight plans by crew;
- Absence of a Quality Assurance System;
- Absence of Safety Management System;
- Lack of supervision of the operator by the Tunisian Authority;
- Error committed by technicians in FQI research;
- Inadequate flight preparation;
- Lack of a flight data monitoring system;
- Error committed by crew in not realizing the incorrect fuel consumption;
- Error committed by pilots for non-compliance with emergency operational procedures.

## 5. Accident Simulation

To verify the effectiveness of procedures applied by pilots during the emergency phase, the flight TUI 1153 was simulated from the start of the take-off until the moment of emergency phase. The flight simulation was carried out in the laboratory at the Faculty of Transport and Traffic Engineering, University of Belgrade, using the Microsoft Flight Simulator 2004 software (Fig. 3).

**Fig. 3.**
*Flight Simulator – Faculty of Transport and Traffic Engineering, University of Belgrade*

To compare the simulated flight with the real flight, the following 4 phases have been identified:

T0 – TAKE-OFF (Table 1 and Table 2);
T1 – CRUISE (Table 3 and Table 4);
T2 – BOTH ENGINES FLAME OUT (Table 5 and Table 6);
T3 – DITCHING (real case) / LANDING (simulated case) (Fig. 4).

**Table 1**
*T0 - Take-off: Real Case*

| | |
|---|---|
| **Airport** | Bari |
| **Model** | ATR 72-202 |
| **Passenger** | 36 |
| **Cabin crew** | 3 |
| **Dummy (Actual) weight** | 19400 (17250) kg |
| **Runway** | 07 |
| **Dammy (Actual) fuel** | 2660 (540) kg |

**Table 2**
*T0 - Take-off: Simulated Case*

| | |
|---|---|
| **Airport** | Bari |
| **Model** | ATR 72-500 |
| **Passenger** | 36 |
| **Cabin crew** | 3 |
| **Gross weight** | 19000 kg |
| **Runway** | 07 |
| **Fuel** | 540 kg |

**Table 3**
*T1 - Cruise: Real Case*

| | |
|---|---|
| **Direction** | AMANO |
| **Model** | ATR 72-202 |
| **Passenger** | 36 |
| **Cabin crew** | 3 |
| **FL – altitude** | 23000 ft |
| **Fuel consumption** | 10 kg/min |
| **Airspeed** | 200 kts |

**Table 4**
*T1 - Cruise: Simulated Case*

| | |
|---|---|
| **Direction** | AMANO |
| **Model** | ATR 72-500 |
| **Passenger** | 36 |
| **Cabin crew** | 3 |
| **FL – altitude** | 23000 ft |
| **Fuel consumption** | 10 kg/min |
| **Airspeed** | 240 kts |

**Table 5**
*T2 - Both Engines Flame Out: Real Case*

| | |
|---|---|
| **Direction** | AMANO |
| **Model** | ATR 72-202 |
| **Passenger** | 36 |
| **Cabin crew** | 3 |
| **FL – altitude** | 1900 ft |
| **Fuel consumption** | 0 kg/min |
| **Airspeed** | 180 kts |
| **Fuel Shown** | 1800 kg |

**Table 6**
*T2- Both Engines Flame Out: Simulated Case*

| | |
|---|---|
| **Direction** | AMANO |
| **Model** | ATR 72-500 |
| **Passenger** | 36 |
| **Cabin crew** | 3 |
| **FL – altitude** | 1700 ft |
| **Fuel consumption** | 0 kg/min |
| **Airspeed** | 130 kts |
| **Fuel Shown** | 0 kg |

**Fig. 4.**
*T3 - Landing (simulated phase)*

The simulation (phases T1-T2 and T3/ LANDING, simulated case) illustrates that, if the pilot carried out the correct procedure prescribed for this situation (procedure “BOTH ENGINES FLAME OUT”), he would had a high probability of landing at the nearby airport of Punta Raisi (Palermo, Italy) avoiding ditching. Specifically, as shown in all simulated cases, the application of the procedure “BOTH ENGINES FLAME OUT” would allow handling the emergency and landing at the nearby airport of Punta Raisi after approximately 80 minutes from the moment of engines shutdown.
In particular, the “BOTH ENGINES FLAME OUT” procedure that was applied in the simulator during the emergency phase and that enabled landing at the airport of Punta Raisi, contained the following settings:

- the aircraft was brought to the speed of “drift down”, which for the ATR 72 in this type of emergency is comprised between 127 and 137 knots;
- the condition lever was first placed on the position of “feathering” and subsequently on the position of “shut off”; in particular, with the first action the propeller blades were put in the flag position to decrease the running resistance of the aircraft, and with the second action the flow of fuel to engines was suddenly blocked;
- the flaps were brought to the 0° position.

## 5. Conclusion

The objective of this paper was to analyze the ATR 72 accident that occurred on the flight TUI 1153 from Bari to Djerba. The Reason’s model was used to identify all the causes and factors that influenced the occurrence of this accident. In particular, as shown in the third section, even if the main cause was the incorrect replacement of the FQI, all the mistakes made by mechanics and pilots before and after take-off had a significant contribution.
Furthermore, through the flight simulation it was shown that if pilots had successfully implemented the procedure “BOTH ENGINES FLAME OUT”, the accident could have been avoided.
In conclusion, being aware of the limits of simulation regarding the flight operating conditions, it has to be noted that the study supported by simulator represents a valid tool for the overall analysis of accidents of this type and therefore can provide important suggestions for the improvement of aviation safety.

**References**

Boeing. 2013. *Statistical Summary of Commercial Jet Airplane Accidents*, Worldwide Operations (1959-2013). Aviation Safety Boeing Commercial Airplanes, Washington, USA. 27 p.

Brady, S.D.; Hillestad, R. 1995. *Modeling the External Risks of Airports for Policy Analysis*. European - American Center for Policy Analysis, RAND, Santa Monica. 80 p.

Čavka, I.; Čokorilo, O. 2012. Cost - Benefit Assessment of Aircraft Safety, *International Journal for Traffic and Transport Engineering*, 2(4): 359-371.

Čokorilo, O.; Mirosavljević, P.; Gvozdenović, S. 2011. An Approach to Safety Management System (SMS) Implementation in Aircraft Operations, *African Journal of Business Management*, 5(5): 1942-1950.

Čokorilo, O.; De Luca, M.; Dell'Acqua, G. 2014. Aircraft safety analysis using clustering algorithms, *Journal of Risk Research*, 17(10): 1325-1340.

De Luca, M.; Dell'Acqua, G. 2014. Runway surface friction characteristics assessment for Lamezia Terme Airfield Pavement Management System, *Journal of Air Transport Management*, 34: 1-5.

Eisenhawer, S.; Bott, T.; Foggia, J.; Brown, S.; Neitzke, K.; Sorokach, M. 2009. Risk-Benefit Analysis of Advanced Air Transportation System Technologies Using Logic Gate Models. In *Proceedings of Eighth USA/Europe Air Traffic Management Research and Development Seminar (ATM 2009)*, Napa, California, USA. 10 p.

Hurst, N.W.; Young, S.; Donald, I.; Gibson, H.; Muyselaar, A. 1996. Measures of Safety Management Performance and Attitudes to Safety at Major Hazard Sites, *Journal of Loss Prevention in the Process Industries*, 9(2): 161-172.

ICAO. 1998. *Human Factors Training Manual*, Doc 9683-AN/950.

ICAO. 2013. *Safety Management Manual*, Doc 9859-AN/474.

Lawrence, P.; Gill, S. 2007. Human Hazard Analysis: A Prototype Method for Human Hazard Analysis Developed for the Large Commercial Aircraft Industry, *Disaster Prevention and Management*, 16(5): 718-739.

Lee, W.K. 2006. Risk Assessment Modelling in Aviation Safety Management, *Journal of Air Transport Management*, 12(5): 267-273.

Oztekin, A.E.; Luxhøj, J.T. 2010. An inductive reasoning approach for building system safety risk models of aviation accidents, *Journal of Risk Research*, 13(4): 479-499.

Reason, J. 1990. *Human Error*. Cambridge University Press, United Kingdom. 302 p.

Shyur, H.J. 2008. A Quantitative Model for Aviation Safety Risk Assessment, *Computers & Industrial Engineering*, 54(1): 34-44.

Waikar, A.; Nichols, P. 1997. Aviation Safety: A Quality Perspective, *Disaster Prevention and Management*, 6(2): 87-93.

Watson, A.; Dolislager, F.; Hall, L.; Raber, E.; Hauschild, V.D.; Love, A.H. 2011. Developing Health-Based Pre-Planning Clearance Goals for Airport Remediation Following a Chemical Terrorist Attack: Decision Criteria for Multipathway Exposure Routes, *Human and Ecological Risk Assessment*, 17(1): 57-121.

# IMPROVEMENT OF ON-TIME PERFORMANCE AT SPLIT AIRPORT

**Mate Melvan**[1], **Matija Bračić**[2], **Stanislav Pavlin** [3]
[1] *Split Airport, Cesta dr. Franje Tuđmana 1270, Split, Croatia*
[2, 3] *University of Zagreb, Faculty of Transport and Traffic Sciences, Vukelićeva 4, Zagreb, Croatia*

**Abstract:** The traffic level at the Split Airport is not equally distributed over the year. A specific characteristic is its extreme seasonality. This has an important impact on Air Traffic Flow Management (ATFM). Congestion and possible delays occur wherever the capacity of airspace sectors is exceeded over a period of time. It is mostly associated with peak traffic hours of the day during the summer schedule. On a day-to-day basis ATFM attempts to plan air traffic demand with the capacity of airports and airspace sectors. Split Airport is Schedule Facilitated airport due to runway constraint. This paper presents software interface between airport capacity and airspace sectors to improve the on-time performance and punctuality of summer schedule.

**Keywords**: on-time performance, airport, network

## 1. Introduction

Aircraft on-time performance represents the most important factor for the regular air traffic flow. In the last few decades the problem of aircraft on-time performance with continuous increase of air traffic becomes one of the main problems in the air traffic system. Certain number of airports with a higher level of traffic (hub airports) as well as the airports with not uniformly distributed traffic due to the lack of capacity makes the limiting factor in the performance of aircraft operations on the ground. Conventional Air Traffic Management (ATM) system in circumstances of air traffic growth will not be able to process aircraft without significant delay. Aircraft on-time performance requires mutual interaction and coordination between all subsystems in air traffic. If a certain delay occurs in any segment of air traffic system that reflects on the entire system (Fig. 1). The limiting factor of aircraft on-time performance is inadequate coordination between airport system (along with all its subsystems) and air traffic control.

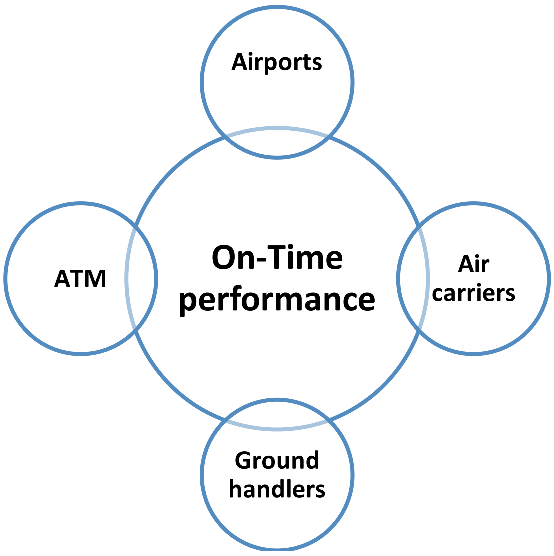



**Fig. 1**
*Aircraft on time performance*

Air traffic system delays have a negative impact on all participants, but the biggest impact is on the passengers as users of transport services. The problem of aircraft on-time performance at the network level could be observed through the aircraft delay in all subsystems of air traffic.

## 2. Aircraft on-time performance

Statistics based on the United States pattern indicates that nearly three quarters of all domestic flights are performed on time. Statistics is based on the reports from top fourteen U.S. airlines and twenty-nine airports by realized aircraft operations that cover most of the United States air traffic (Fig 2). The largest share of aircraft delays represents delay due to late arrival of aircraft from previous flight with 10.43% followed by delay caused by airlines with 7.06%, national aviation system delay with 7.39%, etc.

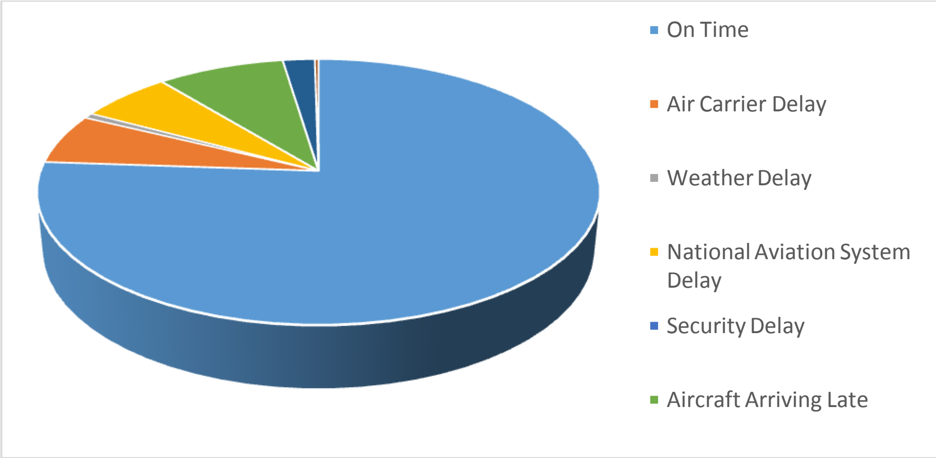



**Fig. 2**
*Aircraft On-Time Arrival Performance within United States air traffic network (January – December 2013)*
*Source: Bureau of Transportation Statistics, Airline Service Quality Performance 234*

At the European level the analysis indicates that the average delay time by the delay flight (for all causes) for both arrival and departure aircraft operations in 2013 amounts to 9.5 minutes according to data generated by air carriers (Fig 3). Departure aircraft operations generated 36.1% of the total aircraft delay which represent increase in the relation to 2012.

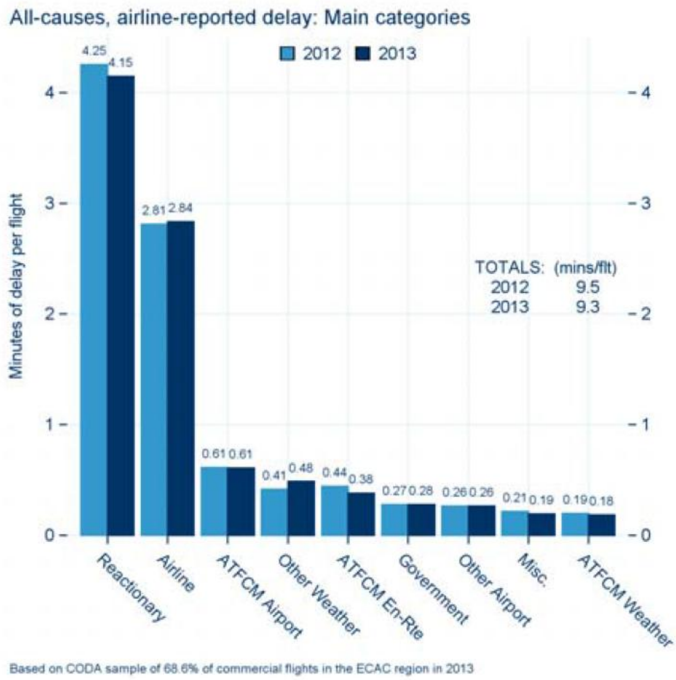



**Fig. 3**
*Primary Delay Causes 2013 vs. 2012 within European air traffic network*
*Source: CODA digest, EUROCONTROL*

The largest share of aircraft delays per single flight for 2013 represent delay due to late arrival of the aircraft and crew from previous flight (reactionary delay) with 4.15 minutes, followed by the delay caused by the airlines with 2.84 minutes, aerodrome airspace capacity with 0.61 minutes and the delay caused by weather with 0.48 minutes, etc.
The Republic of Croatia in terms of delay and aircraft on-time performance follows the trends recorded at the level of the European air traffic network. In the last few years air traffic in the segment of aircraft operations has seen a slight increase. From the perspective of delay within the Croatian airspace the trends show a decrease in recent years. In 2012 the average delay time was 0.28 minutes while in 2013 the average delay time amounted to 0.1 minutes. Delays at the Split Airport indicate that average delay time per delay in 2013 was 20.5 minutes, which represents a decrease of 9% compared to the previous year. Also it is indicated that 75% of delays were reactionary delays.

## 3. Split Airport

Split Airport is a Mediterranean airport located at the central south of Croatia. In 2013 the number of aircraft operations amounted to approximately 18,000. The annual passenger traffic was 1.56 million passengers with a tendency of further growth. According to the number of passenger handled at the airport, it represents the second airport in the Republic of Croatia.

### 3.1 Seasonality

A specific characteristic of the Split Airport is its extreme seasonality, i.e. during the summer season the traffic volume is extremely high, at a monthly level some eleven times higher than during winter months.
Figure 4 shows the distribution of traffic at the Split Airport during calendar year 2013. Aircraft traffic at the airport is most pronounced during summer months (July and August) with an average of 3,500 aircraft operations per month. On the other hand, in the winter months (between November and March) the airport records approximately 400 aircraft operations per month. The traffic at the Split Airport is not uniformly distributed during the year which leads to operational implication in terms of planning and traffic realization.
Another specificity of the Split Airport is daily peak load which is approximately 2% of annual passenger traffic and 1% of annual aircraft operations at the airport. Daily peaks at airport occur during the summer season every Saturday. During peak days at the airport over 26,000 passengers are handled.

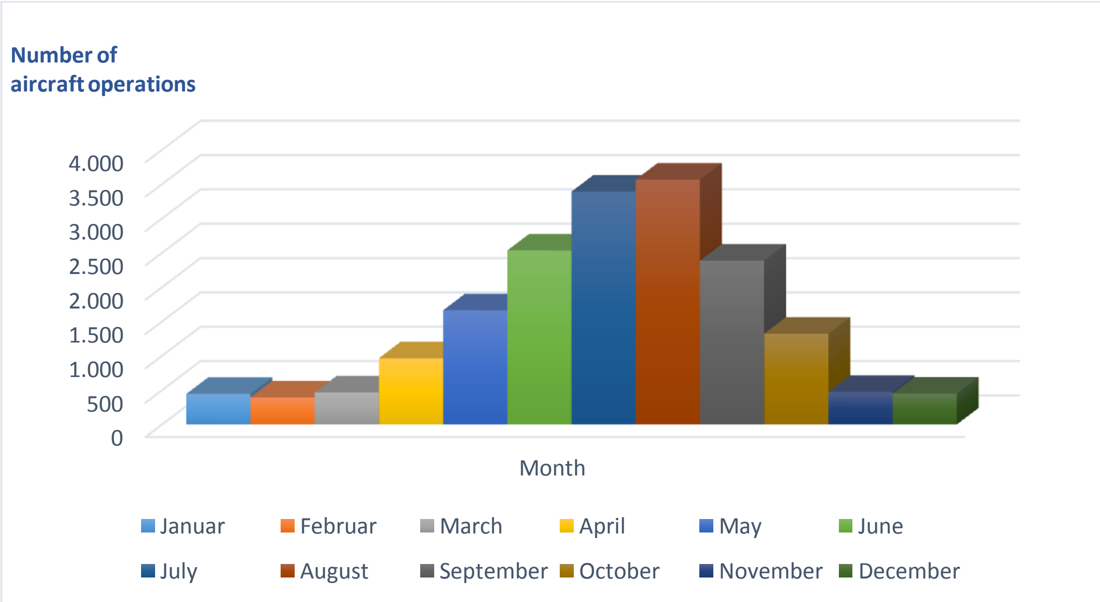



**Fig. 4**
*Aircraft operation at Split Airport (January – December 2013)*
*Source: Split Airport Statistic Report*

### 3.2 Schedule facilitated airport

As the airport with high traffic demand in the summer season, Split Airport encountered infrastructure limitations based on runway capacity for the first time in 2006. Since then, the Split Airport is a Schedule Facilitated Airport (IATA Level 2) for the summer season from March until October. All scheduled air traffic, charter flights, general aviation and business aviation are subject to schedule facilitation made by Traffic Coordinator or Ground Operation Centre.
On the IATA level 2 airport congestion occurs within a certain period of time (day, week, month, and season) and in that period additional coordination between airport and airlines is required in order to create the schedule. The role of airport, together with the scheduled facilitator and airlines is to create such a schedule of aircraft operations at the airport to avoid congestion (Fig 5).

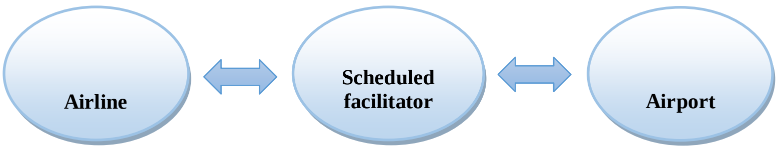



**Fig. 5**
*Schedules Facilitated Airport (IATA level 2)*

This procedure is performed in such a way that the airport advises the schedule time to the airline in which the airline operator voluntarily agrees to avoid congestions or delays caused by the airport constraint. With the schedule-facilitated process the Split Airport increased the annual traffic by 70% in passenger flows from 2006 to 2013.

### 3.3 Network impact

The European Organisation for the Safety of Air Navigation (EUROCONTROL) in 2013 drafted a list of important airports for the European air traffic network. The goal of this list was to define the airports within the European air traffic network whose specific characteristic could have significant impact on the performance of the operations within the network. Table 1 provides a list of the top thirty airports important for the European air traffic network.

**Table 1**
*Top 30 Airports important to the European network, base list summer 2013*

| ADES | AIRPORT_NAME | APT COORD | TOTAL IFR_ARR | DLY_ARR RATIO | SEASONALITY | TOTAL |
|---|---|---|---|---|---|---|
| LGSA | KHANIA/SOUDA | 0.75 | 0.27 | 3.955 | 1.575 | 6.55 |
| LGMK | MIKONOS | 0.75 | 0.27 | 2.825 | 2.25 | 6.095 |
| LGZA | ZAKINTHOS | 0.75 | 0.27 | 2.825 | 2.25 | 6.095 |
| LGSK | SKIATHOS | 0.75 | 0.27 | 1.695 | 2.25 | 4.965 |
| LFMD | CANNES MANDELIEU | 0.75 | 0.27 | 2.26 | 1.575 | 4.855 |
| LGIR | IRAKLION NIKOS KAZANTZAKIS | 0.75 | 0.675 | 1.13 | 2.25 | 4.805 |
| LFKF | FIGARI | 0 | 0.27 | 2.26 | 2.25 | 4.78 |
| LEIB | IBIZA | 0.675 | 0.675 | 1.13 | 2.25 | 4.73 |
| LGRP | DIAGORAS | 0.75 | 0.405 | 1.13 | 2.25 | 4.535 |
| LEPA | PALMA DE MALLORCA | 0.75 | 0.945 | 1.695 | 1.125 | 4.515 |
| LGKF | KEFALLINIA | 0.75 | 0.27 | 1.13 | 2.25 | 4.4 |
| LGKO | KOS | 0.75 | 0.27 | 1.13 | 2.25 | 4.4 |
| LGPZ | PREVEZA LEVKAS AKTIO | 0.75 | 0.27 | 1.13 | 2.25 | 4.4 |
| LGSR | SANTORINI | 0.75 | 0.27 | 1.13 | 2.25 | 4.4 |
| LGKL | KALAMATA | 0.75 | 0.27 | 0.565 | 2.25 | 3.835 |
| LGKR | IOANNIS KAPODISTRIAS | 0.75 | 0.27 | 0.565 | 2.25 | 3.835 |
| EGLL | LONDON HEATHROW | 0.75 | 1.215 | 1.695 | 0.113 | 3.773 |
| LCPH | PAPHOS | 0.225 | 0.27 | 2.825 | 0.45 | 3.77 |
| LEMH | MAHON | 0.675 | 0.27 | 0.565 | 2.25 | 3.76 |
| LTAI | ANTALYA | 0.675 | 0.945 | 0.565 | 1.575 | 3.76 |
| LSZH | ZURICH | 0.75 | 0.945 | 1.695 | 0.113 | 3.503 |
| LFKC | CALVI STE CATHERINE | 0 | 0.27 | 1.13 | 2.025 | 3.425 |
| LGKV | MEGAS ALEXANDROS | 0.75 | 0.27 | 0.283 | 2.025 | 3.328 |
| LDSP | SPLIT | 0.15 | 0.27 | 0.565 | 2.25 | 3.235 |
| LFMN | NICE | 0.75 | 0.675 | 1.13 | 0.675 | 3.23 |
| LIEO | OLBIA COSTA SMERALDA | 0 | 0.405 | 0.565 | 2.25 | 3.22 |
| EHAM | AMSTERDAM | 0.75 | 1.215 | 1.13 | 0.113 | 3.208 |
| LGRX | ARAXOS/PATROS | 0.75 | 0.135 | 0.283 | 2.025 | 3.193 |
| LTFE | MILAS/BODRUM | 0.225 | 0.405 | 0.283 | 2.25 | 3.163 |
| LTFG | ANTALYA-GAZIPASA | 0 | 0.27 | 0.565 | 2.25 | 3.085 |

*Source: Airports being Important to the Network, Action Paper, EUROCONTROL*

Airport entry list important for the European air traffic network is made based on four conditions. For an individual airport to be able to enter the list it must meet at least one of four conditions. The first condition is that a community airport has more than 150,000 Commercial Air Transport Movements per year (CATMs/year); the second condition includes all coordinated (IATA level 3) and scheduled facilitated (IATA level 2) airports with more than 50,000 CATMs/year. The third condition includes non-coordinated and non-schedules facilitated airports with more than 50,000 CATMs/year as well as coordinated and schedules facilitated airports that realized between 1,000 and 50,000 CATMs/year. The fourth condition includes Non-community ECAC airports (European Civil Aviation Conference) with more than 10,000 CATMs/year as well as coordinated and schedules facilitated Non-community ECAC airports.

After the formation of the airport entry list, airport selection is performed based on four Root Criteria. Each Root Criterion has a different weight factor. Summation of individual weight factor gives the total weight factor, based on which airports take a particular place on the list. The first Root Criterion is Airport Coordination Category representing IATA (International Air Transport Association) coordination level for each airport. The maximum weight factor that could be achieved for this criterion is 0.75. The second Root Criterion is the total number of IFR operations (airport class) and the maximum weight factor that could be achieved is 1.35. The third criterion is the proportion of delays in relation to the number of arrival IFR operations for a certain period and the maximum weight factor that could be achieved is 5.65. The last criterion is the seasonality, the distribution of traffic at the airport during the year. According to the season (summer or winter) some of them are subject to peaks. The maximum weight factor that could be achieved for this criterion is 2.25.

From the perspective of the European air traffic the Split Airport belongs to the category of smaller airports but its impact on the network is important. As pointed out, its impact on the network is primarily reflected in the seasonal character of the airport, but also in a certain number of delays that occur during the season. Although the number of delays considering the number of arrival IFR operations is relatively low, it is necessary in the future to create a predictive model that will provide better operation planning on the airport and thus minimize the number of delays.

## 4. Proposed model

The current situation with decentralised management results in inefficient use of the airports capacity. Deployment of the airport and ATC slot consistency will be necessary due to the upcoming airport regulation closing the gap to the SES (Single European Sky) legislation. ATM and Airports have developed different IT technology solutions which continuously and separately give them opportunity for possibly increasing their own efficiency and capacity. The strategy is to contribute to efficient solutions to optimize both capacities of airports and airspace network in the interest of all parties involved in the aviation industry. Airport Collaborative Decision Making (A-CDM) is one of the solutions which has been implemented already at ten airports with good impact on improvement of airport operations. Airport Collaborative Decision Making is a concept that has been designed by EUROCONTROL and represents one of the five measures in the Flight Efficiency Plan published by IATA, CANSO (Civil Air Navigation Services Organisation) and EUROCONTROL. A-CDM is an innovative concept of proactive decision-making in air traffic system. The main goal of A-CDM is to replace the current centralized system of air traffic management with collaborative decision-making in respect to airport airside operations. To establish such a system it is necessary to involve all stakeholders in the air transport system. The A-CDM concept consists of seven core elements: A-CDM Information Sharing, CDM Turn-round Process – Milestones Approach, Variable Taxi Time Calculation, Collaborative Management of Flight Updates, Collaborative Pre-departure Sequence, CDM in Adverse Conditions, Advanced Concept Elements. At the pre-tactical level, A-CDM concept seeks to improve aircraft operation day-to-day planning process. The A-CDM concept at the tactical level strives to ensure accurate information to all participants in the air transport system and generally improve information flow and predictability of events. The objective of the Airport CDM implementation is to improve the overall efficiency of operations at an airport, with a particular focus on the aircraft turn-around procedures which would be achieved through the enhancement of the decision-making process by sharing of relevant up-to-date information.

Since 2013, Split Airport has become a member of EUACA (European Airport Coordinators Association), the association which includes coordinators and schedules facilitators as members who are jointly responsible for allocating slots or advising schedule timings at more than 100 European airports. EUACA voluntarily created and funded a combined database collecting the data from each coordinator and schedules facilitator of the EU and of Norway, Iceland and Switzerland. The EUACA combined database is now holding the data for 190 airports, updated every night, which represents probably the most accurate centralized source of information concerning the traffic expected for the current and upcoming seasons in the European network of airports. EUACA signed in 2010 a memorandum of cooperation with EUROCONTROL in order to provide them with airport planned activity data through this single channel. Since 2012 EUROCONTROL has been using these data for the traffic forecast of the next few days/weeks for the entire airspace under its responsibility (broader than the EU) and this source of information has proved to be more accurate than the previous methodology. Split Airport has established interface between its own database and EUACA combined database to provide information to EUROCONTROL on daily basis. Presently, database sharing is going through the FTP server automatically with all relevant information about the planned flights, flight operator, aircraft type, arrival and departure destinations and times, etc. (Fig 6).

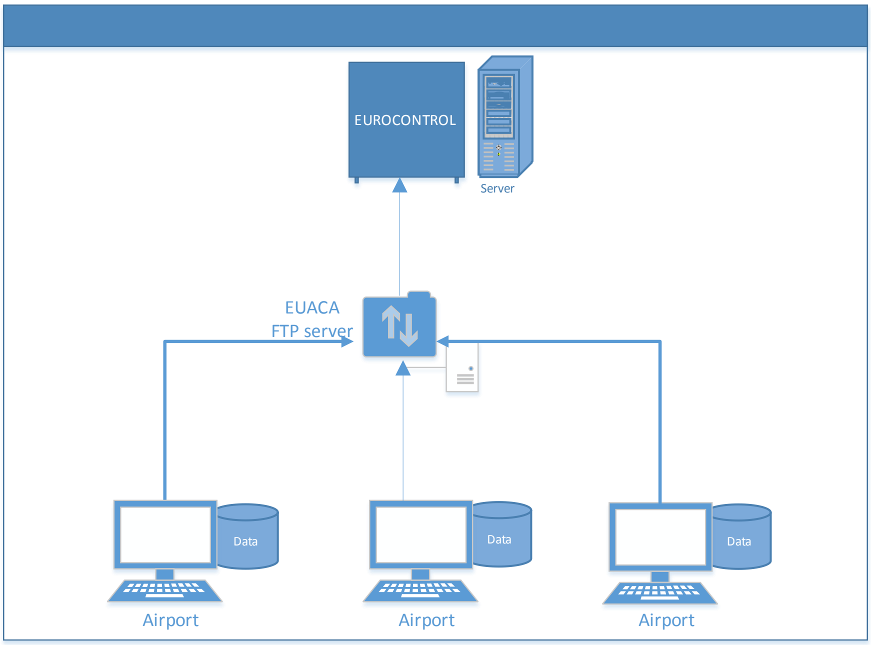



**Fig. 6**
*Proposed database sharing model*

This allows EUROCONTROL more precise planning of resources and provides the airspace capacity where and when it is needed and ensures better predictability through enhancing the overall Network planning processes.The Airport still needs to design and develop dynamic airport software data collection in IATA XML standard which has to be shared through the established interface to achieve joint objectives.

## 5. Conclusion

Delays in the air traffic system are a major problem which reduces or in some cases disables normal air traffic flow. The cause of the mentioned problem is inadequate coordination and cooperation between all participants in the air transport system, especially between airports and ATM. Statistics based on Europe indicates that the average delay in the European air transport network is 9.3 minutes per delayed flight causing for the participants in air transport damage in millions of euro. The aircraft delay significantly affects airports with a high level of traffic as well as airports with non-uniformly distributed traffic. The implementation of slots at the airports represents one of the possible solutions. Practice has shown that the implementation of this solution does not completely solve the problem. Therefore, the work has presented a predictive model based on data-sharing primarily between airports and air navigation service providers at pre-tactical level. Transparent sharing of accurate information on the planned traffic at a particular airport can significantly improve the planning system which EUROCONTROL performed at the level of the European airspace network. Improved planning on the pre-tactical level could significantly reduce the number of delays at the tactical level.

**References**

Bračić, M; Melvan, M; Pavlin, S; Steiner, S. 2014. Implementation Assessment of Airport Collaborative Decision Making at Split Airport, In *Proceedings of the 6th International Maritime Science Conference*, 470-476.

Eriksen, P.; Bretigny-sur, O. 2002. Collaborative Decision Making Information Management in Airports. In *Proceedings of the 21st Digital Avionics Systems Conference (vol.2.), 7B4-1 - 7B4-8.*

EUROCONTROL - Performance Review Report. An assessment of Air Traffic Management in Europe during the calendar year 2013. Available from internet: <https://www.eurocontrol.int/sites/default/files/content/documents/single-sky/pru/news-related/2014-03-06-final-draft-PRR-2013.pdf >.

EUROCONTROL - The Central Office for Delay Analysis (CODA). *Delays to Air Transport in Europe – Annual 2013*. Available from internet: <http://www.eurocontrol.int/sites/default/files/content/documents/official-documents/facts-and-figures/coda-reports/coda-digest-annual-2013.pdf >.

EUROCONTROL - Airport Operations Team. 2013. *Methodology for Airports being Important to the Network.* Action Paper, AOT/03/AP06.

IATA - International Air Transport Association. Worldwide Slot Guidelines. Available from internet: <https://www.iata.org/policy/slots/Documents/wsg-6.pdf >.

Katsaros, A; Psaraki-Kalouptsidi, V. 2011. Impact of collaborative decision-making mechanisms on operational efficiency of congested airports. *Airport Management* 5(4): 351–367.

Koeners, G.J.M.; Rademaker, R.M. 2012. Ground Control Support Functions to Optimize Surface Traffic Flow in a CDM Environment. In *Proceedings of the 31st Digital Avionics Systems Conference, 4E6-1 - 4E6-11.*

Research and Innovative Technology Administration Bureau of Transport Statistic database. 2013. USA. Available from internet: <http://www.transtats.bts.gov/ot_delay/OT_DelayCause1.asp?pn=1>.

Steiner, S; Štimac, I; Melvan, M. 2014. Towards to Collaborative Air Traffic and Airport Management. In *Proceedings of the 22nd International Symposium on Electronics in Transport*, A1.

Šimecki, A; Steiner, S; Čokorilo, O. 2013. The Accessibility assessment of regional transport network in the South East Europe, *International Journal for Traffic and Transport Engineering*, 3(4): 351–364.

# ANALYSIS OF AIRCRAFTS KINEMATIC CHARACTERISTICS FOR PAVEMENT DESIGN

**Mauro D'Apuzzo [1], Bruna Festa[2], Vittorio Nicolosi [3], Antonella Schibani [4], Giulia Collorafi [5]**
[1] *Department of Civil Mechanical Engineering, University of Cassino and Southern Lazio , Cassino , Italy.*
[2] *Department of Civil Architectural and Environmental Engineering, University of Naples "Federico II" , Naples , Italy.*
[3] *Department of Civil and Computer Science Engineering. University of Rome "Tor Vergata", Rome , Italy.*
[4] *Italian Air Force*
[5] *Engineer*

**Abstract:** Runway pavement design and maintenance are a demanding task nowadays both for civil and military airport authorities, and for managing authorities. Mechanistic and empirical mechanistic approaches require a more accurate assessment of actions exerted by aircraft landing gears. Therefore the evaluation of aircraft positioning on runway and the analysis of aircrafts' kinematic characteristics in landing and take-off operations appears fundamental. Following an experimental investigation carried out on an Italian airport, aircraft kinematic data were collected using several video cameras. Landing and take-off procedures were reviewed and a statistical analysis was performed to model touchdown points' longitudinal and transversal distribution. Preliminary comparison with data derived by technical literature has shown a good agreement for the lateral touchdown points' distribution..

**Keywords**: Touchdown points pattern, aircraft kinematic analysis, actions on pavement

## 1. Introduction

Air transportation is one of the fastest growing modes of transport, in 2013, over 33 million departures had taken place worldwide and global air traffic is expected to grow at an annual rate of 5% (from ICAO). At the same time, new airplanes are equipped with landing gears that stress pavements more than the oldest one as from the end of 20$^{th}$ century the manufacturers are focusing on designing airplane gear configurations to optimize the efficiency of the airplane. Because of this change in philosophy and the increasing airplane weights, airfield pavements must be designed to withstand increased loading conditions.

Historical pavement design guides were based on methods of analysis that resulted from empirical research and field performance. Although it may have been possible to adjust these methods to address different gear configurations, if new airplanes, with a unique gear configuration, are added to the traffic mixture at a facility, it may be impossible to assess the impact of them using the empirical design procedures. Therefore newer design procedures, based on layered elastic or finite element analysis, were adopted until the end of nineties to better address the impact of new landing gear configurations and increased pavement load conditions, for example the method introduced by the Federal Aviation Administration (FAA) (AC n.150/5320-6E). The new design processes involve a large number of interacting variables, which are often difficult to quantify, one of these is traffic load action. The "design aircraft" concept, used in the empirical methods, has been replaced by design for fatigue failure expressed in terms of a cumulative damage factor (CDF) using Miner's rule, CDF is expressed as the ratio of applied load repetitions to allowable load repetitions to failure. Load repetition are expressed in term of coverage, which represent the number of times a point in the pavement is expected to be stressed as a result of a given number of aircraft operations (Zhao & Ling 2008). Coverages are function of number of aircraft passes, airplane gear configurations, tire contact area, lateral distribution of aircraft wheel path relative to runway, or taxiway, centerline, and touchdown point. The accurate evaluation of coverages is a critical aspect also for the development of new mechanistic-oriented degradation models for airport pavement manteinance within a conventionale Airport Pavement Management System (APMS) (D'Apuzzo, Giuliana Festa,Mancini and Nicolosi 2012),

Some researches were carried out on lateral distribution of aircraft wheel path, in the following table (Table 1) some of the most relevant studies are summarized.

These studies show that the lateral aircraft distributions are becoming narrower because of development in both equipment for landing aids and aircraft since the first investigations, furthermore most of studies deal with lateral aircraft deviations during taxiway operations while less attention was devoted at runway operations and longitudinal touchdown point position.

In the study presented in this paper an investigation was carried out in two Italian airport for measuring lateral and longitudinal touchdown point distribution following a previous preliminary investigation done at Ciampino airport (D'Apuzzo 2014). The data were analyzed and the results were compared with some indication reported in the literature.

[1] Corresponding author: dapuzzo@unicas.it

**Table 1:**
*Relevant studies on distribution of aircraft wheel path relative to runway, or taxiway, centerline, and touchdown point.*

| Title | Reference | Description |
|---|---|---|
| Evaluation of Existing Airfields for C-5A Operations | Brown & Thompson 1973 | This study preset the results of observations made on some test sections and also makes reference to studies conducted in 1949 and field observations carried out in 1956 for B-47 (Stratojet), B-52, KC-87, and KC-135 aircraft. The concept of wander was introduced and defined as “the maximum lateral movement of a point on the centerline of an aircraft about the centerline (or guideline) on taxiways or runways during operation of the aircraft.” The lateral deviations from the taxiway and runway centerlines, for aircraft considered, was approximately 0.975 m and 10.27 m, respectively, during 75 percent of the time. |
| Field Survey and Analysis of Aircraft Distribution on Airport Pavements | HoSang 1975 | Airplane wander pattern were collected from 9 airports and 12 types of commercial aircraft with maximum gross weights above 18500 kg, including: Boeing 747, 707, 727 and 737; McDonnell-Douglas DC-10, DC-8 and DC-9; Lockheed L-1101; General Dynamics Convair 880 and 580; British Aircraft Corporation BAC-111; and Nihon YS-11. The standard deviations of wander for individual aircraft types, compared at the various airports, varied from 0.91 m to 2.43 m for takeoffs and from 1.22 to 2.74 m for landings. The mean offsets were consistently to the right of the pavement centerline. The mean offset was 0.64 m and 0.975 m respectively on the taxiway large 22.86 m and 31.09 m. The standard deviations on the 22.86 m taxiways ranged between 0.762 and 1.22 m, and those on the 31.09 m taxiway were generally wider, with an overall average of about 1.83 m. Another conclusion of the study was that data were closer to a normal distribution. |
| Wander Patterns for Commercial Aircraft at Denver International Airport | Rufino, Roesler, Tutumluer, Barenberg 2001. | Rufino et al. carried out a new study of wander patterns for three different commercial aircraft (B-777, B-727, and DC-10) operating at Denver International Airport. This study confirmed that the lateral distribution of aircraft on runways follows a normal distribution rather than a uniform distribution. |
| Statistical Extreme Value Analysis of ANC Taxiway Centerline Deviations for 747 Aircraft | Scholz 2003 | This report describes the analysis of 747 taxiway centerline deviation data (9796 events) that were collected from 2000 to 2001 at Anchorage International Airport. Deviations were measured for nose and main gear at two laser locations for each of two 22.86 m straight taxiway segments with shoulder. The mean offset measured range from 0.25 m to 0.04 m. |
| Preliminary Analysis of Taxiway Deviation Data and Estimates of Airplane Wingtip Collision Probability | Cohen-Nir & Marchi 2003 | The report describes analysis performed on taxiing airplane data collected by the Federal Aviation Administration in Anchorage International Airport (ANC) and New York John F. Kennedy International Airport (JFK). Data were collected for B747-100/-200/-300/-400 and SP. The mean offsets range from -0.01 to 0.16 m for (ANC) and from -0.16 to 0.10 for (JFK). The standard deviations of wander range from 0.63 to 0.56 m for (ANC), and from 0.62 to 0.64 for (JFK). The sampled data followed a negative exponential type behavior, particularly the best fit for each data set was achieved with the Logistic distribution. |
| Effect of field aircraft lateral distribution pattern on required thickness by FAA design method. | Wang, Chou and Cheng 2008. | Wang et al. placed several types of sensors in concrete slabs on the taxiway of Taoyuan International Airport to observe the aircraft lateral distribution pattern. They found that aircraft mainly moved quite along the centerline but tended to have a shift to the left of center point. The measured standard deviation of the wandering pattern was 50.4 cm. This standard deviation value is much smaller than what was found in the monitoring project conducted by FAA in 1975 |
| Some remarks on landing kinematic characteristics of aircrafts | D’Apuzzo, Festa, Nicolosi and Schibani 2014 | In this study lateral and longitudinal touchdown point distribution were measured in Rome Ciampino airport. The data on lateral touchdown point showed a mean value with an offset, on the left, of 0.50m, and a standard deviations of 0.61m; no differences were found between different type of aircraft. While the distributions of longitudinal touchdown points of the five most frequent types of aircraft were different. The lateral and longitudinal touchdown point distributions were modelled respectively by a log distribution and a log-logistic distribution. |

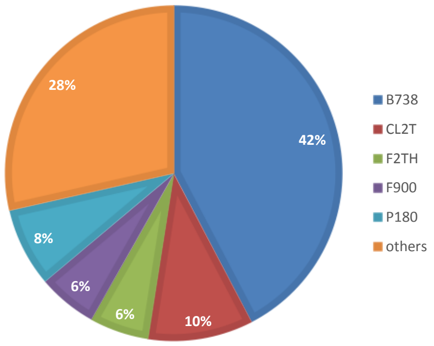
## 2. Experimental data collection and analysis

The experimental study was performed on the runways of two Italian airport: Rome Ciampino G.B. Pastine Airport (RCA), and Naples Capodichino International airport (NCA). The geometric characteristics, the orientation and the traffic volume of the two runways observed are summarized in Table 2. In the traffic spectra of RCA there are about 70 aircraft types, but the 80% of traffic volume observed is due only to 5 aircraft types, as shown in Figure 1a, the same thing happens in the in the NCA airport were the 73% of movements are due to 5 types of aircrafts (Airbus 319, 320 and 321, Boeing 737-700 “73W” and 737-800 “73H” see Figure 1b).

**Table 2:**
*Characteristics of two runways investigated*

| | Rome Ciampino G.B. Pastine Airport | Naples Capodichino International airport |
|---|---|---|
| **Runway length [m]** | 2207 | 2628 |
| **Runway with [m]** | 45 | 45 |
| **Runway orientation** | 15-33 | 06-24 |
| **Traffic volume [movements/year] (ENAC 2013)** | 46266 | 48704 |

a)




b)

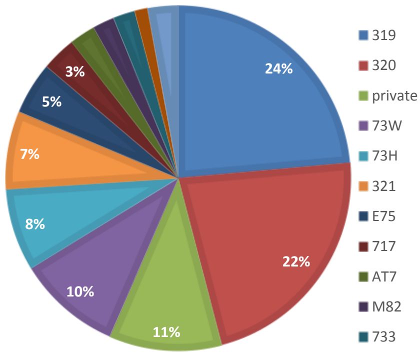
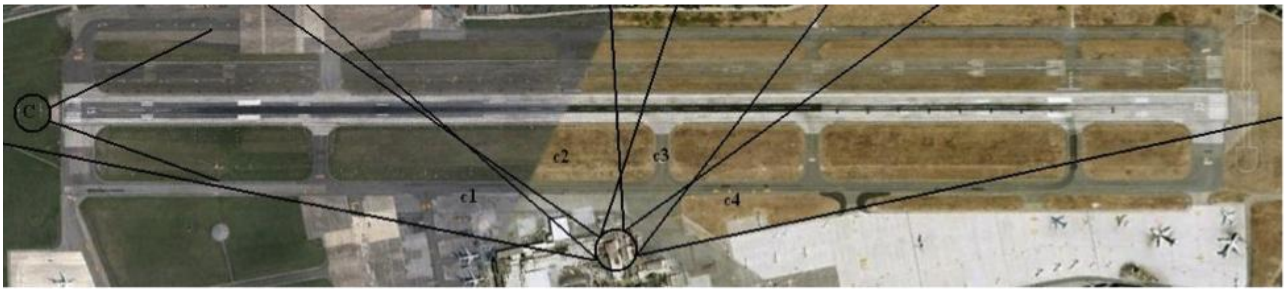



**Fig. 1.**
*Traffic spectra: a) Rome Ciampino G.B. Pastine Airport, b) Naples Capodichino International airport*

Data, on lateral distribution of aircraft wheel path and on longitudinal distribution of touchdown points, were collected by a video detection system. Four high-definition video cameras were placed on the control tower (on the old control tower in NCA), to record landing and take-off operations, and one video camera was placed on the runway end (heading 15/33) of Rome Ciampino Airport, to record aircraft lateral position during landing and take-off operations from 15 to 33 head, which is the direction almost exclusively used (see Figure 2). In the NCA was not possible place a video camera on the runway.




a)


b)

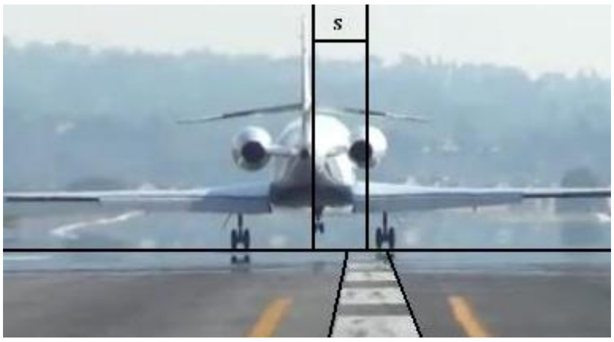
**Fig. 2.**

*Video camera position and view: a) position of cameras in RCA, b) position and view of cameras in NCA.*

### 2.1. Data collection and analysis

In Roma Ciampino airport the data were collected in 15 days and 530 operations (55 departures and 475 landings) were registered. In Naples Capodichino airport data were collected for 10 days and 201 touchdown points were registered. The lateral position of the touchdown points was determined through the video camera placed at the runway end, while their longitudinal distribution was deducted from the elaborations of the film recorded by the video cameras placed on the control tower. The data were recorded for each of the five significant types of aircraft in each airport. The longitudinal position of the touch down points were identified by the distance from the start of the runway threshold marking (see figure 3).


b)




a)

**Fig. 3.**

*System for identify the lateral and longitudinal position of touch down points: a) Cartesian coordinate system used to locate the aircrafts on the runways, b) view of C camera in RCA and measure of lateral position "s"*

### 3. Lateral distribution of aircraft in landing and takeoff operations

The data about later wandering of touchdown points (landing) and starting points (take off) recorded on the runway 15/33 of Roma Ciampino airport were analyzed. First the aircrafts were grouped according the categories suggested by ICAO (ICAO 2009), in order to test if the distributions of lateral position are different for different type of aircrafts. The non-parametric test of Kruskal–Wallis, was carried out for testing whether samples, measured for the 3 groups of aircraft, originate from the same distribution. According to this test there is no evidence of differences between the samples, which means that the lateral wandering measured for all aircraft types can be considered originating from the same distribution, both in landing and take off operations (see table 3).

**Table 3:**
*Results of Kruskal-Wallis test about homogeneity of lateral distributions, in landing and takeoff movements, of different type of aircrafts, grouped according to ICAO in the Rome Ciampino Airport.*

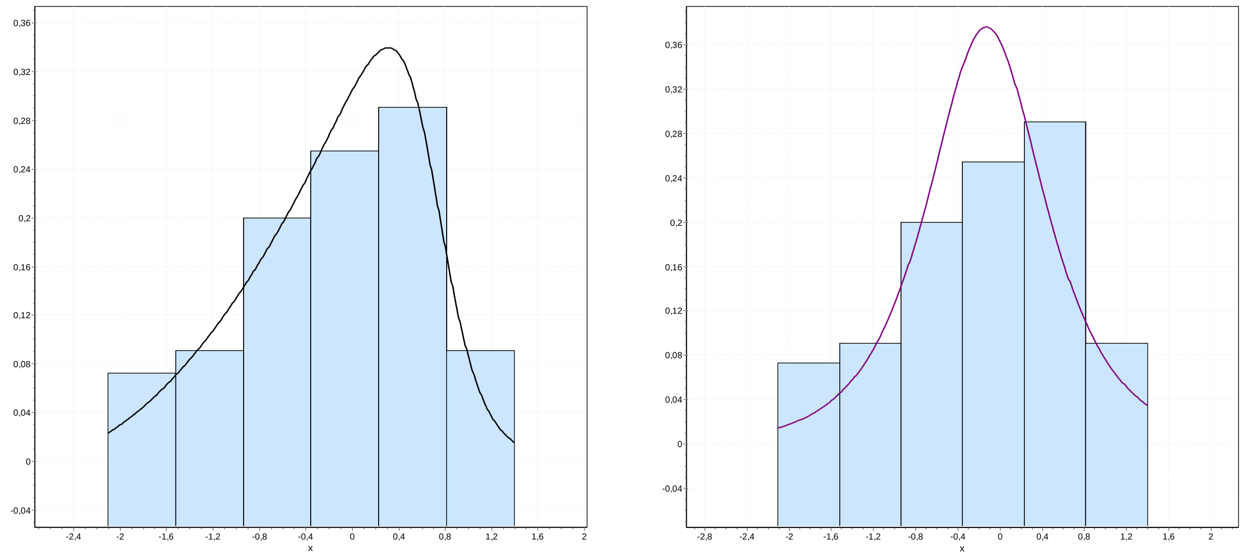
| | **Touchdown points (landing)** | | | **Starting points (takeoff)** | | |
|---|---|---|---|---|---|---|
| **ICAO GROUPs** | **A** | **B** | **C** | **A** | **B** | ***C*** |
| **Sum of ranks, Ri** | 20625 | 20157 | 62503 | 239 | 507 | 794 |
| **Number ni** | 87 | 94 | 273 | 8 | 16 | 31 |
| **Average of ranks** | 237 | 214 | 229 | 30 | 32 | 26 |
| **Sum of all ranks** | 103285 | | | 1540 | | |
| **Degrees of freedom** | 2 | | | 2 | | |
| **Total number of observations across all groups** | 454 | | | 55 | | |
| **Test statistic** | 1.428 ≤ $\chi^2$**(0.05)**=5.991 | | | 1.645 ≤ $\chi^2$**(0.05)**=5.991 | | |

Therefore total aircraft operations data were analyzed (see table 4) and it was found that aircrafts tended to have a shift to the left of center line, as matter of fact the mean position of aircraft axis was - 0.128 m and -0.381m, respectively for takeoff and landing operations, as already observed in others researches (Cohen-Nir & Marchi, 2003; Wang, Chou and Cheng, 2008). The standard deviation of the wandering measured for takeoff operations is close to the values recently measured on some taxiway (see Wang, Chou and Cheng 2008; Cohen-Nir & Marchi 2003), but it is less than values measured in FAA's experiments in 1975 (HoSang 1975). For landing operations it was measured a standard deviation of lateral wander which is close to the minimum value reported by the in FAA's experiments (HoSang 1975).

**Table 4:**
*Mean, standard deviation and median of lateral wonder measure measured on runway 15 of Rome Ciampino Airport for take off and landing operation*

| | **Takeoff** | **Landing** |
|---|---|---|
| **Mean** | -0,12816 | -0,38068 |
| **Standard Deviation** | 0,777259 | 1,276684 |
| **Median** | 0,038 | -0,35315 |

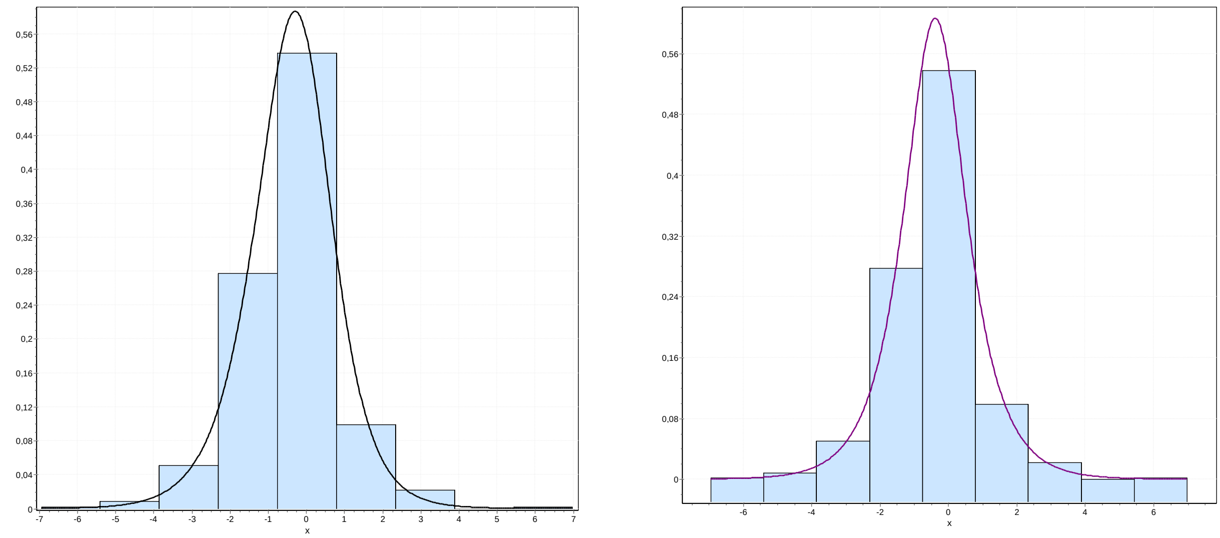
Furthermore about 50 different probability distributions were examined and fitted using the Maximum Likelihood Estimates (MLE). The Goodness of Fit (GoF) tests of Kolmogorov-Smirnov, Anderson-Darling and Chi-Squared were carried out to measure the compatibility of the samples with the theoretical probability distribution function, to compare the fitted distributions and to select one of the models. The analysis shown that the Dagum distribution (4 parameters) describes the data in the best way but the more simple (2 parameters) Hyperbolic Secant distributionis is acceptable too, at a level of significance of 20% (see figures 4 and 5). The Normal distribution, suggested by some authors, does not satisfy the Chi-Squared and Anderson-Darling (level of significance 5%) tests for later wander data of landing operations.



a) b)

**Fig. 4.**

*Fitting of takeoff wander data to: a) Dagum distribution, b) Hyperbolic Secant distributions .*



a) b)

**Fig. 5.**

*Fitting of landing wander data to: a) Dagum distribution, b) Hyperbolic Secant distributions.*

The Hyperbolic Secant distributionis is a continuous probability distribution, placed in the class of logistic distribution, it shares many property with normal distribution (i.e. symmetric, mean=media=mode) but has heavier tails (leptokurtic). The probability density function is:

$$f(x) = \frac{senh\left(\dfrac{\pi \cdot (x-\mu)}{2\sigma}\right)}{2\sigma} \tag{1}$$

and the cumulative distribution function is:

$$F(x) = 2 \cdot \pi \cdot \arctan\left(\exp\left(\frac{\pi \cdot (x-\mu)}{2\sigma}\right)\right) \tag{2}$$

Where

$\mu$ is the mean and $\sigma^2$ is the variance.

## 2. Aircraft touchdown points longitudinal distribution

The longitudinal touchdown points distributions, measured both in RCA and NCA, were analyzed. In the first phase the non parametric Kruskal-Wallis method (since it not assume normal distribution) was used for testing whether touchdown points distribution were different:

a) for different types of aircraft in the same airport;
b) for the two runway

As already done for lateral wandering analysis, the aircrafts were grouped according the categories suggested by ICAO (ICAO 2009), in order to test if the longitudinal touchdown points distributions are different for different type of aircrafts. As ideally more than five data points per group are needed to carry on the test, in the NCA the aircraft type categories A and B were joint together. The results of test (see table 5) showed that there are not statistical differences in the distributions of different aircraft types for the prevailing landing direction; while non homogeneous distribution was found in NCA for the runways 06.

**Table 5:**
*Results of Kruskal-Wallis test about homogeneity of longitudinal touchdown points distributions of different type of aircrafts.*

| | Rome Ciampino Airport | | | Naples Capodichino International airport – runway 24 | | Naples Capodichino International airport – runway 06 | |
|---|---|---|---|---|---|---|---|
| **ICAO GROUPs** | **A** | **B** | **A/B** | **A/B** | **C** | **A/B** | **C** |
| **Sum of ranks, Ri** | 22412 | 24870 | 2253 | 17847 | 65768 | 62 | 884 |
| **Number ni** | 91 | 96 | 20 | 180 | 288 | 8 | 35 |
| **Average of ranks** | 246,28 | 259,06 | 112,65 | 99,15 | 228,36 | 7,75 | 25,257 |
| **Sum of all ranks** | 113050 | | | 20100 | | 946 | |
| **Degrees of freedom** | 2 | | | 1 | | 2 | |
| **Total number of observations across all groups** | 475 | | | 200 | | 43 | |
| **Test statistic =** | 4,01 ≤ □$^2$**(0.05)**=5,99 | | | 0,98 ≤□□□ **(0.05)**=3,84 | | 12,658 ≥□□□ **(0.05)**=3,84 | |

The different distribution in the runway 06 of NCA is caused by the orographic conformation with an hill which influence the landing trajectory of aircrafts. The influence is most significant for the group C than for than A and B, as matter of fact the mean longitudinal position of touch point was 587m for group A/B aircraft and 759m for the group C while in the runway 24 was 390m and 365m.

Furthermore the Kruskal-Wallis test indicated that the longitudinal touchdown point distributions of the two airports were statistically different, but if data are shifted to the mean values (i.e. the $y_1$*=y-□$_1$ and $y_2$*=y-□$_2$ variable are considered) the two distribution are homogeneous. Therefore the longitudinal distributions of touchdown points do not change but they are shifted because the airport may have different landing approach paths based on topography, buildings, or other considerations. Anyway, the data of RC Airport and NC Airport were analyzed separately in order to fit a probability distribution, but all aircraft categories were put together. More than sixty probability distribution were tested and the analysis carried out showed that samples are well modelled by a shifted log-logistic distribution, whose density function is:

$$f(x)=\frac{\alpha\cdot\left(\frac{(x-\gamma)}{\beta}\right)^{\alpha-1}}{\beta\cdot\left[1+\left(\frac{(x-\gamma)}{\beta}\right)^{\alpha}\right]^{2}} \tag{3}$$

where □ is the location parameter, □ is the shape parameter and □ is the scale parameter.

The values of the parameters, for the RCA distribution and NCA distribution are summarized in table 6.

**Table 6:**
*Values of log-logistic parameters distribution from fitting analysis.*

| **Parameters of log-logistic probability density function** | **Rome Ciampino Airport** | **Naples Capodichino International airport – runway 24** |
|---|---|---|
| □ | 3.9885 | 3.7194 |
| □ | 256.83 | 139.91 |
| □ | 179.22 | 210.31 |

### 3. Conclusion

The wander of aircraft patterns represent a basic information in the evaluation of the coverages, which is, in turn, fundamental both in the pavement design criteria and in the maintenance management. In this study a methodology to identify aircraft location during runway operations was presented, it is based on video detection system, which allowed for the determination of both later wander pattern (in landing and takeoff operations) and longitudinal touchdown wander. Lateral centerline deviation data and longitudinal touchdown location data were collected from year 2013 to 2014 on the runways of two Italian airports: Rome Ciampino G.B. Pastine Airport and Naples Capodichino International airport.

The analysis of experimental data collected showed that no statistical differences exists in lateral and longitudinal wonder of different aircraft type (aircraft were grouped as suggested by ICAO).

As far as lateral wander is concerned, this study confirmed some finding of the previous study:

a) aircrafts tended to have a shift to the left (mean was - 0.128 m in takeoff operations and -0.381m in landing operations);
b) standard deviation of the wandering for takeoff operations is close to the values recently measured on some taxiway, but is much smaller than what was found in the 1975 FAA project.

This study showed that lateral distribution of runways operations follows a hyperbolic secant distribution, rather than a normal distribution, as suggested in previous studies. Therefore the distribution is similar to normal distribution but has heavier tails (leptokurtic).

The longitudinal distributions of touchdown points were statistically homogeneous for the two airports examined, but they are shifted, probably because the airports may have different landing approach paths based on topography, buildings, or other considerations. The analysis carried out on the longitudinal wander of touchdown positions showed that sample are well modelled by a shifted log-logistic distribution.

The presented approach seems promising in providing useful information for the evaluation of coverages and the design and maintenance of airport pavement.

Analysis of additional data from other airports should be undertaken to confirm and expand upon the results reported here.

### Acknowledgements

The authors wish to acknowledge the support of Gennaro Bronzone Director of the Naples ENAC Office, Aniello Mattera and Mariella Aiello of GESAC for their assistance during data collection at the Naples Capodichino International Airport, and the Italian Air Force for their assistance during experimental measurements made at Ciampino Airport.

**References**

Boeing 2008. Statistical Summary of Commercial Jets Airline Accidents, WorldwideOperations 1959 – 2008. Boeing. Website: http://www.boeing.com/news /news/techissues/pdf/statsum.pdf. Accessed: Mar. 20, 2010.

Brown, D. N.; Thompson. O. O. 1973. *Lateral Distribution of Aircraft Traffic*. Miscellaneous Paper S-73-56. Vicksburg, MS: U.S. Army Engineer Waterways Experiment Station Soils and Pavements Laboratory.

Cohen-Nir, D.; Marchi, R. 2003. Preliminary Analysis of Taxiway Deviation Data and Estimates of Airplane Wingtip Collision Probability. In *Proceedings of the 82nd Transportation Research Board Annual Meeting, Washington D. C. USA, Paper Number TRB-03-3302*.

D'Apuzzo, M.; Festa, B.; Nicolosi, V.; Schibani, A. 2014. Some remarks on landing kinematic characteristics of aircrafts. In *Proceedings of Sustanability, eco-efficency and conservation in trasportation infrastructure asset management*. Taylor & Francis group (ISBN 978-1-138-00147-3), 385-392, doi: 10.1201/b16730-57.

M. D'Apuzzo, F. Bruna, G. Giuliana, L. Mancini, V. Nicolosi (2012). The evaluation of runway surface properties: a new approach. *PROCEDIA: SOCIAL & BEHAVIORAL SCIENCES*, vol. 53, p. 1193-1202, ISSN: 1877-0428, doi: 10.1016/j.sbspro.2012.09.968

ENAC 2013. *Rapporto e bilancio sociale 2013*. Ente Nazionale dell'Aviazione Civile. Roma Italy.

FAA 2009. Airport Pavement Design and Evaluation. Federal Aviation Administration, Advisory Circular No.150/5320-6E. Available from internet http://www.faa.gov.

Hall, Jr. Jim W.; Ayres, Jr. Manuel; Shirazi H.; Speir, R.; Carvalho, R.; David, R.; Huang, Yih-Ru 2011. *Risk Assessment Method to Support Modification of Airfield Separation Standards*. Airport Cooperative Research Program (ACRP) Report 51. National Academy of Sciences USA (ISSN 1935-9802 , ISBN 978-0-309-21332-5).

HoSang V. A. 1975. Field survey and analysis of aircraft distribution on airport pavement. Federal Aviation Administration Report FAA-RD-74-36. Available from internet http://www.airporttech.tc.faa.gov/Pavement/Downloads/rd74-36.pdf.

ICAO 2009. *Aerodrome design and operations*. International standards and recommended practices – aerodromes annex 14 vol. I fifth edition. International Civil Aviation Organization, Montreal Quebec, Canada, 2009 (ISBN 978-92-9231-332-6).

Rufino, D.; Roesler, J.; Tutumluer, E.; Barenberg, E. 2001. Wander Patterns for Commercial Aircraft at Denver International Airport. In *Proceedings of the 2001 Airfield Pavement Specialty Conference at the 27th International Air Transportation Conference "Advancing Airfield Pavements"*, pp. 158-17.

Scholz, F. 2003. Statistical Extreme Value Analysis of ANC Taxiway Centerline Deviations for 747 Aircraft. Federal Aviation Administration (FAA) and Boeing Company report. Available from internet http://www.faa.gov/airports/resources/publications/reports/media/ANC_747.pdf.

Wang S.Y.; Chou C.P.; Cheng H.J. 2008. Effect of field aircraft lateral distribution pattern on required thickness by FAA design method. In *Proceedings of the 87th Transportation Research Board Annual Meeting, Washington D. C. USA*.

Zhao, H; Ling, J. 2008. Discussion of coverage for airport asphalt pavement design. In *Proceedings of Plan, Built, and Manage transportation Infrastructure in China*, pp. 833-840.

# GROUND ACCESS MODE CHOICE AT JOŽE PUČNIK AIRPORT LJUBLJANA

**Boštjan Žlak**[1]
[1] *Faculty of Maritime studies and Transport Portoroz, Slovenia*

**Abstract:** Understanding the accessibility of an airport and choices airport passengers have to take when deciding on the departure airport is proving to be vital for understanding of the aviation system. Recently a number of papers have been published on the topic of airport catchment area and access mode choice.
Passengers base their decision of departure airport on various factors. Slovenia has just one international airport with regular daily international lines. Jože Pučnik Ljubljana airport is the only Slovenian international airport; however, it is located in the region with several international airports, which together form a multi-airport region within which airport catchment areas overlap. Due to this fact, it is necessary to understand passenger airport choice and factors determining it.
Since there is no existing study on passenger airport choice for the Jože Pučnik airport Ljubljana, this paper attempts to evaluate the access mode choice from the data gathered through a passenger survey.

**Keywords**: access mode choice, airport catchment area, airport competition

## 1. Introduction

The changes in market regulations have caused a change in how airports manage competition. As there was practically no competition between airports while markets were highly regulated, the deregulation of the markets caused open competition between airports. Airports suddenly had to compete for passengers living inside their respective catchment area. Since airports realized their competitive status in the region, they pay close attention of determining factors of passenger access mode choice of departing airport.

Passengers are facing a set of choices to determine their airport of origin. According to Pels (Eric Pels, 2003) and Hess (Hess and Polak, 2005) passengers evaluate the airport services, airline offers and the airport access mode. In addition, they also need to evaluate location of the airport, infrastructural capacity, waiting times at the airport (Ubogu, 2013), parking facilities, etc... Each of choices depends on a set of determining factors such as airfares, accessibility, access time, destination availability, frequency, etc… Combined these factors add to a final preferred choice of departing airport.

In this paper, focus is on access mode choice. We obtained data for preliminary analysis from a survey conducted at Jože Pučnik Airport in 2011 for the purposes of EU Adria-A project, the project designed to contribute to the accessibility and transport reorganization in the Italy-Slovenia cross-border area and used SPSS statistical tool for analysis.

Generally, choice of travel consists of four main dimensions (Hess, 2010) in reference to destination choice, time of travel, mode of travel and route choice. Given the assumption that passengers have chosen the destination and they chose air travel to be their mode of travel, they are now facing two sets of choices. First is the choice of departing airport or port of origin, and the second is access mode choice. In this paper, we are discussing only origin side of the travel and not destination side.

Sometimes passengers have only one option of departing airport, however in many cases passengers depart from a multi-airport area. Therefore, they have option to select an airport according to their preferences. Determining factors of passenger choice include airfare cost, access time, flight frequency, number of operating airlines ect. (Loo, 2008). Planning a journey can be a very complicated matter, since passengers have to take into consideration all of the aforementioned factors. However, this is not all they have to consider. The decision process at some stage has to include accessibility of the airport and access time to airport. As access time is a function of a chosen access mode in a form of public transport (bus, train, taxi, shuttle…) or private transport (car…), accessibility is a function of infrastructure connections to and from airport and with them connected public transport services.
In order to find what are the passengers' preferences, airports, airlines and researchers perform passenger surveys to gather data for forecast modeling. Results of such surveys offer insight into passenger needs and are a base for models for air travel demand forecasting, airport ground access mode choice and airport choice (Cripwell et al., 2009).
Often due to lack of proper infrastructure planning, combined with poor availability of transport services, access to airport is mainly dependent on private car use. Therefore understanding the passenger behavior is essential for proper infrastructure and public transport services planning.

**2. Access mode choice**

Ground access mode choice plays a major role in airport landside development planning. Airport planners need to understand passenger travel patterns in order to be able to plan proper changes to the airport access system and wider changes to the regional transportation system (ACRP, 2008). Larger metropolitan areas and Airports operating in those areas have a large developed network of different transportation modes with plenty of public passenger services ranging from bus, train, subway, shuttle, taxi and other services. Therefore, travelers have a great variety of choices when deciding on their choice of access mode. Their choice is a function of availability, access cost and prior experiences in using certain transport mode, convenience and so on. As Tam reports (Mei Ling Tam et al., 2005) over 70% of passengers arrived at airport Hong Kong using public transport and attributed this phenomena to travel cost, shorter time on public transport mode and lower car ownership. Another study suggests that changes in travel time have positive influence on private car use. Any increase in access time increases likelihood of choosing private car as mode of choice over other access modes. Researchers like Pels (Eric Pels, 2003) pointed out the effects of travel time and travel cost on ground access mode choice as their research shows these are significant variables in airport access mode choice.

When considering smaller metropolitan areas and airports located on outer edges or even outside major cities things change a bit. Often these airports connect to major cities through a very limited transport network, which results in traveler’s diminished motivation to use public transport services since in such cases the frequency of public transport services is expected to be significantly lower or in some cases non-existent. Travelers in those cases have very limited choices. For instance, low frequency of public transport services may cause a mode shift from those services to private car use and vice versa. Conversely, high fees of long-time parking at the airport may divert travelers from using this access mode and shift to either public transport service or to drop-off system, where another person drives the passenger to the airport and later comes to pick them up.

Additionally poor services and weak connectivity of airport to rural areas may in effect cause a passenger leakage from a certain airport to another airport operating in a region. As passenger leakage we understand the willingness of passengers to spend more time to drive to a larger metropolitan area with alternative airport where they have more choices of airlines and departure frequencies with lower airfares (Suzuki et al., 2003). We can safely assume that poor accessibility to an airport and bad public transport services may have a negative impact on airport choice.

**3. Jože Pučnik Ljubljana Airport – case study**

Jože Pučnik Ljubljana airport (LJU) is located 25km northeast of Ljubljana metropolitan area with about 280.000 inhabitants and provides an interesting gateway from direction of Balkan and near East countries towards EU countries (Fig.1). According to Airport Jože Pučnik Ljubljana official statement, the airport has a catchment area of about 4 million potential passengers.

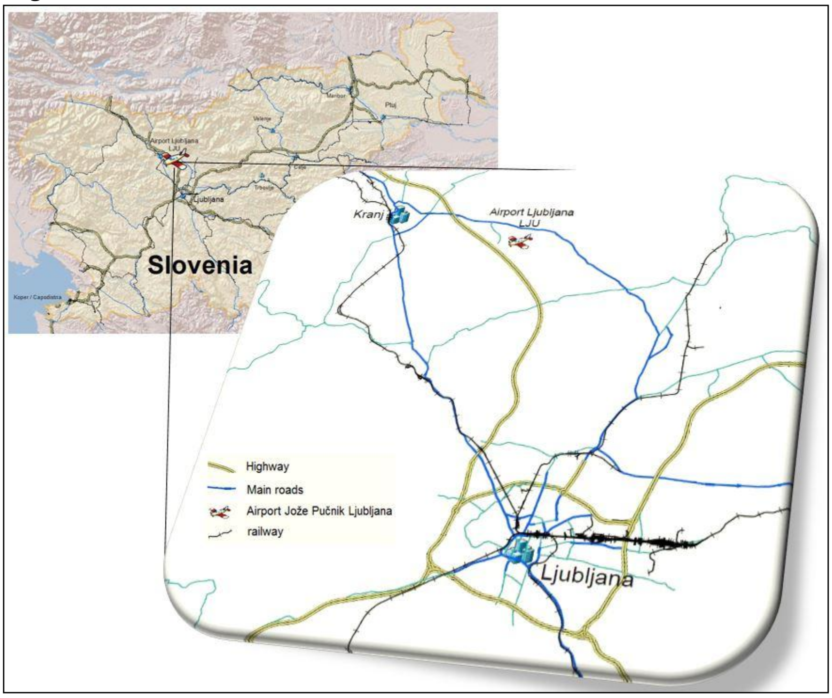



**Fig. 1**
*Location of Jože Pučnik Ljubljana Airport*
*Source: Author*

Like many other airports, due to worldwide economy crisis, it witnessed a steady decrease in passenger numbers from 2008 onwards when it handled 1.65 million passengers, down to just 1.17 million passengers in 2012. In 2013, numbers gradually picked up and airport handled 1.26 million passengers (Fig.2). First half of 2014 also shows an upward trend. Practically entire passenger volume is international, since Slovenia has just one international airport with regular lines and no significant home traffic.

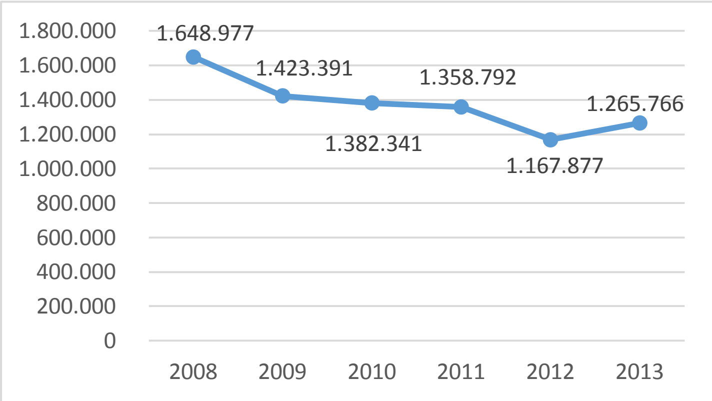



**Fig. 2**
*Passenger traffic at LJU for the period 2008-2013*
*Source: ec.europa.eu/eurostat*

Data for this study was collected through a passenger survey performed in the scope of EU Adria-A project in the spring of 2011. Questionnaires covered demographics information, access mode choice data, travel data and other information. Questionnaires were filled in by the interviewers, and eventually data for 998 passengers was obtained. However, we excluded some of the questionnaires due to incomplete answers. With regard to access mode, a sample of 968 questionnaires offered complete data for our analysis.

Data analysis for trip purpose segmentation shows that 47% of the passengers were business travelers, 39% vacation or pleasure, for 9% of the travelers, trip purpose was visit to relatives or friends, and remaining 5% captured student travel and other purposes (Fig.3). This segmentation is expected due to available destinations and airlines operating at the airport at the time the interviews were conducted.

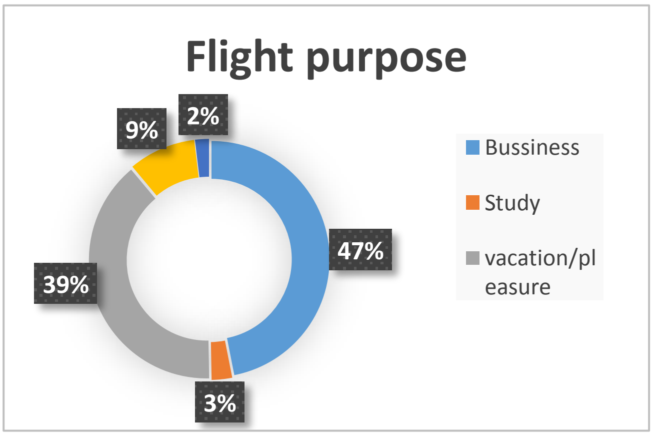



**Fig. 3:**
*Passenger share by flight purpose*
*Source: Authors own calculations*

We expect this segmentation to be different on various airports. For example according to a research performed by Loo B. (Loo, 2008) at Hong Kong International Airport, only 41% of travelers were making a business trip and the rest was for non-business purposes like vacation or visit to relatives. Author attributed this distribution to destinations available at the airport being mostly tourist destinations.
Similarly another research at King Khaled International Airport, Riyadh, Saudi Arabia (Alhussein, 2011) shows 54,3% travelers were making a business trip. It appears the distribution is strongly a function of destinations served at the airport. Airports serving mostly tourist destination are more inclined to serve more non-business travelers than business, which explains the distribution at LJU Airport since it serves various destinations (Fig.4), which are in nature business and non-business.


**Fig. 4:**
*Destination served by Airport Jože Pučnik Ljubljana*
*Source: http://www.lju-airport.si/sl/potniki-in-obiskovalci/info-o-letih/destinacije/*

Not all of the destinations are all year-round, since some of them are more seasonal in nature. Altogether twenty regular lines to European destinations are provided by seven airlines. This segmentation is important since business travelers are less inclined to be sensitive to airfare and ground access cost changes and value time greatly, whereas non-business travelers show higher sensibility to changes in cost than to worse flight frequencies or even greater distance to airport (Eric Pels, 2003).

Upon determining the purpose segmentation, in order to gain some insight into their decision proces, we were interested about the reasons why passengers chose this airport since they also have other options in this region (Airport Trieste, Airport Treviso, airport Venice,…) (Fig.5).

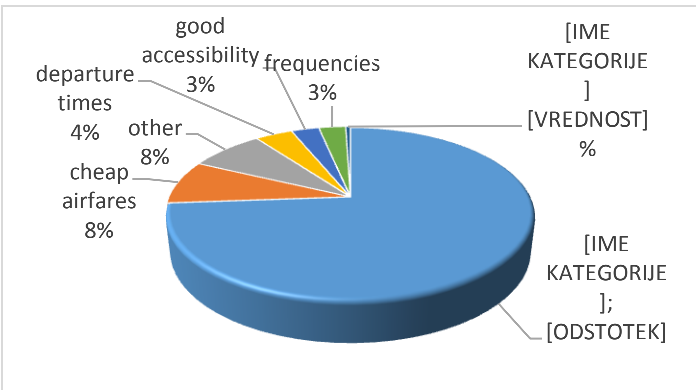



**Fig. 5:**
*Airport choice first reason*
*Source: Authors own calculations*

When passengers were faced with the question about the main reason why they choose this aiport as point of origin a majority of them stated that the main reason was because it was the closest airport (73.7%) and low airfares were only second with 8,2%.

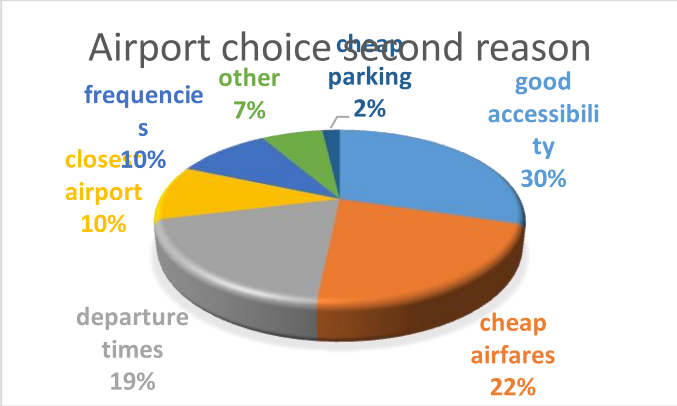



**Fig. 6:**
*Airport choice second reason*

Second choice on why they chose this airport was more evenly distributed between good accessibility (30%), low airfares (22%), departure times (20%) etc.

At this point, it is necessary to emphasize that Airport Ljubljana has a very limited public transport service availability. Railway service is unavailable at Airport Jože Pučnik Ljubljana and bus service, running every day of the week from Ljubljana city center to Airport Ljubljana, guarantees on average only 10 connections a day. Therefore, we expected access mode choice to be heavily in favor of a private car use.

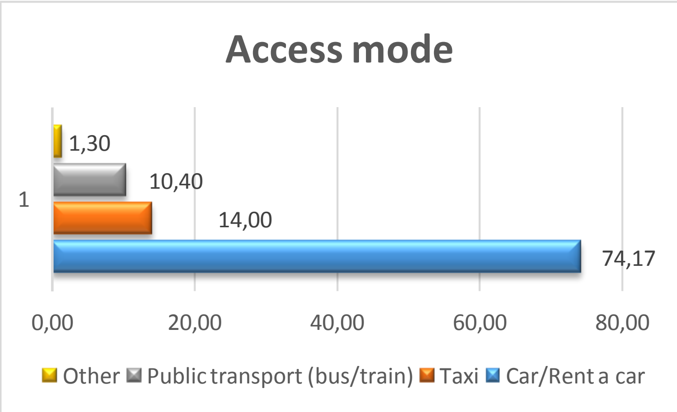



**Fig. 7:**
*Access mode choice*

Analysis shows that we were right, since 74% of passengers used private car as a means to arrive at airport. Interestingly, percentages for non-business travelers hardly differ from business travelers (73% vs 74%). This leads us to believe that passengers don't have enough alternative options in order to replace private car use to access the airport.

## 3. Conclusion

According to various authors presented in this paper, ground access mode choice is an important component in airport landside planning (for example Allhusein, 2011) and variables *travel time* and *travel cost* have a major impact on passenger's mode choice. Results of our preliminary research confirm the findings that the first variable (travel time) has a significant impact on passengers' ground access choice. As much as 74% of all passengers chose Jože Pučnik airport as an airport of origin due to proximity to the airport. It is safe to assume the closest airport is also the airport that passengers can get to in shortest time. Providing that shortest path is the cheapest, indicates the second variable (travel cost) is significant as well. We expected low airfares will have a big influence on airport choice, however data reveals that passengers rate low airfare only second to proximity to airport and attribute it a much lower importance in comparison to the fact that airport is the closest one to their point of origin.
Additionally, the lack of options to use any other access mode is a major factor for acces mode choice of travelers at Jože Pučnik Airport. Within the scope of this preliminary research we did not consider all of the variables which may effect the passengers mode choice and we only observed the ones for which we expected to have the most impact. For example, another aspect we need to look at withinin further research is passenger luggage. Some researchers indicate that the amount of baggage passengers carrie effects their choice on use of public transport modes (Mei Ling Tam et al., 2005) and lastly but not least we need to consider the geographical and geo-political circumstances and the effect these have on airport acces and with it connected access mode choice for Airport Jože Pučnik Ljubljana.

**References**


ACRP 2008. Airport cooperative research program (ACRP) Airport Ground Access Mode Choice Models: A Synthesis of Airport Practice (ACRP SYNTHESIS 5). Project 11-03 ed. WASHINGTON, D.C., 2008: Transportation Research Board.

Alhussein, S. N. 2011. Analysis of ground access modes choice King Khaled International Airport, Riyadh, Saudi Arabia. *Journal of Transport Geography,* 19, 1361-1367.

Cripwell, J. P., Gosling, G. D., BIGGS, D. C., BOL, M. A., BAKER, J., FRANZ, J. D. & BOARD, T. R. 2009. ACRP Report 26: Guidebook for Conducting Airport User Survey. Transportation Research Board.

ERIC PELS, P. N., PIET RIETVELD 2003. Access to and competition between airports a case study for the San Francisco Bay area. *Transportation Research Part A: Policy and Practice*, 71-83.

HESS, S. 2010. Modelling Air Travel Choice Behaviour. *Airport competition: The European Experience*, 151-175.

HESS, S. & POLAK, J. W. 2005. Mixed logit modelling of airport choice in multi-airport regions. *Journal of Air Transport Management,* 11, 59-68.

LOO, B. P. Y. 2008. Passengers' airport choice within multi-airport regions (MARs): some insights from a stated preference survey at Hong Kong International Airport. *Journal of Transport Geography,* 16, 117-125.

MEI LING TAM, LAM, W. H. K. & TAM, M. L. 2005. Analysis of airport acces mode choice: A case study in Hong Kong. *Journal of the Eastern Asia Society for Transportation Studies,* 6, 708-723.

SUZUKI, Y., CRUM, M. R. & AUDINO, M. J. 2003. Airport Choice, Leakage, and Experience in Single-Airport Regions. *Journal of Transportation Engineering,* 129, 212-218.

UBOGU, A. E. 2013. Determinants of Passengers' Choice: A Case Study of Mallam Aminu Kano International Airport (Nigeria). *International Journal for Traffic and Transport Engineering,* 3, 230-242.

# SYNTHESIS OF SIMPLE AIRCRAFT PERFORMANCE MODEL WITH FLIGHT GEAR PILOT INTERFACE

**Alisa Midžić[1], Karolina Krajček[2]**
*[1,2]University of Zagreb, Faculty of Transport and Traffic Sciences, Vukelićeva 4, 10000 Zagreb, Croatia*

**Abstract:** Simple aircraft flight models are often used for performance analysis and air traffic simulation since they are computationally less demanding then full 6DOF models. This paper will show how to define man-in-the-loop simulator in Matlab/Simulink environment. Aircraft motion is based on modified point performance model. Model in Simulink is connected with the Flight Gear simulator for enabling a visual reference and cockpit interface system for pilot. Aircraft dynamics is defined in the vertical plane with elevator, roll angle and engine throttle as control inputs. This Simulink model can be further improved for a more realistic experience. Its main intention is for educational purposes and preliminary analysis.

**Keywords**: point performance model, man-in-the-loop, simulation.

## 1. Introduction

Simple aircraft flight model, unlike six degrees of freedom (6DOF) model, does not take into account moments that act in body axes of the aircraft and therefore defines the motion of the aircraft solely as a motion of its center of mass. During flight, sideslip angle is assumed to be zero. Point mass model are often employed when studying the guidance and control of aircraft (Kinoshita & Imado, 2006). Modeling was conducted in Matlab/Simulink environment using inputs from pilot joystick and simulation has been carried out in open-source simulator FlightGear. Pilot can control elevator angle, roll angle and engine throttle. This flight model can be used for analyzing aircraft trajectories and optimization of flight routes. Similar aircraft flight models can be found in (Javier García-Heras Carretero, 2013) and (Yun & Li, 2013).

## 2. Simple Aircraft Performance Modeling

Aircraft model has symmetric geometry with one piston engine equipped with fixed pitch propeller that flies in low level atmosphere. The aircraft is modeled as rigid body with constant mass and center of mass like in (Yun & Li, 2013). Point performance model implies that aircraft flies in steady flight with all forces in balance, and the sum of moments equals zero. To model aircraft dynamics in Simulink environment it is first required to model the atmosphere, engine and aerodynamics. The relationship between these submodels and input and output interfaces is shown in **Figure 1**.

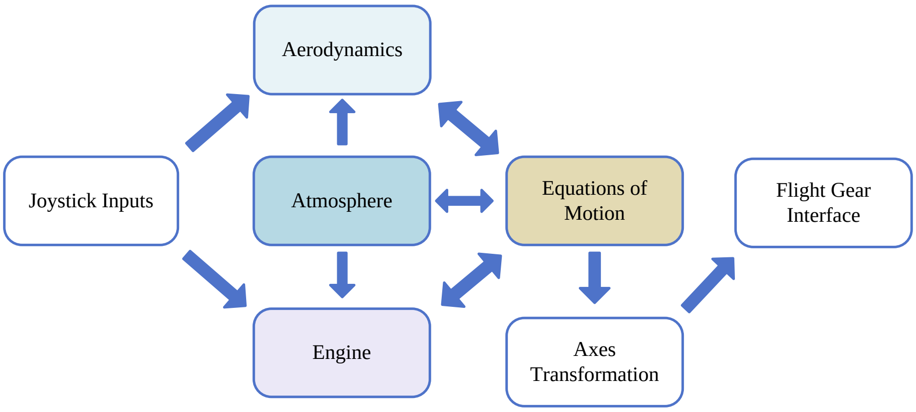



**Fig. 1.**
*Aircraft flight modelling block sheme*

Matlab tool Simulink enables computation of forces and other parameters via pre-prepared blocks or set of functions, or by manually entered equations. The **Figure 2** shows complete model in Simulink where input parameters are provided by pilot joystick interface and the results are presented through flight simulator FlightGear. This model contains required submodels and manually entered blocks for calculation of e.g. angle of attack for steady level flight or unit conversions. Initial state variables are entered in the form of six seven component state vector **X**.

$$\mathbf{X} = \begin{bmatrix} x & y & h & \chi & \gamma & V \end{bmatrix}$$

Where: $x$, $y$ and $h$ are aircraft position in body axis coordinate system ($h$ = -$z$),
$\chi, \gamma$ - velocity yaw and pitch angle and
$V$ – true airspeed.

[1] Corresponding author: karolina.krajcek@fpz.hr

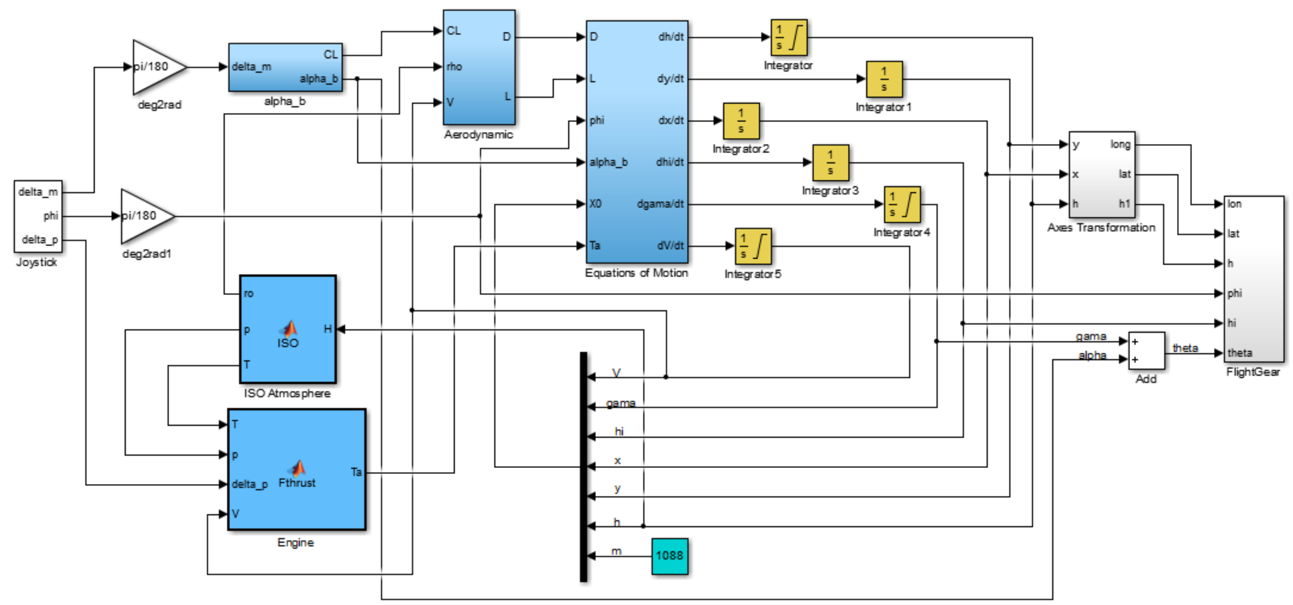



**Fig. 2.**
*Aircraft point performance model in Simulink*

### 2.1. Equations of Motion

Dynamic point performance model is contained in block named *Equations of Motion*. First three equations (1) are used for calculation of disturbances in velocity, pitch angle and yaw angle in velocity coordinate system. Since there is an assumption that there is no wind, there is also no sideslip angle and flight speed has equal value as aerodynamic speed just opposite direction. Next three equations (1) show how to calculate shifts along body axes and local coordinate systems.

$$
\begin{aligned}
& m\frac{dV}{dt} = T\cos(\alpha_T - \alpha_r) - D - W\sin\gamma \\
& mV\cos\gamma\frac{d\chi}{dt} = [L - T\sin(\alpha_T - \alpha_r)]\sin\phi \\
& mV\frac{d\gamma}{dt} = [L - T\sin(\alpha_T - \alpha_r)]\cos\phi - W\cos\gamma \\
& \frac{dx}{dt} = V\cos\gamma\cos\chi \\
& \frac{dy}{dt} = V\cos\gamma\sin\chi \\
& \frac{dH}{dt} = V\sin\gamma
\end{aligned}
\tag{1}
$$

These equations can be derived graphically by resolving forces into components along axes of aerodynamic and velocity coordinate systems. The angle between these coordinate systems is $\mu_A$, but aircraft roll angle $\phi$ can be used for calculation since difference between these angles is so small that it could be neglected (Vrdoljak & Jankovic, 2001). The inputs and outputs for this block are therefore as listed in **Table 1**.

**Table 1**
*Inputs and Outputs of Equations of Motion Block*

| Inputs | | Outputs | |
|---|---|---|---|
| lift and drag | $L, D$ | airspeed | $V$ |
| angle of attack | $\alpha_r$ | position along x and y axes | $x, y$ |
| available thrust | $T_a$ | altitude | $H$ |
| vector of state variables | $X_0$ | velocity pitch angle | $\gamma$ |
| roll angle | $\phi$ | velocity yaw angle | $\chi$ |
| propeller incidence angle | $\alpha_T$ | | |

Instead of writing down these equations manually, blocks *6th Order Point Mass Forces* (*Coordinated Flight*) and *6th Order Point Mass* (*Coordinated Flight*) from Simulink Aerospace Blockset version R2013b can be used, but important advantage of not using ready-made blocks is that this way model can be compiled for any version of Matlab/Simulink since it is not uncommon that when new Matlab version is introduced that some functions from previous version are changed or even no longer exist. And also, these libraries are rarely available without additional fee.

### 2.2. Pilot Joystick Interface

In Simulink, for connecting joystick to a model, block *Pilot Joystick* from *Aerospace Blockset* from (MathWorks, 2010) library has been used. Other blocks can also be used for this purpose such as *Joystick Input* from *Simulink 3D Animation* library, but the outputs of this block are different and not already adapted for aircraft related models. Needed outputs of *Pilot Joystick* block are roll angle $\phi$, elevator deflection $\delta_m$ and throttle position $\delta_p$. Values for roll angle and elevator deflection are given ranging from -1 to 1, so they have to be adapted to fit real possible roll angles and elevator deflections. To be able to use this block, computer has to be fitted with appropriate hardware equipment – properly calibrated joystick.

### 2.3. Atmosphere

Atmosphere modeled is *International Standard Atmosphere* with constant drop rates for atmospheric temperature, pressure and density with increasing altitude. This approximation is valid for altitudes up to an average of 11 kilometers, which is the mean height of the low level atmosphere – troposphere. At sea level temperature, $T_0$ is 15°C, pressure $p_0$ =101325 Pa and density $\rho_0$ =1.225 kg/m$^3$. The equations (2) for temperature $T$, pressure $p$ and density $\rho$ at observed height are as follows:

$$\begin{aligned} T &= T_0 - 6.5\cdot 10^{-3}\cdot H \\ p &= p_0\left(1-2.256\cdot 10^{-5}\ H\right)^{5.256} \\ \rho &= \rho_0\left(1-2.256\cdot 10^{-5}\cdot H\right)^{4.256} \end{aligned} \tag{2}$$

### 2.4. Engine

Engine block contains a function for calculation of available thrust from available power (calibrated for altitude impact). Power depends on the openness of carburetor throttle so position of throttle $\delta_p$ is expressed in unit percentage of power, meaning that maximum openness gives 100% of power for that height and closed throttle does not give power. Important equations (3) are:

$$\begin{aligned} P_a &= \eta_{prop} P_{eng}\delta_P \\ T_a &= \frac{P_a}{V} \end{aligned} \tag{3}$$

Where: $P_{eng}$ [W] – engine power,
$P_a$ [W] – available (propeller) power,
$T_a$ [N] – available thrust and
$\eta_{prop}$ – coefficient of propeller efficiency.

Propeller diameter $D$ = 1.88 m and rotation speed $n$ = 38.2 rev/s.

$$\eta_{prop} = -1.6923\cdot J^3 + 1.4815\cdot J^2 + 0.5670\cdot J + 0.2644 \tag{4}$$

Where $J = V/\left(n\cdot D\right)$.

For this model, data are taken from (Lycoming, 1989) for engine IO-360-A equipped with two-blade fixed pitch propeller and power is calculated with equations (5) as described in (Jankovic, 2002) and (Smith & Dreier, 1977).

$$\begin{aligned} P_B &= -31916 + 0.6783\cdot p_s + 0.003912\cdot p_s\cdot\omega - 12.817\cdot\omega \\ P_A &= 3206.5 + 0.3017\cdot p_s + 0.003785\cdot p_s\cdot\omega + 21.363\cdot\omega \\ p_A &= \frac{P_A - 3922 - 1.638\omega}{0.0034406\omega + 0.41009} \end{aligned} \tag{5}$$

$P_B$ is engine power in Watts for given manifold pressure $p_s$ [Pa] and rotation speed $\omega$ [rad/s] at sea level and in standard atmosphere. $P_A$ [W] is engine power at altitude above sea level at some pressure $p$ and corresponding temperature $T$ in standard atmosphere. $p_A$ is atmosphere pressure for specific $P_A$ power and rotation speed $\omega$ . For some different then standard atmosphere pressure we can calculate engine power $P_D$ as:

$$P_D^{'} = P_B + \left(P_A - P_B\right)\frac{p_D - p_0}{p_A - p_0}$$

Power $P_D$ is engine power when atmosphere pressure is $p_D$ [Pa] and temperature equals standard temperature for that pressure. This equation is valid for static conditions when airspeed is zero. If there is some airspeed total pressure entering the engine must be used instead of $p_D$ but corrected for ram effects. These losses are about 15%:

$$\begin{aligned} p_{total} &= p_D + 0.5\rho V^2 \\ P_D &= P_B + \left(P_A - P_B\right)\frac{0.85 p_{total} - p_0}{p_A - p_0} \end{aligned} \tag{6}$$

Engine power for pressure $p_D$ and some nonstandard temperature $T_D$ is:

$$P_{eng} = P_D\sqrt{\frac{T_N}{T_D}}$$

### 2.6. Aerodynamic

Aerodynamic submodel contains blocks for calculation of lift $L$ and drag $D$. These forces are calculated (7) by multiplying appropriate aerodynamic coefficient, dynamic pressure ( $0.5\rho V^2$ ) and reference area $A$.

$$\begin{aligned} L &= C_L \frac{1}{2}\rho V^2 A \\ D &= C_D \frac{1}{2}\rho V^2 A \end{aligned} \tag{7}$$

Coefficients of lift $C_L$ and drag $C_D$ for the whole body are calculated from equations(8):

$$\begin{aligned} C_L &= C_{L0} + C_{L\alpha}\alpha_e + C_{L\delta}\delta_m \\ C_D &= C_{D0} + KC_L^2 \end{aligned} \tag{8}$$

Elevator angle $\delta_m$ comes from pilot joystick, and elevator angle can be calculated from moment equation (9) in steady flight:

$$\begin{aligned} 0 &= C_{m0} + C_{m\alpha}\alpha_e + C_{m\delta}\delta_m \\ \alpha_e &= \frac{C_{m0} - C_{m\delta}\delta_m}{-C_{m\alpha}} \end{aligned} \tag{9}$$

Aerodynamic coefficient gradients and zero members are used from (Jankovic, 2002) for airplane Cherokee 180 Piper.

### 2.6. Transformation of Axes

Position variables that are calculated in *Equations of Motion* block are expressed in local coordinate system. To be able to use them in FlightGear they have to be transformed into geodetic coordinates of longitude, latitude (in radians) and altitude (in meters above reference height). Model of Earth used for this transformation is WGS84 but it can also be custom made in *Flat Earth to LLA* (MathWorks, 2010) axes transformation block. The relationship between coordinate systems is shown in **Figure 3**.

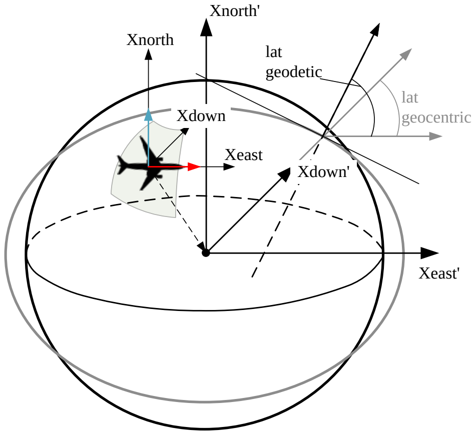



**Fig. 3.**
*Definition of used coordinate systems*

Equations (10) - (13) used for transformation of flat Earth position to geodetic longitude, latitude and altitude are found in literature (MathWorks, 2010). First it is need to transform flat Earth x and y coordinates to North and East coordinates using angle ψ (direction of flat Earth x-axis in degrees clockwise from north). After that, North and East coordinates need to be transformed in geodetic latitude and longitude using radius in the prime vertical $R_N$ and radius in the curvature of meridian $R_M$ (MathWorks, 2010) which depend on equatorial radius and flattening of Earth model used. Altitude is the negative flat Earth z-axis minus height from surface to flat Earth frame.

$$\begin{bmatrix} N \\ E \end{bmatrix} = \begin{bmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{bmatrix} \begin{bmatrix} p_x \\ p_y \end{bmatrix} \tag{10}$$

$$\begin{aligned} R_N &= \frac{R}{\sqrt{1-\left(2f-f^2\right)\sin^2\mu_0}} \\ R_M &= R_N \frac{1-\left(2f-f^2\right)}{1-\left(2f-f^2\right)\sin^2\mu_0} \end{aligned} \tag{11}$$

$$\begin{aligned} d\mu &= \tan^{-1}\left(\frac{1}{R_M}\right)dN \\ d\ell &= \tan^{-1}\left(\frac{1}{R_N\cos\mu}\right)dE \end{aligned} \tag{12}$$

$$\begin{aligned} \mu &= \mu_0 + d\mu \\ \ell &= \ell_0 + d\ell \\ h &= -p_z - h_{ref} \end{aligned} \tag{13}$$

Attitude inputs also need to be transformed to be compatible with Flight Gear. Axes *x* and *z* in FlightGear body-fixed system have opposite directions than in standard aircraft body system. That is why yaw angle has to be put in with opposite sign (shown in **Figure 4**).

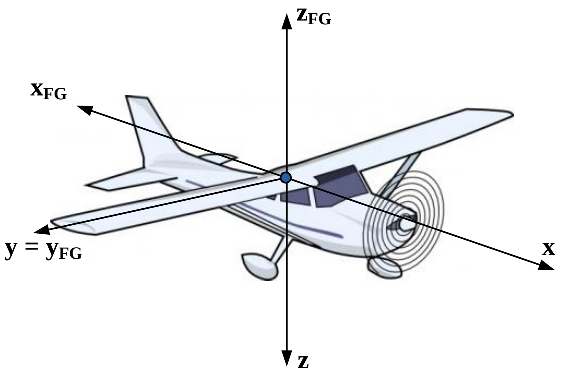



**Fig. 4.**
*Definition of Flight Gear and aircraft body axis*

### 2.7. Connecting to FlightGear

FlightGear is open-source flight simulator available online that is compatible to Matlab environment which means that Simulink models can be run in FlightGear. Aircraft appearance is defined in AC3D geometry file in FlightGear Aircraft directory. AC3D is computer software available for trial use for only 14 days, so if task is to analyze only the precision of dynamic model then it is advisable to use already done aircraft geometry files for visualization since its production is really time-consuming.
Execution of Matlab and FlightGear on the same computer, depending on the complexity of the model, can produce great loads for the processor and cause problems in real-time simulation process. Therefore, it is recommended to run them on two separate computers. Four blocks have to be used: two blocks that describe simulation proces and two for communication between Simulink and Matlab. First it is required to use block *Generate Run Script* to define inital position and geometry model used. Then it is required to set simulation pace (ratio of sim seconds per clock seconds) by using *Simulation Pace* block. Now data can be send to FlightGear. Block *Pack net_fdm Packet for FlightGear* packs position and attitude inputs which means it translates them in C++ language. Next block *Send net_fdm Packet to FlightGear* contains functions that send this packet to remote computer or local host depending on destination IP address. All this, can be programed manually without use of pre-made blocks if the user is familiar with C++ language. Also, if the simulation is still to be run only on host computer then the last three blocks can be replaced with only one block – *FlightGear Preconfigured 6DOF Animation* from *Aerospace Blockset* library.

## 3. Real-time Simulation and Model Verification

Two short man-in-the-loop simulations are described to prove reliability of the model. Initial conditions are:

- $\delta_m = 0$
- $V$ = 113 knots ($V_i$ = 105 knots) and
- H = 5 000 ft.

First simulation has been made with around 75% of power. Results are shown in **Figure 5**. The changes in indicated speed, altitude, flight path angle and angle of attack variables are due to imperfection of human pilot flying skills. Piloting a model is even harder than piloting a plane because there are no forces on the joystick like in real flight, also since there is no automatic trim function programmed any oscillation produces new oscillations. It is evident that the aircraft, to maintain level flight, needs to have positive angle of attack which is in accordance with what is expected for this simple aircraft.

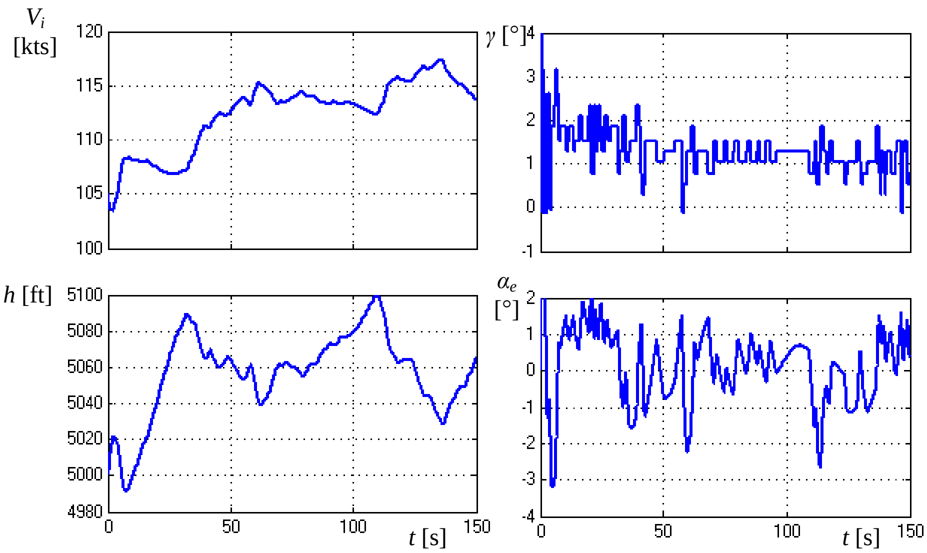



**Fig. 5.**
*Results of first man in the loop simulation for indicated airspeed, altitude, velocity angle and angle of attack*

Purpose of the second simulation is to illustrate longitudinal stability of the aircraft. Initial conditions are the same as in first simulation except that power is now around 60% of maximum available. It is shown in **Figure 6** that if no inputs are used on joystick (aerodynamic surfaces stay in neutral position), aircraft will gain constant speed and angle of attack after approximately two minutes. Compared to short period motion of military aircraft (Thong, 2010) this is realistic since these types of aircraft (for general small aviation and training) have improved stability at the expense of maneuverability since that is more important to aircraft for combat operations. Other simulations have been conducted to test model in flight. Values of state variables are all in expected ranges, and the difference from real flight data are consequence of made assumptions and the accuracy of input data. It is important to emphasize that quality of the model depends on the quality of available aircraft data itself.

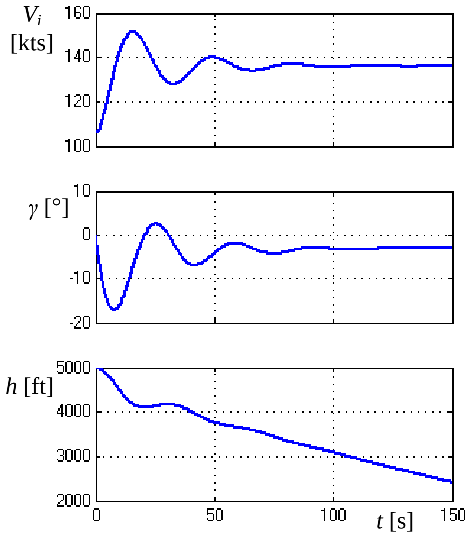



**Fig. 6.**
*Shortperiod oscillations of airspeed, angle of attack and altitude with time*

## 4. Conclusion

In this paper we showed how to connect simple point performance aircraft model in Simulink with FlightGear simulator. This connection enables man in the loop simulation for more realistic trajectories when analyzing aircraft performance or preliminary concepts. Data used for calculating aerodynamic and propulsion forces are based on Piper Cherokee PA-28-180 data from (Jankovic, 2002) and (Lycoming, 1989) but can be easily replaced for other aircraft types. Results shows realistic flight trajectory for two preliminary men-in-the-loop simulations. Further research will include aircraft flight with variable mass and center of mass. Also we will try to include more complex aerodynamics with extractable flaps for improved descent phase of flight.

**References**

Jankovic, S. 2002. *Mehanika leta zrakoplova* (*Aircraft Flight Mechanics*). Faculty of Mechanical Engineering and Naval Architecture. Zagreb.

Javier García-Heras Carretero, F. 2013. Aircraft Trajectory Simulator Using a Three Degrees of Freedom Aircraft Point Mass Model. In *ATACCS '13 Proceedings of the 3rd International Conference on Application and Theory of Automation in Command and Control Systems*, 114-117.

Kinoshita, T.; Imado, F.2006. The Application of an UAV Flight Simulator - The Development of a New Point Mass Model for an Aircraft. In *SICE-ICASE International Joint Conference*, 4378-4383.

MathWorks, T. 2010. *Aerospace Blockset 3*, Revision 16. User's Guide. USA.

Smith, H.; Dreier, M. 1977. A Computer Technique for the Determination of Brake Horsepower Output of Normally-Aspirated Reciprocating Aircraft Engines. *SAE Technical Paper*.

Sorton, E.; Hammaker, S. 2005. Simulated Flight Testing of an Autonomous Unmanned Aerial Vehicle Using FlightGear. In *Infotech@Aerospace*, 13 p.

Thong, C. 2010. Modeling Aircraft Performance and Stability on X-Plane. *The UNSW Canberra at ADFA Journal of Undergraduate Engineering Research*, 3(2): 31 p.

Vrdoljak, M.; Jankovic, S. 2001. Aircraft Performance Obtained from Modified Point Mass Model. *ZAMM - Journal of Applied Mathematics and Mechanics*, 81: 625–626.

Yun, C.; Li, X. 2013. Aerodynamic Model Analysis and Flight Simulation Research of UAV Based on Simulink. *Journal of Software Engineering and Applications*, 6: 43-47.

Zovak, I. 2011. *Planiranje leta borbenog aviona primjenom modela gibanja u ravnotežnom letu* (*Flight Planning of the Fighter Aircraft with the Application of the Steady State Flight Model*). MSc Thesis. Faculty of Mechanical Engineering and Naval Architecture. Zagreb.

*Transport and Environment*

# OPPORTUNITIES FOR IMPROVEMENTS OF THE INTERMODAL TRANSPORTATION CHAIN: A CASE STUDY OF INTERMODAL FREIGHT OPERATION BETWEEN TURKEY AND SERBIA

**Ömer Faruk Görçün**[1]
*[1] Department Head of Logistics at Kadir Has University*

**Abstract:** In order to reduce the external costs of transportation, parties of logistics and transportation such as international organizations, governments, local authorities, transporters and users have been promoting the solutions which will be created more efficient results. Within this framework, opportunities related to rail transportation and reducing the rate of use of road freight transportation are very important factors in a green environment. This paper focus on the economic and environmental impacts of intermodal freight operations, rail transportation and intermodal freight terminals. On the other hand, results which will be obtained by using the intermodal freight terminals and rail transportation have been evaluated comparatively. In this context, several scenarios have been determined and they have been analyzed. In recent years, some concepts such as effectiveness, logistics costs and quick response had become more important with the globalization. These concepts were a result of competition, which have developed quickly. At the same time, not only competition occurs between companies, but also it was seen between global actors. Although these concepts have remained their importance, different subjects have begun to gain an importance, such as environment, human health and external costs of transportation. In order to reduce the environmental effects of the transportation, considering planning the differential and beneficial approaches has become imperative. One of the best ways is intermodality for solving the transportation problems. In this study, transportation alternatives and their effects were analyzed and new transportation alternatives between Turkey and Serbia have been recommended. Intermodal freight terminals which will be established in both of İstanbul and Belgrade city can be creating an effective solution for reducing the economic and external costs. On the other hand, it can provide sustainable and effective transportation system between two countries.

**Keywords**: intermodal freight transportation, terminals, intermodal rail transportation

## 1. Introduction

With the globalization, competition has become a most important concept and it has gained global characteristics. İn this process, some concepts such as logistics service level and logistics costs have become more important. Actually, there is an inverse ratio between these concepts. Companies are looking for ways of the reducing the logistics costs as well as providing the logistics service at a high level. In order to achieve both of these objectives and to gain a competitive edge with the best conditions, most important factors is speed. If companies can increase their speed, they can also increase their competitive advantage. As a result, increasing the speed can provide sustainability, reliability and customer satisfaction. Competition and concepts as mentioned above have remained their importance, different subjects have begun to gain an importance, such as environment, human health and external costs of transportation. In order to reduce the environmental effects of the transportation, considering planning the differential and beneficial approaches has become imperative. One of the best ways is intermodality for solving the transportation problems. Although road vehicles are new and their emissions are low compared to the past, when considering the number of road vehicles which passing from this route, external costs of road transport such as environmental pollution and risks have become serious and they cannot be ignored. On the other hand intermodality has become more important for European policy makers. Intermodal transport in Europe has registered a high rate of growth for many years since the beginning of its services. This growth is partly due to systematic promotion and subsidies received in various EU countries (Ballis and Golias, 2002). The development of intermodal freight is regarded as a key way in which rail can achieve a greater share of the freight transport market, but the limitations of official datasets make it difficult to develop a strong appreciation of the characteristics of existing intermodal flows (Woodburn, 2012). The promotion of environmentally friendly transport modes requires the enhancement of the economic, managerial and technical efficiency of the Intermodal transport terminals as well as of the whole Intermodal transport chain (Abacoumkin and Ballis, 2002). Serbia is an important country for global and regional logistics activities. It is located on the most preferred international road and rail routes between European and Middle East countries. In this study, benefits of intermodal freight transportation and intermodal train operations were analyzed between Serbia and Turkey.

## 2. The Current Situation and Logistics Opportunities in Serbia

Serbia has a great importance according to logistics and transportation actors. It is located in the center of the Balkan Peninsula, which gives it a strategic importance in the field of logistics, transportation and communications in the region. Two important international transport corridors pass through Serbia. First corridor is Pan-European transport corridor X which links nine countries from Germany to Greece and Turkey. In this corridor, road and rail transportation can be operated and its total length is around 2,360 km, of which 874 km is through Serbia. Another important transportation corridor which passes through Serbia is 7$^{th}$ Pan-European transport corridor (Danube waterway) which

[1] Corresponding author: omer.gorcun@khas.edu.tr

links ten states, from Germany to Ukraine and has 44 international ports. In this way, it represents an intersection between the south and the north, the west and the east of the Europe. These corridors intersect in Beograd city.

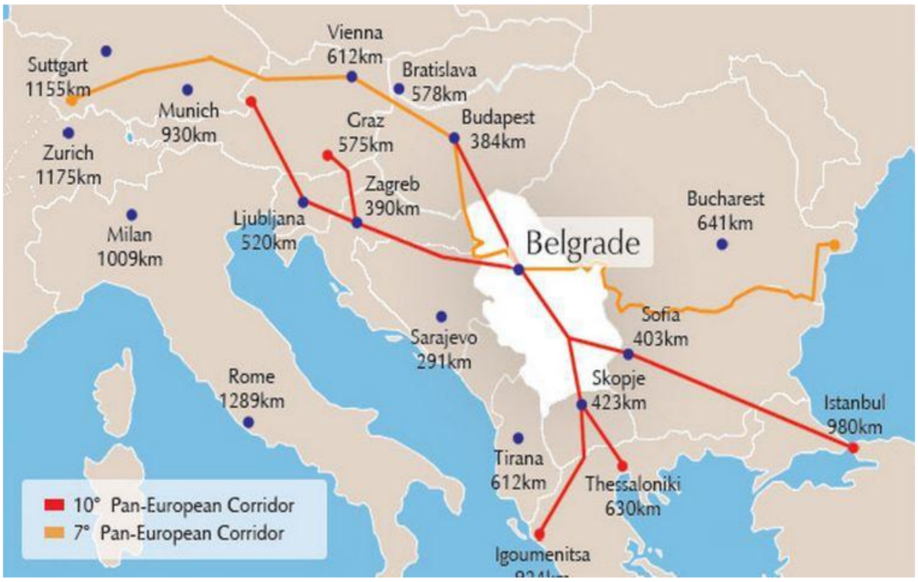



**Fig. 1.**
*Pan European Corridors in Serbia*

Because it is located in the important point which intersect mostly preferred two Pan European transport corridors, this city can be defined as logistics hubs for European and Middle East countries. In contrast, the logistics potential of Serbia cannot be used at adequate levels. Road transportation is mostly preferred transportation mode for both of domestic and international freight transportation. On the other hand, road freight transportation has a very high share, compared to other transportation modes. Road freight transport is a dominant mode of transport in Serbia (Medar and Manojlović, 2011).

According to the 2006 report from the Statistical Office of the Republic of Serbia, the international transport of goods by road (transit excluded) was twice as large as transport by railway and inland waterways measured in transported tones (Statistic Office of the Serbia). This situation is caused by an imbalance in the transportation sector. Unfortunately, railway and intermodal transportation is weak or it is not at the required level. Finally, economic and external cost of road freight transportation have been seen dramatically such as road congestion, accidents, loss of life and economic assets, environmental pollution and noise.

Nowadays, traffic congestion is very high when considering the road freight transport volume. Only number of road freight vehicles from turkey is reached to approximately 400.000 annually. With the traffic flow on the network ranging from less than 2,000 veh/day, to more than 100,000 veh/day. Also, 63% of the Core Road Network had traffic with more than 5000 veh/day in 2010, and a very small (< 10%) of network segments with very low traffic flows, less than 2000 veh/day (Manić, 2012).

According to Turkish authorities, transportation has been done 4884 times by road vehicles from Turkey to Serbia directly. On the other hand, 55.000 Turkish transit road vehicles passed from Serbia in 2013. At the same time, while 3274 times direct expedition have been done by road vehicles from Serbia in Turkey, in total 17.000 road vehicles entered into Turkey through Serbia. However, approximately 85.000 road vehicles of both countries were passed through this route in both of two directions.

When considering the total load between Turkey and European countries, in both of two directions a very busy road traffic which causes from road freight transportation is noticeable. According to statistics that issued by the International Transporters Association (UND), 499,437 road freight vehicles were passed from Turkey's border gate. 265,169 road freight vehicles were existing from turkey, 234,268 vehicles were entered into Turkey for import.

In order to more short distance from other alternatives, transporters use to the 10$^{th}$ corridor which passing from Serbia mostly. In Serbia, routes divides two sections, while the first route continues to west direction, second route continues to north. In order to a portion of road freight vehicles which transported to northern countries such as Romania, Ukraine and etc. Existed from Bulgaria, vehicles which passed from Serbia are reduced to 420,847. In directions of the west, 352.000 road freight vehicles were run between Belgrade and Ljubljana. Routes is separated in different directions, the number of road freight vehicles which carried out on the different parts of the route is decreasing towards the west.

**Table 1**
*Number of Road Freight Operations between Turkey and European Countries*

| Country | Turkish (Export) | | Turkish (Import) | | Foreign (Export) | | Foreign (Import) | |
|---|---|---|---|---|---|---|---|---|
| | Border Gate | | Border Gate | | Border Gate | | Border Gate | |
| | H.Beyli | K.Kule | H.Beyli | K.Kule | H.Beyli | K.Kule | H.Beyli | K.Kule |
| **Albania** | 0 | 203 | | 141 | | 8 | | |
| **Austria** | 334 | 4642 | 208 | 1104 | 312 | 4902 | 179 | 1948 |
| **Belgıum** | 1043 | 3062 | 323 | 1325 | 503 | 1876 | 212 | 632 |
| **Bosnıa** | 39 | 2900 | 4 | 1764 | 17 | 653 | | 202 |
| **Bulgarıa** | 2584 | 9699 | 10042 | 35815 | 2350 | 4468 | 11858 | 24054 |
| **Croatıa** | 30 | 1561 | 17 | 777 | 4 | 344 | 1 | 194 |
| **Czech Rep.** | 1178 | 2240 | 863 | 3439 | 518 | 1193 | 961 | 2441 |
| **Denmark** | 329 | 1613 | 65 | 680 | 241 | 612 | 30 | 125 |
| **England** | 2538 | 7138 | 285 | 1140 | 1065 | 2255 | 43 | 359 |
| **France** | 1749 | 7796 | 1273 | 2427 | 1063 | 3169 | 573 | 878 |
| **Germany** | 7034 | 45265 | | | 8078 | 51101 | 2768 | 5820 |
| **Holland** | 1731 | 6732 | 728 | 793 | 335 | 2971 | 614 | 1009 |
| **Hungary** | 1015 | 2069 | 919 | 2723 | 967 | 3597 | 995 | 7229 |
| **Italy** | 127 | 2047 | 139 | 448 | 72 | 1726 | 708 | 837 |
| **Kosovo** | 0 | 45 | | 100 | | 2 | | 5 |
| **Luxembourg** | 9 | 278 | 1 | 5 | 7 | 36 | 6 | 15 |
| **Macedonıa** | 7 | 870 | 14 | 614 | 1 | 89 | 27 | 124 |
| **Montenegro** | 0 | 193 | 2 | 175 | | 3 | | 17 |
| **Poland** | 3474 | 5286 | 3038 | 6228 | 4538 | 6515 | 2655 | 5947 |
| **Portugal** | 0 | 7 | 27 | 177 | 8 | 2 | 20 | 31 |
| **Romanıa** | 12888 | 13814 | 12027 | 3889 | 12853 | 8425 | 10505 | 4189 |
| **Serbıa** | 130 | 3471 | 26 | 3539 | 15 | 1097 | 34 | 2212 |
| **Slovakıa** | 861 | 1029 | 722 | 1741 | 1096 | 2283 | 593 | 3004 |
| **Slovenıa** | 30 | 1286 | 7 | 162 | 202 | 4342 | 159 | 1006 |
| **Spaın** | 40 | 436 | 591 | 1518 | 361 | 100 | 602 | 656 |
| **Swıtzerland** | 93 | 1988 | 26 | 165 | 113 | 1179 | 11 | 113 |

*Source: Statistics of International Transporters Association (UND) www.und.org.tr*

### 2.1. Impacts of Road Freight Transportation between Turkey and Serbia

Distance of route that is often preferred by road transport operators is 941 kilometers from Istanbul to Belgrade city. Annually, 420,847 heavy road vehicles operate on this route. Even if the average emission value of all of these vehicles is low, traffic intensity causes the serious and intolerable results. According to European Commission, acceptable emission value for per road vehicles is 176 g/km (EC 2007) Total emission value can be calculated as multiplying the route distance and emission factor per heavy road vehicles as seen below;

$$r_{em} = r_{tl} \; x \; r_d \; x \; rw_e \qquad (1)$$

While $r_{em}$ is represents the total emission value, $r_{tl}$ is the total distance of road, $r_d$ is total heavy road vehicles which operate on this route, $rw_e$ is an emission factor per vehicle. When 85.000 heavy road vehicles have been operated in both of two direction, annual emission value can be calculated as below;

$r_{em}$ = 420,847 x 1882 x 176 = 139,398 ton

A calculated emission value that caused by heavy road vehicles is very high. According to an important scientific study realized by the Research Center of Marine and Climate at University of Hamburg, the marginal cost of emission is calculated as 104 dollars per ton. (Tol, 2003) When this value takes into consideration, annual marginal cost of emission can be calculated as 14,497,391 USD. On the other hand, in addition to environmental costs of road transportation, it can cause social and economic costs. There are very large numbers of cost items can be seen in total transportation costs such as energy, transit, driver, loss of value and maintenance costs, etc. In contrary, most important cost item causes from energy consumption of heavy road transportation as below;

$$e_c = r_{tl}\ x\ u_{fc}\ x\ f_{uc} \qquad (2)$$

While $e_c$ is represents the total energy cost, $r_{tl}$ is the total distance of road, $u_{fc}$ is unit energy consumption per road vehicle, $f_{uc}$ is the monetary value per liter of fuel. If these factors are multiplied with each other, the total energy cost can be calculated per heavy road vehicle. As assumed that, unit fuel consumption is 30 liter per one hundred kilometers in optimum conditions. At the same time, unit fuel cost was taken into consideration as the unit sales price without taxes. When the total energy consumption cost calculated, it can be calculated as below;

$e_c$ = 941 x 30% x 0,45 = 127.035 €/vehicle

Annual energy consumption cost can be calculated as multiplying the unit energy cost per vehicles with the number of heavy road vehicles which pass through this route.

$\sum e_c$ = 420,847 x 127.035 € = 53,462,299 USD

In total, annual external and economic costs of road transport can be reached at 67, 96 million euro on the route between Turkey and Serbia. The monetary value of the road transportation cost is very high and it cannot be bearable by Serbian and Turkish policy makers. They should be seeking new transportation alternatives and they should create a new perspective which related to available or new transportation opportunities. In this context, railway transportation may be one of the best ways for solving the transportation problems and it can provide a sustainable transportation system and infrastructure.

### 2.2. Intermodal Transportation Opportunities in Serbia

Intermodal transportation may be defined as the transportation of a person or a load from its origin to its destination by a sequence of at least two transportation modes, the transfer from one mode to the next being performed at an intermodal terminal. Transportation of containerized cargo by a combination of truck, rail, and ocean shipping, dedicated rail services to move massive quantities of containers and trailers over long distances, main transportation mode for the international movement of goods, central piece in defining transportation policy of the European Community, trips undertaken by a combination of private (e.g., car) and public (e.g., light rail) transport, and so on (Crainic and Kim, 2007). Intermodal terminals as Logistics Centers (LC) are complex facilities with multiple functions, including transshipment yards, warehouses, wholesale markets, information centers, exhibition halls and meeting rooms, etc. (Ilin and Simic, 2013). When the intermodal opportunities of Serbia are considered, this country has a great advantage for being an intermodal hub. In contrast, opportunities related to intermodality cannot be used at the required level. According to Freight Service Department, JSC of Serbian Railways, while the share of intermodal transportation is 11.62% the share of conventional transport such as road, rail, and maritime transportation is reached to 88.38% nowadays.

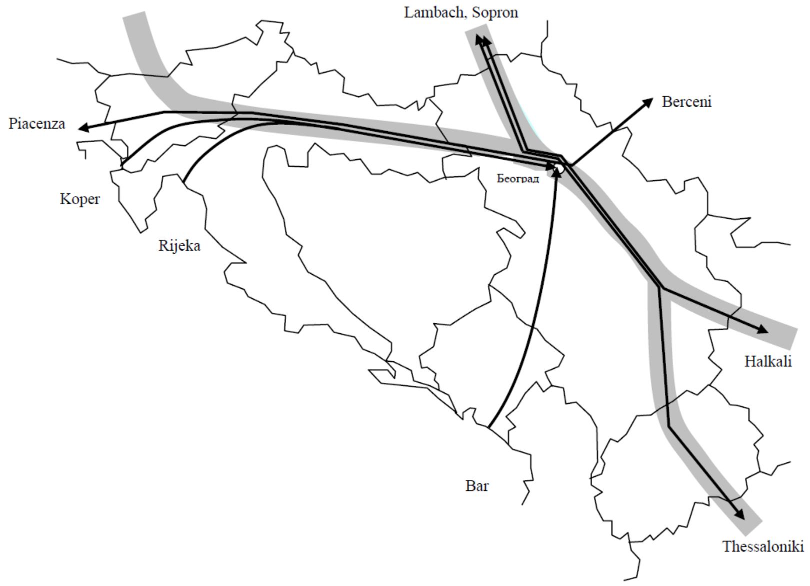



**Fig. 2.**
*International Railway Network in Serbia*

On the other hand, improvements are seen on intermodal transportation in Serbia, although road freight transportation is the dominant transportation mode. Although a large number of intermodal rail freight operations are carried out between European countries and Turkey and it was not continuously and systematically. These operations were carried out because of infrastructure projects such as gazelle project, block train project and etc. But they were not systematic and continuous rail operations. They were shown an individual characteristics and they were organized and planned when needed for such operations. Most important rail operations among them is block container train between Cologne/Germany and Köseköy/Turkey. These railway operations have been started in 2004 in order to create an effective solution for automotive logistics between automotive spare parts suppliers and automobile manufacturers. Nearly 12.000 swap bodies are transported on this route annually. Although, still Cologne – Köseköy line is important, it may not be effective and consistently at the required level. Another important experience related to railway operation between Turkey and European countries is block pipe train between Giengen/Czech Republic and Halkalı/Turkey. These operations were carried out from November 2010 to September 2011 for the gazelle pipeline project. The block train operation runs 130 times and 90.000 tones pipe were transported in the framework of this project.
Halkalı – Sopron rail line is used for container block train operations in the past. Miscellaneous goods have been transported by container train being operated reciprocally between Halkalı and Sopron 4 times in a week. At the same time other rail transportation experiences such as Halkalı – Wien/Austria container block train and Halkalı – Ljubljana/Slovenia were realized in previous time.
All of these projects were considered, Serbia is a transit country and it cannot be shown a dominant characteristic related to logistics flows, although it is an important country which located at the intersection point of rail and road transportation networks. In order to establish a more effective logistics system, Serbia should not be transit country, it should play a key role as the dominant actor related to logistics activities not only at regional but also at global level. Consequently, Serbia is defined as a logistics hub, regional logistics activities can be performed much better and mutual benefits can be gained by parties of logistics such as countries, transporters and users.

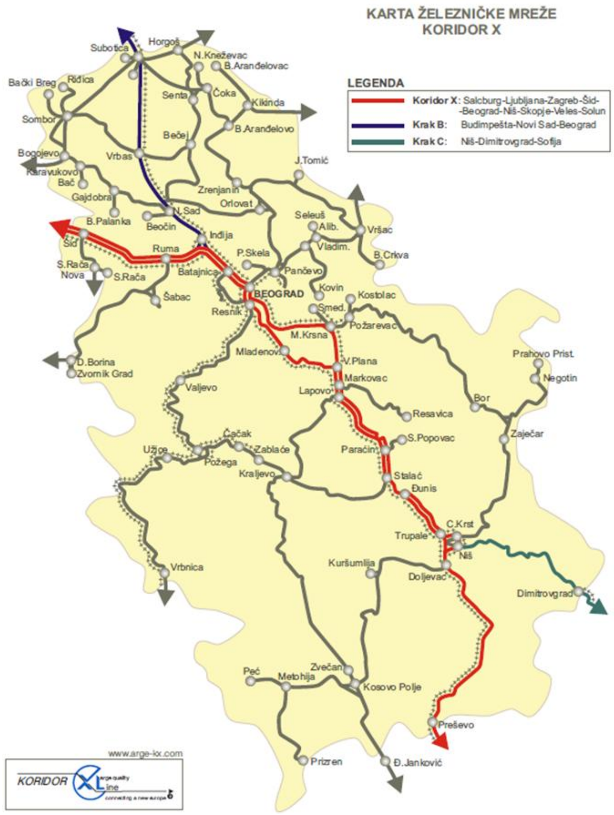



**Fig. 2.**
*Freight transport on Serbian rail network*
*Source: Andrija Ristović, Leading organizer for intermodal transport Freight Service Department, JSC "Serbian Railways" 1st South East Environment Workshop.*

Initially, logistics and freight terminal which can respond to regional logistics needs should be established in Belgrad city. This terminal has two functions such as collecting and distributing. Loads from different European and Middle East countries should be collected to this terminal and they should be consolidated and distributed in there. On the other hand, the high speed freight train may run on this line and it may provide better solution different from conventional trains. In this study proposed that, Belgrad intermodal freight terminal can be a central hub point for Europe and Middle East. When considering the total logistics flows between European countries, total logistics cost can be very high.

Reasons for this can be listed as using level of road transportation is very high, using of rail transportation at low level, because of the transportation operations has been run between two countries such as origin and destination countries, consolidation level and amount of cargo are insufficient according to efficiency, productivity and lower unit logistics costs. Logistics costs are the sum of the transportation costs between each point. In order to transportation operation can be realized with different capacity using level, productivity and unit logistics costs, total logistics cost can be variable and high. If load from different points can be consolidated in intermodal terminals, there may be saved in resources related to logistics activities such as transportation operations and using transportation vehicles and equipment.

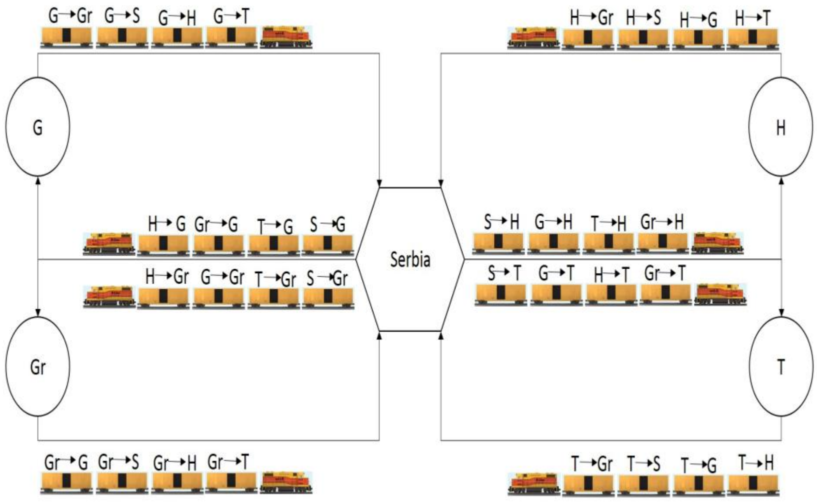



**Fig. 3.**
*Terminal Based Models and Connections between Serbia and European Countries*

**3. Benefits of the Intermodal Transportation between Turkey and Serbia**

When the total amount of cargo which transported between Turkey and European countries and averagely 22 ton cargo can be transported by per road freight vehicles, according to regulations, approximately 9, 3 million ton cargo have been transported in recent year. Daily, 25,366 ton cargo can be moved in both of two directions. Even if 30% of these cargoes can be transported by intermodal freight trains, total economic and environmental costs can be reduced significantly. According to transportation rules 26 ton cargo can be loaded to per 40 feet container and an intermodal freight train capacity is 32 containers. However 832 ton cargo can be transported in one expedition by this train. With the ten intermodal freight rail operations which will be run between Turkey and Serbia in a day, 8320 ton cargo can be transported and total economic benefits can be calculated as seen below;

$$r_{vn} = [C_c \ x \ t_c \ x \ n_{op}] \qquad (3)$$

$$d_{em} = ([r_{vn}] / rv_c) \ x \ r_d \ x \ r_{we} \qquad (4)$$

$$\sum m_{ec} = [(d_{em}) \ x \ e_{mc}] \ x \ n_{day} \qquad (5)$$

While $r_{vn}$ is represents the total transported cargo per intermodal freight train, $C_c$ is container capacity, $t_c$ is number of containers which can be loading per train, $n_{op}$ is number of transport operations in both of two directions in a day, $d_{em}$ is amount of emission which can be reduced in a day, $rv_c$ is capacity per road freight vehicles, $r_d$ is distance of route, $r_{we}$ is emission factor per vehicle, $e_{mc}$ is marginal cost of emission per ton, $n_{day}$ is number of day in a year, $m_{ec}$ is total reduced marginal cost of road transportation annually. On the other hand, with the using of intermodal transport operations by trains, economic loses which causes from road freight transportation can be reduced. National and regional economy can be affected by using the fossil fuels as well as environment, social structure other assets. Road freight transportation is an unproductive and ineffective transportation type. Only, when considering the energy cost of road transportation, it can be seen that are at unsustainable levels. These negative effects cannot be eradicated completely, it may be reduced dramatically with the using of intermodal rail operations. If a portion of cargoes which transport by road vehicles can be shifted to rail transportation, energy cost may be reduced as seen below;

$$\sum re_c = ([r_{tl} \ x \ u_{fc} \ x \ f_{uc}] \ x \ r_{nv}) \ x \ n_{day} \qquad (6)$$

While $re_c$ is represents the reduced energy cost in total, $r_{tl}$ is the total distance of road, $u_{fc}$ is unit energy consumption per road vehicle, $f_{uc}$ is the monetary value per liter of fuel, $r_{nv}$ is the number of vehicles that will be reduced with the using of intermodal freight trains, $n_{day}$ is number of day in a year.

## 4. Conclusion

Economic and environmental impacts of road transportation is very high and it cannot be sustainable. With the using of intermodal freight rail operations, economic and environmental costs of road transport can be reduced dramatically. In this study, the effects of the intermodal freight rail operation and the key role of Belgrade as an intermodal freight terminal have been represented. If only 30% of cargo which transported by road vehicles can be shifted to railway, while environmental costs can be reduced as 4,03,657 dollars, saving on energy cost can be reached at 34,691,526 dollars. As a result, total saving can be reached at 39,395,183 dollars annually.

**References**

Abacoumkin C.; Ballis, A. 2004. Development of an expert system for the evaluation of conventional and innovative technologies in the intermodal transport area, *European Journal of Operational Research* (152): 410–419.

Ballis, A.; Golias J. 2002. Comparative evaluation of existing and innovative rail–road freight transport terminals, *Transportation Research Part A*, 36: 593–611.

Crainic T. G.; Kim K. H. 2007. Intermodal Transportation, C. Barnhart and G. Laporte (Eds.), Handbook in OR & MS, Vol. 14: 467.

European Commission Paper. 2007. Road transport: Reducing $CO_2$ emissions from vehicles,

Ilin, V.; Simic D. 2013. Logistics centres in serbia: framework and potential benefits of ICT application and integration, International Journal of Innovative Ideas, Vol. 13, (2): 2.

Manić A. 2012. Performance assessment of road transport network of the Republic of Serbia in a context of information scarcity, Dissertation for obtaining the Degree of Master in Complex Transport Infrastructure Systems, Technico Lisboa, Portugal p: 21.

Medar, O. Manojlović, A. 2011. Internatıonal road haulage ındustry in Serbia: crıtıcal ıssues analysıs, Journal of Applied Engineering Science, (9): 190.

Ristović A., 2011. Leading organizer for intermodal transport Freight Service Department, JSC “Serbian Railways” 1st South East Environment Workshop.

Statistical Office of the Republic of Serbia. International flows of goods by modes of transport on the territory of the Republic of Serbia, 2006.

Statistics of International Transporters Association (UND). 2013. Statistics related to transportation vehicles,

Tol, R.S.J. 2003. The marginal costs of carbon dioxide emissions: an assessment of The uncertainties, *Centre for Marine and Climate Research, Hamburg University, Hamburg, Germany,* Working Paper FNU-19: 1.

Woodburn A. 2012. Intermodal rail freight activity in Britain: Where has the growth come from?, *Research in Transportation Business & Management* (5): 16–26.

<http://ec.europa.eu/clima/policies/transport/vehicles/index_en.htm>.

<http://webrzs.stat.gov.rs/axd/en/drugastrana.php?Sifra=0005&izbor=odel&tab=49>.

<http://www.und.org.tr/tr/76/istatistikler>.

# ANAGENT-BASED MICRO-SIMULATION FRAMEWORK TO ASSESS THE IMPACT OF RIVER FLOODS ON TRANSPORTATION SYSTEMS: IMPLEMENTATION TRAJECTORY FOR AN ASSESSMENT IN THE BRUSSELS METROPOLITAN AREA

**Ismaïl Saadi[1], Hamed Eftekhar[1], Ahmed Mustafa[1], Jacques Teller[1], Mario Cools[1]**
[1]*LEMA, University of Liège, Chemin des Chevreuils 1, Bât. B52/3, 4000 Liège, Belgium*

**Abstract:** Several studies were made to assess the direct damage but few studies tackled the topic related to the assessment of indirect damage.This paper discusses the development of a large scale agent-based micro-simulation framework to assess the economicdamage caused by river floods. The advantage of this approach is that the economic loss can be assessed at individual level.For that purpose, relationship is established between costs and trips cancelation/duration growth. As the numerical simulations depend largely on the initial demand, the accuracy of the data is fundamental. The model is based upon several disaggregated data related to mobility behaviour of the Belgian citizens. The study area covers the Brussels metropolitanarea which was subjected to several floods in the past and may see the risk increasing in the coming decades.

**Keywords:** agent-based modelling, micro-simulation framework, economic loss, travel demand, flood risk

## 1. Introduction

Since their introduction about two decades ago, activity-based demand models (ABDM) become more popular due to regular improvements and increased capabilities for transportation analysis and forecasting(Vovsha *et al.,* 2002). ABDM are characterized by the presence of behavioural realism, integrity and allowance of interdependencies consistent with higher spatial and temporal resolution(Rasouli and Timmermans, 2014).These models areappropriate to be coupled with dynamic traffic assignment patterns especially those usingagent-based concepts(Hao *et al.,* 2010; Bekhor *et al.,* 2011). For thisreason, Multi-Agent Transport Simulation (MATSim)library is employedthanks to its modular approach,.xml input-output data structure simplifying daily agents ‘plans definition and its powerful mobility simulator (Balmer *et al.,* 2006). Recent approaches were attempted to set up large scale scenarios of traffic simulation through several cities in the world. Although purposes differ between the different scenarios (e.g. evacuation model, air pollution evaluation model, emission model, inundation catastroph model), all of them are related to traffic modelling(Hülsmann *et al.,* 2014; Hao *et al.,* 2010; Lämmel *et al.,* 2010; Murray-Tuite and Wolshon, 2013).
As regards economic impact assessment, several references identified two main components: direct and indirect damages,(see e.g.Ham *et al.,* 2005; Merz *et al.,* 2007; 2010). In addition to direct damages, indirect impact is also significant due to cause and effects relationships. Indirect economic effects of natural disasters like floods are mostly neglectedin existing integrated models (Safarzyńska *et al.,* 2013). Therefore, some effort should be done to tackle this topic especially with the progress in terms of micro-simulation modelling which open new research areas.

## 2. Problem statement

The main topic related to our case study is the accessibility measurement.Indeed, the transportation network links which form the supply sidemightbe subjected to disruptions caused by river floods. Hence disequilibrium occurs since the demand side is no more satisfied. Then traffic demand is auto-organizing itself to reach another equilibrium point.Different questions arise such as: How to quantify the mismatch between supply and demand sides to assess the impacts?What is the necessary period to reach the new equilibrium level?Answering these kinds of questions require a large scale agent-based micro-simulation model.
Besides, in the literature review about flood impact on transportation system, Sohn established an accessibility index integrating distance-decay effect and volume of traffic influence (Sohn, 2006). This index is checked before and after the disaster to allow the impact measurement. Another popular approach consists in link-based capacity-reduction (e.g. Sohn, 2006). This way is more realistic since the link is always physically available but only accessibility parameter is reduced. Sometimes, links could simply be removed from the network before running the model to estimate the impact on travel demand. However, this technique is not the most adapted for smoother analysis because it is synonym of 0% or 100% link capacity which may insert more uncertainties into the modelling process. Indeed, all the values between the intervals are implicitly excluded from the accessibility analysis.

[1]Corresponding author : ismail.saadi@doct.ulg.ac.be

## 3. Data issues

Different types of data are necessary for the modelling purposes. Synthetizingprocess of the population related to the study area is derived from the Belgian Daily Mobility Survey (BELDAM)(Cornelis *et al.*, 2012).It contains also travel diaries which are useful to feed an activity-based demand pattern. When agents are created depending on the considered study area, daily activity plans will be randomly generated and associated to the agents according to a clustering procedure. BELDAM sample represents approximately 0.02 % of the population in Belgium which is not necessarily the most appropriate.

General census survey(2001)is used in order to correct the previous dataset. However, a specific attention should be paid concerning census updating because of changes that have been occured in between 2001 and today.

Other data sources related to mobility behavior called "Enquêtesur les forces de travail" (*EFT*, 2013)contain useful information about modal choice. Data analysis clearly shows that car is the most significant transport mode (see Fig. 1).When performing a multinomial logistic model, important independent variables (education level, age, employment) indicate the most probable mode utilized by each single agent.

The network corresponding to the study area is freely providedfor whole Belgium(*Geofabrik*, 2014). However, extracting sub-regionslimited by the administrative boundaries is not feasible. Therefore, study area is simply specified using a bounding box with coordinates. The network includes 100,467 nodes and 232,715 links. According to researchers, quality is constantly increasing allowing the users to take the advantage of this open-source solution.

Finally, to simulate the flood effects on transportation system, a steady-state inundation map is intersected with the transportation network using ArcGIS Network Analyst Package. Another alternative but more elaborated would be to use inundation data as time series of flooding heights for $(x, y)$ coordinates with a given spatial and temporal resolution. Lämmel has clearly shown how such coupling process is suitable especially with scenarios related to flood events, tsunami (Lämmel *et al.*, 2010).However these data are available with difficulty because of hydrological models dependency.Therefore, to allow combination between mobility simulation and flood scenario, a steady-state inundation map will be considered first.

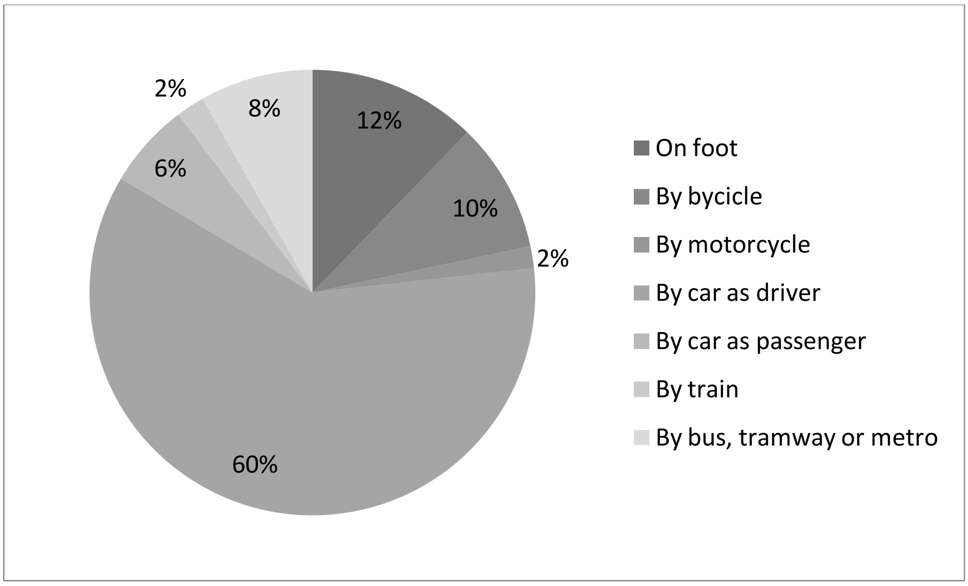



**Fig. 1.**
*Modal split in Belgium (2013)*
*Source: Enquête sur les forces de travail 2013*

## 4. Primary activity locations (residential/work-education locations)

Asafirst attempt, work/education and residential locations (primary activity locations) are randomly localizedinto each municipality. Indeed, asufficient level of accuracy related to locations of agent's activities cannot be provided by data. Therefore, the useof INS municipality codes is the classical approach. Afterwards, the agents are randomly scattered into each municipality according to weights. A similar approach was set up for the Zurich scenario (Axhausen *et al.*, 2008). From data point of view, BELDAM content indicates the address as strings which make unfortunately the location identification process very complicated and unexploitable. Concerning secondary activities,they are generally locatedwith regard to primary activity locations using a multinomial logit (MNL) model.

**5. Activity durations**

Activity durations are derived from a distribution taking into account socio-demographic characteristics. These are generally estimated by extracting them from travel diaries. In this study, only home-to-work trips were considered(see Fig. 2). BELDAM survey provides a trip-oriented structure related to each individual. So before establishing the working activity durations distribution, the structure should be converted into an activity-based oriented structure.After data cleaning process due to encoding mistakes and missing values, the determined distribution is used to assign durations to working activities of individuals.

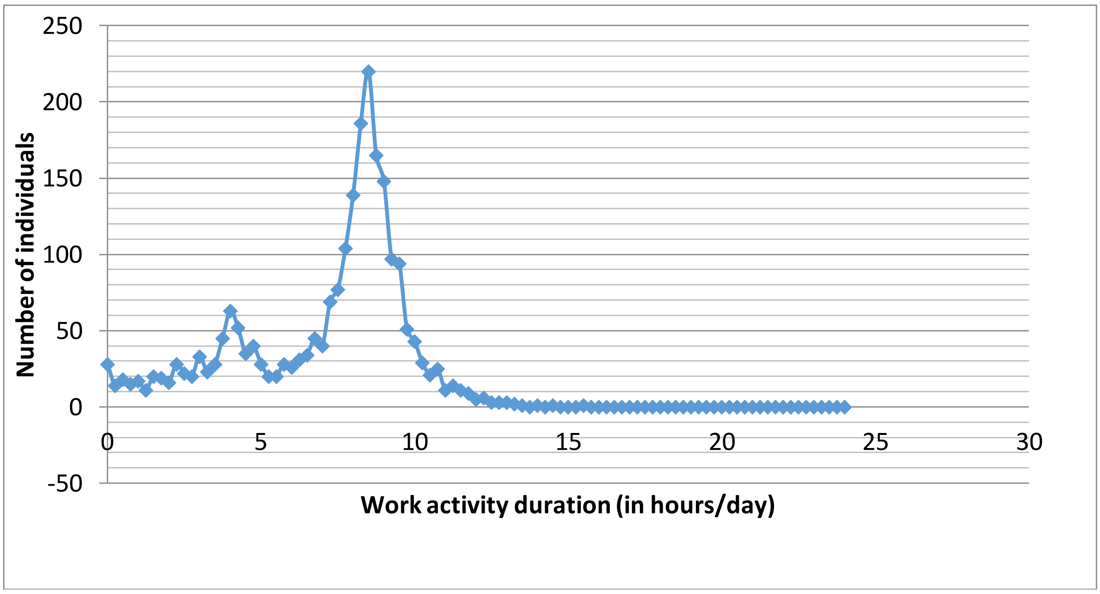



**Fig. 2.**
*Work activity duration distribution*
*Source: BELDAM Survey 2010*

Fig. 2 clearly shows the difference between the individuals working around 8 hours/day which corresponds tothe official number of working hours in Belgium and those who work half-time (about 4 hours/day).Also,there are some who work less than 4 hours, above 8 hoursor in betweendefining a distribution from 0 to about 15 hours/day. Finally, according to a random sampling based on the probabilities associated to their respective values, activity durations are determined.Afterwards, activity start times are deduced from the below distribution (Fig. 3).

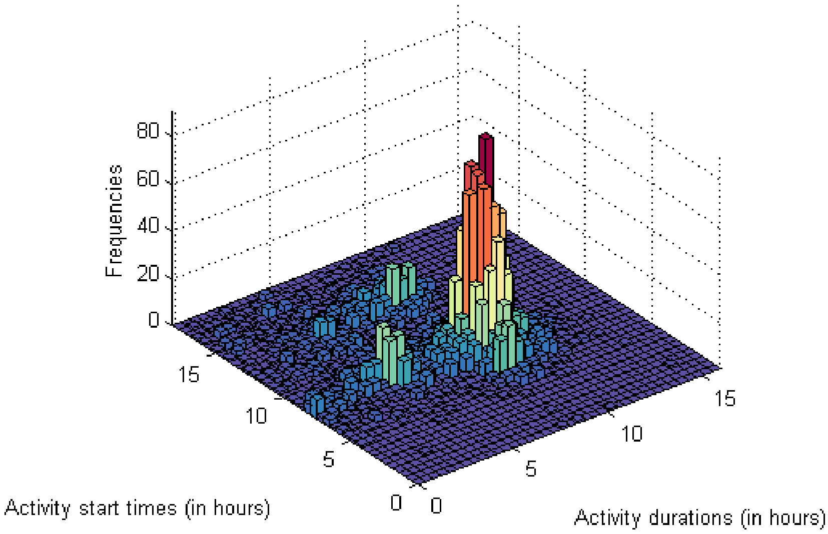



**Fig. 3.**
*Number of individuals'distribution according to activities start times and durations*
*Source: BELDAM Survey 2010*

## 6. Overall modelling framework

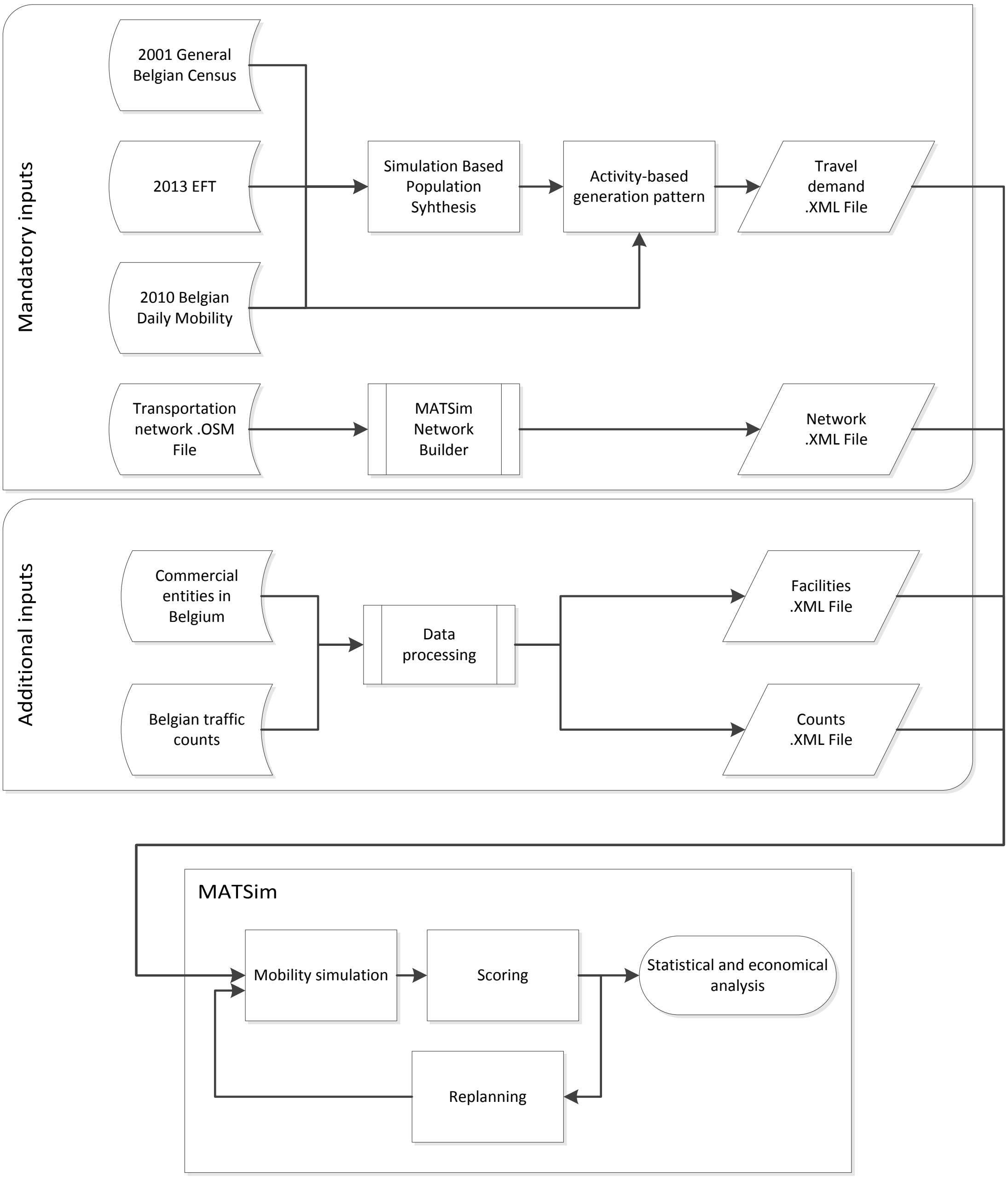



**Fig. 4.**
*General modelling framework*

In order to perform a traffic simulation, MATSim needs at least a network and a synthetic population as inputs(Balmer *et al.*, 2006). As a result, these are considered as mandatory contrary to the additional inputs which are not always available in terms of dataset. But their integration might improve the modelling processespecially validation issues. Indeed, *traffic counts* are generally used to check the simulated counts with reality however their availability depends on countries. In Belgium, the accessibility to such data is more difficult.

*Facilities* correspondto an aggregation of buildings in specific hectares. Activities performed by the agents can be associated with facilities instead of links directly. Facilities can be characterized with opening hours, activity types and locations. However, including the facilities option induces a deep data collection and preparation before being able to integrate them in MATSim. Therefore, only fundamental inputs are considered at the moment.

The *simulation based population synthesis* module is one of the most important in modelling process.It generates an arbitrary number of agents with specific attributes which distributions match those belonging to the real population(Farooq *et al.*, 2013).

Afterwards, agents are assignedwith activity chains through the *activity-based generation pattern*.

When travel demand and network files are converted into .xml format (readable by the library), MATSim uses the *mobility simulation* module to execute agents' plans and"load" the network taking into account spatial and temporal constraints. This step is also called traffic assignment.

In every iteration, daily agents' plans are scored with a utility function(Charypar and Nagel, 2005)to allow the population's optimization process. Indeed, an agent is constantly trying to maximize its utility by avoiding, for example, traffic congestion and late arrivals at scheduled activities.
Given the fact that MATSim is based on an iterative process, a feedback module called "Replanning" is also integrated. The module's role is to perform "minor" changes of agents' plans (e.g. transportation mode change, setting departure times and activity durations, rerouting, etc.). Three kinds of re-planning modules are available for agents. These last make their module choice according to probabilities fixed in the configuration file.
Finally, a statistical analysis tool provides numerous graphics, charts and other outputsrelated to computing performance, traffic counts statistics, absolute and relative errors. They are useful for the case study since they would be usedto assess the impact of floods before and after the disaster.

## 7. Computational issues

Envisaging a simulation integrating several millions of agents would require a high performance computing platform. Therefore, simulations are run usinga multi-core platform to allow parallelization. As stated by Raney et al., two indicators are suitable to assess computational performances: speed-up and real time ratios (RTR)(Raney *et al.,* 2003).Speed-up metric measures the relative improvement in terms of speed execution of a task. It is defined mathematically by the ratio:

$$S(n) = \frac{T(1)}{T(n)}$$

where $n$ is the number of CPU's, $S(n)$ is the speed-up ratio for a given number of CPU's $n$, $T(1)$ is the execution time of one CPU and $T(n)$ is the computation time needed for $n$ CPU's.
Real time ratio translates the speed of time in simulation compared to the reality. In practice, a RTR of 100 means that 100 hours real time traffic can be simulated in one hour by the computer.These ratios constitute important indicators for a micro-simulation especially to estimate the platform capacity. As a result, it's easier to determine the population sample size that should be simulated.

## 8. Discussion and conclusion

The main purpose of this paper is the establishment of a comprehensive framework which will be the basisforafuture research.The literature review groups all the agent-based micro-simulation models set up recently which common point is the use of MATSim (a dynamic traffic assignment simulation tool).It was underlined the necessity to take into account the assessment of economic impact due to river floods at the end of the modelling chain. By performing simulations before and after the disaster, real impact and changes in behavior of travel demand will be assessed.
Actually, the modular approach of the framework allows easy replacement of either the simulation based population synthesis module or the activity-based generation pattern as well as the internal components of MATSim. The efficient MATSim interfaceincluding .xml files make the computing process quicker.
As mentioned previously, xmlfiles readable by MATSim require huge efforts in terms of data preparation. It was shownthat several data sources are used to allow more disaggregation especially those from the 2001 Census which has the highest level of accuracy. However, an updating process should be performed to make data closer to the reality.
Specific approaches were proposed to determine primary activity locations and durations. However,efforts areconducted to improve the methodologies as well as the level of aggregation.
In addition, attention should be paid to calibration. MATSim includes a configuration file (config.xml) which contains all the parameters set by the user before running the simulations. These parameters are related to strategies, scoring utility function, mobility simulator, scenarioand controller.
Validation is amongst the most challenging step since, as mentioned previously, the classical way to validate such a model are traffic counts. However, it was stipulated that the access to such data is very limited in Belgium. Thereforeother validation ways should be tested like modal split which information can be extracted from BELDAM survey 2010or EFT 2013.
Besides, from a pure policy and practice point of view, the implementation of such an agent-based micro-simulation model at Brussels MetropolitanArea with a possible extension through the whole country is absolutely necessary. Particularlywhen policies in terms of transportation system and land use change should be made for future in the context of river floods.

## Acknowledgements

The research was funded through the ARC grant for Concerted Research Actions, financed by the Wallonia-Brussels Federation.

**References**


Axhausen, K.; Balmer, M.; Ciari, F. 2008.A new mode choice model for a multi-agent transport simulation.In: *8th Swiss Transport Research Conference. ETH*.

Balmer, M.; Axhausen, K.; Nagel, K. 2006. Agent-Based Demand-Modeling Framework for Large-Scale Microsimulations,*Transportation Research Record: Journal of the Transportation Research Board*, 1985: 125–134.

Bekhor, S.; Dobler, C.;Axhausen, K. 2011.Integration of Activity-Based and Agent-Based Models.*Transportation Research Record: Journal of the Transportation Research Board*, 2255: 38–47.

Charypar, D.; Nagel, K., 2005. Generating complete all-day activity plans with genetic algorithms. *Transportation*, 32(4): 369–397.

Cornelis, E.; Hubert, M.; Hunyen, P.; Lebrun, K.; Patriarche, G.; De Witte, A.; Creemers, L.; Declercq, K.; Janssens, D.; Castaigne, M.; Hollaert, L.;Walle, F. 2012. *Belgian Daily Mobility (BELDAM): Enquête sur la mobilité quotidienne des belges*, Brussels, SPF Mobilité & Transports.

Farooq, B.; Bierlaire, M.; Hurtubia, R.; Flötteröd, G. 2013. Simulation based population synthesis,*Transportation Research Part B: Methodological*, 58: 243–263.

Geofabrik.2014. Available from Internet: <http://download.geofabrik.de/europe/belgium.html>.

Ham, H.; Kim, T.J.; Boyce, D. 2005. Assessment of economic impacts from unexpected events with an interregional commodity flow and multimodal transportation network model,*Transportation Research Part A: Policy and Practice*, 39(10): 849–860.

Hao, J.; Hatzopoulou, M.; Miller, E. 2010.Integrating an Activity-Based Travel Demand Model with Dynamic Traffic Assignment and Emission Models,*Transportation Research Record: Journal of the Transportation Research Board*, 2176: 1–13.

Hülsmann, F.; Gerike, R.; Ketzel, M. 2014 Modelling traffic and air pollution in an integrated approach – the case of Munich,*Urban Climate*.

Lämmel, G.; Grether, D.; Nagel, K. 2010. The representation and implementation of time-dependent inundation in large-scale microscopic evacuation simulations,*Transportation Research Part C: Emerging Technologies*, 18(1): 84–98.

Merz, B.; Kreibich, H.; Schwarze, R.; Thieken, A. 2010. Review article ‘Assessment of economic flood damage’, *Natural Hazards and Earth System Sciences*, 10(8): 1697–1724.

Merz, B.; Thieken, A.H.; Gocht, M. 2007. Flood Risk Mapping at the Local Scale: Concepts and Challenges. In: S. Begum, M.J.F. Stive, and J.W. Hall, eds. *Flood Risk Management in Europe*. Springer Netherlands, 231–251.

Murray-Tuite, P. Wolshon, B. 2013. Evacuation transportation modeling: An overview of research, development, and practice,*Transportation Research Part C: Emerging Technologies*, 27: 25–45.

Raney, B.; Cetin, N.; Völlmy, A.; Vrtic, M.; Axhausen, K.; Nagel, K. 2003. An Agent-Based Microsimulation Model of Swiss Travel: First Results,*Networks and Spatial Economics*, 3(1): 23–41.

Rasouli, S.; Timmermans, H. 2014. Activity-based models of travel demand: promises, progress and prospects,*International Journal of Urban Sciences*, 18(1): 31–60.

Safarzyńska, K.; Brouwer, R.; Hofkes, M. 2013.Evolutionary modelling of the macro-economic impacts of catastrophic flood events,*Ecological Economics*, 88: 108–118.

Sohn, J. 2006. Evaluating the significance of highway network links under the flood damage: An accessibility approach,*Transportation Research Part A: Policy and Practice*, 40(6): 491–506.

SPF Economie. 2013. <http://statbel.fgov.be/fr/statistiques/collecte_donnees/enquetes/eft/>.

SPF Economie. 2011. <http://statbel.fgov.be/en/statistics/surveys-methodology/census/Census2011/>.

Vovsha, P.; Petersen, E.; Donnelly, R. 2002.Microsimulation in Travel Demand Modeling: Lessons Learned from the New York Best Practice Model,*Transportation Research Record: Journal of the Transportation Research Board*, 1805: 68–77.

# MODELLING VIBRATIONS INDUCED BY UNDERGROUND RAILWAYS IN URBAN AREAS

**Mauro D'Apuzzo[1], Manuela Esposito[2] , Bruna Festa[2], Vittorio Nicolosi[3]**
[1]*University of Cassino and Southern Lazio, Via G. Di Biasio 43, Cassino (FR) 03043, Italy.*
[2]*University of Naples "Federico II", Via Claudio 21, Napoli 80125, Italy.*
[3]*University of Rome "Tor Vergata", Via del Politecnico 1, Rome 01000, Italy.*

**Abstract:** Nowadays the problem of annoyance related to vibrations produced by human activity is increasingly felt, especially in residential areas. In addition, recent development of high-speed railways has made evident the disturbance and environmental damages caused by vibrations. In this study a mathematical prediction model, able to evaluate vibration level induced by underground railways, is presented. The model can by split into several sub-models. In the first one, namely the Generation Sub-Model, the dynamic interaction between rolling stock and railway superstructure is analyzed. In the second one, that can be named Propagation Sub-Model, interaction forces provided by the previous model are applied to tunnel invert and propagation of vibration waves is obtained by means of a finite elements method (F.E.M.) approach.
The validation of the model passed through the comparison between vibration level provided by numerical simulations and on-site measurement carried out in Naples subway, yielding a fairly good agreement.

**Keywords**: Vibrations, subway, mathematical prediction model.

## 1. Introduction

The problem of the environmental nuisance generated by human activities is strongly felt in correspondence of production facilities or other sources of noise and vibrations, such as railways, airports, roads with high traffic density, as they may cause malfunctioning of sensitive equipment, discomfort to people and damage to buildings located in the surroundings. As far as vibrations induced by railway traffic are concerned, there is therefore the need to develop sound and reliable prediction tools in order to evaluate in advance critical situations and to study suitable countermeasures. However it has to be highlighted that, despite the importance of the problem, there is not a common methodology to forecast and control, even if a growing number of case histories, often generated by legal disputes between railway companies and affected communities, is available.
The approaches so far followed to tackle the problem are mainly analytical on one hand and experimental, on the other. However both approaches may show some limits: analytical ones, that try to look through the physics of the phenomenon, may impose restriction on geometrical layout or on materials properties of the specific site to be examined not allowing an immediate solution for most practical situations, whereas experimental approaches, implying prediction tools empirically derived by experimental campaigns, may provide poor estimate for case studies other than those specifically investigated and therefore may suffer from a lack of generality, since it would be too expensive to generate an experimental database which is representative of every possible scenario.
The canonical modeling approach foresees the development of a three-stage model: the *Generation Model* describing the dynamic interaction between the rolling stock and the track, the *Propagation Model* where the propagation through the underlying structures and soils is dealt with and the *Reception Model* where the vibration felt by sensitive receptors are tackled.
In this paper an analytical-numerical prediction model to evaluate the vibration level induced by an underground railway is presented. The model is composed by a *Generation Model* able to estimate the dynamic vertical forces exerted by wheelsets as a results of the travelling of a rail vehicle on a railway superstructure with vertical defects. Dynamic forces are implemented in a more wide mathematical layout considering *Interaction Forces* between the track and the tunnel invert and *Frequency Response Functions* (FRF) of the underlying soils, in order to analyze the dynamic behavior of the system composed by the rolling stock, the defected railway track, the tunnel structure and the surroundings soils as a whole.
The model has been effectively calibrated and validated for the Turin's Underground in previous works (D'Apuzzo 2012) and, following a recalibration and a redesign, a similar attempt has been described in this paper with reference to Naples' Underground line.

## 2. Model's description

The model presented in this paper can be considered as hybrid since the generation sub-model, used to describe the dynamic interaction between the vehicle and the rail, is mainly analytical whereas the propagation one, used to describe the interaction between the rail and the underground soil, is numerical.

### 2.1. Generation model

Vehicles moving onto the rail surface are dynamically excited by the elevation's variations of the profile thus inducing vertical dynamic loads applied by the vehicle's wheelsets to the track. The study of the problem is tackled by solving

[2] Corresponding author: manuela.esposito@unina.it

the differential equation set governing the phenomenon. Vertical dynamic forces are therefore affected by the rail roughness and vehicle modeling assumptions.
As far as the rail defects are concerned, when it is not possible to carry out direct measurements (D'Apuzzo 2014) on the test site, the excitation, represented by the geometrical irregularities of the track, is usually described through a stationary and ergodic random processes and therefore *Power Spectral Density* (PSD) functions of vertical and transversal displacements are used. Examples of PSD models are available in the literature (Esveld, 2001) (Iwnicki, 2006) (Wu et Thompson, 2000). In this study, the PSD suggested by Frederich (Forrest, 1999) has been used:

$$S(f) = \frac{1}{v} \cdot \frac{a}{(b + {}^{f}/_{v})^3} \qquad (1)$$

where:

- *S(f)* is the PSD of the vertical rail defect as a function of temporal frequency, in $m^2$/(cycles/s),
- *f* is the temporal frequency of rail defects, in cycles/s,
- *v* is the velocity of the train, in m/s,
- *a* and *b* are model parameters depending on track age; in this work values equal to $9.93 \text{x} 10^{-1}$ and $6.89 \text{x} 10^{-2}$ have been chosen respectively.

Concerning the vehicle description, a *lumped mass model* where lumped mass and rigid bodies are interconnected between each other through elastic and viscous elements is usually employed. Depending on the detail of the model, it can be very simple (such as a one-degree-of-freedom system) or quite complex (up to several degrees-of-freedom) (Iwnicki, 2006) (Esveld, 2001). The model developed in this study is a ten degree-of-freedom model where the mass and the inertia moments of the car body and of the two bogies and the mass of the four wheelsets are taken into account. Primary and secondary suspension systems have been modeled by means of linear springs and dampers. The wheel-rail contact has been represented by means of a Hertzian spring whose stiffness can be evaluated as reported in (Esveld, 2001) (Aiello 2008).
This model, coupled with a deterministic or a stochastic representation of the road profile, can be used in order to evaluate the PSD or the frequency spectrum of vertical dynamic loads.

### 2.2. Propagation model

As far as the mathematical modeling of the propagation of ground-borne vibrations is concerned, by the advent of high performance computers, various numerical methods emerged, including the *Finite Element Method* (FEM), *Boundary Element Method* (BEM), and their variants (Beskos, 1997). The BEM approach takes into account the radiation damping by means of the use of fundamental solutions but it cannot effectively consider complex layouts in geometrical and mechanical properties of the underlying structures and soils. The FEM approach appears to be more versatile in applications, making easy the description of complex geometrical configurations including embedded structures and multi-layers soils, however, it suffers from the drawback that the soil, which is semi-infinite by nature, can only be modeled by elements of finite size. Consequently, the radiation damping that accounts for the loss of energy due to waves traveling to infinity cannot be accurately modeled (Yang et Hung, 2009). In order to correctly simulate this aspect, very complex boundary conditions have been developed by several authors (Sheng et al, 2005) (Degrande et al., 2005) or combined FEM-BEM approaches have been used, leading to the so-called hybrid methods.
In this study, in order to simulate and underground railway line, the propagation phenomenon is simulated by a transmission sub-model, emplying a FEM technique where nor particular boundary conditions neither BEM elements have been used. All the vibration waves directed to the external surfaces of the finite propagation medium are progressively attenuated within a damping zone (D'Apuzzo, 2007).
The geometry and layout of the 3D FEM model were decided following a calibration phase and according to the stratigraphic characteristics of the test site. The study area extends 5 m in radial direction, corresponding to a tuff uniform deposit with a pozzolana roof, and 10 m in the longitudinal direction starting from a vertical symmetry plane. The study area examined is conveniently wrapped in both directions by means of a surrounding damping area, which shows a gradual increase in the size of the mesh, in order to correctly simulate vibrations propagation (see figure 1).
The cylindrical surface is constrained in both radial and tangential directions, while constraints are also placed on the plane of symmetry along the z-axis in order to fulfill the requirement of symmetry. Mechanical properties of surrounding soils are derived from geotechnical and geophysical measurements, whereas energy dissipation phenomena due to material damping is modeled through the Rayleigh approach (Aiello 2008).
Pulse vertical forces have been applied at tunnel invert level on the plane of symmetry in order to derive the *Frequency Response Functions*, to be introduced in the equation set describing the rolling stock-track-tunnel structure- surrounding soil systems in the frequency domain.

### 2.3 Model's assembling

In order to study the interaction between the subgrade and the railway superstructure and between the rail vehicle and track a mathematical model representing the dynamic behavior of the track excited by the dynamic loads exerted by the vehicle is needed.

The superstructure can be described through continuous or discrete models. In the former case, the model is analytical and is developed using a Winkler beam on a continuous visco-elastic support or layered Winkler beams (Esveld, 2001) (Iwnicki, 2006). In the latter case FEM numerical models of the rail superstructure are usually employed. As previously highlighted, the use of FEM approach does not allow to take into account the infinite length of the railway and therefore the corresponding radiation damping. In this study, the superstructure has been therefore modeled as a Winkler beam resting on a continuous single layer of springs and dampers and *Frequency Response Functions* describing the dynamic interaction at the wheel-rail and at the rail-tunnel invert interface have been conveniently derived. It has to be noticed that it was necessary to evaluate a finite number of discrete interaction points at the rail-tunnel interface in order to tackle the dynamic behavior of the whole vehicle-track-tunnel system.

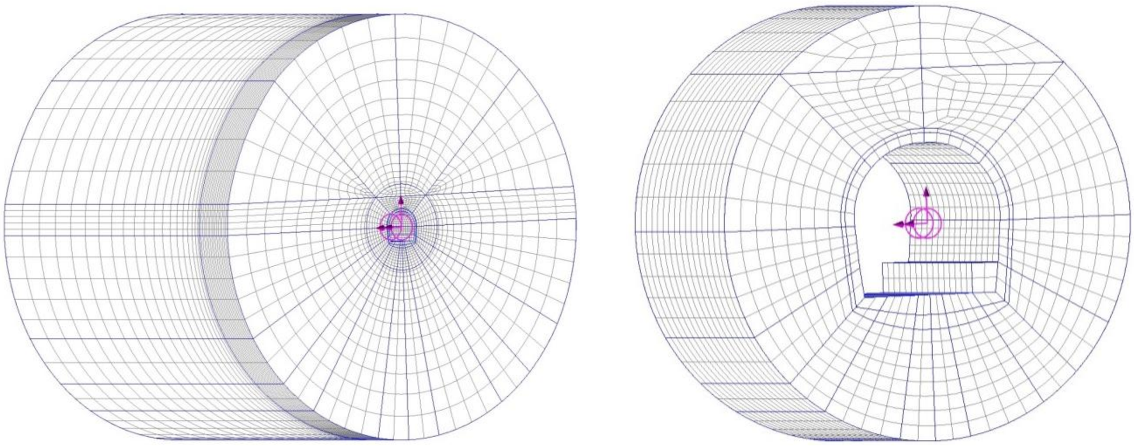


**Fig. 1.**
*3D FEM model developed for the propagation model: complete model(a) and a detail of the study area (b).*

By means of a Fourier Transform, the set of linear differential equations describing the vehicle moving on irregular road profile belonging to a track that is resting on a tunnel invert embedded in a soil formation is transformed into an algebraic set of equation and it is solved, by means of a symbolic math software. Therefore, complex transfer functions between rail profile vertical excitation and dynamic overloads, and between the track and the tunnel invert can be are algebraically derived (D'Apuzzo 2012).

Finally, the rail profile spectral function is combined with the vehicle rail and tunnel invert transfer functions and an Inverse Fast Fourier Transform is performed in order to evaluate the time histories of the vertical loads applied to the tunnel invert surface that will represent the input forces for the propagation sub-model. Since the train speed is well below the critical speed (where resonance due to moving trains can occur) the *Dynamical Reciprocal Rayleigh-Betti Theorem* can be invoked and the dynamic response can be evaluated by analyzing a moving observer with respect to fixed dynamic forces exerted at the wheel-rail interface (D'Apuzzo, 2007) (D'Apuzzo 2012).

The main advantage of this approach is that all the sub-models can be separately evaluated in order to extract the specific FRFs to be implemented in the aforementioned set of equation, providing a substantial decrease of the computational effort (D'Apuzzo 2012). A pictorial representation of the whole system is reported in the figure below together with a list of most significant input parameters.

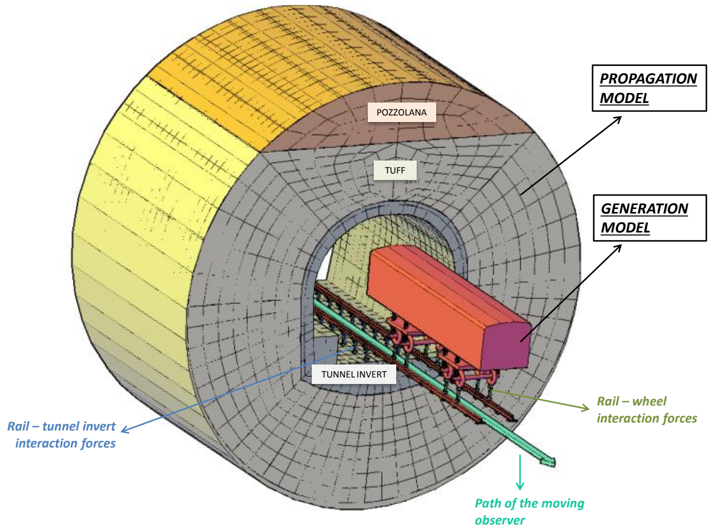



**Fig. 2.**
*Conceptual representation of the whole vehicle-rail-tunnel-soil system.*

| **VEHICLE** | |
|---|---|
| **Parameter** | **Value** |
| Total vehicle length (m) | 35.680 |
| Gauge (m) | 1.435 |
| Wheel diameter (m) | 0.820 |
| Wheel radius (m) | 0.410 |
| Wheel mass (kg) | 600 |
| Vehicle width (m) | 2.850 |
| Empty mass (kg) | 62500 |
| Full load mass (kg) | 92740 |
| Maximum axle load (N) | 115925 |
| Maximum wheel load (N) | 57962 |
| Vehicle speed (m/s) | 11.6 |
| Maximum speed (m/s) | 22.22 |
| **RAIL** | |
| **Parameter** | **Value** |
| Mass per unit length (kg/m) | 60.335 |
| Moment of Inertia ($m^4$) | 0.00003055 |
| Vertical stiffness per unit length ($N/m^2$) | $150 \times 10^6$ |
| Rail pad loss factor | 0.15 |
| Hertzian wheel stiffness (N/m) | $2.503 \times 10^9$ |

**Table 1.**
*Main characteristics of the rolling stock and the rail.*

**3. Test site and experimental campaign**

The whole Line 1 of the Underground of Naples, develops along a 13 km path, with steep gradients (with a maximum slope of almost 5.5%) imposed by topographic restraints and a dual tunnel layout. Their sections are made of cast reinforced concrete; it was excavated in a volcanic tuff formation overlaid by pozzolana.
The railway superstructure consists of a slab track system without sleepers, directly connected to the tunnel invert. The rail is fastened to the tunnel floor through a resilient system that is similar to that used for the Milan metro line. The rail pad is made up of several rubber and cork layers and a finishing layer of sand and epoxy resin mortar. The rail pad spacing is 0.75 m, whereas the rail type is a UIC60.
The rolling stock, made by FIREMA Trasporti, is composed of a 35.680 m long dual wagon system, with the possibility of a composition of three coupled traction units at most. Each wagon is connected, through a secondary suspension system, to two bogies and they are connected to the wheels through a primary suspension system. Main technical data are reported in Table 1.

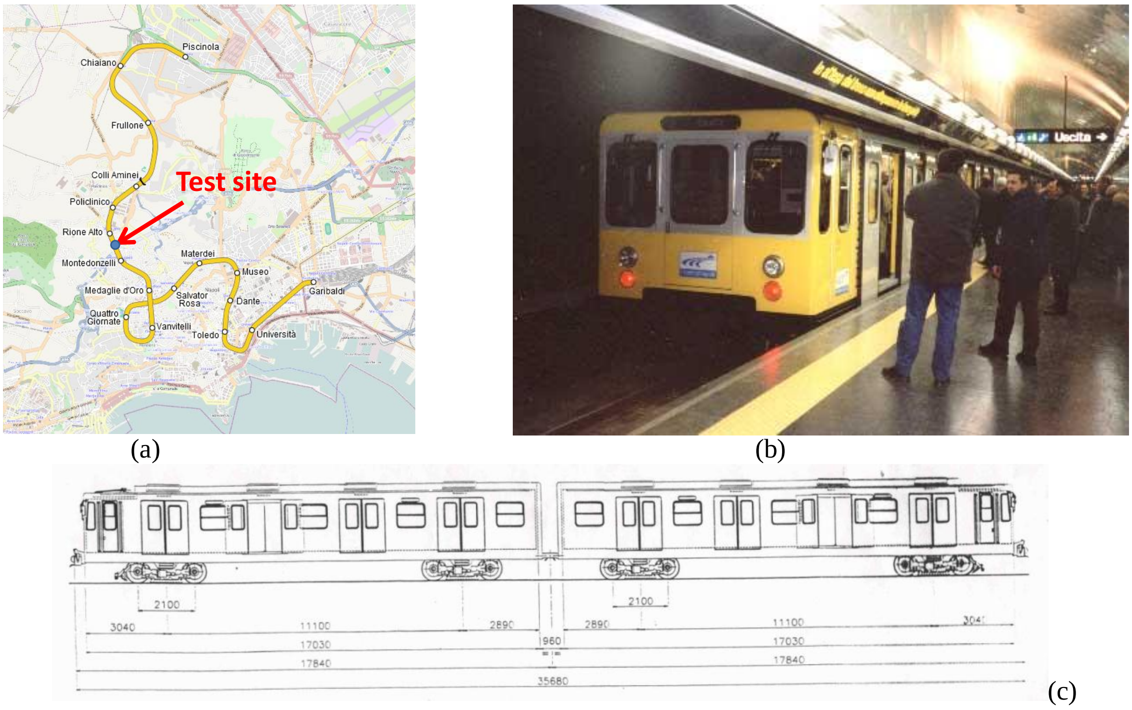



**Fig. 2.**
*Naples' underground Line 1: Plan view of the line and test site (a), trains in operation (b) View of the rolling stock (c).*

Following a corrugation problem occurring on the desceding track a vibration monitoring has been undertaken in the section between Montedonzelli and Rione Alto stations to evaluate the effectiveness of rail grinding. Before and after grinding vibration measurement have been carried out by TECNOIN in the 1996 on behalf of Naples's Underground Company namely MetroNapoli (Silvestri 2007) (Aiello 2008).
The experimental measurements were carried out by using vertical and horizontal velocimeters with a variable frequency response. Vibration measurements expressed in terms of vibration velocity have been carried out on two separate cross-sections in both tunnels and the sensors were positioned at the center of the section, close to the track and on the tunnel lining (see Figure 3). The sampling frequency was 200 Hz.
Therefore, it has been possible to derive the real train speed by analyzing the signal time lags. A statistical analysis has been carried out in order to identify the most significant vibration records that can be used in the following validation of the mathematical prediction model. By analyzing the vibration signals in the time domain, distinct peaks related to the passing of the train wheesets on the rail can be easily detected (Silvestri 2007) (Aiello 2008).

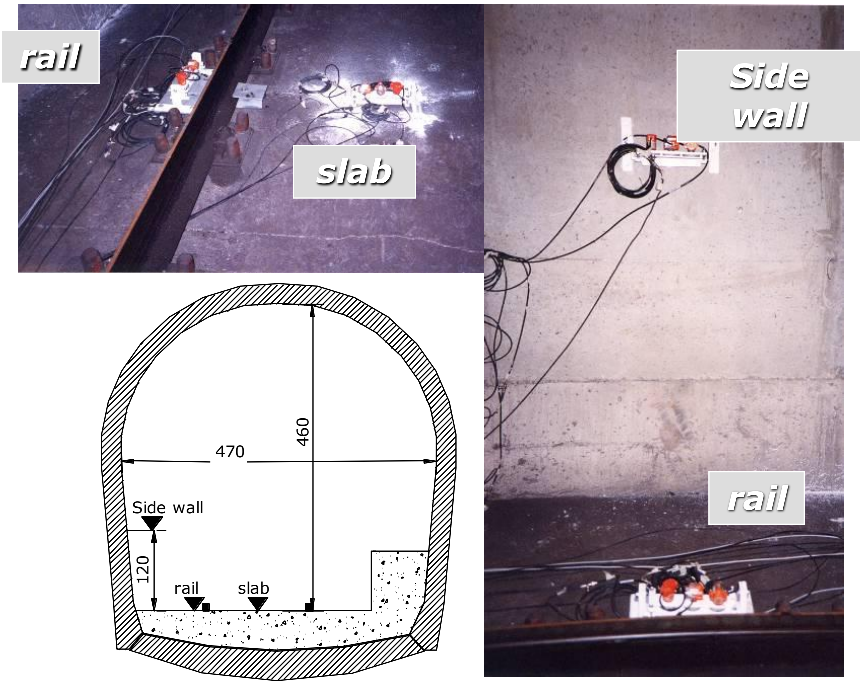



**Fig. 3.**
*A general view of the vibration monitoring test layout in the Naples' Underground Line 1.*

Following a legal dispute between a residential building and MetroNapoli, extensive geotechnical tests and soil characterization were carried out in the 2001 on a location not far away from the test site previously selected. Three geotechnical boreholes were drilled to a depth of 38 m, to carry out standard penetration tests and undisturbed sampling; two cone penetration tests were executed near the boreholes. Down-Hole (DH) tests were also performed in order to evaluate the shear wave velocity of the pyroclastic soils resting on the tuff formation where the tunnel is located. As it can be observed from the following figure (Figure 4) the resulting geomechanical model layout is rather complex. Geomaterials properties in terms of density, Young Modulus, and damping ratio derived from previously described geotechnical investigations and from technical literature where not available, have been implemented in the FEM propagation model in order to derive the relevant *Frequency Response Functions* to be introduced in the equation set describing the dynamic behavior of the vehicle-rail-tunnel-soil system.

## 4. Results analysis

As previously mentioned, the vibration prediction methodology presented in (D'Apuzzo 2012) has been conveniently modified in order to evaluate vibration induced by a Heavy Rail Transit (HRT) operating in Naples' area. A new vehicle model has been developed and railway track model has been re-calibrated, the FEM propagation model has been re-developed in order to take into account new geometrical layout and soil properties. Results provided by numerical simulations have been compared with experimental vibration level, measured by the velocimeter close to the tunnel invert. The comparison has been studied both in time and frequency domain. A maximum frequency of 1000 Hz has used in the numerical simulation therefore according to the Nyquist-Shannon sampling theorem, a limit frequency of 500 Hz can be investigated. Unfortunately, sampling frequency employed in the measurement is 200 Hz, and therefore only the range between 0 and 100 Hz can be analyzed in the frequency domain. Comparison in terms of vertical vibration velocity between the dynamic response provided by the model and the vibration level measured on test site is depicted in the figure 5.

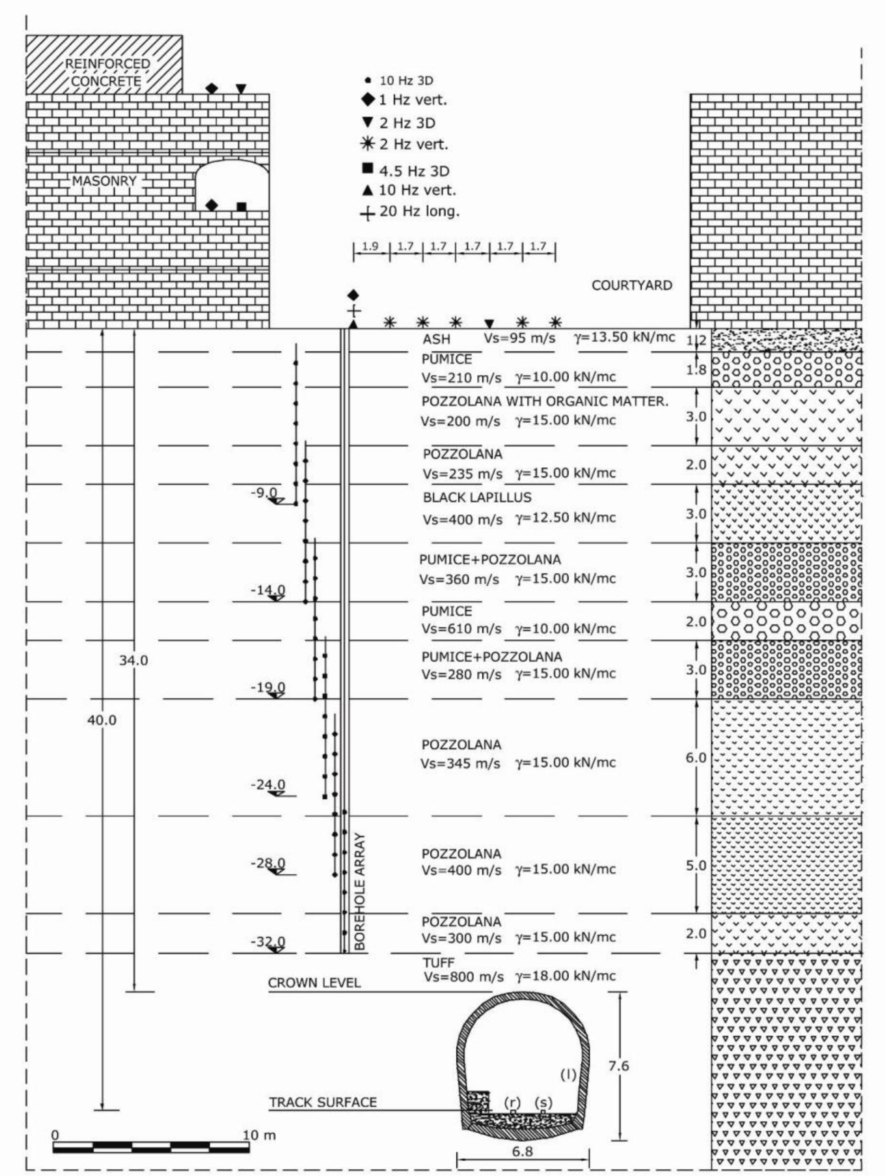



**Fig. 4**

*Test site: cross section with relevant values of various soils. Source: (Silvestri 2007) (Aiello 2008)*

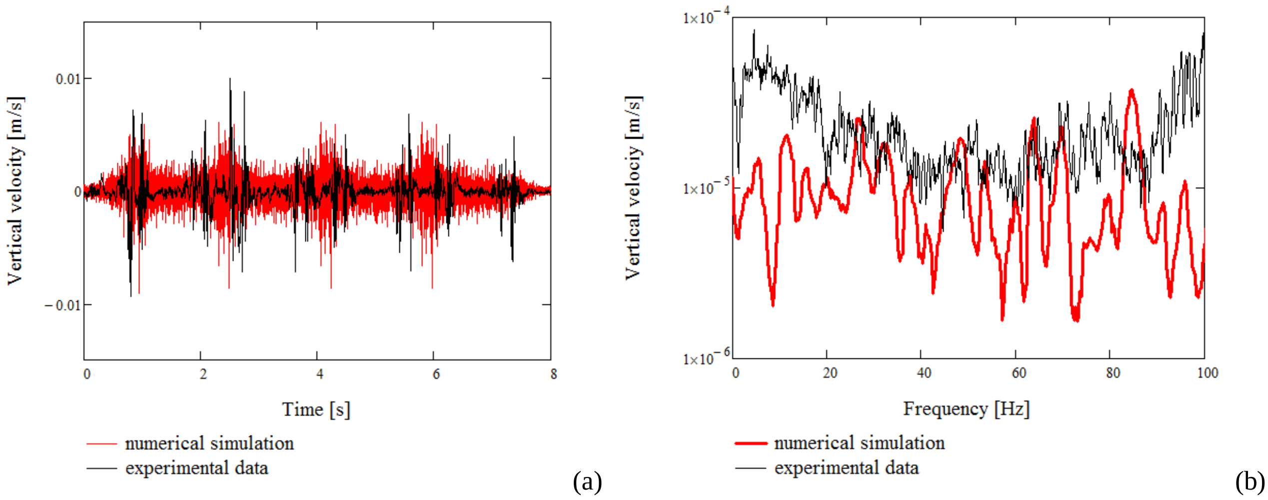



**Fig. 5.**

*Comparison of dynamic response between experimental data and numerical simulation in the time (a) and frequency (b) domain.*

As it can be observed from the figure, a fairly good agreement between numerical simulation and experimental data can be identified in the time domain: a peak in numerical simulation can be observed whenever a bogie transit is recorded by experimental measurements and corresponding maximum vibration velocity amplitudes seem comparable. A noise related to high frequency content is also clearly detectable in the numerical results but, as previously underlined, it may be ascribed to the higher sampling frequency (i.e. the lower integration time step) used in the numerical simulations.
In the frequency domain the comparison is less acceptable, providing the evidence that a more in-depth analysis of vehicle, track and rail roughness parameters should be undertaken in order to improve the agreement between numerical and experimental results.

## 5. Conclusion

In this paper preliminary results provided by a hybrid mathematical prediction model developed to evaluate the vibration level induced by railway traffic is presented. The model is named "hybrid" since it is composed by an analytical generation sub-model and by a numerical FEM propagation sub-model and it is derived from a previous work (D'Apuzzo 2012) where it has been used to effectively evaluate vibration level produced by train of the Turin's Underground.
In order to adapt the model to the Naples' Underground experimental test site, a new vehicle model has been developed and railway track model has been re-calibrated. Furthermore the FEM propagation sub-model has been re-designed in order to take into account new geometrical layout and soil properties.
Results provided by numerical simulations have been compared with experimental vibration level, measured by the velocimeter close to the tunnel invert.
Comparison in the time domain seems fairly satisfactory apart from an high-frequency noise probably related to the lower integration step employed in numerical simulations, whereas in the frequency domain numerical results seem to somehow underestimate experimental data. Basing on these preliminary results, it is acknowledged that further analysis and refinements are needed in order to improve the prediction of the vibration level in the frequency domain.

## Acknowledgements

Prof. Francesco Silvestri of University of Naples "Federico II", Prof. Laura Valentina Socco and Eng. Davide Boiero of Turin Polytechnic, Eng. Vincenzo Aiello, Eng. Elena Bogazzi, TECNO IN staff and Metronapoli staff are gratefully acknowledged.

## References

Aiello V., Boiero D., D'Apuzzo M., Socco L.V. & Silvestri F. 2008.Experimental and numerical analysis of vibrations induced by underground trains in urban environment. *Structural Control & Health Monitoring*, vol. 15, p. 315–348, ISSN: 1545-2255, doi: 10.1002/stc.247.

Beskos, D. E. 1997. Boundary element methods in dynamic analysis: Part II (1986-1996), *Appl. Mech. Rev.*, 50(3), pp. 149-197.

D'Apuzzo M., Nicolosi V. & Bogazzi E. 2012. A Unified Approach for the Prediction of Vibration Induced by Underground Metro. *Procedia: Social & Behavioral Sciences*, vol. 53, p. 62–71, ISSN: 1877-0428, doi: 10.1016/j.sbspro.2012.09.860.

D'Apuzzo, 2007. Some remarks on the prediction of road traffic induced ground-borne vibrations. *Proceedings of the 4th International SIIV Congress*. Palermo. 12th-14th September 2007. ISBN/ISSN: 13 978-88-8207-260-5. PALERMO: GRAFILL S.r.l.

D'Apuzzo, M., , Festa B . and Miraglia E .; 2014. Evaluation of rail defects provided by high performance measuring devices *Sustainability, Eco-efficiency, and Conservation in Transportation Infrastructure Asset Management* Edited by Massimo Losa and Tom Papagiannakis CRC Press pp. 649–Print ISBN: 978-1-138-00147-eBook ISBN: 978-1-315-75712-4 doi: 10.1201/b16730-95.

Degrande, G.; Clouteau, D.; Othman, R.; Arnst, M.; Chebli, H.; Klein, R.; Chatterjee, P.; Janssens, B. 2005. A numerical model for ground-borne vibrations from underground railway traffic based on a periodic finite element–boundary element formulation, *Journal of Sound and Vibration* 293 (2006) 645–666.

Esveld, C. 2001. *Modern railway track: second edition*, MRT-Productions, Zaltbommel, The Netherlands. pp. 1-654.
Forrest, J. A. 1999. *Modelling of Ground Vibration from underground railways. Doctorate thesis*. University of Cambridge.

Iwnicki, S. 2006. *Handbook of railway vehicle dynamics*, CRC Press, Boca-Raton, Florida, 1-526.

Sheng, X.; Jones, C. J. C.; Thompson, D. J. 2005. Prediction of ground vibration from trains using the wavenumber finite and boundary element methods, *Journal of Sound and Vibration* 293 (2006) 575–586.

Silvestri F., Aiello V., Socco V., Boiero D., D'Apuzzo M.; 2007. Experimental measurements and numerical simulations of underground railway ground-borne vibrations. *Proceedings of the 4th International SIIV Congress*. Palermo. 12th-14th September 2007. ISBN/ISSN: 13 978-88-8207-260-5. PALERMO: GRAFILL S.r.l.

Wu, T. X.; Thompson, D. J. 2000. Theoretical investigation of wheel/rail non-linear interaction due to roughness excitation, *Vehicle System Dynamics*, 34, 261-282.

Yang, Y. B.; Hung, H. H. 2009. *Wave propagation for train-induced vibrations. A Finite/Infinite Element Approach*. World Scientific Publishing Co. Pte. Ltd. 471p.

# THE IMPACT OF TRANSPORT POLLUTANTS ON URBAN GROUND AND SOME POSSIBLE MITIGATION MEASURES

**Stojanović Đurđica [1], Veličković Marko[2], Jevđenić Aleksandar[3]**
*[1,2] Faculty of Technical Sciences, University of Novi Sad, Trg D. Obradovića 6, 21000 Novi Sad, Serbia,*
*[3] The Institute for urban planning in Novi Sad, Bul. cara Lazara 3/3, 21000 Novi Sad, Serbia*

**Abstract:** Transport negative impacts on environment are a hot topic among academics and practitioners for years. The most important transport environmental impacts relate to climate change, air quality, noise, soil quality, biodiversity and land take. In the past, the majority of transport environmental research is focused on impacts on air and noise. Just recently, the theory and praxis have started to be more focused on transport impacts on ground. It is particularly important to consider this impact in the cities, where the most of human population live today.
We explored transport environmental impact on urban subsurface, and briefly described the main sources of urban transport ground pollution, quantifying and prevention related problems, and some possible measures to overcome them by using a novel knowledge in theory and praxis. A need for systematic monitoring of urban transport negative impact on land is highlighted. The main opportunities and obstacles to decrease transport impact on ground in Serbian cities are presented and a list of related priority measures is proposed. The measures which support better control and prevention of ground pollution from transport activities and infrastructure have to be an integral part of urban sustainable transport development.

**Keywords**: Urban transport, Urban ground, pollutants, sustainable development, protection

## 1. Introduction

Transport negative impact on environment is a hot topic for years both among academics and practitioners. The most important impacts of transport on the environment relate to climate change, air quality, noise and ground. Transport impacts on ground are numerous and complex. Among the most important on the list are impacts on water quality, soil quality, biodiversity and land use. The common characteristic is that all impacts are long-term and hardly measurable. They are also underresearched, if compared with other transport environmental impacts. Just recently, a significantly rising interest for this topic has been recorded.

The effects of these negative impacts are particularly important in cities, where the majority of human population lives. Transport activities are of vital interest for cities, while transport infrastructure has a significant share in urban land use. There is also a reverse relationship - city needs impact on growing levels of motorization and congestion. As a result, transport has a strong environmental impact in cities.

While a main body of the literature is concerned with transport environmental impacts on air quality and climate change, the impact on soil and groundwater, particularly in urban environment, has been much less underpinned. Fortunately, a rising interest for this topic has been recorded in the recent years.

The objective of this paper is to underpin a need for better exploring of transport impact on ground in Serbian cities and indicate a path for it, relying on some examples of good praxis in European cities. In the second Section, a complex transport impact on ground is presented. Third Section presents a contemporary approach and some measures related with the best practices in decreasing the negative transport impact on urban ground. Fourth Section presents the case study - the current state in Novi Sad related with transport impact on ground, recent research and available database. Some measures are proposed in fifth Section, while final remarks and conclusion are given in the last Section.

## 2. Transport impacts on ground

Transport impacts on ground are various and complex, usually with long-term and cumulative effects. They may be from a point source, e.g. in case of traffic accidents with hazardous materials, or may cause diffuse pollution and widespread disorders. Latter is more important, because it is more common, usually with long-term, and hardly measurable and predictable effects.

Transport is one of the main contributor of urban diffuse pollution, which contains a complex mixture of materials (Pittner and Allerton, 2009). Each pollutant can arise from multiple sources and each source can produce more than one pollutant (Mitchell, *et al*., 2001). Infrastructure construction and maintaining, vehicles, airplanes and ships exploitation, terminal and warehouse operations all impact on pollution of air, water, and soil. Vehicles and roads contribute pollutants through abrasion of road surfaces, corrosion of metal vehicle parts (e.g. brake linings), application of road de-icing materials, and atmospheric emissions (Mitchell, *et al*., 2001).

[1] Corresponding author: djurdja@uns.ac.rs

These impacts can be direct or indirect, e.g. conversion of atmospheric emissions to the runoff. The main sources and pathways of transport pollutants in ground are shown in Figure 1.

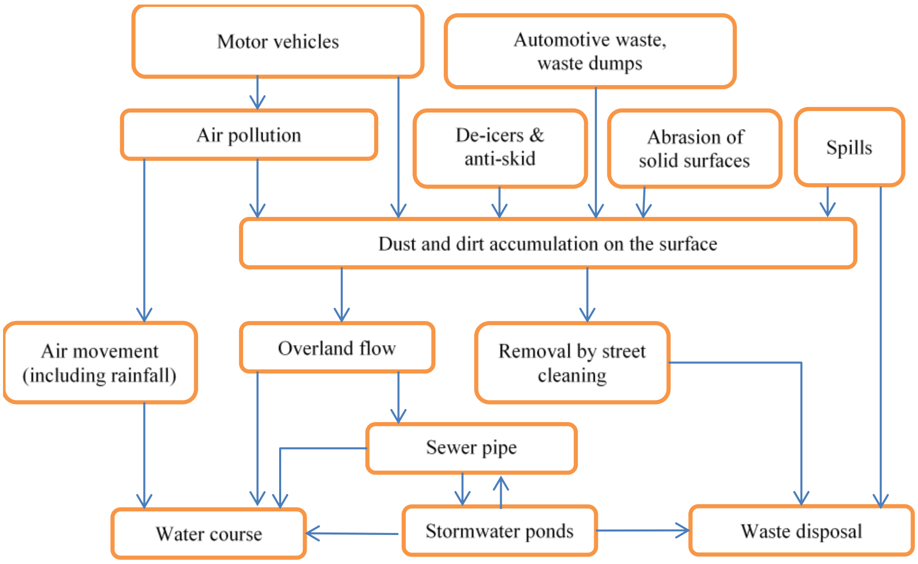



**Fig. 1**
*Transport impacts on ground*
*Source: adapted from Geiger et al. (1987, p. 65)*

Critical transport pollutants are heavy metals, hydrocarbons and salt. The US Environmental Protection Agency (EPA) particularly points out the importance of nitrogen in transport impact on water pollution, which concentrations at outfalls often exceed acute toxicity levels (Lehner, *et al.*, 1999).

Main transport impacts on land use, soil and water will be briefly and separately discussed in the rest of this Section.

**Land use**
Main direct impact is land occupation. Road pavements and other infrastructure associated with vehicular movements (e.g. car parks) can contribute as much as 70% of the total impervious areas in cities and are all recognized as sources of various pollutants to the ground quality (Pittner and Allerton, 2009). Change of land use has an implied negative impact on biodiversity, ecological fragmentation, and ground on a whole.

**Soil quality and biodiversity**
As noted, all described sources and their pollutants impact on roads and land around, and then through the runoffs, on soil and groundwater. Pollutants can be found in all areas surrounding the railroads, ports, airports and other transport infrastructure (Rodrigue, *et al*., 2013).

Transportation also influences natural vegetation and fauna. Many transport routes in urban or suburban areas have required draining land, thus reducing wetland areas and so changing complete ecosystems. Transport may cause introduction of invasive species, or extinction of existing species as a result of changes in their natural habitats (Rodrigue, *et al*., 2013).

**Water quality**
Transport impact on water quality is twofold. First, it impacts negatively on inland waterways and oceans, particularly near the ports, where dredging additionally impacts directly on the bed of a body of water. The second impact is on ground water through the runoffs and it is in the focus of this paper.

Runoff pollution occurs when rain or snowmelt carries away contaminants from rooftops, roads, parking lots, buildings, parks, and other surfaces in cities (Lehner, *et al*., 1999). Streets usually make the majority of impervious surface in the cities, with the highest pollutant loads in most land use categories. Unpaved roads, the dust that collects on paved streets, residential car washing with oil, grease, grit, and detergents also contribute to polluted runoff. According to some estimation, transportation related land use has the second highest level of pollutant concentrations; only piped industrial sources were higher (Lehner, *et al*., 1999).

Industrial salt (sodium chloride), used for deicing, enters into the environment easily in snowmelt runoff. Salt can travel long distances without assimilation and affects terrestrial and aquatic ecosystems (Helmreich, *et al*., 2010).

Driving a car or truck contributes a number of different types of pollutants to urban runoff. Pollutants are derived from automotive fluids, deterioration of parts, and vehicle exhaust. Once these pollutants are deposited onto road and parking surfaces, they are available for transport in runoff to receiving waters during storm events. (Lehner, *et al.*, 1999). Urban stormwater is an important contributor to water pollution because in most areas, stormwater receives no treatment before entering waterbodies.

## 3. Contemporary knowledge and praxis in mitigation of negative transport impacts on urban ground

It is surprisingly how much transport environmental impact on ground is neglected by transport experts. Research usually misses this aspect; even in case when all books are dedicated to this question (e.g. see McKinnon, *et al*., 2010). Practical directions about quantifying related external costs are also scarce, or completely missed, particularly for urban areas. The cited reasons for this situation are (Maibach, *et al*., 2008):

- Complex impact patterns and uncertain valuation approaches for other environmental costs such as nature and landscape, soil and water pollution, costs in sensitive areas.
- No direct relation to infrastructure use and thus to infrastructure pricing, such as costs for infrastructure related nature and landscape and urban areas.
- Difficult allocation to the transport system, such as costs of up- and downstream processes and costs of energy dependency.

Although this explanation is very reasonable, it may be discussed. We are not sure that it was much easier to estimate urban transport impact e.g. on global climate changes and related external costs which belong to local community. However, such estimations are widespread and widely used in the literature and praxis. Further, for local community ground pollution is more important than climate changes. It is clear that increased urbanization through time has led to the modification or deterioration of waterways and groundwater through release of untreated pollutants and resource depletion. Therefore, it is more important to estimate external costs related with ground pollution on local and national level.

For decades, even in most developed countries, water pollution mitigation measures have been focused primarily on water discharges processes from facilities, with less emphasis on diffuse sources. But from recently, this trend has been changed. EPA has started to consider diffuse pollution, particularly urban stormwater pollution, to be the most important source of contamination in national waters, and particularly important for waters near cities (Lehner, *et al*., 1999). In the same time, the sustainable urban road drainage systems have also become a hot topic in European countries (Pittner and Allerton, 2009).

Traditional goal of road drainage is to maximize the rate of flow from the road surface and to separate sub-surface from surface drainage. Thus, it ensures clear safe path for road traffic. Related road surface water drainage techniques involved collecting runoff in collection devices and conveying it to underground closed pipe systems. Historically, many road and other surface water drainage systems in urban areas were combined with foul sewage in a single combined sewer. In case of storms and high quantity of rainfall, overland flows would enter the combined sewer network, leading to spills of untreated sewage into receiving watercourses via combined sewer overflows.

The concentration of pollutants in road runoff can be highly variable and dependent on a wide variety of factors including location, traffic volumes, extent of dry period before a rainfall event, frequency of sweeping and nature of the road surface (Pittner and Allerton, 2009). Therefore, to develop the best measures against diffuse urban pollution, an effective and efficient system of monitoring has to be developed as a necessary precondition.

Sustainable Urban Drainage Systems (SUDS) for roads is a new concept which assumes a system of surface water management designed to drain surface water in a sustainable fashion. Their aim is to mitigate the extent of potential surface water flooding through the control of surface water runoff at source.

UK represents a very good example regarding developed European legislative framework related with SUDS. Recent legislative changes in UK require that surface water runoff from areas constructed after 2007 must be drained by SUDS. It means that all reasonable steps have to be taken to ensure the discharge will not result in the pollution of the receiving water environment. This contemporary approach assumes the use of separate surface and foul water drainage systems to reduce this burden. In fact, SUDS for roads have the fundamental principle to reduce the flow and to dispose of the water to groundwater wherever possible, with minimal impact on environment (Pittner and Allerton, 2009).

This is very important concept considering current climate changes. Our climate is changing and among the identified risks are the dangers of increased frequency, duration and magnitude of storms and increased storm surges. Such extreme events may lead to sewers and drainage systems becoming exceeded. Traditional road drainage systems are usually not designed to cope with extreme rainfall events and flooding occurs in such conditions. Therefore, new solutions have been developed to cope with climate change and expected higher and more erratic rainfall.

## 4. The main opportunities and obstacles to decrease transport impact on ground in Serbian cities - example of Novi Sad

According to urban planning objectives and goals in Novi Sad, soil is a natural resource of vital interest and its long-term usage and reproduction will be obtained by planning usage of surfaces according to its main geomorphological, pedological and hydrological characteristics.

In Novi Sad, in most cases it is implemented closed storm sewer, while some new plans include in some cases combined closed sanitary and storm sewer. In less percent, mostly in the suburban area, the path of the rainoffs includes open storm sewer and infiltration. The treatment of waste water still does not exist, although it is planned for the forthcoming period. It is also planned a systematic monitoring of open flows and groundwater quality as a part of monitoring system in the future and additional regulations which more precisely define maximal concentrations of polluters and conditions for discharging waste water.

According to precipitation data taken from the Meteorological Annual Reports 1950-2013 (source: Republic Hydrometeorological Service of Serbia), both annual average precipitation and maximum daily precipitation have a rising trend in last 5 decades in Novi Sad, which might be a consequence of climate change. This trend, together with the floods in Serbia in spring 2014, support a viewpoint that storms may become more and more extreme and severe in our country. This fact indicates the rising risk of flooding and emphasizes a need that our urban traffic and transportation experts pay more attention on their relation. The increasing of precipitation and climate changes is not included into the traffic planning and development.

Novi Sad has a high level of groundwater, which is continuously monitored. The natural flows are seriously disrupted by anthropogenic activities during city development. The groundwater characteristics are necessary to know in case of implementation of contemporary drainage measures. City plans for environmental protection, e.g. protection of groundwater includes control of polluters and spot pollution in the forthcoming period, but there is any goal, recommendation, or defined measure to include into control urban traffic diffuse pollution.

Novi Sad has recorded an increasing level of motorisation and mobility in the recent period - 4,5% annually, and this trend seems to be continued (JP Urbanizam, 2009). Therefore, the traffic impacts on land use and pollution will be continuously increasing. It is very important in such circumstances to adjust traffic plans so they include the possibilities to monitor and hedge diffuse pollution.

Regarding traffic and transportation, there is developed a comprehensive traffic model NOSTRAM, which allows replication of traffic flows in the past and prediction of future flows and modal shifts. It is considered as a basic starting point for future sustainable traffic and transport development in Novi Sad.

Until now, its outputs were used for some estimation of traffic air pollution and noise mapping by urban planners. The transport researchers also use the model results to underpin urban transport impact on air pollution and climate changes (e.g. see Veličković *et al*., 2014), and, in less extent, urban noise (e.g. Stojanović *et al.,* 2011) in Novi Sad, while research related with impact on ground is neglected.

However, ground pollution of agriculture and non-agriculture land has been measured in the city from recently. The City Council responsible for urban planning, housing and environmental protection have started to continuously monitor the quality of agricultural soil in the city since 2003. Since 2009, the reports have also included some samples from the city green areas (parks and the river beach). Annual reports from 2008-2011 (the last published one was in 2011) reveal that the concentrations of heavy metals Pb, Co, Cu, Cr, Ni, Cd, Mn, Fe and Zn, as well as of PAHs (Polycyclic aromatic hydrocarbons) in surrounding of the roads meet legal standards - if they exist (Vasin *et al*., 2011). It is hard to determine the impact of transport on heavy metals concentration in observed areas, because some of them may be the result of agriculture measures, usage of waste and industrial water and other anthropogenic factors.

Additionally, recently published research about transport impact on ground pollution in Novi Sad (Vasić *et al*., 2012) explores the presence of metal pollutants nearby the crossroads. Their research reveals that regardless of traffic volume, metal concentration in snowmelt, are small, except Na and Zn. The high presence of Na may be explained by de-icing salt. However, some questions may arise about the applied methodology and its limitations; therefore, the research about urban transport impact on ground pollution has to be continued to obtain more comprehensive and more reliable results.

## 5. Priority measures

Transport has a strong environmental impact in cities. Considering sustainable development, urban transport planners have been more focused on land use than on transport impact on urban ground for a long time. Regarding environmental impact on ground, our cities are far away from the best practice in sustainable urban development. As far as we know, this research is one of the first attempts in our country to highlight the importance of sustainable approach in developing urban traffic infrastructure from the ground pollution perspective.

It is not quite clear why urban ground pollution, and particularly diffuse pollution, has been neglected so far in transport literature and policy. Transport is always cited as one of the main contributor of urban diffuse pollution, but it is really hard to find any evidence, assessment, or estimation attempts about its share in numbers. It is the case even with the most developed countries.

The outputs of few available researches on transport impact on pollutants in ground in Novi Sad only open the questions and confirm a need for more serious research and systematic approach. The diffuse transport pollution as an important source is not clearly defined and considered in city plans. Consequently, there are not defined appropriate measures for its mitigation. However, city municipalities, stakeholders and urban planners in Novi Sad have started to recognize the emerged importance of the topic. The research in this paper is an attempt to contribute the discussion about this topic.

Sustainable urban transport policy should include tools for managing surface runoff pollution and water flood risks in a sustainable way. Some important tasks and challenges for urban traffic planners who wish to contribute to the development of local sustainable urban transport plans from the urban ground pollution perspective are related with:

- assessment the role of traffic impact on runoff pollutants in complex urban environment
- selection of indicators which have to be measured
- development of monitoring system
- benchmarking
- long-term audit of contamination sources and estimation of contaminants travel paths and times.

These are complex tasks which require a hard, long-term and interdisciplinary work. Various experts have to be included to obtain a holistic approach to sustainable urban development. Traffic planners have to cooperate with civil engineers, ecologists, hydro geologists, and other experts to develop a qualitative database for sustainable drainage.

Drivers and challenges for applying foreign experiences related with applying contemporary measures for decreasing urban ground pollutions are briefly presented in Table 1.

**Table 1**
*Drivers and challenges for applying contemporary measures*

| Drivers toward sustainable drainage in Serbia | Challenges |
|---|---|
| • EU standards and directives<br>• Climate changes<br>• Following best praxis | • Lack of knowledge about particular transport contribution to urban ground pollution; lack of integrated database and monitoring system of ground pollution in Serbian cities on a whole;<br>• lack of funds to develop an effective and efficient system of monitoring<br>• Lack of legal frameworks and strategic documents in Serbia<br>• Still underdeveloped sustainable urban transport policy and plans in the biggest Serbian cities<br>• Lack of funds to develop sustainable urban road infrastructure<br>• Lack awareness for problem significance<br>• Common challenges in applying SUDS (i.e. lack of space, property rights, maintenance..) |

## 6. Final remarks and conclusion

While transport impacts on air and noise are widely explored, evaluated and economically estimated, the impacts on ground are much less explored. It is hard to find a logical explanation for this, because both kinds of impacts are complex and hardly measurable. Both may be from local to wider coverage and from short to long-term effects. The difference is that air pollution ultimately increases water and soil pollution, through the rainfall.

It is very hard to estimate transport impact on urban ground. Methodology is not so developed as at it is in case of air pollution. Even in cases that soil and groundwater pollution is monitored, there are problems related with methodology to define and evaluate particular transport impact on ground, estimated long-term effects and awareness of topic importance for sustainable urban traffic development.

Traditionally, surface water drainage systems have not been designed with sustainability in mind, particularly regard to water quality, catchment flooding, water resources, site amenity and potential for habitat enhancement. This is still the case with Novi Sad and, very probably, with other Serbian cities, too. The importance of this problem may arise with evidence of increasing precipitation and climate changes in our country.

Although all three problems exist even in the most developed countries, they are particularly expressed in Serbian cities. Overall old fleet, underdeveloped urban logistics planning, combined with lack of funds for research, network maintaining and contemporary measures, lack of awareness about the problem and underdeveloped a comprehensive sustainable policy contribute that urban traffic planners put the soil and groundwater contamination on the last place on the list of priorities. We hope that this paper will inspire both, researchers and practitioners to focus their attention more on this vital urban resource in the imminent period.

**Acknowledgements**

The study was performed within the projects COST TU1206 and TR36030, and partly financially supported by the Serbian Ministry of Education, Science and Technological Development (Project No. TR36030).

**References**


Geiger, W.F.; Marsalek, J.; Rawls, W.J. 1987. Studies and reports in hydrology, *Vol. 1: Planning and design of drainage systems*, UNESCO, Paris.

Helmreich, B.; Hilliges, R.; Schriewer, A.; Horn, H. 2010. Runoff pollutants of a highly trafficked urban road–Correlation analysis and seasonal influences. *Chemosphere*, 80(9), 991-997.

JP Urbanizam – Zavod za urbanizam Novi Sad. 2009. Saobraćajna studija grada Novog Sada sa dinamikom uređenja saobraćaja – NOSTRAM (in English: Traffic Study of Novi Sad with the Dynamics of Traffic Regulation – NOSTRAM), Novi Sad, Serbia

Lehner, P.; Aponte Clark, G.P.; Cameron, D.M.; Frank, A.G. 1999. *Stormwater Strategies*. Community Responses to Runoff Pollution. NRDC. Available from Internet: http://www.nrdc.org/water/pollution/storm/stoinx.asp

Maibach, M.; Schreyer, C.; Sutter, D.; Van Essen, H. P.; Boon, B. H.; Smokers, R.; Bak, M. 2008. *Handbook on estimation of external costs in the transport sector*. CE Delft Solutions for environment, economy and technology, <www.ce.nl>.

McKinnon, A.; Browne, M.; Whiteing, A.; Cullinane S. 2010. *Green logistics: Improving the environmental sustainability of logistics*. Kogan Page Publishers.

Mitchell, G.; McDonald, A.; Lockyer, J. 2001. Pollution Hazard from Urban Nonpoint Sources: A GIS Model to Support Strategic Environmental Planning in the UK, Final Report. Available from Internet: <http://www.geog.leeds.ac.uk/projects/nps/reports.htm>

Pittner, C.; Allerton, G. 2009. *SUDS for roads*, WSP, Edinburgh.

Republic Hydrometeorological Service of Serbia. 1950-2013. *Meteorological annual report – Climatology data*. Belgrade, Republic of Serbia

Rodrigue, J.-P.; Comtois, C.; Slack, B. 2013. *The Geography of Transport Systems*, 3rd Edition. Routledge, New York.

Stojanović Đ.; Basarić V.; Gajić V. 2011. The impact of freight transport on urban noise, In *Proceedings of 3. International conference - Towards a humane city*, Novi Sad: FTN, 27-28 Oktobar, 2011

Vasić, M. V.; Mihailović, A.; Kozmidis-Luburić, U.; Nemes, T.; Ninkov, J.; Zeremski-Škorić, T.; Antić, B. 2012. Metal contamination of short-term snow cover near urban crossroads: Correlation analysis of metal content and fine particles distribution. *Chemosphere*, 86(6), 585-592.

Vasin, J.; Ninkov, J.; Zeremski, T.; Sekulić, P. 2011. *Program praćenja kvaliteta poljoprivrednog i nepoljoprivrednog zemljišta na teritoriji grada Novog Sada tokom 2011. godine* (in English: Quality monitoring program of agricultural and non-agricultural land in the city of Novi Sad in year 2011), Projekat, Institut za ratarstvo i povrtarstvo laboratorija za zemljište i agroekologiju, Novi Sad. Available from Internet: <http://www.environovisad.org.rs/index.php?option=com_content&view=article&id=287&Itemid=102&lang=sr>

Veličković, M. S.; Stojanović, Đ. M.; & Basarić, V. B. 2014. The assessment of pollutants emissions within sustainable urban freight transport development: The case of Novi Sad. *Thermal Science*, 18(1), 307-321.

# HOW FUTURE TECHNOLOGY IN ROAD TRANSPORTATION COULD CREATE A MORE SUSTAINABLE BUILT ENVIRONMENT?

**Athanasios Galanis**[1], **Nikolaos Eliou**[2]
*[1,2] University of Thessaly, Department of Civil Engineering, Pedion Areos 38334, Volos, Greece*

**Abstract:** This paper presents the future technological trends in the area of road transportation both in urban and suburban areas. These new practices have been developed under the framework of "Intelligent Transport Systems" (ITS), targeting in the creation of a more sustainable built environment with numerous benefits in social, financial and ecological features. Uncontrolled traffic leads to traffic congestion both in both urban and suburban areas increasing fuel consumption, CO2 emissions, and traffic noise and travel time. The present transportation trends are the development of new "smart" products, divided in three interdependent paths: smart cars; smart roads; and smart transportation system. Rising of the new era of "Internet of Things" and the implementation of ITS technology in "smart cars" actually means a data flow between the road and the vehicle or among vehicles themselves. Development of such a technology leads to autonomous cars and driverless mobility. By the term "smart road" we describe roads able to communicate each other and with vehicles, transferring data and becoming adaptable in the future transport demand. Finally, the development of a "smart transportation system" could embrace all in process technology from the summary of transport modes mainly in urban areas creating an interface among transport built environment, transport modes and commuters. Implementation of new technology could create numerous benefits in all aspects of our daily transport needs improving the quality of life both in urban and suburban areas and thus preserving the environment for the future generations.

**Keywords**: road, transport intelligent, sustainable, environment.

## 1. Introduction

This paper presents the future technological trends in the area of road transportation both in urban and suburban areas. These new practices have been developed under the framework of "Intelligent Transport Systems" (ITS), targeting in the creation of a more sustainable built environment with numerous benefits in social, financial and ecological features. The art of road construction has been continuously improved since the era of the Roman Empire when was first decided to construct and maintain a road network as a vital part for the preserve of the State. Until today, highways remain more or less as dumb surfaces with virtually no data flowing between vehicles and road infrastructure itself. That is about to change…One of the most important issues of the present global civilization relates directly the transportation problem with extensive urbanization. Both in developed and developing countries people migrate to cities from rural areas to create a family and search for more secure job opportunities. Urbanization changes the economic, social and political setup of a country or region, resulting to numerous major negative impacts in fuel consumption; CO2 emissions; traffic noise; and travel time (Siemens, 2011).

The key word that leads to the solution of those problems is "smart". Implementation of "Intelligent Transport Systems" (ITS) and "Internet of Things" in the transportation sector is to develop new smart products (World Future Society). These are split into three interdependent issues which are the following:

- Smart cars – vehicles
- Smart roads – highways
- Smart transportation system

## 2. Smart Car

Presenting the "Smart Car" technology, we actually describe data flowing between road and vehicle and among vehicles themselves. The development of this idea leads to "driverless mobility" and "driverless cars" (Popular Mechanics). An autonomous car, also known as a driverless car, is an autonomous vehicle capable of fulfilling the human transport abilities of a traditional car. As an autonomous vehicle it is capable of sensing its surrounding area, navigating without human input. Robotic cars exist mainly as prototypes and demonstration systems but are likely to spread wider in the near future.

Autonomous vehicles sense their surroundings using technological assets as radar; GPS; and computer vision. Advanced control systems interpret sensory information identifying appropriate navigation paths, as well as obstacles and relevant signage. Some autonomous vehicles update their maps based on sensory input, allowing vehicles to keep track of their position even when traffic or built environment conditions change or when they enter uncharted areas.

Driverless cars will present numerous benefits in solving traffic congestion problems. Future driverless cars will be electric ones, removing one of the most important obstacles to integrate in our cities. They present zero air pollution and traffic noise creating a cleaner, safer and less congested transport system in urban areas. On the other hand, going

[2] Corresponding author: neliou@civ.uth.gr

driverless may cause legitimate issues like who is actually responsible for a potential traffic accident. Furthermore, could be a driverless artificial intelligence so advanced in order to achieve a damage control? The majority of today forward-collision warning systems can identify the presence of another vehicle; motorist; bicyclist; or pedestrian. Some of them can even apply the brakes in low-speed conditions but none of them can avoid an obstacle. Implementation of driverless technologies will also be blamed for destroying countless jobs, such as: truck drivers; taxi drivers; bus drivers; limo drivers; ambulance drivers; traffic police officers; parking lot attendants etc. If implemented correctly, driverless vehicles may even deal with a fatal blow to automobile insurance industry.

Vehicle design is only the one part of the equation to improve road safety and sustainable mobility. Without reimaging the way we design and maintain highways, driverless cars will only achieve a segment of their true potentials. Combining smart cars (driverless) with smart highways, we can begin to envision a far brighter future ahead. Finally, we will be driving towards a far safer and more resilient society, but in the meanwhile we still have a few bumpy roads to go down.

## 3. Smart Roads

The second part of the “smart transportation” concept is “smart roads”. Such roads can communicate with each other and vehicles, transfer data and become adaptable in the future demands. In the interurban transport network, the Dutch designers “Studio Roosegaarde” proposed five ways to make roads smarter. The “Smart Highway” project proposes energy-efficient concepts that will be tested on a stretch of highway in the Brabant province of the Netherlands. The first concept of the project presents a glow-in-the-dark road that uses photo-luminescent paint to mark out traffic lanes. The paint absorbs energy from sunlight during the day and illuminates the road at night for up to 10 hours. The second concept is the design of a temperature-responsive road paint showing images of snowflakes when the temperature drops below zero, warning drivers to take care on icy roads. There are also two ideas for roadside lighting as concept three and four. The third concept presents interactive street lamps that activate when vehicles are passing through their site through a dimming process. The fourth concept presents “wind lights” that use energy generated by pinwheels activated from nearby passing vehicles causing them to spin around. The fifth concept of the project is the development of an induction priority lane incorporating induction coils under the tarmac in order to recharge electric vehicles. In the urban transport network where most of traffic congestion problems occur, future road markings introduce the idea of the following assets:

- Sensors under the road surface to communicate with the vehicles
- Spotlights under the glass roads in order to transform them into adjustable “driverless mobility systems” (DMS).

Technology provides us the ability to control traffic. In the new world no more traffic signs or signals will exist…or will be optional. The traffic will be approached as a unique flow, just like the water in a tube. The principles of the traffic flow control system will be determined by software, calibrated in such way in order to adopt into transport conditions changes. The “solar surfaces” is one of the most innovative technology assets in the urban transport system. Dynamic traffic lines able to change per second depends the car and its surroundings creating the opportunity for safer roads and more efficient travels. Combining these two specific technologies we can achieve the unity of road environment and vehicles, considering them and treating as one. For example, digital “shadows” predict the future movement of a vehicle, providing the ability to anticipate driverless behaviors. By adding a thin layer of solar powered reprogrammable sensors within the surface of the road, we imagine a future where driverless technology is shared between vehicles and roads. Infusing the urban built environment with information, energy and light, will enable cities adapting into real time changes of urban life. Furthermore, the plasticity will replace the static city cast in concrete, with a future city that dynamically transforms and adapts to the life between the buildings!

Smart road infrastructure and “road intelligence” does not change quickly and easily. The role of infrastructure in supporting high quality traffic with state of the art equipment is necessary. In an integrated approach, vehicles and roads are in synergy offering the elements of automation, eco-driving and energy distribution. In EU, a FEHRL driven program “Forever Open Road” is under implementation (Steve, 2010). This program is promoted in France by IFSTTAR entitled R5G (Road of Fifth Generation), (Jacquot-Guimbal and Hautière, 2012).

## 4. Smart Transportation

In a futuristic transport model we can create a “smart transportation” environment ready to accept and embrace those advanced technological steps which are going to be completely functional in the following decades. With our imagination and inventiveness, we can create this environment using the nowadays technology. Many cities have developed smart mobility programs or are currently in process. In China, more than two hundred million people own a smartphone today – the majority of them in urban areas – and technology start-ups are pushing into development of new applications (apps) meeting the needs of urban dwellers. Because some of these apps are inspired by apps developed by Western countries, localization requires the creation of a Chinese interface, integrating local resources and adapting to local customs. As bloggers on 36Kr.com report (Chinese-language technology blog), some recently developed mobility-related apps for the Chinese market include:

- 摇摇招车/ Yaoyao Zhaoche ("Shake and Call a Taxi") allow users to call a taxi with the simple gesture of shaking the phone. This app links up with car rental companies, because private cars are not allowed to run commercial businesses in China. The cost is about the same as a regular taxi.
- 同行/ iTong Xing ("Ride Together") helps users to find suitable travel mates to share a taxi ride or a ride in a private car. The founder of Xing, Herock, has had long and boring rides in Beijing himself and "wished a comfortable-looking person could hitch a ride and chat with me on the way."
- 嘟嘟快捷租车/ Dudu Cars ("Dudu Express Car Rental") enables users to browse nearby rental cars, and reserve their favorite brands on the phone or the Web site. The smartphone app unlocks the car.
- 租车达人 ("Car Rental Master") works with car rental companies to provide real-time availability information, price comparisons, and online reservations. Users receive a monetary reward for renting a car through this app.

The architects at Superpool in Istanbul also took up the theme of a smartphone based mobility solution in the Audi Urban Future Award 2012. Although there is already an iPhone app for traffic in Istanbul, it only provides information about the current traffic situation, leaving drivers to choose the best route themselves. Superpool's idea is a smart system that helps drivers by calculating the best route on the basis of all information collected from traffic flow. As a smart satellite navigation app it could represent a highly efficient traffic guidance system. The proposals also envisage wide availability of car sharing to solve the traffic chaos and the difficult situation with parking spaces, as well as creating more space in the city for people. Anyone who has experienced the gridlock in the center of Istanbul knows how urgently this city needs intelligent traffic solutions.

In Istanbul the visions are slowly becoming reality. YOYO Car Share is a newly established Turkish company that aims to change traffic, especially in Istanbul. According to its co-founder Berkman Çavuşoğlu, the experience of other cities shows that "each shared vehicle can replace approximately 15 to 20 cars on the road". This relieves the pressure not only on individual traffic but also on public transport, leading to an overall reduction of CO2 emissions. They are optimistic: "When all these advantages are taken into account, private car ownership will be replaced by car sharing. And the change will be faster than what people imagine".

The prerequisites for intelligent car sharing are both the Internet and the smartphone. Only these enable the user to know exactly where an available vehicle is located. However, this has long ceased to be an obstacle. However, the big challenge lies in making it easy to access the vehicles. Getting hold of a vehicle has to be convenient, because everyone ideally wants to have one right on the doorstep. In the final analysis it is autonomous driving, when available vehicles really will roll up at the user's door, producing the breakthrough for car sharing. Nothing could be more convenient…

In pure technical terms the vision of car-sharing in combination with the smartphone and autonomous driving is already feasible: Via a smartphone app a private or shared car is ordered to the user's address. It comes right to the door by electric power from a garage or a distant car park. While this is happening we enter the destination into the app via voice command. The most efficient route is then calculated on the basis of all available traffic information. We get in the vehicle and are driven to the destination. On the way we can answer emails, make phone calls by Skype or stream videos. And if we want to, we can take other passengers with us on the route. At the destination we leave the vehicle, which autonomously looks for a place to park where there is an automatic recharging function on the basis of induction. There it waits for us until we need it again, if we are sharing the vehicle in a pool of users. All city dwellers share vehicles as needed in a community: a convertible for the weekend, an estate car for making a trip to the Ikea furniture store, and a compact van with other passengers on the way to work. The instrument of communication that networks all of this together and can personalize it for us could be the smartphone. This does not seem so very far away…

One small problem might nevertheless remain: the question of status. In some social classes the trend towards sharing extends to houses, apartments and cars. However, for most groups in society the car in particular remains a status symbol and an expression of individuality. A change in thinking in the cultural sphere would have to take place in order to make car sharing accepted and used by a broad mass of people.

## 5. Conclusion

There are several periods identified as the ones structuring the development of smart cars and streets. In the first period (<1980), the autonomous vehicles stated their existence and conquered their independence through the introduction of the first static equipment. In the second period (1980-2000), researchers focused on traffic flow modeling and control strategies giving priority to automobiles in urban streets. Introduction of new sensors allowed real time measurements and dynamic traffic management. In the third period (2000-2010), research focused to on-board autonomous assistance systems. Those systems are now at the commercialization stage. In the present fourth period (2010-2020), research focuses to connectivity through the development of car-to-car and car-to-infrastructure cooperative communication systems. Connectivity leads to new safety functions and the ability to use automation that conceptually appears as extreme autonomous systems. It is certain that it will take time before full automation at high speed can be achieved due to development of highly efficient guiding and collision avoiding systems.

The raise of a new era of “Internet of Things” and implementation of Intelligent Transport Systems (ITS) technology in “smart cars” provides a data flow between the road infrastructure and the vehicles or among vehicles themselves. Development of such a technology leads to autonomous cars and driverless mobility. “Smart roads” are able to communicate both each other and with vehicles, transferring data and becoming thus adaptable in future transport changes. “Smart transportation systems” embrace all in process technological assets from the summary of transport modes mainly in urban areas, creating thus an interface among transport built environment, transport modes and commuters. Implementation of new technology tools and apps could create numerous benefits in all aspects of our daily transport needs improving the quality of life both in urban and suburban areas and thus preserving the environment for the future generations.

**Acknowledgements**

Authors are grateful to the University of Thessaly undergraduate student “Konstantinos Zavitsanos” for his contribution in the literature review of this research.

**References**

Case studies for traffic solutions. Modern concepts and technologies help improve efficiency. Report for Siemens, 2011. Available from Internet: <http://www.mobility.siemens.com/mobility/global/SiteCollectionDocuments/en/road-solutions/interurban/case-studies-for-traffic-solutions-en.pdf>

Chinese-language technology blog, Available from Internet: <http://www.36kr.com/>

How designers plan to create the route 66 of the future, Popular Science. Available from Internet: <http://www.popsci.com/cars/article/2012-10/how-dutch-designers-want-build-super-smart-highway>

In the future, highway will light up and even charge your electric car (maybe). Available from Internet: <http://jalopnik.com/5973219/in-the-future-highways-will-light-up-and-even-charge-your-electric-car-maybe>

Jacquot-Guimbal H., Hautière N., 2012. Recherche et perspective: Les routes de cinquième génération. Administration, 234: 84-89.

Steve P., 2010. The Forever Open Road Programme. In Proceeding of the 1rst International Symposium Highway to Innovation, Tel Aviv, Israel, 2010.

Studio Roosegaarde, Dutch engineers test glow-in-the-dark “Smart Highway”. Available from Internet: <http://mashable.com/2014/04/22/glow-highway-netherlands/>

Studio Roosegaarde, Smart Highway. Available from Internet: <http://www.studioroosegaarde.net/project/smart-highway/>

Superpool – Istanbul, Turkey, Audi Urban Future Award 2012. Available from Internet: <http://www.audi.com/content/com/brand/en/vorsprung_durch_technik/content/2013/08/aufa_superpool_turkey.html>

The internet of cars is a near thing today, World Future Society. Available from Internet: <http://www.wfs.org/blogs/len-rosen/internet-cars-near-thing-today>

The 12 most important questions about self-driving cars, Popular Mechanics. Available from Internet: <http://www.popularmechanics.com/cars/news/industry/the-12-most-important-questions-about-self-driving-cars-16016418?click=main_sr>

Will self-driving cars mean the end of highway fatalities? Popular Science. Available from Internet: <http://www.popsci.com/article/cars/will-self-driving-cars-mean-end-highway-fatalities>

YOYO Car Share, Audi Urban Future Initiative. Available from Internet: <http://audi-urban-future-initiative.com/blog/tagged/berkman-cavusoglu>

# SPEED LIMIT EVALUATION BASED ON POLLUTANT ESTIMATION IN URBAN HIGHWAY: CASE STUDY IN CRACOW

**Daniel Kubek[1], Paweł Więcek[2], Dariusz Grzesica[3]**
[1,2,3] *Institute of Building and Transport Management, Transport Section, Faculty of Civil Engineering, Cracow University of Technology, Cracow*

**Abstract:** The negative impacts associated with the increase of congestion in urban areas are nowadays more and more serious problem for local communities. Rapidly rising level of road traffic affect negatively transportation system performance and causes many environmental nuisance, especially air pollution from car exhaust. The latest reports reveal that the air quality in Cracow metropolitan area is very poor. More than 60% of the days in the year are days with the smog and their number increase every year. One of main source of that pollutant is caused by transportation system. Despite of these facts, the city authorities have not implemented any significant traffic management strategy relating to ecology or decreasing emission of pollutants from transportation system. The article presents an approach to obtain the upper speed limit on Cracow urban highway using traffic measurement, traffic simulation software and pollutant emission models. The presented method can be successfully applied by the municipal authorities because of its relatively low implementation cost and satisfactory effectiveness. The upper speed limit was determined taking into account instantaneous traffic emission of CO2, NOx, VOC and PM and other traffic parameters like flow, density or travel time. The simulation results were obtained considering two scenarios: the first one is no traffic management strategy was provided and the second one is scenario which includes proposed speed limit. The quality of proposed approach was evaluated by the comparison of results reached from different scenarios. The considered in this article topic is relevant due to increased environmental awareness of society and also in the context of sustainable development.

**Keywords**: speed limit evaluation, urban pollutants estimations, traffic simulation software

## 1. Introduction

Cracow is located in the Southern part of Poland. It is second-largest and second-most populated city in the country. Currently, it constitutes a central metropolitan area of Cracow conurbation. The city has been functioning as an administrational, cultural, educational, scientific, economical, service, and tourist centre. Considering its role in the region and in the country, the metropolitan area of Cracow faces serious problems related to congestion. Except the nuisance from road traffic, increase in air pollution is one of the major consequences of existing circumstances. Besides emission from individual economic sectors and industrial branches, rapidly increasing traffic constitutes a serious source of toxic substances discharged into the atmosphere. These compounds constitute major health hazard especially to the residents who inhabit the nearest proximity of busy traffic arteries and ring roads. Nitrogen oxides, benzene oxides, carbon oxides, volatile organic compounds, as well as particulate matters are the most hazardous compounds that have been emitted by road means of transport. Pollution mixtures have a tendency to alteration in time and space as a function of multiple features such as close proximity to road, type and age of vehicles, pattern of traffic and presence of other sources of pollution. Short- or long term exposure to these types of compounds contained in the air may, immediately or in the future, lead to adverse health problems related to respiratory tract, nervous system, cardiovascular system disorders and diseases including cancerous changes (Künzli N.,et al., 2000)(Krzyzanowski M., 2005). Fig. 1 shows average monthly emissions measured throughout a year from a measuring station located approximately 1 km from above mentioned highway. This emission was measured in 2013 against nitrogen compounds, NOx, and suspended dusts, $PM_{10}$. In both cases, average annual emission exceeds allowable levels. For NOx, this figure is twice as big as it is allowed by the standard.

[1] Corresponding author: dkubek@pk.edu.pl

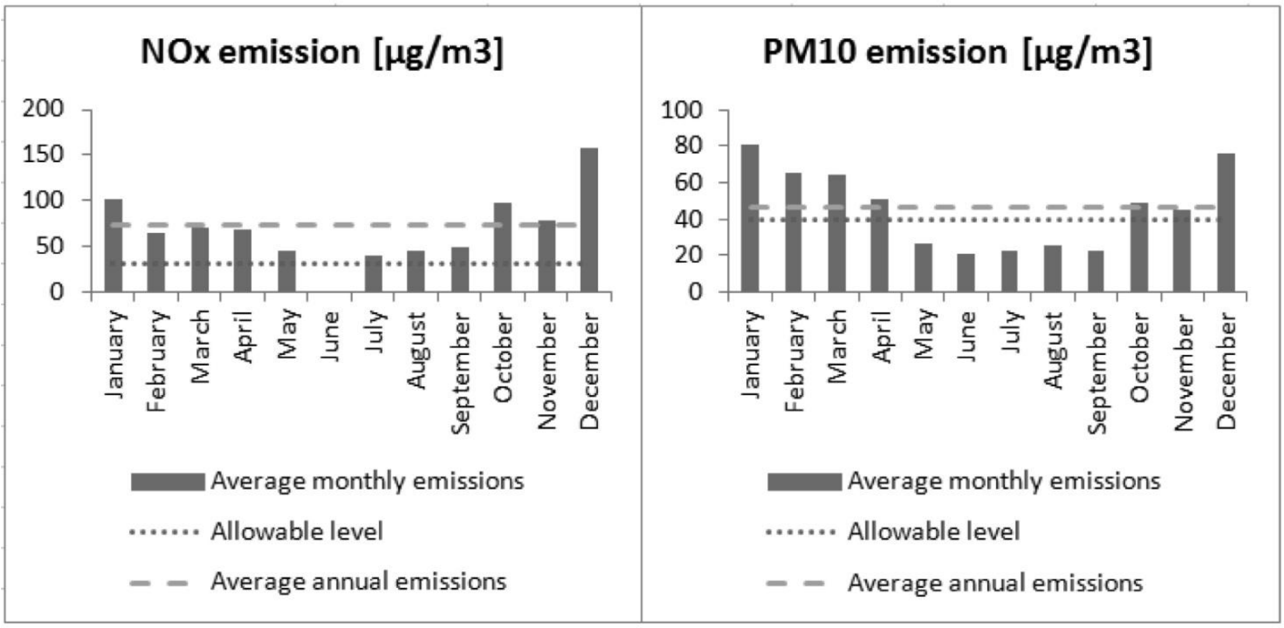



**Fig. 1**
Average monthly emission of NOx and PM10 from measuring station in Bujaka Street in Cracow.
*Source: Own elaboration based on (MSMP 2014).*

### 1.1 Existing solutions

Limitation of traffic in the centre of Cracow is a crucial issue. Current eco-minded solutions related to the traffic have been just planned or implemented at initial stages. Limited traffic zones and metered parking zones as well as access to "Park & Ride" car parks at tram and bus terminuses are the major points in the policy for traffic limitation in the city centre. It should be obtained through implementation of an appropriate traffic arrangement and flow plan in the city. Except that, transport policy includes forming of traffic management intelligent systems which would ensure vehicle traffic flow through implementation of, among others, green wave, time indicator, orientation of outbound traffic from the city centre, and limitation of inbound traffic to the city centre. Additionally, consecutive actions, as part of the transport policy, have been implemented in order to limit pollution of the air at the province level. At the same time, limitation of traffic flow should be obtained through development of integrated public transport, promotion of sub-regional transport interchanges, and increase in transport availability to the areas with the lowest accessibility in the region and modernisation of roads (OMMV 2011).

### 1.2 Conceptual framework for adaptive speed limit in highway

According to the authors of this article, all undertakings carried out by the city administration should be supplemented by pollution monitoring system to enable influence on vehicles' speed in order to minimise the level of pollutions. This system should be applied, particularly, throughout densely built areas in the vicinity of busy traffic routes in the city. Along these routes, pollution concentration may be significantly changed in time as a function of traffic information data.
Researches, which were carried out by foreign scientists, and some relevant works indicate a close relationship between a vehicle's speed and rate of pollution emission (Dijkema M.B.A., et al. 2008) (Coelho M.C., et al., 2005) (Panis L.I., et al., 2006). Therefore strategies and solutions related to traffic management, which purpose is to minimise adverse environmental impact, should be continuously supplemented with additional elements and become the subject for further deliberations and simulations.
The approach, suggested by the authors, which aim is to limit the emission of compounds contained in exhaust fumes through determination of relevant speed of vehicles' flow through arterial roads may constitute a component of eco-minded traffic following Fig. 2.

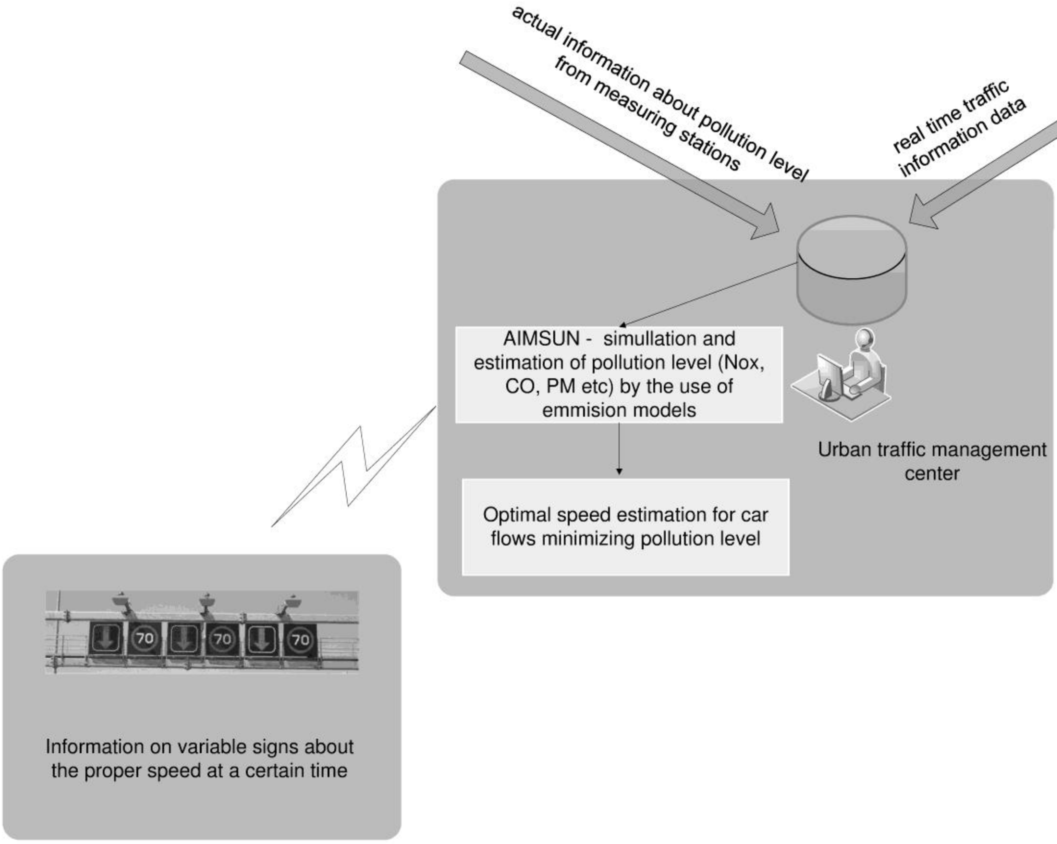



**Fig. 2**
*The conceptual framework of the proposed solution*
*Source: Own elaboration.*

To estimate emission of fumes, AIMSUN (abbreviation for Advanced Interactive Microscopic Simulator for Urban and Non-Urban Networks) software was applied. It is to represent transport systems and micro-simulations of road traffic. Applied simulator contains integrated environmental model that enables estimation of temporary emission of the following four types of fume: Carbon Dioxide ($CO_2$), Nitrogen Oxides (NOx), Volatile Organic Compounds (VOC) and Particulate Matter (PM). Used model for temporary emission has been based onto researches presented in (Panis L.I., et al., 2006) work. It works on the assumption that a vehicle may be operated in one of four following conditions: idling, cruising at constant speed, accelerating, and decelerating. Emission value is calculated at each step of the simulation on the basis of the same equations but considering various coefficients with respect to type of vehicle, type of fuel, and temporary acceleration/braking of each simulated vehicle.

## 2. Case study: Cracow's highway

### 2.1 Research assumptions

Implementation of new strategies for urban road traffic in agglomerations may be significantly difficult as well as costly. Nevertheless, a trial to estimate and foresee all potential effects of a relevant strategy is absolutely necessary for all persons and organisations that have been liable for implementation of new transport policies. Application of road traffic simulator, i.e. AIMSUN, may be helpful in the process of forecasting. Even low investments may enable checking plausibility of effects that have been predicted. A suggestion related to alteration of allowable speed on urban highway has been tested by the authors for Cracow with application of AIMSUN 8.03 software.
Fig. 3 shows a section of A4 highway which goes through Cracow corporate limits. Considering proximity of the highway to buildings, the highway functions as a ring road as well, therefore increase in traffic is significant during peaks. In this case, two peaks have been described: morning and afternoon traffics. For this reason, two periods were taken into consideration. The first 06:30 to 09:30 and the second 15:00 to17:00.

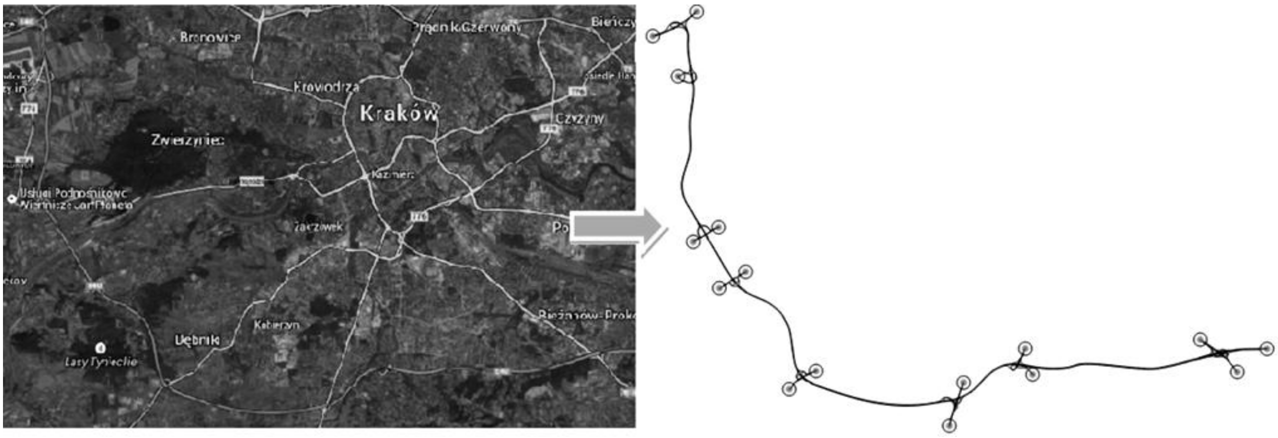



**Fig. 3**
*Aimsun model of Cracow's urban highway*
*Source: Google Maps*

Traffic information data related to the number of vehicles in particular vehicle group were obtained from testings on traffic that were carried out by Municipal Infrastructure and Transport Facility in Cracow in 2012. Three following types of vehicles were applied in a road traffic simulation on the highway: motor cars, trucks, and buses. All outstanding types of vehicles constitute an insubstantial proportion of the total number of vehicles. That is why these have been omitted in the subsequent analysis.

A model of fumes emission, integrated with AIMSUN software, requires entering percentage share of fuel used in vehicles. As per relevant data issued by The Polish Central Statistical Office (CSO 2013), in Poland in 2012 this share included in case of motor cars: petrol engines – 57.7%, Diesel engines – 25.9%, and LPG engines – 14.7%. Trycks and buses were Diesel unit-driven in 100%.

The other crucial parameter of the simulation is driver's reaction time and reaction time at stop. These values have an impact on integrated model of vehicle's flow. In Poland, testing with application of reaction time is rarely carried out. Thus the authors used results of similar testing introduced in (Guzek M. et al. 2010) work, where time to collision was tested. On the basis of those testings, reaction time equal 1 second and reaction time at stop equal 1.2 second were taken.

### 2.2. Tested traffic parameters

As mentioned previously, a model of fume emission enables estimation of emission rate for four types of fumes: $CO_2$, NOx, VOC, and PM. The analysis that is stated below includes checking emission rates when allowable speed on the highway is altered from the current 140 [km/h] to 70 [km/h] including 10 [km/h] interval. Considering the fact that speed has a significant impact on traffic flow, an impact of total speed alteration on the tested network has also been checked. This impact was outlined through determination of the rate of the following 5 parameters alteration: density, flow, stop time, total travel time, and travel time. Density is an average number of vehicles per kilometre for the whole network; flow is defined as average number of vehicles per hour that have passed through the network during the simulation period; stop time represents average time at standstill per vehicle per kilometre; total travel time total is time of travel for all simulated vehicles that have crossed the network; finally travel time is average time a vehicle needs to travel one kilometre inside the network.

### 2.3 Simulation results

For each peak and for each considered speed, computer simulations were carried out (16 simulations in total), where fume emission rates were determined as well as values for traffic flow parameters. Fig. 4. shows charts for emission of fumes under consideration.

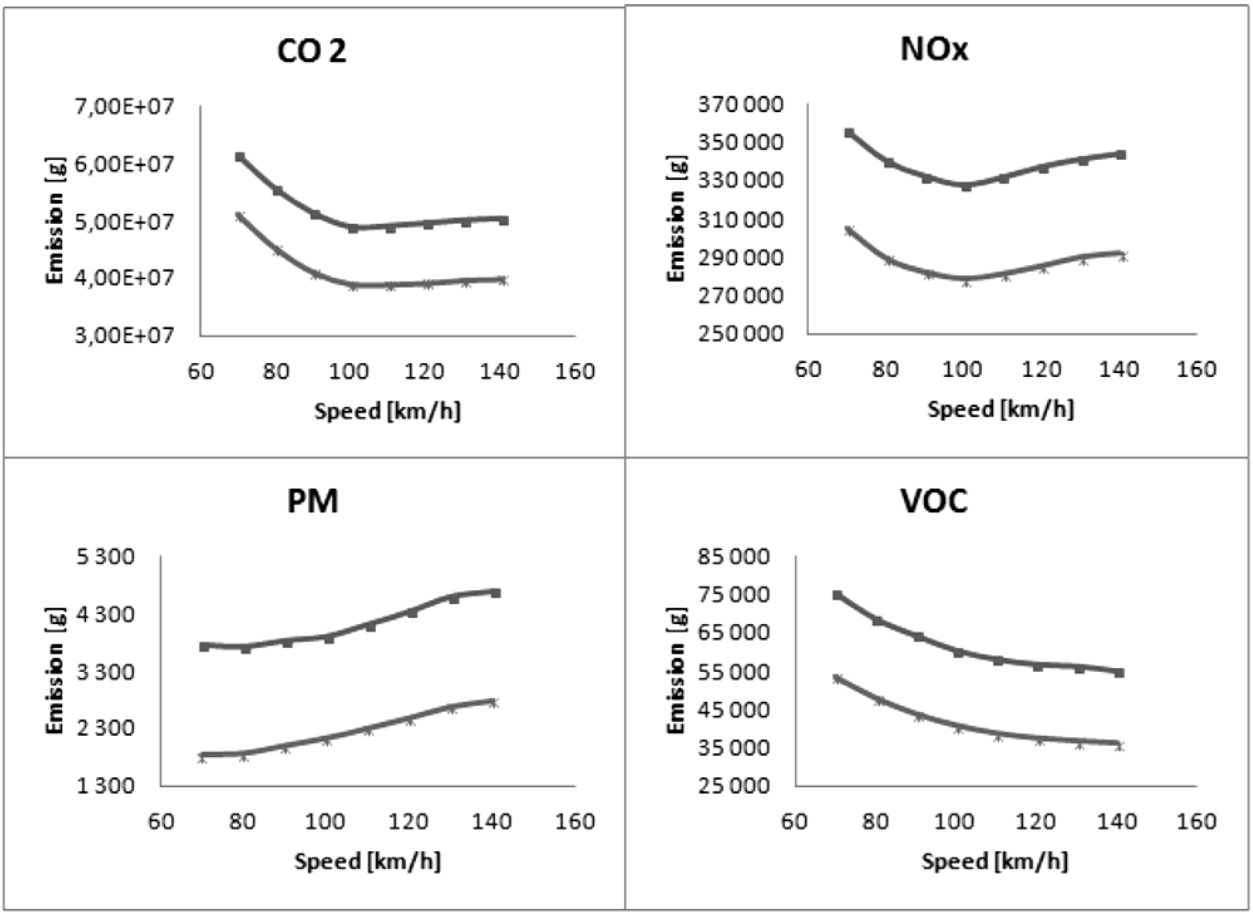



**Fig. 4.**
*Traffic emission of pollutant for Cracow's highway (red line represents emission of afternoon traffic; blue line - morning traffic).*
*Source: Own elaboration based on simulation results.*

The results obtained from the simulations show that reduction of allowable speed limit enables reduction in emission of three of four exhaust fumes. For $CO_2$ and NOx, an optimal speed limit may be indicated within 100 - 110 km/h. For those values the emission is the lowest. As it may be noticed, current speed limit (140 km/h) causes fume emission in a larger extent. PM release rate increases as the speed rises, while in case of VOC the opposite relationship is true. The purpose of testing was to indicate an optimal speed considering alteration of mentioned parameters as well. On the basis of obtained results it may be confirmed that the current value of allowable speed ensures the best values for all considered parameters. Therefore, in the consecutive analyse, these parameters are aggregated to single criteria.
Selection of the most beneficial optimal speed, which would be a compromise on all considered criteria (parameters), is complicated. Expertise and preferential treatment from local authorities are necessary. The question whether eco-minded aspects or enhanced traffic parameters, e.g. the shortest travel time, are more important must be answered. Referring to the fact that this type of analysis must be dedicated to an individual case, the authors applied a simple weighing method, where each identical weight has been attributed to each criterion. The method formulated in this manner comes down to calculating an average value. The most advantageous optimal speed values for particular fume emissions and traffic parameters related to both peaks are shown in Table 1.

**Table 1**
*Optimal allowable speed value depending on particular parameters.*

| Parameter | Optimal speed [km/h] | |
|---|---|---|
| | **Morning peak** | **Afternoon peak** |
| **$CO_2$** | 110 | 100 |
| **NOx** | 100 | 100 |
| **PM** | 70 | 80 |
| **VOC** | 140 | 140 |
| **Traffic parameters** | 140 | 140 |
| **Mean value** | **112** | **112** |

*Source: Own elaboration based on simulation results.*

The results of the analysis show that optimal speed which considers mentioned parameters should equal 112 km/h. It means that, as per applicable traffic regulations, the speed limit on A4 highway in the vicinity of Cracow should equal 110 km/h. Table 2 and Table 3 show potential changes in considered parameters and fume emissions when maximum allowable speed limit equals 110 km/h.

**Table 2**
*Increase/Decrease of traffic parameters values for optimal allowable highway speed equal to 110 [km/h].*

| | Speed [km/h] | Difference [%] | | | | |
|---|---|---|---|---|---|---|
| | | Density [veh/km] | Flow [veh/km] | Stop Time [sec/km] | Total Travel Time [h] | Travel Time [sec/km] |
| **Morning peak** | **140,00** | 4,19 | 13 615,33 | 0,26 | 2 138,01 | 40,82 |
| | **130,00** | 4,24 | 13 618,00 | 0,28 | 2 166,07 | 41,13 |
| | **120,00** | 4,31 | 13 617,00 | 0,26 | 2 198,73 | 41,55 |
| | **110,00** | 4,42 | 13 620,00 | 0,27 | 2 259,35 | 42,21 |
| | **100,00** | 4,61 | 13 616,67 | 0,24 | 2 356,04 | 43,17 |
| | **90,00** | 4,90 | 13 607,00 | 0,24 | 2 496,90 | 44,56 |
| | **80,00** | 5,28 | 13 607,00 | 0,22 | 2 691,51 | 46,43 |
| | **70,00** | 5,81 | 13 590,67 | 0,21 | 2 953,24 | 49,08 |
| **Afternoon peak** | **140,00** | 6,53 | 13 852,00 | 7,48 | 3 156,93 | 49,65 |
| | **130,00** | 6,66 | 13 846,67 | 7,60 | 3 214,65 | 50,23 |
| | **120,00** | 6,70 | 13 854,33 | 7,65 | 3 243,55 | 50,60 |
| | **110,00** | 6,83 | 13 857,67 | 7,52 | 3 307,84 | 51,19 |
| | **100,00** | 7,06 | 13 845,00 | 7,57 | 3 409,25 | 52,17 |
| | **90,00** | 7,46 | 13 823,67 | 7,70 | 3 596,47 | 53,98 |
| | **80,00** | 7,86 | 13 825,00 | 7,53 | 3 807,71 | 55,69 |
| | **70,00** | 8,55 | 13 794,33 | 7,69 | 4 118,60 | 58,74 |

*Source: Own elaboration based on simulation results.*

**Table 3**
*Reduce of emission for particular pollutant with respect to decreased allowable speed (allowable speed equal 110 km/h).*

| | Speed [km/h] | Reduce [%] | | | |
|---|---|---|---|---|---|
| | | CO2 | NOx | PM | VOC |
| Morning peak | 140,00 | 2,39% | 3,77% | 17,31% | -7,12% |
| | 130,00 | 1,93% | 3,08% | 14,15% | -5,47% |
| | 120,00 | 0,63% | 1,47% | 7,34% | -3,39% |
| | 110,00 | 0,00% | 0,00% | 0,00% | 0,00% |
| | 100,00 | 0,20% | -0,87% | -8,10% | 5,10% |
| | 90,00 | 5,24% | 0,30% | -15,13% | 11,69% |
| | 80,00 | 13,79% | 2,84% | -23,66% | 19,40% |
| | 70,00 | 23,61% | 7,79% | -25,06% | 27,71% |
| Afternoon peak | 140,00 | 2,61% | 3,60% | 12,40% | -5,73% |
| | 130,00 | 2,02% | 2,80% | 10,64% | -3,31% |
| | 120,00 | 0,99% | 1,68% | 5,20% | -2,47% |
| | 110,00 | 0,00% | 0,00% | 0,00% | 0,00% |
| | 100,00 | -0,35% | -1,24% | -5,45% | 3,83% |
| | 90,00 | 4,41% | 0,17% | -7,43% | 9,74% |
| | 80,00 | 11,69% | 2,56% | -10,48% | 15,52% |
| | 70,00 | 20,19% | 6,83% | -9,58% | 23,08% |

*Source: Own elaboration based on simulation results.*

The most valuable advantages would be obtained in case of particulate matter emission. For morning peak it would be equal to 17.31%, and for afternoon peak it would be equal to 12.40%. Reduction in $CO_2$ emission would be 2.39% and 2.61%, appropriately, while in case of NOx emission, the reduction would equal 3.77% for morning peak and 3.60% for afternoon peak. As mentioned before, only VOC emission would rise by 7.12% and 5.73%. When speed limit for analysed network is decreased travel time increases but in case of average distribution of this value to statistical traveller (travel time) this growth would equal just under 2 seconds per one kilometre of the road network. Similarly, when considering average density of vehicles, increase in the number of vehicles per one kilometre of the road would equal 0.3. The most relevant issue is that traffic flow would remain virtually unchanged with 0.03% drop.

## 3. Summary

The analysis that has been described in this article shows that the current regulations related to allowable speed on highways in urban areas should be altered; however implementation of constant, unaltered speed may be an inappropriate strategy for traffic management as well. Determination of speed limit should be introduced in an adaptive manner and its value should be determined on the basis of current traffic. The adaptive concept for speed alteration on highways, which has been presented, may be an effective solution for improvement of air quality. It may be also easily integrated with currently implemented local system for traffic control in Cracow as part of transport intelligent systems.

**References**

Central Statistical Office. *Transport.Activity results in 2012*. Zakład Wydawnictw Statystycznych, Warszawa 2013.

Coelho M.C., Farias T.L., Rouphail N.M., 2005. A methodology for modeling and measuring traffic and emission performance of speed control traffic signals. *Atmospheric Environment* 39: 2367–2376.

Dijkema M.B.A., C. van der Zee S., Brunekreef B., T. van Strien R. 2008. Air quality effects of an urban highway speed limit reduction. *Atmospheric Environment* 42: 9098–9105.

Guzek M., Jurecki R., Lozia Z., Stańczyk T.L., Pieniążek W. 2010. Assessment of Drivers' Reaction Times. Tests on the Track and in the Driving Simulator. *Logistics and Transport* 2(11): 64-70.

Krzyzanowski M., Kuna-Dibbert, Schneider J., *Health Effects of Transport related Air Pollution,* WHO 2005.

Künzli N., Kaiser R., Medina S., Studnicka M., Chanel O., Filliger P., Herry M., Horak F. Jr, Puybonnieux-Texier V., Quénel P., Schneider J., Seethaler R., Vergnaud JC., Sommer H., 2000. *Public-health impact of outdoor and traffic-related air pollution: a European assessment*. Lancet 356: 795–801.

MSMP - Air monitoring network of Małopolska Province 2014, (Małopolska sieć monitoringu powietrza 2014) < http://213.17.128.227/iseo/>

Panis L. I., Broekx S., Liu R. 2006. Modeling instantaneous traffic emission and the influence of traffic speed limits, *Science of the Total Environment* 371: 270–285.

OMMV - Office of The Marshal of The Małopolska Voivodeship 2011, *The development strategy of the Małopolska province 2011-2020* (Urząd Marszałkowski Województwa Małopolskiego 2011, *Strategia rozwoju województwa małopolskiego 2011-2020).* Załącznik Nr 1 do Uchwały Nr XII/183/11 Sejmiku Województwa Małopolskiego z dnia 26 września 2011 roku.

# IMPROVING THE NETWORK OF PEDESTRIAN SPACES WITH THE AIM OF CONNECTING THE SAVA'S AND DANUBE'S AMPHITHEATRE

**Aleksnadra Đukić[1], Milena Vukmirović[2]**
[1] *Associate Professor, University of Belgrade – Faculty of Architecture, Blvd. Kralja Aleksandra 73/II, Belgrade, Serbia*
[2] *Associate Researcher, University of Belgrade – Faculty of Architecture, Blvd. Kralja Aleksandra 73/II, Belgrade, Serbia*

**Abstract:** Considering the current trends of urban development of Belgrade with the aim of activating the areas along the banks of the Sava and Danube rivers, there is a need to consider the ways that could achieve connection among them. The establishment of a network of pedestrian directions that connect Sava's and Danube's Amphitheatre and which could be passed through the old city centre is allocated as one of the possible ways of thinking. In accordance with these objectives, Kosančićev venac was analysed as specific polygon located between Sava River and pedestrian zone of Knez Mihajlova Street. The methods that were used in analysis include method of direct surveying of inhabitants based on the Kevin Lynch's determination of the image of the city and a method of Space Syntax, with the aim to measure the intensity of pedestrian movement in the monitored territory. The study was conducted with students of the third year of Bachelor studies on the elected course "Network of pedestrian flows in function of urban redesign" at University of Belgrade – Faculty of Architecture. Summarizing and interpreting the collected data obtained three groups of results. First group of results has enabled the determination of the image of Kosančićev venac, composed of five elements: paths, edges, districts, nodes and landmarks. The second group of results presents an overview of the intensity of pedestrian movement on the observed territory, compared with the content located along the path and with the distribution of elements of the image of the city. The third group covered the results of the individual design proposal on selected pedestrian network segment defined on the basis of established relationships, with the aim to make that place more attractive and visited by pedestrians. This research has indicated the potential of the analysed area for the formation of transverse and longitudinal pedestrian flows. On the one hand they could enable active use of Kosančićevog venac as one of the most important urban ambient of the city, as well as to represent a segment of longer paths that could integrate the connection of the Sava's and Danube's Amphitheatre.

**Keywords**: Pedestrian spaces, network, Kosančićev venac, Belgrade

## Introduction

Danube and Sava rivers are the key characteristics and resources of Belgrade, the capital of Serbia, and they are of historical importance for its economy. Today, with the exception of recreational role of their waterfront, these rivers are largely excluded from active participation as an integral part in city life. Railroads, rail facilities, roads and dense industrial zones represent the barrier between the city and the waterfront and are considered as crucial reasons for this situation. The picture that Belgrade waterfront offers today is made of unacceptable conditions and inadequate use of water and waterfront, and devastated and neglected foreground that opens by looking at the river. In order to define strategy and guidelines for overcoming this problem, Urban Planning Institute of Belgrade is in the process of drafting Belgrade waterfront Study. The first phase was finished in 2009 and resulted in defining the wider zone boundary (zone of influences and broader interests), narrow zone boundary and emphasizing potentials of the waterfront (Urbanistički zavod Beograda, 2009). The overall aim of the strategy is to activate the significant potential of the waterfront and, as an integral part of the city, to build a high quality public spaces that will enhance the appearance and the importance not only of the location but the entire Belgrade as well.

**Fig 1**
*Belgrade waterfront wider zone boundary.*
*Source: Urbanisticki zavod Beograda, 2009*

[1] Corresponding author: adjukic@arfodita.rcub.bg.ac.rs

Belgrade waterfront wider zone boundary was defined in comprehensive process of overlapping all aspects that include natural and artificial features and conditions. This territory covers “the entire aquatorium, as well as part of the territory whose width varies from several tens of meters to several kilometres in the hinterland and has an area of 16950ha” (Urbanistički zavod Beograda, 2009). On the other side the waterfront narrow zone boundary encompasses contents placed in the immediate waterfront environment and thus becomes subject to the narrower interests in terms of resources and potential interventions. The size of this area is 5587ha.

The study concluded that “the most valuable, central part of the waterfront, is the most neglected and that the most attractive locations are almost completely degraded” (Urbanistički zavod Beograda, 2009). This is in the contrast to the potential of the location that include: the value of the waterfront areas in the strictest meaning of the term, inadequate disuse of potentially useful land, vast, diverse and valuable natural areas and large and important cultural heritage of the city, as part of the culture and history of the city in general.

By placing an emphasis on pedestrian environment, as a subject of this research, a light is put on the significance of small scale, often neglected in contemporary projects and development strategies. Scenes that correspond to this angle of perspective put focus on specific advantages of cities proportional to dimensions, senses and walking speed and form a basis for more complex and diverse relations (Vukmirovic, 2014). More concretely, they correspond to improving the quality of the smallest places in such way that people are simply attracted to go there and spend time there. Seen from this perspective, pedestrian movement is of twofold character (Vukmirovic, 2014; Gehl, 2010), since it is perceived as both mode of transport and an opportunity for many other activities. In accordance with this, the basic characteristics of walking and elements that influence it include the speed of pedestrian environment, distances, and the quality of route.

Pedestrian movement has become a very current topic after the 1960-ies. Numerous authors were engaged in these issues with the goal of developing models for designing street furniture (Schubert, 1967; Whyte, 1988); developing models based on relations between the level of services and pedestrian movement (Fruin, 1971; Polus and Schofer, 1983), and forming a rulebook for urban planning (Crawford, 2002) and pedestrian behaviour (Brilon and Grosmann, 1993; Kirsch, 1964). Methods used in this research mainly pertain to direct monitoring, as well as the use of photographs and films of concrete locations filmed during an extended period of time. Numerous investigations have shown that it is possible to foresee the dynamics of pedestrian movement, that it is possible to influence pedestrian flows (via urban design), as well as that pedestrians can move relatively freely and independently only along short distances, while along longer paths movement is conditioned by other pedestrians (Djukic and Vukmirovic, 2012a).

Seen as a part of Belgrade waterfront narrow zone and significant cultural and historical entity, Kosancicev venac was the subject of this research, located between Sava River and pedestrian zone of Knez Mihajlova Street. Considering the current trends of urban development of Belgrade with the aim of activating the areas along the banks of the Sava (“Belgrade on water”, “Beton hala” international competition, etc.) and Danube rivers (“Beko” and “City on water”), there is a need to consider the ways that could achieve connection among them. The establishment of a network of pedestrian routes that connect Sava’s and Danube’s Amphitheatre and which could be passed through the old city centre is allocated as one of the possible ways of thinking. Research presented in this paper is based on a theoretical framework in the domain of urban design, pedestrian movement and quality of the pedestrian environment.

**Methodology**

Having in mind the general aim of the research, the focus is placed on spatial level of the problem, i.e. onto characteristics and elements of pedestrian environment seen from the physical aspect. The research was used the methods that are developed and tested for several years (Djukic and Vukmirovic, 2012b) on University of Belgrade – Faculty of Architecture on elective course “Network of pedestrian flows in function of urban redesign” on bachelor studies and on course “Urban design methods”[2] on Master studies – Integral Urbanism. The study was conducted with students during spring semester of the 2013/2014 school year.

Implemented research methodology encompasses following elements: 1) **SWOT analysis of the location**, presented in a matrix form, is used for dissection of properties and potential of the area. The aim of the analysis was to identify specific of the location viewed from pedestrian perspective, and interpreted as particular factors for its strengths (S), its weakness (W), opportunities (O) for its development and the potential threats (T) it faces. 2) **Analysis the potential of integration of the area in existing network of pedestrian spaces** is intended to determine how the specific space can be connected to other important locations in the city. The focus of this phase of the research is placed on identifying the directions/paths/connections that would extend the existing pedestrian network, while at the same time creates a direct link with the analysed area. **3) Analysis of the legibility of the area** was conducted in the form of surveys of citizens who live or visiting the analysed area. The questionnaire was created based on elements of the image city defined by Kevin Lynch and public spaces and buildings that are specific to the particular location. Summarizing the results of the survey produces an image of the observed location. **4) Analysis of content units in ground floors** in order to determine the relationship between the type, number and rhythm of contents in ground floors of the building alongside the pedestrian routes and the intensity of pedestrian movement. The function of buildings along the street front conditions the character of the viewed path or space, the manner of use and the density of users. For a space to have vitality, it is indispensable to secure a density of user frequency and a density of diversity. **5) Analysis of the intensity of use of**

[2] Both courses are directed by Associate Professor Dr Aleksandra Djukic. Dr Milena Vukmirovic, associate researcher, also participate in these courses as an expert in the field of pedestrian space quality.

**public spaces** was aimed at investigate relations between user movement and the content in ground floors of buildings along the paths of movement. Networks of pedestrian flows are presented in the form of an axis diagram that shows the intensities of pedestrian movement measured at precisely defined locations at specific time intervals (on working days from 8-9h, 12-13h, 16-17h and 20-21h and on Saturdays and Sundays from 12-13 and 16-17h). Established relationships have contributed in determining the ways in which people use certain location and the time.

The presented methodology was applied in research of the territory of Kosancicev venac with the goal to establish current situations pertaining to the manner and intensity of space use from the aspect of pedestrian movement, and define issues that would contribute to its more equal and intensive utilization.

## Results

Following the structure of the research methodologies presented in the previous chapter, results of the research of the territory of Kosancicev venac will be presented.

### Strengths, weaknesses, opportunities and threats of Kosancicev venac

The results of the SWOT analyse is presented in the following matrix.

**STRENGTHS**

Location belongs to the central area of the city
Proximity of Sava River, Belgrade fortress, Kalemegdan park and Belgrade port
Authenticity of the area
Building and cultural heritage site
Proximity of Knez Mihajlova Street and pedestrian zone
Vistas towards New Belgrade
Area is considered as monument of great cultural importance
Important historical, sacral, administrative and public buildings

**WEAKNESSES**

The lack of attractive contents
High level of built heritage protection
Lack of hospitality and hotel facilities
Inadequate access from and to the river
Lack of pedestrian and cycling of infrastructure
Difficult moving on the cobblestones

**OPPORTUNITIES**

Development of tourism
The creation of an authentic space
Increase the attractiveness and recognisability of the location
Transforming Kosancicev venac into pedestrian zone
Establishing the connections between Koscancicev venac and new attractive city sites like – “Belgrade on water”, Savamala, “Beton hala”, “Beko”, “City on water”, etc. (Fig. 2)
Establishing the connections between Sava and Danube’s amphitheatre

**THREATS**

Intensity of traffic in Brankova Street and transit route in Karadjordjeva Street
Railway and Karadjordjeva street seen as a barrier between Kosancicev venac and Sava’s waterfront
Big difference in height between Kosancicev venac and Sava’s waterfront (over 40m) that cause inaccessibility
Poor physical condition of most buildings
Lack of interest of the city and local government for the development of this area

**Fig 2**

*a) Beton hala 1st prize project proposal by Sou Fujimoto Architects, b) “Beko” by Zaha Hadid and c) “City on water” by Daniel Libeskin.*

*Source: Beobuild.rs*

These results are treated as a general input data for detailed analysis of the research area.

### Connecting lines

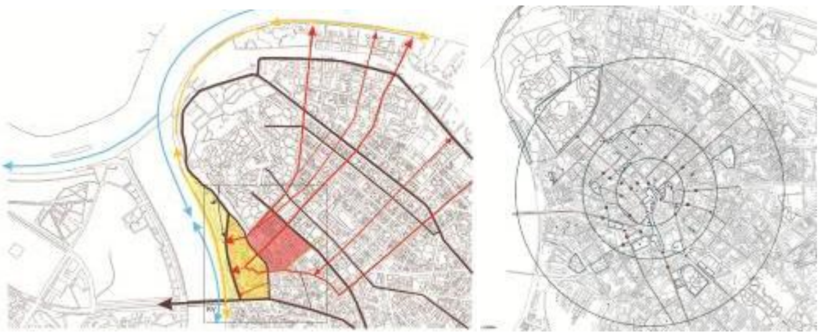


**Fig 3**

*Connecting lines of Sava and Danube’s waterfronts and b) Belgrade city centre – paths and destinations.*

*a) Source: Vasilev M. and Simovic J., 2014 and b) Source: Vukmirovic M., 2013*

Taking into account the geographical position of the location of Kosancicev venac in relation to the Sava and Danube's waterfronts, 6 paths were allocated : 1) path that follow the waterfront, 2) path that extends along the direction of Karadjordjeva Street, Vojvode Bojovica Boulevard and Dunavska Street, 3) path that extends along the direction of Sime Markovica Street, Vojvode Bojovica Boulevard and Dunavska Street, 4) path that follows the directions of Pariska and Tadeuska Koscuska Street and 5) path that extends along the directions of Kralja Petra and Dubrovacka Streets (Fig. 2a). These paths generate direct links between two sides of the city, one that belong to the Sava's waterfront and the other that belong to Danube's waterfront an all of them establish direct connections with the area of Kosancicev venac.

Observing the area in relation to the network of pedestrian spaces in the Belgrade's central area, Kosancicev venac is situated in the area determined by radius of 1000m (Vukmirovic, 2013). Keeping this in mind, certain destination[3] on Kosancicev venac could be directly integrated in the pedestrian space network of the centre of the city by improvement and following the path along Kralja Petra Street and path along Zadarska and Tolicin Venac Street (Fig. 2b).

**The legibility of Kosancicev venac**

The results for this phase of analysis are formed by summarisation of the particular results of the survey of residents and visitors of Kosancicev venac divided into 5 groups depending on which element[4] of "the image" it refer. The study included 400 respondents – pedestrians, surveyed on the territory of Kosancicev venac. Most of the respondents (58%) said that they are living on Kosancicev venac or in the immediate neighbourhood. The other group, characterised as visitors is divided in two subgroups: a) people who like to spend time on Kosancicev venac and people who are just passing through.

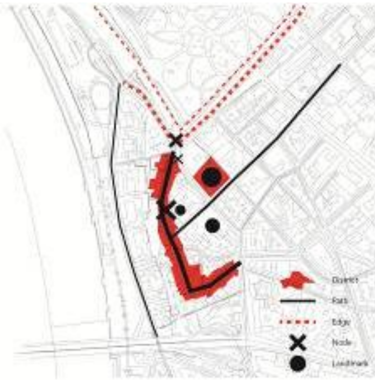


**Fig 4**

*Image of Kosancicev venac.*

*Source: Drobnjak B., 2014*

First group of questions was related to the topic of districts. By interpreting the received answers it was determined that the most visited is the area of the Kosancicev Venac Street, because of its aesthetic and ambient qualities and hospitality facilities, like cafes, restaurants and wine bars. This area is visited during the day, but also during the evening. Other important district that is recognised on this area is Saborna church with its churchyard.

In relation to paths, Kosancicev Venac Street has also been recognised as the most visited and the most pleasant because of its aesthetic and ambient qualities. In addition, Kosancicev Venac Street is marked as public place which need reconstruction and improvement. Kralja Petra and Karadjordjeva Street are also recognised as important paths of the research location.

Pariska Steet is identified as an edge, or clear boundary between the area of Kosancicev venac and Kalemegradn's park on the north. On the south, respondents have allocated Brankova Street as a borderline, as well as barrier between the area of Kosancicev venac and the area of Savamala.

As nodes the most respondents have recognised cafes in Kosancicev Venac Street and crossing of Pariska and Sime Merkovica Street as the most frequently used point. The cafes are seen as popular meeting points. On the other side, crossing of Priska and Sime Markovica Street is the transport hub of two important streets of the city.

[3] Belvedere in Kosancicev venac Street and space of destroyed National library.

[4] Districts, edges, nodes, paths and landmarks.

**Fig 5**
*Urban silhouette of Kosancicev venac.*
*Source: Urbanistički zavod Beograda, 2009*

As the most important landmark of this area and it can be said of the entire city is Saborna Church. Its bell tower dominates at the urban silhouette of this part of the city (Fig. 5). Konak Kaneginje Ljubice, known by its architectural characteristics, is on the second place. Café "Skica" is on the third place, as one of the most known café in this area.
Based on the existing results, the overall image of Kosancicev Venac, defined by its elements, is showed in Fig. 4. Allocated buildings and places could be treated as the basis for upgrading and urban development of this area, because they define its character and authenticity.

**Content units in ground floors**

The analysis of the content units in ground floors of buildings consists of three segments. The results of the first segment of the content analysis show 12 types of activities that could be find on the territory of Kosancicev venac. Identified types of activities includes: services, commerce, restaurants and cafes, transport, business, culture, health, administration, churches, playgrounds, crafts and other activities. In general, this location is mixed-use in its character (Fig. 6a), but there are specific parts of the territory dominated by certain activities (area of Saborna church, part of the Kosancicev Venac Street with cafes, etc.).

**Fig 6**
*Analysis of activities in the ground floors: a) types of activities, b) detail analysis of ground floor activities and c) densities of activities.*
*Source: Jelisavac J., Skerovic J. and Vukovic T., 2014*

The second stage of this analysis encompasses a detailed review of all the existing content on the site. The results show that there are 87 activities placed in the ground floors and oriented towards the public spaces (Fig. 6b). It was concluded that the most of the activities are cafes and restaurants, than shops and, on the third place, cultural contents.
The last stage of this analysis was aimed at determining the rhythm (or the density) and the level of diversity of the activities along the observed streets of Kosancicev venac. Observing the number of contents that occur on 100m, 4 categorises of streets were identifies. Category I (red) with more than 8 contents on 100m, category II (yellow), from 6-8 contents on 100m, category III (green), from 4-6 contents on 100m and category IV (blue) less than 4 content on 100m. Pariska Street is public space with the less number of contents, 2.33 on 100m, and it belongs to category IV. Streets that are the most prevalent in the area of Kosancicev venac belong to category IV (Fig. 6c). Concerning the other streets, 3 of them belongs to category III, 4 belongs to category II and 2 belong to category I. Toplicin venac (14.17) and Pop Lukina (10.95) are streets with the highest number of contents per 100m.

**Intensity of use of public space**

Based on the resuts of analysis of the number of pedestrians on observes public spaces it can be concluded that the highest concentration of users can be found in streets that belong to the parts of the area that are close to the Knez Mihajlova street, as the main pedestrian street in Belgrade. During the whole week, pedestrian movement with the maximal intensity (more than 600 pedestrians per hour) is recorded in this area.

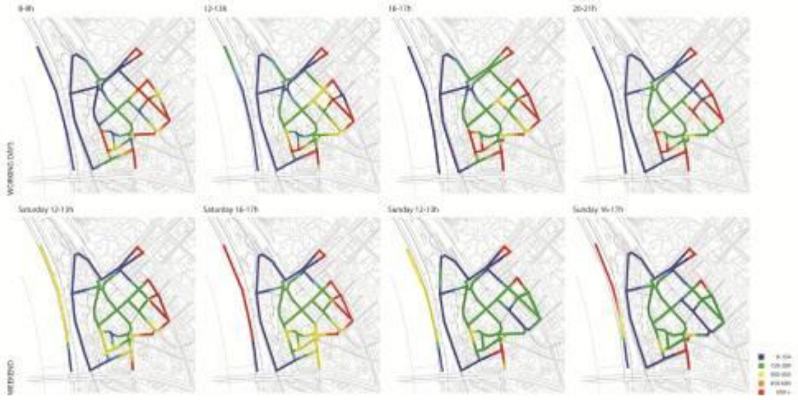


**Fig 7**

*Intensity of pedestrian movement on working days and weekend.*
*Source: Kozovic K., Zivadinovic I., Spasojevic J., Lucic A., Mackic I. and Santa O., 2014*

Maximal intensity of pedestrian movement is noted in Zadarska, Srebrenicka and part of Kosancicev Venac Streets during on working days and on the waterfront (close to the river) on Saturday and Sunday in the afternoon (16-17h). Minimal intensity of pedestrian movement (less than 150 per hour) is recorded in Karadjordjeva and Pariska streets, as well as the routes of Small and Big Stairs – direct connections of Kosancicev venac with the waterfront. On the other streets that belongs to the territory of Kosancicev venac, we recorded the intensity of 150-300 pedestrians per hour.
Based on presented results, it can be concluded that this area should increase the number of contents along the streets and other public spaces, which would contribute to a growth of attractiveness, visits and the number of pedestrians who walk and spend time there.

**Potentials and guidelines for improvement of the area of Kosancicev venac**

In order to activate and improve the quality of Kosancicev Venac as very important part of the Belgrade central area, based on the research results presented in previous chapter, the main problems, value criteria, objectives and vision are defined and presented in the diagram:

| | | | |
|---|---|---|---|
| Lack of direct connections with waterfront and poor quality of existing links<br>An insufficient number of contents in ground floor<br>Overlooked cultural contents and activities<br>Unattractiveness of public spaces<br>Deficiency of urban furniture and equipment<br>Lack of maintenance | **PROBLEMS** | **OBJECTIVES** | Establishment of the logic of pedestrian movement<br>Emphasising cultural and historical heritage<br>Public space revival<br>Increasing the number of contents<br>Creation of multifunctional spaces |
| Pedestrian space network establishing<br>Safety and security<br>Accessibility<br>Comfort<br>Attractiveness | **VALUE CRITERA** | **VISION** | Establishment of direct connection between centre of the city and its waterfront across the area of Kosancicev venac<br>Pedestrian space network that connect main city locations in the centre of the city, grounded on the specific atmosphere of connected urban places and their cultural and historical importance and the multifunctional character of linked public spaces. |

Next to the objectives and established logic raised from the research results, the guidelines for improvement include the following mini urban design strategies considered for the area of Kosancicev venac:

1) An increase in activity along Kralja Petra Street, Big Stairs by creating and inserting new contents and the formation of new path that will directly connect the belvedere on Kosancicev venac with the Sava's waterfront (Fig. 8).

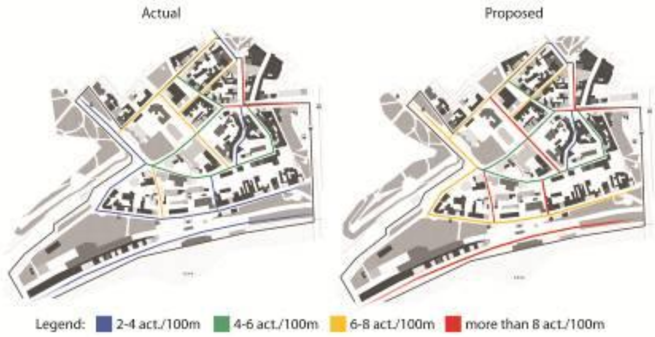



**Fig 8**

*Density of activities in ground floors. a) Actual and b) proposed.*

*Source: Jelisavac J., Skerovic J. and Vukovic T., 2014*

2) Defining the new program for the location based on the specific character of the location and its surrounding. Stimulating the emergence of variety of contents in the ground floor of buildings, especially those ones that arise the attractiveness of the location (Fig. 9).

**Fig 9**

*Program: a) Analysed area and b) Kosancicev venac Street.*

*Source: Jelisavac J., Skerovic J. and Vukovic T., 2014*

3) Equipping urban places in such way that will contribute in forming a pleasant and welcoming atmosphere (Fig. 10): appropriate street regulation, new green areas, new water surfaces, art installations, appropriate paving and textures, accessible ramps and stairs, café gardens, etc.

**Fig 10**

*Ambient of Kosancicev venac.*

*Source: Vasilev M. and Simovic J., 2014*

## Conclusion

According to the data provided by the research, the number of connected lines through the site is satisfied, the existing landmarks are sufficient and easily readable and recognizable but the position of nodes and their connections should be improved. Furthermore, the evaluation of the open public spaces in Kosancicev venac, which was based on selected criteria, opens some questions and traces new directions for the future planning/design process. The important design parameters are identified, detected problems are established and objectives and visions are emphasized.

The identified problems and challenges could present an important input for every plan and action in public space. The research described in this paper highlights the importance of continuous adjustments and upgrading of public spaces, especially in a specific urban environment which is exposed to contemporary currents and trends. Therefore, the comprehensive process of urban regeneration of Kosancicev venac should also be one of strategic aims which trace further urban revival of this area and provides connection between Sava Amphitheatre and historical city centre. On the other side, focusing on improvement of pedestrian spaces adds in highlighting the importance of small urban interventions appropriated on human scale and the manner in which pedestrians experience the urban environment.

The relationships that are noted between the number of pedestrians, number and character of the contents in the ground floor of buildings and the main element of the image of the location, are enabled the establishment of general guidelines and recommendations for future interventions in generating the network of pedestrian spaces, in defining the character of open public spaces and for equipping in accordance with the needs of pedestrians.

**Acknowledgements**

The paper was realized, as a part of the research project “Modernization of Western Balkans” (No. 177009) and the project “Spatial, Environmental, Energy and Social Aspects of Developing Settlements and Climate Change – Mutual Impacts” (No TP36035), PP1: “Climate change as a factor of spatial development of settlements, natural areas and landscapes”, both financed by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

The parts of the research have been done with Bachelor students from Faculty of Architecture, University of Belgrade on elective courses and with students from Master course of integral urbanism during winter and summer semesters 2013/2014.

**References**

Brilon, W. & Grosmann, M., 1993 Methods for the calculation of the capacity and quality of traffic flow in streets. *Straßenbau und Straßenverkehrstechnik*, No. 669

Crawford, J.H. 2002 *Carfree Cities*, Utrecht: International Books.

Đukić A., Vukmirović M. 2012 "Redesigning the network of pedestrian spaces in the function of reduction of CO2 emission. Case study: Pančevo and Vršac", *Spatium*, (27):31-39

Đukić A., Vukmirović M. 2011 "Walking as climate friendly transportation mode in urban environment. Case study: Belgrade", *IJTTE*, (1), 4: 214-230

Fruin, J.J., 1971 Designing for pedestrians: a level-of-service concept. *Highway Research Record*, No. 355, pp.1-15.
Gehl, J., 2010. *Cities for People*. Copenhagen: Island Press.

Kirsch, H., 1964 Capacity and dimensioning of pedestrian footways. *Straßenbau und Straßenverkehrstechnik*, (33).
Polus, A. & Schofer, J.L., 1983 Pedestrian flow and level of service. *Journal of Transportation Engineering*, No. 109, pp.46-56.

Schubert, H., 1967 Planning measures for pedestrian traffic in cities. *Straßenbau und Straßenverkehrstechnik*, No. 56
Urbanistički zavod Beograda, 2009. *Studija "Beogradsko priobalje" - prva faza.* Available at: http://www.urbel.com/documents/priobalje.pdf [Accessed august 2014].

Vukmirovic, M. 2014, *Pesacki prostor i kompetitivni identitet grada*, Zaduzbina Andrejevic, Beograd (eng. *Pedestrian space and competitive identity of the city*)

Vukmirovic M. 2013, *Znacaj i uloga mreze pesackih prostor u generisanju kompetitivnog ideniteta grada*, Univerzitet u Beogradu – Arhitektonski fakultet, Beograd, Doktorska disertacija (eng. *Role and importance of pedestrian space networks in generating the competitive identity of cities*)

Whyte, W.H., 1988 *City: Rediscovering the Center*. New York: Doubleday.

# MODELLING THE INFLUENCE OF VEHICLE TECHNICAL CONDITION ON ROAD TRANSPORT RELATED EMISSION OF HARMFUL GASES IN SERBIA

**V. Momčilović[1], D. Vujanović[2], S. Bunčić[3]**
[1,2,3] *University of Belgrade - Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, 11000 Belgrade, Serbia*

**Abstract:**The national vehicle fleet of the Republic of Serbia is becoming "younger" not long ago, compared to the fleet from the decade of the 1990's. Nevertheless, today's fleet is still characterized by poor maintenance and not very rigid vehicle inspection practices. Although the state opted for maintaining a yearly periodicity of vehicle inspections for passenger cars after initial two years' period (in case of newly purchased vehicles), vehicle inspection stations are literally competing for the market, and therefore became "tolerant to certain negligible faults" in order to attract clients. This is especially true for emission of harmful gases, which, by their opinion, is not directly jeopardizing traffic safety. Meanwhile, there is still a significant import from the EU of used vehicles of questionable (double-tampered) odometer state: firstly increased in order to avoid high import duties and later lowered so as to obtain a better selling price. There is no effective guarantee of the state vehicle is in, mainly because of the lack of records (reliable data) regarding previous owner inspection and maintenance (I&M) habits, avoiding in this way discovery of manipulations with total distance travelled. Therefore, the authors have chosen to model the relations between emissions and I&M practices in Serbia based on thorough analyses. After collecting relevant emission and I&M data for a vehicle category, as required by COPERT 4 emission assessment software, the aforementioned software will be used to estimate national fleet's total emissions of harmful gases. The emission assessment results will be compared to the referent (average) vehicle fleet and appropriately discussed in the paper.

**Keywords:** National Fleet, Emission Assessment, COPERT 4, Inspection & Maintenance, Odometer Tampering

## 1. Introduction

The vehicle technical condition is a factor that besides traffic safety also impacts the environment. The vehicle inspection system should represent a decisive official enforcement measure to impose all vehicle users a minimum level of care and regular maintenance to maintain all the required indicators above minimum or below maximum (allowable) levels. When a vehicle inspection i.e. roadworthiness test is not realised properly and/or completely then users do not feel the obligation of complying with regulations to the letter and problems begin to appear.
For a time, in Serbia large part of newly registered vehicles are used vehicles (i.e. second hand) mainly arriving from EU countries. While importing a vehicle, the importer seeks to minimise import duties and taxes either by presenting realistic mileage (in case of company vehicles with important total mileage values) either by tampering the odometer to increase the mileage (in case of lower to moderate total mileage values). Another problem is that this information from vehicle import duties is not being forwarded to the authority in charge of vehicle inspection and registration. Meanwhile, importers tamper the odometer (once again) lowering the total mileage in order to obtain better market price for the vehicle.

## 2. Background

In the Republic of Serbia, gross domestic product per capita (GDP per capita) fell sharply in 2001, as a result of changes in exchange rates, while subsequently grows until 2008, when it attains its maximum value in the recent past period. The fall of GDP in the aftermath of 2008, first of all, is the consequence of the global economic crisis. However, slight recovery was observed three years later, in 2011, when GDP per capita reached the value of 4,351 Euros. However, as of the next year, this economic indicator falls again to 4,112 Euros. The previous parameter influenced the motorisation rate, which is also related to the population which evenly decreases in time, i.e. continuously follows the so far observed downward trend. The motorisation rate, expressed in number of passenger cars per 1,000 inhabitants, is steadily increasing, while in the period from 1997 to 2012 increased by almost 62%, from 147 to 238 passenger cars per 1,000 inhabitants. This value nevertheless is below the European average, which in 2011 amounted to 477 passenger cars per 1,000 inhabitants.
In order to determine vehicle roadworthiness the regular vehicle inspection is being realised. The periodicity in the Republic of Serbia is annual for passenger cars (after the initial two years period for newly registered cars) while heavy goods vehicles' inspection are half yearly based (every 6 months). Regular or exceptional vehicle inspection is realised in order to:

- Ascertain vehicle technical condition,
- Identify the vehicle to prevent fraud and
- Protect the environment from vehicle related harmful exhaust gases

The VIS while monitoring and scrutinising the vehicle should be uncompromised and pitiless in terms of not complying with legal requirements bearing in mind the cost side (general people standard of living), especially because of its unique preventive role. On the vehicle inspection station it is checked whether the vehicle owns and maintains operational the systems, devices and equipment important for safe and eco-friendly participation in the traffic flow.

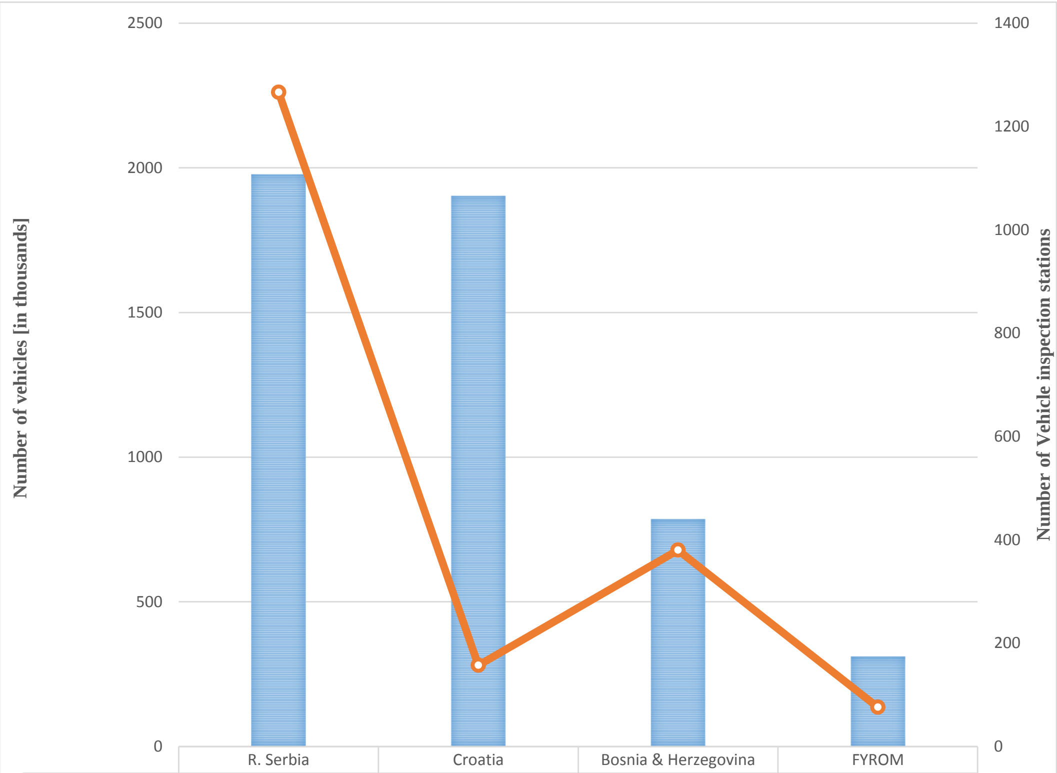
[1] Corresponding author. v.momcilovic@sf.bg.ac.rs

Besides, it is also monitored whether these equipment, devices and systems do comply with local (and pertinent international) regulations. The condition and operation of the following systems is determined: braking mechanism, steering mechanism, lighting and signalling systems, and tires represent some of the most vital systems for the safe traffic and transport. By determining their malfunctions, the vehicles will promptly engage in maintenance in order to eliminate the barriers to participating in the traffic.

Periodical roadworthiness tests enable discovering even small malfunctions or failures that prevent the vehicle from accomplishing safely its primary function of transportation and safe operation. Those failures or malfunctions should not be obligatory critical in view of moving the vehicle, but rather to prevent them to jeopardise other participants and therefore lower traffic safety.

Vehicle inspection is an activity of public interest, which can only be performed by a company that meets the prescribed requirements and that has been authorised to do so.

In our country there are a large number of vehicle inspection stations (more than 1250 in contrary to the neighbouring countries), whose capacities are not properly used, so the owners because of this reason and generally in lack of regular clients searching for confidence and their own safety, finding various ways to attract customers, by discounts or exemptions "while performing vehicle inspections, which should be of concern to the relevant authorities in charge of seamless operation and control (corrective) function of this important institute for safety and roadworthiness tests of motor vehicles. Therefore, the basic idea of the legislator in the implementation of access barriers to this segment of the market need to protect the vehicle inspection system as legally required category by principles of open market, as well as to ensure to existing stations smooth operation and survival, of course with adequate responsibility. Due to this situation, in addition to laws and bylaws and solutions, a comparative analysis of vehicle inspection station network has been realised in order to detect a discrepancy between the former Yugoslav states and to give a basis for effective access control to this segment of the market and thus protect this activity from further degradation. On the following figure 1 is presented the number of registered vehicles per country alongside with the number of vehicle inspection stations which is significantly lower than in Serbia.




| | R. Serbia | Croatia | Bosnia & Herzegovina | FYROM |
|---|---|---|---|---|
| No. of vehicles | 1977253 | 1902630 | 785890 | 311406 |
| No. of VIS | 1266 | 157 | 380 | 76 |

**Figure 1**
*Number of vehicles and vehicle inspection stations per countries from former Yugoslavia region*

So, as a consequence of all the above mentioned, there is a need to prevent further inadequate operation, monitoring and control of vehicles at vehicle inspection stations by letting them to allow some kind of (just to increase profitability of an almost “public service”). Important penalties for infringements should be implemented as well as restrictions (as of banning) from realising such illegal activities.

### 3. Methodology for emission assessment

Based on a survey regarding the operation (vehicle annual activity) and maintenance habits of drivers in vehicle inspection stations, the authors have linked their answers (especially those relating to the mileage, fuel consumption and lack of regular maintenance) to the results from the obligatory measurement of emissions.

In the COPERT 4 model and software, there are two main correction factors of the vehicle emissions taking into account the "vehicle age" on the one hand influenced by engine oil (lubricants) consumption (which will directly influence the $CO_2$ emission) and increased fuel consumption due to the total mileage (vehicle lifecycle).

The first correction factor depicts carbon dioxide emission increase due to ***increasing engine oil consumption (lubricant)*** According to (EEA, 2009) "*New and properly maintained vehicles normally consume small amounts of lubrication oil, due to the oil film developed on the inner cylinder walls. This oil film is exposed to combustion and is burned along with the fuel. Wear due to prolonged engine operation usually increases lube oil consumption, so this should be expected to increase, on an average, with vehicle age. A different vehicle category, the ones operating with 2-stroke engines, consume much more lubricant oil as this is fed in the intake of the vehicle in blend form with the fuel or through a separate injector. A much higher lube oil quantity is needed in this case, which is practically completely combusted in the cylinder. Oil combustion, although a less important factor than fuel combustion, also leads to $CO_2$ production and should be taken into account in the national totals for completeness.*"

According to the table 3-28 from (EEA, 2009) engine oil (lubricant) consumption that is considered to be regular for passenger cars (PC) and light (duty) commercial vehicles (LCV) consuming petrol is from around 0.85 to 1.70 kg / 10,000 km for new, or up to 2.13 kg for old vehicles. As for the diesel PCs and LCVs it is estimated at 0.43 for new, i.e. from 0.85 to 2.13 kg / 10,000 km for old vehicles. Average lube oil consumption for old urban buses is around 8.50, while for new buses is up to 10 times lower with 0.85 kg / 10,000 km. For all heavy goods vehicles (HGV) as the average lube oil consumption it is adopted the value of 1.56 kg / 10,000 km.

If the values of engine oil consumption in the vehicle fleet overcome the above mentioned values the correction factor should be applied. In this sense, it is necessary to assess an average lube oil consumption ($U_k^{lo}$) and express it in [*t*/*km*], i.e. divide it by $10^7$ and put it in the adequate formula instead of the fuel consumption ($U_{k,m}$).

Upon experience of our vehicle fleet and maintenance managers (in private transport companies) lube oil consumption is importantly higher than recommended of 1 *l* at 1000 km, which is almost 10 times higher than the mentioned value (~10 l / 10,000 km). Contrary to this, as importantly raised lube oil consumption in state owned transport companies and vehicle fleets is considered the consumption of around 10% of the fuel consumption on 1,000 km, which in average represents somewhere approximately 30 litres of lube oil per 10,000 km.

The second correction factor to be introduced in the mentioned formula upon need (based on the expert choice – chosen by experience in emission assessment) is the one that correctly depicts the difference between the emissions ***caused by vehicle age***, which is the ***worsening of emission based on the total mileage*** since the entrance in transport operations.

Since the base emission factors used in the formula are referring to the vehicle fleet with the average value of mileage between 30,000 and 60,000 km they have directly incorporated certain factor of "worsening" of the quantities and composition of harmful gases emissions. Nevertheless if the total mileage of petrol driven passenger cars and light commercial vehicles find itself out of the mentioned interval to the emissions of those vehicles should be applied the correction factor that realistically portrays the worsening of the quantities and composition of the emission.

In the European Union it is recommended that this correction factor should not be implemented into the road transport related emission inventory in order to obtain totally complying data on emissions of different member states, because it is established that the vehicles in the EU are not old at all. However, if some important differences appear in the vehicle fleet age or it is expected to sense the influence of this effect in the traffic projections, it should be reconsidered to introduce it as a correction factor regarding the emission based on total mileage, so as to adequately portray the age and technical condition of vehicles.

In the calculation (assessment), it is foreseen the application of a correction factor on base emission factors for petrol driven passenger cars and light commercial vehicles according to the vehicle age. These correction factors are demonstrated by the following formula:

$$c_{D_i} = A_D \times \overline{D} + B \tag{1}$$

where:

$c_{D_i}$ correction factor of the mileage for the given average distance travelled and emission of the i$^{th}$ pollutant,

$\overline{D}$ average mileage of the vehicle fleet (of "old") vehicles that the correction is applied to,

$A_D$ increase in emission per kilometre based on the "age",

$B$ level of emission of the vehicle fleet of new vehicles.

The value B is less than 1 because correction factors have been determined on the vehicle database whose mileage is in the interval between 16,000 and 50,000 km. New vehicles will have lower emission from the emission of vehicles that the emission factors were based on. It is considered important that composition and quantities of emission are worsening after 120,000 km for vehicles of EURO 1 and 2 technology, meanwhile just after 160.000 for the vehicle technology EURO 3 and 4.

The influence of the average speed on the "worsening" of the emission is taken into consideration by combining functions with heavier traffic conditions on urban and extra urban traffic conditions.

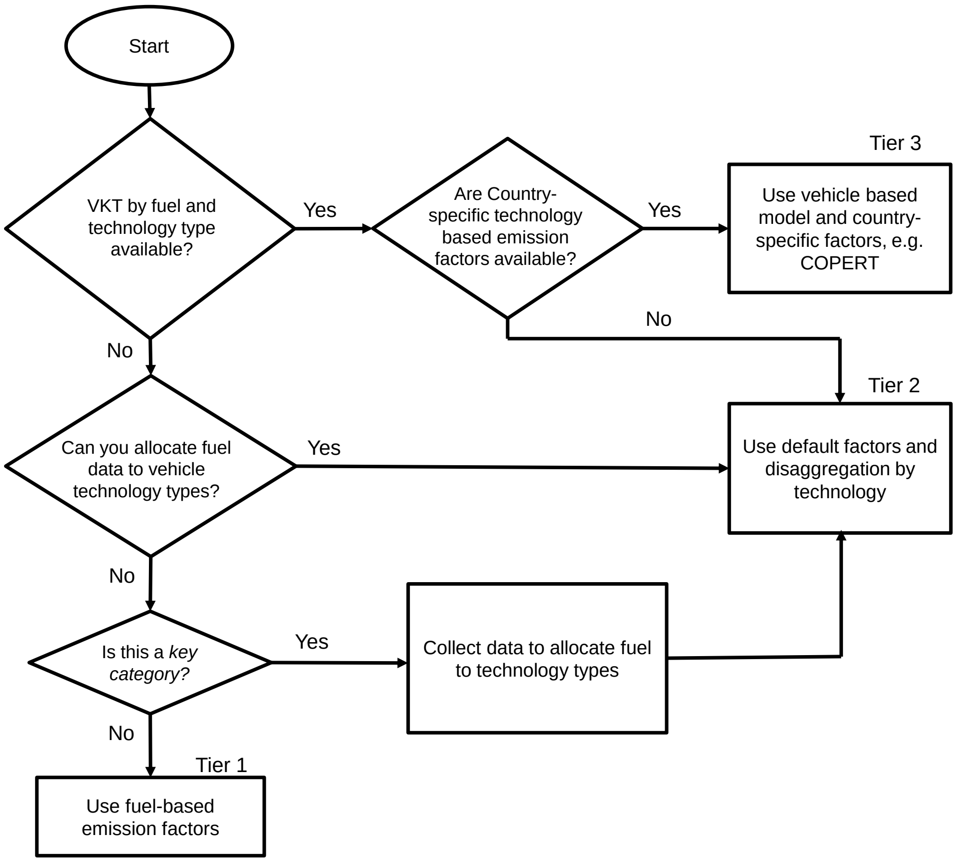



**Figure 2**
*Decision making algorithm regarding determination of $CH_4$ and $N_2O$ emissions depending on the importance of the factor (is it a key vehicle category) and the availability of data*

**4. Research of the emission assessment input parameters**

Research of the indispensable input parameters is conceived in the following phases:
1. The elaboration of a method for determining the relevant average annual distance travelled per vehicle categories defined according to the relevant vehicle classification for the Republic of Serbia;
2. The realization of surveys to determine the values of influential factors onto the value of the relevant average annual distance travelled upon vehicle categories and related indicators influencing road vehicle emissions;
3. The synthesis of the survey results in terms of preparation of input parameters for road transport related emissions evaluation;
4. Comparative analysis of the results of road transport related emissions in the Republic of Serbia from previous years, with recommended international values of input parameters.

Before the realisation of the study, there was a necessity to clean-up and correcting all erroneous data as well as filling out the missing values (by estimation or expert deduction) within the relevant vehicle database, i.e. vehicle register, obtained from the Ministry of Internal Affairs of the Republic of Serbia.

In terms of the method for determining of average distance travelled, an independent survey was based on statistical analysis of the vehicle database, previously adequately structured and cleaned up, due to a certain number of incorrect and illogical data (especially in terms of fuels). On the other hand, some data is missing, one of the most important is the implemented emission control technology (i.e. current Euro standard) used for subdividing vehicles into classes. The major obstacle to high accuracy assessment was the lack of "personal" information in the vehicle database, where crucially important were the Vehicle Identification Number, vehicle registration number which were not delivered because of the potential insight into personal data compromising the right of privacy or hindering the protection of citizens' private personal data.

The surveys consisted of drivers' survey at periodical (regular) vehicle inspections. Drivers' survey aims to quantify and determine their behaviour i.e. annual distance travelled (mileage) and total distance travelled (which allows to determine the average annual distance travelled from the moment of procurement of a given vehicle). In the following tables are given some of the survey results regarding mileage and fuel consumption.

**Table 1**
*Mean annual mileages per purposes of usage and regions in the Republic of Serbia [km] (Survey 2012)*

| Region | Private | Company | Average |
|---|---|---|---|
| Belgrade | 10,528 | 31,219 | 14,809 |
| Vojvodina | 10,753 | 30,597 | 12,726 |
| Southern and east Serbia | 8,869 | 25,286 | 9,684 |
| Šumadija and western Serbia | 9,936 | 23,227 | 10,526 |
| **Republic of Serbia** | **10,085** | **29,347** | **11,952** |

**Table 2**
*Mean total mileages per purposes of usage and regions in the Republic of Serbia [km] (Survey 2012)*

| Region | Private | Company | Average |
|---|---|---|---|
| Belgrade | 110,904 | 115,220 | 111,816 |
| Vojvodina | 139,996 | 167,392 | 142,731 |
| Southern and east Serbia | 146,242 | 219,744 | 149,891 |
| Šumadija and western Serbia | 142,172 | 193,704 | 144,458 |
| **Republic of Serbia** | **136,264** | **150,983** | **137,709** |

The results of the survey (tables 1 and 2) regarding the average annual mileage show that there is an important difference between the individual users that use their cars for private purposes in regard to company vehicles. Company vehicles, not only make a more important mileages but are also more aggressively and less carefully driven. There is also a slight difference between regions that is consistent with the GDP evolution: Belgrade and Vojvodina are more developed than Šumadija and western Serbia (that follows), meanwhile southern and east Serbia is actually the poorer and with lower mileages. Meanwhile regarding total mileages (private) it is a quite different situation, Belgrade has the lowest total mileage because of on one side a shorter lifecycle and on the other probably lower daily mileages (due to urban congested roads, etc.).

**Table 3**
*Mean fuel consumption per fuels and regions [l/100 km] (Survey 2012)*

| Region | Petrol | Diesel | LPG | Average |
|---|---|---|---|---|
| Belgrade | 8.04 | 7.37 | 10.34 | 8.18 |
| Vojvodina | 7.49 | 6.63 | 9.28 | 7.69 |
| Southern and east Serbia | 8.02 | 6.70 | 9.82 | 8.05 |
| Šumadija and western Serbia | 8.03 | 6.71 | 9.31 | 7.91 |
| **Republic of Serbia** | 7.86 | 6.85 | 9.57 | 7.93 |

**Table 4**
*Mean fuel consumption per fuels and purposes of usage in [l/100 km] (Survey 2012)*

| Purpose of usage | Petrol | Diesel | LPG | Average |
|---|---|---|---|---|
| Private | 7.78 | 6.66 | 9.48 | 7.85 |
| Company | 8.66 | 7.95 | 11.00 | 8.65 |
| **Republic of Serbia** | 7.86 | 6.85 | 9.57 | 7.93 |

As regarding the fuel consumption it can be concluded that the lowest consumptions are made by diesel vehicles although higher than factory declared (type-approval) values. On the second place are the petrol cars, and expectedly the worse are the LPG, because of the often unprofessional retrofitting (due to potential cost savings). The fuel price justifies the higher consumption but it still should not be for more than 20% (which it actually is). As for the comparison between company and private the same pattern is shown here: company cars spend around a litre more of each fuel than the private.

The second parameter is introduced in order to detect if the last annual mileage is close to the average value (over the entire present lifecycle) or significantly different, as well as to determine the actual vehicle “age” in terms of past and remaining operation lifecycle until the vehicle write-off. In addition, a significant output is the share of urban and extra urban driving conditions (percentage), which will affect the modelling of the relevant fuel consumption and exhaust emissions. An important indicator is the purpose of vehicle use, which may be for individual (private) or for official (business) purposes. The impact of the last-mentioned is expected to increase the average annual and the total vehicle mileage. In order to determine the vehicle technical condition, in terms of maintenance, the survey comprised a question on average fuel consumption, used as a correction factor of type-approval (declared) value of fuel consumption for different operating conditions (mainly urban, extra-urban, or combined conditions). However, as pointed out and commonly known that even new vehicle average fuel consumption (and thus emissions) sometimes differs significantly from the type-approval values. Among the most important factors influencing differences among declared, average and real-time fuel consumption, there are: driving style (i.e. driver: aggressive / moderate / passive), the traffic flow conditions (free flow / uncongested flow / congested flow) and climatic conditions (average atmospheric temperature / humidity / pressure / precipitation: rain, snow, etc.), as well as a variety of other parameters. All of the above questions are intended to enable driver and vehicle user behaviour modelling and thus directly and indirectly enable relevant vehicle fleet operation modelling in terms of actual average distance travelled, fuel spent and air pollution i.e. harmful and greenhouse gas emissions. In addition, a secondary aim was to determine whether relevant factors differ and how different are the values of respective factors and indicators in the R. of Serbia compared to countries in the region and in the European Union, as to be able to compare values with the reference from this study, as well as in order to observe realized progress in relation to the desired (even designated or committed) one.

The European Commission in 2001 (EC, 2001) has established an internal Transport and Environment Reporting Mechanism (TERM), and the European Environment Agency (EEA) in its report (EEA, 2011) introduced a Core Set of Indicators (TERM CSI) referring to transport that must be monitored by the Member States and that serve as a "measure" of their activities in this regard. One of the last indicators *TERM 34: Proportion of vehicle fleet meeting certain emission standards* (crucial for knowing the age and technological structure of the fleet and possibility for comparisons with other countries) is imposed as interesting for two reasons: a) this indicator is critical for the assessment of vehicle emissions and b) this information is not recorded in the national vehicle database. Since Serbia intends to fulfil all European requirements in order to remove barriers for its entry into the EU, this indicator will have to be accurately determined in the near future, which is not at all simple.

Very significant motor vehicle internal factors will be discussed equally as their impact on vehicle emissions relating to the following:

- Structural (built-in) characteristics, and
- Technical operational (TO) vehicle characteristics.

As far as vehicle structural i.e. "built-in" characteristics there are several problems, which at first glance do not seem problematic. The first problem represent the declared vehicle emission values, which are determined by a standardized type-approval procedure. In the context of this problem, very significant segment is determining the relevant evaluation/measurement procedure of heavy commercial vehicle emissions presented in (TU Graz; TNO; TÜV Nord Mobilität; VTT; AVL; LAT; H.S. Data analysis and Consultancy, 2012).

The second problem is a negative trend of vehicle mass increase, especially regarding passenger vehicles, over the last decade. The indicator of vehicle weight can be observed and analysed individually, but also through other derived indicators as for example per vehicle capacity (kg/passenger seat), per vehicle occupancy (kg/passenger) or even per specific energy consumption unit (kg/energy unit). The mass increase itself influences an increase in rolling resistance, as well as in "engaged work" for starting and stopping the vehicle, which in addition to greater energy consumption affects the increase in air pollution. One of the influential factors causing the mass increase within the same vehicle classes in recent years is a growing number of additional systems (many of which have become part of the vehicle "standard equipment"), built-in for comfort (e.g. air condition, audio-video equipment) or for increasing vehicle safety (increased braking efficiency, traction control, electronic stability program, tire pressure monitoring systems, etc.). The operation of these systems and devices requires additional energy consumption as well (not of the same order of importance for all the devices, but still some increase). Every city should proceed with extreme caution because even in highly developed and environmentally aware cities (as London and Stockholm, that introduced congestion charging), restrictive measures fail to decrease the share of energy- and environmentally-inefficient vehicles accessing those sensitive densely populated areas. Inefficient vehicle owners and fleet managers obliged to pay a fixed time-based fee (daily, monthly or annual) and not the mileage- or entry-based depending on vehicle category (weight), do not share common social interests and neither show willingness or motivation to increase their weight efficiency: load factor in case of goods or occupancy in case of passengers. The weight inefficiency in goods transport is essentially based on two noticed trends: generally low load factor and growing ratio of empty runs of everyday heavier goods vehicles. It is likely that only severe restrictions would influence their current behaviour. Nevertheless, the authors do not anticipate that even introduction of really drastic measures like banning the access for goods vehicles to sensitive urban areas will lead to expected effects, but would most likely increase number of infractions and yet relocate the problem creating new critical areas (Momčilović, et al., 2010).

Another interesting ratio is that of vehicle power and weight. In this respect, (Kageson, 2005) showed that weight and power rating of new passenger cars increased significantly during the 1980s and 1990s. Not only those buying large cars, but also customers of small and medium-size cars were increasingly offered a variety of engine sizes and power ratings. In a study realized for (ECMT, 1995) it was found that average power ratings rose by more than 9 kW between 1980 and 1990 in France, Germany, Sweden and the United Kingdom.

The third important element is the increase in installed engine power (for same vehicle classes) due to the aforementioned increase in the number of power-driven equipment in the vehicle and the growing expectations of drivers in terms of vehicle performance. In the paper (Cvetković & Momčilović, 2011) based on the example of the passenger car Volkswagen Golf is illustrated the trend in mass and engine power (as well as vehicle dimensions) over a 30 years period between models Golf I (from 1974) and Golf V (from 2003). The total vehicle mass and installed engine power have almost doubled. "The amount of embedded electronic components in the current high-class vehicles clearly indicates that this trend will continue and spread in the future as consequence of modern systems' implementation, from initially higher vehicle classes to lower classes on customer request. Electrical installations in modern high-class vehicles reach a length of up to 4,000 m, with a weight over 65 kg and with over 2,000 electrical installations." (Cvetković & Momčilović, 2011)

Another major trend observed in last decade is the increased share of large vehicles in the fleet of passenger cars. More and more users while purchasing a new vehicle opt for a change from conventional passenger vehicle sizes to the larger so-called crossover models, even SUVs that used to be mainly off-road utility vehicles, today represent only a status symbol.

Another significant factor influencing increase in fuel consumption and (excessive) emissions is definitely vehicle technical condition (Bin, 2003), which in our country is often dominantly inadequate due to unaccomplished basic role of vehicle inspection, as an institution, but also due to the irresponsibility of vehicle owners. In this sense, authors (Manojlović, et al., 2011) suggest that in order to manage vehicle technical condition, throughout vehicle operation, responsibility for certain technical interventions should be assigned to the vehicle owner, workshops and vehicle inspection. These technical interventions should be well conceived, with a defined technology, frequency, workers and acceptable realization costs. Properly operating modern vehicles with three-way catalysts are capable of partially (or completely) converting engine-out CO, HC and NO emissions to carbon dioxide (CO2), water and nitrogen. The institute of regular vehicle inspection influences periodical restoration of vehicle's "good condition" (at least once a year in Serbia), meanwhile extraordinary (unannounced) vehicle inspections should influence drivers' responsibility to maintain the vehicle between two vehicle inspections, assisted by modern vehicle On-Board Diagnostic (OBD) systems, which by constantly monitoring vehicle degradation process ensure maintaining of vehicle condition parameters in the permitted range. "*The purpose is to establish a complete system of accountability: monitoring vehicle degradation, timely implementation of activities, traceability in liability determination for failure to accomplish the obligations of the vehicle owner, maintenance workers, vehicle and spare parts manufacturers. Thus, the focus of vehicle technical condition management is relocated in order to assign the responsibility to the driver (vehicle owner) and regular maintenance (technology, skilled workers and original spare parts).*"

On the other hand, at periodical and extraordinary (unannounced) vehicle inspections emission levels should be checked and compared with the permitted limit values according to technology, and Euro emission standards (Bin, 2003). In most European countries the statutory defined and detailed procedures are performed, although there are indications that in some countries it's easier to meet these requirements than in others. Regarding the Republic of Serbia, this, as noted, has not been given special attention, both from the standpoint of the legislator i.e. control bodies, in terms of maintenance network representing the automotive industry in this important task.

One of factors influencing the most vehicle emissions, derived from its operation (method and intensity), is the age, as well as the "current" lifecycle. Older vehicles have generally attained much greater values of distance travelled, and therefore the emission control system on these vehicles reach their end-of-life, which in turn will result in an increase in emissions of these vehicles. On the other hand, greater share of older vehicles in the fleet implicitly suggests that large number of vehicles in the fleet do not meet the modern, more stringent greenhouse gas emission standards (Zachariadis, et al., 2001).

Total vehicle emissions largely depend on the conditions of engine operation. Different traffic conditions (which are linked to the relevant road categories) impose various engine operation conditions, and thus affect the amount of emitted pollutants. In this respect, we distinguish urban, extra urban and highway traffic conditions. Entered data (pre-determined) about vehicle activity and emission factors are assigned to given traffic conditions. Thus, the emissions of certain pollutant $i$ during a "cold start" is first allocated to urban conditions, then to extra urban, and the lowest (negligible) part to highway (it is believed that only a very small number of trips start at or near the highway).

## 5. Conclusions and future research

Definitely, in the survey realised in 2012 by authors and the follow-up analyses an important influence of the vehicle technical condition on the emission of certain pollutants was observed. Besides, emission factors for Serbian vehicle fleet of passenger cars and light commercial vehicles were adequately corrected by lube oil consumption and vehicle age (total vehicle mileage). The same approach was applied for buses and HGVs with adequate values of tolerances regarding the influence of increased fuel and lube oil consumption.

On the other hand, an important pressure has to be done on vehicle inspection stations to measure and respect the legal requirements in view of emission limits for each vehicle technology, which is not the case presently.

As a matter of future research, one of the measures to raise the awareness of the vehicle users is a voluntary on-road remote sensing of CO, HC and NO. It should be conceived as an informative vehicle emission survey comparing those results to the vehicle emission norms (standards) ultimately meant to inform the drivers about their eco efficiency and their legal "status", therefore timely warning drivers about their obligation to maintain their vehicle adequately.

## Acknowledgements

This paper is based on the research on the following technological projects: the project code TR36027 named: "Software development and national database for strategic management and development of transportation means and infrastructure in road, rail, air and inland waterways transport using the European transport network models" and the project code TR36010 named: "Development of the model for managing the vehicle technical condition in order to increase its energy efficiency and reduce exhaust emissions" both supported by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

## References


Bin, O., 2003. A logit analysis of vehicle emissions using inspection and maintenance testing data. *Transportation Research Part D: Transport and Environment,* 8(2003)(3), pp. 215-227.

Cvetković, M. & Momčilović, V., 2011. New technologies and European Union legislation as a basis for vehicle inspection development in the Republic of Serbia, (original in Serbian). In: V. Papić & A. Manojlović, eds. *Towards Sustainable Transport 2011.* Belgrade: University of Belgrade, Faculty of Transport and Traffic Engineering, pp. 109-125.

EC, 2001. *White paper – European transport policy for 2010: Time to decide,* Brussels: s.n.

ECMT, 1995. *Trends in Power Ratings of Cars and Heavy Goods Vehicles,* Paris: ECMT/OECD.

EEA, 2009. *EMEP/EEA air pollutant emission inventory guidebook 2009 - Technical guidance to prepare national emission inventories,* Copenhagen: s.n.

EEA, 2011. *Laying the foundations for greener transport, TERM 2011: transport indicators tracking progress towards environmental targets in Europe,* Copenhagen: Office for Official Publications of the European Union.

Kageson, P., 2005. *Reducing CO2 Emissions from New Cars, A progress report on the car industry's voluntary agreement and an assessment of the need for policy instruments,* s.l.: T&E, European Federation for TRANSPORT and ENVIRONMENT.

Manojlović, A., Milašinović, A. & Papić, V., 2011. Determination of the vehicle inspection system quality, (Original in Serbian). In: V. Papić & A. Manojlović, eds. *Towards Sustainable Transport 2011.* Belgrade: University of Belgrade, Faculty of Transport and Traffic Engineering, pp. 127-139.

Momčilović, V. et al., 2010. *Measures for lowering CO2 emissions from goods transport in urban environment.* New Delhi, India, Transportation Science and Technology Congress.

TU Graz; TNO; TÜV Nord Mobilität; VTT; AVL; LAT; H.S. Data analysis and Consultancy, 2012. *Reduction and Testing of Greenhouse Gas Emissions from Heavy Duty Vehicles - LOT 2, Development and testing of a certification procedure for CO2 emissions and fuel consumption of HDV,* s.l.: European Commission.

Zachariadis, T., Ntziachristos, L. & Samaras, Z., 2001. The effect of age and technological change on motor vehicle emissions. *Transportation Research Part D: Transport and Environment,* Volume 6(2001), pp. 221-227.

*Transport and Economic Growth*

# SOME CONCERNS ON THE ROLE OF HIGH SPEED RAIL SYSTEMS IN TOURISTS' DESTINATION CHOICE

**Francesca Pagliara**[1]
[1] *Department of Civil, Architectural and Environmental Engineering – University of Naples Federico II*

**Abstract:** Tourists have different pictures of their ideal holiday and they are heterogeneous. Market segmentation is the strategic tool to account for heterogeneity among tourists by grouping them into market segments including members similar to each other and dissimilar to members of other segments.
In this respect, an important aspect is the study of the accessibility provided by transport modes which can affect the choice of a given destination for holiday. Transport and tourism are connected, since they represent an important economic activity in most countries around the world. In the last years, major investments on High Speed Rail Systems (HSRS) have been recently carried out all around the world. The objective of this paper is to analyse the factors influencing tourists' destination choice and the impact of HSRS in this choice. Tourists, segmented by different socioeconomic characteristics are analysed together with their different behaviour w.r.t. the choice of the transport mode, using a recent survey carried out for the city of Naples, in the south of Italy.
Results show that several factors affect the choice of this destination and the accessibility provided by the presence of a HSR link connecting Naples with other Italian cities plays its role.

**Keywords**: Tourism market; High Speed Rail, accessibility, destination choice.

## 1. Introduction

An interesting aspect of the tourism market is to identify the reasons explaining tourist's intention to revisit a specific destination. In this respect, HSR might be one of the reasons for tourist to revisit a destination whether this transport alternative is available (Pagliara et al., 2012). Very few contributions are present in the literature. The paper by Seddighi and Theocharous, (2002) analyses the probability of revisiting Cyprus w.r.t. socio-demographic and destination characteristics. In this paper a micro-econometric approach, based on observations of holidaymakers, is proposed. This approach allows the analysis of the characteristics influencing individual travel behaviour. Moreover it provides a conceptual/methodological framework for the understanding of the nature, form and character of the holiday-decision-making processes of tourists. Another research work analyses the variables influencing the probability of revisiting Lisbon by using mixed logit models with bounded parameters. The probability of revisiting Lisbon "*increases significantly with accommodation range, events, food quality, expected weather, beach, overall quality, nightlife, reputation, and safety*" (Barros and Assaf, 2012). From the literature it emerges a lack of contributions analysing the relationship between tourism and transport from a quantitative point of view, and HSR in particular, except for the contributions by Valeri et al. (2012) and Delaplace et al. (2014) for the case study of Rome and Paris and Pagliara et al. (2014) for the case study of Madrid.

With more than 46.1 million tourists per year, Italy is the fifth most visited country in the world, behind France (79.5 million), United States (62.3 million), China (57.6) and Spain (56.7 million). People mainly visit Italy for its art, history, fashion and culture, its beautiful coastline and beaches, its mountains, and old monuments. Italy also contains more World Heritage Sites than any other country in the world. Tourism is one of Italy's growing and most profitable industrial sectors, with an estimated revenue of € 136.1 billion. In 2013 Italy's tourist industry overall turnover was about € 136.1 billion, thus confirming as an ever expanding sector: contributing to over 8.6% of GDP; employing over 1.1 million qualified workers; including more than 153.000 firms throughout Italy's Regions (World Travel & Tourism Council, 2013).
The objective of this paper is to analyse the factors influencing tourists' choice of a destination for tourism purpose and the role of High Speed Rail (HSR) systems in this choice. Section 2 describes the case study of Naples, estimation results of the factors influencing the choice of this destination are highlighted together with the role of HSR systems in this choice. A comparison with the case studies of Madrid and Paris is reported as well in section. Conclusions and further perspectives are reported in section 3.

## 2. The case study of the tourist destination Naples

In the last few years Naples, in the south of Italy, has become one of the favorite destinations for all those Italian and foreign tourists who love spending their holidays in cities of artistic interest. This is also due to the opening of the new HSR link at the end of 2005 (Cascetta et al., 2011).
Unlike cities in which art is stored in museums and daily life happens on the streets, Naples's distinctive mark is its folklore: people living and working among the artistic beauties of the city. Tourism has become a key factor in the city's economy. Buildings, churches, streets, ancient fortresses and castles in the sea, as well as natural caves and places impregnated with mystery and mysticism, all of which makes Naples unforgettable in the minds of all travellers (Pagliara and Papa, 2011).

[1] Corresponding author: fpagliar@unina.it

A Revealed Preference (RP) survey was employed from October the 28th to November the 28th 2013 and from January the 5th to February the 5th 2014. The questionnaire was available on the web and accessible through the following websites: www.tripadvisor.it, www.enit.it, www.turistipercaso.it.
In total, 327 complete questionnaires have been collected and analyzed together with a descriptive and econometric analysis.

The questionnaire is made up of three parts:

1. Collection of socio-economic and demographic characteristics of the tourists (e.g. age, gender, marital status, education, employment, residence, income, travel alone/with a group);
2. Collection of information related to the travel (origin, duration, transport mode, willingness to revisit Naples, budget, etc.);
3. Collection of data concerning the use of HSR.

There is a majority of men (58%) w.r.t women (42%). 59% of the sample is Italian, while the remaining 41% are foreigners. Tourists between 21-35 years old are 38% and between 41-60 are 35%. 45% of the tourists are singles while 55% have a partner. 83% travel in group, mainly with friends (27%) or with the partner (22%). The majority has a degree (60%), 38% went to the high school. For a detailed analysis of the socioeconomic characteristics see Table 1.

**Table 1**
*Socioeconomic characteristics of the tourists*

| **About the Tourist** | % |
|---|---|
| Italian | 59% |
| Foregneir | 41% |
| **Gender** | |
| Man | 58% |
| Woman | 42% |
| **Marital status** | |
| Single | 45% |
| Partnered | 55% |
| **Age** | |
| <21 | 6% |
| 21-35 | 38% |
| 36-40 | 11.3% |
| 41-60 | 34.6% |
| >60 | 10.1% |
| **Level of education** | |
| Junior school | 1.8% |
| High school | 37.6% |
| University degree | 60.6% |
| **Professional status** | |
| Employee | 58.7% |
| Unemployed | 8.3 % |
| Freelance | 5.5% |
| Manager | 5.2% |
| Student | 13.1% |
| Retired | 9.2% |

| **Income** | % |
|---|---|
| <500 | 22.6% |
| 500-1500 | 24.2% |
| 1500-2500 | 22.9% |
| 2500-3500 | 19.9% |
| 3500-4500 | 4.9% |
| >4500 | 5.5% |

| **Travelling** | % |
|---|---|
| Alone | 17% |
| With relatives | 34% |
| With friends | 27% |
| With partner | 22% |

| **Accomodation type** | |
|---|---|
| B&B/hostel | 18% |
| Relatives and friends | 17% |
| Hotel | 51% |
| Other | 4% |

| **Trip arrangement** | % |
|---|---|
| Friends/relatives | 3.4% |
| By myself | 78.9% |
| Travel agency | 17.7% |

*Source: Author's elaboration*

The most chosen transport mode has been HSR as reported in Table 2, followed by plane, train, car and bus (see Table 2).

**Table 2**
*Transport mode*

| Transport mode | Total (%) |
|---|---|
| HSR | 41% |
| CAR | 10% |
| PLANE | 31% |
| BUS | 7% |
| TRAIN (NOT HSR) | 11% |

*Source: Author's elaboration*

Concerning the factors influencing the choice of Naples as destination (see Table 3), tourists are mainly attracted by its gastronomy (29.56%) and by its historical heritage (24. 53%).

**Table 3**
*Factors influencing the choice of Naples as a tourist destination*

| Factors infleuncing the choice of naples | Total | Man | Woman |
|---|---|---|---|
| Gastronomy | 29.56% | 51.06 | 48.94 |
| Good quality of promotion | 8.18% | 38.46 | 61.54 |
| Cultural and artistic heritage | 24.53% | 61.54 | 38.46 |
| Shopping and events | 11.95% | 36.84 | 63.16 |
| Less expensive w.r.t other destinations | 8.18% | 61.54 | 38.46 |
| Presence of relatives and friends | 10.06% | 37.50 | 62.50 |
| HSR | 4.40 | 57.14 | 42.86 |
| Other | 3.14% | 20.00 | 80.00 |

*Source: Author's elaboration*

### 2.1. Modelling the probability of visiting and revisiting a destination by HSR for tourism purpose

The results of the survey have been used to estimate some regression models with the aim of computing the factors influencing the probability of revisiting Naples and the role of HSR on this (Model 1) and the probability of using HSR for visiting cities close to Naples (Model 2) for tourism purpose.
Regression models at multiple variables have been considered, such as (see eq. 1):

$$\boxed{\ln\left(\frac{\pi}{1+\pi}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + .... + \beta_k x_k} \qquad (1)$$

with:

$$\boxed{\begin{aligned}\pi &= \text{Probability}(Y = outcome of interest \mid X_1 = x_1, X_2 = x_2, ....., X_k = x_k) \\ &= \left(\frac{e^{\alpha+\beta_1 x_1+\beta_2 x_2+...+\beta_k x_k}}{1+e^{\alpha+\beta_1 x_1+\beta_2 x_2+...+\beta_k x_k}}\right)\end{aligned}} \qquad (2)$$

where $\pi$ is the probability of the outcome Y, $\alpha$ is the intercept parameter, $\beta_k$ is the vector of the s slop parameters and Xs are the set of predictors.

Therefore:

Model 1 (eq. 3):

$$\boldsymbol{Predicted}\ \mathbf{logit}\ (\mathbf{REVISIT = 1}) = \beta_0 + \beta_{1\,NATION} + \beta_{2\,MARRIED} + \beta_{3\,FREELANCE} + \beta_{4\,LESS-EXP} + \beta_{5\,EVENTS} + \beta_{6\,GASTRONOMY} + \beta_{7\,CULT-ART\_HERITAGE} + \beta_{8\,HSR} \quad (3)$$

where

| | |
|---|---|
| NATION | equal to 1 if the tourist is Italian; 0 otherwise. |
| MARRIED | equal to 1 if the tourist is married; 0 otherwise. |
| FREELANCE | equal to 1 if the tourist is a freelance; 0 otherwise. |
| LESS-EXP | equal to 1 if the tourist has chosen to revisit Naples since it is less expensive w.r.t. other destinations; 0 otherwise. |
| EVENTS | equal to 1 if the tourist has chosen to revisit Naples because of the presence of events (sport. music. etc.); 0 otherwise. |
| GASTRONOMY | equal to 1 if the tourist has chosen to revisit Naples because of its gastronomy; 0 otherwise. |
| CULT-ART_HERITAGE | equal to 1 if the tourist has chosen to revisit Naples because of its cultural and artistic heritage ; 0 otherwise. |
| HSR | equal to 1 if the tourist has chosen to revisit Naples because of the presence of HSR; 0 otherwise. |

Model 2 (eq. 4):

$$\boldsymbol{Predicted}\ \mathbf{logit}(\mathbf{HSR\_NEAR\_CITIES = 1}) = \beta_0 + \beta_{1\,NATION} + \beta_{2\,STATION\_ACCESS} + \beta_{3\,TRAVEL-TIME} + \beta_{4\,SERV\_FREQ} + \beta_{5\,SAFETY} + \beta_{6\,COMFORT} \quad (4)$$

where:

| | |
|---|---|
| NATION | equal to 1 if the tourist is Italian; 0 otherwise. |
| STATION_ACCESS | equal to 1 if the tourist has chosen HSR because of the departure/arrival station accessibility; 0 otherwise. |
| TRAVEL-TIME | equal to 1 if the tourist has chosen HSR because of the reduction of travel time; 0 otherwise. |
| SERV_FREQ | equal to 1 if the tourist has chosen HSR because of the service frequency; 0 otherwise. |
| SAFETY | equal to 1 if the tourist has chosen HSR because it is safe; 0 otherwise. |
| COMFORT | equal to 1 if the tourist has chosen HSR because of the comfort; 0 otherwise. |

In the first model the dependent variable is the probability of revisiting Naples (yes: 1; No: 0). Estimation results are reported in Table 6.

The model presents a good ability of reproducing available data ($\rho^2$=0.462; $\rho^2$adj=0.452). All the parameters are significant and of the expected sign. The variable HSR is significant and has a positive impact on the probability of revisiting Naples (Spagnuolo, 2014). The tourist who has the highest probability of revisiting Naples is Italian, not married, not a freelance, he/she chooses Naples because it is less expensive w.r.t other destinations. Naples offers a lot of events in terms of sport, music and shopping, but the reason why tourists want to come back is the cultural and artistic heritage of the city. In the second model the dependent variable is the probability of choosing HSR for visiting cities close to Naples (yes: 1; No: 0). Table 7 reports the estimation results for Model 2.

**Table 6**
*Probability of revisiting Naples for tourism purpose: estimation results*

| VARIABLE | COEFFICIENT (t-test) |
|---|---|
| (Constant) | 0.324 (8.602) |
| NATION | 0.056 (2.588) |
| MARRIED | -0.118 (-2.986) |
| FREELANCE | -0.094 (-2.090) |
| LESS-EXP | 0.851 (23.142) |
| EVENTS (sport, music, shopping, etc.) | 0.880 (16.872) |
| GASTRONOMY | 0.145 (3.879) |
| CULT-ART_HERITAGE | 0.901 (32.716) |
| HSR | 0.107 (2.310) |

**Table 7**
*Probability of visiting cities close to Naples by HSR for tourism purpose: estimation results*

| VARIABLE | COEFFICIENT (t-test) |
|---|---|
| (Constant) | 0.044 (3.455) |
| NATION | 0.035 (2.058) |
| STATION-ACCESS | 0.266 (10.191) |
| TRAVEL-TIME | 0.530 (20.410) |
| SERV_FREQ | 0.315 (7.530) |
| SAFETY | 0.272 (6.327) |
| COMFORT | 0.229 (7.213) |

Also this model has a good ability to reproduce available data ($\rho^2$=0.493; $\rho^2_{adj}$=0.485). All the attributes are significant and of the expected sign. The tourist that has the highest probability of visiting cities close to Naples by HSR (such, Salerno, Rome etc.) is Italian, he/she chooses HSR because of the accessibility to the departure/arrival station, because of the reduction of travel time, because of the frequency of the service. SAFETY and COMFORT are also factors affecting this choice.

### 2.2 Comparison with the Paris and Madrid case studies

The analysis of the probability to revisit Paris and Madrid has provided very different results (Pagliara et al., 2014).
Estimation results are reported in Table 8. In the case of Madrid, the model presents a high explanatory power indicating that the model fits the sample data pretty well. All the parameters are highly significant (except TRANSP_COST>700€ and the HSR variable, which are not significant) even though they have the expected sign. Indeed, the satisfaction of past experience (FIRST_TIME_MADRID) has a positive impact on the probability to revisit Madrid. In fact those people who already visited Madrid have a higher chance of returning. With reference to the socio-economic characteristics, the Spaniards have a higher probability to revisit Madrid for tourism purposes.
Transport characteristics seem not to have a big impact on the destination choice. Indeed, although the transportation cost has the correct sign, it is not significant, which means that it is not an attribute relevant to determine destination choice. Nevertheless, the quality of promotion of heritage resources is important. The main outcome from the Madrid case study is that tourists will revisit the city irrespective on the presence of the HSR.
In the case of Paris, all the attributes are significant and have the expected sign. Tourists that are willing to revisit Paris are younger than the average. They are French, aged between 18 and 24, and were at the university when they were surveyed. They travel with friends, and they would like to go back to Paris because of its architectural sites, the opportunity of visiting other places from there, and the possibility of visiting relatives. Paris is also a city full of events, and this is a factor influencing the choice to come back, particularly for the youth. The HSR variable is very significant and positive, meaning that for the young people the presence of HSR influences their choice. The variable itself embeds all the characteristics connected with HSRS, i.e. high speed, reduction of travel times, high frequency, reliability, easy access to the station, and so on. Moreover young people know that reduced fares are available so they can benefit from this fact to come back.

**Table 8**
*Variables influencing the probability of revisiting Madrid and Paris for tourism purpose*

| Variable | Madrid Coefficient (t-value) | Paris Coefficient (t-value) |
|---|---|---|
| **AGE_18-24** | - | 0.105 (1.979) |
| **MARRIED** | -0,063 (-2,315) | - |
| **FREELANCE** | 0,104 (2,2040) | - |
| **NATION** | 0,121 (3,633) | 0.192 (3.904) |
| **UNIV** | - | 0.238 (5.111) |
| **FIRST_TIME_MADRID** | -0,083 (-2,809) | - |
| **STAY_RELAT_HOME** | 0,111 (2,131) | - |
| **TRAV_FRIENDS** | - | 0.167 (3.063) |
| **HSR** | -0,015 (-0,552)* | 0.177 (4.167) |
| **TRANSP_COST>700€** | -0,028 (-0,954)* | - |
| **VISIT_RELAT** | - | 0.160 (3.416) |
| **ARCHITECT** | -0,559 (20,409) | 0.434 (9.712) |
| **MULTI_DEST** | - | 0.172 (3.677) |
| **EVENT** | - | 0.0902 (2.036) |
| $\rho^2$ | 0,493 | 0.650 |
| $\rho^2$ adj | 0,485 | 0.595 |

* *Not significant*

The second model intends to identify which variables have an impact on the use of HSR to travel from Madrid or Paris towards nearest cities served by HSR. The variables are described below, and the estimation results are reported in Table 9.

**Table 9**
*Variables influencing the probability of visiting cities close to Madrid and Paris by HSR for tourism purpose*

| Variable | Madrid Coefficient (t-value) | Paris Coefficient (t-value) |
|---|---|---|
| **NATION** | -0,140 (-2,249) | 0,108 (2,417) |
| **INCOME_0-2500Euro** | - | 0,084 (1,954) |
| **TOT_HOLID_7days** | - | -0,425 (-7,443) |
| **STAY_PARIS_5days** | - | 0,238 (4,589) |
| **TOT_HOLID_COST** | - | -0,257 (-2,792) |
| **EASY_2NEARCITIES** | 0,296 (5,101) | 0,289 (5,822) |
| **SAFETY** | 0,273 (4,547) | - |
| **SERV_FREQ** | 0,328(4,317) | - |
| **TICKET_COST** | -0,359 (-6,921) | - |
| **COMFORT** | 0,456 (10,658) | - |
| **STATION_ACCESS** | 0,398 (7,419) | - |
| ρ2 | 0.631 | 0.41 |
| $\rho^2$ adj | 0.594 | 0.392 |

In the case of Madrid, all the attributes have the expected sign and are significant: the probability to reach nearby cities by AVE (EASY_2NEARCITIES), the accessibility of departure/arrival station (STATION_ACCESS), travel comfort (COMFORT), service frequency (SERV_FREQ), and safety (SAFETY) have a positive impact on the probability to use HSR service to visit cities located nearby. The cost of transportation (TICKET_COST) has a negative impact. Foreign tourists are using HSR more frequently than national ones to move to cities close to Madrid by HSR. This fact is confirmed by the negative sign of the variable NATION.

In the case of Paris, the average tourists that will likely visit cities close to Paris by HSR are French with income below €2 500 a month, and are going to stay in Paris less than 5 days out of a trip 7 days long. They will choose TGV because of the easy access to two nearby cities connected by it, and their total cost for the holiday is less than 1000 Euros.

Consequently the role of HSR in the probability of visiting other cities is different in Madrid compared to Paris. In Madrid this is for foreigners while in Paris this is for French people.

### 3 Some concluding remarks

Several factors influence the choice of a tourist, like the presence of architectural sites, the quality of promotion of the destination itself, the presence of events, etc. HSR system also plays a role in this choice, but in a different way. In France HSR is considered as a real alternative transport mode. Therefore as the results of the models show, French tourists choose it for moving around France for their holidays. However in the case study of Spain, tourists have a different perception of HSR because the models show that they are willing to revisit Madrid regardless of the presence of HSR. Despite this trend, the analysis suggest that foreigner tourist in Madrid choose HSR for visiting cities close to Madrid. The results for Madrid are similar to the case of Rome. The Naples case study seem to follow that of Paris, since Italian tourists are affected by the presence of HSR in choosing this destination for their holidays as well as for visiting cities close to it.

Further investigation is necessary to understand the specific role of HSR system on tourism in other countries and, inside the same countries, in other cities and especially intermediate cities. However these findings provide useful information for analysts in their efforts to segment and target specific tourist segments and to identify the way by which HSR can impact tourism and for which reasons. A greater awareness of tourists' characteristics w.r.t. a given destination represents an important input for improving packaging and promotion.

### References

Barros, C. P. and Assaf, A. G. 2012. Analyzing Tourism Return Intention to an Urban Destination, *Journal of Hospitality & Tourism Research*, 36: 216-231.

Cascetta, E.; Papola, A.; Pagliara, F. and Marzano, V. 2011. Analysis of mobility impacts of the high speed Rome–Naples rail link using withinday dynamic mode service choice models, *Journal of Transport Geography*, 19: 635-643.

Delaplace, M.; Pagliara, F.; Perrin, J. and Mermet, S. 2014. Can High Speed Rail foster the choice of destination for tourism purpose?, *Procedia - Social and Behavioral Sciences*, 111: 166-175.

Pagliara, F.; Vassallo, J. M. and Román, C. 2012. High speed vs. air transportation: the Madrid Barcelona case study, *Transportation Research Record: Journal of the Transportation Research Board*, 2289: 10-17.

Pagliara, F. and Papa, E. 2011. Urban rail systems investments: an analysis of the impacts on property values and residents'location, *Journal of Transport Geography*, 19: 200-2011.

Pagliara, F.; Delaplace, M and Vassallo, J. M. 2014. High-Speed train and tourists: what is the link? Evidence from the French and the Spanish capitals. In *Proceedings of the Urban Transport Conference*, 28-30 May, Faro, Portugual.

Seddighi, H.R. and Theocharous, A.L. 2002. A model of tourism destination choice: a theoretical and empirical analysis, *Tourism Management*, 23: 475–487.

Spagnuolo, A. 2014. I sistemi ferroviari ad Alta Velocità e il mercato del turismo: Il caso di studio di Napoli, Master degree dissertation. University of Naples Federico II.

Valeri, E., Pagliara, F. and Marcucci, E. 2012. A destination choice model for tourism purpose. In *Proceedings of the ASRDLF 2012 conference special session on High Speed Rail, Tourism and Territories*, 9th-11th July, Belfort, France.

World Travel & Tourism (2013) report

# INNOVATIVE APPROACHES TO THE PROCUREMENT OF PUBLIC TRANSPORT INFRASTRUCTURE: EVIDENCE FROM RECENT CASES

**Haitam Emhemad[1], Brian Agnew[2], David Greenwood[3]**
*[1,2,3]Faculty of Engineering and Environment, Northumbria University, Ellison Place, Newcastle-upon-Tyne, NE1 8ST, United Kingdom.*

**Abstract:** Following the uprising and governmental changes of 2011, Libya's National Transitional Council in August 2012 handed over the country's direction to the newly elected General National Congress. Although faced with the pressing priorities of re-establishing national stability, a major challenge for the government is the provision of a transport infrastructure that will meet the nation's needs (Libyan General Council for Planning, 2013). It is likely that Libya's prospects for national economic prosperity will be seriously undermined due to its poor transport infrastructure. In fact, the inadequacy of the existing road network predates the recent conflict, and recent examination of road construction procurement and management suggests that the country's approach to developing its transport infrastructure have, to-date, been less than successful. As with many developing economies non-technical issues such as management, planning and financing are the major barriers to the efficient and effective provision of transport infrastructure (see, for example, World Bank, 1988, 1994). The research presented in this paper relates to an examination of the means and methods of approaching these problems in Libya. In particular, procedures and practices in the ministries responsible for road infrastructure are examined. An in-depth investigation has been made into the Libyan planning and legislative context, in order to construct a normative framework for the future design of appropriate procurement arrangements. The ultimate aim of the work is to establish a set of criteria that would be at once beneficial for national development and attractive to the infrastructure providers themselves.

**Keywords**: Procurement, Reconstruction, Transport Infrastructure.

## 1. Introduction:

Continuous financial constraints have urged various governments to privatise public infrastructure in the form of Public Private Partnerships (PPPs) (see, for example, Jones, 2002). PPPs use the skills and, in particular, their managerial abilities of the private sector to build and operate public projects efficiently and effectively over their life time (Leiringer, 2003). By the last decade of the 20th century this had become a popular way of managing public facilities in developed countries, (Li et al., 2005) in. Alliancing is another financial model where Public Finance Initiative (PFI) and PPP are applied to manage public projects. Such an example is the Sydney Harbour Tunnel where the contractor worked in close consultation with the Australian government to manage a project that was delivered well on time and in budget (Shepherd, 2003).

Reliable and trustworthy infrastructure is a vehicle for economic growth. An adequate infrastructure plays a pivotal role in supporting industry, the delivery of services, and the movement of people and goods (Akampurira et al., 2008). Globally, between 1984 and 2008, over 1,300 transportation projects valued at more than US $500 billion have been planned and funded through PPPs (Public Works Financing, 2009), although there appears to have been a visible impact of the 2007/2008 global financial crisis (GFC) upon the public private partnership (PPP) markets globally (Burger et al., 2009).
The objectives of the research reported in this paper:

- To investigate the present situation of Libyan road networks,
- To investigate the issues faced by the transport and housing utilities ministries regarding provision of an adequate road infrastructure in Libya,
- To critically examine the way roads are procured and maintained in Libya;
- To examine the feasibility of alternative methods of procurement in the Libyan context.

## 2. Background:

Libya is situated in North of Africa with area 1,759,540 square kilometres; it is the 17th largest nation in terms of land area of which 90% of land is desert. Libya has a long Mediterranean coastline of 1770kms and a population of less than 6 million producing a population density of 3.2 people/km$^2$ (Daw and El-Bouzedi, 2014). The majority of the population live in cities such as Tripoli and Benghazi but the coastline population accounts for only 10% of the total, the remainder living in locations in the interior. The annual growth rate of the Libyan population is 3.5%.

[1] Correspondence Author: Haitam.ismail@northumbria.ac.uk

During the uprising in year 2011, much of the infrastructure was damaged; there is now an urgent need to upgrade the existing infrastructure (General Council for Planning, 2013). The Libyan economy mainly depends upon oil revenues which accounts for 80% of the total, while 20% comes from the construction sector. Libya is very poor in water resources and 75% of its food is imported the distribution of which justifies the development of an extensive road infrastructure.

## 3. Literature Review:

Schroeder, (1994), examined the management, provision and production of rural roads in Developing Countries focusing on issues related to planning and development but ignored technical problems related to engineering and design. He showed that planning and developments issues cause problems to the local population of rural areas which reduces their potential to increase productive activities in Agriculture and industrial products and increase their market share. He particularly argued that key components of successful rural road planning and management were the incentives created by the Government Departmental arrangements governing the sector. Non-technical problems such as management, planning, and finance were leading issues or barriers to the construction of rural roads as recognised by the World Bank (1988, 1994). Blanc-Brude et al. (2009) studied the costs increase in the traditional procurement methods in the road sector. Their results, based upon the European road sector, suggest that to avoid overruns of time and increasing costs a new procurement and management structure should be introduced to the sector. Involving the private sector in road construction does not increase the overall costs beyond the total costs accrued by the traditional contracts method of procurement. Also involving the private sector removed some of the risks associated with the traditional procurement procedures. A new approach involving Public-Private Partnership (PPP) was advocated. The effectiveness of PPP has been examined by Tanczos et al. (2004) who presented the experiences of private provision of road transport infrastructure in Hungary. Firstly, the researchers noted that though there were many changes made to the institutional structure there were still many threats that need to be addressed. The European trends suggest the reforms in providing about maintenance, control and coordination of the public road system that are deep seated and need to go as far as modifying existing regulations and laws. Above all they inferred that the rules regarding development and implementation of road infrastructure were lacking. The findings also indicated that there were no clear indications about reinvestment for development, operation and maintenance.

## 4. Research Methodology:

The road construction program in Libya is controlled centrally from two main ministries, namely the Ministry of Housing and Utilities and the Ministry of Transport based in Tripoli. The Ministry of Housing and Utilities is responsible for inner city roads, and the trunk roads between the cities are constructed by the Ministry of Transport.
The purpose of the interviews was to determine the method used in procuring new roads in Libya in particular with regard to policy and planning and to identify any problems in implementing the government policy.
A telephone interview was arranged with three project managers from each of the ministries. The project managers who are involved in policy, planning and implementation of ministerial objectives are based in Tripoli, but some work from sub-offices of the respective ministries. Each interviewee was reminded to remain isolated during the interview. The interview questions were first designed in English but the interviews were conducted in Arabic, and the researcher noted all the points from each interview. It is important to mention that this study used a mixed approach while designing the questions, (Byrne, 2002, Kvale, 1996 and Hales and O' Connor, 2008). Three of the questions focused on the quantitative approach, while the rest referred to the qualitative methodology (Davie et al. 2007). There are advantages for each research approach, but it depends upon the complexity of the issues (Creswell, 2003, GPC, 2006). However, as pointed out by Anderson, (Anderson, 2009) the mixed approach has advantages and disadvantages which requires the interviewer to focus on the issues discussed by the interviewees while analysing the responses of qualitative data, if not, this might mislead the researchers and academicians.

## 5. Results and discussions

The results of the survey questionnaire are presented. In some of the questions, the results are presented quantitatively, while in the rest the qualitative approach in analysis is adopted.

### Q1: The respondents work experience:
Two thirds of the project managers interviewed had work in a ministry between 5 and 7 years and the remaining 1/3 had less than 5 years experience.

### Q2: Qualifications of the project managers in both the ministries:
All the respondents were well qualified with either diplomas or degrees in an appropriate subject area. These being BSc (civil Engineering) (33%), Master (civil Engineering) (33%) and Diploma (civil Engineering) (33%).

**Q3: The funding of road building and maintenance:**
A simple question about the funding of roads in Libya was asked from the project managers in both the ministries. The managers in the Ministry of Transport responded that the full amount of money is paid by the central government. They all responded well and further held that 60 percent of the total spend is allocated for new roads, 20 percent for the maintenance of existing roads, and the remaining 20% for emergency projects. The project managers of the Ministry of Housing and Utilities answered very differently. Of the total, one said that in 2008, there was no road constructed but in 2009 only one road was built in the south of Libya. He indicated that in the year 2010 only road maintenance was funded and since 2011 every aspect of constructing roads has stopped. The second project manager stated that existing roads are insufficient and work has stopped due to lack of money. The third respondent maintained that the annual road building program changes because of changes in the Government priority.

**Q4: What is the total number of project managers in both the ministries:**
Under the Ministry of Transport, there were 45-50 project managers working in the roads department and there are approximately another 40 working in the Ministry of Housing and Utilities.

**Q5: What is the role of project managers working on roads infrastructure:**
One of the respondents in the Ministry of Transport stated that he is responsible for carrying out research to know about the people's requirement about roads. The second respondent was responsible to plan the ministerial agenda and the third project manager was following the work assigned to him by his manager in the ministry of transport. The interviewees in the Ministry of Housing and Utilities were employed in a similar way, one of them is looking to the future needs of the roads in Libya while the remaining two managers are planning and supervising ongoing road projects.

**Q6: What are the procedures for engaging construction companies:**
Under the Ministry of Transport, all the construction companies employed between 2008 to 2011 were local and of size small. This lead to problems due to low staffing levels within the companies and low skills base of the workers. In response to this the Ministry sought to engage with construction companies from other countries such as Turkey, Germany and China who are expected to work with a local partner.
All three respondents from the Ministry of Housing and Utilities faced problems from the local companies due to non-completion of the projects and low technical knowhow.

**Q7 what proportion of targets in road construction have been achieved in both the ministries:**
Of the total, 67 percent targets were achieved in the Ministry of Transport compared to 33 percent in the Ministry of Housing and Utilities.

**Q8: As project manager how many miles of roads did you plan to construct annually:**
The project managers in the Ministry of Transport stated that the total mileage of roads planned was not known to them. All three did supervise the ongoing maintenance of roads during 2008 – 2011 between Tripoli and the city of Alkoms, including coast roads. One of the project managers did supervise the roads in Tripoli during 2008 – 2011 and the second one was responsible for planning the roads in the city of Benghazi during 2009 only. The third respondent was responsible for supervising the roads in the city of Zawiyah between 2008-2010.

**Q9: As project manager how many miles of planned roads were constructed under your supervision:**
The respondents of the Ministry of Transport responded that one of them planned the roads between two cities in the south of Libya, during 2008-2009. Under his planned program, 100 miles of roads were constructed. The second respondent planed the roads between Sabha and Murzuq around 130 miles. The third planned the roads between Owbari and Murzuq, around 75 miles during 2008. The three respondents from the Ministry of Housing and Utilities planned roads for the Tripoli city only during 2008-2011 to a total capacity of 135 miles during this time.

**Q10: Do you have management procedures that could meet the future needs of the Libyan people, for roads?**
The Ministry of Transport respondents did mention that Libya does have a robust management procurement procedure as it is not delivering the best roads on time. All the respondents in the ministry of housing and utilities agreed that the management structure in Libya is not enough to meet the future needs of the people.

**Q11: Are you aware that due to the 2011 uprising, the previously planned projects are facing many problems, how you are handling such issues?**
All the respondents clearly pointed out that the up-rising of 2011 badly hit the foreign construction companies. The companies' staff went out of Libya, while, machinery remained in Libya. The ongoing contracts were not completed, and the construction companies did face financial crisis. The work that is remaining to be completed or that was abandoned is being currently being distributed to the same company or to others.
Concerning problems, the respondents indicated that the planning done before the up-rising is not considered now because of the new government has different priorities. In addition, there are lengthy procedures in place, and it takes long time to finalise a contract. This is a major problem for both of the ministries. The changes in ministerial portfolio is a serious issues, as the new minister does not necessarily follow the previous program, instead, he introduces his own agenda.

**Q12: Do you face any threats from technical point of view?**
The responses again are gauged from the total respondents, the threats they are facing were identified in the following order:

- Lack of experience,
- Lack of capacity to do the work,
- Lack of financial resources,
- Technological threats.

**Q13: Which procedure do you use to short-list the companies that applied for road construction contracts?**
Both the Ministry respondents again very clearly answered this question. Both Ministries have an independent department that examines the companies who apply for new contracts in road construction. The short listing criterion includes experience, staff qualifications and bank records. This also includes the examining of completed projects. The company, who meets the criteria set out by the manager eventually, wins the contract.

**Q14: Are you satisfied from the local construction companies?**
Unfortunately, all of the respondents were not satisfied with local construction companies as all the important projects in road construction are being managed by foreign companies. This simply suggests that the technical aspects of the Libyan construction are of poor quality, and this led to the respondents being dissatisfied. All the managers also accepted that for large road projects the Libyan government used to engage international construction companies.

**Q15: Implications**
There are lengthy procedures for finalising roads contracts. A project manager has to wait to follow the complicated steps involved in concluding the contracts. It was also noted that Ministries delay paying the construction companies. This can result in stoppages which sometimes results in work being abandoned by the contractor until the problems are solved by the Ministry.

## 6. Discussion:
It is clear from the responses from both ministries contacted for this initial study that there are severe administrative obstacles to the efficient and effective procurement of roads infrastructure in Libya. An example is the lengthy procedures in finalising construction contracts. This and other aspects are directly impacting the efficiency of constructing and maintaining the roads in Libya. From time to time, the government of Libya has encouraged foreign construction companies to tender for the construction of roads. The government also imposed one conditions that foreign companies would require an agreement with at least one Libyan construction company. Companies were invited from Germany, Brazil, Indonesia, and Turkey. The responses suggest that there were delays that increased the costs. The delays were either from government side, (for example, resulting from delayed payment) or failure to achieve contractual 'close' before the project started (which could, in theory, be due to either contractual party). Therefore, this research envisages that the ongoing construction contracting methods adopted for road infrastructure provision may need to be reconsidered. The problem is also amplified by the Ministries' slef-perception of lack of skills in finalising and managing construction contracts, resulting in the disturbing situation that companies who have won bids were not allowed to start work on time. The Libyan government's bureaucratic channels are quite complicated, and this appears to be directly causing problems with the construction companies' abilities to proceed in a timely way. There is even evidence to suggest that some of the foreign companies had left Libya as a result of experiencing such attitudes and behaviours.

Alongside construction, an important feature of the Libyan roads network is its continued maintenance for. Evidence suggests that maintenance of the current roads is extremely poor. The roads that were destroyed during the revolution of 2011 are still waiting to be repaired for public use. Libya imports foods especially from the neighbouring countries such as Tunisia and Egypt. The unreliable network of roads is not responding to the market demands. The food items remained short in large cities, and this is negatively impacting the economic activities in the whole of Libya.

This research suggests that the current infrastructure is not reliable, and cannot respond to the people's needs in days to come. Therefore, the present government should re-consider the previous five year plan for construction and upgrading the infrastructure as a first priority. The results of this research can be linked to the current staff and their capabilities. It is also noted that the government has, in reality, neglected the training of employees, particularly managers. This is the second decade of the 21st century, technology has grown at the fastest rate, and this is totally absent in the Libyan roads network.

Last but not the least; it is considered that the Libyan government should investigate new ways of procuring its projects. This may include (following the PFI/PPP trend in other countries) privatisation of the road building program, perhaps with related assistance for local construction companies to increase their resources, especially in terms of skilled employees. This aspect has been completely neglected previously, and this is seen in the state of the roads of Libya. The opportunities of trade between the cities are directly related to the suitability of the roads between them. For example, farm produce often does not reach the market in time and the prices even of the local produce are not affordable.

There are many models where by the Libyan government can engage with the private sector (in a PPP for instance) and induce large companies from the global construction, or even oil sectors to enter into novel arrangements for the construction and maintenance of roads, perhaps in exchange for a concession to establish profitable facilities, such as petrol stations, shops, and rest centres. The exact nature of these agreements is difficult to predict at the moment but the opportunity exists to produce a model for partnerships that will be beneficial to the Libyan people and the Libyan economy in general, whilst being more attractive to the supply-side than the previous approach.

## 7. Conclusions:

From above it is concluded that Libyan roads infrastructure is the victim of lengthy procedures. The construction companies have faced many problems because of late payments for the work. It is also inferred that the construction companies who procure the roads could not finish the work due to lack of skills and others. The ministries responsible for constructing roads are using traditional ways in procuring the roads. It can also be concluded that Libyan government is the authority, who allocates the resources. The results indicate that due to poor infrastructure, the economic activities are undermined, and the prices are sky high, and people are suffering.

Therefore, this research recommends strongly that Libyan government should undertake legislation to pass the laws for alternate ways for constructing roads. The private sector involvement is very low, and government needs to encourage the private investors to make investments in strengthening the infrastructure.

**References:**

Akampurira, E., Root, D. and Shakantu, S. (2008), Factors Constraining the Implementation of Public Private Partnerships in the Electricity Sector in Uganda, available at: www.cib2007.com/papers/CIDB2008%20Final%20Paper%20No27.pdf (accessed 2Feb, 2014).

Anderson, V. (2009) Research methods in human resource management. 2nd ed. London: Chartered Institute of Personnel and Development.

Blanc-Brude, F. Hugh Goldsmith, H. and Timo Välilä, T. (2006) Ex Ante Construction Costs In The European Road Sector: Comparison Of Public –Private Partnerships and Traditional Public Procurement , Economic and Financial Report, European Investment Bank.

Burger, P. Tyson, J. Karpowicz, I. and Coelho, M. (2009) “The effect of the financial crisis on public private partnerships”, IMF Working Paper WP/09/144, July.

Byrne, D. (2002) Interpreting Quantitative Data. London: Sage

Davies, M.B. (2007) Doing a Successful Research Project ,Using qualitative or quantitative methods, Palgrave

Daw, M. A. and El-Bouzedi, A. (2014) Prevalence of hepatitis B and hepatitis C infection in Libya: results from a national population based survey, BMC Infectious Diseases, Avalable at: http://www.biomedcentral.com/content/pdf/1471-2334-14-17.pdf , Accessed on 03/02/2014

General Planning Council of Libya, (2006) National Economic Strategy, Cambridge Energy Research Associates (CERA) Harvard University

General Planning Council of Libya, (2013) Report Ministry Activities, Ministry of Transport

Hales, J. and O’Connor, W. (2008) Methodological Review of Research with Large Businesses (tele)

Jones, D. (2002) “Policy development in Australia for public private partnerships – what more is there to do?”, paper presented at a Seminar on Providing Value for Money Through Public Private Partnerships: The Lessons Learnt So Far from Economic and Social Infrastructure Projects, 26 June, available at: www.claytonutz.com/downloads/PPP%20paper%2026% 20June%202002.pdf (accessed 2 Feb,2014

Kvale, S. (1996) An Introduction to Qualitative Research Interviewing

Leiringer, R. (2003) “Technological innovations in the context of public-private partnership projects”, unpublished PhD thesis, Royal Institute of Technology, Stockholm.

Li, B. Akintoye, A. Edwards, P.J. and Hardcastle, C. (2005) “Critical success factors for PPP/PFI projects in the UK construction industry”, Construction Management and Economics, Vol. 23, pp. 459-71.

Public Works Financing (2009) PWF international major project survey. CD-ROM. Westfield, NJ.

Schroeder, L. (1994) Provision and Production of Rural Roads in Developing Countries: A Synthesis of Research Findings. Decentralization: Finance and Management Project. Burlington, VT: Associates in Rural Development.

Shepherd (2003) Project alliancing on a public/private partnership project. Building Australia; June 1999. p. 14–7.

Tanczos, K. Kosztyo, A. Meszaros, F. Bokor, Z. ( 2004 ) Experiences of private provision of road transport infrastructure in Hungary 4th Conference on Applied Infrastructure Research "Regulation and Organizational Models in Infrastructure Sectors – Vertical Dis-/ Integration, Privatization, Concessions and PPP" at Berlin University of Technology,

World Bank (1988). Road deterioration in developing countries: Causes and remedies. A World Bank Policy Study 13370. Retrieved from http:www.ipcb.org/index.php/component/docman/doc_view/741-1988-world-bank-Road/ (Accessed 6 June 2013)

World Bank (1989). Strategic planning and management: A review of recent experience. World Bank Staff Working Papers Number 751.

Retrieved from http://www-wds.worldbank.org/servlet/WDSContentServer/WDSP/IB/1985/09/01/000178830_98101903432616/Rendered/PDF/multi_page.pdf (Accessed 6 June 2013)

# AN ANALYSIS OF THE ECONOMIC COSTS AND BENEFITS OF PUBLIC AND PRIVATE INVESTMENTS ON CYCLING. REVIEW AND CRITICAL ASSESSMENT

**Avgi Vassi[1], Andriani Gogola[2], Efthimmios Bakogiannis[3], Thanos Vlastos,[4]**
[1,2,3,4]*Sustainable Mobility Unit, National Technical University of Athens*

**Abstract:** This paper is trying to identify all the components of economy that affect and are affected by cycling. Their impact on the economy is shown through multiple cases. It also searches the investments on cycling projects as an attempt to document their economic effects and reveal their importance and key lessons learned through this procedure.

**Keywords**: cycling, investments, public investments, private investments, economy

## 1. Introduction

Modern urban agglomerations face several problems that affect quality of life. The growing increase of motor vehicle use is burdening cities with increasing problems and costs related to congestion, accidents, loss of amenity and space, noise, poor quality of life, poor accessibility, pollution, poor urban air quality and energy consumption, having adverse effects on both the natural and built environment. Cycling friendly mobility management schemes are at the center of public debate in many countries. Cycling is slowly becoming a very popular method for travelling and commuting. However, uptake is uneven among cities and countries. This variation in uptake is a reflection of diverging preferences and priorities among both public administrations and citizens. Another important factor is the differences in the availability of necessary infrastructures. These infrastructures include for example the necessary cycling routes, end-of-trip facilities, and connections to other forms of public transportation. In order to achieve the desired modal split, policy makers have focused on the appropriate policies and measures that will promote cycling friendly management schemes. In this effort both the public and private sector can have a significant contribution. Investments in cycling can have a significant impact, by creating favorable conditions and additional incentives towards the uptake of cycling by the population.

## 2. Identification and Impact

The first step of this concerns the identification of all the components of economy that are affected by cycling. The assessment of policy decisions and projects' effects requires identification of: (a) thematic areas of influence, (b) parameters per thematic area and (c) indicators per parameter or thematic area. The identification of the above mentioned elements requires a thorough research in: European Union official policy documents disseminated through official websites, reports and projects written with EU funding through its various operational programmes, academic research papers, reports written by public authorities in their effort to design and implement cycling friendly policies and/ or investments, reports related to cycling projects and/ or investments on cycling written by experts and/ or recognized and reputable authorities, positions expressed by interest groups such as cyclists' federations and cycling industry representatives and data bases and/ or research projects which include quantitative data relevant to the methodology objectives.

*Effects of cycling*

| Area | Parameter | Indicators |
|---|---|---|
| Environment | Energy use | Fuel consumption |
| | Greenhouse gases / Air quality | Emission factors (CO2, CH4, N2O, NOx, NMHC, CO, PM2,5, PM100) |
| | Quality of urban spaces | Traffic noise |
| | | Land consumption |
| | | Attractiveness |
| Transportation | Time costs | Change in transport time by transportation mode |
| | Congestion | Change in motor traffic |
| | Vehicle operating costs | Change in vehicle kilometer by transportation mode |
| | Transit synergies | Change in intermodal transportation |
| | Infrastructure & Maintenance costs | Cost per km of transportation network constructed / maintained |
| Health | Physical fitness | Mortality<br>Morbidity<br>Absenteeism / Work performance / Productivity |
| | Air pollution | |
| | Noise pollution | |
| Safety | Traffic accidents per transportation mode and distance | Minor accidents (<24 hrs hospital) |
| | | Major accidents (>24 hrs hospital) |
| | | Deaths |
| Cycling Tourism | Cycling holidays and day trips by cycle | Number of cycling tourists |
| | | Days of cycling holidays |
| | | Distance cycled per day |
| | | Employment |
| | | Market (local businesses) |
| | | Cyclists' satisfaction |
| Sports & Leisure | Cycling events | Number of cycling athletes, guests, visitors, etc. |
| | | Number of leisure cyclists, participants, etc. |
| | | Employment |
| | | Market (local businesses) |
| | | Cyclists' satisfaction |
| Cycling Industry | Design, manufacture, retail, rental and maintenance of bicycles*, components and equipment, Bicycle Sharing Systems, bicycle parking facilities, etc. | Employment<br>Market |

The step that followed concerned the quantification of indicators. Unit price for indicators was searched in studies, such as feasibility studies, cost-benefit analysis, technical studies, etc., involving cycling projects.
Finally, the key lessons learned from all the review of public and private investments are stated, concerning the need for cities to attract investors on cycling issues.

### 2.1 Environment

Concerning the energy use in the EU, transport depends on oil and oil products for more than 96% of its energy needs as shown in "Energy and Transport in Europe – Statistical Pocketbook", 2010. Also, Europe imports around 84% of its crude oil from abroad. In 2010, the EU's oil import bill was around € 210 billion (EUROSTAT). Cars are the most popular passenger mode across the EU as they represent the 72% of all passenger kilometers. However, the private car, despite its popularity, is hardly ever the most energy-efficient form of transport.
It is generally acceptable that motorised transport causes air and noise pollution which affects health, urban quality of life and enhance the urban economy. Indicative to that assumption are the results from an Amsterdam study pointed out that replacing 30 million car km per year could result in savings of approximately 2.8 million euros as a result of less pollution (I-ce, 2000).
Motorised transport emits $CO_2$, causing the greenhouse effect and with huge cost to the environment. Transport greenhouse gas emissions, including the international aviation and maritime transport, increased by around 34% between 1990 and 2008. Over the same period, energy industries reduced their emissions by about 9% (European Environment Agency). Transport is responsible for about a quarter of the EU's greenhouse gas emissions. This percentage is divided: 12.8% of overall emissions are generated by aviation, 13.5% by maritime transport, 0.7% by rail, 1.8% by inland navigation and 71.3% by road transport (European Environment Agency, 2008).

From the facts above it is pointed out that traffic (from motor vehicles), instead of serving cities, is consuming them. The cost of transport is considered extremely high given the uncertainty surrounding the possibility of climate change and impacts upon biodiversity (Whitelegg, 1993).
Concerning the quality of urban areas, motorised transport is considered responsible for the continuous deterioration of the urban environment (European Commission, 1999). Cycling as an alternative to car use, improves urban quality of life with regards to factors such as pollution and safety (EC, 1999). The traffic noise is one of the biggest environmental problems which are immediately related to urban life. It is more severe and widespread than ever before, and it will continue to increase in magnitude and severity because of population growth, urbanization, and the associated growth in the use of mobile sources of noise. There is growing evidence that noise pollution is not merely an annoyance; like other forms of pollution but it has wide-ranging adverse health, social, and economic effects. The motorised transportation infrastructure (such as roads and car parking) takes up highly valuable city center land and threatens existing open spaces. Reduction in car use and shift to cycling enables better management of the limited and therefore expensive urban space and is also beneficiary for the society: "When walking and bicycling are a natural part of the daily pattern of activity, there is positive spin-off for the life quality and well-being of the individual- and even greater benefits to society" (Gehl, 2010). Bicycle passenger transport is 5 times more space-efficient than car traffic (ECF, 1993). Space savings increase the attractiveness of town centers and furthermore, regained space can be used for productive means (EC, 1999). As an alternative to the car, cycling can increase road capacity which in turn affords time and space savings and improved accessibility, all of which contribute to the urban economy (CTC, 1993).

### 2.2. Transportation

A major aspect of externalities is related to the transportation needs of urban populations. As discussed in the introduction, current urban agglomerations are increasing in every possible index; this includes population size, geographic extent, population density etc. This results in an ever increasing number of automobiles using the existing road infrastructures. Increases in motor vehicle traffic and private motoring have resulted in the steady decline of travel by public transport, bicycle, motorcycle and walking (CTC, 1993).This approach is not sustainable, especially in developed cities, where the availability of road network is given and there is not an option for further expansion, at least not in the commercial city center which is the destination of the majority of commuters. The cost of the lost time using cars is very important as it is measured that in London, 20% of commuters spend more than two hours a day travelling to and from work, which means an add of one working day a week. In Germany, 37% spend one hour a day commuting. In London, Cologne, Amsterdam and Brussels, drivers spend more than 50 hours a year in road traffic jams. In Utrecht, Manchester and Paris, they spend more than 70 hours stuck on roads. (INRIX European National Traffic Scorecard, 2010). Eventually numbers of working hours are lost due to congestion. The bicycle can often produce faster mean journey times than alternatives for trips up to 4-5 km long (Bracher, 1988 cited in Whitelegg, 1993). In addition there is an economic benefit when time savings increase productive time. Incurring the external costs of transport, the chief components of total external costs are accidents, air pollution and climate change. Congestion is less considerable overall, but is the most significant externality in urban areas. Congestion is highly economically inefficient. Average congestion costs in Europe are estimated at approximately 2% of GDP – 120 billion e (UITP, 2003). Replacing car transport with cycling could reduce congestion, improve traffic fluidity and reduce costs of congestion and the number of working hours lost due to traffic jams (EC, 1999).
Road transport is the main contributor to total external costs of transport, accounting for 92% (EEA, 2003). Furthermore, car use accounts for 58% of total costs (EEA, 2001).In addition to all the other costs 13.2% of every household's budget is spent on average on transport goods and services.
It is important to view cycling and public transport not as competitors but cycling should be treated as a complementary 'environmentally friendly' transport mode to P.T. (Holladay, 2002). Cyclists are potentially new customers for P.T., by increasing transit rider-ship through improved accessibility and bicycle parking at transit stations and on transit vehicles. The positive experience of Switzerland showed that a change of 10-15% in the modal share can take place, if cycling and PT are considered allies. There are multiple benefits that derive from this alliance: forms a complete transport chain, enlarges PT's service portfolio, extends PT's catchment area, makes cycling more flexible, and refreshes PT's image as it meets the needs of new customers,. Holladay (2002) states that intermodality can offer useful combined packages which offer a viable alternative to car use, providing convenience and door-to-door travel. On long journeys particularly, a combination of cycling and public transport is a viable alternative. In the Netherlands for example, almost 30% of rail passenger's cycle to a train station and 12% continue their journey by bicycle upon alighting from the train. (I-ce, 2000). Providing bicycle racks on buses has been successful in the U.S., with many transit agencies reporting positive returns on investment and increased ridership.
Regarding infrastructure and maintenance, the construction of cycling infrastructure - even if built to the highest standard - is relatively inexpensive compared to other transport infrastructure: an average $1.5 million per km to plan and construct a separated bicycle path, in the Queensland context for example (Australian Bicycle Council, 2012):
1km of Rail costs the equivalent of 29 kms of bikeway
1km of Motorway/Road costs the equivalent of 110 kms of bikeway
1km of Busway costs the equivalent of 138 kms of bikeway
1km of Road Tunnel costs the equivalent of 324 kms of bikeway
1km of Underground Rail costs the equivalent of 533 kms of bikeway

### 2.3 Health

The cycling's' relation to health is direct. Shayler *et al.*, 1993; I-ce, 2000; BMA, 1992, all conclude that a modal shift from motorised transport to cycling could have significant public health benefits. Increased cycling has the potential to directly improve the health of the individual, in terms of fitness, reduced risk of certain diseases, self-esteem, longevity and quality of life, whilst also indirectly improving the health of society as a whole by reducing atmospheric and noise pollution and reducing road danger.
The economic growth and technological advances have increased the per capita income and decreased the cost of manufactured goods, among them automobiles. Currently in advanced economies the majority of households have access to at least one private vehicle; this results in automobiles being used in every aspect of everyday life (commuting, leisure, etc), further reducing physical activity In OECD 2003 report it is pointed out that "obesity rates have increased in all OECD countries over the past two decades due to poor eating habits and lack of physical activity…more than 20% of people in Mexico, the United Kingdom and Australia are now considered obese" (OECD, 2003). Even households that do not own a private vehicle (either by choice or due to insufficient income) have easy access to the extensive networks of public transportation. It becomes apparent that in modern societies, physical exercise to a great degree has been reduced, both in the work place and in the free time. Both forces in conjunction with the increased wealth and the resulting access to relatively cheap food, have resulted in an increasing portion of the population suffering illnesses related to a sedentary lifestyle such as diabetes, obesity and cardiovascular diseases. The physical inactivity is one of the leading risk factors for health and is estimated to attribute to one million deaths per year in Europe .The aforementioned potential health problems can lead to a significant deterioration of quality of life and a reduction in life expectancy. Furthermore they increase the health costs for the public (and private) health systems, putting a strain on their sustainability. If one takes also into account the changing demographics and the gradual increase in the average age of the population and the prolonged life expectancy, these problems are intensified.
Cycling can be part of the solution to these problems. Frequent use of the bicycle for commuting as well as leisure activities is a very good way to have regular physical activity. This reduces symptoms of a sedentary lifestyle, increases fitness and improves overall health. Therefore it is not a surprise that various stakeholders are recommending the adoption of cycling for everyday activities. The gains for society come in form of reduced healthcare costs, which can mitigate most of the investment costs if a significant modal shift is achieved. Estimates of the direct (health care) and indirect costs of physical inactivity (loss of economic output due to illness, disease-related work disabilities or premature death) are alarming. In the EU, physical inactivity can be estimated to cost about 150-300 euro per citizen per year (ECF, 2013).
Cycling is an efficient and cost effective way of securing an improvement in both public and individual health and reducing health costs (CTC, 1993). Cycling improves societal quality of life and individual psychological well-being, self-esteem and fitness. This could lead to reduced spending on medical treatment, reduced working hours lost due to sickness and reduced spending on workplace absenteeism and increased workplace productivity (CTC, 1993). A study of Amsterdam indicated that a 9% increase in km cycling would create savings from absenteeism and medical treatment of €31.7 million (I-ce, 2000). The International Road Assessment Programme (iRAP), an engineer-led road-safety charity, calculates that road casualties cost 2% of GDP for high-income countries and 5% of GDP for middle- and low-income countries, including medical bills, care, lost output and vehicle damage—$1.9 trillion a year globally. Replacing car kilometers with cycle km generally leads to reduced traffic casualties (CTC, 1993).
Air pollution has severe direct effects on human health. Every year millions of people fall ill because of bad air quality, resulting in reduced quality of life and huge economic losses.
Traffic noise is a major cause not only of hearing loss, cognitive impairment in children, sleep disturbance or annoyance, but also of cardiovascular disease, 1.8% of heart attacks in high income European countries being attributed to traffic noise levels higher than 60dB(WHO, 2011).

### 2.4 Safety

The Royal Society for the Prevention of Accidents, states that every year in UK around 19,000 cyclists are killed or injured in reported road accidents, including around 3,000 who are seriously injured. The numbers only include cyclists killed or injured in road accidents that were reported to the police, as it is known that many cyclist casualties are not reported, even when the cyclist is inured badly enough to be taken to hospital. TRL, in its 2009 report states that around 75% of fatal or serious cyclist accidents occur in urban areas.
The numbers outline a significant societal cost related to cycling. Accidents result in economic cost directly, through the expenses necessary for treating accident injuries. It also creates a negative externality through the cost of lost production and income due to the inability to work. The construction of high quality infrastructures and the adoption of necessary traffic management policies can reduce accidents. This would have a dual positive effect: A reduction in accidents reduces also the accident related costs to society, but most importantly it will improve the attitude towards cycling. People that were discouraged by the increased danger of using a bicycle would be more probable to use it for commuting purposes. This trend would be accompanied with all other positive externalities related to cycling adoption.

A final aspect that has to be mentioned is the fact that increased cycling usage can by itself result in improvements in safety for cyclists. Increased presence of cyclists on roads makes drivers more accustomed to them and can result in the modification of their driving behavior. Being aware that cyclists also occupy the road makes them more careful and alert with regard to them and improves safety for cyclists.
The European research project PROMISING shows that cycling can provide external benefits as well as incurring external costs, for example through savings in public health care due to improved fitness. The external injury accident cost for a car in a town center in Norway per km is estimated at €0,03. The mean external cost of bicycle accidents is estimated to be greater: €0,08 per km of cycling in Norway. Nonetheless the savings on health, taking into account reduction in short term illness and reduced absence from work, have been shown to represent an external benefit of more than €0,6 per km of cycling.

### 2.5 Cycling Tourism- Sports Events- Leisure

The existence of cycling routes can be attractive to a specialized and growing segment of tourism activity. Cycling infrastructure would increase cycling as a sport and leisure activity. This can become a major benefit for the urban environment as a tourist would cycle in order to visit the tourist attractions, but this works in the other way round too as (2) the bicycle itself can be a reason to attract tourists in areas that are not traditional touristic destinations and therefore would increase direct spending in local economies. This would support local businesses, maintain and create new jobs and increase local income. The organization of events could also become an attraction to a large number of cyclists. The economic benefits to the local economy are very important. Furthermore systematic usage of the bicycle for leisure and sports would provide a steady clientele to businesses of the local economy. Usually it is taken into account the average amounts these tourists spend on food, lodging and other goods and services while visiting in order to calculate a figure that represents the total economic value of the visits.
In full compliance, Weigand (2008) argues that the economic impact of bike-related tourism is usually calculated by estimating or counting the number of visitors (and local residents) who are participating in self-guided and organized tours, rides and events. Bike events, such as races and tours, are a subset of bicycle-related tourism that attract visitors to an area, either to participate or to watch. The economic impacts of these activities are easier to quantify because many of the factors are known, such as the number of days in the region, the number of participants, and their origin). In the same study, Weigand states that Maine Department of Transportation estimates that bicycle tourists spend between $25 (day trips) to $115 per day (guided tours) on a combination of retail, services, lodging, food and transportation.
The European Cycle Route Network Eurovelo, Challenges and Opportunities for Sustainable Tourism, Eurovelo, Challenges and Opportunities for Sustainable Tourism, ECF 2012 is a study commissioned by the European Parliament in 2012 estimated that there are over 2.2 billion cycle tourism trips and 20 million over-night cycle trips made every year in Europe. These have an estimated economic impact of €44 billion. A very important example of regional economic impact is North Carolina's Outer Banks, as by a conservative estimate, it generates $60 million in economic activity through bicycle tourism. The one-time investment of $6.7 million on bicycle infrastructure has resulted in an annual nine-to-one return as revealed by Flusche (2009). Concerning the investments in cycling events, a very good example is the San Francisco Grand Prix, which is a one-day race through the city with about 1.3 million participants. This specific race is estimated that returned approximately $3million in economic benefits to the city's $750,000 investment for one day of racing. Another bicycle racing events which can demonstrates economic benefits as well, the Tour de Georgia, a seven-day pro state race generated a direct economic impact of $27.56 million which started in 2007 (Hong, 2007). In 5 years, the race has attracted 2.8 million spectators and resulted in $148 million in economic benefit to the state (Weigand, 2008). The "Registers Annual Great Bicycle Ride Across Iowa" is an annual seven-day bicycle ride across Iowa, that dates back more than 40 years. In 2008, the total direct spending was quantified and (dollars paid for services by travel parties for the event) while in Iowa the figure was estimated at $16,908,642, the total direct, indirect and induced spending for the race was somewhere in between $24.5 and $25.7 million as revealed by Lankford et al. (2008).

### 2.6 Cycling Industry- Businesses

A very big chapter in the cycling economics is the cycling industry, meaning the: design, manufacture, retail, rental and maintenance of bicycles, components and equipment, bicycle sharing systems, bicycle parking facilities, etc. In proximity to large urban centers with significant number of active cyclists businesses could appear that provide higher added-value services and commodities. A good example would be bicycle manufacturers. This could provide an important boost to local economy and increase employment. Furthermore these businesses could expand operations and become a hub for manufacturing, wholesale/distribution, retail and services related to bicycling. In such a case they would become a significant factor in local economic activity, contribute to the increase of local GDP and help maintain and increase employment.
National Bicycle Dealers Association revealed that in 2008, bicycling manufacturing is a $6 billion industry in the US including the retail value of bicycles, related parts, and accessories through all channels of distribution.

The economic benefits associated with riding extend far beyond that number:

- US recreational cyclists spend $46.9 billion on meals, transportation, lodging, gifts and entertainment.
- the spillover effects of all bicycling-related activities could be as large as $133 billion, supporting 1.1 million jobs and generating $17.7 billion in federal, state, and local taxes (Outdoor Industry Foundation, 2006).
- 16,000 bicycle retail shops, employ more than 70,000 people (League of American Bicyclists, 2012).

Three states and one city through their in-depth analyses of the economic impact show from the primary finding of these reports that bicycling and related industry, sales and service activities are a significant economic force and provide a strong source of both direct and indirect revenue and jobs. Indicatively, in the state of Colorado a survey of bicycle manufacturers, retails shops, ski resorts and other bicycle related organizations and Colorado households has been conducted, and has assessed the impact of bicycling in the form of expenditures, production, employment, income and tax revenue. The total annual sales and service revenue is $200 million. Wisconsin has one of the largest bicycle industries in the USA. The total impact of manufacturing, wholesale/distribution, retail, and services related to bicycling is over $556,000,000. Over 3400 jobs are attributed to these industry types. (Colorado Department of Transportation, 2000, Wisconsin Department of Tourism, 2000, Department of Transportation, 2001, Alta Planning and Design, 2006, cited in Weigand, 2008). In 2010, the London School of Economics calculated that in bicycle manufacturing, cycle and accessory retail and employment the benefits were the following

The previous section explored the impact of cycling on local economies through tourism, leisure and sports. This impact can be invigorating to local businesses of many kinds. Beneficiaries can be the stores, the coffee shops and in general any enterprise providing services to cyclists. However this is not the only way the increased adoption of cycling can increase economic activity and income generation. Regarding the relationship between shop turnovers and cyclists, it is well known that it is not based on accessibility by car as proven by (I-ce, 2000). Cyclists shop more frequently and spent more and may be more prone to 'impulse purchases' than motorists (Bicycle Federation of Wisconsin, I-ce, 2000). The connection between bike infrastructures and revenues can be found in Drennen (2009) where it is stated that two-thirds of merchants along San Francisco's Valencia Street, said that new bike lanes had a positive overall impact on their business. A 2009 study of Bloor Street, a commercial street in Toronto, Ontario showed that people who had biked and walked to the area reported that they would spent more money in the area per month than those who drove there.

## 3. Public and Private Investments

Investments for cycling both from private investors and from citizens are necessary. Collaboration between investors and society is a prerequisite to collect enough funding in order that cycling replace as much as possible car trips because the problems that city faces are many and urgent. While the scope of the public sector is not the direct gains for the private sector the direct gains is what mainly is interested for. Public investment for cycling is a way for the private sector to be funded by citizens. Public Investments are a precondition for Private Investments. In the table below, there are all the possible investment for cycling, along with the type of investment (private, public or an alliance of private and public).

*Public, private investments in cycling*

| Investment | Type of investment | Examples |
|---|---|---|
| Shared or separated paths | alliance | "Velocity" 2025 (Manchester UK), Crowfunded Bicycle Lane, Memphis USA |
| Way finding signage, information signage, traffic signals, etc | alliance | |
| Policies regarding traffic management | alliance | |
| Bike-parking and end of trip facilities | alliance | Bicycle End-of-trip Facilites, a guide to Canadian Municipalities |
| Integration of bicycling with public transport | public | Integrated transport system Malmo, Copenhagen |
| Bike sharing | alliance | Barclays Cycle Hire, |
| Education and Promotion | public or private | Promoting cycle for Copenhagen (public), London Cycling Campaign (private) |
| Health insurances' investments | private | Quality Bike Products (QBP) |
| Tourism | alliance | Cycle Tourism (ECF project) |
| Industry | private | Cycling Industry Club, an European alliance of bicycle manufacturers |

*Source: Analysis of environmental aspects, enrichment of individual reports, development of synthesis report (CycleCities project, 2014)*

Regarding the public investment, it was detected through the research that it is really important to the citizens to keep their needs in mind when investing in a specific infrastructure. Concerning the private investment, it is crucial that the public sector will give the right example to the private in order to join their efforts towards the enhancement of cycling culture. It is also very important to use commercial interests to finance cycling projects, as the importance of cyclists as a customer group is very high.

There is a variety of sources that could either facilitate or impede investments. The most prominent among them is the official public policy. Specifically the attitude of public administrations towards cycling, as expressed by the investment and other policy decisions, can have a significant impact (positive or negative) on private attitudes towards using the bicycle. Although public policies are the most common factor, other social or technical/technological factors can also play an important role. For example public investments are more widespread in areas where the local society has a positive attitude towards cycling, giving additional incentive to elected representatives in the pursuit of cycling friendly policies.

## 4. Conclusion

In recent decades there has been an extensive effort to divert people from private cars initially to public transportation and more recently to cycling and walking. Cycling friendly mobility management schemes are at the center of public debate in many countries. In order to achieve the desired modal split, policy makers have focused on the appropriate policies and measures that will promote cycling friendly management schemes. In this effort both the public and private sector can have a significant contribution. Investments in cycling can have a significant impact, by creating favorable conditions and additional incentives towards the uptake of cycling by the population. The uptake of cycling is beneficial in many domains of life: economy, business, health, society…

The uptake of cycling improves the urban environment in terms of pollution but also landscape. Cyclists tend to spend more than car drivers, as they tend to be healthier. Cycling employees (but also students) are more productive. Cycling infrastructures and environmental effects upgrade the quality of urban environment and living as Gehl J. in “Cities for people” states “Planning good cities for bicyclists is handled … in direct relation to discussion on the human dimension in city planning”. Cycling friendly cities are people-friendly cities, welcoming tourists and visitors.

Through a thorough search in the investments that concern all the area of influence of cycling, it is safe to say that public and private domain should cooperate in order to achieve the maximum results. A very good example is Bike Sharing Systems. A private BSS in order to succeed, it is essential to be deployed in a city with some cycle routes, bicycle crossings, etc. But alliances between public and private sector occur more often lately as it is well understood that

“Public investments are a precondition for private investments”

**References**


Australian Bicycle Council. 2012.
< http://www.bicyclecouncil.com.au/files/about/ABC-Communique-2012-Aug.pdf>

Bicycle federation of Wisconsin. The economic impact of Bicycling in Wisconsin.
<http://atfiles.org/files/pdf/Economic-Impact-Bicycling-WI.pdf>

Colorado Department of Transportation. 2000. Bicycling and Walking in Colorado:
< http://www.coloradodot.info/programs/bikeped/building-a-bike-ped-friendly-community/bike-walk-study/technical-report/techstudy1.pdf>

De Hartog., 2010. Do the Health Benefits of Cycling Outweigh the Risks?: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2920084/>

Drennen, E. 2006. Economic Effects of Traffic Calming on Urban Small Businesses
<http://www.bikewalk.org/2004conference/sessions/28_Business_calm/TrafficCalming_summary.pdf >

Energy and Transport in Europe. 2010. Statistical Pocketbook 2010: http://bookshop.europa.eu/en/eu-energy-and-transport-in-figures-pbKOAB10001/

European Commission. 1999, Reclaiming city streets for people. Chaos or quality of life?:
< http://ec.europa.eu/environment/pubs/pdf/streets_people.pdf>

European Cyclist's Federation (ECF). 1993. Cycling in Urban Areas:
< http://www.ecf.com/publications/Download/urbangb.doc>

European Cyclist's Federation (ECF). 2009. The European Cycle Route Network Eurovelo:
< http://ec.europa.eu/enterprise/sectors/tourism/iron-curtain-trail/files/ep_studyeurovelo_en.pdf>

European Environment Agency (EEA). 2003. Indicator fact sheet: TERM 2002 25 EU- External costs of transport:
<http://themes.eea.eu.int/Sectors_and_activities/transport/indicators/cost/external_costs/TERM_2002_25_EU_External_costs_of_transport.pdf>

European Environment Agency (EEA). 2001. Indicator:External costs of transport:
< http://themes.eea.eu.int/Sectors_and_activities/transport/indicators/cost.external_costs/index>

European Environment Agency. 2008. Energy and environment report:
<http://www.eea.europa.eu/publications/eea_report_2008_6>

Flusche, D. 2009. The Economic Benefits of Bicycle Infrastructure Investments:
<http://www.advocacyadvance.org/site_images/content/Final_Econ_Update%28small%29.pdf>

Holladay, D. (2002). Cycling with public transport: combined in partnership, not conflict. In H. McClintock (Ed.) Planning for Cycling: Principles, Practice and Solutions for Urban Planners (pp.110-142). Cambridge, England: Woodhead Publishing Limited.

Hong, S. 2007. Nearly nixed Tour de Georgia pumps up local economy. Bicycle Retailer and Industry News, 16(11), 19-19.

INRIX. 2010.European National Traffic Scorecard:
<http://www.inrix.com/scorecard/default.asp>

Interface for Cycling Expertise (I-ce). 2000. The economic Significance of Cycling: A study to illustrate the costs and benefits of cycling
http://www.ocs.polito.it/biblioteca/mobilita/EconomicSignificance.pdf

Lankford, S., Grybovych, Ed.D., Lankford, J., Winger, C. 2008. The Economic Impact & Spectator Characteristics of RAGBRAI:
<http://www.uni.edu/step/reports/STEP_RAGBRAI.pdf>

League of American Bicyclists. 2012. National Bike Summit State and District Factsheets:<http://www.bikeleague.org/conferences/summit12/factsheets.php >

London School of Economics. 2010. The British Cycling Economy; “Gross Cycling Product”

McClintock, H. (1992). Post-war traffic planning and special provision for the bicycle. In H. McClintock (Ed.) The Bicycle and City Traffic (pp.19-39). London, England: Belhaven Press.

OECD. 2003. Health Data 2003:
< http://www.oecd.org/els/health-systems/oecdhealthdata2003showhealthexpendituresatanall-timehigh.htm>

Outdoor Industry Foundation. 2006. The Active Outdoor Recreation Economy:
http://www.outdoorindustry.org/images/researchfiles/RecEconomypublic.pdf

Rojas, R. 2011. The health risks and benefits of cycling in urban environments compared with car use: health impact assessment study:
<http://www.bmj.com/content/343/bmj.d4521>

Shayler, M., Fergusson, M., Rowell, A.. ( CTC) 1993. Costing the benefits: The value of cycling. Transport Research Laboratory. Crowthorne House, Nine Mile Ride Wokingham, Berkshire RG40 3GA. United Kingdom

TRL. 2009. Collisions Involving Cyclists on Britain's Roads: Establishing the Causes:
<http://www.trl.co.uk/online_store/reports_publications/trl_reports/cat_road_user_safety/report_collisions_involving_pedal_cyclists_on_britains_roads:_establishing_the_causes.htm>

The Royal society for the prevention of accidents. 2013. Annual Report.
<http://www.rospa.com/about/annualreview/info/review2013.pdf>

UITP. 2003. Economics of Transport:
<http://www.uitp.com/Events/madrid/mediaroom/Backgrounders/Economics_of_Transport.Htm>

United Kingdom Department for Transport. 2008. Transport Trends: <http://webarchive.nationalarchives.gov.uk/20100406130654/http:/www.dft.gov.uk/adobepdf/162469/221412/190425/220778/trends2009.pdf>

Weighand, L. 2008. The economic benefits of cycling:
< https://www.pdx.edu/ibpi/sites/www.pdx.edu.ibpi/files/Economic%20Benefits%20of%20Bicycling.pdf>

Whitelegg, J., 1993, Time Pollution:
< http://www.ssti.us/wp/wp-content/uploads/2013/05/Time-Pollution.pdf>

World Health Organization. 2011. World Health Statistics:
< http://www.who.int/whosis/whostat/EN_WHS2011_Full.pdf?ua=1>

# EQUITY APPRAISAL IN SLOVENIA

**Lučka Lorber**[1]
[1] *University of Maribor, Faculty of Arts, Department of Geography, Koroška cesta 160, 2000 Maribor, Slovenia*

**Abstract:** Slovenia prepared the new transport project axis 3 to improve the accessibility for less developed NUTS 3 regions. The proposal of this axis line stream from Austrian border on the north, to Croatian border on the south and one of the main aims is to connect this axis on existing Pan-European highways cross (V. and X. European corridors). The question of the effects of traffic infrastructure on regional development (equity) has been widely addressed in the literature, from both, the aspect of growth as well as the aspect of economic geography. An efficient traffic system ensures utilization of economic and social opportunities, thus affecting the whole society. Local inhabitants and authorities are involving in all phases of planning (holistic approach). The effects of traffic can be divided into 2 groups: direct effects that refer to improved accessibility to the markets due to shorter transportation times and lower costs for the providers and for direct users of transportation means, and indirect effects that are shown at an individual level (larger range of goods, lower prices, land rent) and at the national and regional levels (higher competitiveness, larger mobility, setting up distribution networks, etc.). Improved traffic infrastructure means larger traffic capability as well as higher efficiency and reliability of the traffic structure (lowering the risks carrying out traffic which is especially important in the time of integrating individual production- and service systems into supply chains). Despite a longer route and consequently higher investment- and operational costs, such a route can be more efficient from the aspect of the distribution of the activities in the space, environment protection, equity and suchlike. When it comes to choosing a traffic infrastructure corridor, all of the above results in utilising the multicriteria decision making process which includes multiple, appropriately balanced aspects or limitations to placing such a corridor into the space.

**Keywords:** equity, holistic approach, economic geography

## 1. Introduction

National policy as well as state and private agencies involved in planning transport infrastructure play an important role in planning the transport project axis 3 in Slovenia. The political framework is bound by the existing legislation. while planning is bound by the examples of good practices and the specificity of Slovenian space. Entering the EU in 2004, Slovenia also adopted the EU legislation framework in the area of transport development and building the transport infrastructure. Changes in economic development had the decisive influence on development of transport infrastructure and approaches to planning new transport corridors. De-industrialisation and restructuring the conventional industrial activities, alongside with rapid development of the service sector (Lorber, 1999), set new reference lines in transport infrastructure planning for politicians and transport planners. The awareness of the urgency for the protection of the environment and space is coming to the forefront. With the new rural development paradigm (OECD, 2006), the attitude towards valuation of rural areas has also changed. The transition from the exogenous development paradigm where the key sector was Agriculture, to endogenous sectoral mixed economy changed the needs for mobility of the population in the predominantly rural regions (Lorber, 2005; Lorber and Žiberna, 2014). With the changed rural development paradigm and the care for sustainable development the transport infrastructure planners need to take the limitations into consideration. **By signing the Alpine Convention in 1992, Slovenia, together with other alpine countries, agreed to ensure protection and sustainable development of alpine space and simultaneously to protect economic and cultural interests of local population.** The Alpine Convention covers 33.4% of Slovenia's territory. Regardless of the reasoned need for respecting and implementing the Alpine Convention as well as the need for the realisation of the transport axis 3 project, the realisation of the both of these brings up a conflict of interest between public interest holders (investors in the planned infrastructure), other public interests (maintaining spatial and environmental qualities) and the public. It is a spatial planner's task to find a solution to a conflict of interests that will seek the most optimum solution that will present an acceptable compromise for every interest involved while respecting all the interests involved. The execution of Alpine Convention and its protocols is on the other hand obligation of all governmental institutions, local communities, NGO's, research, expert and other organizations as well as of individuals (Šolar, 2010). Another limitation to the usage of space is posed by NATURA 2000 (covers 37.0% of Slovenia's territory). When planning traffic corridors, the interests of the state as well as its population are made a priority. The interest of the state reflects in the efforts for achieving uniform regional development of the economy and space. The interest of the population in the influence area however, is ensuring equity[2], which is one of the fundamental human rights. Public interest can be divided into 2 groups: direct effects that refer to improved accessibility to the markets due to shorter transportation times and lower costs for the providers and for direct users of transportation means, and indirect effects that are shown at an individual level (larger range of goods, lower prices, land rent) and at the national and regional levels (higher competitiveness, larger mobility, setting up distribution networks, etc.).

[1] Corresponding author: lucka.lorber@um.si

[2] Throughout the document the term equity is employed, rather than the terms justice or fairness. While each of these terms may refer to different concepts in certain contexts, in common usage the terms strongly overlap and are used interchangeably (Bastiaanssen et al., 2014)

Often times, there is an occurrence of conflict between state policy (usually the investor) and the public engagement. There is a wide range of people and organisations that have interest in a particular project and become involved, to varying degrees, in decision-making. These are known collectively as ‘’stakeholders’’ (Cascetta and Pagliara, 2012).
In the remainder of the article we will present the development of the appraisal methods to different approaches to transport infrastructure planning with the emphasis on equity between the participants in planning transport infrastructure. We are aware of the importance of interdisciplinary approach to planning and will, as geographers, present the role of Geography in the new approaches to planning transport infrastructure, as Geography is a science that studies the role and influence of the man in space. While evaluating the legitimacy of building the third development axis we will show the approach by Slovenian politicians, building designers' organisations, local community and the interested civil society in the process of making the project of placing the new transport corridor into space (Oven, 2010). Using SWOT analysis we will present the strengths, weaknesses, opportunities and threats involved in building transport axis 3 for the state and the regions where the planned transport corridor will be aligned. The major features of the policy and planning processes are examined because they both have to reflect the fundamental changes in society and contemporary issues and problems.

## 2. Overview of Slovenian legislation for public involvement in preparation procedures for spatial planning documents

The key documents that will be presented in this chapter as legal foundation for public involvement in preparation procedures for spatial planning documents are: Aarhus Convention, Environmental Protection Act (Official Gazette of the RS, No 41/04 - ZVO-1 in 70/08 – ZVO–1B) and the legislation in force. Public Information Access Act (Official Gazette of the RS, No 51/2006) which is a foundation for public involvement in spatial planning documents preparation procedures will not be presented separately here as it is obligatory for all spatial planners. The Republic of Slovenia had signed and, in May 2004, also ratified the Aarhus Convention on access to information, public participation on decision-making and access to justice in environmental matters. Articles 6 and 7 in greater detail define the possibilities for public involvement in plans, programmes and policies in connection to the environment and space. The relevant clauses are particularly the following:

- the party provides the necessary information to the public and adopts relevant practical and/or other measures to assure transparent and equitable public involvement in preparation of plans and programmes;
- a relevant period of time needs to be ensured for the individual stages in public involvement so as to ensure enough time to inform the public and that the public can prepare and can get efficiently involved in decision making;
- the party ensures public involvement at the beginning of decision making when all options are still open and the public can get involved in an efficient manner;
- the party ensures that the results of public involvement in decison making are considered adequately;
- a relevant public authority will, upon considering the aims of this convention, define the eligible public to get involved;
- the party will adequately aim at ensuring the opportunities for public involvement in preparation of policies, related to space and environment.

The regulatory framework in the area of access to public information in the decision-making processes and protection of their rights is today covered by the Public Information Access Act. Spatial planning regimes before year 2002 mentioned several formal types of public involvement in creating and making decisions on spatial management (public unveiling of spatial implementing regulations and their public consultation) however, the then established forms of public involvement proved not to ensure sufficient influence of the local environment and civil society on decision making on future spatial arrangements. Following the example of other countries, Spatial Management Act, adopted in 2002 (Official Gazette of the RS, No 110/02, 8/03 – amended in 58/03 – ZZK-1), introduced new forms of public involvement in drafting and decision making in spatial management: spatial conferences where in addition to the representatives of organised civil society, subjects whose legitimate requirements affect the planned spatial arrangements were involved. Other already established public involvement instruments were preserved in the Spatial Management Act. Article 5 of the Spatial Planning Act, adopted in 2007 (Official Gazette of the RS, No 33/07 and 70/08 – ZVO-1B) stipulates that , in accordance with the act, everyone has the right to be informed on the procedures of drafting and adopting spatial planning acts and other matters of spatial management. In addition to that, it states that everyone has the right to participate in spatial management procedures with initiatives, opinions and in other ways. Relevant spatial management bodies have to provide everyone with insight into spatial planning documents, and inform the public on spatial planning matters in accordance with the Act. (Šolar, 2008).

## 3. Overview current appraisal of differing methodologies

Cost-benefit analysis (CBA) remains the most common method used to evaluate transport projects. Changes in economic and social development strengthen the civil society and the awareness of comprehensive sustainable spatial development. For this reason, experts from different areas of expertise participated, by way of criticising the CBA method and suggesting new socio-economic indicators and equity, in developing a new multi-criteria analysis (MCA) method to be used in the appraisal of transport projects. From the equity standpoint, the explicit inclusion of equity theories in sustainable mobility inequality indicators (SUMINI) into decision-making processes means upgrading the MCA method (Thomopoulos and Grant-Muller, 2013).

**3.1 Cost benefit Analysis**

Cost-Benefit Analysis (CBA) estimates and totals up the equivalent money value of the benefits and costs to the community of projects to establish whether they are worthwhile. The weaknesses of the method are in the limited range of indicators that are set mainly by engineers and economists. New socio-economic indicators can not be evaluated directly so their influence on the indirect costs and/or benefits are difficult to evaluate. In their paper (Thomopoulos and Grant-Muller, 2013) presented the set of wider impacts that could be captured in principle and are often described as socio- economic impacts, project externalities or more recently as wider socio-economic impacts of transport projects. Examples include environmental impacts, benefits to the natural habitat, visual intrusion, health impacts, settlement cohesion, accessibility, land use planning, agglomeration, labour displacement, habitat fragmentation and equity. The set of wider impacts that should be captured in practice is not formally agreed and is likely to vary with the nature of the project. Mackie and Nellthorp offer a generic description of the wider impacts of transport to justify the commonly anticipated outcome that transport infrastructure projects will generate economic benefits. Their description is based on the argument that: ‘‘changes in transport costs should be reflected in changes in accessibility, which in turn changes the pattern of demand for land. Finally, the level and pattern of prices and outputs is modified’’ (Mackie and Nellthorp 2001, Thomopoulos and Grant-Muller, 2013).

**3.2 Multi Criteria Analysis**

Multi-Criteria Analysis (MCA) is a decision-making tool developed for complex problems. In a situation where multiple criteria are involved confusion can arise if a logical, well-structured decision-making process is not followed. Another difficulty in decision making is that reaching a general consensus in a multidisciplinary team can be very difficult to achieve. By using MCA the members don't have to agree on the relative importance of the Criteria or the rankings of the alternatives. Each member enters own judgements, and makes a distinct, identifiable contribution to a jointly reached conclusion. Alternatives are evaluated on a set of criteria reflecting the decision-maker’s objectives, and ranked on the basis of an aggregation procedure. Scores achieved do not necessarily need to be conveyed in monetary terms, but can simply be expressed in physical units or in qualitative terms (De Brucker et al., 2011).

**3.3 New approach – SUstainable Mobility INequality Indicator (SUMINI)**

Thomopoulos and Grant-Muller in their scientific article ‘’Incorporating equity as part of the wider impacts in transport infrastructure assessment: an application of the SUMINI approach’’ summarised the benefits of SUMINI through the following points:
The theory and practices complement each other and are interrelated which means that the adopted decisions tend to be more transparent and credible.
Using different approaches and theories which can be used when making decisions for different types of equity.
In the light of different equity tapes or principles the decision making process is endogenous as opposed to the so far commonly used exogenous practices.
The SUMINI 's considerable theoretical contribution results from the explicit inclusion of equity theories in the decision making (Thomopoulos and Grant-Muller 2013).

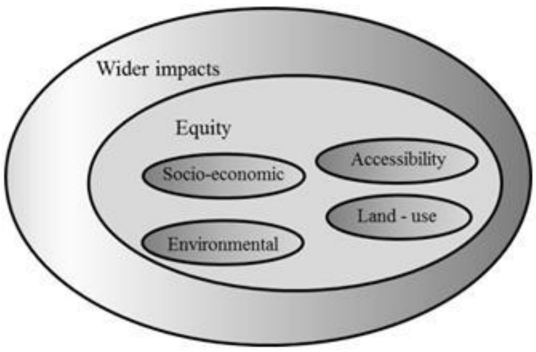



**Fig.1** *Interrelation between wider impacts of transport projects and equity*
*Source: Thomopoulos and Grant-Muller 2013*

These wider impacts include regional development, regional disparities, environment, safety, and labour market. This scope originates from the fundamental policy aims in planning larger transport systems, which is planning a new development transport axis , namely aimed at ensuring equity impacts that result from their socio-economic, environmental or accessibility dimensions. This concept can be supplemented by the range of equity variables that

Litman presented in the table: Equity Evaluation Variables. He defined type of Equity categories, impacts and measurement units. He thus defined the factors that influence the equity analyses, including the type of the included equity, the categorisation of people, which particular impacts are being considered, and how they are measured. These theoretic premises give us a wide enough range of equity impacts to be able to adopt quality decisions and assessments of the proposed transport projects.

**Table 1**
*Equity evaluation Variables*

| Types of Equity | Categories | Impacts | Measurement Units |
|---|---|---|---|
| - Horizontal<br><br><br><br>- Vertical - with respect to - income and social class<br>- Vertical – with respect to – need and ability | - Demographics (age, gender, race, ethnic group, family status etc.)<br>- Income class<br>- Geographic location<br><br>- Ability (e.g. people with disabilities, licensed drivers, etc.)<br>- Mode (personal or public transport)<br>- Vehicle type (cars, trucks, buses)<br>- Industry.(truckers, transit, taxis,<br>- vehicle manufactures.<br>- Trip type and value. | - Price or fare structure.<br>- Tax burdens.<br>- Transportation service quality.<br>- External costs (crash risk, congestions, pollution…<br>- Economic opportunity and development.<br>- Transport industry employment and business opportunities. | - Per capita.<br>- Per vehicle kilometer.<br>- Per passenger kilometer.<br>- Per trip.<br>- Per peak-period trip.<br>- Per euro paid in fare or tax subsidy. |

*Source: (Litman, 2006)*

## 4. Equity appraisal in Slovenia

Slovenia is a small country (20,273 km$^2$), its mean sea level is 557 m. There are two main transport links passing through its territory, namely the Pan-European corridors V. and X. Its population slightly exceeds two million. In the planning of transport corridors, urbanization of the space plays an important role. According to Eurostat, degree of urbanisation is a classification based on the share of local population that classifies Local Administrative Units level 2 (LAU2 units) belonging to the same NUTS 3 region into urban clusters. To classify rural and urban areas, this concept, in addition to population density by municipalities, uses two additional criteria, namely:

- geographical contiguity of local units (municipalities – LAU 2) that comply with the criteria of population density, and
- minimum population threshold.

Based on the said criteria, this concept creates a three-way classification of the areas:

- Densely populated areas
- Intermediate density areas
- Thinly populated areas

According to this definition, majority of Slovenian municipalities belong to thinly populated areas (149 municipalities or 71%), as the only two densely populated areas are the municipalities of Ljubljana and Maribor (1%). The remaining 59 municipalities are classified as intermediate density areas. Some intermediate areas are bound to the two largest urban centres whereas majority of the intermediate areas form individual spatial units.

The fig. 2 shows higher population density alongside the existing motorway cross while the planned new development axis 3 runs mainly through less densely populated areas. Broader influence area of the planned development axis is less developed in terms of transport. Poor access of the inhabitants and business entities to local and regional centres as well as to the motorway cross presents a serious obstacle to development. This means that the population do not have equal development potentials, which affects the socio-economic situation of the area and reflects in demographic indicators. The area is depopulating along with the ageing of the population. When planning transport corridors, we want to achieve maximum positive effects with minimum possible costs. As follows from the theoretical part of the article, it is not necessary that the planned route is also an optimum one. Despite a longer route and consequently higher investment- and operational costs, such a route can be more efficient from the aspect of the distribution of the activities in the space, environment protection, equity and suchlike. When it comes to choosing a traffic infrastructure corridor, all of the above results in utilising the multicriteria decision making process which includes multiple, appropriately balanced aspects or limitations to placing such a corridor into the space.

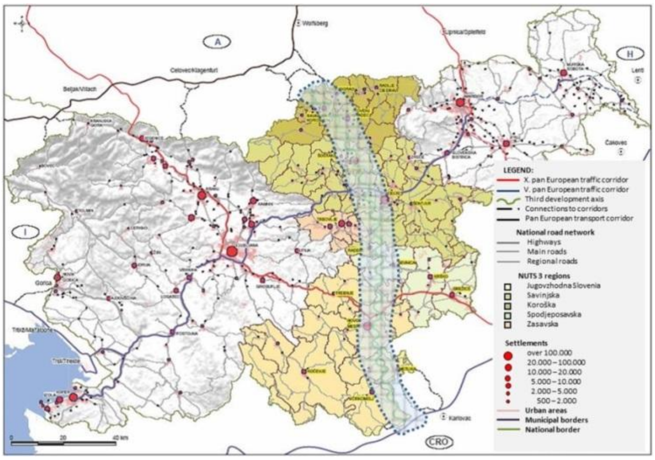



**Fig. 2**
*Corridor for new transport project north – south conection*
*Source: Adapted from Comprehensive Project for the Development of the Area of the Third Development Axis, Ministry of Infrastructure and Spatial Planning, Ljubljana 2006*

The criteria for placing a corridor into space are the following:
1. terrain – representing physical restrictions,
2. environment – representing a value which should be protected; however, when building traffic infrastructure, a certain amount of environmental damage should be taken into consideration,
3. distribution of economic activity in the space to which the new infrastructure should primarily serve,
4. political aspect (political preferences regarding the course of the route).

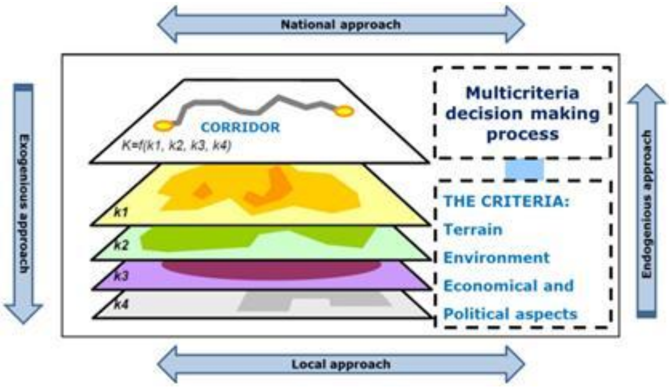



**Fig. 3**
*Multicriteria decision-making process on the siting corridor in space*
*Source: Adapted from Rodrigue (Rodrigue et.al., 2006) cited in Comprehensive Project for the Development of the Area of the Third Development Axis, Ministry of Infrastructure and Spatial Planning, Ljubljana 2006*

The necessary data bases are as follows:
- relief, building register,
- population and job at the level of settlements,
- road network,
- ecologically important areas and Natura 2000, protected natural areas, natural values, cultural heritage, rural values,
- water catchments and water protection areas, artificial lakes and watercourses,
- best agricultural land, forests (reserves and protective forests),
- endangered areas (floods, avalanche-prone and erosion areas),
- areas, intended for development of business activity and house building,
- politically desirable or undesirable areas in regard of building traffic links.

The entire procedure of drafting national spatial plans is public as every individual or the interested public have the opportunity to get acquainted with all the finished phases of the procedure and all the materials that have been prepared. Collaboration with the public in the context of drafting national spatial plans for development Axis 3 the collaboration with the public went on almost entirely according to regulatory provisions. The planned infrastructural object will change the existing relations to a great extent. Living conditions will change. The project will affect the protected areas and will change the visual image of the area alongside the route.

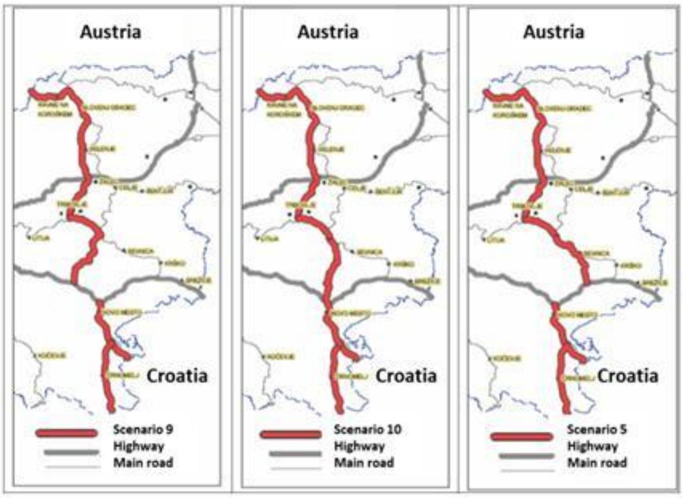



**Fig. 4:**
*The results of multicriteria analysis*
*Source: Comprehensive Project for the Development of the Area of the Third Development Axis, Ministry of Infrastructure and Spatial Planning, Ljubljana 2006*

| | Scenario (i) | | | | | | | | | | Weight (wj) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Indicator (j) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | **9** | 10 | |
| of the spatial impacts | 1,23 | 1,00 | 1,12 | 1,07 | 1,74 | 1,78 | 0 | 1,19 | **2,00** | 1,94 | 33 % |
| of economic efficiency | 1,00 | 1,14 | 1,87 | 1,52 | 1,97 | 1,07 | 0 | 1,99 | **2,00** | 1,70 | 33 % |
| of environmental impacts | 1,60 | 2,00 | 1,13 | 1,13 | 1,27 | 1,00 | 0 | 1,43 | **1,13** | 1,43 | 33 % |
| total score | 1,28 | 1,38 | 1,38 | 1,24 | 1,66 | 1,28 | 0 | 1,53 | **1,71** | 1,69 | 100 % |

SWOT analysis for the area of building third development axis:

| Strenghts | Weaknesses |
|---|---|
| - efficiency of the industry above Slovenian average,<br>- rapid internationalization of business operation,<br>- inclusion into national development programmes and compliance between national and regional development plans,<br>- EU structural help funds. | - high share of the GVA (gross value added) industry, low share of services,<br>- weaker educational structure, above-average unemployment,<br>- consequently, weak entrepreneurial activity,<br>- poor traffic access to traffic hubs in the neighbouring countries, the motorway cross, centres with international importance and in the south the routes even to the centres of national importance,<br>- low density of the traffic network in peripheral regions and insufficient public transport connections within the area, |
| **Opportunities** | **Threats** |
| - endogenous approach to rural development,<br>- new logistics centres;<br>- development of new and old tourist destinations – natural and cultural sources and attractive rural area,<br>- new craftsmanship and industrial zones;<br>- better traffic connection from Austria or Croatia to tourist destinations,<br>- better opportunities for employment in tertiary and maybe also in secondary sector;<br>- shorter route to the Croatian seaside. | - unfavourable demographic trends,<br>- increased competitiveness by other regions,<br>- loss of importance of Luka Koper to Central Europe due to the shorter route to Luka Rijeka,<br>- building a new railway corridor through Slovenia and lowering the possibility of investment in this area,<br>- increased transit through Slovenia,<br>- destroying the ecotypes when building motorways,<br>- heavier pollution of water resources;<br>- devastation of land and smaller number of farms whose primary income comes from agriculture. |

*Source: Adapted from Žibret, 2010 and Comprehensive Project for the Development of the Area of the Third Development Axis, Ministry of Infrastructure and Spatial Planning, Ljubljana 2006*

Building of Axis 3 would mean better accessibility for the population in the area to basic goods, services and activities on both national and regional levels. Regional disparities would lessen. New challenges will be set for both the population as well as for the existing economy. This will stop negative demographic trends and depopulation of space.

## Conclusion

Traffic and transport serve as the means of fulfilment of inhabitant's needs for goods and services. Needs are divided into needs for fulfilling personal needs and needs of the economy. Principally, we distinguish between a number of levels of traffic and transport. Local level comprises traffic and transport flows in urban areas and is predominantly bound to mobility and accessibility of jobs, food- and other material goods supply, access to basic services (healthcare, public administration institutions, cultural goods and sports and recreation goods). When appraising equity, the socio-economic situation of the society serves us as a basis. Underprivileged groups are mostly groups with lower disposable income, older inhabitants, and inhabitants with special needs (the disabled). Different authors address factors that influence general characteristics of these groups. They depend on the size of the countries, racial structure, ethnic structure, use of language, level of education (Karner and Niemeier, 2013; Sánchez et al., 2003; Manaugh, 2013). These differences affect the situation of these groups within society. One common feature of the underprivileged groups is that they tend to be less motorised (under-average ownership of transportation means and smaller number of driving licenses). Thus, they are bound to use public transport. Considering the social status, these groups live in urban neighbourhoods with lower infrastructure level of equipment (lower real estate prices) which are usually far away from the supply centres. Accessibility of all types of goods in urban areas depends mainly on the level of development of transport infrastructure and of all types of public transport. The bases for equity assessment are frequency of public transport and distance between the stops. These depend on general well-being of the environment that is capable of funding these types of transport from public resources. Taking into consideration the size of Slovenia, its geographical features and level of development, we can conclude that the situation in public transport has been getting worse in the recent years. The reason for this is rapid motorisation which affected the decreased use of public means of transport. The economic crisis added to the decreased frequencies of regular passenger lines. De-industrialization and the consequent gentrification of some urban neighbourhoods affected the unequal opportunities of access to goods especially for the socially weak individuals. On regional level however, traffic accessibility of settlements with central functions, accessibility of main transport corridors and transport hubs mean a significant competitive edge. Therefore, when planning new transport development axes – for example the third development axis – we need to consider equity principle for all the inhabitants of the influence area of the planned route. The decision to use Multi-Criteria Analysis in planning the third development axis is correct as it takes in consideration majority of indicators based on which the optimum version of the development axis into space, both direct as well as the indirect ones. With the third development axis, there is a conflict of interests between the state and the inhabitants living along the axis. Building the third development axis brings competitive advantages to the neighbouring countries, which could affect development of the broader region in the long term. From the standpoint of assessing equity of the population in the area of the route, the third development axis is important and useful for greater mobility and accessibility. However, the negative effects of building the third development axis should be resolved in a timely manner, by implementing adequate development policy in the affected areas which could lose their existing competitive advantages in global transport flows (especially the Port of Koper). The final finding is that the majority of general reasons that affect the population equity of access to goods in other, larger countries do not exist in Slovenia. The main limiting factors in building adequate transport infrastructure are: relatively vast protected areas, an uneven and geologically diverse terrain, and lack of public resources to fund the construction. Economic stagnation and the consequent growing number of the population with lower income could have adverse effects on mobility of the population in the near future, which can present a long-term severe threat of growing regional disparities and worsening of demographic structure of the population in less developed Slovenian regions. The legislation provides ample opportunities for inclusion of the interested public and individuals into all phases of planning the transport infrastructure. This enables protection of interests and the right to assure equity to mobility and accessibility for all social groups.

**References**

Bastiaanssen, J.; Lucas, K.; Martens, K. 2014. Transport Equity Appraisal, A literature review, edited by Di Ciommo, F.; Dupont, A. (TEA Cost Action Report). 13 p.

Cascetta, E.; Pagliara, F. 2012. Public engagement for planning and designing transportation systems, SIDT Scientific Seminar 2012, *Procedia - Social and Behavioral Sciences* 87: 103 – 116.

De Brucker, K.; Macharis, C.; Verbeke, A. 2011. Multi-criteria analysis in transport project evaluation: an institutional approach, *European Transport \ Trasporti Europei* 47: 3-24, Available from Internet <http://www.ucalgary.ca/verbeke/files/verbeke/MultiCriteriaAnalysis-MCA-Institutions-Europe.pdf>.

Karner, A. 2012. *Transportation Planning and Regional Equity: History, Policy and Practice*, September 2012, Research Report – UCD-ITS-RR-12-29. 213 p.

Karner, A.; Niemeier, D. 2013. Civil rights guidance and equity analysis methods for regional transportation plans: a critical review of literature and practice, *Journal of Transport Geography*, 33: 126–134.

Laird, J.; Mackie, P. 2009. *Review of Economic Assessment in Rural Transport Appraisal*, Institute for Transport Studies, University of Leeds, Scottish Government Social Research 2009. 96 p. Available from Internet < http://www.scotland.gov.uk/Publications/Recent>.

Litman, T. 2006. *Evaluating Transportation Equity Guidance For Incorporating Distributional Impacts in Transportation Planning.* Victoria Transport Policy Institute, 8 March, 2006, 43 p.

Lorber, L. 1999. The economic transition of Slovenia in the process of globalization. *Geografski zbornik* 39: 133-166. Available from Internet < http://giam.zrc-sazu.si/zbornik/lorber_39.pdf>.

Lorber, L. 2005. Regional development problems of Podravje. *Podravina* 4(8): 107-120. Samobor. Available from Internet < http://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=116624>.

Lorber, L.; Žiberna I. 2014. **European Agricultural Policy and Structural Changes in Agricultural Holdings in Podravje between 2002 – 2012,** Podravina XIII(25): 19-44. Koprivnica. Available from Internet <http://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=186681>.

Mackie, P.; Nellthorp, J. 2001. *Cost-benefit analysis in transport.* In: Hensher, D.; Button, K. (eds.) *Handbook of Transport Systems and Traffic Control*, Chap 10: 143–174. Pergamon, Amsterdam.

Manaugh, K. 2013. *Incorporating issues of social justice and equity into transportation planning and policy*. School of Urban Planning , McGill University, April 9, 2013. Available from Internet <http://tram.mcgill.ca/Teaching/PhD/dissertations/Manaugh.pdf>.

Oven, I. A. 2010. The Third Development Axis – The Spatial Plan of National Importance and Basis for the Placement of Spacial Arrangement. 10.Slovenski kongres o cestah in prometu, Portorož. 106-116.

Rodrigue, J-P.; Comtois, C.; Slack, B. 2006. *The Geography of Transport Systems*, Routledge, New York, USA. 297 p. Available from Internet <http://geonas.at.ua/_ld/0/34_The_Geography_o.pdf>.

Sánchez et al. 2003. *Moving to Equity: Addressing Inequitable Effects of Transportation Policies on Minorities.* Cambridge, MA: The Civil Rights Project at Harvard University.

Šolar, H. 2008. Public participation in spatial plans' preparation procedures, 9. Slovenski kongres o cestah in prometu, Portorož, 22. – 24. oktobra 2008: 133-140.

Šolar, H. 2010. The planning of spatial arrangements of national importance under the conditions of Alpine Convention, 10. Slovenski kongres o cestah in prometu, Portorož, 20. – 22. oktobra 2010: 133-140.

Thomopoulos, N.; Grant-Muller, S. 2013. Incorporating equity as part of the wider impacts.in transport infrastructure assessment: an application of the SUMINI approach, *Transportation* 40: 315–345.

Žibret, B. 2010. The third development shaft or will Slovenia get a benefit of just additional transit? 10. Slovenski kongres o cestah in prometu, Portorož, 20. – 22. oktobra 2010. 117-122.

Sources:
Comprehensive Project for the Development of the Area of the Third Development Axis, Ministry of Infrastructure and Spatial Planning, Ljubljana 2006.

Natura 2000. Available from Internet <http://ec.europa.eu/environment/nature/info/pubs/paper_en.htm>.

OECD 2006. The New Rural Paradigm: Policies and Governance.*OECD Rural Policy Reviews*.

Resolucija o prometni politiki Republike Slovenije (RePPRS), Intermodalnost: čas za sinergijo, 6.6.2006, adapted 2008, Available from Internet <http://www.uradni-list.si/1/objava.jsp?urlid=200658&stevilka=2426 >.

Resolucija o razvoju slovenske logistike in transportni infrastrukturi. Available from Internet <http://www.ef.uni-lj.si/media/document_files/dokumenti/Resolucija-final_1b.pdf>.

Trajnostni mestni prometni načrti - načrtovanje za ljudi, Available from Internet <www.mobilityplans.eu>.

Trajnostno mestno prometno načrtovanje. Priročnik. Available from Internet <http://www.mzip.gov.si/fileadmin/mzip.gov.si/pageuploads/publikacije/>.

# INNOVATIVE FINANCING OF TRANSPORT INFRASTRUCTURE AND ECONOMIC GROWTH

**Jelica Petrović Vujačić**[1]
[1] *University of Belgrade, Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, Serbia*

**Abstract:** The paper analyses the main issues of financing transport infrastructure. These issues are related to the insufficiency of traditional funding schemes for transport infrastructure and attempts to adopt new innovative financing schemes. This become even more urgent at the time of the financial crises. Transport infrastructure investments have a positive impact on economic growth, create wealth and jobs, as well as enhance trade, improve geographical accessibility and the mobility of people. It is imperative to maximize the positive impact on economic growth and minimize the negative impact on the environment.
The main traditional sources of funding for transport infrastructure in the countries of the European Union (EU) include allocations from national and EU budgets, domestic and foreign loans, and official development assistance such as structural and cohesion funds. More recently, the term "innovative finance" has been applied to mechanisms that pay for transportation projects. Innovative finance is a broadly defined term that encompasses a combination of techniques and specially designed mechanisms to supplement traditional financing sources and methods. A number of innovative options that policymakers have at their disposal in order to finance infrastructure investment include: mobilizing additional public revenues; providing financial support measures to attract further private sector involvement in infrastructure; and blending concessional and non-concessional financing for infrastructure financing. The paper elaborates some recent developments of innovative financing of transport infrastructure in the EU and the USA and the lessons learned from these experiences that could benefit less developed countries.

**Keywords**: transport infrastructure, traditional financing, innovative financing, economic growth..

## 1. Introduction

Good quality infrastructure is a key ingredient for sustainable development. All countries need efficient transport, sanitation, energy and communications systems if they are to prosper and provide a decent standard of living for their populations. Unfortunately, many developing countries possess poor infrastructure, which hampers their growth and ability to trade in the global economy. The economic benefits of smart infrastructure investment are long-term competitiveness, productivity, innovation, lower prices, and higher incomes, while infrastructure investment also creates many thousands of jobs in the near-term.

Evaluating how transportation and other infrastructure benefit the overall economy has been the subject of an extensive economic literature. David Aschauer's research found very large economic gains from public capital generally, including but not limited to transportation, suggesting one dollar in output gains for one dollar in increased investment (Aschauer, 1989). Recent researches have highlighted the importance of selecting investments wisely in key areas of the country on the basis of their economic contributions. These researches have also emphasized the importance of maintaining existing assets in a good state of repair.

On the global level, infrastructure invesments are put in context of sustainable development. Faced with common global economic, social and environmental challenges, the international community is defining a set of Sustainable Development Goals (SDGs). The SDGs, which are being formulated by the United Nations together with the widest possible range of stakeholders, are intended to galvanize action worldwide through concrete targets for the 2015–2030 period for poverty reduction, food security, human health and education, climate change mitigation, and a range of other objectives across the economic, social and environmental pillars (UNCTAD, 2014). Reaching the SDGs will require a change in both public and private investment. Public sector funding capabilities alone may be insufficient to meet demands across all SDG-related sectors. Today, the participation of the private sector in investment in these sectors is relatively low. Only a fraction of the worldwide invested assets of banks, pension funds, insurers, foundations and endowments, as well as transnational corporations, is in SDG sectors.

Statistical evidence for United States showed that there is a direct positive link between infrastructure investment and GDP. For instance, for the period 1950-79, growth in public infrastructure contributed almost one- for-one to economic growth. During this period infrastructure investment in core areas such as transportation, water management and electricity generation grew at an average rate of 4% while the overall economic or GDP growth averaged 4.1% during the same period. On the other hand, during the period 1980-2007 growth in public infrastructure investment drastically fell to 2.3% while average annual GDP growth fell to 2.9 percent over the same period (Heintz et al., 2009). In addition to the output and productivity effects, infrastructure investment is believed to create more jobs than other types of investment through direct, indirect and induced jobs. However, the size of indirect and induced effects depends on the magnitude of the infrastructure investment multiplier.

[1] Corresponding author: j.petrovic@sf.bg.ac.rs

An assessment of global infrastructure investments required for the period 2013-2030, and especially transport infrastructure investments (Figure 1), shows that it will cost 57 trillion US dollars to build and maintain the world's roads, power plants, ports, airports, pipelines and the like. The problem with financing infrastructure investments by the banks loans arose with the financial crises since 2008. Banks are not only wary of making long-term loans, they are also reluctant to take as much risk as before. They used to lend 90 per cent of the construction cost of a large projects, such as a toll road in America. Now that figure is down to about 70 per cent. This forces the investors to come up with more of their own cash (The Economist, March, 2014).

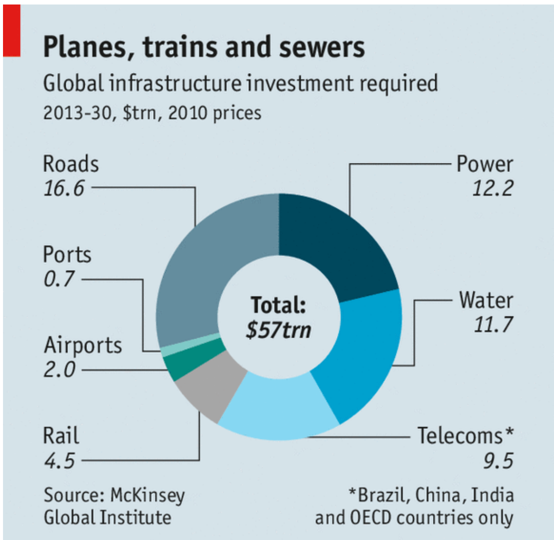



**Figure 1.**
*Global infrastructure investment required 2013-2030.*
*Source: (The Economist. 2014)*

## 2. Investments in transport infrastructure and economic growth

Transport infrastructure investment can contribute to growth by expanding the stock of capital available for use in producing goods and services. With more capital and more efficient production, both real income levels and standard of living can be expected to rise. There is an abundance of literature on the productivity of infrastructure investment. The general conclusion reached is that public capital has an impact on private capital, on labour productivity and hence on economic growth. The key issue in any analysis of these complex interrelationships is the understanding of the extent to which high investments are the cause or the result of economic growth. Transportation investments affect not only the level of economic output but geographic distribution of economic activity.

The relationship between transport investment and economic growth becomes much more complicated when a broader view of economic development is taken, linked to the concept of sustainable growth. This takes into account the effects on the environment, society and the economy both local and national. This reflects, in a way, the changes in transport infrastructure investment policy in Europe, for example, as summarised by Banister and Berechman (Banister & Berechman, 2000). Banister and Berechman are concerned with the relationship of transport infrastructure and economic development and question the effects of the first to the latter. They believe that in developing countries and cities, the relationship does exist and is quite clear, but in developed countries those links are unclear. Specifically for developed countries, they are arguing that additional transportation investment has little impact on the overall accessibility and has as a result the change of business patterns and mode trends and not economic growth.

Certainly accelerated infrastructure investment would provide an opportunity for construction workers to productively apply their skills and experience. These jobs span across a wide variety of different industries. For example, road building not only requires construction workers, but also grading and paving equipment, gasoline or diesel to run the machines, smaller hand tools of all sorts, raw inputs of cement, gravel, and asphalt, surveyors to map the site, engineers and site managers, and even accountants to keep track of costs. Analysis of data from the USA for 2012 annual input-output table and related data from the Bureau of Labor Statistics (BLS) suggests that 68 percent of the jobs created by investing in infrastructure are in the construction sector, 10 percent in the manufacturing sector, and 6 percent in retail trade (National Economic Council, 2014).
All countries face the basic economic problem of allocating scarce resources among competing uses in a way that maximises the net benefits to society. In general, market forces can be relied upon to ensure an efficient and productive rate of capital formation in the private sector. In the public sector, however, market forces are weak and investment objectives are often multifaceted.

This is especially true for transportation infrastructure investments. It often follows that transportation decision makers require additional information about the effects of investments and policies on the environment, business, productivity, economic growth, income distribution and other public concerns in order to ensure that investments yield benefits to the community that exceed the cost of achieving them.

A second type of economic objective concerns the distribution of transportation-generated economic benefits and changes in the incidence of these benefits across locations. There have also been strong, both urban and regional, distributional arguments for investment in transport infrastructure. The regional development policies in the European Union, where powerful and substantial investment has been transferred from rich countries to areas where this investment is needed, are a good example of this. The argument used by the EU is that regional development policy strengthens integration and cohesion in the EU as a whole, while at the same time reducing the disadvantages of poorly connected countries. The policy focus has shifted from priority to rural areas (investment in road and rail extensions to existing motorways in the "peripheral" countries in 1950-70, to economic integration and social cohesion objectives, the Trans-European road network and the high-speed-rail network.

## 3 . Financing transport infrastructure at the time of global crises

Infrastructure financing is to a large extent determined by its economic characteristics which contain a number of inherent challenges to investors: large up-front sunk costs, followed by a low marginal cost for each additional user. Combined with the long average lifespan of infrastructure projects, this poses significant challenges both for private and public investors. The former need to recover their costs, while the latter need to ensure that essential infrastructure services are made available in sufficient amounts and on equitable terms, while also being provided efficiently (EIB, 2010).

To reach the goal of a fully functional and EU-wide multimodal TEN-T 'core network' by 2030, with a high quality and capacity network by 2050, "unprecedented volumes of investment" will be required according to the European Commission. According to EIB estimates, over 1 000 billion EUR of investments will be needed to fulfil the priority targets of the Europe 2020 objectives in the energy and transport sectors.

China is setting up a development bank together with the four other members of the BRICS club of big emerging markets: Brazil, Russia, India and South Africa. The new BRICS development bank will focus on infrastructure lending to poorer countries. China is also pushing to establish another multilateral lender, the Asian Infrastructure Investment Bank, which will concretate on the same thing. With these two new banks, China is exporting a central feature of its development model to the rest of the world. It spent 8,5% of its GDP investing in infrastructure from 1992 to 2011 according to the McKinsey Global Institute. That was more than any other country (The Economist, July, 2014).

Private sector investment in infrastructure helps diversify the financial portfolio of banks, pension funds and other investment institutions and thus lowers overall portfolio risk. In times of crisis, such institutions suffer from losses to the value of their assets. Investment in infrastructure projects provides an interesting solution for such institutions, as it is characterized by low risk and often backed by a government safety net. In the USA, specifically, the costs of borrowing through the issuance of municipal bonds are at historic lows. Bond revenues are the primary source of infrastructure finance at the state and local level—and are also used to match federal funds (National Economic Council, 2014).

Public and private investment are complementary, not substitutes. Synergies and mutually supporting roles between public and private funds can be found both at the level of financial resources (raising private sector funds with public sector funds as seed capital) and at the policy level, where governments can seek to engage private investors to support economic or public service reform programmed. Nevertheless, it is important for policymakers not to translate a push for private investment into a policy bias against public investment. Policymakers need to find the right balance between creating a climate conducive to investment and removing barriers to investment on the one hand, and protecting public interests through regulation on the other. They need to find mechanisms to provide sufficiently attractive returns to private investors while guaranteeing accessibility and affordability of services for all. And the push for more private investment must be complementary to the parallel push for more public investment.

Increasing private investment in SDGs will require leadership at the global level, as well as from national policymakers, to provide guiding principles to deal with policy dilemmas; to set targets, recognizing the need to make a special effort for less developed countries (LDCs); to ensure policy coherence at national and global levels; to galvanize dialogue and action, including through appropriate multi-stakeholder platforms; and to guarantee inclusiveness, providing support to countries that otherwise might continue to be largely ignored by private investors. The potential for increasing private sector participation is greater in some sectors than in others. Infrastructure sectors, such as power and renewable energy (under climate change mitigation), transport and water and sanitation, are natural candidates for greater private sector participation, under the right conditions and with appropriate safeguards.

The many stakeholders involved in stimulating private investment in SDGs will have varying perspectives on how to resolve the policy dilemmas inherent in seeking greater private sector participation in SDG sectors. A common set of principles for investment in SDGs can help establish a collective sense of direction and purpose. The following broad principles could provide a framework.
• Balancing liberalization and regulation.
• Balancing the need for attractive risk return rates with the need for accessible and affordable services for all.
• Balancing a push for private investment funds with the push for public investment.
• Balancing the global scope of the SDGs with the need to make a special effort in LDCs (UNCTAD, 2014).

PPPs in general have decreased substantially since the economic and financial crisis starting in 2008 and so have PPPs in the transport sector. Data from the EIB shows that the market for PPPs has contracted significantly in recent years. From the all-time high point of about 140 PPP deals in 2006 and almost EUR 30 billion worth of PPP deals in 2007, in 2012 the numbers were significantly lower - only 66 deals with financing requirements of less than EUR 12 billion and the transport PPP deals haven't fared much better (EIB, 2012).

## 4 . Innovative financing of transport infrastructure

Traditional government funding sources are insufficient to meet the increasingly complex and diverse needs of the transportation system. Despite record levels of investment in surface transportation infrastructure in recent years, funding is not keeping pace with demands for improvements to maintain the vitality of the transportation system in most countries.

A range of innovative financing solutions to support sustainable development have emerged in recent years, including new financial instruments, investment funds and financing approaches. These have the potential to contribute significantly to the realization of the SDGs, but need to be supported, adapted to purpose and scaled up as appropriate.
It is important to note that many of these solutions are led by the private sector, reflecting an increasing alignment between UN and international community priorities and those of the business community. "Innovative Finance" for transportation is a broadly defined term that encompasses a combination of specially designed techniques that supplement traditional financing methods.

Financing from the EU budget to transport infrastructure projects of European interest is normally seen as a catalyst to the implementation of a project as the total amount of Community aid may not exceed 10 % of the total investment cost. Financial instruments financed by the EU-budget, such as the LGTT, aim to create a 'leverage (or 'multiplier') effect' for the EU funds far exceeding the direct financial contribution itself.

A credit enhancement tool like the Loan Guarantee for TEN-Transport (LGTT) is expected to reduce the risk for private sector actors such as commercial banks and therefore make them more willing to provide long-term lending to demand-based projects. Although factors underpinning the LGTT, such as the cost of debt, are important in the choice of capital structure the key issue today is that governments are not promoting transport projects generally due to constrained public finance. The LGTT is an unfunded mezzanine debt instrument provided to Public-Private Partnerships (PPPs) in the transport sector (within the TEN-T network) to offset the early-period risks associated with demand uncertainty.

The EIB has a portfolio of instruments to finance transport infrastructure projects (Figure 2.). Among these, the LGTT and the equity infrastructure fund, the Marguerite Fund, are the only ones partly financed by the EU budget. Another financial instrument, the Project Bond Initiative,17 was added to the EIB portfolio in 2013 and will also partly be financed by the EU budget.

The LGTT was jointly established in 2008 by the European Union and the European Investment Bank (EIB) where both partners share financial risk. The EIB provides a guarantee in the form of a contingent credit line, which may be drawn upon by the project provider during the first 5 to 7 years of operation, if the revenues generated by a project are not sufficient to ensure repayment of the senior debt, in case the actual revenues from the project fall below the forecasted level (Halvarsson et al., 2014). During the period 2008-2012, the LGTT has been signed by five motorway projects, one maritime project, and one high speed rail project.

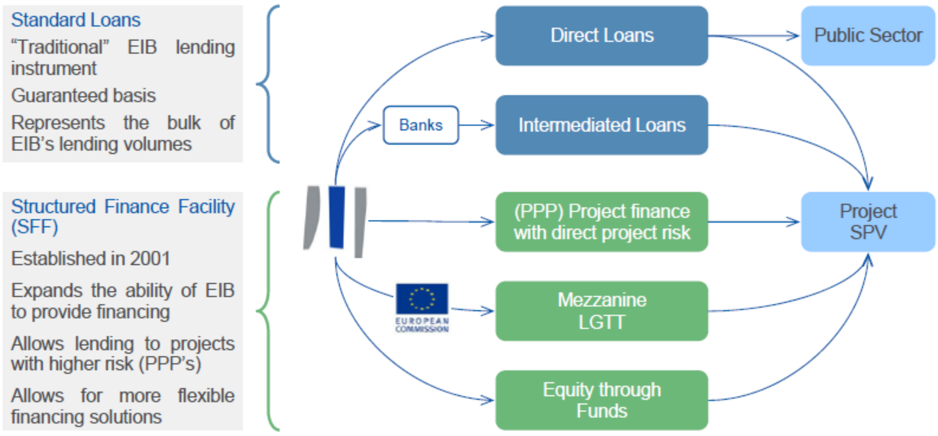



**Figure 2.**
*Overview of financing instruments offered by the EIB to PPP projects*
*Source:(Halvarsson, A. et al.)*

In the USA, through the Federal-aid program, Federal Highway Administration (FHWA) has financed highways primarily through grants that generally cover up to 80 percent of project costs. However, because this approach alone cannot meet the nation's current and future transportation investment needs, the U.S. Department of Transportation (U.S. DOT) innovative finance initiatives respond to the need to supplement traditional financing techniques. Over the last decade, the Federal government has responded to the investment gap by providing new funding techniques that complement and enhance existing grant reimbursement programs. This *Innovative Finance Primer* describes those techniques and provides examples of the techniques as applied by state and local partners.

The primary objectives of innovative finance are to (U.S.DOT, Innovative Finance, 2010):

- Maximize the ability of states and other project sponsors to leverage Federal capital for needed investment in the nation's transportation system;
- More effectively utilize existing funds;
- Move projects into construction more quickly than under traditional financing mechanisms; and
- Make possible major transportation investments that might not otherwise receive financing.

One key to effective use of innovative finance strategies is to recognize what kinds of projects can most benefit from which kinds of tools. As states and private sector sponsors look to applying innovative finance tools, it is important to recognize the potential synergy in combining tools to advance a project. Figure 3. introduces the major categories of innovative financing strategies and aligns those categories with some of the key financial characteristics of candidate projects.

The base of the pyramid represents the majority of highway projects that continue to rely primarily upon grant-based funding because they do not generate revenues, but can benefit from innovative finance tools that enhance flexibility and maximize resources. The mid-section of the pyramid represents those projects that can be at least partially financed with project-related revenues, but may also require some form of public credit assistance to be financially viable. The peak of the pyramid reflects the very small number of projects that may be able to secure private capital financing without any governmental assistance. These relatively few projects may be developed on high-volume corridors where the revenues from user fees are sufficient to cover capital and operating costs.

Advance construction allows a state to begin a project even if the state does not currently have sufficient Federal-aid obligation authority to cover the Federal share of project costs. Under PCAC, a state may elect to obligate funds for an advance-constructed project in stages. With tapered match, the non-Federal matching requirement applies to the aggregate cost of a project rather than on a payment-by-payment basis. Flexible match allows states to substitute private and other donations of funds, materials, land, and services for the non-Federal share of funding for highway projects.
States may use revenue from toll facilities as a credit toward the non-Federal matching share of certain highway projects.

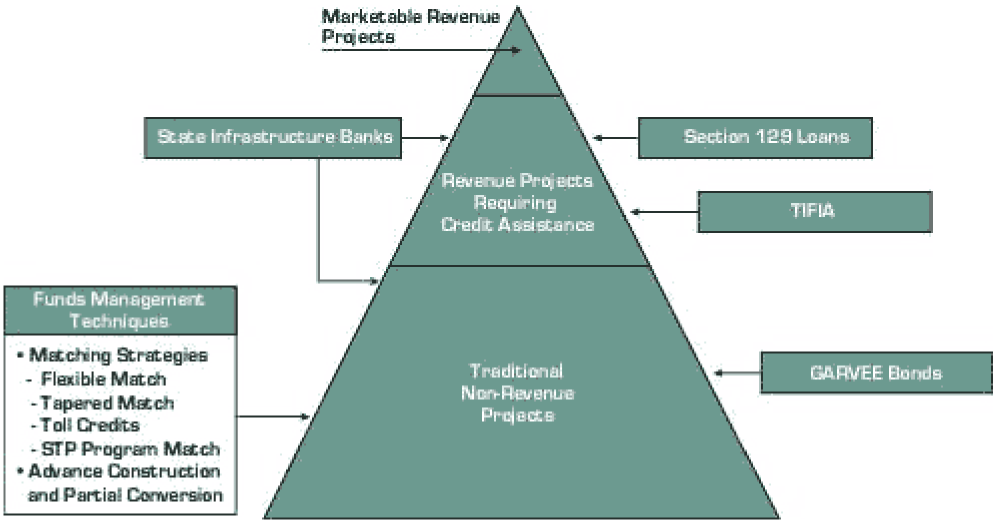



**Figure 3.**
*Project Finance Tools*
*Source:( U.S. Department of Transportation, Innovative Finance, 2010)*

The Transportation Infrastructure Finance and Innovative Act (TIFA) program provides Federal credit assistance in the form of direct loans, loan guarantees, and standby lines of credit to finance surface transportation projects of national and regional significance. TIFA credit assistance provides improved access to capital markets, flexible repayment terms and potentially more favorable interest rates compare to private capital markets for the similar instruments (U.S. DOT, GARVEE Guidance, 2014*)*.

From the Figure 4. We can see the steps associated with GARVEE financing. The proceeds from GRAVEE bonds are used to pay project cost as the construction costs are incurred. Rather than seek reimbursement on a progress basis as contraction costs are incurred, the State Department of Transportation (DOT) will bill FHWA for the Federal share of debt service payments over the term of the bonds.

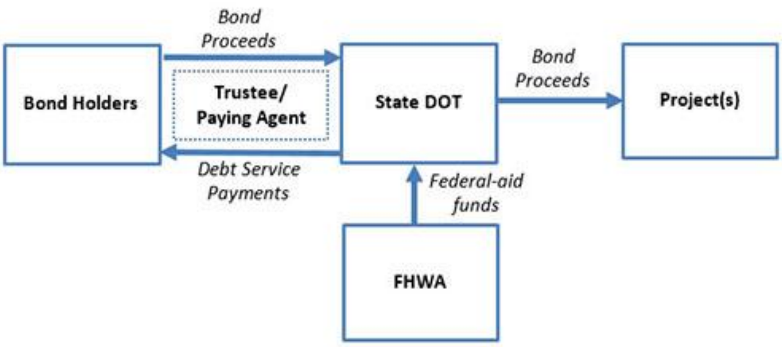



**Figure 4.**
*GARVEE financing*
(*Source: U.S. Department of Transportation, GARVEE Guidance, 2014*)

This “innovative finance” initiative was attempt to meet the increasing gap between transportation capital needs and available resources, without direct increases in Federal grant funding. The initiative also responded to states’ calls for greater flexibility in the use of their Federal aid funds.

FHWA initiated its innovative finance initiative seeking to accelerate projects by reducing inefficient and unnecessary constraints on states’ management of Federal highway funds and expand investment by (i) removing barriers to private investment in surface transportation infrastructure, (ii) encouraging the introduction of new revenue streams, particularly for the purposes of retiring debt obligations, and (iii) reducing financing and related costs, thus freeing up the savings for investment into the transportation system itself.

## 5. Concluding remarks

Due to the importance of infrastructure investment, the attraction of private capital as a partner of public capital is encouraged by innovative financial instruments and techniques. The main concern from investors is a shortage of suitable projects which should be attractive and justified not only for the investor, but also for society as a whole. Multiple stakeholders are working together seeking for innovative solutions of innovative project finance and new approaches for narrowing the gap between capital needs and financial resources. For the economies in transition and less developed economies it is of great interest to find a way to attract investors for their transport infrastructure projects. The lessons learned from the recent experience in innovative financing in developed countries cold be useful.


## Acknowledgements

The work on this paper was supported by Ministry of Science and Technological Development Republic of Serbia, through projects TR 36022 and TR 36006 for the period 2011-2015.

**References**

Aschauer, A.D. 1989. "Is Public Expenditure Productive?" *Journal of Monetary Economics*, 23(2): 77–200.

Banister, D.; Berechman, J. 2000. *Transport Investment and Economic Development*, London : University College London Press. UK and USA. 396 p.

The Economist. 2014. "A long and winding road", March 22nd. Available from Internet: <http://.economist.com/node/21599394/print>.

The Economist, "Bridges to somewhere", 2014, July 19th, p. 66.

European Investment Bank (EIB). 2010. *Financing infrastructure – A review of the 2010 EIB Conference in Economics and Finance*, p.13.

EIB. 2012. *PPP and financing in Europe: Recent trends and EIB involvement.*

Halvarsson, A. et al. 2014. *Ex-post evaluation of the loan guarantee instrument for Trans- European Transport Network projects (LGTT,)* February, Final report, European Commission. 72 p.

Heintz, J., et al. 2009. *How infrastructure investments support the US economy: employment, productivity and growth*, Political Economy Research Institute. January, Available from Internet: <http://www.americanmanufacturing.org/wordpress/wp-content/uploads/2009/01/peri_aam_finaljan16_new.pdf.>.

National Economic Council and the President's Council of Economic Advisers, 2014. *An Economic Analysis of Transport Infrastructure Investment*, The White House. Washington D.C. Available from Internet: http://www.whitehouse.gov/sites/default/files/docs/economic_analysis_oftransportation_investments.pdf>.

UNCTAD. 2014. *World Investment Report 2014: Investing in the SDGs: An Action Plan*. United Nations, New York and Geneva. 186 p.

U.S. Department of Transportation, Federal Highway Administration,. 2010. *Innovative Finance*, Available from Internet: http://www.dot.state.mn.us/planning/railplan/docs/mwrri/append/12_Innovative_Financing.pdf

U.S. Department of Transportation, Federal Highway Administration, 2014. *GARVEE Guidance (Grant Anticipation Revenue Vehicles*). Available from Internet: www.fhwa.dot.gov/ipd/finance/resources/federal_debt/garvee_guidance_2014.aspx

# POTENTIAL APPLICATIONS OF TENSEGRITY STRUCTURES TO BRIDGE CONSTRUCTION

**Yuliya Metodieva Ilieva**[1]
[1] *University of Structural Engineering & Architecture "Lyuben Karavelov" – Sofia, Faculty of Architecture, 175 Suhodolska Str., 1373, Sofia, Bulgaria*

**Abstract:** Tensegrity structures emerged in the twentieth century. They can be defined as self-stressed, stable spatial systems that are at a static equilibrium configuration under no external loads. Tensegrity structures consist of discontinuous compression bars inside a net of continuous tensile cables in such a way that the compression components do not touch each other. They are remarkable, lightweight and impressive in appearance and are strongly attracting the attention of architects and engineers nowadays. The aim of the present paper is to do a thorough research on potential applications of tensegrity structures to bridge construction. The realizations, as well as some conceptual proposals in this field, were collected and analyzed. The advantages and disadvantages of tensegrity structures in bridge engineering were defined. Special attention was given to the geometry configurations of the structures, basic tensegrity modules and technical considerations.

**Keywords**: tensegrity, tensegrity bridges, tensegrity structures, tensegrity systems, tensegrity modules.

## 1. Introduction

Tensegrity structures are lightweight spatial reticulated structures, composed of cables in tension and struts in compression. They can be defined as a subclass of cable structures, but unlike the latter their tensile forces are not anchored. The stability and stiffness of tensegrity structures are ensured by a self-equilibrated and self-stress state among tension and compression members. Their first description was made by Fuller (1962) in his patent "Tensile-integrity Structures". He characterizes them as islands of compression in an ocean of tension. Nowadays the Rene Motro's definition is widely accepted: "*A tensegrity system is a system in a stable self-equilibrated state comprising a discontinuous set of compressed components inside a continuum of tensioned components.*" (Motro, 2003, p. 19). A specific feature of tensegrity structures is that their compression components do not touch each other and do not transfer each other the compression forces which they are subject to.
Tensegrity structures emerged in the twentieth century. The term "tensegrity" reveals the mechanical behaviour of this type of structures as it derives from the words tension and integrity. The American architect and engineer Buckminster Fuller (1962) proposes it. The genesis and development of tensegrity structures, as well some polemic aspects about their invention, were discussed by Ilieva (2014 a). The original patent documents for this type of structures were analyzed by Ilieva (2014 b) in another paper. The concept of tensegrity structures is the focus of multiple research works and received significant interest among scientists and engineers throughout a diverse range of disciplines ranging from architecture, civil engineering, biology, robotics to aerospace engineering as reported by Skelton and de Oliveira (2009). Barbarigos et al. (2010) conclude that although tensegrities have been seen as attractive modular solutions for architects and engineers for decades now, few examples of tensegrity structures have been used for civil engineering purposes. Their variety and advantages are the reason for their big potential for further improvement in architectural and technical aspects as well as their application in construction for creation of roof coverings, bridge structures, interior elements and etc. The prestressing of the construction structures is a means for creating new structural forms as noted by Georgiev (1988).
The aim of the present paper is to do a thorough research on potential applications of tensegrity structures in bridge construction. The realizations, as well as some conceptual proposals in this field, were collected and analyzed. The advantages and disadvantages of tensegrity structures in bridge engineering were defined. Special attention was given to the geometry configurations of the structures, basic tensegrity modules and technical considerations.

## 2. Investigation methods

The method of "analysis – synthesis – assessment" was applied for the purposes of this article. During the analysis the studied object is decomposed in its compound parts. This is so called decomposition or divergence. The synthesis comes after the analysis as a rearrangement of different subparts based on common features. This is so called convergence. Thus the integrity of the object is restored. The assessment of the proposed ideas comes after the analysis and the synthesis. The phases of analysis, synthesis and assessment are repeated many times during the investigation. In the current article the discussion is made from particular to general (induction) and from general to particular (deduction). The comparative analysis is a method of work that was used in each stage of the current investigation. The examined tensegrity bridges were juxtaposed with each other.

[1] Corresponding author: yuliya_ilieva@vsu.bg

### 3. Models and realizations of tensegrity bridges

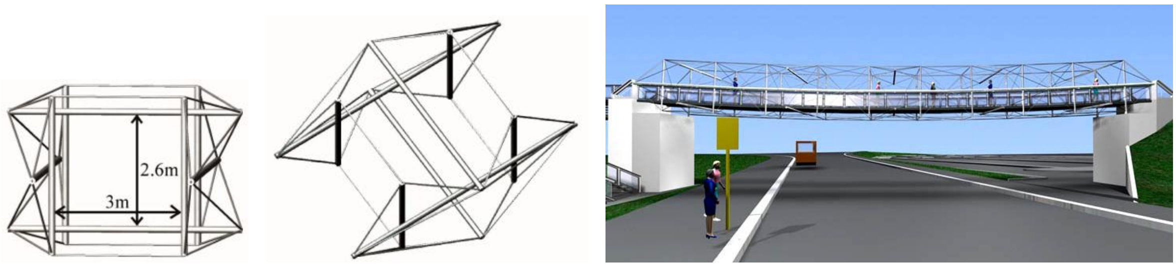
Several realizations of tensegrity bridges and conceptual proposals in this field are considered in this paper. They were collected as a result from a thorough preliminarily investigation and reveal the contemporary state of development of tensegrity bridges.
Jan De Boeck (2013) thinks that tensegrity appeared for the first time in proposals for bridges in 1996. Then Mott MacDonald submitted a conceptual project (Fig. 1) for the London's Millennium Bridge.

**Fig. 1.**
*Presentation of a concept for tensegrity bridge in London*
*Source: (Boeck, 2013, p. 19)*

A step towards a tensegrity bridge was taken in 1998 in Purmerend, the Netherlands as noted by Boeck (2013). It was designed by the architect Jord den Hollander. The bridge (Fig. 2) is used by pedestrians to cross the river. The whole structure comprises 18 spans of 4 m.

**Fig. 2.**
*Bridge in Purmerend*
*Source: (Boeck, 2013, p. 19)*

Andrea Micheletti (2012) together with a research group planned the Tor Vergata footbridge (Fig. 3) in 2005. The constructional project was built near the Faculty of Engineering of the University of Tor Vergata in Rome. It allows pedestrians to cross safely one of the main traffic arteries of the campus. The footbridge is a new, nontrivial example of a modular tensegrity structure. It is composed of five equal modules. The whole structure spans a distance of L=32 m without intermediate supports. Its main axis has a banana shape. The maximum slope at the end modules is 5%.




a) b) c)

**Fig. 3.**
*Tor Vergata footbridge*
*a), b) - Two views of the footbridge module, c) – Side view of the footbridge structure.*
*Source: (Micheletti, 2012, p. 5, 10)*

The Kurilpa Bridge in Brisbane, Australia (Fig. 4) is the largest tensegrity-inspired bridge in the world. It exhibits some tensegrity members, but clearly it isn’t a pure tensegrity structure as noted by Beck and Cooper (2012). The pedestrian and cycle bridge was opened in 2009. It was designed by Cox Rayner Architects and Arup. The constructional project connects Kurilpa Point in South Brisbane to Tank Street in the Brisbane central business district.

The tensegrity bridge comprises three spans – a central span of 128 m and side spans of 57 m and 45 m respectively as reported by ARUP (2010).

**Fig. 4.**
*Kurilpa bridge*
*Source: (http://tensegrity.wikispaces.com/Kurilpa+Bridge)*

In 2010 Barbarigos et al. (2010) proposed a tensegrity "hollow rope" structure for a pedestrian bridge (Fig. 5) to be built in the French speaking part of Switzerland. It spans 20 m over a river and is composed of four identical tensegrity modules. In the paper "Design optimization and dynamic analysis of a tensegrity-based footbridge" Ali et al. (2010) investigated a footbridge that is composed of six ring-shaped tensegrity modules to span 21.6 m.

**Fig. 5.**
*Tensegrity "hollow rope" pedestrian bridge*
*Source: (Barbarigos et al., 2010, p. 1159)*

Briseghella et al. (2010) designed a tensegrity footbridge with an arch deck (Fig. 6). The bridge opening is 30 m.

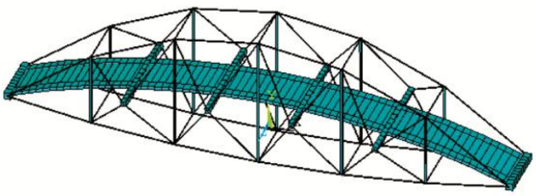
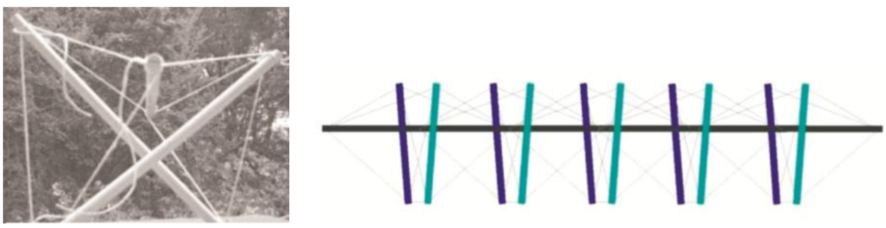


**Fig. 6.**
*Tensegrity footbridge with an arch deck*
*Source: (Briseghella et al., 2010, p. 288)*

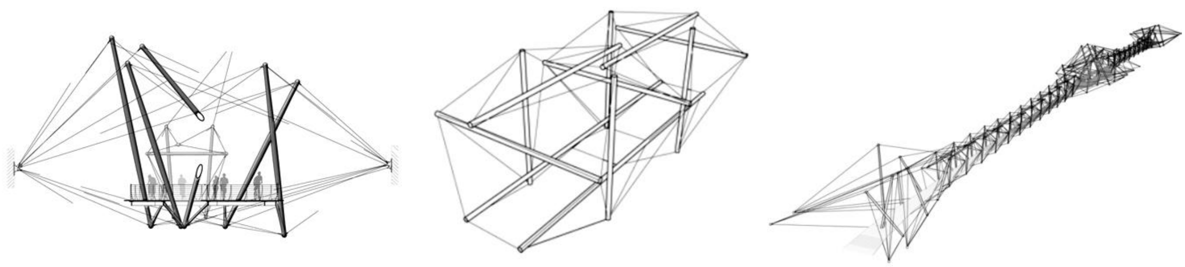
Jan De Boeck (2013) created a bridge like structure comprising a set of 3 strut tensegrity modules, juxtaposed next to each other (Fig. 7). They have a span of 12 m.



a) b)

**Fig. 7.**
*Jan De Boeck's tensegrity bridge*
*a) - Front view of a rotated model of a 3 strut tensegrity module, b) - Sketch illustrating how a 3 strut tensegrity module is multiplied and connected to a bridge like structure.*
*Source: (Jan De Boeck, 2013, p. 65, 67)*

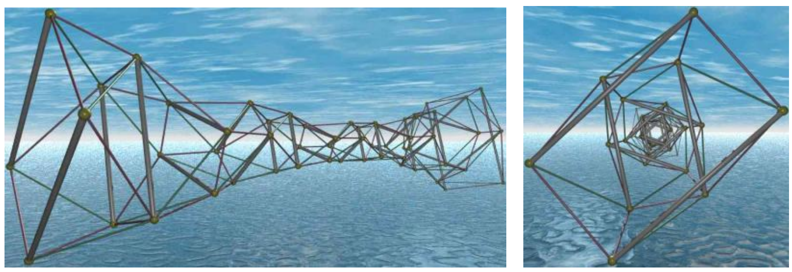
The "Suspended Tensegrity Bridge" (Fig. 8) is designed by Stefano Paradiso and Marco Mucedola (2010). The project concerns a footbridge over the Sesia river, not far from Greggio, Italy. It is situated in the space between the high speed train bridge and the A4 highway bridge and is hung on them. There are no foundation piles.



a) b) c)

**Fig. 8.**
*"Suspended Tensegrity Bridge"*
*a) - Vertical cross section, b) - Model of two tensegrity modules, c) - Sketch illustrating the whole tensegrity bridge.*
*Source: (http://tensegrity.wikispaces.com/Suspended+Bridge+by+Paradiso)*

Tim Tyler's conceptual project for a twisting hexagonal bridge is represented in Fig. 9.



a) b)

**Fig. 9.**
*Tim Tyler's conceptual project for a twisting hexagonal bridge*
*a), b) – Different views of the model*
*Source: (http://hexdome.com/bridges)*

**4. Analysis**

Tensegrity bridges can be assembled of modules that are juxtaposed next to each other along an axis. For the form-finding step of the design process it is necessary to identify the initial topology of the basic constructive tensegrity unit and its initial self-equilibrated self-stress.

The module chosen for the Tor Vergata footbridge (Fig. 3) represents a modification of the so-called expanded octahedron. Tensegrity modules, composed of three pairs of compression members that are mutually orthogonal, were first discovered by Emmerich as noted by Ilieva (2014 b). The structure is composed of 6 struts and 24 cables. In Tor Vergata footbridge four of the parallel struts are vertical. The module enjoys a wide "cross-sectional" space and the smallest number of mechanisms that stiffer the module as stated by Micheletti (2012).

Barbarigos et al. (2010) and Ali et al. (2010) applied the concept of "hollow rope" for their pedestrian bridges (Fig. 5). In their patents Fuller, Emmerich, Snelson, Wemyss and Liapi described one and the same tensegrity module that consists of three struts, spirally twisted as noted by Ilieva (2014 b). These tensegrity systems have important open spaces at their centers. Using their topology, Motro, Maurin and Silvestri (2006) discovered a new family of tensegrity cells, classified by them as "ring modules", because of their hollow tube shape. They also explained a general method for creating these deployable tensegrity cells, starting from any n-sided prism. Module's nodes are situated in 3 parallel layers. The described units consist of diagonal struts, intermediate struts and two kinds of cables: layer cables connecting the base nodes of the straight prism and x-cables, embracing laterally the structure. Design results illustrate that the hollow-rope tensegrity bridge can efficiently meet typical design criteria as concluded by Barbarigos et al. (2010). They examined the structural efficiency index of three module configurations: square, pentagon and hexagon. The results show that that the most efficient module is the pentagon ring and the least efficient is the hexagon.

The tensegrity footbridge with an arch deck (Fig. 6) is composed of four intermediate and two end modules. The intermediate module consists of one horizontal transverse beam, four vertical struts and supporting cables. The end module comprises one horizontal transverse beam, two vertical struts and cables. The cable support system of the bridge includes two upper main cables, two bottom main cables, diagonal cables and transverse cables. The diagonal cables connect strut ends with transverse beams. The transverse cables join the ends of symmetric struts.

The four horizontal transverse beams pull diagonal cables outward to limit transverse and longitudinal displacements of columns and main cables as stated by Briseghella et al. (2010). Besides tensegrity support system, the footbridge also includes a deck and foundations.

Jan De Boeck's bridge like structure (Fig. 7) utilizes the simplest 3 strut tensegrity module. It is positioned in such a way that two of the struts are supporting and the third one is horizontal.

Stefano Paradiso and Marco Mucedola designed their "Suspended Tensegrity Bridge" (Fig. 8) by applying irregular and regular tensegrity modules which are assembled in a sequential order. The regular tensegrity module like the Tor Vergata footbridge (Fig. 3) is a modification of the so-called expanded octahedron. It is composed of two pairs of vertical and transverse struts and three longitudinal compression members.

The tensegrity bridge in Purmerend (Fig. 2) comprises 36 masts, cables and a deck. There is no direct contact between the masts and the deck. The pedestrian deck is hung on the struts by cables.

In structural aspect the Kurilpa Bridge (Fig. 4) is a combination between a conventional cable-stayed bridge and a tensegrity structure. The main masts on the concrete pylons and the major cables are its primary structural elements. Secondary masts are used to ensure additional deck support. Tensegrity principles are applied only to the horizontal spars. Instead of a single mast, each concrete pylon has four asymmetrically opposed stick-like masts that are in good harmonization with the tensegrity members. Cox Rayner Architects and Arup designed a hybrid bridge structure which is very strong. Haig Beck and Jackie Cooper (2012) noted that tensegrity struts and cables work in unison with the major and minor masts and cables to resist twisting and lateral forces arising from patch loads on the deck, wind loads and earthquake loads. Tensegrity elements also provide support for the sun shading canopy.

According to the regulations, concerning the design criteria for footbridges, BD 29/04 (2004) the minimum headroom inside the enclosure shall be 2.3 m (for pedestrians only) and 2.4 m respectively (for pedestrians and cyclists). The minimum total width for a footpath and a cycle track on a bridge and ramps shall be 3.0 m (when they are segregated by a white line, colour contrast or surface texture – pedestrian path 1.5 m and cycle path 1.5 m) and 2.0 m (when they are unsegregated). The tensegrity modules that compose footbridges have to possess an empty space in their interior, sufficient to accommodate the deck and to be comfortable for the people. All models and realizations of tensegrity bridges, which are analysed in this article, meet the normative requirements. The Kurilpa Bridge (Fig. 4) has much larger clear width of its deck structure - 6.5 m.

The type of connection between adjacent modules into a unified structure is an important problem when designing tensegrity bridges. The elongated form of these construction objects determines the longitudinal assembly of their basic units. Strut-to-strut connections and application of supplementary cables are a very common way of joining tensegrity modules. Tor Vergata footbridge (Fig. 3), the tensegrity "hollow rope" pedestrian bridge (Fig. 5) and others can be given as such examples. As a whole, these structures are not pure tensegrity systems, but are composed of tensegrity modules. Nevertheless, when two longitudinal bars are connected strut-to-strut, they may be considered as a single element that has a straight, curved or zigzag form. Thus the compression members don't touch each other and these systems satisfy the requirements definition for tensegrity. As a result of the current analysis, it was also established that two adjacent modules can have common struts, thereby there is no duplication of compression elements (Fig. 6).

The stability, stiffness and load carrying capacity of tensegrity structures can be increased by using additional tensioned members. They impact on the geometrical and mechanical characteristics of the structure. Andrea Micheletti (2012) proposed two types of supplementary elements for Tor Vergata footbridge. V-cables are added laterally the structure to stiffen it against vertical actions. Cross elements on top and bottom sides of the bridge are used to stiffen the structure against horizontal and torsional actions.

The constructional behaviour of tensegrities depends on their shape. For example the main axis of the Tor Vergata footbridge (Fig. 3 c) has a banana form. A comparison with the case of a straight axis reveals that this solution increases the geometric stiffness and reduces by 40 % the maximum displacement as noted by Micheletti (2012). Briseghella et al. (2010) designed a tensegrity footbridge with an arch deck because in comparison with the straight deck it highly decreases positive in plane bending of its longitudinal beams, especially close to the bank, greatly decreases the deck deflection and provides further stiffness to the bridge. Deck prestressing is still used, in order to avoid that high tensile forces are transmitted from the cables to the embankment foundations.

Tensegrity footbridges are generally composed of steel hollow tubes and steel cables. The Tor Vergata footbridge (Fig. 3) is a specific example because its tensile members are of two kinds: cables and traction bars. Traction bars with the same ultimate tensile load have a bigger area and hence a higher stiffness compared with cables. For example two traction bars of 72 mm in diameter are found in each end module of Tor Vergata footbridge; traction bars of 50 mm in diameter are employed as elements composing the tensile paths; traction bars of smaller diameter are employed as cross elements under the deck. The remaining tensile elements are cables of 12 mm and 24 mm in diameter.

Tensegrity footbridges need a deck to allow the passage from one side to the other. It can be integrated in the structure or to be an independent component supported by the structure. We achieve reducing use of material and weight of the global structure when some of bridge's struts play the role of deck's structural elements. For their tensegrity "hollow rope" pedestrian bridge (Fig. 5) Barbarigos L. et al. (2010) provide a steel deck supported on the bottom nodes at both ends of each module. The tensegrity footbridge with an arch deck (Fig. 6) has a deck that is made of two longitudinal beams, two end transverse beams and four transverse beams supported by diagonal cables along the bridge span. A composite steel-concrete deck is adopted to facilitate construction and provide additional rigidity to the overall structure especially for out of plane bending. The thickness of its concrete slab is only 30 mm.

The deck of the bridge in Purmerend (Fig. 2) is suspended at the struts by cables. The 6.5 m clear width composite steel and concrete deck structure of the Kurilpa Bridge (Fig. 4) comprises a reinforced concrete slab made from full depth precast panels joined by in situ concrete stitches which are supported by and act compositely with a series of structural steel cross beams.
The above discussed examples of modular tensegrity footbridges span a distance between 20 m and 32 m without intermediate supports. The Kurilpa Bridge (Fig. 4) has a central opening of 128 m, but it is not a pure tensegrity structure.

## 5. Conclusion

Tensegrity footbridges are remarkable, lightweight and impressive in appearance and are strongly attracting the attention of architects and engineers nowadays. They are perceived as transparent structures. Systems of this type possess a high level of geometrical and structural efficiency. One of the main advantages of tensegrity bridges is their lightness and high strength to weight ratio compared with traditional ones. They utilize the greatest number of tension members and the least possible number of compression members. The former can be made considerably slender to withstand tensional forces than the latter to withstand compression forces. Thus less material and respectively less weight of material is used for bearing a given load. Tensegrity structures are spatial systems that distribute the applied forces into many members. The self-equilibrated and self-stress state among their tension and compression elements and the fact that they do not require anchorage in order to stand a given load make them extremely lightweight. Tensegrity bridges have some disadvantages such as structural design complexity, difficult and complicated way of assembly, high cost, large deformations, vibrations, etc. Although tensegrity systems can be designed to withstand heavy loads, their flexibility makes them to deflect considerably even under small loads. This explains why in bridge engineering they are used mainly for footbridges. The stability, stiffness and load carrying capacity of tensegrity structures can be increased by using additional tensioned members. Tensegrity bridges have big potential for further improvement in architectural and technical aspects in order to be more widely used for civil engineering purposes.

**References**


Ali, N. et al. 2010. Design optimization and dynamic analysis of a tensegrity-based footbridge, *Engineering Structures* 32: 3650–3659.

ARUP. 2010. Kurilpa Bridge, Brisbane. Available from internet: < http: //steel.org.au/media/File/Kurilpa_ Bridge_case_study. pdf>.

Barbarigos, L. et al. 2010. Designing tensegrity modules for pedestrian bridges, *Engineering Structures* 32: 1158-1167.
BD 29/04, 2004. *Design Criteria for Footbridges.* Volume 2, Section 2, Part 8. 17, 21 p.

Beck, H.; Cooper, J. 2012. *Kurilpa Bridge.* The Image Publishing Group, *Australia*. 24 p.

Boeck, J. D., 2013. *Tensegrity bridges, Concept design of pedestrian bridges using tensegrity as load carrying system*. Delft University of Technology, De Boeck Jan MSc. 19, 65, 67 p.
Fuller, B., 1962. *Tensile-integrity structures.* US Patent No. 3,063,521, 14 p.

Briseghella, B. et al., 2010. Tensegrity footbridges with arch deck: static and dynamic behavior. In *Proceedings of the ARCH'10 – 6th International Conference on Arch Bridges*, 287-294.

Georgiev, T., 1988. Defining and classification of the ways for prestressing of the steel structures. In *Proceedings of the Scientific Conference - Higher Military Construction School, Sofia, XI*, 251 - 257 .

Ilieva, Y., 2014 a. Genesis and development of tensegrity structures. In *Proceedings of the 14th International Scientific Conference VSU'2014*, volume III, 55 - 60.

Ilieva, Y., 2014 b. Innovative solutions for designing tensegrity structures. In *Proceedings of the International Scientific Conference DCB'2014*, (in press).

Micheletti, A. 2012. Modular tensegrity structures: the TorVergata footbridge, *Mechanics, Models and Methods in Civil Engineering,* LNACM 61: 375-384.

Motro, R., 2003. *Tensegrity: structural systems for the future*. Kogan Page Science, U.K. 19 p.

Motro, R.; Maurin, B.; Silvestri, C. 2006. Tensegrity Rings and the Hollow Rope. In *Proceedings of the IASS Symposium 2006, New Olympics, New Shells and Spatial Structures*. 470-475.

Paradiso, S.; Mucedola, M. 2010. Suspended bridge by Paradiso. Available from internet: <http://tensegrity.wikispaces.com/Suspended+Bridge+by+Paradiso>.

Skelton, R. E.; de Oliveira, M. C. 2009. *Tensegrity systems.* Springer.

Tyler, T. 2012. Twisting hexagonal bridge. Available from internet: < http://hexdome.com/bridges>.

# MARKET SHARE FORECASTING OF MOBILE TELEPHONY IN SERBIA

**Valentina Radojičić[1], Goran Marković[2], Vesna Radonjić Đogatović[3]**
[1,2,3] *University of Belgrade – Faculty of transport and traffic engineering*

**Abstract:** The analysis of mobile phone diffusion process has an important role for additional development of mobile networks. This paper provides results comparison of mostly used Bass diffusion model to forecast diffusion speed of mobile telephony subscriptions with two popular time series methods (Iterative and Direct methods) for predicting future values, which use explicit relationships between demand and various determining factors. Our goal is to estimate time to reach the saturation level at Serbian mobile market. The obtained results may be a useful managerial tool to analyze diffusion shape of the forthcoming generations of mobile telephony as well as for building appropriate strategic plans for necessary supporting network infrastructure in Serbia.

**Keywords**: Bass model, diffusion model, direct method, iterative method, forecasting, mobile telephony, network infrastructure.

## 1. Introduction

Mobile communications are the most dynamic sector in telecommunication industry. The global number of mobile-phone subscriptions has reached nearly 7 billion at the end of 2013. It is expected that the number of mobile subscriptions will continue to growth in the future. It is evident that mobile communication sector has powerful influence to economic growth in both developed and developing countries and mobile services have become an essential part of national economies. Mobile communications have changed consumer behaviours and transformed the way in which people connect and work, which has the potential to further impact on economic development. Long-term demand forecast is highly important for mobile operator investment decisions as well as for network dimensioning and expansions processes. The decision to adopt new telecommunication service may require considerable time and cognitive effort: however, most of the times, potential users do not have sufficient knowledge to evaluate the advantage implied by a new service or technology compared to investment's risks. Also, it is hard to acquire availability of new service for all users in the short term. The extension of mobile telephony service requires the provision of new generations of mobile phone. Therefore, a forecast will primarily produce accurate estimates of the future demand for these facilities (Fildes 2002; Decker and Gnibba-Yukawa 2010; Graefe and Armstrong 2011).

This paper investigates two approaches for predicting further development of mobile market. Numerous studies in a variety of disciplines suggest that cumulative sale of new service/technology during the time is an S-shaped curve (Mahajan et al. 1990; Mahajan et al. 2000; Krishnan et al. 2000; Geroski 2000; Kumar et al 2007; Michalakelis et al. 2008). Diffusion models have been traditionally used in marketing for capturing the life-cycle dynamics of a new service and as a decision aid in making pre-launch, launch and post-launch strategic choices (Radas 2005; Xiaoping et al. 2012). Since their entrance into marketing, diffusion models have become increasingly complex. This complexity has been driven by the need to enhance the forecasting capability of these models and to improve their usefulness as a decision-making tool for managers. Since in reality sales depend on a variety of external influences, such as the level of product advertising, changes in product price, Gross Domestic Products (GDP) and number of inhabitants, it is of great importance to create the models that can include such external variables. In this paper, we try to compare the Bass diffusion model with two most popular methods in forecasting time series, direct and iterative methods. Relevant literature considering these methods includes (Tiao and Xu 1993; Chevillon and Hendry 2005; Marcellino 2006). (Marcellino 2006) made comparisons between the direct and iterative methods, with the surprising conclusion that in practice the iterative method often performed better.

The rest of the paper is structured as follows. Section 2 describes diffusion models that could be used in theory for market modelling. Section 3 gives a short description of iterative and direct methods. Section 4 presents a short overview of mobile telecommunication market in Serbia. In Section 5, we provide the comparison results for mobile subscriptions forecasting and in Section 6 we performed the evaluations of the used models. Finally, the last section is left for concluding remarks.

## 2. Diffusion models for mobile market

The main purpose of diffusion models in marketing is to estimate new service/product sales after company's basic business decisions are made. Usually, marketing research methods are used for determining the parameters of diffusion model. Different assumptions and consequently different parameters used in these models are reflected to its mathematical form complexity. According to diffusion theory, a new service's sales growth at any time largely depends on the strength of word of mouth from its previous adopters. Similarly, a service sales growth should then depend on the extent to which it receives good word of mouth from its own previous adopters. An edited volume by (Mahajan et al. 2000) covers in depth various topics in diffusion models, such as estimation and applications.

The best known diffusion model used for technology diffusion purposes is the Bass model (Bass, 1969). The Bass diffusion model has been widely used as forecasting procedure of new services/technologies and it was proposed to deal with the problem of initial adopters. An excellent overview of the Bass model, its extensions, and some directions for further research are provided in (Mahajan et al. 1990).

The mathematical structure of the Bass model is derived from a hazard function corresponding to the conditional probability that an adoption will occur at time *t* given that it has not occurred yet. If *f(t)* is the density function of time to adoption and *F(t)* is the cumulative fraction of adopters at time *t*, the basic hazard function underlying the Bass model (Michalakelis et al. 2008) is given by:

$$\frac{f(t)}{1-F(t)} = p + q \cdot F(t) \tag{1}$$

From differential equation (1), with initial condition *F(0)=0*, it could be found the solution of purchase cumulative distribution function *F(t)*, cumulative adoptions *Y(t)*, and non-cumulative adoptions *S(t)* as follows:

$$F(t) = \int_{u=0}^{t} f(u)du = \frac{1-e^{-(p+q)t}}{1+\frac{q}{p}e^{-(p+q)t}} \tag{2}$$

$$Y(t) = mF(t) = m \cdot \frac{1-e^{-(p+q)t}}{1+\frac{q}{p}e^{-(p+q)t}} \tag{3}$$

$$S(t) = mf(t) = pm + (q-p)Y(t) - \frac{q}{m}Y^2(t) \tag{4}$$

This model has the following 3 key parameters: the parameter of innovation or external influence (*p*), the parameter of imitation or internal influence (*q*) and the market potential (*m*). Parameter *q* reflects the influence of those users who have already adopted new technology (i.e. word-of-mouth communication from previous adopters), while *p* captures the influence that is independent from the number of adopters (i.e. external communication). The sale at time *t* is *S(t)*, where *m* refers to the market potential for the new technology. The size of the market potential *m* is probably the most critical element in forecasting matters and a reliable estimation of it should be necessary.

Bass model could be used to predict the timing and magnitude of the peak sales and the shape of the diffusion curve. But, the most applications of the Bass model are used to make plans and decisions before the service/technology has been introduced to the market. Usually, no sales data exist that are required to estimate *p* and *q*. Managers do not have an intuitive estimate of *p* and *q*. In such a case, Bass parameters could be evaluated in two manners, usually. One way is to use analogies with other similar services or diffusion process. The second way is analytical using comparative procedure with some other countries where a service/technology already exists, using ordinary least squares (OLS) multiple regression (Michalakelis et al. 2008).

However, if historical data sales are available, the necessary parameters could be estimated by curve fitting procedure. Usually, the market potential should be estimated by taking into account different influence factors such as economy and social development of a particular area, presence of competitive technologies, operator's infrastructure investment strategies, etc.

## 3. Iterative and direct methods for mobile market forecasting

In (Buys et al. 2009) it is noticed that strong correlation between the number of mobile phones per inhabitant and the Gross Domestic Product (GDP) per capita exists. In (Andrianaivo and Kpodar 2012) and (Lee et al. 2009) the impact of mobile phone penetration on GDP per capita growth is investigated. It has been shown that countries with a proportionately higher share of 3G connections enjoy an improved GDP per capita growth compared to countries with comparable total mobile penetration but lower 3G penetrations (Jenny and Isaac 2010). Hence, in order to estimate the future mobile telephony demand, the expected general development and especially the economic development of a country has to be considered. In our research, it is performed using the most popular methods in forecasting time series, throughout direct and iterative methods. These methods use the explicit relationships between dependent variable (mobile phone demand) and various explanatory variables. The choice of explaining variables is dependent on availability of reliable data. The population growth and living standard could be the most common explanatory variables for estimation of mobile user penetration. Typical measure of living standard is Gross Domestic Product (GDP).

The iterative method is performed throughout several steps.

Step 1: Selection of the explanatory factors which have significant influence on the dependent variable.

Step 2: Calculation of the correlation coefficients between the dependent variable $Y$ and each of the explanatory factors.

Step 3: Selection of the explanatory factor with the highest correlation coefficient is made.

Step 4: For the explanatory factor $x_k$ chosen in step 3, the first iteration is performed. The parameters $a_1$ and $b_1$ are calculated by solving the following system of equations:

$$a_1 n + b_1 \sum_{i=1}^{n} x_{ki} = \sum_{i=1}^{n} Y_i \tag{5}$$

$$a_1 \sum_{i=1}^{n} x_{ki} + b_1 \sum_{i=1}^{n} x_{ki}^2 = \sum_{i=1}^{n} x_{ki} Y_i \tag{6}$$

where $n$ is the number of statistical data.

Based on that, a regression function is determined:

$$Y_1 = a_1 + b_1 x_k \tag{7}$$

Step 5: Residuals are calculated in the following way:

$$r_1 = Y_1 - (a_1 + b_1 X_1) \tag{8}$$

Step 6: Calculation of the correlation coefficients between the dependent variable $Y$ and remaining explanatory factors.

Step 7: Selection of the explanatory factor with the highest correlation coefficient is made (based on calculations in step 6).

Step 8: For the explanatory factor $x_j$ chosen in step 7, the second iteration is performed. The parameters $a_2$ and $b_2$ are calculated by solving the following system of equations:

$$a_2 n + b_2 \sum_{i=1}^{n} x_{ji} = \sum_{i=1}^{n} r_{1i} \tag{9}$$

$$a_2 \sum_{i=1}^{n} x_{ji} + b_1 \sum_{i=1}^{n} x_{ji}^2 = \sum_{i=1}^{n} r_{1i} \tag{10}$$

Based on that, a regression function is determined:

$$r_2 = a_2 + b_2 x_j \tag{11}$$

Step 9: Based on the regression functions from the first and second iteration, the forecasting formula is determined:

$$Y_2 = Y_1 + r_2 = a_1 + a_2 + b_1 x_k + b_2 x_j \tag{12}$$

Step 10: Iterations are performed for all remaining explanatory factors. Number of iterations depends on the number of chosen explanatory factors.

In direct method forecasting formula is derived directly from a system of equations. The number of equations depends on a number of chosen explanatory factors. Generally, this model assumes that the variable $Y$ depends on $m$ explanatory factors:

$$y = a + b_1 x_1 + \cdots + b_m x_m \tag{13}$$

$$na + b_1 \sum_{i=1}^{n} x_{1i} + \cdots + b_m \sum_{i=1}^{n} x_{mi} = \sum_{i=1}^{n} y_i \tag{14}$$

$$a \sum_{i=1}^{n} x_{1i} + b_1 \sum_{i=1}^{n} x_{1i}^2 + b_2 \sum_{i=1}^{n} x_{1i} x_{2i} + \cdots + b_m \sum_{i=1}^{n} x_{1i} x_{mi} = \sum_{i=1}^{n} x_{1i} y_i \tag{15}$$

$$\ldots$$

$$a \sum_{i=1}^{n} x_{mi} + b_1 \sum_{i=1}^{n} x_{1i} x_{mi} + b_2 \sum_{i=1}^{n} x_{2i} x_{mi} + \cdots + b_m \sum_{i=1}^{n} x_{mi}^2 = \sum_{i=1}^{n} x_{mi} y_i \tag{16}$$

where parameters $a$, $b_1$, $b_2$, ..., $b_m$ have to be calculated.

It is well known that development over time for a mobile services growth could be described through an S-shaped curve and this type of curve has to be used only for long-time forecasting. However, it is possible to perform the more accurate short-time forecast of each phase of service life cycle separately. This statement gives

us opportunities to additionally applied these two models for prediction the time to reach the market`s saturation level.

**4. Mobile communication market in Serbia: a short overview**

The mobile telecommunication market in Serbia is under the regulation by national Republic Agency for Electronic Communications (RATEL 2012). Currently, there are 3 competitive mobile operators at the telecommunication market in Serbia (Telekom Srbija, Telenor and Vip mobile). They have started their operation in 1998, 2006 and 2007, respectively. However, the beginning of mobile telephony in Serbia was even before, in 1995, with introducing analog NMT network by operator Mobtel Srbija, which has been thereafter sold to Telenor in 2006.

During the last years, mobile market in Serbia has been rapidly increased, from about 3,3 million of users in 2003 to more than 10 million of subscriptions in 2011, which is compound annual growth rate (CAGR) of about 20% (RATEL 2012). The penetration rate or mobile density in 2012 was 126.19% and the total number of active subscriptions was about 9,14 millions.

Obviously, total number of mobile subscriptions significantly exceeds the number of inhabitants due to utilization of multiple subscriptions per mobile user. There are various reasons such as using different tariff packets by individual user, double subscriptions for private and business purpose, multiple subscriptions due to SIM card locked by mobile operator etc.

Total revenues from mobile market in 2012 were about 850 million of Euros (RATEL 2012). In 2012 total outgoing traffic on the mobile network amounted to 11.72 billion minutes of calls. The annual average of traffic per user in 2012 was 1 282 minutes or approximately 3 minutes and 30 seconds daily. Some relevant indices for mobile market in Serbia are summarized in Table 1.

**Table 1**
*Mobile phone market`s indicators*

| Year | Total number of mobile subscriptions [millions] | Total revenues [million of Euros] | Total outgoing traffic [million of minutes] | Mobile density [subscriptions per 100 inhabitants] | Prepaid [%] | Postpaid [%] |
|---|---|---|---|---|---|---|
| 2003 | 3.28 | 255.50 | 827 | 44.52 | 91 | 9 |
| 2004 | 4.34 | 295.50 | 1176 | 57.94 | 89 | 11 |
| 2005 | 5.51 | 400.00 | 1823 | 73.50 | 87 | 13 |
| 2006 | 6.64 | 581.50 | 2594 | 88.60 | 84 | 16 |
| 2007 | 8.45 | 838.84 | 3771 | 112.70 | 80 | 20 |
| 2008 | 9.62 | 913.43 | 5985 | 128.27 | 77 | 23 |
| 2009 | 9.91 | 826.74 | 8203 | 132.20 | 74 | 26 |
| 2010 | 9.92 | 769.20 | 8839 | 132.24 | 67 | 33 |
| 2011 | 10.18 | 846.70 | 10147 | 142.99 | 64 | 36 |
| 2012 | 9.14* | 850 | 11719 | 126.19* | 59 | 41 |

** number of active users in last 3 months of 2012*
*Source: RATEL -2012*

A dominant position at the Serbian mobile market has operator Telekom Srbija, followed by Telenor and lastly Vip mobile. Figure 1 illustrates historical market share of mobile operators in terms of the number of subscriptions.

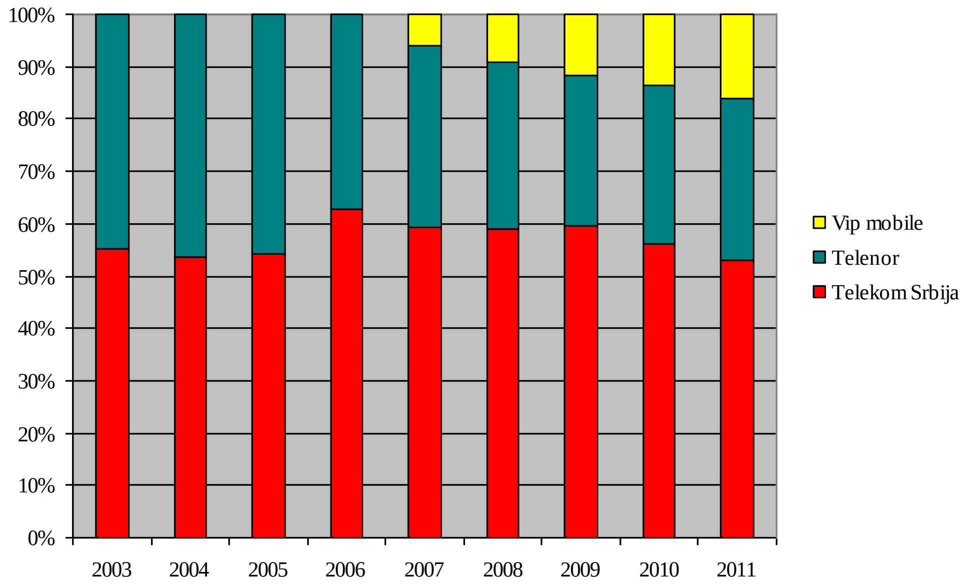



**Fig. 1.**
*Mobile communication market share between operators*

All existing operators have the licences for offering mobile network services in accordance with GSM/GSM1800 and UMTS/IMT-2000 standards.

## 5. Forecasted results for mobile telephony in Serbia Subtitle

In order to evaluate the diffusion process and the capability of the considered models to capture it, the actual penetration data of Serbian mobile market are used. Our goal is to forecast time to reach the saturation level which is estimated and recommended as reference value from ITU-T (170) (RATEL 2012). In our analysis, we used a data set consisting of the number of mobile phone subscriptions in Serbia given by the first column of Table 1. For earlier data, before the year 2003, we used the available historical data obtained from different sources, because they were not available in RATEL`s official annual report.
Firstly, we consider the results obtained by Bass diffusion model. In order to forecast the number of mobile subscriptions we have to estimate the Bass diffusion parameters *p*, *q* and *m*. The market potential *m* is estimated based on the total number of inhabitants and the expected number of mobile phone subscriptions per 100 inhabitants. Based on expected data the total number on inhabitants could not be changed noticeably in future short-term period compared to current value which is about 7.2 millions. According to the reference value for the mobile density in Republic of Serbia is assumed to be 170 (RATEL 2012), we estimated that the total market potential will be about 12.5 millions of subscriptions. This estimation considers only traditional mobile phone subscriptions (2G and 3G services), excluding potential future M2M applications. By assuming this value for market potential, the Bass parameters of innovation and imitations are obtained based on the available statistical data set for mobile subscriptions given by Table 1. After fitting the Bass curve to best capture the real data set for mobile subscriptions we obtained the following values of parameters p=0.006 and q=0.46. Based on these parameter values, we obtained the estimated values for the cumulative number of mobile subscriptions up to the year 2012 and forecasted number of subscriptions in future short term period (see Figure 2).

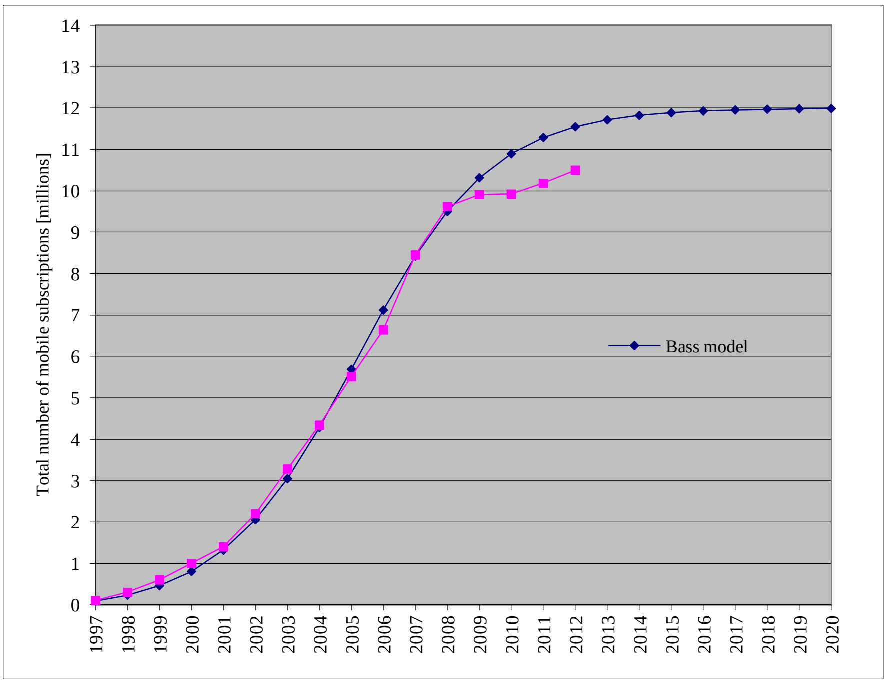



**Fig. 2.**
*Real data sales and Bass diffusion estimated values for mobile subscriptions in Serbia*

Because of noticed mismatching between the real data sales and Bass diffusion curve during several last year we assumed that various economic parameters have significant influence to diffusion speed of Serbian mobile market growth. It was the reason to choose direct and iterative forecasting methods that take into consideration these factors.
To perform iterative and direct methods we assumed two explanatory factors: GDP and the number of inhabitants. Historical data from 2001 to 2012 are used for forecasting because both explanatory factors show high deviations in period before 2001 due to former economic and political circumstances.

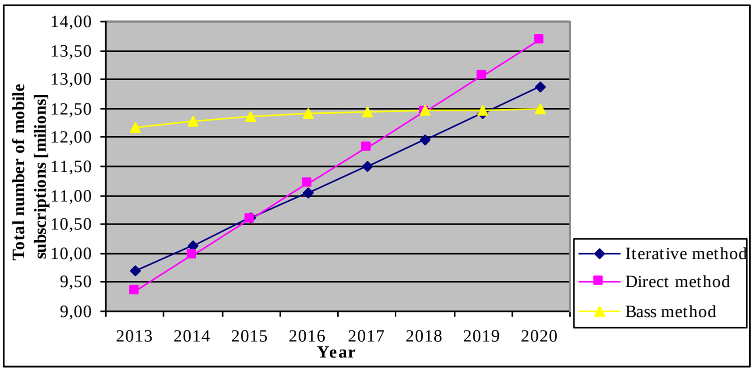



**Fig. 3.**
*Forecasted results comparison of mobile subscriptions obtained by Bass, Iterative and Direct methods*

Using the results shown by Figure 3, it could be seen that time to reach market`s saturation level according to Bass model, Direct method and Iterative method is 2017, 2018 and 2019, respectively.

## 6. Evaluation of forecasted results

In order to evaluate the precision of the obtained results, we performed the MAE (Mean Absolute Error), MAPE (Mean Absolute Percentage Error) and Durbin-Watson tests. Two standard error measures: mean absolute percentage error (MAPE) and the mean absolute error (MAE) are employed to quantify the deviation in the units of the data. As a rule, smaller the value of MAPE or MAE, better fit of the model is achieved. A scale to judge the accuracy of the model based on the MAPE measure as shown in Table 2.

**Table 2.**
*A scale of judgment of forecast accuracy (Lewis 1982)*

| MAPE | Judgment of Forecast Accuracy |
|---|---|
| Less than 10% | Highly Accurate |
| 11 to 20% | Good Forecast |
| 21 to 50% | Reasonable Forecast |
| 51% or more | Inaccurate Forecast |

Durbin-Watson test should be used for checking the existence of systematic errors (Montgomery et al. 2001). The test assumes calculating a value

$$DW = 2 - 2\frac{w}{v}, \tag{16}$$

where parameters $w$ and $v$ are following:

$$w = \sum_{i=1}^{n-1}(y_i - \hat{y}_i)(y_{i+1} - \hat{y}_{i+1}) \text{ and } v = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{17}$$

The Durbin-Watson (DW) statistic lies in the range 0-4. A value of nearly 2 indicates that there is no first-order autocorrelation. An acceptable range is 1.50 - 2.50. If the DW statistic gives unsatisfactory test result, suitable adjustments must be made to the data and the analysis should be repeated. The forecasting results evaluations are given by Table 3.

**Table 3.**
*Statistical indicators of results precision*

| Test | Bass model | Iterative method | *Direct method* |
|---|---|---|---|
| MAE | 0,412144366 | 0,923365875 | *0.804900358* |
| MAPE [%] | 8,975140872 | 18,92530842 | *14.98781653* |
| Durbin-Watson | 0.22 | 1.3 | *1.49* |
| *Number of observations* | 16* | 12* | *12** |

**For Bass model we used time series from 1997; For Iterative and Direct methods we used time series from 2001.*

The Bass model has a lowest forecasting error (MAE) and Iterative method has a highest error (MAE). The Bass model has highly accurate forecast (MAPE), but direct and iterative methods have good forecast (MAPE), too. Durbin-Watson test results are less than 1.5 for all applied forecasting models, which indicates the presence of positive autocorrelation.

According to forecasting results evaluations we found that the Bass model is more accurate than direct and iterative methods. However, because of high correlation between the mobile phone density and GDP as well as the number of inhabitants it could be concluded that direct and iterative methods should be applied to forecast the diffusion dynamic during the phase of rapid growth and time to reach market`s saturation level.

## 7. Conclusion

Long-term demand forecasts are of great importance for investment decisions, network dimensioning and service improvement. In this paper we have estimated future trend and analyzed the pattern and adoption rate of mobile subscribers in Serbia. We have analyzed forecasting results derived from Bass diffusion model as well as two regression models.

The diffusion models have proved pertinence in marketing for capturing the life-cycle dynamics of a new service as well as for decision support in making pre-launch, launch and post-launch strategic choices. The Bass model is the most widely used in the actual decision-making procedure, because of its simple parameters setting. However, in the case of Serbian market, it could be observed that certain mismatching between the real data and Bass diffusion values exists during the last period. We take into account that socio-economic factors have a strong impact on the diffusion rate at the mobile market. Therefore, we assumed that regression models with GDP and the number of inhabitants as explanatory factors could be used for modelling of mobile telecommunication market dynamic. For this purpose, we have applied iterative and direct methods with the explanatory factors previously defined.

The actual penetration data of Serbian mobile market has been used for evaluation diffusion and regression procedures. It is shown that time to reach the market`s saturation level in Serbia according to Bass, direct and iterative methods is estimated to be occurred in the years 2017, 2018 and 2019, respectively. For validation of considered forecasting models we have used standard statistical error measures (MAE and MAPE) as well as Durbin-Watson test.

These results may be exploited for analysing the diffusion shape of new generations of mobile telephony. In addition, appropriate strategic plans for necessary supporting network infrastructure in Serbia may rely on this or related research. Some aspects for further research should include the examination of the diffusion patterns between pre-paid and post-paid subscriptions as historical evidence reveals that users' choices often vary.

**Acknowledgement**

This paper is partially supported by the Serbian Ministry of Education, Science and Technological Development (research project: TR-32025).

**References**


Andrianaivo, M.; Kpodar, K. 2012. Mobile Phones, Financial Inclusion, and Growth, *Review of Economics and Institutions*, University of Perugia Electronic Press, 3(2), Article 4.

Bass, Frank M. 1969. A New Product Growth for Model Consumer Durables, *Management Science* 15(5): 215-227.

Buys, P.; Dasgupta, S.; Thomas, T. S; Wheeler, D. 2009. Determinants of a Digital Divide in Sub-Saharan Africa: A Spatial Econometric Analysis of Cell Phone Coverage, *World Development*, 37(9): 1494–1505.

Chevillon, G.; Hendry D.F. 2005. Non-parametric Direct Multi-Step Estimation for Forecasting Economic Processes, *International Journal of Forecasting*, Elsevier 21(2): 201-218.

Decker, R.; Gnibba-Yukawa K. 2010. Sales Forecasting in High-Technology Markets: A Utility-Based Approach, *Journal of Product Innovation Management* 27(1): 115-29.

Fildes, R. A. 2002. Telecommunications demand – a review, *International Journal of Forecasting* 18: 489-522

Geroski, P. A. 2000. Models of technology diffusion, *Research Policy*, Elsevier, 29(4/5): 603-625.

Graefe, A.; Armstrong, J. S. 2011. Comparing Face-to-Face Meetings, Nominal Groups, Delphi and Prediction Markets on an Estimation Task, *International Journal of Forecasting* 27(1): 183-195.

Lee, S. H.; Levendis, J.; Gutierrez, L. 2012. Telecommunications and Economic Growth: An Empirical Analysis of Sub-Saharan Africa, *Applied Economics* 44(4): 461/469.

Jenny, A.; Isaac, M. 2010. Mobile Phones and Economic Development in Africa, *Journal of Economic Perspectives* 24(3): 207-232.

Krishnan T. V.; Bass, F.M.; Kumar, V. 2000. Impact of a Late Entrant on the Diffusion of a New Product/Service, *Journal of Marketing Research*, 37(2): 269–278.

Kumar, A., Baisya, R.K; Shankar, R.; Momaya, K. 2007. Diffusion of Mobile Communications: Application of Bass Diffusion Model to BRIC Countries, *Journal of Scientific& Industrial Research* 66(4): 312-316.

Lewis, C. D. 1982. *International and Business Forecasting Methods*, London: Butterworths.

Mahajan, V.; Muller, E.; Bass, F.M. 1990. New Product Diffusion Models in Marketing: A Review and Directions for Research, *Journal of Marketing* 54(1): pp. 1-26.

Mahajan, V.; Muller, E.; Wind, Y. 2000. New-Product Diffusion Models: From Theory to Practice", in *New-Product Diffusion Models*, eds Mahajan, V.; Muller, E. and Wind, Y., Kluwer Academic Publishers, Boston, p 3-24.

Marcellino, M.; Stock, J.H.; Watson, M.W. 2006. A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Time Series, *Journal of Econometrics* 135: 499-526.

Michalakelis, C.; Dimitris, V.; Thomas, S. 2008. Diffusion models of mobile telephony in Greece, *Telecommunications policy*, 32(3-4): 234-245.

Montgomery, D. C.; Peck, E. A. and Vining, G. G. 2001. Introduction to Linear Regression Analysis. 3rd Edition, New York, New York: John Wiley & Sons.

Radas, S. 2005. Diffusion Models in Marketing: How to Incorporate the Effect of External Influence, *Privredna kretanja i ekonomska politika*, 15(105): 30-51.

RATEL: An Overview of Telecom Market in the Republic of Serbia in 2012. Available from internet: <http://www.ratel.rs/upload/documents/Pregled_trzista/An_Overview_of_Telecom_Market_in_the_Republic_of_Serbia_in_2012.pdf >

Tiao, G. C.; Daming X. 1993. Robustness of maximum likelihood estimates for multi-step predictions: the exponential smoothing case, *Biometrika* 80(3): 623–641.

Xiaoping, L.; Linjie, L.; Yi, L. 2012. Research of Forecasting the Number of 3G Users Based the BASS Model, In *Proceedings of the 2nd International Conference on Computer Application and System Modeling*, 1095-1098.

# EVALUATION METHODS FOR SPECIFIC PROJECTS CASE STUDY OF TOURISTIC RESORTS ROAD NETWORK

**Draženko Glavić**[1], **Vladan Tubić**[2], **Marijo Vidas**[3]
[1,2,3]*University of Belgrade-Faculty of Transport and Traffic Engineering, Belgrade,*

**Abstract:** The Feasibility study for touristic resort low volume road network is an atypical problem for evaluation. Cost Benefit Analysis-CBA, traffic forecast, and traffic analysis for this purpose need to be modified. Road network of touristic resort is a low volume road network and primarily serves to meet the transportation needs of the touristic resort visitors and to allow the uninterrupted operation of the touristic resort. On this type of network only induced or generated traffic exist. Existence of other types of traffic such as normal and diverted are almost negligible. This fact imposes a special approach in traffic forecasting. The paper presents theoretical and methodical support for CBA evaluation of low volume road network of touristic resorts and presents - two case studies i.e. for touristic resort Golija and touristic resort Stara planina.

**Keywords**: Low volume touristic resort road network, feasibility study, traffic forecast, CBA.

## 1. Introduction

Preparation of project documentation on the stage of preliminary or conceptual design requires the parallel creation of Pre-Feasibility and Feasibility Study.

Usual methodology for conducting feasibility study applied to the classical road construction are not applicable in the case of specific projects. In the case of low volume road network of touristic resorts, it is necessary to modify and adjust methodology to specific characteristics of the project.

The case of evaluation of low volume road network of touristic resorts is specific project that differ from construction of road or city bypass or from other typical road projects i.e. reconstruction or rehabilitation. To evaluate tourist resorts road network a series of new methodological approaches needs to be taken to determine the functional, economic and financial indicators and criteria.

This paper, will briefly show the specifics of evaluation on the two case studies (Glavic, 2009) and (Glavic, 2010). The specifics of analysis are:

- Analysis and forecast of traffic flows,
- Level of Service and capacity analysis,
- Cost-Benefit Analysis.

## 2. Analysis and forecast of traffic flows

Traffic is composed of various components. The three basic components according to (Kuzovic, et al., 2010) are normal, diverted and generated or induced traffic. The normal traffic is in a function of the natural-normal growth of influencing factors on traffic forecast (GDP, motorization, population, mobility). In the case of a tourist resort network in existing situation (resort and resort road network do not exist) there is no any kind of activity, so normal and diverted traffic are equal to zero. Therefore, for the purposes of the feasibility study of the tourist resort road network the traffic forecasts mostly is in a function of generated or induced traffic. Normal and diverted traffic usually do not exist.

Although, if in the perspective the considered network will become a part of a wider road network that connects and other places and not just resort, then normal and diverted traffic can be expected. If the considered network has only one function, and that is serving the future resort, then only induced traffic will exist.

### 2.1. Forecast of induced AADT

Due to the lack of any information bases, such as published publications about traffic counting data or other studies as a sources, for purposes of determining the AADT in base year (Kuzovic, et al., 2007), it had to be accessed directly detailed analysis of the capacity of each functional object and facility of the future resort.

The consequence of the above is that the only source of data on the basis of which is determined AADT for base year is Plan of detailed regulation and studies and designs documentation of the tourist resorts. Within this documents are shown all future facilities with activities and functions of the resort.

[1] Corresponding author: drazen@sf.bg.ac.rs

### 2.2. Analysis of functional facilities capacity of the future resort

Tourist resort is designed as a multi-purpose, integrated complex with a program focusing on the exclusive, highly commercial tourism (Glavic, 2009) and (Glavic, 2010). Facilities for recreation and sport, public services and utilities, transport and technical infrastructure are also provided. A specific feature of resort will provide a planned sports and recreation, entertainment, commercial and cultural facilities. Functional activities and capacities are:

• Tourist beds in total:
  - Main hotel with convention center;
  - Small hotels;
  - Apartment complex ;
  - Multi-family apartments;
  - Single-family houses ;
  - Motel ;

• Daily visitors, of which:
  - skiers
  - non-skiers

• Official beds

• Employees

• Different agencies and services in the resort complex include:
  - Special services - mountain rescue service,
  - Fire service,
  - Service for road maintenance,
  - Traffic-bus station and parking lots,
  - Organized public transport;
  - Craft services - technical base of ski resorts (workshops, warehouses, garages, etc.),
  - Services for sports equipment,
  - Catering, repairs, maintenance facilities and utility equipment,
  - Health-clinic and pharmacy,
  - Welfare-kindergarten;
  - Culture and entertainment center (intended for the presentation of the values of the Nature Park, Nature Park Museum, presentations and recreation facilities in the area, education programs, research desk, gift shop, internet center, and others.),
  - Administration - central reception, the travel agencies, banks, post office, police station, etc.

• Shopping facilities include: commercial services, outlets, food and drinks, which will usually be placed within the objects (ground level in hotels and in tourist "villages").

• For sports and recreational facilities, the complex of tourist resorts are planned various activities such as: special summer recreation and entertainment complex with pool and waterslide for adults and children, tennis and squash courts, bowling, mini-golf, playground, etc.; in winter are planned lifts for the ski school (ski area for beginners and games in the snow) children's sleighs, ice rinks for children and adults; the spa and recreation center (with wellness treatments. Within the complex are provide the starting point / terminals and transit checkpoints for lifts, alpine, snowboard and Nordic ski trails (in the winter), or excursions and hiking trails (in the summer).

Generated or induced traffic is consequence of the development of entirely new activities or increase existing activities. In the case of the mountain resort all the generated traffic is related to the development of entirely new activities. New activities caused by tourism activities are consequences of:
- Winter tourism (hotel accommodation and daily tourism)
- Summer tourism (hotel accommodation and daily tourism)
- Transport of Employees
- Logistics needs of the resort (delivering food and other supplies, garbage collection, resort maintenance etc.).

Based on these functional elements and its capacity according to the following categories:
• Tourist beds
• Official bearings
• Day-trippers
• Employees

The total number of resorts simultaneous users is determined. This data is the starting point for determination of AADT for base year.

Based on these functional content, capacity of resorts facilities, next step is the preparation of the table of influencing factors for traffic forecast. An example of such a table is given below:

**Table 1**
*The influential parameters on the prognosis of future trends.*

| **Resort functions** | Vehicle Type | **Distribution** | **Occupancy** | **Capacity utilization** | **Period** |
|---|---|---|---|---|---|
| Tourist beds<br>Official bearings<br>Day-trippers<br>Employees<br>Logistical needs | PC | - | - | - | - |
| | BUS | - | - | - | - |
| | HT | - | - | - | - |

*Source: Authors*

### 2.3. The traffic growth rates

For the average annual growth rate of traffic, the plans for future development of the tourist resort is consulted, as well as the value of the growth rate in other relevant similar studies in region, together with data on GDP, mobility, motorization, the gravitational area, the number of inhabitants and other influencing factors. The forecast is done for the best case and worst case scenario, such as the inclusion of negative factors of the impact of the global economic crisis and/or low occupancy of touristic resorts.

### 2.4. Total traffic flows

As previously mentioned, if the roads within the tourist resorts will become part of a future municipal or state road, there is the potential for normal and diverted traffic. Since this fact is not a typical case for tourism resort's then normal traffic and diverted traffic is not considered, and the total traffic flows equals the induced traffic (Kuzovic, et al., 2010).

Using the determined AADT for base year and traffic grow rates, the traffic forecast is conducted (Kuzovic, et al., 2007).

## 3. Analysis of the capacity and level of service

Methodology for determining the quality of transport services, as an integral part of the Guidelines for the preparation of feasibility studies of roads, provides values of certain indicators of the level of service (LOS). The relevant procedure is carried out in several steps in order to assess the traffic conditions for the traffic flows on the new resorts road network.
Traffic conditions are evaluated in accordance with a HCM-2000 (TRB, 2000), based on the elements of the situational plan and longitudinal profile, as well as the characteristics of traffic.

Results of the analysis must indicate at high level of service for tourist resorts users. High LOS is goal since the purpose of this type of road network is to have satisfied costumer, i.e. satisfied traveler on resort road network.

## 4. Cost-benefit analysis

For the analysis of the economic benefits (HEATCO, 2006) it was necessary to collect certain information and input data from following sources:
- Technical and operational characteristics were obtained from the preliminary design.
- Data on traffic flows are derived from forecasts of traffic flows.
- Data on vehicle speeds are determined in the traffic analysis.
- Data from the results of field investigations and interviews, which was conducted in the relevant studies,
- Data from the HDM-4;
- Data from the vehicle manufacturer specifications,
- Data on market prices of vehicles, tires, spare parts, fuel, salaries, etc .

After defining the input data, next step is determination of the direct benefits per road sections. Benefits were determined on the basis of differences in the Road Users Costs-RUC, between modern asphalt road conditions and gravel road conditions.

1) For the calculation of benefits arising from the Vehicles Operation Costs VOC, the HDM-4 model is used (PIARC, 2000). The Basic operational and economic parameters (prices) are adapted to local conditions.
2) To calculate the travel time cost TTC in the 20-year period direct analysis is applied.
3) For the calculation of the expected direct economic benefits, difference between the RUC of network without investment and network with investments in the 20-year initial period of exploitation is determined (UNECE, 2003).
4) The economic construction costs were determined as the financial price less VAT for the project (WB, 2005).
5) Indicators for evaluation from socio-economic aspects EIRR and ENPV were determined using the appropriate software.
6) Indicators EIRR and ENPV are subjected to the sensitivity analysis with respect to possible variations in the economic construction costs and economic benefits (EC DG Regional policy, 2002).
7) The justification, from the socio-economic aspect, is determined by comparing the values EIRR with OCC and comparing the values ENPV with zero.

Analysis of the following RUC of the considered road networks are conducted

1) Economic vehicles operation costs VOC
2) Economic travel time costs TTC for passengers in cars and buses.

### 5. Case studies

Following two case studies are presentation of above described methodologies

#### 5.1. Case study for Golija touristic resort road network (Glavic, 2009)

The following figure shows planned network of roads for Golija touristic resort.

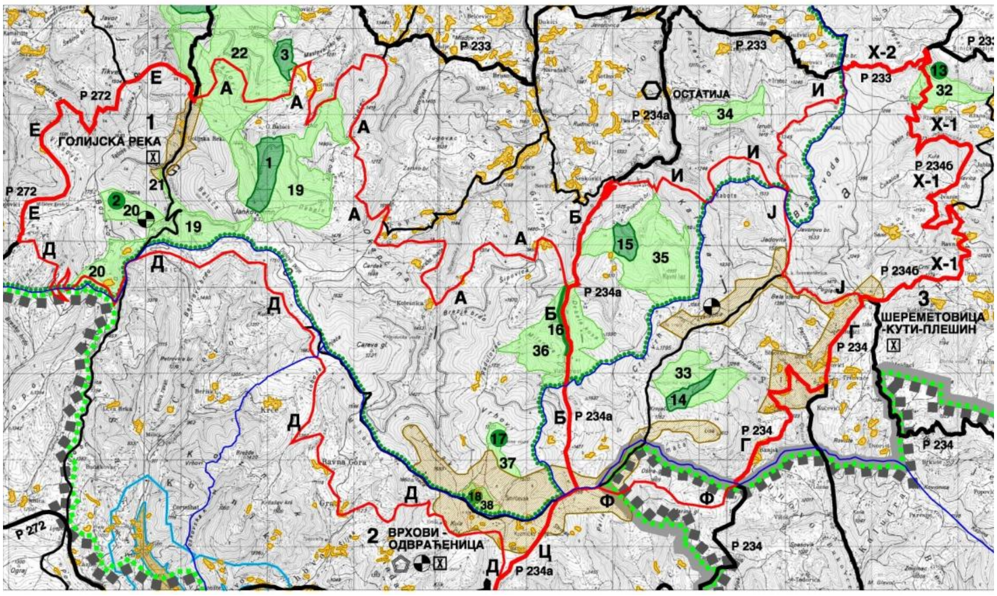



**Fig. 1:**
*Planned road network of the Golija touristc resorts*
*Source: Spatial Plan - "tourist road Golija")*

1) Traffic flows per years of the phase realization of the project

| Years | PC | BUS | HV | TOTAL |
|---|---|---|---|---|
| 2012 | 363 | 12 | 35 | 409 |
| 2014 | 1102 | 55 | 48 | 1205 |
| 2019 | 1704 | 88 | 61 | 1853 |
| 2031 | 2429 | 126 | 84 | 2640 |

2) Basic information about the project

| | |
|---|---|
| Design speed | 60km/h |
| The width of the traffic lanes | 2x3.25m |
| Shoulder width | 1.50m |
| Sidewalk width | 1.50m |
| Minimum horizontal curve radius | 30m |
| Minimum cross slope | 2.5% |
| The maximum cross slope | 4% |
| Maximum gradients as steep as | 10% |
| The minimum vertical curve radius | 250m |

3) The costs of the investment and investment dynamics per years

| Financial construction costs (RSD) | Economic construction costs (RSD) | Investment dynamics per years (RSD) | |
|---|---|---|---|
| 4,817,305,000 | 3,853,844,000 | 2009 | 963,461,000 |
| | | 2010 | 1,348,845,400 |
| | | 2011 | 1,541,537,600 |

4) The expected economic benefits

The expected benefits obtained after the economic evaluation of the analyzed networks and analysis of their operating costs are:

| Economic Benefits (RSD) | |
|---|---|
| 2012 year | 2031 year |
| 133,281,381 | 745,524,241 |

5) Value of the basic of economic evaluation indicator EIRR is

- EIRR = 8.37%

6) An assessment of the economic feasibility of investing from the socio-economic aspects

Investment in the project with respect to the values of obtained EIRR for the project of Golija touristic resorts low volume road network has a moderate economic justification.

7) Conclusions and recommendations of the Feasibility Study

Functional evaluation demonstrate need for the implementation of this project. The economic evaluation shows economic justification for the investment. It is recommended to start with constructing the road network for touristic resort Golija.

**5.2. Case study of Stara planina touristic resort road network (Glavic, 2010)**

The following figure show planned network of roads for Stara planina touristic resort.

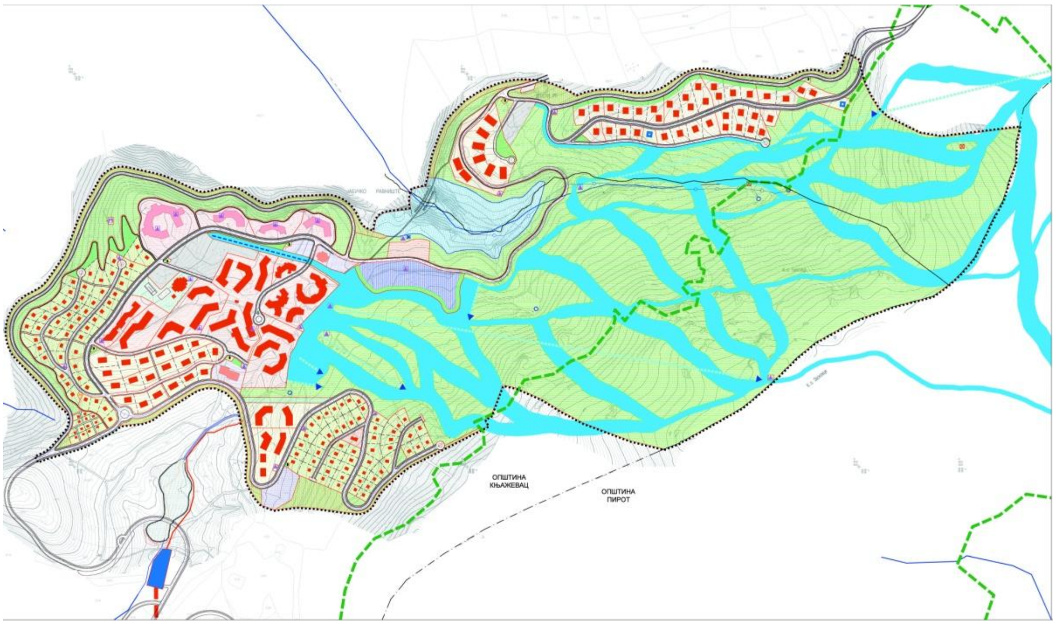



**Fig. 2:**
*Planned road network of the Stara planina touristc resorts*
*Source: Spatial Plan - " Stara planina touristc resorts "*

1) Traffic flows per years for sections S1, S2, S3

| Year | AADT (vpd) | | | | | |
|---|---|---|---|---|---|---|
| | P C | BUS | L T | MT | HT | TOTAL |
| 2012 | 747 | 51 | 36 | 103 | 5 | 942 |
| 2031 | 1,408 | 81 | 57 | 163 | 8 | 1,717 |

2) Basic information about the project

| | |
|---|---|
| Design speed | 40km/h |
| The width of the traffic lanes | 2x3.5m |
| Shoulder width | 1.50m |
| Sidewalk width | 1.50m |
| Minimum horizontal curve radius | 45m |
| Minimum cross slope | 2.5% |
| The maximum cross slope | 4% |
| Maximum gradients as steep as | 10% |
| The minimum vertical curve radius | 250m |

3) The costs of the investment and investment dynamics per years

| Financial construction costs (RSD) | Economic construction costs (RSD) | Investment dynamics per years (RSD) | |
|---|---|---|---|
| 182,385,000 | 149,555,700 | 2011 | 149,555,700 |

4) The expected economic benefits

The expected benefits obtained after the economic evaluation of the analyzed networks and analysis of operating costs are:

| Economic Benefits (RSD) | |
|---|---|
| 2012 year | 2031 year |
| 27,763,723 | 53,983,809 |

5) Values of the basic parameters of economic evaluation

- EIRR = 21.51%
- ENPV= 170,643,101 (RSD)

6) An assessment of the economic feasibility of investing in the project from the socio-economic aspects:

Investment in the project with respect to the values of obtained EIRR=21.51% and ENPV=170,643,101RSD shows that Stara planina touristic resorts road network has a high economic justification. Positive results are obtained from sensitivity analysis. Sensitivity analysis show economic justification even for worst case scenario.

7) Conclusions and recommendations of the Feasibility Study for road network of Stara planina touristic resorts

Procedures carried out in functional evaluation shows need for the implementation of this project. The economic evaluation shows economic justification of the investment. It is recommended to start constructing tourist resort Stara planina road network (sections S1, S2, S3).

### 5.3. Sensitivity analysis

Specificity for projects of tourist resorts is in fact that future traffic flows are directly related only to the success of well operating of tourism resort, i.e. occupancy of resorts.

For this reason, there is a need to analyze the worst case scenarios within the sensitivity analysis. Worst case scenarios means the situation of low usage of resorts capacity, significantly lower than the adopted within the traffic forecasts.

In such cases, the worst case scenario in the sensitivity analysis would have the following limits:
• For economic cost up to +10%
• For economic benefits -30 to -50%

## 5. Conclusion

Traffic forecast for tourist resort low volume road networks, is totally different from the usual traffic forecast for other roads networks. Difference is in the fact that only induced traffic flow are analyzed and that AADT is determined in the function of capacity of touristic resorts facilities.

Traffic analyze of tourist resort road networks, which was conducted in accordance with the relevant methodology, show that the future tourist resort road network has a high Level of Service for tourist resorts users.

Economic evaluation procedures, is conducted trough analyze of economic costs and economic benefits. Economic benefits is in the function of road user's costs-RUC, i.e. TTC and VOC.

Specificity of Cost Benefit Analysis – CBA for the cases of tourist resorts road network is not to analyze the RUC of network with and network without investment (because network without investment practically does not exist). Instead of this approach modified methodology is applied and solution is find in the direct analysis of differences in RUC for paved new road and for unpaved gravel road.

The assessment of economic feasibility is determined on the calculated values of the internal rate of return-IRR and net present value-NPV.

**REFERENCES**


Glavic D. et al. 2009. *Pre-Feasibility Study with the General Design for "Turistic resorts Golija"*, CEP.

Glavic D. et al. 2010. *Feasibility study of road construction C1, C2, C3 for Tourist resort "Jabučko ravnište" in Stara Planina*. IMS

Kuzovic, L. et al. (2010). Induced Traffic and Its Treatment in the Evaluation of Motorway Projects. *Promet-Traffic & Transportation*, 22(6): 459-465.

Kuzovic, L. et al. (2007). Model for Forecasting and Benefits Assessment from Induced Traffic. *Journal of road and traffic engineering*, 54(4): 17-22.

TRB (2000). *Highway capacity manual HCM-2000*. Washington DC: National Research Council - Transportation Research Board

HEATCO. (2006). *Developing harmonized European approaches for transport costing and project assessment*. In Deliverable 5 proposal for harmonized guidelines

HDM (2000). HDM-4: The highway development and management series.
Available from Internet: http://www.piarc.org/en/knowledge-base/road-assets-management/HDM-4-Software/

UNECE. (2003). *A Set of Guidelines for socio-economic cost-benefit analysis of transport infrastructure project appraisal*. New York,

World Bank. (2005). Transport Notes: *A Framework for the Economic Evaluation of Transport Projects*

EC DG Regional policy. (2002). *Guide to cost-benefit analysis of investment projects*.
Available from Internet: http://www.strukturalni-fondy.cz/upload/1082642530guide02_en.pdf

*Transport Infrastructure*

# ECONOMIC EFFICIENCY OF LARGE-SCALE INFRASTRUCTURE PROJECTS IN THE CZECH REPUBLIC

**Jana Korytárová[1], Vít Hromádka[1], Eva Vítková[3], Lucie Kozumplíková[4]**
*Brno University of Technology, Faculty of Civil Engineering, Department of Structural Economics and Management, Veveří 95, Brno, 602 00, Czech Republic*

**Abstract:** This article examines the corporate economic efficiency of large-scale government investment into road infrastructure of the Czech Republic. The research objective is to analyse and evaluate the achieved level of economic efficiency of road and highway projects in the Czech Republic, and at the same time the expediency of the investment resources spent on their realisation. This paper identifies and assesses the approach used in the Czech Republic to evaluate these types of projects. Based on a sample of case studies, a comparison is made between the results achieved from the point of view of the costs and gains of administrators as well as the users of the road infrastructure, and the environmental impact. CBA (cost-benefit analysis), sensitivity analysis and mathematical simulation using the Monte Carlo method have been used. The conclusion discusses the analyses and mathematical simulations with the aim of evaluating the manner in which these projects are being analysed from the point of view of economic efficiency. Further important factors which have a fundamental influence on efficiency have also been researched, and possible obstacles to evaluation or to properly achieving the economic efficiency of infrastructure projects are identified.

**Keywords**: Economic Efficiency, Road Infrastructure Projects, Czech Republic, Analysis, Simulation.

## 1. Introduction

The territory of the Czech Republic is characterised by high population density, which places great demand on quality of road infrastructure. However, a significant part of the road infrastructure in the Czech Republic is currently in unacceptable technical as well as capacity condition. Growing living standards accompanied with lifestyle changes increase the number of people using passenger car for their private travel and increasing consumption of commodities burdens roads with truck traffic more and more. As the time passes demands for speed of transport grow too as the time becomes more and more precious. All of the abovementioned and other here unmentioned factors place increasing pressure on the required quality and capacity of the road infrastructure. The road infrastructure administrator (the Road and Motorway Directorate in the Czech Republic managing motorways, speedways and 1st class public roads in the country) monitors and analyses this demand and addresses road structure issues. However, the resources, mostly coming through the State Fund for Transport Infrastructure, are considerably limited and therefore financing of individual projects is governed by predefined criteria and conditions for their use. These include legislative and technical preparedness of the projects, with the key role performed by the compulsory project evaluation with regard to economic efficiency by HDM-4 software and standard investment economy indicators respecting the principles of the Cost-Benefit Analysis (CBA). The present article focuses on efficiency analysis of selected sample projects planned to be implemented by the Road and Motorway Directorate and defined expected efficiency of road infrastructure projects in the Czech Republic. The analysis uses data from practical studies evaluating economic effectiveness of road infrastructure projects. The purpose of the present article is to map economic effectiveness of road infrastructure projects and define their expected effectiveness in the Czech Republic. Another output is also specification of the mean contributions of the analysed projects to their total contribution.

## 2. Contemporary references in the literature

Gellert and Lynch (2003) categorise megaprojects into four types: infrastructure (e.g. dams, ports and railroads); extraction (e.g. minerals, oil and gas); production (e.g. heavy military hardware such as fighter aircraft, chemical and manufacturing plants); and consumption (e.g. tourist resorts, malls and theme parks). This paper is focused on megaprojects in transport, especially road and highway infrastructure.

It is very important to pay close attention to the total impact of these projects on society as well as their efficiency, because the impacts of these projects (technical, financial, economic, legal and environmental) are huge. The substantial cost and the range of direct and indirect impacts these projects have on the community, environment and budgets all attract high levels of public attention and/or political interest.(Capka, 2004)

Megaprojects are characterised by complexity, uncertainty, ambiguity, dynamic interfaces, significant political or external influences and time periods stretching to a decade or more (Floricel & Miller, 2001). At least three features associated with megaprojects are notable: the large sum of resources; the high human, social and environmental impact; and extreme complexity (Capka, 2004; Flyvbjerg, Bruzelius, & Rothengatter, 2003).

The complexity of megaprojects is brought about by a number of contributing factors such as tasks, components, personnel and funding, as well as numerous sources of uncertainty and their interactions (Mihm, Loch, & Huchzermeier, 2003; Sommer & Loch, 2004).

[1] Corresponding author: hromadka.v@fce.vutbr.cz

The economic valuation of large-scale projects and megaprojects in transport, especially road and highway, infrastructure in the Czech Republic, is based on the essential principles of CBA (Cost-Benefit Analysis). The CBA approach is described in detail in the European Commission Methodology (Florio, 2002).
In order to express a project's real utility for society, it is very important to describe in detail and quantify the benefits for the society and the impacts caused by the realisation and operation of such project.
Ongoing improvement and more realistic social cost-benefit ratios are still key areas of research (Priemus, 2010). In the Czech Republic, the HDM-4 model (which is explained in Chapter 3) is used for the economic valuation of important projects in road and highway infrastructure. One very important aspect which must be taken into account is the quality of supplies. Many studies on this topic that focus on the construction sector highlight examples of unethical behaviour by companies in that industry (Cetković & coll., 2012). Nonetheless, this concerns issues of public procurement and the realisation phase of the project, which is not the subject of research within this paper.

## 3. Methods of Analysis

The underlying concept for the economic evaluation of transport projects is based on the principles of the CBA and WLC (Whole-Life Costs) approaches (Korytárová, 2010).
One very comprehensive approach to evaluating investment projects in transport infrastructure is the HDM-4 model. The model was developed from 1993 to 2000 at the University of Birmingham, United Kingdom, with financial support from the World Bank. The HDM-4 model (Highway Development and Management) is focused on evaluating investment projects in the area of transport infrastructure, especially highway projects. It consists of a sophisticated software tool for analysing roads and highways and investment decision-making in this area.
The evaluation of the economic efficiency of the construction of roads and highways in the Czech Republic is nowadays based on the Implementing Guidelines for Evaluating the Efficiency of Road and Highway Construction in Investment Objectives, issued by Ministry of Transport of the Czech Republic. The Czech System for the Evaluation of Roads is in principle based on the HDM-4 model, and supplemented with information specifically calibrated for the Czech Republic. The main criteria for evaluating the efficiency of these projects are the net present value (NPV), internal rate of return (IRR) and the benefit-cost ration (BCR).
The essential principle of the economic efficiency of transport infrastructure project evaluation is based on CBA. Here all relevant socio-economic benefits and costs are compared with respect to the value of money over time. The calculation formula for evaluating roads and highways is drawn from the following groups of impacts:

- Infrastructure costs
- User costs
- Others costs
- Additional external costs

Infrastructure costs mainly include investment costs for realising the construction and costs for maintenance and repairs.
User costs mainly involve costs connected with the operation of vehicles, e.g. fuel, tire wear, the repair and maintenance of vehicles, as well as costs connected with commercial vehicles (drivers' wages, insurance, depreciation and overhead costs).
Other costs are divided into three parts: the evaluation of passengers' time, the evaluation of time during the transport of goods, and losses from vehicle accidents.
Additional external costs represent the external impacts of transport on a project's surroundings, mainly on the environment. In this category are calculated losses due to noise and air pollution caused by traffic.
All analyses conducted in the HDM-4 model are dependent on the relevant inputs. The basic input is information about the present state and future development of transport for the zero and investment variant. The zero variant represents the "no project" alternative, whereas the investment variant supposes that the evaluated project is implemented. Based on information about traffic flow, its intensity, its structure and changes due to realisation of the project, it is possible to determine the differences between the two variants and to apply suitable methods to derive net benefits or costs arising from the project realisation and operation. Next, these evaluated benefits and costs are compared in terms of the evaluation criteria (NPV, IRR and BCR). From this comparison, it is finally possible to recommend the project's implementation or rejection.

## 4. Case Study

The case study includes introduction, analysis and economic evaluation of major projects in the area of road infrastructure in the territory of the Czech Republic. The projects are assessed and compared especially with regard to their economic efficiency. As already mentioned in chapter 3, evaluation of effectiveness and efficiency of major road infrastructure projects (implemented by the Road and Motorway Directorate of the Czech Republic, including new developments, reconstructions and modernisation of motorways, speedways and 1st class roads) in the Czech Republic is compulsory and is done with the help of the HDM-4 software, considering change of total construction costs between the investment and the zero variant in economic efficiency evaluation.

In the case of the HDM-4 software assisted analysis the costs are perceived in the social context. In addition to the costs of infrastructure building and maintenance, repairs and reconstructions costs connected with operation of means of transport are also considered (including fuels, lubricants, tyre wear etc.), as well as costs connected with travel times, traffic accidents and the external costs including air pollution with emissions and noise. The present case study includes together 9 construction projects, some of them including more variants of implementation. Thus the present paper works with 15 project variants. The individual projects or variants are described in Chapter 4.1 below.

### 4.1 Projects and project variants addressed by the present case study

The project of speedway R52 concerns reconstruction of the existing road I/52 in the section between Pohořelice and Austrian border. Five variants have been defined and three have been evaluated. Variant I considers a four-lane road of R 25.5 category – new construction category R 25.5/100, variant II proposes a four-lane speedway with a separating middle strip between the two directions making maximum use of the existing I/52 road, and the combined variant III is identical with alternative R52_25_5 without implementation of the section of I/52 road along the Nové Mlýny dam (between the Upper and the Middle reservoir), which would be left as it is.

Project D11 involves construction of a section of D11 motorway and R11 speedway from the current end of D11 motorway outside Hradec Králové to the Polish border. Part of the R11 speedway project is divided into three variants. Variants I and II assume construction of a R 25.5/120 category speedway in full and half profile (2+1), while variant III proposes speedway construction in category R 21.5/100.

Project I49 involves relaying of road I/49 addressing the unacceptable traffic condition of the section of the existing I/49 road from the end of Vizovice municipality up to and including the direction and elevation break between Vizovice and Lhotsko municipalities in the north-eastern part of the Czech Republic.

Similarly project I34 considers relaying of the existing I/34 road between Lišov and Vranín municipalities in the southern part of the Czech Republic. The new road is designed in category S11.5/80.

Project R35 involves new construction of speedway R35 in the section between Opatovice and Mohelnice in the north-eastern part of the Czech Republic. The speedway is designed with a middle direction-dividing strip in category R25.5/120. The project has been designed in two variants, variant I including a multilevel crossing at Řídký, and variant II not including this crossing.

Project R7 considers construction of a new speedway no 7 in the section Slaný – Bitozeves. The R7 speedway will connect the capital city of Prague with the town of Chomutov. The project includes two variants, apart from a few exceptions only differing in the selected road category (variant I – R 25.5/100, variant II – R 21.5/100).

Project R3 concerns the continuous section of the R3 speedway between Třebonín municipality and the Austrian border.

Project I38 considers implementation of a new road connecting road I/38 on the northern city outskirts (exit in the direction Kolín, Čáslav) and road I/34, or II /150 on the east (exit towards Svitavy and Žďár nad Sázavou). Road I/38 will be built in category S 11.5/70.

Project I20 concerns road construction "I/20 Losiná, bypass" on the south-eastern end of the city of Pilsen. This construction is part of the I/20 road relaying and D5 - Seč building construction project. The building construction project is designed in two stages, the present I20 project being part of stage I.

### 4.2 Evaluation of economic efficiency of the projects

The purpose of the present paper is assessment of economic efficiency and effectiveness of major road infrastructure projects. The general assessment is based on the sample projects briefly described in Chapter 4.1 above. All of these projects are in their preparatory - pre-investment stage in the context of which their financing from the State Fund for Transport Infrastructure is to be decided. Financing of the projects is conditioned, inter alia, by their preparedness in the sense of complete documentation for building permit proceeding and their economic effectiveness. Economic effectiveness analysis of each of these projects was made in the past with compulsory use of HDM-4 software, described in detail in Chapter 3 above. The basic outputs of the economic evaluation include standard indicators of investment economy, i.e. Net Present Value (NPV), Internal Rate of Return (IRR) and Benefit-Cost Ratio (BCR). The evaluation also includes sensitivity analysis assessing the impact of input factor change (especially investment costs and user benefits) on the economic effectiveness indicators. The output of the present paper is specification of the expected level of effectiveness of road infrastructure projects in the Czech Republic. This expected value is calculated with the help of Oracle Crystal Ball software, primarily designed for prediction of development of dependent quantities by simulations using the Monte-Carlo or Latin Hypercube methods. The expected revenue rate is characterised with the mean, median and standard deviation values. As the specification of the expected revenue rate used data from 15 projects or their variants the size of the sample cannot be considered statistically conclusive. Nevertheless, the results can be used at least for general deduction of standard efficiency of road infrastructure projects in the Czech Republic. Regarding the nature of the factors expressing economic efficiency of projects the expected value was specified on the basis of the BCR indicator for its relative nature and transparent calculation. To make the outputs of the analysis more interesting the assumed contributions of the individual projects to total economic benefit were also specified, albeit on the basis of arithmetic mean only.

### 4.2.1 Input values for analysis

As mentioned above, the inputs for the analysis and specification of the expected value of BCR were taken from economic efficiency studies of the individual projects listed in Chapter 4.1. For the sake of completeness the following is the formula [1] for specification of the Benefit-Cost Ratio.

$$BCR_{(m-n)} = \frac{NPV_{(m-n)}}{C_m} + 1 \quad [1]$$

Where:

$BCR_{(m-n)}$ Benefit-Cost Ratio

$NPV_{(m-n)}$ Net Present Value (discount rate r)

$C_m$ Discounted investment costs of the construction

Table 1 below shows a basic survey of outputs of the economic efficiency studies of the individual investment projects considered for our analysis. The table includes BCR values, plus, for the sake of increased informative value, the assumed investment cost values and net present values, not used in the subsequent analysis any more.

**Table 1**

*Survey of outputs of economic efficiency studies of the selected projects*

| Project | Variant | Investment costs | NPV | BCR |
|---|---|---|---|---|
| R 52 | I | 8 007 155 201 | 4152,77 | 1,642 |
| | II | 7 031 215 000 | 3896,69 | 1,692 |
| | III | 6 566 704 201 | 3398,63 | 1,682 |
| D 11 | I | 35 448 888 000 | 6463,10 | 1,235 |
| | II | 32 024 729 000 | 7908,46 | 1,315 |
| | III | 26 333 135 000 | 10149,48 | 1,481 |
| I 49 | I | 169 918 548 | 124,05 | 1,760 |
| I/34 | I | 1 048 853 000 | 2952,78 | 3,965 |
| R35 | I | 39 305 741 083 | 43226,70 | 2,577 |
| | II | 39 218 971 871 | 44953,49 | 2,645 |
| R7 | I | 8 618 157 000 | 1 308,46 | 1,202 |
| | II | 9 778 568 000 | 1 413,12 | 1,194 |
| R3 | I | 8 944 373 666 | 3229,64 | 1,394 |
| I38 | I | 1 163 771 970 | 1909,79 | 2,750 |
| I20 | I | 1 054 785 000 | 83,249 | 1,086 |

*Source: Studies of economic efficiency of selected projects, in-house processing*

Table 2 shows a survey of individual benefits in the form of operation cost saving of the users, cost savings related to travel times, traffic accidents and external cost saving connected with mitigation of environmental impact (pollutant gas emissions, noise).

**Table 2**
*Contributions of individual project benefits to total benefit from the projects*

| Project | Variant | Contributions to benefits | | | |
|---|---|---|---|---|---|
| | | Operation cost saving | Travel time | Accidents | External costs |
| R 52 | I | 15,62% | 79,33% | 5,03% | 0,01% |
| | II | 17,43% | 77,00% | 5,56% | 0,01% |
| | III | 15,55% | 78,46% | 5,85% | 0,15% |
| D 11 | I | 14,00% | 72,20% | 13,10% | 0,70% |
| | II | 14,50% | 71,40% | 13,40% | 0,70% |
| | III | 16,10% | 70,60% | 11,50% | 1,80% |
| I49 | I | 34,30% | 51,40% | 1,50% | 12,70% |
| I/34 | I | 24,70% | 68,00% | 2,00% | 5,30% |
| R35 | I | 17,50% | 68,90% | 9,60% | 4,00% |
| | II | 18,10% | 68,50% | 9,60% | 3,90% |
| R7 | I | -0,65% | 75,38% | 26,89% | -1,62% |
| | II | -10,79% | 87,60% | 24,36% | -1,17% |
| R3 | I | 19,30% | 61,80% | 17,50% | 1,30% |
| I38 | I | 21,83% | 69,29% | -0,11% | 8,99% |
| I20 | I | 17,80% | 77,64% | -1,13% | 5,69% |

*Source: Studies of economic efficiency of selected projects, in-house processing*

### 4.2.2 Results of analysis

The analysis involved specification of the expected BCR value for road infrastructure projects in the Czech Republic. The expected BCR value is considered as a random quantity and was specified by simulation using the Monte-Carlo method used by Crystal Ball software including complementing statistical characteristics. The resulting BCR value can be characterised with the following probability distribution, whose curve shape approximates the Maximum Extreme probability distribution. The probability distribution is shown in fig. 1.

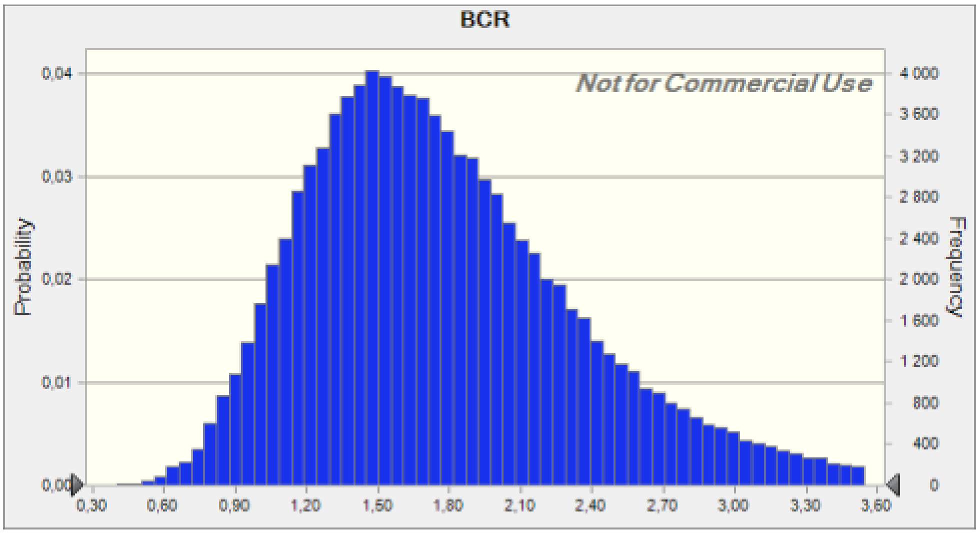



**Fig. 1.**
***Probability distribution of random quantity Benefit-Cost Ratio***
*Source: In-house processing using data from economic efficiency studies and Crystal Ball software*

The statistical characteristics of the random quantity based on simulation by Monte-Carlo method are shown in Tab. 3.

**Table 3**
*Statistical characteristics of random quantity Benefit-Cost Ratio*

| Statistics | Forecast values |
|---|---:|
| Trials | 100 000 |
| Base Case | 0,00 |
| Mean | 1,80 |
| Median | 1,70 |
| Standard Deviation | 0,62 |
| Variance | 0,39 |
| Skewness | 1,12 |
| Kurtosis | 5,22 |
| Coeff. of Variability | 0,35 |
| Minimum | 0,35 |
| Maximum | 7,02 |
| Range Width | 6,67 |

*Source: In-house processing using data from economic efficiency studies and Crystal Ball software*

Table 4 summarises the results of specification of the expected BCR value including calculated mean.

**Table 4**
*Results of specification of the expected BCR value*

| Expected value | Results |
|---|---:|
| Mean | 1,80 |
| Median | 1,70 |
| Standard Deviation | 0,62 |
| Average Value | 1,84 |

*Source: In-house processing using data from economic efficiency studies and Crystal Ball software*

The next part of the analysis compares the mean individual project shares in the total benefit of the construction projects as a whole. The arithmetic mean calculation is based on the data from Table 2. No simulation was performed in this case. The results of the comparison are shown in Table 5.

**Table 5**
*Mean individual project shares in total benefit*

| | Shares in benefits | | | |
|---|---|---|---|---|
| | Operation | Travel time | Accidents | External |
| **Average Value** | 15,69% | 71,83% | 9,64% | 2,83% |

*Source: In-house processing using data from economic efficiency studies*

**5. Discussion**

The present paper includes analysis and evaluation of expected efficiency of road infrastructure projects in the Czech Republic. Specification of the expected Benefit-Cost Ratio value was based on outputs of economic efficiency studies performed for major Czech projects in this area. The evaluation in total included fifteen variants of nine projects. Although the sample was not statistically conclusive, the analysis provides relevant and interesting data. The first part of the analysis assesses the expected value of BCR of the projects in question, specified by simulation with the help of the Monte-Carlo method. The resulting value is 1.84, which shows a very good assumed revenue rate of the road infrastructure projects in the Czech Republic considering the minimum required value of 1. For comparison the mean BCR value was also calculated as arithmetic mean, with the result four hundredths higher, which is a consequence of a different probability distribution consideration. Specification of mean shares of individual project contributions to the total project benefits as complementary information provided by the analysis speaks in favour of travel time saving (about 72 %), with the lowest contribution (albeit not negligible) of the benefits connected with environmental impact. The individual project shares will certainly also differ due to the different characteristics of each of the projects. With regard to the environmental impact the highest benefits will be represented by the bypass projects taking traffic away from town and city centres. Projects increasing transport capacity will bring benefits in the form of reduced travel times and numbers of serious traffic accidents. Projects increasing speed of traffic will probably generate higher operation costs mainly connected with fuel consumption and tyre wear.

## 6. Conclusion

The purpose of the present article was to map economic efficiency of road infrastructure projects and to define expected effectiveness of these projects in the Czech Republic. Individual outputs also include specification of mean shares of individual project benefits to their total benefit. As mentioned above, the evaluated number of sample projects is not statistically conclusive, but still the results of the analysis show a relatively high level of economic efficiency of the prepared road infrastructure projects. It needs to be noted, though, that this efficiency is assessed ex ante, i.e. the values are expected and depend on predicted cost development, technical condition of the roads and traffic intensities, connected with a considerable level of uncertainty. Another step of the present research may be extension of the portfolio of the assessed projects and a more detailed specification of the relation between the criteria indicators and the input values, especially traffic intensity development.

## Acknowledgements

This paper has been written with the support of The Ministry of Education, Youth and Sports Program COST CZ project LD14113 Effectiveness of Megaproject in the Czech Republic.

**References**

Capka, J. R., (2004). Megaprojects – They are a different breed. *Public Roads Magazine*, 68, http://www.fhwa.dot.gov/publications/publicroads/04jul/01.cfm (accessed on 24.05.11).

Cetković, J.; Knezevič, M.; Rutešić, S.; Hanák, T.; Melović, B. (2012) Credit rating evaluation in the example of construction industry. Technics Technologies Education Management. 2012. 7(1). p. 285 - 293. ISSN\~1840-1503.

Floricel, S., & Miller, R. (2001). Strategizing for anticipated risks and turbulence in large-scale engineering projects. *International Journal of Project Management*, 19, 445–455.

Florio, M. (2002) *Guide to Cost-Benefit Analysis of Investment Projects*. [PDF dokument]. European Commision. Available from: http://www.strukturalni-fondy.cz.

Flyvbjerg, B., Bruzelius, N., & Rothengatter, W. (2003). *Megaprojects and risk: An anatomy of ambition*. Cambridge, UK: Cambridge University Press.

Gellert, P. K., & Lynch, B. D. (2003). Mega-projects as displacements. *International Social Science Journal*, 55, 15–25.

Korytárová, J.; Hromádka, V. (2010) Building Life Cycle Economic Impacts. In *Management and Service Science* (MASS), 2010 International Conference. China, Wuhan, MASS 2010. P. 1 - 4. ISBN 978-1-4244-5326-9.

Mihm, J., Loch, C., & Huchzermeier, A. (2003). Problem-solving oscillations in complex engineering projects. *Management Science*, 49, 733–750.

Priemus, H. (2010). Mega-projects: Dealing with Pitfalls, *European Planning Studies* Vol. 18, No. 7, July 2010.

Sommer, S. C., & Loch, C. H. (2004). Selectionism and learning in projects with complexity and unforeseeable uncertainty. *Management Science*, 50, 1334–1347.

# 3D ROAD SURFACE MODEL FOR MONITORING OF TRANSVERSE UNEVENNESS AND SKID RESISTANCE

**Martin Slabej[1], Michal Grinč[2]**
*[1,2]Research centre University of Žilina in Žilina, Univerzitná 8215/1, 010 26 Žilina, Slovakia*

**Abstract:** Qualitative characteristics of pavement in wide range reflects the pavement serviceability, which is a summary of the characteristics of the pavement, providing a fast, smooth, economical and especially safe driving of motor-vehicles. The target factor of pavement serviceability and safety of roads represents the quality of their surface properties. Within the Research Center of the University of Zilina are dealing among others with scanning the pavement surface in order to monitor the individual parameters of pavement serviceability. This paper deals with the creation, analysis and evaluation of 3D- road surface model in terms of two properties pavement serviceability – the rut depth and its texture. Measurements are realized on an experimental field with accelerated straining on the pavement (loading wheel). The long-term accent should be mentioned predictions functions of selected pavement serviceability parameters.

**Keywords:** pavement serviceability, 3D scanner, profile, rut depth, texture

## 1. INTRODUCTION

A requirement for quality and safe road infrastructure resonates at present more frequently considering the obligations of EU member states to reduce the number of accidents and fatalities on the roads. Pavement diagnostics is together with the collection of traffic data and updating process of current database of the road network, the main input of pavement management processes. It is the most challenging process of the entire PMS in terms of financial, time, technological and personnel resources. As new technologies arise, they entail not only the simplification the data collection process but also increase in quality of measurement results and even measurement of data which were previously possible to obtain only through destructive methods. The ability to measure more parameters, measure them faster, more accurately and safely, enables us to obtain the desired volume of data needed for complex PMS decision-making processes - estimation of optimal intervention time and choosing of optimal technology for maintenance repair and rehabilitation. This article is focused on the collection of data regarding pavement structure using the interpretation and collection of data regarding the quality of road surface by scanning its surface with a laser beam.

## 2. Testing field and measurement equipment

The test field has a length of 6 meters and a width of 2.2 meters. The pavement structure was designed as a pavement for a road with traffic load class TLC III. It is a flexible pavement with bitumen concrete surfacing. The wearing base layer is made of asphalt concrete (AC) 11; CA 35/50; 40 mm thick. The base course layer is made of asphalt concrete (AC) 16 P, CA 35/50; 80 mm thick. The road base is a mechanically bound aggregate MSK 31.5 GB; 180 mm thick. Sub-base is gravel ŠD; 31.5 (45) GC; 200 mm thick.

**Fig. 1.**
*Accelerated Pavement Testing Facility of the University of Zilina design of the facility with 3D scanner*

The laser scanning technology allows focusing in detail on the pavement surface and it's near surroundings in a coordinate system. In many cases, it is a more detailed measurement than it is provided by other technologies. Among other advantages of this method a self-moving machine can be included, which does not need any other external mechanisms, that can make difficult to scan remote surfaces and thereby negatively affect the measurement itself [1]. The system allows free movement of the object during scanning. It is also possible to see an image of the scanned surface in real time. On the test field (Fig. 1) were realized the measurements using a smaller type of hand-held 3D scanner - ZScanner 800 (Fig. 2) with a maximum resolution of 0.1 mm. The disadvantage of handheld equipment is time consuming measurement and the evaluation.

[1] Corresponding author: martin.slabej@rc.uniza.sk

**Fig. 2**
*Measurement of 3D hand scanner – Zscanner 800*

## 3. TRANSVERSE EVENNESS

The evaluation of evenness road in transversal direction consists in measurement of transverse evenness as aberrance from theoretical condition. From this reason is in terminology important to differentiate these two basic terms. Transverse road evenness is unevenness of road surface in vertical direction on the traffic direction. It expresses as the difference between existing and theoretical transversal profile of road. The measurement and evaluation of transverse evenness is realized because of determination quality of road pavement from the aspect of permanent deformations, which are well-known as road rut (Čelko et al., 2000).
By measurement of transverse evenness are valuating these parameters:
- ***rut depth /RD/*** – vertical distance between connection of apex wave and the lowest point of wave,
- ***permanent deformations /PD/*** –vertical distance between first and last point of measured profile and the lowest point of wave (fig.3),
- ***water depth /WD/*** – vertical distance between horizontal flat ground in the position the lowest point of wave and the lowest point of wave (Fig.3).

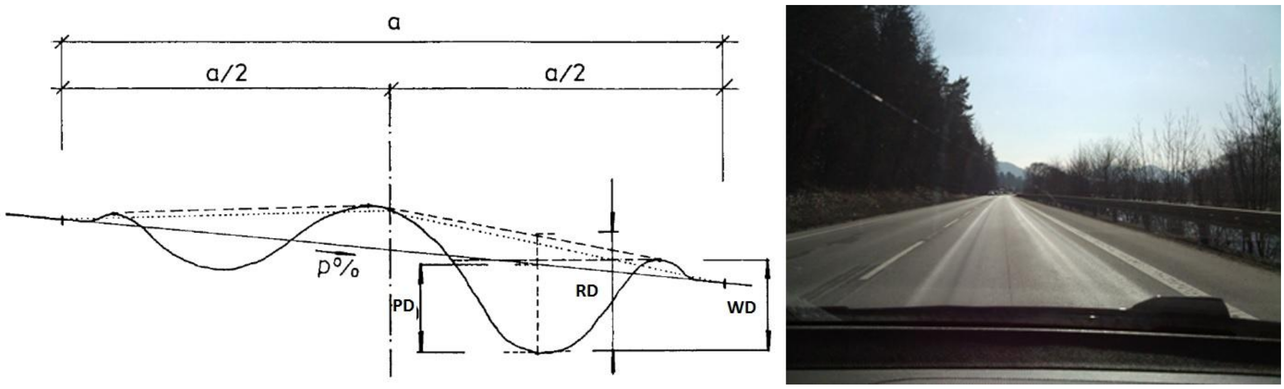



**Fig.3**
*Elementary characteristics of transverse evenness and pavement rut in traffic lane*

The creation of road ruts (Fig. 3) is as a result of mostly two dominant aspects; in first case overlimited traffic load and excessive traffic density and in second case parking and staying heavy trucks on unassimilated surface of pavement – especially inappropriately elected surface (Decký et al., 2010, Kováč et al., 2012). From this reason is very helpful to use the testing field to develop for more detail the reason of creation of road ruts. (Fig. 1).

## 4. TEXTURE

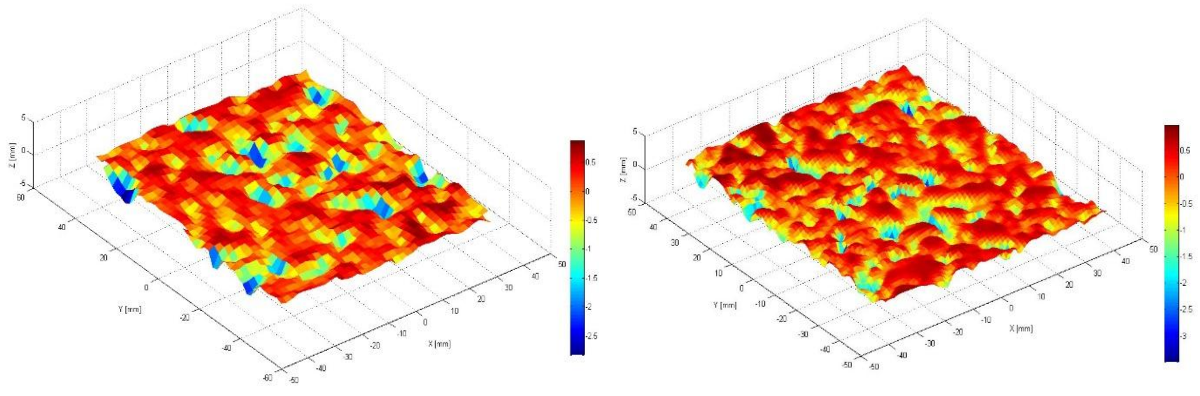
Texture of pavement surface has great impact on anti-skid features of the pavement, maybe the greatest. It is the morphological layout of material of the pavement surface. It is usually described by surface profile which is defined by two coordinates. They are a combination of bumps described by wavelength (horizontal projection of bumps) and amplitude showing vertical projection of bumps in terms of given range. The surface texture influences plenty characteristics of car-pavement collaboration including friction by wet weather, noise, water spraying, rolling resistance, tyre wear and damage of the car. In the point of view of skid resistance microtexture and macrotexture have their special meaning (Decký et al., 2012).

Microtexture reflects tiny prominences on aggregate grains and describes how the grains are smooth or rough and therefore the friction between tyre and pavement surface rises. It is characterised by wavelength range from 0,001 to 0,2 mm and amplitude range from 0,0 to 0,2 mm [4]. Due to the range there is created impression of rough surface but microtexture is usually too soft to recognize it visually. Microtexture of aggregate surface issues elementary friction level and is important on dry surface by low speed up to 40 km/h. Another important meaning lies in an interruption of continual water film and creation direct contact of tyre with pavement surface [4]. Values of microtexture are partially influenced by the ability of aggregates to keep sharp edges and so maintain rough surface which should resist to smoothing caused by truck traffic at longest. Microtexture is partly depended on composition of an asphalt mixture as mineralogical structure of aggregates, max grain size, percentage of small aggregates, and content and type of asphalt binding.

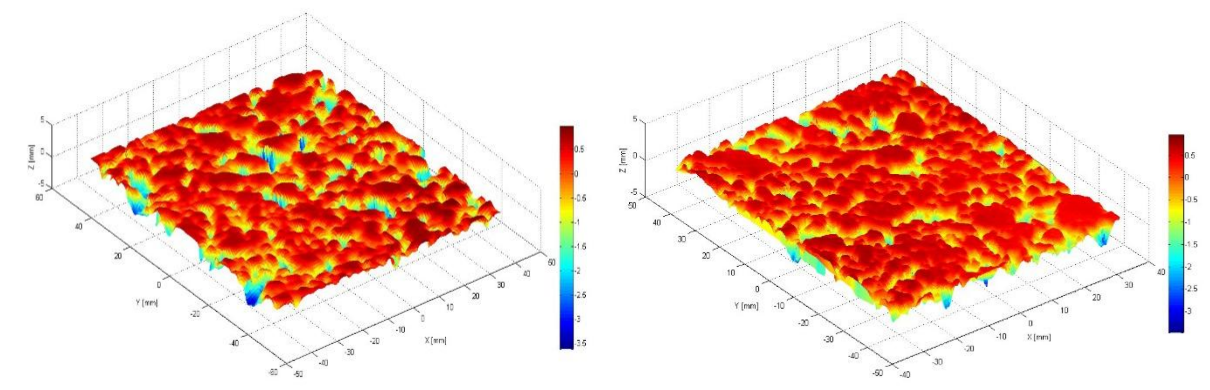
Macrotexture of pavement surface is responsible for basic drain ability of pavement. It represents irregularities on pavement surface and describes a way in which single aggregate grains are ordered. It is characterized by wavelength range from 0,25 to 10 mm and amplitude range from 0,2 to 10 mm [4]. It is important for fast water diversion from surface of wet pavement because the water acts as lubricant and it shows in the friction between tyre and pavement. Macrotexture plays serious role by middle and higher speeds of vehicle (over 40 km/h). A good macrotexture can be get by suitable proposal of aggregates-mortar rate. It can be also achieved with proper combination of methods of final surface modification.

## 5. ANALYSIS OF MEASURING DATAS

The scale and accuracy of scanning are important for correct formulation of the requested level of texture and for implementation of scanned data. Scanning with the help of 3D scanner enables formulation of surface in three proportions and therefore complex information about the pavement surface is available. To compare accuracy of the process, the scanning was done in four resolutions: 2, 1, 0,5, 0,2 mm. The values were selected with a view to potentialities of the device and also to the macrotexture defined above.



*Pavement surface by the scanning in 2 mm resolution* *Pavement surface by the scanning in 1 mm resolution*



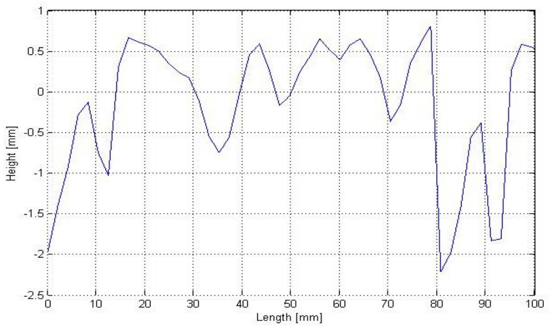
*Pavement surface by the scanning in 0,5 mm resolution* *Pavement surface by the scanning in 0,2 mm resolution*

**Fig. 4**
*The accuracy of the scanning surface*

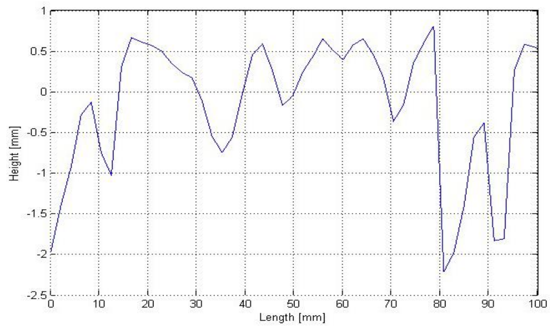
From the scanned surface can be displayed profile by an elected resolution. It is clear that increasing accuracy of resolution scanning the profiles show tinier projections. Appropriately chosen scan accuracy is important for further processing of the measured data (Slabej and Kotek, 2013).
Based on profiles from measured surface the mean profile depth (MPD) was calculated and it is used to represent state of surface macrotexture. The values of the mean profile depth. The values of average depth profile are changing gradually with precision scanning. As the accuracy of scanning was rising the MPD value also rised. When the accuracy was incremented 0,1mm the MPD value incremented 0,01 unit and dark columns are the scanned values.

**Table 1.**
*Values mean profile depth surface*

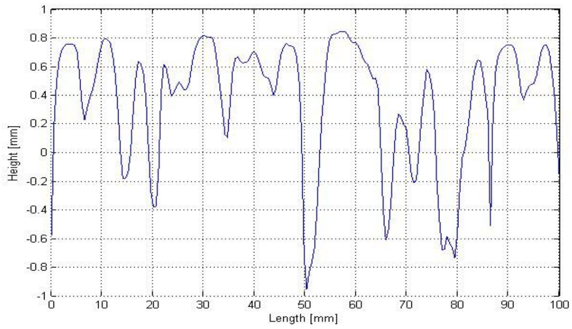
| Accuracy (mm) | 2 | 1,9 | 1,8 | 1,7 | 1,6 | 1,5 | 1,4 | 1,3 | 1,2 | 1,1 |
|---|---|---|---|---|---|---|---|---|---|---|
| MPD (mm) | 0,604 | 0,612 | 0,619 | 0,627 | 0,634 | 0,642 | 0,649 | 0,657 | 0,664 | 0,672 |
| | | | | | | | | | | |
| Accuracy (mm) | 1 | 0,9 | 0,8 | 0,7 | 0,6 | 0,5 | 0,4 | 0,3 | 0,2 | 0,1 |
| MPD (mm) | 0,679 | 0,686 | 0,694 | 0,701 | 0,708 | 0,716 | 0,732 | 0,748 | 0,764 | 0,780 |



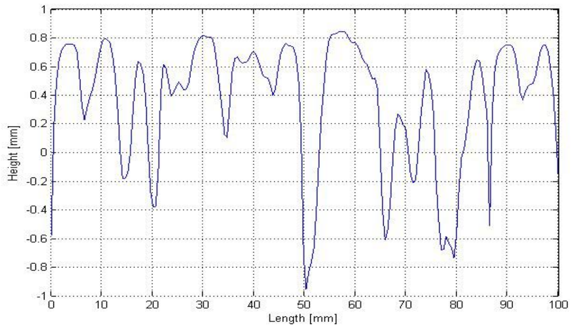

*Texture profile scans at a resolution 2 mm*




*Texture profile scans at a resolution 1 mm*




*Texture profile scans at a resolution 0,5 mm*




*Texture profile scans at a resolution 0,2 mm*

**Fig. 5**
*The texture profile of the scanning surface*

Scan unevenness is sufficient to scan at a lower resolution, seeing that the transverse unevenness is in the order of millimeter values. Scanning of cross section with a length of 1.6 meters and a width of 8 cm was realized with the scanning resolution accuracy of 1mm. For better determination of analyzed profile and mainly due to the relatively difficult lighting conditions are used in the measurement template bounding observed cross section (Fig. 6).

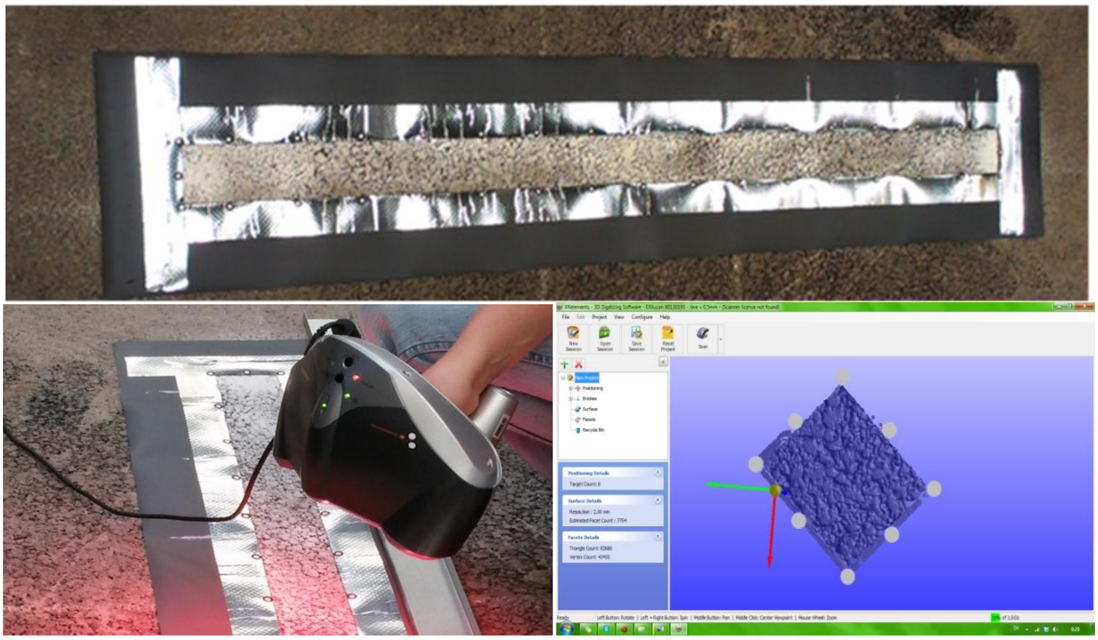


**Fig. 6**
*Template bordering the observed cross section with analyzing the surface in SW "VXelements"*

The schema on Fig. 7 shows one of the evaluation of the cross section on the test field. As is immediately obvious, the transverse unevenness is not present, respectively intercepted unevenness is inappreciable value.

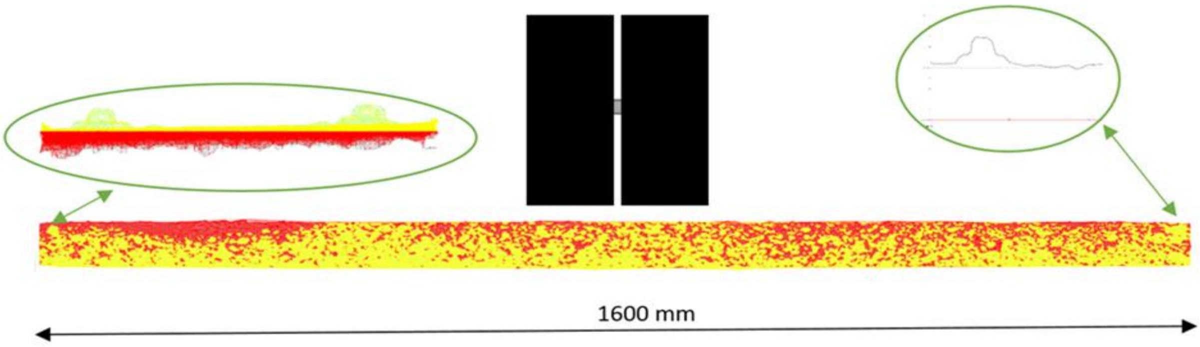



**Fig. 7**
*Scanned cross-section of the analyzed road*

During the detailed and enlarged view of the edge of the profile are evident nails, which have their legitimate role in repeated measurements and will serve as a permanent points through which it will be possible to measure an identical profile in any timeframe. In the middle of profile is shown double axle, which simulates the real effects of crossover of the truck on the experimental field. Our priority effort will monitor changes in the surface properties of the pavement after a defined amount of axle loads. On the ground of this data will then be possible to determine the functions that will predict the development of analyzed serviceability parameters of pavement´s surface, depending primarily on the traffic load (Přikryl et al., 2011, Sivaneswaran et al., 2004).

## Conclusion

One of the most important and the most significant factors which directly influences economic factors as well as safety traffic on the roads is a parameter of serviceability. It is important not only from the perspective of properly elected maintenance, reconstruction, or planned road restoration. Constant road´s monitoring and analyzing helps overall diagnostics of the road network.

This article introduces basic characteristics and two important parameters of serviceability skid resistance and transverse unevenness. It also points out on the necessity of the further monitoring and analyzing the above mentioned parameters. They will create features and models that will later play an important role in the future predictions and thereby directly help to develop and attempt to improve our road network.

**References**

Čelko J. a kol. 2000. The surface properties of pavements, in Slovak: Povrchové vlastnosti vozoviek (Prevádzková spôsobilosť vozoviek), EDIS Žilina, ISBN 80-7100-774-9.

Decký M., Drusa M., Pepucha Ľ., Zgútová K. 2013. Earth Structures of Transport Constructions. Pearson Education Limited, Edinburg Gate, Harlow, Essex CM20 2JE. Edited by Martin Decký, p. 180, ISBN 978-1-78399-925-5. Printed and bound in Great Britain by Ashford Colour Press, Gasport Hampshire.

Decký M., Gavulová A., Pitoňák M., Putirka D., Vangel J., Zgútová K. 2010. Designing and budgeting asphalt pavements, in Slovak: Navrhovanie a rozpočtovanie asfaltových vozoviek, BTO print, Žilina, ISBN 978-80-970388-0-9

Kováč M., Remišová E., Čelko J., Decký M., Ďurčanská D. 2012. Diagnostics of pavement´s serviceability parameters, In Slovak: Diagnostika parametrov prevádzkovej spôsobilosti, Žilinská univerzita v Žiline, EDIS, ISBN 978-80-554-0568-1.

Přikryl M., Kutil L., Žák J. 2011. 3D measurement technology of laser scanning, in Czech: 3D měření technologií laserového skenování, Švédsko, silnice 41 (väg 41) Bergham–Gullberg. Conference Asfaltové vozovky, České Budějovice, 77-81

Sivaneswaran N., Pierce L.M. and Mahoney J.P. 2004. Transition from Manual to Automated Pavement Condition Surveys, Washington State's Experience, 6th International Conference on Managing Pavements

Slabej M., Kotek P. 2013. Pavement surface characteristics influencing a pavement serviceability, In: TRANSCOM 2013: 10-th European conference of young research and scientific workers: Žilina, June 24-26, 2013, Slovak Republic. - Žilina: University of Žilina, ISBN 978-80-554-0696-1. - p. 289-294.

*The research is supported by European regional development fund and Slovak state budget by the project "Research centre of University of Žilina", ITMS 26220220183.*

# PREDICTION OF THE PAVEMENT DEVALUATION BASED ON THE REPEATED GPR MEASUREMENTS

**Michal Grinč[3], Martin Slabej[2], Andrea Porubiaková[3]**
*[1,2,3]Research centre ŽU in Žilina, Univerzitná 8215/1, 010 26 Žilina*

**Abstract:** Today, a considerable large emphasis is put on pavement diagnostic methods, which are non-invasive, simple, safe, time efficient and non-intrusive from the viewpoint of interference with traffic flow of particular road. It is important to know how does the road construction behave when is loaded by the traffic and how does its degradation evolve. For this purpose the test field with real road construction was built. The main aim of the test field is to simulate traffic load in laboratory condition. From this point of view, very convenient is use of the GPR method in terms of the pavement and road investigation. At the beginning, two Horn antennas were used with central frequencies of 1 GHz and 2 GHz on the test field to verify its construction and to map its state of the art. Secondly, the same configuration was carried out for the repeated measurements to examine damage caused by the simulated traffic load. In the last step, the prediction of the test field degradation should be evaluated. Here, the first 3D measurement is presented.

**Keywords**: GPR, horn antennas, prediction model

## 1. INTRODUCTION

The most challenging process of the pavement management processes is the pavement diagnostics including the collection of traffic data and updating process of current database of the road network. This issue was addressed by a number of research teams around the world (AASHTO, 2010a, AASHTO, 2010b, Decký, 2003, Timm and McQueen, 2004) and each progressive automated data collection method tends to be very marketable. The issue of the level of quality was also engaged in a number of studies, for example. (Decký *et al*., 2013, Sivaneswaran *et al*., 2004, Timm and McQueen, 2004) in principle, they are showing that the automation of data collection does not necessarily distort the measured data. An overview of current methods and automation procedures of data collection was process by the American Association of State Highway and Transportation Officials (AASHTO) in its publications (AASHTO, 2010a, AASHTO, 2010b).

New technologies allow the simplification process of the data collection together with increase in quality of measurement results. Nowadays, technologies allow carrying out geophysical survey in many different settings and manners. The 3D alignment of survey starts to be preferred instead of 2D alignment. The main advantage of the 3D alignment is obvious. An interpreter can see geophysical anomalies in 3D and therefore interpret geophysical anomalies with more precision. Another advantage is the distribution of an anomaly in space and bigger area surveyed indeed.

This study is focused on the pavement condition and prediction of its degradation within time. For this purpose the test field constructed at the University of Zilina is used. For evaluation of many different measurements carried out on the testing field, measured subject has to be thoroughly known. The test field is going to be investigated by GPR system with two horn antennas with middle frequencies of 1 GHz and 2 GHz in 3D alignment. The investigation will be done in many cycles in order to find out the best degradation model and to predict the future degradation of the pavement.

This test field is part of the pavement accelerated testing facility. The research links up with pavement research performed by company VUIS-roads in the 90's. The general principle of APT testing is to apply artificially inducted load similar to real life traffic load in a compressed time period, thus providing an expedited means of evaluating factors associated with traffic-pavement interaction.

Degradation model characterizes supposed changes of particular parameter or characteristic pavement condition index in relation to time, or repeated loading. Standard practice for estimation of pavement degradation and definition of such model is to repeatedly measure and asses pavement characteristics in given time intervals on particular road section, and, after their statistical processing, ascertainment of their relationship, in most cases, with traffic load, time, or other factors affecting a given parameter. The implementation of accelerated tests on pavement test field (Fig. 1.) is currently a priority objective of the research activities in the immediate future.

## 2. ACCELERATED PAVEMENT TESTING FACILITY – THE TEST FIELD

The test field has a length of 6 meters and a width of 2.2 meters. The pavement structure was designed as a pavement for a road with traffic load class TLC III. It is a flexible pavement with bitumen concrete surfacing. The wearing base layer is made of asphalt concrete (AC) 11; CA 35/50; 40 mm thick. The base course layer is made of asphalt concrete (AC) 16 P, CA 35/50; 80 mm thick. These layers are connected by penetrating coating PS; 0.5 kg/m2.

[3] Corresponding author: michal.grinc@rc.uniza.sk

The road base is a mechanically bound aggregate MSK 31.5 GB; 180 mm thick. Sub-base is gravel ŠD; 31.5 (45) GC; 200 mm thick. Conformity of all supplied materials has been confirmed by tests affirming the quality elaborate supplied by the constructor; quality of particular layers was confirmed through quality tests performed during the construction as prescribed in the test plan.

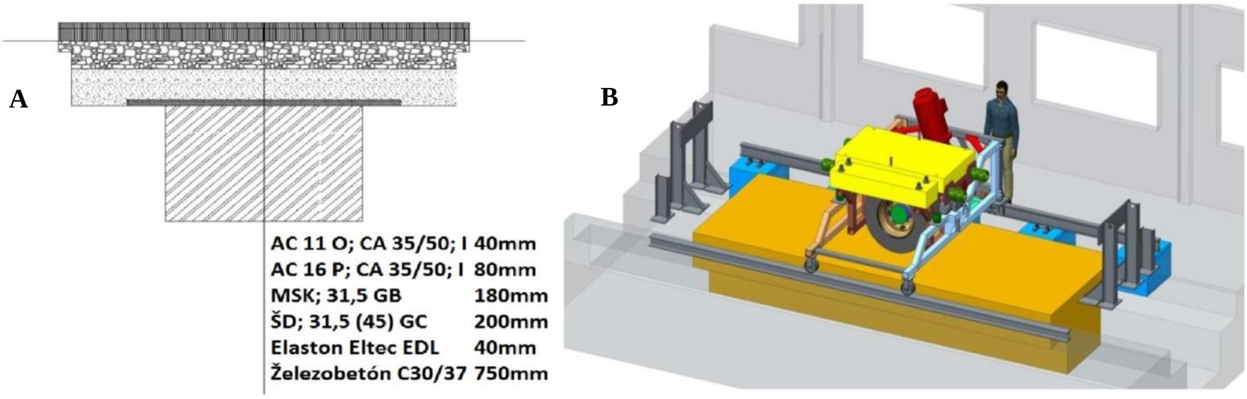



**Fig. 1.**
*Accelerated Pavement Testing Facility of the University of Zilina – A. cross section of pavement test field and B. design of the facility*

### 3. Diagnosis road construction layers using GPR (Ground Penetrating Radar)

The geophysical survey is getting more and more popular for its effective and economic reasons. As the time goes by, new geophysical equipment and interpretation techniques simplify geophysical measurements and make them more accurate. Among current methods of geophysical survey, the most effective, in terms of road construction diagnostics (except of FWD), is GPR survey. The reasons for growing popularity of the GPR method are its non-intrusive nature, high resolution, ideal depth range (depending on the transmitting antenna), low cost and high speed of a survey. Moreover, it can be carried out in full traffic. Data acquisition is performed in situ by GPR device, consisting of transmitter and receiver antenna, control unit and a computer. The measurement is carried out directly on the surface of a studied structure. In our case, it was used device from GSSI - SIR 20 with two Horn - type antenna (center frequency of 1 GHz and 2 GHz). Transmitter and receiver antenna is hanged above surface at a distance of about 0.45 m.

#### 3.1. PRINCIPLES AND METHODOLOGY OF GPR MEASUREMENT

GPR is a relatively fast method of geophysical survey. It is based on emitting of a high-frequency electromagnetic signal in regular time impulses into the rock, soil or anthropogenic environment and its subsequent registration of passing and reflected waves from bodies and interfaces. The position system (e.g., GPS, odometer) is an important part of the measurement system that identifies the position of each measurement within the measured distance (Matula, 2013).

Several GPR time records (A - scan) at regular intervals along specified profile form B - scan or a radargram. A radargram shows continuous record of measurements along a profile. Radargrams collected along given profiles in x and y direction can be stack together and form C – scan or 3D radargram. Data processing is carried out in a specialized software system. Modified data is then interpreted and graphically processed in the end. Thicknesses of layers, possible delamination, also built objects, inhomogeneities and other hidden faults can be calculated from the resulting travel times.

In this study two horn antennas (1 GHz and 2 GHz) are used with the SIR – 20 system. Since the test field has compact size, grid of 1,4 x 2,6 m is enough to cover whole area of interest. The antennas are hanging on the rail system and moving by hand. The distance passed is recorded by the incremental cell clip on the wheel that run in the rail. This system allows us to measure under comfortable condition in the 3D configuration.

### 4. Processing and results

The before mentioned test field was investigated by the GPR system SIR-20 (GSSI) in a 3D alignment. In this study, both antennas, with center frequency of 1 and 2 GHz were applied. The step of measurement of 0.01 m was used in both cases. The time record length differs, 20 ns were used for 1GHz antenna and 30ns for 2GHz antenna. The total measured length of the profile was 1.4 x 2,6 m with 0,2 m grid spacing. Manufacturer's specified depth range for this type of antenna is about 0,9 m for 1 GHz and 0.75 m for 2 GHz antenna, which is sufficient for diagnosing of road condition and construction layers. Radargram was treated by special software package ReflexW and Radan 7.0.

The raw data were firstly processed in the following manner; 1 Static correction - setting of the 0 value (the surface of the test field); 2 1D filtration – background; 3 Car-box filtration; 4 Selective amplification of the signal. Results are represented by a series of pictures shoving 3D model at different depth niveau (Fig. 3. and Fig. 5.). These views were defined based on the 2D radargram interpretations. (Fig. 2. and Fig. 4.).

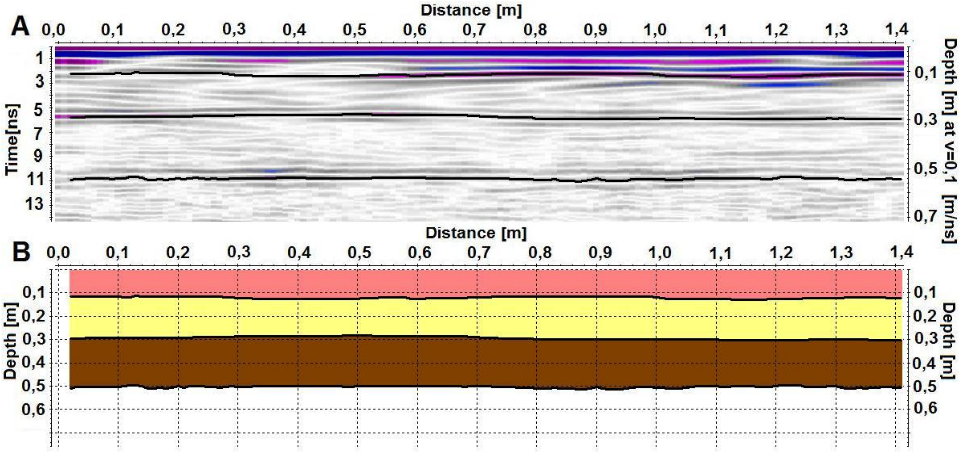



**Fig. 2.**
*A. The resulting radargram obtained by 1GHz antenna, black line indicates boundaries between the layers of the test field. B. The resulting depth section of the test field.*

**Table 1**
*Depths of the test field layers with trigger interval of 0,1 m, their average value and dielectric constant.*

| Profile | layer 1 | layer 2 | layer 3 | layer 4 |
|---|---|---|---|---|
| **distance [m]** | **depth [m]** | **depth [m]** | **depth [m]** | **depth [m]** |
| 0,0 | ? | 0,12 | 0,29 | 0,50 |
| 0,1 | ? | 0,11 | 0,29 | 0,50 |
| 0,2 | ? | 0,11 | 0,29 | 0,50 |
| 0,3 | ? | 0,11 | 0,29 | 0,50 |
| 0,4 | ? | 0,11 | 0,29 | 0,50 |
| 0,5 | ? | 0,11 | 0,29 | 0,50 |
| 0,6 | ? | 0,11 | 0,29 | 0,50 |
| 0,7 | ? | 0,10 | 0,29 | 0,50 |
| 0,8 | ? | 0,10 | 0,29 | 0,50 |
| 0,9 | ? | 0,10 | 0,29 | 0,50 |
| 1,0 | ? | 0,10 | 0,29 | 0,50 |
| 1,1 | ? | 0,10 | 0,29 | 0,50 |
| 1,2 | ? | 0,10 | 0,29 | 0,51 |
| 1,3 | ? | 0,11 | 0,29 | 0,51 |
| 1,4 | ? | 0,11 | 0,29 | 0,50 |
| **Aver. depth[m]** | **?** | **0,11** | **0,29** | **0,50** |
| **Diel.constant** | **?** | **8,1** | **8,6** | **10,3** |

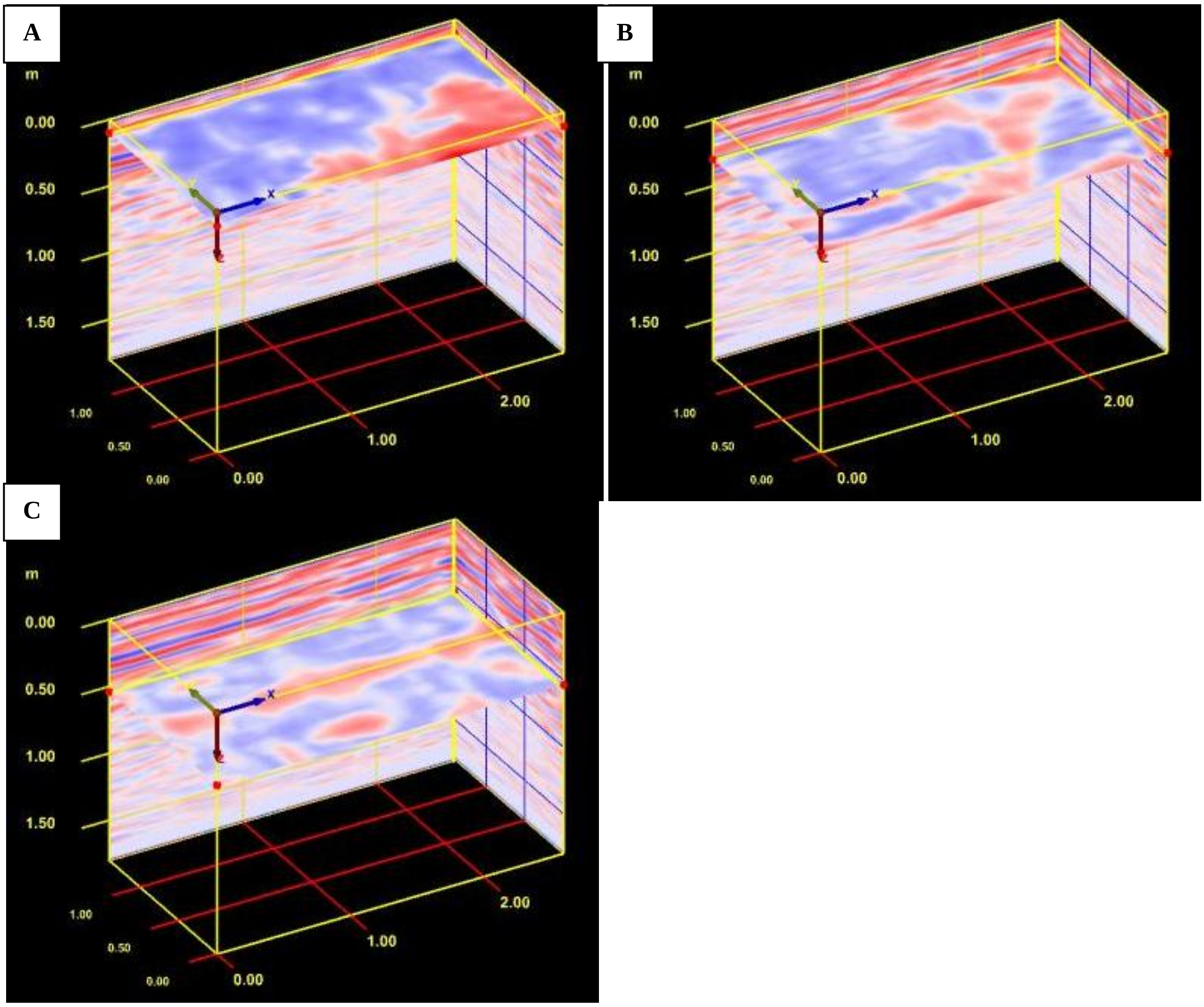



**Fig. 3.**
*3D visualization of GPR results at different depth levels; A. 0,1 m B. 0,3 m C. 0,54 m*

The before mentioned 2D interpretations were carried out for identification of the test field layers. These two profiles are perpendicular to the drive way (axial load) of the test field device. Considering the constructing layers of the pavement, both antennas show similar results but different values for dielectric constants. The treated radargrams was converted to depth section after the raw data processing. Every pavement interface was determined by analysis of the amplitude in the time travel section (Fig. 2.A and Fig. 4.A) and on this basis it was converted into the final depth section (Fig. 2.B and Fig. 4.B). The resulting GPR interpretation of the test field is almost identical with the project documentation. It can be observed intact subhorizontal layers of the experimental road section. The values of the layers thicknesses, as well as the velocity of propagation of the electromagnetic signal in the studied environment and calculated values of the dielectric constant are shown in the table (Table 1 and Table 2).

Results based on interpretation of measurements carried out with 1 GHz antenna show three of four layers. The very first one cannot be distinguished, which might be caused by the lower sensitivity of the 1 GHz antenna on very shallow anomalies. Both interpretations, 2D as well as 3D, show subhorizontal layers which are not disrupted or bend. This result was expected because the test field is very new. Very similar results were obtained by processing data carried out with 2 GHz antenna. In this case all four layers can be identified. Here, the very similar results can be observed as it is in 1 GHz antenna case. All layers appear to be subhorizontal without disruption or bend. Degradation of the pavement can be expected in future, when the axial load will be applied in many cycles. This degradation should be also observed in the radargrams.

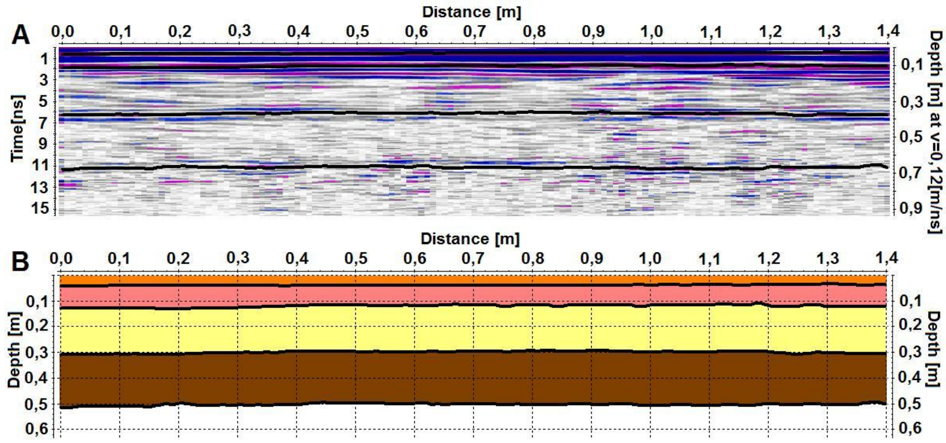



**Fig. 4.**
*A. The resulting radargram obtained by 2GHz antenna, black line indicates boundaries between the layers of the test field. B. The resulting depth section of the test field*

**Table 2**
*Depths of the test field layers with trigger interval of 0,1 m, their average value and dielectric constant*

| Profile | layer 1 | layer 2 | layer 3 | layer 4 |
|---|---|---|---|---|
| distance [m] | depth [m] | depth [m] | depth [m] | depth [m] |
| 0,0 | 0,04 | 0,13 | 0,31 | 0,51 |
| 0,1 | 0,04 | 0,13 | 0,31 | 0,51 |
| 0,2 | 0,04 | 0,13 | 0,31 | 0,50 |
| 0,3 | 0,04 | 0,13 | 0,30 | 0,50 |
| 0,4 | 0,04 | 0,12 | 0,29 | 0,50 |
| 0,5 | 0,04 | 0,12 | 0,30 | 0,50 |
| 0,6 | 0,04 | 0,12 | 0,30 | 0,50 |
| 0,7 | 0,04 | 0,12 | 0,30 | 0,50 |
| 0,8 | 0,04 | 0,13 | 0,30 | 0,51 |
| 0,9 | 0,04 | 0,12 | 0,30 | 0,50 |
| 1,0 | 0,04 | 0,12 | 0,30 | 0,51 |
| 1,1 | 0,04 | 0,12 | 0,30 | 0,51 |
| 1,2 | 0,04 | 0,12 | 0,30 | 0,50 |
| 1,3 | 0,03 | 0,12 | 0,30 | 0,51 |
| 1,4 | 0,04 | 0,13 | 0,31 | 0,51 |
| **Aver. depth[m]** | **0,04** | **0,12** | **0,30** | **0,50** |
| **Diel.constant** | **4,6** | **4,9** | **13,3** | **13,6** |

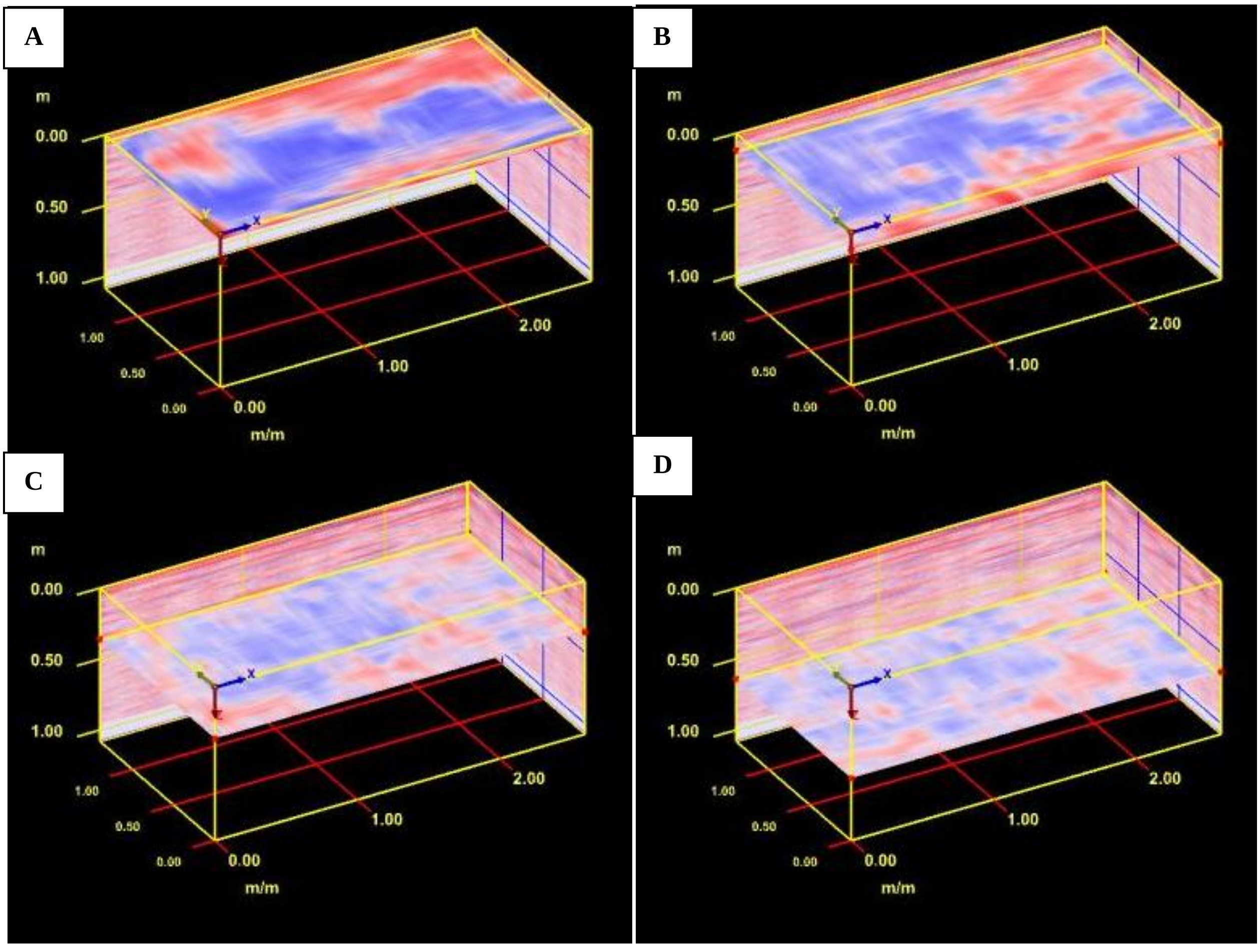



**Fig. 5**
*3D visualization of GPR results at different depth levels; A. 0,04 m B. 0,12 m C. 0,36 m D. 0,60 m*

Despite of above mentioned, methods of GPR have not yet become a standard tool for the road administrators in the diagnosis and maintenance of roads. It is also obvious that these methods, when they are used in proper manner, can save some financial expenses in pavement management processes.

## 5. Conclusion

Progressive methods for the assessment of road surface quality, are meant to facilitate the fulfilment of one of the central objectives of the research activity 3.1 – "research and development in the field of monitoring and assessment of transport infrastructure" in the framework of the Research centre founded under the auspices of the University of Zilina.

One of the main subjects of this activity is the development and verification of new condition diagnostic and monitoring methods of the transportation infrastructure. The comprehensive output of the research project will be a development of infrastructure diagnostic systems and methodologies for automated data collection. Consequently, it will therefore be possible to objectively evaluate variables and in-variable parameters of a road.

However the 3D GPR survey is more time consuming than classical 2D survey, it gives better picture about the distributions of anomalies in the investigated area. Moreover repeating measurements under the same conditions after certain cycles of the pavement load allows observing its degradation not only at the top of the pavement but also within the whole column. On the basis of evaluation of serviceability - its individual parameters, it will be possible to create degradation models and trend lines in order to predict the future development of road network's technical parameters; this will ensure the safety of traffic and fluid traffic flow while minimizing the cost during the whole life cycle of the road.

## Acknowledgements

The authors are grateful to the European regional development fund and Slovak state budget for supporting the project “Research centre of University of Žilina”, ITMS 26220220183.

VÝSKUMNÉ CENTRUM
ŽILINSKEJ UNIVERZITY V ŽILINE

**References**


AASHTO. 2010a. Standard Practice for Quantifying Cracks in Asphalt Pavement Surfaces from Collected Images Utilizing Automated Methods

AASHTO. 2010b. Standard Practice for Collecting Images of Pavement Surfaces for Distress Detection

Decký, M. 2003. Dynamic interaction of the vehicle with the road from the aspect of its longitudinal equality. Habilitation thesis, In Slovak: Dynamická interakcia vozidla s vozovkou z aspektu jej pozdĺžnej rovnosti. Habilitačná práca (Inaugural dissertation), *KCS, SvF, Žilinská univerzita*, *Slovakia*.

Decký, M., Drusa, M., Pepucha, Ľ., Zgútová, K. 2013: Earth Structures of Transport Construction; Editor: Decký, M., *Pearson Education Limited*, ISBN: 978-1-78399-925-5

Matula, R., 2013. Non-destructive diagnostics road construction roads using georadar. PhD. Thesis. In Slovak: Nedestruktivní diagnostika konstrukcí vozovek pozemních komunikací georadarem. Dizertačná práca (Dissertation thesis), *Univerzita Pardubice, Pardubice*, 148 s., 2013.

Sivaneswaran, N., Pierce, L.M. and Mahoney, J.P. 2004. Transition from Manual to Automated Pavement Condition Surveys: Washington State's Experience, In *Proceedings of the 6th International Conference on Managing Pavements*

Timm, D. H., McQueen, J. M. 2004. A Study of Manual vs. Automated Pavement Condition Surveys, *Highway Research Centre, Alabama.*

# THE ISSUE OF URBAN STATIC TRAFFIC ON SELECTED EXAMPLES IN BRATISLAVA IN THE CONTEXT OF ECONOMIC SUSTAINABILITY

**Daniela Spirkova[4], Julius Golej[2], Miroslav Panik[3]**
*[1,2,3]Institute of Management, Slovak University of Technology in Bratislava, Vazovova 5, 812 43 Bratislava, Slovakia*

**Abstract:** Urban static traffic is an essential component of urban transport infrastructure which capacity of disproportionately increasing number of car traffic in the city is increasingly undersized. The issue of parking in Slovakia is a long-term problem, whether in existing buildings or with new development projects where parking costs are only kind of forced expenditure on which developers often want to save money. The main objective of this paper is the analysis and critical assessment of the current state of static traffic in the largest city - in Bratislava, one of the most densely populated conurbation in the Slovak Republic. Achieving the objective is implemented through selected examples, whether in the existing residential area - the largest housing estate in Central Europe - Petrzalka, on new residential development projects or densification of existing inner-city residential built-up area. All these examples have a common denominator - the lack of concepts and solutions of static urban traffic. Authors try to bring in recommendations based on the concepts implemented in selected foreign examples as a potential solution.

**Keywords**: transport infrastructure, urban static traffic, economic impacts, real estate development.

## 1. Introduction

European cities are important centers of housing, economy and culture. In those cities are concentrated industry, transport, miscellaneous services, education and tourism. There are more jobs than in rural settlements. They are home to 70% of the EU population and make up over 80% of EU GDP. Mobility within towns is increasingly complex and inefficient. Many European towns and cities are affected by chronic traffic congestion, which causes an estimated annual cost of 80 billion EUR (Spirkova et al., 2013). Transportation has become the biggest polluter of the environment in cities. Slovakia is not an exception too, where emissions from transport are growing every year. While in 1990 emissions were at 8 percent, today the rate of environmental load reaches 17 percent. From the transport point of view, Bratislava is major national and international hub which favorable position within Europe was historically predetermined to become a crossroad of continental transport systems.

## 2. Problem identification

Actual problems of transport in Bratislava is the fact that to Bratislava daily commute to work (eg. Government institutions, university departments and schools) from 160,000 to 200,000 people which increases demands for parking spaces. Catchment area is the whole Bratislava region, part of Trnava region, but now also the border areas of Austria and Hungary, which has to be regarded as a natural fact for the large city such as the capital city of Slovak Republic.

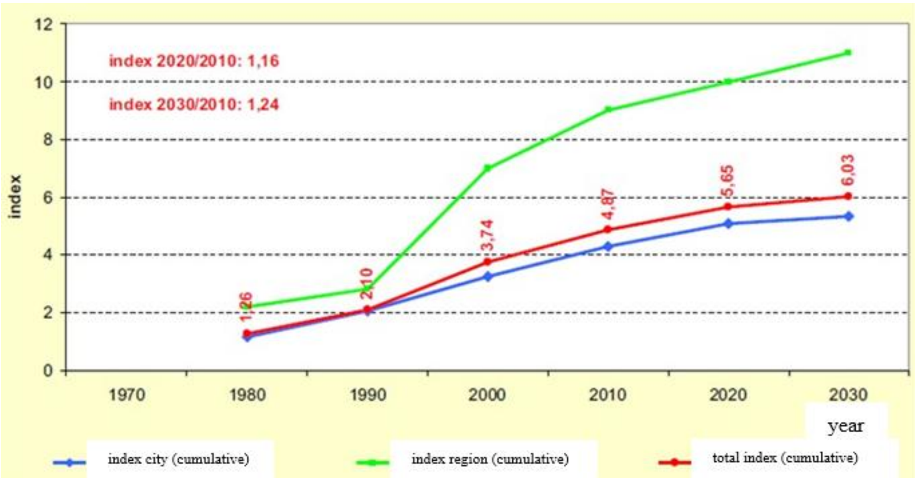



**Fig. 1.**
*Prognosis of growth of car traffic in the city area*
*Source: Methodology of traffic capacitive impact assessment of investment projects, 2014*

[1]Corresponding author: daniela.spirkova@stuba.sk

Problem, that Bratislava was not ready for, is a major building boom which caused a further increase in both passenger and freight transport in the city center. This means that the constantly increasing level of motorization (Fig. 1) brings to Slovakia and especially in densely populated urban areas around the capital - Bratislava even higher space requirements. That becomes the most valuable quantities especially in the inner-city environment. This fact makes new demands on the urbanization of our cities, the professionalism of solutions to traffic problems and high standards to ensure a quality environment. One of the most serious current problems closely related to urban space in Bratislava is the traffic situation, especially the issue of static traffic.
As shown in Figure 2, a growing trend is also in the number of registered passenger cars in Bratislava, suggesting that the current capacity requirements do not conform to design demands of the Seventies of the last century, which basis was used for building today's static traffic. In the terms of processed prognosis, an overall increase in the number of passenger cars up to 275 thousand in year 2030 will be reflected in increased demand for parking places, causing an increase of differences in meeting the needs of static traffic in all parts of the city. Areas for static traffic will no longer be possible to solve with organizational measures on communication network and open spaces, as these options are mainly in the city center fundamentally exhausted. Also the current static transport equipment fails to meet the current needs of the resident population. The problem is starting also with capacity of communications, particularly hub points which are connecting service roads and accesses to civil equipped facilities of mass capacity. Conflicts between vehicular traffic and the preferred cycling and pedestrian movement also occur more frequently. Their solution is very challenging, since the state is given by the overall concept of bad urban conception that does not allow separation of automobile traffic from the use of pedestrian and bicycle movement and even in the position of a recreational movement or transport service. Finally, just static traffic complicates the situation of firefighters in their intervention of fires including possible evacuation of people using a high-rise technology. The most difficult situation is a dense concentration of apartment buildings, where are not enough parking spaces. We can document the model cases on Bratislava districts - Petrzalka and Raca. Cars are often parked on the sidewalks, on both sides of the road and significantly limit the width of the routes, in many cases, firefighters, on their way to extinguish the fire, must get out of the firetruck and manually move parked vehicle.

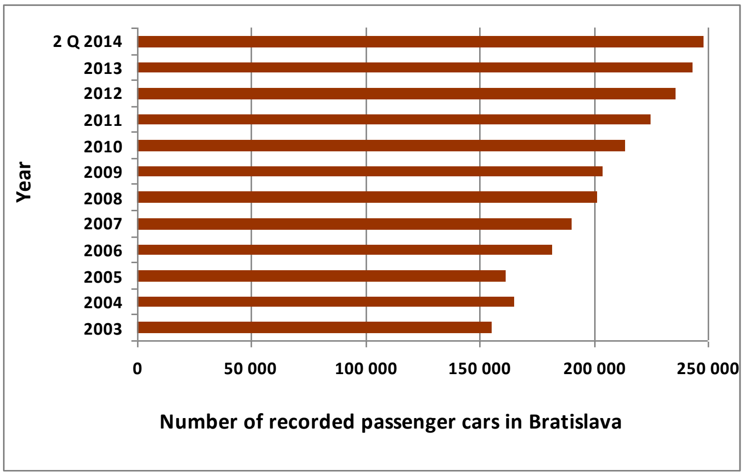



**Fig. 2.**
*Evolution of the number of registered passenger cars in Bratislava in the last decade*
*Source: processed with the materials of Ministry of Interior of the Slovak republic, 2014*

Another important issue in the field of comprehensive transport solutions in Bratislava is the fact that there is insufficient traffic data database and the city does not have sufficient details of current conditions of its urban road network. City of Bratislava is lacking the scheduled surveys and their results, which would be able to determine the disproportion of the current state and predict its development (Methodology of capacitive traffic impact assessment of investment projects, 2014).

## 3. Methodology and materials

The methodology of the design capacity and parking space is completely covered by the Slovak technical standard STN 73 6110: Designing local roads in 2004, updated with the changes (STN 73 6110/Z1) from 2014. These standards replace the original construction standards of the years 1977 and 1986, by which the majority of parking areas have been designed and exist in this form until today.

The formula for calculating the number of stands in the solved area:

$$\mathbf{N = 1{,}1} \times \mathbf{O_o} + \mathbf{1{,}1.\ P_o} \times \mathbf{k_{mp}} \times \mathbf{k_d} \quad (1)$$

$O_o$ - basic number of parking places of residents
$P_o$ - basic number of parking places

$k_{mp}$ - regulatory coefficient of urban location:

- historical core 0,05
- central urban area (inner circle) 0.3
- wider city center (medium range) 0.8
- local centers in city districts 0.6
- specifically defined zones (public sports facilities, shopping centers, ...) 0.7
- other city territory 1.0

$k_d$ - coefficient of impact of transport work division

| | | | | | |
|---|---|---|---|---|---|
| individual car transport: other transport | 35:65 | 40:60 | 45:55 | 55:45 | 60:40 |
| coeficient $k_d$ : | 0,8 | 1,0 | 1,2 | 1,3 | 1,4 |

**Table 1**
*Key indicators in the design of parking places*

| Type of Property | Purpose entity | Parking place on purpose entity | The number of stands of short-term (%) | The nuber of stands of long-term (%) |
|---|---|---|---|---|
| *Parking stalls - multi-storey buildings:* | | *(1 Parking place /20 inhabitants) ** | | |
| • 1 bedroom apartments | apartment | 1 / apartment | - | 100 |
| • 2 bedroom apartments | | 1.5 / apartment | - | 100 |
| *Service (shops, shopping centers):* | number | | | |
| • employees | | 4 (5)* | - | 100 |
| • visitors to 1 h | | 10 (10)* | - (100)* | - |

*** data according to norm STN 73 6110 from year 2004*
*Source: STN 73 6110/Z1 Planning of local communications, 2014*

Key indicators in Table 1 are derived for the degree of motorization 1: 2.5, annual performance of vehicles 10,000 km, the city of 50,000 residents with a central zone and division of transport work of individual automobile transport in the ratio of 40:60 to other transport (rail transport + bus service + urban transportation + cyclists). In special cases, it needs to be taken into account the movement of pedestrians and cyclists. Database data in this paper is drawn from the Statistical Office of SR and Methodology of trafic-capacitive assessment of impact of investment projects (2014). Calculation examples of design capacities of parking areas in the contribution are realized on the selected model examples - Bratislava Petržalka and Rača. In the city district of Petržalka is presented a model example solutions of parking places within the densification of housing complex by newly realized housing estate development projects.In the city district of Raca is presented a model example of solving poorly constructed parking places in the initial neighborhood.

**Table 2**
*Number of people in analyzed areas*

| | |
|---|---|
| Population in Bratislava | 612,682 |
| The estimated population of Bratislava - commuters | 300,000 |
| Population in Petrzalka | 111,021 |
| Population in Raca | 20,251 |

*Source: Statistial Office of the Slovak Republic, 2013*

## 4. Model examples

### 4.1. City district of Petrzalka - calculation of parking places within the densification of housing complex (Fedinova Street)

Parking currently is among one of the most serious problems in the biggest Slovak housing complex - Petrzalka. Today there live approximately 111 000 citizens. In the construction of Petržalka in the Seventies, no one thought about a sharp increase in car traffic. According to current estimates of Petržalka magistrate, today in this city district is about 30,000 missing parking spaces. Per dwelling it is 0.6 parking space - according to a survey of static and dynamic traffic survey in the territory of Petržalka.
The main issue is the unlimited use of public space, where drivers often park their cars on sidewalks, crosswalks, green areas, playgrounds and sports fields. This trend is deeply rooted and drivers park at these sites because of habit and financial convenience to avoid paid parking (Schlosser, 2014). Petržalka static traffic faces several challenges. One of them is proprietary settlement of land, of which the principal amount is owned by city.

Another challenge is finance, as an example, for one parking space is required form 1,200 to 1,500 EUR. Families also currently have two or three cars. Problem is also parked company cars in the city district, and sometimes unfulfillable requirements of residents to park their vehicle in front of their residence.
Another major problem is the new development projects. With their implementation, design of capacity of parking places is often undersized. In the following example - Fedinova street in Petržalka, the calculation according to the above methodology and comparison with the current state of the project.
New development project (apartment block) was carried out in the housing complex between the existing blocks of flats on the Fedinova street in Petrzalka.Housing complex was filled with multipurpose building, which forms 78 new flats and commercial premises situated in the basement (Fig.3). Multifunctional building has 14 floors and the same height as neighboring apartment buildings.On each floor, there are 4 two-bedroom and 2 one-bedroom apartments with an area of 30-52 square meters. Investor also built along 50 new parking spaces on both sides of the building (more on Fig. 3).


**Fig. 3.**
*Multifunctional building and parking places in the surrounding area*
*Source: authors*

*Calculation*
For the analyzed problem (according to the above rules) applies:
Oo = 0; Po = 52 x 1.5 + 26x1 + 1x4 + 1x10 = 118; kmp = 0,6; kd =**1,0** (2)

Then the total number of parking spaces is the:
N= 1,1 . 0 + 1,1 . (52x1,5 + 26x1 + 1x4 + 1x10) . 0,6 . 1,0 = **77,88** (3)

Calculation of parking places, which we made according to the methodology of STN shows us that the capacity of parking places needed for the multipurpose building should be 78 places. In fact, the investor states, that he built 50 parking places by which he greatly undersized (by almost a third) the necessary capacity in the already very crowded housing area.

**4.2. City district of Raca (Kadnarova street)**

Static and dynamic transport is one of the biggest problems that plague the urban neighborhood of Raca. In recent years, the situation is dramatically deteriorating. Not all sites are equally problematic, therefore they require an individual approach. Raca was separate municipality until 1946 known as Racisdorf. According to the last census in 2013, it had 20,251 citizens. Rača is divided into three main parts: Krasňany, Rača and Východné. Part Krasnany was typical low residential buildings neighborhood with a relatively low density of built-up areas. Residential buildings which had 3-4 floors were built mostly in the period of 1950-1965.During the construction were created numerous inner blocks, which were later used as green space for relaxing. Part Stara Raca is characterized especially by the construction of detached houses and houses built in line with the common courtyard. This type of construction was typical for the early 19th century, when in the common yard lived bigger families and their close relatives. Ensuring of static traffic at Rača was adequate. The turning point came in the early Eighties, when the density of construction started to increase significantly. With the advent of new residential projects (but also administrative), parking problem got significantly worse.
The critical situation is in Raca, part Krasnany at Kadnarova street. Kadnarova street is on the one hand a typical example of the original housing and on the other hand, the image of the current building development. Kadnarova street, with a length of nearly 1.6 km, is one of the longest urban streets in Rača. At the beginning of the street is very low initial construction and at the end of the street are new modern houses. In the Sixties, there were built garage parking places that do not conform current requirements. A total number of 90 built garage places which belong to four apartment houses. In each four-story apartment building is located 92 two-room apartments, which were originally intended as a uniform state apartments for teachers. Outdoor parking areas were not built. Overall, for the 368 flats, there are available 90 parking places, which means for one garage belongs 4 cars that is insufficient for today's needs.

The consequence of this is status is parking on both sides of road which results in the narrowing of the road. If two cars are at the same time passing trough the street, they must do difficult maneuvers which creates a traffic jam. The problem is also partial or complete parking on the pavement which is limiting pedestrians. Residents of apartment buildings park partially or completely on sidewalks, roadsides and urban greenery. Critical situations are shown in Fig.4.

**Fig. 4.**
*Static traffic on the Kadnarova street*
*Source: authors*

In assessment of parking spaces deficit in terms of the applicable norm STN 736110, we reached the following results.

The number of required parking spaces we get by $\mathbf{N = 1{,}1} \times \mathbf{O_o} + \mathbf{1{,}1.\ P_o} \times \mathbf{k_{mp}} \times \mathbf{k_d}$ (4)

For the studied problem (area 1) (by the above rules) applies: Oo = 0 ; Po = 368 x 1.5 = 552; kmp = 0,6; kd =**1,0** (5)

Then the total number of required parking spaces is: N = 1,1.0+1,1.368.1,5.0,6.1 = **364,32** (6)

At the analyzed area, according to the parking standards, there should be number of 364 parking places, while today there is 90 places. Deficit of parking places is 274 parking places.

## 5. Economic impacts of static solutions of transport in Bratislava

It can be assumed that main tools for managing the long-term parking will be annual fees. Many parking places in the exposed districts of Bratislava, including Petrzalka, are occupied by vehicles with registration plates from other regions. That is why one of the suggestions for local residents would be bonification advantage of realization of city parking card.That card wolud give (among other benefits) to residents an advantage parking in the locality. The annual commercial rate will apply to motorists from other regions, who will for longer park in other city districts. Another option are the fees associated with the number of cars per household. Fees would be set depending on the number of cars of each household has. Household with several cars had one parking place and with every other parking place the fee would grow. Another solution is based on the idea which with the new strategy of static traffic is dealing with (from 2012) - it is charge for public space. In practice, this means that parking will not be probably free. The first prerequisite is to divide the city to parking zones, which do not have to necessarily correspond to the boundaries of districts. City will cancel the rental of parking places. Places will only be reserved for disabled persons, rescue units, taxis and diplomats. The most comprehensive and also most urgent issue in the context of static traffic is short-term parking. In addition to visitors, the fact is that within the city are moving in cars also its residents. Since districts would have according to different traffic situation in each zone a certain autonomy, several solutions are offered to residents of Bratislava. In less critical, especially peripheral parts, the short term parking will be free, while in exposed zones may be free parking limited with time interval, such as 30 minutes or certain day season.Another possibility, which can be considered is to allow all-day parking for a fee that exceeds the rate for public transport, that will be financially motivating for people to use public transport at the expense of individual automobile transport (Zverkova, 2012).

## 6. Conclusion

Economic development of society brings besides positives also the range of secondary impacts.One of them is expanding requirements on mobility.The key challenge for urban areas is implementation of sustainable mobility concepts and transport policies which also contribute to competitiveness of regions and cities in the significantly enlarged EU. One important aspect in the area of urban traffic is "the parking problem" which has huge impact on the quality of urban life. Parking problems greatly affect urban mobility, safety, pollution, accessibility and traffic flow.The question of handling the car mobility within urban agglomerations appears in the area of conflict between concerns of economy, society and environment. Inner city space has become too valuable and it has to be used other way than for parking facilities. Many cities currently work on these issues of parking and are trying to find sensible solutions and approaches.Many new solutions were already implemented.

Some of them are parking management schemes with flexible, new private car parks with use of new technologies and also "park & ride" systems. However, the basic problem still persists which is mainly due to lack of integrated approaches.
Bratislava is one of the European metropolis, which seeks for the solution for several decades.The list of projects that have an ambition to contribute to the solution of traffic problem in Bratislava is quite long. Solving problems with static traffic in Bratislava could be implemented using comprehensive regulation through traffic signs. In practice this means charging for parking at a time of increased congestion and the designation of paid parking zones with the road signs. Paid parking zones could improve the environment and conditions for non-motorised road users, as well as improving the quality of transport services. One of the solutions to the problems with static traffic could be building a semi-recessed and recessed parking, garage houses, increase recessed parking with one or two floors. Addressing of static traffic in different districts of the city lies in cooperation with the magistrate of the capital. Cooperation includes the selection of appropriate areas that would capacitively mean an increase in parking areas for individual districts. Solution of this complex issue could be also implemented through PPP projects. Another solution is a free parkings not only on the borders of the city, but also outside, for example in Malacky, Pezinok or Senec. The condition is to create high-quality service suburban bus line, which will operate at appropriate intervals.

Summary of newly suggested strategy of static traffic in Bratislava:
1. The division into city parking zones
2. Long-term parking:
   Residents (permanent residence in Bratislava) = basic annual charge:
   - one car in the household - basic charge,
   - second car in the household - times the basic charge,
   - third car at home - commercial charge,
   - in other parts of the city for free or for long-term parking a symbolic fee.

Visitors = ban of long-term parking without paying the annual fee:
   - fee for parking - commercial charge (times the basic rate higher than for the resident)

3. Short-term parking
   Residents:
   - less exposed districts - free,
   - exposed zone - free at certain intervals or charged with a symbolic rate.

   Visitors:
   - commercial charges,
   - time limitation of parking,
   - garage houses,
   - free outside parking or parking with price of full day public transport ticket.
4. Public transport:
   - improvement and densification with service lines
   - integrated (intercity buses, trains, public transport)
   - elimination of competition between the various components of the integrated system
   - strengthening of the rail transport (trams, trains)
5. Automatic control system for the whole territory of Bratislava.

Another solution of problem of static transport in Bratislava we see in the construction of smart parking spaces by installing smart parking sensors placed directly on the parking places.Drivers would be allowed to easily find a free parking place and would contribute significantly to the reduction of emissions in the city. One example is the city of London, where self-governing body in the central London district of Westminster approved the construction of smart parking spaces (this concept has proved in the past, when were on three busy streets, Savile Row, Jermyn Street and St John's Wood High Street, installed 189 parking sensors). Part of the London parking plan is installing very simple smart parking sensors on all existing parking places in central London.The sensors will be able to identify whether a parking place is available and inform the driver via a separate mobile application.
We can conclude that to the reduction of traffic significantly contribute intelligent parking systems which improve the ecological situation, dynamics, availability, and quality of life in the historic parts of town and it can be specifically adapted to the conditions of the locality.Part of smart parking system should definitely be also integrated parking management systems.

**References**

Fabor, T. 2009. Transport in Bratislava – Current problems and vision. *Eurostav Jounal vol..3.*

Schlosser, T. 2014. *Proposal for an amendment of the Methodology of traffic capacitive impact assessment of investment projects.* 11 p.

Spirkova, D., Ivanicka, K., Zubkova, M.2013. *Economic aspects of sustainable development of built environment in Europe*. Publisher MCDP International UG, Germany. 232 p.

STN 73 6110 *Planning of local communications*. 2004. Slovak Institute for Standardisation. 65 p.

STN 73 6110/Z1 *Planning of local communications*. 2014. Slovak Institute for Standardisation. 2 p.

Statistial Office of the Slovak Republic. 2013. Available from Internet: http://portal.statistics.sk/showdoc.do?docid=82242.

*Methodology of traffic capacitive impact assessment of investment projects.* 2014. The Annex to Decision of the Mayor of the capital city of the Slovak Republic, Bratislava. 16 p.

Zverkova, S. 2012. *Parking in Bratislava waiting for major changes.* Available from Internet: http://www.petrzalka.sk/wp-content/uploads/2012/03/Parkovanie_v_Bratislave_cakaju_velke_zmeny.pdf.

# INNOVATIONS IN CONCRETE PAVEMENTS FOR A SUSTAINABLE INFRASTRUCTURE

**Mauricio Pradena[1], Lambert Houben[2]**
[1] *Civil Engineering Department - Universidad de Concepción, Chile*
[1,2] *Section Road Engineering, Delft University of Technology, the Netherlands*

**Abstract:** Concrete pavements (CPs) are durable and they do not need periodic invasive maintenance interventions. Nevertheless, CPs are hardly chosen when only initial costs, instead of life-cycle costs, are considered in the evaluation. Nowadays, there are innovations in Jointed Plain Concrete Pavements (JPCPs) that reduce initial costs about 25% with respect to alternatives with equivalent structural capacity. This paper addresses the question if the innovations early-entry saw-cutting of joints, joints without seals and shorter joint spacing (without dowels bars) are able to maintain the traditional life-cycle performance of CPs. All these innovations affect the joints of the JPCP, and these ones the JPCP performance. Accordingly, the objective of the present paper is to analyse the effects of the joints behaviour on the performance of the JPCPs innovations. The joint behaviour is characterized by the joint activation and opening, the joint capacity to transfer traffic loads and the joint deterioration. The calculations of the joints activation and opening are made with a model developed by the authors. For the estimation of the joint transfer capacity; the results of finite-element software are used. The analysis is completed with field data of the JPCPs innovations. The innovations analysed contribute to a sustainable infrastructure as they can maintain, and even improve, the traditional life-cycle performance of CPs with lower initial costs. Nevertheless, for the design hypotheses to be valid, it is necessary to assure the joints activation and to limit the joints opening to 1.2 mm. With this purpose, for the analysed conditions, it is recommended to cut the joints at least at 30% of the JPCP thickness.

**Keywords**: Innovations, Concrete, Pavements, Sustainable.

## 1. Introduction

A technical, economic and environmental analysis over the road life cycle shows that Concrete pavements (CPs) are a sustainable paving alternative. In fact, CPs are durable and they do not need periodic invasive maintenance interventions. Nevertheless, CPs are hardly chosen when only initial costs, instead of life-cycle costs, are considered in the evaluation. However, in order to reduce construction costs, nowadays there are JPCPs innovations as Early-Entry Saw-Cutting (EESC) of joints, Joints Without Seals (JWS) and Shorter Joint Spacing (ShJS). The question is, Are these JPCPs innovations able to maintain the traditional favourable life-cycle performance of CPs?. All these innovations affect the joints of the JPCP, and these ones the JPCP performance. Accordingly, the objective of the present paper is to analyse the effects of the joints behaviour on the performance of the JPCPs innovations EESC, JWS and ShJS. The joint behaviour is characterized by the joint activation and opening, the joint capacity to transfer traffic loads and the joint deterioration. The calculations of the joints activation and opening are made with a model developed by the authors. For the estimation of the joint transfer capacity; the results of a finite-element analysis tool are used. The analysis is completed with field data of the JPCPs innovations.

## 2. Innovation in jointed plain concrete pavements

### 2.1. Early-Entry Saw-Cutting (EESC) of joints

Early-Entry Saw-Cutting (EESC) consists in a shallow cut (up to 30 mm depth) made with light equipment that allows the saw-cutting 3-5 hours after concrete placement. EESC was introduced to the paving industry in 1988 by a concrete pavement contractor looking for a method to cut the joints shortly after the surface is finished, in order to eliminate the need to return the next day to cut the joints (McGovern, 2002). EESC relieves internal concrete stresses avoiding the ‘wild’ cracking of the pavement and it is postulated that the saw-cut can be shallower at an early age, taking advantage of the significant changes in moisture and temperature conditions at the surface of the slab to help initiate the crack below the saw-cut (Zollinger et al, 1994) .

### 2.2. Short Joint Spacing (ShJS)

When the concrete slabs are designed and constructed such that only one set of truck wheels rest on a single slab, the slab tensile stresses are reduced also because shorter slabs produce less slab curling. This result in thinner concrete pavements (70 to 100 mm less than traditional AASHTO design) (Roesler, 2013) and savings in initial construction costs about 30% (Covarrubias, 2008). Between the pavement design features are: short joint spacing (< 3 m), slab thickness 80 mm to 200 mm; granular base with limited fines (≤8% passing 75 μm) and minimum 15 cm thickness; thin saw-cut at joints (2-3 mm thick); no joints sealing; lateral confinement with curb, shoulder or vertical steel pins; no dowel or tie bars (Roesler, 2013).

[1] Corresponding author: m.a.pradenamiquel@tudelft.nl

### 2.3. Joints Without Seals (JWS)

The function of the joint seals is keeping the joint free of water and incompressible. The introduction of water can drag fines from the base (pumping), so less support and it could produce JF as well as cracks on the slabs. The coarse incompressible materials in the joint have the potential to produce considerable pressure against the edges of the joint, so spalling or even splitting cracks. The costs of sealing transverse contraction joints it is estimated between 2 and 7 percent of the initial construction cost of a JPCP (Hall, 2009). And when the cost of keeping the joints sealed for 10 years is added, the JPCP with sealed joints ends up costing up to 45% more than the one with unsealed joints (Shober, 1987). The problem is that the joint seals are not working well enough, not keeping the joint free of water and long-term Joint Faulting (JF) data shows a strong correlation with annual rainfall (Jung et al, 2011). In fact, the average service life of the joint seals is less than 10 years. They commonly have adhesive and/or cohesive failures. But also they can suffer from hydraulic pressure from tires (Jung et al, 2011). Considering the problems of the joint seals and their costs, there is increased interest in eliminating joint sealants. The innovation of JWS consists in a saw-cut as narrow as possible and a base with limited fines. In effect, a saw-cut ≤ 3 mm impede the introduction of coarse material in the joint and when the amount of fines in the underlying base/soil layer is limited; the water cannot drag fines, so no pumping, and no production of JF for this concept. The JWS can be applied to JPCPs with traditional slabs or ShJS.

## 3. Joint behaviour on the performance of CPs innovations

The expectations of the costumer must be considered in the Life Cycle (LC) of the road. For that reason, the design must be studied together with the maintenance, the impacts upon the users and the residual value of the pavement (Pradena and Echaveguren, 2008). Therefore, the evaluation needs to consider the effect of the joint performance on the JPCP performance, so at the end upon the users and the costs. For instance, JF is the major contributor to the JPCP's roughness, and joint spalling could increase it as well. In this paper the JPCP roughness is quantified by the International Roughness Index (IRI) and the joint behaviour is characterized by the joint activation and opening, the joint capacity to transfer traffic loads and the joint deterioration.

### 3.1. Joint activation and opening

The calculations of the joints activation and opening are made with a model developed by the authors. In this section a brief description of the model is made, more details can be found in the work of Houben (2010a, 2010b) and Pradena and Houben (2012). In the JPCP, the occurring tensile stresses as a product of the restricted deformation follow from Hooke's law, but they are affected by the viscoelastic behaviour of the concrete (relaxation) (Houben, 2010a).

$$\sigma(t) = g*R*E(t)*\varepsilon(t) \quad \text{(MPa)} \tag{1}$$

where $E(t)$= time-dependent modulus of elasticity of the concrete (MPa); $\varepsilon(t)$ = total time-dependent JPCP tensile strain due to shrinkage and thermal effects (-); R = relaxation factor (viscoelastic JPCP behaviour) (-) ; g = enlargement factor (-)

$$g = \frac{h}{h - jd} \quad \text{(-)} \tag{2}$$

The greatest tensile stresses occur in those weakened cross-section. When the JPCPs thickness is $h$ (mm) and the saw-cut depth $jd$ (mm), the relative joint depth $rjd$ is:

$$\text{rjd} = \frac{jd}{h}*100 \quad (\%) \tag{3}$$

The tensile stresses build up over the so-called breathing length ($L_{a1}$). This is that part of a long structure that exhibits horizontal movements due to temperature changes (or another varying influencing factor) (Houben 2010b).

$$L_{a1} = \frac{E_{cm}*\varepsilon}{\gamma * f} \quad \text{(m)} \tag{4}$$

Where $E_{cm}$ = average modulus of elasticity (MPa) at the moment of the crack; $\varepsilon$= maximum total obstructed deformation of the pavement (-); $\gamma$= volume weight of the concrete (kN/m$^3$); $f$ friction between the concrete slab and the underlying base (-).

The cracks occur when the tensile stress ( $\sigma$ ) exceeds the tensile strength ($f_{ctm}$). Because of the initial crack width $w_{1i}$ , a reduction $\Delta\sigma_1$ of the maximum tensile stress occurs. The spacing $L_{w1}$ between 2 primary cracks is equal to 2 times the breathing length.

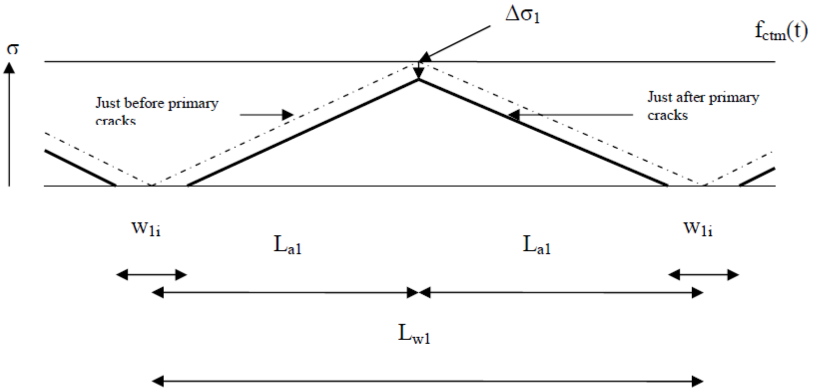



**Fig. 1**.
*Development of the tensile stresses and cracks initiation in the pavement*
*Source: Houben, 2010b*

Eq. 5 allows the calculation of the initial crack width $w_{1i}$ of the first occurring primary cracks.

$$w_{1i} = \frac{1000000 * E_{cm}(t) * \varepsilon(t)^2}{\gamma * f} \qquad \text{(mm)} \tag{5}$$

The change $\Delta w_n(t)$ and the development of the crack width $w_n(t)$ follows from:

$$\Delta w_n(t) = \frac{1000000 * E_{cm}(t) * \varepsilon_n(t)^2}{\gamma * f} \qquad \text{(mm)} \tag{6}$$

$$w_n(t) = w_{1i} + \Delta w_n(t) \qquad \text{(mm)} \tag{7}$$

The model considers in an indirect way the effect of the slab length in the cracks width development. However, the magnitude of the slabs length reduction in ShJS needs to apply a Correction Factor (CF) to the crack width under the joints of the traditional JPCP resulted of the modelling. AASHTO and MEPDG models relate directly the joint opening with the slab length, i.e. if the short slabs length of ShJS is 50% of the traditional slabs length, the crack width of the ShJS would be 50% of the value of the one in the traditional slab. The authors of the present paper have found in field measurements a reduction of crack width of 40% (i.e. CF=0.6), when the short slab length is 50% of the traditional slab length (Pradena and Houben, 2014a) that is the case modelled in this paper, in particular slab length of ShJS 2m, i.e. 50% of 4 m (traditional slab length). A CF 0.6 is equivalent to apply a safety factor of 1.2 to the calculations of the crack width made with the simplified expressions of AASHTO and MEPDG. The modelling takes into account shallow saw-cut associated to EESC (RJDs 20% and 25%) and saw-cut of RJD 30% in order to have a point of comparison. The construction at the day or days when the temperature is highest in a location, i.e. the ‘warmest moment of the year’, produces the widest cracks under the joints (Houben, 2010a), so the most unfavourable conditions for the transfer of traffic loads at the joints. Hence, this is the construction time chosen for the modelling. The temperature is 35°C at the ‘warmest moment of the year’, the period of evaluation is 8640 hours, i.e. practically 1 year; the concrete grade C28/35 and friction 1.

### 3.3. Joint capacity to transfer traffic loads

The joint capacity to transfer traffic loads can be quantified by the Load Transfer Efficiency (LTE). In the present paper the LTE related with joint opening is obtained from the experimental verification of joint load transfer of the 3D finite-element analysis tool EverFE (Davids and Mahoney, 1999). The wider crack widths under the joints are used to the relation with the LTE, because they are the ones that control the JPCP design.

### 3.4. Joint deterioration

The evaluation of the effects of the JWS on the performance of JPCPs considers as a reference the joint seals behaviour and the fact that the costs associated to the joints seals needs to be justified by enhancing the performance of the JPCP. The evaluation takes into account JPCPs applications with different objectives (accessibility or mobility), time in-service (years and/or Equivalent Single Axle Loads, ESALs) and climatic conditions (rainfall and freeze-thaw action)

## 4. Results and evaluation

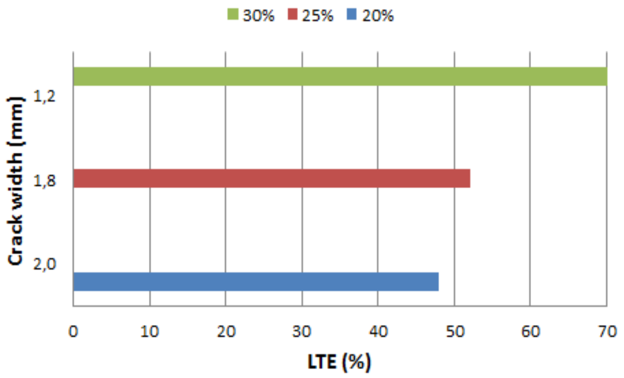
### 4.1. Joint activation

For RJD 20% and 25% the joint activation is 50%, hence the Effective Slab Length (EfSL) is not short anymore. Consequently, not only one set of truck wheels rest on a single slab and there is not a slab curling reduction. Because the EfSL is longer than the designed slab length of 2 m, the effective stresses produced for the traffic and the slab curling are higher in the JPCPs in-service than the ones considered originally in the design. This situation produces deterioration of the pavement and eventually cracks on the slabs. Moreover, where there are joints that remain uncracked, there is a waste of money and time, saw-cutting the joints, preparing the joints to receive the seal, installing the seal and dowels bars at uncracked joints.

For the RJD 30% the joint activation is 100%, hence the EfSL is 2 m, i.e. the slabs are working as they were designed.

### 4.2. Joint opening and capacity to transfer traffic loads

In undowelled JPCPs, a LTE 70% or higher is generally considered appropriate to a good performance. According to the experimental verification of the joint load transfer of the finite-element software EverFE, the crack width under the joints must be 1.2 mm as maximum for a LTE ≥ 70% (Davids and Mahoney, 1999). For the analysed conditions that is produced when the saw-cut is at least 30% RJD (Fig. 2a).




**(a)**

**(b)**

**Fig. 2.**
*Crack width – LTE for different RJD (a) and ShJS at a bus lane at Concepción City downtown in Chile (b).*

At the University of Illinois, U.S.A. an Accelerated Pavement Testing (APT) of JPCPs with short slabs (without dowels bars) was performed. The APT has shown that the LTE converged to values over 70%. In fact, part of the conclusions of that study were: the smaller slab sizes maintained a medium to high LTE over the accelerated loading period for all slab thicknesses without the development of any JF. And the fatigue performance of short slabs, in terms of allowable number of ESALs, significantly exceeds the allowable traffic on the equivalent thickness of traditional JPCPs (Roesler et al, 2012). This good performance is observed in different projects in Chile as well. For instance the Fig.2b shows a JPCP with ShJS of a bus lane in the main avenue of Concepcion City downtown in Chile, where no JF was detected after 7 years in-service.

### 4.3. Joint deterioration

In U.S.A more than 100 sections with and without seals have been investigated in different climates, including zones with rainfall levels over 1500 mm/year and 8 states with freeze-thaw action. According to the comparative analysis of joints deterioration, the joints seals would not enhance pavement performance (Hall, 2009). But the most remarkable experience in U.S.A. is the one of the Wisconsin Department of Transportation (WisDOT) that has investigated for 50 years joint filling/sealing in urban and rural areas, for various traffic levels and truck loadings, type of bases, soils and joint spacings, with and without dowels bars. The results have always shown that sealing does not enhance pavement performance.

Even they have concluded that the pavements with unsealed joints performed better than the pavements with sealed joints and the pavements with shorter joint spacings performed better than the pavements with longer joint spacings (Shober, 1987). In 1995 only Wisconsin reported that it had dispensed with joint sealing entirely, reporting savings of 6.000.000 US dollars annually with no loss in pavement performance (Shober, 1997). In 2000, 3 states (Alaska, Hawaii and Wisconsin) reported they do not apply joint sealing (Jung et al, 2011). Hawaii is the wettest state of U.S.A. with an average annual rainfall of 1785 mm, and Alaska has oceanic climate in the occidental coast and continental and arctic climates in the rest of the state, hence rainfall and freeze-thaw effects in the joints.

Austria, Belgium and Spain have achieved a suitable service life for up to 30 years with JWS and undoweled JPCPs for country roads with light truck traffic (Burke and Bugler, 2002). Austria and Belgium are countries with, at least, rainfall over 500 mm/year in all the country, and big part of the territory with more than 1000 mm/year. Both countries include extensive regions with temperatures below zero as well (freeze-thaw). In Chile JPCPs with ShJS, synthetic fibers and JWS have been built for LVRs with less than 1.000.000 ESALs in areas with rainfall over 500 mm/year and 1000 mm/year respectively. Although there is not enough experience in Chile yet, the JWS in LVRs can be considered feasible due to the LVR objective (accessibility), the low speed, the experience of Austria, Belgium and Spain, and due to the smaller joint opening in JPCPs with ShJS, even more when fibers are included (Pradena and Houben, 2014b).

Measurements of JF and IRI were made on JPCPs with JWS in Guatemala after 3, 8, 15 and 22 million ESALs. The JF values were always less than 2 mm and the IRI values less than 2.4 m/km (Salgado, 2011). Guatemala is a mountainous country with level of rainfall over 1000 mm/year in almost the whole country, and zones with more than 1500 mm/year.

Chile has experience with JWS at different climatic conditions, time in-service and JPCPs applications. For instance, applications as bus corridors in areas with moderate rainfall (between 500 and 1000 mm/year) have shown good performance. Local streets with JWS have not shown joints affected by spalling, according to the classification of the FHWA (FHWA, 2003), after 2.5 years to 6 years in-service in areas with more than 1000 year/mm of rainfall and even potential freeze-thaw action (Pradena and Houben, 2014b). In addition, the successful cases of Guatemala can be assimilated to the ESALs of local streets and bus corridors. Even more, a possible extrapolation of JWS to more exigent cases can be suggested considering stricter specifications for saw-cuts (width ≤ 2.5 mm), base fines content (≤ 6%) and ShJS to obtain smaller joint openings. This suggestion is based specially in the levels of IRI similar to a new JPCP in the rural roads with JWS at Guatemala, the 50 years of excellent experience of WisDOT (even better than with sealed joints) and the fact that nowadays is possible to make saw-cuts of 3 mm or less, instead of the JWS 3-6 mm wide considered in the WisDOT experiences (Pradena and Houben, 2014b).

## 5. Conclusions

The innovations analysed contribute to a sustainable infrastructure as they can maintain, and even improve, the traditional life-cycle performance of CPs with lower initial costs. Nevertheless, for the design hypotheses to be valid, it is necessary to assure the joints activation and to limit the joints opening to 1.2 mm. With this purpose, for the analysed conditions, it is recommended to cut the joints at least at 30% of the JPCP thickness. JWS with thin blade (≤ 3 mm) and limited fines in the base (≤ 8% passing 75 μm) can be used in streets, LVRs, parking lots and any JPCP application where the seals are not working well enough. Even an extrapolation to more exigent cases could be possible with stricter specifications for saw-cuts (width ≤ 2.5 mm), base fines content (≤ 6%) and the use of JPCPs with short joint spacing to obtain smaller joint openings.

## Acknowledgements

The following organizations and persons are acknowledged: The 14th International Scientific Geo-Conference SGEM and its publisher, the National Highway Laboratory of Chile, the Service of Housing and Urbanism (Bio Bio Region, Chile) and the MSc. Marcos Diaz (Building Engineer, Consultant and Academic, Chile).

**References**


Burke M.P.; Bugler J.W. 2002. The Long-Term Performance of Unsealed Jointed Concrete Pavements. In *Proceedings of the 81st Annual Meeting TRB.*

Covarrubias J.P. 2008. TCP Design (Thin Concrete Pavement). Available from Internet: < http://go.worldbank.org/PXC5K2PVA0>.

Davids W.G. & Mahoney J.P. 1999. Experimental Verification of Rigid Pavement Joint Load Transfer Modeling with EverFE. *Transportation Research Record 1684*.

FHWA. 2003. Highway Functional Classification: Concepts, Criteria and Procedures. Available from Internet: <http://www.fhwa.dot.gov/planning/processes/statewide/related/highway_functional_classifications/fcauab.pdf>
Hall K.A. 2009.Performance of Sealed and Unsealed Concrete Pavement Joints. Available from Internet: <http://www.fhwa.dot.gov/pavement/concrete/pubs/hif09013/hif09013.pdf>

Houben, L.J.M. 2010. Transversal Cracking in Jointed Plain Concrete Pavements for Dutch Climatic Conditions. In *Proceedings of the 7th International DUT-Workshop on Design and Performance of Sustainable and Durable Concrete Pavement*.

Houben, L.J.M. 2010. Model for Transversal Cracking in Non-Jointed Plain Concrete Pavements as a Function of the Temperature Variations and the Time of Construction. In *Proceedings of the 7th International DUT-Workshop on Design and Performance of Sustainable and Durable Concrete Pavement.*

Jung Y.S., Neshvadian K. & Zollinger D. 2011. Seal No Seal. Available from Internet: <http://sealnoseal.org/PDF/UtahTTI/SealNoSealutahworkshop.pdf>.

McGovern M. 2002. The Latest on Early-Entry Sawing. Available from Internet: < http://cement.org/tech/pdfs/CT023Saw.pdf >.

Pradena M.A.; Echaveguren T.B. 2008. Quality Management in Road Maintenance Based on the Approach to the Client Principle (in Spanish). *Revista de la Construcción, 7(2).*

Pradena M.A.; Houben L.J.M. 2012. Early age deformation and internal stresses in jointed plain concrete pavements for Dutch conditions. In *Proceedings of the 2nd International Conference on Microstructural-related Durability of Cementitious Composites*.

Pradena M.A.; Houben L.J.M.. 2014. Sustainable Pavements: An Analysis of the Crack Width in Jointed Plain Concrete Pavements with Short Slabs. In *Proceedings of the 14th International Multidisciplinary Scientific GeoConference SGEM.*

Pradena M.A.; Houben L.J.M.. 2014. Sustainable Pavements: Unsealed Joints in Jointed Plain Concrete Pavements. In *Proceedings of the 14th International Multidisciplinary Scientific GeoConference SGEM*.

Roesler J.R. 2013. Thinner Concrete Pavement Technologies. Available from Internet:< http://www.cptechcenter.org/ncc/TTCC-NCC-2013.cfm>.

Roesler J.R.; Cervantes V.G.; Amirkhanian A.N. 2012. Accelerated Performance Testing of Concrete Pavement with Short Slabs, *International Journal of Pavement Engineering, 3(6).*

Salgado M. 2011. Analysis and Prediction of the Behaviour of Semi-Rigid Short Slabs Concrete Pavements (in Spanish). Available from Internet: <http://www.icpa.org.ar/publico/seminario_piarc/sesiones01.html>

Shober S.F. 1987. Are pavement joint sealants always necessary?. Available from Internet: <http://www.concreteconstruction.net/paving/are-pavement-joint-sealants-always-necessary.aspx>.

Shober, S.F. 1997. The Great Unsealing: A Perspective on PCC Joint Sealing, *Transportation Research Record,* 1597.

Zollinger, D. G.; Tang T.; Xin D. 1994. Sawcut Depth Considerations for Jointed Concrete Pavement Based on Fracture Mechanics Analysis, *Transportation Research Record, 1449*.

# *Land Use Development, Spatial and Transport Planning*

# EVALUATION OF SPATIAL STRUCTURE DETERIORATION OF LAND PLOTS CAUSED BY A NEARBY MOTORWAY, BASED ON A SELECTED EXAMPLE USING GIS TOOLS

**Cezary Kowalczyk [1], Jacek Kil[2], Małgorzata Dudzińska[3], Katarzyna Kocur-Bera[4]**
*[1,2,3,4]University of Warmia and Mazury in Olsztyn, Faculty of Geodesy and Space Management, Poland*

Abstract: Construction of motorways and ring roads is a serious planning challenge. By their expansive nature, such investments have a great impact on the environment, landscape as well as on the land use planning process. Their planning and construction frequently cause social and spatial conflicts in the areas through which they run. With modern GIS techniques it is possible to conduct analyses of the impact of such roads on the spatial structure of rural areas and to quickly review different variants of the course of a road.
This article evaluates the change in the spatial structure of plots caused by motorway construction. At the first stage, information was gathered and analysed on the current state. Next, the surface and spatial impact of the new motorway was evaluated. Finally, the evaluation results were used to determine the extent of morphological structure changes to the plots. The results were presented on maps and graphs.
Employing GIS tools allows to unify procedures, to reduce the time needed for the planning and design process and to analyse different locations of the road in order to minimise the negative spatial effects.

Keywords: motorway, morphological structure, land plots

## 1. INTRODUCTION

Poland is a country situated in Central Europe. Due to its special position, it is a transport junction on routes from Eastern and Western Europe and an important logistical node in the flow of goods. A consequence of Poland being a part of the Soviet bloc during the period between 1945 and 1989 is the poor development of its road network, especially motorways and expressways. After Poland's accession to the European Union, the number of cars on Polish roads grew rapidly. Little progress in road construction occurred until recent years. Quick and convenient transport facilities and a sufficiently dense road network of decent quality are the foundation of economic development and proper operation of a national economy. It is also a way of overcoming the "economic exclusion" of the poorest regions.

There were 358 km of motorways in Poland in 2000. In 2005 there were 552 km and in 2008 it reached 765 km (Rocznik Statystyki Międzynarodowej, 2012). However, Poland hosting the Euro 2012 resulted in motorways expanding to about 1400 km. Experts forecast that in or around 2017 Poland may be one of the European leaders in terms of the total length of motorways and expressways. The total length of motorways and expressways in Poland now exceeds 2600 km. Motorways are the safest roads. The accident rate on motorways in the European Union are several times lower than on urban and suburban roads (Lenart, 1998). There were 44 fatal casualties on motorways in Poland in 2012, i.e. 3.09 person/100 km of roads (Wypadki drogowe w Polsce w 2012 r.).

However, construction of motorways and expressways also has negative sides. Several controversies, especially at the planning stage, are created by their course and by their impact on the environment. Local communities living in the vicinity of motorways – both those existing and being constructed – fear the negative consequences of their construction and use. This is associated with increasing social awareness of the consequences of development of motor transport. However, pro-environmental education brings a lot of good because community expectations can be fulfilled only if all the entities which participate in carrying out road construction projects are able to recognise the extent and consequences of the environment degradation and take corrective and preventive measures.

## 2. MATERIALS AND METODS

### 2.1. Study Area

Construction of a motorway affects agricultural space, which results in losses in agricultural land. The extent of such losses may be determined by an analysis of the variability of land use and soil quality class as well as the planning of access roads to land plots along the planned motorway. The adopted measure of the multifarious effect of a motorway on agricultural land is the change in the land value, which is determined only based on diversity of its usefulness for agricultural production. The value is a measure of valuation of usability of land for agricultural production (Bacior, 2012).

[1] Corresponding autor: cezarykowalczyk@gmail.com

Bacior developed a simplified method which enables a preliminary assessment of the impact of a motorway during the road course planning phase. Such a preliminary assessment includes an analysis of such data as:

- diverse quality of soil along the motorway course,
- distribution of roads and motorway flyovers,
- the land area which can be accessed only by crossing the motorway zone,
- parameters of the layout of the plots cut through by the motorway,
- green vegetation belt deployment.

These data formed the basis for determining the changes caused by motorway construction on the land production values. They were later used to comprehensively estimate the effect of the motorway on agricultural land (Bacior, 2012). This method comprehensively estimates the effect of a motorway on agricultural land, including (Wilkowski, 1995; Harasimowicz, 1998) loss of land taken for the motorway, deterioration of the production capabilities of the land situated near the motorway and of the parameters of the layout of the farms crossed by the motorway (Bacior, 2012).

The course of a motorway has a fundamental effect on the spatial layout of farms. Cutting land from the farmsteads of its owners, the degraded parameters of plot layout or their shape may largely decrease their productivity. This is a consequence of longer access roads, limited use of production means and inability to make full use of machines. A reduced farm area may lead to a reduction of its livestock and, in consequence, to leaving some livestock stands in a shed unused (Bacior and Harasimowicz, 2005).

The issue of the impact of a motorway on the natural environment was dealt with by Badora. He used a multivariate analysis to assess spatial conflicts arising as a result of a motorway construction. Kozłowski pointed out that the higher the natural environment value and its devastation and degradation was, the stronger the conflict was between human activities and the natural environment (Kozłowski, 1997).

Badora identifies the following criteria of evaluation of the natural value of the land adjacent to a motorway:

- diverse lay of land,
- density of the flowing water network,
- area of standing water,
- area of meadows and pastures,
- share of protected soil (class I – III and organic soils),
- forest area,
- areas of existing and planned nature protection, taking into account their rank in the national system of nature protection,
- share of NATURE 2000 and ECONET-PL areas,
- landscape diversity,
- sozofloristic value of protected and rare plant stands. (Badora, 2004)

In his method, Badora defined the conflictability of a motorway towards the natural environment as the sum of the natural environment valuation indexes and the environment transformation indexes. Applied to agricultural land, it means that the more the space is transformed, the larger the outlay required to restore it to its original state. The author rightly points out that due to previous decrease in its natural value, agricultural land is the best type of land for motorway construction. But the potential for conflict with economic activities may be much higher. (Badora, 2004).

An evaluation of the extent of “deformation” of the spatial structure of the near-motorway area was also made by Wilkowski (Wilkowski, 1995). He developed a method based on determination of the impact of a motorway on agricultural farms by means of a multi-criterial matrix of causes and effects. The matrix consists of 23 parameters which specify the factors which may bring about important or notable effects that a motorway has on agricultural farms. The interrelations between various factors associated with a motorway construction are used to construct a cause and effect matrix. There are three main types of interactions directly associated with the motorway existence: motorway construction, operation and maintenance.

In the part which deals with the effects of a motorway construction, the author proposes taking into account five groups of effects:

I. Loss of agricultural land,
II. Disruption of the spatial structure of farms,
III. Decrease in crop yield in the zone adjacent to the motorway,
IV. Limitations in the use of land previously used by vegetable gardens, cultivation of berry plants, meadows and pastures,
V. Disruption of existing technical infrastructure.

This method may be helpful in evaluating the impact of a motorway on agricultural land, in developing the optimum programme of management and agricultural activities in rural areas, where the road construction project will be carried out, as well as in preparing opinions on various potential routes of the motorway course through the agricultural land of different villages.

The issue of disrupting the functional structure of agricultural land, in a simplified manner, was dealt with by Chmielowiec and Kaszycki (Chmielowiec and Kaszycki, 1997). They grouped farms in acreage groups and assessed the change percentage in the following ranges:

- less than 50% change,
- 50-90% change,
- more than 90 % change.

The issue of determining the impact of a motorway on an agricultural space was also dealt with by Marcinkowska. Her paper presents a new approach to assessing the extent of disruption of the spatial structure of agriculture, exacerbated by a motorway running nearby. The proposed assessment involves the use of indexes which describe the extent of disruption of rural area. The indexes were divided into 7 groups and 4 types. The groups concern the type of effect exerted by a motorway construction, whereas the types concern a spatial element (Marcinkowska, 1999). The principles of selection of corrective actions were also developed depending on the extent of transformation of different spatial elements, as well as a list of corrective measures. The developed method can be helpful in defining corrective measures to improve the spatial structure of agriculture which deteriorated as a result of the motorway construction and in establishing the urgency and sequence of repair management and agricultural work.

### 2.2. Input Data

A 15 km fragment of the A-1 motorway between Drzonowo and Klęczkowo was the object of the study. The analysed part of the motorway is situated in the communes of Lisewo, Płużnica and Stolno in the province of Kujawsko-Pomorskie. It runs through typical agricultural area, at a certain distance from densely built-up areas, passing by the main settlement centres. It has 10 flyovers, with an average distance of 1500 m from each other.

The analyses were conducted on a group of plots of land, situated near the motorway. Table 1 presents the basic data on the communes through which the motorway runs. The population size in the communes is similar and ranges from 5,026 (commune of Płużnica) to 5,331 people (commune of Lisewo).

The largest area is occupied by the commune of Płużnica (119 km$^2$); with a comparable number of plots of land, the average plot area is 2.8 ha. Taking the average area and perimeter of a plot, it is possible to determine the width-to-length ratio of an average plot in a commune. The best ratio of a side length (width-to-length) of an average plot was found in the commune of Stolno: 1:2.5, whereas the worst one (oblong plot) is in the commune of Płużnica (1:11). This may be caused by the great density of roads.

**Table 1**
*Data on the communes crossed by the motorway*

| Commune | Total area (km$^2$) | Population | Number of plots | Average plot area (m$^2$) | Average plot perimeter (m) |
|---|---|---|---|---|---|
| Lisewo | 86.31 | 5331 | 4514 | 21198 | 669,3 |
| Stolno | 98.43 | 5201 | 4371 | 25084 | 705,9 |
| Płużnica | 119.33 | 5026 | 4729 | 28603 | 714,8 |

*Source: own compilation*

### 2.3. Metods

The study was based on an analysis of the morphological structure of the area situated in the immediate vicinity of the motorway. Morphological analysis is a common method applied in many areas of science (Bitner et al., 2009; Bitner, 2011; Pigol et al., 1993).

The morphological structure, as used in this paper, describes only the plot structure, which is changed as a result of subsequent divisions or consolidations. The term “morphological” applies only to the geometric properties of the structure and it has this meaning when used in mathematics and physics (Bitner, 2011).

The basic unit covered by the analysis is a plot of land, which is the smallest unit of the country area. The basic units can be subsequently transferred by geodetic divisions.

One of the reasons for such divisions may be the construction of a motorway. The right of way exercised in the physical space, causes considerable changes in the natural environment and in the morphology of an area. Plots on a considerable portion of the area are divided, which results in the creation of a new spatial structure. Owing to an analysis of the plot mosaic, it is possible to assess how the changes caused by the motorway may be regarded as a significant disruption of the spatial structure. This assessment applies only to the morphological properties of an area.

An initial state was taken for an analysis of diversity of the plots of land, assuming the normal distribution of the perimeter and area of the plots of land on a larger area than the analysed one. If such a structure is cut through by a road, this will result in disruption, whose effect can be observed with varying intensity, depending on the distance from the source of disruption.

Fig. 1 shows the course of the motorway against the background of the mosaic of plots of land. Parallel lines were drawn along the selected section of the motorway, at a distance of 400 and 800 m from it. These distances were chosen in order to analyse the areas situated in the vicinity of the motorway, with the assumption that the plots of land cut through by the motorway will not be taken into account in the calculations for the neighbouring line by more than 5% (the ratio of the number of plots repeated on the neighbouring lines to the number of plots cut through by a line). A considerable distance between the neighbouring lines would bear the risk that areas originally (before the motorway was constructed) dissimilar.

The shape index was taken as the basic index which reflects the area morphology. An analysis of the shape of flat objects is used frequently in geographic studies of settlements. The shape index was determined from such fundamental parameters of a figure as its area and perimeter. The only measure of shape, which employs a plot area ($P_{dz}$) and its perimeter ($O_{dz}$) is the **k** index proposed by Kostrubiec, which is calculated from the formula:

$$k=\frac{O_{dz}^{2}}{P_{dz}}-4\pi \tag{1}$$

The shape index **k** is a measure of “compactness” of a figure. It is equal to zero for a circle and grows when the figure becomes more elongated. The **k** index grows to infinity for infinitely narrow rectangles (Bitner et al., 2009). The shape index for a rectangle with the side ratio of 1:2 is equal to 5.44, and for a square – 3.44.

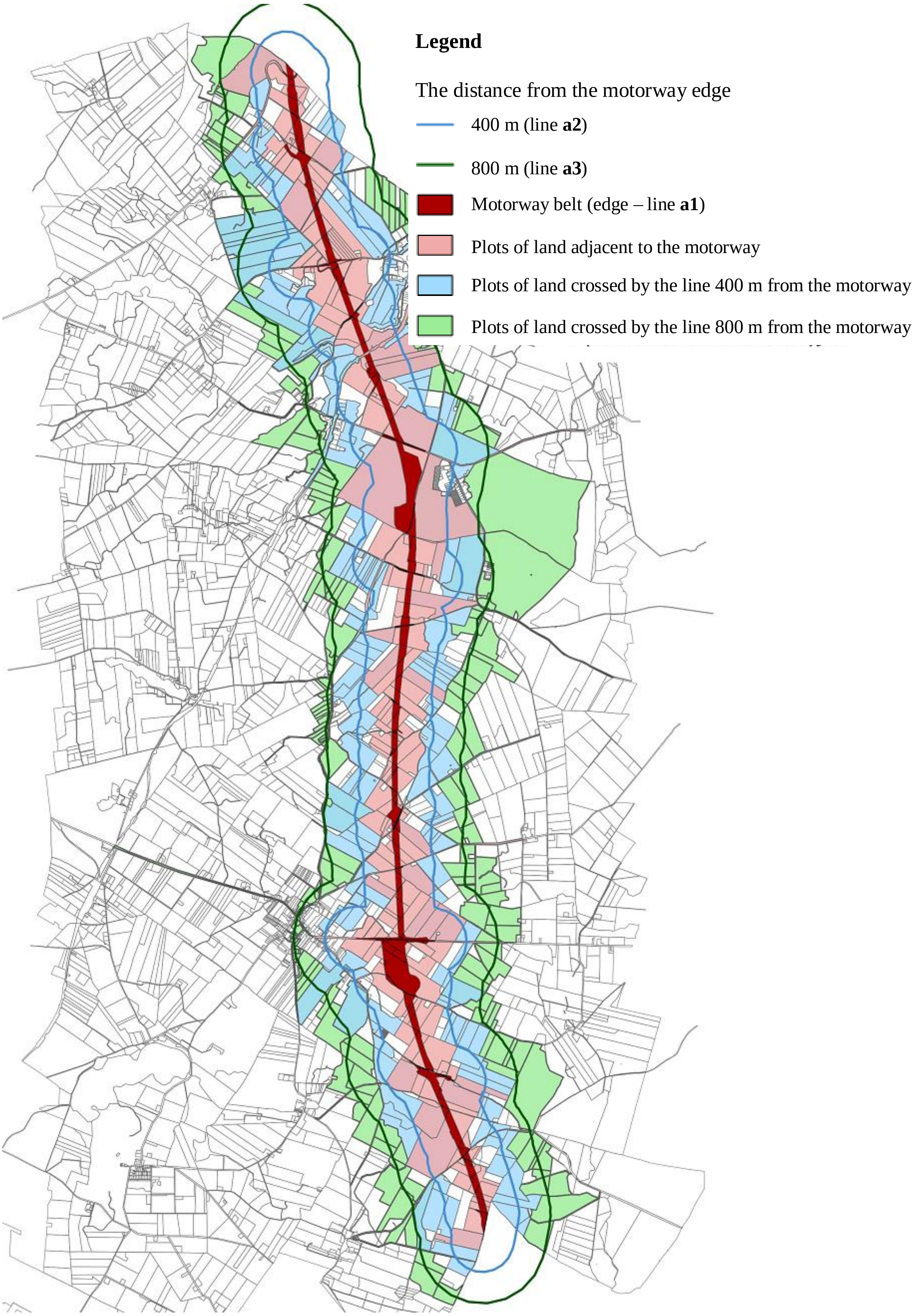



**Fig. 1.**
*The course of the A1 motorway and the plots in the area of the analysis*
*Source: own compilation*

## 3. Results

A spatial analysis was conducted using the GIS software (http://www.quantum-gis.pl); the analysis involved drawing two parallel lines along the section with a length of 15 km, with the distance from the motorway edge of 400 m (line **a2**) and 800 m (line **a3**). The motorway edge is marked **a1** (Fig. 1). The next step involved exporting data containing information about each plot of land cut through by each of these lines: plot number, area, perimeter, position relative to the cardinal directions (E- East and W-West). The data from the analysis of the digital map provided the basis for calculating the shape index; the statistical description of the results is presented in Table 2.

**Table 2**

*Descriptive statistics of the plots crossed by lines at different distances from the motorway*

| Line | Direction | Number of plots | Shape index | | | | | Plot area ($m^2$) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Average ($k_{avr}$) | Median | Minimum | Maximum | STD | Average ($P_{avr}$) | Median | Minimum | Maximum | STD |
| **a1** | E | 161 | 18.11 | 3.97 | 1.00 | 180.90 | 32.48 | 31349 | 12329 | 4 | 484898 | 51880 |
| | W | 208 | 19.09 | 3.51 | 1.00 | 277.00 | 38.81 | 28498 | 9520 | 18 | 569609 | 64647 |
| **a2** | E | 173 | 24.93 | 5.14 | 1.00 | 225.80 | 44.04 | 41440 | 20682 | 363 | 484898 | 56790 |
| | W | 267 | 25.44 | 6.53 | 1.00 | 277.00 | 45.03 | 40112 | 19930 | 284 | 569609 | 64128 |
| **a3** | E | 124 | 31.87 | 4.14 | 1.00 | 502.00 | 65.80 | 66147 | 31934 | 360 | 1378303 | 151451 |
| | W | 221 | 28.33 | 4.69 | 1.00 | 276.10 | 51.82 | 35364 | 13158 | 54 | 308360 | 48258 |

*Source: own compilation*

These results show that the average shape indexes of plots situated in the immediate vicinity of the motorway is lower, which means that their elongation is smaller. This is shown in graph 1 and 2. The results for plots situated on the east side are comparable with those on the west side of the motorway. The plots situated at the distance of 800 m from the motorway have the shape index of 28.33 (west side) and 31.87 (east side), i.e. a side length ratio of 1:8 (west side) and 1:9 (east side).

Conversely, the plots situated in the vicinity of the motorway (near line **a0**) have a shape which is more advantageous for agricultural production. The side ratio of the plots situated on the west side is 1:6 and those on the east side – 1:5. This shows that the shape of the plots situated along the motorway is more advantageous than of those situated farther away from it.

In order to better illustrate the changes in the morphology of the area, the shape index is presented on the diagram of dispersion (diagram 1 and 2), separately for the two sides of the motorway.

This result leaves no doubt: the plots situated farther away from the motorway have a worse side length ratio, i.e., they have a higher shape index.

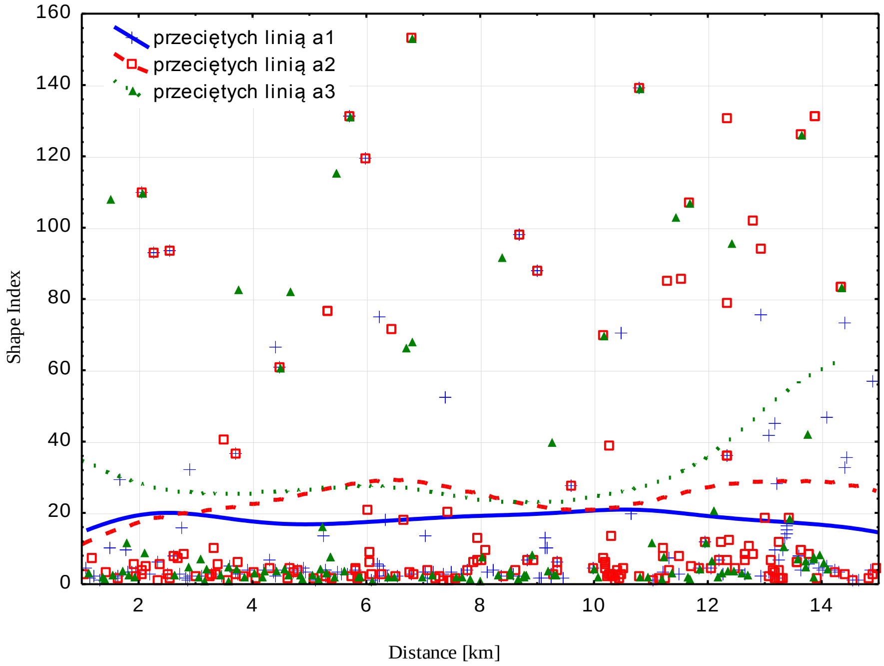



**Diagram 1.**
*Dispersion of the shape index for different groups of plots on the east side (a line smoothed out by the least-square method, weighted with distances, was fitted in)*
*Source: own compilation*

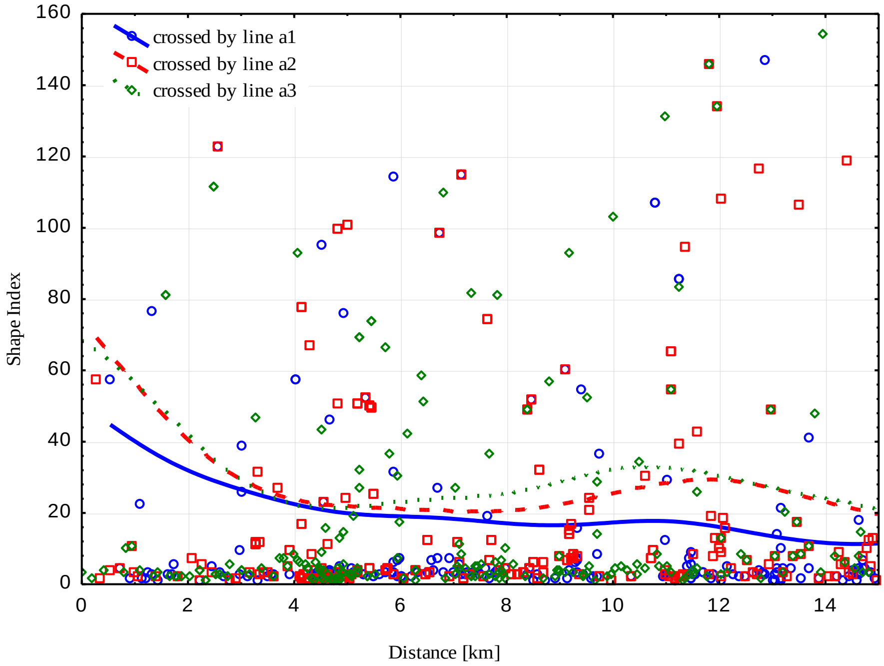



**Diagram 2.**
*Dispersion of the shape index for different groups of plots on the west side (a line smoothed out by the least-square method, weighted with distances, was fitted in)*
*Source: own compilation*

Table 2 shows an analysis of a change of the plot area depending on the distance from the motorway. The plots crossed by line **a3** have the average area of 3.5 ha (west side) and 6.3 ha (east side), whereas the plots situated in the vicinity of the motorway – 2.8 ha (west side) and 3.1 ha (east side). The dispersion of the plot area for plots crossed by the lines is shown in Fig. 3 and 4. The X presents the distance from the starting point and the Y axis – the area expressed in $m^2$.

Curves smoothed out with the least-squares difference weighted with the distances were fitted into the data set. The result confirms previous observations: the area of plots along the entire section of the motorway are smaller near the motorway.

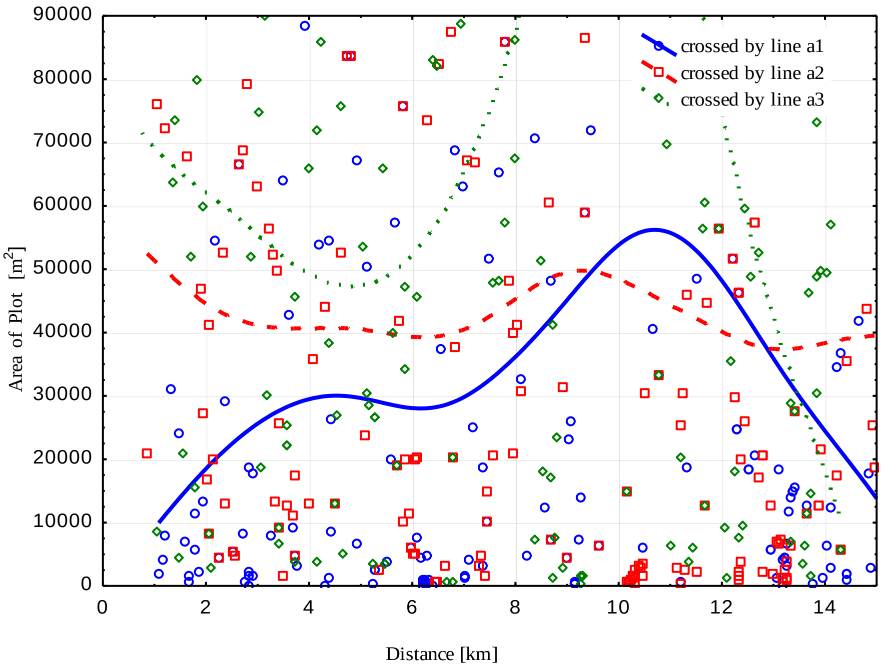



**Diagram 3.**
*Dispersion of the physical area in groups situated on the east side (a line smoothed out with the least square difference method, weighted with distances, was fitted in)*
*Source: own compilation*

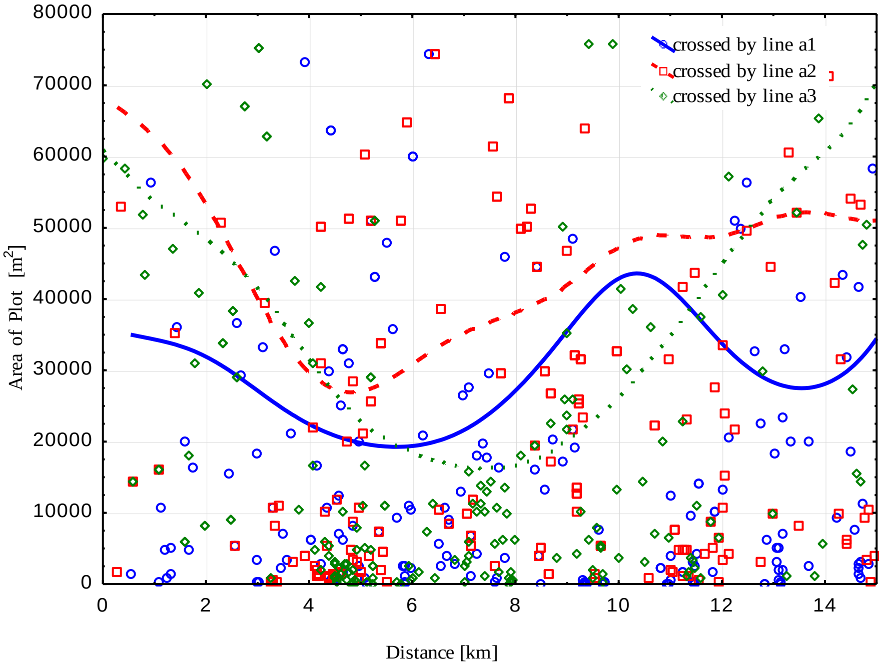



**Diagram 4.**
*Dispersion of the physical area in groups situated on the west side (a line smoothed out with the least square difference method, weighted with distances, was fitted in)*
*Source: own compilation*

In order to determine how the shape index changed as a result of the construction of the motorway, Table 3 determines the shape variability. Change of the shape $\gamma$ was determined from the ratio of the difference between the average plot shape index for the line drawn farther away and the line closer to the motorway, and the shape index for the plots cut through with the line drawn farther from the motorway.

$$\gamma=\frac{(k_{\check{s}r\_an}-k_{\check{s}r\_a(n-1)})}{k_{\check{s}r\_an}} \tag{2}$$

The results are shown in Table 3. The plots situated farther from the motorway have a higher shape index (up to 40% difference when the average shape indexes for plots cut through with the extreme lines are compared).

**Table 3**
*Change of the shape index on the lines under examination*

| Change between lines | Direction | Change of shape γ |
|---|---|---|
| a3 – a1 | E | 43% |
| | W | 33% |
| a2 – a1 | E | 27% |
| | W | 25% |
| a3 – a2 | E | 22% |
| | W | 10% |

*Source: own compilation*

Change of the area **δ** was determined from the ratio of the difference between the average plot area for the line farther away from the motorway and the average plot area for the line close to the motorway, and the average plot area for the line farther away.

$$\delta=\frac{(P_{\check{s}r\_an}-P_{\check{s}r\_a(n-1)})}{P_{\check{s}r\_an}} \tag{3}$$

The average plot area crossed by the extreme lines changes by up to 53% for the plots situated on the east side and up to 19% for the plots situated on the west side of the motorway. The results are shown in Table 4.

**Table 4**
*The index of average change of the plot area on the lines under study*

| Change between lines | Direction | Area change δ |
|---|---|---|
| a3 – a1 | E<br>W | 53%<br>19% |
| a2 – a1 | E<br>W | 24%<br>29% |
| a3 – a2 | E<br>W | 37%<br>-13% |

*Source: own compilation*

## 4. Conclusions

The morphological analysis of the spatial structure, changed as a result of a motorway construction, provided grounds for determining:

- the relationship between the location of the motorway and a change of the plot size;
- typical shape of plots situated in the area. The analysis revealed that, statistically, the plots situated farther from the motorway are more elongated rectangles than those situated closer to the motorway.

The ratio of the plot side lengths determined after the motorway was constructed, was 1:6 on the west side and 1:5 on the east side. Compared with the initial values (the side length ratio: 1:8 on the west side and 1:9 on the east side), this result is very good in terms of the plot shape. However, the average plot size decreased by approximately 40% (53% for the plots on the east side and 19% for the plots on the west side of the motorway).

A decrease in the average plot size may have a negative effect on the agricultural production on those plots.
The relatively small length of the motorway section under analysis prevents one from being able to formulate any general hypotheses regarding changes in the agricultural land morphology. However, the structure of the analysed areas was considerably regular. Therefore, it may be supposed that these findings have universal value, which will be confirmed in other studies of areas affected by motorways.

**References**

Bacior, S. 2012. Oddziaływanie autostrady na grunty rolne na przykładzie odcinka Brzozówka – Nowa Jastrząbka (The impact of the motorway on agricultural land in the example section Brzozówka - Nowa Jastrząbka), *ACTA Scientiarum Polonorum, Administratio Locorum* 11/2012: 5-14.

Bacior, S.; Harasimowicz, S. 2005. Oddziaływanie autostrady na przydatność rolniczą gruntów i ich wartość na przykładzie odcinka autostrady A-4 między Brzeskiem a Tarnowem (The impact of the motorway on the suitability of agricultural land and the value of the example section of the A-4 between Brzesko and Tarnów) . *Geodezja. kartografia i aerofotogrametria. Wydawnictwo Politechniki Lwowskiej* 66: 9-19.

Badora, K. 2004. Autostrada – środowisko przyrodnicze. Studium konfliktów przestrzennych na przykładzie opolskiego odcinka autostrady A-4 (Highway - the natural environment. Study of spatial conflicts on the example of Opole section of the A-4). *Wydawnictwo Uniwersytetu Opolskiego. Opole.* 178 p.

Bitner, A. 2011. Charakterystyczny kształt działek ewidencyjnych na terenach zurbanizowanych – analiza struktury morfologicznej miast (The characteristic shape of parcels in urban areas - Analysis of the morphological structure of cities). *Acta Sci. Pol. Geodesia et Descriptio Terrarum* 10(1)/2011: 23-32.

Bitner, A.; Hołysz, R.; Fiałkowski, M. 2009. From complex structures to complex processes: Percolation theory applied to the formation of a city. *Physical Review* E 80 (3), 037102: 1-4.

Chmielowiec, S.; Koszycki, L. 1997. Niektóre aspekty oddziaływania autostrady A-4 na grunty rolne i leśne w województwie krakowskim (Some aspects of the motorway A-4 on agricultural and forest land in the province of Krakow). *Zeszyty Naukowe Akademii Rolniczej w Krakowie nt. Metodyka oceny oddziaływania autostrady na grunty rolne i leśne.* Kraków.

Golachowski, S.; Kostrubiec, B.; Zagożdżon, A. 1974. Metody badań geograficzno-osadniczych (Research Methods geo-settlement), *PWN, Poland,* 207 p.

Harasimowicz, S. 1998. Ocena oddziaływania autostrady na grunty rolne (Assess the impact of the motorway on agricultural land). *Przegląd geodezyjny* 6: 6-12.

Kozłowski, S. 1997. W drodze do ekorozwoju (On the way to sustainability). *PWN, Poland,* 288 p.

Lenart, W. 1998. Strategiczna ocena oddziaływania polskiej sieci autostrad na środowisko (Strategic Impact Assessment Polish motorway network on the environment). *Probl. Ocen Środ.* 1: 8-13.

Marcinkowska, I. 1999. Rozprawa doktorska nt. Określenie koncepcji i zasad kształtowania struktury przestrzennej obszarów wiejskich w związku z budową autostrad (Doctoral dissertation on. Term concepts and principles shaping the spatial structure of rural areas in connection with the construction of highways). *Maszynopis, Olsztyn, Poland,* 95 p.

Pignol, V.; Delannay, R.; Le Caër, G. 1993. Characterization of topological properties of 2D cellular structures by image analysis. *Acta Stereol.* 12: 149–154.

*Rocznik Statystyki Międzynarodowej (Yearbook International Statistics) 2012.* Główny Urząd Statystyczny, Poland, 614 p.

Wilkowski, W. 1995. Wielokryterialna metoda oceny wpływu autostrady na gospodarstwa rolne (Multi-criteria method for assessing the impact of the highway on farms). *Prace Naukowe Politechniki Warszawskiej. Geodezja* 33: 5-19.

Wypadki drogowe w Polsce w 2012 roku (Road accidents in Poland in 2012). 2013, *Komenda Główna Policji, Warszawa, Poland,* 90 p.

# MODERN CONCEPT OF STREETS AND SPACE DESIGN - SHARED SPACE

**Ana Trpković[1], Branimir Stanić[2]**

*[1,2] University of Belgrade, Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, 11000 Belgrade, Serbia*

**Abstract:** Shared Space is a relatively new and unconventional traffic and urban design concept. It could be defined as a "street or place designed to improve pedestrian movement and comfort by reducing the dominance of motor vehicles and enabling all users to share the space rather than follow the clearly defined rules implied by more conventional designs". The development of Shared Space could be related to the name of Hans Monderman, Dutch traffic engineer, who presented his radical ideas of "Naked Streets" in the early 1980s in Holland. Successful achievements of his projects spread this new opportunity for street redesign through the world. Over the time, these ideas have evolved into the Shared Space concept. In fact, streets that encourage sharing of the space are not new. Many historic streets operate as Shared Space, particularly in historic city centers and residential zones. Actually, what it is about? It is clear that there is no equality in space sharing in traditional way of street operation. There is a significant dominance of motor vehicles compared to other users (pedestrians, cyclist, etc.). Shared Space concept, as its name implies, tends to make balance in street space sharing between motorized and non-motorized users. It seeks to change the way streets operate by reducing the dominance of vehicular movements through traffic design and regulation measures. Redesigning parts of the street network, using this model experienced great expansion in the last ten years through various projects throughout Europe and in the world. This paper will present the basis of the space sharing concept and its design and implementation demands, as well as traffic engineer role in this process of street and space redesign. The best European and world practice will be mentioned and also possibilities for applying Shared Space in local conditions will be considered.

**Key words:** Shared Space, street redesign, urban space, traffic design;

## 1. Introduction

"Shared Space" represents a relatively new concept of physical planning. Its beginnings can be linked to Hans Monderman (19 November – 7 January 2008) who was a Dutch road traffic engineer and innovator. The forerunner of this concept was the "Naked Streets" project initiated and implemented by Monderman in his native Friesland in the north of the Netherlands in the early 1980s. Within this pioneer project the possibility of increasing efficiency and safety of traffic operation was investigated precisely through its deregulation. The initial standpoint was based on the hypothesis that by removing traffic signs and signals, as well as usual user segregation elements, the users will be affected, their attention and mutual interaction would increase, having as the final result a safer and more pleasant traffic and urban environment. Each user is treated as an intelligent individual in direct contact with other user. Leaning on intelligence and sense of each individual, instead of traffic signs and signals, Monderman found that the traffic efficiency and safety improved when the street and surrounding public space was redesigned to encourage each person to negotiate their movement directly with others. Surprisingly good results of implementing this idea in above mentioned Dutch little town led to expansion of the initial idea about Naked Streets towards a new "Shared Space" concept. What in fact is "Shared Space"? This concept can be defined in a number of ways and it can be said that it represents a complex process, and not only a set of design measures. One of the definitions determines Shared Space as *a street or place designed to improve pedestrian movement and comfort by reducing the dominance of motor vehicles and enabling all users to share the space rather than follow the clearly defined rules implied by more conventional designs* (Local Transport Note, DfT, 2011). The idea of Shared Space, in fact, is not so new. Many historic parts of cities operate in this way, although not specially designed as "Shared Space". Actually, within this concept equality of rights of users to use street space is created/given back. This approach tends to reduce impact of motorized traffic on non-motorized users, primarily pedestrians. In other words, it aims at achieving balance in the use of space, what requires strategically planned implementation of a series of various (design and control) measures, oriented towards predefined goals. It is clear that physical planning based on the principles of this concept requires involvement of experts in various fields (traffic engineers, urban designers, town planners, landscape architects, accessibility/mobility specialists, maintenance team managers, lighting engineers...), thus in this sense multidisciplinary is one of its most significant characteristics. Synergy of these various professions is oriented towards the final goal of creating an inclusive, vibrant and convivial environment – space for people, what is the basic idea of this concept.

## 2. Basic Principles of the Shared Space Design

Design of Shared Space represents a complex process which, through redesign of the existing urban space, increases its functionality and establishes a more balanced (more equal) relationship between its users. As already mentioned, this process requires multidisciplinary and planned approach to each stage: from the project initialization (by interested parties), designing, implementing, and finally to maintenance and monitoring. At the very beginning the following questions should be answered: What is required, why it is required and how will we deal with it?

[1] Corresponding author: a.trpkovic@sf.bg.ac.rs

Generally, not a single space will be redesigned into SS just to make a change for change's sake, but to do so there must be justified reasons (traffic, spatial, social, economic), which will be established precisely from the answers to these questions. In this respect, the obtained answers show to us the vision, purpose and finally the actions to be undertaken in order for this concept to be established.

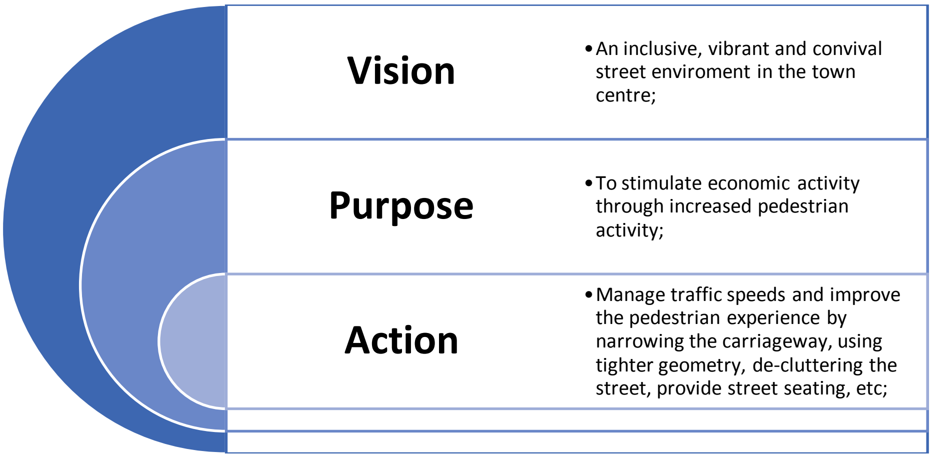



**Fig. 1.**
*An example of Vision, Purpose and Action*
*Source: (LTN, Dft, 2011)*

Traffic engineers have key and unavoidable role in each stage of this process, from location planning and selection to implementation of concrete measures, choice of materials, maintenance and follow-up of the newly created traffic-spatial conditions. Of course, active involvement of professionals in other fields is necessary in order for the newly created space to fulfill all its functions in the best possible way.

### 2.1 Shared Space Design Requirements

Shared Space concept can be applied to different traffic and spatial conditions; however, surely there are certain limitations and requirements. When speaking of such physical planning, one primarily thinks about streets and intersections (squares). Well designed streets and intersections can offer opportunities for recreation, social interaction and physical activity and also could have a positive impact on local economic performance. If a street does not perform well for people wishing to spend time in it, it is an indication that its place function is too low (LTN, 2011). The relationship between place and movement is presented with the place/movement matrix, a concept introduced in the Manual for Streets (DfT, 2007) - Figure 2.


**Fig. 2.**
*Place/movement matrix*
*Source: Manual for Streets, DfT, 2007*

This place/movement matrix shows how movement function and place function, depend on trip type. Spatial function significantly increases within Shared Space, but not always at the expense of movement function. In fact, it is extremely important for the streets to retain movement function, if we really want to have Shared Space.

In some cases it is possible that even movement function increases within the scope of such spaces, owing to total decrease in vehicle delays.
It is precisely the significance of movement function in urban spaces that makes it necessary to take into consideration the most important traffic parameters, flow and speed, in the function of Shared Space. Traffic flow (q) and vehicle speed (v) are necessary design inputs that need to be taken into account. In this regard it is necessary to define some boundary values of these parameters suitable for this concept. If we first take a look at traffic demand, some experiences and research work indicate that the flow (Q) exceeding 100 veh/h is recognized by pedestrians as the boundary value. Above this value they do not feel comfortable and equal, so they just want to cross instead of occupy the space. It is clear that this is a very low value of Q, but it does not mean that Shared Space cannot be applied also to the network parts with considerably higher vehicle intensities. Good results of application of this concept have also been noticed in the environments with higher flows and the value shown in no way presents the upper limit for reshaping the space into Shared Space, but it shows the level of willingness of pedestrians to use all of the street space. In this regard, it is exactly vehicle speed (v) that has significant impact on willingness of pedestrians to share space with drivers. The higher the permitted speed, the pedestrians are less willing to share. So, for the Shared Space desirable design speed[3] is less than 30 km/h, and preferably less than 25km/h.
Taking into consideration the defined traffic restrictions, but previous experiences as well, application of this concept is particularly suitable at: historic places (like city centers or other places rich in cultural or/and historic heritage); places with high pedestrian flows; places where pedestrian desire lines are various; but also in other different traffic and urban conditions. Shared space has also been applied to some arterial routes, restoring their traditional place functions. Home Zones and some country lanes, particularly those with a Quiet Lanes designation, tend to operate as shared spaces, too. *(Local Transport Note, DfT, 2011)*

## 3. Shared Space Design Measures

With introduction of Shared Space concept, character of chosen street/space changes. Precisely because of that it is needed, first of all, to understand how such space operates in the existing state. In this regard each design is preceded by a serious research work on gathering data on traffic, attitudes of users and interested parties, existing patterns and habits. This data should serve as indication and assistance to the design team when choosing adequate measures for transforming a street into Shared Space. In addition, the collected data, to some extent, enables making comparison of various spaces and behaviour of their users. Designing Shared Space generally falls into three categories:

- physical and operational;
- behavioral;
- materials, implementation and maintenance.

Each individual space has its requirements as regards key activities or use: pedestrian movement, vehicle movement, provision of comfortable space for pedestrians, enabling various events, space for social gathering, parking, public transport (locations of stops), service function (e.g., delivery of goods). Depending on the dominant function(s) and needs, balanced use of space by users is established by applying adequate measures. Unconventional physical and control measures, along with adequate choice of materials for their implementation, lead to change in the space itself, but also in the behavior of users. Their mutual interaction and eye contact are becoming a very important factor for efficient and safe operation of traffic process. The most important and most frequent control-design measures applied to reshaping the conventional public spaces into Shared Spaces are:

**Relieving the environment** by way of: deregulation, i.e. elimination of conventional standard traffic signals, removal of curbs, changes in use of surface materials. Precisely these measures make this space radically different in terms of traffic.

**Designing the space for low speeds**: by use of different base courses and/or by changing surface appearance; reducing traffic signals and other control measures, by placing street furniture at unexpected and unusual places (e.g., in the middle of the street), bicycle parking lots, by planting trees, visually narrowing and reducing sight distance and by using a less comfortable geometry.

**Courtesy crossing** is usual name for the crossings within such street spaces. The reason why such crossings are called courtesy crossings can be found in the fact that in Shared Spaces usually there are uncontrolled crossings (although in some cases presence of controlled crossings is necessary as well) and drivers are not required by law to give way to pedestrians, but they do so out of “courtesy”. Designing of courtesy crossings usually implies use of materials of various texture, carriageway narrowing to create spatial limitation for vehicles, by placing bumps (unless at-grade space is in question), by use of various vertical barriers (e.g. poles), etc.

[3] Design speed is a target speed that designers intend most vehicles not to exceed and is dictated primarily by the geometry of tracked vehicle paths within the street.

Location and frequency of the crossings themselves depend on the pedestrian desire lines that can be obtained through analysis of the data on existing state, but also from forecast of attractions within a newly developed space.

**Specially shaped transition zones (Gates)** – Transformation of a conventional street into Shared Space can be marked by various measures used individually or in combination, such as: road width narrowing, visual narrowing, specially designed portals, bumps, changes in materials, other ways of marking (particularly when there is significant difference in street/road categories between zones).

**At-grade surface** should not be a designing goal by itself, but it should be the response to the desired level of sharing. When we want to emphasize attractiveness of a public area and when parts of a space are understood predominantly as pedestrian and bicycle spaces, it is suitable to initiate design of such surfaces. In such cases motor vehicles still tend to remain in the middle of the street, what can be intensified by use of tonal contrast on the surface material or by using street furniture. Other ways of physical segregation between motorized and non-motorized traffic are minimized.

**Fig. 3.**
*Example of a Shared Space gateway feature and Courtesy crossing*
*Source: Local Transport Note, DfT, 2011*

**Design of parking and loading areas** in Shared Space streets requires careful consideration, particularly when at-grade space is in question. These functions should be organized in such way not to hinder other traffic participants (particular care should be taken about persons with special needs). On the other hand, parking can be used as a means of space segregation, speed control, etc. When organizing parking within such Shared Spaces, special care should be taken not to harm the esthetics of the redesigned space by parking places marking.

**Public transport within Shared Space** – When reshaping space the existing Public Transport routes and/or locations of stops should be taken into account. It is necessary to make consultations with bus operators during the planning stages and in accordance with the potential to locate stops on the zone fringes and to make route adjustments, if any, taking care of users' needs, but also about the changes to be brought by such redesign.

**Materials** – Special emphasis should be placed on the significance of materials and urban furniture to be used in such spaces. Choice of high-quality and suitable materials does not mean that they must be expensive. On the contrary, it should be borne in mind that such spaces need to be maintained, thus when selecting equipment and surface base course this should be taken into account. Here function of materials is multiple, street furniture should be multipurpose, if possible, besides its basic function to have another use value as well. Concept of such spaces to their users should offer socially pleasant surrounding by providing certain number of places where they can sit and rest and in this way make such street space more pleasant.

**Lighting -** Lighting is an important feature in Shared Space schemes, and hence it becomes unavoidable part of such projects. From the point of view of traffic, lighting has an important role in safe operation of traffic, thus in this unconventional traffic urban environment the role of lighting gains in importance both because of esthetic and safety reasons.

**4. World and Local Practice Short Overview**

The concept of Shared Space urban shaping has its roots in the Netherlands. It originated by evolution of the Woonerf concept and afterwards Hans Monderman's Naked Streets project. One of the first and most important *Shared Space* research projects was conducted between 2004 and 2008 by seven project partners from five countries of the North Sea Region. The programme was supported by the European Interreg IIIB North Sea Programme and by various local authorities involved. The five countries involved were the Netherlands, Belgium, Denmark, Germany and the United Kingdom. Seven pilot projects were substantially completed during the project timescale, bringing a range of physical changes to a variety of differing examples of urban and rural contexts. The contrasting contexts of each of the projects have provided the opportunity to explore different aspects of *Shared Space* principles in a variety of forms. (Interreg IIIB project Shared Space, 2008) This programme has provided opportunities to test the principles of *Shared Space* in a wide range of different conditions, and to generate transnational exchange of knowledge between five countries and seven municipalities. It has also facilitated understanding of *Shared Space* across the rest of Europe and much of the rest of the world through the unusually high levels of publicity generated. As a result, *Shared Space* is now an established set of principles in many countries. From the European continent the Shared Space concept spread to other parts of the world and found its use in Australia, New Zealand, Japan, USA, Canada,etc. Even in cities where car dominance is significant and traffic network is oriented to motorized modes this idea has found it application. Many cities implement shared spaces or other mixed traffic concepts in order to revitalize city centers and provide alternatives to car-centric road designs.

**Fig. 4.**
*Examples of application of Shared Space concept: UK, New Zealand, USA, Japan*.

As far as our country is concerned, Shared Space as a concept of traffic-related physical planning of urban space is almost non-existent. Some individual traffic calming programs or designing the zones with "calmed" traffic have been carried out in previous years. The measures applied are short-term and their number is very small. In addition, these projects most frequently were not based on systematic and strategic approach, and were without clearly defined goals. Studies of relevant traffic parameters are minimized, if any were conducted at all. Consequences of such approach are poorly implemented programs with almost unpredictable effects. Also monitoring of applied measures mainly comes down to safety aspect, which is important, but neither the only one. There is an urgent need for planned reconstruction of urban areas throughout the country which were devastated and abandoned in last decades. Thus, application of Shared Space concept should be seriously reconsidered through this process.


**Fig. 5.**
*Shaping street space in Belgrade*

## 5. Conclusion

Shared Space represents a new traffic and urban design concept of reshaping streets, junctions and public spaces. Basic idea is removing the segregation between motorized and non- motorized modes and creating an integrated space without conventional traffic signalization and markings. Instead of traffic signs or signals, curbs and road markings, traffic flows are controlled by social interactions and supported by infrastructure design measures. Shared Space views public spaces, first and primarily, as spaces for people. It means that public spaces should be design to people – not to restrict them. Within this concept, the dominating traffic function become equal to other public space functions through implementation of unconventional traffic design and regulation measures. In this paper are presented some of the most important and mostly used traffic measures for redesigning public spaces. The popularity of Shared Space increase in recent years through the Europe and World as it is seen as a chance for reducing car dominance in cities. Now, we have many examples of successfully redesigned urban spaces which operate as shared space. Unfortunately, in our country there isn't urban area which is planned and designed according to the rules and principles of this concept. Some public spaces are reshaped with traffic calming measures, but these actions represents separate and short time improvements of local conditions. Generally, design of Shared Space is a complex process which, through redesign of the existing urban space, increases its functionality and establishes a more balanced relationship between its users. Traffic engineers have key and unavoidable role in every stage of this multidisciplinary process which final goal is creating an inclusive, vibrant and convivial environment – space for people, what is, in fact, the basic idea of this concept.

**References**

Department for Transport (DfT). 2002. *Inclusive Mobility – A Guide to Best Practice on Access to Pedestrian and Transport Infrastructure*. London, UK. Available from Internet: http://www.dft.gov.uk

Department for Transport (DfT). 2005. *Home Zones – Challenging the Future of our Streets.* London, UK. Available from Internet: http://www.dft.gov.uk

Department for Transport (DfT). 2007. *Manual for Streets. Thomas Telford.* London, UK. Available from Internet: http://www.dft.gov.uk

Department for Transport (DfT). 2011. *Local Transport Note 1/11 – Shared space.* London, UK. Available from Internet: http://www.dft.gov.uk

Hamilton-Baillie, B. 2008. Shared Space: Reconciling People, Places and Traffic, *Built Environment*. Vol. 34, No. 2,161-181.

Hamilton-Baillie, B. 2006. *What is Shared Space?*Available from Internet: www.hamilton-baillie.co.uk.

Interreg IIIB project Shared Space. 2008. *Shared Space - Final Evaluation and Results*. Netherlands.

Monderman, H.; Clarke, E.; Hamilton-Baillie, B. 2006. Shared Space - The Alternative Approach to Calming Traffic. *Traffic Engineering and Control*. Vol. 47, No. 8, 290-292.

Trpković, A.; Stanić, B. 2012. Shared Space - Present State in Europe. In *Proceedings of the X International Conference of Traffic Control Techiques (TES), Subotica.*

Trpković, A.; Stanić, B. 2008. Traffic Deregulation in Small and Medium Towns – *Naked street* project example. In *Proceedings of the VIII Regional Conference of Traffic Control Techiques (TES), Sombor.*

# PEDESTRIAN INFORMATION SYSTEM IN THE FUNCTION OF EMPOWERMENT OF CITY IDENTITY

**Milena Vukmirović[1], Eva Vaništa Lazarević[2]**
[1] *Research Associate, University of Belgrade – Faculty of Architecture*
[2] *Full Professor, University of Belgrade – Faculty of Architecture*

**Abstract:** The purpose of this paper is in presenting the potential of immediate pedestrian environment that contribute in the construction of the identity of the city. The proposed approach is based on the conclusions of the research on the communications potential of the city and its unique and redundant artefacts. Special attention of this paper is given to pedestrian information system that belongs to the category of urban furniture and the effects of its implementation in cities. This research puts in evidence the importance of urban furniture in generating distinctive physical identity of the city. In addition, the contemporary approach in designing and improving urban furniture highlights the potential for redundant artefacts to contribute in creating an smart city and the way the city is used by its inhabitants and visitors. The outcomes of this paper are in emphasizing the importance of the elements of the physical environment in the city that are in the direct contact with pedestrians. In addition the pedestrian information system can be seen as navigational system that help pedestrians determine where they are and where they need to go to reach a destination. On the other side it has its own identity that can also contribute to the identity of the city as whole.

**Keywords**: Pedestrian information system; Pedestrian environment; Identity of the city; Unique and redundant artefacts

## Introduction

An emphasis on the pedestrian environment has the intention to highlight the significance of small scale, often neglected in contemporary projects and development strategies[2]. Visions, which correspond to this angle of perspective, put focus on specific advantages of cities proportional to dimensions, senses and walking speed and form a basis for more complex and diverse relations. More concretely, they correspond to improving the quality of the smallest places in such way that people are simply attracted to go there and spend time there. At the same time, urban design and planning on human size has the task to encourage intensity of pedestrian movement as part of integral urban policy with an aim to develop lively, safe, sustainable, healthy, as well as legible cities.

Legibility is a term used to describe “the ease with which people can understand the layout of a place” (Linč 1974). It became popular during the sixties as an urban planners’ reaction on destructive influences of modernism on American cities and urban decline. For those authors3 the legibility could be achieved by the renewal of social and symbolic functions of open public spaces. In accordance with this, the loss of the human dimension of space was the reason why they observed the city from the perspective of city’s inhabitants and everyday users. On the basis on the concept of legibility and having in mind that “every citizen have deep associations related to the particular area of the city and associated with personal memories and meanings”, Lynch (1974) has recognised five elements4 that define the image of the city and that help people to find their way around and understand how a place works. Cities where those elements can be clearly recognized offer much more than visual pleasure and emotional security, because they increase the depth and intensity of human experience (Madanipour 1996). On the other side, Golledge (1999) has found that the familiarising with the environment is the best achieved by movement that include physical recording of the parts of the network structure and cognitive processing and memorising the knowledge achieved during the movement and way finding. In this context, singling out and analysing specifics of pedestrian movement and elements that influence it determine the scope of pedestrian space. This scope is related to the scene covered by pedestrians’ sight, where at the same time the most intensive contact is established by ground floors of objects and impressions of space located in the distance. Besides the stated, pedestrian environment is to be observed from perspectives much wider than those of traffic, since it covers important elements concerning living conditions and human capabilities in a city (Gehl 2010). Having in mind the conclusions dealing with the relation of pedestrian vs. environment, walking is most intimate with environment and it enables much more articulated processes of interpretation and memory.

Communication of the city is seen in relations to models that are based on non-verbal communication and models and territory of synergetic inter-representative networks (SIRN). The first approach, based on non-verbal communication of urban environment (Rapoport 1990), has enabled separation of elements of pedestrian environment that influence the creation of its *identity*. They encompass fixed (walls, columns, objects etc.), semi-fixed (urban furniture, details, etc.) and changeable elements (people and their non-verbal communication). On the other side, the SIRN theory has enabled the determination of the face of pedestrian space, based on quality and quantity of information that it carries. By questioning the concept of city legibility, defined by Kevin Lynch and presented by five elements that form “the image of the city”, theory of SIRN, Portugali (2011) recognizes “the face of a city”.

[1] Corresponding author: milena.vukmirovic@urbanlab.org.rs
[2] M. Vukmirovic, Pesacki prostor i kompettitivni identitet grada, Beograd: Zaduzbina Andrejevic, 2014
[3] Lynch, Jacobs, Cullen, etc.
[4] Paths, edges, districts, nodes and landmarks

The face of the city is formed by elements – external representations, which have the capacity to cover and transmit the information. Information is divided into Shannon's, which include informational capacity of channels of communication and semantic, which are in relation to the meaning, a way of experiencing space, having in mind the specifics of a recipient.

Thus, besides Lynch's, the elements of the face of the city also include other buildings, streets, parks, etc. The difference between them lies in the level (quantity) of information they contain, and are therefore more or less important for the legibility of the city. It follows that the face of the city is the message. Established relations are presented in the form of pedestrian space communication model (PSCM) with clearly demonstrated phases of communication process and mutual relations between pedestrian space experiences, characteristics of physical structure and identity of a city. Formation of PSCM is based on existing marketing models of identity and image communication applied to products, companies or places.

Special attention of this paper is given to pedestrian information systems as a form of urban furniture and equipment. Observing them as redundant artefact defined by Portugali, the city could transmit the message of its technological progress and developmental objectives, and, at the same time, improve the legibility, the quality of everyday environment and life in the city. On the other side this kind of information system is used for pedestrian orientation in the city, which could facilitate movement of inhabitants and visitors, as well as their participation in the overall transport system of the city. The paper is composed of four parts. The first part explains the importance of coordinated small urban interventions, i.e. urban furniture in the improvement of the city identity presented in the form of pedestrian space communication model. Methodology and research polygons selection is in the second part. The research results presented in the form of review of the situation in five European cities are in the third part of the paper. The fourth part of the paper deals with the review of the contemporary forms of pedestrian information systems as integrated programmes of transportation, information and identity projects, its characteristics, as well as technologically advanced elements of urban furniture that aim to improve peoples' understanding and enjoyment of the city. The concluding part will summarise the role of integrated pedestrian information system in empowerment of legibility and overall identity of the city.

## PEDESTRIAN SPACE COMMUNICATION, REDUNDANT ARTEFACTS AND PEDESTRIAN INFORMATION SYSTEM

By observing communication from a cultural aspect, every cultural phenomenon can be observed from the communication perspective, i.e. as a way of communicating. Focusing on defining the relationship between pedestrian spaces and identity of the city, pedestrian space communication model (PSCM) is formed, on the ground of the knowledge from the marketing and branding models, non-verbal communication of urban environment and theory of synergetic inter-representative networks (SIRN). PSCM is based on existing marketing models (Balmer & Gray 2000; Kavaratzis 2008) of identity and image communication applied to products, companies or places. Observing communication of a city as threefold system covers *primary, secondary and tertiary communication* (Fig 1). Primary communication encompasses communication effects of city activities, by which the communication itself is not the aim of these activities. It is divided into four broader areas of intervention: landscape strategies, infrastructure projects, organizational and administrative structure, and city's behaviour. Secondary communication represents formal, intentional communication which is found in well-known marketing instruments, such as in-door and out-door branding, public relations, graphic design, use of logos etc. Tertiary communication refers to communication marked as word of mouth, which is amplified through the media. With respect to its character the whole process of branding and the other two types of image communications are intended to encourage and reinforce positive tertiary communication (Kavaratzis 2008). This is particularly the case of the city residents and visitors, which are the most important target audiences in city branding and the most important participant in place marketing.

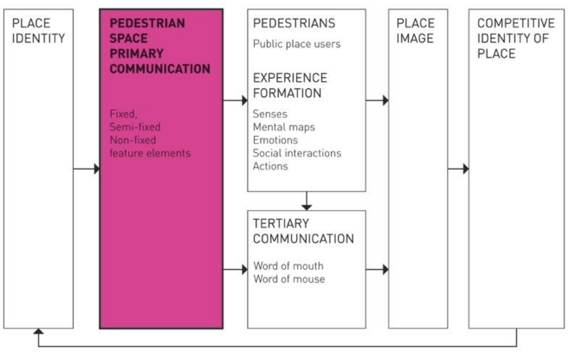



**Fig. 1**
*Pedestrian space communication model (PSCM)*
*Source: Vukmirovic, 2014*

Pedestrian space communication model covers certain specific phases place communication based on real place identity – identity of pedestrian spaces. Identity of a city is presented by primary – non-verbal communication that encompasses fixed, semi-fixed and changeable elements of pedestrian space. Interacting with the immediate surrounding the experience of urban environment i.e. a city is created. The experience will to a large extent participate in formation of individual image, which is built into common image of that space, but the city in general.

Besides the stated, depending on intensity of formed experience, there is a possibility of establishing tertiary communication, both by direct communication with other people and through social media. Tertiary communication also has a role in formation of common image of a city. Depending on type of the established image and degree level of image within population, communication process will result in formation of competitive identity of a city. When observed in this way, identity of a city and competitive identity can be seen as synonymous. By establishing this analogy, competitive identity can be seen as final and starting result of a process identity-positioning-image, because it placed in the process of city's identity becomes more competitive.

On the basis of the theory of synergetic inter-representative networks, Potugali (2011) recognizes the face of a city, which is formed by elements – external representations, which have the capacity of covering and transmitting information. Information is divided into Shannon's, which include informational capacity of channels of communication and semantic, which are in relation to the meaning, a way of experiencing space, having in mind the specifics of a recipient. Determining the pedestrian perspective to a city has enabled determination of the face of a city, perceived from the pedestrian perspective. Having this in mind, it is concluded that pedestrian space encompasses both classes of urban artefacts – unique and redundant. Unique artefacts are urban forms that carry the high value of Shannon's information. They can be differentiated in terms of character of semantic information that defines the source of uniqueness of these urban elements. On the basis of this they are divided into geometrical, symbolic and legendary unique artefacts. Redundant artefacts encompass those objects, which because of their multiplicity (repetitiveness) form categories with different features. Suitable way to differentiate them is establishment of categories in relation to ratio. By applying this four subtypes are identified: *urban furniture* (open space equipment), *objects*, *urban sceneries*, and *street network*.

In this context, pedestrian information system is observed as form primary communication of the pedestrian environment in the area of landscape strategies and as a type of semi-fixed element of pedestrian space. Following the definition of redundant artefacts it can be seen as specific open space equipment, or urban furniture.

**Methodology and research polygons**

Pedestrian information systems are used for spatial orientation of pedestrians, which facilitates pedestrian movement of citizens and visitors, and their joining with the overall traffic system of the city. For research purposes, we divided it into three groups depending on the type of information: 1) information for the orientation purposes, 2) information for the purposes of establishing a connection with other types of public transport and 3) others, complementary forms of information. The case study method was used in analysis of the actual state of the pedestrian information system in the monitored cities. Method of comparative analysis is used in order to summarise the results obtained using case studies.

**Identified polygons**

European cities, which were the polygons of the research, were selected on the basis of the set of criteria that includes (Vukmirovic 2013): 1) ESPON categorisation, 2) possess car-free cities affiliation, 3) sustainable transport and quality of open public spaces provides a platform for city brand creation, 4) the success of cities in achieving positive reputation by establishing and developing the appropriate network pedestrian space and 5) have developed a network of pedestrian spaces in the city centre. The identification of the cities that were the subject of the analysis was carried out by a system of elimination, in accordance with the mentioned criteria. Considering the fact that ESPON categorisation has 5 categories of cities, the 5 cities were identifies. The cities that were selected are: 1) **Paris** (Global node) – has developed pedestrian space network in the central area, ranked in the first 10 cities which have high quality of open spaces, highly ranked by the consultant offices in relation to the city brand; 2) **Munich** (MEGA 1) - has developed pedestrian space network in the central area, in the opinion of its citizens is the first in relation to the quality of open spaces, the third compared to the beauty of the streets and buildings, and the second in terms of recreational facilities, is among the top ten cities in relation to the quality of life, city offer and city brand capital; 3) **Vienna** (MEGA 2) - has developed pedestrian space network in the central area, in the opinion of its citizens is among the top ten cities in relation to the quality of open space, the beauty of streets and buildings and recreational facilities, according to the Mercer criteria Vienna is considered as the best city to live on a world scale, while the lists of other consulting companies very highly ranked it as well (e.g. Vienna is the second world city compared to the criterion of vitality by EIU ranking); 4) **Malmo** (MEGA 3) - has developed pedestrian space network in the central area, is very highly rated according to thoughts of their inhabitants in relation to the quality of open space (2), the beauty of streets and buildings (2) and recreational areas; and 5) **Ljubljana** (MEGA 4) - has developed pedestrian space network in the central area, in the opinion of its citizens is well characterized with respect to all three criteria and a special weight was provided by EIU survey, which puts Ljubljana on the 13th place in relation to improvements in cities and the use of sustainable modes of transport

**REVIEW OF THE PEDESTRIAN INFORMATION SYSTEMS IN FIVE EUROPEAN CITIES**

**Paris**

In Paris, pedestrian information system is used in order to spatially orient pedestrians and to facilitate pedestrian movement of citizens and visitors, and their participation in the overall public transport system. There have been observed four types of information for the purpose of informing pedestrians in Paris (Fig. 2): 1) information for the orientation purposes, 2) information for the purposes of establishing a connection with bus transport, 3) information for the purposes of establishing a connection with underground, metro transport and 4) others, complementary forms of information. All these types of pedestrian information elements have a distinctive visual identity, which is reflected in the characteristic and precise graphic and product design (form, font, colour, tags, etc.).

For the purpose of pedestrian orientation the street signs and vertical elements are used. They are placed at important street intersections or large open public spaces as table with the names of important pedestrian destinations. The tables are placed horizontally and are oriented such that indicate the direction which should navigate to these destinations. On some more frequent places these elements could be improved by the addition of maps of the immediate area.


**Fig. 2**

*Elements of the pedestrian information system in Paris.*
*Source: Flickr.com*

In order to facilitate the movement at the level of the whole city, multiple modes of transport are combined. This refers to the movement transvers between waking and other types of public transport, which are determined by the elements of information that are placed at bus stops and entrances to metro stations. In addition to a clearly defined system of information (visual identity: poles, stations, metro signs, etc.), in relation to the indication of the type of transport at these locations, pedestrians can be accurately informed about the immediate and wider environment over the map (bus stops) or three-dimensional spatial imaging (metro), as well as the arrival time and specifics of the transport route.

The types of information that belong to other complementary groups include signs which point to the construction sites, damages on the direction of movement, etc. It can be concluded that for this type of pedestrian information elements there is somewhat established visual identity. The recognition of the pedestrian information system that is used in Paris and its importance to the general appearance of the city is supported by the places where some of these signs may be purchased as souvenirs.

**Munich**

Pedestrian information system is partially developed in Munich and includes three types of information: 1) information for the orientation purposes, 2) information for the purposes of establishing a connection with other types of public transport (U and S-Bahn, trams, buses and public garages) and 3) others, complementary forms of information (Fig. 3).


**Fig. 3**

*Pedestrian information system in Munich: exiting and planned.*
*Source: http://flickrhivemind.net*

Signs about street names are placed on poles and oriented in the line with the direction of particular paths. They are recognisable by the traditional font in white colour placed on the blue background. Signs with information for other modes of transport are on vertical poles and placed on the sites of the tram and bus stops, or entrances of the U-Bahn and S-Bahn.

There are also other types of information on these locations (within the stations) in the form of city maps, maps with public transport lines, a digital display that shows the time of arrival of the transport, watch, etc. - Which allow easy reference and information users, enabling easy user information and connection with the other types of transport, as well as other parts of the city. The third group of information includes various types of information and warnings that point to the existence of construction sites, route damages, and so on.
The research of the existing information system in Munich indicates that the city doesn't have integrated pedestrian information system. Concept called „Digitale Signaletik“ (Eng. “Digital signage”), which is in the development phase, can be considered as a segment of developing this idea, based on advanced technologies.

**Vienna**

Here is no integrated pedestrian information system in Vienna, but pedestrians could get three types of information in public space: 1) information for the orientation purposes, 2) information for the purposes of establishing a connection with other types of public transport (U-Bahn, trams, buses and public garages) and 3) others, complementary forms of information (Fig. 4).
Signs with street names are located at the beginning of the street, placed on the facade of the building. These signs include the name and the number of the district, which is also important in orientation in Vienna. Street name is written in a traditional font, in black on a white background. The signs, which mark the tram and the bus stops, have retained traditional look. A map of the city clearly shows the location where the user is located, other movement options and the particular paths and directions nearby. Sign that mark the entrance to the underground is placed on the pole, while the name of the station is located above the entrance. The third group of information includes various types of information and warnings.


**Fig. 4**

*Elements of pedestrian information system in Vienna.*
*Source: Flickr.com*

**Malmo**
Malmo has “Program for street information boards and advertising systems” (Gatukontorets informatörer varit 2007) that defines the elements used for informing in the city (Fig. 5). The need for these elements emerges as a consequence of the development of the city and greater touristic visits, in order to facilitate the orientation and movement in the city and to strengthen the identity of the place. The Program and rules include: 1) the information tables that contain information on the history and attractions, plan of the surroundings and the map of various services such as a toilet, phone, etc.; 2) signs which mark the ambience with special purposes such as picnic areas, roasting grill areas, beach, etc. and are complemented by an ordinance of the terms of use of space; 3) signs used in parks and landscaped areas for information about specific trees and plants; 4) bulletin board where people can advertise and leave messages about events in a city or neighbourhood; 5) information on the reconstruction or construction of roads and trails and warnings related to building sites with the name of the department or institution responsible for the implementation of these activities.


**Fig. 5**

*Elements of pedestrian information system in Malmo.*
*Source: www.malmo.se*

In addition to the general elements of information, there is information that is used on the tram and bus stops. They refer to the lines of public transport, starting and final destination, and information about the arrival time. For the purposes of public information graphical rules that must be followed are determined. Graphic design includes fonts, colours and official logo of the city.

**Ljubljana**

Pedestrian space network in the centre of Ljubljana hasn't integrated information system. Three types of information in public space are identified in the city: 1) information for the orientation purposes, 2) information for the purposes of establishing a connection with bus transport and 3) others, complementary forms of information (Fig. 4). All these elements have a distinctive visual identity, which is reflected in the way of design and precise graphic rules (font, colour, tags, etc.). For the purpose of pedestrian orientation tables with street and important destination names are placed at major intersections or large open spaces. The tables are placed horizontally, and are oriented in that way to indicate the direction to these destinations.

A three-dimensional model of the central part of Ljubljana is places on Prešeren Square (Fig. 6a). This model represents one of the symbols Prešern square and Ljubljana at whole, but it is also very important for the orientation in the city. In order to orient and inform pedestrians, the company Trimo is organizing an international competition for the design of multifunctional urban installations. Some of the winning solutions are implemented and installed in urban places of the city. Information pavilion, designed by Rok Grdis is among them. It is a very noticeable structure consisting of six segments of differently shaped, coated with red glossy metal. On the inner side there are lamps and panels for setting information. The pavilion is set in Tivoli park (Fig. 6b)

**Fig. 6**

*a) 3D model of Ljubljana and b) Information pavilion in Tivoli park.*
*Source: www.flickr.com and www.trimo-urbancrash.com*

Summarised results showed that there is no integrated pedestrian space information systems in analysed cities and the existing elements are used for this purposes. In all cases they represented elements for information in order to orient users of open space include: 1) street signs with distinctive visual identity and 2) tables with the names of important locations, which indicate the moving direction and destinations. They have a distinctive visual identity that is characteristic for each analysed city. The only example could be found in Munich, but it is still in development phase, titled "Digital signage". Second type of information includes elements that are placed within the standpoint of public transport. They offer the information of the position of station and have distinctive visual identity and they are easy to detect. Other information on these sites includes timetable information and destinations that appear on digital displays. Third type of information stands out complementary facts about building sites, route damages, etc.

**Review of the contemporary forms of pedestrian information systems**

In addition to the exiting pedestrian information systems that are analysed through the evaluation of particular elements in selected European cities, the elements that are technologically advanced than others are observed in Paris and Munich. Paris is characterised by elements of pedestrian information system created as a result of the project called Intelligent Urban Furniture[5]. On the other side, the concept of integrated pedestrian information system is developing in Munich. The implementation of both projects is in initial, experimental stage.

**Intelligent pedestrian information system in Paris**

The project "Intelligent urban furniture of Paris" has been initiated in early 2011 and led by Anne Hidalgo, actual Mayor of Paris, and Jean-Louis Missique. The idea was organized as a public call, with professionals who have been invited to contribute in order to conceive Paris as "the city of tomorrow". The main focus of the call was to give participants complete freedom to suggest solutions, which would make Paris more pleasant and efficient. Observed as a city of creation, invention and revolution, the project was aimed to achieve an image of a city that would contribute to "a city of innovation, a great showcase of new technologies and services" (Missika 2013). There were more than 50 participants to the call while 40 projects made it to a final selection, for which the prototypes have been produced. By analyzing intelligent elements of urban furniture in Paris it has been established that out of 40 projects, 25[6] belong to pedestrian space.

[5] *Mobilier Urbain Intelligent*

[6] On the basis of their dominant function it can be concluded that they incorporate elements which belong to the following sub-categories: (1) commercial open space contents, (2) seating spaces and elements, (3) cafe gardens, (4) pedestrian information system, (5) other methods of informing pedestrians, (6) advertising points, (7) internet and multimedia services, (8) lightning, (9) public space art and (10) multi-sensory effects (Vukmirović 2013, 503-509).


**Fig. 7**

*a) Totem digital, b) Decodeur urbain and c) nAutreville.*
*Source: Intelligent Urban Furniture of Paris*

Projects of pedestrian information system include: “Mobulles par Connecthings”, “Borne muséographique pour cimetières”, “Décodeur urbain”, “Totem digital”, “Plan tactile ViaDirect City”, “Oxialive colonne numérique”, “Smart borne sur candelabra”, “Pelle intéractive Tracetel”, “EchoLink panneaux intéractifs” and “nAutreville”. Element of urban furniture called “Mobulles par Connecthings” has a relay function for digital information. It is meant to inform the visitors and local inhabitants on useful information and entertaining contents. It is accessed by digital code (bar code), which is read by the mobile phone. Stickers with “smart etiquettes” are placed on the existing elements of urban equipment. “Borne muséographique pour cimetières” is a multimedia terminal in which there are information that concern localities of cemeteries. This element of urban furniture enables direct insight into content on the stated locations and overtakes of maps with locations of tombs by using mobile phones. “Décodeur urbain” has the characteristics of an over-sized smart phone (Fig. 8b). It serves in some ways to “decode” the city Practically, it is a type of urban guide and the very application is designed so as to enable easier way to get to know the city and specific locations, which is achieved when a user searches the content, depending on his/her own preferences. “Totem digital” is a big screen placed on a vertical carrier, which gives information in real time (Fig. 8a). It uses all the advantages of digital technology and transmits excellent picture quality, even in strong sunlight. In real time it displays news from Paris as an official Twitter feed of the City of Paris. “Plan tactile ViaDirect City” is a two-sided terminal organized as Via Direct software solution that quickly helps orientation and provides the ability to access information important for moving in the city. Urban furniture named “Oxialive colonne numérique” equipped by defibrillator and a touch screen specially suited for people with reduced mobility. “Smart borne sur candelabra” are info-terminals placed on candelabras. Used as clocks, short-term supplement for electronic vehicles on 2 or 4 vehicles and informing pedestrians. They are supported by Android platform. “Pelle intéractive Tracetel” is a prototype of an Interactive digital touch screen, placed on the vertical carrier. It gives out information on the location, most important contents in museums and other cultural institutions.

Multimedia device called “EchoLink panneaux intéractifs” displays and creates information in and about the immediate environment. It also enables communication among the users. “nAutreville” is an interactive transparent panel which centralizes the informative space (Figure 8c). By turning it on a vertical axis by 360$^{o}$ it enables gathering of information in the surrounding in the Argumented (City ID 2000) Reality form. It includes contents such as: real time window on the immediate environment which is in GPS system and it showcases information on the screen, transparency enables optimal intelligence in urban environment, it can be rotated, which gives the user the possibility to “scan” the city, flexibility and user-oriented information system portrays specificity of every area and adaptation to different profiles (tourists, inhabitants etc.). Other methods of informing pedestrians could be recognized in project “Concept Abribus”. This prototype of urban furniture is the example of future intelligent public transportation stops. In such places citizens can inform on different city contents by following digital information that can be selected. The idea is to secure some kind of “Paris within your reach”. Interface is designed as an open system in a form of a platform, which is attractive and accessible to all. It consists of an info-panel, a bench, a canopy and digital advertising panel. Among other functions and contents there is a touch-screen by which one can check out local information and adverts, ATM, remote-controlled defibrillator by GPRS, free Wi-Fi etc.

**Digital signage in Munich**

"Digital Signage" is a concept, presented in 2011, for a dynamic control system using the example of Munich. The concept is based on the intelligent use and combination of different technologies for orientation in the city. Control systems should help people to orient themselves in complex buildings or areas (Fig. 8.). The intelligent use of mobile devices, screens in public space or other technical possibilities can help to improve the orientation and by linking these technologies more individualized routing is ensured.

**Fig. 8**

*Digitale signaletik – combination of different technologies for orientation in Munich.*
*Source: Design made in Germany 2011*

In this concept, four elements form the basis (Design made in Germany 2011): 1) poles and 2) signpost as **static**, and 3) smart phones and 4) eyeglasses as **mobile** elements. The static elements serve as "landmarks": places where you can orientate yourself. The mobile elements are perfect for individual navigations.

**Bristol Pedestrian Sign System**

Considering the issues of the image of the city, its identity to residents and visitors, and the brand that it offers as critical factors of urban development, as a part of a Creative Bristol[7] campaign, local authorities developed an initiative called Bristol Legible City. It was a pioneering project of this kind in UK.
The program of Bristol's urban development was defined "in the late 1990s as the city was embarking on a series of major regeneration projects in the centre" (RUDI 2003). As one of "the major problems facing the city was communicating of the changes to the public, which was made worse by the confusing layout of Bristol city centre, in turn exacerbated by poor signage and low levels of useful public information" (RUDI 2003), the team of Bristol City Council's planning department[8] has developed the Legible City Concept. The aim of the concept was "to improve people's understanding and experience of the city through the implementation of identity, information and transport projects integrated with artist's work" and 'to make better sense of the city for both residents and visitors, helping to connect areas and amenities together and enhance the city's identity'. In practice, it consists of a series of information and sign making projects which aim to guide visitors and residents around the city.

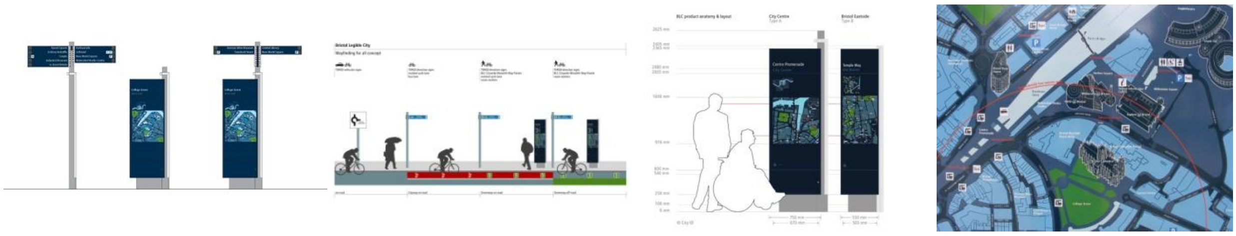


**Fig. 9**

*Bristol monolith panels and map.*
*Source: RUDI, 2003*

Bristol pedestrian sign system was the first major project of the Bristol Legible City initiative. Considered as the most comprehensive system in Europe, it has been designed specifically to encourage walking and to aid wayfinding in a complex city centre environment. The system encompasses9 directions signs and monolith shaped map panels to aid people's understanding of the city centre. The monolith panels (Fig. 9) help the user identify where they are and help them plan their journey through a combination of text and map information. Each panel is double sided, with the same type of information repeated on each face, but orientated towards the direction in which you are looking (City ID 2003). The panels are titled by the name of the street and which neighbourhood the user is in. The maps have innovative features forming a comprehensive mapping system. Two types of map are provided – a diagrammatic map of the city centre and a more detailed 'heads-up' map of the immediate area. The diagrammatic map shows the principal city centre neighbourhoods and the main attractions. In the map (Fig. 9) development, a detailed street map was used as a base that included pavements, crossings and steps, which are important information for pedestrians. A wheelchair user can easily see that the short cut on the map contains steps and should be avoided. In addition to the detailed local walking map[10], an abstract city map indicates where you are in the context of the whole city centre. Its is easily possible to navigate from one side of the centre to the other by heading in the general direction to the next panel and taking on new directions. The monolith panels are located at car parks, other points of arrival and at major junctions and spaces throughout the city centre.

[7] The city's bid to be Capital of Culture 2007
[8] Working in conjunction with a range of partners including the Chamber of Commerce and South West Regional Development Agency.
[9] In total there are nearly 40 monolith panels and 60 direction signs, with over 700 fingers. It was introduced in Spring 2001. Additional signs have been added in Summer 2002 and the system is being extended into Clifton during 2003.
[10] The maps on the panels are just the one segment of a mapping system in Bristol. All developed and stored digitally, the maps contain layers and a language that is being adapted for different uses – a walking map, arrival maps and maps for brochures and communications.

The focus and location of the system has been designed around the development of a primary pedestrian route network, termed the ‘Blue Route’ (City ID 2003), this links the main city centre neighbourhood areas together. Direction signs are interspersed with the monolith panels, at minor junctions, to provide continuity.
Having in mind that it was the first project in the scope of BLC initiative, pedestrian information system was created for a large number of people and groups, as well as needs to be adaptable for other projects (City ID 2000). The idea was generated on the base of one identity framework and visual language that could be used for all other segments. They include following design elements: a typeface, colours, icons and pictograms, a mapping system and product design. As a central element of BLC visual language is the ‘Bristol Transit’ typeface[11] that is used consistently on all information provided by the initiative, and creates a distinctive signature for Bristol. In addition to the typeface, a set of icons and colours for Bristol Legible City has been designed to be instantly understood by the widest audience. The dominant colour is blue, chosen for its clarity and its reflection of the city’s character and heritage – waterfront and maritime history, fountains, bridges blue glass and steel. Other colours are following standard conventions (red for Hospitals and light blue for parking).
Bristol pedestrian sign system, following the vision of creating a well-connected environment with unique and coherent identity, is defined in that way that it could be expanded on other types of street furniture and objects[12], i.e. redundant artefacts. Beside this, it contributes to the creation of an identity that is nationally and internationally recognised of all that is the best about Bristol. The detection of positive effects of the implementation of this initiative, has led to its development and application in other UK cities[13], which aim to improve the quality of their pedestrian spaces, as well as to enhance their identity.

## Conclusion

The results of the case studies of 5 European cities have shown that, while attempting to improve quality of pedestrian spaces and to define themselves as pedestrian cities, they have not develop an integrated pedestrian information system. It is still represented by the particular elements for information in order to orient users of open space that include: 1) street signs with distinctive visual identity and 2) tables with the names of important locations, which indicate the moving direction and destinations. These elements have distinctive visual identity that is characteristic for each analysed city, but it is not interconnected with other elements and objects in public space. It as also observed that the elements that are technologically advanced are detected only in Paris and Munich. Paris is characterised by elements of pedestrian information system created as a result of the project called Intelligent Urban Furniture. On the other side, the concept of integrated information system is developing in Munich. The implementation of both projects is in initial, experimental stage.
However, in order to contribute in the strengthening of identity and image of the city, the pedestrian information system need to be integrated and visually defined as is the case of Bristol Pedestrian Sign system. Based on a Legible City concept, guided by the mental structuring of a city’s image through form, it incorporate the ability of the user to recognize the parts and synthesize a coherent whole. On the one hand, this system makes the orientation easier and on the other, contributes to the strengthening the city's identity as the emphasis on its unique artifacts, as well as through the action of the system itself considered as the system of redundant artefacts. Given the criteria of legibility reflected on this way, the language of design becomes of high importance and every element of the system must be carefully designed and integrated to contribute to recognition of the city.

## Acknowledgements

The paper was realized, as a part of the research project “Modernization of Western Balkans” (No. 177009) and the project “Spatial, Environmental, Energy and Social Aspects of Developing Settlements and Climate Change – Mutual Impacts” (No TP36035), PP1: “Climate change as a factor of spatial development of settlements, natural areas and landscapes”, both financed by the Ministry of Education, Science and Technological Development of the Republic of Serbia.

[11] Bristol Transit typeface is adapted from the award winning Berlin information system from the award winning Berlin information system.
[12] Telephone kiosks, information booths, design of bus timetables, etc.
[13] London, Bath, Oxford, Somerset, Southampton, Birmingham, etc.

**References**

Balmer, J & Gray, E 2000, 'Corporate Identity and Corporate Communications: Creating a Competitive Advantage', *Industrial and Comercial Training*, pp. 256-261.

City ID 2000, *Bristol Legible City*. *From here to there*, Bristol.

City ID 2003, *Bristol Legible City*. *You are here*, Bristol.

Design made in Germany 2011, *Digitale signaletik*, viewed august 2014, http://www.designmadeingermany.de/2011/

Gehl, J 2010, *Cities for People*, Island Press, Washington.

Golledge, RG 1999, 'Human Wayfinding and Cognitive Maps', in *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes*, Johns Hopkins Uniersity Press, Baltimore.

Kavaratzis, M 2008, 'From City Marketing to City Branding: An Interdisciplinary Analysis with Reference to Amsterdam, Budapest and Athens ', Doctoral dissertation, University of Groningen, Groningen.

Linč, K 1974, *Slika jednog grada*, Građevinska knjiga, Beograd (eng. *The image of the city*)

Madanipour, A 1996, *Design of Urban Space: An Inquiry into a Socio-spatial Process*, John Wiley and Sons, Chichester.

Portugali, J 2011, *Complexity, Cognition and the City*, Springer, Berlin, London, New York.

RUDI 2003, *Bristol Legible City initiative*, viewed august 2014, http://www.rudi.net/node/15845

Rapoport, A 1990, *The Meaning of the Built Environment: A Nonverbal Communication Approach*, The University of Arizona Press, Tuscon.

Vukmirovic, M. 2014, *Pesacki prostor i kompetitivni identitet grada*, Zaduzbina Andrejevic, Beograd (eng. *Pedestrian space and competitive identity of the city*)

Vukmirovic M. 2013, *Znacaj i uloga mreze pesackih prostor u generisanju kompetitivnog ideniteta grada*, Univerzitet u Beogradu – Arhitektonski fakultet, Beograd, Doktorska disertacija (eng. *Role and importance of pedestrian space networks in generating the competitive identity of cities*)

# HIERARCHY AREAS IN THE COMMUNICATION NETWORK IN POLAND

**Katarzyna Kocur-Bera**[1], **Małgorzata Dudzińska, Cezary Kowalczyk, Jacek Kil**
*University of Warmia and Mazury in Olsztyn, Faculty of Geodesy and Space Management, Poland*

**Abstract:** Communication networks are used to transport people and goods. This article attempts to identify the structure of the road network in Poland and valuable areas in this structure (their hierarchy). To accomplish these objectives it uses the results of general traffic measurements. The knowledge about the location of the most important nodes or areas in the communication network plays a vital role as far as safety improvements and limitation of transport negative influence on the environment and living conditions are concerned. The study carried out on the road network in Poland indicated that there are eight important node regions, the so-called "valuable areas": three regions at the national level and five regions at the voivodeship level. The national regions include Górny Śląsk (Upper Silesia), the Warsaw Region and the Northern Region, while at the voivodeship level the group includes the sub-regions of Wrocław, Cracow, Bydgoszcz, Poznań and Lublin. According to the possessed knowledge, it is possible to preliminarily identify the space around the researched communication networks/regions where the environment ought to be analysed.

**Keywords:** hierarchy in communication networks, threat of space from transport, cluster analysis

## 1. INTRODUCTION

The present communication networks and those being planned to be built are essential for communication between units and social groups. The route is conditioned by environmental, economical and social factors (Pickup and Town, 1983), political and military factors. The sustainable development of transport is connected with economical, technological, social and cultural development considering environmental conditions (Tica et al., 2011). Moreover, it means that the location of transport networks influences space in social, natural and economic way. Transport networks affect space in a very complex manner. Besides transport functions, i.e. providing conditions for the traffic of people and flow of goods, energy and information, road networks also perform the following functions: service function – within the scope of production and consumption sector support; integration function – developing the necessary relationships and shaping the bonds within varied systems; and acceleration function – consisting in both speeding up the development and determining conditions for economic stimulation (Reilly et al., 2009; Budner, 2003; Kocur-Bera, 2010a). A poor quality of road links restricts the synergistic effect and does not make full use of the potential existing in the economy, education, science and culture. Lack of good territorial cohesion also contributes to the persistence of significant developmental differences between individual regions; moreover, a poor condition of road infrastructure is considered to be an important barrier to the development of entrepreneurship. A developed road network increases the possibilities for making use of competitive advantages, such as geographical location or market size. It affects the improvement of trade with other countries; ensures appropriate production and service support; improves the quality of passenger transport operations; allows gaining benefits resulting from the expansion of a single market and creates additional possibilities for attracting foreign capital as well as enhancing the mobility of workforce (*Programme For The National Road Construction For 2008-2012*). The effects may as well have negative influence, such as: greenhouse gas emission (Török, 2013), excluding land from farming and taking lands that are envirinmentally essential, and exposing ecosystem to destruction (Coffin, 2007), degradation (Carl et al., 2006), and soil, water and air pollution (Petrowic et al., 2008; Steiner and Boziciewic, 2008; Kecman et al., 2011; Dudzińska, Kocur-Bera, 2014), emitting vibrations and noises (Shukla et al., 2009), destroying cultural image (Fu et al., 2010; Jaarsma and Willems, 2002, Serrano et al., 2002), etc. According to Central Statistical Office of Poland, transportation sector is responsible for emitting over 28% of nitric oxide, 27% of carbon monoxide and 15% of dust in Poland (CSO, 2008). In accordance with the assumption of *National Transport Policy for 2006-2025 of Poland*, it is requied to reduce the negative influence of transport on natural environment and living conditions by: (1) accomplishing the integration of the transport policy with ecological policy; (2) increasing competitiveness of types of transport other than road and air transport; (3) using "the user pays" principle that leads to proper competition conditions that consider environmental and spatial aspects; (4) maximal consideration of ecological issues; (5) considering the sustainable transport development principle; (6) promoting solutions connected with spatial and functional integration of transport subsystems; and (7) promoting innovative solutions of transport system. The tasks ought to be accomplished during designing communication networks as well as considering the present roads. The aim of the research is to indicate the hierarchy of communication networks in Polish transport networks. Transport networks are characterized by features such as: hierarchy, accessibility, topology, geometrical features, etc. (Yerra and Levinson, 2005). Regional science explores hierarchies of places or hierarchies in the system of cities (Beckman 1958; Beckman and McPherson 1961), the hierarchy of roads have been analysed with the use of GPS data (Levinson and Zhu, 2011). In the article the networks have been analyzed considering the number of running cargo and passengers vehicles as with the traffic flow one can associate many negative influences on the space. The influences include: soil and water pollutions, noises or susceptibility to rare threats caused by transporting dangerous substances (Kocur-Bera, 2011b).

[1] Corresponding author: Katarzyna.Kocur@uwm.edu.pl

The knowledge about the location of the most important nodes or areas in the communication network plays a vital role as far as safety improvements and limitation of transport negative influence on the environment and living conditions are concerned.

## 2. MATERIALS AND METHODS

### 2.1. STUDY AREA

Poland is situated in central-eastern Europe. Since 2004, Poland has been a member of the European Union and, together with Finland, Estonia, Latvia and Lithuania; it forms its eastern border. Poland is one of the largest European countries: it covers an area of 312 679 km$^2$ and has a population of 38.1 million. In Poland, the structure of transportation is mainly based on car transport. In 2010, this type of transport constituted as much as 70.4% of all transportation operations and is steadily growing. The other forms represent a minor part of the structure of transportation: railway transport – 15.4%, air transport – 0.1%, pipeline transport – 7.6%, inland shipping – 0.3%, and marine shipping – 6.2% (CSO, 2010). Public roads in Poland are divided into national, regional, district and communal roads (*Act on Public Roads*, 1985; Kocur-Bera, 2011a).

The network of national roads includes highways and express roads, international roads and other communication pathways that ensure the integrity of national road system, access roads to public border crossings, roads that constitute ring-roadsnof large urban agglomerations and roads with military functions. The network of voivodeship roads consists of roads that connect cities and roads that are important for given voivodeship and military roads (not included in the national road system) (*Act on Public Roads of 21 March 1985)*.

In 2009, the total length of roads throughout Poland amounted to 384,830.0 km. In urban areas it amounted to 64,983.9 km, while in non-urban areas to 319,846.1 km.

According to their functions within the road network, the length of national roads amounted to 18 607.9 km, voivodeship roads to 28 461.1 km, district roads to 126,599.2 km and communal roads to 211,186.3 km (CSO, 2010). Poland has a limited network of highways: in 2010 their length was 854.7 km, which constituted app. 2 km per 1 000 km$^2$ and 2 km per 100 000 citizens while the mean values for 27 EU member-countries were 13 km and 15 km, respectively (CSO, 2010).

In 2009, the density of road networks throughout Poland amounted to 123.1 km per 100 km$^2$. Generally, roads in Poland are general access roads and limited access roads (motorways and expressways). Road communications availability in Poland amounts to the level of 75 % of the average value for the European Union (*Programme For National Road Construction For 2008-2012*).

The density of roads in the individual Polish voivodeships is variable. The highest density of roads is found in the voivodeship of Śląskie (20.70 km/100 km$^2$) and it is higher by 38% than the mean road density in Poland (15.05 km/100 km$^2$). The density of roads is also higher than the average value in Poland in such voivodeships as Dolnośląskie, Opolskie, Lubuskie, Małopolskie, Kujawsko-Pomorskie, and Świętokrzyskie. Except for Kujawsko-Pomorskie, these voivodeships are located in the south-west and western part of Poland. Industry is the main branch of the economy in this region. The voivodeships where the average density of roads approximates the mean value for Poland include: Wielkopolskie, Pomorskie, Zachodniopomorskie and Mazowieckie. The first three are situated in the north and north-west of Poland. The fourth, Mazowieckie, is located in the central region and includes Warsaw, the capital of Poland. In the other voivodeships (Łódzkie, Podkarpackie, Wamia and Mazury, Lubelskie, Podlaskie) the density of the road network is below the national mean value (see Fig. 1).

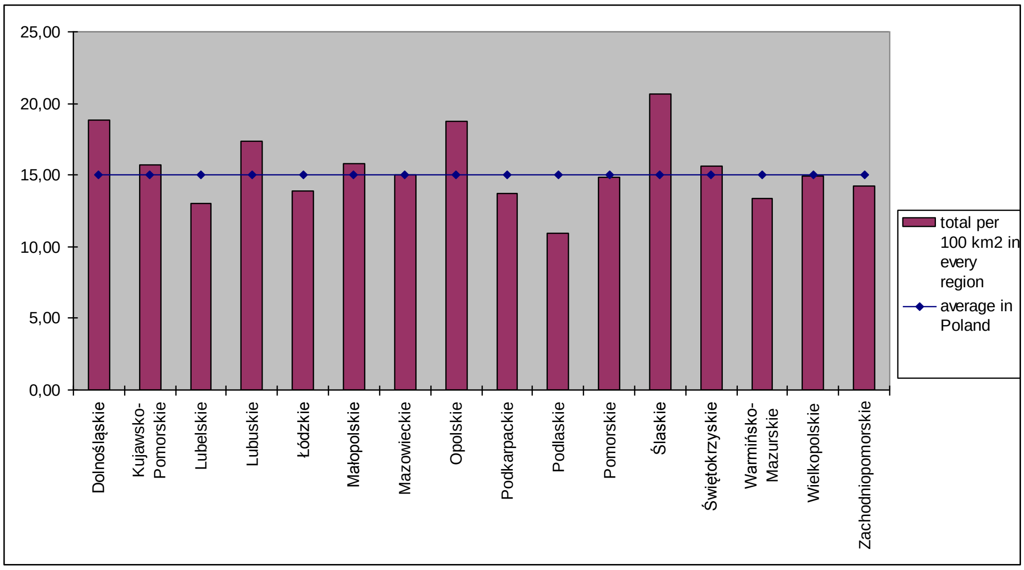



**Fig. 1.**
*Index of length for national/regional roads in relation to the average value in Poland per 100 km$^2$.*
*Source: own study*

### 2.2. DATA COLLECTION

For research purposes, results were acquired of direct measurements taken under the General Traffic Measurement (GTM) performed in 2010 for the General Directorate for National Roads and Motorways in Warsaw, Poland. Periodically, every 5 years a General Traffic Measurement takes place on national and regional roads. The latest measurement was performed in 2010.

The primary objective of the General Traffic Measurement is to obtain basic traffic parameters and characteristics for all sections of national and regional road networks, excluding roads located in towns and cities. While developing the method for performing the General Traffic Measurement, the following factors were taken into account: (1) the possibility of obtaining all the necessary data output, (2) the required accuracy of results, (3) the cost of measurement performance, (4) the degree of risk at a level which allows obtaining reliable results. The selected method for performing the General Traffic Measurement provides the possibility of obtaining all the necessary data output while maintaining the highest accuracy at an acceptable cost and permissible degree of risk (*Guidelines for the Organisation and Performance of the General Traffic Measurement on National Roads in 2010*). The measurement results are used by numerous institutions, *inter alia* in the management, maintenance and planning of the road network development, in analyses related to environmental protection and in the feasibility studies and design work in the field of road engineering.

In line with the recommendations of the United Nations Economic Commission for Europe (UNECE) in Geneva, vehicles are registered under seven categories: motorcycles and scooters; cars and other passenger vehicles with up to 9 seats; light commercial vehicles with a weight of up to 3.5 tonnes; vehicles with a weight exceeding 3.5 tonnes, without trailers; vehicles with a weight exceeding 3.5 tonnes, with trailers; buses; agricultural tractors; self-propelled machinery, bicycles and mopeds.

The basic criterion for dividing the road network into measuring sections is the quantitative homogeneity of traffic on a given section of the road. The limits of measuring sections are located on intersections and nodes where such traffic changes are possible. Moreover, the limits of measuring sections are also found at other points where the traffic does not always undergo significant changes but where it is necessary for other reasons, for example, a State border (*Guidelines for the Organisation and Performance of the General Traffic Measurement on National Roads in 2010*).

The basic parameter, calculated based on the General Traffic Measurement for all sections of the national and regional road network, is the average daily traffic in a given year. It is defined as a number of vehicles passing through a given section of a road within 24 consecutive hours, on average during one year.

The following types of measuring points are distinguished on sections of national roads: (1) TYPE A – units for continuous measurement of traffic; (B) TYPE B – sections on the non-urban roads with standard daily traffic fluctuations for a given road within provinces; (3) TYPE C – exit roads from large urban agglomerations; (4) TYPE D – arterial roads through smaller towns or villages; (5) TYPE E – other sections on which direct measurement is taken; (6) TYPE F – sections on which measurement is not taken and the traffic volume gets assigned to a comparable one, from another point; (7) TYPE G – sections of dual-carriageway roads and single carriage roads with four lanes; (8) TYPE H – sections of other roads on which direct measurement is taken; (9) TYPE I – as for the inter-regional roads [Guidelines for the Organisation and Performance of the General Traffic Measurement on National Roads in 2010]. On sections of regional roads, the following are distinguished: (1) TYPE P – basic sections on which direct measurements of traffic are taken on a full-time basis; (2) TYPE M – passages through towns and villages, where direct measurements of traffic are taken on a full-time basis; (3) TYPE W – other sections on which direct measurements of traffic are taken on a limited time basis; (4) TYPE T – sections of roads, on which direct measurement is not taken (*Guidelines for the Measurement of Traffic on Regional Roads in 2010*).

For all measuring points, the average daily traffic is calculated using the formula:

$$SDR = \frac{M_R * N_1 + 0{,}75 M_R * N_2 + M_N * N_3}{N} + R_N (vehicles/day) \tag{1}$$

where:
SDR – average daily traffic of motor vehicles, in total;
MR – average daily traffic on working days (from Monday to Friday, between 06:00 AM and 10:00 PM);
0.75 MR – average daily traffic on Saturdays and holidays (between 06:00 AM and 10:00 PM).
For bicycle traffic, it is possible to assume a coefficient other than 0.75, based on the results of additional and independently-taken measurements of bicycle traffic;
MN – average daily traffic on Sundays and holidays (between 06:00 AM and 10:00 PM);
RN – average night-time traffic (between 10:00 PM and 06:00 AM);
N1 – the number of working days in a year (in 2010 – 253);
N2 – the number of Saturdays and holiday evenings in a year (in 2010 – 52);
N3 – the number of Sundays and holidays in a year (in 2010 – 60);
N – the total number of days in a year (in 2010 – 365).
Calculation of the MR, MN and RN values is varied depending on the types of measuring points.
For Type F and Type H points:

$$M_R = \frac{1}{3}\left(\frac{X_2 + X_6}{2} + \frac{X_3 + X_8}{2} + \frac{X_1 + X_4}{2}\right) \tag{2}$$

$$M_N = \frac{1}{2}\left(\frac{X_5 + X_7}{2} + X_9\right) \tag{3}$$

$$R_N = \frac{1}{2}\left(X_{10} + X_{11}\right) \tag{4}$$

For Type G points:

$$M_R = \frac{1}{3}\left(X_2 + X_4 + X_8\right) \tag{5}$$

$$M_N = \frac{1}{2}\left(X_5 + X_9\right) \tag{6}$$

$$R_N = X_{11} \tag{7}$$

where:
X1, X2, .......... X11 the total volume of motor vehicles traffic in subsequent measurement days (*Guidelines for the rganisation and Performance of the General Traffic Measurement on National Roads in 2010*).

### 2.3. METHODS

In order to indicate the hierarchy areas in communication nodes/areas in communication networks, the cluster analysis (CA) and Ward's method. The method enables indication of the most valuable/the most important points of communication networks, indicating communication areas (a corridor of high average daily traffic) in the regional and sub-regional scales. The algorithm, through analogy, enables indicating areas exposed to negative influence of transport on the environment and living conditions. It does not regard means that compensate the influence of transport, for instance: wildlife crossings or acoustic screens. This direction of the research will be developer in further researches.

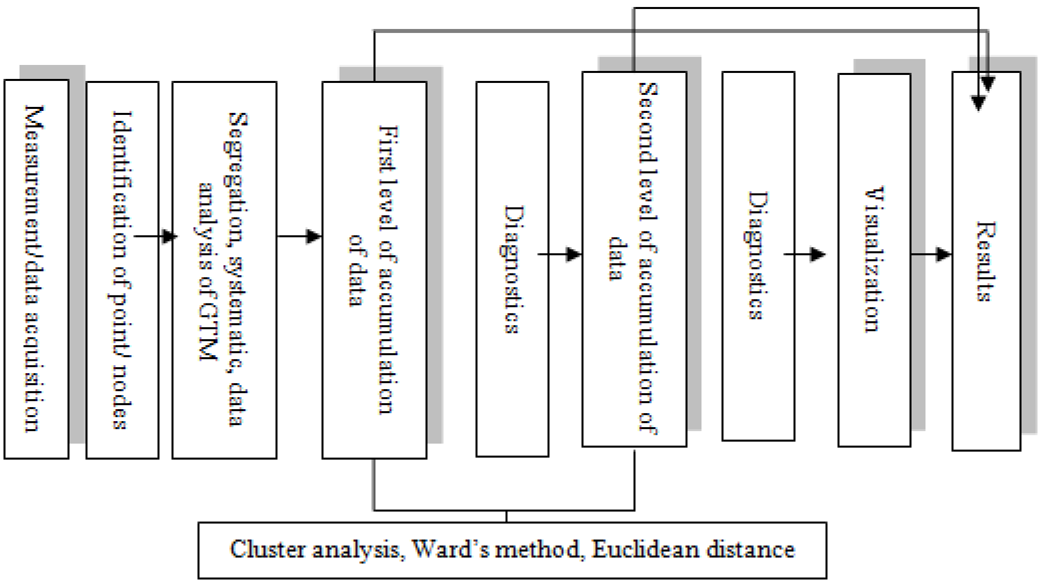



**Fig. 2.**
*Concept of cluster analysis (algorithm).*
*Source: own study*

### 3. RESULTS - GEOGRAPHICAL LOCATION OF DOMINANT NODES

39 % of all nodes are located on national roads and represent 63 % of the total traffic, while 61 % of nodes are located on regional roads and represent 37 % of the total traffic. Calculations were made using the licensed software STATISTICA, Version 10. Table 1 presents measurement data with the highest traffic flow. On this basis selected areas vat the regional and sub-regional levels (see Fig. 3). The regions correlate with the administrative divisions of Poland into voivodeships and counties. The most significant regions in terms of communications include the area of Katowice (the geographical region named the Upper Silesia), Warsaw (the Warsaw Region) and Gdynia (the geographical region called the Northern Region). These regions overlap with the geographical territories and industrial zones in Poland. Upper Silesia (Górny Śląsk) is a node region located in the voivodeship of Śląskie with the highest density of roads in Poland (20.7 km/100 km$^2$). They are concentrated around such cities as Katowice, Sosnowiec and Chorzów. The Warsaw Region is situated in the voivodeship of Mazowieckie. The density of roads approximates the national average (15.04 km/100 km) and they are concentrated around the capital city. The Northern Region is situated in the voivodeship of Pomorskie (Pomeranian). The density of roads is slightly below the Polish average (14.85 km/100km$^2$) and they are concentrated around Trójmiasto (Tricity), which includes Gdańsk, Sopot and Gdynia.

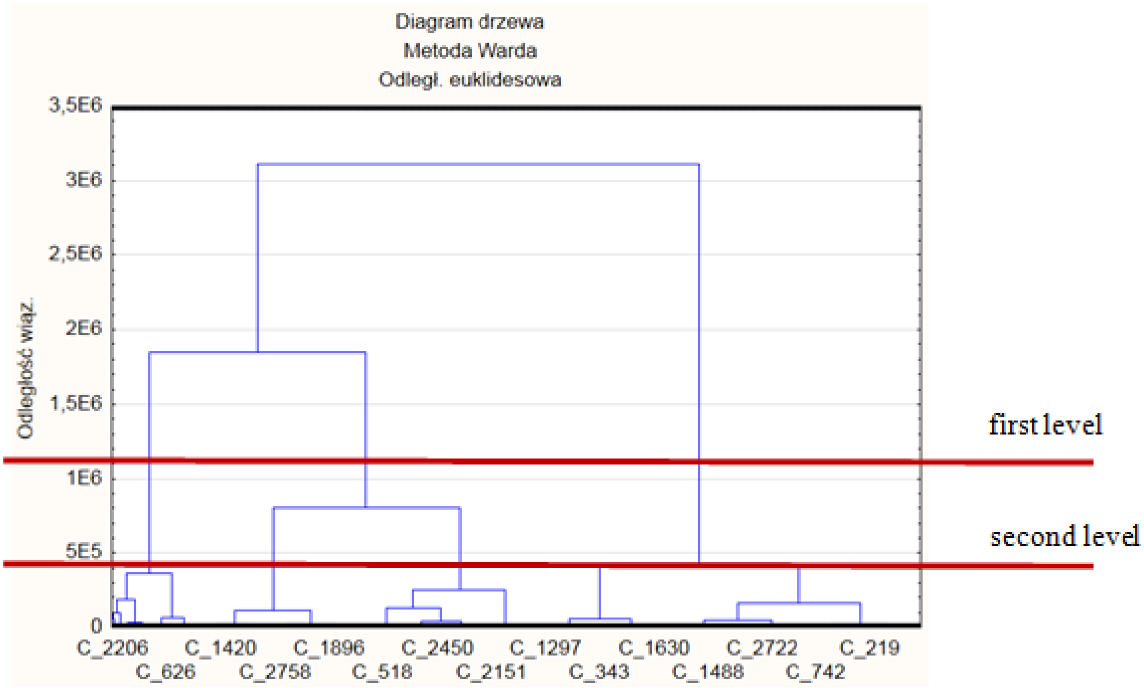



**Fig. 3.**
*Dendrogram for the 2791 general data traffic measurement and 8 variables (characterizing traffic flow).*
*Source: own study*

**Table 1**
*The first level of date accumulation together with description of points*

| Rank | Name of point | General traffic measurement (GTM) | Name of Region |
|---|---|---|---|
| 1 | Wrocław-Węzeł Bielany | 62 187 | Lower Silesia |
| 2 | Raszyn-Janki | 61 240 | The Warsaw Region |
| 3 | Warszawa-Marki | 56 852 | The Warsaw Region |
| 4 | Raszyn (Przejście) | 56 340 | The Warsaw Region |
| 5 | Łomianki-Warszawa | 55 602 | The Warsaw Region |
| 6 | Warszawa-Mysiadło | 54 812 | The Warsaw Region |
| 7 | Osowa-Węzeł Matarnia | 53 736 | The Northern Region |
| 8 | Gdynia Port-Wielki Kack | 51 578 | The Northern Region |
| 9 | Czeladź-Sosnowiec | 51 258 | Upper Silesia |
| 10 | Węzeł Sośnica-Chorzów | 54 615 | Upper Silesia |
| 11 | Chorzów-Katowice | 50 770 | Upper Silesia |
| 12 | Sosnowiec-Katowice | 104 339 | Upper Silesia |
| 13 | Katowice (Przejście) | 75 020 | Upper Silesia |

*Source: own study*

At the voivodeship level, hierarchical analysis has revealed a tendency towards the formation of local node sub-regions which are located in the following voivodeships (provinces): Dolnośląskie, Małopolskie, Kujawskopomorskie, Wielkopolskie and Lubelskie (see: Table 2). Traffic flow is concentrated around five cities: Wrocław, Kraków, Bydgoszcz, Poznań, and Lublin.

**Table 2**
*Second level of date accumulation*

| Rank | Name of voivodship | Name of point | Traffic flow | Name of sub-regions |
|---|---|---|---|---|
| 1 | Dolnośląskie | Wrocław Bielany – (Węzeł) | 62187 | Lower Silesia |
| | | Węzłe Kąty Wrocławskie-Węzeł Bielany | 39164 | |
| 2 | Opolskie | Nogowczyce-Granica Województwa | 31830 | No |
| | | Lewin Brzeski-Gradków | 28606 | |
| 3 | Lubuskie | Świebodzin (Obwodnica) | 21547 | No |
| | | Gorzów Wielkopolski (Obwodnica)-Deszczno | 15619 | |
| 4 | Małopolskie | Kraków(Piekary)-Kraków(Tyniec) | 41520 | Sub-Region Kraków |
| | | Balice I-Balice Ii (Lotnicko) | 34234 | |
| 5 | Kujawsko-Pomorskie | Lubicz-Lubicz | 26726 | Sub-Region Toruń-Bydgoszcz |
| | | Osielsko-Bydgoszcz | 22912 | |
| 6 | Świętokrzyskie | Ćmińsk-Kielce | 22167 | No |
| | | Suchedniów-Występa | 21001 | |
| 7 | Wielkopolskie | Dębina(Węzeł)-Krzesiny (Węzeł) | 39500 | Sub-Region Wielkopolski |
| | | Tarnowo Podgórne-Poznań | 35424 | |
| 8 | Zachodnio-pomorskie | Rurka-Droga Wojewódzka 142 | 22862 | No |
| | | Droga Wojewódzka 142-Szczecin | 22111 | |
| 9 | Łódzkie | Droga Nr 8-Rokszyce | 44159 | No |
| | | Rzgów-Tuszyn | 29364 | |
| 10 | Podkarpackie | Rzeszów-Kraczkowa | 29703 | No |
| | | Sanok (Przejście) | 24795 | |
| 11 | Warmińsko-Mazurskie | Węzeł Elbląg Wschód-Pasłęk | 18486 | No |
| | | Bartoszyce (Przejście) | 17119 | |
| 12 | Lubelskie | Kalinówka-Świdnik (Krępiec) | 35790 | Sub-Region Lublin |
| | | Grabów-Lublin | 22171 | |
| 13 | Podlaskie | Choroszcz-Białystok | 21163 | No |
| | | Augustów-Przejście 2 (Ogrodniki) | 18752 | |

*Source: own stady*

In Dolnośląskie, this zone is centred around Wrocław. The roads from the south-east towards Opole and from the west towards Legnica are the most loaded with traffic flow. Wrocław, together with the neighbouring cities, is located within the Sudetes Industrial Area.

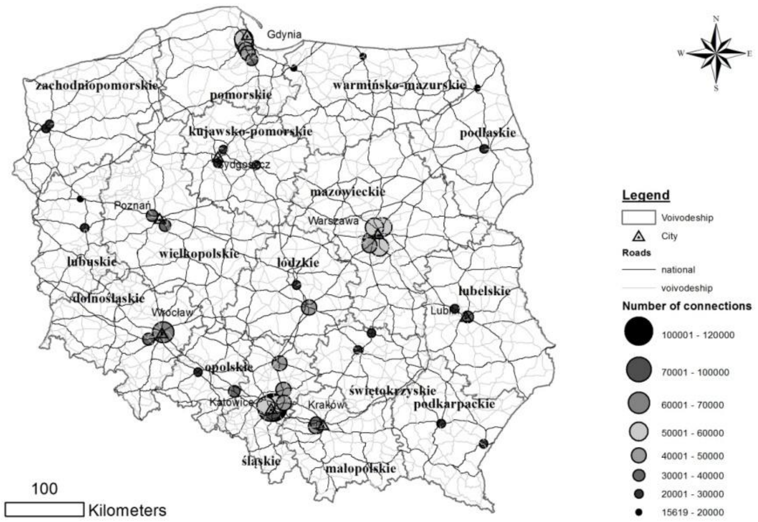



**Fig. 3.**
*The spatial distribution of the most important nodes in the road network in Poland.*
*Source: Kocur-Bera 2014.*

In Małopolskie, this zone centers around Cracow. The roads from the south towards Zakopane and from the west towards Katowice are the most loaded with traffic flow. The city, together with the neighboring towns, forms the Cracow Industrial Area.
In Kujawskopomorskie, there is a zone around Bydgoszcz, the roads from the north towards Grudziądz and from the east towards Toruń are the most loaded with traffic flow.
Within the voivodeship of Wielkopolskie, there is a zone around Poznań. The roads from the south towards Luboń and from the west towards Września and Swarzędz have the heaviest traffic flow.
In Lubelskie, this zone centres around Lublin. The roads from the north towards Puławy and from the south-east towards Zamość and Dorohusk (the border crossing with Ukraine) are the most loaded with traffic flow.
The spatial analysis did not reveal any node sub-regions in the following voivodeships (provinces): Opolskie, Lubuskie, Świętokrzyskie, Zachodniopomorskie, Łódzkie, Podkarpackie, Warmińsko-Mazurskie and Podlaskie. In these voivodeships (provinces), the most important nodes were separated from each other and they did not form a cohesive structure, i.e. a local sub-region.

## 4. Conclusions

The hierarchy of communication networks have been analysed in the article. The authors' algorithm DPA have been used. The studies of the road network in Poland were based on traffic flow measurements that were taken on the national and voivodeship (province) roads. Analysis allows us to identify hierarchy areas in communication network, based on the most important nodes found in a given region and on the neighborliness matrix. The study carried out on the road network in Poland indicated that there are eight important node regions: three regions at the national level and five regions at the voivodeship level with increased possibility of negative influence of transport on natural environment and living conditions. The national regions include Górny Śląsk (Upper Silesia), the Warsaw Region and the Northern Region, while at the voivodeship level the group includes the sub-regions of Wrocław, Cracow, Bydgoszcz, Poznań and Lublin. With information on their position, it is possible to provide special supervision for public security or impact on the natural environment through exhaust gas emissions, vibrations, noise and polluted runoff from roads, ect. (Kocur-Bera, 2010b). The algorithm facilitates the identification of initial needs of the natural environment analyses.

**References:**


*Act on Public Roads of 21 March, 1985* (Journal of Laws No. 14, Item 60, as amended), Poland. Available from internet: <http://isap.sejm.gov.pl/DetailsServlet?id=WDU19850140060)> .

Beckmann, M.J. 1958. City hierarchies and the distribution of city sizes. *Economic Development and Cultural Change 6: 243–248.*

Beckmann, M.J., McPherson, J. 1970. City size distributions in a central place hierarchy: An alternative approach. *Journal of Regional Science 10: 25–33.*

Budner, W., 2003. *Lokalizacja przedsiębiorstw.* (Location of companies). Wydawnictwo Akademii Ekonomicznej w Poznaniu, Poland.

Carl, S. C., Chen, L., Gao, W., Chen, W., Yan, M., 2006. Impact of planting grass on terrene roads to avoid soil erosion. *Landscape and Urban Planning 78: 205–216.*

Coffin, A. W., 2007. From road kill to road ecology: A review of the ecological effects of roads. *Journal of Transport Geography 15: 396-406.*

CSO, 2010. *Transport – activity results in 2009*. Statistical Information and Elaborations, Warsaw ISSN 1506-7998, Poland. Available from internet: < http://www.stat.gov.pl>.

CSO, 2008. *Ochrona Środowiska* (Enviornmental protection), Poland. Available from internet: <http://www.stat.gov.pl>.

Dudzińska, M., Kocur-Bera, K., 2014. Land consolidation as the driving force behind ecological and economic development of rural areas. ***9TH INTERNATIONAL CONFERENCE "ENVIRONMENTAL ENGINEERING",*** *Section: Technologies of Geodesy and Cadastre*; eISSN 2029-7092/eISBN 978-609-457-640-9, Poland. Available from internet: <http://dx.doi.org/10.3846/enviro.2014.204>

Fu, W., Liu, S., Degloria, S. D., Dong, S., Beazley, R., 2010. Characterizing the "fragmentation–barrier" effect of road networks on landscape connectivity: A case study in Xishuangbanna, Southwest China. Landscape *and Urban Planning 95: 122–129.*

Guidelines for the Measurement of Traffic on Regional Roads in 2010. The Ministry of Infrastructure, Department of Roads and Motorways, Warsaw, Poland; November 2009. Available from internet: <http://www.gddkia.gov.pl/pl/987/gpr-2010>.

*Guidelines for the Organisation and Performance of the General Traffic Measurement on National Roads in 2010*. [Appendix to Order No. 59 of the General Director for National Roads and Motorways of 12 October, 2009, Poland. Available from internet: <http://www.gddkia.gov.pl/pl/987/gpr-2010>.

Jaarsma, C. J. , Willems, G. P.A., 2002. Reducing habitat fragmentation by minor rural roads through traffic calming. *Landscape and Urban Planning 58: 125–135.*

Kecman I., Antić K., Babić M., Pavelkić V., 2011. Serbian road transport and pollutant emissions: an indicator based review and comparative analysis. *International Journal for Traffic and Transport Engineering, 2:70–82* UDC: 502.3:504.5:621.43.064;

Kocur-Bera, K., 2010a. Geoinformacja w zarządzaniu siecią transportową część 1. (Geo-information in transport network management – part 1). *Acta Scientiarum Polonorum, Administratio Locorum, Real Estate Management*, 9 *(3): 55-62,* Poland.

Kocur-Bera, K., 2010b. Specyfika wybranych oddziaływań sieci drogowej na otaczającą przestrzeń (The specific character of selected effects of the road network on the surroundings). *Acta Scientiarum Polonorum, Administratio Locorum, Real Estate Management, 9 (2): 89-100,* Poland.

Kocur-Bera, K., 2011a. Specyfika finansowania sieci drogowych (The specificity of financing road networks). *Acta Scientiarum Polonorum, Administratio Locorum, Real Estate Management, 10 (2): 59-68,* Poland.

Kocur-Bera K., 2011b. Wykorzystanie platformy GIS w zarządzaniu kryzysowym. (Use of the GIS platform in crisis management). *Acta Scientiarum Polonorum, seria: Administratio Locorum 10 (4): 27-41,* Poland.

Kocur-Bera, K., 2014. Scale-free network theory in studying the structure of the road network in Poland. *PROMET - Traffic & Transportation, Vol. 26 (3): 235-242*. Available from internet: <http://www.fpz.unizg.hr/traffic/index.php/PROMTT/article/download/1316/1195>.

Levinson, D., Zhu, S., 2011. The hierarchy of roads, the locality of traffic, and governance. *Transport Policy Volume 19 (1): 147–154.*

*National Transport Policy for 2006-2025 of Poland.* Ministry of Infrastructure, Warsaw, Poland. Available from internet: <http://pkbwl.mir.gov.pl/files/0/1795243/PolitykaTransportowaPastwanalata20062025Rzdowa.pdf>.

Pickup L., Town S., 1981. *The Role of Social Science Methodologies in Transport Planning*, TRRL Supplementary Report SR 698.

Petrovic, J., Ivkovic I., Vujacić I., Zezelj S., 2008. Environmental and economic effects of natural gas and biodiesel buses in public urban transport – the case of Belgrade . *The 7th International Conference Environmental Engineering.*

*Programme For The National Road Construction For 2008-2012*. [Appendix to Resolution No 163/2007 of the Council of Ministers of 25 September, 2007, Poland. Available from internet: <http://www.mi.gov.pl/2-48203f1e24e2f-3961.htm>.

Reilly, M.K., O'Marac, M.P., Seto, K.C., 2009. From Bangalore to the Bay Area: Comparing transportation and activity accessibility as drivers of urban growth. *Landscape and Urban Planning 92* (2009) 24–33.

Serrano, M., Sanz, L., Puig, J., Pons, J., 2002. Landscape fragmentation caused by the transport network in Navarra (Spain) Two-scale analysis and landscape integration assessment. *Landscape and Urban Planning 58:113–123.*

Steiner, S., Bozicevic, J., 2008. Subsidiarity in Strategic Transport Planning. *International Conference: Transportation and Land Use Interaction*, University Politehnica of Bucharest.

Shukla A.K., Jain S.S., Parida M., J. B., 2009. Performance of FHWA model for predicting traffic noise: A case study of Metropolitan city, Lucknow (India). *Transport,24(3): 234-240.* Available from internet: <http://dx.doi.org/10.3846/1648-4142.2009.24.234-240>.

Tica S., Filipović S., Živanović P., Bajčetić S., 2011. Development of trolleybus passenger transport subsystems in terms of sustainable development and quality of life in cities. *International Journal for Traffic and Transport Engineering, 1(4): 196 – 205*, UDC: 629.313:629.34.

Török A., 2013. Environmental comparism of road and railway transport: a case study in Hungary. *International Journal for Traffic and Transport Engineering, 4(2): 210 – 219.* Available from internet: <DOI: http://dx.doi.org/10.7708/ijtte.2014.4(2).07>

Yerra M. B., Levinson D.M., 2005. The emergence of hierarchy in transportation networks. *Ann Reg Sci* (2005) 39:541–553, Available from internet: <DOI: 10.1007/s00168-005-0230-4>.

# HISTORY OF THE PUBLIC TRANSPORT NETWORK AND VEHICLE STOCK OF BUDAPEST FROM 1968

**Gyula Gaal**[1]
[1] *Research fellow, KTI Institute for Transport Sciences Non-Profit Ltd.*

**Abstract:** As it is being prepared, soon the suburban bus lines of the Budapest Transport Limited[2] will probably be operated by the interurban public coach transport companies. Hence, BKV's history will end from the point of view of the company's current status. The subdivision of the company to bus and rail-guided departments was planned to occur in 2013. The public transport of Budapest wasn't examined in the aspect of the network and vehicle stock in its 45 year history. The aim of the article is to introduce the history of the BKV via these coherent, homogeneous series of data.

**Keywords:** BKV, Budapest, capital city, public transport, network

## 1. Introduction, factors that affect the examination of the topic

Since 1 January 1950 the territory of Budapest is 525 km$^{2}$. Although, there hadn't been a change in the size of the city area, the change in the number of inhabitants, the ratio of built-in areas and the density of the different parts of the city affects the requirements for public transport,.The other thing that represents stability is the fact that at least until the end of 2013 BKV had the same five subsections as in the beginning, since 1 January 1968. From this point of view the data can be compared on the same basis[3].

Unification of the database had to be done by the author. The more place an exact data can be found, the more chance there is to see different values.Just one example: in year 1968 the route length of trams and cog-rail is 417.7 km according to source (*Koroknai, Sudár, 1987.*), while 366.5 km according to source (*Yearbook of Budapest Transport Stock Company, 1998.*) while both sources are reliable. In such cases the one was chosen which was part of a longer data series used the same methodology. Such differences can be found in the number of vehicles as e.g. it is often not obvious whether the numbers mirror the passenger trams only or cargos as well. As much as possible, the data shown below contains only passenger vehicles.

The aim was to have the last day before the formation of BKV as base date of data (31. December 1967). In the case of many predecessor companies of BKV network data weren't available. These data are considered as same with the data of the BKV's first day, 1 January 1968. The number of vehicles should be also the same on both days. From the data of the 2000s the number of buses contains those buses either which are operated by subcontractors. Let's see the numbers (based on the statistical data of *Koroknai, Sudár, 1987.; BKV, 1998.; SKV, 1968.; KSH, 1990.; KSH, 1991.; KSH, 1995.; BKV, 1995-2012a.; BKV, 2012b.*[4]) after this preamble!

## 2. Length of track and length of network

The length of the track means the total length of the traffic-used parts in the case of tram, suburban rail (HÉV) and metro lines, not depending on how many platforms they have. It contains the tracks to the terminals and remises. It does not contain the tracks of bypasses, intermediate terminals, factory sidings etc.

In the case of rubber wheel vehicles the length of network means the total length of bus lines. The roads that are used by several different lines are counted only once.On *Fig. 1,* the change of the length of tracks and networks are shown.

[1] Corresponding author: gaal.gyula@kti.hu
[2]BKV was founded 1. January 1968. 1. January 1996 the owner, Municipality of Budapest modified it to a joint stock company. Since 6. February 2006 till todays it is operating as a public limited company.
[3]Please note that the metro belongs to BKV since year 1973 [2], while the underground (today it is known as Metro line 1) from the beginning.Due to the reason of easier elaboration, I don't deal with the ship lines, the chair-lift and the funicular.
[4] I thank the colleagues of BKK Centre for Budapest Transport PLC. for making available and permitting the use of year 2012's data (status of 28. February 2013) for me in order to complete the figures.

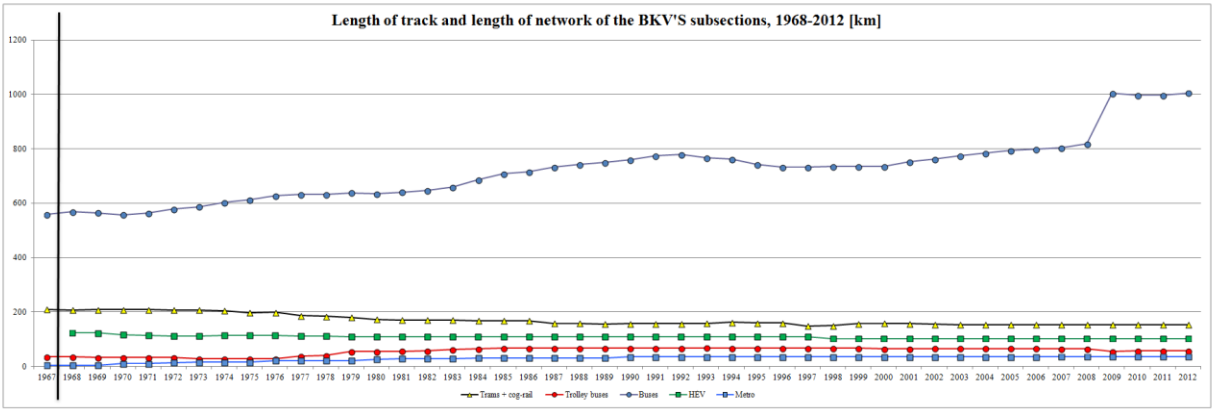



**Fig. 1**
*Length of track and length of network of the BKV's subsections, 1968-2012*
*(Source: own editing)*

The trams' length of track – as we expected it because we know that tram lines were cancelled– decreased significantly (but it is hard to recognize this in the shadow of the obvious dominancy of the bus lines). The length of network increased dramatically in the case of the bus sector. Since 2009 statistics contain the night network as well. After a small increase the network length of trolleybus lines stagnates. The building of metro lines can also be seen.

### 3. Length of routes and number of stops

Either from the length of routes' point of view (*Fig. 2*) the biggest change occurred at trams: after a strong decrease and fluctuation we hope to keep the positive change. In the case of buses there was a small decrease in the 1990s which was followed by a dynamic increase in the 2000s. Since 2009 the numbers contain the day auxiliary and night bus service network – this is the reason for the jump in this year. After the increase in the case of trolleybuses continuity can be seen.The statistical methodology of suburban rail (HÉV) is significantly different therefore it was left out from the figure.

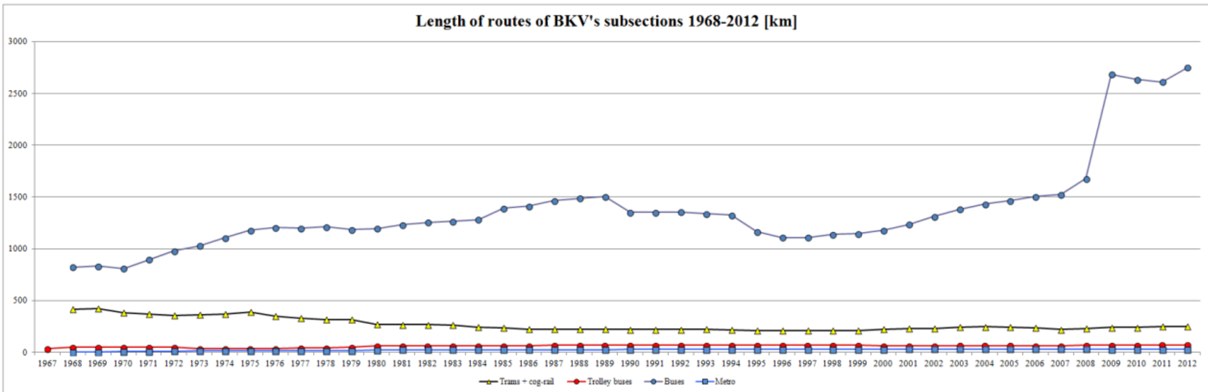



**Fig 2**
*Length of routes of BKV's subsections 1968-2012*
*(Source: own editing)*

The number of tram stops is decreasing continuously, while the bus stops' is essentially increasing (*Fig. 3*). According to the tendencies, trolleybus and metro subsections have similar trend lines from the length of track/network, length of routes and number of stops points of view.Although metro line 4 is under construction, new stops are not yet used, so the reason for the increasing of metro stops in 2011 is just the modification of the counting methodology.This is a good example to illustrate that a constant status can be shown in many different ways which unfortunately worsens the comparability of the data.

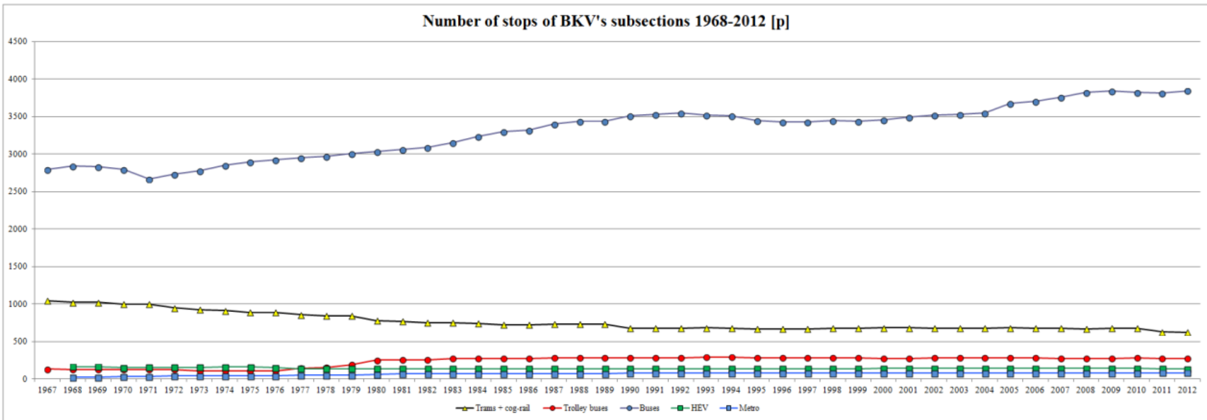



**Fig. 3**
*Number of stops of BKV's subsections 1968-2012*
*(Source: own editing)*

The average distance of stops were analyzed by mathematical methods (1):

$$\frac{l_v \cdot 1000}{N_m} = h_m \qquad 1)$$

where:

- $l_v$: length of routes [km]
- $N_m$: number of stops [p]
- $h_m$: average distance of stops [m]

This shows us an extremely hectic picture (*Fig. 4*). Bringing the night bus services in the statistics can be easily recognized again. The reason for the jump is on one hand that the length of routes is way longer than it was with the older methodology, and on the other hand one stop is one stop, not depending on how many different lines use them. The number of stops which are used only by night buses is very small, so the number of bus stops doesn't increase, consequently the average is getting worse.

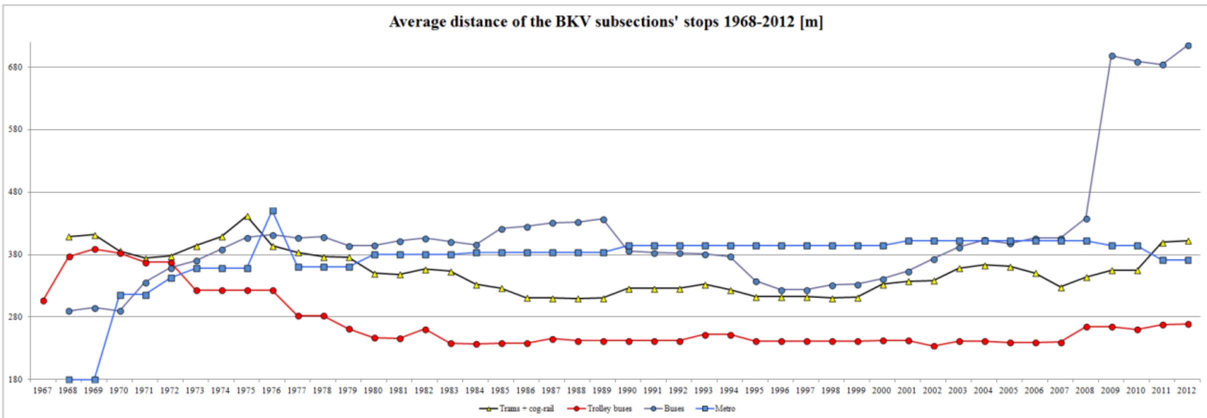



**Fig. 4**
*Average distance of the BKV subsections' stops 1968-2012*
*(Source: own editing)*

## 4. Acquisition and total number of vehicles

If we are talking about a 45-year-long period, it is hard to decide to count with the date of purchasing or with the date of first use of the new vehicles if the aim is to have a unified database (*Gerlei et al., 2008.*) but author aimed the latter. The numbers of new[5]and total vehicles can be seen on the following figures[6]. The number of bought buses decreased dramatically (*Fig. 5*) In the 2000s this was amended with buying used buses and with subcontracted operating.

[5]The new vehicles contain the bought used ones as well, e.g. Van Hool buses, TW6000 trams.
[6]At year 2012 the status on 28. February 2013 is shown.

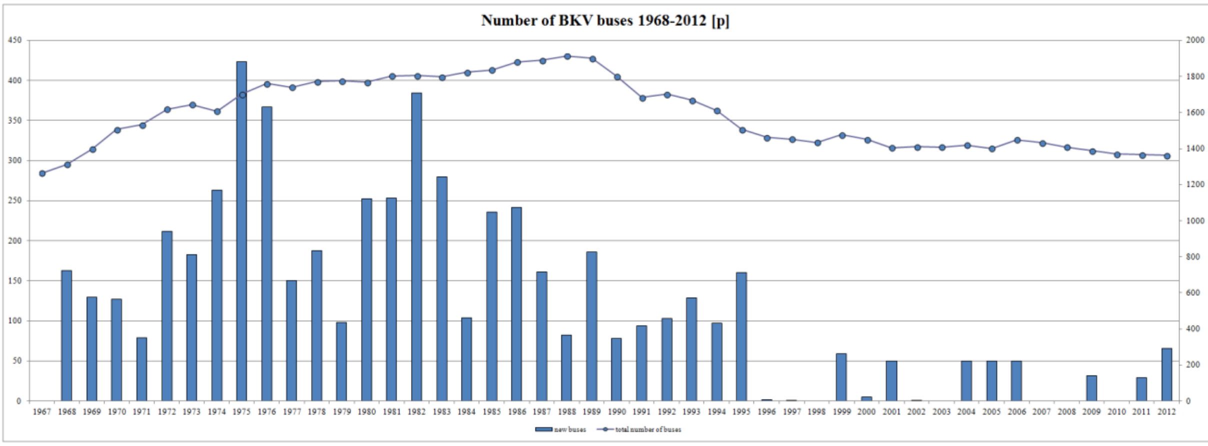



**Fig. 5**
*Number of BKV buses 1968-2012*
*(Source: own editing)*

As for decades Hungary was the biggest bus manufacturer country in Europe. Therefore it is interesting to compare the produced number of Ikarus buses with the number of the new vehicles bought by BKV. On *Fig. 6* the produced numbers of those Ikarusbus types are shown,which were used by BKVthe most often[7] (*Gerlei et al., 2008.*). Anyway, the two graphs fitting quite well. As it can be seen from *Fig. 6.*less and less amount of buses were sold although the Hungarian bus production is still in operation.

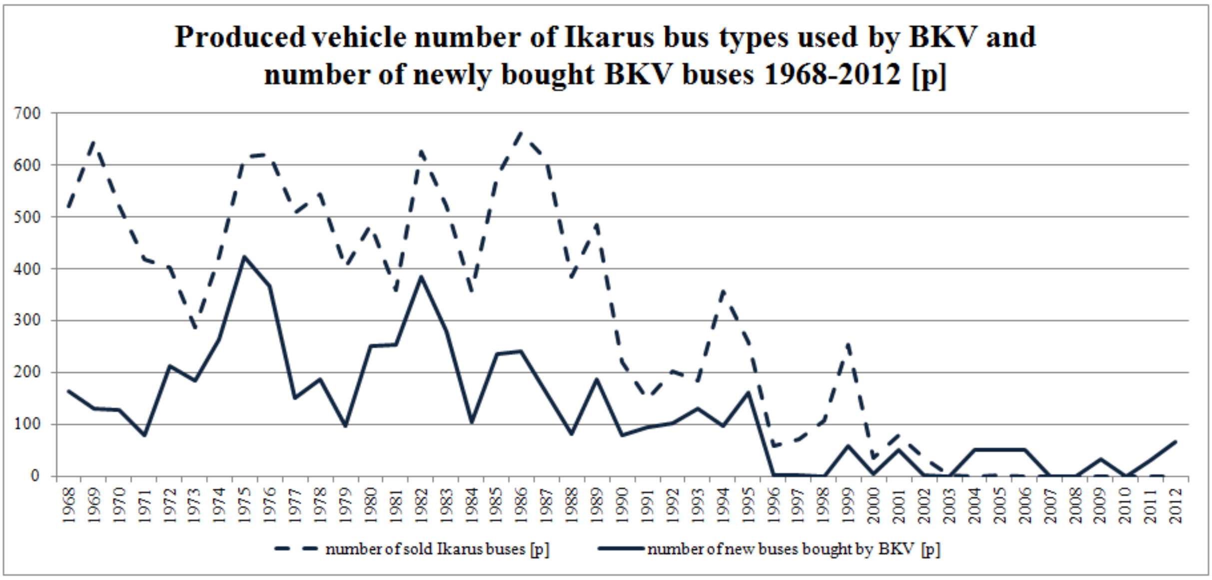



**Fig. 6**
*Produced vehicle number of Ikarus bus types used by BKV and number of newly bought BKV buses 1968-2012*
*(Source: own editing)*

The most of trolleybuses arrived at the end of the 1970s. After the short increase at the end of the 1980s there almost weren't any new purchase while the total number is more or less constant since the 1990s (*Fig. 7*).

[7] These models are: Ikarus 311, 620, 630, 180, 556, 260, 280, 263, 405, 415, 435, 412.

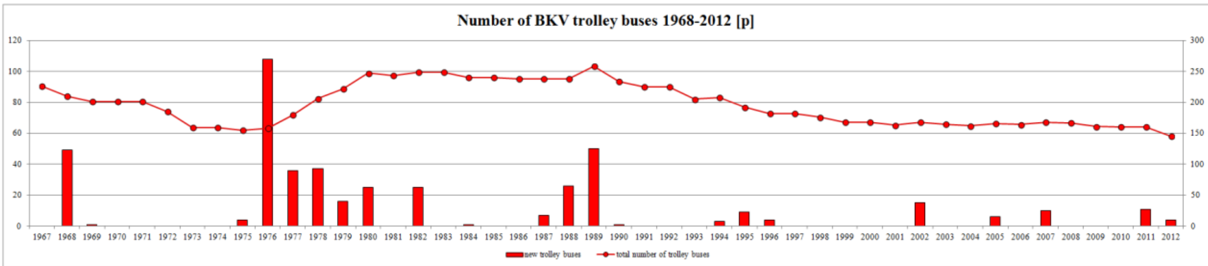



**Fig. 7**
*Number of BKV trolley buses 1968-2012*
*(Source: own editing)*

In the 1970s a few trams arrived regularly, which was followed by 1980 and 1984 by a high number of new trams. Since then most years passed without new acquisitions (*Fig. 8*). The total number got better in 2001 because of the used TW6000 trams bought from Hannover but this is the only year in the decreasing tendency since 1984. However, I have to add that the new trams' capacity is increasing constantly (e.g. one Combino can transport twice as many people as an old industrial articulated while the number of trams halves).

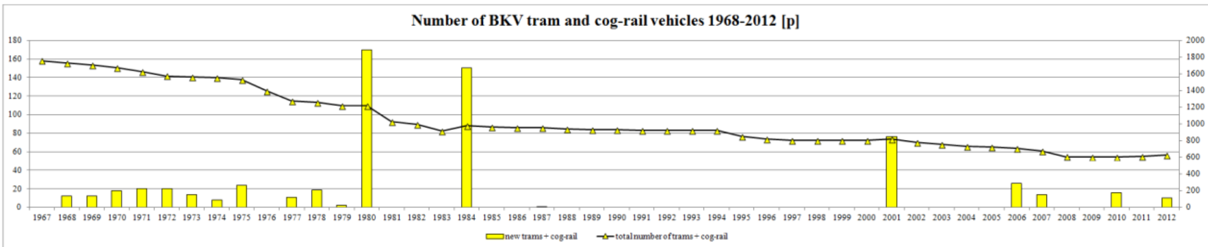



**Fig. 8**
*Number of BKV tram and cog-rail vehicles 1968-2012*
*(Source: own editing)*

Metro vehicles haven't arrived until 2012 since the 1990s (except year 2000) (*Fig. 9*). Since 2012 the new Alstom Metropolis trains are released and a positive change is brought by the inauguration of metro line 4 at spring 2014 which means 60 new metro coaches (*BKV PLC. Announcements*).

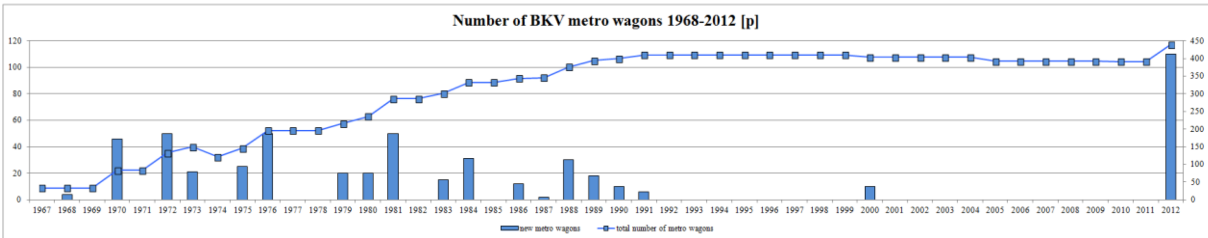



**Fig. 9**
*Number of BKV metro wagons 1968-2012*
*(Source: own editing)*

Between 1971 and 1983 there were nine years, during which Budapest got new HÉV vehicles. From 1984 till today none (*Fig. 10*). Since 1997 not even one was excluded from traffic.

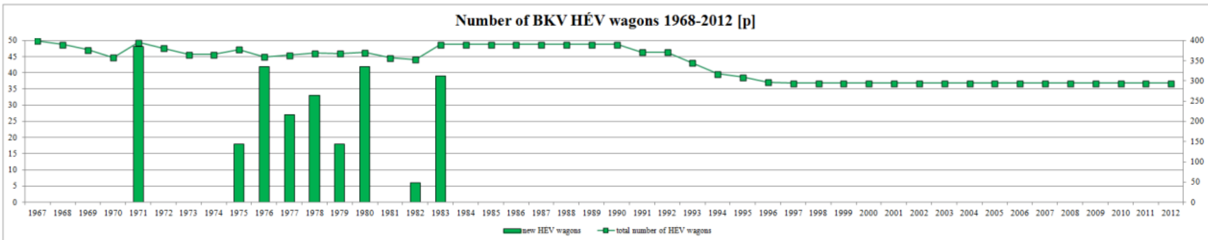



**Fig. 10**
*Number of BKV HÉV wagons 1968-2012*
*(Source: own editing)*

**5. Indicators**

*Fig. 11* shows how the proportion of the subsections' number of vehicles changed from 1969 until 2012 compared to 1968. Increase (significant) occurred only in the case of metro due to the fact that in 1968 only a short part of the underground was in operation. The meanwhile-built metro lines are many times longer than this. In the case of buses it can be seen that while the length of routes increased significantly, these growing routes are served by a decreasing number of vehicles (see below). 1989 must be mentioned as this was the year when the decrease of the number of buses began.

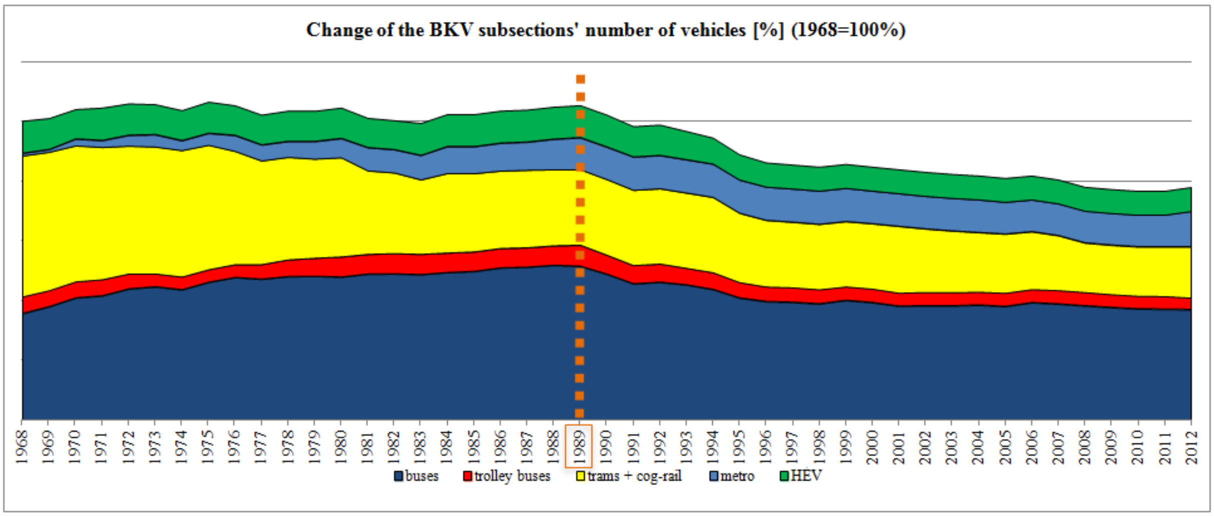



**Fig. 11**
*Change of the BKV subsections' number of vehicles 1968-2012*
*(Source: own editing)*

Does the situation differ if the whole number of vehicles and the complete length of the networks' routes are together analysed? The different subsections' number of vehicles were multiplied with the given years' length of routes (2) (according to the lacks of HÉV data, it is missing from the figure) so the BKV subsections' transport performance is presented in vehicle kilometers as:

$$N_{BKV} \cdot l_v = P_{BKV} \quad 2)$$

where:

$N_{BKV}$: total number of vehicles without HÉV[p]
$l_v$: total length of routes without HÉV [km]
$P_{BKV}$: transport performance of BKV without HÉV [vkm]

The resulted performance indicator shows how the transport performance of BKV changed: it can be stated that the longer and longer length of routes is served with a decreasing number of vehicles (*Fig. 12*).

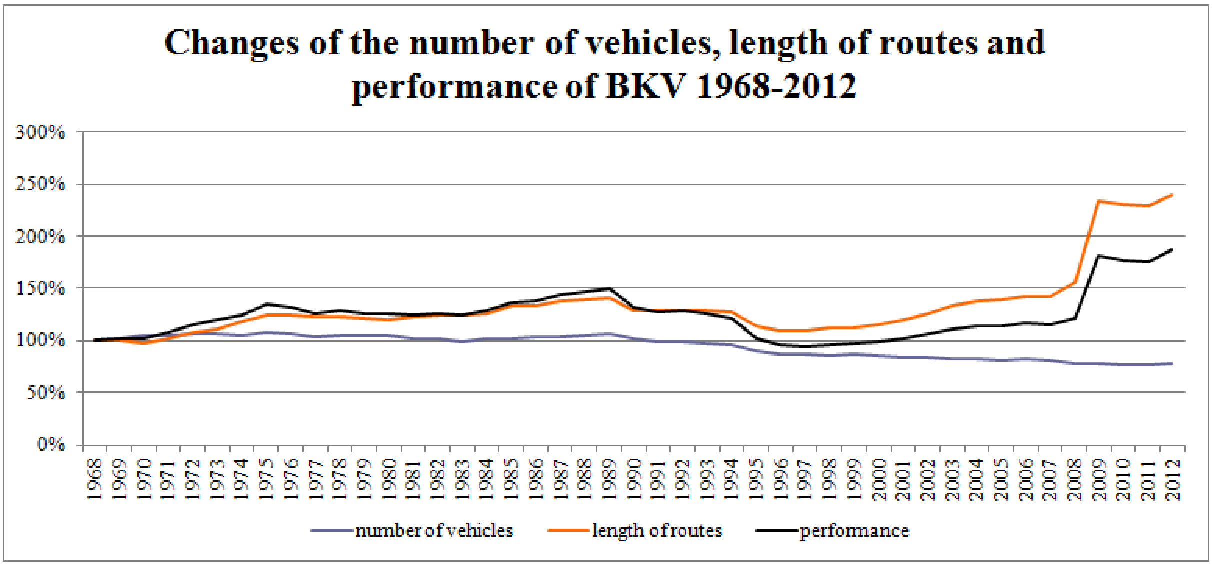



**Fig. 12**
*Changes of the number of vehicles, length of routes and performance 1968-2012*
*(Source: own editing)*

### 6. Conclusion

There are many, even more serious topics which are to be analysed in connection with Budapest's public transport. People have to face several problems from the everyday technical and age problems of the vehicles through the anomalies in connection with the constructing of metro line 4 till the organizational controversies. These are very sensitive questions and their analysis wasn't an aim of this article.

Now hypotheses are verified. It is obvious that on one handbus lines dominate the system while the rate of rail-guided departments decreased dramatically. That shows us the lack of using environment friendly technologies. The developing network is served by a decreasing number of vehicles which are even getting older and older, hopefully the tendency of the past years' renewing of buses will be kept. Also a positive fact is that in 2013 CNG buses were bought, however the infrastructure must be constructed.

## References

*Annual reports of BKV PLC.* (1995-2012). (In Hungarian: BKV Zrt. Évesjelentések (1995-2012)) http://www.bkv.hu/hu/kiadvanyaink/ Downloaded: 29. December 2013. 14:19

*BKV PLC. Announcements.* (In Hungarian: BKV Zrt. Közlemények.) http://www.bkv.hu/hu/kozlemenyek Downloaded: 16. June2014. 15:46

Gerlei T., Kukla L., Lovász Gy. 2008. *The secular history of Ikarus.* (In Hungarian: Az Ikarus évszázados története.) Maróti, Budapest.

Hungarian Central Statistical Office Budapest City Directorate: *Statistical Yearbook of Budapest.* (In Hungarian: Budapest statisztikai évkönyve) Statisztikai Kiadó Vállalat, 1968.

Hungarian Central Statistical Office: *Transport, Post and Telecommunication Yearbook 1989.* (In Hungarian:Közlekedési, posta és távközlésiév könyv 1989.) Központi Statisztikai Hivatal, Budapest 1990.

*One and a half century of the capital city's public transportation. Volume III.1919-1985.* (In Hungarian: A főváros tömegközlekedésének másfél évszázada III. kötet 1919-től 1985-ig.) (Ed. Ákos Koroknai, Kornélia Sudár). BKV, Budapest, n.a.

*Parameters book of the public transport of Budapest. Valid from 1. January 2012.* (In Hungarian: A budapesti közösségi közlekedés paraméterkönyve. Érvényes 2012.január 1-jétől.) Budapesti Közlekedési Központ Zártkörűen Működő Részvénytársaság, http://www.bkv.hu/upload/M_166/rek4/2394.pdf Downloaded: 16. June2014. 15:48

*Statistical Yearbook of Budapest 1990.* (In Hungarian: Budapest statisztikai évkönyve 1990.) Központi Statisztikai Hivatal Fővárosi Igazgatósága, 1991.

*Statistical Yearbook of Budapest 1994.* (In Hungarian: Budapest statisztikai évkönyve 1994.) Központi Statisztikai Hivatal Budapesti és Pest megyei Igazgatósága, 1995.

*Yearbook of Budapest Transport Stock Company 1968-1998.* (In Hungarian: A Budapesti Közlekedési Részvénytársaság Évkönyve 1968-1998.) BKV Rt., n.a.

# THEORETICAL CONSEQUENCES OF TRAVEL TIME BUDGET

**Adam TOROK**[1]
[1] *Budapest University of Technology and Economics*

**Abstract:** This paper conclude the findings of constant travel time budget. The travel time expenditures are constant only at the most aggregate statistical level. Nevertheless, individuals' travel time expenditures do show patterns that can be partly explained by measurable characteristics such as individual and household characteristics, attributes of activities at the destination, and characteristics of residential areas. If we consider that travel time expenditures are constant at the aggregate level over time and space the transport development as service development or as vehicle improvement would lead us to some question. This paper is focusing on these questions. Consequently, further research on travel time budget is justified.

**Keywords**: travel time budget, development, improvement

## 1. What is travel time budget?

Over the last forty years of travel demand analysis, time has been a variable of central importance to our understanding of the demand for travel (*Pas, 1998*). A frequently-studied time related measure is the amount of time allocated to travel. The concept of a "travel time budget" (TTB) refers to the idea that individuals' average daily travel time tends to be relatively constant independently from time and space. The behavioral hypothesis is that people have a certain (generally non-zero) amount of time that they are willing (or may even want) to spend on travel. Proponents of a travel time budget generally go beyond the suggestion of an individual-specific budget, however, to the observation that the actual size of that budget, as an average taken at a regional or national scale, is relatively stable across time and space. At the extreme, the TTB is viewed almost as a universal constant: 1.1 – 1.3 hours per traveler per day (*Zahavi and Ryan, 1980; Zahavi and Talvitie, 1980*), about 430 hours per person per year (*Hupkes, 1982*), 50 minutes to 1.1 hours per person per day (*Bieber et al., 1994*), 1.1 hours per person per day (*Schafer and Victor, 2000*), or 1.3 hours per person per day *(Vilhelmson, 1999*).

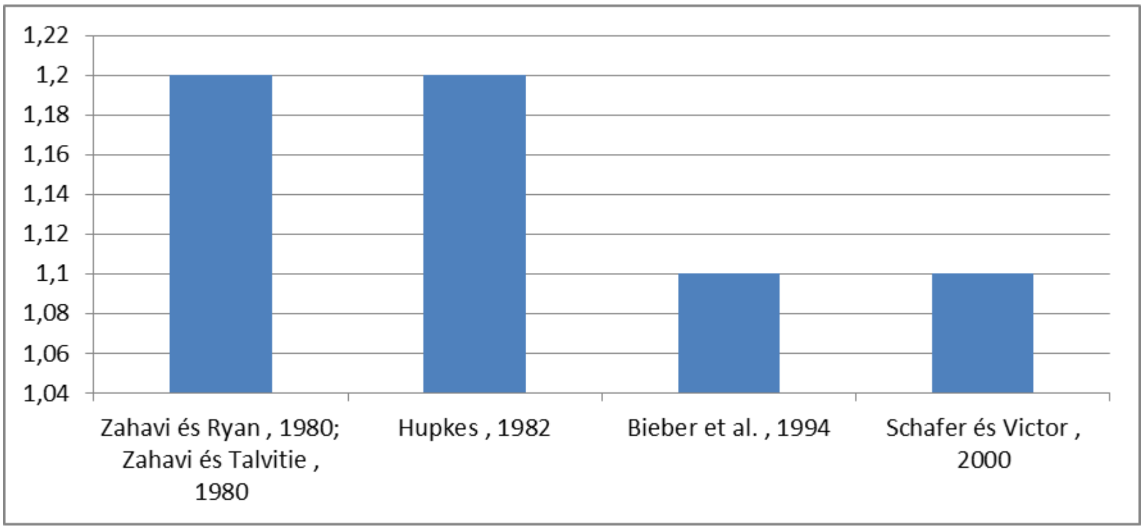



**Fig. 1.**
*Comparative analysis of Travel Time Budget*

First, Zahavi proposed that, on average, humans spend a fixed amount of their daily time budget traveling—the travel time budget (TTB). Zahavi posited a constant TTB for travelers—people, who make at least one motorized trip per day. We generalize the concept to a per-capita basis and find a similar constancy. Time-use and travel surveys from numerous cities and countries throughout the world strongly suggest that TTB is approximately 1.1 hours per person per day (*Fig. 1*).

[1] Corresponding author: atorok@kgazd.bme.hu

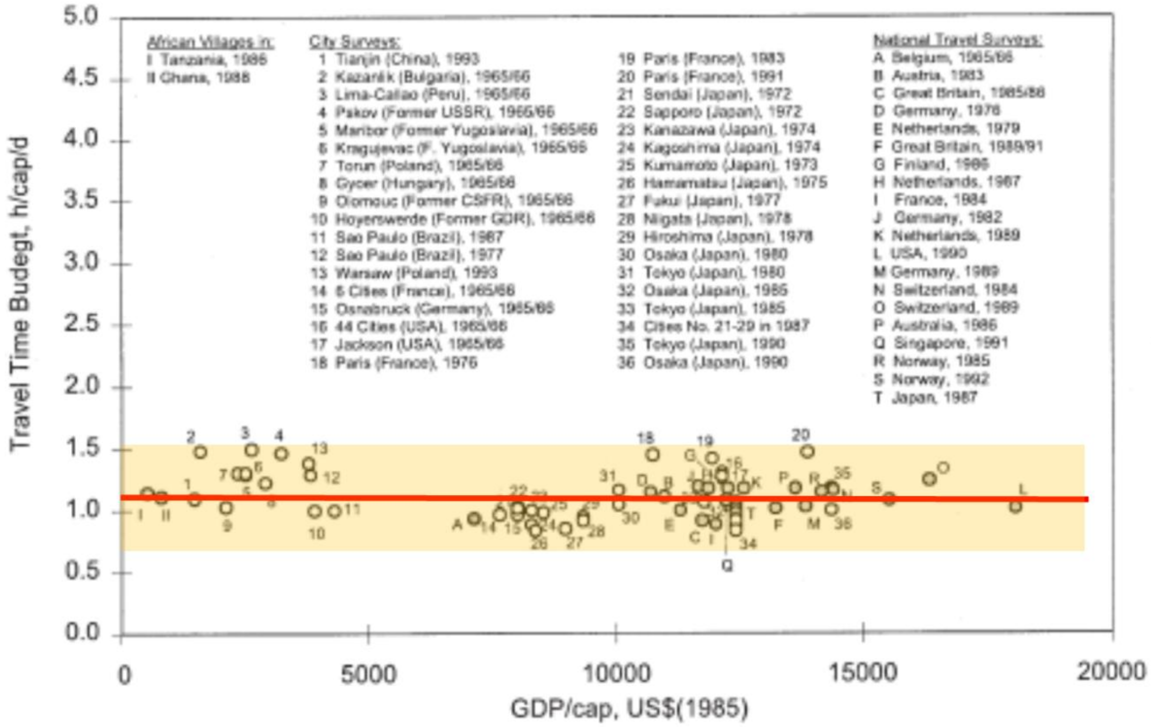



**Fig. 2.**
*Constant Travel Time Budget Worldwide*
*source: Schafer A., Victor D. G.; 1999*

While the TTB is constant on average, many variations are evident when examining the behavior of small populations and individuals. According to the Hungarian Statistic Office in 2008:

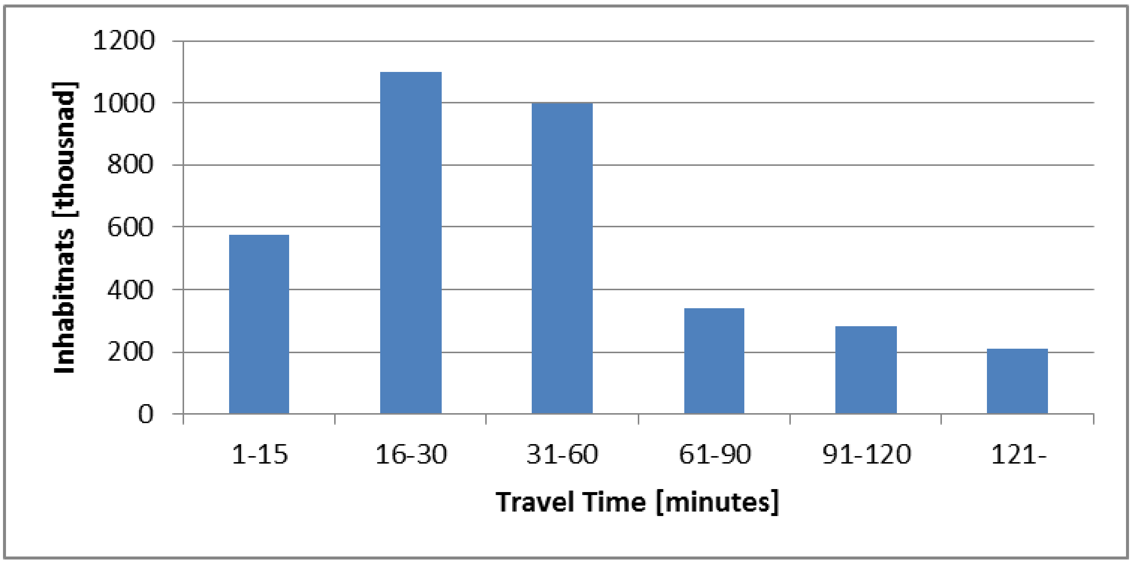



**Fig. 3.**
*Constant Travel Time Budget in Hungary*
*source: Hungarian Statistical Office*

## 2. Continuous Improvement

On the other hand, the TTB idea appears, at least at first glance, to clash with one of the most fundamental tenets of conventional travel behavior theory: that travel time is a disutility to be minimized. The travel time minimization principle underlies a great deal of policy-making as well as virtually all regional travel demand forecasting models, and is used to justify monetizing the benefits of transportation improvements on the basis (primarily) of travel time savings. But obviously, under a TTB, travel time is not minimized but is kept constant. If that is true, then, for example, the typical travel demand model is asking the wrong question. Rather than assuming the individual to be asking, “What is the least amount of travel I can do in order to accomplish a given set of activities?”, the individual instead should be viewed as asking, “What is the most attractive set of activities/ destinations I can achieve, given a certain travel time budget?” (*Mokhtarian P.L., Chen C.; 2003*). Therefore people shift to faster travel modes as their total mobility rises. If people using the same travel time over time and the technological advance gives us the opportunity to use faster and faster vehicles in order to fulfill our mobility demands that would lead to travel further and further.

## 3. Merging the ideas

The position of the TTB concept in the transportation planning and modeling profession is paradoxical. One reason is the common observation that at the aggregate level, when travel speeds increase over time – whether due to improvements in technology or additions of capacity to the system – travel distances tend to increase so as to keep travel times approximately constant (*Zahavi and Ryan, 1980; Hupkes, 1982; Marchetti, 1994; Barnes and Davis, 2001*). This links the TTB concept to the induced demand debate (e.g., *Noland and Lem, 2002*), with one extreme arguing that, at least from energy and air quality standpoints, it is useless at best and counterproductive at worst to add network capacity (or, presumably, to implement any operational efficiencies that increase overall speeds, as *Taylor, 2002* notes), since people will simply take advantage of the improvement to travel more.

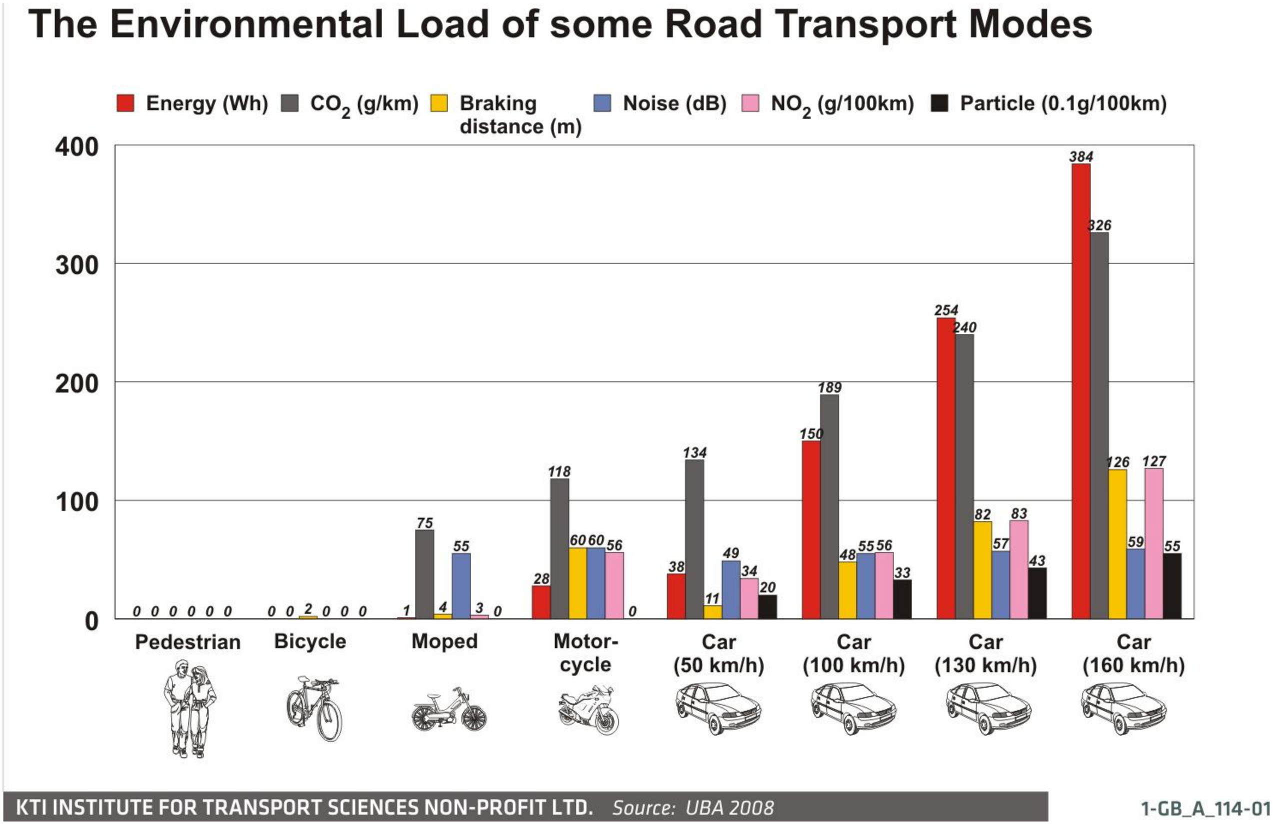



**Fig. 4.**
*Environmental load of road transport modes*
*source: KTI – Institute for Transport Sciences*

If these are true than in the industrialized regions of the world, the traffic volume from automobiles declines as they are replaced by faster modes for instance aircraft and high-speed trains, and that emissions of carbon dioxide from automobiles stabilize after 2020 without any policy intervention. But in developing countries, where mobility is lower and slower, automobility continues to rise; failure to begin regulation of these emissions in the near future could bequeath an impossible task of stabilizing carbon emissions in these countries later in the century, unless zero-carbon transportation fuels are rapidly and widely introduced into the market. Globally, although policy attention today is focused on automobiles, in the longer term emissions from aircraft will be more important.

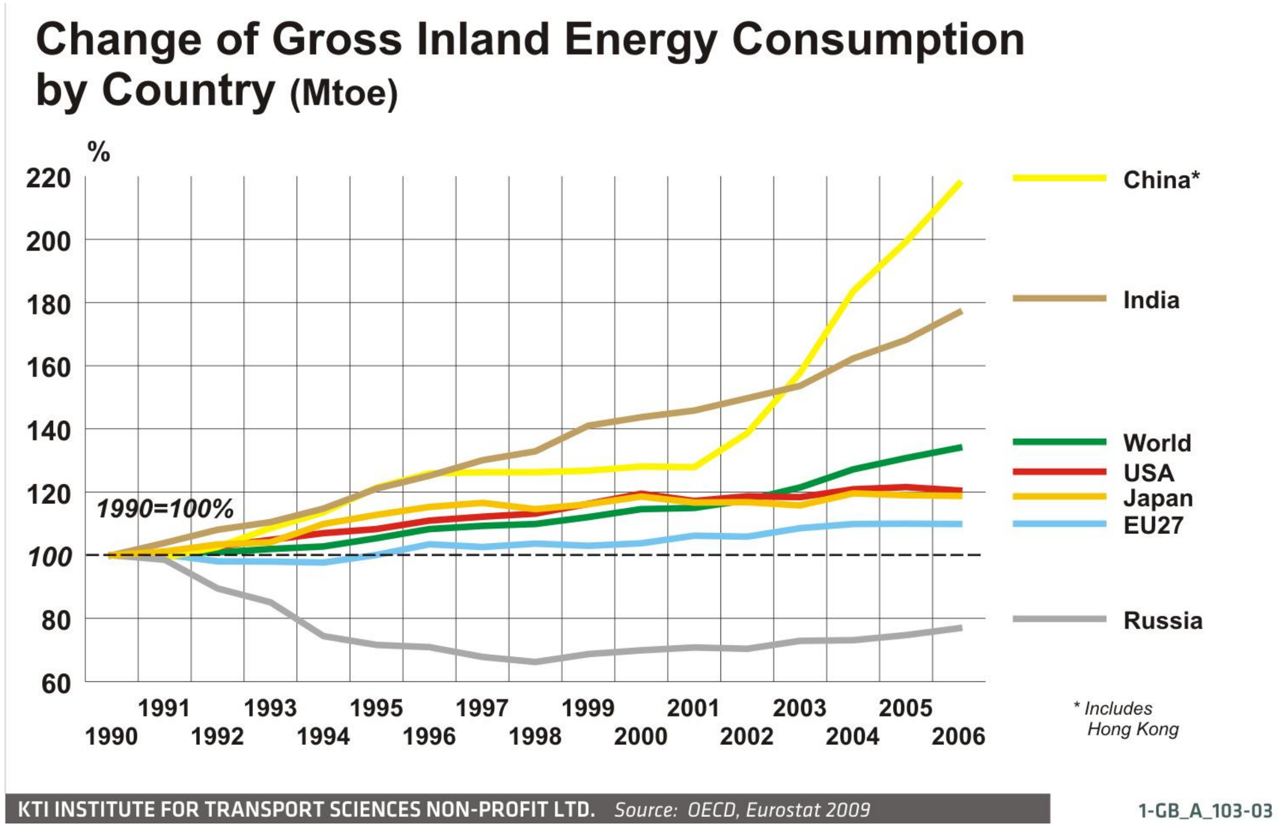



**Fig. 5.**
*Gross energy consumption by Country*
*source: KTI – Institute for Transport Sciences*

**Acknowledgements**

The author is grateful to the support of Bólyai János Research fellowship of HAS (Hungarian Academy of Science)

**References**

Barnes G., Davis G. 2001. Land Use and Travel Choices in the Twin Cities, 1958-1990. Report No. 6 in the Series: *Transportation and Regional Growth Study*. Available at: www.cts.umn.edu/trg/publications/pdfreport/TRGrpt6.pdf, accessed May 28, 2014.

Bieber A., Massot M. H., Orfeuil J. P. 1994. Prospects for Daily Urban Mobility. *Transport Reviews* 14(4), 321-339.
Schafer A., Victor D. G. 2000. The Future Mobility of the World Population. *Transportation Research A* 34(3), 171-205.

Hupkes G. 1982. The Law of Constant Travel Time and Trip-Rates. *Futures,* February, 38-46.

Mokhtarian P. L., Chen C. 2003. TTB or Not TTB, that is the Question: A Review and Analysis of the Empirical Literature on Travel Time (and Money) Budgets, *Transportation Research Part A: Policy and Practice* 38(9-10):643-675, doi: 10.1016/j.tra.2003.12.004

Marchetti, C. 1994. Anthropological Invariants in Travel Behavior. *Technological Forecasting and Social Change* 47, 75-88.

Noland R. B., Lem L. L. 2002. A Review of the Evidence for Induced Travel and Changes in Transportation and Environmental Policy in the US and the UK. *Transportation Research D* 7, 1-26.

Pas E. I. 1998. *Time in Travel Choice Modeling: From Relative Obscurity to Center Stage*. In: Theoretical Foundations of Travel Choice Modeling. Tommy Garling, Thomas Laitila, and Kerstin Westin (Eds.). Elsevier, New York, 231-250.

Schafer A., Victor D.G. 1999. Global passenger travel: implications for carbon dioxide emissions, *Energy* 24:657-679

Taylor B. D. 2002. Rethinking Traffic Congestion. *Access* (the magazine of the University of California Transportation Center) 21:8-16.
Vilhelmson B. 1999. Daily Mobility and the Use of Time for Different Activities: The Case of Sweden. *GeoJournal* 48, 177-185.

Zahavi Y., Ryan J. M. 1980. Stability of Travel Components over Time. *Transportation Research Record* 750, 19-26.
Zahavi Y., Talvitie, A. 1980. Regularities in Travel Time and Money Expenditures. *Transportation Research Record* 750, 13-19.

# ANALYSIS OF LAND USE AND MOBILITY SCENARIOS FOR THE REDUCTION OF TRANSPORT ENERGY IN THE URBAN AREA OF CATANIA

**Giuseppe Inturri**[1]**, Matteo Ignaccolo**[2]**, Michela Le Pira**[3]**, Valentina Mancuso**[4]**, Salvatore Caprì**[5]
[1,2,3,4,5] *University of Catania, Department of Civil Engineering and Architecture*

**Abstract:** A consolidated approach for pursuing sustainable urban mobility consists of reducing transport energy through three key strategies: a land use distribution lowering the need and length of motorized mobility; adoption of measures favouring low impact transport modes; incentives for energy efficient vehicles' fleets.
A simple land use and transport model is presented in this paper to calculate a commuting transport energy indicator able to address the delivery of sustainable urban and transport plans. The model is based on a mathematical description of the transport system and a transport mode choice model following ideal simple rules based on the distance from home to workplace and on zonal transit network accessibility. For each transport mode energy consumption, capacity and load factor are considered. Assignment of flows between home and work zones is formulated as a standard transportation problem to minimize the required transport energy.
The model is applied to the urban area of Catania, a medium-sized town in Italy, for three scenarios: 1) improving walking/cycling facilities and accessibility to the existing transit network; 2) adding four BRT lines and one metro line to the transit network; 3) relocating a fraction of residents around the stations of the metro line. These scenarios reflect some of the main measures of the urban mobility and land use plans which are currently under discussion. For each scenario we compute the "Transport Energy Dependence" indicator, the travelled distance and modal share for each transport mode at the urban and traffic zone scales.
First results show the methodology is suitable to different aims: to evaluate the potential impact of changes in land use, transport and vehicle technologies policies in terms of transport energy consumption; to define a minimum transport energy as a reference value to be compared with the actual consumed transport energy; to define transport energy requirements for the approval of land use and transport plans.

**Keywords**: sustainable mobility; energy efficiency; transportation problem; land use and transport planning; transport energy.

## 1. Introduction

Urban areas are home to 70% EU population and generate 85% of the Union's GDP (EC, 2009). In 2011 transport used 1/3 of all energy and 70% of all oil in EU (EC, 2013). Urban areas produce 25% of all $CO_2$ transport-related emissions (EEA, 2013). Meeting the climate change targets and reducing energy oil dependency is primarily a role of local communities. Cities have to multiply their efforts of achieving the 60% reduction in greenhouse gas emissions as required by the European Commission (EC, 2011). A radical change in the approach towards sustainable mobility is needed to ensure a more energy-efficient and climate proof transport system. An integrated land use and transport policy is one of the most long-term, strategic and effective way to reduce carbon emissions and improve security of energy supply.
Some of the available models make aggregated statistical correlations among urban density and fuel consumptions. Others are based on sophisticated transport demand behavioural approaches to transport mode choice and transport energy consumption. However these models are of scarce utility for professionals and policy decision-makers, who wish to incorporate transport energy issues while choosing adequate urban and transport planning measures.
To this aim we propose a model to evaluate the potential impact of changes in land use, transport and vehicle technologies policies in terms of transport energy consumption. The model calculates a "Transport Energy Dependence" (TED) indicator at the level of the single traffic zone and of the whole urban area. It is not the actual energy consumed, but the amount of energy each person would use for commuting in an ideal condition of trip distribution (commuting to the nearest available workplace) and transport mode choice (the most energy efficient mode of transport is used, according with its availability and the length of the trip). TED indicator is an ideal, easily measurable, minimum transport energy that can be used as a lower limit against which to measure the actual transport energy, or to make comparisons among different neighbourhoods of the same urban area or among different cities; besides it can define the transport energy standard required for the approval of urban and transport plans or to effectively control urban land use and transport system development.

## 2. State of the art

The urban mobility patterns and related energy consumptions are affected by the dimension, density, design and transport level of service of the city, as well as by socio-economic features.
Newman and Kenworthy (1989) found that transport related fuel consumption is reduced by urban density. The idea is that higher density increases the probability of shorter trips' length and then walking and public transport use. Karathodorou et al. (2010) estimated a fuel demand model and found that urban density affects fuel consumption, mostly through variations in the car stock and in the distances travelled, rather than through fuel consumption per kilometre.

[1] Corresponding author: email ginturri@dica.unict.it

Compact, transit accessible, pedestrian oriented, mixed use development patterns and land reuse epitomize the application of the principles of Smart Growth (www.smartgrowth.org/network.php) and other popular movements and planning approaches for urban sustainable development, such as New Urbanism and Transit Oriented Development (La Greca et al., 2011). Planning for high density has two main goals in the context of transport energy consumption: reducing trip length and total mobility by concentrating residential, employment and services areas (Cervero, 1988); changing the modal split to reduce the share of the private car use in relation to public transportation, walking and cycling (Barrett, 1996).

Though it is widely accepted that a compact city is more energy efficient and less polluting because its dwellers can live closer to shops and workplaces and can walk, bike, or take transit, some authors (Neumann, 2005) question whether compact city is a form of sustainable urban development in all point of views. According to Hall (2001), internationally, "Travel is much more strongly linked to fuel prices and income" than population density (p. 103). Actually, the estimation of the real level of energy consumption in the transport system is affected by the complexity of the relationships linking all the factors mutually interacting in the urban system. The effectiveness of a land use policy towards higher densities and mixed use is affected by the willingness of the population to accept high levels of density and by their social attitude towards a change in modal split from the private car to public transportation and non motorized modes. Mode choice behaviour is also affected by the level of service and accessibility provided by low energy transport modes, by car ownership rate and by the transport demand measures adopted to limit the use of cars. Besides, a non-compatible planning of mixed land use might result in higher values of vehicle kilometres travelled to reach activities that may be located far away (Mindali et al., 2004).

In this respect it seems interesting the concept of transport energy specification that Saunders et al. (2008) introduced as a tool to incorporate transport energy into urban planning. It is calculated as a function of the land use and the transport system of an urban area with the aim to ensure development occurs within a defined design boundary. The transport energy specification is not bound to the complex modelling described above. It is a simple indicator of the minimum transport energy used if people would select the most energy efficient mode of transport available according with simple rules based on the distance between land use locations. In synthesis it is not an estimate of the real transport energy consumption but a calculation of a transport energy dependence.

Based on Saunders' approach , we build a new methodology to calculate a transport energy dependence (TED). While Saunders' indicator is designed to be used at neighbourhood level, our index is applied to a whole urban area. Saunders calculates the transport energy dependence in a small enclosed area, where workers are assigned to their closest work locations up to they are saturated, then anyone else is assigned to the city centre by default. We use an optimal assignment model to distribute flows of workers to their workplaces in order to minimize the total transport energy used in the whole urban area. As in the work of Saunders, we use a mode choice model based on simple rules, but it is also conditioned by the accessibility of all different transit system which are explicitly represented; the option of intermodal trips is included.

## 3. Methodology

The proposed methodology is based on three interacting models: land use, transport and energy (Fig. 10). It computes the minimum transport energy required in an urban area, given a set of fixed land uses, transport options and energy variables that represent a planning scenario.

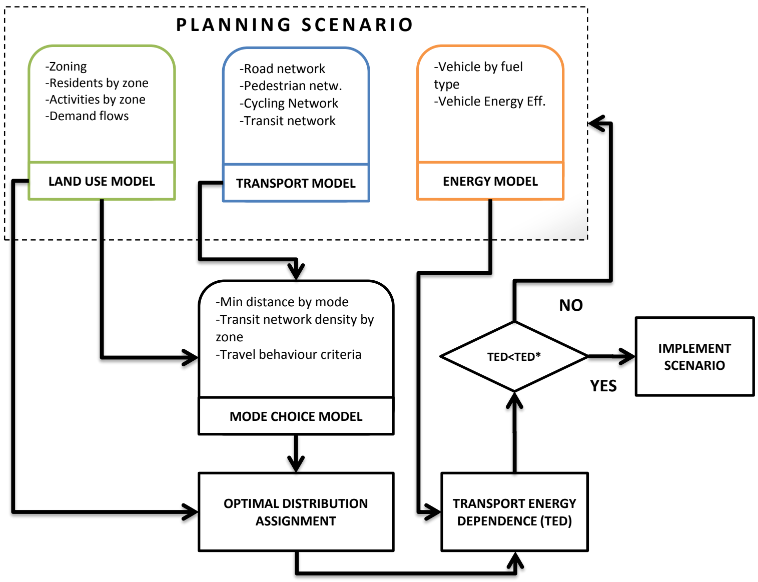



**Fig. 10**
*TED model flow chart*

The land use model specifies the number of residents and activities present in each traffic zones the urban area has been subdivided in. Each traffic zone is assumed to generate a demand flow to all other zones as a function of the number of residents and a fixed trip weekly frequency for each type of destination (workplace, retail, services, etc.). The transport model provides a mathematical description of the road, transit, pedestrian and cycling networks and calculates the shortest distance between each OD pairs for each transport mode. Then the transport mode choice model assigns a transport mode to each OD pair following a set of simple fixed rules based on the length of the trip and transit accessibility. Then the optimal distribution of demand flows among ODs is calculated by solving a standard transportation problem (Hillier and Lieberman, 2001) to minimize the total transport energy in the urban area. Let $Z$ be the total transport energy and $x_{ij}$ the number of workers living in zone $i$ and commuting to zone $j$, the transportation problem is to minimize $Z = \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij}$, subject to $\sum_{j=1}^{n} x_{ij} = o_i$, $\sum_{i=1}^{n} x_{ij} = d_j$, $x_{ij}>=0$, $\sum_{i=1}^{n} o_i = \sum_{j=1}^{n} d_j$ ,where $o_i$ are the workers living in zone $i$, ($i$ = 1, 2, ..., n), $d_j$ the workplaces in zone $j$, ($j$ = 1, 2, ..., n) and $c_{ij}$ the energy consumed travelling from zone $i$ to zone $j$ by the selected transport mode.

Then, for each planning scenario $s$, the total amount of transport energy of the urban area, the so called “Transport Energy Dependence” $TED_s$ (MJ) can be calculated with the following expression:

$$TED_s = \sum_o \sum_d t_{od} \cdot l_{od} \cdot \frac{e_v}{c_v \cdot LF_v}$$

being

$t_{od}$ number of trips assigned from zone $o$ to zone $d$ to minimize $Z$ (passengers)
$l_{od}$ shortest distance between zone $o$ and zone $d$ (km)
$e_v$ unit energy consumption of the transport mode chosen (MJ/km)
$c_v$ capacity of the vehicle (spaces)
$LF_v$ load factor (passengers/spaces)

For a given scenario, $TED_s$ is the minimum energy that in a week would have been consumed if, given a fixed distribution of population and urban functions, and transport mode options, every person could work in the nearest workplace and would choose the best available transport mode for each distance range to be travelled. Of course the model can be easily extended including other weekly trips made by the population for other reasons (school, shopping, entertainment, etc.).

If TED (MJ/week) is divided by the total number of urban population $N$ and expressed in MJ/person/week, it can be used to compare different cities or the same city over time according to its forecasted demographic evolution. When TED is divided by total travelled distance and expressed in MJ/pax-km it can represent a clear indicator of the energy efficiency of the transport system in the urban area. When different scenarios have to be compared within the same urban area, $TED_i$ (MJ/person/week) can be computed for each traffic zone $i$ according to the number of residents $N_i$.

## 4. Case study

### 4.1. General features of the urban area of Catania

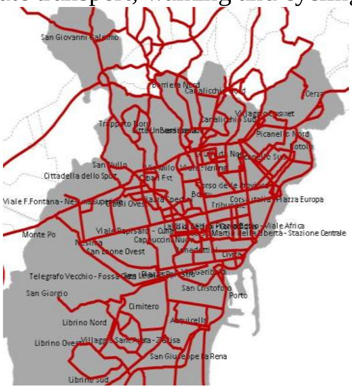
The methodology has been applied to the urban area of Catania, a medium-sized city (300,000 inhabitants) located in the eastern part of Sicily, Italy. The city is part of a greater Metropolitan Area (750,000 inhabitants), which includes the main municipality and 26 surrounding urban centres, some of which constitute a whole urban fabric with Catania. The case-study can be considered as quite significant for similar urban areas, which are characterized by medium population density (about 7,000 people per km$^2$ in the urban area of Catania) and size (about 45 km$^2$), a strong attraction towards the central business district from the surrounding areas, heavy car traffic volumes along radial routes mainly for commuting purposes (about 20,000 vph during the peak hour of the average working day) and a limited public transport ridership. The main city contains most of the working activities, mixed with residential areas. Even if several attraction polarities (hospitals, main schools, shopping centres) are spread over the whole territory, the transport demand pattern is mostly radial. The modal share for public transport is about 15%, while the amount of travelled kilometres by bicycle is negligible (even if increasing), thus leading to a modal share of individual transport close to 85%. Regarding transport supply, the urban road network is about 700 km long, and it consists of a highway ring connecting all radial highways to the other Sicilian towns, few internal major roads (as ‘extensions’ of the main highways and rural roads) and a internal road network, which lacks a complete hierarchical organisation. The urban public transport consists of an extensive regular bus services network (about 250 km long), and a short metro line, mainly underground, of less than 4 km, which is going to be extended to 11 km within a couple of years. Traffic congestion, limited public transport utilisation, little diffusion of cycling and walking for systematic trips, inefficiency of the parking management, absence of city logistics measures are the main critical issues for the transport system of Catania.

### 4.2. Transport system

The transport model consists of:

- Transport demand: only commuting flows are considered; a frequency of 5 home-to-work trips/week is assigned to all employed population.
- Transport supply: the road network is composed of 516 nodes and 1122 links; transit network considers 49 bus lines, 4 BRT lines and 1 metro line. The transport network is implemented within PTV VISUM software package used to compute the shortest path between all origin and destination pairs by all modes of transport. When more transit systems are available, the shortest path is a combination of them, thus including the option of transit intermodality.

Fig. 11a shows the coverage of the transit network operated with regular bus lines, Fig. 11b is the road network available for private transport, walking and cycling.



Bus network (a)

Road Network (b)

**Fig. 11**
*Bus network (a) and road network (b) in Catania*

The land use model considers a subdivision of the urban area in 50 traffic zones (Fig. 12). For each zone the following information are available: households, employed population, non-working population, number of activities and number of workplaces.

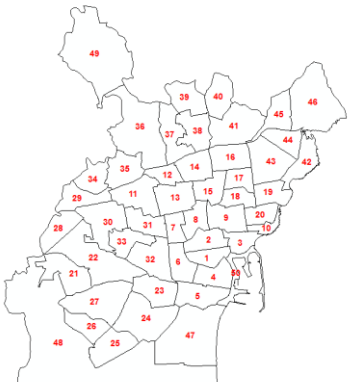


**Fig. 12**
*Zoning of the urban area of Catania*

### 4.3. Transport mode choice model

Transport modes considered are: walking, cycling, regular bus transit, bus rapid transit, metro and private car. The shortest distance among each pair of traffic zones is calculated; distances are measured on the road network and are assumed to be the same for walking, cycling and car modes. Then for each OD pairs of zones, the following simple mode choice model is adopted:

1. if the distance to the first available workplace is less than 500 m, then walking is the chosen mode;
2. if the distance to the first available workplace is less than 1000 m, then cycling is the chosen mode;
3. if conditions 1 and 2 do not occurs and the zone transit network density (both in origin and in destination zones) overcomes a “transit network density threshold” (calculated as described soon after), then transit is used;
4. if none of the previous conditions occurs, then private car is used.

There is not an explicit representation of the internal pedestrian and cycling paths within each zone, so the “equivalent radius” of each traffic zone is calculated as the radius of the circle having the same area. As far as the number of workers do not exceed the available workplaces of a zone , they are assumed to reach their works by walking or cycling within their origin zone, if the equivalent radius is less than 500 m or 1000 m respectively; otherwise, motorized modes are taken into consideration.

The transit network density threshold is calculated as the one corresponding to the maximum distance a person is willing to walk to access a transit stop, assuming a uniform distribution of the lines along a grid street pattern within the zone, according with the scheme in Fig. 13. In Table 1 the threshold is calculated for the bus, BRT and metro considering an increasing willingness to walk to the stop as a function of the transit system performance (Dittmar et al., 2004). For each zone and each transit system the transit network density is calculated as the ratio between the total extension of the transit lines crossing the zone and its area. It is expressed in km/km$^2$.


**Fig. 13**
*Transit density threshold to walk to stop*

**Table 1**
*Transit density threshold*

| Transit network | Maximum walking distance | Transit density threshold |
|---|---|---|
| Regular Bus Transit | 300 m | 6.67 km/km$^2$ |
| Bus Rapid Transit | 600 m | 3.30 km/km$^2$ |
| Metro Transit | 800 m | 2.50 km/km$^2$ |

We refer to the study of Kenworthy (2003) for the unit energy consumption used for each mode of transport.

**Table 2**
*Unit energy consumption*

| Mode of transport | Unit energy consumption |
|---|---|
| Private Car | 3.30 MJ/pax-km |
| Regular Bus Transit | 1.17 MJ/pax-km |
| Bus Rapid Transit | 0.69 MJ/pax-km |
| Metro Transit | 0.48 MJ/pax-km |

**4.4. Scenarios**

TED is calculated to test the land use and transport planning strategies described in the following scenarios. Scenario 0 represents the present situation and the baseline for evaluating other scenarios. Scenario 1 represents a short term transport policy, where the catchment area of the bus system is enlarged by improving the easiness to access the transit stops, through better pedestrian infrastructures, pedestrian safety measures and car traffic calming measures meant to favour walking to transit. Using the concept of equivalent distance (Wibowo et al., 2005), it is assumed that all measures adopted will increase the maximum willingness to walk to the nearest transit stop from 300 m of the scenario 0 up to 400 m in the scenario 1. The consequence is that the bus network density threshold is reduced from 6.67 to 5 and the number of zones overcoming the limit is increased from 10 of scenario 0 to 26 in scenario 1 (Fig. 14).

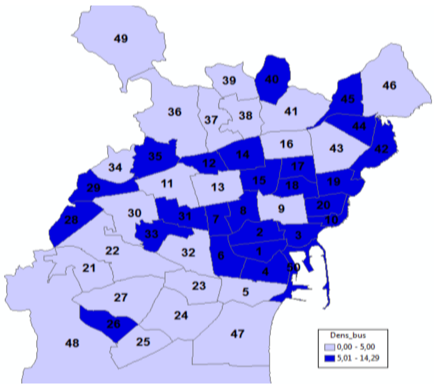



**Fig. 14**
*Distribution of bus network density in scenario 1*

Scenario 2 is a medium-long term transport policy where the transit network is enhanced by the introduction of four BRT lines and the extension of one metro line (**Fig. 15**). Again, the transit density threshold are lowered according with Table 1.

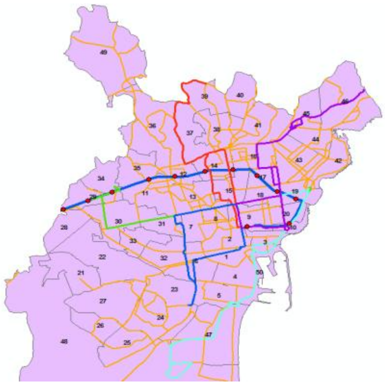


**Fig. 15**
*Transit network in scenario 2: regular bus (yellow), metro (thick blue), BRT (other colours)*

Scenario 3 represents a long term spontaneous land use change. It assumes a 10% increase of density in the zones directly served both in origin and destinations by the metro line as a consequence of the re-locations of households from all other zones attracted by the accessibility improvement. The scenario may simulate the effect of adopting the so called Transit Oriented Development policies.

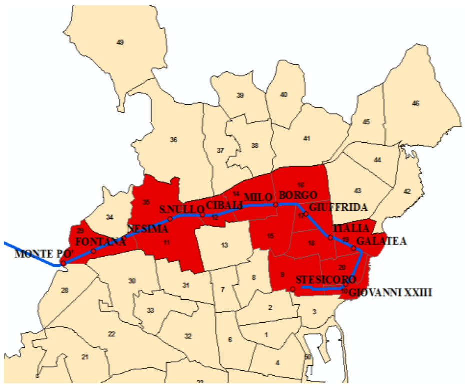


**Fig. 16**
*Household densification along the metro line in scenario 2*

## 5. Results

The improved accessibility of bus transit stops in scenario 1 reduces the TED per person by more than 16.6% against the scenario 0; the introduction of the BRT network and the metro line of scenario 2 determines a 24.2% reduction, while the relocation of land uses in scenario 3 produces a 27.1% reduction, always with respect to scenario 0 (Fig. 17).

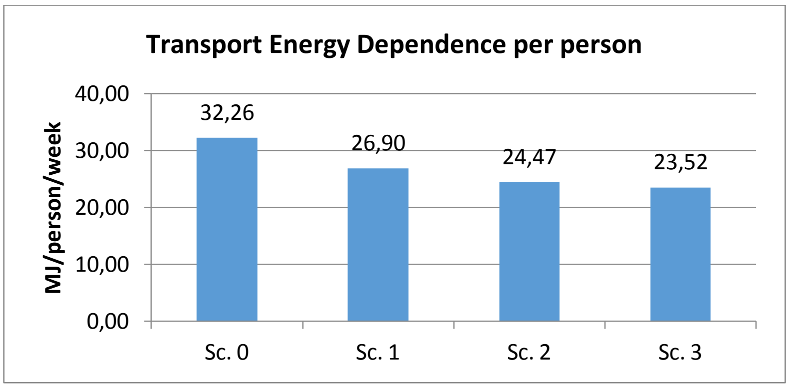



**Fig. 17**
*Transport Energy Dependence for the analysed scenarios.*

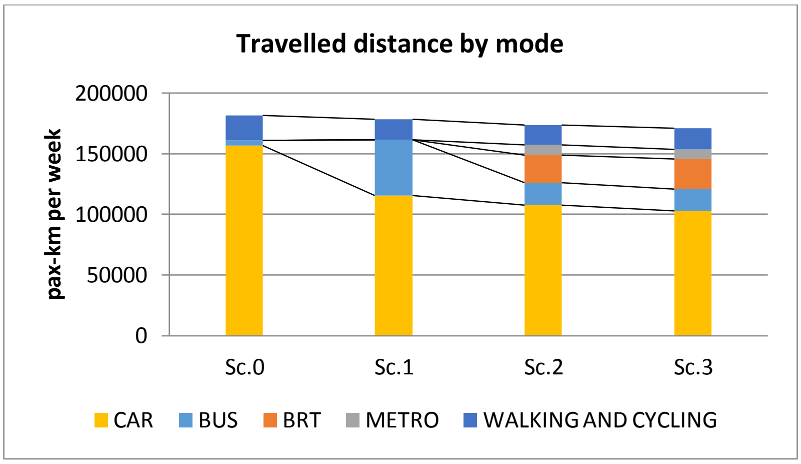



**Fig. 18**
*Modal share for the analysed scenarios*

As it can be seen in Fig. 18 and Table 3, scenario 1 records only a small reduction in travelled distances (-1.7%) but a quite high reduction of TED (16.6%) because of the great increase in bus share (from 2.3 % to 25.6%) that lowers the inefficient use of cars (from 86.4 to 64.8%), though also non-motorized modes are lightly reduced (from 11.3 to 9.5%). As expected, in scenario 2, BRT and metro contribute to an increase in public transport use (from 25.6 to 28.3%), both for the better connectivity of higher performance systems and new available intermodality options. Also walking and cycling have a small increase (1%) probably as a consequence of a better assignment of flows of workers to workplaces. TED keeps on lowering because of the better energy efficiency of BRT and metro compared with other motorized modes and for a further reduction in car share (3.0%). The relocation of land use activities around the metro stops in scenario 3, brings small increases in transit share (1.6%). The total reduction of TED is 5.9% against the scenario 2.

**Table 3**
*Results of scenarios: modal share (pax-km percentage), total passengers-km and TED values*

| | Sc.0 | Sc.1 | Sc.2 | Sc.3 |
|---|---|---|---|---|
| BUS | 2.3% | 25.6% | 10.7% | 10.5% |
| BRT | 0.0% | 0.0% | 13.0% | 14.6% |
| METRO | 0.0% | 0.0% | 4.6% | 4.8% |
| TOTAL TRANSIT | 2.3% | 25.6% | 28.3% | 29.9% |
| CAR | 86.4% | 64.8% | 61.2% | 60.1% |
| WALKING AND CYCLING | 11.3% | 9.5% | 10.5% | 10.1% |
| ALL MODES | 100.0% | 100.0% | 100.0% | 100.0% |
| PAX-KM | 181597 | 178574 | 175628 | 170904 |
| TED (MJ/PERSON/WEEK) | 32.26 | 26.90 | 24.47 | 23.52 |
| UNIT TED (MJ/PAX-KM) | 2.88 | 2.44 | 2.26 | 2.23 |

In Fig. 19 TED per each traffic zone in scenario 0 is shown. This indicator is useful to evaluate to what extent transport energy dependence of a zone is affected by its centrality, its connection with the transit network, walking and cycling permeability and its balance between residents and workplaces.

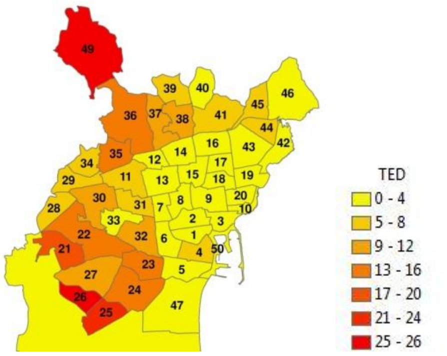



**Fig. 19**
*Scenario 0: TED per traffic zone and per person*

## 6. Conclusion

A simple land use and transport model has been presented to calculate a commuting transport energy indicator to support the delivery of sustainable urban and transport plans. The model is based on a mathematical description of the transport system, where transport mode choice follows ideal simple rules based on distance from home to workplace and transit network accessibility.

The model has been applied to the urban area of Catania and the impacts of different planning scenarios has been tested, reflecting some of the main measures of the urban mobility and land use plans which are under discussion.
The results have shown the methodology is suitable to evaluate the potential impact of changes in land use and transport policies in terms of transport energy consumption. This could easily lead to the definition of transport energy requirements for the approval of land use and transport plans and to support a sustainable development of our cities.

**References**

Barrett, G. 1996. In: Jenks, M., Burton, E., Williams, K. (Eds.), *The Transport Dimension*, *The Compact City: A Sustainable Urban Form?*, E&FN Spon Publishers, London and New York, pp. 171–180

Cervero, R. 1988. Land use mixing and suburban mobility. *Transportation Quarterly* 42 (3): 429–446

Dittmar, H.; Ohland. 2004. *The new transit town: best practices in Transit Oriented Development.* Island Press. Washington, D.C.

European Commission. 2009. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions 'Action Plan on Urban Mobility' (COM(2009) 490 final of 30 September 2009).

European Commission. 2011. White Paper 'Roadmap to a Single European Transport Area —Towards a competitive and resource efficient transport system' (COM(2011) 144 final of 28 March 2011)

European Commission. 2013. EU Transport in figures. Statistical Pocketbook 2013. ISBN 978-92-79-28860-9. doi:10.2832/19314.

European Environmental Agency. 2013. A closer look at urban transport. TERM 2013: transport indicators tracking progress towards environmental targets in Europe. European Environment Agency, Report 11/2013

Hall, P. 2001. *Sustainable cities or town cramming?* In Planning for a sustainable future, edited by A. Layard, S. Davoudi, and S. Batty. London: Spon.

Hillier, F.S.; Lieberman, G.J. 2001. *Introduction to Operations Research.* McGraw-Hill, New York.

Karathodorou, N.; Graham, D.J.; Noland, R.B. 2010. Estimating the effect of urban density on fuel demand. *Energy Economics* 32: 86–92

Kenworthy, J.R. 2003. Transport energy use and greenhouse gases in urban passenger transport systems: A study of 84 global cities. In: *International Sustainability Conference, 17 - 19 September, Fremantle, Western Australia*

La Greca, P; Barbarossa, L; Ignaccolo, M.; Inturri, G.; Martinico, F. 2001. The density dilemma. A proposal for introducing smart growth principles in a sprawling settlement within Catania Metropolitan Area *Cities,* 28 (6)

Mindali, O.; Raveh, A.; Salomon, I. 2004. Urban density and energy consumption: a new look at old statistics. *Transportation Research A* 38: 143-162.

Neuman, M. 2005. The compact city fallacy. *Journal of Planning Education and Research* 25:11-26

Newman, P.; Kenworthy, J.R. 1989. *Cities and Automobile Dependence: An International Sourcebook*, Gower, Aldershot

Wibowo S. S. and Olszewski P., 2005. Modeling Walking Accessibility to Public Transport Terminals: Case Study of Singapore Mass Rapid Transit. *Journal of the Eastern Asia Society for Transportation Studies*, Vol. 6,

Saunders, M.J.; Kuhnimhof, T.; Chlond, B.; Rodrigues da Silva, A.N. 2008. Incorporating transport energy into urban planning. *Transportation Research Part A* 42

*Transport Policy*

# ESTIMATION OF RAILWAY PROJECTS DURATION AT INITIATING PHASE: A STATISTIC APPROACH

**Federico Fiori[1], Edoardo Favari[2], Andrea Felice Grassi[3]**
[1] *Department of Civil and Environmental Engineering, Politecnico di Milano, Italy*
[2] *Politecnico di Milano, Italy*
[3] *Politecnico di Milano, Italy*

**Abstract:** In Large Transportation Projects (LTPs) estimation of duration provided in tender and initiating phase are often contradicted in executing phase, due to a mix of technical and contractual problems arising during the project execution. This paper present a research that apply a statistic approach to this problem and, determining the key features of a LTP and a set of historic data coming from the railway field, present a quantitative solution. The first part of the paper makes an analysis of existing literature on this topic, then the methodology is discussed. In the second part, a statistical model is presented, to show how the set of identified relevant variables are used to estimate the duration according to project features avoiding other goals such as political goals of cutting the project duration. In fact, the common sense of technical advisers is that the duration of a project is often given by non-technical goals, but the drivers are more referred to a specific deadline such as elections or large public events (World Cups, Universal Exposition etc.), asking to cut and cut time and, as a result, projects become more expensive, hard to manage, and face technical problems in the operating phase. In the third part, this method is calibrated on a set of input coming from real railway LTPs of the last ten years. The aim of this paper is to let the decision makers know that, given a new project, the past experience says that it will have a given duration according to its features, and to ask for a shorter project duration would be risky and cause project failure, or let the project cost rise. It also allow to define a guideline to determine which contractual form is the best according to the project.

**Keywords**: Railway Project Management, Contract Management, Project Duration Estimation

## 1. Introduction

In this paper the Authors investigated a new way for determining the duration of a Large Transportation Project (LTP) *ex ante*, depending on some project features both technical and managerial. When completed, the research will allow to perform simulation at the pre-tender phase of project, enabling the contracting authority to tune the project features at the be to reduce duration, costs and litigation in the executing phase.
The point of view of the Authors is that there are a set of drivers in LTPs that, if optimized, can significantly improve the projects' performance in terms of duration and costs. These drivers are technical, such as the length of a railway project, the number of stations, etc., but also managerial, such as the type of contract adopted for contracting etc. These drivers, put together in a mathematic model, can provide information on expected project performances, and, by simulating scenarios, lead the contracting authority to optimize these drivers of a given project.
The Authors performed an analysis on existing literature, even if on the specific topic there is nothing really fitting, and everything currently available is referring to classic risk management, which is not suitable for the purpose of this paper.
A duration model is defined, including all the variables available in the database used for the analysis, which was a railway projects database. Thanks to the database, the model is also calibrated, and the accuracy of the new method is calculated by MAPE (Mean Absolute Percentage Error), giving very good results.
The work is at the very beginning, so the perspective are also discussed at the end, and seem to be potentially very relevant.

### 1.1 Purpose and Objectives

A deep analysis of the existing literature and of the work-world, has pointed out the current lack of reliable models and tools, really able to face complexity and uncertainty of Large Transportation Projects and, on other hand, able to provide a close to reality duration estimate, since project planning phase, with low effort in term of time and resources.
The existing and commonly used techniques for project scheduling and project duration estimate have several limitations if matched with the context, the environment and the public procedures of LTPs. As a consequence, the study has been focused on implementing an approach, created on purpose, to overcome limitations and hurdles of both existing scheduling techniques and general behavior of contractors and contracting authorities dealing with LTPs planning at tender phase.

[1] Corresponding author:federico.fiori@polimi.it

In order to assure the decision makers to know at planning phase that, given a new project, it will last a certain time according to its internal and external features the following prerequisites have to be balanced and matched:

- few data already known at general planning phase (pre-bid phase) but very representative of LTPs duration, in order to avoid problems of lacking project specification;
- the most important variables of an LTP, including ones linked to uncertainty, peculiarities, contingencies and complexity should be accounted;
- overcome biases and approximation due to interrelations between single project activities, that means avoiding bottom-up analysis (i.e. project duration is not estimated as the sum of activities duration on the critical path);
- avoid complex, time-consuming and usually inaccurate risk quantification and risk assessment tasks;
- reliable approach based on objective data and not on experts elicitation, in order to avoid experts biases and to provide an outside view;
- useful for any decision makers (e.g. contractors, contracting authorities etc.) in order to avoid inadequate bid, legal dispute, penalties and any problem, linked to extra-time and cost, causing project failure;
- no need of integrating the approach with existing software and procedures.
- user friendly approach, immediate implementation and easy analysis of the outputs.

The statistical approach proposed in this paper has been conceived to successfully answer each prerequisite listed just above.

### 1.2 Methodology

The first step of the research introduced in this paper, has been the focus on the way planning of LTPs (but following reflections can be spread to public construction projects in general) is approached by both contractors and contracting authorities; simultaneously the investigation of pathologic time and cost overrun for such projects all over the world has been accomplished. The main topics leading to projects failure seem to be a diffused contractors and contracting authorities behavior related to public tender process, poor project planning due to unsuitable project specifications for bidding, lack of suitable estimation tools (i.e. experience shows that unsuitability is caused by innate approximation features as well as challenging applicability) able to face the context (i.e. uncertainty due to contingencies, technical and operational conflicts and hurdles, external influences etc.), lack of aggregate and objective historical data and furthermore the so called myths of “think positive” and of “schedule compression”. In order to overcome these problems the proposed approach has been based on the prerequisites listed in paragraph 1.1. Since the background and the desired findings have been defined, the research went on selecting the most appropriate modeling approach and the most representative variables to be accounted, as illustrated in section 2. In order to set the best fitting technique, different approaches have been investigated and some others are currently under studying. The collection of a set of input data coming from real railway LTPs of the last ten years has allowed the calibration and the validation of the proposed model, as shown in section 3.

### 1.3 Bibliographical review and state of the art

Until nowadays research devoted to estimate large engineering project duration has been carried on in order to provide best tuning for scheduling and network techniques (i.e. CPM, PERT, Stochastic simulation methods) or investigate time-cost relationship models.

Some authors (Jacob 2003, Vandevoorde and Vanhoucke 2006, Lipke et al. 2009, Caron et al. 2013) basing on the Earned Value Management approach, have proposed formulations and statistical analysis to forecast the expected duration of a project at completion; unfortunately this approaches are valid only if employed during project execution and are not suitable at planning phase.

Speaking about methods able to provide time estimation at project planning phase, many studies have been focused in the past just on modeling relationship between project duration and project costs defining best mathematical forms fitting this relationship. Some researches proceeded in defining optimized time-cost activity tradeoff able to point out constraints, to be matched with traditional scheduling and network techniques.

Many authors proposed prediction models able to estimate construction duration of civil engineering project (almost all in building construction field) according to project cost considered as unique variable. Despite several authors established that construction duration is determined by numerous internal and certain factors and also derived a hierarchy of these factors (i.e. Kumaraswamy and Chan 1995), very few went on investigating the relationship between these variables and the duration of the project. Even if it is true that almost everything in a construction project can be approximated into a cost value, this approach is too reductive especially if related prediction models must be useful for real employment at project planning phase. Such models are based on the unique assumption that more cost means more duration or in other words same cost means same duration; the limitation of neglecting other useful information can lead to erroneous duration outcomes. Furthermore at planning phase this rough approximation is unsuitable to describe complexity and uncertainty features of large engineering projects and of a construction site at all.

Some studies have been addressed to define piecewise functions according to time-costs relationship for each project activities. Recently, some studies, mainly from China and Korea (e.g. Zheng, D.X.M. et al. 2004), have proposed discrete formulations and probabilistic distribution (i.e. Weibull distribution) for fitting time-cost relationship and sophisticated algorithm to define on a case-by-case basis which is the best mathematical formulation.
According to time-cost approach to estimate project duration, several attempts to perform a bi-criterion time-cost analysis have been carried on by authors of project management discipline with reference to construction projects in different country; these models has been calibrated in order to provide the mathematical form and its coefficients, fitting local conditions and environment. Kumaraswamy and Chan, 1995 conducted a wide survey in order to investigate time-cost relationship for both building and civil engineering projects in Hong Kong. During the survey they investigated main factors affecting construction project duration and stated that besides the project characteristic macro variables many significant factors influence construction durations. They recommended that future research should be conducted to identify significant variables. An attempt to evaluate influence of other variables apart from costs on construction projects duration, has been carried on by Kaka and Price, 1991; as other authors before, they proposed an empirical time-cost formulation calibrated through linear regression of data from buildings construction and small roadwork projects performed within the period 1984-1989 in UK. The following stage of the analysis was the classification of sample projects according to type of project, type of contract and type of competition (bid, non-bid etc.). Seven groups were created and tested for highlighting difference among estimated and actual duration relationships; curiously the type of competition was found to have no effect on the cost-duration relationship.
The time-cost approach is not the only developed and available in specialist literature. As a matter of fact some authors, neglecting time-cost relationship has proposed models able to predict construction duration works implementing a matched approach between probabilistic simulation and on purpose tuned scheduling techniques; an example is a recent study of Nguyen L. et al., 2013 defining a model to predict building construction duration. Typical construction activity sequences in building construction were first identified, and then the statistical distributions of controlling activities on the works sequences were surveyed; this model conduct to knowledge of construction time forecast by introducing the concept of controlling (in other words essential) activities. Many studies present tuning for computer simulation approaches, that is to say risk simulation techniques based on Montecarlo method or other stochastic methods, to be employed through scheduling techniques as PERT or CPM where the sum of duration of all activities obtained through simulation represents the total project duration.
A deep bibliographical research has been performed but very few of the studies reviewed directly concerned LTPs and project duration estimate at initiating phase. A relevant study for this purpose was conducted by Irfan et al., 2010; they defined both a deterministic and probabilistic approach to evaluate relationship between duration, cost and contract type of some kind of works in highways construction (pavement construction and rehabilitation, highway asset maintenance, traffic facilities installation and bridge replacement) in the state of Indiana (U.S.A.). Even if this study is really devoted to transport infrastructure field, it cannot describe the complex multiplicity of Large Transportation Projects; as a matter of fact no variables except from cost and contract type have been considered. Real employment of this approach highlights a quite wide approximation; the model lead to the same result if two different projects have the same cost and the same contract type. Moreover the approach assumption of "more cost means more duration" lead to wrong result when any technology or technical choice is able to reduce construction time at the expense of cost increase. Furthermore no own features of each project able to influence duration can be represented just by estimated cost at planning phase.
The research introduced in this paper, conceived at Politecnico di Milano (Grassi A.F. 2013) and still ongoing can overcome limitations of the existing approaches providing a useful planning tool for real employment at project planning phase and created on purpose for LTPs.

## 2. The model: statistical approach and variables

In order to face LTPs, providing a model able to estimate their duration at project planning phase, more than a couple of representative features of project duration need to be taken into account. On this assumption the only way to perform the analysis is to investigate sophisticated econometrics techniques; the most common modeling techniques universally applied for prediction analysis are:

- Linear or Non-linear Regression Model
- Discrete Choice Model
- Simulation Based Estimation
- Limited Dependent Variable – Model for Duration Data

In selecting a prediction modeling technique three main factors of analysis should be considered. The first one is related to the attributes of the object whose prediction must be estimated; this concerns the type of explanatory variables (i.e. if variables value increase, model output must have a reasonable accordance). The second factor concerns the nature of the input data of the explanatory variables (e.g. if the data is cross-sectional, time-series or panel, if the dependent variable is discrete or continuous, if sufficient condition-based, age-based, hybrid model data) and which data of explanatory variables are available to be analyzed (e.g. performance data, indirect duration data).

The third factor is related to the modeling techniques on its own (e.g. if a deterministic or probabilistic model is preferred; the specific statistical technique to be used, the measure of goodness-of-fit used to validate model results).

### 2.1. Modeling technique

In accordance with existing literature the most suitable techniques identified for potential application in duration estimate modeling and analysis are Regression Models and Models for Duration Data. Duration models represent a probabilistic approach for predicting the likelihood of a continuous dependent variable passing beyond or "surviving" at any given unit of time. The variable of interest in the analysis of duration is the length of time that elapses from the beginning of some event either until its end or until the measurement is taken, which may precede termination (Green 2012). Duration models allow to perform, besides a probabilistic approach, also a regression like approach matching simplicity of regression models while relaxing their constraints in term of input data and mathematical analysis. In each contexts duration models to measure the duration of certain states need to be applied, it is required to define how the durations are distributed, and above all how they are affected by relevant covariates. This last feature is crucial for the application of a prediction model to estimate complex states depending on several variables as the one discussed in this paper that is LTPs duration. The most common and ease approach to perform an estimate analysis through duration models is the estimation of a parametric model by means of Maximum Likelihood. Thanks to this approach, given the hypothesis that the durations follow some definite probability law, it is possible to estimate the parameters of that law, factoring in the influence of covariates.
According to the findings of specific literature, in order to introduce influence of covariates in duration models, it is possible to assume a linear relationship between the logarithm of the survival time T and the characteristics of the covariates X:

$$\ln(T) = X\beta + z \quad (1)$$

where $\beta$ is a vector of coefficients and z is an error term.

Equation 1 leads to Weibull, generalized gamma, log-normal or log-logistic models for T; in survival analysis, this model is called Accelerated Failure Time Model (AFT model) and provides an alternative to the commonly used probabilistic hazards models, defining a direct relationship between time duration (failure time) and covariates, that are the explanatory factors of the item whose duration is investigated. AFT models can therefore be framed as linear models for the logarithm of the survival time; for this reason they are called probabilistic regression-like approach. AFT models are fully parametric models that means a probability distribution must be specified. In order to be used in an AFT model, a distribution must include a scale parameter; the logarithm of the scale parameter is then modeled as a linear function of the covariates. In other words a probabilistic distribution must be defined for the estimated durations.
The log-logistic distribution provides the most commonly used AFT model as well as the Weibull distribution even if gamma and log-normal distribution can be applied. According to what has been stated before AFT model seems a valid solution for describing both complexity and uncertainty of the object of such an analytical simulation. As a matter of fact it is able to overcame limitations of linear and non-linear regression analysis providing the same simplicity in outcomes comprehension, in immediate interpretation of covariates influence as well as ease in covariates update or substitution. Furthermore with AFT model is possible to perform an enough precise and robust regression analysis for the purpose of estimate LTPs duration; obviously estimate duration is not expected to be exact but the introduction of probabilistic elements into a deterministic duration model, as stated by some authors, allows to match outcomes precision and robustness. The following expression is the functional form of the model which rules the analysis:

$$T = e^{\beta_0 + \sum_{i=1}^{n} \beta_i X_i} \quad (2)$$

It is composed by the following four elements:

- dependent (endogenous) variable T: it is the outcome of the model that is the estimated duration;
- independent (explanatory) variables $X_i$: they are quantitative or dummy variables not changing over time;
- constant term $\beta_0$: it is a coefficient defined through the calibration of the model;
- calibration coefficients $\beta_i$ of the covariates: they are numerical values related to the covariates defined thanks to the calibration of the model; they represent the influence of the covariates on the dependent variable T.

It is important to underline the basic assumption of AFT models about time invariance of the quantitative explanatory variables to introduce in the model. Furthermore, for econometric applications, the crucial difference between an endogenous and an explanatory variable is that the explanatory variables are not systematically affected by changes in the other variables of the model, especially by changes in the endogenous variables. Speaking about AFT models, the assumption of exogeneity of all explanatory variables can be relaxed but the choice of the factors conditioning LTPs duration has been carried on in order to create a model where explanatory variables can be considered exogenous at a large scale.

### 2.1. Choice of the explanatory variables

It is important to point out that the introduction in the AFT model of several explanatory variables (mainly non-dummy quantitative variables) could lead to less accurate outputs; in other words adding explanatory variables is not consistent with estimate expectation but could be source of disturbance because those variables may be not truly exogenous at all. This circumstance does not mean necessarily that the model cannot elaborate a factor at all, but it can be required a different quantification and classification of that factor only. In brief the choice of the best consistent combination of explanatory variables to introduce in the AFT model is not a priori known but should be investigated through an iterative model calibration process, in order to evaluate several alternatives and obtain the best model validation possible. With reference to the frame proposed, it is important to underline that four main inputs in the selection of the explanatory variables have been respected:

- exogeneity of the explanatory variables at large scale;
- positive duration dependence of the explanatory variables;
- reduce the number of the explanatory variables as much as possible to introduce in the model just the exogenous variables really able to explain the output variable;
- univocal knowledge of the explanatory variables at project planning phase according to patent lack of project specifications available at tender phase.

The first three assumptions of the previous list guarantee more realistic and robust outputs while the last one is a required assumption for the approach introduced in this paper.
It is possible to state that the duration of an LTP depends on four macro categories:

- technical and dimensional features of the project;
- peculiar features of the construction site;
- project environment contingencies causing delay or time advance on planned schedule;
- contractual features of the project.

Since a deep analysis of the first three points of the list above have been already accomplished for different kind of LTPs (Grassi 2013), in this section just the impact of the contractual features of a project is discussed.
According to the authors' experience, a good contract is a key to avoid conflicts during the project, since the earliest to the latest stage, and so a critical driver to allow a project to stay inside the give duration, or to delay and delay project conclusion. That's why the authors wanted to introduce in the model some variable regarding the contract features of the project, in order to define how contract feature are involved in the project duration. It is almost impossible to objectively define quantitative variables enabling to capture all the feature of a contract: a contract is written is a literature work, written in natural language, even if a technical language, and, as far as it is not an equation, there are no numbers to be taken and put in another equation. So the authors' choice has been for dummy variables, activating factors in the regression model when a contractual feature is present or not in the project contract. There are almost infinitive features that could be taken into consideration in contracts. Due to the fact that our analysis starts from a database of railway and underground projects, the Authors would have had also the opportunity to trace a lot of features that Public-Private Partnership contracts have, but, frankly speaking, this would have required a huge level of effort in analyzing the contract, defining which feature would be relevant or not for the purpose, and calibrating all these variables; probably this would be a good topic for a doctoral dissertation but for this paper was out of reach. Instead the authors concentrate on few feature, considering that the important contribution of this paper is to demonstrate that this approach is possible and it is also relevant and applicable in real projects. At the end, considering the information available in the database, it has been decided to put in the model just the three high level features of contracts, meaning:

- Fixed-price contracts (lump sum): a total price for a defined product or service to be provided is set since the beginning, meaning that the risk of unforecast activities is transferred to the contractor;
- Cost-reimbursable contracts: the contractor is reimbursed according to actual costs incurred plus a fee representing the profit; in this case the risk unforecast activities stays in the contracting authority scope;
- Time and Materials (T&M) contracts: actually the less used in the railway field, and in case used only for little part of the project scope.

This categorization has been pretty easy to make, because is very evident if an activity is paid by fixed-price, cost-reimbursable or T&M. The only issue for modeling is that most of the revamping projects are hybrid, willing to pay part of the scope by fixed-price (generally speaking the new part to be built) and part of the scope by cost-reimbursable (generally the part of revamping, or all the stages of the project necessary to pass from the existing system to the new, and revamped, one. In this case it has been decided to implement not dummy variables, but fractional variable, according to the percentage of project scope paid by fixed-price, and the complementary for the part paid by cost-reimbursable.

For example, for a project whose total amount is 10M€, having 8M€ paid by fixed-price and the rest by cost-reimbursable, it has been put 0.8 for the fixed-price factor and 0.2 for the cost-reimbursable factor in the duration model. This approach, pretty easy to implement, gives a good return in terms of Error Analysis and allows to take into account into the model potentially an unlimited number of contractual features. In this paper, it has also been consider the possibility that a given project could have a longer duration if the contract is fixed-price or cost-reimbursable: the Authors understand that is a naïve approach, because the topic must be investigated in deep according to hundreds of contractual features available, and in addition to the hundreds of project features paired, one by one, to each contractual feature, but they wanted to include and investigate *in nuce* in this paper, and the authors' willing is to investigate this possibility in the next few years by a PhD research. The aim of the contractual analysis for project duration is to understand which is the level of risk that is wise to transfer from the contracting authority to the contractor in order to reduce at maximum the overall cost of a project, that are deeply linked to the project duration. This kind of analysis, that this paper just investigate only at the very early stage, from the author's point of view will be a key topic for contractual risk management that is becoming more and more critical for Large Transportation Projects worldwide.

## 3. Model calibration and validation

Once duration model specification and input data collection has been performed, the analysis requires the calibration and the validation of the model. The calibration of the model proposed has been performed through a statistical analysis software. Validation, the verification of model outcomes comparing real duration with estimated duration, has been carried out through sample data not processed for calibration.
In order to accomplish a sample calibration, the most representative explanatory variables have been identified thanks to appropriate collation of the railway projects data. Technical variables (e.g. length of the line), project features variables (e.g. high speed line) has been identified in addition to project cost and project contract type. Similarly the most frequent contingencies causing time delay (e.g. economic) have been considered.
Since the AFT prediction model is a parametric model, a distribution function for the dependent variable must be defined; as stated in specific literature this selection cannot be *a priori* accomplished. When comparing distributions from alternative families, an individual chi-squared test statistic should be calculated for each distribution:

$$\chi^2 = -2[\mathrm{LL}(0) - \mathrm{LL}(\beta_c)] \quad (3)$$

where LL(0) represents the restricted log-likelihood function and $\mathrm{LL}(\beta_C)$ represents the log-likelihood function at convergence. Existing literature state as most appropriate measure of statistical fit, for hazard-based duration models and so for AFT models, the $\chi^2$. Using this statistics test, the Weibull, Exponential, Log-Logistic and Lognormal distributions have been tested. The best calibrated model has been identified on the basis of the goodness-of-fit measure (i.e. higher value of $\chi^2$), the intuitiveness of the signs of the calibration coefficient and the minimization of the collinearity; the Log-Logistic distribution seems the most suitable for the given input data.
The available dataset has been divided into two datasets; a training dataset for calibration coefficients calculation and a validation dataset. Several attempt of model validation demonstrate that even if the population of the calibration dataset decreases, the outcomes are satisfying anyway.
An important test of model reliability is the analysis of calibration outcomes that is to verify the sign of each calibration coefficient in order to check if the shown positive or negative duration dependence is realistic according to input data.
For the sample calibration attempts performed, calibration coefficients are consistent to what expected and allow a quite immediate interpretation. As a matter of fact all explanatory variables have positive duration dependence (i.e. positive coefficient sign); positive duration dependence (i.e. the growing of the variable contribute to increase duration) is obvious for variables concerning technical features and for variables related to contingencies causing time delay. However a further verification is to evaluate the size of the coefficients for the dummy variables; these values define the impact on project duration estimate.
To further evaluate the efficiency of the proposed model, forecasting accuracy should be investigated; at this purpose the Mean Absolute Percentage Error (MAPE), estimated as follows, is a proper indicator:

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}|PE_i| \quad (4)$$

where $PE_i = 100 \cdot (X_i - F_i)/ X_i$ is the percentage error for observation i depending on real duration $X_i$ and estimated duration $F_i$.
According to the dataset available, a satisfying forecasting accuracy has been reached since the MAPE measure is equal to 2,8413%. The MAPE value indicates that on average, the forecast underestimates or overestimates true values by less than 3%. This is the demonstration that the proposed model is reasonably accurate and it is really able to predict LTPs duration.
Figure 1 shows model validation findings through the comparison between real project duration and estimated duration of some projects of the validation dataset.

| | Duration | fitted |
|---|---|---|
| 1 | 1050.00 | 1050.79 |
| 2 | 980.00 | 1140.46 |
| 3 | 250.00 | 252.38 |
| 4 | 1432.00 | 1421.73 |
| 5 | 510.00 | 513.90 |
| 6 | 910.00 | 906.51 |
| 7 | 900.00 | 868.57 |
| 8 | 1550.00 | 1471.16 |
| 9 | 230.00 | 227.22 |
| 10 | 1332.00 | 1297.73 |
| 11 | 400.00 | 393.46 |
| 12 | 715.00 | 716.04 |
| 13 | 790.00 | 869.96 |
| 14 | 1080.00 | 1076.67 |
| 15 | 300.00 | 304.20 |
| 16 | 1592.00 | 1664.37 |
| 17 | 455.00 | 455.69 |

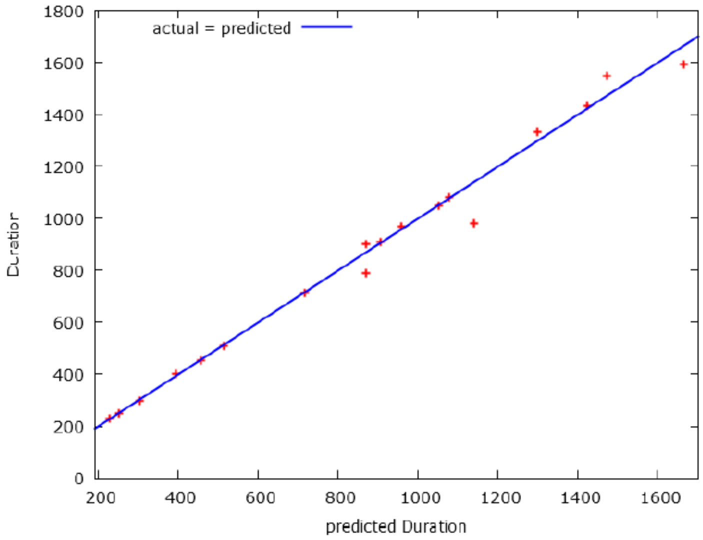



**Fig. 1.**
*Real duration vs Estimated duration*
*Source: Elaboration of the authors*

## 4. Further development

Model calibration and validation according to the dataset available, concerning railway projects, have shown the prospective of such a model to estimate realistic LTPs duration. The regression-like approach provided by AFT models (with Maximum Likelihood analysis) has shown encouraging accuracy and ease in outcomes interpretation so it can be stated that it is an appropriate modeling approach. As a matter of fact linear and non-linear regression approaches and Least Square Estimation models have been tested with the same input data but resulting outcomes are not so accurate. However in order to improve the quality of the econometric and statistical analysis, several sophisticated but complex approaches (e.g. the method of instrumental variables) are currently under investigation. This could lead to find the best model calibration possible for employment purpose. Any other development of this research pass through the availability of extensive datasets whose influence is critical for model calibration; as a matter of fact extensive data collation can improve output accuracy.
Moreover the approach introduced in this paper has been conceived for application to each large engineering project and so the widening of the analysis is currently ongoing.

## 5. Conclusion

The aim of this paper is to demonstrate that it is possible to approach project risk management in a statistic way. The Authors performed a literature review, than defined the regression-like model to perform the risk analysis regarding project duration linked to some selected project feature, and applied it on an real projects database, calibrated the tool and validate the tool.
The authors think that this approach, if fully investigated and calibrated, could give a big hand to project planners to design their project at the best, taking into account which features are affecting project duration and, in perspective, many others. In addition, when mature, this tool will allow to make simulation and, given a project, could advise the contracting authority to decide how to design the project in terms of contract feature and many others. The main hindrances to complete the research and have a usable tool is the availability of data: according to the Authors experience, no transportation agency or network manager have these data already collected and organized in a database: data are there, but a huge work for ordering them is still to be done.

## 6. References

Caron F. et al. 2013, *"A Bayesian approach to improve Estimate at Completion in Earned Value Management",* Project Management Journal.

Favari E. 2013, "*Large Transportation Projects Management – a non-linear approach"*, Politecnico di Milano, Doctoral dissertation.

Ford K. M. et al. 2012, *"Estimating Life Expectancies of Highway Assets"*, Report 713, National Cooperative Highway Research Program, TRB, USA.

Grassi A.F. 2013, "*Project Management for large transport infrastructures projects: an innovative tool to estimate construction site duration and a fitting approach in monitoring and controlling project execution performances"*, Politecnico di Milano, MSc Thesis.

Green W. H. (2012), *"Econometric Analysis – International Edition, 7th ed."*, Pearson.

Irfan M. et al. 2010, *"Planning-stage estimation of highway project duration on the basis of anticipated project cost, project type and contract type",* International Journal of Project Management.

Jacob D. 2003, "*Forecasting project schedule completion with earned value metric's"*. The Measurable News 2003(March):1. 7–9.

Kaka A. and Price A. 1991, *"Relationship between value and duration of construction projects"*, Construction Management and Economics, 9 1991, pp. 383–400

Kumaraswamy M.M. and Chan D.W.M. 1995, *"Determinants of construction duration"*, Construction Management and Economics, 13 (1995), pp. 209–217

Lipke W. et al. 2009, *"Prediction of project outcome: the application of statistical methods to earned value management and earned schedule performance indexes"*, International Journal of Project Management.

Nguyen L. et al. 2013, *"Simulating Construction Duration for Multistory Buildings with Controlling Activities"*, Journal of Construction Engineering Management.

Vandevoorde and Vanhoucke 2006, *"A comparison of different project duration forecasting methods using earned value metrics",* International Journal of Project Management.

Zheng, D.X.M. et al. 2004, *"Applying a genetic algorithm-based multi-objective approach for time–cost optimization"*, J. Constr. Manage. 130 (2), 168–176.

# MARITIME ACCESSIBILITY – KEY FACTOR IN SETTING TRANSPORT CORRIDORS. CASE STUDY: ROMANIAN PORTS TO BLACK SEA

**Sorin ILIE[1], Gabriela MITRAN[2], Viorel NICOLAE[3]**
*[1,2,3]University of Pitesti, Faculty of Mechanics and Technology, Automotive and Transport Department*

**Abstract:** The issue of globalization as a result of changes in the global economy which have increased in recent years, has created an environment in which the transport sector must constantly adapt. This process manifests itself mainly at the intercontinental level, involving the transport and distribution of goods and raw materials over long distances more efficiently in terms of duration and costs of transport and, at the same time, with reduced negative implications on the environment. The transportation mode for which a balance is achieved in relation to the limitations above mentioned is shipping, but it requires very good connections with continental modes, namely road transport and rail transport, which is why lately at european level a major attention is given to the development of transport corridors. This paperwork aims to study the influences which ensuring the accessibility of shipping has on the establishment of the road and rail transport corridors. The fundamentation of the conclusions according to which, for a proper functioning of the overall transport system, as a large technical system, it is necessary to provide connections and connectivities between the networks of the component subsystems, was realized in a frame of a case study that analyzes the accessibility of the Black Sea Port Constanta from Romania. It is highlighted that by its geographical position, Romania ensure the connection by sea between Central / Eastern Europe and Central Asia / The Far East, through the Port of Constanta, and in order to take transport flows at intercontinental level, are necessary land transport infrastructures, with high capacities, whose functionality in the intermodal transport network to be established so as to ensure the development of a sustainable transport. Thus, the in order to increase the accessibility of this port, on Romanian territory is mapped the route of the The Rhine - Danube Corridor (Strasbourg / Mannheim - Prague - Vienna / Bratislava - Budapest - Arad - Constanta) for road and rail transports. In conclusion, performing an intermodal transport, which allows the interconnection of different transport modes to realize a goods movement, with the specific advantages of each transport mode, like flexibility of road transport, high capacity of rail transport, low costs of shipping and high speed of air transport, is possible by ensuring the accessibility of each of these transport modes.

**Keywords**: maritime accessibility, freight transport, transport corridor, intermodal transport, goods movement, connectivity of transport networks.

## 1. Introduction

Of all time maritime transport represented a key factor that contributed to the economic development worldwide. Nowadays the globalization phenomena we face, is dependent on maritime transport system, a system that facilitated the unprecedented growth of economic and trade exchanges worldwide. At European Union level, 74% of the trading of goods with foreign countries is carried out by sea, either as monomodal system or as part of a multimodal transport system *(European Commission, 2013)*.

According to the document *"Ports 2030 - Gateways for the Trans European Transport Network"* prepared by European Commission, the quantity of merchandise manipulated in EU ports in 2011 was 3.7 billion tons, and the corresponding forecasts to a pessimistic evolution scenario show an increase with 50% of that amount until 2030. Satisfying this need for mobility of goods and, at the same time the enrollment in the requirements of a sustainable transport, necessitates the development of transport infrastructures on land, of large capacity, to ensure accessibility to shipping. In this respect, recently, within the European Union was planned the development of a multimodal transportation network (Trans European Transport Network, TEN-T). This has two main components, *Core Network* and *Comprehensive Network*, organized in the form of corridors. To configure *Core network*, in the first phase were identified the main transport hubs (for goods and passengers). Then, at drawing links between nodes represented by hubs, it was aimed the attraction of long distance traffic flows, in the conditions of increasing the effectiveness and efficiency of transport activity, supporting territorial cohesion, contribution to reducing emissions of greenhouse gases and air pollution, as well as sustainable land use *(European Commission, 2014)*.

On the Romanian territory have been established routes of two corridors: *The Orient/East - Med Corridor* (connects the maritime interfaces of the North, Baltic, Black and Mediterranean Seas) and *The Rhine - Danube Corridor* (connects the central regions around Strasbourg and Frankfurt via Southern Germany to Vienna, Bratislava, Budapest and finally the Black Sea). According to the European Commission *(European Commission, GIS dynamic maps, 2014)*, the road route of corridor *Rhine - Danube* connects the following cities: Arad - Timisoara - Sibiu - Pitesti - Bucharest - Constanta, but the current government of Romania supports as priority the next route: Arad - Timisoara - Sibiu - Brasov - Bucharest - Constanta *(Guvernul Romaniei, 2014; Guvernul Romaniei, 2013)*.

In this paper the authors propose a study that highlights the accessibility of Constanta Port (hub in the TEN-T Core Network) in the hypotheses of the two alternative routes mentioned, depending on accessibility indexes based on the travel distance, on travel time and on total cost of travel *(Ortuzar and Willumsen, 2011; Popa, 2009)*.

[1] Corresponding author:sorin.ilie@upit.ro; sorin_ilie_dis@yahoo.com

## 2. Methodology

*Accessibility*, as component element in a closed loop system consisting of **land use**, through hosted *socio-economic activities*, and **transport**, represents the measure of traffic deployment within this complex system. Changes at the levels of transport or land use systems influence the accessibility, both directly and indirectly (Figure 1).

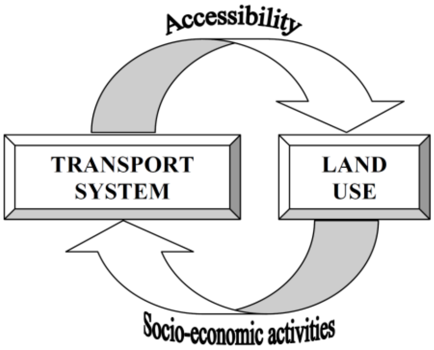



**Fig. 1.**
*The relationships between accessibility and land-use / transport systems*
*Source: adapted from Raicu, 2007*

Starting with accessibility studies introduced by Hansen in 1959, over time there have been intensive concerns regarding this concept, through which are highlighted methodologies to accessibility quantification and available tools, as well as its influence in the transport planning and land use processes *(Geurtz, 2006; Salze, 2011)*. Taking into account the complexity given by dynamics in time and space of accessibility, and the behavioral influence of individuals on this, the methodologies for quantifying accessibility include mathematical models with a high degree of abstraction imposed by the need to formalize complex processes of social, economic and behavioral natures.
In the specialized literature, the indicators through which is assessed the accessibility are related, on the one hand to the transport supply performance experienced by the supplier, and on the other hand to the performance experienced by the user. In practice, these indicators take into account the travel costs, the constraints that arise in carrying out the travel and the utility of available transport modes *(Koopmans et al, 2013; Hull et al., 2012; Popa, 2009)*.
The costs of travel can be used in evaluating the accessibility as argument of function of impedance that occurs in performing the trip between an origin and a destination point. The accessibility of a point of interest is inversely proportional with the impedance function specific to the selected route.
The mathematical expression of the accessibility of a point of interest *j* relative to all origin points *i* in a given space *S* has the following form:

$$A_j = \sum_{i \in S} \frac{1}{f(x_{ij})} \tag{1}$$

where:

- $f(x_{ij})$ *is the function of impedance (the function of difficulties in carrying out the travel),*
- $x_{ij}$ *is the variable of impedance function,* which can be:
  - *travel distance (* $d_{ij}$ *);*
  - *travel time (* $t_{ij}$ *);*
  - *generalized cost of travel (* $c_{ij}$ *), which, depending on the analysis purposes, may have different components (value of time, monetary cost, tolls, comfort, convenience, safety).*

The values of accessibility indicators play a major role in the planning process of transport networks at all levels of analysis (local, national, regional) and in all stages:

- analysis of the current situation and defining the dysfunctions;
- analysis of produced effects by the studied alternatives (multi-criteria analysis);
- verifying the achieving of proposed objectives;
- identifying negative consequences.

Also, the usefulness of assessing accessibility occurs in post-implementation analyzes, with the aim of comparing the obtained results after implementation with the planned ones.

## 3. Case study

### 3.1. Background

In the frame of this paperwork is studied the accessibility of Constanta Port (harbor on Black Sea in Romania) through the national network of public roads in Romania, in the context of TEN-T network planning on the territory of this country. Due to its geographical position, Romania ensures the connection by sea between Central / Eastern Europe and Central Asia / The Far East, through The Port of Constanta. The connection with Central and Western Europe is made through the border point Nădlac (Arad). In 2012 in the Port of Constanta were transported (received and shipped) 50 million tons of merchandise *(NC Maritime Ports Administration SA Constanta, 2013)*, ranking this harbor in the top 20 ports in the EU (Figure 2).

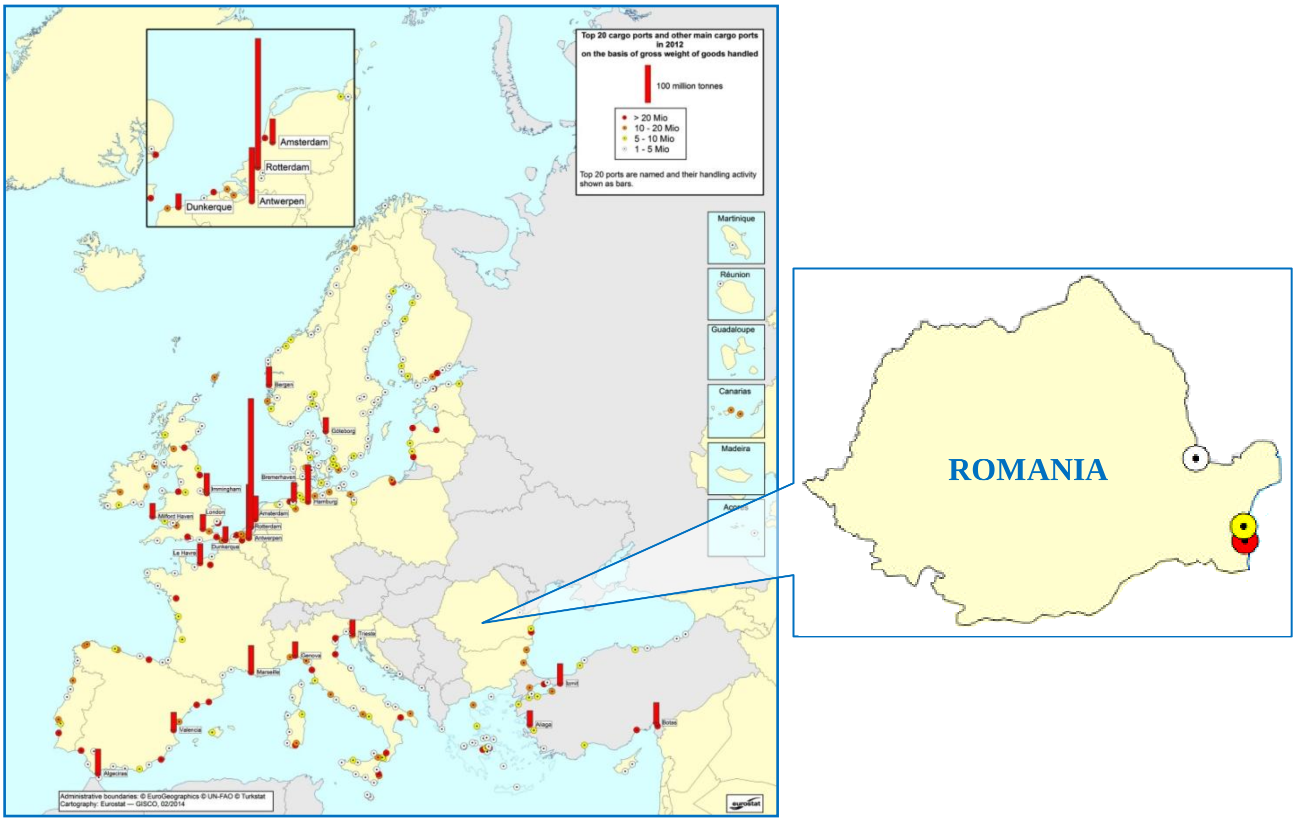



**Fig. 2.**
*Main european cargo ports in 2012 by gross weight of goods handled: Europe and Romania*
*Source: Eurostat, 2014*

Within this application are taken into consideration two hypotheses in which could be found the route of high capacity road corridor to serve the transportation pole Port of Constanta (Figure 3).

(i). *Arad* - *Timisoara* - ***Sibiu*** - ***Pitesti*** -***Bucharest*** - *Constanta (1st scenario);*
(ii). *Arad* - *Timisoara* - ***Sibiu*** - ***Brasov*** - ***Bucharest*** - *Constanta (2nd scenario).*

### 3.2. Transport model

The accessibility of Constanta Port, as destination point from Central and Western Europe, has been studied based on mathematical model presented in Chapter 2. The argument of impedance function was:

- Hypothesis 1: travel distance;
- Hypothesis 2: travel time;
- Hypothesis 3: generalized cost of travel.

In the expression of generalized cost were considered the value of time spent in traveling and the infrastructure use fee (one of the route variants comprises a section which is in awarding process to be made in concession).
For estimating the value of travel time specific to traffic relations on long distance, that have as origin the Western border of Romania through crossing point Nădlac (Arad) and as destination the Port of Constanta, has been developed a transport model *(Ortuzar and Willumsen, 2011)*. The application was developed using *Visum* software, specialized in modeling of traffic flows, offering results that will be input data to transport planning.

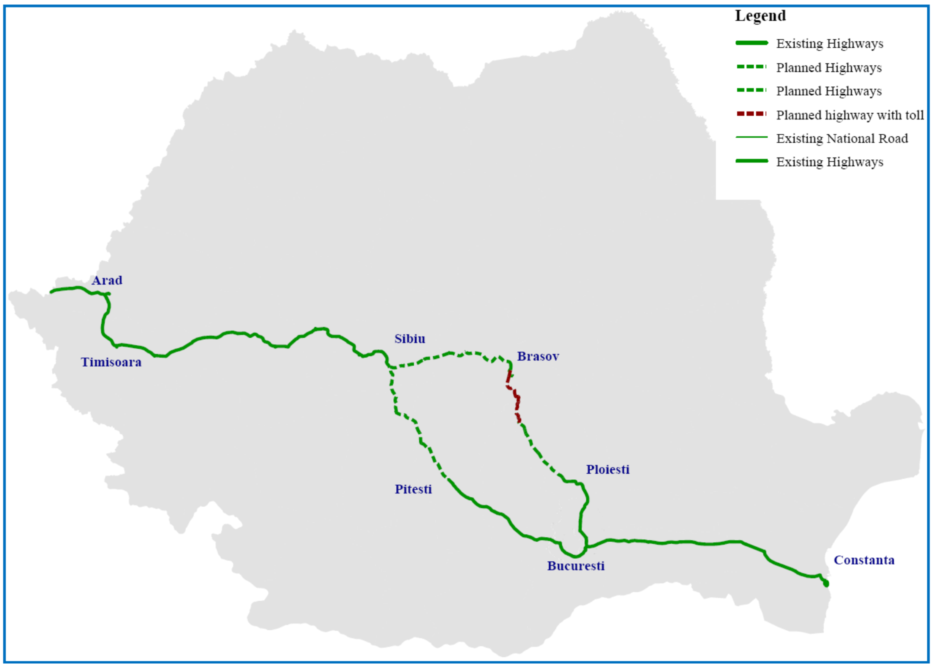



**Fig. 3.**
*The studied TEN-T corridor routes*

In this model, the transport network was formalized as a graph with arcs and nodes in which were introduced following parameters: length, free flow speed, number of lanes, transport capacity, vehicles categories which can use that road segment (cars, light duty vehicles, heavy duty vehicles), etc. Following traffic modeling, considering in network graph the two route variants, were obtained the data presented in Table 1.

**Table 1**
*Travel time for freight vehicles*

| Scenario | Highway length $d_{ij}$ [km] | Travel time $t_{ij}$ [hours] |
|---|---|---|
| Scenario 1 | 837.7 | 11.03 |
| Scenario 2 | 931.2 | 12.66 |

### 3.3. Calculating accessibility

#### 3.3.1. Hypothesis 1

In this assumption, is analyzed the accessibility of the transport node in terms of **distances** from the origin points of travels $\left(x_{ij} = d_{ij}\right)$. Relation (1) became:

$$A_j = \sum_{i \in S} \frac{1}{f(x_{ij})} = \sum_{i \in S} \frac{1}{f(d_{ij})} = \frac{1}{d_{ij}^2} \tag{2}$$

a) Scenario 1: $A_{j1} = \dfrac{1}{837.7^2} = 1.425 \cdot 10^{-6}$

b) Scenario 2: $A_{j2} = \dfrac{1}{931.2^2} = 1.153 \cdot 10^{-6}$

Taking into consideration the distance traveled in the network between the two points of origin and destination, the accessibility provided within scenario 1 is with 24% higher than that offered by the route of scenario 2.

#### 3.3.2. Hypothesis 2

Within this assumption, is studied the accessibility of the transport node regarding **travel time** between the points of origin and destination of travel. Substituting $x_{ij} = t_{ij}$ in relation (1), is obtained:

$$A_j = \sum_{i \in S} \frac{1}{f(x_{ij})} = \sum_{i \in S} \frac{1}{f(t_{ij})} = \frac{1}{t_{ij}^2} \quad (3)$$

a) Scenario 1: $A_{j1} = \frac{1}{11.03^2} = 8.219 \cdot 10^{-3}$

b) Scenario 2: $A_{j2} = \frac{1}{12.66^2} = 6.239 \cdot 10^{-3}$

In terms of total duration of travel, within scenario 1 the accessibility is with 32% greater than in scenario 2.

**3.3.3. Hypothesis 3**

In the frame of this hypothesis is computed the transport node accessibility taking into consideration the generalized cost of travel $x_{ij} = c_{ij}$. The components of travel generalized cost which were taken into account are: *the value of time spent in traveling* and *the access fees on toll roads*. The relation (1) particularized for this case is:

$$A_j = \sum_{i \in S} \frac{1}{f(x_{ij})} = \sum_{i \in S} \frac{1}{f(c_{ij})} \quad (4)$$

The travel difficulties function has the form:

$$f(c_{ij}) = c_{ij}^2 \quad (5)$$

in which:

$$c_{ij} = VOT \cdot ALV \cdot t_{ij} + TOLL \ [EUR] \quad (6)$$

where:

- *$VOT$ is the value of time, in $[EUR/ton/hour]$;*
- *$ALV$ is the value of average load per vehicle, in $[ton]$*
- *$t_{ij}$ is the travel time from origin "i" to destination "j", in $[hour]$;*
- *$TOLL$ is the value of road toll, in $[EUR]$;*

The values considered for these parameters are:

- $VOT = 1.27 \ EUR/ton/hour$, according to *Transport Master Plan for Romania*, published in 2014 *(Ministerul Transporturilor, 2014)*;
- $ALV = 14.7 \ tons$, according to data published by *Eurostat* for Romania in the document *"Road freight transport by journey characteristics"* for year 2012 *(European Commission, Eurostat, 2014)*;
- $TOLL = 4.3 \ EUR$ for heavy duty vehicles on second route (Arad - Timisoara - Sibiu - Brasov - Bucharest - Constanta). On this route, the section Comarnic - Brasov (60 km) is in awarding procedure in order to be realized in concession. Studies published by the Ministry of Transport indicate the value of 4.3 euros as access fee for heavy freight vehicles.

The accessibility indexes for point of interest Port of Constanta for a heavy freight vehicle that has as origin the Romanian border Nădlac, in the two considered scenarios related to the routes of the highway corridor, are:

a) Scenario 1: $A_{j1} = \frac{1}{(1.27 \cdot 14.7 \cdot 11.03 + 0)^2} = 2.358 \cdot 10^{-5}$

b) Scenario 2: $A_{j2} = \frac{1}{(1.27 \cdot 14.7 \cdot 12.66 + 4.3)^2} = 1.726 \cdot 10^{-5}$

Analyzing the data obtained, it is found that the accessibility in scenario 1 is with 36 % higher than the one offered by scenario 2.

## 4. Conclusions

In this paper the authors briefly presented a case study in which has been analyzed comparatively the on road accessibility of the transport node Constanta Port, provided by the alternative routes of Rhine - Danube Corridor on Romanian territory (*scenario 1:* Arad - Timisoara - Sibiu - Pitesti - Bucharest - Constanta, and *scenario 2:* Arad - Timisoara - Sibiu - Brasov - Bucharest - Constanta), having as objective taking over long distance traffic flows (transit over Romania). In order to express the accessibility, have been considered the accessibility indicators related to impedance function, in cases in which its argument is *distance*, *duration*, and respective *generalized cost of travel*. In all the three studied cases, the accessibility provided by first route is greater than that offered by second one.

Analyzing the values of these differences, it is observed that as the argument of impedance function is more complex, taking into account not only the spatial distribution of the points of origin and destination, but also the characteristics of the transport network (travel time) and socio-economic characteristics of users (value of time), the gap between the accessibilities of the two routes is greater.

*(i).* When the argument of impedance function is *travel distance:* $A_{j1}$ is with **24 %** higher than $A_{j2}$;

*(ii).* When the argument of impedance function is *travel time:* $A_{j1}$ is with **32 %** higher than $A_{j2}$;

*(iii).* When the argument of impedance function is *generalized cost of travel:* $A_{j1}$ is with **36 %** higher than $A_{j2}$;

In conclusion, the accessibility of points of interest in the frame of transport networks, expressed through accessibility indicators, is a complex parameter, which harmonizes information on transport networks, land use, travel behavior of users, regional development policies (social equity, environment protection, etc.), providing valuable information in the planning process of transport networks.

## Acknowledgements

This work was funded through the project *PERFORM - Sustainable performance in doctoral and postdoctoral research*, ID: *POSDRU/159/1.5/S/138963*, co-financed by the European Social Fund – Investing in People, within the Sectoral Operational Programme Human Resources Development 2007-2013.

**References**

European Commision 2013. Ports 2030. Gateways for the Trans European Transport Network. Available from Internet: < http://ec.europa.eu/transport/modes/maritime/ports_en.htm>.

European Commision 2014. GIS dinamic maps. Available from Internet: <http://ec.europa.eu/transport/infrastructure/tentec/tentec-portal/main.jsp>.

European Commision 2014.The planning methodology for the trans-European transport network (TEN-T). Available from Internet: < http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52013SC0542&from=EN>.

European Commision. Eurostat 2014. Available from Internet: <http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/File:Main_european_cargo_ports_in_2012_by_gross_weight_of_goods_handled2.png>.

Geurtz, K. 2006. *Accessibility, land use and transport.* Eburon, Delft, 245 p.

Guvernul Romaniei 2013. Available from Internet: <http://www.autostrada-transilvania.ro/img/iurmu_autostrazi_romania_r2_74486600.pdf>.

Guvernul Romaniei 2014. Hotărârea nr. 24/2014 privind instituirea Programului naţional "ROMÂNIA 100". Available from Internet: <http://lege5.ro/Gratuit/gm4dqmrxgy/hotararea-nr-24-2014-privind-instituirea-programului-national-romania-100->.

Hull, A; Silva, C.; Bertolini, L. 2012. *COST Action TU1002 - Accessibility Instruments for Planning Practice*. ISBN13: 978-989-20-3210-8, 370 p.

Koopmans, C.; Groot, W. ; Warffemius, P. ; Annema, J.A. ; Hoogendoorn-Lanser, S. 2013. Measuring generalised transport costs as an indicator of accessibility changes over time, *Transport Policy*, Elsevier, Volume 29: 154–159.

Ministerul Transporturilor. 2014. Masterplan general de transport pentru Romania. Volumul 2, Partea C: Ghid pentru evaluarea Analizei Cost-Beneficiu Economice si Financiare si pentru Analiza de Risc. Available from Internet: <http://www.ampost.ro/fisiere/pagini_fisiere/Ghid_ACB.pdf>.

NC Maritime Ports Administration SA Constanta 2013. Annual report 2012. Available from Internet: < http://www.portofConstanta.com/apmc/portal/static.do?package_id=st_rap_anual&x=load>.

Ortuzar, J.D.; Willumsen, L. 2011. *Modelling Transport*. 4$^{th}$ Edition, John Wiley and Sons, Ltd., Publication, ISBN: 978-0-470-76039-0, 608 p.

Popa, M. 2009. *Economia transporturilor.* Politehnica Press Publisher, Bucuresti, Romania, 325 p.

Raicu, S. 2007. *Sisteme de transport*. AGIR Publishing House, Bucharest, Romania, 484 p.

Salze, P.; Banos, A.; Oppert, J.M.; Charreire, H.; Casey, R.; Simon, C.; Chaix, B.; Badariotti, D.; Weber, C. 2011. Estimating spatial accessibility to facilities on the regional scale: an extended commuting-based interaction potential model, *International Journal of Health Geographics,* doi: 10.1186/1476-072X-10-2.

# TRAMWAY AND LRT IN ITALY: THE CASE OF FLORENCE AND GENERAL ISSUES

**Giovanni Mantovani**[1]
[1] *Independent Professional, Rome, Italy*

**Abstract:** The introduction recalls some aspects of the remarkable Tramway Renaissance in the last three decades. It took place also in Italy, although not so widely as in France. After a general picture of the Italian pre-existent and new networks, the case of Florence is deeply examined, as a most significant one. In this City the old tramway network was closed in 1958, but a wide light rail network has been conceived about 15 years ago and it has been preferred to an underground Metro. The first line is in operation successfully since 2010, two more lines are under construction and extensions are planned. This network is described, giving some details on the experience of the first line and discussing the question of the second line passing through the historic centre. The paper ends remembering the reasons to prefer the tram, when suitable to the required line capacity, and some conditions to have good results.

**Keywords**: Tramway, Light Rail, Italy, Florence.

## 1. Introduction

The last three decades saw the so-called Renaissance of Tramway in some European countries that had almost abandoned what had been the main system of urban public transportation since the end of 19th century. In these countries the decline was caused, from the 30’s to 50’s of the last century, by the concurrence of different factors: the street space demanded the private motorization, the obsolescence of tracks and vehicles, the development of metro systems in greater cities, the technical progress of buses and trolleybuses, and some more. But in the 70’s and 80’s the growing traffic congestions and other adverse effects of the widespread motorization brought to the awareness of urban mobility problems and strategic need of an effective mass transit. In this frame the tramway, but a new tramway that often we call light rail transit, has found a new important role.

What are the main features of the new tramway? As regards vehicles:
- High capacity (not only to reach high line capacities, but also to cut operation costs, by means of a lower incidence of the driver cost per passenger)
- Modular construction (that means reduced costs, improved quality and maintainability, easiness of revamping)
- Low floor (essential for disabled passengers, comfortable for all people, advantageous for operation thanks to the reduction of stop time)
- Modern traction equipment (three-phase asynchronous motors and power electronics, that together give better performances, energy saving and recovery, limited volume, simpler maintenance)
- Use of diagnostic, communication and information systems (aimed to manage promptly and efficiently the rolling stock, to dynamically optimize the line operation and to assist the passengers)
- Optional running without overhead power line (by means of ground power supply, or energy accumulation onboard; to be used where the overhead line has an unacceptable visual impact, as in historic areas or near monuments, or interferes with rows of trees).

As regards lines:
- Innovative track types (aimed to an easier and cheaper construction, better geometric quality, easier maintenance, reduction of noise and vibration generated by tram circulation)
- Reserved lanes (protected, if necessary, and advisably without any other circulation, even public)
- Advanced traffic light priority systems (able to give always a very high probability of no stop, without significant penalization to the conflicting traffic)
- Centralized monitoring and control systems (based on communication among the control centre, the onboard equipment and the devices installed along the line).

High capacity, exclusive right of way, efficient priority at crossings, centralized control and some other features bring to a Light Rail Transit (called *Metrotramvia* in Italian), which is a convenient alternative to Metro or Light Metro when line capacities not higher than about 8,000 p/h are required.

It must be pointed out that the modern tramway can be flexibly adapted to the territory. The best performances at reasonable construction costs are obtained when the alignment runs on wide streets, where a clear separation from other vehicles and pedestrians is easily feasible and allows a good commercial speed and regularity, but short tunnel or viaduct sections can be realized to overcome areas unable to host the track, with a relatively small increase of the entire line construction cost. And it is also possible to go through pedestrianized zones: the passing of trams at intervals of some minutes along a well defined path is safer for pedestrians than that of other vehicles, with random paths and times.

[1] Corresponding author: giomant12@gmail.com

Thus, also a pedestrianized zone (as a historic centre) can be served by trams, just fixing a proper speed limit on that short section, with small effects on the commercial speed of the complete line. A first good example of flexibility was in 1994 the line “A” of Strasbourg: 9.7 km, of which 1.2 km in tunnel and about 0.5 km in quasi-pedestrianized streets in City centre.

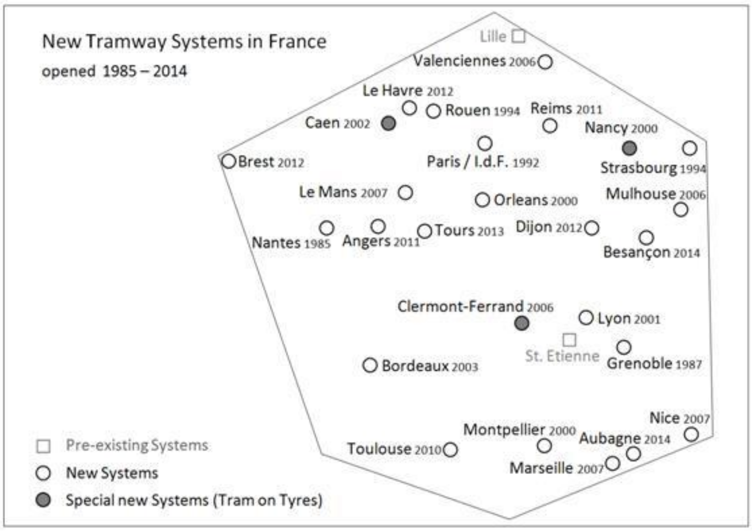



**Fig. 1.**
*New Tramway Systems in France, 1985-2014*
*(from official data)*

A good indicator of the Renaissance is the number of cities in which new tramways or LRT’s have been built from 1971 to 2014: about 120, plus some others under construction. The majority is in Europe (65) and France stands out, with 26 new systems (about 660 km, 64 routes in totality) opened from 1985 to 2014 (fig. 1). These figures are approximate, because it is often difficult to classify a system and to evaluate if it must be considered as a new one.
In order to have a general view of the tramway presence in the world, one must consider the other 240 systems never discontinued (mostly located in central and eastern Europe) and in many cases modernized and extended. An important example is that of some German cities where the tram circulation on the streets of central areas was transferred in new underground sections, connected with the maintained surface sections (a transition from *Strassenbahn* to *Stadtbahn*).

**2. Tramways and Light Rail Transit in Italy**

After the closure of the Cagliari network, in 1973, four urban tramway networks remained active and the Renaissance has been late. Only in 2003 trams began to run again in a City that had abandoned them, Messina. Other projects took a long time to be defined and realized, some of them cancelled. Today the new systems in operation are 7, each of them with only one route at the moment, and the total length of tracks is 58 km (fig. 2). A very quick review of the networks – pre-existent, new ones and cancelled projects – follows.

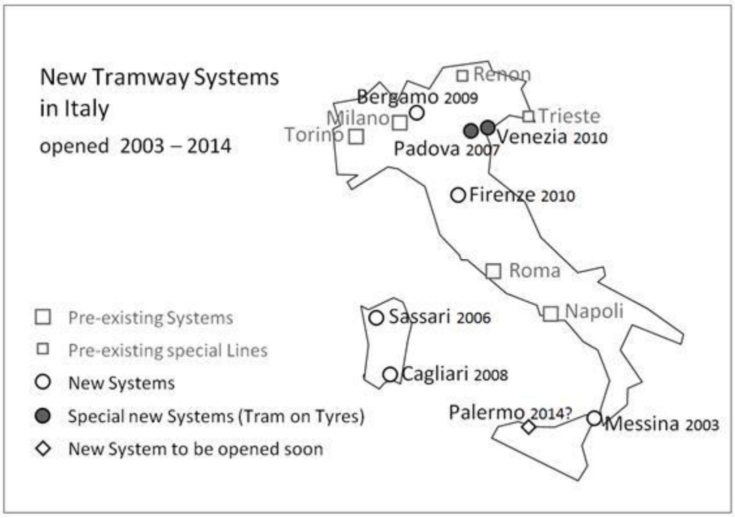



**Fig. 2.**
*New Tramway Systems in Italy, 2003-2014*
*(from official data)*

### 2.1. Pre-existent Systems

- **Torino**. At present 9 routes. A controversial LRT line (route 3) was realized in the 80's. An important step was, in the last decade, the modernisation and extension (with LRT characteristics) of tram line 4, now a 18 km line crossing the underground metro in City centre; in the meantime the construction and the activation of the metro line caused some reductions of the tram network.
- **Milano**. At present 16 urban and 1 interurban routes. Also in Milano the network has suffered some reduction due to the developments of the underground system (now 4 metro lines are in operation). Four tram lines with LRT characteristics have been built in the last two decades in the outskirts (routes 15, 7, 4 and 31).
- **Roma**. At present 6 routes. The network, once the largest in Italy, suffered progressive reductions and reached its minimum extension in 1980. In 1990 a new line was activated (route 2) and another one in 1998 (route 8) which, after 67 years, brought again trams at the border of historic centre. Both lines present some characteristics of LRT.
- **Napoli**. At present 3 routes. A short extension was recently opened, plans exist to reactivate a line closed for metro works and to realize various extensions in the outskirts.
- **Trieste**. After the closing of the urban network in 1970, only a special line survives. It connects the City centre with a hilly suburb and a section is operated with the aid of a funicular.
- **Renon**. A historic line connecting some mountain villages near Bolzano.

### 2.2. New Systems

- **Messina**. One line, completed in 2003. It crosses the City, with a length of 7.7 km. The rolling stock consists of 15 bidirectional *Cityway* trams supplied by Alstom Ferroviaria, 22.5 m long.
- **Sassari**. One narrow gauge, single track line, completed in 2006 and subsequently extended along an existing regional railway line, reaching a total length of 4.3 km. Operated with 4 *Sirio* bidirectional trams supplied by AnsaldoBreda, 27 m long. Further extensions are planned, also in form of tram-train.
- **Padova**. One Tram-on-Tires line, Translohr type, completed in 2007. A standard tramway had been chosen and contracted, but after a change in the municipal administration the project has been cancelled and replaced with the Tram-on-Tires solution, presumably expecting simpler construction and flexibility in operation. The line is 10.3 km long and runs through the centre. The rolling stock consists of 16 bidirectional *TE3* vehicles, 25 m long, provided with batteries to overcome a section of 600 m without overhead line. Other lines are planned.
- **Cagliari**. One narrow gauge, single track line, 6.3 km long, obtained by converting the first section of a regional railway and opened in 2008. It is operated with 9 *06T* bidirectional trams (length: 29,5 m) supplied by Škoda; other trams will be supplied by CAF. The doubling of the track is in progress and a 1.8 km branch using a viaduct will be soon activated. Also a time-sharing tram-train service on another section of the railway will be soon activated. Plans exist to extend the network in the town.
- **Bergamo**. One suburban line, 12.5 km long, built on the alignment of a dismantled local railway. Operated with 14 *Sirio* bidirectional trams supplied by AnsaldoBreda, 32.1 m long. A similar line, on the alignment of another dismantled railway, is planned, together with a line running through the centre of town.
- **Firenze**. One line, 7.7 km long, operated with 17 bidirectional *Sirio* trams supplied by AnsaldoBreda, 32.1 m long. Details on the case of Florence are given below.
- **Venice**. One Tram-on-Tires line, Translohr type, opened in 2010 and extended in 2014. It is 8 km long, it runs in mainland, connecting Mestre with Marghera, and it is operated with 20 *TE4* bidirectional vehicles, 32 m long. Next year a branch will be activated from Mestre to Venice, passing on the bridge across the lagoon.
- **Palermo**. Three lines are in construction, connecting suburban districts to interchange points at the main railway stations of the City. The first of them, 5.5 km long, will be soon activated. The rolling stock consists of 20 bidirectional *Flexity Outlook* trams, supplied by Bombardier, 32 m long.

### 2.3. Cancelled projects, two cases

- **Bologna**. Three underground metro lines were originally planned at the end of 80's by the Municipality, but after a change in the local government a tram network was preferred and admitted to a State financing. The realization of this project was delayed for various reasons and another local government change caused the cancellation of tramway, substituted with an optically guided trolleybus (*Civis* of Irisbus) network. The following government came back to rails, designing a LRT line, mostly underground, on the southern and eastern axes of the City. Thus the guided bus system was limited to the western axis, but various problems arose about its realization that was stopped and also the light rail project was stopped for lack of money. The trolleybus project was modified (*Cristalis* instead of *Civis*; automatic guide only when approaching and leaving stops) and works are still in progress. Twenty-five years to have a (modern, certainly) trolleybus in only one part of the City is emblematic of how not enlightened politics and other factors can damage the development of public transport: not only wasted money, but also big delay of benefits, which is itself a social cost.

- **Verona**. A tramway network has been planned and financed at the end of 90’s and then works have been contracted, but an intricate series of circumstances cancelled the tramway and gave way to a hybrid (electric-diesel) guided trolleybus network. Works are now at the initial steps.

## 3. The case of Florence

### 3.1. The old tramway

The first Italian electric tramway, inaugurated in 1890, ran from Florence (Piazza San Marco) to Fiesole, utilizing the Sprague system (fig. 3). From 1898 onwards, with the conversion to electricity of other already existing lines and the construction of new lines, a wide tramway network progressively developed in the City (fig. 4). Its decline began in the Thirties and, accentuated by damage suffered during the War, continued even after the Municipalization of the service (1946), until its elimination (1958) (fig. 4) as buses and trolleybuses gradually prevailed. Both the decay of trams and tracks, and, surely, pressure exerted by petrol and bus lobbies played against the continuation of tram system, as did the widespread opinion that public transportation was doomed to a marginal role due to the strong advance of private motorization (even the trolleybuses went out of use in 1973).

**Fig. 3.**
*The first electric Tramway in Italy,*
*from Firenze to Fiesole, activated in 1890*
*(G.M.’s collection)*

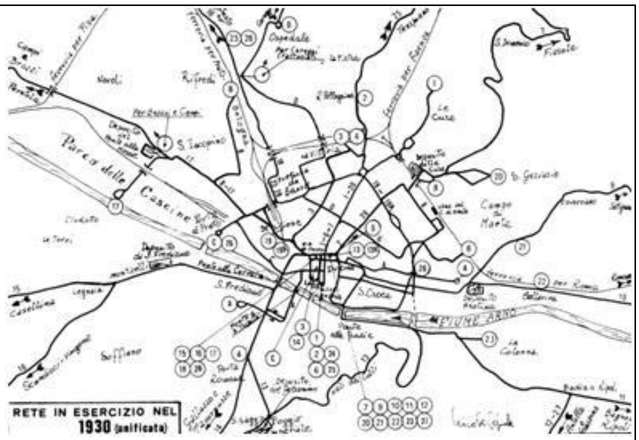



**Fig. 4.**
*The tramway network of Firenze in 1930*
*(Cefaratti, Centocinquant’anni di trasporto pubblico a Firenze, Calosci 2007)*

### 3.2. The plans for the new tramway

Also in Firenze the growing awareness of urban mobility problems led soon town planners to consider the possibility of realizing a modern rail system. As in other cities, two opposing solutions were examined: light metro (to be built underground) and tramway. In 1995, with the approval of a town plan revision, the solution based on a fast tram system (a light rail transit, we could say now), which could gain access to state funding (as could a metro system), prevailed. The realization of a tram system could also benefit from an agreement made with the Italian railway company which entailed economic contributions and engineering services as compensation for the troubles given to the City because of the construction of the high-speed railway line. A mobility plan was approved in 1999 defining the network, which would consist of three lines:

**1**, diametrical West to East, from Scandicci to Rovezzano, with a route which skirts the City centre running along part of the ring roads.

**2**, radial North-West to centre, from the Peretola airport to Piazza Beccaria, crossing the historic centre.

**3**, another diametrical, North to East, from the hospital district of Careggi to Viale Europa running along part of the ring roads (partly joining up with line 1).

Implementation procedure began at once on part of line 1, from Scandicci to the main railway station (Santa Maria Novella), with public financing from different sources.

### 3.3 First line description

The first line (fig. 5) is 7.69 km long, plus a branch of 0.57 km to the depot, which is near Scandicci terminus. Besides the two termini there are 12 stops, with a distance variable from 312 to 1023 m, according to the density of residences and other urban functions. The Firenze terminus is in front of a side entrance of the main railway station, allowing a quick interchange with trains. The Scandicci terminus is near the A1 motorway, where a large car and bus interchange park is planned.

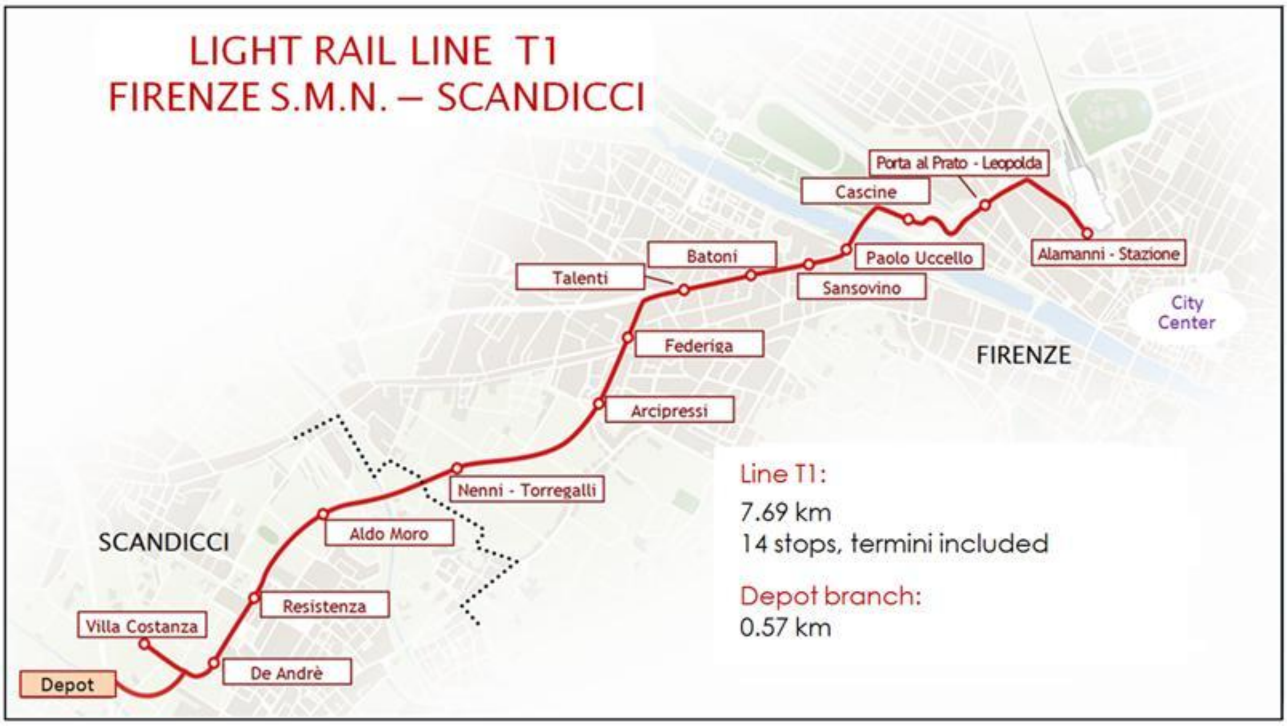



**Fig. 5.**
*Schematic map of the first line*

Along the line five important structures have been realized: a bridge across the Arno river (124 m, reserved to trams, pedestrian and bicycles), the widening of a pre-existing bridge across the Greve river and, to eliminate critical crossings, two subways and a viaduct. The line is basically level, there are significant slopes (up to 7%) only on the ramps connecting to subways and viaduct. The main characteristics of the line are given in Table 1.

**Track**
- Gauge: 1435 mm
- Rails: grooved, Ri60R13 profile (NP4aS in sharp bends)
- Sleepers: Bi-bloc, embedded in platforms
- Fastenings: Nabla

*Elastomeric mats are used in sections where high attenuation of vibration was required.*
*Different construction, as embedded rail, in some short sections*

**Overhead line**
- One 120 mm² wire, tensioned on most of the line
- Line voltage: 750 V d.c.

**Power system**
- Two primary power deliveries: 15 kV a.c., 12 MVA each
- Five power substations, 2x 1,5 M W

**Telecommunications**
- Proprietary fiber optics network
- Tetra radio system, with two base stations

**Central control room**
- Monitoring and control of operation (tram running, line devices, stop equipment)
- Remote control of power substations
- Surveillance of all stops by means of a CCTV system

**Depot/workshop**
- Total surface: 49,700 m² (in Vingone zone, Scandicci)
- Present capacity of the sheltered tram parking: 16
- Main buildings: power substation, rolling stock workshop, line equipment workshop, operator offices and control room

**Lane paving**
- All stops, sections in historic streets: porphyry setts
- Section in *Cascine* Park: grit bound by a special synthetic resin
- Avenues out of Florence downtown: grass (*sedum*)
- Crossings and some limited sections: asphalt

**Rolling stock**
- Number of trams: 17
- Supplier: AnsaldoBreda SpA
- Type: Sirio Firenze
- Architecture: bidirectional, 5 bodies on 3 boogies ($B'_o2B'_o$), entirely low floor (0.35 m on ToR)
- Length: 32,1 m
- Width: 2,4 m
- Tare weight: 39,8 t
- Wheel diameter: 660 mm (new)
- Power: 4x106 kW
- Traction equipment: IGBT inverter
- Maximum speed: 70 km/h
- Average acceleration: 1,0 m/s² (up to 40 km/h)
- Average service deceleration: 1,1 m/s²
- Average emergency deceleration: 2,4 m/s² (using electromagnetic shoes)
- Sitting places: 42 (plus 8 folding seats)
- Standing places: 160 (at 4 passengers/m²)
- Wheelchair places: 4
- Air conditioning: independent equipment for driver cabins and passengers

**Table 1.**
*Main characteristics of the first line*

It is interesting to note that no other vehicle is admitted on the tram lanes (except emergency vehicles and only on paved stretches) and that all crossings, vehicular and pedestrian, are provided with traffic lights controlled by a centralized tram priority system based on arrival forecast. Thus safety is ensured and also good commercial speed and regular headways.

**Fig. 6.**
*The new bridge across Arno river (downstream from "Ponte della Vittoria" (G.M.)*

**Fig. 7.**
*The depot in Scandicci Vingone (G.M.)*

### 3.4. Problems during first line construction

Diverse difficulties have been encountered during the works, which consequently lasted longer than expected (5 years instead of 3, in round figures). The project underwent numerous modifications, owing to new requirements of local administrations and intended operator, which caused partial stops of works and redesigning. Another typical cause of delay was the diversion of unexpected utilities. Archaeological surprises gave a strong contribution too: the remains of an Etruscan house, a river port quay of 18th century and walls of uncertain origin have been found and required investigations and preservation works. Another delay (and cost increase) came from the disposal of contaminated soil

found in digging. Limited stoppage of some works was also due to the request of shopkeepers to have no hindrance to customers access during periods of greater commercial activity.
An aspect to consider is the trouble given from works to everyday life of people resident or having an activity in the involved streets. All the possible was made to reduce this trouble, but there were continuous complaints, exploited for political reasons.

### 3.5. First line operation

The line, called T1, is at present operated every day from 04:30 (first run from Scandicci) to 00:30 (last run from Firenze), with headways which in the workdays range from 3'40" in peak hours to 7'50" in early morning and late evening. Thus a maximum line capacity of 3,300 passengers/hour*direction (at a standing passengers density of 4/m$^2$) is offered. The time from terminus to terminus is about 23 minutes, giving a commercial speed near to 20 km/h. A good result, fruit of exclusive lane and efficacious traffic light priority. Regularity and quick solution of circulation troubles are ensured by the centralized control system.
The line, inaugurated in February 2010, met a great success since the first days of operation. The present patronage reaches 44,000 passengers in working days, against a forecast of about 30,000. It is interesting to note that the load diagram during the day is quite flat, which is an important goal for public transport, seldom reached; this means that the T1 acquired many non-systematic users, users that are usually difficult to attract to public transport
A significant effect on the modal split was obtained, as confirmed by an investigation: 24% of passengers declared that they were using a car or a motorcycle before the tramway activation. Therefore, less pollution and less traffic congestion have been ensured.

**Fig. 8.**
*Passengers at Firenze terminus of line T1 (G.M.)*

**Fig. 9.**
*Inside view of a Sirio Firenze tram (G.M.)*

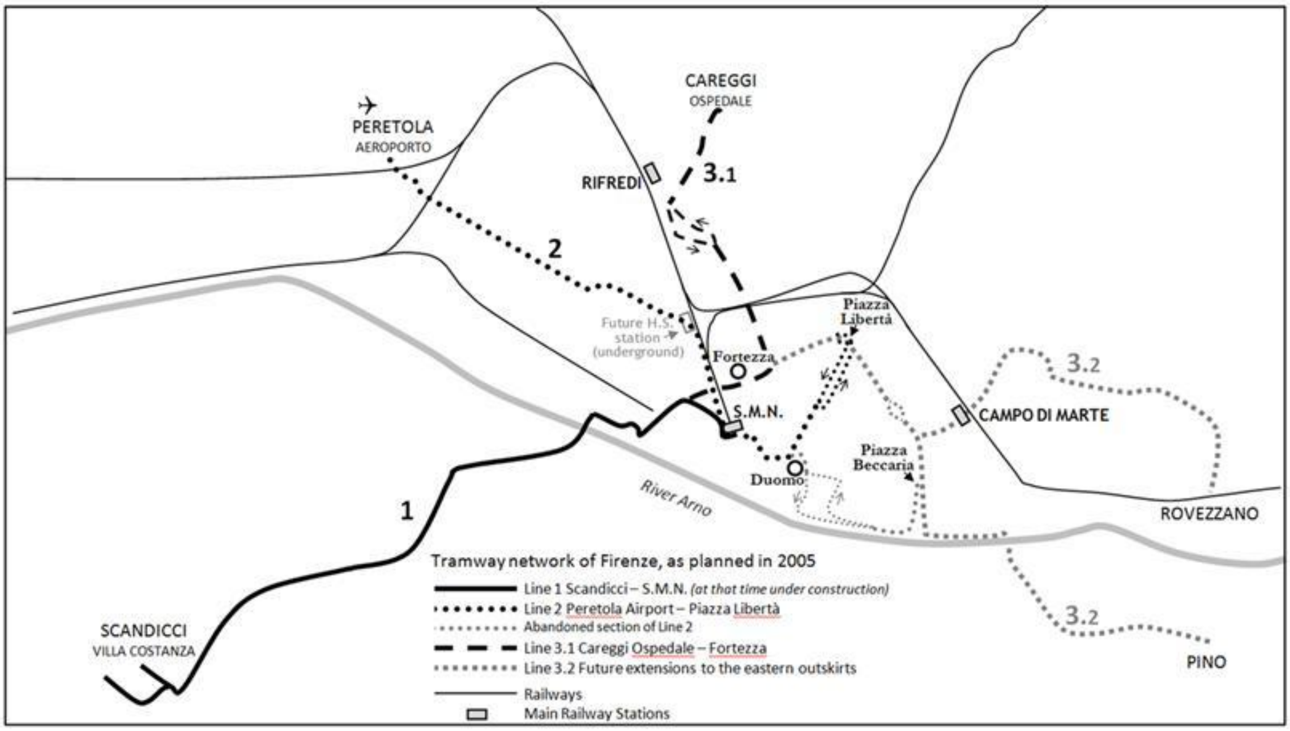



**Fig. 10.**
*Tramway network of Firenze, as planned in 2005*

### 3.6. The difficult gestation of the second and third line

Fig. 10 shows the Firenze network as it was planned in 2005, when a project finance contract was signed to build the extension of the network and to operate it entirely. We can note some differences compared to the 1999 plan, namely:

- Line 1 was limited to the main railway station (S.M.N.)
- Line 2 was diverted, in downtown, to Piazza Libertà, as it was established that the streets between Duomo and Piazza Beccaria were impassable for modern tramways.

Line 3 was divided in two construction phases: 3.1, from Careggi to Fortezza (to be connected with line 1) and 3.2, from Fortezza to Rovezzano and Pino, to be realized later. At that time line 3.2 was not financed and only today it is likely that money will be provided for a part of it. It should be considered that line 3.2 is just an administrative definition and does not identify an operation route.

The construction of lines 2 and 3.1 was started a few months ago, because many difficulties arose in the past years. Here we deal only with examples of major alignment problems that had to be solved and with public disapproval manifestations.

- The connection between line 1 and 3.1 has been modified in order to allow a joint service from Scandicci to Careggi without leaving unserved the main railway station (S.M.N.) and the border of City centre. Instead of connecting directly lines 1 and 3.1, the new tracks will be linked to the present terminus of line 1, passing on the other side of the station. Moreover, the terminus of line 1 (a stop, in future) in front of the main station has already been moved to the side of the station to ease passengers interchange with trains and to leave space to traffic in front of the station.
- Some solutions were examined to solve the intricate problem to serve the future High Speed station with line 2 and to insert its tracks in critical streets between the H.S. station site and the main station (S.M.N.). It was chosen the passing on one side of the H.S. station, with a stop near an access, and, by agreement with the national railway company, to continue the line inside a disused railway area, within the framework of the requalification project of that area. Trams will go out through a passage to be opened in a building which must be kept for historic reasons.
- The interference of trams (lines 2 and 3.1) with other traffic, considered critical in some junctions, has been solved by means of three subways (leaving trams in surface) and one viaduct (for trams). One of the subways is however still under discussion.
- The separation of the two tracks in different streets, foreseen for a section of line 3.1, appeared to be not functional; furthermore there were doubts about inserting tramway in one of those streets, subject to primary flow of traffic. Thus a new double track alignement was studied.
- The more complex problem regards the passing of line 2 in City centre. For a stretch of about 300 meter in the monumental area of the Duomo (not in front of it but just in its vicinity) it was foreseen to reserve the street to trams and pedestrians, using interlaced track to leave more space to pedestrians. Furthermore, in the same stretch, in order

to minimize the visual impact it was decided not to use the overhead line but to let trams run by means of traction batteries (it was 2007, today other solutions would be perhaps preferred). Unexpectedly in 2009 a full pedestrianization of this area was established, causing the cancellation of the entire section of line 2 from main station (S.M.N.) to Piazza Libertà through the City centre. This generated a great problem, because the central area of Firenze, inside the ring roads (the “Viali”), is very large and without tramway it should be served by local buses, with passengers discomfort due to changing from tram to bus and v.v.. The devised solution was to route the line through the ring roads to Piazza Libertà and there to turn on the right and enter in the central area for about 600 m, but only a small part of the centre will be thus served and, moreover, this solution is not coherent with a future continuation of line 2 towards Rovezzano or Pino. Afterwards an underground central section was also considered, on the model of German *Stadtbahnen*; the tunnel should be deep, owing to geotechnical conditions and density of historic buildings, and over 3 km long, due to possible locations for access ramps. This results in high costs, archeological risks, long construction times, reduced accessibility of trams. The problem seems to be still open.

An important issue is the public opinion response to tramway projects. The plan to realize lines 2 and 3.1 gave rise to the opposition of citizens groups, based on various kinds of worry: less space for cars, visual impact, obstacles to pedestrians, accidents, noise and vibration, troubles during the works and so on, in many cases just the fear of a changement. These worries were unfounded, but somebody used them to incite people for political reasons and no-tram committees were formed. The Municipality made a broad and good information campaign and in 2008 it held a referendum; the no-tram vote prevailed by a very narrow margin, but the quorum was not formed and however the referendum was only consultative. So the project went on. Even though the first line success and the cancellation of the line 2 central section reduced the contrariety, some manifestations continued up to now, backed by prejudices and opportunism.

## 4. Some short conclusions

Experiences in the world, especially in France, confirm that modern tramway (with full characteristics of LRT, when possible) can give a good answer to mobility demand, in medium-sized cities as primary public transport system and in greater cities as complement to metro. As a matter of fact:

- It appears functional and economically convenient compared to both buses and metro, for line capacities ranging roughly from 2,500 to 8,000 p/h, keeping in account also the construction costs.
- It can be built much faster than a metro and this allows to reach public benefits sooner.
- It uses electric energy and the power consumption per km∗p is lower than with buses or trolleybuses.
- It is more accessible than metro (especially deep metro) and keeps passengers in visual contact with the City.
- For not long journeys, up to about 5 km, the typical door-to-door time is lower than with metro.
- It is more comfortable and attractive for passengers than bus or trolleybus.

Often the fixed alignment is mentioned as a disadvantage, but it is irrelevant if lines are defined according to good (and respected) town plans. Also noise and vibration are questioned, but it is not the case of good trams on good track.

But the experiences, especially in Italy, give us also some lessons:

- Exhaustive feasibility studies must be carried out before deciding if and how to realize a line. A system vision and an interdisciplinary approach are needed; all aspects (technical, town planning, economic, social) must be considered. Particularly, tram lines must be defined in the framework of a renovation of the entire public transport network.
- High quality is required for the line design and all external conditions must be well known before designing, so as to avoid modifications during the works. Particularly, utilities diversion must be planned accurately and archeological risks must be investigated.
- Urban requalification is an important complement to the tram line.
- Politicians must not steal technicians job and technicians must not steal politicians job. A new town council must not change a project in progress, if correctly developed.
- Work sites must be carefully planned, in order to reduce troubles to residents and shopkeepers as much as possible.
- Great attention must be paid to have consensus of citizens and so-called stakeholders. A wide and wise concertation is necessary from the very beginning and efficacious information campaigns must be carried out in all phases.

**References**

Dambrosio M. 2008. La tranvia a Firenze nei piani e nei progetti (1967-2001), *Opere*, 19: 74-87

Matulli G. 2013. *La tranvia e la città*. Edizioni Polistampa. 78 p.

Mantovani G. 2014. Il ritorno del tram a Firenze, *La Tecnica Professionale,* 21(2): 32-41

# THE PLACE AND ROLE OF THE RESCUE COORDINATION CENTRE (RCC) IN SEARCH AND RESCUE OPERATIONS

**Svetislav Šoškić[1], Jovica Ćurčić[2], Ljubinko Karapetrović[3], Aca Randjelović[4], Midhad Krluć[5]**
[1,2,3,4,5]*University of Defence, Military Academy of The Republic of Serbia*

**Abstract:** The International Aeronautical and Maritime Search and Rescue Manual, hereafter the IAMSAR Manual, published by the IMO and the International Civil Aviation Organization (ICAO) is for maritime SAR based on the Hamburg Convention and for aviation on the Chicago Convention. The IAMSAR Manual contains practical guidelines for the organisation of maritime and aeronautical SAR, mission coordination, operations of search and rescue units (SRUs) and provision of SAR-related training.
RCC (rescue coordination center) - a unit responsible for promoting efficient organization of search and rescue operations within a search and rescue region. Rescue Coordination Centre (RCC) means a designated coordination centre that acts as the national Search and Rescue Region's coordination centre and the international point of contact for search and rescue (SAR) services and that takes care of the continuous maintenance of coordination and communication preparedness and coordinates SAR operations within its Search and Rescue Sub-Region (SRS).
Definition of tearms Search: An operation, normally coordinated by a RCC or RSC, using available personnel and facilities to locate persons in distress. (IAMSAR, Volume 1) Definition of terms Rescue: An operation to retrieve persons in distress, provide for their initial medical or other needs and deliver them to a place of safety.
Work should contribute to the establishment of thinking for RCC in Republic of Serbia due to the growing need of interventions in the prevention and elimination of consequences of natural disasters and better organization of search and rescue using the standards prescribed in the the world.

**Key words:** IAMSAR, SAR, RCC (rescue coordination center), natural disasters

## 1. Introduction

All the participants of river transport are obliged by regulations of different conventions and laws (e.g. the Danube Convention, the Budapest Convention, Agreement on the Sava River Basin, Law on Navigation and Ports on Inland Waterways, etc.) to come to aid to persons or a ship in distress.
Rescuing people, vessels, and material properties on water is connected with the search. It is normally conducted after successfully performed search. Search and rescue on water is a complex operation which involves many factors in order to be carried out effectively. It will not be a success without permanently ready service which can organize a rescue action if the need occurs. Many of our surrounding countries have search and rescue services on water at a national level (Romania, Croatia, Montenegro, etc.). Republic of Serbia has also organized rescue and search on water services, but the issue is at what level, whether those services are permanently engaged, equipped, etc.

## 2. Organising Search and Rescue in Republic of Serbia

According to the articles 73., and 74. of the Law on Navigation and Ports on Inland Waterways search and rescue of people and things in jeopardy on waterways is conducted by the ministry in charge of internal affairs in cooperation with port authorities (Sector for Water Navigation and Security of Water Navigation/Ministry of Transport). On minister of the internal affairs orders other state organs, business companies, legal entities, and entrepreneurs that perform their activities on internal navigable waterways (INW) are engaged with their available personnel and equipment.

Ministry of internal affairs (henceforth MIA RS) has defined units at a national level in the Sector for the State of Emergency that are intended for these activities of search and rescue on water. In addition to these, some other parts of the units subordinate to the MIA RS are engaged (police – divers, river police, fire fighting rescuing and helicopter units, etc.). Units of MIA RS are partially equipped with gear, movable equipment and vessels for efficient search and rescue on water.

In the Law on the States of Emergencies terms such as “search and rescue of people on the occasion of traffic accidents” and “protection and rescue from the floods and accidents on water and under water” are defined which involve conducting protection, search and rescue of people and material belongings on water and under water. This Law also defines the jurisdiction of state organs, which according to the article 11. of the Law on MIA RS is given the authority to organize reconnaissance, informing and timely warning and alarming at the territory of Serbia, as well as taking part during accidents in water transport. Ministry of Defence (henceforth MD) and Serbian Armed Forces (henceforth SAF) are given the jurisdiction, according to the article 12. of this Law, to employ certain organizational parts of MD and SAF in providing support and aid during search and rescue when other forces and means of the search and rescue system do not suffice, and on the orders of the Ministries – organizational units in charge of the states of emergencies. Within the Civil Protection Management/SES MIA RS (articles 98. – 102. of the Law on the State of Emergency) a unit for rescue on water is formed.

Based on the Law on the State of Emergency of the Republic of Serbia, Minister of the Internal Affairs has formed a Book of Regulations on the organization and way of employing the civil service specialized units. Based on this book of Regulations the unit for search and rescue is given the authority to:

- conduct affairs and tasks on rescuing lives and health of people and animals as well as material and cultural properties on rivers, lakes and flooded areas;
- transport people, animals and material belongings across rivers and lakes;
- rescue and find drowned people from rivers, lakes, wells, and other water areas;
- rescue material property threatened by water.

For the needs of help and support of the MIA RS in search and rescue on water , and based on the third mission of the SAF "supporting the civil authorities in facing security threats" Minister of Defence makes the decision to engage the necessary units of the SAF on that task. Units which can be engaged in search and rescue on water in the Republic of Serbia are mainly those which are daily in touch with the tasks on water (River Flotilla), but also other units of the SAF (aviation, special units, etc.). Through Operational Centre of the System of Defence of the R.S., and through the coordination of soldiers on duty, certain units of the R.S. for putting to practice tasks of search and rescue on water are alarmed.

In case of declaring the state of emergency in the R.S. in joint operational command of the GHQ SAF a team in charge of the tasks and obligations of the third mission of the SAF (search and rescue on water) is formed, which through the Operational Centre of the System of Defence coordinates with the other units of the SAF. Permanently engaged forces for search and rescue in the SAF, which could be used in the search and rescue on water, are only the aviation units (helicopters). The SAF are currently setting other forces to be employed in the third mission, which the SAF could organize as permanently or periodically engaged forces.

Organizing reconnaissance and intelligence on inland waters of the R.S. is partially done through River Information System (RIS), which represents a concept of coordinated telecommunications services and information systems in providing support to the inland navigation and its link to the other means of transport. It includes the systems for locating and tracking ships, showing electronic navigational maps, providing information on navigable waterways, as well as other services. The system is under the jurisdiction of the Navigable Waterway (Ministry of Transport). At the beginning of 2013, the whole navigable waterway of the Danube River was covered by this system. However, the question of coverage of the RIS of the Sava River, the Tisa River and the Danube-Tisa-Danube canal remains.

## 3. Organising Search and Rescue on water

### 3.1. Management of search and rescue on water service

After looking into system-normative documents (laws, rules, regulations, etc), functioning and current state of affairs (at a state level there is no developed national Plan for Search and Rescue on Water) a possible improvement of organization and establishing search and rescue service at a national level has been suggested.
It is necessary to form an office (department, section, group, etc) within one of the responsible ministries of the Serbian government.

The work of that office would be based on a national plan for search and rescue (its definition and application), because it would include several state services in a coordinated action of search and rescue, as well as the possibility of cooperating with neighbouring countries in order to establish a common service the main advantage of which would be:

- better supervision of a certain area,
- a greater possibility for the services to complement each other,
- decreased individual expenses and
- a possibility to expand the searching area to neighbouring countries.

The management of the Office would produce a Plan which would contain:

- the area of its jurisdiction,
- means, equipment and personnel for the service to function,
- make contracts with the beneficiaries of the means, other parties, search services in the neighbouring countries,
- divide the area of supervision within a certain number of sectors and establish a centre for coordination within each of them, and if necessary a rescue sub-centre and stations for alarming , and make a plan for the professional development of the staff.

### 3.2. Centre for coordination and search and rescue on water

It would be the main central unit of the search and rescue on water service. The manager of the centre would be responsible for its work, and the centre needs to have:

- a detailed Plan of search and rescue on water,
- proper facilities in order to function,
- means and equipment for search and rescue, especially for communication and
- trained personnel.

The work of the centre is based on the plan and contains regulations and data on:

- procedures,
- the work of certain departments within the search and rescue on water service,
- means of communications and the way they are used,
- staff that takes part in the search and rescue tasks and
- the information gathered.

The centre must possess reliable equipment for receiving aid calls, equipment for maintaining communication via land or radio equipment with the rescue units, rescue centres within sectors (sub-centres), main stations for RIS, centres for supervising air transport, meteorological devices, centres for coordination with the surrounding areas, ships, medical institutions, hydro-meteorological institution, etc.

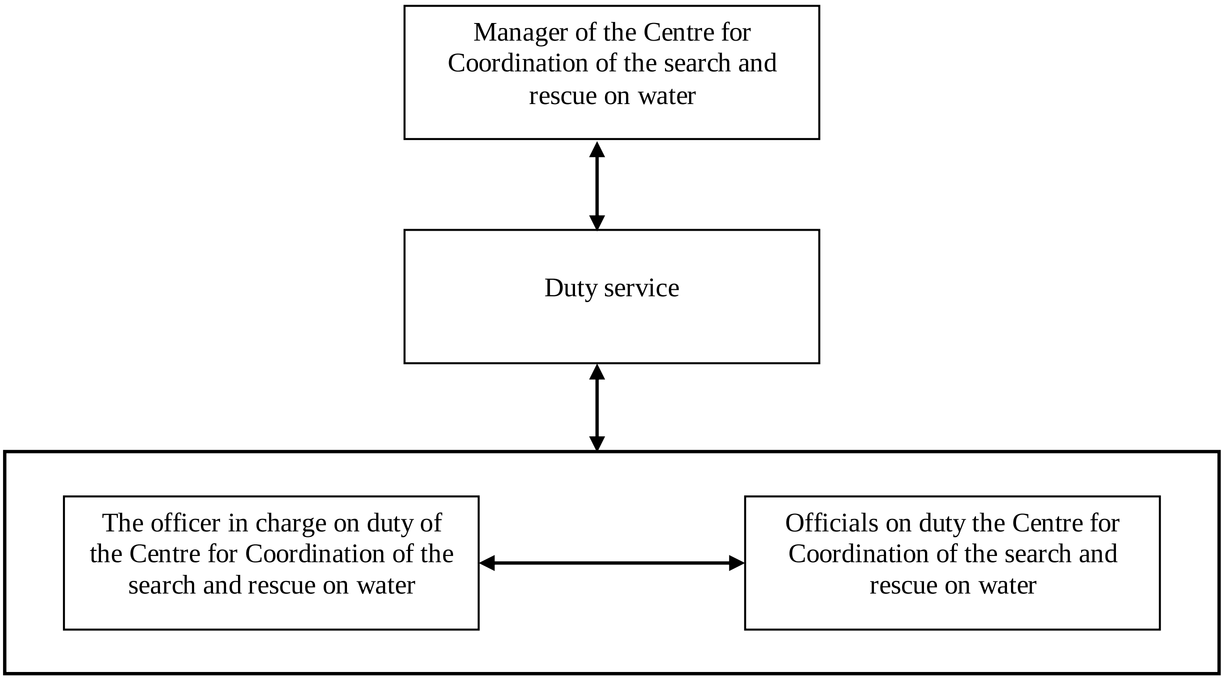



**Fig.1.**
*Preposition for overall organization of the Centre for coordination of search and rescue on water*

### 3.3. Centres within the sectors (sub-centres)

They are founded within the sectors of the navigable waterways, and are responsible for particular territory and a section of the navigable waterway. They are directly linked with the Centre for Coordination and the neighbouring centres within sectors. Provided with the necessary equipment and staff, they just carry out the tasks as instructed by the Centre for Coordination. Sub-centres would have the role to alarm because more than often they are able to be the first to notice a signal for help. They would have main stations of RIS included which would maintain the connection between persons in distress and the Centre for Coordination or a direct link with the search units in the field and enable the communication support during search and rescue on water. Sub-centres can be integrated within the port authorities.

### 3.4. Training the management and search and rescue on water units

The effectiveness of the search and rescue on water team is based on the management and the units of the service being well trained. The training programme is produced by the management of the service and it includes only the management, the centres within the sectors and rescue units. Within the rescue units, the focus of the training would be on gaining the practical experience with the means for search and rescue on water, providing first aid and coordination. The standards and criteria for every action would be formed, the procedure of search and rescue on water and evaluation papers would be defined, based on which the work of the rescue units would be valued. During the training, different means and helping devices should be used in order for the training to be more useful and of higher quality.

The training programme would be based on:

- basic programme (gaining knowledge and skills on procedures and equipment),
- training programme (in order to gain experience and coordination) and
- training (applying the gained knowledge, evaluating the units and applying the lessons learned on previous training).

Search and rescue drills would be planned from more simple to more complex ones (one unit and the coordination of several units, with a scenario made in advance). Both the management and the sectors within the sectors would be required to take part in those drills. Every project of the drill and reaction to the scenario could be of use to the management of the drill and the commanders of the units in making decisions in real life situations of search and rescue on water.

The means and equipment for performing a search and rescue task include buildings, offices, information – telecommunication means, vessels, aircraft, search equipment and rescue equipment. The search and rescue service can have only a part of this equipment but it is obliged to make special contracts with other organizations (state, business, or legal) so that, if the need occurs, they could take part in the search and rescue task immediately and with all the necessary equipment and personnel.
It is necessary for all the participants of the search and rescue task to be compatible (well-acquainted with the procedures, the way of using telecommunication means and providing first aid).

### 3.5. Rescue units

Within a search unit we could include every ship, boat or aircraft which takes part in search and rescue on water.
Special search units are specialized units of the search and rescue on water service which are directly responsible to the management of the search and rescue on water service, formed and equipped for search and rescue on water tasks.
Special units are well provided with food supplies, medical supplies and other necessities, and have well trained staff, for acting in all sorts of situations for which they are trained and intended. The management of the search and rescue on water service has to evaluate which units must be on full or partial alert, and which units are formed upon demand. Members of special units must be physically fit, of good health, well acquainted with the area of their action, trained for dealing with available equipment and providing first aid.
Special units can be formed within the search and rescue service and these may include the river rescue, the diving rescue and the aircraft rescue unit.
Special – specialized units can be deployed along the whole navigable waterway within a sub-centre. These units would need to have a defined area of responsibility.

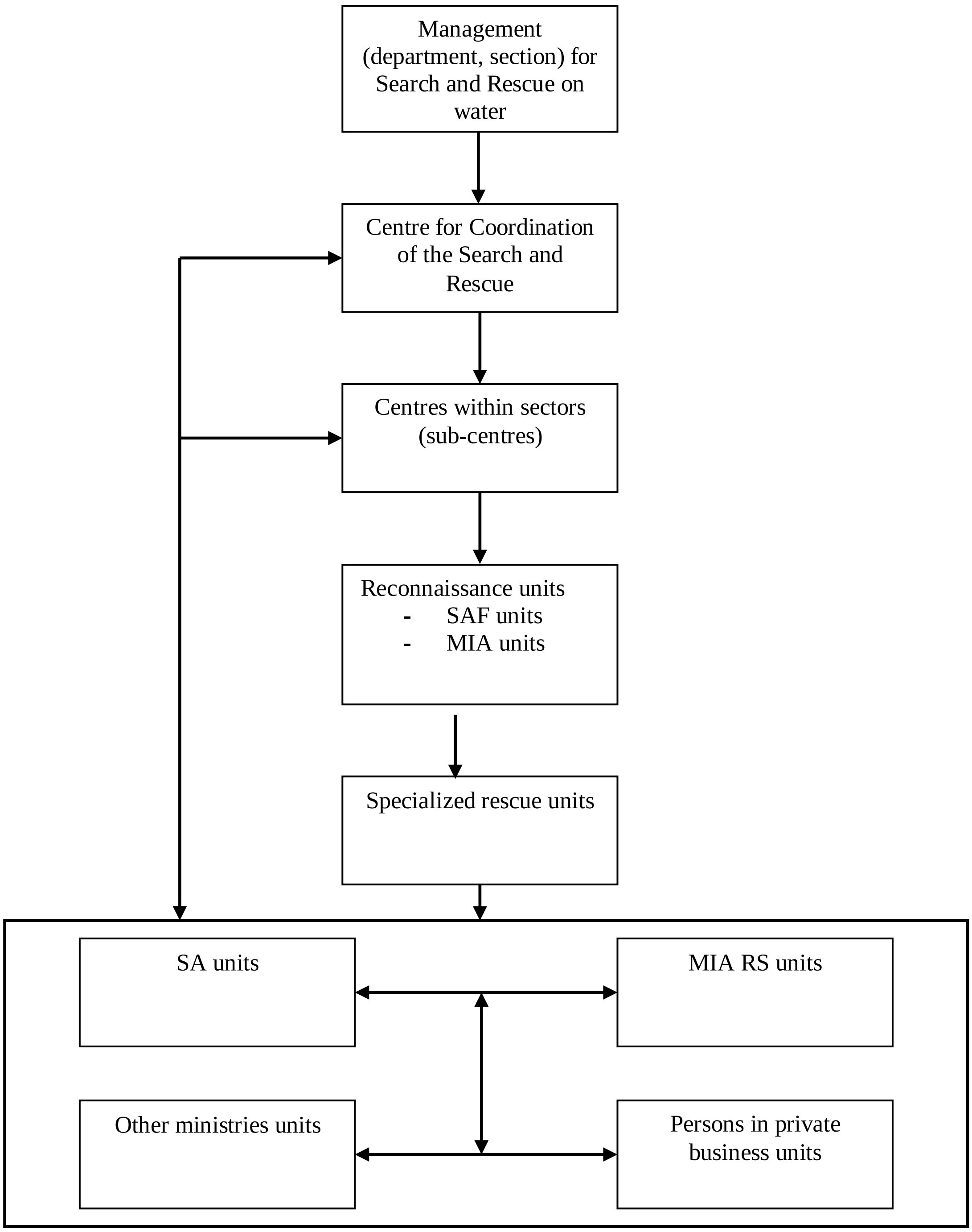



**Fig.2.**
*Preposition for Scheme Overall organisational search and rescue on water*

## 4. ACTING UPON AN AID CALL – PUTTING SEARCH AND RESCUE ON WATER INTO PRACTICE

The purpose of a search is to determine the exact position of persons in distress and it is of vital importance that this be done as soon as possible, as time passes the chances of survival decline rapidly. If we know the exact location of the distressed, we only need to put the rescue task into practice.

### 4.1. Coordination of the search and rescue on water procedures

Successful management of the search and rescue on water procedures demands coordination between the organizations and units taking part.

The beginning of the search starts with an aid call and is based on the objective conclusion that there has been an accident. An aid for help (which can be made in several ways) can be in the form of a message containing a call or the confirmation of the call made directly to the principally responsible operational centre of the management of the search and rescue service (Centre for Coordination), or indirectly to the centres of the sectors (sub-centres of the main stations). Alert messages are considered to be: a radio message MAYDAY, a signal that there are casualties made by DSC device, an SOS, or some other visual way which indicates that there is some danger and that persons need some aid.

Having received a signal for a help, the centre which has received the call does all the preparations necessary for the search and rescue action to begin. If the exact position is unknown, search forces are directed to the area from which the call has been made, all the interested parties are informed, as well as the neighbouring sub-centres and the main operational centre of the search and rescue on water management service, nearby ships and the full operational ability is established.
Search and rescue can begin even if there is a possibility that an accident may have happened even though no alert message has been received, and based on objective and reasonable assumption that an accident has occurred. In those circumstances the centre for coordination determines the levels of danger:

- the level of uncertainty – there is a doubt that an accident has happened based on facts,
- the level of alertness – there is a doubt that an accident has occurred because after the level of uncertainty no contact in any way or by anyone has been made with the ship. After the level of alertness has passed, the centre for coordination undertakes certain actions so that search and rescue can begin.
- the level of casualties – it is declared when certain information on casualties has been received and a demand that a quick action for rescuing people in distress is needed, or after futile attempt to contact a ship on the alert.

Search and rescue is run by:

- a coordinator in the centre for coordination (Mission Coordinator) – an officer on duty or authorised official in the centre at the time of the reception of the call;
- the commander of the scene of the accident (On Scene Commander) – the commander of a special unit which has been determined by a coordinator at the centre to harmonise the action on the search area;
- the coordinator of the surface search (Coordinator Surface Search) – the commander of the most suitable ship at the place of search is determined by the coordinator or the commanders of all the ships taking part in the search.

In case of accident (failure) of a ship and its call for aid, the commanders of others ships are obliged, if they have understood the call, to come to the scene of the accident and to maintain contact via radio with the ship in distress. At the scene of the accident the coordinator of the surface search (CSS) is established – if the ship is a trade ship or if the commander of the scene of the accident (OSC) – if a specialised ship is involved (including warships).

It is desirable for the OSC (or CSS) to be established as soon as possible and that it informs the closest centre of the sector (the main station) upon taking the duty, and that it supplies regular reports on the development of the situation in equal time terminals.

When approaching the scene of the accident, the strictest discipline and regulation procedures in maintaining radio connection must be adhered to, and if there is a language barrier, INTERCO should be applied.
In order to coordinate the search successfully, OSC makes an overall plan – map of the area which is being searched, but it must also be done by the other ships taking part in the search.

### 4.2. Procedures of the ship in distress

A ship in distress must send a signal to other ships and centres of the sector that it requires help which includes:

- the identification of the ship,
- the position,
- the nature of the accident and the type of help required and
- every other announcement which may facilitate the search: weather conditions nearby, the direction and force of wind, visibility, the presence of navigational hazards (ice, etc); the time of abandoning the ship; the number of crew members remaining on the ship; the number of seriously injured persons; the number and kind of search means which have been activated, etc.

People in distress must endeavour to remain on the ship as long as possible, and if they are on boats and rafts they must keep as close as possible to the scene of the accident for easier evacuation. It is necessary to appoint a lookout who could observe the possible appearance of a ship or a helicopter. At the sight of the rescuers he should endeavour to draw attention with all available means (visual and sound signals).

### 4.3. Procedure of the ship providing aid

Upon receiving a call or message about the accident, the commander (captain) of the ship has to forward it to other ships and nearby riverside stations, and then confirm that the message has been received and give the following information to the ship in distress: identity, position, speed, estimated time of arrival, the real course towards the ship in trouble. Also, it is necessary to maintain communication with the ship in jeopardy, constant service of listening, and organise nonstop service of observation.

Ships heading to the position of the ship in trouble are obliged to prepare themselves for the rescue, both of the ship and men on it. The importance of proper and timely preparations on the ship is invaluable for successful implementation of rescuing measures.

The initial phase is completed when the ships have finished the search in the most probable area, and then the CSS considers the most efficient way of continuing the search.

Although the search is not completed, the CSS directs his actions to repeating the search in the same area, taking into account additional adjustments, and if necessary, expands the most probable area, and upon receiving new information determines again which the new probable area is.

It may happen that the abandoned ship, carried by wind and the course of the mainstream, is located sooner than the rescue boat, so the search should be carried out against the wind, i.e. upstream. At the site of the shipwreck and the sunken ship, parts of the equipment and/or oily traces of petrol can be seen, the rescue boat will probably be located downwind, that is downstream from that wreckage.

It must be stressed that during these search operations Regulations on Avoiding Collision are fully applied because the commander of the ship must always bear in mind the safety of its own ship and crew.

After successful search there comes the rescue. It is necessary to decide on the best possible rescue plan and send the most suitably equipped ships to the site of the accident. Those are the ones with nets for climbing the ship, rescue boats or rafts, or well equipped crew members for entering water and providing first aid to the survivors.

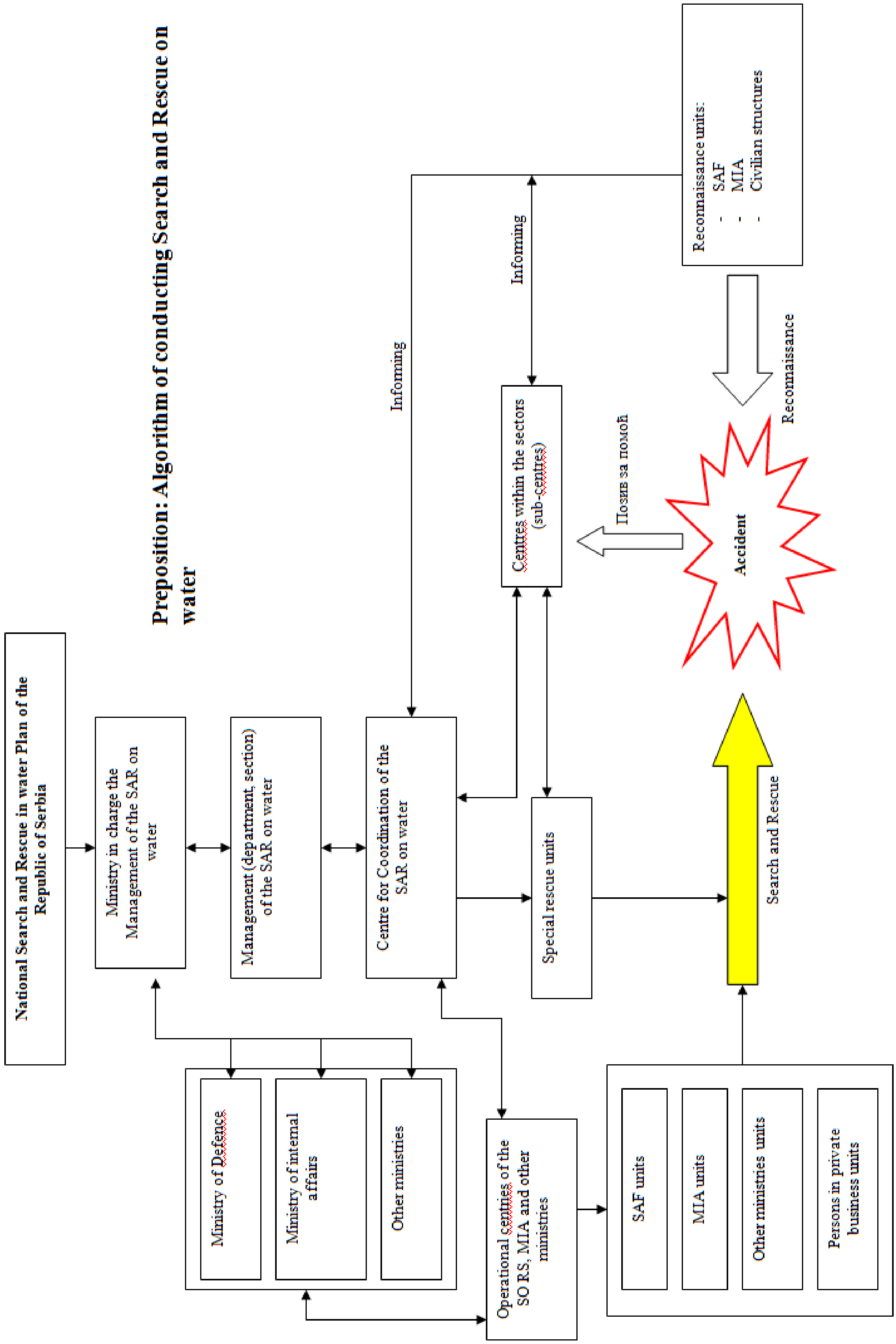
The injured can be transported, if possible, to the ship providing medical help, and then be questioned provided they are able to supply additional information about the accident. All the ships must be informed that the search has ended, as well as the nearest centre of the sector (the main base) that the search is finished by providing details on:

- the names and locations of the ships transporting survivors, and the numbers of survivors on an individual ship,
- the state of the survivors (their health),
- the need for providing medical health and
- the state of the object and whether it presents a danger for navigation.

If the search is unsuccessful, the CSS continues the search until all hope of rescue is lost.

When deciding to end the search, probability is taken into account:

- that the survivors have been there in the area of the search,
- of locating the objects or persons defined by the goal of the search,
- that the survivors might still be alive taking into account the temperature, wind, weather conditions, the speed of the mainstream and
- the remaining time that the units conducting the research could stay at the site of the accident.




**Preposition: Algorithm of conducting Search and Rescue on water**

## 5. CONCLUSION

Rescuing people on rivers, lakes, and canals is every act or activity undertaken in order to preserve a ship (boat) or some other property which is in danger on navigable waterways or any other water surface. Bearing in mind the importance of human lives, rescuing people is always obligatory.

However, we see that in the Republic of Serbia search and rescue on water service is currently not at a satisfactory level. There are good examples of organising these services in some of the neighbouring countries, most of all in the Republic of Croatia and the Republic of Montenegro. In order for the search and the rescue on water service to be fully developed at a national level some system-normative documents must be formulated, centre and sub-centres must be formed, equipment and personnel must be provided, special units must be established, etc. This is all conditioned by material – financial issues, but the state itself must find a way to fund this important safety aspect on all water surfaces at the territory of the Republic of Serbia, because as it has already been emphasised, the inland navigable waterways of the Republic of Serbia stretch to 1,680 km and a very important international navigable waterway runs through the Republic of Serbia (the Danube River – Corridor 7).

The safety on all water surfaces for human lives is a very important factor, as well as organising search and rescue on water service. River Flotilla, as a unit of the Serbian Armed Forces should have a significant role in forming search and rescue units at a national level. Units of the River Flotilla, with some minor investments and modification of the equipment and devices, could contribute in raising the search and rescue on water in the Republic of Serbia to a higher level.

**REFERENCES:**


Government of Serbia, National Strategy for EU Accession / Offices of the Government of Serbia, Belgrad

Government of Montenegro, National Search and Rescue Plan (OG Montenegro No. 4/06), Podgorica, 2006.

Manual search and rescue merchant vessels (MERSAR), IMO, London.

IAMSAR Manual, IMO, London.

Law on Emergency Situations R. Serbia, Belgrad.

Government of Serbia, Regulations on the organization and method of use of specialized civil protection units, Belgrad

Group of authors Theoretical basis GMDS system, Kotor, 1997.

SOLAS, International Maritime Organization, London, 2001.

STCW 95, International Maritime Organization, London, 1996.

International Ship and Port Facility Security Code, IMO, London, 2003.

P.C.Smith and J.J.Seaton; GMDSS for navigators, Butterworth-Heinemann Ltd, London, 1994.

MCA; Personal survival at sea, Southampton.

Available from internet: www.imo.org

Available from internet :www.safety-at-sea.co.uk

# EUROPEAN DIVERSITY – REGIONAL INNOVATION POTENTIAL IN TRANSPORTATION

**Merja Hoppe[1], Martin Winter[2]**
*[1,2]Zurich University of Applied Sciences, Institute of Sustainable Development, Technikumstrasse 9, Winterthur, 8401, Switzerland*

**Abstract:** Transportation is essential for European competitiveness. While the transportation system guarantees location quality for the economy and citizens, the transportation industry provides jobs and strengthens the economy in general. Recently the transportation sector has had to face challenges coming up with new developments. Investments had to be limited due to the financial and economic crisis, new competitors are arising on the global market, climate change requires action affecting transportation and European regional disparities need to be overcome – also and especially when it comes to transportation. The change requires new strategies to adapt to and also to actively shape the process. The European Commission aims to support the transportation sector by promoting innovation on a regional level. To ensure successful effects of funding, financial support should not be spread but focus on regions of highest innovation potential in transportation. The European FP7-funded project METRIC has developed a method to measure innovation potential in transportation applying it to the regions in Europe. The results should serve as a basis for funding distribution and should also identify certain types of regional innovation patterns.
This paper describes the analytical approach to measuring the innovation potential in the transportation sector, also showing first results. While the methodological approach has been developed the results have to be considered preliminary. The project will run until March 2015; the work is still in progress. Already at this stage certain pattern and types of regional innovation can be distinguished showing that enabling factors, including funding, are not always connected with innovation output and at the same time are no preconditions for innovation. Besides this, some regions show a connection between funding and innovation in transportation allowing us to identify the success factors for innovation in a later step.

**Keywords**: transportation, regional innovation, transformation, innovation funding, European disparities

## 1. Transportation and innovation as an European challenge

European regions are diverse in many regards, including transportation. Transportation systems, industries and services serve as the backbone of the economy and society. The transportation sector is important in supporting the European economy in general, providing sound living conditions and jobs for European citizens as well as creating added value. Compared on a global scale the European transportation system can be considered to be of high quality. With a number of world-leading companies in the transportation industry and high-quality infrastructure and services, Europe shows a competitive advantage. Beyond this good position, certain challenges are emerging. These challenges are related to recent changes in frame conditions on a global, European and national level:

1. There is a change going on in the transportation system (not limited to the EU but worldwide) related to investment and maintenance cost. Due to the financial crisis, lower tax revenues and increasing cost for infrastructure, investments in transportation had to be limited. At the same time, a high quality of mobility and accessibility has to be provided to strengthen location quality and support the economy and job creation.
2. The competitive landscape in the transportation sector is changing due to a shift of demand, technical development and production towards Asian and other upcoming competitors.
3. Climate change requires actions to address energy consumption and CO2 emissions in transportation on a political level.
4. European socio-economic disparities need to be overcome to ensure intra-European integrity and competitiveness. This is in fact closely related to transportation providing the basic infrastructure for regional economy and accessibility of jobs as well as employment opportunities in related industries.

The European transportation sector not only has to successfully adapt to these changes but needs to develop strategies to actively shape the process of change. Related to challenges linked to changing frame conditions, innovation is seen as a key issue in successfully managing a process of transformation. By developing new products, services and mobility concepts, regional unique selling points could be developed strengthening economic competitiveness. Thus both regional disparities in transportation infrastructure and industries are crucial for European development. The European Commission aims to address this by supporting the transportation sector through promotion of innovation. Funding has to be effective. Therefore, especially in times of limited budgets of public authorities, funding should focus on regions not only in need of initial financial support but showing promising pre-conditions. Best results from investments in innovation can be expected in regions providing a basis, a potential for innovation, which can be further developed ~~with~~ by giving initial monetary investment support.

To achieve this overall goal of identifying regional innovation potential for funding in transportation, the European FP7-funded project METRIC focuses on analyzing Europe to identify regions with innovation potential in the transportation sector. Based on this, types of regions should be characterized to provide a basis for funding decision on a regional level. An indicator-based approach has been developed in order to measure regional innovation potential based on available statistical data. Besides indicators reflecting innovation itself, ~~also~~ variables linked to socio-economic frame conditions will also be integrated in the analysis. This is based on the idea that economic structure, industrial clusters or labor market aspects are related to innovation potential. The identified regional innovation potential should serve as a basis for focusing funding in transportation and to realize potential, leading to increased regional competitiveness in the transportation industry.

The paper describes the approach, method and preliminary results of the innovation potential analysis, which is work in progress. Using statistical data provides empirical evidence, which at the same time is limited due to data availability. Thus, further steps of the project will include additional qualitative analysis and a plausibility check of results based on case studies.

## 2. Measuring regional innovation potential

Related to the aim of identifying regions with a high innovation potential and to promote these regions individually, the methodological approach was based on the EU conception for regional innovation. Two aspects had to be considered basic prerequisites:

I. Dealing appropriately with regional disparities within the EU, and
II. Developing a method to measure innovation potential in transportation on a regional level.

I. Regional disparities: Due to the socio-economic differences within Europe, regions should not be comparatively analyzed as a whole; regions might not be able to catch up with leading European transportation clusters, even if they might have potential for innovation on a lower, but viable, level. Regional disparities can be taken into account by separating different groups with each reflecting a certain socio-economic level. This allows a comparison of every region in Europe within a group of equal or even similar preconditions. The European Commission used the Gross Domestic Product (GDP) per head as a basis to distinguish different regional economic levels and to determine the necessity of funding in every European region (European Commission, 2013a). In METRIC the method was used to distinguish between three types of European regional levels based on the GDP (figure 1).

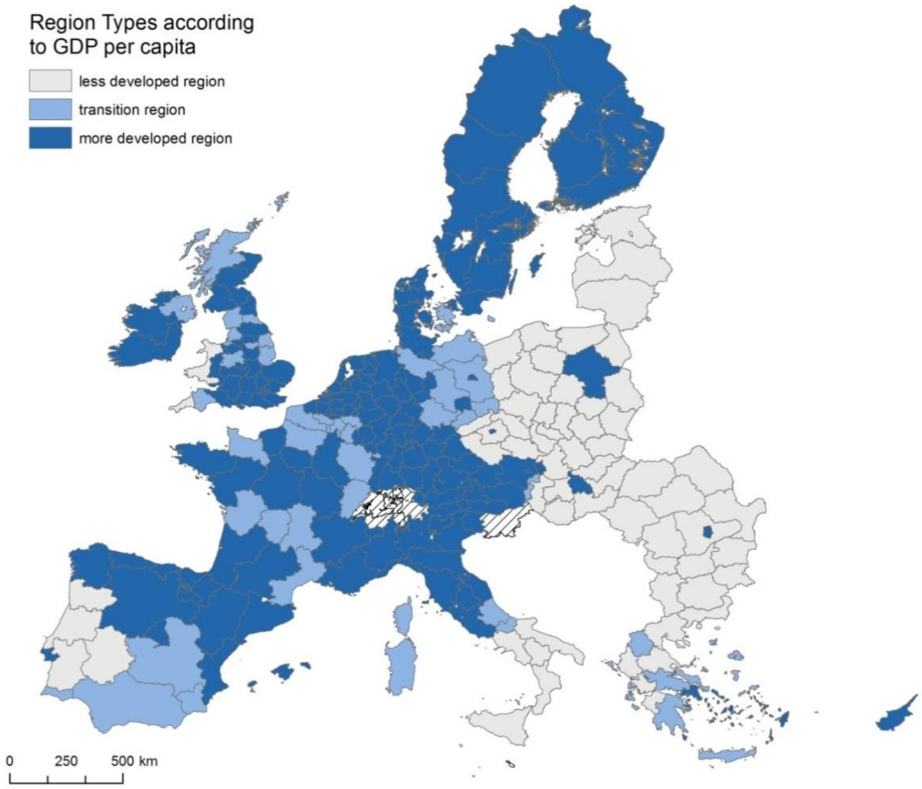



**Figure 1.**
*Typology of NUTS2 Regions (data source: European Commission 2013)*

According to the classification method, “more developed” regions report over 90% of the EU-average of the GDP, the “transition” category between 75% and 90% of the EU-average and “less developed” regions lower than 75% of the EU-average. (European Commission, 2013a) The categorization was used to compare regions within their group in the following innovation potential analysis.

II. Measuring regional innovation potential in transportation: when developing a method one has to face the challenge of limited availability of innovation-specific data for the transportation sector, especially when it comes to regionalized data. METRIC reverted to existing innovation indicators used by the European Commission, and modified and adapted the indicators to transportation.

The European Commission developed the so-called "Innovation Union Scoreboard" (IUS) as a tool to measure general innovation on a country or NUTS 1 level. A set of 25 indicators and eight innovation dimensions allows us to measure and compare innovation (European Commission, 2013b). For regional scales of NUTS 2 level, the European Commission has created the "Regional Innovation Scoreboard" (RIS). As not all data used for the IUS indicators are available on a regional level, the scoreboard uses 12 indicators out of the 25 indicators of the IUS. The RIS subset covers the areas of human resources, education, research & development, funding & investments, patents as well as economic effects (European Commission, 2012).

As the METRIC project aims to measure regional innovation focused on transportation, the RIS-approach, measuring innovation in general, had to be modified to adapt it to transportation. Since the EU has no tool focused on transportation, the METRIC project created a "Regional Innovation Scoreboard Transportation" (RIST), based on RIS. The new scoreboard includes 10 different indicators using the same structure and concept as RIS. Each indicator was selected from the Eurostat database with particular attention to the fact that the innovation potential is to be measured. Every indicator is related to the transportation sector in order to get the branch-specific innovation potential score. RIST (similar ~~as~~ to IUS and RIS) uses indicators classified in three groups related to different aspects of innovation:

1) The **Enabler** indicator group "captures the main drivers of innovation performance external to the firm" (European Commission, 2013b, p. 5) using data on human resources, research systems and funding:

1. Share of government R&D spending on transportation in percentage.
2. Share of highly-educated (tertiary level) persons employed in the transportation sector
(including all manufacturing and services).[1]
3. Share of persons employed in science and technology in the transportation sector
(including all manufacturing and services)
4. Employment in technology and knowledge-intensive sectors in the transportation sector

2) The **Firm Activities** group of indicators reflects the innovation efforts including firm investments, entrepreneurship as well as intellectual assets (European Commission, 2013b). The firm activities of the RIST group consist of the following indicators:

1. Average number of patents per year per 100.000 employees in the transportation sector.
2. Share of innovative enterprises (which introduce product/ process innovation every year) in the transportation sector.2
3. Share of enterprises that have introduced new or significantly-improved products that were new to the market.
4. Share of enterprises that have introduced new or significantly-improved products that were only new to the firm.

3) The **Output** indicator group is about the economic effects of innovation (European Commission, 2013b) and is characterized by the following indicators:

1. Growth value added in the transportation sector.
2. Growth employment in the transportation sector.

Data for the single indicators have been collected for the 261 European NUTS 2 regions. To meet the above-described needs of comparability of the heterogeneous European regions they have been divided into the three groups of less developed (65), transition (48) and more developed regions (148) (see chapter 2). For each single indicator values were calculated within each of the three groups as share of the groups' average – which is in line with the original method of IUS and RIS. Values of each indicator have been standardized to allow a better comparison. Finally, the summation of all indicators builds a general score of "innovation potential in transport" as shown in figure 2. Consequently, the assumption is that the higher the innovation potential of a region, the bigger the score. While the general score provides a general ranking about the sum of innovation aspects, single indicators allow the user to analyze specific aspects of the innovation potential in detail – serving as a basis to support the further interpretation of the preliminary results described in this paper.

The above listed indicators can not only be analyzed one by one or as a total score (figure 2); the indicators were also grouped according to the above described classification of 1) Enablers, 2) Firm activities and 3) Output (see figures 4, 5 and 6 in the Annex).

[1] Indicator 2., 3. and 4.: share in percentage of the total of persons employed in transportation sector.
[2] Indicator 1., 3. and 4.: share in percentage of the total of enterprises (min. 10 employees) in transportation sector.

**4. Distribution of innovation potential in transportation across Europe**

The result of the general innovation potential score in figure 2 provides a classification of scores in quartiles. The 25% of highest ranking regions within each group (more developed, transition and less developed regions) are labeled as "Top", the next 25% are classified as "High", followed by the next 25% in the "Moderate" and the last quarter in the "Low" group. The general score reflects the heterogeneity of Europe in transportation, although the different economic levels have already been considered by distinguishing between different groups according to their level of development. All categories of Top, High, Moderate and Low potential are spread all over Europe - although some innovation clustering structures can be vaguely discerned. To identify spatial patterns of innovation potential, the results were generalized.

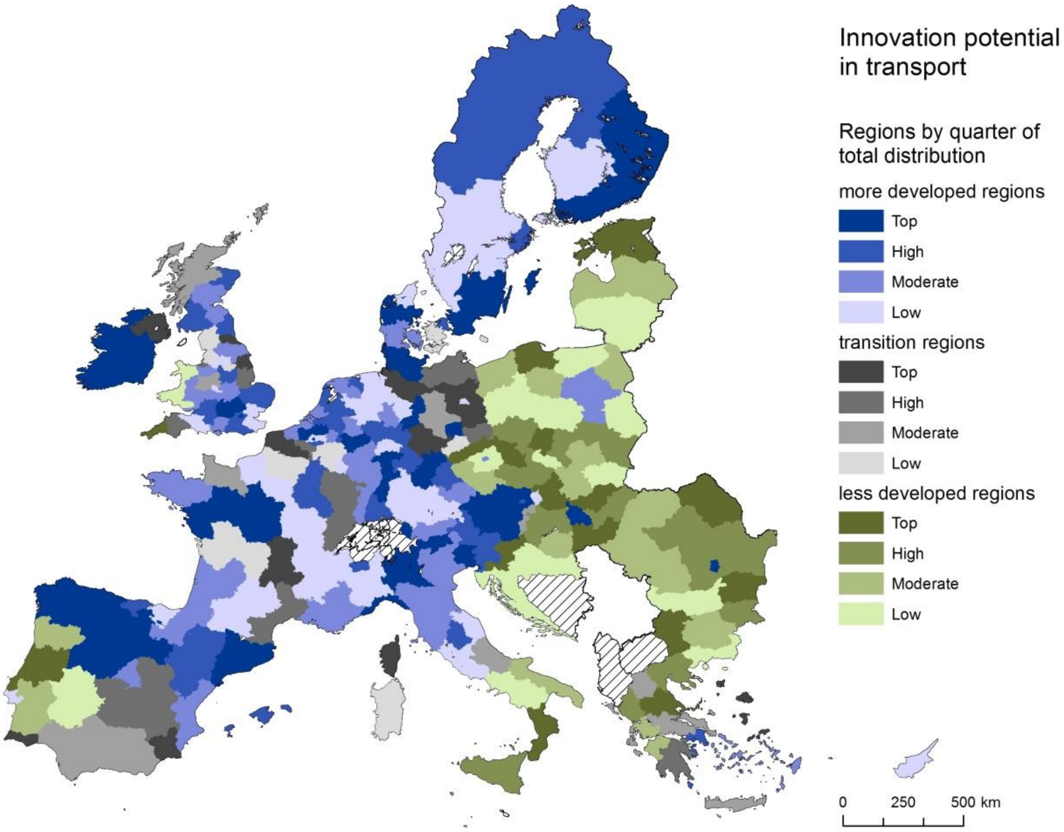



**Figure 2.**
*Innovation potential in transportation sector (data source: Eurostat 2014)*

To build categories, the main innovative regions have been identified and grouped in table 1. The threshold for considering a region in the table was that a NUTS 2 region would need to belong to the top 25% of all regions with at least 8 of 10 indicators (75%). This was to eliminate regions scoring high on few indicators which might lead to a general high score caused by singularities, such as a single strong company in a peripheral region; as innovation potential is understood as a broadly based characteristic of a region these regions should not be considered (e.g. North- and East-Finland; Galicia etc.). The identified highly innovative regions were clustered into virtual axis, corridors or greater regions to provide a generalized overview as listed in table 1 and illustrated in figure 3. Cross marks in the table are set to emphasize the specific strength of the conglomerates. It should be mentioned that findings are based on preliminary results and will have to be further analyzed to provide additional empirical evidence and to refine the results; selection and clustering of regions might be modified slightly, as well as regional characteristics.

**Table 1.**
*Results of Scores in the Regional Innovation Scoreboard in Transportation*

| Index | Indicator \ Conglomerate of NUTS 2 regions | Benelux – NW Germany | Axis Turin - Vienna | Triangle Warsaw – Prague – Budapest | Axis Nantes – Paris | Axis Barcelona – Strasbourg – Dresden | Axis London - Newcastle | Greater Bucharest | Estonia | Ireland |
|---|---|---|---|---|---|---|---|---|---|---|
| Enabler | Share of government R&D spending on transportation sector | X | X | X | X | | | | X | X |
| | Highly educated persons employed in transportation sector | X | | X | | X | X | | X | |
| | Share of persons employed in science and technology in transportation sector | X | X | X | X | X | | | X | X |
| | Employment in technology and knowledge-intensive sectors in transportation sector | | | X | | X | X | X | X | |
| Firm activities | Average number of patents per year per 100.000 employees in transportation sector | X | X | X | X | X | | X | | X |
| | Share of innovative enterprises in transportation sector | | X | | | X | X | X | X | |
| | Share of enterprises that have introduced new or significantly improved products that were new to the market in transportation sector | X | X | X | | | X | | | |
| | Share of enterprises that have introduced new or significantly improved products that were only new to the firm in transportation sector | X | X | | X | X | X | X | | |
| Output | Growth value added in transportation sector | X | X | X | X | X | X | X | | |
| | Growth employment in transportation sector | X | X | | X | | | X | | X |

Based on the above-described general score for regional innovation and the subscales of table 1 certain characteristics for types of regions concerning their type and dynamic of innovation can be identified. Regions scoring high in innovation differ substantially concerning their performance in firm activities, innovation output and enabling conditions, as measured by the indices.

- Regions along the axis Benelux-NW-Germany and Turin-Vienna as well as the Eastern triangle Warsaw-Prague-Budapest show high values in the enabler group, which reflects high education level of regional labor force and R&D spending of government. In these regions enabling conditions seem to lead to strong regional output effects of growth value added and employment in transportation; first additional analysis also showed that general economic and wealth aspects like purchasing power of inhabitants and low unemployment rates are related to this. Also firm activity related outputs are directly linked to innovation. Patents as well as intra- and extra-firm product innovation are above average in these regions. Northwest-Germany is the leader according to the scores. In the Eastern triangle, Budapest serves as leader of this region in terms of highest scores on most indicators compared to Prague and Warsaw, especially with regard to firm activities. An exception is Prague leading in innovation output. The axis Turin-Vienna is especially strong in firm activities, which is also related to growing employment and added value – while enabling (e.g. highly-qualified employees in Vienna) is weaker.

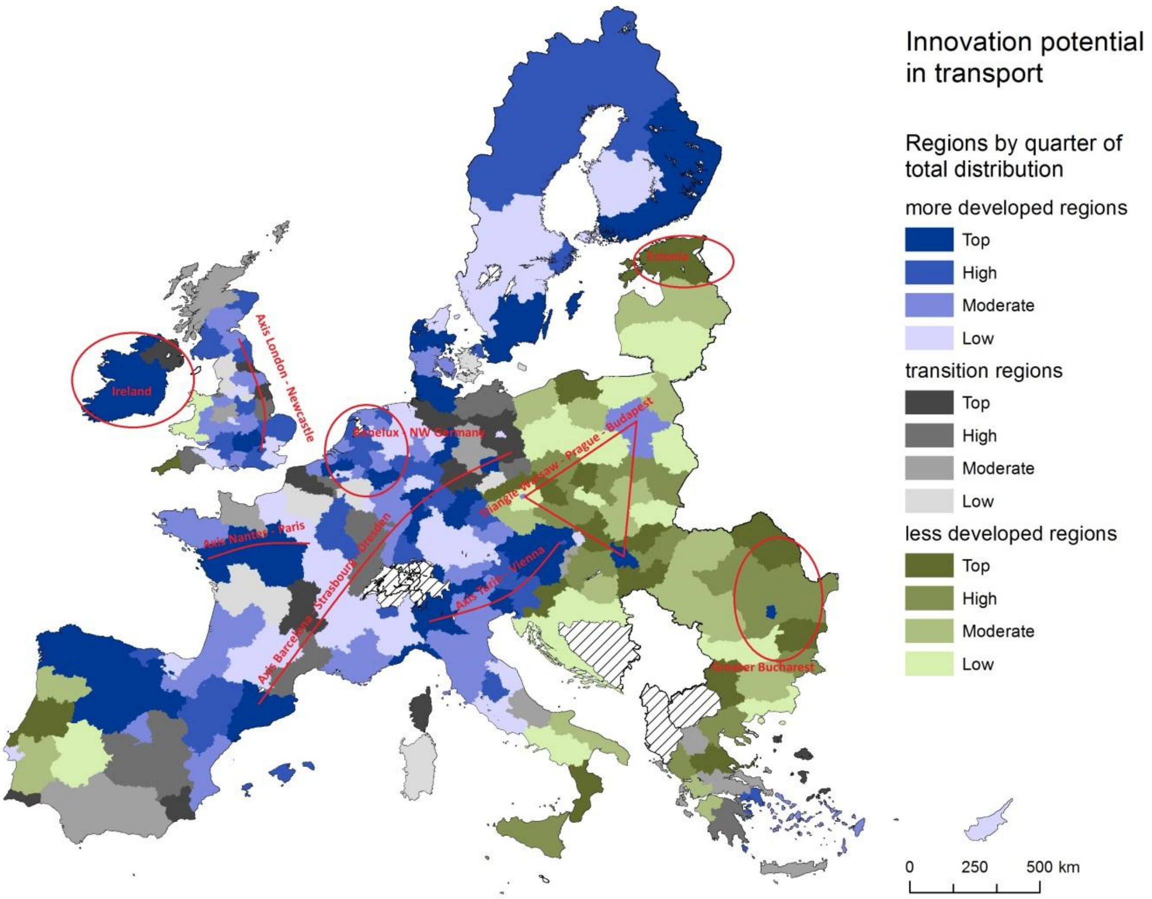



**Figure 3.**
*Axis of Innovation potential in transportation sector (data source: Eurostat 2014)*

- In contrast to the first group, Estonia reaches high scores in education and governmental R&D investment enabling innovation, without showing outstanding innovation output – enabling factors do not (yet) have effects on job and economic growth; at least the region is strong concerning innovative enterprises. Although the region might have potential for future development of innovation in transportation, the recent specialization of companies might not reach the high-tech or high innovation level if compared to the whole of European regions. Ireland shows similar low innovation activities, even though it scores high in enabling factors. The share of government spending in R&D in transport is high, which leads to the assumption that innovation funding might lead to job growth, without producing innovation.

- The Greater Bucharest region shows positive innovation effects, especially with regard to the labor market, without a noticeable supporting environment like R&D subsidies or highly-skilled labor force. Nevertheless the region's firm activities are remarkable concerning patent and product innovation, also related to employment in knowledge-intensive sectors. This might be explained by certain specializations of companies in the transportation sector – without having the industry embedded in a high-tech regional economy.

- Axis Barcelona – Strasbourg – Dresden / Axis Nantes – Paris appear similar as to the first group. The regions show a kind of "classical" innovation pattern with high shares of employment in science and technology, patent applications and product innovation which is reflected in growth value added and employment growth in transportation (the latter with the exception of regions in the corridor Barcelona, Strasbourg, Dresden). In general these regions are characterized by economic wealth measured by the purchasing power of inhabitants. With the exception of the French part they do not receive above-average governmental investment on in R&D. The axis Barcelona – Strasbourg – Dresden is strong in firm activities like patents and firm innovation, employment in knowledge intensive sectors including R&D, engineering, high-tech production etc. The French axis is not really strong in firm activities (especially in innovation to market), but in innovation enabling (education and R&D funding) as well as in economic success reflected in output indicators. It should be mentioned that Paris is not the leader with highest scores along this axis, Nantes and the surroundings of Paris are stronger in transportation innovation – which can be explained by suburbanization characteristics of the sector.

- The axis London – Newcastle differs from the latter described type of regions as they are especially strong in innovative firm activities, which are related to growth added value in transportation, but not significantly reflected in job growth in the sector or patent applications. Also general regional wealth level is not considerably higher when compared to the previously described traditional innovation regions. Greater London is the leader in this group with the highest score.

As the description and interpretation of the preliminary results show, further supporting work has to be done to provide empirical evidence for the correlation between funding, enabling factors and innovation production as well as positive innovation effects. Limited by data availability, no additional statistical evidence can be provided. Thus, besides additional deeper analysis of the given data set, the METRIC project will combine the data based results with qualitative information – leading to an estimation of regional innovation potential in transportation at the end.

## 5. Conclusion

Besides identifying European regions with high innovation potential in transportation – some of them are already performing well – the preliminary stage of the METRIC analysis reveals also open questions and additional challenges.
The first results show an innovation landscape as heterogeneous as Europe itself. Many regions are strong in firm activities without scoring high in the enabling indicators; other factors than those analyzed enable innovation in these regions or their "enablers" do not score especially high, still supporting innovation sufficiently.
Other regions are strong in both or even all of the innovation indices. These regions seem to follow the "ideal" path of good regional frame conditions providing a good basis for innovation activities – which in turn is leading to beneficial outcomes like job growth.
On the other hand, some regions supported by a high level of innovation enabling benefit from job growth without showing remarkable innovation activity; investment seems to lead to economic growth in transportation, without stimulating real innovation activities.
These contradictory findings point out that there is no single innovation path, but a more complex connection between enabling factors, actual innovation activity and beneficial, measurable outputs of innovation.
Thus, in the next steps of the project, the interdependencies, relationships and types of different innovation aspects will be analyzed in more detail. The aim is to identify regional types of innovation landscape. As regional innovation potential seems to be related to specific characteristics and regional processes in the transportation sector, they have to be considered for recommendations on supporting funding by the European Commission. Besides additional qualitative analysis this step of verification will also include a critical reflection on the used indicators. Limited available data is a serious challenge which needs to be overcome.
Besides this, socio-economic frame conditions will be included in the analysis to provide additional knowledge about framing, supporting and enabling factors for regional innovation as well as providing links to outcomes of high innovative capacity on the regional level.

## Acknowledgements

The results presented in this paper have received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 605428'.

**Appendix**

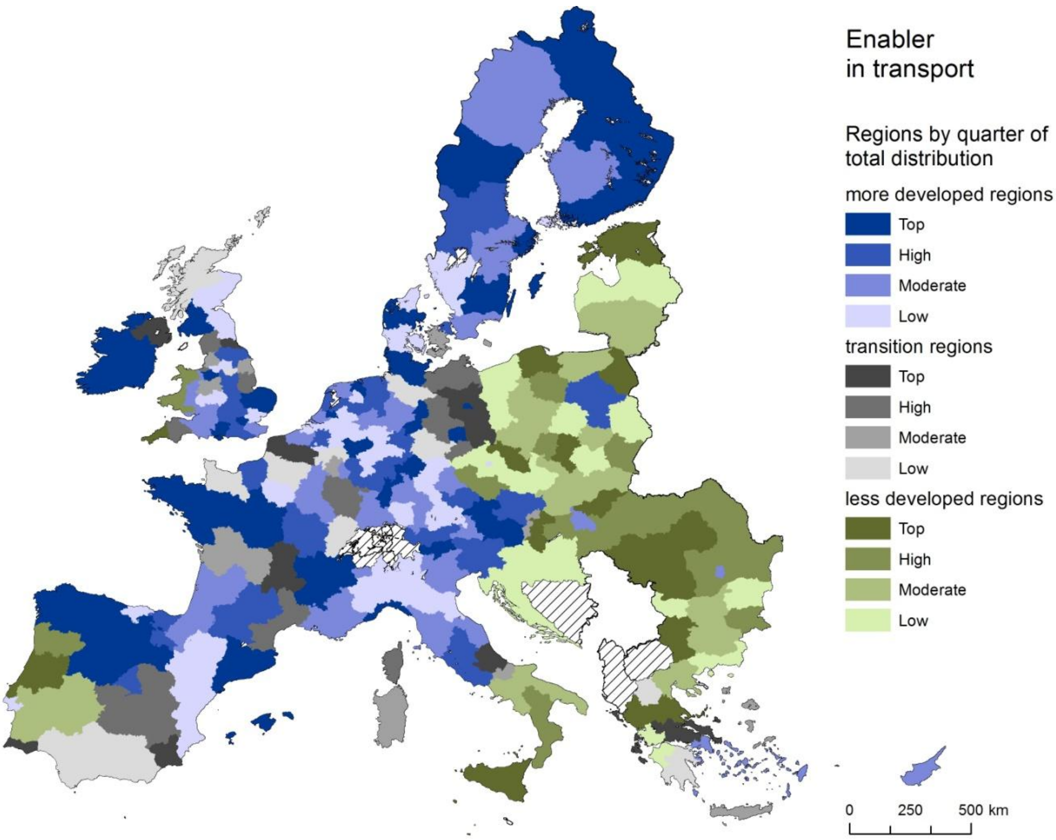



**Figure 4.**
*Enabler group of the Regional Innovation Scoreboard for Transportation (data source: Eurostat 2014)*

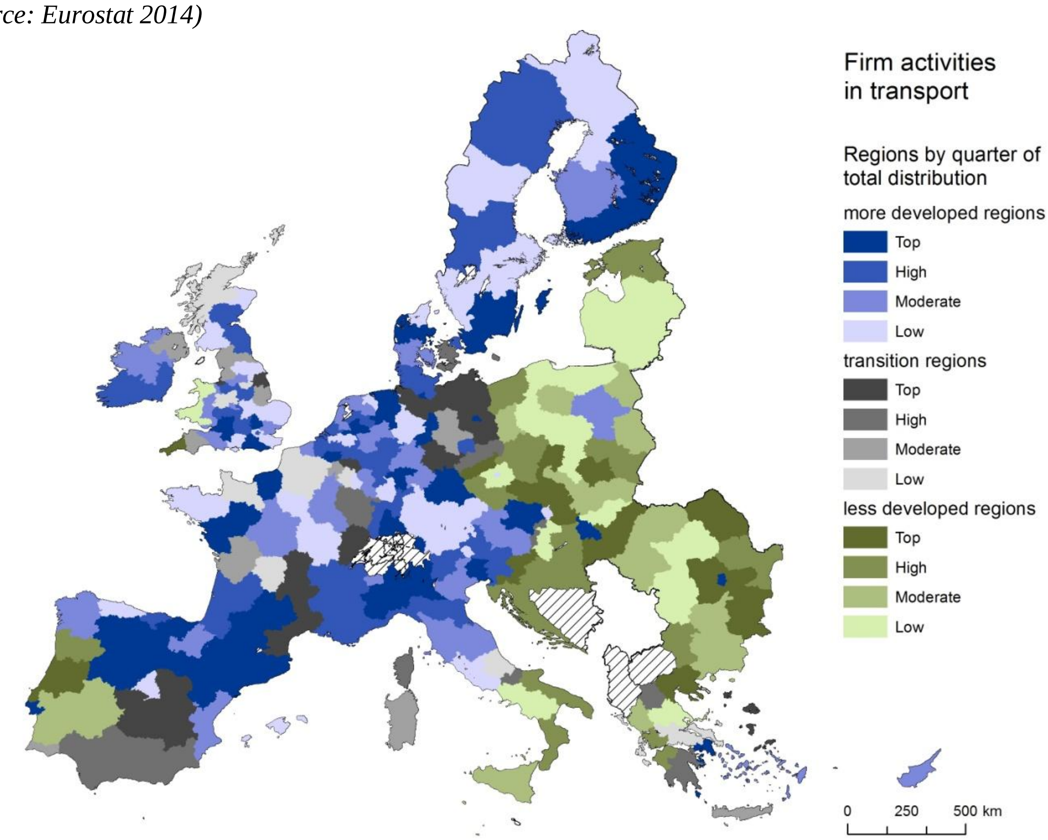



**Figure 5.**
*Firm activities group of the Regional Innovation Scoreboard for Transportation (data source: Eurostat 2014)*

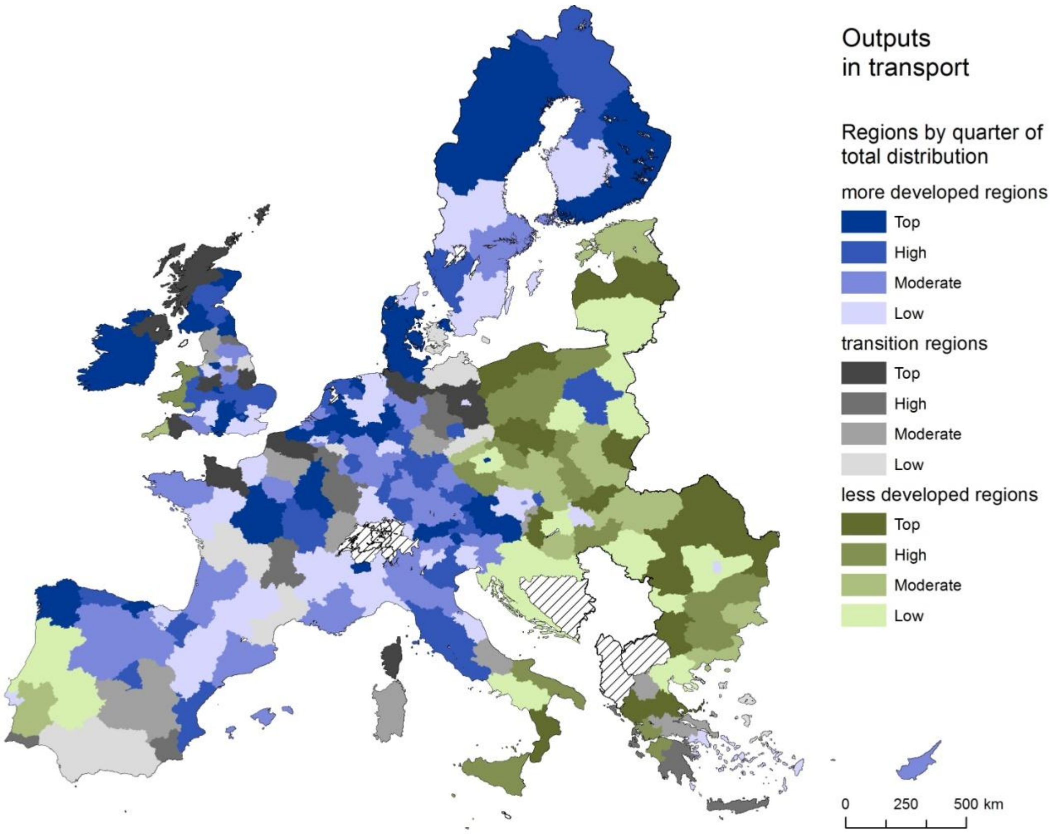



**Figure 6.**
*Outputs group of the Regional Innovation Scoreboard for Transportation*
*(data source: Eurostat 2014)*

**References**

European Commission. (2012). Regional Innovation Scoreboard 2012. Brussels: European Commission.

European Commission. (2013a). Refocusing EU Cohesion Policy for Maximum Impact on Growth and Jobs: The Reform in 10 points. Brussels: European Commission.

European Commission. (2013b). Innovation Union Scoreboard 2013. Brussels: European Commission.

European Commission. (2014). *COMMISSION IMPLEMENTING DECISION of 18 February 2014.* Brussels: European Commission.

# INNOVATION PERFORMANCE OF THE TRANSPORT SECTOR AT REGIONAL LEVEL

**Rob Konings[1], Erik Louw[2]**
*[1,2]Delft University of Technology – Faculty of Architecture, OTB Research Institute for the Built Environment, Delft, The Netherlands*

**Abstract:**Transport is a key enabler of economic and social activity and a major industry in the European economy. However, transport is also a source of environmental concerns and other negative externalities. Improving the efficiency of the transport system, mitigating its negative impacts and increasing the competitiveness of the transport industry are the key challenges of EU policy. Innovation is essential for the European transport strategy to achieve these challenges. Acknowledging the fact that regions are important engines of economic development, regional innovations can be important drivers for growth and improvement of the transport sector in Europe. The innovation process and relative performance of the transport sector at regional level is, however, not extensively studied.
In the framework of the European FP7-funded project METRIC a method has been developed to measure and explain the innovation performance of the transport sector at regional level.
This paper outlines this method and presents the results of applying this method to 251 European regions. The method is based on structural equation modelling. This modelling technique enables to get scores for the innovation performance of each region and to derive estimations that express the relative importance of the explanatory factors for innovation performance.
The analysis showed that the high performing regions were predominantly found in Germany and Sweden, while the low performing regions were located in Finland, Poland, Romania, Bulgaria, Hungary, Spain and United Kingdom. However, the scores of regions on innovation performance are rather different when the performance is measured for the transport manufacturing and transport service sector individually. Among the factors that may explain transport innovation performance of regions funding possibilities appear to be highly important, while the relevance of innovation milieu is outmost limited.

**Keywords**: transport innovation, innovation indicators, regional analysis, structural equation modelling

## 1. Introduction

Transport is a key enabler of economic and social activity and a major industry in the European economy. The transport industry employs more than ten million people, accounting for 4.5% of total employment in the EU, and representing 4.6% of gross domestic product (GDP). Manufacture of transport equipment provides an additional 1.5% of employment and 1.7% of GDP (European Union, 2013).
The European transport system is one of the world's best qualified systems, but nevertheless faces major challenges. Growth in passenger and freight transport demand has increased considerably in the last decades and is predicted to rise further. However, overall infrastructure capacity is reaching its limits and (financial) possibilities to expand transport infrastructure significantly are also limited. The challenge is to make more efficient use of existing infrastructure, while meeting higher requirements for safety, security and reliability as well as user convenience (European Union, 2012).
The unabated growth in transport demand also endangers the target of achieving a 60% reduction in greenhouse gas emissions by 2050 from the 1990 level. The increase in total emissions that result from transport growth will likely offset possible reductions of new efficient technologies. Development of more energy-efficient technologies and stimulation of the use of more sustainable transport solutions are therefore needed.
Last but not least, the leadership position of the European transport manufacturing (vehicles and other transport equipment) is challenged. The emerging role of in particular Asian countries endangers the European competitiveness and hence employment.
Innovation is considered essential for the European transport strategy to tackle these challenges. Innovations should lead to more efficient, environmentally sound and more intelligent transport in Europe and to more competitiveness of the European transport industry in developing and selling such new, smart transport solutions.
The conditions for realising innovations and being successful in implementation will vary in Europe, not only at country level but also at regional level. In view of limited budgets of the European Commission to support innovation, the need to spend these budgets effectively, the acknowledgement that regions are important engines of economic development and the fact that regions in general can increasingly apply for EU-finance (e.g. through the smart specialisation strategy program) there is an interest to identify the transport innovation potential of regions. The FP7-funded project METRIC focuses on analysing this transport innovation potential of regions. As part of this METRIC-project this paper describes a method to measure and explain transport innovation performance of regions and presents the results of applying this method to 251 European regions.

[1] Corresponding author: j.w.konings@tudelft.nl

## 2. Framework for measuring and explaining transport innovation potential

### 2.1 Conceptual model

Innovation performance in general is measured on a regular base across the European Union. The most extensive measurement tool is the Union Innovation Scoreboard (UIS), which provides an annual benchmark of the innovation performance at the country level of the EU member states and other European countries. Innovation performance is measured using a composite indicator – the Summary Innovation Index – which summarizes the performance of a range of different indicators. IUS distinguishes between 3 main types of indicators – Enablers, Firm activities and Outputs – and 8 innovation dimensions, capturing in total 25 indicators.
The other measurement tool, the Regional Innovation Scoreboard (RIS), accompanies the UIS. The RIS provides a comparative assessment of innovation performance across 190 regions of the European Union, Norway and Switzerland, but it is less frequent and detailed due to lack of innovation data at the regional level. The RIS 2014 uses data for 11 of the 25 indicators in the UIS. Similar as in the IUS where countries are classified into 4 different innovation performance groups, Europe's regions have also been classified into Regional Innovation leaders (34 regions), Regional Innovation followers (57 regions), Regional Moderate innovators (68 regions)and Regional Modest innovators (31 regions) (Hollander et al., 2014).
In order to measure the transport innovation performance of regions it would have been logic to adopt the approach of UIS and RIS. The limited availability of data at regional level, however, becomes even more problematic when regional data needs to be collected at sectoral level, i.e. the transport sector. For several of the indicators in RIS the data are not available at the level of the transport sector. On the other hand data regarding relevant indicators to measure transport innovation performance are only available at country level. In view of these limitations the RIS measurement framework has been adopted to the framework as presented in figure 1.

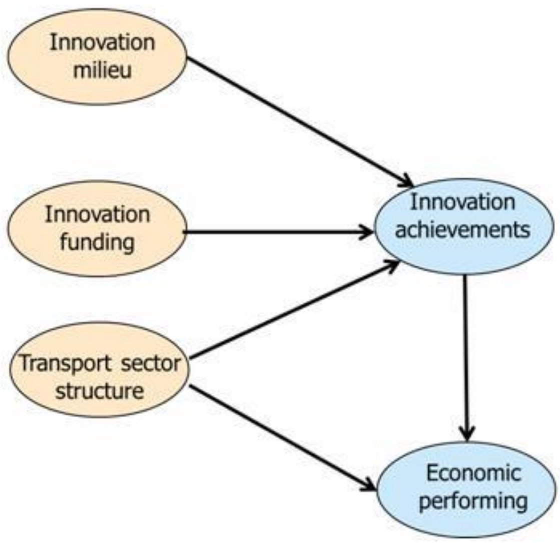



**Figure 1.**
*Measurement framework of regional transport innovation performance*

For the measurement of regional transport innovation performance two dimensions have been distinguished (see also the Regional Innovation Scoreboard 2012, Methodology report: Hollanders et al., 2012):

- **Innovation achievements**: captures the *output of innovation activities* in the transport sector in the region;
- **Economic performing**: captures *the economic effects of innovation activities* in the transport sector in the region.

Based on literature review we have defined major aspects that are often quoted in literature as relevant factors to encourage innovation activities and to succeed in the implementation of innovations.
These drivers are considered as explanatory factors for the innovation achievements. In addition, it is assumed that the economic performing of the transport sector can be explained by its innovation achievements. Innovation may induce better economic performing, but the level of economic performing of the transport sector will obviously also be determined by other factors. These other effects on the regional economic performing of the transport sector are supposed to be caused by the characteristics of the transport sector in general (see figure 1).
Taking into account the availability of data at regional (NUTS 2) level indicators have been defined to reflect (or actually enable measurement of) innovation achievements and economic performing as well as each of the explanatory factors.

### 2.2 Indicators: data

The next step was to define indicators to enable the measurement of ‘Innovation achievements’, ‘Economic performing’, ‘Innovation funding’, Innovation milieu’ and ‘Transport sector structure’. The availability of data played a key role in this process. Data was needed at the NUTS2 regional level. To a great extent the data could be obtained from Eurostat statistics. Other sources were the Joint Research Centre of the European Commission and the Cluster observatory. However, not all data were available at the NUTS2 level. In particular data of the Community Innovation Survey (CIS) from Eurostat were only available at country level. Where needed the data have been regionalized according to the regionalization technique proposed in the RIS 2014 ((Hollanders et al., 2014: 37). Furthermore, the number of missing values in the data have been reduced by using different imputation techniques (see Konings and Louw, 2014). Regions with a substantial number of missing values, however, were dropped. Hence 251 (of 284) regions were kept for the analysis. The data for the transport sector covered both the *manufacturing sector* (i.e. vehicles and other transport equipment) and the *transport service sector*. The following indicators were used:

Innovation achievements:
- Share of innovative enterprises in transport sector 2010.
- Share of highly innovative enterprises in transport sector 2010.
- Average number of patents per year (2006-2008) per 100.000 employees (2008-2010).
- Share of enterprises that have introduced new or significant improved products that were new to the market as share of total population 2010.
- Share of enterprises that have introduced new or significant improved products that were only new to the firm as share of total population 2010
- Turnover from innovations 2010.

Economic performing:
- Growth of value added in the transport sector 2008-2011.
- Growth of employment in the transport sector 2008-2011.
- Growth of labour productivity in the transport sector 2008-2011.

Transport sector structure:
- Average firm size 2010.
- Share of manufacturing (or services) in total employment of the transport sector 2008.
- Level of transport specialization 2008.

Innovation funding:
- Product and/or process innovative enterprises that received any public funding as share of the total population 2008.
- Share of public R&D as % of GDP 2008.
- Share of business R&D as % of GDP 2008.
- Share of government R&D spending on transport 2008.

Innovation milieu:
- Specialisation in Transport research as % of total FP7 EC funding 2008.
- Cluster Quality 2008.
- Population density 2008.
- Share of product and/or process innovative enterprises engaged in any type of co-operation as share of total population 2008.
- Share of high educated persons in total transport employment 2008.
- Share of persons employed in science and technology in total of all manufacturing activities and in transport services 2008.
- Share of employment in technology and knowledge-intensive sectors in transport services 2008.
- Educational level total labour force 2008.

### 2.3 Structural equation modelling (SEM)

SEM is a statistical technique to investigate complex causal relations through combining the technique of regression analysis, path analysis and factor analysis. For the purpose of our analysis SEM is an appropriate tool, because it enables a comprehensive measurement of the innovation performance of regions in the transport sector and it enables to get insights in the relative importance of the explanatory factors for the transport innovation performance.

The influence of the explanatory factors is revealed in the so called structural model of SEM, while scores on the factors, i.e. ‘Innovation achievements’ and etc., are obtained through the so called measurement model of SEM. Usually both models are run simultaneously in SEM, and hence all estimations are obtained in one time. Special software exists (e.g. AMOS in SPSS) to run SEM-models. This approach, however, could not be applied here. The limited number of observations (251 regions) related to the relative large number of parameters to estimate (caused by the number of factors and their relations and indicators) caused difficulties in fitting the model. A possibility to overcome this estimation problem was to construct factors by their indicators before the AMOS analysis was undertaken. Hence a two-step approach was taken in which the scores of factors are input for the SEM-model (i.e. the structural part of the model).

Different methods exist to compute the factor scores, all having some advantages and disadvantages. Based on decision rules and recommendations provided in literature about multivariate data analysis (see e.g. Hair et all., 1998) we have chosen for the method of summated scales (= somscores). This means that each indicator gets an identical weight in the construction of the factor. In order to apply the summated scales all indicators have been standardized by means of transforming their original values into Z-scores.

### 3. Results of the SEM analysis

The type of activities in the transport sector of a region can be very diverse. The activities are in particular different between transport manufacturing and service companies, and hence differences are also very likely in their innovation behavior and performance. A region may show a good innovation performance in transport manufacturing, while the transport service companies may have a weak innovation performance. It is important to note that the innovation performance of the total transport sector can be a result of very different performances in the two subsectors. Therefore a SEM analysis is performed for the total transport sector, but also for the transport manufacturing and transport service sector individually.

#### 3.1 Model estimations for the total sector

The estimations for the total sector are shown in figure 2, i.e. the standardized path coefficients, which express the strength of the causal relations and the R squares for ‘Innovation achievements’ and ‘Economic performing’, which express the (local) explanatory strength of the model. The double arrows (expressing correlations between the explanatory factors) have been included to obtain a pure effect estimation of each explanatory factor only.

‘Funding’ is the most important determinant (0,69) for the ‘Innovation achievements’, followed by ‘Sector structure’ (0,30). The role of the ‘Innovation milieu’ should be interpreted as negligible as the path coefficient is close to zero. In view of theoretical notions this is a remarkable result. Specific attention was given to the composition of indicators in this factor. Several variants in the composition of indicators - in which fewer indicators were included – were tested in different model runs. The results, however, were robust. That is to say, they did not lead to relevant changes in the path coefficient that describes the strength of the influence of ‘Innovation milieu’ on ‘Innovation achievements’. Therefore we kept the original composition of indicators for the factor ‘Innovation milieu’.

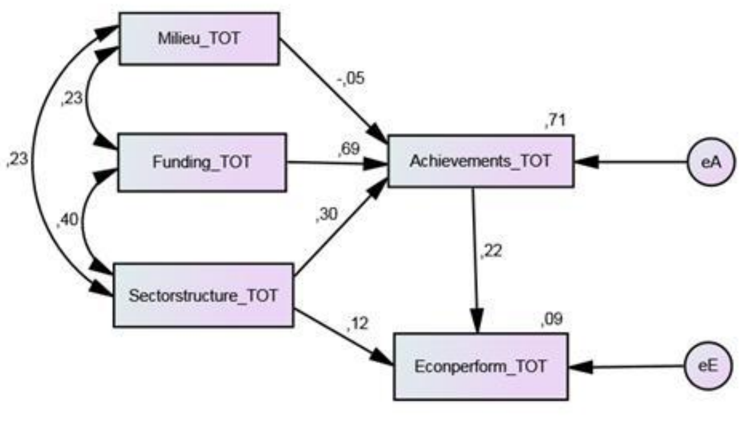



**Figure 2.**
*Parameter estimations for the total transport sector*

Figure 2 also shows that ‘Innovation achievements’ are a significant determinant for the ‘Economic performing’ of the transport sector of a region. In addition, there is a direct effect of ‘Sector structure’ on the ‘Economic performing’ (0,12) as well as an indirect effect that is induced via ‘Innovation achievements’. In the right corner, on top of the endogenous factors the R-square is plotted. It seems that the ‘Innovation achievements’ are quite well explained by the factors ‘Funding’, ‘Sector structure’ and ‘Milieu’.

On the other hand the 'Economic performing' of the transport sector can, to a very low extent, be explained by this model. Apparently, there are others factors relevant for the 'Economic performing' which are not included in this model.

In figure 3 the relative scores of the regions on 'Innovation achievements' are mapped. In this map two categories represent the scores of regions above the European average ('above average' and 'far above average') and two categories reflect the scores below the European average ('below average' and 'far below average').
The high performing regions are predominantly in Germany, and, remarkably, regions in Portugal and Sweden also score well. On the other hand regions in France, Spain, Norway, Hungary and in particular in Finland, Poland, Romania and Bulgaria are at the lower end of the scores.

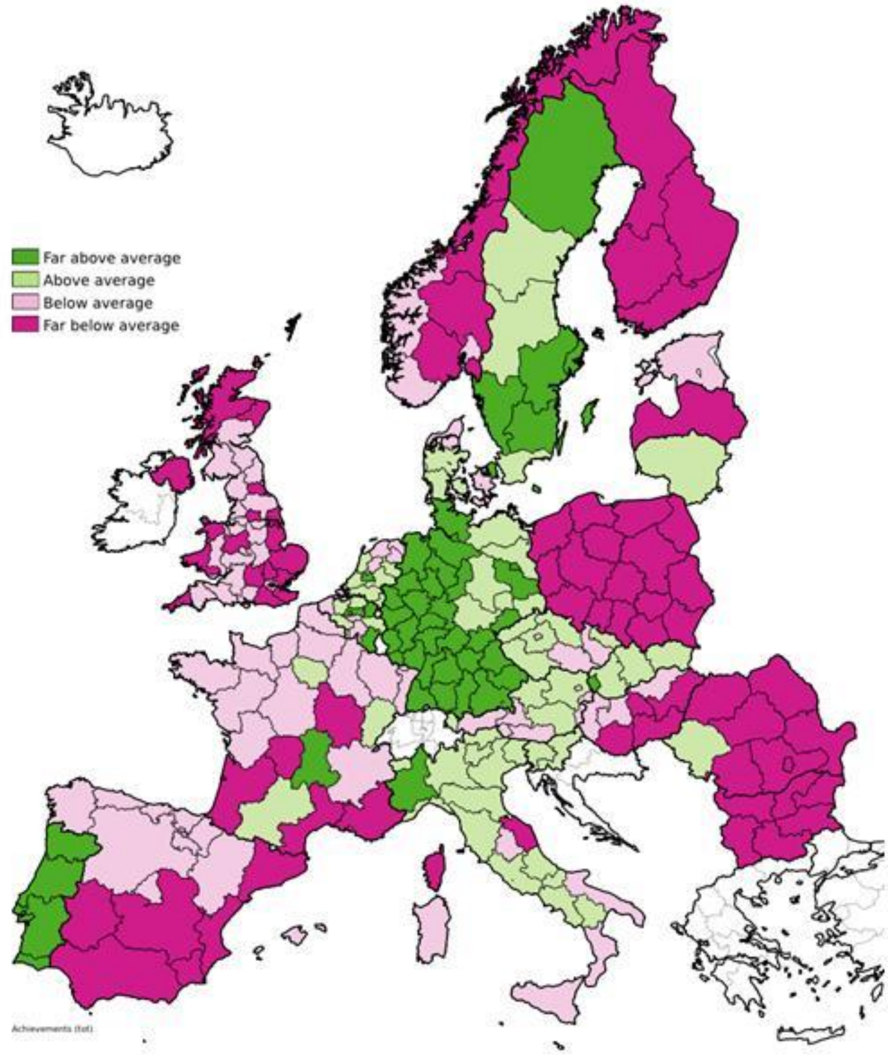



**Figure 3.**
*Scores of regions on 'Innovation achievements' in the total transport sector*

**3.2 Model estimations for the transport manufacturing sector**
Figure 4 depicts the result of the analysis for the transport manufacturing sector only. Both 'Funding' and 'Sector structure' seem to be important for the 'Innovation achievements' in this subsector. The role of the Sector structure may point at the relevance of 'mass' to achieve strong innovation performances in the manufacturing sector. Namely, the presence of large companies and a high share of transport manufacturing companies go along with high scores on 'Innovation achievements' of regions. The importance of 'Innovation milieu' for 'Innovation achievements' appears also for the manufacturing sector marginal. The variance in 'Innovation achievements' which is explained by the model (0,61) is nevertheless substantial.

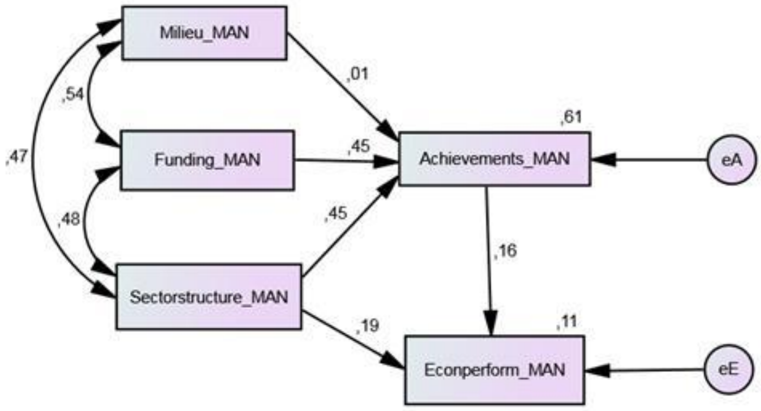



**Figure 4.**
*Parameter estimations for the transport manufacturing sector*

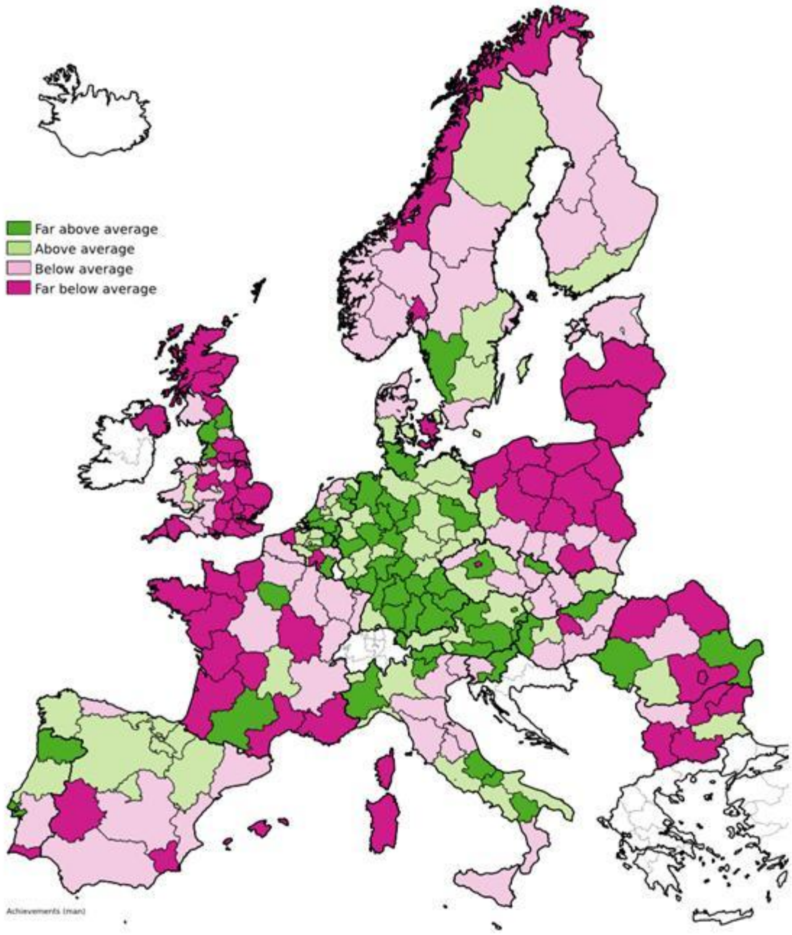



**Figure 5**.
*Scores of regions on 'Innovation achievements' in the transport manufacturing sector*

The relative scores of regions on 'Innovation Achievements' are mapped in figure 5. Many regions that have a high innovation performance are found in Germany, but also a lot of other countries have some regions that perform well, such as France, Italy, Portugal, the Netherlands, Belgium, Austria, Czech Republic, Hungary, United Kingdom Romania and Sweden. The pattern of high performing regions is partly framed by the location pattern of the car manufacturing industry. Several regions that have a weak innovation performance in transport manufacturing are located in France, United Kingdom and Poland. Compared to the map that covers the 'Innovation achievements' in the total transport sector, the strong performing regions in the manufacturing sector on one hand and the weak performing regions on the other hand are less geographically clustered.

### 3.3 Model estimations for the transport service sector

The results for the SEM-analysis for the transport services sector only (see figure 6) are rather different from the manufacturing sector, but resemble the results for the total sector. This can be explained by the fact that there are much more transport service activities (in terms of employment and companies) than transport manufacturing activities. Hence the strengths of relationships found in the services sector put their mark on strengths of relationships observed in the total transport sector.
'Funding' appears to be by far the most important determinant for the 'Innovation achievements' in the services sector. There is a modest role of 'Sector structure' in influencing the 'Innovation achievements', but the quality of the 'Innovation milieu' here also appears not important. The power of the model to explain 'Innovation achievements' is, however, relatively good (comparable to the models for the total and the manufacturing transport sector). Note that the model of the services sector is not able to explain the 'Economic performing' of the transport services. There appears to be a weak and even negative relation between 'Innovation achievements' and 'Economic performing'. The influence of the 'Sector structure' on the 'Economic performing' is also negligible. We cannot explain why these results are different from the models for manufacturing and the total transport sector.

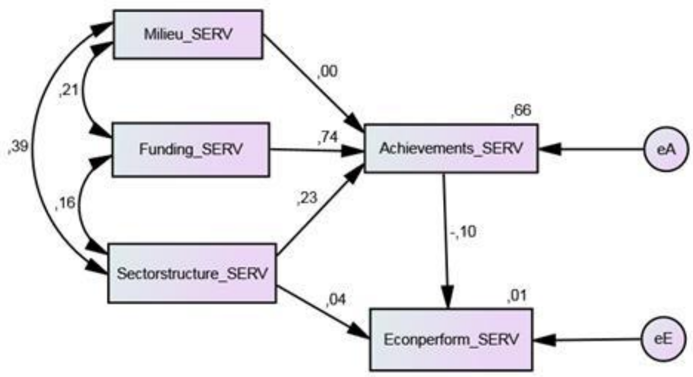



**Figure 6.**
*Parameter estimations for the transport services sector*

Figure 7 illustrates the regional scores on 'Innovation achievements' in the services sector. Regions that are in a similar range of performance are largely clustered geographically. High performing regions are found in Germany, Sweden and Portugal. Italy has for example many moderate performing regions and United Kingdom very modest performing regions. The lowest performing regions are strongly clustered in Spain, Norway, Poland, Slovakia, Hungary, Romania and Bulgaria.

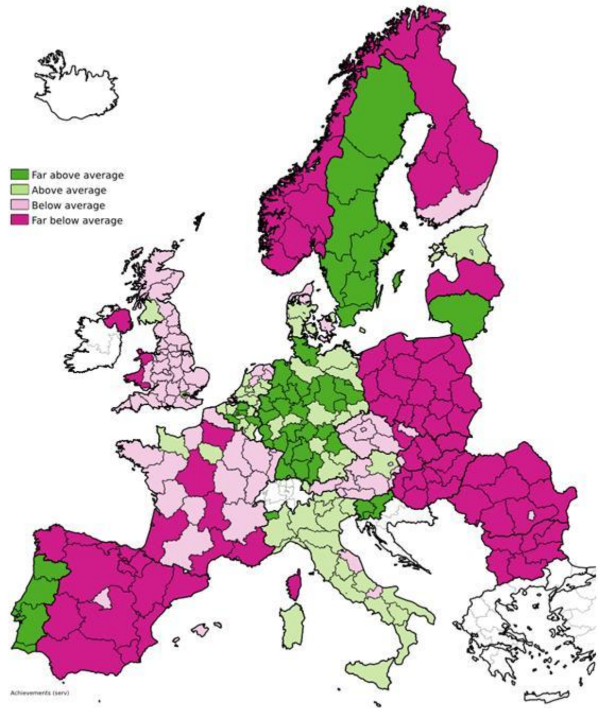



**Figure 7.**
*Scores of regions on 'Innovation achievements' in the transport services sector*

## 4. Conclusions

In this paper we have measured the innovation performance of the transport sector at regional level. We used an approach that enabled a comprehensive measurement that took into account that there is not one unambiguous indicator to measure performance, but rather a set of indicators since performance actually has several dimensions.

A major observation from our analysis is that the relative innovation performance of regions is strongly determined by the definition of the transport sector. First the innovation performance of regions was measured for the total transport sector, i.e. the transport manufacturing and service activities together. This definition is obviously relevant to map the innovation performance of regions in general. However, because manufacturing and service activities are so different, and most likely, also the innovation behaviour of these subsectors, but also because of large differences between regions in the relative size of the transport manufacturing and service sector it made sense to also look at the innovation performance of these subsectors separately.

As far as the results for the total transport sector are concerned, the high performing regions were predominantly found in Germany and Sweden, while the low performing regions were located in Finland, Poland, Romania, Bulgaria, Hungary, Spain and UK. Since the transport services sector is much greater (in terms of employment and number of companies) than the manufacturing sector, the scores of regions in the services sector resemble the scores for the total transport sector. The map of innovation performances of regions in the manufacturing sector gave a more diversified pattern, i.e. more regional variation within the different countries.

In order to explore explanations for the transport innovation performance of regions and their relative performances a Structural Equation Model (SEM) was used. From general innovation theory relevant factors to explain innovation performance were related to the innovation achievements, i.e. funding, innovation milieu and transport sector structure. Several indicators were used to enable a comprehensive measurement of these factors.

The results of this SEM-analysis indicate a relative great importance of funding possibilities for the innovation achievements. This observation was valid for both the analysis for the total transport sector and for the analyses of the manufacturing and service sectors. However, the role of funding appears to be even more important in the service than manufacturing sector. In general the results suggest that the availability and possibilities for innovation funding are determined at national level rather than regional level, since the scores of regions on funding show hardly variation between regions within countries. The importance of the transport sector structure for innovation achievements is in the manufacturing sector greater than in the service sector. A striking result is that the role of innovation milieu for the innovation achievements appears to be negligible for both the manufacturing and service sector. Since this does not fit to theoretical expectations it gave rise to critical reviews and reformulations in the construction of this factor (i.e. the composition of its indicators), but this did not lead to relevant changes in the outcome.

In this METRIC-project we are faced with the fact that we can only use secondary data, and as the experiences with RIS have shown, there is limited availability of data. In principle we needed comparable data as used in RIS, but even more detailed, because it was needed at sectoral level (i.e. the transport sector). Therefore it took great efforts to obtain regional data for all indicators that were initially proposed to be included in the SEM-model. For several indicators the data appeared too incomplete to keep them in the analysis, while for other indicators, a lot of data imputation (e.g. regionalisation of data) was needed. It is uncertain to what extent this leads to biasedness and has affected our model results. Considering the great emphasis the European Commission places to innovation policy and research it is recommended to extend and improve the data collection to support research in this field.

**Acknowledgements**

The results presented in this paper have received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 605428’.

**References**


European Union 2012. *Innovating for a competitive and resource-efficient transport system*, www.transport-research.info, Luxembourg.

European Union 2013. *Employment in the EU transport sector*, www.transport-research.info, Luxembourg.

Hair, J.F.; Anderson, R.E.; Tatham, R.L.; Black, W.C. 1998. *Multivariate data analysis*, Fifth edition. New Jersey: Prentice Hall.

Hollanders, H.; Derbyshire, J.; Lewney, R.; Tijssen, R.; Tarantola, S.; Rivera Leon, L. 2012. *Regional Innovation Scoreboard 2012, Methodology report*, Brussels.

Hollanders, H.; Es-Sadki, N.; Buligescu; B.; Rivera Leon, L.; Griniece, E.; Roman, L. 2014. *Regional Innovation Scoreboard 2014*, Brussels: European Union, Directorate-General for Enterprise and Industry.

Konings, R.; Louw, E. 2014. *Transport research and innovation performance of regions, D3.2 – report of the METRIC-project*, Delft., 56 p.

# MAPPING OF REGIONAL TRANSPORT RTD FRAMEWORKS IN EUROPE

**Vladislav Maraš[1], Zoran Radmilović[1], Eleni Anoyrkati[2], Sunil Maher[2], Rob Konings[3], Merja Hoppe[4], Martin Winter[4], Ana Condeco[5], Aris Christodoulou[5], Slobodan Mitrović[1]**

[1] *University of Belgrade, Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, 11000 Belgrade, Serbia*

[2] *Coventry University Enterprises Limited, Priory Street, Coventry, CV1 5FB, United Kingdom*

[3] *Technical University of Delft, Stevinweg 1, Delft, 2628 CN, Netherlands*

[4] *Zurich University of Applied Sciences, Institute of Sustainable Development, Technikimstrasse 9, Winterthur, 8401, Switzerland*

[5] *JRC, Joint Research Centre of the European Commission, Rue de la Loi 200, Brussels, 1049, Belgium*

**Abstract:** Transport is a key enabler of economic and social activity, but also the source of environmental concerns and other negative externalities. The efficiency of a transport system affects the costs and environmental impacts of the growing volumes of passengers and freight. According to the White Paper (2011), innovation is essential for the development of a European transport strategy in order to achieve the identified challenges. Therefore, this paper presents the results of a mapping process of regional research and innovation activities across the European transport sector. It is based on the intermediate results of FP7 project METRIC ("Mapping European Regional Transport Research and Innovation Capacities"). Particular attention has been given to the examination of prioritized objectives in R&I infrastructure in different EU countries, with specific emphasis on the area of regional transport research.

The mapping process was based on efforts to collect significant amount of useful indicators and indexes relating to European regions (quantitative data), as well as relevant policies, initiatives, strategies, clusters, actors, etc (qualitative data). The sets of quantitative data are mainly taken from EUROSTAT and Cluster Observatory websites. Qualitative data was determined from numerous other sources, such as relevant web sites, reports, papers, etc. Lessons have been drawn from specific regional cases of transport research and innovation policy governance. In this respect, we elaborated and researched the state of regional research and innovation activities, policies and programmes and their most recent trends in European regions at NUTS2 level.

Furthermore, we also detail the importance of various transport sectors for a selection of NUTS2 regions. The work undertaken also included an analysis on how the priorities of innovation and RTD strategies are formulated, determined what type of innovation is the focus of transport sector, and how this varies across European regions.

**Keywords**: transport sector; research and innovation activities; mapping process; European NUTS2 regions; innovation priorities

## 1. Introduction

Transport is the backbone of the European economy, accounting for about 7 % of GDP and more than 5 % of total employment in the EU (METRIC DoW Annex 1). EC MEMO, 2012 indicate that "increased cooperation and prioritization in transport research and innovation" could provide benefits for Europeans as well as for the European industry. These processes require the development and implementation of 'innovation roadmaps' that would help accelerate the generation of new knowledge as well as the exploitation of existing cutting edge research results. Therefore, new and innovative technologies are essential for reducing Europe's dependence on oil and production of GHGs, maintaining our mobility in a sustainable manner, improving the competitiveness of European transport industry, increasing productivity, preserving the jobs, raising standards of living and supporting economic growth in Europe.

Regions are considered as the key actors of growth models driven by innovations. Their role in innovation systems is rather complex. OECD (2011) shows that: i) regional innovation systems follow varied development paths; ii) heterogeneity can sometimes be more pronounced within countries than between countries; iii) while R&D and patenting are mostly concentrated in key regions in top OECD innovative countries, new regions are emerging as knowledge hubs; iv) regional collaboration and networks are becoming increasingly relevant for innovation; v) firms carry out both technological innovations (new products and processes) and non-technological innovations (such as new business models and organizational methods); and vi) design and creative industries are strongly shaped by regional factors and are vital for regional competitiveness. Therefore, to understand the differences in the status of transport researches in European regions, it is essential to take into account and analyze different type of regions or a selection of regions with different levels and importance given to transport researches. This is also the same with any other type of researches.

This paper presents the results of the METRIC project efforts to map the data on existing RTD activities in European regions. Particular emphases have been put on the transport sector. Efforts in this mapping process have been directed to the collection of as much as possible relevant quantitative and qualitative data. Success in the process of data collection influenced the structure and the content of the intermediate METRIC result. These results refer to the process of building the METRIC database, as well as putting together the relevant deliverables and papers.

Our mapping process was based on efforts to collect significant amount of useful indicators and indexes relating to European regions (quantitative data), as well as relevant policies, initiatives, strategies, clusters, actors, etc (qualitative data). These sets of quantitative data are mainly downloaded from Eurostat and Cluster Observatory websites. Due to the complexity of the data and these databases, all these sets were processed in a way that the data can be easily interpreted at the regional level and, as such, included in the METRIC database - entitled “FP7 METRIC Regional Transport RTD & Innovation Repository”. While building the repository, we have also developed a tool for graphical presentation and elaboration of the data contained in the repository. A user of the tool is not only able to determine relevant data, but also perform basic mathematical operations with the data, export it to Microsoft Excel for possible additional handling of data or download the graphs with the required data. Furthermore, qualitative data were taken from numerous resources, including relevant web sites, reports, papers, etc. Database of such qualitative data on research and innovation in European regions has already been developed by European Commission and is available at the Regional Innovation Monitor Plus web site (http://ec.europa.eu/).

The rest of this paper is organized as follows. Chapter 2 of this paper deals with regional dimension of innovation systems, as well as concepts of national and regional innovation systems. Chapter 3 reveals the significance of transport or transport sectors and relevant research activities for some of the selected regions. This chapter also deals with responsibilities for science, research and innovation policy in European countries. Chapter 4 mutually compares priorities of analyzed regional research and innovation strategies and regional operational programmes with particular focus given to transport sector. Chapter 5 concludes the paper.

## 2. Innovations and regional systems

As recent empirical studies indicate that geographical proximity is a prerequisite for learning and innovation (Autant-Bernard and Massard, 2009), it is obvious that spatial dimension should be of primary interest for regional policy makers. However, Autant-Bernard et al., 2013 also state that even where they exist, the effects of proximity are never exclusive, and interact with effects far more remote nationally and internationally. Thus, the geographical proximity effects have to be combined with other dimensions, such as organizational, institutional or cognitive proximity. The same authors conclude that geographical proximity, per se, is neither a necessary nor a sufficient condition for learning to occur. However, it generally facilitates interactive learning by strengthening the other dimensions of proximity (Boschma, 2005).

In addition, Peri (2005) estimates that the likelihood of patent citations decreases by 75% when crossing the regional border (slightly higher than Jaffe et al.’s (1993) estimations, which indicated a decline between 50 and 60%). Only 20% of average knowledge is learned outside the average region of origin, and only 9% is learnt outside the country of origin. Bottazzi and Peri (2003) explain patent production in European regions by the research performed at different geographical levels: within the region, within regions situated at less than 300 km distance and within regions situated at a distance between 300 and 600 km. According to their results, doubling R&D spending in a region would increase the output of new ideas in other regions within 300 km only by 2–3%, while it would increase the innovation of the region itself by 80–90%. Using French data, Autant-Bernard (2001) produce a weaker parameter of spatial decay. For public research, inter-regional spillovers are four times as small as intra-regional spillovers. For private research, the human capital effects are divided by two when moving from the region to its neighbours. According to Eurostat figures, only 27 of the 260 regions spend the equivalent of over 3% of their GDP on research and development whilst over 40% of the EU’s total R&D expenditure of around EUR 200 billion is generated in these regions.

In general, regions represent political governance systems, below the national but above the local level of public administration. The concepts of innovation system at national level, as well as at regional levels are given in Fig. 1 and Fig. 2. Cooke, 1998 state that regional innovation systems are not being present in many countries. Even if they exist, they may be diverse in nature and dominated by a major industry. Such an industry is usually characterized with its strong regional supplier linkages and connections to public or private research laboratories and higher education. Alternatively, Cooke, 1998 also indicate that, as with many industrial district regions, there may be few direct links to research laboratories but many to various intermediaries and service providers, including regional and local public providers.

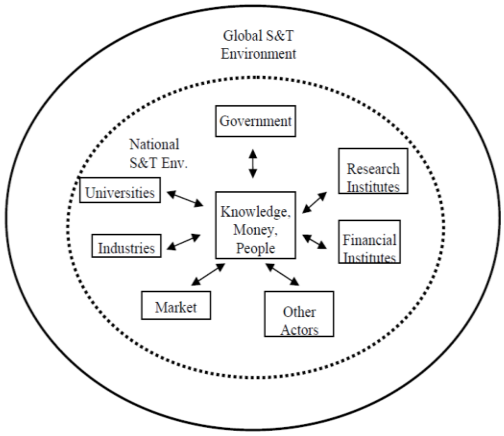



**Fig. 1.**
*National Innovation System*
*Source:http://www.unescap.org/tid/publication/indpub2507_chap3.pdf*

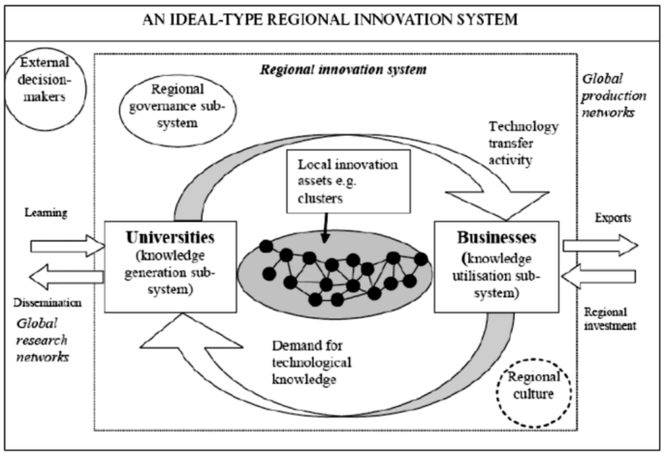



**Fig. 2.**
*An ideal-type Regional Innovation System (RIS)*
*Source: (Engelhardt, 2013)*

### 3. Mapping process and the database

This chapter is based on desk based analysis of relevant literature and on on-going work being undertaken by the European Regional Innovation Monitor (RIM), ERAWATCH, Eurostat, and other relevant databases and reports. RIM is an initiative of the European Commission's Directorate General for Enterprise and Industry, which has the objective to describe and analyze innovation policy trends across EU regions. RIM covers EU-20 Member States: Austria, Belgium, Bulgaria, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Ireland, Italy, the Netherlands, Poland, Portugal, Romania, Slovakia, Spain, Sweden and the United Kingdom. ERAWATCH offer a unique policy monitoring and benchmarking service on the full range of research and innovation policies in the EU and beyond. Eurostat's Community Innovation Survey (CIS) is the main data source for measuring innovation in Europe.

The chapter analyses and synthesizes trends in the regionalization of European research and innovation policy and the implications this has for the interactions between national and regional policy governance based on a review of academic and policy literature. Lessons are drawn from specific regional cases of transport research and innovation policy governance and policy. We also introduce and explain the responsibility for science, research and innovation policy in European countries. In this respect, we performed an elaboration of the state of regional research and innovation policies and programmes with a focus given to transport sector.

#### 3.1. Selected NUTS2 regions

European regions (majority is NUTS2 regions) which have been taken into account are introduced here. Several of these regions are related to the whole country, i.e. specific country constitutes the NUTS2 regions. These countries are identified as follows: Cyprus, Estonia, Latvia, Lithuania, Luxemburg, Malta, Slovenia[1] and Iceland. In addition, Serbia is consisted of five NUTS2 regions, but due to the lack of available data, the country is analyzed at a national level.

Transport as a whole or its related sectors are regarded as very important for the development of some of the considered NUTS2 regions. Detailed analysis, performed in METRIC project, as well as other relevant materials (EC, 2013; TRIP, 2013), revealed the significance of transport or transport sectors and relevant research activities in some of the selected regions. The selected and analyzed NUTS2 regions are the following[2]: Austria (Lower Austria; Upper Austria-automotive sector), Belgium (Southwest Planning Region), Croatia (Northwestern Croatia), Cyprus, The Czech Republic (Praha), Denmark (Hovedstaden; Zealand- transport and eco-mobility), Estonia, Finland (Etelä-Suomi; Länsi-Suomi), France (Île-de-France-other transport technologies, i.e. other than aeronautics and space and automobile; Rhône-Alpes-other transport technologies and automobiles), Germany (Bayern-NUTS1-automotive industry (Audi, BMW, MAN, Knorr-Bremse); Lower Saxony-NUTS1-automotive sector (80 % of all R&D capacities are concentrated in the automotive sector), aeronautics, the shipbuilding and the railway-technology), Greece (Attiki), Hungary (Central Hungary), Italy (Lombardy-sustainable mobility (modal integration – road, rail and metro, car pooling, car sharing, road pricing, etc); Lazio-accessibility to environmentally-friendly transport services); Latvia-relatively specialized in transport technologies other than automobiles and aeronautics; Lithuania-other transport technologies (other than automobiles and aeronautics), transport, logistics and e-systems; Luxemburg-space-satellite telecommunications and automobiles, Malta-development of the transport infrastructure; The Netherlands (Noord-Brabant-maintenance

[1] Slovenia is considered as a single region at the NUTS 2 level. Still, for the purposes of the EU cohesion policy, two cohesion regions were introduced in Slovenia Vzhodna Slovenija (Eastern Slovenia) and Zahodna Slovenija (Western Slovenia).
[2] Analyzed regions and relevance of transport or transport sector in these regions are given in parenthesis.

(aircraft, cars, ships); Groningen-transportation, storage and communication), Poland (Mazowieckie; Slaskie; Opolskie); Portugal (Lisbon- automotive, aeronautics, logistics and transportation); Romania (Bucharest-Ilfov), Slovakia (Bratislava), Slovenia (motor vehicles), Spain (Catalonia), Sweden (Stockholm), United Kingdom (London-NUTS1- major national and international transportation hub; West Midlands-NUTS1-automotive, rail and aerospace; Wales-NUTS1), Iceland- transport in sparsely populated areas; traffic in cold climates; air traffic management), Norway (Oslo og Akershus- other transport technologies), Serbia, Switzerland (Espace Mittelland; Zurich).

### 3.2. Research and innovation policy – responsibility

Responsibility for science, research and innovation policy in European countries might be allocated to different levels of policy-making. In some European countries, development of research and innovation policies and programmes is a task of national authorities. In other countries, regions are in charge of setting up frameworks for research and innovation. There are also countries where these responsibilities overlap between national, regional or community levels. This is clearly presented in Table 1. This table also lists relevant documents which constitute a framework for research and innovation activities at national or regional level. Relevant sets of data are obtained from METRIC research efforts and taken from relevant web sites of European Commission, Regional Innovation Monitor and Regional Policy – INFOREGIO.

**Table 1.**
*Responsibility level for R&I policies and strategies in European countries*

| Country | Responsibility level | Relevant documents |
|---|---|---|
| Austria | national level | ✓ 2012 National Reform Programme;<br>✓ National Reform Programme, 2008-2013;<br>✓ Becoming an innovation leader', published in 2011;<br>✓ National Energy Strategy from 2010. |
| Belgium | communities and regions | ✓ The Flemish Science and Innovation policy – Flemish region;<br>✓ "Marshall Plan 2. Green" – Walloon region;<br>✓ Regional Plan for Innovation - Brussels-Capital |
| Bulgaria | national level | ✓ National Strategy of Scientific Research 2020 |
| Croatia | national level | – |
| Cyprus | national level | ✓ Strategic Development Plan for 2007–2013;<br>✓ Revised National Reform Programme of the Republic of Cyprus |
| Czech Republic | national level | ✓ National Research, Development and Innovation Policy 2009–2015 – RDI |
| Denmark | national level | ✓ Strategic Research Principles and Instruments, 2012;<br>✓ RESEARCH 2015 and RESEARCH 2020 |
| Estonia | national level | ✓ Estonian Research, Development and Innovation Strategy 2007–2013 (RDI);<br>✓ Knowledge-based Estonia |
| Finland | national level | ✓ Research and Innovation Policy Guidelines 2011–2015 |
| France | national programme | ✓ The National Research and Innovation Strategy for period; 2009–2012 |
| | national research programmes with a regional focus | ✓ The competitiveness clusters;<br>✓ The higher education and research poles (PRES);<br>✓ The thematic networks for research (RTRA) |
| Germany | federal level | ✓ High-Tech Strategy 2020 – one of priorities is communication and mobility; |
| | 16 German states – Bundesländer | ✓ **Entrepreneurial Regions" (*Unternehmen Region*);**<br>✓ "Regional Research and Innovation Strategies on Smart Specialisation-RIS3";<br>✓ "Inno-Regio Programme" and its follow-up programmes, "Innovative regional growth centres" and "Interregional alliances" |
| Greece | national level | ✓ Strategic Development Plan for Research, Technology and Innovation; 2007-2013 |
| | research policies at regional level | ✓ developed by most of Greek regions |
| Hungary | national level | ✓ Strategy Programme of the Hungarian Transport Research – Transport Mobility, Transport Economics, Infrastructure, Energy and Environment, Traffic Safety and Security, Intelligent Transport Systems;<br>✓ National Development Concept – 2012-2020, priorities include automotive |

| | | |
|---|---|---|
| | | and logistics sectors;<br>✓ Government’s Mid-Term Science, Technology and Innovation Policy Strategy – 2007-2013;<br>✓ New Hungary Development Plan including the Transport Operational Programme – 2007-2013;<br>✓ New Szechenyi Plan – 2010-2020 |
| | Regional innovation policies per se do not exist | |
| Ireland | national level | ✓ National Development Plan – 2007-2013: priorities in the transport sector are to improve the inter-urban road network and public transport services, as part of an integrated approach to meeting transport needs;<br>✓ Strategy for Science, Technology and Innovation – 2006-2013;<br>✓ Innovation Ireland – 2014-2020;<br>✓ Statement of Strategy – 2011-2013: main principles, priorities and goals for transport policy;<br>✓ Smarter Travel Plan 2009–2020:<br>✓ the Framework Policy for Sustainable Transport;<br>✓ National Cycle Policy Framework *Transport 21* |
| Italy | important documents at national level | ✓ National Reform Programme for innovation, Growth and Employment;<br>✓ Triennial National Research Programme 2011–2013:<br>✓ one of the identified thematic priorities for the national research policy is Sustainable mobility |
| | regions have full autonomy in RDTI policy | |
| Latvia | national level | ✓ Guidelines for Development of Science and Technology for 2009-2013 |
| | planning regions have neither the level of responsibility nor the funding capacity to develop their own research policy | |
| Lithuania | national level | ✓ National Innovation Strategy – 2010-2020;<br>✓ National Higher Education and Research Policy (transport is not included in the list of priority research areas) |
| | there are no functional regional research policies or programmes | |
| Luxemburg | national level | ✓ CORE programme 2008-2013; transport research does not have a priority status |
| Malta | national level | ✓ National Strategic Plan for Research and Innovation – 2011–2020 |
| | no regional research programmes | |
| The Netherlands | national level is responsible for the research policy | ✓ National innovation strategy - "Naar de top": logistics is one of nine identified “top sectors”;<br>✓ Strategic Agenda for Higher Education, Research and Science – published on 1 July 2011;<br>✓ National Planning Strategy up to 2020 and for the period 2020–2030 |
| | provinces develop regional innovation policies | ✓ RAAK (Regional Attention and Action for Knowledge circulation) programme – specific policy programme that focuses on regional development |
| Poland | national level has a significant role to play in research, development and innovation | ✓ National Programme for Scientific and Research & Development (R&D) Activities<br>✓ National Development Strategy for 2007–2015<br>✓ National Strategic Reference Frameworks for 2007–2013<br>✓ State Transport Policy for 2006–2025 |
| | regional level | ✓ regions are increasingly playing their role in regional innovation policies in terms of responsibilities and financial means<br>✓ major tasks of regional self-authority are regional roads and transport management and water transportation |
| Portugal | national level | ✓ Portuguese Strategic Programme for Entrepreneurship and Innovation, end of 2011<br>✓ Pluriannual Funding Programme<br>✓ National Strategic Reference Framework 2007-2013 |

| | | |
|---|---|---|
| | regional level | ✓ no scope for regionally defined policies aiming at research and innovation promotion, with Azores and Madeira regions as possible exceptions |
| Romania | central government is responsible for the RTDI programmes | ✓ Research Development and Innovation National Strategy – 2007-2013<br>✓ National Research Development and Innovation Plan II – 2007-2013 (NP II)<br>✓ Reform Action Plan for R&I – adopted in 2011<br>✓ National Strategy for Sustainable Development - Horizons 2013–2020–2030, part: Action Plan for implementing the Sustainable Transport Strategy in the period 2007–2013, and 2020–2030<br>✓ National Strategic Reference Framework 2007-2013 – part: Sectoral Operational Programme-Transport |
| | no regional research programmes | |
| Slovakia | national level | ✓ State S&T Policy up to 2015<br>✓ 2010 New Model of Financing Science, Technology<br>✓ "Minerva 2.0" strategy<br>✓ Innovation Strategy – 2007-2013<br>✓ Innovation Policy – 2011-2013<br>✓ Transport policy up to 2015<br>✓ Strategy of the Development of Transport up 2020 |
| | regional R&D and innovation programmes | ✓ no explicit regional R&D and innovation programmes and policy initiatives, but regional governments approve regional innovation strategies<br>✓ other transport technologies and manufacture of electrical motors are among main technology sectors (particular in automotive sector) |
| Slovenia | national level | ✓ Research and Innovation Strategy of Slovenia 2011-2020 (RISS); National Higher Education Programme 2011-2020 (NHEP); National Research Framework Programme Slovenian Competitiveness 2006–2012 |
| | necessity of regional policy for R&D is not explicitly mentioned in the official strategic documents | |
| Spain | national level | ✓ National Science and Technology Strategy–ENCYT<br>✓ National Scientific Research, Development and Technological Innovation Plan – National R&D&I Plan, Strategic Plan for Infrastructure and Transport – PEIT, 2005–2020<br>✓ The National Strategy for the Electric Vehicle, 2010-2014<br>✓ Energy Saving and Efficiency Strategy Action Plan, 2008–2012 |
| | Spanish regions are designing their new innovation strategies aligned with smart specialization | |
| Sweden | research policy is decided at national level | ✓ national strategy for transport related RDD – Research, Development, Demonstration<br>✓ Multimodal Strategy and Action Plan for the Use of Intelligent Transport Systems up to 2015<br>✓ National Plan for Sweden’s Transport System for 2010–2021<br>✓ Strategy on Sustainable Development of the Baltic Sea, 2010–2015<br>✓ Vehicle industry research programmes – since 1994<br>✓ Social and Economic Significance of the Automobile, BISEK programme<br>✓ Second generation biobased motor fuels<br>✓ Energy efficient reduction of exhaust emissions from vehicles<br>✓ Governing transitions towards Low-Carbon Energy and Transport Systems for 2050 (LETS 2050)<br>✓ Forska&Väx |
| | programmes that are aimed for regional development | ✓ VINNVÄXT programme<br>✓ Regional growth programmes<br>✓ transport technologies, motor vehicles are the areas with intensive patenting |
| United Kingdom | national level | ✓ Science and Innovation Investment Framework – 2004–2014<br>✓ The Innovation and Research Strategy for Growth, published-end of 2011<br>✓ The Future of Transport White Paper<br>✓ Technology Strategy Board Innovation Platforms: Low Carbon Vehicles and |

| | | |
|---|---|---|
| | | Intelligent Transport Systems & Services |
| | regional level | ✓ some of the regions in UK are granted a special status and more competencies towards self-governance<br>✓ automobiles is one of key science and technology areas based on the numbers of scientific publications and patents<br>✓ UK has a world-class reputation in aerospace research |
| Iceland | national level | ✓ Iceland 2020<br>✓ The Icelandic national transport policy, 2003-2014<br>✓ Building on Solid Foundations<br>✓ Science and Technology Policy for Iceland 2010-2012<br>✓ National Transport Plan 2009–2012 |
| | there are no regional research programmes | |
| Norway | national level | ✓ Innovation White Paper "Climate for Research"<br>✓ National Transport plan 2010-2019<br>✓ Intelligent Freight Transport – SMARTRANS, 2007-2013<br>✓ TRANSNOVA, 2009-2011<br>✓ Norwegian Centres of Expertise (NCE) programme, launched in 2006<br>✓ Arena programme, organized in 2002 and evaluated in 2011<br>✓ Freight Transport and Logistics, 2007-2010<br>✓ Sustainable Urban Transport, 2007-2010<br>✓ Risk and safety in transport – RISIT, 2002-2009<br>✓ Climate and Transportation |
| | regional level | ✓ The Research Council's programme for R&D and innovation aims to transfer more responsibility for R&D and innovation activities to the regions (VRI programme-Programme for Regional R&D and Innovation, introduced 2007) |
| Serbia | National research programmes – nation-wide programmes | ✓ National strategy for science and technological development, 2010-2015<br>✓ Strategy of development of rail, road, water, air and intermodal transport in the Republic of Serbia, 2008-2015 |
| | almost no regional RTDI programmes | ✓ research governance is to a limited extent transferred to the Autonomous Province (AP) of Vojvodina ("Basic Directions of Technology Development of AP Vojvodina") |
| Switzerland | national research programmes launched at the federal level | ✓ National Transport Policy "Making Mobility Sustainable and Fully Integrated"<br>✓ "Sustainable Development Strategy" – 2012–2015<br>✓ The Swiss Energy Transports research programme<br>✓ The Accumulators / Supercapacitors research programme |
| | no regional research programmes, neither programmes that are initiated by the federal government (e.g. national research programmes with a regional focus) nor programmes that are started by the cantonal authorities (regional research programmes) | |

In addition, relevant transport research and innovation programmes may be found in two analyzed regions in Germany[3]. These are:

- Bavaria – 12 technology-specific innovation policy programmes; 9 technology-open programmes (end of 2012), among them, four programmes are in the area of mobility and renewable energies: Program "e-mobility"; Demonstration projects for energy use of biomass; Promoting innovative energy technologies and energy efficiency (BayINVENT); Research and development projects in the field of biotechnology and genetic engineering (BayBio);
- Niedersachsen (Lower Saxony) – 13 programmes aiming at the promotion of research, development and innovation primarily for companies, universities and research institutions; 12 of these programmes do not aim at strengthening specific technologies, but the innovation performance of companies in general; one programme, "Aviation funding guidelines", was only aimed at promoting R&D in a field of aviation research.

[3] In Germany, the smallest administrative level is not the "Bundesland" (e.g. Bayern) but the "Regierungsbezirk" (e.g. Oberbayern). Those can / must create partly own research and innovation programs.

## 4. Innovation and RTD priorities

With our efforts, we have also aimed to answer, to some extent, the following questions:

- What objectives are prioritized in R&I infrastructure in different EU countries?
- What type of innovation is the focus of transport sector, and how does this vary across European regions?
- To what extent is transport research and innovation across European regions coordinated with other relevant programmes and strategies (i.e. where are synergies, overlaps and gaps)?

To answer these questions, we mutually compared the priorities of regional innovation strategies (RIS) of analyzed regions and relevant regional operational programmes (ROPs). Activities defined in the innovation strategies may be or, in some cases, are implemented through the projects within the regional operational programmes. Therefore, innovation strategy should be compatible with the regional Operational Programme (e.g. Regional Innovation Strategy for Malopolska Region). Furthermore, priorities given to projects funded within ROP are often based on the specialization fields identified in regional innovation strategies (e.g. ROP for the Śląskie voivodship, ROP for the West Pannon). Obviously, compatibility of regional innovation strategies and regional operational programmes is an important segment of successful strategic planning of R&I activities at regional level. Therefore, in this chapter, particular attention is given to the transport related priorities and relevant compatibility of elaborated strategies and programmes.

Of the 44 analyzed regions, a total of 7 (15.9%) regions (i.e. 6 NUTS1 or NUTS2 regions – Upper Austria, Bavaria, Lower Saxony, Közép-Magyarország, Stockholm, West Midlands, and one country – Estonia) were identified that explicitly define transport research as a research priority in their respective R&I strategy. These transport research related priorities differs among each other and go from the general approaches such as transport, transport logistics, transport technology and mobility, to more specific transport mode or technology like automotive industry, innovative future technology such as aerospace technology or fuel cell technology. These priorities also cover transport infrastructure aspects (development of region's transport infrastructure; effective infrastructure such as roads, railways, airport and ports).

On the other side, transport related priorities are included in the Operational programmes (OPs) of analyzed regions at much higher levels. Out of these 44 analyzed regions, regional OPs are identified for 35 of them. In total, 28 regions (80 %) included transport related issues among the priorities of their OPs. These priorities comprised the following aspects and its related areas:

- **Accessibility** (public transport and transport accessibility; quality of the public transport; regional public transport; improvement of the accessibility of the larger region, assuring housing and transportation to meet the population growth; improvement of co-operation with regions and improvement of co-ordination between different types of public and goods transportation; improvement of regional accessibility and level of well being; improvement of accessibility to environmentally-friendly transport services; development of an effective, integrated, safer, environmentally-friendly, high-quality accessible transport system);
- **Urban issues** (environment and sustainable urban development; strengthening the urban transport network; urban development and secondary transport networks; urban public transport);
- **Traffic safety** (implementation of engineering measures for traffic safety; improvement of safety standards to reduce accident rates and congestion in the TEN-T network);
- **Mobility** (facilitation of the population mobility; increasing sustainable mobility by means of modal integration (road, rail and metro) and by encouraging the widespread use of forms of transport which have a reduced environmental impact (car pooling, car sharing, road pricing, etc.); improvement of mobility in metropolitan areas);
- **Transport networks** (improvement of logistics connections, including ICT connections; stimulation of innovation within and between logistic chains and attracting new logistic businesses; developing combined transport; extending regional and trans-European road links; integrating regional roads into the international road networks; adaption of Trans-European Transport network (TEN-T) to growing traffic intensity; integration of transport and spatial development plans);
- **Transport infrastructure** (specific infrastructures for sustainable growth; improvement of infrastructure in the areas of rail connections, transport in urban areas and intermodal integration; improvement of technical parameters of transport infrastructure; development of regional water transport infrastructure (improvement of utilization potential of seaports; improvements of ports and its motorway, including actions for a better traffic management)**;** development of land transport infrastructure; development of infrastructure capacities of international airports; support to the development of cycling tracks; connections of the regional transport network with the national and international networks, in particular with the TEN-T; improvement of the regional, as well as local transport infrastructure; construction of an intermodal railway station at an airport; widening of motorways and reserving lanes for public transport; sustainable transport and strategic infrastructure);
- **Energy savings / renewable energy** (support to the promotion of energy savings and renewable energies; supporting energy and electricity projects; promotion of more efficient energy consumption; maintenance (aircraft, cars, ships));

- **Environmental impact of transport** (encouragement of environmental-friendly transport; application of environmentally-friendly solutions to increase transport accessibility of the region; improve the safety and increase mobility of the region's residents; development of multimodal and environmentally-friendly transport; reduction of the negative impact of road traffic congestion; reduction of atmospheric and noise pollution levels).

In addition, increased levels of R&D activity and knowledge transfer between research institutions and businesses in transport technologies can also be found among priorities of OPs. It is obvious that several areas could not be clearly assigned to one of the transport related aspects. Areas listed under the Accessibility, Urban issues and Mobility are particularly interlinked. The similar case is with Transport Networks and Transport Infrastructure aspects. Environmental impacts are clearly visible in any other aspect.

This analysis obviously indicates the difference in importance given to transport research in R&I Strategies of European regions and to transport and its aspects in Regional Operational Programmes. Therefore, in order to overcome these gaps and by taking into account the importance of transport for European industry, it is clear that transport research and innovation activities should be specifically elaborated and much more incorporated in the future R&I strategies and programmes at regional level in Europe.

## 5. Conclusion

This paper provides a snapshot of innovation capabilities in EU28 and four associated states, i.e. Iceland, Norway, Serbia and Switzerland. Particular emphases were given to the regional perspective of the European transport sector, by making use of diverse data and information sources. It addresses quantitative and qualitative innovation data and different regional indicators and indexes that supplement each other.

The METRIC database, developed within the project, is designed to present the data related to the mapping European regional research & innovation capabilities of NUTS2 regions in EU28 Countries and several non-EU countries (Norway, Switzerland, Iceland). Collected data is stored in the database and could be searched by online application. There are two search modes – basic and advanced. This database and relevant tools aim to assist a number of stakeholders (i.e. regional agencies, etc) in exploring the innovation capacity profile of their (or other) regions. Also, this tool could be used to feed into JRC's i3S platform, assist the EC in getting a snapshot of the innovation for their cohesion fund policy documents, etc.

Further, we analyzed a selection of the programmes and research strategies identified from research work undertaken within the METRIC project. Relevant research and innovation strategies as well as regional operational programmes, approved by the European Commission, have also been elaborated. Transport related priorities in these strategies and programmes have been identified and compared. In this way, we were able to determine what type of innovation is the focus of transport sector, and how this varies across European regions. Significant difference in importance given to transport research in R&I Strategies of European regions and to transport and its aspects in Regional Operational Programmes have been pointed out. Therefore, we may conclude that innovation strategies should take into account or be updated in order to be compatible with the relevant regional Operational Programme, as well as with other regional related documents, such as development strategies, RIS3 requirements, etc.

All these mapping efforts are intended to provide inputs to measurement, comparison and explanation of transport research and the innovation performance of regions. They are also expected to contribute to the process of identification of types of innovative regions and provide insights into appropriate indicators and regional dimensions of transport innovation roadmaps. All these activities are among objectives of the METRIC project.

It is clear from this deliverable and the results of the mapping process of regional RTDI frameworks that transport innovation are, in most cases, not autonomous, but evolve as response to the structures and purposes of societal milieu in EU member states, associated states and their regions in which the transport system operates. Further analysis of the transport innovation processes requires models of change that describe the evolution of an entire European transport system, and that are consistent with the special characteristics of the regional transport sector. However, improved data sources are needed to study how transport related research and innovation programmes are coordinated and how they could be interlinked across Europe at national as well as regional levels. It is obviously a direction for future researches in this area. This task will require a high level of cooperation among research projects, relevant authorities, research organizations and industry.

## Acknowledgements

The results presented in this paper have received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 605428'.

**References**


Autant-Bernard, C. 2001. Science and knowledge flows: evidence from the French case, *Research Policy*, 30: 1069–1078.

Autant-Bernard, C., Massard, N. 2009. *Report on the underlying mechanisms of knowledge diffusion*. IAREG deliverable 4.2b.

Autant-Bernard, C., Fadairo, M., Massard, N. 2013. Knowledge diffusion and innovation policies within the European regions: Challenges based on recent empirical evidence, *Research Policy*, 42: 196-210.

Boschma, R. 2005. Editorial: Role of proximity in interaction and performance: conceptual and empirical challenges, *Regional Studies*, 39: 41–45.

Bottazzi, L., Peri, G. 2003. Innovation and spillovers in regions: evidence from European patent data, *European Economic Review,* 47: 687–710.

Cooke, P. 1998. *Introduction*: *origins of the concept*, in H. Braczyk, P. Cooke and M. Heidenreich (eds) Regional innovation systems, UCL Press, London, pp. 2-25.

Engelhardt, A.K. 2013. *Benchmarking Regional Innovation in the light of the Europe 2020 Agenda*: *The case of Twente*, Master thesis, University of Twente.

European Commission. 2011. *White paper: Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system*.

European Commission. 2012. *MEMO*: *Research and innovation for Europe's future mobility*.

European Commission. 2013. *Innovation performance in EU Member States and Associated countries: Innovation Union progress at country level*. Directorate-General for Research and Innovation, Directorate C – Research and Innovation, Unit C.6. – Economic analysis and indicators.

Jaffe, A.B., Trajtenberg, M., Henderson, R. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations, *Quaterly Journal of Economics* (August), 577–598.

METRIC project. 2013. *Annex* 1 –“*Description of work*”.

OECD. 2011. *OECD Reviews of Regional Innovation*: *Regions and Innovation Policy*. OECD Publishing.

Peri, G. 2005. Determinants of knowledge flows and their effect on innovation, *Review of Economics and Statistics*, 87: 308–322.

TRIP, Transport Research and Innovation Portal. 2013. *Compendium of Transport Research Funding – A Guide to European, International and National Programmes*. Communicating Transport Research and Innovation, European Commission’s Directorate General for Mobility and Transport (DG MOVE).

http://ec.europa.eu/enterprise/policies/innovation/policy/regional-innovation/monitor/.

http://www.unescap.org/tid/publication/indpub2507_chap3.pdf.

# HOW TO IMPLEMENT EUROPEAN REGULATIONS OF HOMOLOGATIONS TO VEHICLE MARKET IN SERBIA, CASE OF EXTERIOR MIRRORS

**Vojislav Tomić[1], Boban Nikolić[2], Saša Milanović[3], Miloš Milošević[4]**
[1,2,3,4]*Faculty of Mechanical Engineering, Center for engines and motor vehicles, Niš*

**Abstract:** Because of vehicles importance and their impact on society, a huge number of regulations are applied on them. One of the important regulations that has been applied many years in the world and recently in the Republic of Serbia represents the control of the vehicles conformity, as a means of vehicle control before putting them into traffic. Control of vehicles conformity is often associated with the control of vehicles homologation of assemblies, devices and parts for all motor vehicles. Through the brief analysis, this paper presents recent implementations of a unique European vehicle homologation system on to M 3 categories of vehicles in Serbia. Motor vehicles type M 3 is designed and constructed primarily for the transport of persons and their baggage (the bus). A special part of the research belongs to M 3 homologation inspection of driver's main exterior mirrors according to the Regulation R 46. Based on the submitted producer documentation, Regulation 46 and on the basis of legal acts of the Republic of Serbia, the ultimate goal is achieved, i.e. we successful implement European regulations of homologations to driver's main exterior mirrors. As a final verification of Regulations implementation, the new vehicle gets the certificate of approval.

**Keywords**: Homologation, vehicle M 3, exterior mirrors

## 1. Introduction

Because of importance of vehicles and their impact on society, a huge number of regulations is applied on them. The common goal of all regulations is to increase the safety of all traffic participants, protect the environment and human health as well as energy savings (Majerczyk, 2011). One of the important regulations that has been applied many years in the world and recently in the Republic of Serbia represents the control of the vehicles conformity, as a means of control of the vehicle before putting them into traffic. Control of vehicles conformity is often associated with the control of vehicles homologation of assemblies, devices and parts for all motor vehicles, with the provisions of Regulation and technical specifications (Ministarstvo za Saobraćaj i infrastrukturu RS, 2014), (Klisura et al., 2011). The homologation is obligatory for the road motor vehicles and trailers covered by the vehicle categorization for homologation pursuant to the provisions of UN ECE regulations [United Nations Economic Commission for Europe, 1999], (Klisura et al., 2011). The first agreement on uniform vehicle homologation was adopted in Geneva on 20.03.1958. This agreements was accepted and ratified by the former SFRY decree on the ratification of 15.04.1962. From then until now, motor vehicle homologation service covers the whole vehicle approval process, including all relevant individual European Community (EC) directives and equivalent Economic Commission for Europe (ECE) automotive regulations. The special Directives and regulations on motor vehicles, their trailers, systems and components regulate the vehicle type-approval in the European Union (European Commission, 2014).
The aim of this paper is to get practical directions how to implement EU-Directives and UNECE-Regulations on vehicles in Serbian traffic system. The complete procedure of the control of vehicles and documentation according to ECE and UN directives in the Republic of Serbia is the responsibility of the Agency for Traffic Safety. As this area is widely applied in all categories of motor vehicles, in this paper we will consider only the vehicles of M3 categories. Motor vehicles type M3 is designed and constructed primarily for the transport of persons and their baggage. A motorbus M3 is a vehicle for the transport of passengers with more than nine seats, including the driver seat (Ministarstvo za Saobraćaj i infrastrukturu RS, 2010). In this paper, emphasis is placed on the control of the part of the assembly of a vehicles, i.e. the inspection control of the mirrors for individually manufactured new vehicles type M3 (the buses). The new EU Directive 2007/38/EC requires that by 1st April 2009, all trucks over 7.5 tones and vehicles over 3.5 tones with the larger cab must comply with the expanded field of vision. In Figure 1 is shown the difference between the new and the old directive. Figure 1 clearly shows how new directive increases safety criteria of viewing by increasing the viewing angle of exterior mirrors compared to the old one.

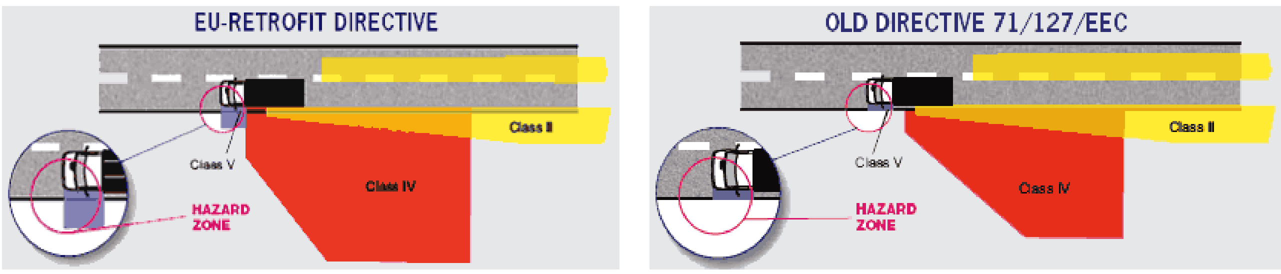



**Figure 1:**
*Comparative surface of exterior mirrors area with old and new directives*

[1] Corresponding author: Vojislav@masfak.ni.ac.rs

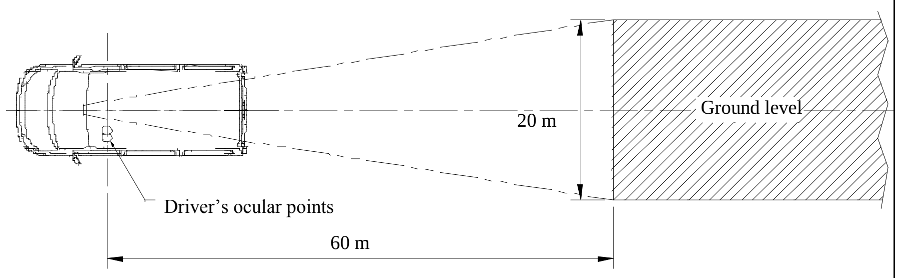
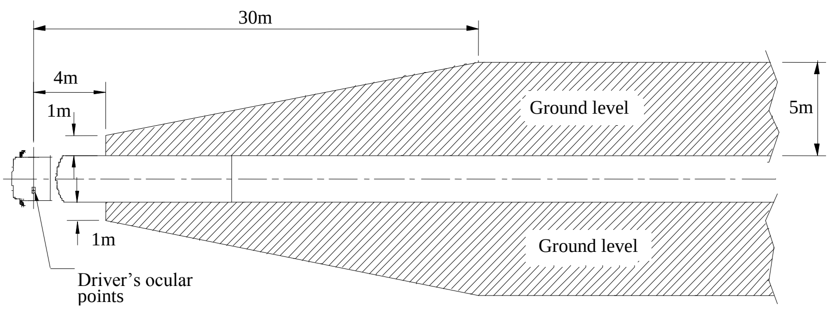
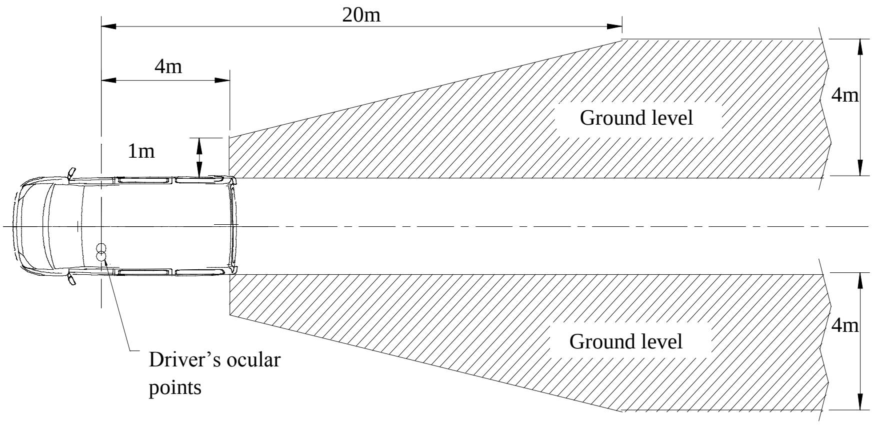
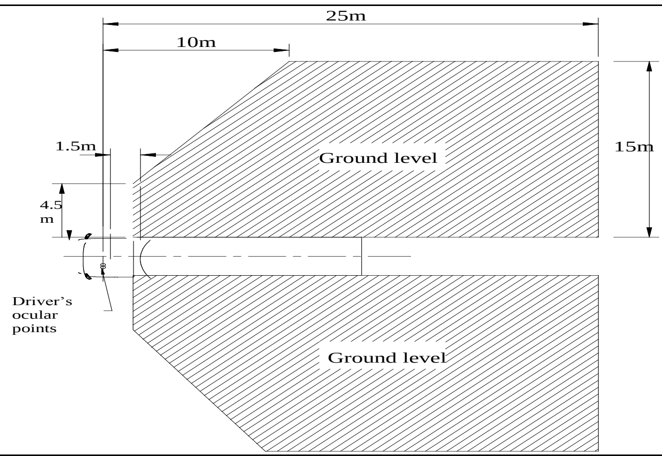
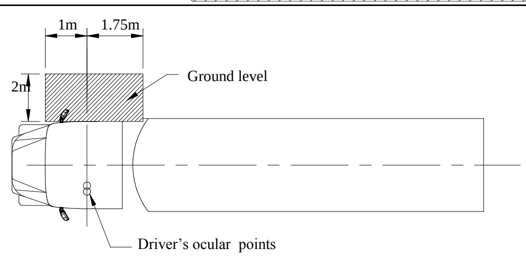
## 2. Installing guidelines for exterior mirrors with individually manufactured vehicle

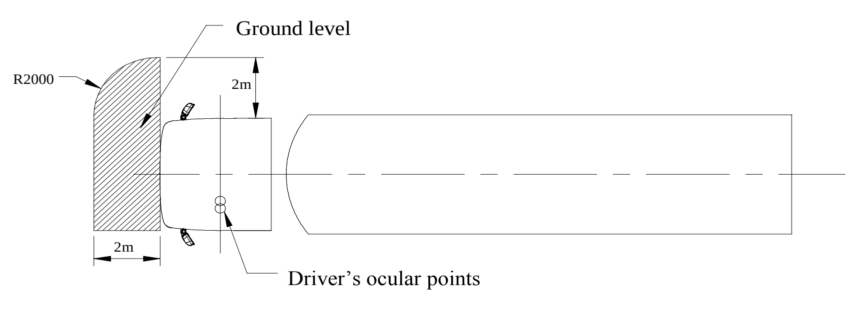
In Europe, total seven ECE (Economic Commission for Europe) regulations and 5 corresponding EC directives deal currently with the structural and seat design for buses and coaches (Mayrhofer, E., 2011), (Hammer, G., 1994). In the Republic of Serbia, this Regulation defines rules of implementations for equipment, parts, assemblies and devices on vehicle. For that reason, control of new vehicles manufactured individually according to the regulations includes the following verification:
1) the dimensions, 2) the weight and the axle load, 3) the stability, 4) the level of stress state of the critical points of the supporting structure, 5) the dynamic characteristics of vehicles (max. velocity, acceleration, climbing), 6) the braking system (ECE Regulations no. 13, 13-H and 78), 7) the noise level (ECE Regulations no. 9, 41, 51 and 63), 8) mounting devices for lighting the road (ECE Regulation no. 48, 53, 74 and 86), 9 ) the tachometer (ECE Regulation No. 39), 10) the vehicles for the transport of dangerous materials (ECE Regulation No. 105), 13) ... 22) installation of safety glazing materials (ECE Regulation No. 43) and 23) installation of mirrors (ECE Regulations no. 46 and 81).
In the flowing part of the paper, we will be focus to the rules of installing mirrors for new individually manufactured vehicles. Installation rules of mirrors on a vehicle types (L - motorcycles, M - passenger cars, N - cargo vehicles and T - special vehicles) is defined by Regulations for division of motor vehicles and technical requirements for vehicles in traffic on the roads (Ministarstvo za Saobraćaj i infrastrukturu RS, 2012). In these Regulations, in the Article 75 there are detailed guidelines by which controller should follow controlling of mirrors on the vehicles that were first time released in traffic. The driver's mirror on a motor vehicle must be installed and designed as a minimum: 1) one driver's mirror on a tractor types T1 and passenger car type M with no more than two seats, as well as on a motorcycle that was manufactured or first time registered after January 1st, 1970. 2) two driver's mirrors on a motorcycle if the motorcycle first registered in the Republic of Serbia after July 1, 2011. and the tractor types T2 to T5 as well as the tourist train; 3) two driver's mirrors on passenger car with four or more seats, where one seat is located inside the body, and the other outside the body on the left side of the vehicle; 4) two driver's mirrors on the bus, passenger car pulling camping trailer, and trolley on both the outside sides of the front part of the vehicle and one driver's mirror inside the body of the vehicle; 5) two driver's mirrors on cargo vehicles from both the outside sides of the front part of the vehicles (Ministarstvo za Saobraćaj i infrastrukturu RS, 2012).
The driver's mirror must be homologated and built according to uniform technical requirements. The driver's mirror must be installed on that way, to enable to the driver monitoring road and traffic behind the vehicle or set of vehicles, when in a vehicle is the maximum number of persons, or when the vehicle is loaded. The driver's mirror must be connected to its support with a joint, so that it can be placed in any position for the purpose of monitoring road and traffic behind the vehicle and to remain in that position at normal shocks during travel. The driver's mirror mounted inside the body of vehicle types M1 must be located at the point where the driver from his seat can adjust it by hand.
For all the details about the installation of the mirrors it is necessary to analyze in detail the ECE Regulation i.e. UN regulation no. 46 and 81, and some of the typical guidelines drawn from these regulations are given later in this paper.

## 3. Installation of exterior mirrors guidelines by UN Regulations

*"Devices for indirect vision"* means devices to observe the traffic area adjacent to the vehicle which cannot be observed by direct vision. These can be conventional mirrors, camera-monitors or other devices able to present information about the indirect field of vision to the driver.
Subject to the type/size of vehicle, there are currently 6 class of mirror (tab 1.) which can be fitted to a commercial vehicle as per the table below. Class I (Interior rear view mirrors), Class II (Main exterior mirrors), Class III (Main exterior mirrors), Class IV (Wide angle mirrors), Class V (Close proximity mirrors) and Class VI (Front mirrors).

**Table 1.**
*Field of vision by class of views*

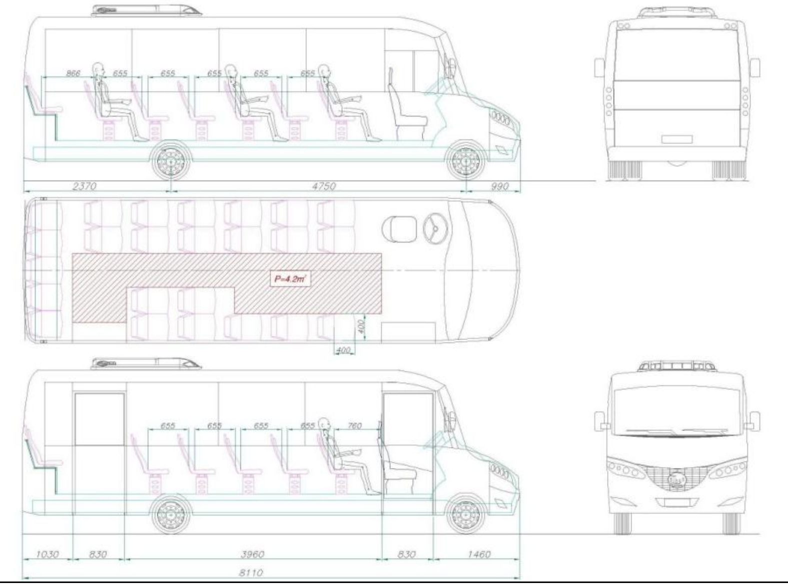
| Field of vision | Class of views, case by mirrors |
|---|---|
| Field of vision of Class I mirror<br><br>The field of vision must be such that the driver can see at least a 20 m wide, flat, horizontal portion of the road centred on the vertical longitudinal median plane of the vehicle and extending from 60 m behind the driver's ocular points to the horizon. |  |
| Field of vision of Class II mirror<br>The field of vision must be such that the driver can see at least a 5 m wide, flat, horizontal portion of the road, which extends from 30 m behind the driver's ocular points to the horizon. In addition, the road must be visible to the driver over a width of 1 m. |  |
| Field of vision of Class III mirrors<br><br>The field of vision must be such that the driver can see at least a 4 m wide, flat, horizontal portion of the road, which extends from 20 m behind the driver's ocular points to the horizon. In addition, the road must be visible to the driver over a width of 1 m, which 4 m behind the vertical plane passing through the driver's ocular points. |  |
| Field of vision of Class IV wide-angle mirrors<br><br>The field of vision must be such that the driver can see at least a 15 m wide of the road, which extends from at least 10 m to 25 m behind the driver's ocular points. In addition, the road must be visible to the driver over a width of 4.5 m, which is point 1.5 m behind the vertical plane passing through the driver's ocular points. |  |
| Field of vision of Class V close-proximity mirror<br><br>The field of vision must be such that the driver can see a flat horizontal portion of the road along the side of the vehicle, bounded by the following vertical planes |  |

| | |
|---|---|
| Field of vision of Class VI front mirror<br><br>The field of vision shall be such that the driver can see at least a flat horizontal portion of the road, which is bounded by: a transverse vertical plane 2000 mm in front of the plane and 2 000 mm outside in the outermost side of the vehicle opposite to the driver's side. |  |

## 4. Control Example for installing mirrors in new vehicles M3 / CA - Class II

For the tested/controlled vehicle whose characteristics are shown in Table 2, the verification of installation of the mirrors was done on the basis of the documentation provided by the manufacturer of the uncompleted vehicle and the documentation submitted by the manufacturer of the completed vehicle. Also control verification of the mirrors implementations and homologations was done in order to determine compliance with the requirements of the regulations on traffic safety and security and the technical characteristics of the vehicle. Tests/controls, which have been performed and that the results presented in this paper, Table 3, were carried out in accordance with the requirements of relevant ECE regulations mentioned above, the Regulation on the examination vehicles, the Regulation on the division of motor vehicles and trailers, and the rules of technical profession by use of the appropriate measuring equipment.

**Table. 2**
*Basic characteristics of the vehicle M3*

| **COMPLETED VEHICLE** | |
|---|---|
| Type of vehicle/Category:<br>BUS M3 / CA - Class II |  |
| Completed vehicle manufacturer:<br>Niš-ekspres a.d. Niš<br>Blagoja Parovića Street no. 1, 18000 Niš | |
| Brand/ Model/Commercial typ:<br>NIBUS / 95.07 | |
| Control:<br>Faculty of Mechanics, University of Niš,<br>Center for engines and motor vehicles | |
| Seats:<br>Number of seats: 25 (twenty-five) | |
| Standing area:<br>Total free area $P_u = 4.20\ m^2$<br>Available free area intended for passengers to stand $P = 3.34\ m^2$<br>Number of standing areas: **13 (thirteen)** | |

In this paper, the only one segment of control of the vehicle is shown, i.e. the control of mirrors. While testing mirrors of the vehicle it was found that they have the homologation marks: e9 03 * 4362 and IV e9 03*4353. On the basis of the homologation and submitted documentation it was found that the manufacturer of mirrors is "Magna Donnelly Espana".

**Table 3.**
*Mirror installation*

| ECE Regulations 46 (description) | | Required values of ECE 46 | | Examination/control results | |
|---|---|---|---|---|---|
| | | 4 m distance | 30 m distance | According to Regulations requirements | Compliant with the manufacturer's documentation of the completed vehicle |
| Field of view[1)] | Regulations 15.2.4.2.1. (left side – outside mirror on the driver's side) | min 1 m | min 5 m | yes | yes |
| | Regulations 15.2.4.2.2. (right side – outside mirror on the passenger side – opposite the driver's) | min 1 m | min 5 m | yes | yes |

[1)] *Field of view refers to the horizontal part of the road, limited to the driver's (passenger's) side with the plane which is parallel with the medial longitudinal vertical plane and passing through the farthest point of the vehicle.*

Based on the documentation submitted by the manufacturer of the completed vehicle (homologation documentation from mirrors manufacturer, schematic view, built-in measures and technical documentation of installed mirrors) and beside on his statement that the installation of the mirrors performed according to ECE 46, it's approached to:

• Checking the data/marks of installed mirrors and their comparison with the data from the homologation documentation of the manufacturer of the mirrors.

• Checking the installation based on the submitted technical documentation.

Based on above aforementioned, the control of the driver's mirrors field of view is performed, and the test results are shown in Figure 2. The plane A, shown in Figure 2, is a vertical plane parallel to the vertical longitudinal plane of vehicle and passes through the farthest points on the right side of the vehicle. The plane B, is the vertical plane parallel to the vertical longitudinal plane of the vehicle and passing through the farthest points on the left side of the vehicle.

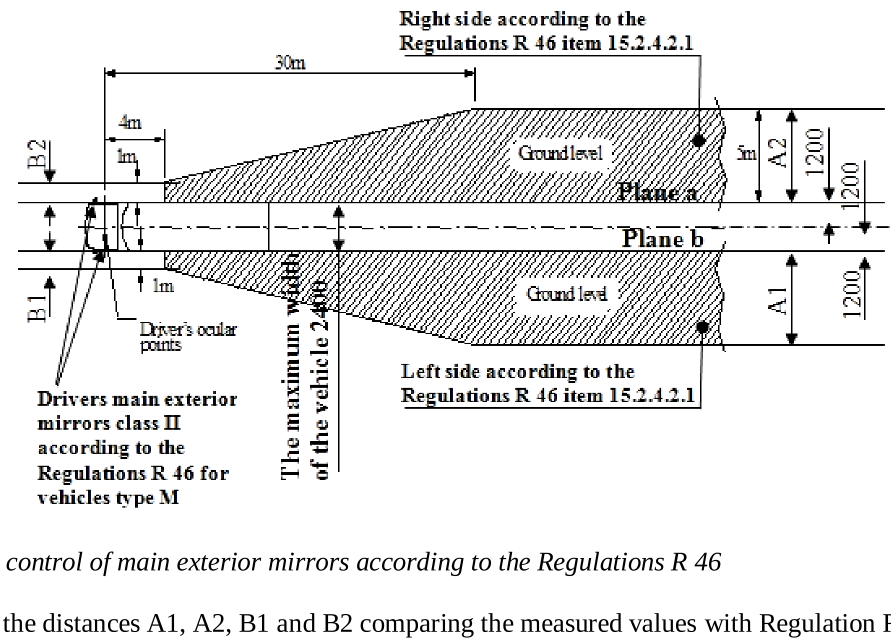



**Figure 2.**
*Inspection results control of main exterior mirrors according to the Regulations R 46*

During control of the distances A1, A2, B1 and B2 comparing the measured values with Regulation R46, it is concluded the following:

The distance A1 on Figure 2 should be at least 5 m according the Regulations R46, but it is measured 12 m;
The distance A2 on Figure 2 should be at least 5 m according the Regulations R46, but it is measured 10 m;
The distance B1 on Figure 2 should be at least 1 m according the Regulations R46, but it is measured 2.2 m;
The distance B2 on Figure 2 should be at least 1 m according the Regulations R46, but it is measured 2 m;
All dimensions satisfy the required values defined in the Regulation R46.

## 5. Conclusion

This research provides an analysis of potentials and obstacles in the process of Serbian implementation of European Union Regulations of Approval for Motor Vehicle Equipment and Parts. The Serbia vehicle homologation project is in the development phase. The control system of institutes for technical vehicle inspections in entire Serbia is an important link in the chain of homologation implementation. The idea of this paper is to give the methodology of control-inspection M 3 type vehicles using the multi-modal method access. A special part of the research belongs to control driver's main exterior mirrors according to the Regulation R46, and to their implementation in real problem. Finally, using these Regulations the aim is reached, a exterior mirrors control-inspection is done, and the certificate of approval is given for NIBUS / 95.07 (Fig 3). The paper results gives practical result of ECE Regulations implementations, because of that this paper provides a scientific contribution to the science literature in this area of research.


**Figure 3:**
*New Nis-Express bus, NIBUS / 95.07*

**References**

European Commission, 2014. Directives and regulations - motor vehicles. Available from Internet:
< http://ec.europa.eu/enterprise/sectors/automotive/documents/directives/motor-vehicles/index_en.htm>

Hammer, G. 1994. Technical Requirements for Coach Safety from the Perspective of the EU and the ECE. *5th Coach Forum of the TÜV Southwest*. 35-40.

Klisura, F. 2010. *Contribution to a Survey of the impact of stations for vehicle technical inspection to the traffic safety in Bosnia i Herzegovina*. Master's Thesis.

Klisura, F.; Mustafić, I.; Barut, M.; Jašarević, S. 2011, The effect of vehicle homologation in B&H to the improvement of vehicle age through a six-month review of technical inspection in FB&H, *15th Trends in the Development of Machinery and Associated Technology*. DOI: http://www.tmt.unze.ba/zbornik/TMT2011/250-TMT11-324.pdf, 1041-1044.

Majerczyk, A. 2011. Technical requirements and methods of testing retrofit LPG systems for motor vehicles in accordance with the UN / ECE 115 regulation, *Journal of KONES Powertrain and Transport*. DOI: http://ilot.edu.pl/kones/2011/3_2011/2011_majerczyk_technical_requirements.pdf, 18(3): 227-235.

Mayrhofer, E.; Steffan, H.; Hoschopf, H. 2014. Enhanced coach and bus occupant safety, *European Commission 5th Framework*. DOI: http://ec.europa.eu/transport/roadsafety_library/publications/ecbos_final_report.pdf , 05-0351.

Ministry of Transport and Infrastructure RS, 2010, The law on road traffic safety, Regulation on homologation of vehicles, spare parts, devices and vehicle equipment, *The Official Gazette of the Republic of Serbia*, № 41/09, 53/10.

Ministarstvo za Saobraćaj i infrastrukturu RS, 2012, Pravilnik o podeli motornih i priključnih vozila i tehničkim uslovima za vozila u saobraćaju na putevima, (Regulation of motor vehicles and trailers, and technical requirements for vehicles in road traffic) Službeni glasnik Republike Srbije br. 40/2012, 102/2012, 19/2013 i 41/2013, Vlada Republike Srbije. Available from Internet:
< http://www.paragraf.rs/propisi/pravilnik_o_podeli_motornih_i_prikljucnih_vozila_i_tehnickim_uslovima_za_vozila_u_saobracaju_na_putevima.html>

Ministarstvo za Saobraćaj i infrastrukturu RS, 2014, Pravilnik o ispitivanju vozila (Regulations of Vehicle Testing), Službeni glasnik Republike Srbije br. 8/2012 i 13/2013, Vlada Republike Srbije. Available from Internet:
< http://www.amss-cmv.co.rs/pdf/pravilnik%20o%20ispitivanju%20vozila%20sa%20izmenam.pdf>

Ministarstvo za Saobraćaj i infrastrukturu RS, 2012, Pravilnik o podeli motornih i priključnih vozila (Regulation of motor vehicles and trailers), 2012, Službeni glasnik Republike Srbije, broj 40/2012, Available from Internet:
< http://www.kbs.rs/userfiles/files/news/p%20vozil%20kbs%202012.pdf>

United Nations Economic Commission for Europe 1999, Available from Internet:
< http://www.unece.org/trans/main/wp29/wp29wgs/wp29gen/wp29classification.html>

# *Human Factors*

# PEDESTRIAN BEHAVIOUR IN URBAN AREA

**Roberto Murrau[1], Francesco Pinna [2]**
*[1,2]Department of Civil Engineering, Enviromental and Architetural DICAAR - University of Cagliari (Italy).*

**ABSTRACT:** The pedestrian behavior is influenced by several factors, including: characteristics of the user, numerousness of group, road infrastructures and environmental factors.
These factors were studied by means the collection of data carried out in the city of Oristano (Sardinia-Italy) on eleven sidewalks and five crosswalks.
The objective was to study the pedestrians behavior, researching the link between independent variables and the dependent variables that, for sidewalks was only the pedestrian speed while for crosswalks were the speed of crossing, the crossing time, the waiting time and the total time.
The regression models were constructed by using ten sidewalks and four crosswalks so ignoring one for each. In the construction, were considered more variables that gradually were excluded on the basis of the p-value.
The models thus detected were deemed significant according to their coefficient of determination and were validated with data from the sidewalk or crosswalk excluded from the construction of the same.
Both for sidewalks that crosswalks were found some reliable models.
The models construction is useful to improve the understanding of the pedestrians behavior and then obtain useful indications to design pedestrian infrastructures with characteristics closer to the real pedestrians behavior.
The present study aims to give greater importance to pedestrians, analyzing how they relate with the urban context in which they live and how it conditions their behavior, so as to design infrastructure in which they feel an integral part and main actors of the urban scene, giving them the respect they deserve and a new sense of belonging to the city in which they live.

**Keywords:** Pedestrian behaviour, sidewalk, crosswalk, urban area.

## 1. THE PEDESTRIAN MOBILITY IN URBAN AREA

In the past, we only move by walking, and for this reason roads were built for mankind and paths were limited to walking distances; with the advent of the car and the increased covered distances, even the city began to grow: streets, have been conceived and designed primarily for motor traffic, to the detriment of pedestrian component. Only recently a turnaround is emerging, with greater consideration of the quality of pedestrian spaces and of vulnerable road users.

## 2. Individualization of the area

The area of study is the city of Oristano (Sardinia-Italy).
Oristano presents considerable problems with regard to pedestrian infrastructures, like insufficient sidewalks width, improper placement of street furniture, inadequate positioning of parking lots, poor lighting; ect. The choice of survey sites was based on roads where pedestrian movement is more considerable, to collect more data.
The downtown roads taken into consideration, all convergent to the main city square, are five, all characterized by activities and services for citizens.
All roads have one lane and one travel direction, with the sidewalk on each side, parkings arranged on one or two sides depending on the road width and without traffic lights. The sidewalks characteristics are:

| Road | Side | Width [m] | Pedestrian Number | Mean Pedestrian speed [m/s] | Parking |
|---|---|---|---|---|---|
| Contini | Dx | 0.95 | 556 | 0.89 | No |
| Contini | Sx | 0.95 | 346 | 0.94 | Si |
| Figoli | Dx | 1.10 | 1023 | 0.98 | Si |
| Figoli | Sx | 0.80 | 944 | 1.03 | No |
| Mazzini | Dx | 1.50 | 599 | 0.99 | Si |
| Mazzini | Sx | 1.45 | 295 | 1.00 | No |
| Mazzini (Large) | Sx | 5.90 | 804 | 0.98 | Si |
| Tharros | Dx | 1.55 | 719 | 0.94 | No |
| Tharros | Sx | 1.35 | 1099 | 0.93 | Si |
| Tirso | Dx | 1.90 | 1360 | 0.94 | Si |
| Tirso | Sx | 1.90 | 1155 | 0.97 | Si |

[2] Corresponding author: fpinna@unica.it

The pedestrian crossings characteristics are:

| Road | Width [m] | Pedestrian Number | Vehicles Number | Mean Pedestrian speed [m/s] | Waiting Time [s] | Parking |
|---|---|---|---|---|---|---|
| Contini | 4.5 | 690 | 1073 | 0.82 | 1.07 | Si |
| Figoli | 6 | 516 | 1066 | 0.87 | 1.65 | Si |
| Mazzini | 9.5 | 641 | 1483 | 0.99 | 0.91 | No |
| Tharros | 5.5 | 679 | 1262 | 0.95 | 0.89 | No |
| Tirso | 7 | 727 | 722 | 0.94 | 0.95 | Si |

The Level of Service, calculated with the HCM method is generally very low.

## 3. Data collection

Reliefs were made on sidewalks located near the square and on the first pedestrian crossing. The data was collected using video recording for each section and in good weather conditions, on weekend, in the shops opening hours. The next step was to collect all the data relating to single pedestrian: for each street were recorded separately the two sidewalks and the pedestrian crossing.

For sidewalks, has been taken into consideration every pedestrian pace to determine the travel time and to calculate speed. For pedestrian crossings was considered every pedestrian passed from one part to another so to determine waiting time, crossing time and to calculate pedestrians speed. Pedestrians age was determined subjectively by the operator. Each pedestrian was cataloged according to age in five categories and was distinguished the individual pedestrians from groups different for the number of people composing them.

## 4. Data processing: the models

It is studied the pedestrians behavior on pedestrian infrastructure, first with respect to sidewalks, then pedestrian crossings. Regression models are constructed using the experimental variables; in particular the dependent variable $Y_i$ is linked to the independent variables $X_i$, by:

$Y_i=ß_0+ß_1X_1+ß_2X_2+\ldots..+ß_mX_m+\mathcal{E}$

Where $ß_0$, $ß_1$, ..., $ß_m$ are the regression coefficients and $\mathcal{E}$ is the error component of the model.

Performing various tests and combining variables among them, is evaluated the level of significance, discarding the variable when the p-value exceeded 5%.

Determined the behavior pedestrian model, the validation is carried out to verify if the model is sufficiently accurate. The validation is carried out by checking the study data with other not used during the estimation of parameters and calculating the coefficient of determination.

### 4.1. Analysis of variables

The dependent variables describe user behavior and are strongly influenced by the characteristics of pedestrians and infrastructure. It is chosen, for sidewalks, the pedestrian speed and for pedestrian crossings the pedestrian speed, the crossing time, the waiting time and the overall time. Four dependent variables are considered:

- pedestrian speed, determined, for sidewalks, as the time to travel a section of known length and, for pedestrian crossings, as the time to pass from one side of the road to the opposite [m/s];
- crossing time, ie the time taken by pedestrian to switch from one side of the road to the opposite [s];
- waiting time, ie the time taken by pedestrian before to decide to cross [s];
- overall time, sum of waiting time and crossing time [s].

The independent variables that could affect the dependent variables performance, are:

- pedestrian age distinguished with following values: 1 - age class from 0 to 10 years old, 2 - age class from 11 to 18 years old, 3 - age class from 19 to 40 years old, 4 - age class from 41 to 65 years old; 5 - age class over 65 years old;
- group composition, distinguished with the following values: 1 - individual pedestrians; 2 - groups of two pedestrians; 3 - groups of three pedestrians; ect..
- length of walked paths: for sidewalks the length was known while for pedestrian crossings are the roadways width;
- sidewalk width [m];
- parking presence: 0 if absent and 1 if pedestrian is not visible by moving vehicles cause parking.

In addition to these are found other variables such as: pedestrian density, pedestrian space, total pedestrian flow, flow of pedestrians passed outside from sidewalk or pedestrian crossing, number of people belonging to each age class, capacity of the sidewalk, number of obstacles on the path, width, traffic flow on the pedestrian crossing, lighting, function of sidewalk. The variables are calculated for each pedestrian and at intervals of one minute and fifteen minutes.

## 4.2. SIDEWALKS MODELS

### 4.2.1. Relationship between speed and all independent variables

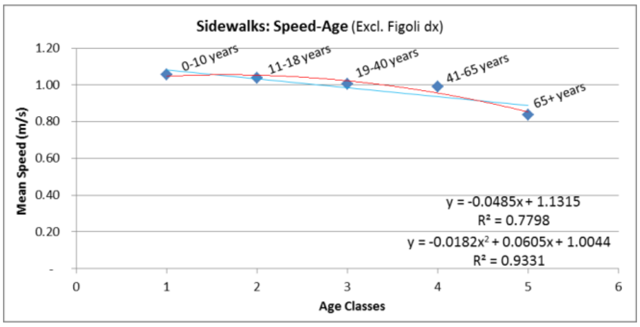
The first tests are performed with data of each passage, and after, data per minute with all variables described above; some of these are later excluded depending on the p-value; appreciable results are not obtained because the $R^2$ coefficient obtained in the construction phase is low, with values between 0,30 and 0,40. For this reason, the tests are performed using data at intervals of 15 minutes.

Even if the variables considered have a p-value less than 5%, it is not been possible to obtain an acceptable $R^2$ coefficient. It happen that, by including a few variables, others lost significance as for example, a few age classes, sidewalk width.

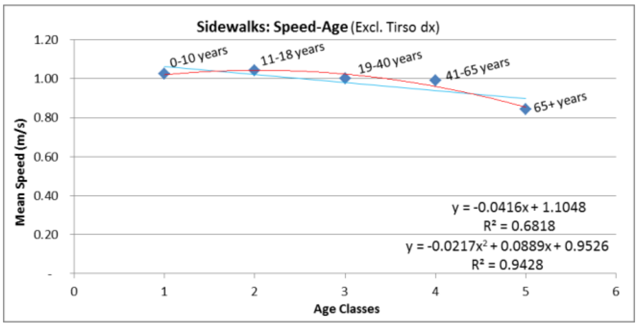
### 4.2.2. Relationship between speed and age classes

It is considered the relationship between number of pedestrians in each age class and speed, then differentiating speed only for each age class, regardless the number of pedestrians. In both cases, it is not possible to obtain a valid model. This is due to the fact that, in the case of non-homogeneous groups (made up of people of different ages), the velocity trend is strongly affected by the slower pedestrians.

To overcome this problem, it is considered the group as a single pedestrian with age equal to the average of its components, but it is created a multitude of data with a low significance level. So, it is considered exclusively the passage of individual pedestrians; in this way we have obtained the average speed of each age class. It is found that the relationship between speed and age classes can be expressed both by a linear model both by a polynomial model (second order). All models have good coefficients of determination in construction and in validation. In validation phase is ruled out a sidewalk at a time, noting that models obtained showed some difference. Graphs 1 and graph 2 show this.




**Graph 1.**



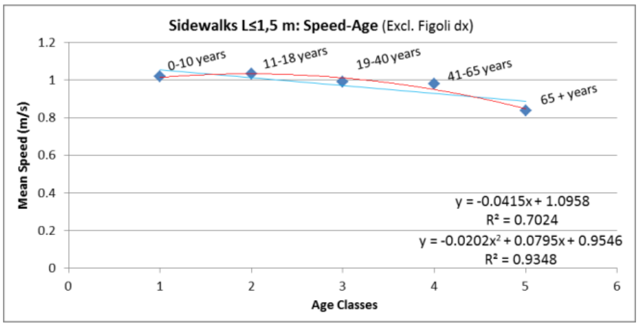

**Graph 2.**

The models validation for the right sidewalk of Figoli road have provided coefficients $R_{Lin}^2$=0,60 and $R_{Pol}^2$=0,87; for the right sidewalk of Tirso road are obtained $R_{Lin}^2$=0,82 and $R_{Pol}^2$=0,92. Differences between coefficients of determination both in construction and both in validation can be attributed to the different width of sidewalks that influence user behavior.

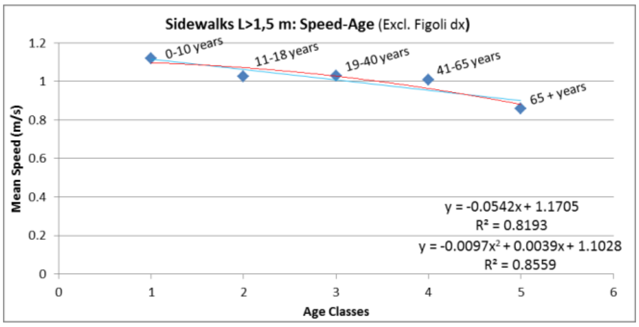
For this reason sidewalks with L≤1,5 m and those with L>1,5 m are distinguished and are constructed linear and polynomial models, subsequently validated with sidewalks excluded in construction. All models have good coefficients of determination both in construction and both in validation. In validation phase was ruled out a sidewalk at a time, noting that models obtained showed regression coefficient with difference smaller. Graph 3 and graph 4 show this.

In validation, in the case of models cited above, are obtained for sidewalks with L≤1,5 m coefficients $R_{Lin}^2$=0,60 and $R_{Pol}^2$=0,90; while for sidewalks with L>1,5 m are obtained $R_{Lin}^2$=0,82 and $R_{Pol}^2$=0,92.

The best models to describe relationship between speed and age classes are those in polynomial form obtained differentiating sidewalks width because they provide the best coefficients of determination in construction and in validation.



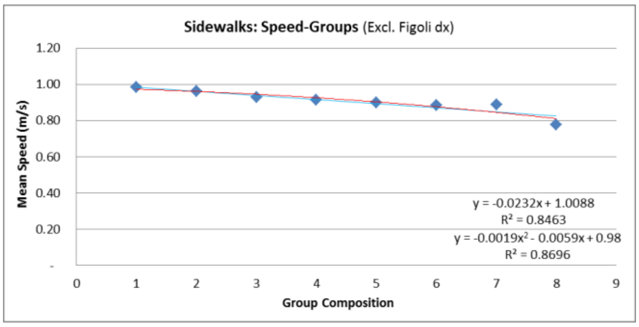

**Graph 3.**



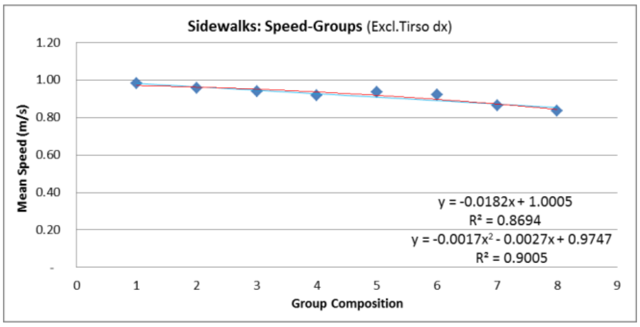

**Graph 4.**

### 4.2.3. Relationship between speed and groups composition

The relationship of the speed compared to groups, considers the number of people belonging to the group without regard to their age. In construction models is ruled out a sidewalks at a time used for subsequent validation and it is found that the relation between the speed and the groups can be expressed both by a linear model, both by a polynomial model (second order). All models had good coefficients of determination in construction, while in validation they presented coefficients sometimes insufficient. Graph 5 and graph 6 show this.

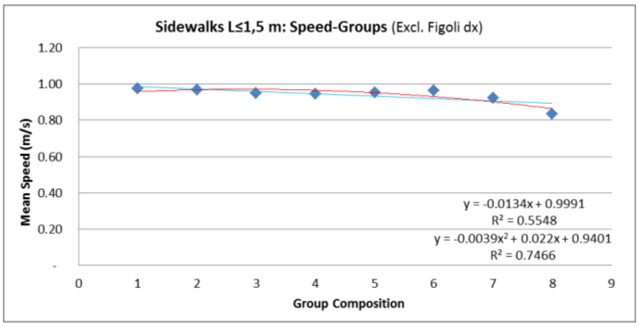



**Graph 5.**

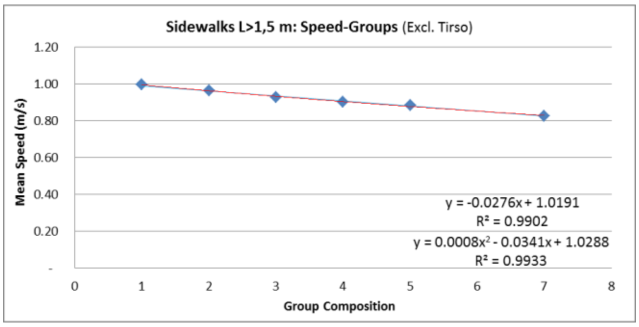



**Graph 6.**

The models validation for the right sidewalk of Figoli road, have provided coefficients $R_{Lin}^2$=0,57 and $R_{Pol}^2$=0,72; while for the right sidewalk of Tirso road are obtained $R_{Lin}^2$=0,53 and $R_{Pol}^2$=0,38. These results, in validation, it are possible to due primarily to the influence of the numerousness group compared to the sidewalk width and also at the presence of groups of more than five pedestrians, very few, so that provide an average speed not statistically significant. As first step, sidewalks with L≤1,5 m and those with L>1,5 m are distinguished and are constructed linear and polynomial models, subsequently validated with sidewalks excluded in construction. In validation phase is ruled out a sidewalk at a time. Coefficients of determination in construction phase of models, for L≤1,5 m decrease while those with L>1,5 m improved considerably. Graph 7 and graph 8 show this.

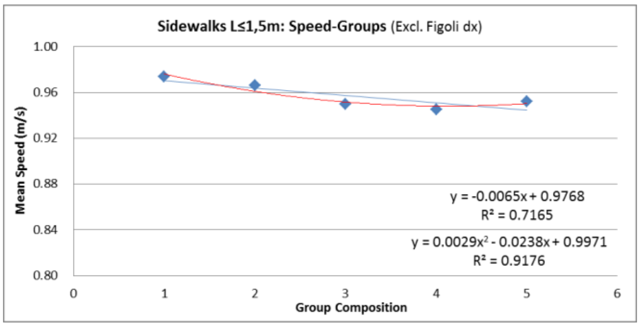



**Graph 7.**

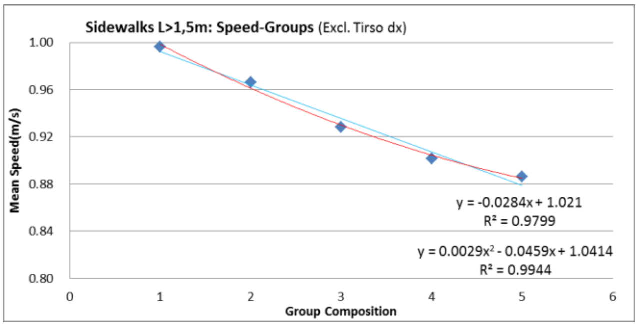



**Graph 8.**

In validation, for models cited above, for sidewalks with L≤1,5 m are obtained $R_{Lin}^2$=0,67 and $R_{Pol}^2$=0,66; while for sidewalks with L>1,5 m are obtained $R_{Lin}^2$=0,60 and $R_{Pol}^2$=0,65.
Later, models are constructed, distinguishing sidewalks width but excluding groups with more than five components cause their low percentage. The construction models (linear and polynomial) are validated with sidewalks excluded in construction. In validation phase was ruled out a sidewalk at a time. The coefficients of determination in construction phase of models, in the case of L≤1,5 m improved while those with L>1,5 m remained almost equal. Graph 9 and graph 10 show this.




**Graph 9.**




**Graph 10.**

The validation, in the case of models cited above, have provided for L≤1,5 m coefficients $R_{Lin}^2$=0,53 and $R_{Pol}^2$=0,77, while for L>1,5 m are obtained $R_{Lin}^2$=0,97 and $R_{Pol}^2$=0,93.
In general it is noted that in models, the average of speed of groups decreases with increasing the component number.
The best models to describe relationship between speed and groups are those obtained differentiating sidewalks width; for sidewalks with L≤1,5 m the best model is in polynomial form because it provides highest coefficient in construction and in validation, while for L>1,5 m both models (linear and polynomial) can be used because provide good coefficients but the linear form is easier.

### 4.3. MODELS FOR PEDESTRIAN CROSSINGS

#### 4.3.1. Relationship between dependent variables and independent variables

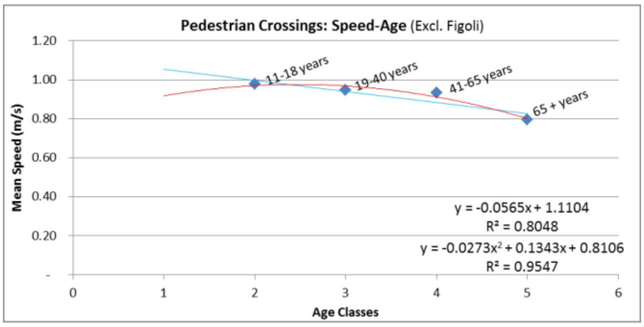
The tests are carried out using data at 15 minute intervals.
Using speed as dependent variable it is found that some variables, initially relevant for the model construction, such as traffic flow, pedestrian flow and age classes were discarded for the low p-value. The resulting model has an $R^2$ acceptable, but in validation doesn't not reveal an appreciable value of the coefficient of determination. This result probably is due to the fact that the speed contains information already present in other variables such as time, so crossing time is used as dependent variable. Models re constructed excluding from each, one hour in two pedestrian crossings at a time, used for model validation. It is obtained:
$Y=1,1397+0,8890X_1+0,5207X_2$ con $R^2=0,96$
where $X_1$ is the length of pedestrian crossing and $X_2$ is the parking presence in the forward direction.
The regression coefficients of the global model are always within the confidence intervals for estimates of regression coefficients of individual models; individual models, in the validation, always provided satisfactory results, so we can assume that the global model provides reliable results. The model is valid for pedestrian crossings with a length included between 5,5 m and 7,5 m.
It is not found a valid model for the dependent variable "waiting time", using the same analysis methods.
This probably derives from the fact that variables identified are not sufficient to describe the waiting time.

#### 4.3.2. Relationship between speed and age classes

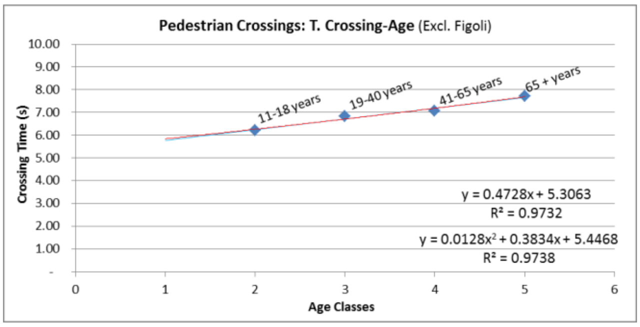
Models between crossing speed and pedestrians age are not reliable, so we differentiate the age in five classes considering only individual pedestrians. Even in this case, indeed, is not possible to find a link with groups having age non-homogeneous. All age classes of individual pedestrians, excluding the class 0-10 years old (no passages of individual pedestrians) are considered calculating the average speed of each class. In models construction is ruled out a pedestrian crossing at a time, used for the subsequent validation; the relationship between speeds and age classes can be expressed both by a linear model both by a polynomial model. All models had good coefficients of determination both in construction both in validation. The graph 11 show this.
In validation are obtained $R_{Lin}^2$=0,85 and $R_{Pol}^2$=0,93.
The best models to describe relationship between speed and age classes is the polynomial form because it provides highest coefficient in construction and in validation.

#### 4.3.3. Relationship between crossing time and age classes

Even for the crossing time is searched a relationship with age classes, only with data relating to individual pedestrians. Such relationship can be expressed both by a linear model both by a polynomial model. In validation phase is ruled out a pedestrian crossing at a time. The graph 12 show this.
In validation, for pedestrian crossing of Figoli road, it is found that $R_{Lin}^2$=0,81 and $R_{Pol}^2$=0,82. The two models (linear and polynomial) overlap, therefore the crossing time increases linearly with age. Therefore the linear form is used because easier.



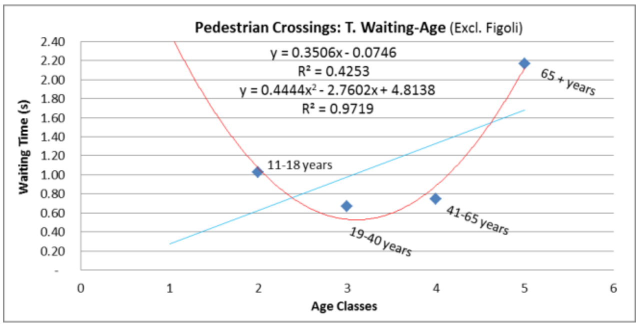

**Graph 11.**

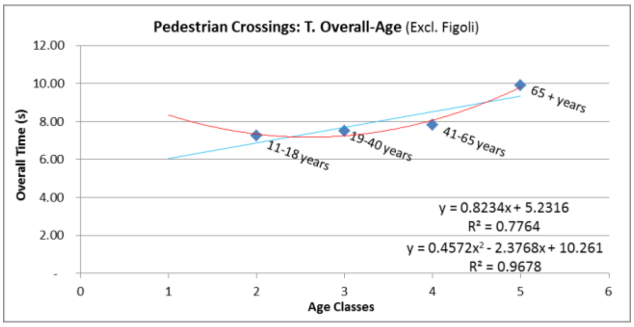



**Graph 12.**

#### 4.3.4. Relationship between waiting time and age classes

Considering the waiting time, in the same manner of previous tests, ie distinguishing the data by age classes of individual pedestrians, is found a polynomial model. In validation phase is ruled out a pedestrian crossing at a time. The graph 13 show this.




**Graph 13.** **Graph 14.**

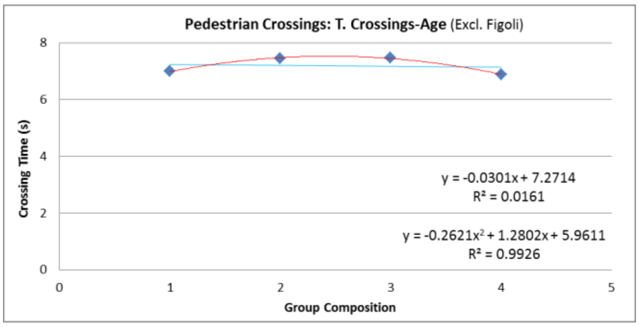
In validation, with the pedestrian crossing of Figoli road, is obtained $R_{Pol}^2$=0,88; similar results are obtained using other roads. Therefore, the polynomial model is suitable to describe the relationship: waiting time-age. It is observed that pedestrians who spend more time to cross are also those who spend more time to make the decision to cross.

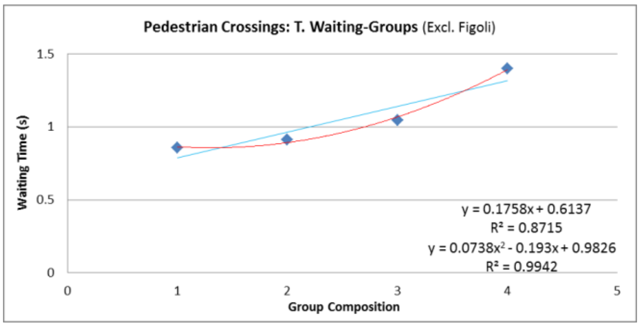
#### 4.3.5. Relationship between overall time and age classes

Studying the relationship between overall time and age classes, two models are obtained: linear and polynomial. In validation phase is ruled out a pedestrian crossing at a time. The graph 14 show this.
In validation, with the pedestrian crossing of Figoli road, are obtained $R_{Lin}^2$=0,77 and $R_{Pol}^2$=0,99.
The polynomial form is used because provides highest coefficients of determination.

#### 4.3.6. Relationship between the dependent variables and pedestrians groups

In the relationship between speed and groups, two models are found: linear and polynomial. Both during construction, provided sufficient coefficients of determination: $R_{Lin}^2$=0,65 and $R_{Pol}^2$=0,75, but in validation phase ruling out a pedestrian crossing at a time, low coefficients are obtained. Therefore models found are not consider satisfactory to describe such relationship.
The relationship between crossing time and groups is described only by the polynomial form while the relationship between waiting time and groups composition can be described both by a linear form both by a polynomial form. In validation phase is ruled out a pedestrian crossing at a time. Graph 15 and graph 16 show this.
In validation, with pedestrian crossing of Figoli road, for the relationship "Crossing Time-Groups" it is obtained $R_{Pol}^2$=0,58, while for the relationship "Waiting Time-Groups" are obtained $R_{Lin}^2$=0,40 and $R_{Pol}^2$=0,25.




**Graph 15.** **Graph 16.**

The relationship between overall time and groups composition can be described only by the polynomial form that provided a coefficient of $R_{Pol}^2$=0,99. In validation however, with the pedestrian crossing of Figoli road, is obtained a coefficient not adequate, $R_{Pol}^2$=0,05.
Therefore for pedestrian crossings, in the case of independent variable "group composition", all models have a high $R^2$ under construction, but only the crossing time in polynomial form has also given positive results in validation.
With regard at the "Overall Time", the polynomial form provides a valid result in construction but not in validation, probably for the interference of waiting time. Similar results are found considering also the others pedestrian crossings.

This result is probably due to one of characteristics of pedestrian crossings behavior that consists in the fact that, while waiting to cross, more pedestrians gather on the edge of the sidewalk, forming a platoon in crossing phase, where it is difficult to determine the waiting time with precision.

## 5. CONCLUSIONS

The research is still under investigation but first results allow some considerations about pedestrian behavior.
In the case of sidewalks, width influences pedestrian behavior so models are distinguished according to the width. To describe the relationship between speed and age classes, the best models for L≤1,5 m and L>1,5 m are those in second order polynomial form because they provide the best coefficients of determination in construction and in validation.
To describe the relationship between speed and groups, the best model for L≤1,5 m is a second order polynomial form that provides the best coefficient of determination both in construction and both in validation while for L>1,5 m both models (linear and polynomial) can be used because both provide good coefficients of determination but the linear model is preferred as simpler.
In general, we observe that speed decreases with increasing age class, similarly, speed decreases with increase of component group. In both cases this reduction of speed is greater on sidewalks with L>1,5 m compared to those with L≤1,5 m, probably because individual pedestrian and groups are less affected by the space available.
For pedestrian crossings, the relationship between speed and age classes is described by a model in second order polynomial form that provides the highest coefficients of determination in construction and in validation.
The relationship between crossing time with age classes, is described both by linear form and both by polynomial form because models have similar coefficients of determination in construction and in validation; linear and polynomial models tend to overlap, so the linear model is considered more suitable to describe the pedestrians behavior as simpler.
The relationship between waiting time and age classes is described by a model in second order polynomial form that provides the highest coefficients in construction and in validation.
Finally the relationship between overall time and age classes is described by a model in second order polynomial form because it provides the highest coefficients of determination in construction and in validation.
In conclusion in relationships between dependent variables and age classes: speed decreases with increase of age classes; crossing time increase with increasing of age classes; waiting time increase considerably with increasing age classes; also overall time consequently increase with increasing age classes; this emphasize that pedestrians who spend more time to cross are also those who spend more time to make the decision to cross.
Analysis for pedestrian crossings as a function of composition groups have shown that, between all dependent variables, only crossing time gives acceptable results.
Relationship between crossing time and composition groups can be expressed only by a model in second order polynomial form with high coefficients of determination in construction phase of model and just sufficient in validation phase.
This means that group size affects the time used by pedestrians to cross.
For pedestrian crossings, finally, is determined a multiple regression model only for the crossing time that is directly proportional to length of pedestrian crossing and parkings presence in forward direction of vehicular current, in fact pedestrian starts to cross positioning himself behind the vehicle in park, which acts as a protective barrier for pedestrian to obtain a better visual contact Vehicle-Pedestrian: this reduces the path characterized by less safety.
Models do not take into account the interference between pedestrians that walk in opposite directions, gender, luggage, ect.. Future researches will take into account these additional variables.

**References**

Buchanan P., Heuman D., *Measuring the benefits of pedestrian improvements*, Fifth Conference on walking in the 21st Century, Copenhagen, Giugno 2004

Giustini M., *Gli anziani e la mobilità pedonale*, Istituto Superiore di Sanità, Convegno Nazionale, 1 Febbraio 2013

Ilango T., Rastogi R., Chandra S., *Behavior and perception of pedestrian walking in groups*, TRB Annual Meeting, 2011

International Transport Forum, *Pedestrian Safety, Urban Space and Health*, OECD, 2011

Jianhong Y., Xiaohong C., Nanjing J., *Impact analysis of human factors on pedestrian traffic characteristics- walking speed revisited*, TRB Annual Meeting, 2011

Kadali B., Vedagiri P., *Modelling pedestrian road crossing behavior under mixed traffic condition*, Paper n.3 ISSN 1825-3997, 2013

NZ Transport Agency, *Guide to undertaking community street reviews*, Giugno 2010

NZ Transport Agency Research, *Predicting walkability*, report 452, Agosto 2011

Obaidat M., Al-Maised H., *A knowledge-based system for pedestrian's roadway crossing behavior through video cameras*, Jordan Journal of Civil Engineering vol.1 n.2, 2007

Rahman K., Noraida A.G., *Analysis of pedestrian free flow walking speed in a least developing country: a factorial design study*, Novembre 2012

Transportation Research Board, *Highway Capacity Manual*, Ottobre 2000

Transportation Research Board, *Improving pedestrian safety at unsignalized crossings*, NCHRP report 562

US Department of Transportation, *Pedsafe: pedestrian safety guide and countermeasure selection system*, publication n.FHWA-SA-04-003, Settembre 2004

# DEVELOPMENT OF OPTIMAL REPAIR TIME FOR THE REPAIR OF MUNICIPAL TRANSIT VEHICLES CLUTCHES

**Solomon Olasunkanmi Odeyale[1], Jubril Temidayo Jolaoso[2]**
*[1,2] Mechanical Engineering Department, Olabisi Onabanjo University, Ibogun Campus, Nigeria*

Abstract: Time study offers the much needed organizational growth in terms of effective labour utilization, methods improvement, wage calculation, task assignment and employee welfare by analyzing maintenance operations into elements that make up the entire repair work. A typical fleet management company is faced with repairing on routine the clutches of its fleet. A time study was conducted by carrying out analysis of a video recording of the repairs of the clutches of a case study municipal transport fleet. The analysis revealed that the current repair process has 15 major repair process, 124 elemental procedures of work of which; 56 were Operation; 17 were Inspection; 29 were Transportation; 0 were Storage and 22 were Delay. The current repair process sees a clutch repair to be completed in 3:34hours. However, a clutch repair process was proposed using methods engineering. The proposed method has 18 major repair elements, 130 elemental procedures of which; 70 is Operation; 18 is Inspection; 27 is Transportation; 4 Storage and 11 Delays. The proposed method was seen to be completed in 3:11hours and far lesser delays and increased inspection which enhance the quality of repair work. A time study of the current repair process revealed that tool and spares procurement occupied the chunk of the delays in the current repair process, it further revealed that the gearbox disassembly process amounts to greater time, energy and man-hour wastage hence, a gantry was proposed to be mounted on the bus walkway to aid the lifting-off of the gearbox. On tool and spares procurement, tools and spares were proposed to be procured firsthand before the actual repair work is commenced. This would reduce the “to and fro” movement to the store. The proposed method also factored ergonomics into its sequence of repairs by allowing rest time for fatigue, personal, standing and basic allowances as against the current process where breaks can be taken at technician’s wish.

**Keywords**: maintenance, clutch repair, fleet management, time study, ergonomics, methods engineering.

## 1. Introduction

Automobile servicing and maintenance companies use a largely manual but semi automatic machines in carrying out their major functions of repairs and overhauling, where most of their work processes are done manually by their workers. Lack of standardized procedure and optimal time for repair works can be said to be responsible for inefficiencies and a shortfall in productivity particularly to servicing/maintenance operations. It affords engineering managers to either indiscriminately specify shorter repair times, hence creating a time constraint making the technicians in a haste to complete the task ahead of them. Thereby making them skip some essential sections of the repair work or make the managers over stretch the time for a repair work and in this case afford the technicians a liberty of time and hence so much time will be expended thereby reducing effectiveness in labor utilization. This research work seeks to determine how time and motion studies affect productivity in the clutch repair process and in addition to determine how the establishment of work and time standards improve the quality and reliability of automobile transmission system. The optimal repair time to be developed will be such-that the constraint of time and wastage shall be eliminated whereby technicians shall be able to carry out a clutch repair in a standardized sequence in an optimal time. The main goal of this research is to determine the optimal repair time for municipal transit vehicle clutches. In pursuing this goal, the following objectives are sought for;

i. To establish accurate repair time standards for municipal transit vehicle clutches.
ii. To eradicate/merge non-productive elemental sequence in the repairs of clutches.
iii. To minimize costs by specifying direct/indirect materials or tools to provide repairs; time and spare.
iv. Improving the work process in terms of service time, number of process and production layout by proposing the new work process and recommending optimal workforce/technicians for clutch repairs.

Code of ethics in engineering practice stressed the importance of adherence to standards in carrying out any engineering tasks; Standards such as the routine/sequence of a repair work or the sequence of manufacturing. As most organizations are striving to be world class, having a time standard for their repair and maintenance work will give them a strong footing to compete and will align them to the latest trends such as the idea of lean manufacturing and methods engineering. The idea of lean manufacturing is to eradicate process(es) that does/do not contribute to the economics and profitability of the firm. Time of course is crucial to either a manufacturing or servicing company as time is used as a reference to set targets; as quality is reigns supreme so is the timeliness of maintenance. This research work however seeks to adopt the idea of lean manufacturing and methods engineering by identifying all elemental procedures in a clutch repair exercise, eliminating time wasting and fruitless elements and developing a standard repair time for such activity to be carried out while taking cognizance of the levels of skills of technicians, facility layout, available tools and other inconsistencies and giving room for allowances incumbent on such activities according to the principles of ergonomics and work study.

[1] Corresponding author: solomon.odeyale@gmail.com

## 2. Literature Review

Today, most multinational businesses and industries are, by necessity, restructuring themselves in order to operate more effectively in an increasingly competitive world. The public service sector is however not an exception. Fundamental tools required to increase productivity include: methods, time study standards and work design. Work measurement is a term widely used in industry which covers several different ways of finding out how long a job should take to complete. Standard times are used by industries for manpower planning, estimating labor costs and calculating the productivity of workers, scheduling, budgeting and designing tasks allocated to people. A great number of research activities, with a base in task analysis, have been developed during last years with the aim of integrating methods and techniques from engineering and ergonomics in order to estimate the time needed to perform a task in a planned but not yet existing product or workplace. For organizations that operate without standards a 60% performance is typical. When time standards are established, performance improves to an average of 85%, a 42% increase (Niebel and Frevailds, 1999). Establishing time standards is a step in the systematic development of new work-centers and the improvements in methods used in existing work-centers. Areas such as planning, control, training, and scheduling are closely related to standards functions. To operate effectively, all of these areas depend on time and operational procedures.

### 2.1 Fundamentals of Work Measurement, Methods, Motion and Time Study

The fundamental purpose of work measurement is to set time standards for a job. Methods engineering is concerned with the selection, development, and documentation of the methods by which work is to be done. It includes the analysis of input and output conditions, assisting in the choice of the processes to be used, operations and work flow analyses, workplace design, assisting in tool and equipment selection and specifications, ergonomic and human factors considerations, workplace layout, motion analysis and standardization, and the establishment of work time standards. A primary concern of methods engineering is the integration of humans and equipment in the work processes and facilities. Motion study has the greatest potential for savings. We can by eliminating the task or combining the task with some task. We can rearrange the elements of work to reduce the work content and we can simplify the operation by moving part. Thus, among the techniques for motion study are as follow:

(i) Process charts
(ii) Flow diagrams
(iii) Operation charts
(iv) Flow process charts
(v) Multiple activity charts

The techniques of time study start with the last motion technique and it shows the close relationship between motion study and time study. The techniques of time study are:

(i) Stopwatch time study
(ii) Expert opinion standards
(iii) Predetermined time standards
(iv) Work sampling time standards

Motion and Time study technique can be used widely for variety of research and in this research it shall be applied in determining optimal repair time for municipal transportation vehicles. Motion and time study helps management determine how much is produced by workers in a specific period of time, therefore making it easier to predict work schedules and output. Motion and Time Study is a scientific method designed by two different people for the same purpose, to increase productivity and reduce time. The two methods evaluate work and try to find ways to improve processes. Now, existing and emerging industries use Motion and Time Study to cushion their quest for improved productivity and input utilization. Industries use it to measure and simplify work in order to reduce costs. (Foster, 2003)
This research shall reveal how the Maintenance and Engineering Management can spearhead the achievement of dramatic productivity impact by implementing work measurement to plan, schedule, manage and control maintenance activities. Methods study is an analysis of ways of doing work. The mnemonic SREDIM (a common-sense heuristic or general problem solving strategy) represents the method study stages**:** S-elect the tasks to study**,** R-ecord the facts about it**,** E-xamine these**,** D-evelop a new method, I-nstall/implement it and M-aintain it. On the other hand, work measurement involves assessing the time a job should take to do. Similar steps are involved as to method study;1. Select the tasks, 2. Record the facts, 3. Analyze them, 4. Calculate basic and standard times for the task 5. Agree the method and its related time.

### 2.3 Review of Past Work in Work study, Method Study and Time Measurement

Abdul and Aliza (2010) carried out a research to improve production capabilities for Small Medium Enterprise (SME) industry. Their research focused on SME, which produce chili sauce. They applied the principle of Time and Motion technique to improve work process at SME, they identified the problems in the production work process and improved it in terms of production time, number of process and production layout by proposing an efficient work process to SME. Their research used systematic observation; process chart and stopwatch time study as research methodology. Muhamad and Law (2002) made a study on work improvement in a car manufacturing company. They identified problems at the metal finish line and proposed a recommendation to improve the efficiency of the current situation. Based on their observations and the collected data, online work-in-progress (WIP) was identified as a major problem and this was said to be caused by insufficient movements due to material handlings and unbalanced workload. Oke (2006) presented a case study in the development and application of a time study model in an aluminum manufacturing plant involved in the production of kettles, frying pans, and cooking pots of diverse categories. He asserted that the three products have similar production processes, then, the production process was broken down into jobs and tasks by the use of differential calculus. Grisselle et al (2002) conducted a research to find the repair time standard for the brakes of transit vehicles using a case study. Their report described in details the procedure followed by the maintenance technicians for changing the brakes of the buses. As typical to all time studies, the major repair sequence was analyzed and every elemental procedure involved in the repair work was listed. They identified about 261 elemental procedures involved in the repair of transit vehicles and further determined the standard time it would take to complete each of the procedures. Eswaramoorthi et al.(2010) applied the principle of lean manufacturing/maintenance to improve the productivity index of an assembly system. Their study proposed an integrated cost model for a typical assembly process to determine cost per part more precisely by considering seven types of "contributing factors". This procedure is performed under different task time conditions to configure the assembly system in terms of cost per piece and to decide the adaptable layout. Mitsunobo Fujita (2008) conducted a study on the statistical method for deriving standard work time of refinishing vehicles using Multiple Linear Regression Analysis. He used a method for deriving standard work times for refinishing damaged vehicles from a relatively small number of time studies, based on a statistically appropriate model. An example application for deriving standard work times from actual refinishing work was used as a case study. His study identified 168 elemental procedures involved in vehicle refinishing.

## 3.0 Methodology

A Clutch is a machine member used to connect the driving shaft to a driven shaft, so that the driven shaft may be started or stopped at will, without stopping the driving shaft. The understanding of Clutch failure modes is critical to effective diagnosis of clutch system and the overall transmission system of the automobile. Hence, this section highlights major clutch failure modes and the interaction of the clutch assembly parts that can lead to clutch failure. Wearing away of the Guide Tube, Spline, Main shaft bearing ; Worn Flywheel Bearing/spigot bearing, ridged heat damage and Flywheel surface; Friction lining dust contamination, Worn/seized Release Arm pivots, to greased components; Leakage of transmission and engine oils; Incorrect alignment of clutch Disc or damage caused during incorrect fitment of transmission; Incorrect or insufficient clutch faulty hydraulic systems adjustment, worn clutch cables and faulty hydraulic systems Haven identified the failure modes of clutches, a right diagnosis can then be easily embarked upon as the major fault modes have been identified. Common clutch problems includes; uneven finger height, clutch cover pressure plate heat damage, scored surface inside release bearing, axial wear of the spline boss in the clutch disc hub assembly, broken damper spring, disc hub splines worn (tapered wear). Others are; Diaphragm clip misplacement, bent drive & normal drive strap, drive plate fracture and clutch cover lining disintegration.

### 3.1 Clutch Repair Process

The repair sequence for a clutch of an automobile can be understood by considering Figure 1 below. The process starts with the disassembly of the propeller shaft/drive shaft from the U joint securing it to the gearbox to removing the gearbox to access the clutch and further to the flywheel.

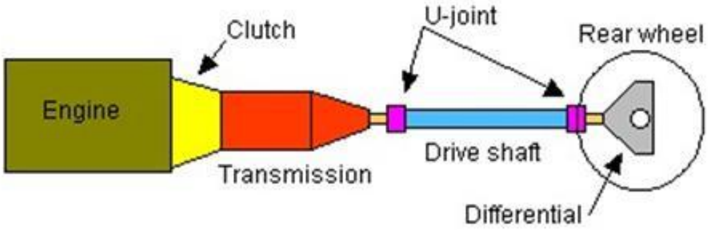



**Fig. 1.**
*Automobile Transmission System*

A case-study clutch repair activity is taped, watched, analyzed and categorized under major processes. These major processes are then reviewed and durations are attached to respective activity. The video coverage was done with the consent of the technicians but so much effort was put into not distracting them. The major processes involved in clutch repairs are listed below. Each major elemental procedure is then further analyzed into sub-elemental procedures making up each major process. An extensive thorough analysis of each elemental procedure is further done to categorize them appropriately. Grisselle et al (2002) classified repair process elements into five categories to capture all elemental procedures involved. The categorization is as follows:

Operation Inspection Transportation Storage Delay

These categorizations are adopted for the clutch repair process under study. The existing sequences of clutch repair processes are; 1) Bus parking 2) Open engine hood 3) Disengage battery terminals 4) Propeller disassembly 5) Gearbox disassembly 6) Clutch disassembly from engine 7) Flywheel disassembly from engine 8) Mount flywheel to engine 9) Mount clutch to engine 10) Mount gearbox to engine input shaft 11) Mount propeller to gearbox 12) Inspect and retighten all bolts 13) Return battery terminals 14) Housekeeping 15) Examine clutch pedal and test-drive

**Table 1**
*Summary of the Current Clutch Repair Process*

| S/N | Major repair procedures | TWE* | Operation | Inspection | Transportation | Storage | Delay | TNE* |
|---|---|---|---|---|---|---|---|---|
| 1 | Bus parking | 465 | 1 | 3 | 1 | 0 | 1 | 6 |
| 2 | Open engine hood | 605 | 1 | 0 | 0 | 0 | 1 | 2 |
| 3 | Disengage battery terminals | 23 | 3 | 0 | 1 | 0 | 1 | 5 |
| 4 | Propeller disassembly | 310 | 1 | 0 | 2 | 0 | 1 | 4 |
| 5 | Gearbox disassembly | 2431 | 10 | 0 | 7 | 0 | 9 | 26 |
| 6 | Clutch disassembly from engine | 680 | 2 | 1 | 1 | 0 | 1 | 5 |
| 7 | Flywheel disassembly from engine | 972 | 1 | 2 | 2 | 0 | 1 | 6 |
| 8 | Mount flywheel to engine | 1112 | 2 | 1 | 2 | 0 | 2 | 7 |
| 9 | Mount clutch to engine | 1748 | 10 | 1 | 2 | 0 | 1 | 14 |
| 10 | Mount gearbox to engine input shaft | 2458 | 14 | 2 | 4 | 0 | 3 | 23 |
| 11 | Mount propeller to gearbox | 273 | 3 | 1 | 2 | 0 | 0 | 6 |
| 12 | Inspect and retighten all bolts | 92 | 0 | 4 | 0 | 0 | 0 | 4 |
| 13 | Return battery terminals | 62 | 2 | 1 | 2 | 0 | 0 | 5 |
| 14 | House keeping | 173 | 4 | 0 | 0 | 0 | 1 | 5 |
| 15 | Examine clutch pedal and test-drive | 626 | 2 | 1 | 3 | 0 | 0 | 6 |
| | **Total elemental procedure** | | **56** | **17** | **29** | **0** | **22** | **124** |

*TWE* = Time for the work elements(s)* *TNE* = Total numbers of elements*

**Table 2**
*Summary of Major Repair Process*

| Major repair process | Number of elemental procedure | Time taken(seconds) |
|---|---|---|
| Bus Parking | 6 | 465 |
| Open Engine Hood | 2 | 605 |
| Disengage Battery Terminals | 5 | 23 |
| Propeller Disassembly | 4 | 310 |
| Gearbox Disassembly | 26 | 2431 |
| Clutch Disassembly From Engine | 5 | 680 |
| Flywheel Disassembly From Engine | 6 | 972 |
| Mount Flywheel To Engine | 7 | 1112 |
| Mount Clutch To Engine | 14 | 1748 |
| Mount Gearbox To Engine Input Shaft | 23 | 2458 |
| Mount Propeller To Gearbox | 6 | 273 |
| Inspect And Retighten All Bolts | 4 | 92 |
| Return Battery Terminals | 5 | 62 |
| House Keeping | 5 | 173 |
| Examine Clutch Pedal And Test-Drive | 6 | 626 |
| **Total Elemental Procedure** | **124** | |
| **Total Time In Hours** | **12030 sec** | |
| **Total Time In Hours** | **3.34hrs** | |

**4.0 Results and Discussion**

**4.1 Results of Time Study of Proposed Clutch Repair Process**

Having carried out a time study of the current clutch repair process in existence in the case study, this paper proposes an optimal and a more effective and efficient clutch repair process. The proposed repair process is based on expert opinions and recommendations as well as by an informed understanding of the nature of the repair work in consonance with time study principles whilst factoring ergonomics to play a decisive role. As against 15 major procedures identified in the current repair process, the proposed method identifies the following repair processes; 1) Parking the bus 2) Transmission inspection & troubleshooting 3) Open engine hood 4) Tool procurement 5) Spares procurement 6) Electrical isolation and cable disconnection 7) Propeller disassembly 8) Gearbox disassembly 9) Clutch disassembly from engine 10) Flywheel disassembly from engine 11) Mounting flywheel to engine 12) Mounting clutch to engine 13) Mounting gearbox to engine input shaft 14) Mounting propeller to gearbox 15) Reassembly inspection 16) Electrical cables reconnection 17) Housekeeping 18) Examining transmission and test-driving. Analytical analysis and the time-study of the elemental procedures of respective major repair elements are further presented in the subsequent sections of this research.

**Table 3**
*Time-study summary of the Proposed Clutch Repair Process*

| S/N | Major repair procedures | TWE* | Operation | Inspection | Transportation | Storage | Delay | TNE* |
|---|---|---|---|---|---|---|---|---|
| 1 | Parking the bus | 225 | 1 | 0 | 1 | 0 | 0 | 2 |
| 2 | Transmission inspection & troubleshooting | 525 | 0 | 4 | 0 | 0 | 0 | 4 |
| 3 | Open engine hood | 5 | 1 | 0 | 0 | 0 | 1 | 2 |
| 4 | Tool procurement | 660 | 0 | 0 | 1 | 3 | 0 | 4 |
| 5 | Spares procurement | 600 | 0 | 0 | 2 | 1 | 0 | 3 |
| 6 | Electrical isolation and cable disconnection | 25 | 3 | 0 | 1 | 0 | 1 | 5 |
| 7 | Propeller disassembly | 310 | 2 | 0 | 1 | 0 | 1 | 4 |
| 8 | Gearbox disassembly | 1371 | 15 | 1 | 6 | 0 | 3 | 25 |
| 9 | Clutch disassembly from engine | 360 | 2 | 1 | 1 | 0 | 1 | 5 |
| 10 | Flywheel disassembly from engine | 785 | 1 | 2 | 1 | 0 | 1 | 5 |
| 11 | Mounting flywheel to engine | 335 | 4 | 1 | 0 | 0 | 0 | 5 |
| 12 | Mounting clutch to engine | 942 | 12 | 1 | 1 | 0 | 1 | 15 |
| 13 | Mounting gearbox to engine input shaft | 1254 | 19 | 2 | 4 | 0 | 1 | 26 |
| 14 | Mounting propeller to gearbox | 273 | 3 | 1 | 2 | 0 | 0 | 6 |
| 15 | Reassembly inspection | 80 | 0 | 3 | 0 | 0 | 0 | 3 |
| 16 | Electrical cables reconnection | 62 | 3 | 0 | 2 | 0 | 0 | 5 |
| 17 | House keeping | 173 | 3 | 0 | 1 | 0 | 1 | 5 |
| 18 | Examining transmission and test-driving | 650 | 1 | 2 | 3 | 0 | 0 | 6 |
| | **Total** | **8635** | **70** | **18** | **27** | **4** | **11** | **130** |

*TWE* = Time for the work elements(s)*   *TNE* = Total numbers of elements*

**Table 4**
*Further Time-Study Summary of the Proposed Process*

| Major repair process | Number of elemental procedure | Time taken(seconds) |
|---|---|---|
| Parking the bus | 2 | 225 |
| Transmission inspection & troubleshooting | 4 | 525 |
| Open engine hood | 2 | 5 |
| Tool procurement | 4 | 660 |
| Spares procurement | 3 | 600 |
| Electrical isolation and cable disconnection | 5 | 25 |
| Propeller disassembly | 4 | 310 |
| Gearbox disassembly | 25 | 1371 |
| Clutch disassembly from engine | 5 | 360 |
| Flywheel disassembly from engine | 5 | 785 |
| Mounting flywheel to engine | 5 | 335 |
| Mounting clutch to engine | 15 | 942 |
| Mounting gearbox to engine input shaft | 26 | 1254 |
| Mounting propeller to gearbox | 6 | 273 |
| Reassembly inspection | 3 | 80 |
| Electrical cables reconnection | 5 | 62 |
| House keeping | 5 | 173 |
| Examining transmission and test-driving | 6 | 650 |
| **Total** | **130** | **8635** |

Grisselle et al, (2002) identified 5 types of allowances during brake repairs and benchmarked a percentage of normal time for respective allowances. Since clutch repair operations exclude intermittent noise, this shall be left out in the computation of the total allowances thereby leaving Personal, Basic, Standing and tediousness allowances accumulating to 13%.

**Table 5**
*Types of allowance and associated percentage of normal time*

| Type Of Allowance | Percentage of Normal Time |
|---|---|
| Personal | 5 |
| Basic Fatigue | 4 |
| Standing | 2 |
| Tediousness | 2 |
| Total | 13% |

Computing total allowance;
Normal time = 8635seconds
Allowance = 13% of Normal time
= 13% * 8635
= 1122.55seconds
Total repair time therefore = Normal repair time + Allowance
= 8635 + 1122.55
= 9757.55sec (3hrs 11minutes)

## 5. Discussion

The results of this time study were presented on Tables 3, 4 and 5. However, contained in Table 6 below is at a glance comparison of the existing and proposed repair processes in terms of number of elemental procedures therein. Number of operations in the proposed repair process (70) is higher than that of the existing repair process (56) indicating that the time allocated for clutch repair in the proposed repair process is chiefly spent on actual productive work thereby ensuring optimal labour utilization on strictly productive tasks. Table 6 further reveals lesser delays and transportation and improved inspection.

**Table 6**
*Proposed and Existing Repair Process Comparison In Terms of Number of Elemental Procedures.*

| Process Categorizations | Number Of Elemental Procedure For Existing Repair Process | Number Of Elemental Procedure For Proposed Repair Process |
|---|---|---|
| Operation | 56 | 70 |
| Inspection | 17 | 18 |
| Transportation | 29 | 27 |
| Storage | 0 | 4 |
| Delay | 22 | 11 |

The current method sees a clutch repair operation to be carried out in 3:34Hrs. Despite time-lead, ergonomics and allowances of the technicians are not well factored. And efficiency and effectiveness of the technicians are not assured throughout the day for the daily task. Major repair elements are 15 in the current method. This reveals that the repair elements are muddled up which can ideally before further simplified. Searching of tools and attachment occupies a whopping 22 individual delays amounting to serious time wastage. Due to frequent procurement of the parts as at when needed the repair time is further stretched. Another factor which stretch the repair time is human habits (e.g., incessant interaction among technicians, foul plays etc). This causes delay and interruption in the work due to frequent visits to the inventory house. On material handling; an enormous amount of time is wasted during the assembly and disassembly of the gearbox due to its weight and constricted space. There are also delays due to interruptions. Technicians are distracted due to difficulty in locating the tools and new parts. The tediousness caused to the technician increases as a result of routine visits to the parts store and the laborious nature of some of the repair elements. The efficiency of the technicians drop after first of such repairs is carried out and their ability to perform at close to the same efficiency in other daily repair works is hampered. Also many of the elements are repeated due to lack of standardization.

In the proposed method; the time study estimates a clutch repair work to be carried out in 3:11hrs. This represents a better process management and reflects the advantage of material handling and tool and part procurements. As well as factoring the allowances incumbent on this kind of repair work. The major repair elements of the proposed method of repair are 18, which reflects a further simplification of the repair process and reduces complexities. As all the required tools are procured firsthand before repair work is initiated, delays are reduced to 11. However, intermittent delays during repair work can not be entirely eliminated owing to human nature and the nature of the job. Also as all the parts required are procured firsthand before repair work is initiated, time is further well managed and wastages associated to to & fro movement for part procurement are eliminated. The human habits noticed in current method are not completely eliminated by this proposed method. However, it is reduced due to less frequent visit to the tool and part procurement store. On material handling, a mini gantry and chain-block is used. This makes the job less-tedious and saves a great deal of time that qualified as delays in the current repair process. The proposed standard method allows the technician to work continuously with minimum distractions and lesser physical energy usage. The proposed method also reduces the tediousness caused to the technician decreases as the operation is standardized and alternative material handling tools are introduced. The efficiency (utilization of technicians) increases as the time study eliminates routine work, standardizes the process and eliminates far-reaching physical energy usage, hence preserving their energy for other daily tasks. Elements are organized in such way that redundancy of operations is minimized.

## 6. Conclusion

The principle of Time study was adopted for a clutch repair task. The task was broken down into elemental procedures, critically reviewing each work elements, determining their duration in time domain, merging unproductive tasks and recommending improvements. The time study of the existing repair process revealed that there are; 15 major processes, a total of 124 elemental procedures of which 56 is operation, 17 is inspection, 29 is transportation and 22 is delays carried out in 3hrs 34mins. The proposed method however, identified that some processes cannot be merged, hence they were simplified into 18 major processes, 130 elemental procedures of which 70 is operation, 18 is inspection, 24 is transportation, 4 is storage and 11 is delay carried out in 3hrs 11mins with allowances of personal, basic fatigue, standing and tediousness all factored. The proposed method saw a reduction in numbers of delays and hence a shortened repair time. The principal finding of this work is that the establishment of accurate and consistent standards improves execution of the procedures required to complete the task of a clutch repair and can be used as a bench mark in maintenance planning, wage calculation, efficiency measurement as well as in the training of technicians. The Time study of the current repair process revealed lack of standardized procedure, no formal allowances and a lot of delays that can be avoided. The current process sees clutch repair to be carried out in 3hrs 34 minutes. However, a proposed repair process for a typical clutch repair operation for the brand of buses understudied can be carried out in 3hrs 11minutes. The time covers all elements of the repair process with the ergonomics of the repair factored to accommodate incumbent allowances. There is a 23minutes difference in duration between the current and the proposed repair process which qualifies for time saving and wage deployment. The proposed repair process is flexible enough to allow technicians to embark on other tasks even without taking further breaks as the method avoids tediousness and far lesser physical energy usage.

**References**

Abdul T.B. ; Aliza A. 2010. An Impact Time Motion Study on Small Medium Enterprise Organization. Universiti Tun Hussein Onn Malaysia. Johor, Malaysia.

Eswaramoorthi M.; Mohanram P.V. ; Prasad P.S 2010. Developing an Effective Strategy to Configure Assembly Systems Using Lean Concepts", International Journal Of Lean Thinking Volume 1, Iss 2.

Foster 2003. "Operational Management – Time and Motion Study.Ppt". <www.freequality.org>

Grisselle C.; Vikramaditya B.; Yasmin P.; Rajesh C. 2002 "Repair Time Standards for Transit Vehicles" Hodson, William K. Maynard's Industrial Engineering Handbook 4th Edition. McGraw-Hill, Inc.

Muhamad Z.M.; Law C.Y. 2002. Work Improvement At A Car Manufacturing Company. Journal Teknologi, 36(A): 55–69 Universiti Teknologi Malaysia.

Niebel B. ; Freivalds A. 1999. Methods Standards & Work Design Time Study. Boston: WCB McGraw Hill.

Oke, S.A. 2006. A Case Study Application Of Time Study Model In An Aluminum Company. Pacific Journal of Science and Technology. 7(2):153-162.

# IMPACT OF TRANSPORT DISADVANTAGE ON EDUCATION OF HIGH SCHOOL POPULATION OF THE CITY OF ZAGREB

**Slaven Gašparović**[1]
*[1] University of Zagreb, Faculty of Science, Department of Geography, Marulićev trg 19/II., Zagreb, Croatia*

**Abstract**: Today, transport disadvantage is very outspread phenomenon which can influence individual's need and opportunities. According to scientific literature, youth are often considered transport disadvantaged part of the society. Their dependance on other people in the aspect of transportation greatly influence their everyday life. In many ways crucial segment of youth's life is education. The aim of this paper is to investigate the influence of transport disadvantage on some segments of education of high school population of the City of Zagreb. The population of interest comprised 826 high school students of high schools in the City of Zagreb. The research was based on travel time, due to the fact that transport disadvantage is a function of accessibility and mobility. Students' attitudes regarding the influence of transportation on education, as well as the correlation between travel time and academic achievement or the pupils' absence from the school, respectively, were investigated in this paper. Research was conducted and analysed using quantitative, as well as qualitative methods. Data were obtained through questionnaire survey and analysed via regression and correlation methods. In order to deepen the research of the specific problems in high school population, additional research was performed in eight focus groups of high school students of the City of Zagreb. Data analysis showed significant correlation between travel time and academic achievements of pupils, as well as correlation between travel time and absence from the school. It was also found that travel time negatively affected attitude of pupils regarding the influence of transportation on their education. These results are strongly supported by states collected through focus group research.

**Keywords:** transport disadvantage, high school population, travel time, education, City of Zagreb

## 1. Introduction

One of the objectives of any country's social policy should be ensuring "social equity". This implies enabling people to perform various life functions and to participate in various life activities, in order to avoid their exclusion. Transport plays a fundamental role in achieving social equity (Minogue, 1998; Foley, 2004). Transport is considered to be one of the fundamental factors in human life, which impacts access to life functions and their successful performance. Within transport, two elements have a critical role: mobility and accessibility. These form the fundamental requirements of today's globalised society (Hoyle and Knowles, 1998). In certain situations, mobility and accessibility can be hindered, limited or even impossible, and as such this jeopardises the use of transport services, and with that the fulfilment of daily needs. If individuals or entire social groups find themselves in such a situation, they will be exposed to the process of transport disadvantage (Gašparović and Jakovčić, 2014). In today's society, automobiles are considered the main means of transport. As such, automobiles are also an important factor that influences the definition of transport disadvantage for people. Not having, or not being able to drive a car (due to legal or other restrictions) is often listed as the main factor leading to transport disadvantage (Murray and Davis, 2001; Clifton and Lucas, 2004). Furthermore, other factors are considered to be the person's financial standing, and physical characteristics (such as gender or disability). In line with this, certain authors have defined a range of groups they deem to be at a transport disadvantage. One of these groups is children and youth (e.g., Murray and Davis, 2001; Stanley and Stanley, 2004; Dodson, et al., 2004; Hurni, 2006; Hurni, 2007). Though some authors list children and youth as a transport disadvantaged social group, it should be noted that not all age groups of youth are equally exposed to the issue of transport disadvantage. High school pupils from the ages of 15 to 18 years are considered the most exposed. Small children and primary school aged children have a lesser need for mobility than high school pupils. High school pupils are almost always required to travel a greater distance to their schools and extracurricular activities, and the locations where they spend their leisure time (i.e. evening outings) than younger children (Hopkins, 2010; Horton, et al., 2011). For that reason, the social group considered in this paper is the group of high school pupils.

School and extracurricular activities represent a crucially important segment in the life of young people in many aspects (attainment of knowledge, socialisation, identity building, etc.) (Hopkins, 2010). The issue of travelling to school is relatively poorly covered in the scientific literature, since the majority of papers examine the topic of travelling to work (van Goeverden and de Boer, 2013). High school pupils travel to work every day, and this therefore poses the question of whether transport impacts the pupil's school activities and, if yes, how it does do. Therefore, the objective of this paper was to determine the influence of transport disadvantage on educational activities of high school pupils in the City of Zagreb. Since transport disadvantage is a function of mobility and accessibility, the focus of the study was placed on determining the influence of time distance on pupil's academic success and on missing classes. In order to further corroborate the results, this paper also provides the views of high school pupils concerning the influence of transport on their school activities.

[1] Corresponding author: slaveng@geog.pmf.hr

## 2. Theoretical framework and research methodology

Considering their inability to drive an automobile, high school pupils will largely have limited mobility and accessibility to certain activities, since they will depend on transport from other persons (e.g. parents, friends...), the use of public transport, or walking or cycling. High school pupils will particularly be affected by the restricted participation in educational activities. High school pupils living at the city periphery or in rural areas will be most affected by the issue of accessibility. Considering that high schools are most often situated near the city centre, this problem may arise due to the distance of the pupil's home from the town centre, and with that the associated travel time and travel costs will be increased. These issues may also reflect on the pupil's academic success. Academic success is an important factor in the life of every pupil, as it forms the precondition for future educational and professional success, and of life overall (Babarović, et al., 2010). Academic success depends on a series of factors, such as cognitive capacities, the personality of the individual and environmental factors (e.g. socioeconomic status of the family, properties of the teaching process, teachers and the school the pupil attends) (Babarović, et al., 2010; Maras and Rodek, 2012). The factor of the school the pupil attends will impact academic success based the location of the school, i.e. the distance of the school from the pupil's home. This influence is often negative. Lin et al. (2013) determined that the distance from the school impacts the academic success of the pupil, as those living further from schools have poorer academic success than those pupils living near the school. The same conclusion was reached by Kamaruddin et al. (2009) and Raychaudhuri et al. (2010). Differences in academic success can also appear between pupils living in the city and those living in rural areas (Owoeye and Yara, 2011) due to the greater distances the pupils from rural areas must travel to school. Also, a greater distance to school can also lead to a lower share of youth taking part in secondary education (i.e., for the United Kingdom, see SEU, 2003; for Australia, see Currie, 2007). Furthermore, the selection of high schools to attend also in part depends on the level of transport services, i.e. the distance to the school. In some countries, a difference is seen in the percentage of pupils attending high school between those youth living at the city periphery and those living near the city centre. Time spent travelling to school can also impact the time available for studying and free time, in the sense that pupils living far from school will have less time to study and less free time than those pupils who spend little time travelling to and from school. The level of services of public transport is often significantly lower in the city periphery than in parts of the city nearer the centre (lower frequency of vehicles, poorer organisation of transport lines, etc.) (Gašparović, 2014). Youth living at the city periphery or outside the city (particularly in more remote rural areas), due to the distance to be crossed and the poorly developed transport services, will have fewer opportunities for participating in extracurricular activities, such as foreign language classes or other extracurricular education activities. With that, these youth may not have these additional education opportunities.

This paper is based on the methods of surveying and interviewing. A survey was conducted in seven high schools in the City of Zagreb in April 2013. A total of 1053 pupils were interviewed, which is just over 3% of the total number of high school pupils of the City of Zagreb in the 2012/2013 academic year (30,970 pupils[2]). After processing the questionnaires, 826 pupils remained (only those pupils having permanent residence in the boundaries of the City of Zagreb and pupils without a driver's licence were included). The questionnaire provided general information on the participants (including their academic success, number of absences from class, and means of travel to school) and the opinions and stance of pupils regarding the accessibility of school activities, and any problems that they might meet regarding the accessibility of those activities. With respect to gender, 429 female (51.9%) and 397 male (48.1%) pupils participated in the survey.

In order to obtain more in-depth information on the issue of transport disadvantage among the high school population, this study also included conducting interviews with focus groups. The focus group research was conducted in December 2013 and January 2014 in two high schools in the City of Zagreb. Pupils in each school were divided into four groups based on their age and gender. Group I was made up of female pupils in years 1 and 2, Group II of male pupils in years 1 and 2, Group III of female pupils in years 3 and 4, and Group IV of male pupils in years 3 and 4. As such, the interviews were held with these eight focus groups (four groups per school). Within each group, the dichotomy of pupils based on their place of residence was pronounced (half of pupils living near the city centre and half living nearer the city periphery). Each group was comprised of between 8 and 10 pupils.

[2] Based on the data of the City Office for Education, Culture and Sport of the City of Zagreb.

In conducting the research, the *Code of Ethics of Research with Children* (2003) was fully abided by, and guided by the idea that ethic issues are an exceptionally important segment in the planning and execution of research, particularly when youth are included in the research (Cohen, et al., 2007). A permit for the research was obtained from the Ministry of Science, Education and Sport of the Republic of Croatia and the principal of each school. Consent for interviewing pupils within the focus groups was also provided by the pupils' parents. The survey questionnaire was anonymous and completely voluntary.

In this paper, sections of the questionnaire and interviews conducted in the focus groups relating to the evaluation of the influence of transport problems on school activities and the difficulties pupils face in trying to resolve these problems. The data collected in the survey were processed using the software package SPSS Statistics 20.0 using the statistical correlation method (Pearson and Spearman correlation coefficient) and regression analysis.

The segment of researching the influence of transport on pupils' academic success and absence from classes was carried out on the pupils of years 2, 3 and 4, with a total of 640 pupils included in the analysis. The reason for this is the fact that the analysis was conducting on the pupil's academic success and absences from class achieved in the previous academic year.[3] Therefore, pupils in year 1 (186 pupils) were excluded from this analysis considering that they attended primary school in the previous year. Primary schools are located in such a way that pupils are required to travel a much lesser distance to school. A large portion of primary school pupils travel to school on their own on foot, and therefore do have absences from class due to transport.

## 3. Influence of transport on the school activities of high school pupils

The selection of the means of travelling to school will depend on many factors, though primarily on the distance between the pupil's home and the school (van Goeverden and de Boer, 2013). In the City of Zagreb, high school pupils most often use public transport (90.8%) or they walk or cycle to school (8.6%), while only 0.6% of pupils are driven to school by someone (Gašparović, 2014). The problems facing pupils in their travels to school are primarily dependent on the means of travel to school, and this will ultimately depend on a series of other different factors (e.g. congestion, frequency of public transport, etc.).

Transport, as the link between the place of origin and the destination, will be the key factor connecting the two components of accessibility. Therefore, the goal was to determine the extent to which transport impacts the school activities of pupils, and to that aim, the pupil's opinions on that issue were examined. Almost two-thirds (65.3%) of the high school pupils in the City of Zagreb believe that transport affects their school activities (Table 1). However, transport was found to relatively rarely influence school activities, considering that 82.2% of high school pupils stated that transport only rarely or occasionally impacts their school activities, with a weak to moderate influence. Just under one-fifth (17.8%) of pupils claim that transport has a more frequent effect (often or almost always) on their school activities, with a high to very high influence.

**Table 1**
*Frequency and strength of the influence of transport on the school activities of high school pupils in the City of Zagreb*

| FREQUENCY/STRENGTH OF THE INFLUENCE | NUMBER OF PUPILS | SHARE (in %) |
|---|---|---|
| INFLUENCE | 539 | 65.3 |
| *rare / poor* | *186* | *34.5* |
| *occasionally / moderate* | *257* | *47.7* |
| *often / high* | *80* | *14.8* |
| *almost always / very high* | *16* | *3.0* |
| NO INFLUENCE | 287 | 34.7 |
| **TOTAL** | **826** | **100** |

*Source: survey questionnaire, 2013*

[3] Academic success from the previous academic year is used in research by, for example, Vranković et al., 2011.

In line with the influence of distance from activities on the daily life of individuals, there also arose the need to investigate the influence of the travel time to school on the stance of the pupil on the influence of transport on their school activities. The results obtained indicate an association between the time of travel to school and the stance of the pupil on the influence of transport on their school activities. It was observed that the stance of pupils on the frequency of influence of transport on the daily life increased with the time necessary for travel to school. An analysis was conducted only on those pupils who expressed the attitude that transport impacts their school activities. The results indicate a slight though statistically significant association between the attitudes on the influence of transport on school activities and the travel of time to school ($\rho = 0.324$; $p < 0.01$).[4] Pupils who spend more time travelling to school believe that transport more often influence their school activities than those pupils who spend less time travelling to school. This justifies the assumption that distance from activities negatively impacts the daily life of individuals, and that pupils who travel longer to school have greater problems with their school activities than those pupils who live near to school. The ways that transport impacts school activities varies, and certain factors are more prominent (Table 2).

**Table 2**
*Ways in which transport impacts the school activities of pupils*

| WAY OF INFLUENCE | NUMBER OF PUPILS | SHARE (in %) |
|---|---|---|
| ABSENCES FROM CLASS | 230 | 42.7 |
| LOSS OF TIME | 146 | 27.0 |
| TRAFFIC CONGESTION | 55 | 10.2 |
| FREQUENCY OF PUBLIC TRANSPORT LINES | 54 | 10.0 |
| WEATHER CONDITIONS (winter) | 32 | 5.9 |
| DISTANCE | 8 | 1.5 |
| EFFECTS ON FATIGUE AND CONCENTRATION | 8 | 1.5 |
| CROWDS IN PUBLIC TRANSPORT VEHICLES | 3 | 0.6 |
| ASSISTANCE IN LEARNING | 3 | 0.6 |
| **TOTAL** | **539** | **100** |

*Source: survey questionnaire, 2013*

The greatest influence of transport on the school activities of pupils is seen in absences (42.7%). The reasons for these absences may vary, though pupils most often expressed traffic congestion and the frequency of public transport lines as the main reasons.

> *Traffic congestion is a problem on the way to school, and then I am late. It is a bigger problem when school is in session in the morning.*
> [Where is the traffic the worst?]
> *On Ljubljanska* (Zagrebačka Avenue) *between Getro and Remiza.*
> (male pupil, 16 years, Rudeš)

For pupils a general problem is the loss of time due to travel (27.0%). Pupils believe that this loss of time is translated into reduced time for learning or rest, and also to increased fatigue and reduced concentration in class (8%).

> *When I have school in the morning, and when I come home, if school is finished at 2 pm, I get home at 3 pm, I need to do some homework, and I can't, I eat something and I'm dead tired, I have to nap for at least two hours, so I'm tired. Then sometimes I study until 1 or 2 in the morning, and I could be done earlier if I was closer.*
> (female pupil, 15 years, Savica)

Traffic congestion is a generator of many problems, and pupils also take this factor into consideration in assessing the impacts on school activities (10.2%) and often independently stress this as a problem. The consequence of traffic jams will be fully associated with the above stated influences: tardiness and additional loss of time due to travel.

[4] Variables relating to the assessment of the frequency of influence of transport on specific activities are coded as follows: 1 – No influence; 2 – Poor influence / rarely influences; 3 – Moderate influence / occasionally influences; 4 – High influence / often influences; 5 – Very high influence / almost always influences

*It affects me because I have to catch an earlier bus. For example, when I have school in the morning, then I leave at 6:45 am instead of 7:15 am, as the traffic is heavy and I am constantly looking at my watch to see if I will make it on time. This way I arrive earlier and have time for a coffee and I know I will make it to school. But that's why I have to leave home earlier.*
[Are you often late if you take the 7:15 am bus?]
*Yes, often.*
(female pupil, 15 years, Miroševec)

The frequency of public transport lines is a problem that some pupils face (10.0%). The sparse frequency of lines, above all bus lines, can affect the organisation of travel to school or can result in tardiness.

*My bus only runs every half hour and I get to school at 7:40 am and I don't know what to do. The next bus would have me arrive at 8:05 am and then I would be late. The buses should run more often.*
[Where do you live in Zagreb?]
*I live in Šestine.*
[And what is the situation when you have school in the afternoon?]
*When I have school until 8 pm, I don't get home until 9 pm because there is no bus.*
(female pupil, 16 years, Šestine)

Transport will depend, to a certain extent, on weather conditions, and therefore for some pupils, that factor will affect their school activities (5.9%). This primarily refers to unfavourable weather conditions that are typical during the winter months, when snow often hinders normal traffic flow. However, rain is also known to slow traffic throughout the city, and therefore rain and slow can cause traffic congestion that then influences traffic flow and, as a result, can lead to possible tardiness of pupils to school.

*If it is snowing, then I have problems getting to school because I live in Šestine.*
[Do you think you would have better grades?]
*No, my grades wouldn't be better, but I would be late less often. I will soon be getting my driver's licence, so then I will drive to school and I won't be late anymore.*
[Do your teachers accept this as a reason for being absent?]
*The teachers don't really accept this, and I have had a few unjustified absences due to transport.*
(male pupil, 18 years, Šestine)

The spatial distance from the public transport stops or the general spatial distance of the pupil's residence from the school is another reason listed by pupils as having an effect on school activities (1.5%). Though overcoming the spatial distances will depend, among other things, on the mode of transport selected by the pupils, its efficacy and other factors previously discussed, this variable will still have a significant influence on the travels of certain pupils to school.

*Another way it impacts me is that I have to walk about twenty minutes to the tram stop, then take the tramway, and then walk further to school. And I never know when the tram is going to come, so I always leave earlier and then I arrive at school earlier.*
[How long does it take you to get to school?]
*An hour.*
(female pupil, 15 years, Remetinec)

Congestion as a factor affecting certain activities need not only be viewed as traffic congestion (0.6%). This variable may also reflect the capacity of transport vehicles in public transport, which ultimately can result in not boarding a vehicle and waiting for the next, and in tardiness to school.

*I travel from Dubrava and I don't think that* (transport) *strongly affects grades, instead it has a greater impact on arriving at school, since the morning crowds on the tram are huge, and sometimes I can't even board, and then I have to walk 15 minutes to the final station. In doing so, I waste time and then am late for school.*
(male pupil, 18 years, Dubrava)

To this point, these were primarily comments on the negative impacts of transport on pupils' school activities. However, the time spent in public transport while travelling to school is used by some pupils for studying or reviewing learning materials (0.6%).

*I study in the tram and bus.*
(male pupil, 18 years, Jarun)

Considering that academic success depends on a series of factors, in this analysis, the variable of academic success has been attempted to be placed in a mutual relationship with the available data from the survey questionnaire (Table 3). For that purpose, the following predictors were used: pupil gender, pupil age, time spent weekly on studying, motivation, importance that good grades has in the pupil's life, assessed time of travel to school and means of travelling to school.[5] The influence of transport on academic success was expressed using a regression analysis. Prior to doing so, it is worthwhile determining whether there is an association between the previously listed predictors and the dependent variables, in this case, the grade point average. For the analysis, the possible association between travel time and the use of public transport with academic success will be of crucial importance.

**Table 3**
*Association of certain predictor variables with academic success*

| PREDICTORS | r | p |
|---|---|---|
| AGE | .064 | .064 |
| GENDER | .221** | .000 |
| TIME SPENT STUDYING | .219** | .000 |
| TRAVEL TIME | -.139** | .000 |
| ABSENCES DUE TO TRANSPORT | -.136** | .000 |
| PUBLIC TRANSPORT | -.092* | .020 |
| | **ρ** | **p** |
| SIGNIFICANCE OF ACADEMIC SUCCESS | .294** | .000 |

r = Pearson's correlation coefficient
ρ = Spearman's rank correlation coefficient
p = likelihood of the random appearance of the value
* value is statistically significant at a risk level of 5% ($p < 0.05$)
** value is statistically significant at a risk level of 1% ($p < 0.01$)
*Source: survey questionnaire, 2013*

The results obtained indicated the assumption of the presence of an effect of distance of the pupil's home from the school on academic success, measured in travel time. Namely, there is a statistically significant correlation ($r = -0.139$; $p < 0.01$) between these variables. The correlation is negative, indicating that pupils who travel longer to school may have poorer academic success. Though this correlation may fall under the category of negligible associations, it systematically arises, and as such may indicate a problem of the influence of traffic or transport disadvantage on the academic success of certain pupils.

Further, the analysis indicates that some pupils who use public transport to get to school have poorer academic success than other pupils. However, this correlation is even weaker than the previous one ($r = -0.092$; $p < 0.05$) and may also fall under the category of negligible association, though one cannot ignore the fact that it arises regularly. This correlation can be justified with the fact that pupils who take public transport to school also live further away, and as a rule spend more time on travel than pupils who walk, cycle or are driven to school. The negative statistically significant correlation with academic success is also shown for the total number of classes missed due to transport ($r = -0.136$; $p < 0.01$). This will be discussed in greater detail further in the section on the influence of transport on missed classes.

After establishing a correlation between the dependent variables and predictors, the regression analysis was conducted to determine in detail the influence of individual (independent) variables on academic success as the dependent variable (Table 4). Backward stepwise regression was employed.

[5] For the predictor "means of travel to school", only the results for pupils who travel to school using public transport will be used. The reason is that the regression analysis is used to analysis this issue, which does not tolerate a small sample, and since 55 pupils of years 2, 3 and 4 walk to school, 4 are driven by car, and 1 pupil drives a moped, this small number is not suitable for the regression analysis.

**Table 4**
*Standardised beta-coefficients of variables that provide to be significant predictors of the effect on academic success in the regression analysis performed using backward stepwise regression.*

| PREDICTORS | β | R | $R^2$ | $R^2$-corrected |
|---|---|---|---|---|
| AGE | 0.096 | 0.398 | 0.158 | 0.150 |
| GENDER | 0.134 | | | |
| TIME SPENT STUDYING | 0.146 | | | |
| IMPORTANCE OF ACADEMIC SUCCESS | 0.238 | | | |
| TRAVEL TIME | -0.083 | | | |
| ABSENCE FROM CLASSES DUE TO TRANSPORT | -0.067 | | | |

β = standardised regression coefficient
R = coefficient of the multiple correlation
$R^2$ = coefficient of the multiple determination
$R^2$-corrected = corrected coefficient of the multiple determination
*Source: survey questionnaire, 2013*

The results indicate that the predictors explained 15.0% of the variance of the academic success (R = 0.398; $R^2$-corrected = 0.150). The largest share of the regression analysis was explained by motivation, i.e. the importance of achieving better academic success. The travel time to school had a negative β coefficient, i.e. was negatively correlated with the criterion. This indicates that individuals spending more time on travelling to school have poorer academic success. Also, the predictor relating to absences due to transport also had a negative β coefficient, indicating that individuals with more absences due to transport had poorer academic success.

It should certainly be noted that the regression analysis included variables that were available for testing via the survey questionnaire. Due to the anonymity of the survey, sensitive personal questions and respecting the privacy of high school pupils (which are still mostly minors), and the social status of pupils (family's financial situation, parent's education status, number of family members, number of automobiles in the family, etc.) were not tested. It is very likely that these factors largely impact the academic success of pupils, and as such would likely change the relations in the regression analysis. These facts certainly open the possibility for further research on the issues of the impacts of transport and transport disadvantage on the academic activities and academic success of pupils.

Mastering skills and attaining competencies will certain largely depend on the pupil's attendance of classes. Furthermore, attendance of classes will also impact academic success. Frequent absences from classes disturb the continuity of learning and will certainly reflect on the pupil's work (Markuš, 2009). The emphasis in research is placed on the total number of classes missed due to transport, as it is assumed that each class missed is harmful for the pupil's education process, regardless of whether these are justified or unjustified absences.

The influence of transport on absences from school is primarily visible in pupils who spend more time travelling to school. Namely, the correlation between the number of absences from classes due to transport and the time pupils spend in travelling to school is positive and statistically significant ($r = 0.401$; $p < 0.01$). This is a correlation that has the strength of a truly significant correlation, and indicates an increased number of absences from schools with increasing time spent travelling to school. Therefore, it is possible that pupils who travel longer to school will have more absences from class, which is expected. However, these absences will have a consequence on the education process of pupils, who will not have the same learning conditions as pupils who have been absent from class significantly less due to transport. Such a situation can also be seen in the pupil's academic success, which was tested statistically. In line with this, a slightly negative correlation was seen between the variables absences from class due to transport and academic success, though it should be emphasised that the correlation is statistically significant and appears systematically ($r = -0.136$; $p < 0.01$). This indicates the possible negative impact of absences due to transport, i.e. the situation that academic success is reduced due to an increasing number of absences due to transport. It is necessary to consider that absences for other reasons, and not only transport, will also affect academic success, though these data were not available for the purposes of this study.

The previously stated result indicates that academic success may be dependant, to a certain extent, on transport. A difference will certainly be seen between pupils who travel further to school and pupils who arrive at school quickly. Pupils requiring longer to travel to school will, to a certain extent, indicate the issue of the influence of transport on their grades and academic success. These are pupils who live in parts of the city near the periphery, and who need 60 to 90 minutes for the one-way trip to school, meaning that they spend approximately the same amount of time returning. These pupils will also most often mention traffic congestion or the sparse frequency of public transit lines as the main reason for the long trip to school. It should be noted that there were no significant differences found between males and females in this segment of the research.

*I live in Trnava, and my bus often runs irregularly. When I get to Kvaternik square, the bus there runs regularly, so I have to leave much earlier because I never know when I will arrive.*
[Do you believe that transport affects your learning and your grades?]
*I think it does in comparison to those who live closer, they always have time for everything, and we are always tight with time, we have to leave earlier to get to school, meaning we have to wake up earlier.*
[How do you try to resolve that problem?]
*Because I spend more time travelling, then I study at night and I study in the bus, because I travel for half an hour from Trnava to Kvaternik square.*
(female pupil, 18 years, Trnava)

*I travel an hour and a half to school from Kraljevečki Novaki and yes, I think it does impact my school and I think I would certainly have better grades if I lived closer.*
[How do you explain that?]
*I don't have time to do my homework. When I have school in the morning, I have to catch the bus at 6:10 am, and I have to get up at about 5 am. The bus does not run often, only every half hour.*
(female pupil, 17 years, Kraljevečki Novaki)

*I think that the main problem in the influence of transport on school is the lost time for studying.*
[Where do you travel from and how long does it take you to get to school?]
*I live in Stenjevec and it takes me an hour and 20 minutes to get to school, especially in the morning because of the congestion on Bologna Avenue, and I also have to walk 10 minutes to get to the bus. If I travelled by car I would certainly arrive faster, because I wouldn't have to walk to the bus, wait for the bus, then the tram, and then walk to school.*
[Do you think this affects your grades?]
*I think that I have less concentration because I'm tired, and because of that lower grades.*
(male pupil, 19 years, Gajnice)

*Travel is not so important, but it does have some influence. Not like I'm about to fail the academic year, but there are some nuances of influence.*
(male pupil, 16 years, Borovje-Žitnjak)

Pupils living relatively near their school, particularly if they are able to walk and therefore do not need much travel time, claim there is no influence of travel or distance, or transport, on academic success. Such an attitude should not be surprising, since these pupils travel to school about ten minutes on foot, which leaves more time for studying and doing homework, and for rest. Again, there were no differences in the attitudes between male and female pupils.

[How do you travel to school?]
*I walk to school.*
[Do you think that this affects your grades?]
*No, I don't think it does at all.*
(female pupil, 16 years, Kozjak)

[How long does it take you to travel to school?]
*It takes me about 15 minutes on the bus from Remete.*
[Do you think that impacts your school activities?]
*I don't think it does.*
(male pupil, 18 years, Remete)

The influence of transport on school activities is also visible in the segment of lost time. This will have a direct influence on pupils who travel longer to school, and again the influence of longer travel time and frequency of public transit lines will be evident (pupils living in neighbourhoods at the periphery and who travel 60 minutes or more to school). Though some pupils believe that transport does not directly impact their academic success, this is not true, as the time lost in travel will directly affect the organisation of their daily activities. Some pupils study at night, which results in fatigue, a lack of concentration and lower grades. Both male and female pupils expressed virtually identical stances.

> *It* (transport) *affects me quite a bit because I live in Sesvete and I have to combine bus and tram. The bus runs only every 45 minutes, and if I miss that bus, then I automatically miss the next bus at Kvaternik square and then I will likely be late for school.*
> [How does the frequency of bus lines affect your school activities?]
> *I have to leave much earlier,* [how much earlier?] *and hour and a half earlier, and today, for example, I didn't have time to learn the lessons for Croatian.*
> (female pupil, 15 years, Sesvete)

> *I think it's a little easier for those who live closer, but we can overcome this. It's not critical, but I'm sure that there are some nuances of a difference. I don't think it will affect whether we pass with a 4 or a 5, it has a greater affect on the loss of time.*
> (female pupil, 16 years, Mikulići)

> *It takes me an hour and a half to get to school and I lose a lot of time. We probably get less sleep because of the long trip to school.*
> (male pupil, 16 years, Branovečina)

In addition to the assessment and time for studying, transport also affects pupils' absence from classes. The greatest problem will be for those pupils living further from school, particularly if they use public transport to get to school. No differences were observed among the sexes on this issue. A larger number of unjustified absences may lead to the passing of disciplinary measures against the pupil for the violation of their duties and non-fulfilment of their obligations. Pupils most often listed traffic congestion, crowds in the public transport vehicles and frequency of transport lines as the main reasons for being late for school.

> *Tardiness is a problem, because they will not justify those absences. Some class teachers will, others will not. I had one unjustified absence last year, but my friend had four because of transport. They tell us to bring a confirmation letter from ZET* (''Zagreb Electric Tram'' company).
> [And do you bring it?]
> *Who's going to go and ask for one!?*
> (female pupil, 15 years, Sesvete)

> *You can't board the bus because it's full, so you wait for the next one, and it's full too, and then you're late for school.*
> (male pupil, 15 years, Rudeš)

## 4. Conclusion

One of the major preconditions of future success on academic and professional level as well as in life are grades and academic success. It depends on a variety of objective and subjective factors. In this context, the influence of transport disadvantage is analysed in relation to academic success of pupils.

The results obtained confirmed that travel time influenced pupils' academic success. Although this relation is relatively weak, it is still statistically significant. Pupils traveling longer to school had lower grades and academic success than those who travel less. Beside, these pupils believe that transport has more impact on their school activities. In this group problems related to transport were being late for school, loss of time due to traveling, traffic congestion and frequency of transport lines. In addition, traveling time has impact on pupils' absence from classes. Number of pupils' absence from classes due to transport is increasing with traveling time to school. More importantly, pupils' grades and academic success are decreasing with increasing number of absence from classes due to transport.

As most of the high schools of the City of Zagreb are located in broader region of city centre, pupils living at the city periphery will travel longer to school than those living closer to city centre. In this case, pupils will be transport disadvantaged, and their dependence on public transport (which most of the pupils use for traveling to school) will become apparent. Pupils living at the city periphery have more problems related to insufficient development of public transport system and to lower frequencies of public transport compared to frequencies in city centre. These pupils have to change two or three modes of public transport in order to reach the school which often leads to being late to school because of bad connections between public transport lines used.

Logically, spaces further from city centre cannot have equal frequencies of public transport as the ones in the centre. Still, transport disadvantaged spaces need to be analysed, so more efficient organisation of public transport can be worked on in order to improve quality of life. In this way, differences in transport disadvantage could be reduced in a sense of easier accessibility to life opportunities, and transport and social equity.

**Acknowledgements**

This research is part of the doctoral dissertation “Impact of transport disadvantage on everyday life of high school population of City of Zagreb” defended on 14th of July 2014 at University of Zagreb, Faculty of Science, Croatia.

**References**

Babarović, T.; Burušić, J.; Šakić, M. 2010. Psihosocijalne i obrazovne odrednice školskog uspjeha učenika osnovnih škola: dosezi dosadašnjih istraživanja (Psychosocial and educational determinants of primary school achievement: grasp of previous research; In Croatian), *Suvremena psihologija* 13(2): 235-256.

Clifton, K.; Lucas, K. 2004. Examining the empirical evidence of transport inequality in the US and UK. In: *Running on empty: Transport, social exclusion and environmental justice* (ed. Lucas, K.). The Policy Press. Bristol. 15-36.

Cohen, L.; Manion, L.; Morrison, K. 2007. *Metode istraživanja u obrazovanju (Research Methods in Education; In Croatian).* V. edition. Naklada Slap. Jastrebarsko. 464 p.

Council for children of the Government of the Republic of Croatia and National institute for protection of family, maternity and youth. 2003. Etički kodeks istraživanja s djecom (Code of Ethics of Research with Children; In Croatian). Available from Internet: <http://www.dijete.hr/en/documents/doc_view/232-etiki-kodeks-istraivanja-s-djecom.raw?tmpl=component>.

Currie, G. 2007. Young Australians: No Way to Go. In: *No Way To Go – Transport and Social Disadvantage in Australian Communities* (ed. Currie, G.; Stanley, J.; Stanley, J.). Monash University ePress. Clayton. 08.1-08.14.

Dodson, J. et al. 2006. Transport Disadvantage in the Australian Metropolis: Towards new concepts and methods. In *2nd Bi-Annual National Conference on the State of Australian Cities: proceedings*, 03-2 – 03-23.

Foley, J. 2004. Overview. In: *Sustainability and social justice* (ed: Foley, J.). Institute for Public Policy Research. London. 1-12.

Gašparović, S. 2014. *Utjecaj prometne marginaliziranosti na svakodnevni život srednjoškolske populacije Grada Zagreba (Impact of transport disadvantage on everyday life of high school population of City of Zagreb; In Croatian)*. PhD thesis. University of Zagreb, Faculty of Science, Department of Geography. Zagreb. 273 p.

Gašparović, S.; Jakovčić, M. 2014. Transport disadvantage: example of high school population of City of Zagreb, *Geoadria,* in press.

Hopkins, P. E. 2010. *Young people, place and identity*. Routledge. London. 320 p.

Horton, J.; Kraftl, P.; Tucker, F. 2011. Spaces-in-the-making, childhoods-on-the-move. In: *Children and Young People's Spaces* (ed. Foley, P.; Leverett, S.). Palgrave Macmillan. Basingstoke. 40-57.

Hoyle, B.; Knowles, R. 1998. Transport geography: An introduction. In: *Modern Transport Geography, Second revised edition* (ed. Hoyle, B.; Knowles, R.). John Wiley & Sons. Chichester. 1-12.

Hurni, A. 2006. *Transport and Social Disadvantage in Western Sydney: A Partnership Research Project*. University of Western Sydney and Western Sydney Community Forum. Sydney. 104 p.

Hurni, A. 2007. Marginalised groups in Western Sydney: The experience of sole parents and unemployed young people. In: *No Way To Go – Transport and Social Disadvantage in Australian Communities* (ed. Currie, G.; Stanley, J.; Stanley, J.). Monash University ePress. Clayton. 10.1-10.11.

Kamaruddin, R. et al. 2009. The Quality of Learning Environment and Academic Performance from a Student's Perception, *International Journal of Business and Management*, 4(4): 171-175.

Lin, J.-J.; Huang, Y.-C.; Ho, C.-H. 2013. School accessibility and academic achievement in a rural area of Taiwan, *Children's Geographies,* DOI:10.1080/14733285.2013.812308, published on-line 28.06.2013.

Maras, N.; Rodek, J. 2012. Utjecaj nekih stratifikacijskih čimbenika na školski uspjeh (The influence of some stratification factors on school success; In Croatian), *Školski vjesnik – časopis za pedagoška i školska pitanja,* 61(1-2): 41-57.

Markuš, M. 2009. Psihosocijalne determinante školskih izostanaka (Psychosocial determinants of absenteeism; In Croatian), *Napredak*, 150(2): 154-167.

Minogue, K. 1998. Social justice in theory and practice. In: *Social justice: from Hume to Walzer* (ed. Boucher, D.; Kelly, P.). Routledge. London. 253-266.

Murray, A. T.; Davis, R. 2001. Equity in regional service provision, *Journal of Regional Science*, 41(1): 577-600.

Owoeye, J. S.; Yara, P. O. 2011. School location and academic performance of secondary school in Ekiti State, Nigeria, *Asian School Science*, 7(5): 170-175.

Raychaudhuri, A. et al. 2010. Factors affecting Student's academic performance: A case study in Agartala Municipal Concial Area, *Bangladesh e-journal of sociology*, 7(2): 34-41.

Social Exclusion Unit. 2003. *Making the Connections: Final Report on Transport and Social Exclusion*. Office of the Deputy Prime Minister. United Kingdom. 147 p.

Stanley, J.; Stanley, J. 2004. *Improving Public Transport to meet Comunity Needs: A Warrnambool Case-study*. Bus Asociation Victoria and Warrnambool Bus Lines. Melbourne. 92 p.

van Goeverden, C. D.; de Boer, E. 2013. School travel behaviour in the Netherlands and Flanders, *Transport Policy*, 26(1). 73-84.

Vranković, B.; Vuk, R.; Šiljković, Ž. 2011. *Kvalitativna analiza ispita vanjskog vrjednovanja obrazovnih postignuća učenika osmih razreda provedenih 2008. godine: geografija i integracija nastavnih sadržaja iz geografije i povijesti (Qualitative analysis of the exams of the external evaluation of educational achievements of eighth grade pupils performed in 2008: Geography and integration of teaching content in Geography and History; In Croatian)*. Nacionalni centar za vanjsko vrednovanje obrazovanja. Zagreb. 270 p.

# PEDESTRIAN BEHAVIOR AT SIGNALIZED INTERSECTIONS IN IZMIR, TURKEY

**Pelin Onelcin**[1], **Yalcin Alver**[2]
*[1,2] Ege University, Civil Engineering Department, Izmir, Turkey*

**Abstract:** This paper investigates the walking speed and delays of pedestrians at two signalized intersections where the vehicle speed limits are different. Data was collected during afternoon and evening peak hours on November 15, 2013 and on December 6, 2013. Observational surveys were conducted using a video recording technique. Pedestrians were categorized according to their gender, group size, items they were carrying and age. The results showed that individuals walked fastest when the group size is taken into consideration. The smallest 15$^{th}$ percentile walking speed was seen in the oldest age group (over 60 years old). Pedestrians experienced high delays both at roadsides and at medians. The factors affecting the pedestrian walking speed were analyzed by ANOVA.

**Keywords**: pedestrian delay, pedestrian walking speed, signalized crosswalk.

## 1. Introduction

Pedestrian behavior at signalized intersections is very important for traffic engineers to design more efficient and safe intersections. The share of the same road space by pedestrians and vehicles in different time intervals lead to a complexity at signalized crosswalks (Lee and Lam, 2008).

Pedestrian walking speed is used to design safe and efficient signal timings. In each country, pedestrian characteristics show variations. Thus, the signal timing design should be specified to the country. In Turkey the design walking speed is given as 1.4 m/s by Turkish Standards Institution (TSI, 2012). Nevertheless, this walking speed is overrepresented for many of the intersections in Turkey, particularly where the number of elderly pedestrians is high.

The factors affecting the pedestrian walking speed at signalized crosswalks were analyzed by multi factor analysis of variance (ANOVA). Pedestrians' age was categorized in three groups; the young group (10-19 years old), the adult group (20-59), and the old group (over 60 years old). Group size was categorized in two groups; individuals, and pedestrians who walked within a group of two and more people. Another factor that was concerned was the pedestrians carrying items. The difference in walking speed between pedestrians who carried items and who did not was analyzed.

There are three phases in one signal cycle at the observed intersections. These are green phase, all red phase, and red phase. Pedestrians who arrive during red phase usually experience high delays. In roads separated by a median, pedestrians experience delays not only at roadside but also at medians. In this study both type of delays were observed during red phase and green phase, and the delay data presented here involves both illegal crossings (during red phase) and legal crossings (during green phase).

In this paper the aim is twofold. First, the authors would like to find the walking speed of pedestrians who cross at signalized intersections where the speed limits are different and to recommend a walking speed value for signal timing planning. Second, the authors would like to determine the roadside and median stopped delays that the pedestrians experience at the signalized intersections.

This paper has the following structure. The second part gives a brief literature review about pedestrian walking speed and pedestrian delay. The third part explains the field study and presents the results of the observations. The fourth part is a general discussion about the findings of the study. And the final part is the conclusion.

## 2. Literature Review

In Highway Capacity Manual 2010 the pedestrian walking speed is assumed to be effected by density, gender, size of platoon, percentage of elderly population, handicapped pedestrian population and child pedestrian population, and an average walking speed is given as 1.2 m/s (HCM, 2010).

In Madison and Milwakee, Wisconsin 1,947 pedestrian crossings were observed at eleven intersections. ANOVA analysis showed that age had the most significant effect among all factors on walking speed. Pedestrians under the age of 30 were the fastest among the observed pedestrians with the 15$^{th}$ percentile walking speed of 1.27 m/s. Pedestrians who crossed in groups of 2-4 walked 0.12 m/s slower compared to the pedestrians who crossed individually. Authors recommended a 1.2 m/s walking speed at locations with normal pedestrians' population, and a 0.88 m/s walking speed where almost all pedestrians are over age of 65 (Gates et al., 2006).

[1] Corresponding author: onelcinpelin@gmail.com

Tarawneh (2001) searched for the contributing factors such as age, gender, distance crossed and the group size to pedestrian walking speed at signalized intersections in Jordan. At 12 intersections 3,500 pedestrian crossings were observed. Age was categorized in five groups and the highest walking speed ($15^{th}$ percentile speed of 1.07 m/s) was observed in 21-30 age range. Pedestrians who walked individually or in couples were faster than the pedestrians who walked in groups of three or more. Males walked faster than females with 1.12 m/s and 1.11 m/s $15^{th}$ percentile walking speeds, respectively. The average pedestrian walking speed was found to be 1.34 m/s. The author proposed a $15^{th}$ percentile walking speed of 1.11 m/s for Jordan. In the case of high ratio of older pedestrians it was recommended to use 0.97 m/s of $15^{th}$ percentile walking speed.

Bennett et al. (2001) conducted a survey at four signalized intersections in Melbourne, Australia. Observations showed that $15^{th}$ percentile speed varied from 1.18 to 1.59 m/s. The walking speed was obtained as 1.24 m/s when all data from the observations were combined.

Pedestrian delay estimation is based on uniform arrival rates and fixed pedestrian timing in Highway Capacity Manual (2010). Nevertheless this assumption does not apply to two stage crossings.

Virkler (1998) indicated that there are times when pedestrians do not comply with signal rules to minimize their delays. In Brisbane, Australia at 18 crosswalks pedestrians who crossed during green signal, pedestrians who entered the crosswalk during the clearance interval, pedestrians who entered the crosswalk during the red interval, and the delay of all pedestrians were observed. By modifying the equation developed by Braun and Roddin, they used 69% of the clearance interval for entering the crosswalks.

Li et al. (2005) developed a model for pedestrian delays in developing cities like Xi'an, China. The authors noted that the models currently used in developed countries do not take into account the delay that pedestrians may encounter during the green signal. They found that the average delay of pedestrians during the green signal was 1.9 s. In Turkey, pedestrians experience delays during green signal as well. Thus, delays are observed to determine the local pedestrian behavior at signalized intersections in Turkey.

## 2. Field Study

### 2.1. Sirinyer and Bostanli Intersections

Two intersections in different parts of Izmir, Turkey were observed using a video recording technique. The locations of the intersections are demonstrated in "Fig. 1". Sirinyer intersection is three-legged and Bostanli intersection is four-legged with signalized crosswalks and the main roads are all four-laned.

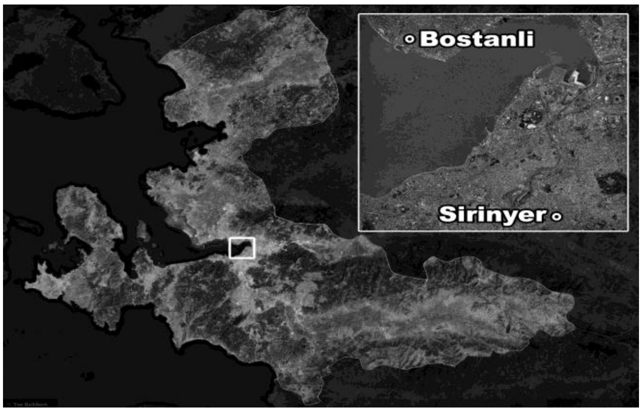
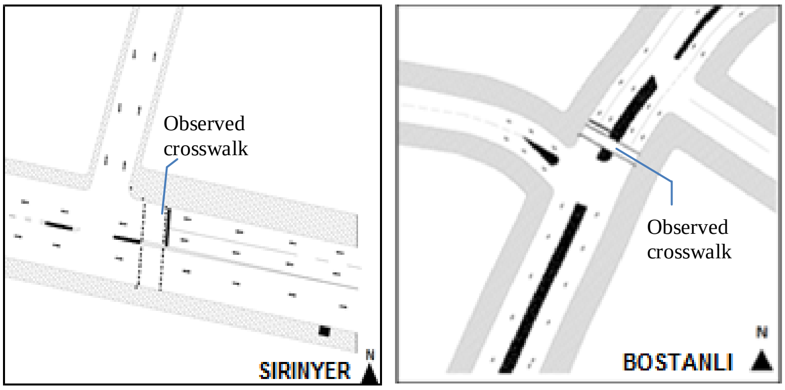



**Fig. 1.**
*Observed intersections in Izmir, Turkey*
*Source: Google Earth*

Sirinyer has a high vehicular traffic throughout the day. The left turning drivers usually did not yield the right of way to the pedestrians and thus pedestrians experienced delays during the green phase as well. The length and the width of the crosswalk are 13.6 m and 3.45 m, respectively. The cycle length is 100 s which consist of 15 s green signal and 85 s of red signal. In Fig.2 (a) the drawing of the intersection is demonstrated.

Bostanli intersection is located in a commercial area where there are many cafes and restaurants that attract people. The cycle length is 108 s at the observed signalized crosswalk. The green signal duration is 30 s and the red signal duration is 78 s. The length and the width of the crosswalk are 18.7 m and 3.78 m, respectively. Pedestrians can wait at refuge island which has a 3.75 m length and 3.78 m width. Roads are two-laned in each direction however one lane is used as a parking lane thus only one lane allows the traffic flow. Fig. 2(b) demonstrates the drawing of the Bostanli intersection.




**(a)Sirinyer** **(b)Bostanli**

**Fig. 2.**
*Drawings of the surveyed intersections*

### 2.2. Data Collection

Data was collected at the above mentioned signalized intersections using a video recording technique. Two hours of recordings were conducted at each intersection during afternoon (12:30-13:30) and evening (17:00-18:00) peak hours on weekdays and then were extracted in office. Sirinyer intersection was observed on November 15, 2013 and Bostanli intersection was observed on December 6, 2013. In total, 840 pedestrian crossings were observed.

## 3. Results

Data obtained from video recordings are summarized below. Table I and Table II shows the average walking speed, standard deviation, and the 15th percentile walking speed by gender, age group, group size and items carrying.

**Table 1**
*Walking speed statistics by gender, age group, group size and items carrying at Sirinyer intersection*

| Factor | Level | Count | Walking Speed (m/s) | Standard Deviation | $15^{th}$ Percentile Walking Speed (m/s) |
|---|---|---|---|---|---|
| Gender | Female | 322 | 1.28 | 0.268 | 1.04 |
| | Male | 158 | 1.31 | 0.231 | 1.04 |
| Age | <20 | 88 | 1.34 | 0.21 | 1.13 |
| | 20-59 | 345 | 1.29 | 0.267 | 1.05 |
| | >60 | 47 | 1.14 | 0.244 | 0.90 |
| Group Size | Individual | 313 | 1.31 | 0.276 | 1.05 |
| | 2+ | 167 | 1.25 | 0.222 | 1.05 |
| Items Carrying | Yes | 183 | 1.27 | 0.245 | 1.05 |
| | No | 297 | 1.30 | 0.269 | 1.05 |

In Table I it can be seen that male pedestrians walked faster than females. The youngest age group (10-19) walked fastest among all age groups. The average walking speed decreased as the pedestrian age increased. Individuals walked faster than the pedestrians who walked within a group of two and more pedestrians. However, the $15^{th}$ percentile speed was the same both for individuals and for pedestrians who moved in groups (1.05 m/s). Pedestrians who did not carry any items walked faster than the pedestrians with items. However, again the $15^{th}$ percentile walking speed was observed as 1.05 m/s both for pedestrians who carried items and for those who did not. The smallest $15^{th}$ percentile walking speed was observed for the oldest age group (0.90 m/s).

At Bostanli intersection, in contrast to Sirinyer intersection, female pedestrians walked faster than male pedestrians. Table II shows that walking speed decreased as the pedestrian age increased. Pedestrians who moved within a group walked slower than the individual pedestrians. Pedestrians who carried items walked slower than the pedestrians who did not carry any items. The 15th percentile walking speed was identical (1.10 m/s) for males, for pedestrians who moved within a group, and for pedestrians who carried items. The smallest 15th percentile walking speed was observed for the oldest age group (1.04 m/s).

**Table 2**
*Walking speed statistics by gender, age group, group size and stuff carrying condition at Bostanli intersection*

| Factor | Level | Count | Walking Speed (m/s) | Standard Deviation | 15th Percentile Walking Speed (m/s) |
|---|---|---|---|---|---|
| Gender | Female | 322 | 1.28 | 0.268 | 1.04 |
| | Male | 158 | 1.31 | 0.231 | 1.04 |
| Age | <20 | 88 | 1.34 | 0.21 | 1.13 |
| | 20-59 | 345 | 1.29 | 0.267 | 1.05 |
| | >60 | 47 | 1.14 | 0.244 | 0.90 |
| Group Size | Individual | 313 | 1.31 | 0.276 | 1.05 |
| | 2+ | 167 | 1.25 | 0.222 | 1.05 |
| Items Carrying | Yes | 183 | 1.27 | 0.245 | 1.05 |
| | No | 297 | 1.30 | 0.269 | 1.05 |

Pedestrian delays by gender, age group, group size and items carrying are given in Table III and Table IV for Sirinyer and Bostanli intersections, respectively. Table III shows that pedestrians at roadsides experienced higher delays compared to the delays they experienced at medians. At Sirinyer intersection, the median did not provide an adequate waiting area for pedestrians. On the other hand, pedestrians who started crossing from the roadside usually had to wait at the median during the green phase because the left turning vehicles were moving on the crosswalk during the same time interval. Thus, pedestrians experienced delays at medians as well. The smallest delay was observed at median among females whereas the highest delay was observed at roadside among pedestrians who did not carry any items.

At Bostanli intersection, in contrast to Sirinyer intersection, pedestrians experienced higher delays at medians. At this intersection, pedestrians could wait at a waiting area before they started their second stage crossing. There was also a special case in Bostanli. Although the roads were two-laned, one lane was used as a parking lane. When pedestrians entered the crosswalk they were still protected by the parking vehicles until they reached the second lane. The usual behavior at this intersection was that pedestrians waited for a short time (around 7 s) at roadside and speeded up, and then they arrived at the refuge island. In the second stage of the crossing they waited longer because in this case pedestrians faced the traffic flow first. That is why the median delays were longer than the roadside delays at this intersection. The highest delay was observed at median among pedestrians who moved within a group of two and more pedestrians. The smallest delay was observed at roadside among the young pedestrians.

**Table 3**
*Delay statistics by gender, age group, group size and items carrying at Sirinyer intersection*

| Factor | Level | Delays at roadside (s) | Delays at median (s) |
|---|---|---|---|
| Gender | Female | 16.65 | 5.55 |
| | Male | 17.20 | 6.29 |
| Age | <20 | 24.03 | 17.55 |
| | 20-59 | 24.02 | 14.08 |
| | >60 | 18.89 | 10.5 |
| Group Size | Individual | 23,39 | 14.89 |
| | 2+ | 24.05 | 13.48 |
| Items Carrying | Yes | 20,97 | 15.65 |
| | No | 25,02 | 13.29 |

**Table 4**
*Delay statistics by gender, age group, group size and items carrying at Bostanli intersection*

| Factor | Level | Delays at roadside (s) | Delays at median (s) |
|---|---|---|---|
| Gender | Female | 6.37 | 8.45 |
| | Male | 6.84 | 8.76 |
| Age | <20 | 4.26 | 6.94 |
| | 20-59 | 6.63 | 9.45 |
| | >60 | 8.25 | 7.2 |
| Group Size | Individual | 6.34 | 7.96 |
| | 2+ | 7.01 | 9.49 |
| Items Carrying | Yes | 7.76 | 7.08 |
| | No | 6.10 | 9.08 |

The walking speed data were analyzed using ANOVA to find the factors affecting the pedestrian walking speed. The analyses were conducted for two intersections where the speed limit is 30 km/h (Bostanli intersection) and where the speed limit is 50 km/h (Sirinyer intersection). This allowed determining whether the lower vehicle speed had a relaxing influence on pedestrians not to hurry while crossing because they did not perceive any collision risk. The ANOVA analyses were performed in SPSS v20 using the Univariate General Linear Model command (IBM, 2011). The confidence level was set to 95%.

The ANOVA on the pedestrian walking speed revealed significant effects of items carrying, $F(1,839)=4.291$, $p<0,05$; group size, $F(1,839)=13.235$, $p<0,0001$; and age $F(2,839)=29.341$, $p<0,0001$.

Vehicle speed had a significant effect on walking speed, $F(1,839)=50.625$, $p<0,0001$. At Bostanli intersection where the vehicle speed limit is 30 km/h, pedestrians walked faster with an average walking speed of 1.41 m/s. At Sirinyer intersection the average walking speed was found to be 1.29 m/s.

Gender did not reveal any significant effect on walking speed when all data was combined. Interactions between the factors did not have any significant effect either.

Table V shows the walking speed statistics of the combined data. Male pedestrians had a higher walking speed when all data was combined. The 15$^{th}$ percentile walking speed was lowest for the old pedestrians. The average walking speed of all pedestrians was found to be 1.34 m/s and the 15$^{th}$ percentile walking speed was found to be 1.10 m/s. This is 21.4% lower than the walking speed which is proposed by Turkish Standards Institution (TSI).

**Table 5**
*Walking speed statistics by gender, age group, group size and stuff carrying condition at both intersections*

| Factor | Level | Count | Walking Speed (m/s) | Standard Deviation | 15$^{th}$ Percentile Walking Speed (m/s) |
|---|---|---|---|---|---|
| Gender | Female | 502 | 1.33 | 0.262 | 1.05 |
| | Male | 338 | 1.36 | 0.263 | 1.10 |
| Age | <20 | 143 | 1.41 | 0.228 | 1.13 |
| | 20-59 | 602 | 1.35 | 0.263 | 1.10 |
| | >60 | 95 | 1.16 | 0.224 | 0.94 |
| Group Size | Individual | 558 | 1.37 | 0.274 | 1.10 |
| | 2+ | 282 | 1.28 | 0.228 | 1.05 |
| Items Carrying | Yes | 279 | 1.29 | 0.243 | 1.05 |
| | No | 561 | 1.36 | 0.269 | 1.13 |

## 4. General Discussion

Pedestrian walking speeds and delays are the main concern of this study. At two signalized intersections in Izmir, Turkey two hours of observations were conducted. In total, 840 pedestrian crossings were observed.

At both intersections traffic flow was separated by medians. Pedestrian delays at both roadside and at median were extracted from the video recordings. Pedestrians experienced higher delays at roadside in Sirinyer than in Bostanli. This can be explained in terms of risk avoiding behavior of pedestrians in roads with higher speed limits. Pedestrians preferred waiting instead of crossing without high delays because they may have perceived higher degree of risk. On the other hand, where the speed limit was low, such as in Bostanli, pedestrians may have perceived lower degree of risk and they started crossing without experiencing high delays. This also had an effect on walking speed. Pedestrians who entered the crosswalk without encountering high delays, especially during red phase, tended to complete their crossings by walking faster. This can be called as “wait less, walk faster” situation.

The factors affecting the walking speed were analyzed by ANOVA. Gender did not reveal a significant effect on walking speed. On the other hand, all the remaining factors; age, group size, items carrying and vehicle speed, had significant effects.

In Sirinyer pedestrians had lower walking speeds compared to Bostanli. Pedestrians had greater delays at this intersection but they completed their crossings without speeding up during the green phase.

The effect of speed limit was compared at two intersections where the speed limit is 30 km/h and 50 km/h. It is generally expected that pedestrians walk faster when the speed limit is high. However, the results were different than expected. Pedestrians preferred waiting before crossing and then walked calmly without hurrying in Sirinyer where the speed limit is 50 km/h. The exact opposite situation was observed in Bostanli where the speed limit is 30 km/h.

The walking speed changed with age such that the young pedestrians walked fastest and the old pedestrians walked slowest. The $15^{th}$ percentile walking speed was found to be 1.10 m/s when all data was combined. This is 21.4% lower than the walking speed proposed by TSI. The signal timings arranged based on the walking speed proposed by TSI leads to short green phases, and consequently an increase in the number of illegal crossings.

The recommended $15^{th}$ percentile walking speed for Izmir is 1.05 m/s. In case of high ratio of elderly pedestrians it can be taken as 0.94 m/s.

## 4. Conclusion

In this study the effects of gender, age, group size and items carrying on pedestrian walking speed at signalized crosswalks were analyzed, and the pedestrian delays at roadsides and medians were presented for two intersections with different speed limits in Izmir, Turkey. The results showed that pedestrian delays were lower and their walking speeds were higher at the intersection where the speed limit was low. Further investigations at signalized intersections with various speed limits will help to understand whether a relation between the walking speed, delay and vehicle speed exists or not.

**References**

Bennett, S.; Felton, A.; Akcelik, R. 2001. Pedestrian movement characteristics at signalized intersections. In *Proceedings of the 23rd Conference of Australian Institutes of Transport Research (CAITR 2001), Monash University, Melbourne, Australia, December 10-12.*

Gates, T.J.; Noyce, D.A.; Bill, A. R.; Ee, N.V. 2006. Recommended walking speeds for Pedestrian clearance timing based on Pedestrian characteristics. *Transportation Research Record: Journal of the Transportation Research Board,* 1982: 38-47.

Highway capacity manual.2010. *Transportation Research Board*. National Research Council, Washington, DC.

IBM Corp. Released 2011. IBM SPSS Statistics for Windows, Version 20.0. Armonk, NY: IBM Corp.

Lee, J.Y.S.; Lam, W.H.K. 2008. Simulating pedestrian movements at signalized crosswalks in Hong Kong. *Transportation Research Part A*, 42(10): 1314–1325.

Li, Q.; Wang, Z.; Yang, J.; Wang, J. 2005. Pedestrian delay estimation at signalized intersections in developing cities. *Transportation Research Part A,* 39(1): 61-73.

Tarawneh, M.S. 2001. Evaluation of Pedestrian speed in Jordan with investigation of some contributing factors. *Journal of Safety Research*, 32(2): 229-236.

Turkish Standards Institute. 2012. Urban roads - Design criteria on sidewalks and pedestrian areas (TS12174).

Virkler, M.R. 1998. Pedestrian compliance effects on signal delay. *Transportation Research Record: Journal of the Transportation Research Board*, 1636: 88-91.

# GENERAL OSCILLATORY MODEL OF VEHICLES FOR PASSENGER TRANSPORT

**Dragan Sekulić [1], Vlastimir Dedović [2], Dušan Mladenović [3]**
[1,2,3] *University of Belgrade, Faculty of Transport and Traffic Engineering, Belgrade, Serbia*

**Abstract:** This paper defines a generalized oscillatory model of land transport passenger vehicles. The model is built in the ADAMS/View software, and is intended for simulation studies of the oscillatory behavior of road and rail vehicles. By the use of this model it is possible to carry out analysis of the user's oscillatory comfort (driver and all/each passenger). Different '*oscillatory comfort zones*' on the platform of the vehicle can be mapped by considering the comfort parameters of each user. The model can be used also for the analysis of other problems of vertical vehicle dynamics (for example the evaluation of vehicle active safety). The paper presents some of the results of simulations made - the change of acceleration on the seats of three users and changes of the wheel vertical reactions of an intercity bus. Oscillatory excitation of the model is carried out using the recorded signal of roughness of an asphalt-concrete pavement in good condition.

**Keywords:** general oscillatory model of vehicles, vertical vehicle dynamics, simulation, ADAMS/View

## 1. Introduction

The rating of the oscillatory behavior of ground vehicles for passenger transport can be carried out by using experimental and simulation studies. Simulation studies are performed by using oscillatory models and appropriate software packages (Genta, 1997; Pečeliūnas, et al., 2005; Iwincki, 2006; Sekulić and Dedović, 2011). Models are specifically designed and often tied to the vehicle under consideration and have limited application. By generalizing such models a wider range of applications can be achieved. The purpose of this paper is to define and develop a more generalized model that could be used to analyze the oscillatory behavior of road and rail vehicles for transport of passengers.

Ground transport means for passenger transport can be described for the purpose as a system of rigid bodies attached by elasto-damping joint elements. In order to achieve a practical solution by modeling, it is necessary to introduce various assumptions and simplifications so that a model could enable an analysis of such aspects of vehicle behavior, significant for the purpose of research. General oscillatory model should include elements (rigid bodies and joints) that have a major impact on the oscillatory behavior of the vehicle and users' oscillatory comfort, for example, elastically supported vehicle mass (vehicle body), the seats of users, suspensions, axles and wheels. Some simplified models of road and rail vehicles for passenger transport, presented as a system of interconnected rigid bodies, are shown in Fig. 1 and Fig. 2.

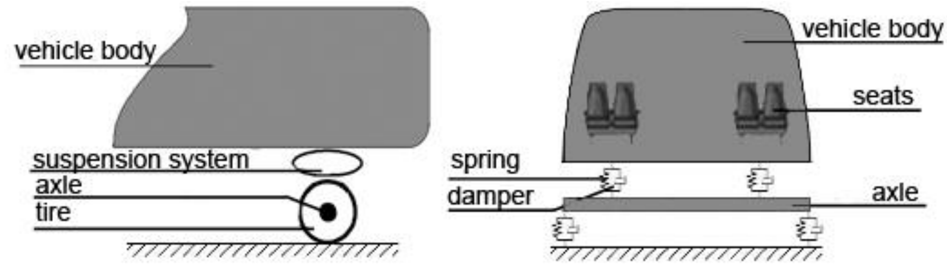



**Fig. 1.**
*Simplified model of road vehicle as a system of interconnected rigid bodies*

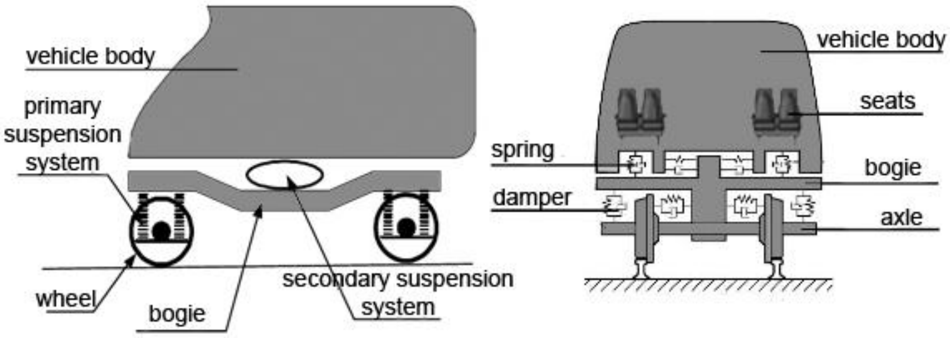



**Fig. 2.**
*Simplified model of rail vehicle as a system of interconnected rigid bodies*

[1] Corresponding author: d.sekulic@sf.bg.ac.rs

Further in this paper, the similarities and differences between individual elements of road and rail passenger transport vehicles have been presented. Comparative analysis provides a basis for the design of the general oscillatory model which can be used for investigation of the oscillatory comfort of each user. Comprehension of user's comfort parameters allows mapping of the oscillatory comfort on the vehicle platform (Sekulić, 2013). General oscillatory model, presented in this paper, is built in the ADAMS/View modulus of the MSC.ADAMS software.

## 2. Comparative analysis of road and rail vehicle components

### 2.1. Elastically suspended mass of the vehicle

Road and rail vehicles for passenger transport are characterized by having a suspended mass (vehicle body) within which there are seats for passengers, with or without suspension system. Fig. 3(a) and Fig. 3(b) show the elastically suspended vehicle mass with seats, for a bus and for a rail passenger vehicle. Vehicle body itself can be modeled as a rigid or as elastic (Sekulić, 2013; Tianfei, et al., 2010). In this paper, the assumption is made that the vehicle body is rigid. Fig. 4 shows an elastically supported rigid vehicle body with seats, modeled in the ADAMS/View software. Geometric parameters (e.g. length, width and height) and mass parameters (masses and moments of inertia) in the ADAMS/View software working environment can be easily changed and adapted to a particular type of vehicle by using the commands intended for their setting.

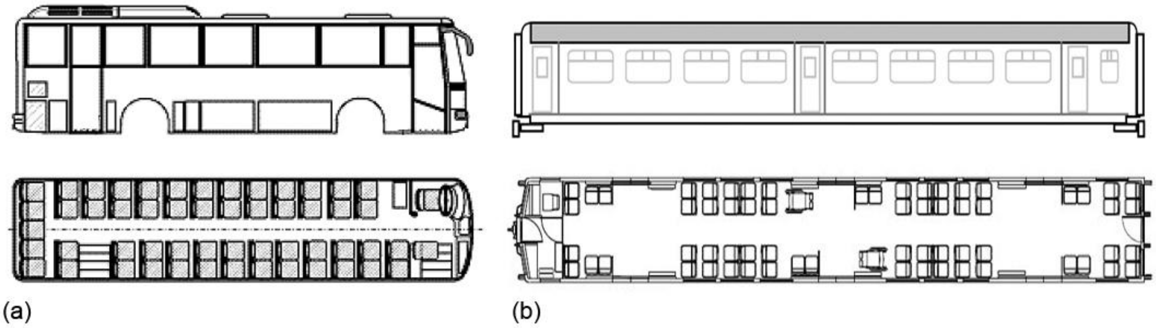



**Fig. 3.**
*Vehicle body with seats a) bus and b) rail passenger vehicle*

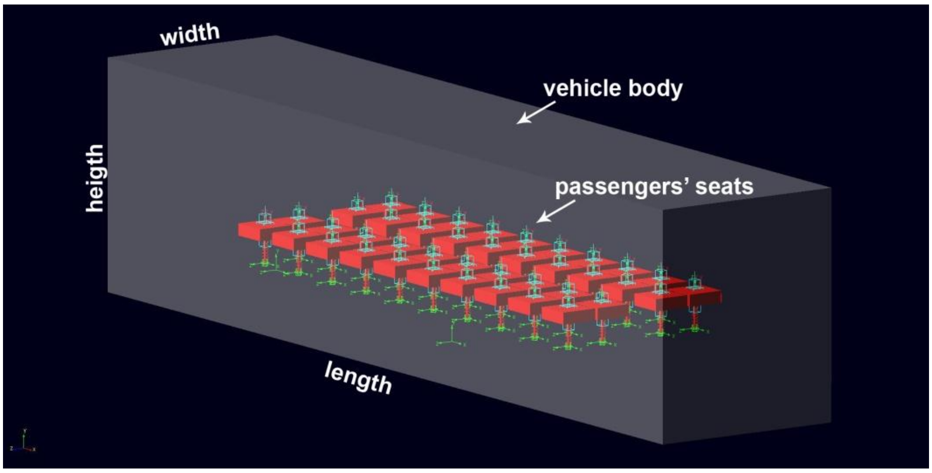



**Fig. 4.**
*Elastically suspended mass of the vehicle as a rigid body, equipped with passenger seats*

### 2.2. Passengers and drivers' seats

It is known that the vibrations of the vehicle are transmitted to user's body through the seat. The seats are usually rigidly mounted, usually to the floor of the vehicle, and the attenuation of vibrations is performed by their elastic cushions only. The driver's seat, unlike the passenger's one, has its own system of suspension that allows an improved comfort. In the ADAMS/View software, user's seats are modeled as rigid bodies connected to the floor of the vehicle body through translational joints. These joints allow a vertical motion only. In order to evaluate the effect of vibration on the comfort of users, the markers can be placed on seats with aim to collect the signals of translational acceleration in the *x*, *y* and *z*-axes direction.

Fig. 5 and Fig. 6 show the driver's and the passenger seats of a bus and of a rail passenger vehicle. Fig. 7 shows the model of the seat formed in the ADAMS/View software.


**Fig. 5.**
*The seats of a) bus driver and b) bus passengers*


**Fig. 6.**
*The seats of a) rail vehicle driver and b) rail vehicle passengers*

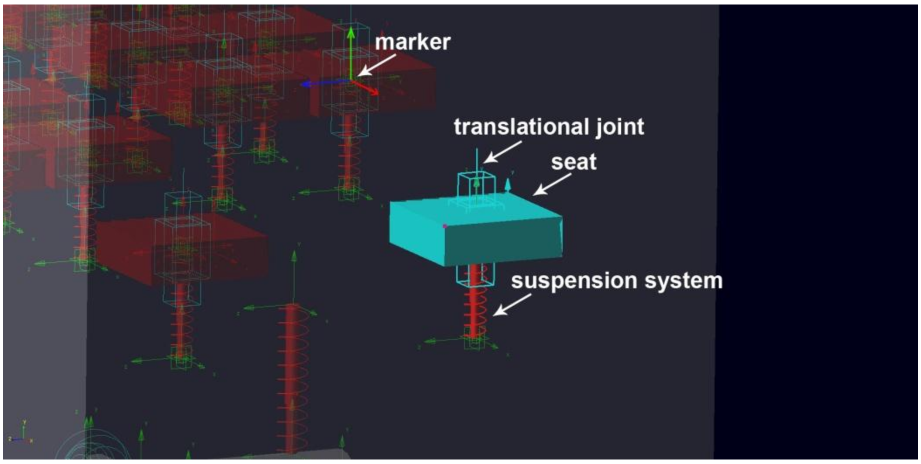



**Fig. 7.**
*Model of the seat in the ADAMS/View software*

### 2.3. Axles, wheels and suspension systems

Regarding the road passenger transport vehicles, elastic suspension system may be designed with rigid axles or with independent wheels. Road surface roughness (oscillatory excitation) is transmitted to the axles through elastic wheels (tires). Fig. 8 shows an example of the front and rear rigid axle of a bus.

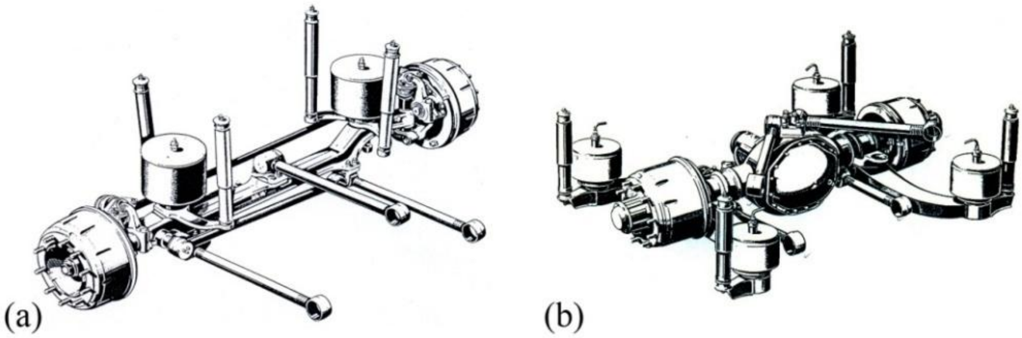



**Fig. 8.**
*Bus rigid axles a) front and b) rear*

Tires have a significant impact on the oscillatory behavior of road vehicles. In this respect, their elastic properties in radial direction are particularly important.

The railway passenger vehicles have rigid axles (Fig. 9(a)), connected to the bogie by the axle housing (Fig. 9(b)). One bogie usually have two rigid axles (Fig. 9), and a rail passenger vehicle generally has two of them (Fig. 10). Unlike road passenger vehicles, rail vehicles have solid wheels. Track irregularities (oscillatory excitation) are transferred to the axle through rigid wheels. Oscillatory excitation is transmitted from axles to the bogie through the elements of primary suspension. The bogie is connected to vehicle body by secondary suspension system with two air springs (Fig. 9).

While designing a general oscillatory model, the essential difference between road and railway vehicles is in the transmission of oscillatory excitation (elastic and rigid wheels) and it has to be taken into account.

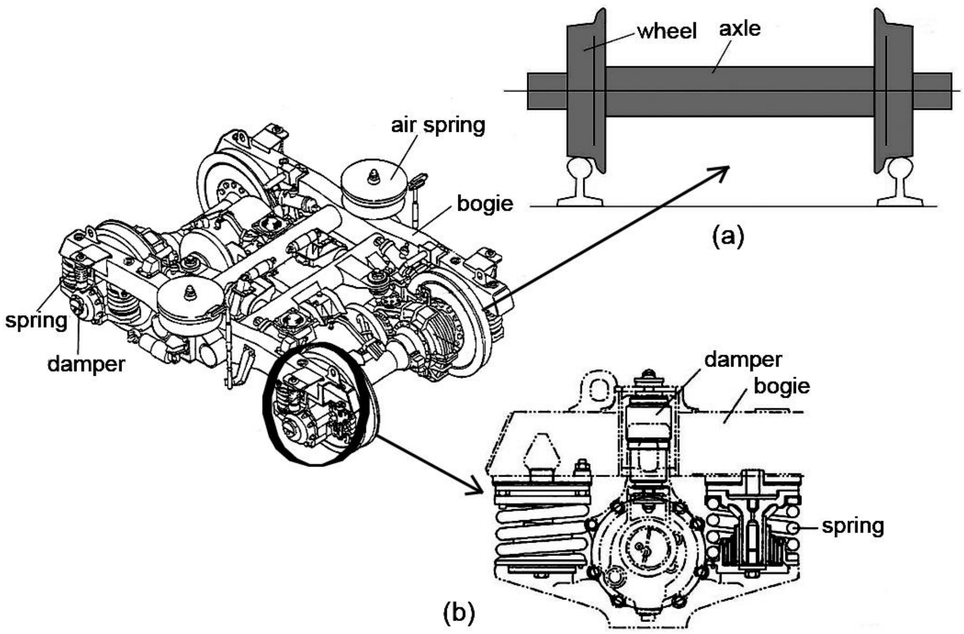



**Fig. 9.**
*Bogie of rail passenger vehicle (a) rigid axle with wheels and (b) the axle housing*

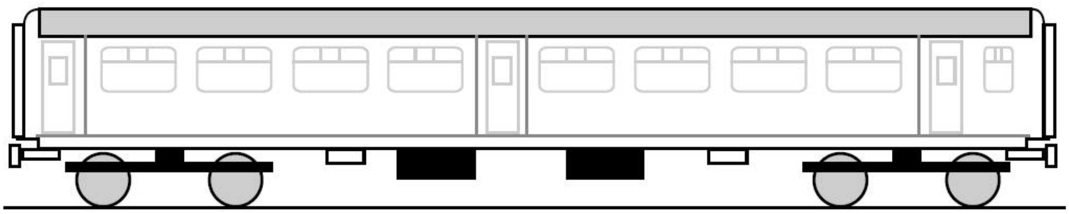


**Fig. 10.**
*The railway passenger vehicle with two bogies*

The axle housing contains the elements of the primary suspension system - steel springs and hydraulic shock absorbers (Fig. 9(b)). The springs and dampers are connected to the frame of the bogie. When defining a general vehicle oscillatory model, two (or more) axles of the bogie are replaced by a single '*equivalent axle'*. The equivalent axle means one axle with the same oscillatory characteristics as the set of axles replaced. Equivalent spring stiffness and equivalent shock absorbers damping can be, according to the scheme shown in Fig. 11, calculated by using expressions (Eq. (1) and Eq. (2)).

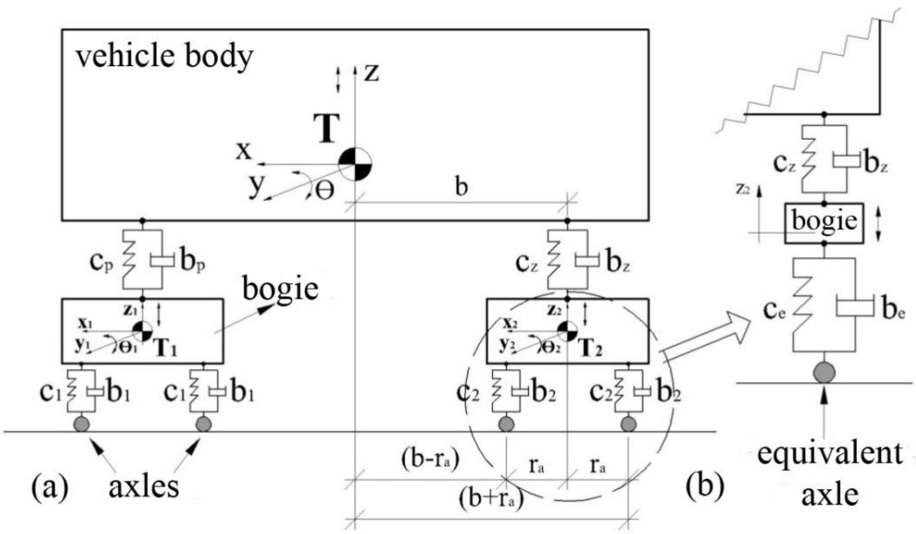



**Fig. 11.**
*Schematic representation of a) elastic suspension of a rail vehicle and b) equivalent axle*

$$c_e = c_2 \frac{(b-r_a)^2}{b^2} + c_2 \frac{(b+r_a)^2}{b^2} \quad (1)$$

$$b_e = b_2 \frac{(b-r_a)^2}{b^2} + b_2 \frac{(b+r_a)^2}{b^2} \quad (2)$$

Road passenger vehicles elastic tires characteristics (stiffness and damping in radial direction) as well as the characteristics of springs (stiffness) and hydraulic shock absorbers (damping) of the primary suspension system of rail passenger vehicle are modeled in the ADAMS/View software by *SPRING-DAMPER* elements (Fig. 12).

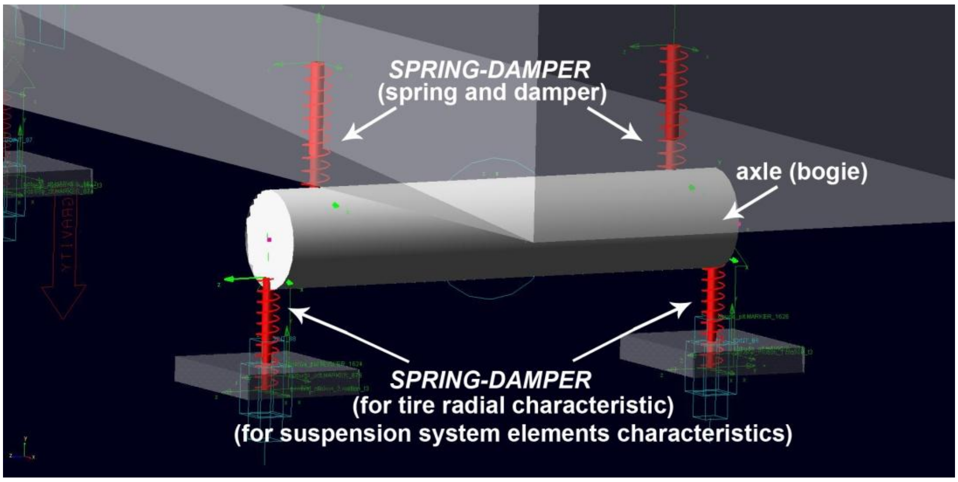



**Fig. 12.**
*SPRING-DAMPER elements used for modeling the characteristics of radial tires, elastic and damping elements of the vehicle suspension system*

For the purpose of this analysis, it can be assumed that the tire is characterized by linear radial stiffness (Miege, 2004). Damping force in the tire is small and often neglected. Within rail vehicle, the characteristics of elastic elements (steel springs) and hydraulic shock absorbers of the primary suspension systems are generally nonlinear (Iwincki, 2006). Characteristics of springs and shock absorbers, whether linear or nonlinear, can be defined by analytical expressions using ADAMS/View function Builder tools. *SPLINE* function enables introduction of springs and shock absorbers characteristics explored through laboratory testing, already written in some of the text files.

In the case of springs and dampers linear characteristics, the force in the *SPRING-DAMPER* element is linearly dependent on the relative displacement and relative velocity of supports of this element (markers I and J, Fig. 13).

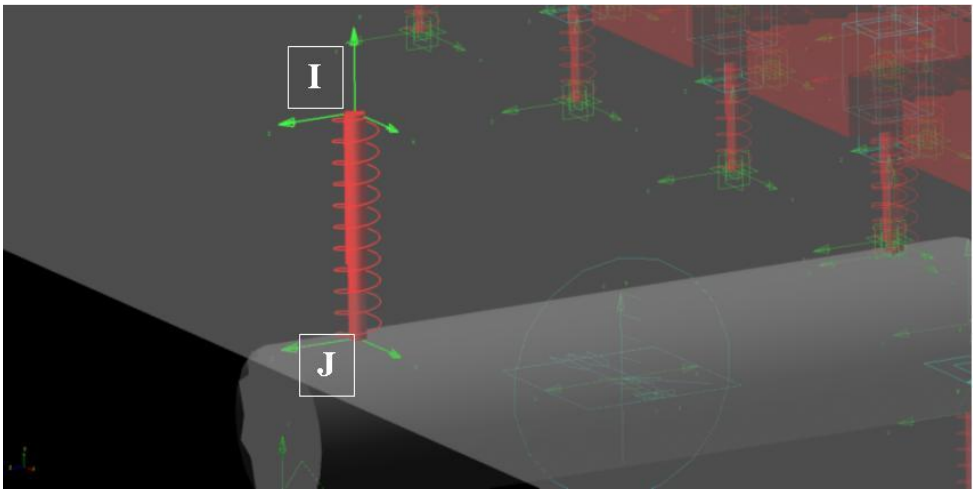



**Fig. 13.**
*SPRING-DAMPER element with markers*

The force generated by *SPRING-DAMPER* element is defined by Eq. (3)

$$F = \begin{cases} -b(dz/dt) - c(z - L_p) + F_p & za \quad z \neq L_p, \\ F_p & za \quad z = L_p, \end{cases} \tag{3}$$

where in:

*z* - distance between the element supports;
*dz/dt* - relative velocity of element supports motion;
*b* - shock absorber damping coefficient;
*c* - spring stiffness;
$L_p$ - preloaded length of springs (reference length of spring);
$F_p$ - preloaded spring force (reference force in the spring);

The dominant oscillatory motions of solid axle of the vehicle with elastic tires are vertical displacement and angular displacement around the longitudinal CG axis of the axle. The dominant oscillatory motions of the rail vehicle bogie are vertical displacement and angular displacements around bogie longitudinal and transverse CG axis. Reduction to the equivalent axle during the development of the general oscillatory model allows ignoring the angular displacement around the transverse bogie CG axis (Fig. 11). In the ADAMS/View software, axle of the road vehicle or rail vehicle bogie is represented as a rigid body with the appropriate mass and moments of inertia (see Fig. 12). Using two joints - *Inline Primitive Joint* and *Parallel axes Primitive Joint*, the motion of the axle rigid body within the general oscillatory model is limited to vertical motion and angular motion around its longitudinal CG axis (Fig. 14).

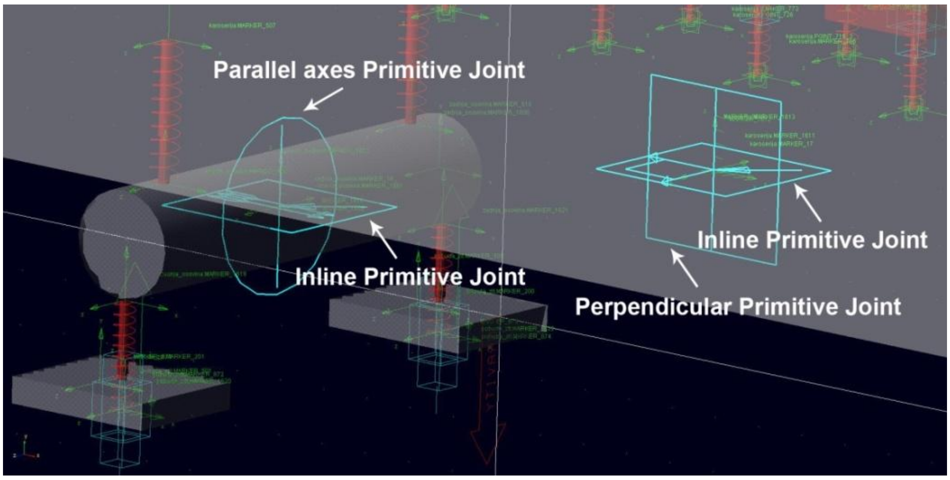



**Fig. 14.**
*Joints introduced to restrict the motion of rigid bodies (vehicle body and axle/bogie) of the general oscillatory model*

Modern buses are equipped with an air suspension system. The elastic element of such a system is the air spring support. Fig. 8 shows a typical bus pneumatic suspension system. The body is suspended by the air springs and hydraulic telescopic shock absorbers at the front and rear axles of the bus. Front axle usually has two air springs and four telescopic shock absorbers and the rear axle has four air springs and four telescopic shock absorbers (Fig. 8(a) and Fig. 8(b)). The axles are also connected to the structure by corresponding connection rods to prevent unwanted motions.

The body of the vehicle, when considered as a rigid, has six degrees of freedom (DOF). However, due to the type of joints, some motions of the body can be neglected (eg. angular motion around the vertical axis of the vehicle center of gravity (yaw)). The dominant oscillatory motions are the vertical displacement (up and down), the angular motion around the transverse CG axis (pitching) and angular motion around the longitudinal CG axis (rolling) (Dedović, 2004). In order to take into account dominant oscillatory motions already mentioned, the body of the general oscillatory model of vehicle is connected to the fixed part *GROUND* (road/track surface) through two joints - *Inline Primitive Joint* and *Perpendicular Primitive Joint* (Fig. 14). The combination of these two joints allows translation of the body in vertical direction and angular body motion around the longitudinal and transversal CG axis.

To introduce the oscillatory excitation in the general oscillatory model (recorded or modeled longitudinal roughness of road or track), four rigid bodies are defined as so-called '*false masses*', connected to the road/track surface (part *GROUND*) by translatory joints (Fig. 15). The oscillatory excitation can be introduced to translatory joints by using the function *CUBSPL* (Renguang, 1997). It should be noted that the defined rigid bodies of the oscillatory model, i.e. false masses, in case of a rail vehicle, represent the left and right solid wheels of the equivalent front and rear rigid axle.

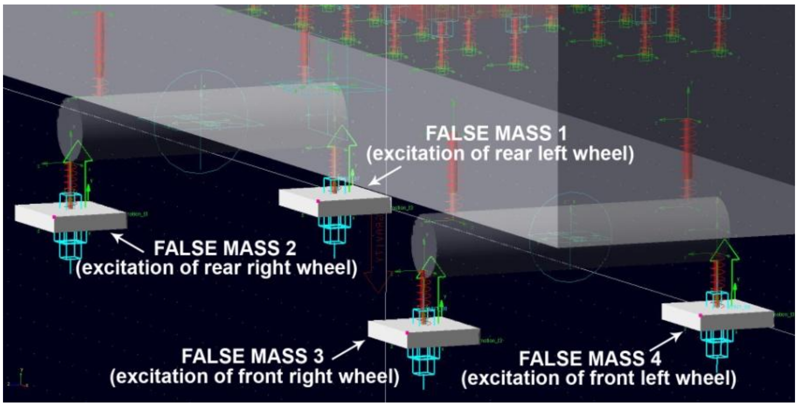



**Fig. 15.**
*False masses and translational joints*

### 3. General oscillatory model

Comparative analysis shows that among the elements that influence the oscillatory behavior of road and rail vehicles, in terms of vehicle dynamics, there is a great similarity, which allows defining and building a single general oscillatory model. Fig. 16 shows such a vehicle model built in ADAMS/View software.

The values of the vehicle geometric parameters (wheelbase, height, width and length of the body, the users seat position), the vehicle oscillatory parameters (stiffness and damping characteristics of each individual suspension), as well as the values of parameters of masses (seats, body, axles, moments of inertia of the body and axles, etc.) can be easily changed in the software.

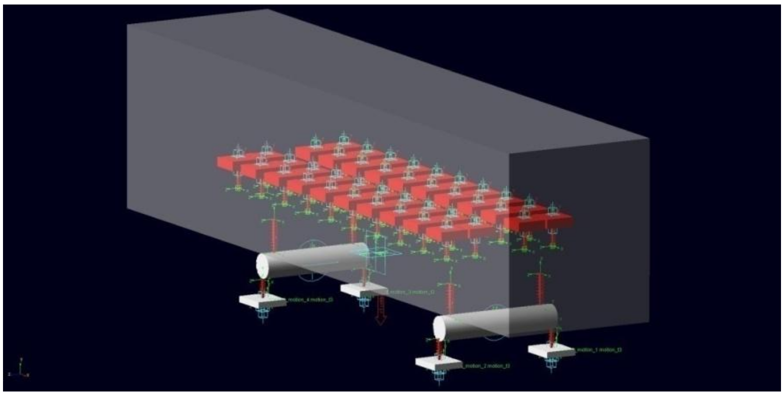


**Fig. 16.**
*General oscillatory model for the analysis of vehicle oscillatory behavior*

Based on the general oscillatory model, using the appropriate commands for tuning different parameters of the vehicle, it is possible to form oscillatory models for specific buses. An example is a double-decker tourist bus, Fig. 17. Fig. 18 shows the oscillatory model of this bus, allowing the analysis of oscillatory comfort on the passenger seats installed on both vehicle decks.

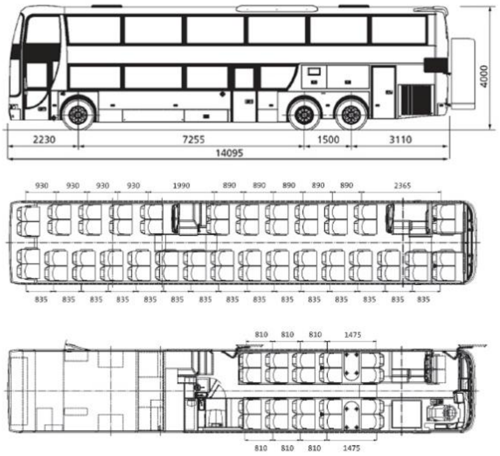



**Fig. 17.**
*Double-decker bus*

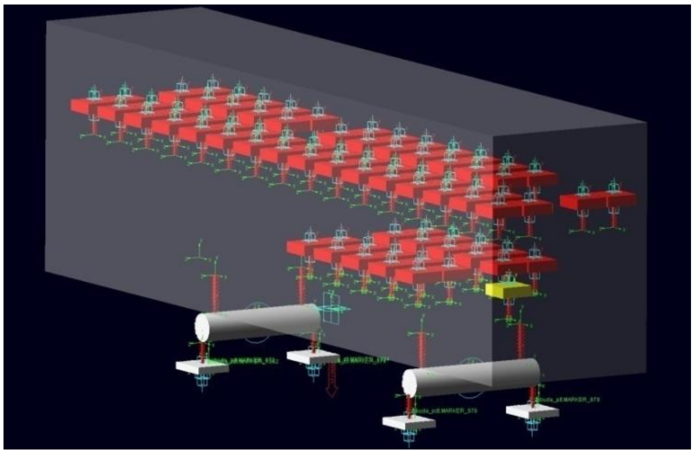


**Fig. 18.**
*Oscillatory model of double-decker bus*

With minor modifications, the general model can be improved in order to approach the real vehicle. Paper (Sekulić, 2013) presents the built of an original spatial oscillatory model of intercity bus IK 301 with 65 DOF, with a system of elastic suspension modeled more precisely. This original model has been used as a specific groundwork for determination of the '*approximately equal oscillatory comfort zones*' defined and more closely explored in this work.

### 3.1. Application of the general oscillatory model with the IK 301 bus

Fig. 19(a) shows the seat layout in the intercity bus IK 301. The bus is equipped with driver and co-driver seats, as well as 53 passenger seats. Fig. 19(b) shows the general oscillatory model with seat layout that corresponds to the one on the IK 301 bus. All the parameters of the model (geometry, oscillatory and mass parameters) are adjusted according to the relevant parameters of the actual bus IK 301. The values of the parameters are derived from the available literature (Mladenović, 1997; Nijemčević, et al., 2001; Simić, et al., 1979).

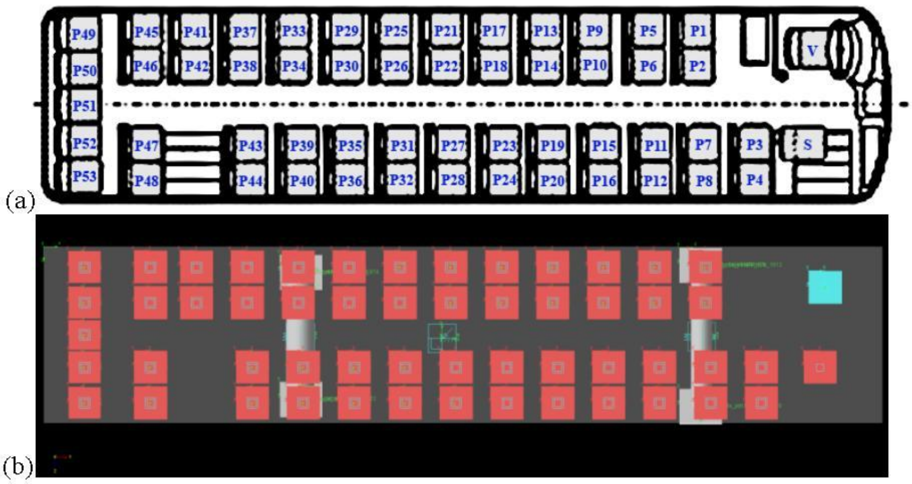



**Fig. 19.**
*Seat layout of a) intercity bus IK 301 and b) IK 301 oscillatory model*

Elements of the IK 301 oscillatory model (total number of rigid bodies, joints introduced and forces) are listed in table 1. Table 2 lists an overview of the joints introduced between the rigid bodies of the bus IK 301 oscillatory model and DOFs excluded by these joints.

**Table 1**
*Elements of the bus IK 301 oscillatory model*

| **Overview of the bus oscillatory model** | **Values** |
|---|---|
| Rigid bodies | 62 |
| *Joints between rigid bodies:* | |
| Inline joint | 3 |
| Parallel axis joint | 2 |
| Perpendicular joint | 1 |
| Translational joint | 59 |
| Total number of joints | 65 |
| *Force elements in model:* | |
| *SPRING-DAMPER* elements | 63 |
| Motions introduced - *CUBSPL* function | 4 |

**Table 2**
*Summary of joints of the oscillatory model bus IK 301 and the degrees of freedom which the introduced joints exclude*

| | | **Rotational DOF removed** | | | |
|---|---|---|---|---|---|
| | | 0 | 1 | 2 | 3 |
| **Translational DOF removed** | 0 | | Perpendicular Primitive Joint | Paralel Axes Primitive Joint | |
| | 1 | | | | |
| | 2 | Inline Primitive Joint | | | Translational Joint |
| | 3 | | | | |

The expression (4) represents *Gruebler*'s equation used to determine the number of degrees of freedom (*DOF*) of a mechanical systems composed of multiple interconnected rigid bodies (Mechanical Dynamics, Incorporated, 2001).

$$DOF = (\text{number of movable parts} * 6) - \sum_i (\text{joint}_i * \text{constraints}_i) - \sum_j \text{motion}_j \qquad (4)$$

According Gruebler's equation, the number of degrees of freedom of the particular IK 301 bus oscillatory model is equal to the:

$$DOF = (62*6) - (3*In\_J*2 + 2*PA\_J*2 + 1*Pe\_J*1 + 59*Tr\_J*5) - 4*Tr\_M = 62$$

where in:

*In_J* - Inline Joint;
*PA_J* - Parallel Axes Joint;
*Pe_J* - Perpendicular Joint;
*Tr_J* - Translational Joint;
*Tr_M* - Translational Motion;

### 3.2. The oscillatory excitation of the bus IK 301 model

Fig. 20 shows an actually recorded signal of the roughness of a asphalt-concrete pavement in good condition as a function of time. Roughness is registered on the road section of 161 m and at the speed of the of 80 km/h (RoadRuf Software, 1997). Measuring vehicle has recorded roughness on two tracks.

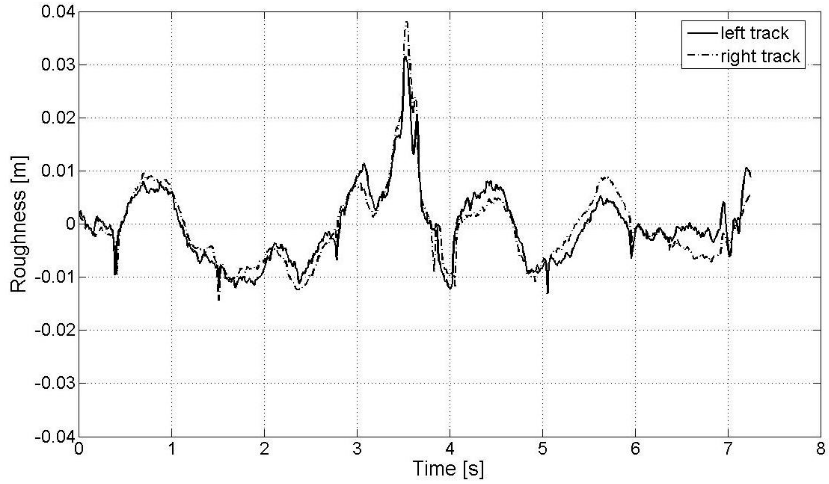



**Fig. 20.**
*Bus oscillatory excitation - good asphalt-concrete, speed of 80 km/h*

The signal of road roughness is introduced as the oscillatory excitation to translatory joints of the oscillatory model by using *CUBSPL* functions (Fig. 15).

### 3.3. Simulation and some results

For the purpose of numerical integration the Gear Stiff (*GSTIFF*) integrator is selected, with a formulation *I3*. *GSTIFF* integrator uses a backwards differentiation formula and *Newton-Raphson* algorithm for the numerical integration of the differential equation (Mechanical Dynamics, Incorporated, 2000). Simulation time was set to 7 seconds. Acceleration signals were selected every 0.001 second. The simulation is carried out so the bus oscillatory model has been set in equilibrium position by using the command "*Find static equilibrium*", and then dynamic simulation was performed.

In Fig. 21 the accelerations obtained by the simulation for three different locations of the bus occupants have been shown - on the driver seat and on two passenger seats (passenger22 and passenger53), in directions of *x*, *y* and *z*-axis. Occupants suffer dominant acceleration along the *z*-axis. Four seconds after the beginning of simulation, vertical accelerations have the highest intensity for the entire simulation process. The reason for this is the left and right wheels of the bus pass over the bump whose height is approximately 4 cm (Fig. 20). Passenger in the back of the bus (seat 53) has the highest peak value of the vertical acceleration and it is approximately -6.0 m/s$^2$.

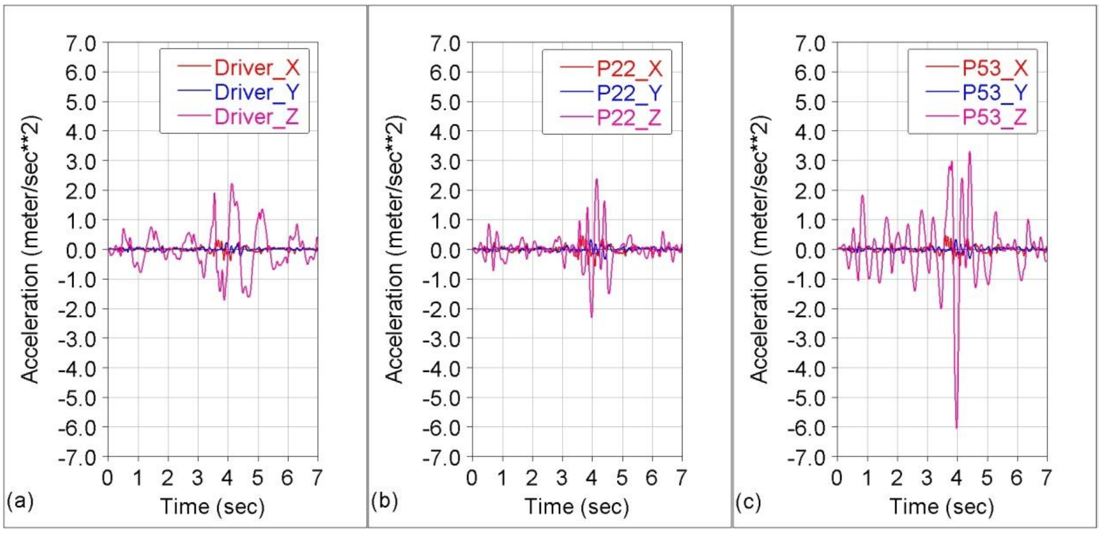



**Fig. 21.**
*Changing the acceleration of a) driver b) passenger22 and c) passenger53*

Fig. 22 shows the change of vertical reactions between road and bus wheels. All wheels during the simulation remain in contact with the ground.

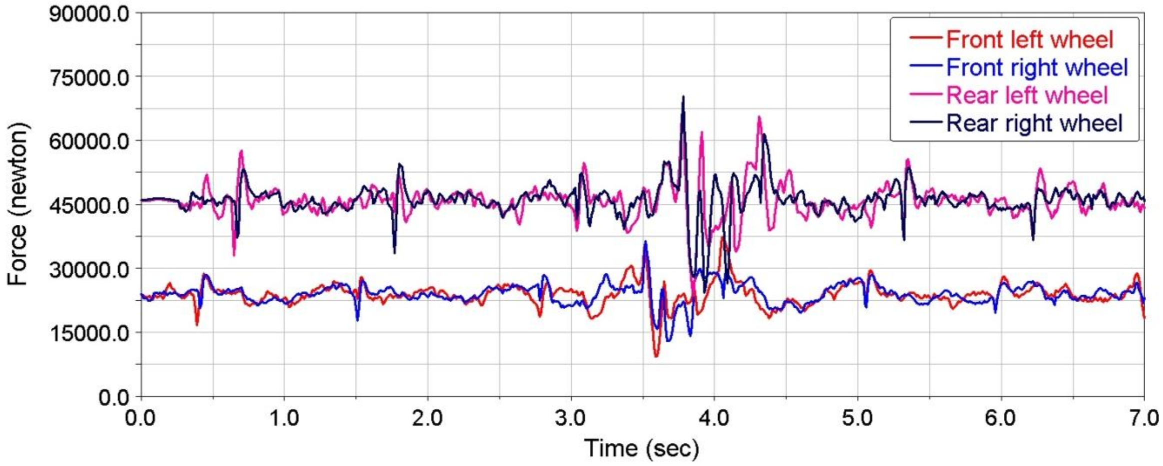



**Fig. 22.**
*Vertical reactions between road and bus wheels*

**4. Conclusion**

On the base of a comparative analysis the paper shows that, in context of vertical dynamics, there is a considerable similarity between road and rail passenger vehicles, with all their main elements: vehicle body, user seats, axles and suspension system. These elements have a dominant influence to the oscillatory behavior of the vehicle. Therefore, it is possible to define a general oscillatory model for simulation study of the oscillatory behavior of the vehicle, as well as its passive and active safety. The authors have designed such a general oscillatory model in *ADAMS/View* module of the *MSC.ADAMS* software.

The values of vehicle masses parameters, as well as geometric and oscillatory parameters, can be changed by using the corresponding commands in the ADAMS/View software. This enables the adjustment of oscillatory model for the various kinds and types of vehicles.

A general oscillatory model has been used in this paper to simulate the dynamic behavior of the vehicle and the parameters that characterize a modern intercity bus were applied.

It has been shown that the actually recorded signal of road roughness can be used for the oscillatory excitation of the model, which is significant for more accurate simulation results. The model oscillatory excitation was the record of the actual roughness of a good asphalt-concrete pavement at a speed of 80 km/h.

It has been proven that the general model presented allows studying the oscillatory comfort of each user of the vehicle (driver and each passenger). In this paper, as examples, the vertical and horizontal accelerations for three vehicle users were presented (driver, one passenger in the middle and one in the back of the bus), as well as vertical reactions between road and bus wheels. The information on accelerations that the vehicle users were exposed to, obtained by simulation, is an important prerequisite for mapping the comfort of the vehicle platform. Using the data resulting from simulations, the '*approximately equal oscillatory comfort zones*' has been defined.

By using the presented model and the same methodology, besides accelerations and vertical reactions, one can analyze other relevant oscillatory parameters related to displacements, velocities, accelerations and forces in the system (e.g. deformation of elastic elements of the suspension system, radial deformation of the tire, vertical displacement and acceleration of axles, forces in elastic and damping elements, etc.).

**Acknowledgements**

Support for this research was provided by the Ministry of Education, Science and Technological Development of the Republic of Serbia under Grant No. TR36027. This support is gratefully acknowledged.

**References**

Dedović, V. 2004. Dinamika vozila [*Vehicle dynamics*], Belgrade: University of Belgrade, Faculty of Transport and Traffic Engineering. 244 p. (in Serbian).

Eriksson, P.; Friberg, O. 2000. Ride comfort optimization of a city bus, *Structural and Multidisciplinary Optimizatoin,* 1(20): 67-75.

Genta, G. 1997. *Motor vehicle dynamics: modeling and simulation*. World Scientific Publishing Co. Pe. Ltd. Singapore. 539 p.

Iwincki, S. 2006. *Handbook of Railway Vehicle Dynamics*. Taylor&Francis Group, LLC. USA. 552 p.

Mechanical Dynamics, Incorporated 2001. *Advanced ADASM/View Training Guide,* Mechanical Dynamics, Ann Arbor, Michigan. 370 p.

Mechanical Dynamics, Incorporated 2000. Building Models in ADAMS/View, Available from Internet: <http://www.mme.uwaterloo.ca/~me321/BuildingModels.pdf.>.

Miege, A. 2004. *Tyre Model for Truck Ride Simulations,* Cambridge: University of Cambridge, Departement of Engineering. 57 p.

Mladenović, D. 1997. An analysis of constructions parameters impact on the oscillatory behaviour of the bus. Master thesis, Faculty of Transport and Traffic Engineering, Belgrade. 306 p.

Nijemčević, S.; Dragojlović, D.; Zečević, S.; Milosavljević, B. 2001. Technical-selling book. Ikarbus AD, fabrika autobusa i specijalnih vozila, Beograd, 237 p.

Pečeliūnas, R.; Prentkovskis, O.; Garbinčius, G.; Nagurnas, S.; Pukalskas, S. 2005. Experimental research into motor vehicle oscillations in the case of changeable deceleration, *Transport,* 5(20): 171-175.

Renguang, D. 1997. Available from Internet: <http://www.mscsoftware.com/support/library/conf/adams/na/1997/uc970025.pdf>.

RoadRuf Software, 1997. *Software for Analyzing Road Profiles,* Ann Arbor: The University of Michigan Transportation Research Institute.

Sekulić, D. 2013. Investigation of passengers' oscillatory comfort in the bus with respect to the seat position and quality, Ph.D. thesis, Faculty of Transport and Traffic Engineering, Belgrade. 235 p.

Sekulić, D.; Dedović, V. 2011. The effect of stiffness and damping of the suspension system elements on the optimization of the vibrational behavior of a bus, *IJTTE - International Journal for Transport and Traffic Engineering,* 4(1): 231-244.

Sekulić, D.; Dedović, V. 2011. Intercity bus users vibration comfort analysis through an oscillatory model with seven DOF using ADAMS/View software, *Journal of Applied Engineering Science*, 3(9): 401-410.

Simić, D.; Savičević, A.; Ninković, D. 1979. Comparative investigation of oscilatory parametars of driver's seat, *MVM - Motorna Vozila Motori* (24-25): 7-68.

Tianfei, M.; Ankang, Z.; Dengfeng, W.; Shumingi, C.; Junchao, D. 2010. Ride Comfort Analysis of Commercial Vehicle Using Flexible Multi-body and Finite Element Methods, In *Proceedings of the 2nd International Conference on Mechanical and Electrical Technology,* 295-299.

*Urban Transportation Planning*

# AN OVERVIEW OF BUS AND MINI BUS PUBLIC TRANSPORT SERVICE QUALITY IN SOUTH AFRICA: A CASE STUDY OF JOHANESBURG

**Krishna K. Govender**[1]
[1]*University of KwaZulu-Natal, Pietermaritzburg, South Africa and Regenesys Business School Johannesburg, South Africa*

**Abstract:** This paper reports on commuters' perceptions of bus and mini bus taxi service using an alternate to the most popular and commonly used SERVQUAL instrument, namely RECSA. From face-to-face interviews using a structured questionnaire to survey a convenience sample of 690 commuters at specifically selected mini-bus taxi and bus ranks using a commuter intercept survey, it was ascertained that overall the perceived quality of public bus transport services exceeded that of minibus taxis, despite the minibus taxis being the dominant mode of public transport. All the RECSA dimensions of transport service quality influenced the respondents' perception of public bus service quality, whereas only three, namely reliability, affordability and extent of the service, influenced their perception of the minibus taxis service quality.

To improve public road transport service quality, service providers should among others, implement scheduling systems to improve the punctuality of the service, invest in communication systems, introduce a comfort rating system, improve the arrival times at destinations and reduce journey length. They should also improve the condition of the minibus taxi shelters, increase the frequency of the service on certain routes especially during peak periods on weekdays, take commuters closer to their destination through modal integration and elimination of transfers, improve commuter safety, more especially by changing driver behaviour through focused safety and driver training programmes, and improve affordability and provide value for money service.

The findings could also serve to inform public transport policy makers and providers, since a public transportation model and policy based on improving the perceived service dimensions is likely to increase the demand for public bus and minibus taxi transport and, reduce the use of private motor cars, thereby addressing the public road transport conundrum in Johannesburg in particular, and South Africa in general.

**Keywords:** public transport service; minibus taxi service; public bus service; transport service quality.

## INTRODUCTION

Several researchers (Javid, Okamura, Nakamura and Wang, 2013) argue that the rapid increase in urban population and automobile ownership and usage have resulted in urban transportation problems in developing countries. According to several South African studies (National Planning Commission, 2011; Mokonyama, 2012), the quality of public transport service in South Africa requires urgent improvement, since it affects mostly poorer members of the community who rely on it for daily commuting. Several studies in developing countries (Pucher, Korattyswaroopam and Ittyerah, 2004; Ngatia, Toshiyuki and Fumihiko, 2010; Mashiri, Moeketsi and Baloyi, 2010; Finn and Mulley, 2011) confirm that the poorer members of the community are often faced with inadequate transport service, poorly arranged schedules, the absence of facilities – including bus stops and shelters, and the infrequency of services, particularly at off-peak times, which severely compromises the convenience of these services. As was revealed from the National Household Travel Survey (2003) 71% of train users, 55% of taxi users and 54% of bus users were dissatisfied with the level of crowding (Gauteng Province 2009). In addition, 74% of bus users, 64% of taxi users and 53% of train users were also unhappy with the facilities at stops, ranks and stations. Furthermore, there are excessive delays for public transport with an average waiting time of 40 to 65 minutes (Gauteng Province 2009). From the aforementioned, it is safe to state that South African commuters are clearly unhappy with public transport service quality.

The aforementioned implies that public passenger transport service organisations (in South Africa) must conduct perception studies on an ongoing basis in order to enable them to meet the needs of passengers, and McKnight *et al*. (1986) further asserts that the perception studies and results must be used effectively to deliver quality service that meets the needs of passengers.

In light of the above, this paper reports the response to the following key research questions with respect to bus and mini-bus taxi service, namely: What are the commuters' perceptions of the quality of the service? To what extent would their perceptions of the service influence their future demand for buses and minibus taxis? What is the importance of each service quality dimension in choosing a public transport mode? To what extent do the service quality dimensions influence passenger transport mode choice? What is the importance of the service quality dimensions in choosing a mode of passenger transport?

## TRANSPORT SERVICE QUALITY

Service quality in the public transport sector has remained an elusive and a much neglected area of study. Data regarding the quality and performance indicators of public transportation services are vaguely determined and, in fact, are practically nonexistent (Simona 2010), and much of the debate has centred around the system itself: spatial designs, systems configurations, city network developments, government policies, and engineering services. Generally, service

[1]Corresponding author: krishnag@regenesys co.za

quality has remained a challenge for the majority of public transport organisations, because of the challenge inherent in measuring service quality (Zeithaml and Bitner 2000). Furthermore, although a major challenge confronting public transport in general is that of service quality, it (service quality) is a complex area of study and measuring service quality in public transport in particular, is made even more difficult by the subjective nature of service (McKnight *et al.* 1986; Parasuraman, Zeithaml & Berry 2000).

In order to get a sense of the challenges facing public transport service quality in general, some reference to studies from the developing and developed world is provided. For example, research conducted in Singapore in 2012 revealed that there are four important dimensions that should be considered in public transport service, namely, connections, extent of service, liveable cities, and inclusivity, all of which enhances commuters travelling experience (Land Transport Authority, 2013). A study in Scotland by the Department of Transport identified more than 30 different service attributes ranging from the punctuality and reliability of the service to the cleanliness of stations. These attributes were considered important to passengers and caused dissatisfaction if not delivered to a satisfactory standard, with public transport organisations deeming it important to try their best to meet passengers' reasonable needs at all times (Samson & Thompson, 2007).

From a European on public transport study conducted in 2009 (Simona, 2010), it was evident that data regarding quality indicators of public transportation services were vaguely determined and practically non-existent. With respect to the developing countries, a study conducted in Kenya in 2010 (British Youth Council, 2012) showed regulations, improvements in comfort levels (from less overcrowding) and, safety of passengers as important service attributes which led to more attractive public transport. In Lagos Nigeria, the following service quality attributes contributed to improving public transport system: cheaper fares, travel time was reduced, waiting time at bus stops fell during peak and off-peak hours, improved safety and reliability of the new system, reduction of externalities: fuel consumption for vehicles used along the corridor, demonstrating the undeniable advantage of formal public transport (International Association of Public Transport, 2010). As regards the perception of service, only 33.5% of the Nigerian respondents felt that buses in their area were reliable whereas, while a marginal more (35.8%) disagreed (British Youth Council, 2012:2).

Some researchers (Beirao and Cabral, 2007; Eboli and Mazulla, 2007) assert that evaluating and measuring transport service quality remains challenging and important, since transport service quality comprises abstract and intangible constructs such as comfort and safety. Generally, most of the research on service quality has been conducted using the renowned SERVQUAL (Parasuraman, et al. 1988) which uses the RATER (responsiveness, technology, empathy, reliability) dimensions of service quality, or a modification thereof because of inter-alia, features such as the simultaneous measurement of expected and perceived quality, and the ease of interpretation of its results (Barabino, Deiana and Tilocca, 2012). Despite the aforementioned researchers citing and acknowledging that limited studies exist, inter-alia, Prasad and Shekhar (2010), Too and Earl (2010) where attempts were made to adapt and adopt SERVQUAL to evaluate perceived service quality among rail and bus passengers respectively (Barabino et al. 2010: 241; Randheer, Al-Motawa and Vijay, 2011), the SERVQUAL instrument has not been without criticism (Buttle, 1995; Lages and Fernandes, 2005), inter-alia, its appropriateness for measuring service quality across service institutions. In light of the aforementioned, and due to the complexities of measuring service quality in public transport, this paper reports on a study conducted to explore commuters' perceptions of bus and minibus taxi service in terms of the RECSA service quality dimensions of McKnight *et al.* (1986). Table1 below is self-explanatory and depicts the five service quality dimensions of both the RATER and RECSA service quality models. It is apparent that although RECSA incorporates some elements of RATER, it is more appropriate for measuring transport service quality.

**Table 1:**
*Service Quality Dimensions and Attributes*

| RECSA Service Quality Dimensions | Cluster of Service Attributes | RATER Service quality Dimensions | Service Attributes |
|---|---|---|---|
| Reliability | Arriving on time<br>Notification of delays<br>Waiting away from home<br>Delays en-route | Reliability | The ability to perform the promised service in a dependable and accurate manner |
| Extent of the service | Total hours of service<br>Service on weekends<br>Service on public holidays<br>Service on weekdays<br>Service in the evening | Assurance | Knowledge and courtesy of employees and their ability to convey trust and confidence |
| Comfort | Guaranteed seat<br>Smooth ride<br>Sheltered waiting areas<br>Air conditioning | Tangibility | The physical evidence of the service, eg. the appearance of the personnel and physical facilities, and equipment used to provide the service |
| Safety | Low probability of accidents<br>Low probability of falling<br>Low probability of assault | Empathy | Caring, individualized attention provided to customers |
| Affordability | Alternatives –season tickets<br>Cheap fares<br>Value for money | Responsiveness | The readiness and willingness to help customers in providing prompt timely services |

*Source: Adapted from McKnight et al. (1986); Parasuraman et al. (1988)*

## RESEARCH METHODOLOGY

A descriptive research design was used for collecting data since several researchers Anastas (1999) and Kirshenblatt-Gimblett (2006), argued that descriptive studies collect a large amount of data for analysis that can yield rich findings and lead to important conclusions and recommendations. All (800 000) daily commuters in the City of Johannesburg comprised the population of interest (Johannesburg Development Agency, 2012), however primary data was collected from commuters who only used public buses and minibus taxis. Geographical sampling was used, in terms of which all identified bus terminals were grouped into a homogeneous cluster for example, terminals in townships were grouped together to create one cluster, terminals in the suburban areas were grouped together to create another cluster, and terminals in the CBD were grouped together to create a final cluster to ensure the homogeneity of the subjects in each of the three clusters (Cooper & Schindler, 2001). All the homogenous clusters were included in the sample. The survey was conducted during peak periods in the morning and the afternoon, over a period of one month, at the following Johannesburg bus and min-bus taxi terminals: Central Business District; Johannesburg Suburbs; South of Johannesburg. As a result of the complexities involved in the process of sampling in the public transport context, the choice of a probability sample is always a challenge (McKnight, *et al*. 1986; Cooper and Schindler, 2001), as was demonstrated in the research conducted by the Burbidge City Bus Company. In the aforementioned study, non-probability sampling techniques were used to select the sample, thus the selection of bus terminals using a map and choosing a sample from the list of bus terminals concerned was deemed to be a realistic and acceptable solution to overcoming the sampling challenge (Cooper and Schindler, 2001).

A structured questionnaire was developed and used in the survey and the same questions were asked of all commuters but, mini-bus taxi or public buses, were substituted where necessary. A five point Likert scale (See Appendix A) was used since it is most commonly used when semantic differentials and multiple Likert type questions in a questionnaire on a scale (1-5), and in order to determine if the scale is reliable (Nagel, 2007). The questions were divided into five categories, each of which was used to measure the service quality constructs, and separate scales were constructed for buses and minibus taxis, respectively based on the five dimensions, namely Reliability, Comfort, Extent of the Service, Safety and Affordability. The five scales were then combined into a cumulative index which was termed the Perception of Service Quality Index (PSQ). Each scale was weighted equally in the index, so that it was possible for index scores to range from 0 to 50. In view of the fact that respondents had been asked their opinions about the importance of each of the five categories, it was also possible to create a cumulative index in which the individual categories were weighted according to their perceived importance for each individual. The result was termed the Importance-Weighted Perceptions of Service Quality Index.

The Statistical Package for the Social Sciences (SPSS) was used to analyze the data (Coakes and Steed, 2003). The Jarque-Bera Test was used to assess whether the PSQ Index data followed a normal distribution. For both buses and minibus taxis, the null hypothesis of normality was rejected with p-values of 0.000017 and 0.00015 respectively, which meant that it was not possible to use a t-test to compare the mean perceived service quality of buses and that of minibus taxis and therefore a non-parametric test the Mann-Whitney test was used to compare the medians. For the individual service quality dimensions, except Service, as well as the overall index (and weighted index), it was possible to draw a definite conclusion that the perceived quality of bus transport exceeded that of minibus taxis by a significant margin. Multiple linear regression analysis was used to determine the relationship between the Perceived Quality of the Service Index for buses and mini-bus taxis, and the demographic variables. The study used the test of significance of 95% confidence level, with a 5% margin of error. Generalised Linear Regression analysis was also used to test if the data fitted the model (Bollen, 1989; Byrne, 2010).

## RESEARCH FINDINGS

The measurement scales were assessed for internal consistency using the Cronbach's alpha which coefficient reflects how closely related a set of items are as a group. The values were all above 0.64, indicating an acceptable level of internal consistency for all the scales (Byrne, 1989). It became apparent that only 29.8% of the respondents used public buses while the vast majority (70.2%) used minibus taxis as a preferred mode of transportation. A large percentage (30.3%) of public transport users were aged 19 years, and 28.7% were between the ages of 25 and 34 years. Although public buses and minibus taxis were used predominantly by students and scholars, their preferred mode of transport was the minibus taxi. The aforementioned finding reaffirms what has been reported by some researchers such as Thomas, Ryneveld and Pascarel (2010) and (Ndebele, 2011) and in national reports (South Africa, 2012).

With regard to service quality, there were significant differences in the respondents' perceptions across age groups, especially respondents in the " under 19 years" , "20-24 years" , "25-34 years", "35-50 years" when these groups were compared with respondents in the "51 years and above" group. In addition, there were significant differences with respect to the overall service quality perceptions among the respondents in the different education and income groups. Age, gender and level of education were all statistically significant at the 5% level. Age had a negative coefficient, indicating that, on average, younger people have a more favourable perception of the minibus taxis than older people. Gender had a positive coefficient, indicating that on average, women perceived minibus taxis service more favourably than men. On the other hand, educational level had a negative coefficient, indicating that less educated people viewed minibus taxis favourably compared to more highly educated people.

In view of the fact that the indices did not follow a normal distribution according to the Jarque-Bera test, the non-parametric Spearman's correlation method was used in preference to the Pearson's method. The correlation coefficient between the Perceived Quality of Service Index for buses and the Perceived Quality of Service Index for minibus taxis (for those who offered opinions on both modes of transport) was 0,068 and not statistically significant (p-value = 0,41). This suggests that the respondents' opinion of bus service quality was independent of their opinion of minibus taxi service.

The generalised linear regression model was also considered, since it caters for categorical and continuous explanatory variables and a response variable that may or may not necessarily be normally distributed. Since, the average of several variables was taken to calculate the overall service quality, it is a well-established fact that the response variable will be normally distributed. The fitted model for public buses was: Overall service quality=$\beta_0$+ $\beta_1$ *age+ $\beta_2$*gender+ $\beta_3$*occupation+ $\beta_4$*education+ $\beta_5$*income+$\varepsilon$

It was evident that age, education and income were all significant (at the 5% level) in influencing the overall service quality provided by the busses, since their p-values are all less than 0.05. However, gender and occupation were not significant in influencing the respondents' perception of the overall service quality provided by the bus. There were however differences at the 5% level of significance, among the education groups when comparing all the respondents in the group with some form of education and their perception of the overall service quality, since the p-values were less than 0.05. It was further evident that there were significant differences at the 5% level (since the p-values were less than 0.05) in the education group between respondents with less than matric, matric and matric plus a tertiary qualification groups when comparing all those groups to respondents in the group with some other form of education, with respect to the overall service quality. There were also differences in those who were in the “0-R1000” and “R5001-R6000” income bracket, when comparing to these groups of respondents to those that were in the “Above R6000” group, with respect to their perception of the overall service quality provided by public busses.

The Generalized Linear Model fitted for mini-bus taxis was: service quality=$\beta_0$+ $\beta_1$ *age+ $\beta_2$*gender+ $\beta_3$*occupation+ $\beta_4$*education+ $\beta_5$*income+$\varepsilon$ . It became evident that age, education and occupation are all significant at the 5% level in influencing the overall service quality provided by the minibus taxis, since their p-values are all less than 0.05. However, similar to busses, gender and income were not significant in influencing the overall service quality provided by the minibus taxis. It was also evident that there were significant differences at the 5% level, (since the p-values were less than 0.05) in the age between respondents in the “under 19yrs”, “20-24 yrs”, “25-34 yrs”, “35-50 yrs” when comparing all these groups to respondents in the “51 yrs and above” group with respect to their perception of the overall service quality provided by mini-bus taxis. There were also differences between those who were in the “0-R1000” income group and those that were in the “above R6000” income group with respect to the overall service quality.

## PERCEPTIONS OF SERVICE QUALITY

Table 2 presents the summary statistics in respect of the importance attached to each of the five dimensions of quality of service, on a scale of 1 (less important) to 5 (very important). Table 6 reveals that all five dimensions were considered very important, and equally so.

**Table 2**
*Importance of Service Quality Dimensions*

| | Buses | | Minibus Taxis | |
|---|---|---|---|---|
| | **Mean** | **Median** | **Mean** | **Median** |
| Reliability | 4,3 | 5,0 | 4,4 | 5,0 |
| Comfort | 4,4 | 5,0 | 4,2 | 5,0 |
| Service | 4,3 | 5,0 | 4,3 | 5,0 |
| Safety | 4,3 | 5,0 | 4,3 | 5,0 |
| Affordability | 4,4 | 5,0 | 4,3 | 5,0 |

It also became evident that, buses were perceived as being more reliable than minibus taxis, and those who used public buses more often tended to have a higher opinion of the quality of the bus transport service, while those who used minibus taxis as their primary mode of transport did not do so because they had a high opinion of the quality of the minibus taxi experience, since the coefficient for the mode of transport used the most often was not significant. Therefore, it can be concluded that those who use minibus taxis tended to do so for reasons other than their opinion of this mode of transport, for example, they may use the minibus taxis for the convenience of this mode of transport as opposed to safety and reliability.

With regard to the punctuality of the service and communications, it became evident punctuality was an important factor which influenced the respondents’ perception of the overall service quality, while communication was not as important. Despite the study findings, communication is likely to influence the overall service quality, since some literature (Mashiri, *et al.* 2010) revealed the importance of information and communications technology systems to improve communication with passengers. It also became evident that timetables were important for both bus and minibus taxi commuters, which finding is also supported by the literature which showed that timetables were important to the scheduling process, because the dissemination of information to passengers is critical to the successful operation of public transport services, and in maintaining and stimulating demand (Mashiri, *et al.* 2010).

The study also revealed that timely arrival at the destination (Reliability) is an important factor influencing passengers’ perception of the transport service. Furthermore, the time taken to arrive at the destination is important in explaining the utilisation of minibus taxi service. However, contrary to what is reported in the literature, adherence to routes did not emerge as an important factor influencing the commuters’ perception of either bus or minibus taxi transport service. For example, some researchers (Shaibani, 2005) maintained that adherence to the route is important because it improves the predictability of public transport while enabling passengers to plan effectively.

The Extent of service was one of the top five important factors that correlated positively with the perception of service on the part of both bus and minibus commuters, and taking passengers to their exact destination was shown as being an important factor in influencing passengers’ perceptions of the service. Public transport accessibility and availability were also perceived as being important factors in commuters’ perception of public transport service. Availability of the service in the evenings and on public holidays influenced the respondents’ service quality perceptions. Driver friendliness also emerged as one of the important factors that correlated highly with the respondents’ perception of service of the both bus and minibus taxi commuters.

With regard to the Comfort of service, buses were perceived as being more comfortable than minibus taxis. Furthermore, seat availability emerged as an important factor for evaluating the service of both bus and minibus taxis, while the condition of the bus shelters was less of a factor for the bus commuters than for the minibus taxi commuters. Regarding the Safety of service, it emerged that buses were perceived as being safer than minibus taxis. The safety of the service also emerged as an important factor in the commuters' perceptions of the service offered by both the buses and minibus taxis. Finally, buses were perceived as being more Affordable than minibus taxis, and the affordability of service emerged as being an important factor influencing the respondents' perceptions of the service offered by both bus and minibus taxis.

**Recommendations, Limitations of the Study and Future Research**

Public transport must be differentiated from other modes of transport, such as private vehicles, by providing a better, superior service than that offered by the other modes of transport, and by making the commuters aware of the service being provided without compromising convenience, comfort or safety. The marketing and repositioning strategy of public transport in Johannesburg, and the country at large, should focus on the RECSA dimensions as being the most important service quality dimensions which influence passengers' perception of the service and future service utilisation.

Public transport organisations and policy makers should place more emphasis on the punctuality of the service by implementing scheduling systems in order to improve the punctuality of service by implementing service planning software, such as, amongst others, MICROBUS and Mentor Streets Schedule Software Suites, which have been successfully implemented in other countries. Timetables should be implemented, since they are used as a point of reference by passengers, and are important and relevant to the overall service quality.

Since commuters will utilise public transport provided if it arrives at their destination timely, public transport organisations should seek methods to improve the timely arrival at the destination, such as increasing the frequency of service on each route, and during weekends, evenings, and on public holidays. The minimization or even elimination of transfers, and the proper integration of the public transport services should be the focus of public transport policy makers because such a strategy would completely overhaul the public transport service and, as a consequence, increase the utilisation and the demand for public transport.

Public transport drivers should also be compelled to attend customer service training, and quarterly assessments on the impact of the training on customer service should be conducted. In addition, they should be thoroughly screened prior to being employed, a qualification in customer service, experience in driving, a clean criminal and traffic offences record, and knowledge of the rules of the road should be introduced as the minimum requirements for a professional driver's permit. Journey length is often the leading cause of dissatisfaction on the part of the commuters, and the findings reveal that time taken to arrive at the destination is important for the all commuters. Research on journey length will be valuable to both academic and professionals of passenger transport.

In terms of the limitations of the study, it should be noted that the data was only collected from commuters in Johannesburg, and not throughout South Africa. Although data collection can be costly, it is often even more costly to make erroneous decisions or arrive at conclusions and generalizations based on inadequate information, weak data, insufficient data (sample too small for use in extrapolation), all of which could result in the research losing credibility. Thus, for greater generalizability data from a larger sample selected from across all provinces in South Africa could shed more light on the commuters' perceptions across the country.

**References**

Anastas, J.W. 1999. *Research design for social work and the human services - flexible methods: descriptive research*. 2nd ed. Columbia University Press Publishing, New York.

Barabino, B.; Deiana, E.; Tilocca, P. 2010. Measuring service quality in urban bus transport: a modified SERVQUAL approach. *International Journal of Quality and Service Sciences*, 4(3): 238-252.

Beirao, G.; Cabral, J.S. 2007. Enhancing service quality in public transport systems. *Transport Policy*, 14: 478-489.

Bollen, K.A. 1989. *Structural equation with latent variables*. Wiley Publishing, New York.

British Youth Council. 2012. *Public transport & young people in Suffolk: youth parliament report*. Available from Internet: <http://www.byc.org.uk>.

Byrne, B.M. 2010. Structural equation modeling with EQS: basic concepts, applications, and programmes. 2nd ed. Erlbaum Publishing, New Jersey.

Buttle, F. 1995. SERVQUAL: review, critique, research agenda. *European Journal of Marketing*, 30(1): 8-32.

Cape Regional Chamber of Commerce & Industry. 2011. *Cape Town's public transport crisis*. Railways Africa Publishing, Cape Town.

Coakes, S.; Steed, L. 2003. *SPSS Analysis without anguish*, Version 11.0 for Windows, Milton: Miley Press: 135 – 145.

Cooper, R.D. ; Schindler, S.P. 2001. *Business research methods.* McGraw-Hill Publishing, New York.

Dinmgra, C. 2011. Measuring public transport performance; Lessons for developing countries. *Sustainable Urban Transport Technical Document No. 9*. GIZ: Bonn, Germany.

Eboli, L.; Mazulla, G. 2007. Service quality attributes affecting customer satisfaction for bus transport. *Journal of Public Transport*, 10(3): 30-42.

Finn, B.; Mulley, C. 2011. Urban bus services in developing countries and countries in transition: A framework for regulatory and institutional developments. *Journal of Public Transportation*, 14(4): 89-107.

Gauteng Province. 2009. *Gauteng land transport framework*. Available from Internet: <http://www.roadsandtransport.gpg.gov.za>.

Hunt, H.K. 1979. Conceptualisation and measurement of consumer satisfaction and dissatisfaction. *Journal of Marketing Science Institute*, 77–103: 1–490.

International Association of Public Transport. 2010. *Public transport in sub-Saharan Africa-major trends and case studies.* Trans-Africa Consortium Publishing, Brussels.

Javed, M.A.; Okamura, T.; Nakamura, F.; Wang, R. 2013. Comparison of commuters' satisfaction and preferences with public transport: A case of wagon service in Lahore. *Jordan Journal of Civil Engineering*, 7(4): 461-472.

Johannesburg Development Agency. 2012. *Fast facts – Johannesburg.* JDA Publishing, Johannesburg. Available from Internet: <http://www.jda.org.za>.

Kirshenblatt-Gimblett, B. 2006. *What is research design: The context of design.* New York University, New York.

Lages, L.F ; Fernandes, J.C. 2005. The SERVPAL instrument: A multi-tem instrument for measuring service personal values. *Journal of Business Research*, 58(11): 1562-1572.

Land Transport Authority. 2013. *Land transport master plan 2013*. Department of Transport Publishing, Singapore.

Mashiri, M.A.M.; Moeketsi, P.N.; Baloyi, V. 2010. *Increasing public transport market share in South Africa: The options.* Available from Internet: <http://www.thredbo-conference-series.org>.

McKnight, C.E.; Pagano, A.N.; Paaswell, R.E. 1986. *Using quality to predict demand for special transportation, in behavioural research for transport policy.* VNU Science Press Publishing, Netherlands.

Mokonyama, M. 2012. Public transport transformation. *African Renaissance Conference on Public Transport*, 1-17.

Nagel, F. 2007. *Cronbach's alpha*. Softpedia Publishing, Romania. Available from Internet: <http://webscripts.softpedia.com>.

National Household Travel Survey. 2003. *Key results.* Department of Transport Publishing, Pretoria.

National Planning Commission. 2011. *Transportation*. The Presidency Publishing, Pretoria.

Ngatia G.J;, Toshiyuki O.; Fumihiko, N. 2010. The structure of users' satisfaction of urban public transportation service in developing countries: the case of Nairobi. Available from Internet: < http: www.easts.info/**public**ations/journal_proceedings/100237.pdf>

Parasuraman, A.; Zeithaml, V. ; Berry, L.L. 1988. SERVQUAL: A multiple-item scale for measuring customer perceptions of service quality. *Journal of Retailing*, 64(1): 12–40.

Pucher, J.; Korattyswaroopam, N. 2004. The crisis of public transport in India: Overwhelming needs but limited resources. *Journal of Public Transportation*, 7(4): 1-20.

Randheer, K.; Al-Motawa, A. ; Vijaya, P/J. 2011. Measuring commuters' perception of service quality using SERVQUAL in public transportation. *International Journal of Marketing Studies*, 3(1): 21-34.

Richard, P. 1983. *Urban Mass Transportation*. Digital Network Publishing, Boston.

Samson, R. ; Thompson, P. 2007. *Passenger focus: Priorities for improving services in Scotland*. Transport Scotland Publishing, Scotland.

Zeithaml, V.A. ; Bitner, M.J. 2000. *Services marketing: Integrating customer focus across the firm*. 2nd ed. Irwin McGraw-Hill Publishing, New York.

# MACRO-MODELLING OF URBAN REGIONS

**Kenneth Scerri[1], Dora Lee Borg[2]**
*[1,2]Department of Systems and Control Engineering, University of Malta, Msida, MSD 2080, Malta*

**Abstract:** With the increase in the urban population witnessed in all major cities worldwide, efficient urban traffic planning, modeling and control are indispensable. In this paper we propose a novel macro-model for urban networks based on queuing theory and cast in a state-space representation. The model for each junction incorporates a switching mechanism capturing both the nonlinear dynamics of normal traffic conditions and the linear evolution of vehicle queues during junction block-back scenarios. Moreover, while competing models suffer a quadratic increase in computational cost with every added junction, our model exhibits only a linear increase in dimensionality and thus significantly lowers its computational demands. The simulation results from the proposed model are validated against a standard micro-modelling simulation package. These results will demonstrate through a Monte Carlo simulation that given correct model parameters, the proposed macro-model can well capture the complex vehicle dynamics of an urban region, thus paving the way for an accurate analysis of the network and for the development of efficient urban traffic control strategies.

**Keywords**: Junction Modelling; Urban Network; Macro-modelling; Junction Block-Back

## 1. Introduction

Continuous migration towards major cities around the world has brought an escalation in the number of inhabitants of urban areas. In some urban environments, mobility demands are exceeding the infrastructure capacity leading to increased traffic congestion. Expansion of the current road infrastructure is a possible solution, yet not always viable due to land-use restrictions and financial limitations. An alternative solution to urban congestion is a more efficient use of the current infrastructure through the adoption of dynamic and intelligent control strategies adaptable to prevailing traffic conditions. Such control strategies frequently require a dynamic and computation efficient representation of the salient vehicle dynamics at troublesome urban junctions and their neighbourhoods.

Most model-based control strategies are developed on macroscopic models due to their advantageous mathematical representation which lends itself well to classical control strategies, coupled with their low computational demand. Over the years, various macroscopic models for urban regions have been introduced. In 1963, Gazis and Potts (1963) introduced the store and forward model on which the TUC system is based. This model represents the urban network as a graph of links and junctions, where each link is modelled using conservation theory. A model proposed by Kashani and Saridis (1983) and later extended by Van den Berg *et al.* (2003), models different traffic scenarios by updating a discrete-time model in small simulation time steps. Daganzo (1994) proposed the cell transmission model. This model is based on the kinematic wave equation, with average traffic flows and densities modelled on separate sections of each link. Although robust, this model is hampered by higher computational requirements with negative repercussions on real-time implementations. Recently, Pecherkova *et al.* (2008), proposed a novel computationally efficient model of an urban micro-region, which describes each junction through the dynamic evolution of the queue lengths, vehicle intensities and road occupancy for the links leading to the junction.

The work in this paper expands on this later model of urban traffic dynamics. The computational efficiency of the proposed model is improved to obtain a linear increase in the model dimension with every added junction, as opposed to the exponential increase of the original model. To further capture the true dynamics of traffic at urban intersections, a novel macroscopic representation describing block-back at up-stream junctions is developed. The system switches between these two models based on the queue lengths of the interlinks. To validate this model its accuracy when compared to a commercially available micro-simulator is studied and documented in this work.

This paper is divided into 4 sections. Section 1 provided a brief motivation, an overview of competing models and the main aims of this work. Section 2 presents the computationally efficient model being proposed, together with the novel mathematical representation that captures junction block-back. In Section 3, an example based on two adjacent intersections is given, together with the results from a commercially available micro-simulator describing the same urban region. Finally, Section 4 identifies the main results and draws some concluding remarks.

## 2. Model Development

The state space model of an urban region proposed in this work is based on the model introduced by Pecherkova et al. (2008). Traffic flow is represented using non-linear dynamics and each link of the controlled region is represented by the following quantities:

- Queue length ($\zeta_k$) – the number of cars waiting on each link to pass through the intersection at the start of each cycle in unit vehicles [uv]

[1] Corresponding author: kenneth.scerri@um.edu.mt

- Input intensity ($\gamma_{I_k}$) – the rate of incoming unit vehicles on each link per cycle [uv/c]
- Occupancy ($\phi_k$) – the portion of time the detector is occupied by a vehicle [%]

with $k$ being the index of each cycle. The duration of the full traffic light cycle can be set to any reasonable value for each junction in the network.

### 2.1. Modelling Traffic at a Single Junction

The dynamics of the queue length on each link can be described by the principle of conservation (Homolova and Nagy, 2005), where the queue in the next cycle ($\zeta_{k+1}$) depends on the previous queue ($\zeta_k$), the input intensity ($\gamma_{I_k}$) and the outgoing intensity ($\gamma_{S_k}$) from the link through the relationship:

$$\zeta_{k+1} = \zeta_k + \gamma_{I_k} - \gamma_{S_k}(\zeta_k, \gamma_{I_k}, \mu_k)\mu_k + w_{1_k} \qquad (1)$$

where the outgoing intensity is given by the non-linear relationship:

$$\gamma_{S_k}(\zeta_k, \gamma_{I_k}, \mu_k) = S - Se^{-\frac{\zeta_k + \gamma_{I_k}}{S\mu_k}} \qquad (2)$$

with S being the saturation flow determined by the physical properties of the intersection, $\mu_k$ being the ratio of green light at each link per cycle and $w_{1_k}$ being a white zero-mean Gaussian noise process describing the random variations from the mean behaviour.

The input intensity to the junction is modelled as a Markovian process with known mean and standard deviation which is given by:

$$\gamma_{I_{k+1}} = \gamma_{I_k} + w_{2_k} \qquad (3)$$

with $w_{2_k}$ again being a white-noise, zero-mean Gaussian process.

The occupancy at each cycle period is defined as a random process with the mean given through a linear relationship between the previous occupancy and the previous queue length with known standard deviation given by:

$$\phi_{k+1} = \kappa\zeta_k + \beta\phi_k + w_{3_k} \qquad (4)$$

with $w_{3_k}$ also a white-noise, zero-mean Gaussian process.

Based on the junction dynamics given by equations (1) to (4), the dynamic equation of a state space model representing each link at a junction can be given by:

$$\begin{bmatrix} \zeta_{k+1} \\ \gamma_{I_{k+1}} \\ \phi_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ \kappa & 0 & \beta \end{bmatrix} \begin{bmatrix} \zeta_k \\ \gamma_{I_k} \\ \phi_k \end{bmatrix} + \begin{bmatrix} -\gamma_{S_k} \\ 0 \\ 0 \end{bmatrix} \mu_k + \begin{bmatrix} w_{1_k} \\ w_{2_k} \\ w_{3_k} \end{bmatrix} \qquad (5)$$

Such dynamics may be expanded based on the number of arms of an intersection as shown in (Pecherkova *et al.*, 2008). Also, the measurement equation represents the type and number of detectors implemented on the links of the controlled intersection. In this work, only loop detectors will be used and thus each detector is represented by the measurement equation:

$$\begin{bmatrix} \gamma_{I_k} \\ \phi_k \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \zeta_k \\ \gamma_{I_k} \\ \phi_k \end{bmatrix} + \boldsymbol{v}_k \qquad (6)$$

where $\boldsymbol{v}_k$ represents zero-mean, Gaussian measurement noise capturing sensor errors. Note that such loop detectors may be placed at strategic locations on both the roads approaching and leaving the junction.

### 2.2. Expansion to Multiple Junctions

With every added junction, the original work proposed to append extra states to the states space dynamics of equation (5). Such a methodology results in a quadratic increase in the computational time with every added junction. To overcome this problem, a novel modification is proposed in this work to reduce the computational demand while fully preserving the model accuracy. Thus, each junction in the urban region is modelled by a separate state space model with the input intensities of the interlinks connecting adjacent junctions modelled as inputs to the state space model. To feed these inputs, the measurement equation is modified to include the output intensities of the roads connecting the junctions. This reduces the dimensionality of the model with a linear increase in the computational demand with every added junction. This methodology is illustrated based on an urban region with two signal-controlled T-junctions as depicted in Figure 1.

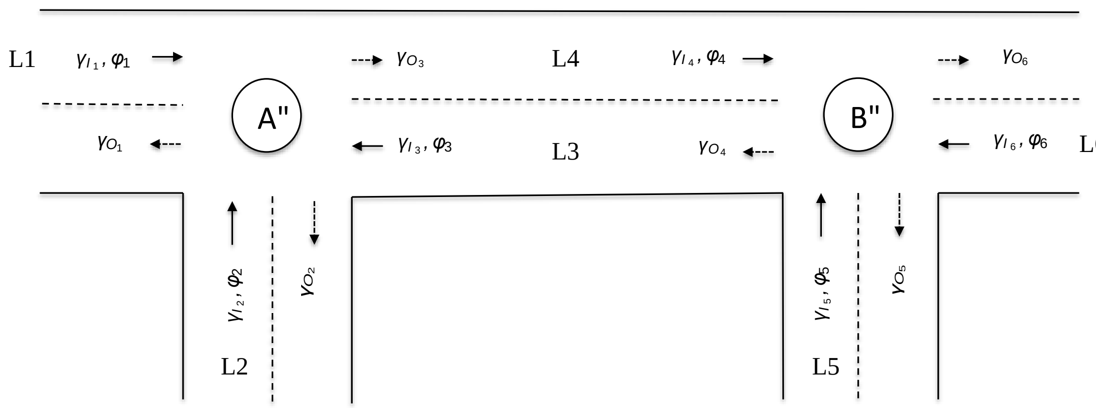



**Fig. 1.**
*Schematic Representation of an Urban Region*
*Source: Produced by Authors*

The region illustrated in **Fig. 1** is assumed to be equipped with a strategic detector on the input links placed a some meters away from the stop line to provide measurements of the input intensity and the occupancy and an output detector on the output links to give the output intensity. The interlinks (L3 and L4) connecting the neighbouring junctions are equipped with both detectors. Note that, such a sensor arrangement is not a requirement for this implementation. Either the input or the output sensors can be removed if the junction turning ratios are assumed to be known. Nevertheless, the cost saving obtained from such an implementation is partially off-set by the added computational effort required to estimate the intensities which are not measured and a possible reduction in accuracy due to inaccurate or time-varying turning ratios.

Based on the proposed implementation, the state dynamics of junction A are given by:

$$\boldsymbol{x}_{\boldsymbol{A}k+1} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \kappa_1 & 0 & 0 & 0 & 0 & \beta_1 & 0 & 0 \\ 0 & \kappa_2 & 0 & 0 & 0 & 0 & \beta_2 & 0 \\ 0 & 0 & \kappa_3 & 0 & 0 & 0 & 0 & \beta_3 \end{bmatrix} \boldsymbol{x}_{\boldsymbol{A}k} + \begin{bmatrix} -\gamma_{S_1} & 0 & 0 & 0 \\ 0 & -\gamma_{S_2} & 0 & 0 \\ 0 & 0 & -\gamma_{S_3} & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \boldsymbol{u}_{\boldsymbol{A}k} + \boldsymbol{w}_{\boldsymbol{A}k} \tag{7}$$

where $\boldsymbol{x}_{\boldsymbol{A}_k} = \left[\zeta_{k_1}\ \zeta_{k_2}\ \zeta_{k_3}\ \gamma_{I_{k_1}}\ \gamma_{I_{k_2}}\ \phi_{k_1}\ \phi_{k_2}\ \phi_{k_3}\ \right]^T$, $\boldsymbol{u}_{\boldsymbol{A}_k} = \left[\mu_{k_1}\ \mu_{k_2}\ \mu_{k_3}\ \gamma_{I_3}\right]^T$ and $\boldsymbol{w}_{\boldsymbol{A}_k} = \left[w_{k_1} \ldots w_{k_8}\right]^T$.

The measurement equation now also includes the output intensity at each link of the junction, which depends on the turning ratio $\alpha_{ij}$ of the outgoing intensities going from the *i*th arm to the *j*th arm. Therefore the measurement equation for junction A is given by:

$$\boldsymbol{y}_{\boldsymbol{A}k} = \begin{bmatrix} 0&0&0&1&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0 \end{bmatrix} \boldsymbol{x}_{\boldsymbol{A}k} + \begin{bmatrix} 0&0&0&0\\ 0&0&0&0\\ 0&0&0&1\\ 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0\\ 0&\alpha_{21}\gamma_{S_2}&\alpha_{31}\gamma_{S_3}&0\\ \alpha_{12}\gamma_{S_1}&0&\alpha_{32}\gamma_{S_3}&0\\ \alpha_{13}\gamma_{S_1}&\alpha_{23}\gamma_{S_2}&0&0 \end{bmatrix} \boldsymbol{u}_{\boldsymbol{A}k} + \boldsymbol{v}_{\boldsymbol{A}k} \qquad (8)$$

where $\boldsymbol{y}_{\boldsymbol{A}_k} = \left[\gamma_{I_{k_1}}\, \gamma_{I_{k_2}}\, \gamma_{I_{k_3}}\, \phi_{k_1}\, \phi_{k_2}\, \phi_{k_3}\, \gamma_{O_{k_1}}\, \gamma_{O_{k_2}}\, \gamma_{O_{k_3}}\right]^{\mathrm{T}}$ and $\boldsymbol{v}_{\boldsymbol{A}_k} = \left[v_{k_1} \dots v_{k_8}\right]^{\mathrm{T}}$.

The same modeling procedure is used for junction B where the state equation is given by:

$$\boldsymbol{x}_{\boldsymbol{B}k+1} = \begin{bmatrix} 1&0&0&0&0&0&0&0\\ 0&1&0&1&0&0&0&0\\ 0&0&1&0&1&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ \kappa_4&0&0&0&0&\beta_4&0&0\\ 0&\kappa_5&0&0&0&0&\beta_5&0\\ 0&0&\kappa_6&0&0&0&0&\beta_6 \end{bmatrix} \boldsymbol{x}_{\boldsymbol{B}k} + \begin{bmatrix} -\gamma_{S_4}&0&0&1\\ 0&-\gamma_{S_5}&0&0\\ 0&0&-\gamma_{S_6}&0\\ 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0 \end{bmatrix} \boldsymbol{u}_{\boldsymbol{B}k} + \boldsymbol{v}_{\boldsymbol{B}k} \qquad (9)$$

where $\boldsymbol{x}_{\boldsymbol{B}_k} = \left[\zeta_{k_4}\, \zeta_{k_5}\, \zeta_{k_6}\, \gamma_{I_5}\, \gamma_{I_6}\, \phi_{k_4}\, \phi_{k_5}\, \phi_{k_6}\right]^T$, $\boldsymbol{u}_{\boldsymbol{B}_k} = \left[\mu_{k_4}\, \mu_{k_5}\, \mu_{k_6}\, \gamma_{I_4}\right]^T$ and $\boldsymbol{v}_{\boldsymbol{B}_k} = \left[v_{k_1} \dots v_{k_8}\right]^T$.
The measurement equation is given by:

$$\boldsymbol{y}_{\boldsymbol{B}k} = \begin{bmatrix} 0&0&0&0&0&0&0&0\\ 0&0&0&1&0&0&0&0\\ 0&0&0&0&1&0&0&0\\ 0&0&0&0&0&1&0&0\\ 0&0&0&0&0&0&1&0\\ 0&0&0&0&0&0&0&1\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0 \end{bmatrix} \boldsymbol{x}_{\boldsymbol{B}k} + \begin{bmatrix} 0&0&0&1\\ 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0\\ 0&0&0&0\\ 0&\alpha_{54}\gamma_{S_5}&\alpha_{64}\gamma_{S_6}&0\\ \alpha_{45}\gamma_{S_4}&0&\alpha_{65}\gamma_{S_6}&0\\ \alpha_{46}\gamma_{S_4}&\alpha_{56}\gamma_{S_5}&0&0 \end{bmatrix} \boldsymbol{u}_{\boldsymbol{B}k} + \boldsymbol{v}_{\boldsymbol{B}k} \qquad (10)$$

where $\boldsymbol{y}_{\boldsymbol{B}_k} = \left[\gamma_{I_{k_4}}\, \gamma_{I_{k_5}}\, \gamma_{I_{k_6}}\, \phi_{k_4}\, \phi_{k_5}\, \phi_{k_6}\, \gamma_{O_{k_4}}\, \gamma_{O_{k_5}}\, \gamma_{O_{k_6}}\right]^T$ and $\boldsymbol{v}_{\boldsymbol{B}_k} = \left[v_{k_1} \dots v_{k_8}\right]^T$.

### 2.3. Capturing Junction Block-Back

To obtain an accurate model of the traffic flow through an urban region, the inclusion of junction block-back dynamics is crucial. Junction block-back occurs when the vehicle queue on a link connecting neighbouring junctions reaches its maximum capacity. Thus, an otherwise unsaturated junction starts to exhibit reduced traffic flows due to junction blocking as a result of the queue build-up from its neighbour.

In this work we propose to model this behaviour based on the observation that when a junction is approaching block-back, the outgoing intensity from the input links of the junction are no longer dependent on the saturation value of the links but are limited by the buffer space available in the interlink. Thus, considering the urban region depicted in **Fig. 1**, the condition for block-back of junction A is given by:

$$\zeta_{k_4} + \left(\alpha_{13}\gamma_{s_1}\mu_{k_1}\right) + \left(\alpha_{23}\gamma_{s_2}\mu_{k_2}\right) - \left(\gamma_{S_4}\mu_{k_4}\right) > \zeta_{k_{4max}} \qquad (11)$$

where $\zeta_{4_{\max}}$ is the maximum queue possible in link L4. If this condition is satisfied, the outgoing intensities of L1 and L2 are therefore given by:

$$\gamma_{S_1} = a_{14}(\zeta_{4_{max}} - \zeta_4) + \alpha_{13}\gamma_{S_1}\mu_1 \quad (12)$$
$$\gamma_{S_2} = a_{24}(\zeta_{4_{max}} - \zeta_4) + \alpha_{23}\gamma_{S_2}\mu_2 \quad (13)$$

where $a_{14}$ and $a_{24}$ are the ratios of cars from L1 and L2 that travel to L4 respectively.

### 3. Comparative Example

To analyse the accuracy of the modelling strategy being proposed, it was tested on the urban region schematically depicted in **Fig. 1** and compared to the results obtained from the micro-simulation package, *Aimsun*, commercially available from *Transport Simulation Systems (TSS) Ltd.*. The macro-model presented in Section 2, was implemented on *Mathwork*'s *Matlab*. The simulation step-time for the macro-simulation was taken as the traffic light cycle time of 90s. Input intensities, fixed traffic light timings and model parameters are as given in **Table 1** and **Table 2** for junctions A and B, respectively. Note that, the saturation values, $S_i$ and turning ratio, $\alpha_{ij}$ for any given junction can be obtained directly from measured data, while the model occupancy parameters κ and β may be obtained through linear regression based on the same data (Pecherkova *et al.*, 2008).

**Table 1** **Table 2**
*Macro-Model Parameters for Junction A* *Macro-Model Parameters for Junction B*

| Parameter | Value |
|---|---|
| Mean intensity on L1 | 15veh/c |
| Mean intensity on L2 | 10veh/c |
| Saturation on L1 ($S_1$) | 45veh/c |
| Saturation on L2 ($S_2$) | 38veh/c |
| Saturation on L3 ($S_3$) | 30veh/c |
| $\mu_1, \mu_2, \mu_3$ | 0.33, 0.22, 0.45 |
| $\kappa_1, \kappa_2, \kappa_3$ | 3,2.1,4.6 |
| $\beta_1, \beta_2, \beta_3$ | 0 |
| $\alpha_{12}, \alpha_{32}$ | 0.3 |
| $\alpha_{21}, \alpha_{23}$ | 0.5 |
| $\alpha_{31}, \alpha_{13}$ | 0.7 |
| $w_{A1}, w_{A2}, w_{A3}, w_{A6}, w_{A7}, w_{A8}$ | $N(0,0.1)$ |
| $w_{A4}, w_{A5}$ | $N(0,1)$ |
| $\boldsymbol{v}_A$ | $N(0,0.1)$ |

| Parameter | Value |
|---|---|
| Mean intensity on L5 | 30veh/c |
| Mean intensity on L6 | 40veh/c |
| Saturation on L4 ($S_4$) | 36veh/c |
| Saturation on L5 ($S_5$) | 38veh/c |
| Saturation on L6 ($S_6$) | 42veh/c |
| $\mu_1, \mu_2, \mu_3$ | 0.33, 0.22, 0.45 |
| $\kappa_4, \kappa_5, \kappa_6$ | 3,2.1,4.6 |
| $\beta_4, \beta_5, \beta_6$ | 0 |
| $\alpha_{45}, \alpha_{65}$ | 0.3 |
| $\alpha_{54}, \alpha_{56}$ | 0.5 |
| $\alpha_{46}, \alpha_{64}$ | 0.7 |
| $w_{B1}, w_{B2}, w_{B3}, w_{B6}, w_{B7}, w_{B8}$ | $N(0,0.1)$ |
| $w_{A4}, w_{A5}$ | $N(0,1)$ |
| $\boldsymbol{v}_{\mathbf{B}}$ | $N(0,0.1)$ |

**Fig. 2** shows the simulated input intensities on all roads leading to the two junctions as inputted to both simulators. Note that, the intensities on links 1 and 2 (input to junction A) as significantly lower then the intensities on links 5 and 6 (input to junction B). It is thus envisaged that while junction B will probably experience high queue lengths, vehicles arriving at junction A should travel through the intersection with lower delays unless block-back conditions arise.

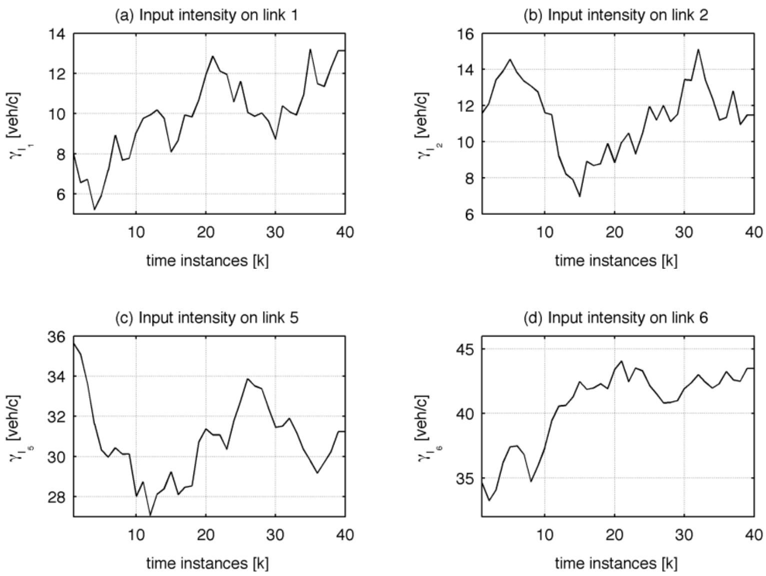



**Fig. 2.**
*Input Intensities on Links 1, 2, 5 and 6.*
*Source: Produced by Authors*

A single run of the simulation obtained the queues shown in **Fig. 3** for the proposed macro-model and in **Fig. 4** for the micro-model. Note that, while the micro-model required 2.03 seconds to obtained these results; the corresponding macro-model results were obtained in 0.82 seconds; 59% reduction in computational time for the same machine. On comparing the queue length accuracy, similar queue dynamics were witnessed from both; with queues on links 1, 3 and 4 exhibiting low stable values while queues on links 2, 5 and 6 show queues of linearly increasing length. This was expected due to the high input intensities on the links of junction B and the low traffic flow saturation value of link 2.

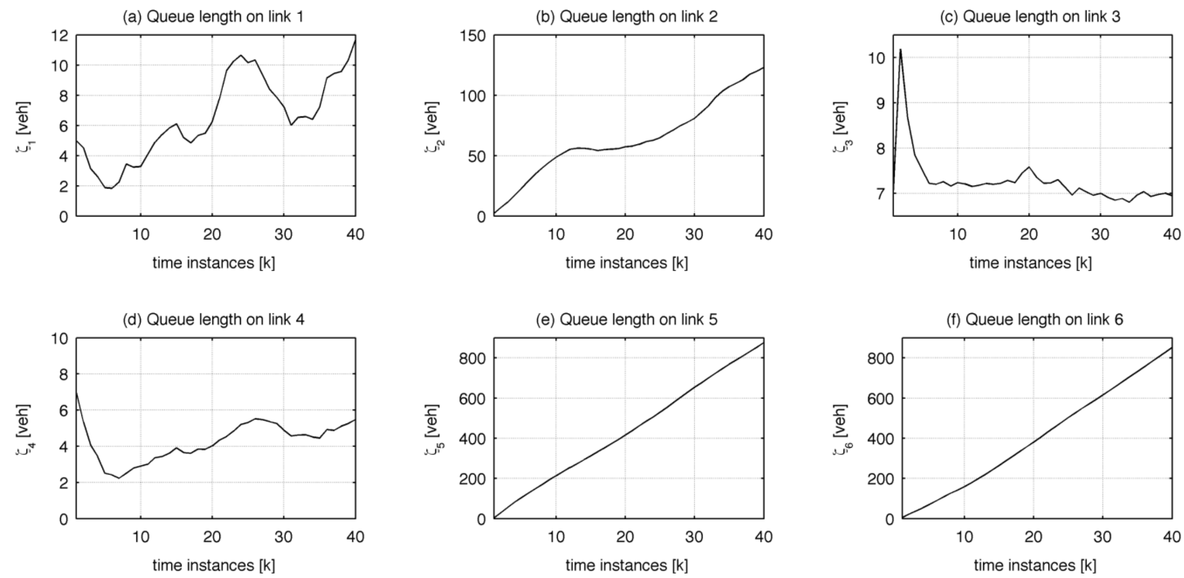



**Fig. 3.**
*Queue Lengths on All Links as Given by the Proposed Macro-Model*
*Source: Produced by Authors*

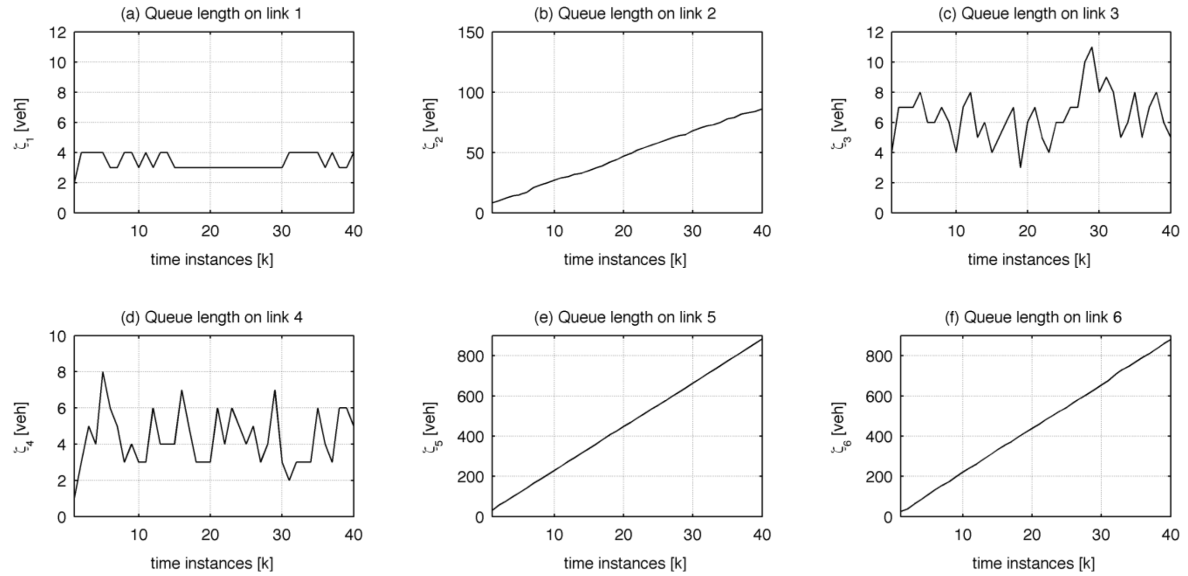



**Fig. 4.**
*Queue Lengths on All Links as Given by the Micro-Model*
*Source: Produced by Authors*

Although the behaviour shown in **Fig. 2** and **Fig. 3** is generally similar, some minor differences can be noted. These discrepancies between the proposed macro-model and the micro-model, are attributable to the different outgoing intensities obtained by each simulator. Note that, for the macro-model the outgoing intensities are given by the non-linear dynamics of *equation 2*, which is shown graphically in **Fig. 5**. **Fig. 5** also shows the outgoing intensity as measured on AIMSUN for multiple runs at the same junctions. Clearly, for large input intensities, the saturated output flow is similar, thus leading to the similar queue length on the heavily loaded roads of links 4 and 6. Nevertheless, for low and medium traffic flows, the output intensity for each simulator can vary significantly thus accounting for the differences in queue lengths on the first four links. Nevertheless, irrespective of these differences, both models capture the salient stable or increasing queues at any link on the network.

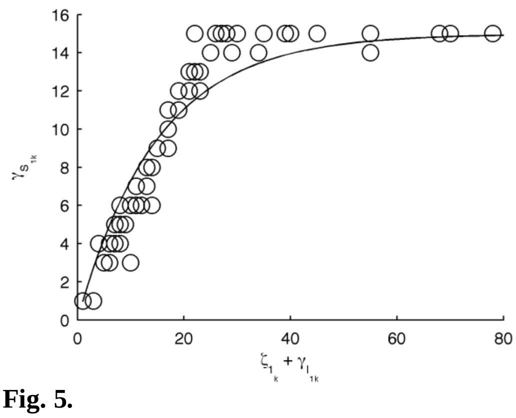



**Fig. 5.**
*Output Intensity for Different Queue Lengths as Used by Both Models (Circles represent the micro-model output intensities for different queue lengths and input intensities. The full line represents the macro-model output intensity relation as given by equation 2.) Source: Produced by Authors*

To further validate the accuracy and study the differences between the two simulators, Monte Carlo runs of 50 realizations were performed. **Fig. 6** and **Fig. 7** give box plots for the queue lengths obtained from the macro- and micro-model, respectively based on 50 realizations of the input intensities with statistics as given in **Tables 1** and **2**. Clearly, the traffic in the interlinks is predicted by both simulators to remain within low values, while a significant queue build-up is obtained on the outer links with link 5 and 6 being more congested.

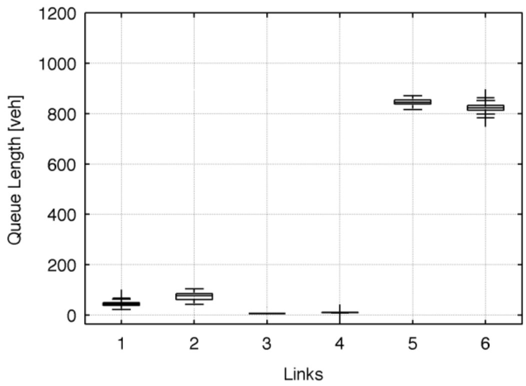



**Fig. 6.**
*Queue Lengths from the Macro-Model*
*Source: Produced by Authors*

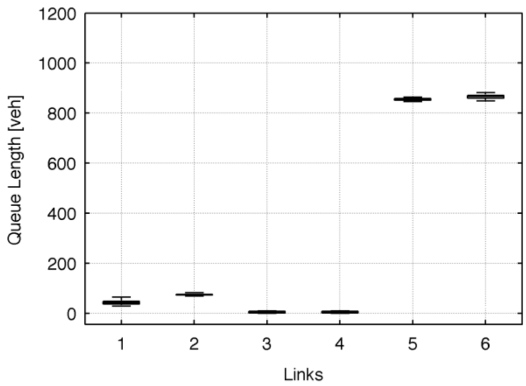



**Fig. 7.**
*Queue Lengths from the Micro-Model*
*Source: Produced by Authors*

## 4. Conclusions

Heavy traffic in our cities has a severe detrimental effect on our health, finances and standard of living. Transport engineers and policy makers are thus continuously seeking novel solutions to alleviate these problems though the use of simulation studies. Such studies require accurate and computational efficient traffic models to capture the dynamics of the full urban network. Towards such an aim, the macro-model proposed in this work is shown to capture accurately complex traffic dynamics with significant computational improvements over the industrial standard micro-models. The state space formulation of this model also lends itself well to the development of a wide range of control and model estimation techniques, thus paving the way for the development of novel and intelligent traffic lights systems in urban environments.

**Acknowledgements**

The research work disclosed in this publication is partially funded by the *Master it!* Scholarship Scheme. The Scholarship is part-financed by the European Union (European Social Funds).

**References**

Gazis, D.C.; Potts, R.B. 1963. The oversaturated intersection. In *Proceedings of the Second International Symposium on Traffic Theory*, London, UK, 221–237.

Kashani, H.; Saridis, G. 1983. Intelligent control for urban traffic systems. *Automatica*, 19(2), 191-197.

Van der Berg, M.; Hegyi, A.; De Schutter, B.; Hellendoorn, J. 2003. A macroscopic traffic flow model for integrated control of freeway and urban traffic networks. In *Proceedings of the 42nd IEEE Conference on Decision and Control.* Maui, Hawaii, 2774-2779.

Daganzo, C. F. 1994. The cell transmission model: a simple dynamic representation of highway traffic consistent with hydrodynamic theory. *Transportation Research*, 28B(4), 269-287.

Pecherkova ́, P.; Dunık, J.; Flıdr, M. 2008. Modelling and simultaneous estimation of state and parameters of traffic system. *Robotics, Automation and Control*, 1(1), 319–336.

Homolova ́, J.; Nagy, I. 2005. Traffic model of a microregion. In *Preprints of the 16th World Congress of the International Federation of Automatic Control. IFAC*, Prague, 1–6.

# THE EFFECT OF UPGRADING PUBLIC TRANSPORT QUALITY ON MOBILITY CULTURE IN TOWN – BASED ON THE KRAKOW CASE STUDY

**Wieslaw Starowicz[1], Lidia Zakowska [2]**
[1] *Cracow University of Technology, Faculty of Civil Engineering, Transport Chair (L-32)*
*24 Warszawska St., 31-155 Cracow, Poland*

**Abstract:** This work presents the current approach to the sustainable mobility development in cities. The aim of this study is to show how the process of upgrading the quality of public transport influence the culture of mobility in urban districts. Based on the Krakow city case study, the quality of public transport service is characterized and evaluated. The mobility in Krakow town and metropolitan area is described with a special emphasis on its degree of sustainability. This analysis leads to evaluation of the effect which the assessed quality of public transport services may bring on the road density in the city district.
The aim of this work is to open the scientific discussion on how to create the sustainable public transport system that supports the idea of new transport culture, involving proper instruments.
As a result of this study, several conclusions were drawn and some directions for the future successive actions were indicated. Among selected actions, an important issue to solve is to find out which preconditions are required in order to upgrade, to increase the quality of all services which contribute to the public transport quality.
Final recommendations stress that the process of development of the urban public transport offer has to be studied and evaluated based on those requirements, which are in line with the most important expectations of current and future users.

**Keywords**: Public transport, mobility, transport quality, indicators, sustainability.

## 1. Introduction

All metropolitan areas suffer from similar problems resulting from congestion, safety risks for travelers, time losses and degradation of the environment (air and water pollution etc.). The costs of social and environmental losses can be significantly reduced by using specific management methods and techniques based on principles of sustainable transport.

The European Commission emphasizes the importance of ensuring sustainable urban mobility for all citizens, including the disadvantaged. This includes innovative organization schemes such as clean and safe vehicles and means of transport with lower levels of pollution, new high quality public transportation modes and rationalization of private transport, communication infrastructure, and integrated town planning and transport taking into account their relationship with growth and employment (Decision no 1982/2006/EC). There is a clear need to develop integrated, safer, greener and smarter pan-European transport systems for the benefit of all citizens, society and climate policy including environmental aspects (Zakowska, 2013) as well as natural resources. This research is based on multi-criterion analysis and evaluation regarding transport economics, quality, safety, accessibility, and efficiency, environmental protection and equity.

Shaping of sustainable mobility in cities consists on conscious resignation of travelling by private transport in favor of ecological transport (including public transport). Extremely important factor of this formation is determination and acceptation of selected and defined measurement indicators of progress or regress on the way to achieve sustainable mobility. Its systematic defining allows to monitor if transport behavior of cities' inhabitants are consistent with the concept of sustainable mobility and what is the shaping rate and efficiency. In the literature one can find different research approaches to sustainable mobility in urban areas (Bryniarska & Starowicz, 2010). Frequently, among numerous proposals, there are sustainable transport indicators - but not sustainable mobility indicators. Those indicators are significant as far as mobility point of view is concerned since this feature (besides accessibility) plays an important role in the development of sustainable transport.

## 2. Sustainable mobility indicators

International team led by W. Lerner (Ciaston-Ciulkin, 2014) has worked out a set of sustainable mobility universal indicators. The Lerner's team created 11 indicators describing rate (level) of sustainable mobility in urban areas and they conducted research on 66 cities all over the world. Indicators proposed by authors may be assigned in two groups: first five indicators measure level of sustainable mobility and following six indicators measure effects of rules applied in the city to build new culture of mobility. Bipolar scale has been defined for each indicator and its maximum and minimum value represented relatively the highest and lowest values of indicators noted in the set of analyzed cities.

[2] Corresponding author: lzakowsk@pk.edu.pl

Depending on value of specific indicator city was given certain amount of points. While summing up those points value of sustainable mobility indicator was calculated. Total amount of points that potentially may be scored is 100. Table 1 presents full list of applied criteria and obtainable points.

**Table 1.**
*List of indicators measuring sustainable mobility after the Lerner's team*

| Number | Indicator name | Amount of points obtainable |
|---|---|---|
| **Sustainable mobility factors (32,5 points at maximum)** | | |
| **1.** | Ecological travels (pedestrian, bicycle, public transport) in modal split<br>**Maximum 7,5 points.** | Highest indicator (7,5 points)<br>Lowest indicator (0 points) |
| **2.** | Strategy of sustainable mobility<br>**Maximum 10 points.** | Alternative energy sources (2 points)<br>Sustainable development (2 points)<br>Intermodal solutions (2 points)<br>Transport infrastructure (2 points)<br>Traffic restrictions (2 points) |
| **3.** | Car sharing system<br>**Maximum 5 points.** | Lack of system (0 points)<br>Planning to introduce system after 2011(1 point)<br><50 cars/ 1 million inhabitants (2 points)<br>51–100 cars/ 1 million inhabitants (3 points)<br>101–200 cars/ 1 million inhabitants (4 points)<br>>201 cars/ 1 million inhabitants (5 points) |
| **4.** | Urban bike rental system<br>**Maximum 5 points.** | Lac of system (0 points)<br><100 bicycles/ 1 million inhabitants (1 point)<br>101–500 bicycles / 1 million inhabitants (2 points)<br>501–1000 bicycles / 1 million inhabitants (3 points)<br>1001–5000 bicycles / 1 million inhabitants (4 points)<br>>5001 bicycles / 1 million inhabitants (5 points) |
| **5.** | Smart cards application<br>**Maximum 5 points.** | Lack of system (0 points)<br><0,1 card/ inhabitant (1 point)<br>0,1–0,25 card/ inhabitant (2 points)<br>0,25–0,5 card/ inhabitant (3 points)<br>0,5–1 card/ inhabitant (4 points)<br>>1 card/ inhabitant (5 points) |
| **Sustainable mobility impact indicators (maximum 67,5 points)** | | |
| **6.** | Death toll in road accidents per 1 million of inhabitants<br>**Maximum 15 points.** | Lowest indicator (15 points)<br>Highest indicator (0 points) |
| **7.** | Air pollution - $C0_2$ emission from transport [kg/inhabitant]<br>**Maximum 15 points.** | Lowest indicator (15 points)<br>Highest indicator (0 points) |
| **8.** | Motorization rate [number of cars/1 million inhabitants]<br>**Maximum 7,5 points.** | Lowest indicator (7,5 points)<br>Highest indicator (0 points) |
| **9.** | Average travel speed [km/h]<br>**Maximum 7,5 points.** | Lowest indicator (7,5 points)<br>Highest indicator (0 points) |
| **10.** | Satisfaction on transport system [points]<br>**Maximum 15 points.** | Average of 3 independent experts' assessments<br>Additionally public ballot |
| **11.** | Average travel time to work [minutes]<br>**Maximum 7,5 points.** | Shortest indicator (7,5 points)<br>Longest indicator (0 points) |

*Source: own study based on The Future of Urban Mobility, Towards networked, multimodal cities of 2050, Arthur D Little future lab, No.1: Future of Urban Mobility.*

Cities' hierarchy, depending on location and indicator value has been presented in the figure 1. Interesting conclusions came out from the study. Average sustainable mobility indicator in analyzed cities was 64,4%. It says that analyzed cities on average used approximately 2/3 their possibilities when using best practices and achieving the best results of all aspects. Only 15 % cities has obtained 75 points in total, and the best results have been observed in Hong Kong and Amsterdam (above 80 points). It is worthy to note that there are mainly European cities in the group of 15% cities of the highest level of sustainable mobility (7 cities per 10), at the same time they are poorly represented in the group of the lowest indicator of sustainable mobility (1 per 13).

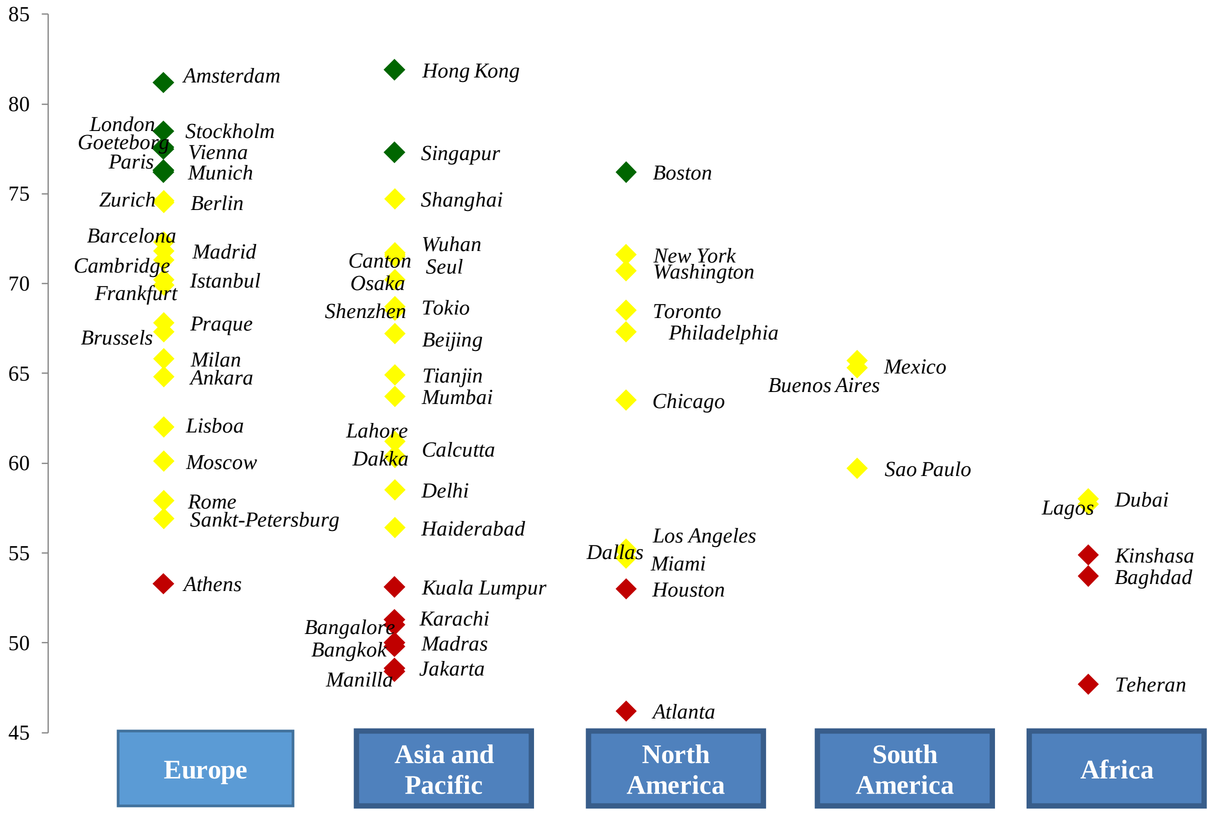



**Fig. 1.**

*Cities' hierarchy, depending on location and sustainable mobility indicator value*

*Source: The Future of Urban Mobility, Towards networked, multimodal cities of 2050, Arthur D Little future lab, No.1: Future of Urban Mobility.*

As results of above quoted studies prove, the level of sustainable mobility in cities is very diverse. Developing cities of Asia, Africa and South America generally struggle with its low level. Effects of activities in high developed American cities towards new culture of mobility may be considered as unremarkable. In those cities an intense development of individual transport is observed (80% private motor vehicle in urban model split). The level of sustainable mobility in high developed Asiatic cities does not differ meaningfully from the North American cities. The most advanced in developing of sustainable mobility are European cities. Only in this group average indicator of sustainability is above-average and is 70%.

Package of indicators measuring level of sustainable mobility in the urban area of Wroclaw (Municipal Holding of Krakow, 2013) Poland, may be considered as interesting. Full list of indicators consist of nine groups of factors. The key factors are seen mobility, economic and environmental factors, since they are of significance when shaping of sustainable mobility is concerned.

All indicators have been presented synthetically in the table 2.

**Table 2.**
*List of indicators measuring of the level of sustainable mobility in the urban area of Wroclaw, Poland*

| Number | Group of factors related to: | Proposed indicators: |
|---|---|---|
| **1.** | Sustainable mobility | • Modal split (positive estimation for growth of green travels and reduction of individual transport).<br>• Number of works (institutions, plants etc.) with mobility plans.<br>• Number of works (institutions, plants etc.) with mobility consultants.<br>• Indicators monitoring changes in public awareness about sustainable mobility fi. percentage of population supporting restrictions towards motor vehicles traffic; percentage of population supporting implementation of priorities for public transport at the cost individual vehicles; percentage of population supporting development of bicycle paths;<br>• Indicators monitoring changes of transport behavior of owners of private cars who resigned daily travels to work;<br>• Indicators monitoring effectiveness of city promotion of sustainable transport, fi. number of participants (receivers) of promotional events, including educational actions etc. |
| **2** | Public transport | • Level of inhabitants satisfaction on public transport system as far as all criteria's submitted to annual estimation.<br>• Technical condition of tram infrastructure (tramways) – reduction of tramway's of poor quality share.<br>• Average travel time in public transport.<br>• Length of roads with separated lanes for buses.<br>• Number of intersections with system for tramway priority at traffic lights.<br>• Standards of maximum load in the vehicles of public transport in the peak hours quantified by the size of filling up the surface of standing room. |
| **3** | Motor transport | • Technical condition of streets (reduction of streets length share of poor condition).<br>• Indicator of congestion / monitoring transport speed of private vehicles in the peak hours. |
| **4** | Non-motorized transport | • Level of inhabitants satisfaction on conditions for pedestrians.<br>• Level of cyclists satisfaction on conditions for riding a bicycle.<br>• Length of streets in restricted traffic zone/ vehicle-exclusion zones. |
| **5** | Freight transport | • Share of freight transport by logistic center. |
| **6** | Parking policy | • Number of eliminated parking places at roadways and pavements, substituted by places at parking plots, including paved and unpaved. |
| **7** | Safety of moving | • Absolute death toll different groups of users.<br>• Number of robberies and violence towards public transport passengers. |
| **8** | Economy | • Transport infrastructure expenditure in the city budget (share)<br>• Annual expenditure for investment in the public transport, bicycle infrastructure, road works and maintenance – from the city budget and other sources (fi. EU funds)<br>• Amount of subsidies for public urban transport from the city budget (difference between costs and income)<br>• Relation of expenses on public transport comparing to expenses on roads.<br>• Relation of expenses on transport infrastructure to expenses on its maintenance. |
| **9** | Environment | • Number of buses fulfilling at least EURO 5 standards, which regulates |

| | | |
|---|---|---|
| | | permissible level of pollutants emission in exhaust gases.<br>• Number of registered electric or hybrid vehicles.<br>• Level of $CO_2$ emission from transport in the city. |

*Source: own study on the base A. Rudnicki: Report to the project document: Wroclaw Mobility Policy. (in Polish: Koreferat do projektu dokumentu Wroclawska Polityka Mobilnosci)*

Admittedly, group of social factors has not been isolated but factors monitoring safety of moving (fi. number of car accidents or their victims). Remaining factors measuring level of sustainable mobility have been classified into groups of factors related to the functioning of public, individual, non-motorized, freight transport and parking policy.

Full list of indicators measuring progress of balancing mobility consist of few dozen of items. Five of them have been indicated as the key indicators:

- transport modal split in the city,
- average travel time in the public transport,
- death toll in car accidents (mainly pedestrians and cyclists),
- users' appraisal of urban transport system functioning,
- social awareness towards sustainable mobility.

## 3. HIERARCHY AND EVALUATION OF EUROPEAN CITIES ACCORDING TO THE LEVEL OF SUSTAINABLE MOBILITY

Researches on the level of sustainable mobility in the group of Polish and European cities have been undertaken by Ciaston-Ciulkin (Ciaston-Ciulkin, 2014). Proposed indicator of sustainable mobility is the sum of points obtained in the evaluation of seven partial indicators, similarly to the study of Lerner (Ciaston-Ciulkin, 2014), measuring the level of modal split sustainability among available transport modes:

- share of pedestrian and bicycles travel in the total number of travels [%] – stimulant,
- share of individual transport in the total number of travels [%] – destimulant,
- share of public transport in the non-pedestrian travels [%] – stimulant,

and results of activities devoted to sustainable mobility:

- travel time to work – defined as a percentage of working people travelling to work not longer than 30 minutes [%] – stimulant,
- motorization rate – defined as number of private vehicles per 1 000 inhabitants – destimulant,
- social factor – defined as death toll in car accidents per 1 000 inhabitants – destimulant,
- ecological factor – defined as air pollution from nitric oxide [number of days per year with exceeding permissible norm 120 $\mu g/m^3$] – destimulant.

For each of seven analyzed partial indicators, six ranges of equal scope were created, where extreme values presented the smallest and the biggest noted observations in the group of all cities. Depending on the range where the value of specific indicator was located, analyzed city received suitable value of points (from 0 to 5). This value was calculated into mobility indicator, assuming that 35 points equals 100%. Value of sustainable mobility indicator illustrates maximal amount of points scored by the city. The research was conducted in 29 European cities, 10% of the group consisted of Polish cities (Warszawa, Krakow, Gdansk). Access to the data, which was taken under consideration (in Municipal Holding of Krakow, 2013), was a definite criterion while selecting cities for the research.

In the synthetic meaning sustainable mobility indicator for all 29 analyzed cities is 55% per 35 points at maximum; cities average obtained 19 points. Definitely Scandinavian, German language and Spanish cities scored the highest number of points. Paris was also listed in this group. The lowest, in all analyzed aspects, Italian cities, Luxembourg and Brussels have been scored. British cities (except for London) and Polish cities – except for Kraków have been located below average.

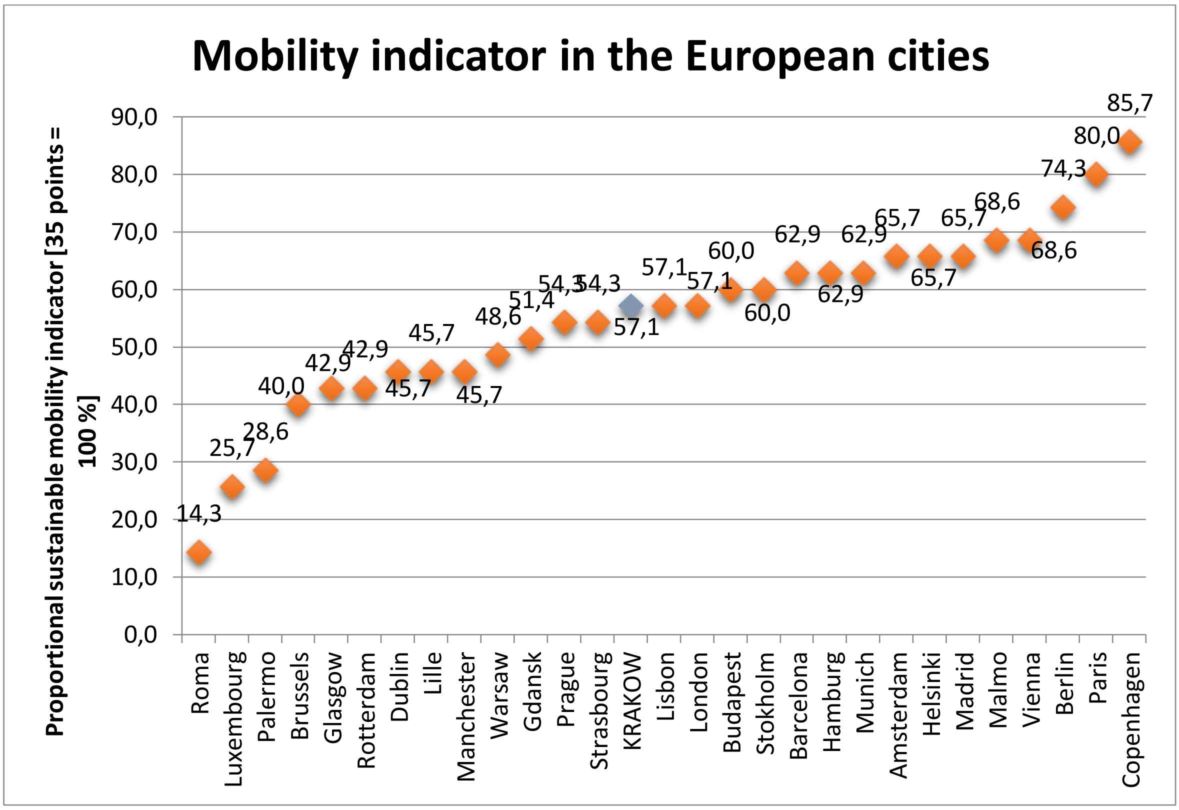



**Fig. 2.**
*Hierarchy of European cities according to the sustainable mobility indicator*
*Source: own study.*

## 4. RESEARCH ON RELATION BETWEEN SUSTAINABLE MOBILITY INDICATOR AND QUALITY OF TRANSPORT SERVICES

Activities to improve quality of transport services result in change of mobility culture in cities assuming that direction of those measures is consistent with expectations of city's inhabitants. Quality of services should be shaped in a manner satisfying its receivers.
This is why while studying impact of transport services quality on the process of shaping new culture of mobility in cities inhabitants' opinions on the functioning of public transport (gained in the quality of live researches conducted in 75 cities of the European Union, Croatia and Turkey in 2009, see EC Eurobarometer, March 2010) have been taken under consideration.
Participants of the research expressed their satisfaction on the quality of public transport services on the five-stages scale (very satisfied; satisfied; dissatisfied; very dissatisfied; no-opinion). Presented below is percentage share of satisfied and dissatisfied respondents.
It is easy to notice that quality of public transport services is appreciated mostly by inhabitants of European cities which are generally characterized by above average value of sustainable mobility indicator as well as its specific components. Quality of provided public transport services is highly appreciated by not less than 4/5 inhabitants of cities of the highest indicator of sustainable mobility (among others Copenhagen, Amsterdam, Paris, Malmö). Definitely inhabitants of Helsinki are of the highest appreciation for quality of public transport, only 7 % is not satisfied with its operation. Only few more dissatisfied inhabitants are in Vienna – 9%, Strasbourg – 10%, Hamburg – 13% and Stockholm – 13%. The lowest appreciation for transport services quality is among inhabitants of Italian cities, which were the weakest as far as sustainable mobility indicator is concerned. In Palermo, only 12% of inhabitants are satisfied with quality of public transport services, in Rome – 35 %. Also Budapest (48%) and Istanbul (60%) are among cities in which quality of public transport services is of very low appraisal.
In the group of cities in which quality of public transport services is appreciated by less than ¾ inhabitants are cities of moderated sustainable mobility indicator. In this group there are two Polish cities: Warszawa and Gdansk – 68%. The biggest group of inhabitants satisfied with quality of transport service in Polish cities are in Bialystok – 76% and in Krakow – 77%. It is worthy to underline that in both cities (EC Eurobarometer, March 2010) passengers perceived significant improvement of quality comparing to 2006 (improvement by 18,5% and 24,6%).

Results of the descriptive analysis, demonstrating reliance between sustainable mobility and quality of public transport services are confirmed also by statistical analysis. Conducted study on dependency indicates strong relation between both variables. Calculated positive correlation coefficient has the value 0,56 which indicates that usually high quality of public transport responds to high sustainable mobility indicator and inversely. This reliance is illustrated in the figure 3.

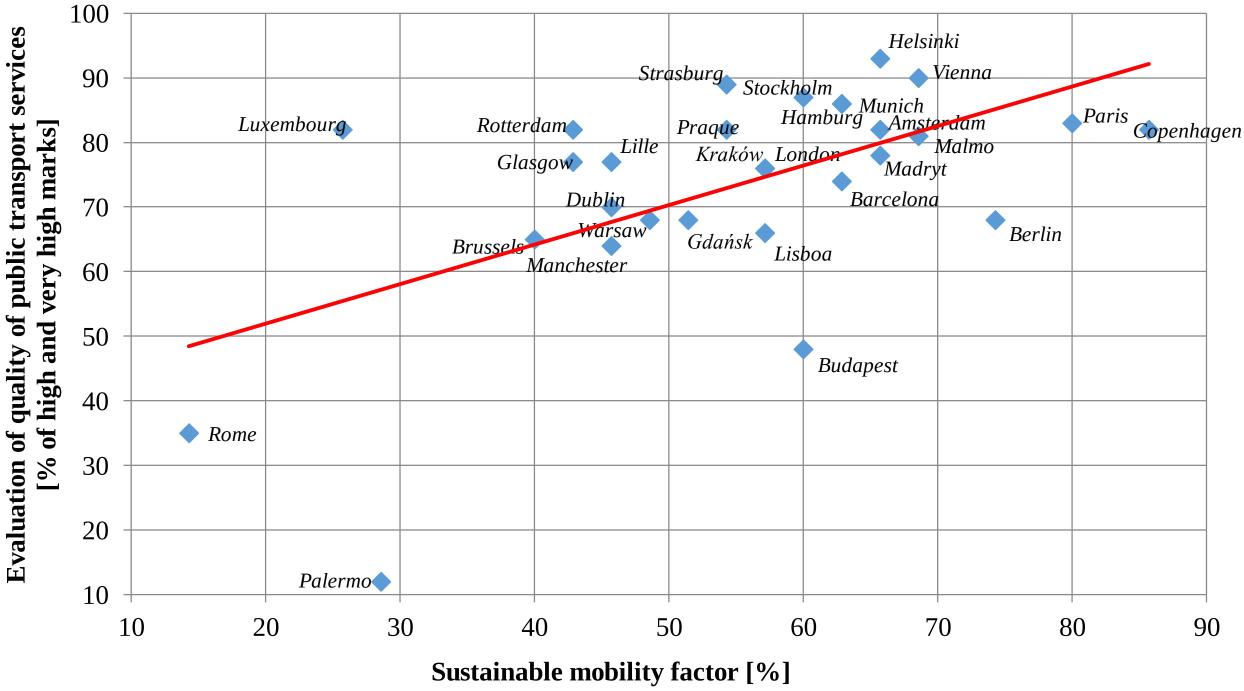



**Fig. 3.**
*Dependence of sustainable mobility factor size on the quality of public transport in cities*
*Source:* (Ciaston-Ciulkin, 2014): *Quality of public transport service as an instrument of change of mobility culture in cities. (in Polish: Jakosc usług transportu publicznego instrumentem zmian kultury mobilnosci w miastach), doctoral dissertation written under the supervision of W. Starowicz, Szczecin University*, Faculty *of* Management *and* Economics *of* Services, *2014*

## 5. LEVEL OF SUSTAINABLE MOBILITY IN CRACOW AND QUALITY OF PUBLIC TRANSPORT SERVICES

Analyzes of different researches (Rudnicki Edt., 2010) focused on quality of public transport services, provided for local, national or international needs concludes with the high level of transport services quality in Cracow (Starowicz, 2013).

Analyzing the trend line (fig. 4) of dynamics of quality change evaluation of transport services in the years of 2002 - 2012, on the base of results of annually conducted social researches concerning communal services, it is not difficult to notice that quality of public transport services is regarded as growing.

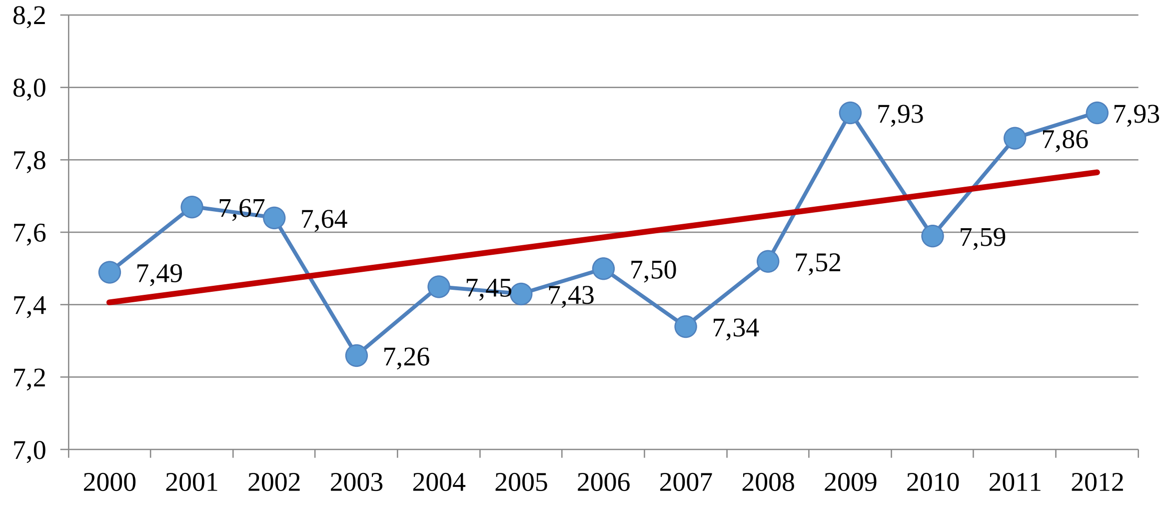



**Fig. 4.**
*Quality of public transport services perceived by city inhabitants in the years 2000-2012. Difference in the evaluation of public transport services in Krakow in relation to the average of seven biggest Polish cities [%]*
*Source: own study based on the social researches ordered by the Municipal Holding of Krakow (Krakowski Holding Komunalny S.A)*

One of the main expected effects of improving of public transport services quality is sustainable transport behavior meaning decrease of number of travels by individual transport.
Growth of sustainable transport behavior in Krakow is a result of intense involvement of city authorities. The city applies many tools to shape a new culture of mobility, including investment, financial as well as promotion and education instruments.

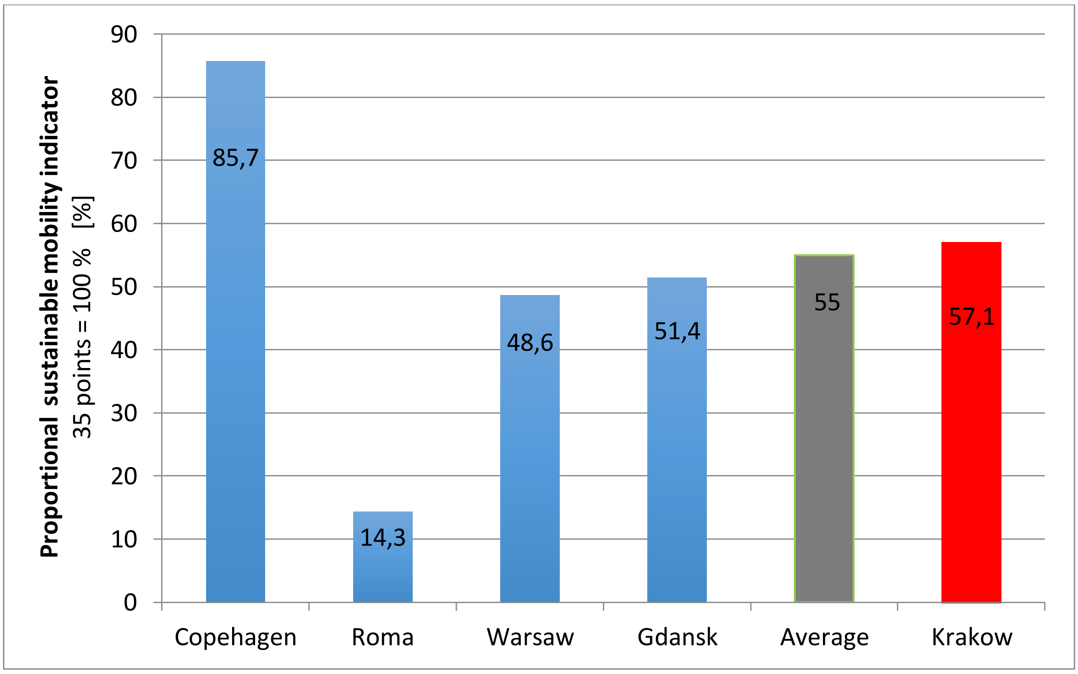



**Fig. 5.**
*Comparison of values of sustainable mobility indicators in selected Polish cities against extreme indicators in the European cities [%]*
*Source: own study*

Conducted research on the level of sustainable mobility in Krakow proves that it is above average in the group of analyzed European cities and the highest among Polish cities. The level of sustainable mobility is 57,14% and is 4% higher than average value, 75% higher than the lowest value of indicator describing sustainable mobility in Rome (figures 2 and 5).

In relation to the highest value of the Copenhagen indicator, Krakow indicator consists 2/3 of its value which leads to the conclusion concerning necessity to undertake strong involvement on the way towards new culture of mobility and to achieve level similar to the model city.

## 6. Conclusions

Krakow is an example of the city which activities in the scope of urban transport system development are intentionally focused on the shaping a new culture of mobility. Especially public transport is promoted, expanding offer of services and improving their elements. Resignation of travel by private vehicle in favor of public transport results from the fact that quality of services meets expectations of passengers. It is important that improvement of transport services quality in Krakow is meeting expectations of growing number of passengers every year.

Priorities for mass transport vehicles are being continually expanded through the following solutions:

- separate street lanes for buses, also permitting taxi and special vehicle traffic (approx. 20 km),
- rails integrated into the street surface with low separators from automobiles traffic (approx. 10 km),
- traffic management systems giving priority to streetcars at intersections with traffic signals,
- restriction in access to the downtown for personal vehicles (parking zones, parking fees, new traffic patterns),
- efficient traffic management system (priority for public transport, anti-congestion measures, maintaining or restoring punctuality and regularity in the functioning of mass transport),
- joint tram/bus stops enabling quick transfers.

The application of priorities in mass transport in Krakow area has resulted in:

- restoration of punctuality and regularity in transport circulation,
- increase in the attractiveness of mass transport,
- increase in passenger counts, but reduction in the volume of travel by personal vehicles,
- significant reductions in personal vehicle traffic in the downtown area.

The priority for city tram based transport, together with flexible bus system in metropolitan area is essential for sustainable transport development in European cities.
All those actions and strategies leads to more sustainable transport and as a consequence to better life quality in the city and in the whole metropolitan area.

**References**

Bryniarska Z., Starowicz W.. 2010. The results of public Transport system research in selected Polish cities (In Polish: *Wyniki badan systemów publicznego transportu zbiorowego w wybranych miastac*h), Wydawnictwo SITK RP w Krakowie, Krakow, 2010. (ISSN 1231-9155)

Ciaston-Ciulkin A. 2014. *Quality of public transport service as instrument of change of mobility culture in cities*. (in Polish: *Jakosc usług transportu publicznego instrumentem zmian kultury mobilności w miastach)*, doctoral dissertation under supervision of Starowicz, Uniwersytet Szczecinski, Wydzial Zarzadzania i Ekonomiki Usług, 2014.

CIVITAS CARAVEL Project. Available from internet: www.civitas-caravel.org

European Union, Regional Policy: *Survey on perception of quality of life in 75 European cities,* Flash Eurobarometer, March 2010.

Exerted statistic data of the Eurostat website, or Available from internet: www.EPOMM

Municipal Holding of Krakow (Krakowski Holding Komunalny S.A. *http://www.khk.krakow.pl/khk/badania-spoleczne/ ( 05.07.2013).*

Rudnicki. A. Edt. 2010. Innovations for sustainable Transport in Cities. Experiences from the EU Project Civitas-Caravel. *(in Polish: Innowacje na rzecz zrownowazonego transportu miejskiego. Doswiadczenia z realizacji projektu Unii Europejskiej Civitas-Caravel). SITK RP, Oddzial w Krakowie, 2010.* (ISBN: 978-83-86219-53-7).

Starowicz, W. 2013. Krakow as a model for sustainable transport development. *Transport Net.Sustainable Cities. Building cities for the future*. Ed. Climate Action, London, 2013 (ISBN: 978-0-9570432-8-2).

Statistics on Krakow city Transport. Available from internet: www.zikit.krakow.pl

*The Future of Urban Mobility, Towards networked, multimodal cities of 2050, Arthur D Little future lab, No.1: Future of Urban Mobility*. Available from internet: *http://www.adlittle.com/downloads/tx_adlreports/ADL_Future_of_urban_mobility.pdf* (access: 02.07.2014).

Zakowska, L. 2013. Sustainable safety of transport users in relation to alcohol abuse in Poland. *International Cooperation On Theory And Concepts in Traffic, ICTCT* Extra Workshop, Stellenbosch, South Africa. April 2013.

# URBAN MANAGEMENT & MOBILITY IN STRATEGIC PROJECTS IN EASTERN SERBIA

**Uroš Radosavljević[1], Danilo Furundžić[2]**
*[1,2] University of Belgrade Faculty of Architecture, Belgrade, Serbia*

**Abstract:** Climate change in general and recent floods in particular in Serbia urges for creative, efficient and effective adaptation and mitigation measures to be implemented in urban and rural areas, especially in the field of transport and urban and rural tourism development. The experience show that 'hard' measures in the field of transport are costly and take time to implement, and that they have to be followed with 'soft' measures which cover the issues of management and organization in transportation sector and raising travel awareness towards more sustainable modes of transport among citizens and decision makers. This shift calls for complex conception of new adaptive planning, governance and management approaches and instruments to sustainable urban and rural development and mobility in an integrated way that is context sensitive. This paper will present contemporary theoretical and practical European approaches in the field of sustainable urban and rural development and mobility management and their applicability in Serbia aiming to identify possible directions for improvement and use specific management instruments that might work in this context. Our methodology is based on the case study framed in the elaboration of strategic project *'Integral Plan for the Sustainable Development of Rural Tourism in Eastern Serbia'* in Golubac, Kladovo, Majdanpek and Negotin done within the umbrella of larger GIZ *project 'Municipal economic development in the Danube Region'*. We argue that it is possible to work within the community, with ordinary people, households, vine producers, and with local government representatives, incrementally, trough site visits talking with people, participative interactive workshops and formal presentations in order to build trust, raise awareness and recommend sustainable measures and projects in the field of mobility and travel awareness for urban and rural tourism development aiming at identifying possible improvement directions in Serbian context.

**Keywords**: urban management, mobility management, strategic projects, community.

## 1. INTRODUCTION

Climate change represents potentially catastrophic risks to the economy and society around the planet (IPCC, 2007a). Attention has recently started to turn towards adaptation measures that climate change impacts carry, since the changes that have happened are inevitable in the future, even if the measures for reducing the effects of climate change are realized (IPCC, 2007b). While the national governments will have to take a leading role in the creation of comprehensive policies for climate change mitigation, strategies and measures will have to be developed at the local and regional level, where the impact will be felt at most.

Lalović, Živković & Radosavljević (2011) argues that the paradigm of sustainability puts people and human actions more than ever in focus within the context of climate change. The question of urban governance, management and planning of sustainable development is aimed towards adaptation and mitigation measures of climate change. The subject of discussion is narrowing to the questions of effectiveness of the decision making process which will result in future in real human behavior change in accordance with desired global and local effects.

Policies to mitigate climate changes in the field of urban and rural tourism development will have to find a balance between potentially conflicting objectives, such as objectives related to local and national economic entrepreneurial development on one hand, and the protection of natural environment and tourist destinations on the other. Also, there are direct and indirect effects and impacts of certain economic and transport activities on tourist destinations vulnerable to climate change. On the other hand, it is important to emphasize that the decisions that are made about the impacts of tourism on climate change have important implications for local, national, global, and values for the generations that come after us. All these aspects and impacts must be taken into account to arrive at balanced and effective policies that do not violate any of the above-mentioned common sectors.

These challenges of climate change and impacts it has on human and natural settlements and life including economic and technological transformation calls for strategic responses and different approach in sustainable development.

The transport sector plays an important role in enabling processes of urban development by facilitating movement of goods and people and improving accessibility to them. It is the case worldwide within urban areas and especially in rural areas, where transport and dispersed land allocation with low density development have enormous direct impact on land resources and consumption of energy, both for transport needs in terms of distances traveled, mainly done with the high usage of the private car and with high costs for infrastructure and energy needed for maintenance of desired quality of life (Radosavljević, Lalović & Đorđević 2013). Low density development has also high economic, environmental, and social impacts and costs in terms of traffic congestion and time spent commuting; air, land, and water pollution; health conditions and even obesity among population; decreased or no accessibility for particular social groups, namely the poor and low income.

[1] Corresponding author: yros@arh.bg.ac.rs

In such a context, urban growth possesses represent considerable planning, governance, and management challenges for the variety of urban sectors, such as housing, sanitation, water, and transportation. Previous urban planning experience in many countries has proved that the common practice and urban design solutions have a negative impact on overall energy consumption, and existing planning methodology make it difficult to identify and locate resources and capacity for sustainable urban development, climate change mitigation and adaptation measures, and sustainable transport within it. At the same time, many cities around the world and increasing number of citizens and government officials began advocating a smarter approach to land use planning, which should include: a compact community development and urban groups, more transportation choices through increased accessibility and inter-modality in cities, mixed land use, and the preservation and expansion of green and agricultural land (Lalović, Živković and Radosavljević, 2011). This approach achieves a more ecological, economic, and social benefits that contribute to the quality of life in cities and rural areas, and at the same time serves for the reduction of energy consumption and greenhouse gas emissions.

## 2. Context – Sustainable Development of Rural Tourism Project

German International Cooperation (GIZ) wanted to empower and foster individual municipalities of Eastern Serbia to apply jointly for local economic development programs in the project Municipal economic development in the Danube Region, which lasted from 2010-2012. With the project's support, capacity building activities were carried out at the national and regional level in Serbia with Ministry of Economy and Regional Development, Ministry of Agriculture, and the Regional Development Agency of Eastern Serbia (RARIS). In this way, GIZ intention was to initiate strategic networking among bordering municipalities in Eastern Serbia. Four bordering municipalities in the Lower Danube Region, namely Negotin, Golubac, Kladovo, and Donji Milanovac participated with the Project Authentic ethno from Eastern Serbia for the sustainable rural tourism development. Faculty of Architecture was selected as a consultant agency in order to facilitate the process in the project and produce Integrated Plan for Sustainable Development of Rural Tourism (Radosavljević et al. 2012a). We will present and discuss main critical aspects of the process of project formulation that have led to sustainable measures and projects in the field of mobility management and travel awareness in the context of urban and rural tourism development.

### 2.1. Collaborative approach for the mobilization and inclusion of key stakeholders

First problems and constrains that have appeared in the process of the project drafting were related to previous relations of municipal representatives and low level of communication including conflicting attitudes how rural tourism and mobility should be developed in a sustainable way. This problem was mostly evident between Negotin municipality on one side and Kladovo, Donji Milanovac and Golubac municipalities on the other.

We found actors who barely communicated with each other, although formally supposed to cooperate on the basis of the signed Memorandum of Understanding. We realized soon that the cause of poor relations is a consequence of the process of goals formulation and the structure of the project, and the other one is tender for hiring expert planning team, which was conducted by the municipality Negotin. The other three partner municipalities in the project were not engaged enough nor consulted in the process of the formulation of the project goals and about the terms of the reference for consulting agency. All these aspects led to the situation that they felt having secondary role towards the leading role of the Negotin municipality.

In addition, we noted that the diversity of their individual knowledge and vision of the approach to the rural tourism development and sustainable transport within it was based on the conflicting views and interests of attracting investment and tourists to their respective municipality. Additionally, these obstacles were based on low personal and political relationships which created an atmosphere full of antagonism and rhetorical speech, which even questioned the semantic meaning of concepts, such as discussion about the interpretation of the concept of rural tourism, distrust towards planning team, and doubts about mobility issues. This situation is typical in the Serbian context of planning and management of urban development due to our experience since there is rarely a simple situation in which actors agree from the beginning of the planning process.

The specific approach for trust building that the expert team from the Faculty of Architecture have chosen for the establishment of dialogue and communication between the opposing actors and institutions from four municipalities was to shift the focus from the traditional problem solving of poor relations, towards mutual agreement that is in the interests of all partners to find ways to implement the project in cooperation with each other. The argument we have used was that the traditional spatial and urban plans, state incentives and a number of strategies from the national level were not sufficient as simulative and operational framework for the sustainable development of rural tourism and mobility in Eastern Serbia and that there are no clear mechanisms for the implementation of the above mentioned plans. In that respect, the first step that followed was to bring up issues to the representatives of the municipalities and their tourism organizations their own interdependence and the need for efficient collaborative management and promotion of their territory at the regional level. We presented them the advantages and benefits available, as well as disadvantages and losses that arise unless they cooperate on a competitive market.

We started from the premises that if we want to have real governance by the people and integrated plan that builds from the bottom to the top and from the community towards government, we have to include different groups of relevant stakeholders, as well as to ensure their participation, attention and commitment in the various stages of the planning process.

There were several goals of involvement of various relevant individual and group local stakeholders and organizations in the process. We wanted to directly collect data and information about their values, interests and needs, as well as specific knowledge they individually possess. It was also important to establish personal contact with stakeholders in order to gain the necessary confidence. The confidence has worked as a channel through which we firstly expanded awareness about the need for cooperation. Secondly, later on in the process, we have eventually expanded the idea about individual measures related to the climate change mitigation and adaptation measures.

After we have established communication between actors of the public sector in the network, the following step we have made together with them was to map, mobilize and involve private sector and vilalgers using several key elements and methods in the process. We have used several different methods and techniques, such as information, consultation, and inclusion of key stakeholders in order to gain feedback and suggestions from them on solutions in the desicion making process throughout diferent stages of the strategy formulating process.

We have recognized leadership qualities of individual municipal representatives and their enthusiasm related to the project, thanks to which they have mobilized stakeholders. It has been one of the channels for the dissemination of information and further involvement of local stakeholders during the site visits, workshops and informal meetings that have been organized. The second channel of communication was local media. Intensive field visits of different villages and households including wine cellars in villages Rogljevo and Rajac, in the period from October 2011 till Jun 2012 had multiple roles. Beside ordinary data collection process, talking with individual local entrepreneurs, farmers and rural householders to determine the values that the villagers have individually was essential. In this way, stakeholders were engaged in the planning process.

A series of workshops were held in four municipalities and in Rogljevo and Rajac wine cellars villages in the period from October 2011 till Jun 2012. Stakeholders had the chance to react and express their ideas, suggestions and remarks on proposed preliminary proposals done by the expert team from the Faculty of Architecture. They have also influenced design guidelines and final project during workshops, regular meetings and voting for best solutions during exhibitions.

### 2.2 Management mobility measures for increased accessibility and tourist's comfort

We will present main outcomes of the project and formulated management mobility measures.

Depending on the different types of tourists and their needs, and of specific and typical forms of rural tourism that already exist or could be formed, we have recognized the potential for increasing mobility in four main groups of management mobility measures:

- Connectivity, both external and internal;
- Signalization, mainly as signposts;
- Modes of transport, mainly as non motorized transport like increasing the use of bicycles and horse transportation, and pedestrian movement; and,
- Services, both direct in transport and information and indirect, such as tourist facilities.

The potential of improved mobility and connectivity for all four groups of management mobility measures has been recognized by relevant stakeholders as a key instrument that can support integrated approach of urban management and fulfill goals of sustainable development in four municipalities, as follows:

- Supports the access to urban and rural economic centers and individual villages in tourist cluster of four destinations in Eastern Serbia;
- Improves the quality of life, health and safety of residents and tourists;
- Contributes to mitigation of climate change, flooding and impacts caused by road transportation onto the living environment;
- Optimizes the use of expensive infrastructure and significantly reduces the need of high infrastructure investments.

This means that the potential benefits have been identified both for tourists and local residents. It also mean that with minimal investments, a general increase can be expected for the use of existing transport networks and modes of transport in order to achieve a safe and comfortable transportation, accessibility and mobility for municipal inhabitants and goods on one hand, and various types of tourists on the other hand.

In order that tourists fully experience everyday rural and natural way of life, some of the integrated rural tourism and mobility measures were developed. Experts from the Faculty of Architecture proposed measures that were integrating existing habits of local villagers both in terms of their tourist offer and mobility issues. In such a way, residents and local authorities suggested urban management and mobility measures, such as tourism packages and products like a typical day going to the harvest in a field by foot or tractor; a tractor ride with rural hosts to another village, town or cultural monument; or pedestrian tour by forest guards in national parks and protected nature reservations.

On the other hand, the potential use of existing village roads and trails, beside their primary function for vehicle transportation, has been identified for other tourist purposes and activities in accordance with specific tourist's requirements. Some of the proposed and accepted measures were mapping, signposts, and the use of gravel country roads and forest trails for country biking and other recreational purposes, such as pedestrian thematic routes.

In the context of tourist's visits and tours of cultural and historic heritage and natural environment in the region and throughout the area, where accommodation in the village is the base center where a tourist comes back after the tour, a good information system is needed for the quality of tourist's stay and effective and efficient use of both tourist's and municipal resources and time. In this sense, the potential for improving the information that tourists might use are recognized for a variety of basic services, such as information on nearby banks, post offices, shops or the nearest health facility, as well as services related to the ease of easily navigate in space, arriving to the destination by car, bicycle, water and pedestrian transportation.

Such an information system is not only a database available on the Web site of municipalities and local tourism organizations. It can additionally include printed maps with various types of tourist information and combined tourist packages for different tours depending on the time available; travel awareness activities that include educated rural hosts to help domestic and foreign tourists and provide quality information in general and even provide local transport in particular; basic traffic signals and system in the form of info-boards for pedestrian and bicycle paths.

Specific mobility measures such as thematic hiking and biking trails are planned for realization in few phases and could be further arranged at the local level by involving local residents and owners of individual parcels, experts in specific areas depending on the topic route, tourist experts, etc. Basic information on the maps should include, except the basic data about the natural, cultural, and tourist info centers and basic service, data that contain information about the distance, slope, time of movement form the bottom to the top of the destination and vice versa on the hilly areas such as eastern Serbia, and so on.

When it comes to signage, it has been agreed on the level of four municipalities that tourist information signs for the pedestrian and bicycle paths and larger boards with description of specific places of interest and destinations should be placed the theme trails and places of interest. They are supposed to be placed either in terms of individual ambient sites as well as the cultural, historical and environmental or natural systems; sometimes as a individual plant species along the trail, on other occasions signs for track-passing between fenced plots etc. The aim of signage is done in order to facilitate tourists for their better orientation in space and provision of complete comfort on one side, but it is meant also for increasing accessibility for local residents and villagers in their everyday activities.

Alternative modes of transportation, such as vehicles with alternative energy source, bicycles and walking, village carts, horse riding, as well as mobility measures which restrict access of cars to protected and unprotected rural environmental entity, protected natural reservation areas, represent, together with the regulation of parking for buses and cars, not only the potential but also the commitment to environmentally clean development of rural tourism, the protection of nature and the complete experience of tourists.

Another important aspect of mobility measures was seen in the provision of services for individual tourists in terms of bicycle repair shops, pit stops, refreshment and resting areas, especially for cyclists in the spring and autumn period. This measure is gaining importance, especially since it is associated to the expected increase in the number of bikers or cycle tourists, how they are called, which cycle along the Danube bike trail, that is a part of the Eurovelo 6 route from the Atlantic Ocean to the Black Sea. In that respect, it has been agreed among municipal representatives and villagers that these specific bicycle oriented places should be located next to the corridor along the Danube. Another level is that they can be located in villages in the Danube's hinterland. It does not mean that each time a new facility has to be erected, but that soft urban management and mobility measures might be implemented. Stakeholders accepted that existing rural accommodation could be used as a first phase and that the villagers themselves could be further trained to repair bikes.

Residents, municipal representatives and experts from the Faculty of Architecture placed full attention for the Rajac and Rogljevo wine cellars due to it's huge national heritage value. It has been acknowledged that it was of extreme importance to make a mobility plan and movement independently and further incorporated into a formal plan, such as a plan of detailed regulation or plan of designated tourist area. Mobility plan should include possibilities of alternative modes of transportation, pedestrian paths with a suitable slope for walking all on one hand. On the other hand, plan should also as agreed contain restrictive measures, such as bus restrictions and individual cars access for tourists coming from neighboring villages of Rajac and Rogljevo the vine cellars.

This mobility management measures would significantly reduce unfavorable effects to the surrounding protected ambient of vine cellars. Such measures are high on the priority agenda of local villagers in the vine cellars. This approach involves the use of the urban management and mobility scheme, in which, on one hand, the owners and tenants of vine cellars and emergency vehicles have constant daily access; and on the other hand, maintenance services, subcontractors and suppliers have time managed limited access, and tourists-visitors have partly prohibited or restricted access with their own car. This means that a provision of a designated area for bus parking and individual tourist's cars in the neighboring villages is needed. It also means that with such a restricted access, a provision of additional modes of transport to the vine cellars is a must. Measures includes activities that already exist, but need additional further support from the municipality of Negotin; and activities that are going to be realized in future, that are costly, but achievable. As we mentioned before, such activities might include driving on a villagers tractor, small electric vehicle, or even if the opportunity occurs in the future, physical elements like a sloping elevator. Vine cellars of Smedovac could also have above mentioned restrictions and possibilities.

## 3. Conclusion

We have argued in this paper that it is possible to work within the community, with ordinary people, households, vine producers, and with local government representatives, incrementally, trough site visits talking with people, participative interactive workshops and formal presentations in order to build trust, raise awareness and recommend sustainable measures and projects in the field of mobility and travel awareness for urban and rural tourism development aiming at identifying possible improvement directions in Serbian context.

Due to the inclusion of all relevant stakeholders and joint dialogue between them, we have raised the awareness of the importance of climate change among them and adaptation and mitigation measures which included mobility management measures and sustainable development of rural tourism.

As a consequence, all stakeholders agreed that the proposed measures and guidelines in the Integrated Plan for Sustainable Development of Rural Tourism in Eastern Serbia (Radosavljevic et al. 2012a) and the Guide for the regulation of rural tourism households (Radosavljevic et al. 2012b) incorporate recommendations based on the principles of rational use of land and the use of renewable energy. Those principles included measures on: 1) alternative forms of transportation for residents and tourists, especially using non-motorized transportation modes and reduction of air pollution; 2) use of traditional ways of building in harmony with nature and the environment at the village level and use of environmentally friendly materials; 3) establishment of ecological networks and biodiversity for ecotourism development; and, 4) travel awareness and training activities for quality tourism offer and sustainable development.

We conclude from our research that planners and urban designers can support local communities in the process of adapting to climate change through mapping and involvement of relevant actors and stakeholders from the very beginning, and identify and overcome regulatory and institutional barriers in the process of collaborative goals definition and integrated measures.

We have also found that in order for the collaborative action on the dissemination of knowledge on climate change and adaption and mitigation measures to work in Serbian context, several conditions have to be met. First, prior to developing awareness of climate change, a necessary prerequisite in Serbia at the regional and local level is the collaboration between often conflicting stakeholders, primarily the public sector. Collaboration is essential, since only trough the establishment of a dialogue and raising awareness process, dissemination of information and knowledge can enhance transformative ideas about the importance and impact of climate change. That leads to the next step, in which methods, measures, and tools for mitigation and adaptation to climate change are discussed and formulated.

**Acknowledgements**

This paper is part of the research conducted through the Scientific Project “Spatial, Environmental, Energy and Social Aspects of Developing Settlements and Climate Change – mutual impacts” (TR 36035), financed by the Ministry of Education, Science and Technological Development of the Republic of Serbia from 2011 to 2014.

**References**


IPCC, 2007a. *Climate change 2007: Synthesis report*. Contributions of Working Groups I, II, and III to the fourth assessment report of the Intergovernmental Panel on Climate Change. Geneva: IPCC. Available from Internet <http://www.ipcc.ch/publications_and_data/ar4/syr/en/contents.html>

IPCC, 2007b. Summary for Policymakers. In: *Climate Change 2007: Mitigation. Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. Available from Internet<http://www.ipcc.ch/publications_and_data/ar4/syr/en/contents.html>

Lalović K., Živković J., & Radosavljević U. 2011. Urban Governance Challenges in Planning and Implementation of Climate Change Adaptation and Mitigation Measures. In *Proceedings of the International Academic and Professional Conference "Architecture and Urban Planning, Civil Engineering, Geodesy - Past, Present, Future*. Banja Luka: Faculty of Architecture University of Banja Luka. pp. 2333-2345.

Radosavljević, U., Nedeljković, S., Đukanović, Z., Bobić, A., Milojkić, D., i Simić, I., 2012a. *Integrated Plan for Sustainable Development of Rural Tourism*. Belgrade: Faculty of Architecture University of Belgrade.

Radosavljević, U., Nedeljković, S., Simić, I., Bobić, A., i Milojkić, D., 2012b. Guide for the regulation of rural tourism households. Opština Negotin.

Radosavljević, U., Lalović, K. & Đorđević, A. 2013. Sustainable Urban Development & Concept of Mobility Management in Belgrade. In N. Martins (Ed.), *Proceedings of the conference Sustainable Urban & Transport Planning – SUTP 2013* Belgrade: United Nations Development Programme (UNDP) Serbia. pp. 78-90.

# EXAMPLES OF ALTERNATIVE TRANSPORTATION SYSTEMS IN URBAN ENVIRONMENTS

**Danilo S. Furundžić[1], Uroš Radosavljević[2]**
*[1,2]University of Belgrade Faculty of Architecture, Belgrade, Serbia*

**Abstract:** For the first time in human history, more than half of the world's population is living in cities, with every third citizen living in heavily populated urban informal settlements. In the next 50 years, more than 3 billion people, or approximately two-thirds of future urban population, could be living in slums. What will transport in urban environments look like in the future? Decades ago, people imagined modern metropolises as cities in science fiction movies, with flying cars and buildings above the ground. In reality, futuristic urban transport pod car systems are likely to be succeeded by rickshaws and bicycles. While we aspire to start living in ultramodern cities of the new millennium, alternative transportation systems that are appearing in overcrowded metropolises around the world are bringing us a taste of reality. Urban transport planners have started to resolve this topic, considering alternative transportation systems in urban environments. Slums are not only a social and urban phenomenon, but also a product of the lack of mass public transport. This paper will try to demystify urban transport phenomenon and present the seven W's (Who, What, When, Where, Why, for Whom, and hoW) on the topic. Specific urban environments conditions, that have generated alternative forms of public transport, can be recognized as a base for classifying different implemented key-study examples of alternative transportationsystems in urban environments. In this way this paper can be used as a basic survey paper on the topic.

**Keywords**: transportation, mobility, alternative transportation, informal settlements, slums.

## INTRODUCTION

Cities, as mayor drivers of social and economic development, encourage the migration of people. It is estimated that by year 2050 more than two-thirds of the world's population, approximately 6.2 billion people will be living in cities. [Pinderhughes 2004] The increasing growing number of city inhabitants creates pressure on rational usage of city resources, especially in the domain of land usage. The planning profession had engaged the problem of intent city expansion into the surrounding suburbia and rural areas. This territorial expansion created the need for more frequent and longer trips to central urban areas. Cities had suddenly and uncontrollably grown, in such a way that a sufficiently system of public transport could not be developed, and individual transport became essential. Car-oriented cities were created. [Vugt, Lange, Meertens 1996]

With the number of cars increasing, the level of traffic was increasing as well, creating delays and congestion on existing infrastructure grids. The solution for the emerging problem requested new infrastructure capacities that demanded new land resources and urged dislocation of city activities, furthermore enhancing the growing demand for privately owned motor vehicles.

The growing number of cars, as well as the increasing traffic congestion, led to significant problems like pollution with carbon dioxide, noise pollution and reduced public safety. Experts agree that the universal presence of the privately owned motor vehicle as the main mode of transporting people and goods breaks the quality of life in the cities. Individual transport, in modern urban communities, has become untenable. Experts from various fields are trying to deal with the given problem in different ways. Reducing the usage of motor vehicles, as well as the promoting and developing public transport and alternative non-motorized individual transport. Countries leading the changes in alternative transportation systems are economic superpowers like China, Japan, USA and the countries of Western Europe. [Pinderhughes 2004]

## 1. TRANSPORT

Transport is an infrastructural system that efficiently transfers people and goods. It helps the development of trade, and in a long run provides socio-economic development and vitality of the city. [Korica 2008] Transport infrastructure consists of a set of different modes of transport: air, water, road, rail, cable etc. The system can be divided into infrastructure, vehicles and operations. The base of the infrastructure is the network (of highways, tunnels, bridges, streets etc.) that accepts and facilities (stations, pumps, maintenance centers, garages, parking lots, individual vehicles etc.) that regulate transport. These two combined ensure normal operation, and are equally important in all aspects of the system.

Road transport takes 80% of all traffic of goods and passengers. It is leading in the length of its network (24 million km or 70% of the world transport system), and is in the general economics a necessity. Only in year 2012 70.520.493 vehicles were produced of which 52.726.117 individual cars. The automotive industry has an average annual turnover of 1.889.840 million euros, employs over 8 million people, and 50 million people are indirectly dependent on it. Transportation plays a critical role in the creation of the urban center and its forms, and is the main factor of how successfully a society develops. Forms of transport greatly vary depending on the level of development of the country, socio-economic conditions and geographic characteristics.

[1]Correspondingauthor: dfurundzic@gmail.com

The transport infrastructure affects the city's physical structure, spatial arrangement of the surroundings routs of travel and generally the way the urban area is used. [Pinderhughes 2004] All transport activities that appear with the city urban area is considered city traffic.

## 2. CITY TRAFFIC IN DEVELOPING COUNTRIES

Underdeveloped countries and developing countries represent a large group of countries with a lower living standard, underdeveloped industrial base, and low per capita income. City traffic in such countries is very direct, mainly consisting of travel form home to work. According to research in 1990, 40-50% of all trips in such urban environments were to work, traveling to school 20-35%, and only 15% of travel was for some other business-social purposes. [Armstrong-Wright 1997] Transport in these countries is usually based on mass public transport, walking and individual small motorized and non-motorized vehiclesdriven by human or animal power. [Pinderhughes 2004] According to a study conducted by the World Bank, public transport makes up 50-70% of all city traffic, where leading the way are small motorized and non-motorized vehicles (such as bicycles, motorcycles, tuk-tuks, rickshaws etc.) as the cheapest and most convenient mean of transportation. These vehicles are an imperative in the area of human activity, as their role is to facilitate the transport of people and goods in a safe, reliable and cost-effective manner. [Enoch 2012] Still such vehicles are considered inadequate modes of transport because they cause traffic jams and congestion.
Although the use of cars is growing in recent years, car traffic is a negligible part in the transportation system of Africa, Latin-America, India and Southeast Asia. The population in these countries depends on mass public transport, but things are changing. Most are experiencing rapid population growth and urbanization, and many have fast-growing economies. Enhanced mobility has positive effects on the economy allowing, more efficient transport of goods and people, improving the performance of work and income. Although the benefits are enormous, the costs can be substantial as wellThe number of privately owned motor vehicles is increasing in almost all developing countries. China is a good example of this, as one of the largest car manufacturers in the world. The Government of the Republic of China stimulates, through a variety of projects and initiatives, car purchases. Despite the fact that China has the most polluted air in the world, urban planners and government leaders are creating regulations and laws that favors the usage of cars. This phenomenon is connected not only to the increase in personal income but also to the idea of maintaining the economic development of the country making internal consumer demands.

## 3. CITY TRAFFIC IN DEVELOPED COUNTRIES

Contrary to transport used in developing countries, people in developed countries use fewer types of transport. Developed countries have the largest stake in the global use of cars and a very small percentage of small motorized and non-motorized vehicles. Traffic in these countries is very diverse, as travel form home to work takes only a third of all trips, where two-thirds are social and recreational trips. The population is facing a big problem of dependence on privately owned motor vehicles. The use of cars has risen over the past 5 decades. The number of motorized vehicles increased from 75 million to 675 million between 1950 and 1990. Almost 80% of these vehicles is primarily used for individual transport. In the period between 1970 and 1990, the total number of kilometers flown per passenger increased by 90% in Western Europe. [Armstrong-Wright 1997]
There is a misbalance in energy usage. For example, although USA accounts for less than 5% of the world's population, it spends more than one third of the energy for transport, at the global level. In 2000 in USA there were more respective owners of cars than registered voters.A resident of North American consumes five times more energy on transport than the average Japanese, and nearly three times more energy than a citizen of Western Europe. [Pinderhughes 2004] Europeans have a greater tendency to use alternative forms of transportation such as bus, train or bicycles. Cars in Europe take between 11 - 26 % of all trips, and walking and cycling between 33 - 50 %. Still from 1970 to 1995 car traffic in Europe has doubled, for example the number of private cars in the UK has increased by 93 % from 1971 to 1992. Although, western Europe is understanding the sustainable issues of car dependence, and is making its green strategy, the new members of EU that are in economic transition are busting their economy by stimulating car purchases. Between 1995 and 1996, car sales in Poland grew by 40 percent, in Warsaw in 1996, there were about 400 cars per 1,000 people and motor traffic is becoming a major source of air and noise pollution. [Wrigh 1997]

## 4. INDIVIDUAL TRANSPORT

There are many reasons for choosing car over other means of transport and alternatives, such as: speed, time, cost, flexibility, as well as social reasons such as comfort, status and security of transport. However, cars have multiple negative consequences, especially from an environmental and economic point of view. They contribute to the pollution of the environment (during production polluting the environment, and afterwards by using fossil fuels as an energy source creating harmful gases) directly or indirectly, and thus affecting humane health. For example, from the International Energy Agency (IEA) the total $CO_2$ (carbon dioxide) emissions in 2010 by sectors were as such: electricity and heat production 41,2%, energy industry own use 5,2%, industry and construction 20,4%, households 6,2%, other sectors 4,6% and traffic 22,3%, where $CO_2$ emissions from cars took an astonishing 18,5%.[Wrigh 1997]The transportation sector is nowadays the fastest growing, and is producing an enormous amount of greenhouse gases.It is expected that emission of $CO_2$ rises for 57% by 2030.A total of over 500 million cars that use gasoline, combined with thousands of power plants that use coal, make the main driver of climate changes. [Brown 2001]Cars account for more than a quarter of the world's production of carbon dioxide in whole, through production and by using fossil fuels. Accordingly, the reduction of emissions from this sector will have a major impact on air quality, health and global warming.

Noise pollution is also one of the most important problems in modern transport. Noise at the edge of highway can reach a level of over 80 decibels-dB, which is a very high degree of nose disturbance, and with long-term exposure, can cause permanent harm to hearing and neuro vegetative systems. It is estimated that 170 million people that live in the so-called gray areas of noise pollution in suburban areas (with 55-65 dB) are addressed.[Brown 2001]From the social point of view, the use of a car is threatening the quality of urban life, via contamination, pollution, reduced efficiency and expensive cost of transportation. In addition, there is an increasing number of traffic accidents, reduced public safety and general alienation of people in traffic.

Opposite of the common opinion, excessive use of cars may cause downside economic effects as well, as a country can easily badly invested in unsustainable transport infrastructure systems and make irreversible debts. The economic impact is also felt at the local level. Maintaining a vast transport network is expensive. Frequent traffic jams can stifle the development of the local economy and thus reduce business growth. Economic efficiency of the transport system is reflected in the cost of infrastructure, transport costs, time costs, and so on. Transportation costs of using cars are 30 - 40% higher compared to the cost of public transport. Time costs are another important factor. To solve the existing traffic congestion by building new roads to city residents to save time is not possible. One cannot achieve the expected benefits. Although the use of a car makes it easy to find and execute jobs, the money that is spent for the fuel is forcing people to leave the periphery and move closer to the urban center of the city, recanting from the comfort of a single family house. Congestion reduces efficiency of the economy because employees spend more time on the way to work. Facilities for passenger traffic have a high cost of building, and occupy a large part of central urban space. Usually traffic takes 20% of city land use from the entire city and in some areas like downtown up to 40% of the total surface area, making it undoubtedly an important item in the cityland economy.
The main objective of traffic is improving overall city mobility in accordance with the social, economic and environmental development goals. In urban areas with high urban activity, traffic congestion cannot be avoided by increasing the capacity of the traffic network, with new construction of roads and parking. Based on previous experience of city development and theory, it is a fact that a synchronized urban transport system of roads and parking spaces that would allow unlimited use of cars is not possibility to build.

## 5. PUBLIC TRANSPORT

Through comparison of the advantages and disadvantages of both modes of transport, individual and public, it is possible to draw some conclusions. Private transport has advantages of privacy, comfort, direct speed, has independence and freedom of movement, has more possibility for destinations and traveling routs, can be paused, is flexible, and has additional possibility for transporting goods. Disadvantages are in price (fuel, tolls, taxes), of parking, usage stress (traffic jams and congestion), security regarding accidents, isolation, and is not urban efficient.Public transport has advantages of price, it runs automatically, is less stressful to use, one can relax, has possibility of quality time while driving (reading, completion of certain obligations, etc.), lower concentrations of pollution, is social active, is more urban efficient, and has a high velocity of transportation (using special lanes for public transport etc.). Disadvantages are in lost time while waiting, comfort, security, lack of specific paths and traveling routes, need of having another form of mobility, occasional unreliability, possible overcrowdings, and has no additional possibility for transporting goods.
Public transport is a transport service available to the all public. From aspects of environmental protection, health and economic prosperity, the use of mass (large-scale) public transport has many advantages compared to transport with privately owned motor vehicles. Benefits of public transport compared are multiple. Internal combustion engines are on top of that 3 to 4 times less efficient than electric motors. Globally for example rail transport produces 35 times less carbon dioxide than cars, and spends an average of 4 times less energy. [Maletin 2003] Metro consumes 0.03 kg of oil per passenger - per kilometer, a car 0056 kg, what is very close the consumption of air travel at 0076 kg. Less energy-demanding and more environmentally friendly than other modes of transport, public transport comes to less pollution per person, and also reduces traffic congestion. Investment in public transport is therefore the future for city commuting.
In most major cities in Europe and the USA public transport systems are developed and regularly maintained. However, in the cities of Africa and Asia, as well as in the cities of Eastern Europe, the system is not developed. Statistics show an increase in the use of public transport, over the last ten years in many countries. Between 2004 and 2008 there was a large increase in the use of mass transportation in Europe andUSA of about 11-12% and even 20% in the UK. A recent survey in the USA shows that the cost of driving, added up with the cost of making infrastructure for driving, the cost of regulation air pollution, noise, time cost of traffic jams and accidents (about 125 million), is in reality 3 times higher than the price that are paying users of highways via tolls and petrol. That means that driving individual owned motor vehicle is constantly draining our economy. Some economics concluded that it would be up to 5 times cheaper to run a city without a car, and that such an outrage idea is implementable. [Beirao& Cabral, 2007]

## 6. KEY-STUDY EXAMPLES OF INNOVATED TRANSPORT

The purpose of a road is to transport people, not necessarily cars. With this in mind alternative transportation systems try to respect this statement and move more people on the existing roads build for cars.From the perspective of the quality of urban life of the population, it is generally accepted that traffic has a significant impact on the social, economic and natural environment, and that overcrowded, polluted and noisy cities are not adequate for life. As the problems associated with the use of cars are piling up, with the fact that the growth in motorized transport has negative effects on the economy, and that there is a collapse of mobility in existing megalopolises, the traffic planning is increasingly discussing about alternative innovated transport.
Studies show that public and non-motorized modes of transport offer significant advantages over individual and classical public transport in urban communities. Besides being cheaper to produce, purchase and maintenance, they provide more transport opportunities, emit lower exhaust emissions and reduce congestion. There are numerous examples of cities in the world testifies to the strong initiative to reduce the negative impact of transport on the environment.

### 6.1. STREETCAR(PORTLAND)

Portland became the first city in USA that introduced a modern tram network in order to improve the existing system of public transport based on the bus and light rail. These trams were shorter, lighter, and narrower and were more flexible to maneuver in streets. A tram networks could be build faster than light rail, commuter rail or subway, and was sufficiently green. Most streets supported the weight of the tram route without additional construction work, thereby reducing the investment cost and timeframe. Setting the tram tracks in relation to the underground infrastructure such as water supply and transmission remained a problem that had to be addressed. Nevertheless, by 2001 funding was carried out by the state through a public-private partnership, and the new tram system was build. Cost of construction of the first phase amounted to $ 56.9 million, including the purchase of seven trams.The tram system was called: Streetcar, and the lines were design in such a way to provide an alternative to driving for people who are already in the city. Stations are arranged to cover every part of the city and are set in accordance with the characteristic urban points in Portland. The line of the tram connects northwestern and southwestern of the city, so they can take in all passengers coming by light rail from the suburbs. The land use is minimal, as a greater part of its route, the Streetcar share the right lane with cars next to the parking lanes. At tram stations street parking is replaced with recessed curbs, where trams stop only on passenger demand. To ensure the flow of traffic, trams are equipped with a wireless system that allows them to influence the traffic signs. This system also allows the tram to offer real-time arrival information to passengers waiting at the tram stop. The popularity of the tram ride is on the rise. Although planned to transport up to 3,000 passengers a day, after 6 years of use, this number had grown to as many as 10,000 passengers a day.

### 6.2. CRGO (DRESDEN, GERMANY)

In Dresden tram traffic was applied in order to reduce air pollution and traffic congestion caused by heavy trucks. Dresden CarGo tram was designed to transport parts and materials from Volkswagen Friedrichstadt logistics centers outside of Dresden in a new sustainable factory in the center of Dresden. CarGo tram commutes on the existing city passenger tram rail network. Each tram carries 60 tons of material, which is equivalent to a truck with three trailers. What is interesting is that the tram tracks lead directly to the factory where old internal factory wagon tracks are integrated in the network. Trams were scheduled to be running 21 hours a day and carrying about 2,000 tons, in this way replacing transport more than 100 trucks per day.The success of the Dresden CarGo tram encouraged planners across Europe to see whether similar systems can work in other places, especially within old cities, where there is a problem of mixing passenger and cargo transport. As well as in urban areas where there are several companies that deliver goods to the same location. In AmsterdamCityCargo pilot program began in 2007. It is estimated that 53 trams and 600 electric trucks replaced about 1,200 delivery trucks before used in Amsterdam. This reduction in truck traffic resulted in the reduction of air pollution and noise, but also was a great money saver.

### 6.3. SCHWEBEBAHN(WUPPERTAL, GERMANY)

The Schwebebahn system is used in Wuppertal in Germany since 1901. They are an important addition to the rail transport technologies because they save city space. The system can be placed in narrow rows, or across rivers or along existing roads without contributing to traffic congestion. Supporting structures are constructed of steel bearing components, and the railline is attached to the underside of the structure, so the vehicles are hanging on wheels that move on rails. Schwebebahn in Wuppertal is the oldest example of AirTram system in the world. Hanging 8-12 meters in the air, making a route of 15 kilometers and caring nearly 80,000 people a day this system is the primary public transportation system in the city of 500.000 inhabitants. This form of transport is very secure, as only five accidents and one victim had the system encountered for more then a century of existence.In the new millennium,thiAirTram it is getting a makeover with the modern Monorails and or SkyBus systems.Unlike conventional rail systems, these systems are separated from other traffic and pedestrians, usually appearing in strong commuting areas like airports (New York, London, Paris, Singapore) as well as downtown centers (Beijing, Venice, Tokyo). They avoid red lights, intersection turns, and traffic jams. Unlike subways passengers enjoy sunlight and views, expensive and noisy ventilation systems are not necessary, and most importantly they obtain electricity from the track structure, eliminating costly and unsightly overhead power lines and poles.

### 6.4. BRT (CURITIBA, BRAZIL)

The Brazilian city of Curitiba adopted the practice of planning that integrates and focuses on the development of sustainable transport. After an initial unsuccessful plan of expanding the road infrastructure due to limitations of finance, the institute for research and planning implemented draconic measures to reduce the negative impact of private transport. Curitiba Master Plan reduced the spread of the city, introduced laws on zoning and land-use, created a transit-oriented city, reduced turnovers on individual transport, ensured effective public transport and engaged customers to use it.

Curitiba, a city of more than 1.5 million people had a big problem with traffic congestion. The public transport consisted of buses operated via private companies, making large traffic jams. Due to the lack of funds for constructing the subway, the city decided to improve public transport creating a new transport system BRT (Bus Rapid Transit) that has an independent bus networked, based on a bus that uses a priority lane that is integrated in the pedestrian zone. Thepriority lane allows the bus to operate independently of the surrounding traffic and is “free”, organized and rapid such as any metro system, but costs much less. With a total length of about 45 miles, and making about 2 million transfers a day the Curitiba BRT is considered a success story, and is planned to be introduced in other Brazilian cities.

It is believed that the BRT in the future will become a major alternative to the existing public transport, especially in cities with limited budget and free land space, as the ones in developing countries. Financial resources required to build 10 BRT lines are equal to the construction of 1 metro line and have almost the same passenger frequency. Analysts believe that BRT can be a good candidate for the existing road network to be adapted because of lower costs compared to other transit systems. The BRT system has in such ways been successfully applied in France, Turkey, and the USA.

In Asia TransJakarta bus system is the largest BRT in the world, and is significantly reducing car traffic. Before this system, public transport in Jakarta was unsatisfactory. Congestion in the city streets, increasing the number of car users has led to degrading air quality in the city. Practice shows that 14% of travelers have changed the car for the BRT system.

## 7. NON-MOTORIZED TRANSPORT

Cities are increasingly turning to the use of non-motorized transport because of the benefits it brings to the environmental and economic aspects. Today, for a variety of reasons (health, environmental, financial, and others) cycling is increasingly popular, and is the most outspread and effective form of non-motorized transportation. The bicycle originated in Europe in the 19th century as a mean of recreation and since then has spread worldwide. The concept has not changed, except that today's bicycles are much easier and safer to ride.

Despite personal preferences related to cars, bicycles are not only a feasible alternative to motor vehicles, but are already an essential component of the transportation system in most developing countries, slowly gaining popularity in developed countries as a sustainable urban transport system. Cycling has become a common part of city traffic. Even at the begging there was a problem sharing traffic by cyclists, horse-drawn carts and pedestrians. Increasing levels of motorization and appearance of cars enhanced this problem. In most countries according to current traffic regulations the bicycle is considered a transport vehicle, and is therefore equal to other vehicles in traffic. This means that there are legislations on the required age of the driver, rules of movement on public roads and technical requirements. Compared with motorized vehicles, bicycles have many advantages, especially for commuting over short distances. They are a non-polluting mean of transportation that actually improves environmental quality by reducing air pollution, noise and congestion. They are much more convenient than public transport, because passengers do not have to wait, and can bypass traffic jams and save public money, and are very, very economical. Even the most expensive bicycles are much cheaper than cars. On average, a new car costs 6 times more than a new bicycle. A new highway will cost over 60 million dollars per kilometer, compared with 9 million per km of bicycle paths or tracks. Long-term savings are even greater when taking into account the savings associated with ongoing costs for fuel, repairs, registration and car insurance. Benefits provided by cycling are: zero emissions, harmless to other drivers, low purchase price and maintenance, easy access to urban areas, cheap infrastructure network, exercise and stress reduction, ability to transport goods and passengers, health etc.

Although the bicycle is a popular form of transport safety prevents most people who own bikes to ride them regularly. Most city streets are designed keeping in mind only the cars, and cycling in thestreet can be dangerous. The transport system has to have an independent network, as the bicycle is not considered an equal member of city traffic as other vehicles in practice. In reality, the most important preconditions for implementing cycling as a form of equal city transport is the existence of adequate infrastructure for the safe operation of bicycle traffic in urban areas. One must have built the infrastructure for bicycle traffic, which includes lanes and roads designed exclusively for bicycles, overpasses and underpasses for bikes at dangerous intersections, secure parking, and as important specialized space at work for changing into the work clothing.

Cycling as a mode of transportation is widely spread in many parts of the world. Asia continues are the dominant market. Between 1997 and 2002 bicycles made 52% of traffic, taking precedence over all other vehicles, including cars and public transportation in Beijing. More than 15% of travelers in Japan used bicycles to travel to work or to commute to high-speed trains, with safe parking before boarding the trains that transport them daily. Many European cities are also bicycle dependant. For example, the Dutch government made a priority the safety of cyclists and implement a wide range of plans and measures to improve security. Today as many as 50% of all trips in the Dutch cities are done on bicycles.[Pinderhughes 2004]

Bicycles are an individual form of transport, where rickshaw are the bicycle public alternative. They are vehicles on bicycle wheels that does not pollute the environment and are used for the transport passengers as well as goods. Rickshaws are used in many parts of India, across Southeast Asia and many other developing countries. They have an important role in the traffic system, not only functional, but also socialas they maintain a certain level of income of the most poor, and therefore are very popular among travelers.Rickshaws provide a cheap, reliable transportation option that is tailored to the traveler in urban communities. They are particularly useful for women, children, the elderly and the disabled who cannot with ease and comfort use public transport. For example, local rickshaw drivers in Mumbai daily bring and take children to school with great reliability and personal care. Despite these essential benefits, the city authorities in many developing countries, starting with a ban rickshaws, on the basis that the vehicle is moving slowly causing traffic jams or simply that it is not in line with the vision of the city that is modernizing, developing and progressing.[Pinderhughes 2004]

## CONCLUSION

During recent years, there is an increase of the number ofpeople living in cities, and as a consequence of this, cities are enlarging, unplanned and planned.The degree of physical expansion of newly built parts of cities often exceeds the capacity of basic public servicesandurban infrastructure.This includes the city transport infrastructure, as well as the existing traffic network. The demand of roads is growing so rapidly that the economy is struggling to supply the demand.With the increase in population there are a growing number of vehicles.Excessive car use is leading to constantcity traffic jams, congestion and pollution, degrading the quality of urban life. People who live in the suburbs often have a faster and cheaper way to get to the center, because public transport does not cover emerging areas of the city. The city center is in this way depreciated. The use of private motor vehicles leads to a number of problems, and people are commuting across larger distancesdaily. The mobility of citizens has a negative trend, which is opposite to the social demand of the new millennium. Instead of better connected, people are becoming more isolated.

Rising energy prices, rapid depletion of non-renewable resources and climate changes affect both the lives of the people and the economies of the cities across the world.The idea that the era of individual transport is on the exhale is not new.Trying to predict the future is always risky, the need for alternative transportation systems in urban environments completely independent from cars and the existing public systemwill always exist. However, the concept of sustainable development, almost certainly, will not allow the artistic vision of urban congestion, such as those in science fiction movies.Although it will take more research to look at many aspects of future transportation systems in urban environments before choosing the right path, it is not difficult to conclude that the new target is a transportation system that is automated, safe, energy efficient and environmentally friendly. One can argue if the existing transportation system has become unsustainable, but nobody can argue that the need of steady improvements of this system and promotion of alternative transportation is a must. It is not important if the changes come via modernized public transport and green automatic individual transport of the future, or non-motorized individual transport of the past, as long as the changes enable the scale of city mobility to continue constantly growing.

**REFERENCES**

Armstrong-Wright A. 1997. Public transport in Third Worl Cities.Departman of Transport. London

Black, J. A., Paez,A., &Suthanya, P. A. (2002) *Sustainable urban transportation: performance indicators and some analytical approaches.* J. Urban Plann. Dev: 184-209.

Bajin, D. (1995). *Saobraćaj u gradskojulici – ekološkiaspekt*. Beograd: ZadužbinaAndrejević

Brown L., Flavin C., French H. 2001. *State of the world*.W.W.Nortn. New York

Beirao, G. and Cabral (2007) "*Understanding attitude towards public transport and private car: A qualitative study*", Transport Pollicy, Volume 14.

www.sciencedirect.com/science/article/pii/S0967070X07000522 (downloaded 14.01.2014.)

Depolo, B., V., (2006), *Indukovanaizgradnjaiinvesticije u saobraćaju*, Beograd: SaobraćajnifakultetUniverziteta u Beogradu

Enoch P. 2012.*Sustainable Transport, Mobility Management and Travel Plans.*Ashgate Press. UK

Jakobsson; C. (2007).*Instrumental motives for private car use*. U Gärling, T. &Steg, L. (Ur.), *Threats to the Quality of Urban Life from Car Traffic: Problems, Causes, and Solutions*. Amsterdam: Elsevier.

King, J. (2012) *How Does Car Pollution Affect the Environment & Ozone Layer?*,SFGate.

www.homeguides.sfgate.com/car-pollution-affect-environment-ozone-layer-79358.html (downloaded 02.09.2014.)
Korica R. 2008. *Infrastruktura, saobracaj, urbanizam, arhitektura.*ArhitektonskiFakulteUniverziteta u Beogradu.Beograd.

Lund, H. & Clark, W.W.(2008). *Sustainable energy and transportation systemsintroduction and overview.*Utilities policy, Vol 16, 59.-62.str.

Maletin, M. (2004) *Planiranjesaobraćajaiprostora*, Građevinskifakultet, Beograd:Orion art

OECD Proceedings (1996) *Towards sustainable transportation.* Paris: OECD Publications. http://www.oecd.org/greengrowth/greening-transport/2396815.pdf (downloaded 12.12.2014)

Pinderhughes R. 2004. *Alternative urban futures –planning for sustainable development in cities throughout the world*, Rowman and Littlefield Publishing. New York. USA

VugtVanM., Lange VanP., Meertens M. 2002. Commuting by car or public transportation? A social dilemma analysis of travel mode judgements.*European Journal of Social Psychology*, Vol. 26: 373-395

Wright, A. (1997), *Alternative Urban Futures*, USA: Rowman&Litterfiel Publishers, Inc.

*Transport Logistics*

# DECISION-MAKING SUPPORT FOR INTERMODAL FREIGHT TRANSPORTATION

**Kristin Huebner[1], Melanie Weber[2]**
*[1,2]German Aerospace Center (DLR), Institute of Transportation Systems, Berlin, Germany*

**Abstract:** In case of major events or disasters rescue and emergency forces basically use their internal information and communication systems. But when increasing traffic endangers emergency measures or the transportation of people and goods, emergency forces often rely on their experience of infrastructure and routing traffic. Their situation's estimation incorporates a limited view of scale and field of operation, rather than an overall, wide-scale view involving various stakeholders. Therefore, at an earlier stage the DLR developed and tested a demonstration portal which supports decision-making by providing consolidated information and operation recommendations towards individual motor car traffic based on traffic simulation. The portal provides e.g. aerial image data, maps, isochrones of journey time and current regional traffic information. In the event of the collapse of road traffic and the necessity to sustain the transportation of goods, intermodal freight transportation is essential. Therefore, in a first prototype implementation the existing portal is extended by other modes of transportation as well as required reloading sites for the region of Brunswick. Vehicle types, follow-up and trans-shipment models are defined to enable simulation-based recommendations for operation. The portal's expansion provides a wide range of information regarding passenger and goods transportation. These include data of infrastructure and points of interests of the different modes of transportation, average data concerning transported goods and contact data of emergency forces for coordination. The portal presents a prototype of the consolidation of intermodal infrastructure information. In future, the integration of real-time intermodal traffic data and simulation is aimed at as well as real-life application by authorities.

**Keywords**: Emergency Measures, Intermodal Freight Transportation, Information Portal.

## 1. Introduction

In 2012, in Europe 91 disasters were reported whereas worldwide a total of 552 disasters was summed up. One third of the disasters were technological disasters and two thirds were categorized as natural disasters. About 70% of the reported technological disasters were transport accidents whereas almost half of the reported natural disasters were caused by flooding or extreme temperature. In total about 140,000 people were affected worldwide in 2012. (IFRC, 2014)

Disasters like the flood in Germany in June 2013 or large-scale accidents, which require longtime blocking of transportation routes, constrict the reliability of the transportation system. From this it follows disruption in the commercial transport and limitation of mobility of the general public. Moreover, rescue and emergency forces also depend on the reliability of the transportation infrastructure to manage emergency logistics as well as the recovery of the transportation infrastructure. One major challenge within humanitarian logistics is the lack of resources in terms of transportation capacity (Kovács et al., 2009). Such a situation asks for a tool which enables authorities to have an overview of the current traffic situation and which provides traffic forecast and further relevant operation information. (Bieker et al., 2012) Currently, there are no common automated traffic information and traffic decision support systems which support management decisions. Decisions have to be made by manual estimations, based on experience and the available knowledge about traffic situations (Touko Tcheumadjeu et al., 2012). Within the project VABENE – "Traffic Management for Large Scale Events and Disasters" and its following project VABENE++ which focus on traffic management for large scale events and disasters, solutions and tools for a decision-making support for emergency forces like the police force, fire service or technical relief organisations and other authorities with security responsibilities is developed. The projects aim at the efficient management of emergency logistics and the nearby traffic flow under extreme conditions and at enabling response teams to rapidly reach locations where they are needed. (DLR, 2014)

Decision-making solutions strongly support emergency forces during their rescue operations and field actions by providing consolidated reliable information and operation recommendations. It aims at facilitating communication, coordination and action. At an earlier stage the DLR developed a demonstrator for decision-making support – the EmerT-Portal (Emergency mobility of rescue forces and regular Traffic portal). The portal provides information related to passenger transport with individual motor car traffic and can be accessed by a protected web interface. It provides amongst others aerial image data, maps, isochrones of journey time and current regional traffic information for a sample region. In the event of the collapse of road traffic and the necessity to sustain the supply of an affected region and the transportation of goods as well as the provision of rescue equipment, freight transportation and other modes of transportation have to be taken into consideration. Intermodal transshipment of goods has to be displayed to allow information and operation recommendations if one transportation network collapses. Therefore, the implementation of freight transportation in addition to passenger transport as well as the extension towards all existing modes of transportation of the EmerT-portal is essential for an aspired future practical application of the portal by emergency forces. It allows a holistic solution for decision-making support. This paper presents the extension of the EmerT-portal which aims at the integration of further modes of transport to emphasize the relevance of intermodal transport within a disaster or major event. Information will be aggregated and not only focused on one mode of transportation.

[1] Corresponding author: kristin.huebner@dlr.de

## 2. Previous State of EmerT-Portal

### 2.1. Technical Realization

The EmerT-Portal is an aggregate of different systems for the detection, processing, management, storage, visualization and basic provision of information primarily of the areas transportation, public security and infrastructure. The main component of the portal is a traffic data platform (TDP) which was developed within the DLR at an earlier date. The platform provides functions and services based on Postgres-/PostGIS databases. The TDP is complemented with systems for traffic simulation and forecasting, differentiated routing as well as a web application for visualization.

### 2.2. Features/ Information

EmerT is a web-based decision support application for real time traffic situation, prognosis and traffic simulation during exceptional situation like disasters, major events, incidents or emergencies. This application previously supported three demonstration regions in Germany for which information and decision support can be displayed: Cologne, Bonn and Munich. The portal can also be used for information exchange and cooperation between traffic management centers, emergency rescue services, police and emergency call centers. (Bieker et al., 2012), (Bonert, 2011)
One main feature of the system is the provision of traffic state estimation and forecasting by using a mesoscopic traffic simulation - “SUMO – Simulation of Urban Mobility” (Krajzewicz et al., 2012). The simulation also can be used to identify traffic bottlenecks what supports the decision maker to take appropriate traffic management actions. However, current traffic data is needed to achieve accurate simulation results. Therefore, input data can be provided by ground-based sensors like induction loops and Floating Car Data (FCD).

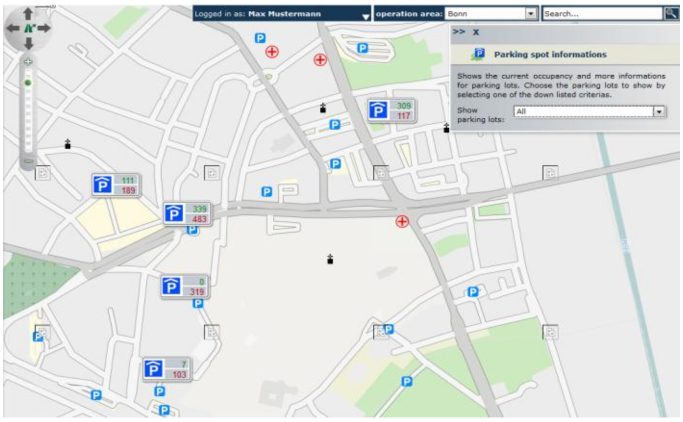



**Fig. 1.**
*EmerT-Portal – POI – parking space*
*(EmerT-Portal, 2014)*

Together with the current state of traffic, further information provided by traffic cameras and airborne monitoring can also be displayed within the portal. (Bieker et al., 2012)
The portal also provides a range of functions in terms of “routes & monitoring” such as a routing planner and isochrones about the journey time in particular for first responders. Furthermore, specific area information can be displayed. These include points of interest like fire departments, police stations, gas stations or parking areas (cp. Fig. 1). The information about parking areas can be extended towards the display of currently and general available parking lots. During a major event this helps to narrow down the operation or rescue area, to identify assembly areas but also to evaluate the dimension of the event.
Fig. 2 shows the layout of the portal which provides the mentioned information in a structured way. The portal contains a main menu where one can choose the displayed information field, it also contains a news feed, a personalized task bar, an operational area and a function box where more detailed information concerning the information field are presented (e.g. traffic state differentiated by color). Within the map a map editor is implemented which allows blocking streets and adapting speed limits temporarily. This feature allows a basis for rerouting traffic. (Bieker et al., 2012) However, all presented information are related to individual road traffic and do not comprise intermodal transport.

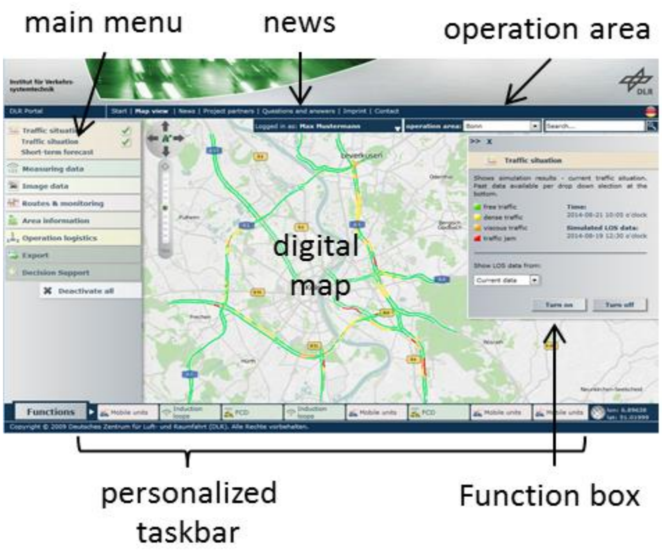



**Fig. 2.**
*EmerT-Portal - Layout*
*(EmerT-Portal, 2014)*

### 2.3. Aim of Extension of Portal

The extension of the EmerT-Portal aims at the integration of further modes of transport to emphasize the relevance of intermodal transport within a disaster or major event. Information will be aggregated and not only focused on one mode of transportation. Further, not only passenger transportation but also freight transportation needs to be considered in case of a disaster to provide rescue equipment and sustain the affected local and regional provision with goods and its economy.

## 3. Method

The basis for an extension of decision-making information displayed within the EmerT-Portal is presented by the underlying data which eventually is implemented in the TDP. An implementation of decision-making support for intermodal freight transportation requires data concerning the transportation infrastructure, the different transport modes as well as data concerning freight transportation flow. As the portal is a demonstrator for rescue forces and emergency managers, only a limited demonstration region was chosen. As shown in Fig. 3, the demonstration region of the intermodal extension covers an area within about 33 km to 83 km around the city centre of Brunswick, Lower Saxony, Germany.

Within the chosen research area, an extensive data collection through desk research, questionnaires, data requests and telephone interviews was conducted. Data sources were, amongst others, statistical databases, infrastructure and transportation operators as well as ministries like the Federal Ministry of Transport and Digital Infrastructure (BMVI). The BMVI provided the “Forecast of nationwide transport relations in Germany 2030” which is based on year 2010 (DLR Clearing House Transport, 2014). For the simulation other sources like the open source service Open Street Map (OSM), the navigation service provider Navteq, data from the traffic model Validate of the company PTV as well as FCD data and data from vehicle traffic counters like induction loops was used. Eventually, the data was collected, inspected, consolidated and integrated into the TDP.

### 3.1. Data concerning Traffic Infrastructure

To display different modes of transportation within EmerT, the transport network of the different modes needs to be integrated. In addition to the already displayed street network, the rail network, network of waterways and airports were implemented. The rail network contains e.g. the routes, the number, course and direction of the tracks, railway stations, crossings, depots, shunting yards and section block signals whereas the network of waterways implies the course, direction and dimension of the waterways, locks, harbours and bridges. Airports are only considered as origin or destination and not displayed in detail.

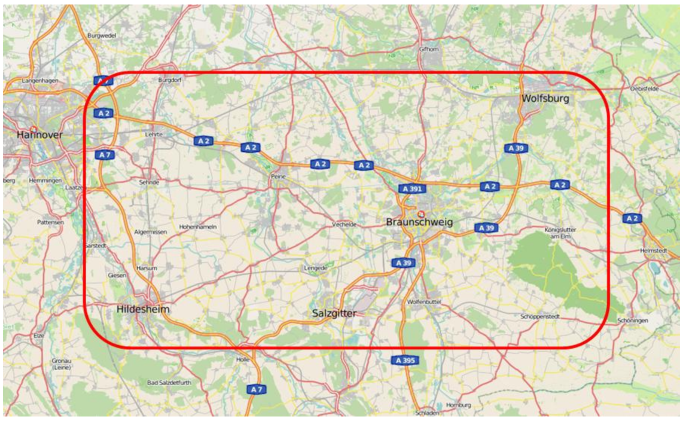



**Fig. 3.**
*Demonstration region Brunswick, Germany*
*(OpenStreetMap, 2014)*

### 3.2. Data about Transport Operation

Relevant transport operation data contain the maximum and average velocity on different routes of each network. Related to the railway system it includes the distinction between routes that are cleared for freight traffic, routes that are cleared only for passenger trains and routes that combine both kinds of traffic.

### 3.3. Data concerning Logistic Infrastructure

The implemented logistic infrastructure contains amongst others reloading sites, trans-shipment centres, freight depots, intermodal transport terminals and yards. Besides, relevant industries in the sample region of Brunswick are considered as shipping points.

### 3.4. Data of Transportation Flow

Transportation flow data is presented by the demand of transportation within each mode of transport. Therefore, origin-destination (OD) matrices are required that display what amount of a specific good is transported by which traffic carrier from location A to location B taking route AB. Hereby, a comparable granulation of the demand data between the different modes of transport is pursued which as a first approach is achieved within the mentioned forecast of the BMVI. It provides a large database of OD relations for 25 different goods over all modes of transport and corresponds to the geographical structure of NUTS (Nomenclature of Territorial Units for Statistics) 3- regions established by Eurostat (cp. Fig. 4). These OD relations were implemented to give an overview of the nationwide freight transport (DLR Clearing House Transport, 2014a), (DLR Clearing House Transport, 2014b).

### 3.5. Specific Data for the Simulation

For the simulation different and already consolidated data is required. Within the simulation database networks, infrastructure data and traffic demand data are merged.
To extend the EmerT-Portal towards the new exemplary region of Brunswick, the region's infrastructure networks have to be obtained. The Brunswick road network is a combination of several sources. The basis forms a map provided by Navteq (Version 2010/ Q2) which is complemented by road-bound traffic that was generated/ recorded within a former project in the same area. The waterways which are not an applicable network within Navteq maps are exported in form of polygons and then converted into knots and edges to create a network of waterways. The railway network is exported from OSM respectively Open Railway Map as the data density and amount of additional information like route numbers is higher.
In opposition to road traffic, trains and ships do not follow a car-following model. Therefore, new following models for the additional modes of transports were developed. In order to model a train-following scheme, signals are exported from OSM.
The basis for the road-bound traffic demand forms the traffic model Validate. The rail traffic demand is extracted from timetables provided by national train companies like Deutsche Bahn AG which are further supplemented by generated freight traffic which e.g. is extracted from the forecast of the BMVI (Deutsche Bahn, 2011).

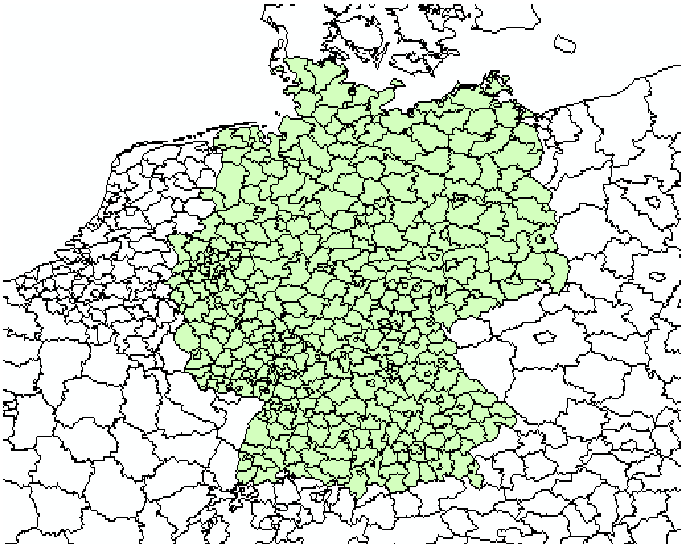


**Fig. 4.**
*NUTS 3 regions (transport cells)*
*(ArcMap, 2014)*

## 4. Extended Decision-Making Support with the EmerT-Portal

### 4.1. Features/ Information

As mentioned, the extension of the EmerT-portal aims at the integration of further modes of transport to emphasize the relevance of intermodal transport within a disaster or major event. Therefore the new extensions can be divided into two areas.

One implementation is the expansion of the information for the region Brunswick by the above presented data. As presented in Fig. 5, the implementation of various layers allow a separate but also combined display of the transportation infrastructures, logistic infrastructure, points of interests and transfer points between the different modes of transport. The selection of what information should be displayed is realized by an adapted function box with selection boxes (cp. Fig. 5). The infrastructure and logistic information are complemented by additional information about e.g. accessibility, available technology, average freight turnover and contact information. Those additional information can be helpful for rescue forces and decision makers for evacuation or the provision of goods within a region.

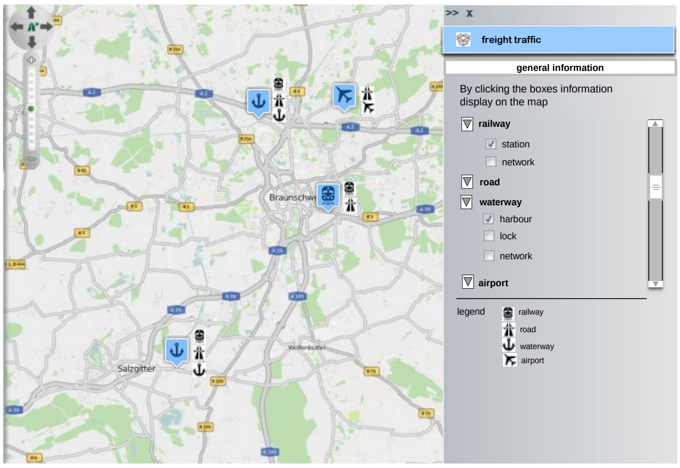



**Fig. 5.**
*Extended information on different modes of transport*
*(EmerT-Portal, 2014)*

In addition to local information for the region Brunswick, a nationwide extension towards the display of average freight traffic between NUTS 3-cells (cp. Fig. 4) is implemented. The average freight traffic can be displayed separately or combined for all modes of transport and apportioned for the different goods transported. Hereby, individual transport cells for the demonstration region Brunswick and bordering transport cells can be selected to get an overview of the regional freight transport.

The second extension of the EmerT-Portal is a demonstrator of the freight transport simulation. It allows a comparison of the daily (freight) traffic and the traffic during a major event or emergency. At a late date a generation of individual scenarios is pursued but the first approach contains a fictitious example of the Brunswick region which is presented below.

### 4.2. The Brunswick Scenario

The initial situation of the scenario relates to an accident of a truck loading chemicals on Wolfenbuettler Strasse on level of a railway bridge (cp. Fig.6). The vehicle catches fire and remains stuck below the railway bridge where it burns out completely. As a result parts of the railway bridge are destroyed. The driver can be rescued unharmed. A large area around the accident scene is blocked in a long-term due to leaking chemicals and the damaged bridge. All traffic – road, tram, railway - is interrupted. The chosen disaster relates to big ratio of technological disasters/ transport accidents and the identified challenge of limited transport capacity (IFRC, 2014), (Kovács et al., 2009).
This scenario arises two questions: What routes do rescue forces have to take to reach the location of the accident within the given period of time? How can, in a long term, traffic (passenger and freight) be managed and changed onto the other modes of transport and what impact does it have on the overall freight transport?
The Brunswick scenario is simulated before and after the emergency and the simulation results are intended to be presented within a short film which is provided at an additional function "scenario simulation" within the portal. Supplementary, the function allows general information about the scenario and decision support for decision makers.


**Fig. 6.**
*Extract of Brunswick scenario*
*(EmerT-Portal, 2014)*

### 5. Results and Discussion

Since the relevance of intermodal transport decisions increases, especially during disasters or major events when road traffic is limited or shut down, the extended EmerT-Portal provides emergency forces with necessary infrastructure and logistics data for road bound modes of transport.
However, the provided information within the portal have to be considered carefully as they only present an estimation of the transportation demand. This results from the limited existence of discretionary logistics and transportation data. A cooperation with transport (infrastructure) operators as well as logistics service providers is restrained due to data protection and business competition issues. Therefore, the obtained data concerning transport demand on the different modes of transport vary high in granularity and precision.
The transport flow on waterways so far is an estimation that bases on the amount of ships per lock per day. As no detailed distribution over the course of the day is available, an equal distribution with consideration of the lock closure times is assumed. Similarly, the transport flow for rail services in total is estimated by adding an assumed demand of rail freight traffic to the extracted passenger rail service traffic from available timetables. On top an overall freight demand of Germany is narrowed to a freight demand within the region of Brunswick. This combined estimation of transportation flow is likely to be deficient. To receive a realistic, applicable view on intermodal transportation and routing the precision and completeness of data is essential and the granularity should be comparable between the different transport modes.
Further, the applied concepts for modeling freight transport allow a demonstration of an entire transport chain including the handling of goods. Even so, handling processes are modelled at a basic level and detailed handling information like handling times are estimated. Again, a more complex, realistic modelling is necessary for a future application of the information portal, same as the transferability of the portal's information and features towards a different region. Though, the extension of the portal presents an important step towards a more holistic traffic and transportation overview, it is limited to the exemplary region of Brunswick. Moreover, the practical applicability is restricted as an expansion towards other regions is not feasible.

This results mainly from missing transportation flow data which presents a regions demand of transport and at the same time the pressure on the transportation infrastructure which displays a basis for decision-support and routing recommendations. The available transportation flow overview of Germany allows only a limited application.
In the end, the extension of the EmerT-portal by integrating information concerning all ground-based modes of transport as well as logistics and freight transportation data is still on the demonstration level.
Nevertheless, even at this stage it can be seen as a major achievement for rescue forces especially in ad-hoc disaster situations entailing extensive evacuations. In this case, an instantaneous knowledge about transport capacities or the volume of the flows of goods helps to assess the overall impact on the mobility of an entire region. It enables better evacuation routes by utilizing passenger and freight transport and balances the street network load by contemplating freight transport as well as individual evacuation traffic.

## 6. Future Research/ Outlook

Aiming at the practical application of the EmerT-portal during rescue operations, the limitation on one exemplary region should be overcome and an extension towards other regions and eventually the whole country should be implemented. Therefore, overall infrastructure and logistics data is required as well as data about the transportation flow/ demand. An extension towards other regions requires a higher availability of infrastructure and transportation data. Today a first basis is presented by an overall German transportation flow which, for a practical need, has to be narrowed to regional and local level.
Further steps will be the implementation of freight transport-related recommendations for actions, respective an intermodal routing recommendation. This would allow emergency forces to ship humanitarian goods and equipment faster and unerring to where it is required and at the same time secures the efficient transportation of goods into an affected region to keep up the provision of the people. However, the implementation of intermodal routing recommendations demands more realistic underlying freight transportation and handling models.
Further interest focuses on the creation and evaluation of complex scenarios for training purposes as well as on the integration of passenger transport respective public transport into the portal to eventually receive a holistic view of an emergency in terms of the traffic situation. A big challenge is the development of interfaces and operational applicability of the EmerT-portal in terms of the technical link to existing information systems for emergency forces. Therefore, in the future the EmerT-Portal will be enhanced towards a data portal to establish interoperability with existing control and command centre systems. Interfaces will follow Open Geospatial Consortium (OGC) -standards or widely accepted conventions like Tile Map Service (TMS) or Open Source Geospatial Foundation (OSGeo). The development will further focus on the extension and optimization of the provided data and services.

**References**

Bieker, L.; Behrisch, M.; Ruppe, S. 2012. EmerT – a web based decision support tool for Traffic Management. In *Proceedings of the 19th ITS World Congress, Vienna, Austria.*

Bonert, M. 2011. EmerT – Emergency Mobility of Rescue Forces and Regular Traffic. Handout

Deutsche Bahn. VDV (Eds.). 2011. *Eisenbahnatlas Deutschland*. 8th ed. Cologne. Germany. Schweers + Wall

DLR. 2014. VABENE++ Traffic Management for Large Scale Events and Disasters. Available from Internet: http://www.dlr.de/vabene/en/, last accessed: August 21st, 2014.

DLR Clearing House Transport. 2014a. Forecast of nationwide transport relations in Germany 2030. Available from Internet: http://daten.clearingstelle-verkehr.de/276/, last accessed: August 22nd, 2014.

DLR Clearing House Transport. 2014b. Verkehrsprognose 2030 - Kurzbeschreibung der Güterverkehrsmatrizen. Available from Internet: http://daten.clearingstelle-verkehr.de/276/3/Datensatzbeschreibung%20G%C3%BCterverkehr.pdf, last accessed: August 22nd, 2014.

EmerT Portal. http://vabene.dlr.de/Emert/EmerT_Login.html, last accessed: August 21st, 2014.

DLR. 2014. VABENE++ Traffic Management for Large Scale Events and Disasters. Available from Internet: http://www.dlr.de/vabene/en/, last accessed: August 21st, 2014.

IFRC. 2014. World Disaster Report – 2012 disaster data. Available from Internet: http://worlddisastersreport.org/en/data/index.html, last accessed: September 14th, 2014.

Kovács, G.; Spens, K. 2009. Identifying challenges in humanitarian logistics. *International Journal of Physical Distribution & Logistics Management*. 39 (6): 506-528.

Krajzewicz, D.; Erdmann,J.; Behrisch, M.; Bieker, L. 2012. Recent Development and Applications of SUMO - Simulation of Urban Mobility. *International Journal On Advances in Systems and Measurements*. 5 (3,4): 128-138.

OpenStreetMap. 2014. Available from Internet: http://www.openstreetmap.de, last accessed: August 21st, 2014.

Touko Tcheumadjeu, L.C.; Ruppe, S.; Ali, M.; Bieker, L. 2012. EmerT – Supporting traffic parameter estimation from low cost and low resolution uncalibrated web cameras, In *Proceedings of the 19th ITS World Congress, Vienna, Austria.*

# ADVANCED LOGISTICS FOR SUSTAINABLE URBAN AREAS

**Meng Lu**[1]
[1]*Dutch Institute for Advanced Logistics, The Netherlands*

Abstract: An overarching concept for logistics and mobility may significantly contribute to a sustainable intelligent transport system in such a way that other sectors and the European economy as a whole will benefit. This requires an integrated view on a wide range of topics such as passenger mobility, the supply chains and flows of goods, related passenger, freight and traffic flows, land use planning, policy development, technological development, and economic, legal, and financial issues. Moreover, proper understanding of the behaviour and policies of individuals, companies and governments is needed. Urban logistics is defined as all movements of goods (as distinct from people) in to, out from, through or within the urban area made by light or heavy vehicles, including also service transport and demolition traffic, shopping trips made by private households and waste (reverse logistics). The paper presents an innovative solution for creating a reliable, robust and sustainable system for mobility and logistics in urban areas by taking a holistic view. The movement of people and goods share (information and physical) infrastructure. The paper presents some of the good practices of ITS applications in the field of urban freight transport in several European logistics clusters. In addition, it addresses the needs for cooperation between logistics clusters to create a systemic improvement in urban logistics in Europe for sustainable urban areas, to better establish the contribution of urban logistics to the economy and to determine the role of the urban authorities.

**Keywords**: logistics, business model, Intelligent Transport Systems, smart city, policy

## 1. Introduction

There is an on-going migration trend from rural to urban areas, leading to higher population densities and growing economic activity in those locations. Within the OECD countries, this evolution is very clear: in 1950, 50% of the population lived in cities, 77% in 2000 and it is expected that by 2020, this will rise to 85% [OCED, 2003]. Urban transport infrastructure has to accommodate ever larger numbers of commuters and commercial traffic volumes. In addition to the increase in the absolute number of consumers, transformations in the retail landscape (more chain retailers and shopping centres, less independent local stores) as well as new business models and technological innovation related to transportation and e-commerce (e.g. mail order outlets) reshapes or increase the demand for urban logistics solutions. Metropolitan areas are endowed with often global transportation hubs enjoying constant growth due to globalisation trends. This puts additional strain on the local infrastructure due to their incoming, outgoing, and transit passenger and cargo flows. These trends towards ever larger and volatile traffic volumes lead to situations in which the capacity of the local transport infrastructure is exceeded, especially during peak hours. Negative effects of transport activities on the quality of life in urban areas such as congestion, traffic safety, air pollution and noise are increasing public resistance to additional invasion. Changed priorities in regard to infrastructure development and environmental protection often further tighten the requirements raised towards transportation activities. In order to sustain an efficient functioning of the economic system and preserve quality of life in metropolitan areas, new solutions for the movement of cargo and passengers in the urban context have to be found. [SoCool@EU, 2014]

Sustainable transport system (interns of green, efficient, safe and secure) is a must for the common European market for enabling the free movement of goods and people. In this system, new transport solutions for reducing negative impact on residents and the environment, new mechanisms and ICT-based ITS (Intelligent Transport Systems) solutions for execution and control of urban logistics need to be considered. The paper proposes a holistic to deal with the urban freight transport. It takes all shipment, actors, e.g. shippers, receiver, logistics service providers (LSPs), local authorities, people and vehicles, as components of an integrated logistics system.

The paper is structured as follows: In Section 2 urban freight transport will be discussed in-depth and lessons learned will be summarised. Section 3 proposes a novel ITS solution for urban logistics. Policy initiatives for establishing smart city will be addressed in Section 4. Some recommendations will be provided in Section 5, and finally Section 6 draws conclusions.

## 2. Urban Freight Transport: Lessons Learned

Urban freight transport is defined as all movements of goods (as distinct from people) in to, out from, through or within the urban area made by light or heavy vehicles, including also service transport and demolition traffic, shopping trips made by private households and waste (reverse logistics) [Lindholm, 2012]. The OECD (Organisation for Economic Co-operation and Development) Working Group on Urban Freight Logistics [OECD, 2003] defines urban goods transport as the delivery of goods in urban areas, including the reverse flow of waste.

[1]Corresponding author: lu@dinalog.nl

Urban freight transport is fundamental to the economic vitality [Allen et al., 2000; Munuzuri et al., 2005]. Urban freight transport includes (1) all types and sizes of goods vehicles and other motorised vehicles used for (core) goods collections and deliveries at premises in the urban area; (2) all types of goods vehicle movements to and from urban premises including goods transfers between premises, ancillary goods deliveries to urban premises, money collections and deliveries, waste collections and home deliveries made from urban premises to customers; and (3) service vehicle trips and other vehicle trips for commercial purposes which are essential to the functioning of urban premises. [Allen et al., 2000]

Compared with transport of people (through integrated public transport networks financed by the authorities), freight transport is very different and rather immature. Freight transport almost completely managed by private Logistics Service Providers (LSPs) and/or shippers. Inefficient movements of underutilised vehicles (e.g. vehicles running with a low load factor or empty) have negative impact on urban traffic. Novel logistic solutions are required to better manage the flow of goods into, and out of, urban areas.

Urban logistics practices seem to be dominated by failing (and often subsidised) initiatives and typically concern very local approaches [Quak, 2008]. Local and regional governments impose strict rules on freight vehicle access to cities in order to address environmental and social concerns. However, many of the measures significantly increase transport costs and the complexity of the urban freight transport planning and emission. Local regulations, such as time windows, are often not harmonized between cities, resulting in vehicle utilization problems, inefficient transport operations, increased emissions and significant additional costs for carriers and shippers [Quak & Koster 2009].

Best practices of urban freight transport can be summaried as follows:

- Consolidation of distribution. Consolidation refers to combining different loads and carriers onto the same vehicles [Benjelloun, Crainic & Bigras, 2009].
- Triple-helix cooperation. This reflects the cooperation between authorities, industry and academia in each logistics cluster.
- Entrepreneurships. Companies, especially SMEs, may have good opportunity to create new business.
- Horizontal collaboration. Appropriate business models and the (economic, environmental and social) impacts are based on substantially increased horizontal collaboration between shippers on one side, and between LSPs on the other side.

Previous research shows that better coordination and consolidation of the urban freight distribution resulting in fewer vehicles, better utilization of these vehicles and less emissions [Taniguchi, Yamada & Tamagawa, 1999]. In order to achieve these goals, new organizational models for the management of freight movements within the urban area need to be developed. Through the use of coordination and consolidation by the logistics industry, a more efficient (both in terms of costs and environmental impacts) urban transportation system will be achieved. Note that the logistics industry includes:

- Shippers, such as producer (raw material suppliers, base industry, manufacturing industry), trading companies, and retail companies
- Logistics service providers
- Transport companies
- Terminal operators
- Physical infrastructure providers (including ports) and operators
- Support industry (finance, law, ICT, equipment/vehicle/vessel manufacturers, infrastructure construction companies)

From the experiences concerning issues in urban freight logistics in some European logistics clusters (The Netherlands, Belgium, Germany, Sweden, Denmark and Turkey), the following lessons can be learned:

- It is difficult to obtain and manage in a timely and effective way up-to-date data describing the state of an urban traffic network. Due to unavailability of data, before-and after-evaluation cannot be carried out. Main reasons are: 1) vehicles travelling through the network belong to different carriers that provide service to different suppliers and retailers. Carriers, retailers and suppliers are reluctant to share any (demand, capacity, prices) data; 2) detailed information of the network is not available to all actors (e.g. infrastructure, road congestion, planned road construction activities); and 3) lack of identification and proper measurement of the relevant performance measures.
- Lack of awareness, knowledge and innovative policies. Local authorities often focus on the infrastructure of urban mobility for people. Government investment is mainly focusing on public transport, while investment in goods transport is simply left to the private sector. This might be due to the fact that logistics is a very fragmented sector, in which a broad variety of companies and associations play a role. There is no clear and

univocal voice in the area of research and innovation on logistics, while actually there is a strong need for this. Traditionally multidisciplinary efforts were not sufficiently put concerning the planning of energy, water, transport, logistics and communication facilities for public health and safety in a systematic way.

- Insufficient collaboration and appropriate business models. Besides aforementioned issue of data availability, lack of transparency and trust are essential obstacles. Transparency and trust are considered as basis for collaboration. All the companies in a partnership would work as a team and that they are together for the benefit of each other in the concept of mutuality and solidarity. In addition, companies in a collaboration needed a similar culture, similar business objectives and a desire to make collaboration work. The increasing competitive environment will force companies to find efficiency improvements and the main drivers are cost reduction, customer demands and the need to maintain or improve service levels. One of the main barriers to collaboration was finding the right partners.

## 3. ITS for Further Improving Urban Logistics

To improve the consolidation of distribution, and the coordination between the different industry actors (such as shippers and LSPs) based on sound business models, while meeting the objectives and restrictions set by municipalities, both technical and non-technical solutions are needed.

Today ITS (Intelligent Transport Systems) is quickly progressing globally. However, deployment of ITS in the domain of freight transport and logistics is rather limited, and insufficient attention has been given by authorities to this area, compared with the domain of mobility for passenger transport. ITS applications will substantially contribute to sustainable city distribution. The use of ICT (Information and Communication Technologies) can improve infrastructure and develop more efficient ways to provide services to citizens and businesses.

Overhauling infrastructure and services of a smart distribution requires a lot of technologies, such as telecommunications, wireless networks, a smart grid, sensors, facial-recognition systems, integrated transport and logistics system, crowdsourcing and data aggregation. The potential drivers for sustainable freight transport and logistics directly relate to the needs of the logistics industry. These are, inter alia, accessibility, accountability, affordability, availability, competitiveness, reliability, seamlessness, security and safety, and transparency.

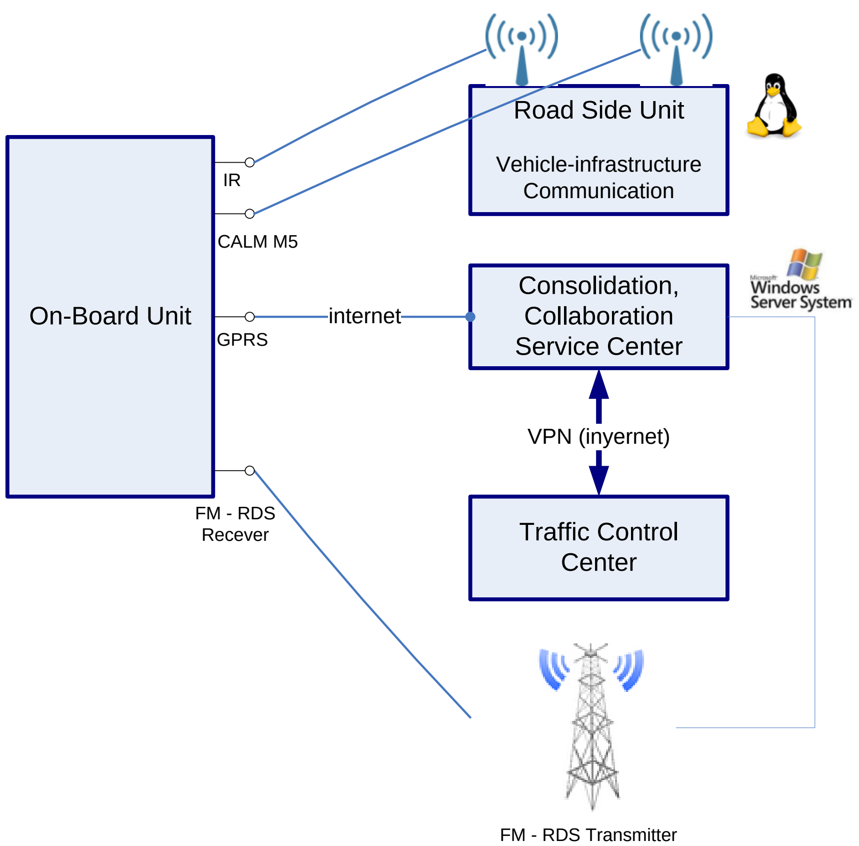



**Fig. 1.**
*Illustration of the service system architecture for city distribution through consolidation, collaboration and coordination*
*Source: own*

The novel concept proposed in this paper is to collaborate, consolidate and coordinate (cross-chain control) in supply chain and freight operations, in order to create a systematic approach for the improvement of urban communities, including, e.g.

- improve planning through demand control (at strategical, tactical and operational level through e.g. destressing the supply chain and pre-trip and on-trip planning)
- establish automating payment systems for services (such as public transport)
- improve return and waste management, and dispose of waste efficiently
- alert first responders in case of emergency
- offer better product availability and service level
- increase vehicle load factor
- reduce empty running
- reduce pollution and noise
- alleviate congestion
- ameliorate intra-modal conflicts in urban areas
- free up space at stores

From a technical perspective, this can be realised through a highly integrated logistics service system architecture (see Figure 1). The Service Centre is connected with in-vehicle systems (On Board Unit - OBU) through GPRS (General Packet Radio Service) via internet, and in addition with the Traffic Control Centre. The concept of Cooperative ITS needs to be introduced. Infrastructure based vehicle-vehicle and vehicle infrastructure communications will be applied. The OBU has various interfaces such as FM-RDS (Frequency Modulation - Radio Data System), GPRS, CALM M5 (Continuous Air interface for Long and Medium distance - Microwave 5 GHz) and IR (infrared).

An appropriate business model is one of the key elements for improving city distribution through high-level consolidation and collaboration. Tangible and substantial benefits for the industry need to be demonstrated (as an incentive for a mind change). The application of the proposed system requires a high degree of horizontal collaboration between shippers, LSPs and transport operators.

## 4. Policy Initiatives for Smart City

Besides the abovementioned technical innovation, policy innovation for urban freight transport is one of the key facilitators, as public bodies (city, regional, national, international) also play an important role for supporting innovation. Main policy initiatives that are implemented are: for instance, vehicle restriction, time-windows, road pricing, licensing and regulation, parking and unloading, carrier cooperation, (company driven) vehicle routing improvement, technological vehicle innovation, consolidation center, road infrastructure development, standardisation of load-units, transport reorganizing, transport auction and intermodal transport. However, these are not always successful, and do not have social, environmental and economic impacts of sufficient size.

There is a strong need for cooperation between logistics clusters to create a systemic improvement in city logistics in Europe for sustainable urban areas, to better establish the contribution of city distribution to the economy. Logistics clusters (defined as triple helix: industry, authorities and education & research institutes) are key entities for constituting crucial research and innovation initiatives for providing significant contributions to a sustainable intelligent transport system. A better cooperation between retailers, logistics service providers, research institutes and authorities will help to improve city distribution. This has been amply demonstrated in The Netherlands, facilitated through a national logistics foundation (funded by the Dutch government). It is expected that similar logistics clusters will collaborate and operate by, for instance

- exchanging knowledge and experience between clusters;
- facilitating retailers and logistics service providers (especially SMEs) for enhancing and improving their businesses, and/or generating new businesses
- transforming and converting research results to real world applications

Such a comprehensive system will stimulate the process to realise sustainable urban freight transport, and to rapidly improve urban logistics in a systematic way.

Smart urban freight transport will further contribute to sustainable smart city initiatives, which require multidisciplinary efforts concerning the planning of energy, water, transport, logistics and communication facilities for public health and safety. In addition, a better use of available infrastructure or a reduction of infrastructure needed offers important societal benefits, first and foremost from an environmental point of view, but also by using energy more efficiently. Reducing flows through improved collaboration and consolidation, e.g. by means of city distribution centres followed by last-mile distribution by standardized electric vehicles, represents an enormous yet achievable challenge. At the same time, governments can speed up implementation by appropriate regulation or by providing smart incentives that stimulate the transport sector and logistics service providers to behave in a socially desirable direction. Government may also influence decisions to locate or relocate businesses by offering attractive yet environmentally responsible opportunities, again by providing proper incentives.

## 5. Recommendations

In the past decades, master plans have been proposed and implemented in different countries [OECD, 2003]. For instance, in Japan, quantitative targets were set as indicators of the achievement of the policies objectives (such as loading rate of all trucks to 50% during the beginning of the 21st century). In The Netherlands, the Platform for Urban Distribution (Platform Stedelijke Distributie) was initiated by the national government. It harmonises municipal legislation and promotes best practices on a national level, based on local experiences. In the United Kingdom, the national government has affirmed its commitment to bringing about more sustainable urban distribution operations.

In the past decades, a vast amount of research (including research done in national and European projects) has been carried out in the area of urban freight transport. It seems that sustainable logistics becomes a hot topic for chatting, especially by people in the social science area and by politicians. Researchers are continuously seeking for national and European funding for (follow-up) research and pilots, although barriers have been analysed and best practices have been collected.

The term "sustainable" has been used widely and frequently in the policy framework. There are trade-offs between the three dimensions of sustainability. It has been realised that urban freight transport policy requires consultative planning (PPP), and policies should be formulated so as to enhance developments in the private sector (see Figure 2). However, there are still no clear answers to the following basic questions: 1) What needs to be done to create a systemic improvement in urban logistics? 2) How can we better establish the contribution of urban logistics to the economy? 3) What should urban authorities do?

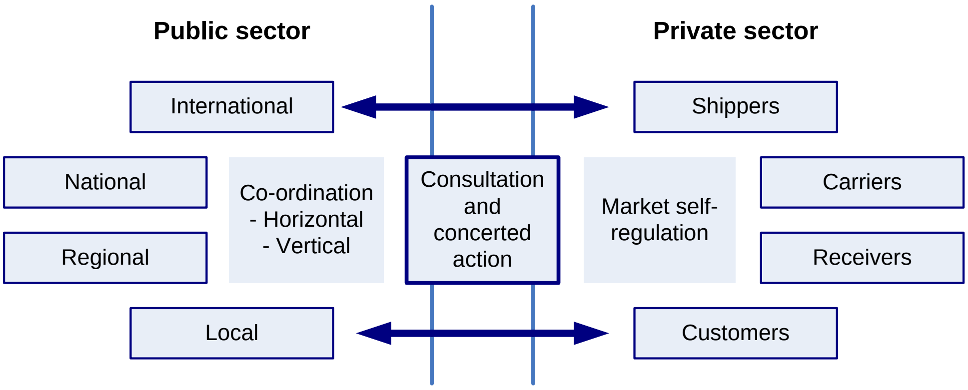



**Fig. 2.**
*Consultation position related to governmental co-ordination and market self-regulation*
*Source: [Van Binsbergen & Visser, 2001]*

Recommendations on solutions dealing with challenges of urban freight transport are (extended based on [OCED, 2003]):

- Collaboration, consolidation and coordination are the essential keys for achieving sustainable urban freight transport.
- To increase awareness of the importance of urban freight transport is a must
- To exchange knowledge and experience through logistics clusters is an efficient approach for disseminating best practices and for creating a systemic improvement in urban logistics. International co-operation is needed.
- Regulations need to be harmonised, standardised, stable, easy to be enforced and cost-effective.
- Infrastructure capacity should be used more imaginatively and adequate logistic facilities need to be provided
- Information and communication technologies can substantially support sustainable urban freight transport, for which the knowledge and experience in the ITS area is very valuable.
- Cleaner, low noise and more energy-efficient vehicles are optimised alternatives for urban transport.
- Safety and security of goods transport should be emphasised.
- Reverse logistics needs to be further developed

## 6. Conclusions

The key elements for sustainable freight transport and logistics can be categorised as follows: infrastructure (physical, information, organisational and finance), (new) business models, and education and training. Logistics is not only a matter of how to transport, but also concerns developing a (supply chain) view on what to transport. There is a strong need for Intelligent Transport Systems (ITS) development in the domain of freight transport and logistics, and especially for the implementation of ITS. The proposed concept and ITS solution for urban freight transport will also help governments to improve their policy making concerning the improvement of accessibility of the urban areas, reduction of congestion and pollution caused by urban freight distribution. Collaboration, consolidation and coordination are the essential keys for urban freight transport. Logistics clusters can play an important role for exchanging knowledge and experience. This,the improvement of urban logistics can be accelerated. Policy initiatives are the important elements for realising sustainable urban freight transport.

## Acknowledgments

The paper presents part of the results of a EU-funded project SoCool@EU (Sustainable Organisation between Clusters Of Optimised Logistics @ Europe), DG RTD (Research and Innovation), Regions of Knowledge Programme (FP7-REGIONS-2011-1). The author especially thanks the consortium partners, and participating third parties for their valuable suggestions and support.

**References**

Allen, J., S. Anderson, M. Browne, Jones, P. 2000. *A framework for considering policies to encourage sustainable urban freight traffic and goods/service flows*. Transport Studies Group, University of Westminster, London.

Benjelloun A., T.G. Crainic, Bigras, Y. 2009. *Toward a taxonomy of city logistics projects*. Procedia - Social and Behavioral Sciences 2(3), 6217-6228, Elsevier.

Van Binsbergen, A., Visser, J. 2001. *Innovation steps towards efficient goods distribution - Systems for urban areas*. TRAIL Thesis Series No. T2001/5, Delft.

Lindholm, M.E. 2012. *Enabling sustainable development of urban freight from a local authority perspective*. PhD thesis, Chalmers University of Technology, Göteborg.

Munuzuri, J., J. Larraneta, L. Onieva, Cortes, P. 2005. Solutions applicable by local administrations for urban logistics improvement. Cities, 22 (1), 15-28.

OECD. 2003. *Delivering the goods - 21st century challenges to urban goods transport*. OECD Working Group on Urban Freight Logistics, Paris.

Quak, H. 2008. *Sustainability of urban freight transport retail distribution and local regulations in cities*. PhD thesis, Erasmus University Rotterdam, Rotterdam.

Quak, H.J., De Koster, R. 2009. Delivering goods in urban areas: How to deal with urban policy restrictions and the environment? *Transportation Science*, 43(2), 211-227.

SoCool@EU Consortium. 2011. *Description of Work*, SoCool@EU (Sustainable Organisation between Clusters Of Optimised Logistics @ Europe), SoCool@EU Consortium. (Restricted)

SoCool@EU Consortium. 2012. *Deliverable D2.1: Cluster analysis report*, SoCool@EU (Sustainable Organisation between Clusters Of Optimised Logistics @ Europe), SoCool@EU Consortium.

SoCool@EU Consortium. 2014. *Deliverable D4.1: Implementation of joint action plan*, SoCool@EU (Sustainable Organisation between Clusters Of Optimised Logistics @ Europe), SoCool@EU Consortium.

Taniguchi, E., Yamada, T., Tamagawa, D. 1999. *Modeling advanced routing and scheduling of urban pickup/delivery trucks*. World Congress on Intelligent Transport Systems, Toronto.

# SYNCHROMODALITY FOR ENABLING SMART TRANSPORT HUBS

**Meng Lu**[1]
[1] *Dutch Institute for Advanced Logistics, The Netherlands*

**Abstract:** Synchromodality is a concept that takes a holistic view of (freight) transport, including and integrating all available modes, new logistics and transport concepts, facilitating infrastructures, (ICT) technologies, services, new policies, and governance. The basic idea is the use of alternative transport modes in a flexible way, depending on temporary circumstances as well as product and supply chain characteristics. Synchromodality requires a superb information infrastructure and data exchange between all parties involved (e.g. shippers, main ports, inland hubs, smart corridor operators, customs, and inspection authorities). The objective consists in the investigation of possibilities to use the existing transport infrastructure and resources more efficiently by developing sophisticated synchromodal transit chains through improved information availability and exchange and better planning and scheduling. The objective implies a holistic analysis and step forward to the application of the synchromodal concept, including the investigation of the societal costs and benefits of a synchromodal view on (freight) transport, the identification of the most important barriers for implementation and transition, and the development of solutions to overcome the indicated barriers. The paper will address the following issues: 1) in-depth analysis of implementation lessons (positive and negative) from intermodal and co-modal studies and pilots (identifying most important implementation and transition); 2) development of strategies for a synchromodal transport system with associated new logistics/service concepts, ICT systems, infrastructures (terminals) and possible vehicle concepts.

**Keywords**: logistics, supply chain, synchromodal, multimodal, co-modal

## 1. Introduction

Synchromodality is a concept that takes a holistic view of (freight) transport, including and integrating all available modes, new logistics and transport concepts, facilitating infrastructures, (ICT) technologies, services, new policies, and governance. Synchromodal transport was first proposed in The Netherlands [TNO, 2010; ECT, 2011; EVO 2011].

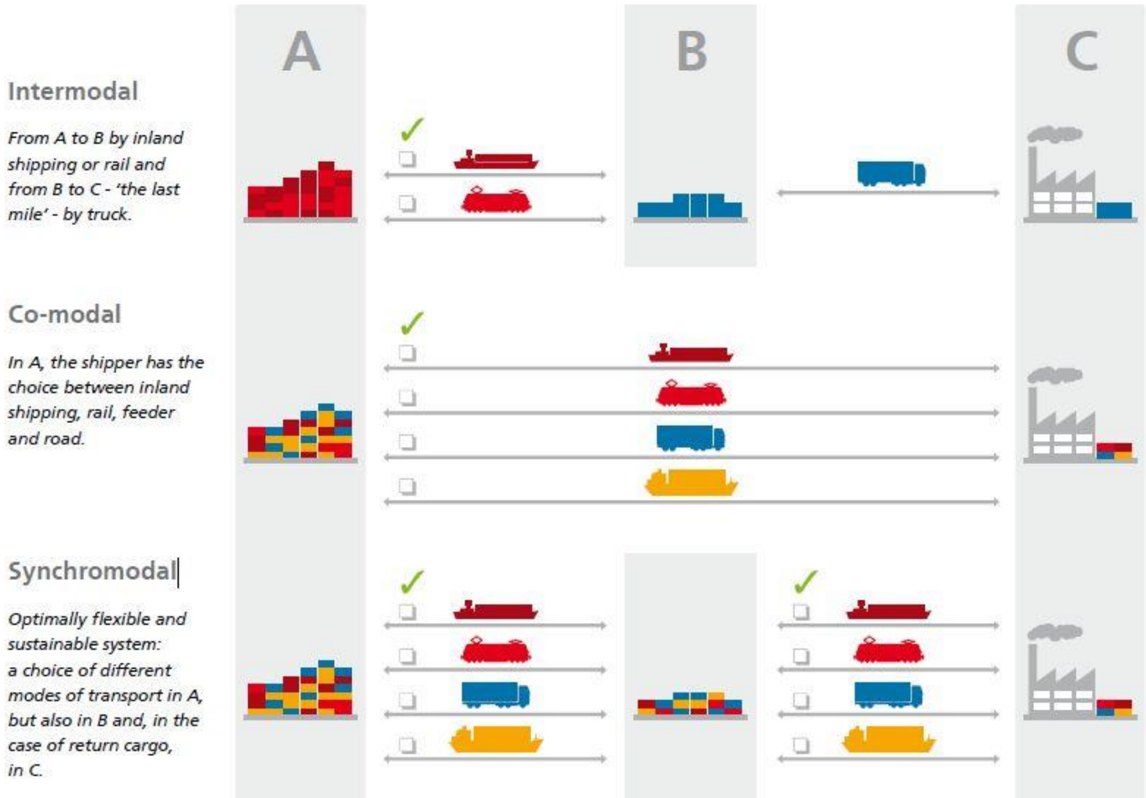



**Fig. 1.**
*Illustration of intermodal, co-modal and synchromodal*
*Source: (ECT, 2011)*

The basic idea is the use of alternative transport modes in a flexible way, depending on both temporary circumstances and product and supply chain characteristics. Synchromodality requires a superb information infrastructure and data exchange between all parties involved (e.g. shippers, main ports, inland hubs, smart corridor operators, customs, and inspection authorities).

[1] Corresponding author: lu@dinalog.nl

The objective consists in the investigation of possibilities to use the existing transport infrastructure and resources more efficiently by developing sophisticated synchromodal transit chains through improved information availability and exchange, and better planning and scheduling. The objective implies a holistic analysis and a step forward towards the application of the synchromodal concept, including the investigation of the societal costs and benefits of a synchromodal view on (freight) transport, the identification of the most important barriers for implementation and transition, and the development of solutions to overcome the indicated barriers. The paper will address the following issues:

1) in-depth analysis of implementation lessons (positive and negative) from intermodal and co-modal studies and pilots (identifying most important implementation and transition);
2) development of strategies for a synchromodal transport system with associated new logistics/service concepts, ICT systems, infrastructures (terminals) and possible vehicle concepts.

In Section 2 the main on-going national and European projects related to synchromodal transport are reviewed. Section 3 presents the status of studies and pilots on synchromodality, including lessons learned and challenges. Section 4 proposes a roadmap for the implementation of and transition towards a synchromodal transport system at the European level, including strategies and associated concepts, such as related to logistics/services, ICT, infrastructure and vehicles. Finally, conclusions are drawn in Section 5.

## 2. Development and Implementation of Synchromodal Transport

This section provides an overview of the development and implementation of synchromodality. The selected Dutch national projects, co-funded by Dinalog, are summarized in Table 1. And Table 2 presents related (on-going) EU-funded projects and initiatives.

**Table 1**
*Overview of Dutch national projects related to synchromodality*

| No. | Description of selected national projects |
|---|---|
| 1 | *Ultimate - Efficient multimodal hinterland networks* (R&D)<br>It develops an Extended Gate principle for inland transport, which will be managed efficiently and with a minimum of paper documentation. The seamless flow of goods from seaports to inland locations will be a major enabler to prevent negative external effects from interrupting the transport, such as congestion in seaports and on motorways, and to enhance the competitiveness of multimodal inland nodes for warehousing and value added activities.<br>www.dinalog.nl/en/projects/r_d_projects/ultimate__efficient_multimodal/ |
| 2 | *Cargo Driven Intermodal Transportation* (R&D)<br>It develops new concepts and facilities to add value to cargo with minimum costs and a minimum use of land and road infrastructure, and with minimum emissions through: managing cross-dock operations at the cargo level; managing container flows at the cargo level; and creating value with information flows at the cargo level. Smart and competitive intermodal transport and logistic solutions are essential.<br>www.dinalog.nl/en/projects/r_d_projects/cargo_driven_intermodal_transportation/ |
| 3 | *Sense and Response* (demonstration)<br>To transport high-tech and pharmaceutical products, the ability to avoid or quickly respond to disruptions is critical. The project develops and demonstrates innovative concepts and value-added services related to supply chain control in the global distribution of high-sensitive products, the underlining techniques and tools for real-time monitoring that provide the visibility required for joint supply chain control and sense and response mechanisms, as well as the business value.<br>www.dinalog.nl/en/projects/demo_projects/sense_and_response/ |
| 4 | *DaVinc3i - Dutch Agricultural Virtualized International Network with Consolidation, Coordination, Collaboration and Information availability* (R&D)<br>It develops innovative concepts for information architectures, co-modality, coordination and cooperation in logistics and commerce. This requires cooperation and harmonisation within the ornamental plant cultivation industry, so that with the correct information, fragile flowers and plants arrive at the right place at the right time. It contains four main research elements: quality controlled logistics, transport management, supply chain information and business models.<br>www.dinalog.nl/en/projects/r_d_projects/davinc3i/ |
| 5 | *SALOMO - Situational Awareness for LOgistic Multimodal Operations in container supply chains and networks* (R&D)<br>Critical hubs in our infrastructure struggle to efficiently plan an optimal use of resources and an efficient and flexible throughput of cargo. The project aims to empower the terminals and their inland connections by facilitating better decisions and planning through increased situational awareness, as well as better trained staff able to deal with dynamic circumstances. Modern supply chains become longer and more complex due to increasing globalisation, labor off shoring, and growth.<br>www.dinalog.nl/en/projects/r_d_projects/salomo/ |

| | |
|---|---|
| 6 | *SIEGG - Secure Information Exchange Extended Gate* (Demonstration)<br>It demonstrates a new ICT based method, called the SIEEG functionality, to increase logistics efficiency. The SIEEG functionality enables logistical agents to plan, manage and synchronize available capacity (people, modalities and space). In addition, it benefits the efficiency of the logistical process and the transnational transportation of goods, specially to/from areas with heightened security standards.<br>www.dinalog.nl/en/projects/demo_projects/sieeg/ |
| 7 | SPoT (Demonstration)<br>It aims to reinforce the hubs in control of the logistic chain and to enable synchro modal transport. This helps to support an efficient, sustainable and reliable handling of the increasing flow of goods from mainports of Antwerp and Rotterdam which are transported via strategic located inland terminals to the European hinterland. Thereby, the current Dutch capacity of various modalities and current infrastructure will be optimized, the transport service level through The Netherlands is increased and pricing is more attractive compared to competition.<br>www.dinalog.nl/en/projects/demo_projects/spot/?highlight=SPoT+ |

*Source: Own development (based on information received from the web sites)*

**Table 2**

*Overview of EU-funded projects related to synchromodal transport topics*

| No. | Description of selected EU-funded projects |
|---|---|
| 1 | *SoCool@EU - Sustainable Organisation between Clusters Of Optimised Logistics @ Europe*<br>It aims to create an open European platform of excellence in the area of logistics and supply chain management in connection with hubs and gateways. It will enable research-driven regional clusters to collaborate and exchange experiences for increasing sustainability and competitiveness of logistical services and synchromodal transport operations.<br>http://www.socool-logistics.eu/ |
| 2 | *WINN - European Platform Driving KNOWledge to INNovations in Freight Logistics*<br>It aims to establish a broad collaboration framework for an open network that will be accessible and benefit for all logistic clusters, networks and companies operating at European level that will constitute the European Technology Platform (ETP) on Logistics. The new platform, which was given the name Alliance for Logistics Innovation through Collaboration in Europe (ALICE), was launched in June 2013.<br>http://www.winn-project.eu/ |
| 3 | CO3 - Collaboration Concepts for Co-modality<br>It aims to establish a business strategy enabling companies throughout the supply chain to set up and maintain initiatives to manage and optimise their logistics and transport operations by increasing load factors, reducing empty movements and stimulating co-modality (through horizontal collaboration between industry partners).<br>http://www.co3-project.eu/ |
| 4 | *CASSANDRA - Common Assessment and analysis of risk in global supply chains*<br>It aims to improve supply chain visibility and business execution as well as the efficiency and effectiveness for government supervision. CASSANDRA will develop a data-sharing concept, which can be visualized as a data pipeline that connects entities and gathers and distributes data according to predefined conditions.<br>http://www.cassandra-project.eu/articles/data-sharing.html |

*Source: Own development (based on information received from the projects that the author involved and the web sites)*

## 3. Lessons Learned and Challenges

An increasing share of global transport is containerised, reducing handling costs and facilitating the use of multiple modes of transportation. The flexibility needed to be independent of a specific mode of transport substantially increases the coordination and planning issues along the transit chain. Challenges for academia, private companies and governments are to use the existing infrastructure and resources more efficiently by developing sophisticated synchromodal transit chains through better availability and exchange of information, and better planning and scheduling. There are strong needs from the European industry to build more efficient and cost effective green corridors.

### 3.1 Lessons learned

Synchromodality is the optimal alignment (synchronization) in and between networks of shippers, carriers and available infrastructure. A strong growth in transport requires a better use of transport facilities and infrastructure, and the reinforcement each other, as this is the most economically attractive. At the macroeconomic level, the objectives of synchromodality could be determined as follows:

- maintain an efficient transport network that can handle the growing transport demand from shippers, both present and in the future;
- consider sustainability as a critical decision variable which makes economic growth and decrease of $CO_2$ emissions go hand in hand by far-reaching sustainable transport;

- Physical and digital infrastructure should be used to optimum quality and reliability

At the microeconomic level, these objectives are translated into operational action. Within and between the networks of the transport system shippers and carriers can work together to enable flexible switching between modes and load flows, and pooling, in order to achieve fuller vehicles, shorter lead times, more reliable services, an increased level of service, fewer transport movements and a lower integral cost for transportation. Also for the individual organisation, synchromodale alignment and bundling will creatre in this way increased cost efficiency, more sustainable transport and a better use of resources. [Het Topteam Logistiek, 2011]

A synchromodality pilot was run successfully in The Netherlands [Lucassen & Dogger, 2012]. The route was chosen between Rotterdam, Moerdijk and Tilburg. A synchromodality framework was developed (see Figure 2).
More formalised collaboration improved cooperation and control (to develop joint and real-time planning of containers in a network). Strategic locations for inland terminals and network design options needed to be considered. The way parties cooperate in the network also influences existing business models and ownership of assets. Ownership of assets has a large influence on business interests and can result in sub-optimisation. A modal split can also be stimulated by putting penalties on parties in the network. Influence of government on the transport system can be substantial, while interests of municipalities and the government are not always aligned with the interests of businesses. Cooperation between authorities and companies is necessary. A flexible and dynamic transport system is needed to give users the opportunity to benefit from switching. In current systems no formalised processes for switching exist and inflexibilities like cut off times do occur. Issues like the need for customs transit declarations, and the fact that insurance is generally done per modality and that container documentation is generally organised for container groups, make switching very difficult. The role of the shipper is important because mode free booking is a prerequisite for using a synchromodal transport concept. Shippers are mainly interested in receiving a reliable service against low costs. Price incentives can also be used to stimulate synchromodal transport. The benefits are expected to be with the transport executors who can benefit from better utilisation of their infrastructure and transport assets. The synchromodal transport concept promotes flexibility in all ways and never enforces a single view on the network and cooperation. [Lucassen & Dogger, 2012]

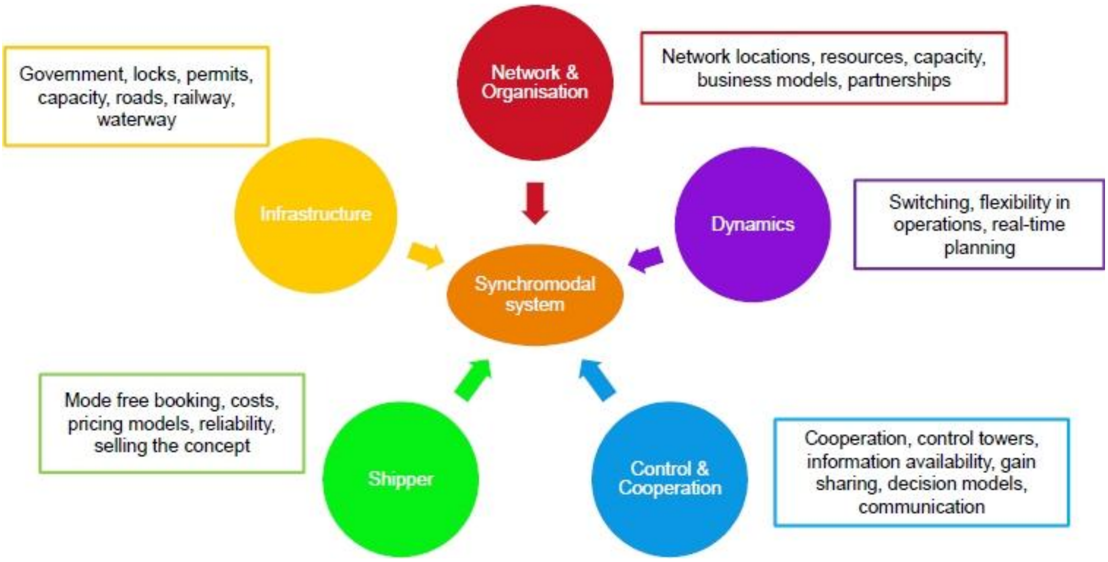



**Fig. 2.**
*A framework of synchromodal transport system*
*Source: (Lucassen & Dogger, 2012)*

Although synchromodality is plausible and promising, the synchromodal transport system is somehow complex (compared with the current system), and requires an adequate physical and information infrastructure, willingness of participants throughout the various supply chain disciplines, and as well a mind shift of the various actors the chain. The following pre-requisitions and/or risks for the implementation and transition should not be ignored:

- understanding of the concept of synchromodality in the area is insufficient;
- insufficient (triple-helix) entities throughout the various supply chain disciplines are willing to participate;
- a mind shift is a must for realising synchromodality, however, this is very difficult to achieve;
- key partners need to take the initiative, but are often confronted with obstacles, such as lack of a proper systematic approach, and inadequate support in terms of collaboration and financing.

### 3.2 Challenges for Further R&D

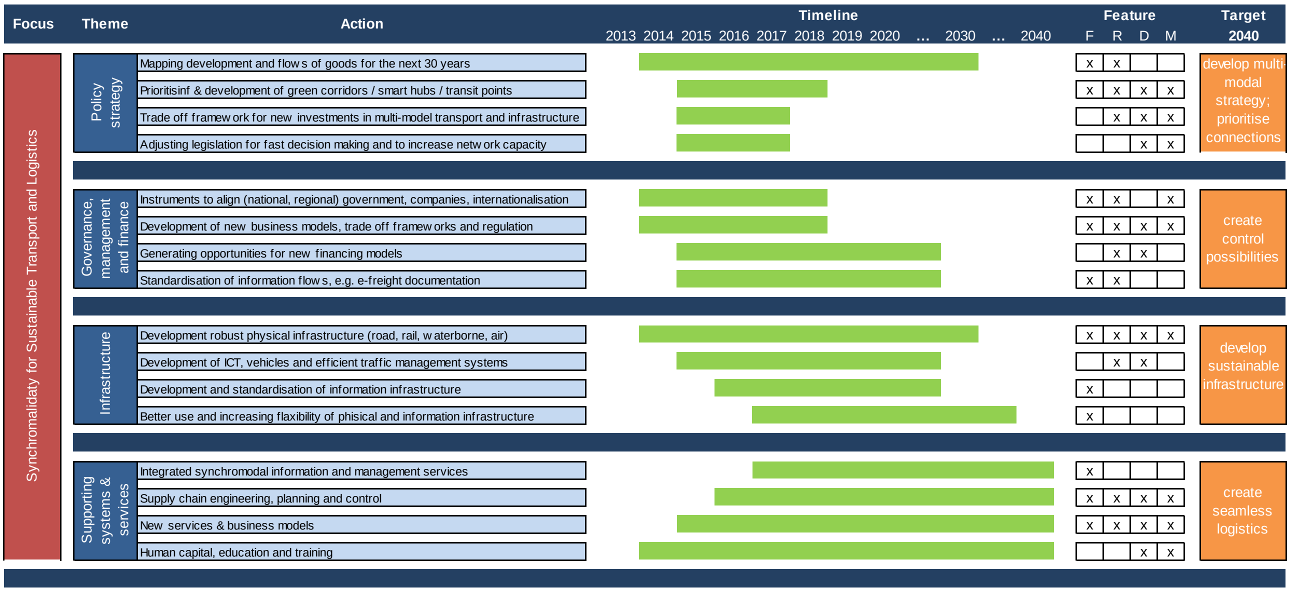
The main challenges for further research and innovation can be summarised as follows:

- Identify the most appropriate freight transport markets for the synchromodal concept (i.e. defined by transport distance, value of transported goods, hetero/homogeniety and density of the flows, type of packaging);
- promote the deployment, assessment, and evaluation of a synchromodal transport system (based on developed information infrastructure and data model), to optimise logistics (including the supply chain);
- bring together companies in, for example, different clusters (or regions) to develop efficient "back-hauls" for synchromodal shipments; logistics clusters could support and facilitate to identify possible products for "back-hauls", to avoid empty shipments, and/or to substantially increase load factors;
- investigate the societal costs and benefits of a synchromodal view on (freight) transport, implement seamless intermodal transport and a freedom of choice between (combinations of alternative) modes;
- explore the most important barriers for implementation and transition in different regions;
- develop instruments (concepts, services, physical infrastructures, technologies, legislation, governance) to overcome barriers.

In addition, from a scientific perspective, there is also a discussion as to what extent synchromodal planning is done offline or online. It is important to mention that many goods are delivered following a very fast service pattern, despite, in many cases, the absence of the necessity to do so. This opens up the prospect for adequate offline multi-modal planning of the transport movements in the transit chain. Shipment sizes (after consolidation) should be such that all resources in the transit chain are used as efficiently as possible while maintaining a requested service level. An online adaptation related to real-time problems (as in synchromodality) of these offline-prepared plans seems very difficult in the transit chain due to the very limited number of degrees of freedom to switch modes in practice. This depends largely on the underlying transit chain and the number of available transportation modes between the different nodes.

## 4. Roadmap of synchromodality at European Level

The development of strategies for a synchromodal transport system in Europe is associated with various themes, such as policy, governance, management and finance, physical and information infrastructure, new technologies, as well as supporting systems and logistics services. Figure 3 presents an ambitious roadmap of synchromodal transport at the European level.



*Note: F – Fundamental; R – R&D; D – Deployment; M – Market take-up / Pilots*

**Fig. 3.**
*A roadmap of synchromodality implementation and transition in Europe*
*Source: (own, adapted based on [Het Topteam Logistiek, 2011] )*

## 5. Conclusions

Synchromodality takes a holistic view on transport. In a synchromodal transport system, one could choose between different modalities at any time, based on real-time information and conditions, e.g. a low water level, the emergency of the load, and the traffic and transport situation. Synchronised services requires seamless processes in seaports and seamless information exchange and service design.

Synchromodality could be flexibly deployed in all modes of transport on the basis of cooperation between modalities (unlike competition between modes): water (inland waterways or short sea), rail, air and/or road. It could pave the way for effective policies (legislation, investments in physical infrastructures, investments in supporting technologies, instruments to overcome implementation barriers), as well as create new business opportunities based upon value-added logistics services.

It must be emphasised that synchromodal transport needs to be matched with public policies to enable effective multi-modal implementations. When appropriate policy instruments are implemented, the impact of the implementation of the research results will be a better functioning and performing, and thus more sustainable and cost-effective transport system.

## Acknowledgements

The paper presents part of the result of a EU-funded project SoCool@EU (Sustainable Organisation between Clusters Of Optimised Logistics @ Europe), DG RTD (Research and Innovation), Regions of Knowledge Programme (FP7-REGIONS-2011-1). The author especially thanks the consortium partners, and participating third parties for their valuable suggestions and support.

**References**

De Langen, P.; Van der Horst, M. 2008. Coordination in hinterland transport chains: a major challenge for the seaport community. *Maritime Economics and Logistics*, vol. 10, no. 1-2 , pp.108-129.

ECT. 2011. *The future of freight transport: ECT's vision on sustainable and reliable European transport*. ECT (Europe Container Terminals), Rotterdam.

European Commission. 2011. *White Paper: roadmap to a single european transport area – towards a competitive and resource efficient transport system.* EC COM(2011) 144, Brussels.

EVO. 2011. *EVO Logistics Yearbook: Your guide to the world of logistics and supply chain management*. Zoetermeer. ISBN: 978-90-79470-00-6.

Gorris, T.; Groen, T.; Hofman, W.; Janssen, R.; Van Meijeren, J.; Oonk, M.; et al. (2011). *Implementatieroadmap Synchromodaliteit*. TNO, Connekt & Dinalog, commissioned by the Ministry of Infrastructure and Environment. The Hague.

Het Topteam Logistiek. 2011. *Partituur naar de top - Adviesrapport Topteam Logistiek.* Het Topteam Logistiek. The Hague.

Lucassen, I.M.P.; Dogger, T. 2012. *Synchromodality pilot study: identification of bottlenecks and possibilities for a network between Rotterdam, Moerdijk and Tilburg*. TNO 2012 P10128 (ver. 1.0), Delft.

Oonk, M.E. 2014. Smart logistics corridors and the benefits of Intelligent Transport Systems. In *Proceedings of Transport Research Arena 2014*, Paris

SoCool@EU Consortium. 2011. *Description of Work*, SoCool@EU (Sustainable Organisation between Clusters Of Optimised Logistics @ Europe), SoCool@EU Consortium. (Restricted)

Tavasszy, L.; Van der Lugt, L.; Janssen, G.; Hagdom - Van der Meijden, E. 2010. *Verkenning Synchromodaal Transportsysteem*. TNO, commissioned by the Ministry of Infrastructure and Environment, The Hague.

TNO. 2010. *Verkenning synchromodaal transportsysteem*. TNO (Toegepast Natuurwetenschappelijk Onderzoek), TNO-034-DTM-2010-04431 Delft.

# A LONGITUDINAL SURVEY STUDY OF IZMIR COMMUTER RAIL SYSTEM (IZBAN)

**Sen Samet[1], Alver Yalcin[2]**
*[1,2]Ege University Civil Engineering Department, Izmir, Turkey*

**Abstract:** Before Izmir Commuter Rail System (IZBAN), most of the trip makers along the railway were using the city buses, minibuses or private cars. After IZBAN was put into service, some people changed their previous trip behaviors and they started travelling by IZBAN. Therefore a big travel demand in IZBAN occurred. In this study, the characteristics of passengers and their trip behaviors are determined based on the longitudinal data conducted via two wave trip surveys. Just after one year from IZBAN's opening, the first wave of the surveys was carried out among 539 passengers at six stations during morning peak hours between 07.00 am - 09.30 am. The second wave was carried out among 669 passengers at the same six stations two years after the first wave during the same morning peak hours. As a result of this study, the respondents' socio-economic characteristics, the distribution of trips by region, the impact of IZBAN on transport modes, the change in travel time and travel cost, current trip data and satisfaction data were obtained. These data enabled to compare two waves and explain the changes in socio-economic factors and trip behaviors. Also by using binary logit model, the factors that affected willing of car use were found. In both waves, 10% of the respondents stopped driving their own cars and they started to take IZBAN. This is an important development in solving traffic problems. Converting from private car to public transportation means less traffic congestion.

**Keywords:** Commuter rail system, comparative study, longitudinal survey, public transportation, binary logit model

## 1. Introduction

Central business districts (CBD) are becoming more attractive since the metropolitan areas have been growing economically and socially. Commercial areas, schools, recreational areas and health care centers lead to an increase in ridership to the downtowns of the cities. Commuter rail systems are invested so as to provide faster, safer, more economic and more comfortable connection between downtowns and suburban areas. Briefly, commuter rail systems are the railway systems that connect suburban areas to the city centers. Generally, residential buildings and population are great in particular parts of a city. Commuter systems enable to create homogeneous structured areas and they can create other dynamic centers. Finally, population density can be decreased in city centers.In this study, two wave trip surveys for Izmir Commuter Rail System (IZBAN) users were carried out. These kinds of surveys are called *longitudinal surveys*. The defining characteristic of a longitudinal study is that individuals are measured repeatedly through time (Diggle et al., 2002). There are four important advantages of longitudinal study in transportation planning and research. The first reason, the effects of transportation policies, which are usually in terms of changes in trends, can be detected more easily using repeated observations. Second, the history and experience of a person can be very important in explaining current behavior. Third, longitudinal surveys offer, mainly, the possibility of focusing on the variable(s) of interest and removing the possible bias of various time independent observed and unobserved factors. The fourth, an established longitudinal sample provides rapid and inexpensive access for specific policy-relevant issues (van Wissen and Meurs, 1989).
According to the two wave questionnaire surveys, the following results were obtained:

- Respondent's socioeconomic factors
- Respondent's previous mode choices before the opening of IZBAN
- Respondent's current trip information
- Service quality of IZBAN

Building a commuter rail system was a necessity because of a high traffic demand to the city center, an increase in auto-ownership ratio and insufficient road network. The following positive results will occur if the car drivers are attracted to the commuter rail system:

- Decrease in traffic congestion
- Increase in trip capacity of public transit
- Lower accident probability
- Decrease in environmental problems

In section two, a literature review is presented about longitudinal studies in transportation research. In section three, IZBAN is introduced. In section four, longitudinal survey methodology is explained. In section five, the two-wave survey results are considered and compared. And finally in section six, conclusion is mentioned.

## 2. Literature Review

Many studies have been done about longitudinal research in transportation planning. In this section, some researches are mentioned. While research methods are refined and the limitations of cross-sectional research are documented, there has been an increasing interest in longitudinal research methods for investigating passenger's trip characteristics (Bissell et al., 2008)

[1] Corresponding author: yalcin.alver@gmail.com

Hensher (1986) published a paper about the dimensions of automobile demand. The major objective of his study was to obtain reliable forecasts of the variables which drove the fundamental energy equation: “energy consumed = efficiency of technology (liters/100 kilometers)*utilization rate (kilometers per period)”. So it was necessary to view the levels of vehicle usage and vehicle fuel efficiency as outputs of the broader household decision process. For this reason, he conducted four wave household surveys. He investigated the individual (household) choice behavior. From longitudinal surveys, the respondents' vehicle data (make, model, body type etc.), household data for every member (age, driving license, occupation, income etc.) and other data (housing rent, daily trips, major household activities etc.) were obtained. At the end of that study, they developed an econometric model system which jointly modeled the household's choice of vehicles and utilization level by using longitudinal data (Hensher, 1986).

A ten wave longitudinal survey in Seattle was conducted by The Puget Sound Transportation between 1989 and 2002. It was designed as a general purpose urban household study and mainly focused on transportation. Information obtained from the survey has assisted in long range transportation forecasting and analysis used to inform decisions regarding highway and road construction, public transportation development, carpooling and parking policies (Ortuzar et al., 2010).

In Japan a series of seven panel surveys were carried out between 1987 and 2005. The first survey had seven waves and it was used to evaluate trip behavior changes caused by an automated railway (AsTram) connecting a newly developed suburban area to the CBD in Hiroshima. The largest survey had nine waves between 1989 and 1997. It was used to measure the effects of a newly-opened shopping center in Kofu on shopping behavior. The last one had only two waves (2003-2005) and was used to evaluate the behavioral changes of registrants of special transport services in Okayama (Ortuzar et al., 2010).

"The role of commuter rail in Istanbul's public transportation" was studied as a thesis subject in 2005 at Fatih University, Turkey. For the thesis, a survey was conducted. One of the objectives of this study was to determine the underlying reasons of passenger reduction and operational problems of Istanbul's commuter rail system. Moreover a question about the passenger's mode choice before commuter rail was asked. According to the survey results, the most common mode used before commuter rail was the mode of minibus (Isik, 2005).

Massachusetts Bay Transportation Authority (MBTA) conducted some studies: highway traffic counts, passenger counts, surveys and before & after studies. The passengers were asked which public transportation mode they used before Greenbush commuter rail. Furthermore, the passengers were asked about their trip origins. The most significant question was: how did you make your trip before Greenbush rail?" According to the survey data, there were some comparisons about trip prices and trip times of the passengers. In addition, the impact of new lines on other modes was considered. After the new line was opened, the most affected mode was commuter boat service. Greenbush commuter rail attracted passengers from express bus lines, boat lines and other rail lines (Guptill, 2010).

## 3. Izmir Commuter Rail System (IZBAN)

Izmir is the 3rd largest metropolitan city of Turkey with 4,061,074 population in 2013 (TSI, 2013). Izmir Metropolitan Municipality is responsible for all province area of 12,007 km$^2$.

Day by day, traffic problems are getting worse in Izmir. Building a new rail system has become essential because of the over capacity of available transportation modes and high ratio of private car usage. Petrochemical, steel and other industrial zones, historical places and universities are located in Izmir. These places produce a high ridership demand. For this reason it is very difficult to solve traffic problems with rubber-wheeled vehicles which have low capacity. *Izmir Metropolitan Municipality* (IMM) and *Turkish State Railways* (TSR) prepared a project to solve traffic problems and to cover high demand of ridership in north-south line of the city. Actually this project was a rehabilitation of old rail lines which were built at the end of 19th century by French and British companies. The system was opened on 6th of March, 2011 officially.

IZBAN is the biggest commuter rail system in Turkey with the connection to the Adnan Menderes International Airport. It has 80 km length. There are 32 stations along the line. IZBAN has north and south axles. The north axis has 20 stations between Aliaga-Halkapinar. Average travel time is 60 minutes in this axis. The south axis has 12 stations between Alsancak-Cumaovasi. Average travel time is 30 minutes in this axis. Furthermore, IZBAN has two transfer stations to the Izmir Metro (Subway line of Izmir City). These transfer stations are called Halkapinar and Hilal. Fig. 1 shows the map of IZBAN.


**Fig. 1.**
*Izmir Commuter Rail System and Izmir Metro Map*

## 4. Longitudinal Survey and Survey Results

Due to high response ratio and to gather more detailed data, face to face survey technique was used in this study (Owens, 2002). For this study, two-wave longitudinal survey was conducted. The first wave of the surveys was carried out between April 10, 2012 and May 11, 2012. The second wave was carried out between November 18, 2013 and January 15, 2014. The surveys were conducted at six stations between 07.00 am and 09.30 am. The reason for choosing morning peak-hours was to capture work and school trips. Fig. 1 shows the stations in bold dots where the surveys were conducted. These stations were; Sirinyer, Alsancak, Bayrakli, Karsiyaka, Cigli and Menemen. The survey includes four main parts. These parts are: personal information, trip modes before IZBAN, trip characteristics of IZBAN and general questions about IZBAN.

In the first part, personal information was collected. The passengers were asked about age, occupation, physical disability status, monthly income, driving license, auto-ownership, and education level. In the second part, the main question was: "Which transportation mode (public transportation, private vehicles, other modes) did you use for your same trip before IZBAN?". The third part was "trip information after IZBAN" in which the passengers were asked about their trips' origin-destination zones. After this question, they were asked about at which station they took off the train. Furthermore, trip purposes of the passengers were asked. The other question was about how they arrived at IZBAN station and how they continued their trips until they reached their destination points. Also travel times and travel costs of the passengers were asked. The last question was about passengers' weekly usage frequency of IZBAN.

The last part comprised a detailed satisfaction questions. The passengers were asked if they were satisfied with IZBAN or not. The answer range was; very dissatisfied, dissatisfied, neutral, satisfied and very satisfied.

In the first wave, a survey was conducted among randomly chosen 539 passengers at six IZBAN stations in peak hours between 07.00 am - 09.30 am while the second wave of the survey was conducted among randomly chosen 669 passengers at the same IZBAN stations during the same peak hours. Surveys were completed addressing the questions to each respondent personally, which is called face-to-face survey method. Details about the numbers of surveys for each wave are shown in Table 1.

**Table 1**
*Numbers of Surveys for Each Wave*

| Station | First wave | Second wave |
|---|---|---|
| Sirinyer | 99 | 121 |
| Alsancak | 66 | 100 |
| Bayrakli | 95 | 121 |
| Karsiyaka | 102 | 121 |
| Cigli | 101 | 109 |
| Menemen | 76 | 97 |
| Total | 539 | 669 |

In the first part of the surveys passengers were asked to specify their socioeconomic characteristics. In Table 2, detailed socioeconomic factors are shown. According to the socioeconomic factors, there was a slight increase in age group of 18 and younger by 4% while the ratio of age 30-39 decreased sharply by 6%. In the first wave the ratio of private sector workers was 47% as the ratio in the second wave was 51%.

There was an important change in monthly income groups as well. The ratio of respondents who had +2000 Turkish Lira (TL) income increased by 5% in the second wave. It showed that people who had a higher income had a tendency of using public transportation. While in the first wave, the ratio of respondents who had driving license was 67%, in the second wave the ratio was 64%.

One of the biggest goals in public transportation was to promote the car drivers to use public transportation instead of their cars. In this study, the ratio of respondents who had their own cars increased by 5% pleasingly. It meant that while the car ownership ratio was increasing, people were being more aware of public transportation and they preferred it.

**Table 2**
*Detailed socio-economic factors*

| | | **Percentage (%)** | |
|---|---|---|---|
| | | **1st wave** | **2nd wave** |
| **Gender** | Female | 33 | 30 |
| | Male | 67 | 70 |
| **Age** | <18 | 11 | 15 |
| | 18-29 | 39 | 33 |
| | 30-39 | 23 | 22 |
| | 40-49 | 16 | 18 |
| | 50-64 | 9 | 11 |
| | +65 | 2 | 1 |
| **Disabled** | Yes | 1 | 2 |
| | No | 99 | 98 |
| **Occupation** | Public servant | 17 | 18 |
| | Private sector | 47 | 51 |
| | Own job | 6 | 4 |
| | Retired | 1 | 1 |
| | Student | 25 | 25 |
| | House wife | 2 | 1 |
| | Not working | 1 | 0 |
| **Monthly income (TL)** | <Minimum wage | 29 | 28 |
| | Minimum wage-1000 TL | 14 | 16 |
| | 1000 TL-1500 TL | 23 | 18 |
| | 1500 TL-2000 TL | 18 | 17 |
| | +2000 TL | 16 | 21 |
| **Education level** | Primary school | 9 | 11 |
| | Secondary school | 6 | 8 |
| | High school | 40 | 34 |
| | Associate degree | 5 | 5 |
| | Bachelor's degree | 37 | 37 |
| | Master's degree | 3 | 5 |
| **Driving license** | Yes | 67 | 64 |
| | No | 33 | 36 |
| **Car ownership** | Yes | 22 | 27 |
| | No | 63 | 61 |
| | Share with family | 15 | 12 |

In the second part of the surveys passengers were asked about their previous transportation modes before IZBAN. In the first wave 85% of the respondents were making their same trips by public transportation vehicles (city bus, minibus), 10% of them were driving their own car for the same trip while 5% of them were using other modes. Also in the second wave, it was found that 34% of the respondents were new users of IZBAN.

In the third part, the passengers were asked about their trip purposes. In the first wave 71% of the respondents were making *work* trip while 67% of them were making *work* trip in the second wave. In both waves, the ratio of *school* trips was the same by 25%.

One of the sharpest results was the change in average travel time. While in the first wave the average travel time was 45 minutes, in the second wave it was 60 minutes. Furthermore, the average travel cost of the respondents increased by 9% because of the increase in trip fares. Also, in the third part of the survey, the passengers were asked about their weekly usage frequency. In the first wave, the ratio of the passengers who took IZBAN once a week was 4% while the ratio was 8% in the second wave. The ratios of the respondents who took IZBAN 2-3 times, 4-5 times and 6-7 times a week had very little changes.

The respondents were also asked about how they arrived at IZBAN station. In both waves, the ratio of arriving by buses was 38%. The ratio of arriving by walk was 51% in the first wave while the ratio was 49% in the second wave. The next question was about how they reached their destinations after taking off IZBAN. The bus usage increased by 5% and the ratio became 22%. The ratio of walking after IZBAN was 54% in the first wave, while the ratio was 47% in the second wave. In the first wave of the surveys, there was only one transfer station between IZBAN and Izmir Metro. The station is called *Halkapinar.* In the first study, 19% of the respondents were taking Izmir Metro at Halkapinar station. Five months ago from the second wave, the second transfer station between IZBAN and Izmir Metro was opened. That station is called *Hilal.* So that, the passenger density at Halkapinar transfer station decreased. 10% of respondents were taking Izmir Metro at Halkapinar while 4% of them were taking Izmir Metro at Hilal station. The total ratio of passengers who transferred to Izmir Metro decreased by 5% in the second wave. Table 3 shows the details of the current trip information of the respondents.

**Table 3**
*Detailed current trip information*

| | | Percentage (%) | |
|---|---|---|---|
| | | **1st wave** | **2nd wave** |
| **Trip purpose** | Work | 71 | 70 |
| | School | 25 | 25 |
| **Mode choice to IZBAN** | Walk | 51 | 49 |
| | City bus | 38 | 38 |
| | Private car | 4 | 4 |
| **Mode choice after IZBAN** | Walk | 54 | 47 |
| | City bus | 17 | 22 |
| | Izmir Metro | 19 | 14 |
| **Weekly usage frequency (day)** | 1 | 4 | 8 |
| | 2-3 | 12 | 10 |
| | 4-5 | 45 | 44 |
| | 6-7 | 39 | 39 |
| **Transfer to Izmir Metro** | Halkapinar | 19 | 10 |
| | Hilal | Not available | 4 |
| **Average travel time** | | 45 min | 60 min |
| **Average travel cost** | | 1,79 TL | 1,95 TL |

In the last part of the surveys, the passengers were asked seven questions about satisfaction of respondents, these questions are:

- Question 1: Information system satisfaction at transfer and IZBAN stations
- Question 2: Information system satisfaction in the train
- Question 3: Transfer satisfaction to IZBAN or from IZBAN
- Question 4: Station's physical condition satisfaction
- Question 5: Headway satisfaction
- Question 6: Comfort satisfaction in the train
- Question 7: Satisfaction for the disabled passengers

As it can be seen from Table 4, the passengers were satisfied with train's comfort and information systems in the trains in both waves. Also the ratio of transfer satisfaction was 67% in the first wave while the ratio was 56% in the second wave. This decrease was caused because of the construction of new transfer bridge which connects IZBAN and Izmir Metro at Halkapinar during the second wave. The ratio of headway satisfaction was 49% in the first wave. But in the second wave the ratio was 40%. Day by day, the ridership was increasing. For this reason the train was getting more crowded and this reason affected the headway satisfaction. Moreover, the ratio of disabled people's satisfaction decreased from 47% to 43%.

**Table 4**
*Satisfaction results*

| | Very dissatisfied(%) | | Dissatisfied (%) | | Neutral (%) | | Satisfied (%) | | Very Satisfied (%) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1st | 2nd | 1st | 2nd | 1st | 2nd | 1st | 2nd | 1st | 2nd |
| **Q1** | 5 | 6 | 25 | 29 | 6 | 5 | 62 | 59 | 2 | 0 |
| **Q2** | 1 | 2 | 15 | 14 | 5 | 5 | 74 | 77 | 5 | 2 |
| **Q3** | 3 | 7 | 20 | 22 | 8 | 14 | 67 | 56 | 2 | 1 |
| **Q4** | 6 | 9 | 21 | 25 | 5 | 8 | 62 | 57 | 6 | 1 |
| **Q5** | 8 | 16 | 34 | 38 | 8 | 5 | 49 | 40 | 1 | 1 |
| **Q6** | 1 | 3 | 12 | 11 | 5 | 6 | 77 | 78 | 5 | 2 |
| **Q7** | 4 | 10 | 25 | 29 | 22 | 18 | 47 | 43 | 2 | 0 |

## 5. Modelling Methodology and Model Results

Modelling part of the study aimed to find the socio-demographic characteristics of the passengers who made the current trips by their private cars before IZBAN. Binary logit model was chosen for modelling. Private car users before IZBAN were used as the dependent variable, and the answers to questions related to socio-demographic characteristics and current trip information were used as the independent variables for binary logit models. The utility function of the model is as follows:

$U_{in} = V_{in} + \varepsilon_{in} \qquad U_{jn} = V_{in} + \varepsilon_{jn}$

Where; $V_{in}$, contains the independent variables of the utility function and $\varepsilon_{in}$ is the error term. In the binary logit model $\varepsilon_n = \varepsilon_{jn} - \varepsilon_{in}$ is accepted to be logistically distributed. In other words, the error term is independent and identically distributed. The probability of choosing alternative *i* in the logit model is shown in the following equation:

$$P_n(i) = \Pr(U_{in} \geq U_{jn}) = \frac{1}{1+e^{-\mu(V_{in}-V_{jn})}} = \frac{e^{\mu V_{in}}}{e^{\mu V_{in}+V_{jn}}}$$

$P_n(i)$ = the probability of choosing alternative i, $\mu$=positive scale parameter.

Only the data of the respondents who have driving license and who have private car or share a car with family members were used. So that, 199 surveys from first wave and 254 surveys from second wave were used for modelling. According to the binary logit model results, *self-employed* (passengers who have their own job) variable, *income5* (passengers who have an income of more than 2000 Turkish Liras) variable, *car owner* (passengers who have their private cars) variable and *weeklyusage2* (passengers who use IZBAN 2-3 days a week) variable are significant for the first wave and they all have positive effects on preferring private car usage while *civil servant* (passengers who work for government) variable, *age2* (passengers who are between 18-29) variable, *income3* (passengers who have an income between 1100-1700 Turkish Liras) variable, *car-owner* (passengers who have their private cars) variable, *education4* (passengers who still study at university or graduated from university) variable, *education5* (passengers who still study master’s degree or finished their master’s degree) variable and *weeklyusage1* (passengers who use IZBAN 1 day a week) variable are significant for the second wave. *Civil servant* variable, *age2* variable and *income3* variables have negative effects. *Car-owner* variable, *education4* and *education5* variables and *weeklyusage1* variable have positive impacts on preferring private car usage. Both waves have only one same significant variable (*Car owner*). But the other significant variables are consistent in both waves. High income (*Income5*) increase the car usage rate in the first wave where low income (*Income3*) decrease the car usage rate in the second wave. The combined results for estimated binary logit models are shown in Table 5.

**Table 5**
*Results of binary logit models*

| | First wave | | | Second wave | | |
|---|---|---|---|---|---|---|
| | **Variable** | **Coefficient** | **Prob.\|z\|>Z*** | **Variable** | **Coefficient** | **Prob.\|z\|>Z*** |
| | Constant | -2.8923*** | .0000 | Constant | -1.8355*** | .0003 |
| | Self-employed | 2.0035*** | .0004 | Civil servant | -1.0982*** | .0077 |
| | Income5 | 1.2807*** | .0017 | Age2 | -.9889** | .0300 |
| | Car owner | .8639* | .0614 | Income3 | -1.3002** | .0250 |
| | Weeklyusage2 | 1.2372** | .0228 | Car owner | .8326* | .0654 |
| | | | | Education4 | .7556** | .0471 |
| | | | | Education5 | 1.2422** | .0345 |
| | | | | Weeklyusage1 | .9176** | .0459 |
| **Sample size** | 199 | | | 254 | | |
| **Log likelihood function** | -82.23146 | | | -115.91622 | | |
| **Restricted log likelihood** | -99.85671 | | | -132.71192 | | |

**6. Conclusion**

In this study, the respondents' socio-economic characteristics, the distribution of trips by region, the impact of IZBAN on transportation modes, the changes in travel time and travel cost and satisfaction data were obtained based on the longitudinal data conducted via two wave trip surveys among IZBAN users at the train stations. After obtaining the data of IZBAN users, some comparisons were made. Car ownership rates were 22% and 27% in the first and second waves, respectively. Although car-ownership rates of the passengers increased, the passengers had a tendency of using IZBAN. Also in both waves, before IZBAN 10% of the respondents were making their same trips by their private cars. After IZBAN, they left their private cars and they started to use IZBAN. Both results are quite pleasant because as more cars are removed from the streets, the less traffic congestion will be.

In both waves, 38% of the respondents were arriving at IZBAN station by city buses. This showed the importance of feeding bus system. It can be seen from the satisfaction results that people wanted an economic, fast, safe and comfortable public transportation system.

Results showed that younger users (*age2*), civil servants who have fixed working hours and low income groups (*Income3*) tend to use IZBAN where higher income groups (*Income5*), self-employed users who have various working hours tend to use their private cars more.

**References**

Bissell, L.; MacKay, J.; Vogt, A. 2008. Using panel studies in travel research: Issues of participation, participant attitudes and best practices, Canada.

Diggle, P.; Heagerty, P.; Liang, K.Y.; Zeger, S.L.2002. *Analysis of longitudinal data*, Oxford Statistical Science Series 25, 2nd ed., Great Britain, 369 p.

Guptill, R. 2010. Greenbush commuter rail before and after study, Cental Transportation Planning Staff for Massachusetts Bay Transportation Authority.

Hensher, D.A. 1986. Dimensions of automobile demand: A methodological overview of a research project, *Environment and Planning A*, 18(10): 1339-1374.

Isik, A. 2005. Role of commuter rail in Istanbul's public transportation: A case study of Haydarpasa-Gebze line, Turkey

Ortuzar, J.D.; Armoogum, J.; Madre, J.L.; Potier, F. 2010. Continuous mobility surveys: The state of practice, *Transport Reviews: A Transnational Transdisciplinary Journal*, 31(3): 293-312.

Owens, L. K. 2002. Introduction to survey research design, Survey Research Laboratory Fall Seminar Series, Available from Internet: <http://www.srl.uic.edu>

TSI(Turkish Statistical Institute). 2013. Available from Internet: <http://www.tuik.gov.tr/UstMenu.do?metod=temelist>

VanWissen, J.G.; Meurs, H.J. 1989. The Dutch mobility panel: Experiences and evaluation", *Transportation*, 16(2): 99-119.

# RAILWAY INDUSTRIAL TRACK AS THE LAST MILE IN SUPPLY CHAIN MANAGEMENT

**Borna Abramović[1], Nikolina Brnjac[2], Jasmina Pašagić Škrinjar[3]**
[1] *University of Zagreb, Faculty of Transport and Traffic Sciences, Vukelićeva 4, HR – 10000 Zagreb, Croatia*

**Abstract:** Today's transport of goods has renaissance in using again railway as part of transport chain. In the past because of inert railway the very big amounts of goods go to road transport. Of course, in these days, it is very ecologically to do modal shift from road to railway. Therefore, the railway is again one of the key player in transport chain. Supply chain management is integrated approach for solving all movements of good form point of origin to point of destination. Specially, last five years is very intensive research and use of supply chain management. Normally, every transport have last and first mile, so it is very pragmatic to optimize this step in whole transport chain. With the greater role of railway in whole supply chain management the last mile become very significant. In today's railway, last mile is on more than 85% of transport goods on railway industrial track. Therefore, it is very important to optimize every aspect of using railway industrial track. Key aspects are (1) legal, (2) organizational, (3) technical, (4) technological and (5) economic. All those five aspects are crucial for using of railway industrial track as the part of supply chain management. In this article, it will be present case study of using the railway industrial track in Croatia and interaction and optimization of key aspects.

**Keywords**: railway industrial track, supply chain management, goods

## 1. Introduction

Today's world economy it's not possible without transport. For business is very important to have stable, safety and cheap transport. In last decade was many optimization methods use to cut down cost of transport. One of most complex method for optimizing transport process is Supply Chain Management. Main goal of Supply Chain Management is to create a chain of suppliers that focuses on maximizing value to the ultimate customer. Supply Chain Management is the integration of the activities that procure materials and services, transform them into intermediate goods and final products, and deliver them through a distribution system. Because of that we can talk today that is no longer competition between companies but competition is between supply chains.

Standard Supply Chain has: (1) supplier, (2) manufacturer, (3) distributer, (4) retailer and (5) customers. For transport way of thinking it's important which transport mode it uses for connection in the most cost effective way. Two most cost effective transport modes are water (maritime and inland water) and railway. On area where is not possible to use water transport as a rule in transport chain is using railway. Unfortunately for railway because of too long transit times, regardless of transport cost, very often is use road transport. Therefore, for railway operators the crucial thing is to find practical ways to cut down the transit times. One of the methods to cut down transit time is using of Railway Industrial Track for first and last mile. Considering the technology in railway freight transport, the Railway Industrial Track is the most frequent starting and end points of the technological process. The adjective "industrial" in the name of these tracks originates from the times of the advent of railways and the establishment of the railway stations. The development of the railway was closely connected with the expansion of the industrial production method which required mass transport of goods. The first Railway Industrial Track were constructed in order to connect the stations with big industrial and production plants, mines, ports and quays, where mass goods, especially coal, timber and iron ore were loaded and unloaded. By using the Railway Industrial Track the railway operators provide "door-to-door" service, thus being competitive to road traffic, and the user can directly load/unload their goods at their premises, in fact, comprehensive transport service is offered. When railway use industrial track they can be very important player in chain supply management.

## 2. Last Mile in Supply Chain Management

Supply Chain Management is an integrated approach to planning, implementing and controlling the flow of information, materials and services from raw material and component suppliers through the manufacturing of the finished product for ultimate distribution to the end customer. It includes the systematic integration of processes for demand planning, customer relationship collaboration, order fulfilment/delivery, product/service launch, manufacturing/operations planning and control, supplier relationship collaboration, life cycle support, and reverse logistics and their associated risks. These processes, which employ a combination of people, systems and technology, can be performed by the firm itself or in collaboration with external supply chain partners.

Supply Chain Management is strategic in orientation and recognizes that the competitive strength of a firm is not only determined by its products but also by the operations and activities that place the products into customers' hands and provide supporting services. Efficient and effective Supply Chain Management enhances firm performance and adds value by increasing asset utilization to gain competitive market advantage. The responsiveness and efficiency of a company's supply chain arising from its design and management is integral to the firm's ability to successfully compete in the global marketplace.

[1] Corresponding author: borna.abramovic@fpz.hr

Because of that we can talk today that is no longer competition between companies but competition is between supply chains.

Standard Supply Chain has: (1) supplier, (2) manufacturer, (3) distributor, (4) and retailers (5) customers. Between each element comes to the transport process. Sometimes the transport process can be only internal, for example when the same company distributor and retailers, but much more naturally and with external transport process. In the literature usually greater emphasis is on the last mile because represents final link to the customer. But in the transport process first mile is as important as the last mile.

The Last Mile is a common distribution problem where near the destination a high capacity long distance transport system is facing high costs and complexity. This is mainly related to the need to break down the size of the transport unit (less economies of scale) and where congestion impairs reliability as well as capacity. Three major elements impact the effectiveness of last mile logistics:

- Terminal operations,
- Drayage operations, and
- Warehousing.

The efficiency of terminal operations in terms of capacity, turnover and gate access, will impact last mile logistics because a significant segment of freight flows are originating or bound to a terminal (maritime, rail or barge).

The standard last mile problem for drayage (movement to and from a terminal) is congestion on local and regional roads, commonly due to a lack of capacity. This particularly takes place at bottlenecks such as highway interchanges, bridges and tolls. Since containers are carried on chassis. Deliveries (movement from a distribution centre), often done in less-than-truckload fashion, are also a fundamental part of the process. Solution for this problem is in using the Railway Industrial Track.

Warehousing is a distribution centre that can be used for capabilities of performing deliveries on time. So it’s very important element in last mile.

The economics of transporting cargo theoretical distribution costs on the initial/final and transport sequence costs based on different response of these costs in the transport process. In fact, the costs per unit of transport (km) decrease with increasing distance at which the transport is performed. Practically, the size of the expenses incurred during the initial/final works, in cargo transport does not depend on the distance of transport. Therefore, the initial/final costs behave fixed according to distance of transport. If the distance is greater, then the cost of the initial/final work is distributed to a larger number of units and their size per unit decreases. In contrast, the cost of transport sequence increased with increasing distance at which the transfer takes place. Therefore, the cost to transport sequence usually said to be relatively variable distances behave.

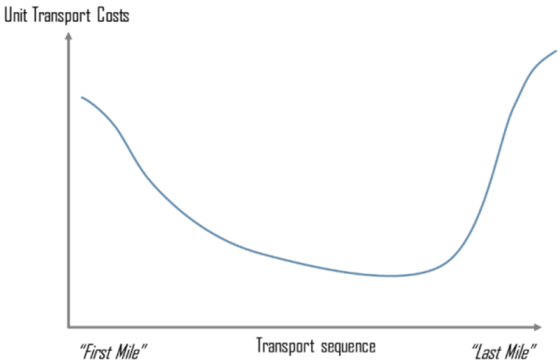



**Fig. 1.**
*First and Last Mile Unit Cost*
*Source: Rodrigue, J-P. et. al. 2013*

### 3. Supply Chain Management and Railway Industrial Track

In order that elements of Supply Chain were interconnected transport is needed. The connection of the elements and the transport is shown in Figure 2. Therefore, without transport, the internal as well the external, no meaningful Supply Chain can be made. Meaning of Supply Chain is that in certain point of time and space provide customers a certain amount of goods. The most important task is to optimize Supply Chain costs from supplier to customers while matching the demand and the supply. Just matching the demand and the supply has some difficulties:

- Uncertainty in demand and/or supply,
- Changing customer requirements,
- Decreasing product life cycles,
- Fragmentation of supply chain ownership,
- Conflicting objectives in the supply chain, and
- Conflicting objectives even within a single firm.

Interesting relationship is between marketing/sales, production and distribution. From one side Marketing/Sales wants to have more Finished Goods Inventory (FGI), fast delivery, many package types and special wishes/promotions then Production wants to have bigger batch size, depots at factory, latest ship date, decrease changeovers and stable production plan but distribution wants to have full truckload, low depot costs, low distribution costs, small number of stock keeping units (SKUs) and stable distribution plan.

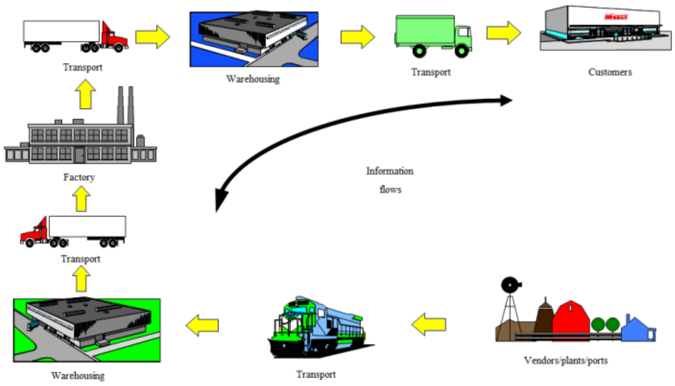



**Fig. 2.**
*Elements of Supply Chain*
*Source: Ballou, R. H. 2003*

In fact, the chain that connects the above requirements is transport. Therefore transport has an important role stabilizing Supply Chain. Transport has different modes and the Supply Chain chooses for one of transport mode that can provide transport to the default time with certain quantity of goods for a certain amount of money. Thus, there is a healthy competition between modes, but also healthy competition between different service providers of each transport mode.
Considering that there are three essential factors for choosing the transport mode: (1) the quantity, (2) the time and (3) price. Their mutual mix will give the transport mode selection. Of course, each transport mode has its advantages and disadvantages. Supply Chain Manager must have decision to choose a transport mode and then operator for having transport services. In the process of transport in the Supply Chain there is needed to transport large quantities of goods in a relatively short time. In land transport choice should fall on the railways. Railway in Europe, in the last few decades, had experience of the challenges of liberalization and for that reason must necessarily be closer to the user. Very effective way of approaching the customer's is use of Industrial Railway Track.

## 4. Advantages of using Railway Industrial Track

Today, with the intensive demands for transport service the transport chain plays the main role. The transport chain represents a connection between the “origin” and the “destination” for the purpose of transport service using adequate means, routes and technology of operation, thus in fact forming the optimal transport system. In surface transport, road transport dominates as opposed to railway transport, although both road and railway transport should be complementary. The historical fact is that railway transport is mass transport, i.e. it is always connected with big industrial complexes – users. The users have realized the advantages of railway freight transport and they constructed their own tracks within their premises. In fact, the tracks that diverge from the stations or from open railway lines, and serve for the needs of one or several users are called Railway Industrial Track. Railway Industrial Track are privately owned by various companies, mainly big users in the area of seaports, river ports, oil companies, logistic centres and production activities where there is demand for mass transport of freight. Within the company area the user may be only one, but regularly the tracks fork out into groups intended for handling activities. If the length of the basic Railway Industrial Track amounts to several hundred metres or several kilometres, it is often called an industrial or railway line, i.e. the area that includes the Railway Industrial Track is called a railway complex.
It is precisely the dominance of road traffic that allows provision of “door-to-door” service and owing to the existence of Railway Industrial Track the railways has covered the method of providing full transport “door-to-door” service and is thus trying to keep its place on the market.
The Railway Industrial Track is very important both for the users and for the railways. This is necessary in order to strengthen the competitiveness on the transport market and to return to the leading position in land freight transport which objectively does belong to the railways because of numerous advantages. The significance of Railway Industrial Track can be observed from the users’ aspect and from the railway operator’s aspect.

The advantages of Railway Industrial Track to users:

- Direct delivery of wagon consignments to the place of production activities within the company,
- Shorter time of goods handling,
- Minimal damaging of goods,
- Avoiding of significant handling costs,
- Saving on handling costs by using adequate proper machinery, and
- Total lower transport costs due to the selection of railways as cost-affordable land operator.

Advantages of Railway Industrial Track for the railways:

- Possibility of starting direct trains,
- Disburdening of station handling tracks,
- Shorter time of wagon turnover,
- Direct connection of internal and international railway transport, and
- Formation of a solid and permanent contact with the big users.

Owing to the advantages of Railway Industrial Track, habits are created of using railway transport, and this is the main precondition for the survival on the transport market.

**5. Technological Process at Railway Industrial Track**

The technological process represents a recipe how to successfully create a transport service. Successful transport service can be created if there is excellent technological process. In fact, the technological process contains activities that need to be carried out in order to transport the freight efficiently and effectively. The technological process may be divided into macro and micro process. The macro technological process includes all aspects of freight transport, whereas micro technological process includes activities required to successfully create the transport service at a concrete place of work, for a concrete freight. Macro technological process has the following phases:

1. Research of transport market,
2. Tasks of initial end activities,
3. Activities at railway station,
4. Transport activities, and
5. Activities of organizing and managing the transport process.

The study of transport market is the basic origin of creating the transport service. The most important part of the transport market research is to determine the demand. The demand is the payment capability of the users to use the transport service. The study of demand is a scientifically founded process which can be defined at the level of the railway station, railway line and network in general. The railway station is the origin-end points of the transport process and they are in fact generators of goods for transport. Of special interest here is the Railway Industrial Track which are used by the users traditionally related to railway traffic. In the first place these are ports and quays, industrial complexes and big economy centres. The railway line has its capacity and therefore it is necessary to study the supply – demand relation. Since the transport of goods can be performed from any railway station to any other railway station, then it is necessary to look at the research from the network aspect, in this step the freight flows are of special significance.

The initial activities understand acceptance of goods for transport. Before the very physical acceptance of goods for transport, it is necessary to order the wagon(s) and to apply for the customs, sanitary and phytopathological inspection. After that a consignment note is filled in and the price of the transport service determined. The transport process is followed by the final activities, which means delivery of wagons to Railway Industrial Track, check of the transport price and if necessary, a claim.

At railway station the setting of wagons on Railway Industrial Track is organized and after loading/unloading the wagons are included in the train. Also, if Railway Industrial Track is not fitted with a scale the wagons are weighed. At railway station usually the wagon technical inspection is carried out as well.

In the transport process the wagons may be, as required, assembled or disassembled in trains, their order and composition can be changed. If required, during the transport process the additional freezing can be done, as well as repair of freight and possible removal of buffer wagons.

The management activities are in fact responsible for solving the technological process at a tactical level. This means the development of technical regulations of loading/unloading, scheduling, regulation of work with customs and other bodies, development of transport conditions for special shipments, and human resources management in the field. The term transport process management means that one can detect at any moment the location of a wagon or a shipment, and realize constant connection with the users, informing them about the transport procedure. On the other hand, transport process management is the operative level which accompanies and controls the entire transport process. Today, the user wants to have, and this is possible technologically and technically, the information about the entire flow of the transport process.

The micro technological process regulates: activities, places, contractors and time necessary for successfully performed loading/unloading of goods into and from the wagons. In order to develop this, it is necessary to know the following elements:

1. Opening hours of the railway station and the Railway Industrial Track,
2. Place of handling of individual types of goods,
3. The handling mechanisation means,
4. Setting wagons for loading and unloading,
5. Ratio of company areas with Railway Industrial Track,
6. Regulations of working with goods, and
7. Regulations of work with special goods.

Entering into the company circle, the Railway Industrial Track eliminate the need for indirect participation of road vehicles in delivery, handling and turning out and allow dispatch/acceptance of wagon shipments from the area of the corporate subjects.
Loading or unloading of goods on Railway Industrial Track is performed within the area of the user, in 95% carried out by the user, who in this case may be the sender or the receiver. During loading the authorized railway worker supervises it or in agreement with the operator this is done by the users themselves.

## 6. Case Study: in Croatia

### 6.1. Condition of Railway Industrial Track

Today 870 industrial tracks of total construction length of over 500 km are connected to the railway network in the Republic of Croatia. In 2009 there were 314 basic Railway Industrial Tracks registered on the network, which means that 556 tracks have been abandoned, or as much as 63.91% and because of poor maintenance there is threat of further abandoning of Railway Industrial Tracks. When 146 temporary closed and inactive tracks are taken into consideration, the number of unused basic tracks rises to 702 and the percentage amounts to a high 80.69%. Thus, today in the Republic of Croatia there are 168 basic industrial tracks at which manipulative freight activities are performed. This means that only 53.50% of registered tracks are used. There is also substantial fall in the number of active users whose average annual rate in the last five years amounted to -1.84%, and the share of active users in their total number amounts to 41.89%.
The abandoning of industrial tracks has been caused to a great extent by the changes in economy because of which the existing network of industrial tracks does not match any more the actual customers' needs.
The average static load of wagons on industrial tracks in the Republic of Croatia amounts to 42 tons/wagon and is greater than the one at railway station tracks where it amounts to 31 tons/wagon. Based on this, one could conclude that industrial tracks are in good technical condition. However, this refers mainly to industrial tracks of big customers, whose share in the scope of work of the entire network is about 75%. If 168 active industrial tracks are taken into consideration, then about 150 industrial tracks remain at which the condition is rather poor. Currently, these industrial tracks do not contribute significantly to the railway freight transport as the tracks of big customers, but in no way should they be neglected. At most of them better results would be achieved in better working conditions.
The majority of tracks were constructed 30 years ago and more, when the railways dominated in the freight transport. Then even the tracks constructed in difficult terrain conditions were connected to the network. Today, it is precisely these tracks that present the biggest problem since they require more complex and more expensive maintenance. A group of such industrial tracks features too low permitted axle load which makes it impossible to use maximal capacity of the wagons, thus increasing their number. It is often the case that their value of permitted axle load is not in harmony with that on the connection railway line. At time of the construction, the designed parameters of the railway lines and of the industrial tracks were compatible. After the reconstruction of the railway lines that have been actual on the HŽ network in the last decade, their permitted axle load has been raised, and on industrial tracks it remained at the previous level. This reduces the possibility of maximal usage of wagons and this eventually reflects on the operation of the entire network. The biggest problems occur during the need for reconstruction of single tracks since this refers to high costs that sometimes, apart from the reconstruction of the permanent way and substructure of the tracks include also the reconstruction of structures such as bridges, viaducts, water ducts and cuttings. Their reconstruction requires large financial investments. This refers also to the construction of new tracks. When the potential user realizes how expensive it is to build the tracks, they often give up and orientate themselves to the usage of other modes of transport.
Current and investment maintenance of industrial tracks in the customer's area is entirely financed by their owners who face high costs. The place which represents the demarcation point between the user and HŽ-Infrastructure regarding financing of maintenance costs is the deviation point or derailer which separates the industrial tracks from the connection railway station or railway line. The maintenance of the deviation point or derailer and the respective signalling and safety devices, and the catenary on the electrified tracks is financed by HŽ-Infrastructure Ltd.
The industrial tracks have to be maintained according to adequate standards and technical regulations of the railway sector of the Republic of Croatia. If due to neglect of this requirement (by the owner) the condition of the industrial tracks threatens the safety of traffic and people, HŽ-Infrastructure Ltd has the right to stop the delivery of wagons to the industrial tracks. Regular monthly inspections of the technical condition of all the tracks are carried out by the Section of railway line maintenance from HŽ-Infrastructure Ltd.

### 6.2. Analysis of Survey Questionnaires

The research was carried out with the goal of establishing the need for the modernisation of the existing and the construction of new industrial tracks, with emphasis on the need for construction of industrial tracks in the area of business and free zones that are 2 kilometre distance from railway line. The establishment of the need to build industrial tracks was carried out through a survey of business entities in the subject zones. The questionnaire consisted of 16 questions designed to show the indicators of business entities' operations and the need for the modernisation of the existing and the construction of new industrial tracks. Since a large number of business entities did not respond to the e-mailed survey questionnaire and even after a telephone conversation, field research was further carried out. Field research was carried out in cooperation with HŽ Cargo Ltd.

The survey has included 518 business entities operating in business and free zones, of which 134 responded to the survey, the respond ratio was 26%. The highest number of survey participants, 75% even, deals in production of finished goods and the highest number of business entities, 32% of them, use up over 25 tons of cargo per week. In the survey was established that 86% of participants use solid matter. When we look transport units 19% of business entities use containers for packing their products, 80% use palettes, 19% of them transport bulk freight and 47% transport single items. On figure 3 is shown which transport modes use business entities for goods delivery.

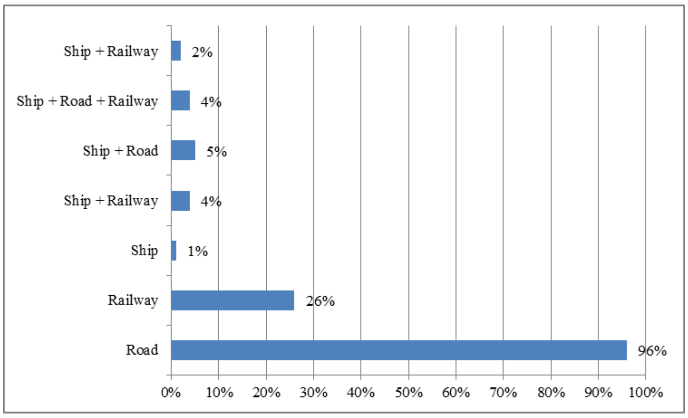



**Fig. 3.**
*Transport modes which use business entities for goods delivery*
*Source: Authors*

To the question "Who performs the reloading?" 90% of business entities answered that they are reloading the goods themselves, for 48% of them it is performed by the operator and for 7% by a broker. Response analysis established that the business entities performing the goods' reloading themselves are those which transport the goods on palettes and in trucks. Reloading of goods is performed by the operators mostly for those entities operating in wood industry. To the question "Who organises the transport?" 82% of business entities answered that their own logistics organises the transport and for 48% of them it is organised by a freight forwarder/logistician. The result of answers to this question is interesting both for the operators and the companies dealing in logistics in the Republic of Croatia, considering that a large number of business entities is organising the transport by themselves. On figure 4 is shown which transport distance for goods delivery.

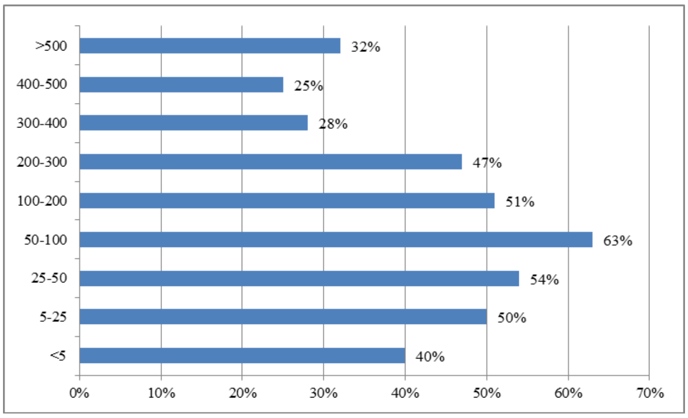



**Fig. 4.**
*Transport distances for goods delivery*
*Source: Authors*

The analysis showed that the survey participants, 34% of them, believe that construction and use of an industrial track would help the development of their company. The survey questionnaire further included the question “Your opinion about the railway transport and proposal for better cooperation?” Herein, the persons filling out the questionnaire could indicate their opinion about the railway and their proposals for better cooperation between the railway operator and their companies. It can be concluded from the responses that the majority of business entities operating in the zones are not satisfied with the services provided to them by the railway. They believe that the service quality does not correspond with the price. A part of the business entities had a bad experience in transporting goods by railway and does not want to use the services of railway transport. Some say that they were forced to stop transporting their products by railway because their customer failed to receive the goods on time. Other part of survey participants believes that an industrial track within the area of their company would benefit the company's development, because they are transporting cargo by railway and the extra reloading at the railway yard incurs costs. They would invest part of the funds in construction of an industrial track because they are aware the investment would be worthwhile in the future.

## 7. Conclusion

Supply Chain Management is an integrated approach to planning, implementing and controlling the flow of information, materials and services from raw material and component suppliers through the manufacturing of the finished product for ultimate distribution to the end customer. Meaning of Supply Chain is that in certain point of time and space provide customers a certain amount of goods. The most important task is to optimize Supply Chain costs from supplier to customers while matching the demand and the supply. Because of that we can talk today that is no longer competition between companies but competition is between Supply Chains. Standard Supply Chain has: (1) supplier, (2) manufacturer, (3) distributor, (4) and retailers (5) customers. In order that elements of Supply Chain were interconnected transport is needed. Therefore transport has an important role stabilizing Supply Chain.

The Last Mile is a common distribution problem where near the destination a high capacity long distance transport system is facing high costs and complexity. This is mainly related to the need to break down the size of the transport unit (less economies of scale) and where congestion impairs reliability as well as capacity.

There are three essential factors for choosing the transport mode: (1) the quantity, (2) the time and (3) price. Their mutual mix will give the transport mode selection. Of course, each transport mode has its advantages and disadvantages. Supply Chain Manager must have decision to choose a transport mode and then operator for having transport services. In the process of transport in the Supply Chain there is needed to transport large quantities of goods in a relatively short time. In land transport choice should fall on the railways. Railway in Europe, in the last few decades, had experience of the challenges of liberalization and for that reason must necessarily be closer to the user. Very effective way of approaching the customer's is use of Industrial Railway Track.

The management on the Railway Industrial Track can be improved in several ways that depend on the nature of the problems of individual tracks. Generally, one could say that the basic assumption of good organization of work is the rationalization of the technological process which has to insure timely delivery and turning out of the wagons from the users’ areas. In fact, based on the management on the Railway Industrial Track there are macro, as strategic and tactical and micro, as operative technological processes. In producing them all the stakeholders should participate, which means in this case the customer, operator, and railway infrastructure manager. In this way a win-win position is realized for all the transport service participants.

The analysis of Railway Industrial Track with respect to their technical condition and ownership of the same, as well as the analysis of part of the issues related to the construction of new Railway Industrial Track, suggests that it is necessary to consider several possible proposals on the manner of encouraging the modernisation and the construction of Railway Industrial Track.

The conducted research shows that Railway Industrial Track are a logical extension of the railway infrastructure, as well as an important part of the railway traffic system and the traffic system as a whole and can be use efficiently and effectively in Supply Chain. A higher quality railway transport system improves the economic picture of a particular region, saves money for transport infrastructure, reduces energy consumption, protects the environment and increases traffic safety. On account of all this, the Railway Industrial Track issue should be approached very seriously and much more dedication and money should be invested into solving the issue, all with the goal of finding a quality solution for the modernisation, the construction and the management of Railway Industrial Track.

Having all advantages in mind for the Railway Industrial Track they can be excellent solution of last mile issue and of course one of necessary chain in Supply Chain Management.

**References**

Abramović, B., Karatur, D., Lacković, Z. 2013. Modelling train cost in HZ Cargo d.o.o. In Proceedings of the 21nd International Symposium EURO - Zel 2014, 6-13

Abramović, B., Lovrić, I., Stupalo, V. 2012. Analysis of intermodal terminals service quality in the Republic of Croatia, Promet- Traffic & Transportation 4(24): 253-260

Abramović, B.; Tuškanec, M. 2010: Industrial tracks used for freight rail transport. In Proceedings of 13th International Conference on Traffic Science (ICTS), CD.

Ballou, R. H. 2003: *Business Logistics: Supply Chain Management*, Prentice Hall, New Jersey, 816

Berndet, T. 2001: Eisenbahngüterverkehr (Railway Freight Traffic), Teubner, Stuttgart - Leibzig - Wiesbaden, 296

Bogović, B. 2006: Prijevozi u željezničkom prometu (Transport in Railway Traffic), University of Zagreb, Faculty of Transport and Traffic Sciences, Zagreb, 280

Brnjac, N.; Abramović, B., Štefičar, S.; Sučić, M. 2013: Importance of industrial tracks in the Republic of Croatia. In *XV Riunione Scientifica della Società Italiana di Economia dei Trasporti e della Logistica (SIET)*, CD

Rodrigue, J-P.; Comtois, C.; Slack, B. 2013: *The Geography of Transport Systems*, Routledge, New York, 432

Heyl, J.; Heizer, J.; Render, B. 2006: *Principles of Operations Management*, Pearson - Prentice Hall, New Jersey

Supply Chain Operations Reference-model (SCOR). Available from Internet: <https://supply-chain.org/our-frameworks>

# THE ADVANCED PREDICTION METHOD OF INVENTORY COSTS IN COAL MINE INDUSTRY

**Paweł Więcek [1], Dariusz Grzesica[2], Daniel Kubek[3]**
[1,2,3] *Institute of Building and Transport Management, Transport Section, Faculty of Civil Engineering, Cracow University of Technology, Cracow*

**Abstract:** The growing needs in terms of providing entities competitiveness on an energetic market force them to seek a new solution. It can help to increase the effectiveness of their activity in which the key element is an optimal inventory policy. The inventory cost determines a reasonable management of enterprise finance and also affects product availability for customers, which is relevant in maintaining appropriate level of service. Such costs are usually influenced by many factors and relations between them. Therefore, it is difficult to obtain an accurate prediction results by the use of classic approaches. The article presents an advanced prediction method of inventory costs in order to determine a proper financial policy and to find opportunities of their reduction also. For the preparation of costs forecast an auto regression and moving averages models have been applied. The advantage of the proposed approach is high accuracy and possibility of the use with a limited data set that are available.

**Keywords**: inventory management, prediction methods, logistics

## 1. Introduction

One of the crucial aims of logistics is to maintain as low as possible inventory level in the warehouses and others nodes of logistics networks. Zero inventory is the ideal state in the inventory management, it can not only reduce costs but also promote the development of the new products (Li at al., 2008). To achieve this goal, the availability of many data and information about economic processes are strongly required. They can indeed affect the company to a big extent. In general information about the past of process and its future are both necessary for enterprise (retrospective and prospective data). Therefore, in a field of logistics there is a strong need to perform forecasts of demand, costs, inventories or customer preferences etc. The purpose of this predictions is to mitigate uncertainty level in processes of planning and decision making. Unrecognized future situation and trends may act adversely to the company in terms of economic efficiency and customer level of service. Decisions in the area of inventories level control and their maintenance costs can bring expected benefits only on the basis of reliable prediction models. In order to maintain the proper level of inventories the appropriate methods of forecasting based on time series are commonly used. Many of them are based on regression method. The regression analysis forecast is one quantitative analysis method in the commonly used economic forecast. During the observation and analysis of economy development history and present situation, the accurate mathematical model which reflect relations in processes is need to be established. Then will calculate and forecast the change of dependent variable according to the change of independent variable in the future, thus will carry on the forecast to the trend of future economy development (Bin, Cang, 2009) The most popular method applied in forecasts of inventory level is exponential smoothing. According to current logistics forecast research both at home and abroad, regression analysis and exponential smoothing are of easy computation but unsatisfactory computational accuracy (Wang, 2009). In (Kim and Ryan, 2003), a simple exponential smoothing forecasting techniques is also used to quantify the cost impact in a simple supply chain. The more advanced prediction method based on the relationship between the elements of the time series are autoregressive and integrated moving average models (ARIMA). Thanks to them, there is no need to study the external factors affected the level of stocks in mining companies because ARIMA models recognize the internal structure of analyzed time series. It can allow to predict future values with higher accuracy.

## 2. Inventory costs

Of the many types of logistics costs the storage costs have significant importance on the overall financial condition of enterprises. Irrational inventory management policies can contribute to a significant increase of inventory costs. Basically we can distinguish costs associated with inventory creation, maintenance and shortage. These groups also contain the cost incurred as a result of delayed delivery, the penalties for lack of products, the lost sales opportunities cost and costs of product obsolescence, etc. Very often, the level of inventory costs incurred in the enterprise are determined by the impact of external factors (prices of external logistics services including warehousing, transportation, services, interest rates on loans to finance inventories) (Krittanathipa et al., 2013). In practice, it is rare that the businesses consider inventory costs separately with division into their individual components. Therefore, keeping accurate forecasts of inventory costs allows to continuously improvement of logistics processes and making more effective decisions in this field.

[1] Corresponding author: dkubek@pk.edu.pl

### 3. Data collection and analysis

Because of its random nature most of the surrounding us business processes are considered as a stochastic processes. Stochastic process is associated with the concept of a random function $Y(t, \xi)$ where: $t, t_i \epsilon T, i = 1,2, \dots, n ; \xi \epsilon \Omega$. Set T is a set of real numbers while set Ω is an elementary events space. If we assume that *T* describes passage of time on the axis, then a random function Y (t, ξ) can be called a stochastic process. It is said that the time series used to build forecasts are realizations of the stochastic process. The time series shows the course of the process in one of the many possible cases.

ARIMA Models (Auto - Regressive Integrated Moving Average) are advanced forecasting tools, suitable for modeling processes characterized by high dynamics and the difficulty in identifying relations which are shaping course of the process. ARIMA models are based on autocorrelation phenomenon what means the correlation of variable values forecast with the values of the same variable delayed in time. The main characteristic feature of ARIMA models is the fact that the value of the predicted variable at time *t* is a linear combination of values of the same variable from previous periods *t - 1, t -2, ..., t - p* increased by a certain value of the random component. Among such models we can distinguish three basic types:

- autoregressive models (AR)
- moving average models (MA)
- mixed autoregressive and moving average models (ARMA).

In a general way autoregressive model of order p can be represented as follows:

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + e_t \qquad (1)$$

where:
$Y_t$, $Y_{t-1}$, $Y_{t-2}$, $Y_{t-p}$ - the value of the variable, respectively, at the time t, t - 1, t - 2, t - p,
$\varphi_0$, $\varphi_1$, $\varphi_2$, $\varphi_p$ – model parameters,
$e_t$ – value of random component at the time period *t*,
*p* - order lag.

Delay parameter *p* determines how far back in time should we reach to determine the value of the forecasted variable at time *t*. If the random components of the past are correlated we have to deal with the process of moving average (MA), which is expressed by the formula:

$$Y_t = \vartheta_0 - \vartheta_1 e_{t-1} - \vartheta_2 e_{t-2} - \cdots - \vartheta_q e_{t-q} + e_t \qquad (2)$$

where:
$e_t$, $e_{t-1}$, $e_{t-2}$, $e_{t-q}$ – residuals of the model respectively at the time period *t, t - 1, t - 2, t - q*,
$\vartheta_0$, $\vartheta_1$, $\vartheta_2$, $\vartheta_q$ – model parameters,
*q* – order lag.

In order to better adaptation of the model to the historical data the connection of (AR) and (MA) parts of the model is often made. It is known as the ARMA model having both a p and q parameters. This combination may be represented as follows:

$$Y_t = \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \cdots + \varphi_p Y_{t-p} + \mathrm{e_t} - \vartheta_0 - \vartheta_1 e_{t-1} - \vartheta_2 e_{t-2} - \cdots - \vartheta_q e_{t-q} \qquad (3)$$

In this model, the predicted value of the variable in a particular time period *t* depends on the values of the past, as well as residues of the model in the previous time instants.

To apply ARIMA models the stationarity condition of the examined time series representing economic process is required. By stationarity, we can understand the constant time variance and the expected value of the process. To investigate the stationarity of inventory costs data from energy company the extended statistical Dickey Fuller test has been applied. Otherwise it is known as a unit root test, taking into account possible autocorrelation of the random component. The null hypothesis (H0) assumes lack of stationarity of examined variable and the alternative hypothesis (H1) indicates that the variable is stationary. Null hypothesis is rejected if the calculated value of the ADF test statistic is less than the critical value read from the appropriate tables for the adopted level of significance. Detailed description of test can be found in (Charemza, Deadman, 1997), (Dickey, Fuller, 1979).

When the stationarity condition is not satisfied the examined time series should be reduced to stationarity by calculating the following differences: $\Delta y_t = y_t - y_{t-1}$ - the first differences, $\Delta^2 y_t = \Delta y_t - \Delta y_{t-1}$ - the second differences. If the calculation of first differences is sufficient to achieve stationary states it is said that the series are integrated of order one I (1). Fig. 1 presents a summary of the monthly inventory cost in the company of the energy sector in the years 2004-2010.

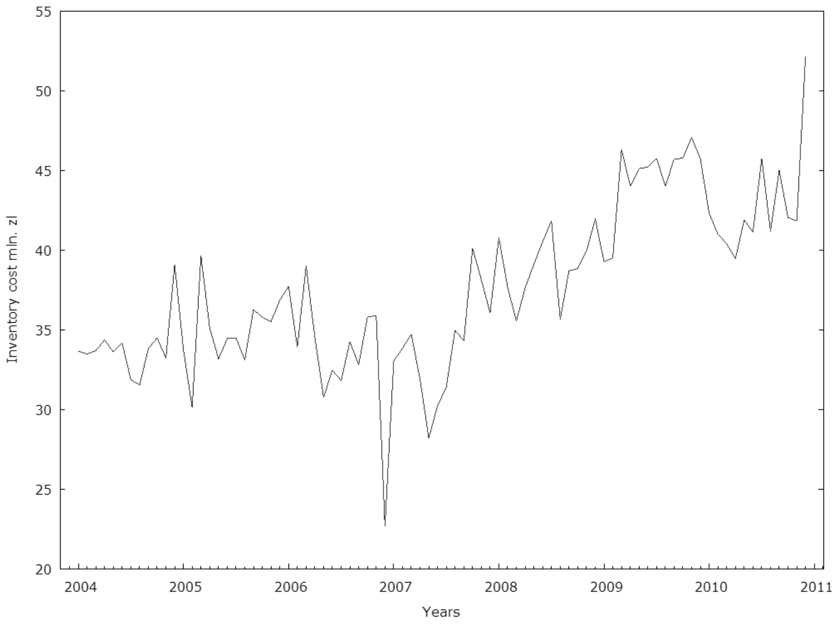



**Fig. 1.**
*Inventory costs between 2004 and 2010*

For the above data the extended stationary Dickey Fuller test was carried out to determine the degree of integration of the series. The inference was made for the significance level $\alpha = 0.05$. In this case, the p-value (talking about the level of significance for which the null hypothesis can be rejected) at a level less than 0.05 indicates stationarity of the series. In this example, the p value was 0.63 so that the series are nonstationary. Therefore it should be reduced to stationarity by the operation of differentiation. In addition, in order to stabilize the variance the values of time series have been squared . The specified in this way time series is shown in Fig. 2

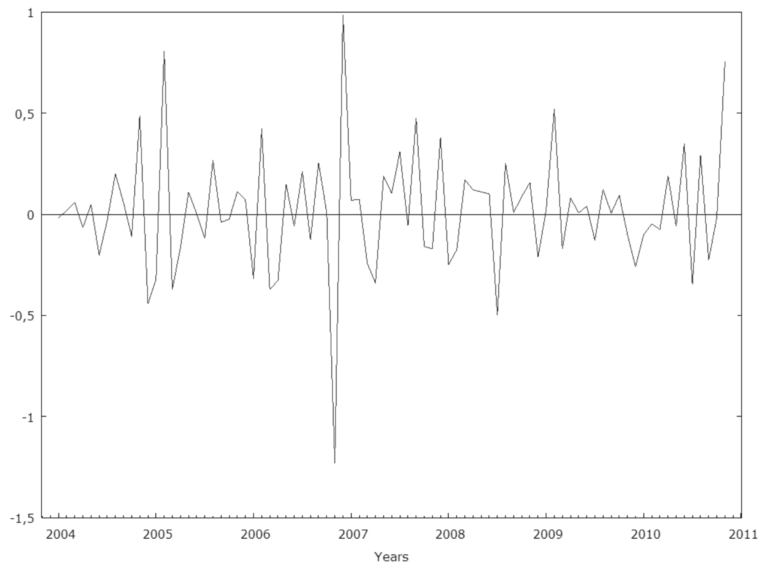



**Fig. 2.**
*The time series after squered and differentiation operation*

Thus obtained series was again subjected to a statistical unit root test. Obtained p-value equal to $4{,}33 \cdot 10^{-22}$ is less than 0.05. Therefore, after applying the operation of differentiation the considered time series was reduced to stationarity and it is integrated of order one I(1).

**4. Model parameters**

The initial step of forecasting by the use of autoregressive methods is the identification process of the ARIMA model. In order to this an analysis of the autocorrelation function (autocorrelation function ACF) and partial autocorrelation (Partial autocorrelation function PACF) of time series is made based on the methodology of Box Jenkins [3] (Fig. 3). According to the methodology of Box and Jenkins, the number of statistically significant parameters of the function PACF suggests Select the number of delays (parameter q) in part of MA ARIMA model. The number of statistically significant parameters of the function ACF suggests to choose the number of delays (parameter p) in part of AR ARIMA model.

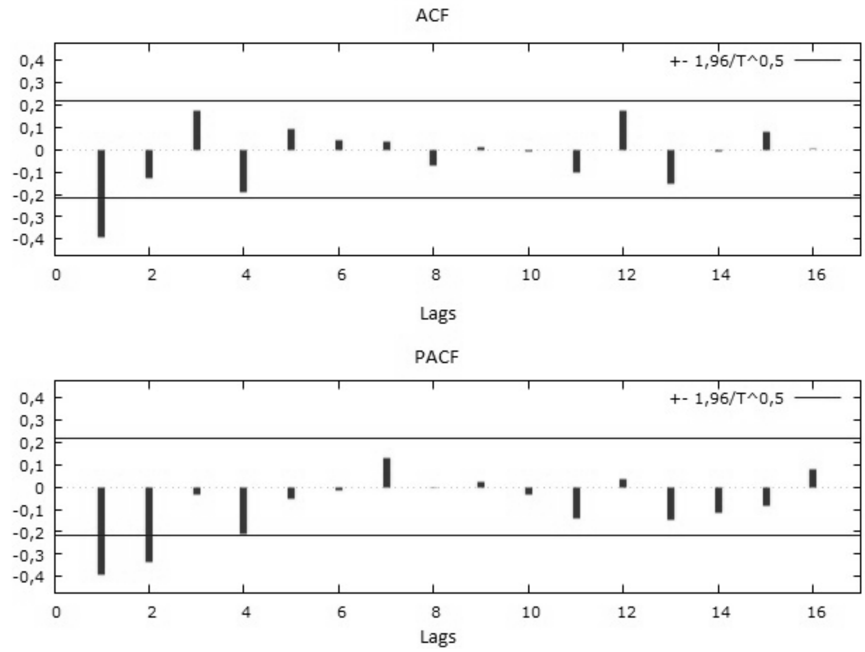



**Fig. 3.**
*Autocorrelation (ACF) and Partial autocorrelation functions (PACF)*

For the autocorrelation function significantly different from zero coefficient is for lag number one, remaining are not significantly different from zero, so you can skip them. In the case of partial autocorrelation function significantly different from zero coefficient are for delays 1 and 2. At the border of significance is a parameter for delay 4. On the basis of previously conducted reasoning can be assumed that the most appropriate model will be ARIMA (1,1,4). Therefore, the selected prediction model of inventory costs for that company can be summarized as follows:

$$Y_t = \varphi_1 Y_{t-1} + \mathrm{e_t} - \vartheta_0 - \vartheta_1 e_{t-1} - \vartheta_2 e_{t-2} - \vartheta_3 e_{t-3} - \vartheta_2 e_{t-4} \quad (4)$$

Estimates of the parameters was carried out by maximum likelihood, which allows to verify the adopted model. The maximum likelihood estimator is the value of the parameter that maximizes the likelihood function; it is often easier to work with the log-likelihood, and this causes no problems because the value that maximizes the likelihood function also minimizes the log-likelihood [8]. To estimate the parameters of the ARIMA (1,1,4) model the data from 79 values of the variable have been used, while the positions from 80 to 82 were used to verify proposed method . Finally, the prediction model can be formulated as follow:

$$\mathrm{Y_t} = -0{,}434 Y_{t-1} + 0{,}175 e_{t-1} + 0{,}416 e_{t-2} - 0{,}104 e_{t-3} + 0{,}051 e_{t-4} \qquad (5)$$

The obtained model can be used to prepare predictions. The original and the adjusted data are shown in Fig. 4.

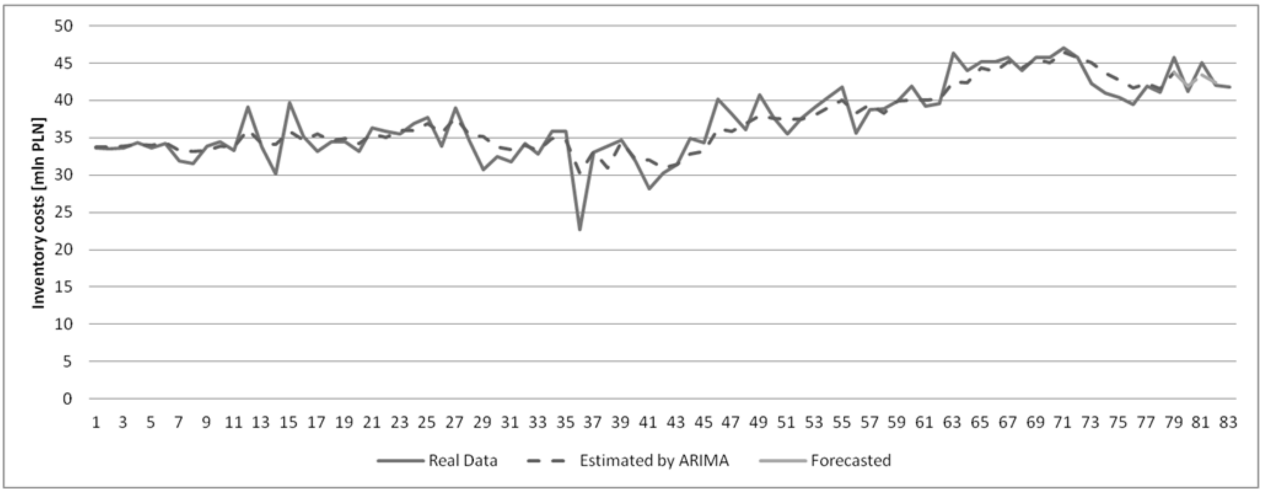



**Fig. 4.**
*The original and predicted values of the inventory costs in the considered company*

In order to verify the correctness of ARIMA model parameters selections the autocorrelation analysis of the model residues were conducted. The lack of significantly different from zero autocorrelation coefficients for subsequent lags indicate that the model parameters were correctly selected. Fig. 5.

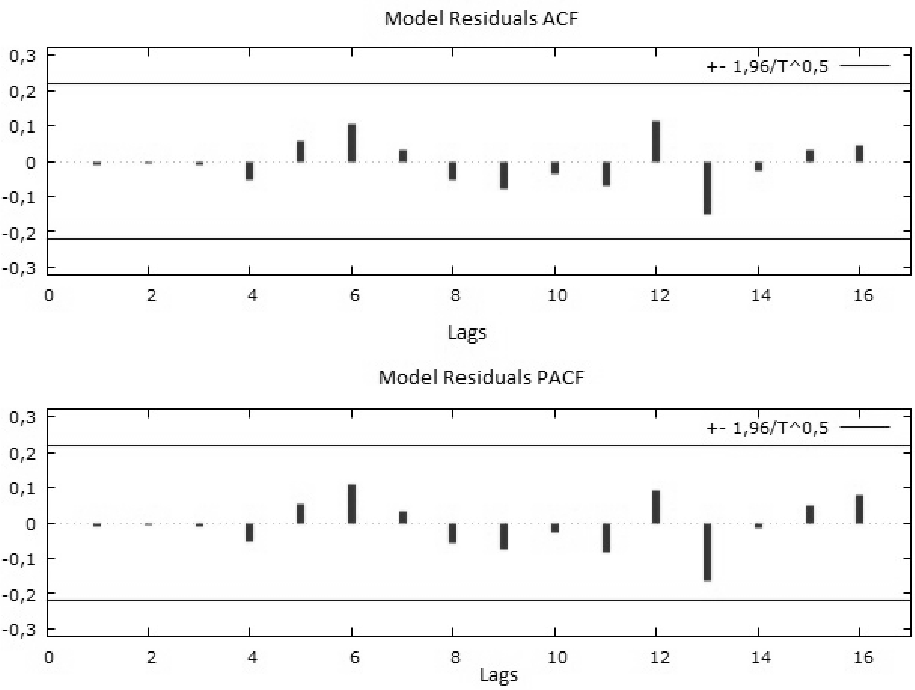



**Fig. 5.**
*Autocorrelation (ACF) and Partial autocorrelation functions (PACF) for residuals of the model*

Despite the relatively good fit of the model to historical data the predictions expired both underestimated and overestimated can be observed. In order to determine the quality of the analysis, the average prediction error has been calculated by the formula:

$$S_p = \sqrt{\frac{1}{m}\sum_{t\in I_{ep}}(y_t - y_t^P)^2} \qquad (6)$$

where:
$y_t$ – original value of the variable
$y_t^P$ – predicted value of the variable
$m$ – number of observations

The calculated average prediction error is equal to 1.866. To determine the effectiveness of prediction model, the relative measure of prediction accuracy ex post was calculated (factor V), assuming normal distribution of the random component.

$$V = \frac{S_p}{y_{t \in I_{ep}}} \qquad (7)$$

where;

$S_p$ - average prediction error (formula 6)
$y_{t \in I_{ep}}$ – the average of predicted data

It allows to compare the obtained forecasts and determines the percentage share of prediction error in real value of forecasted variable. Generally it is assumed that if the factor V satisfies the following conditions:

- $V \leq 3\%$, forecast are very good accuracy,
- $3\% < V \leq 5\%$, forecast are good,
- $5\% < V \leq 10\%$ predictions can be acceptable,
- $10\% < V$, to predictions are unacceptable.

For the numerical example shown in the work the factor V = 4% . Presented fig. 5 shows a lack of autocorrelation between the residues in the modeled data, therefore it can be assumed that the ARIMA model has been correctly estimated and can be used to determine the predictions for the next subsequent time periods in company. In order to evaluate the fit of the model to empirical data the determination coefficient $R^2$ has been calculated. Its value is $R^2$= 0.75 On this basis it can be concluded that the estimated model is relatively well matched to the actual data.

## 5. Conclusions

ARIMA models are the among of the most effective forecasting methods. On the basis of the graph of the autocorrelation and partial autocorrelation functions it is necessary to carry out inference about the proper number of model parameters ($p$ and $q$) , assuming that the considered time series is stationary. ARIMA models can be successfully used to predict the costs in mining companies, enabling quick responses to the market changes. Early information about the size of future storage costs can serve the company to make optimal decisions concerned with the development of strategies in the field of materials and storage. Thus, it is able to provide the availability of products for consumer in the required time. On the basis of this numerical example, it can be concluded that ARIMA models provide high quality of prepared forecasts. The calculated measures of prediction accuracy and high value of determination coefficient fully confirm this thesis. The application of this type advanced forecasting tools allow enterprises and managers to reduce the risk in planning and decision-making processes in the field of logistics and it helps to increase overall effectiveness of companies.

**References**

Bin, X.; Cang, Z. 2009. The Application of Multiple Regression Analysis Forecast in Economical Forecast: The Demand Forecast of Our Country Industry Lavation Machinery in the Year of 2008 and 2009, *Knowledge Discovery and Data Mining, 2009. WKDD 2009. Second International Workshop*: 405-408.

BuHamra, S.; Smaoui, N.; Gabr, M. 2003. The Box-Jenkins analysis and neural networks: prediction and time series modelling. *Applied Mathematical Modelling*: 805-815

Charemza, W.W., Deadman, D.F. (1997): New Econometric. *PWE*, Warsaw.

Dickey, D.A., Fuller, W.A. (1979): Distribution of the Estimators for Autoregressive Time Series with a Unit Root. „*Journal of the American Statistical Association*", No. 74.

Kim, H.K.; Ryan, J.K. 2003. The Cost Impact of Using Simple Forecasting Techniques in a Supply Chain. *Naval Research Logistics*: 50.

Krittanathipa V., Cha-uma S.,, Suwandeeb S., Rakkarnc S., Ratanamaneichatd C.,The Reduction of Inventory and Warehouse Costs for ThaiTraditional Wholesale Businesses of Consumer Products, Procedia - Social and Behavioral Sciences 88 ( 2013 ) 142 – 148

Li, T.; Zhang, S.; Xiong, F. 2008. A Novel Inventory Level Forecast Model Based on Artificial Neural Network, *MultiMedia and Information Technology, 2008. MMIT '08. International Conference*: 82-85.

Montgomery, D. C.; Jennings, C. L.; Kulahci, M. 2008. Introduction to Time Series Analysis and Forecasting. *Wiley*.

Wang, F. 2009. Demand forecast of the Logistic park based on the Curve of growth theory, *Intelligent Computing and Intelligent Systems, 2009. ICIS 2009. IEEE International Conference, Volume: 3*: 57-61.

# CRITICAL FACTORS OF ICT ADOPTION IN KEY LOGISTICS SECTORS: PROPOSED HYPOTHESES AND MODELS

**Vladimir Ilin**[1]

[1] *University of Novi Sad, Faculty of Technical Sciences, Department of Traffic Engineering, Novi Sad, Serbia*

**Abstract:** The main objective of this paper is to propose critical factors of Information and Communication Technologies (ICT) adoption in key logistics sectors. This study proposes three discrete variable models in transportation, warehousing and freight forwarding. The variety of proposed key ICT factors is vast in the literature. Also, many factors are difficult to quantify and measure, such as increased customer's satisfaction. Therefore, the special focus is required when selecting the right criteria as an indicator of ICT adoption in logistics. Technology-Organization-Environment (TOE) framework was used as a starting point of this research. The technological aspect proposes Global Positioning System (GPS), Geographical information systems for transport (GIS–T), Transportation Management System (TMS), barcode, Radio Frequency IDentification (RFID), Warehouse Management System (WMS), Electronic Data Interchange (EDI), Extensible Markup Language (XML), Freight Forwarding System (FFS), Local Area Network (LAN), Wireless Local Area Network (WLAN), Internet, Global System for Mobile communications (GSM) as influenced factors. The organizational aspect proposes Size, Number of Branches (NoB) and Skill of Employees (SoE) as influenced factors. The environmental aspect proposes Collaboration with Partners (CwP), Competitive Pressure (CP) and government pressure (GP) as influenced factors. Logistics regression is suggested as a statistical method to test mentioned variables. This paper offers a new methodology for the research of ICT adoption in key logistics sectors.

**Keywords**: information and communication technologies, logistics, technology-organization-environment, logistics regression.

## 1. Introduction

Information and Communication Technologies (ICT) provide support for logistics processes and solve many logistics problems. Synergistic relationship between ICT and logistics systems enables optimization of manufacturing, procurement, warehousing, transportation and freight forwarding. ICT intrigue evolution and reorganization of logistics processes making them more customers oriented. Increasingly, the traditional way of doing business in logistics without ICT support slowly fades (Ilin and Groznik, 2013; Ilin and Simić, 2012).

Transportation, warehousing and freight forwarding are major logistics sectors and important cost factors in logistics services. ICT often provide support for logistics activities in mentioned sectors, including optimization of processes, increased transparency of material and information flows, enhanced decision-making processes, reduced delivery and lead times, increased high-quality data processing, improved customer oriented services, improved tasks execution, decreased total costs, etc. When integrating ICT solutions into logistics company it is essential to generate a long-term strategy. Therefore, critical factors of ICT adoption in transportation, warehousing and freight forwarding need to be identified.

The main idea of this paper is to design the methodology for determination of critical factors of ICT adoption in transportation, warehousing and freight forwarding. Technology-organization-environment framework was proposed as a starting point of the research. Several steps are identified as important when designing how the new methodology should looks like. Firstly, the survey-based approach for collecting data is suggested. Secondly, hypotheses are formulated and explained. Thirdly, mathematical models for transportation, warehousing and freight forwarding sector are proposed, but not tested. Logistics regression is proposed as statistical method to test discussed variables in created models. Finally, the questionnaire for data collection from all of three mentioned logistics sectors is designed.

## 2. Review of ICT application in transportation, warehousing and freight forwarding

There are many different types of ICT solutions which have application in logistics. This paper targets three logistics sectors: transportation, warehousing and freight forwarding and applicable technologies in each one. There are several crucial sectors in logistics, but according to the Serbian Business Registers Agency (www.apr.gov.rs, 2013) the most of registered logistics companies are oriented towards three previously mentioned logistics services. The focus is sited on Global Positioning System (GPS), Geographical information systems for transport (GIS–T), Transportation Management System (TMS), barcode, Radio Frequency IDentification (RFID), Warehouse Management System (WMS), Electronic Data Interchange (EDI), Extensible Markup Language (XML), Freight Forwarding System (FFS), Local Area Network (LAN), Wireless Local Area Network (WLAN), Internet, Global System for Mobile communications (GSM), respectively. These technologies integrated into logistics processes improve the functioning of logistics systems (Fig. 1). Identified technologies in transportation, warehousing and freight forwarding are explained in Table 1. Benefits of ICT application in logistics systems are pointed out in Table 2.

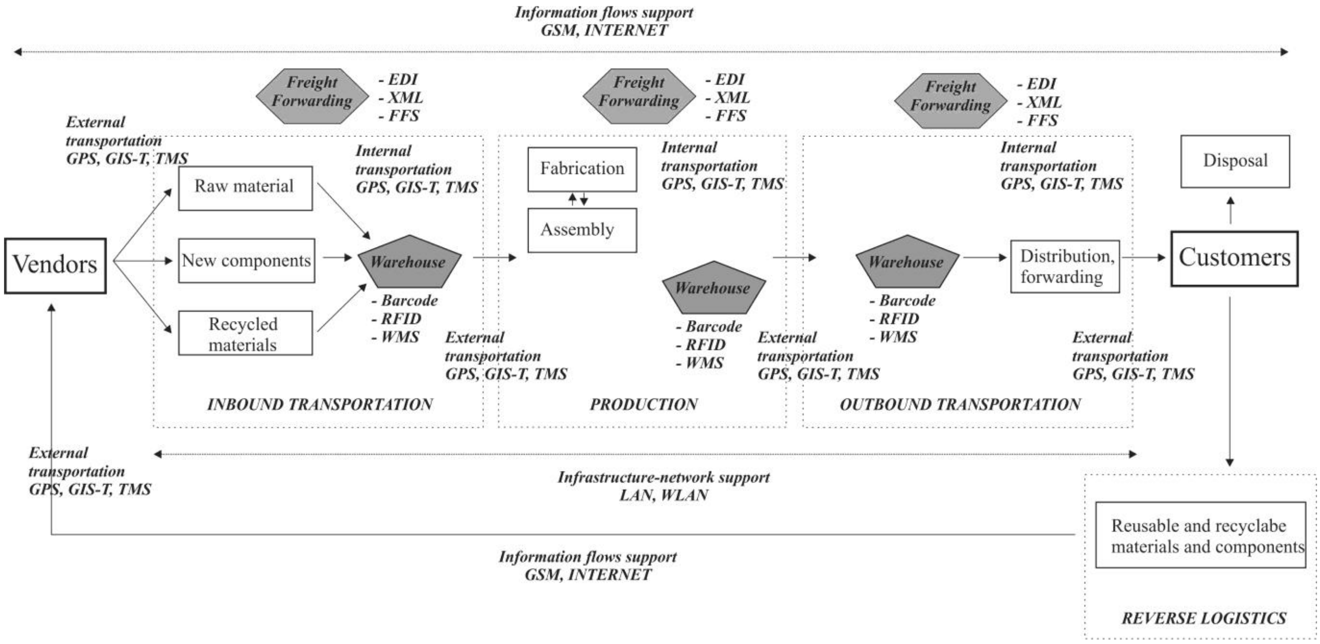



**Fig. 1**
*ICT application in key logistics sectors (Sarkis et al., 2004, edited by author)*

**Table 1**
*Review of ICT application in key logistics sectors*

| | ***LOGISTICS SECTOR*** | | |
|---|---|---|---|
| | ***TRANSPORTATION*** | ***WAREHOUSING*** | ***FREIGHT FORWARDING*** |
| ***ICT*** | 1. GPS is a space-based satellite navigation system that provides information about location and time, anywhere on or near the Earth, where there is an open line of sight to four or more GPS satellites (Wang et al., 2011) | 1. A barcode is an optical representation of data, which shows information about the object to which it is attached (Galo and Manduchi, 2011) | 1. EDI is a technology that enables inter-organizational information and documents exchange among partners without human intervention using structured and agreed-upon national and international message standards (Ngai et al., 2008). |
| | 2. GIS–T are interconnected hardware, software, data, working people, organizations, and institutional arrangements for collecting, storing, and communicating particular types of information about the Earth (Fletcher, 2001). | 2. RFID is one of the most promising, rapidly developing and easy-to-use technologies, which use radio-frequency (RF) signals for automatic identification of objects and items (Amit, 2009). | 2. XML is a toolkit for creating, shaping and using markup languages. It defines a set of rules for encoding documents in a format that is both human-readable and machine-readable (Ray, 2003). |
| | 3. TMS facilitates tasks as transportation planning, performance measurement, control over vehicle loading and management of routes, distances and freight payments (Gilmore and Tompkins, 2000). | 3. WMS manages and optimizes operational and administrative activities along the warehousing process, which involves receiving, inspecting, labeling, storing, sorting, packing, loading, shipping, issuing documents and managing inventory (Banzato, 1988). | 3. FFS refers to software solutions (e.g. CargoWise One) which improve and facilitate electronic documents exchange and optimize freight forwarding operations related with imports, exports and transshipments. |
| | 4. LAN is a computer network that interconnects computers within a limited area such as office building using network media (Behrouz, 2003). | | |
| | 5. WLAN refers to wireless technology that allows an electronic device to transmit data wirelessly (using radio-waves) over a computer network (Roshan and Leary, 2003). | | |
| | 6. GSM is a standard set developed by the European Telecommunications Standards Institute (ETSI) to describe technologies for the second generation (2G) digital cellular networks (Akhila and Kumar, 2009). | | |
| | 7. Internet is a global system of interconnected computer networks that use the standard Internet protocol suite (TCP/IP) to link several billion devices worldwide (Poirier and Bauer, 2001). | | |

**Table 2**
*Benefits of ICT application in key logistics sector* (Ilin et al., 2013, edited by author)

| *ICT systems* | *Technologies* | *Benefits and opportunities in logistics* |
|---|---|---|
| *Transaction system* | *EDI* | Provides exchanging of electronic documents which reduce bureaucracy, streamlining and logistics costs. |
| | *XML* | The same function as EDI, but XML is also a metalanguage for electronic document management and web publishing (Nurmilaakso, 2008). |
| | *INTERNET* | Enables maintaining of connections with partners through e-mail, sharing resources and services, such as invoice and logistics oriented applications (World Wide Web). |
| *Management system* | *TMS* | ERP-T: software solutions in transportation; TMS facilitates tasks as transportation planning, performance measurement, control over vehicle loading and management of routes, distances and freight payments. |
| | *WMS* | ERP-W: software solutions in warehousing; WMS manages and optimizes operational and administrative activities along the warehousing process, which involves receiving, inspecting, labeling, storing, sorting, packing, loading, shipping, issuing documents and managing inventory. |
| | *FFS* | ERP-FF: software solutions in freight forwarding; FFS solutions are mostly used for e-commerce and e-business (e.g. e-procurement). |
| | *LAN* | Provides wired computerized information system including network interface and technical control in wired-supported sections of logistics company. |
| | *WLAN* | Provides wireless computerized information system including network interface and technical control across logistics company. |
| *Control system* | *Barcode, RFID* | Support various logistics activities, such as automatic identification of objects and items, labeling, picking, vehicle loading and unloading. |
| | *GSM* | Supports maintenance of connections between subjects (employees, partners, customers) in logistics processes. |
| | *GPS* | Supports routing and tracking. |
| | *GIS-T* | Supports routing and tracking, but also traffic analysis, travel-forecasting models with land use information and estimation of pollution. |

## 3. Proposed methodology: TOE framework and formulated hypotheses

A review of literature suggests that the technology, organization, and environment (TOE) framework (Tornatsky and Fleischer, 1990) may be a good starting point for analysis of ICT adoption in key logistics sectors. The TOE framework explains adoption of innovations in different industries (Oliveira and Martins, 2010). Therefore, it could be a very useful scheme to identify the most significant technologies in transportation, warehousing and freight forwarding as well as organizational and environmental factors.

The TOE framework is a comprehensive tool for studying the adoption of technology in an organization. It identifies three aspects of an organization that influence the adoption of new technologies including organization, technology, and environment. The technological aspect depicts the technologies that are relevant to the organization in its pursuit of the business objectives. The organizational aspect is defined by several descriptive measures including size and scope, managerial structure and internal resources. The environmental aspect describes the macro area in which an organization conducts the business, with business partners, competitors and the government. Various factors categorized in these three groups are deemed to affect the decision of an organization towards their adoption of latest technologies (Duan et al., 2012).

Proposed hypotheses regarding the technological aspect (Table 3) and the organizational aspect (Table 4) are assumed to be positively related to the adoption of ICT in transportation, warehousing and freight forwarding. There is only one negatively assumed hypothesis regarding the environmental aspect (Table 5).

### 3.1. Technological aspect

**Table 3**
*Proposed hypotheses regarding the technological aspect*

| *Technologies* | *Survey-based methodology* | *Proposed hypotheses* |
|---|---|---|
| GPS | "Does a company use GPS technology for tracking and tracing vehicles?" | H1. GPS technology is positively related to the adoption of ICT in transportation. |
| GIS-T | "Does a company use GIS-T technology for capturing, storing and analyzing data during transportation?" | H2. GIS-T technology is positively related to the adoption of ICT in transportation. |
| TMS | "Does a company use TMS software solutions during transportation?" | H3. TMS software solutions are positively related to the adoption of ICT in transportation. |
| BARCODE | "Does a company use barcode technology during warehouse operations?" | H4. Barcode technology is positively related to the adoption of ICT in warehousing. |
| RFID | "Does a company use RFID technology during warehouse operations?" | H5. RFID technology is positively related to the adoption of ICT in warehousing. |
| WMS | "Does a company use WMS software solutions during warehouse operations?" | H6. WMS software solutions are positively related to the adoption of ICT in warehousing. |
| EDI | "Does a company use EDI-based frameworks to exchange standardized data?" | H7. EDI framework is positively related to the adoption of ICT in freight forwarding. |
| XML | "Does a company use XML-based frameworks to exchange standardized data?" | H8. XML framework is positively related to the adoption of ICT in freight forwarding. |
| FFS | "Does a company use FFS software solutions during freight forwarding?" | H9. FFS software solutions are positively related to the adoption of ICT in freight forwarding. |
| LAN | "Does a company use LAN technology during logistics operations?" | H10. LAN technology is positively related to the adoption of ICT in transportation, warehousing, and freight forwarding. |
| WLAN | "Does a company use WLAN technology during logistics operations?" | H11. WLAN technology is positively related to the adoption of ICT in transportation, warehousing, and freight forwarding. |
| GSM | "Does a company use GSM technology during logistics operations?" | H12. GSM technology is positively related to the adoption of ICT in transportation, warehousing, and freight forwarding. |
| INTERNET | "Does a company use Internet-based frameworks for e-business?" | H13. Internet is positively related to the adoption of ICT in transportation, warehousing, and freight forwarding. |

### 3.2. Organizational aspect

**Table 4**
*Proposed hypotheses regarding the organizational aspect*

| *Organizational factors* | *Survey-based methodology* | *Explanation of formulated hypotheses* | *Proposed hypotheses* |
|---|---|---|---|
| Size | "How many employees are working for the company in the country?" | Mostly, larger companies have multiple levels of bureaucracy and this can impede decision-making processes regarding new ideas and projects (Hitt et al., 1990). Nevertheless, ICT technologies can be very useful tool in larger organization with multiple sectors that have to coordinate simultaneously. In today's business environment, large volume of business dictates more employees and strong ICT support. | H14. Size of the organization is positively related to the adoption of ICT in transportation, warehousing, and freight forwarding. |
| NoB (number of branches) | "How many branches does a "parent" company have in the country?" | The number of branches measures the geographical rather than operational firm scope (Zhu et al., 2003). Still, ICT support can be very useful for maintaining the stability of logistics systems and customer's demands when connecting all branches. | H15. NoB is positively related to the adoption of ICT in transportation, warehousing, and freight forwarding. |
| SoE (skills of employees) | 'What is the percentage of employees with a university degree in a company?'' | High education is very significant precondition when introducing new technology in the company. In logistics industry employees with a college or university degree usually organize daily, weekly and monthly activities and ICT technology can be very helpful tool when planning as well as performing tasks. | H16. SoE is positively related to the adoption of ICT in transportation, warehousing, and freight forwarding. |

### 3.3. Environmental aspect

**Table 5**
*Proposed hypotheses regarding the environmental aspect*

| *Environmental factors* | *Survey-based methodology* | *Explanation of formulated hypotheses* | *Proposed hypotheses* |
|---|---|---|---|
| CwP (collaboration with partners) | "How many partners does a company have in the country?" | The number of partners should indicate at company's strength and tendency for innovation and investment in new technology. | H17. CwP is positively related to the adoption of ICT in transportation, warehousing, and freight forwarding. |
| CP (competitive pressure) | "How high the pressure from the competition is?" | By using an innovation, companies might be able to alter the rules of competition affect the industry structure, and leverage new ways to outperform rivals, than changing the competitive landscape (Oliveira and Martins, 2010). | H18. CP is positively related to the adoption of ICT in transportation, warehousing, and freight forwarding. |
| GP (government pressure) | "How high the pressure from the government is?" | Government pressure may result with the increased level of ICT adoption in companies if there is a strategic plan along with supporting laws which support ICT business. Currently, economic situation in Serbia, including laws, is still nontransparent which discourages domestic and especially foreign companies to make large investments. | H19. GP is negatively related to the adoption of ICT in transportation, warehousing, and freight forwarding. |

## 4. Proposed models in transportation, warehousing and freight forwarding

### 4.1. ICT adoption in transportation

$$P(ADOPTION = 1) = \Lambda(\beta + \sum_{i=1}^{13} \beta_{Xi} \times X_i + \varepsilon)$$

Where Λ(*) stands for the probability density function of the logistic distribution. ADOPTION is a dependent variable, $X_i$ stands for independent variables GPS, GIS-T, TMS, LAN, WLAN, INTERNET, GSM, SIZE, NoB, SoE, CwP, CP and GP, respectively. $\beta_{Xi}$ represents the estimated coefficients between the independent variables and dependent variable, while ε is the measurement error.
The logistic regression evaluates a model with no adoption (ADOPTION=0) and adoption (ADOPTION=1) of ICT technology in the transportation sector.

### 4.2. ICT adoption in warehousing

$$P(ADOPTION = 1) = \Lambda(\beta + \sum_{i=1}^{13} \beta_{Yi} \times Y_i + \varepsilon)$$

Where Λ(*) stands for the probability density function of the logistic distribution. ADOPTION is a dependent variable, $Y_i$ stands for independent variables Barcode, RFID, WMS, LAN, WLAN, INTERNET, GSM, SIZE, NoB, SoE, CwP, CP and GP, respectively. $\beta_{Yi}$ represents the estimated coefficients between the independent variables and dependent variable, while ε is the measurement error.
The logistic regression evaluates a model with no adoption (ADOPTION=0) and adoption (ADOPTION=1) of ICT technology in the warehousing sector.

### 4.3. ICT adoption in freight forwarding

$$P(ADOPTION = 1) = \Lambda(\beta + \sum_{i=1}^{13} \beta_{Zi} \times Z_i + \varepsilon)$$

Where Λ(*) stands for the probability density function of the logistic distribution. ADOPTION is a dependent variable, $Z_i$ stands for independent variables EDI, XML, FFS, LAN, WLAN, INTERNET, GSM, SIZE, NoB, SoE, CwP, CP and GP, respectively. $\beta_{Zi}$ represents the estimated coefficients between the independent variables and dependent variable, while ε is the measurement error.
The logistic regression evaluates a model with no adoption (ADOPTION=0) and adoption (ADOPTION=1) of ICT technology in the freight forwarding sector.

**5. State of the art: Questionnaire**

# QUESTIONNAIRE

**1. Company's profile**

*1.1. Type of services:*
☐ *transportation* ☐ *warehousing* ☐ *freight forwarding*
*1.3. Year of establishment:* ___________
*1.4. Yearly turnover:* __________

**2. Technology aspect**

*Transportation*
*2.1. Does a company use GPS technology for tracking and tracing vehicles?* YES☐ NO☐
*2.2. Does a company use GIS-T technology for capturing, storing and analyzing data during transportation?* YES☐ NO☐
*2.3. Does a company use TMS software solutions during transportation?* YES☐ NO☐
*Warehousing*
*2.1. Does a company use barcode technology during warehouse operations?* YES☐ NO☐
*2.2. Does a company use RFID technology during warehouse operations?* YES☐ NO☐
*2.3. Does a company use WMS software solutions during warehouse operations?* YES☐ NO☐
*Freight forwarding*
*2.1. Does a company use EDI-based frameworks to exchange standardized data?* YES☐ NO☐
*2.2. Does a company use XML-based frameworks to exchange standardized data?* YES☐ NO☐
*2.3. Does a company use FFS software solutions during freight forwarding?* YES☐ NO☐
*(FFS refers to software solutions which improve and facilitate electronic documents exchange and optimize freight forwarding operations related with imports, exports and transshipments.)*
*All sectors*
*2.4. Does a company use LAN technology during logistics operations?* YES☐ NO☐
*2.5. Does a company use WLAN technology during logistics operations?* YES☐ NO☐
*2.6. Does a company use GSM technology during logistics operations?* YES☐ NO☐
*2.7. Does a company use Internet-based frameworks for e-business?* YES☐ NO☐

**3. Organization aspect**

*3.1. How many employees are working for the company in the country?* ___________
*3.2. How many branches does a "parent" company have in the country?* ___________
*3.3. What is the percentage of employees with a university degree in a company?* ___________

**4. Environment aspect**

*4.1. How many partners does a company have in the country?* ____
*4.2. How high the pressure from the competition is?*
*4.3. How high the pressure from the government is?*

strongly disagree ←→ strongly agree
4.2: 0 1 2 3 4
4.3: 0 1 2 3 4

## 6. Final remarks and conclusion

When determining the critical factors of ICT adoption in various logistics sectors it is necessary to identify different ICT solutions relevant to targeted logistics sector. In this paper, TOE framework is proposed as a starting point for the selection of critical factors. The most of the effort is focused towards the technological aspect which proposes some of the ICT solutions that were not previously considered. Proposed ICT solutions in three models are mostly interested in Serbian logistics sector according to author's opinion. All hypotheses regarding the technological aspect are positively related to the ICT adoption in different logistics sectors. Similarly, all hypotheses regarding the organizational aspect are positively related to the ICT adoption in different logistics sectors. The environmental aspect is different from the technological aspect and the organizational aspect regarding hypotheses formulation and positive impact of ICT adoption in logistics sectors. Government pressure is projected to be significant inhibitor of ICT adoption in logistics sector, but thorough research (survey-based methodology) and statistical analysis (logistic regression) will confirm or reject that hypothesis. The organizational and environmental aspects are analyzed before from the similar perspective (Nurmilaakso, 2008; Oliveira and Martins, 2010), but in this paper hypotheses are proposed in accordance with the current Serbian economic situation. It would be very interested to state additional hypotheses for transportation, warehousing, and freight forwarding, respectively. Transportation sector may be interested with the consideration of the migration from GPS technology towards GIS-T technology. Warehousing sector may be interested with the consideration of the migration from barcode technology towards RFID technology. Freight forwarding may be interested with the consideration of the migration from EDI-based framework towards XML-based framework (Oliveira and Martins, 2010).

In each of three proposed models there are thirteen independent variables. In order to test all of thirteen hypotheses sufficient statistical sample is needed. Where is the limit for sufficient statistical sample statistical analysis should imply. However, factor analysis need to be performed before deciding which factors are relevant for the research. Therefore, each model can have less than thirteen independent variables. Also, models can be easily reorganized by selecting some other technologies and by including them into models. Although proposed methodology require thorough research (using designed questionnaire) to collect enough data, future results may show which technologies and organizational and environmental factors are more or less important in transportation, warehousing, and freight forwarding.

## Acknowledgements

This research is supported by the Ministry of Science and Technological Development of the Republic of Serbia, project No. TR 36030.

**References**

Akhila, S; Kumar, S. 2009. Handover in GSM networks. In Proceedings of the 5th International Conference on MEMS NANO, and Smart Systems, 142-145.

Amit, R. 2009. RFID: the next generation auto-id technology, Microwave Journal, 3(52): 58-76.

Banzato, E. 1988. Warehouse Management System, Sistema de Gerenciamento de Armazens. Iman, Sao Paulo. 97 p.
Behrouz, A.F. 2003. Local Area Networks. McGraw-Hill. 618 p.

Duan, X.; Deng, H.; Corbitt, B. 2012. Evaluating the critical determinants for adopting e-market in Australian small-and-medium sized enterprises, Management Research Review, 35(3/4): 2040-8269.

Fletcher, D.R. 2001. Geographic Information Systems for transportation: a look forward. Available from Internet: <http:// onlinepubs.trb.org/onlinepubs/millennium/00047.pdf >.

Galo, O.; Manduchi, R. 2011. Reading 1D barcode with mobile phones using deformable templates. IEEE Transaction on Pattern Analysis and Machine Intelligence, 9(33): 1834-1843.

Gilmore, D.; Tompkins, J. 2000. Transport plays key role in supply strategy, ID Systems, 8: 16-17.

Hitt, M.A.; Hoskisson, R.E.; Ireland, R.D. 1990. Mergers and acquisitions and managerial commitment to innovation in M-form firms, Strategic Management Journal, 11(1): 29-47.

Ilin, V.; Groznik, A. 2013. Leveraging ICT application in logistics, choice or necessity?, Journal of Road and Traffic Engineering, 59(2): 13-18.

Ilin, V.; Simić, D. 2012. Information and communication technologies shaping the intelligent logistics systems. In Proceedings of First International Conference on Traffic and Transport Engineering, 337-343.

Ilin, V.; Veličković, M.; Bašić, A.; Mirčetić, D. 2013. ICT impacts on reverse logistics: framework and opportunities. In III International Symposium Engineering Management and Competitiveness, 451-456.

Ngai, E.W.T.; Lai, K.H.; Cheng, T.C.E. 2008. Logistics information systems: the Hong Kong experience. International Journal of Production Economics, 113(1): 223-34.

Nurmilaakso, J.M. 2008. Adoption of e-business functions and migration from EDI-based to XML-based e-business frameworks in supply chain integration, International Journal of Production Economics, 113(2): 721-733.

Oliveira, T.; Martins, F.M. 2010. Understanding e-business adoption across industries in European countries, Industrial Management and Data, 110(9): 1337-1354.

Poirier, C.C.; Bauer, J.M. 2001. E-supply chain. Berret-Koehler Publishers, Inc. 229 p.

Ray, T.E. 2003. Learning XML. O'Reilly Media, Inc. USA. 387 p.

Roshan, P.; Leary, J. 2003. 802.11 Wireless LAN Fundamentals. Cisco Press, Inc. USA. 312 p.

Sarkis, J.; Meade, L. M.; Talluri, S. 2004. E-logistics and the natural environment. Supply Chain Management: An International Journal, 9(4): 303-312.

The Serbian Business Registers Agency. 2013. Available from Internet: <http://www.apr.gov.rs/>.

Tornatsky, L.; Fleischer, M. 1990. The Process of Technology Innovation. Lexington Books, Lexington, M.A.

Wang, W.; Sun, B.; Zhou, K.; Zhao, X.; Wang J. 2010. The use of GPS for the measurement of details in different areas, In Proceedings of the 3th International Symposium on Electronic Commerce and Security, 59-62.

Zhu, K.; Kraemer, K.; Xu, S. 2003. Electronic business adoption by European firms: A cross-country assessment of the facilitators and inhibitors, European Journal of Information Systems, 12(4): 251-268.

# INTEGRATION OF ROBUST SHORTEST PATH WITH PICKUP AND DELIVERY VEHICLE ROUTING PROBLEM

**Daniel Kubek** [1]
[1] *Institute of Building and Transport Management, Transport Section, Faculty of Civil Engineering, Cracow University of Technology, Cracow*

**Abstract:** The high changeability of vehicle speed in the urban network increases the level of uncertainty as to the estimated travel times. The rush hours, location of road sections, weather conditions or traffic incidents are just some of the factors which influence vehicle speeds. These and other factors are accountable for a significant difference between the estimated speed and the actual value. As a consequence, the designated routes of vehicles based on incorrect and obsolete data may be non-optimal or even infeasible. Such situation may lead to an increase in transportation costs borne by the delivery company operating in an urban area. Moreover, non-optimal routes could create augment costs in the entire transportation system e.g. due to noise and pollutant emission, increased road occupancy or damage to the touristic image of the city. The article presents an approach based on integration of the shortest path problem with the pickup and delivery vehicle routing problem. Through the integration of both methods it is possible to obtain exact solutions to robust vehicle routing problems within a reasonable time. The analysis and the characteristics of the robust model are presented on a real example - the road network of Cracow. Data and characteristics of traffic were obtained from road traffic simulator Aimsun 8.0 and the optimization was performed using CPLEX 12.5 environment.

**Keywords:** robust optimization, vehicle routing problem, shortest path problem

## 1. Introduction

The functioning of modern cities is inseparably related to the functioning of transport. A dynamic development of cities leads to an equally dynamic uncontrolled development of transport. This development generates numerous problems involving the inhabitants, e.g. crowdedness, air pollution or noise. Presently, the problems of the functioning of cities are closely related to the application of modern tools and methods as well as to an economic practice known as city logistics. The article focuses on one of the aspects of city logistics, namely on urban cargo transport. The aim of city logistics in the scope of cargo transport is to coordinate the flow of goods in the system. Further in the article the concept of city logistics vehicles will be understood as all vehicles rendering transport services within the city in the scope of the city logistics tasks.
One of the most significant urban problems is increased traffic in urban areas, which makes it necessary to allow for longer delivery times while planning distribution. The delivery times are frequently determined based on average moving speed for individual sections, which are not known precisely at the planning stage or which data source is unreliable due to e.g. a measurement or estimation error. Moreover, a distinctive characteristic of urban traffic is a high changeability of speed within the city road network. There are numerous factors that affect the length of time needed to traverse a given section of a road, such as rush hours, location of a road section, weather conditions, random incidents or traffic incidents. Those and other factors make speed estimations for the given sections different from the actual value and for this reason vehicle routing based on such data may prove non-optimal or infeasible. Such changes may result in an increase of costs for a company that performs services in the scope of goods transport for a given area (city) as well as in an increase in costs for the entire transport system (Adamski A., et al. 2014). This includes e.g. external costs of transport, such as flue gas emissions, emitted noise, increased road occupancy, or city image issues. This article presents the problems of city logistics vehicle routing problem with consideration to the aspects of reverse logistics and to data uncertainty levels, e.g. data on average speeds over individual sections of the city road network. The routes for city logistics vehicles were obtained based on the robust shortest path problem (RSPP) integrated with the pickup and delivery vehicle routing problem. The presented approach aims to minimise the above-mentioned problems. The approach has been presented based on the actual road network of the city of Cracow.

## 2. Proposed framework

Modelling of actual systems with the application of discrete optimisation in solving vehicle routing problems frequently assumes that:
– costs are fixed (deterministic models) (Karaşan O. E., et al. 2001),
– costs are changeable in time (dynamic models) (Dellaert N., et al. 2013),
– costs are assumed with a certain probability (stochastic models) (Cheng J.,2013).

The last approach assumes that costs may be determined with a certain dose of uncertainty. The uncertainty is determined by the probability of the occurrence of a cost value, e.g. the cost of travelling over a section of a road. Stochastic models assume that uncertainty (here the probability of the cost occurrence) happens in accordance with an assumed probability distribution, which is known beforehand.

[1] Corresponding author: dkubek@pk.edu.pl

The latter condition is not entirely correct as there is not a 100 percent certainty that the assumed probability distribution of costs exactly corresponds to the actual change in the road network sections. Moreover, to obtain a reliable probability distribution it is necessary to perform numerous observations of the object, and still the distribution will remain an approximation of reality (Ben-Tal A., et al. 2009).
This inconvenience is dealt with by the fourth approach to the problems, i.e. by robust optimization. It assumes that some data may change, yet the changes are not defined by a specific probability distribution but by a family of distributions. An additional advantage of this model in comparison with the stochastic approach is a lower demand for computing capacities (Bertsimas D., et al. 2003).

### 2.1. Research review

Mathematical modelling of actual systems frequently encounters difficulties related to data uncertainty, its disruption or incompleteness or measurement error (Stavros A. Zenios, 2009). The first publication related to the problems of data uncertainty in mathematical programming appeared at the beginning of the 70s, where the model was aimed at finding the worst-case solutions (Soysters A.I., 1973). More publications on robust modelling followed as late as in the 1990s and since then the popularity of the issue has been increasing (Gabrela V., et al. 2014).
According to the author's knowledge, scientific studies on robust mathematical models describing the problem of vehicle routing so far have mainly concentrated on the description of the uncertainty related to the customer demands (Moghaddam B.F., et al. 2012), or on the creation of models robust to data uncertainty by applying fuzzy logic (Peng J., et al. 2006). In the paper (Agra A., 2013) an attempt was made to create a robust model of the routing problem with time windows. This approach concerned sea transport, where the main assumption was that there exist all possible connections within a certain set of sea ports (customers). Transfer of the assumption to the urban transport system environment is not possible as a connection between any two points in a road network is a set of some individual road sections. A road network for the sake of routing problems is expressed as a directed graph consisting of a set of points (intersections) and a set of arcs (road sections). The grid of connections between the points is frequently expressed as a sparse zero-one matrix. Most routing problem models assume that the incidence matrix is a full matrix. The condition is possible to meet by using the shortest path model. Thus, a new full incidence matrix will be obtained where any connection between any two points will be a set of consecutive road sections. However, each section of the network is characterised by its own travel time uncertainty and thus the application of the shortest path problem to calculate travel times would come down to averaging uncertainty.
Some points in the set are characterised by a certain level of demand, whereas the remaining points are traversing points. There are mathematical models (Prins C., et al. 2004) which assume that certain sections in a network may be so called traverse sections thus eliminating the issue of data preparation through the shortest path problem. This approach more precisely reflects the actual road network and the manner in which customer service is performed by a city logistics vehicle. However, a disadvantage of this model while programming is a high level of required computing, which means a longer calculation time. A study (Donati, A.V., et al. 2003) presented integration of a robust model of the shortest path problem with the of time dependent vehicle routing. The above model assumed that each section of the network is characterised by interval travel times and the values are different and dependent for a given number of time intervals. Robust shortest paths were determined based on the interval travelling time and, further, the optimal travel routes. Unfortunately, the assumption of discrete travel times may undermine the FIFO principle, i.e. the vehicle that left earlier in a given time interval should reach point B sooner than a vehicle that left later in a later time interval, in the same route from A to B. The above problems indicate a lack of methods by means of which it is possible to determine robust paths for city logistics vehicles.

### 2.2. Integration of robust shortest path with pickup and delivery vehicle routing problem

The above assumptions make us conclude that the right direction of scientific research is to create effective methods, which would allow to determine such routes for city logistics vehicles that would be robust to possible fluctuations while travelling within a road network. The proposed model structure has been presented in Fig. 1 and may be without much difficulty implemented in the architecture of Intelligent Transportation Systems within the framework of the Advanced City Logistics System proposition (Adamski A., et al. 2014). The approach presented below assumes that by means of appropriate tools it is possible to determine a travel time forecast for a given road network. Each predictive analysis involves a level of uncertainty. Another source of uncertainty is road incidents. Based on a historical analysis of such incidents it is possible to indicate the road sections where the probability of a road incident is higher. For such sections a higher level of data uncertainty must be indicated in the model. Road traffic data obtained in this way will be sent to a traffic simulator, e.g. AIMSUN where it is possible to receive information related to travel times that may occur e.g. on the following day. Based on the data robust paths may be determined between all points, i.e. customers that are to be served the following day. This article focuses on the problem of simultaneous pick-up and delivery of goods from and to customers. Pick-up of goods from customers has been considered, which means inclusion of reverse logistics aspects thanks to which it is possible to additionally reduce the number of city logistics vehicles travelling within the city transportation system.

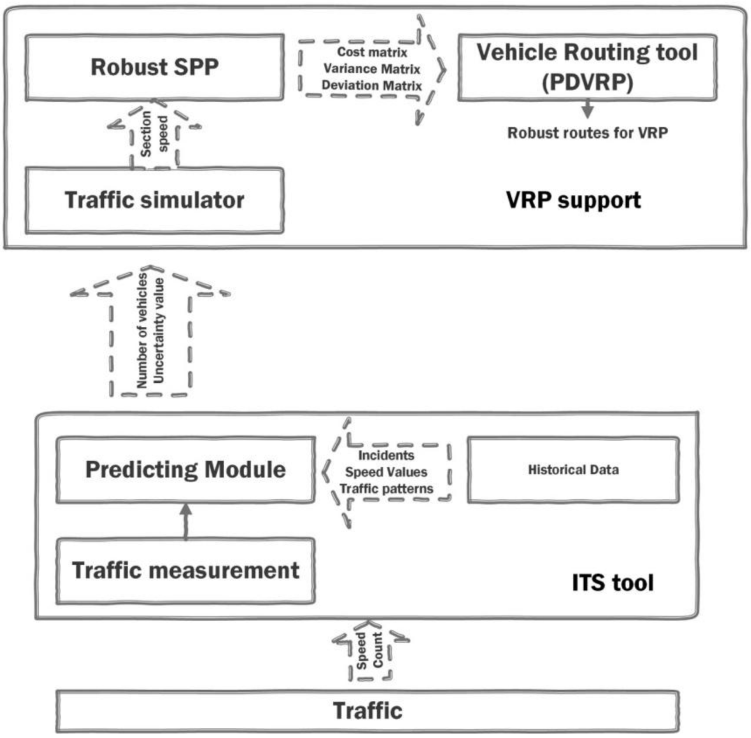



**Fig. 2**
*Proposed framework for Robust Pickup and Delivery Routing Problem.*
*Source: Own elaboration.*

The applied mathematical model of the RSPP problem was presented in the study (Bertsimas D., et al., 2004), and its application in relation to the actual road network is included in the work (Kubek D., 2014a). Robustness of the SPP model is defined by introducing a set of uncertainties, which comprises uncertainty variables. The values of the variables are determined by parameter Γ, which is responsible for the level of the solution conservatism. Practically, it means determining for which number of criterion function cost matrix coefficients there are deviations from the average value. If J indicates the number of cost matrix elements (the number of sections in a network), the Γ parameter will produce values within a range $[0,|J|]$ . For $\Gamma=0$, the model will come down to the deterministic version, whereas if $\Gamma=|J|$, the model will come down to selecting the worst-case scenarios.
Thanks to the determination of robust paths it is possible to integrate the obtained results with the model of simultaneous pickup and delivery vehicle routing problem – hereinafter SPDVRP). Once the problem has been formulated it becomes possible to solve the problems by means of the exact method. Such integration of both models enables to obtain a goods transport schedule which would be robust to possible traffic fluctuations. The analysed case uses the model presented in the work (Subramanian A., et. al.2010). The proposed approach was tested on an actual road network of the city of Cracow.

## 3. Case study

The presented RSPP model was implemented to determine an optimal route in the actual road network of the city of Cracow, with consideration to traffic intensity within the network. Due to the fact that at present there is no access do any actual data related to the current traffic, the travel time changeability characteristic was obtained using traffic simulator Aimsun 8.0. This software enables to perform e.g. a microsimulation of road traffic. Fig. 3 presents the analysed network and its model in Aimsun 8.0 program.
Traffic intensity data at entrances into the network were entered based on the measurement of traffic conducted in 2012 by Infrastructure and Transport Authority in Cracow. The measurement used in the analysis was conducted on a workday during the morning peak time. Using the traffic research and the functionality of the simulator it is possible to obtain changeability characteristics for individual sections of the network. The results obtained from the simulator indicate (Tab. 1) that there are sections in the network where the differentiation of travel times is not high but there are also sections where the travel time changeability is high. It is proved by a high variance and standard deviation of travel time. The model uncertainty for those sections will be higher.

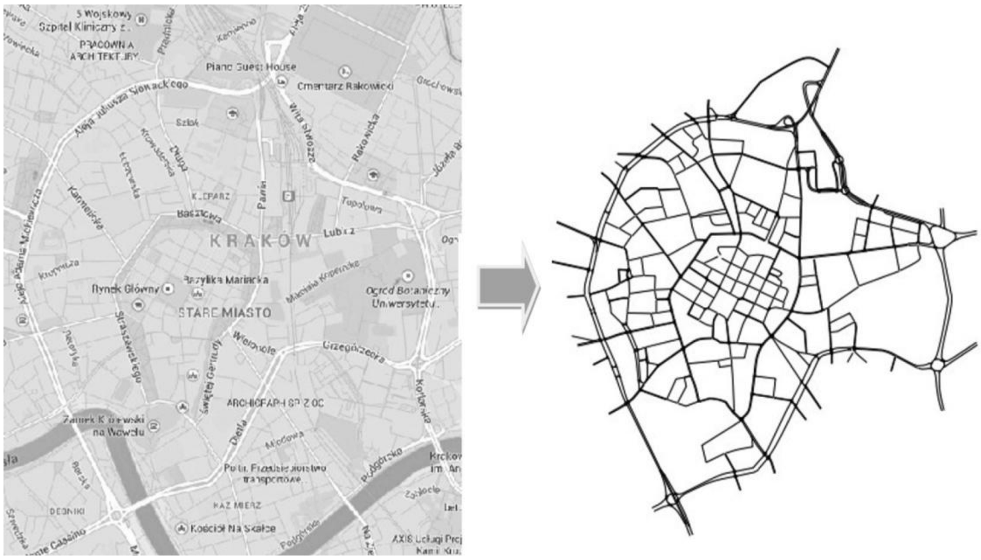



**Fig. 3**
*The analyzed area and its AIMSUN's model- centre of Cracow.*
Source: Own elaboration based on Google Maps

**Tab. 1**
*A part of simulation results from AIMSUN 8.0.*

| ***No. of section in Aimsun*** | ***Traverse time [sec.]*** | | | ***Deviation*** | ***Variance*** | ***Changeability in respect to mean value*** | |
|---|---|---|---|---|---|---|---|
| | ***minimal*** | ***maximum*** | ***mean*** | ***[sec.]*** | ***[-]*** | ***min*** | ***max*** |
| 334 | 12,9141 | 15,8212 | 13,8458 | 0,7071 | 0,5001 | 6,73% | 14,27% |
| 335 | 6,4594 | 7,1783 | 6,8817 | 0,2181 | 0,0476 | 6,14% | 4,31% |
| 336 | 6,4800 | 11,1999 | 7,5323 | 1,2396 | 1,5365 | 13,97% | 48,69% |
| 337 | 3,5750 | 4,5807 | 4,1084 | 0,2852 | 0,0813 | 12,98% | 11,50% |
| 338 | 6,9739 | 9,6968 | 8,9955 | 0,7765 | 0,6030 | 22,47% | 7,80% |
| 339 | 7,7784 | 16,7394 | 11,4883 | 2,2885 | 5,2373 | 32,29% | 45,71% |
| 340 | 9,3962 | 9,9366 | 9,6292 | 0,1644 | 0,0270 | 2,42% | 3,19% |
| 341 | 9,6172 | 10,5652 | 10,2042 | 0,2872 | 0,0825 | 5,75% | 3,54% |
| 343 | 10,0987 | 18,0915 | 12,1300 | 2,2391 | 5,0133 | 16,75% | 49,15% |
| … | … | … | … | … | … | … | … |

*Source: Own elaboration based on simulation results.*

The model of determination of robust routes was implemented in optimisation environment CPLEX 12.5, where the obtained solution was determined using the exact optimisation method (Branch and Bound algorithm). The presented road network was defined as a directed graph consisting of 191 points and 710 directed arcs. Time travel uncertainty was defined as the difference between the minimum and maximum value of travel time. A big difference in the values means that a given section of the road is characterised by high travel time changeability. The level of conservatism was determined as $\Gamma=5$, which means that a given solution will be characterised by a low level of conservatism. The number of potential customers equals 13 and each customer has been assigned a random demand and supply of transported goods within a range [10,100 kg]. The loading capacity of the city logistics vehicles was determined as 3500 kg. The maximum number of vehicles is 2. The optimisation criterion in both variants is minimisation of travel times.
For each possible pair of customers the shortest deterministic paths were defined (based on average values) and the shortest robust paths according to the model (Bertsimas D., et al., 2004). The average calculation time for each pair for the robust model was 1 second. For the sake of comparison it was also calculated what the time between the points would be if the fluctuations in travel time were actually executed and the path were defined based on the deterministic model. The obtained consecutive data were sent to the routing module, where a sequence of visited customers was built. The obtained results are presented in Fig. 4. The customers were marked as square points. The solid line indicates the first route and the dotted line the second route.



**Fig. 4**

*Deterministic solution (left figure) and robust solution (right figure) for simultaneous pickup and delivery vehicle routing problem.*

*Source: Own elaboration based on simulation results.*

As may be seen, not only does the robust solution change the routing from one point to another, but it also changes the sequence of customers and the assignment of a given customer to a vehicle. It is a vital question for companies rendering transport services in a city area, as it determines the level of customer service. If all the assumed deviations from average travel times were executed, the operational time gain for both vehicles would equal 7.79%.

## 4. Conclusions

The presented problems address a crucial question which is the city cargo services. This aspect is significant for the entire transport system of the city and, above all, for the city inhabitants and transport companies. The presented approach shows the advantages of applying robust optimization, thanks to which it is possible to describe the actual object i.e. city traffic in a more detailed manner. It must be pointed out that the analysed example does not consider time windows, i.e. the time at which the customer would prefer the service to be rendered. This fact determines the level of customer service, which is a key factor for transport companies. In view of this, further studies of the proposed model should be aimed at the customers' flexible time windows. Flexible time windows are such in which delays in the service are possible, as well as waiting times before the service commencement. An additional asset of the proposed method is integration of forward logistic and reverse logistic and decomposition of the robust pickup and delivery vehicle routing problem (RPDVRP) into two simpler problems: robust shortest path problem (RSPP) and pickup and delivery vehicle routing problem (PDVRP). Thanks to this approach, it is possible to obtain a solution in a rather reasonable time.

**References**

Adamski A., Kubek D., 2014. *HITS: Advanced City Logistics Systems*. Human Factors of a Global Society: A System of Systems Perspective. CRC Press USA

Agra A., Christiansen M., Figueiredo R., Hvattum L.M., Poss M., Requejo C., 2013. The robustvehicleroutingproblemwithtimewindows. *Computers & Operations Research* 40: 856–866

Aimsun 8 Dynamic Simulators Users' Manual

Ben-Tal A., El Ghaoui, L., Nemirovski, A.S. 2009. *Robust optimization*. Princeton University Press

Bertsimas D., Sim M. 2003. Robust discrete optimization and network flows. *Mathematical Programming* 98: 49-71

Bertsimas D., Sim M. 2004. Price of Robustness. *Operations Research* 52(1): 35-53

Cheng J., 2013. Distributionally robust stochastic shortest path problem, *Electronic Notes in Discrete Mathematics* 41: 511–518

Dellaert N., Woensel T., Kok T. 2013.Dynamic shortest path problems: Hybrid routing policies considering network disruptions, *Computers & Operations Research* 40(12): 2852–2863

Donati, A.V., Montemanni, R., Gambardella, L.M., Rizzoli, A.E., 2003. Integration of a robust shortest path algorithm with a time dependent vehicle routing model and applications. *International Symposium on Computational Intelligence for Measurement Systems and Applications, CIMSA 2003*: 26 - 31

Gabrela V., Murata C., Thiele A. 2014. Recent advances in robust optimization: An overview. *European Journal of Operational Research* 235(3): 471–483

Institute for City Logistics – Uniwersytet Kyoto: <http://www.citylogistics.org/>

Kara ̧san O. E., Pinar M. Ç., Yaman H. 2001. The robust shortest path problem with interval data. *Technical report*, Bilkent University

Kubek D., 2014. Robust shortest path problem in urban environment (Wyznaczanie odpornych ścieżek w warunkach miejskich). *Logistyka* 4: 3965-3971

Moghaddam B.F., Ruiz R, Sadjadi S.J., 2012.Vehicle routing problem with uncertain demands: An advanced particle swarm algorithm. *Computers & Industrial Engineering* 62: 306–317

Peng J., Shang G., Liu H., 2006. A Hybrid Intelligent Algorithm for Vehicle Routing Models with Fuzzy Travel Times. *Computational Intelligence, Lecture Notes in Computer Science* 4114: 965-976

Prins C., Bouchenoua S. 2004. A memetic algorithm solving the VRP, the CARP and general routing problems with nodes, edges and arcs. *Recent Advances in Memetic Algorithms*. Studies in Fuzziness and Soft Computing. Springer: 65–85

Soyster A., 1973. Convex programming with set-inclusive constraints and application to inexact linear programming. *Operation Research* 21: 1154-1157

Stavros A. Zenios.2009. Robust Optimization. *Encyclopedia of Optimization*. Springer US

Subramanian A., Uchoa E., Ochi L.S. 2010. New Lower Bounds for the Vehicle Routing Problem with Simultaneous Pickup and Delivery, *Experimental Algorithms, Lecture Notes in Computer Science* 6049: 276-287

# IMPORTANCE OF MAINTENANCE MANAGEMENT'S EVALUATION ON VEHICLE FLEET ENERGY EFFICIENCY

**Davor Vujanović[1], Vladimir Momčilović[2], Nebojša Bojović[3], Stevo Bunčić[4]**
[1,2,3,4]*University of Belgrade, Faculty of Transport and Traffic Engineering, Vojvode Stepe 305, Belgrade, Serbia*

**Abstract:** In this paper, the importance of evaluating manager's efficiency on fleet energy efficiency increase was presented in the framework of maintenance management. For the fleet operation to become more energy efficient it is necessary to efficiently manage its maintenance. Fleet maintenance management influence both the vehicle maintenance process, as well as the core transport process, but also the environment. In this sense, the effective maintenance management requires an integrated consideration of vehicle maintenance process, transport process and their environment. Since the effects of the implementation of certain measures of maintenance management can be measured by various indicators, the paper defines suitable parameters from the transport process, from the vehicle maintenance process and from the environment. Defined indicators can be used to measure the manager's efficiency. For this purpose, a model of indicators was developed with calculated relative weights. By using the developed model "Overall management score" (S), which shows to what extent the managers were effective in the maintenance management. Based on the results achieved by the model implementation in a company with vehicle fleet, the importance of evaluating managers is highlighted to increase the energy efficiency of the fleet.

**Keywords:** maintenance management, vehicle fleet, performance indicators, managers' evaluation, energy efficiency

## 1. Introduction

The objective of the research in this paper were companies with own vehicle fleets for freight transport that for the supplied transport service acquires profit. The profit in observed companies is highly influenced by transport and maintenance costs, among others. The objective of considered companies is to accomplish all planned transportation tasks with as low as possible transport and maintenance costs. Efficient maintenance management will influence an increase in fleet's energy efficiency and thus the observed decrease in total costs.
Numerous authors believe that in order to efficiently manage the maintenance it is necessary to coordinate and jointly observe the core and the maintenance process. This integrated approach was mainly researched in the field of industrial production, such as energy production plants, paper, chemical products, metal products and so on (Ashayeri et al., 1996), (Brah & Chong, 2004), (Brandolese et al., 1996), (Marquez, 2007), (Marquez & Gupta, 2006), (McKone et al., 2001), (Waeyenbergh & Pintelon, 2002), (Zhu et al., 2002), (Zhu & Pintelon, 2001). In these papers different models and systems were recommended with the objective of increasing efficiency and productivity of industrial equipment and machines.
However, vehicle operation and maintenance processes vary compared to the same processes referring to industrial equipment and machines in the production. Since vehicles are mobile resources which are additionally subject to external environmental factors (e.g. road infrastructure condition, traffic conditions, weather conditions, territorial distribution of specialized workshops in the surroundings, etc.) and which interact with the natural environment there is a need for a different approach in maintenance management, in relation to the machine in the production. On the other hand, the maintenance process impacts the environment in terms of quality of performed maintenance interventions on vehicles. This is reflected in the proper vehicle operation, i.e. vehicle reliability and safety, the harmful exhaust gas emissions, fuel and oil consumption and so on. (Yamamoto et al, 2004).
In this sense, efficient fleet maintenance management is necessary to simultaneously and jointly observe: 1 the transport process as the primary (core) process that makes profit to the enterprise; 2. the vehicle maintenance process, as a logistical support to the transport process, where maintenance interventions restore the "readiness for operation" of vehicles that were in the previous state of "unreadiness for operation"; 3. the environment, that incorporates safety aspects and environmental protection from maintenance impact, as well as external factors such as condition of road infrastructure, traffic conditions, weather conditions and so on.
In this manner defined integrated approach to maintenance management requires the application of the "Process based maintenance" concept, (Zhu et al, 2002), (Zhu & Pintelon, 2001). This concept, among other things, involves defining the indicators required to measure the effects of the application of specific measures in the maintenance management, tracking the indicator values in relation to the adopted limit values and decision-making by managers in the event of unauthorized deviation from the limit values.

[1] Corresponding author: d.vujanovic@sf.bg.ac.rs

In case of several indicators evaluating effects of the implemented improvement measures, it is necessary to determine which of the indicators has a greater importance for achieving a defined company objective. In addition, certain indicators are interdependent, the increase of the value of one indicator may differently affect the value of other indicators. In this sense, the considered area represents a classic example of Multiple Criteria Decision Making - MCDM.

Accordingly, for purposes of measuring the maintenance management efficiency of managers it is necessary to define appropriate indicators, as given in the chapter 2 of this paper. In chapter 3 is presented the methodology for evaluating the maintenance management efficiency of managers. As part of the methodology presented represents a model containing the defined criteria of significance in relation to the company objectives. In addition, this chapter describes the evaluation method of managers. The obtained results from the implementation of the model in a company with its own fleet are presented in chapter 4. Meanwhile, in the chapter 5 are presented the conclusions and future research possibilities.

## 2. Indicators of vehicle fleet maintenance management

As earlier noted, for an efficient fleet maintenance management it is necessary to observe integrally the transport process, the vehicle process and the environment. Since for such a comprehensive approach to maintenance a “Process based maintenance” is suitable, it will be analysed and suitable indicators will be defined in these areas. However, in order to define suitable indicators for maintenance management it is necessary to know the interactions that exist between the transport process, the vehicle maintenance process and their environment (Vujanovic et al, 2012).

Considered companies achieve profit carrying freight, or performing transport services. For this purpose, the company uses its own vehicle fleet usually of heterogeneous structure consisting of commercial vehicles of various construction-operation (CO) groups. On the basis of submitted transportation requests from clients, the executives recognize relevant attributes and sort requests by type, destination and time of delivery of freight and transport tasks form. The list of all planned transport tasks that the fleet should realize in a certain period of time represents an operation plan - OP (Milosavljevic et al, 1996), (Momčilović et al, 2007).

However, during the transport tasks realization the vehicles undergo the process of technical condition degradation and such vehicle will eventually submit different maintenance requests (Figure 1). These requests initiate a necessary maintenance intervention (preventive and corrective) to be realized on vehicles. Time periods when the vehicle is immobilized because of the maintenance interventions realization are periods when the vehicle cannot work, it is “unready for operation” and will not be available for the realization of transport tasks. After the realization of maintenance interventions the technical condition of the vehicle is upgraded and returned to its regular state “ready for operation” and becomes available for the transport process.

The maintenance process interacts with the environment in terms of vehicle safety and environmental protection (harmful exhaust gas emission, old spare parts, old tires, old motor oil and gear oil, etc.). Likewise, the environment interacts with the maintenance process, i.e. submission of maintenance requests because of the bad condition of road infrastructure (potholes), traffic conditions (urban congestion), climatic conditions, etc.

To improve the fleet maintenance management it is necessary to constantly measure the effects achieved by the implementation of certain measures with the help of suitable indicators (Figure 1). In this regard, the review of the literature and experience of the authors define a set of suitable indicators used for evaluating manager’s efficiency in fleet maintenance management (Table 1).

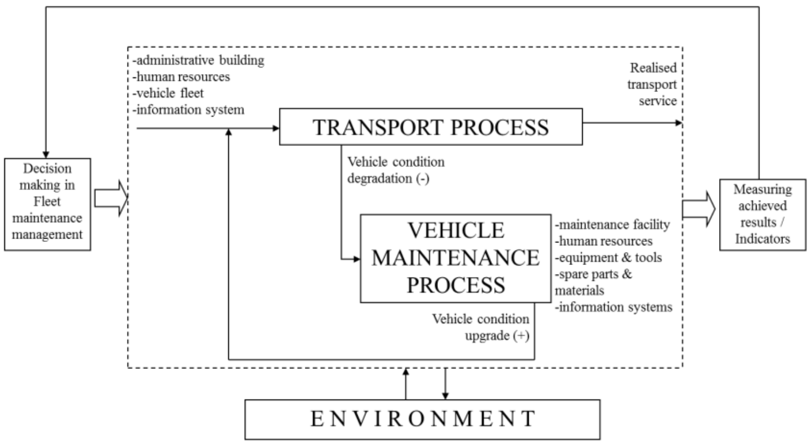



**Figure 1.**
*Interdependence of transport process, maintenance process, and environment*
*source: (Vujanović et al, 2012)*

As suitable indicator in the transport sector many authors (Li et al., 2009), (Milosavljević et al., 1996), (Papić et al., 1999) use the *Operation plan (OP) realisation percentage* (hereinafter referred as T1). The indicator T1 represents the percentage of realised transportation tasks in relation to the total number of planned transport tasks (in ton-kilometres), according to the requested OP in the observed period.
According to many authors (Kamakaté & Schipper, 2009), (McKinnon, 1999), (Ruzzenenti & Basosi, 2009), and (Vujanović et al., 2010) significant increase in vehicle fleet energy efficiency lies in better use of vehicle cargo capacity. In this sense, a suitable indicator affecting fleet energy efficiency within the transport process would be *Vehicle Payload Utilisation* (hereinafter referred as T2). Indicator T2 is the ratio of the carried consignment mass (or volume) and vehicle payload capacity (or available cargo compartment volume) in observed period.
In papers (Haghani & Shafahi, 2002), (Papić et al., 1999) it has been demonstrated how efficient allocation of vehicles to maintenance and application of adequate maintenance strategies influence a decrease of the necessary vehicle fleet for the realization of the OP. It will lead to important savings in energy and material, so as a relevant indicator might be used *Vehicle Fleet Utilisation Rate* (hereinafter referred as T3). Indicator T3 represents the ratio of number of vehicles necessary for OP realisation and total (inventory) number of vehicles in certain period.
According to literature (Mobley et al., 2008), for the analysis of the equipment reliability, an indicator of *Mean Time Between Failures* (MTBF) can be used, hereinafter referred as M1. M1 is calculated from the ratio of vehicle working hours (or kilometres travelled), and number of failures in the observed period. With the improvement of the indicator M1, higher vehicle reliability can be expected. This will equally enhance vehicle safety.
The indicator *Mean Time to Repair* (MTTR) represents a measure of system's maintainability according to (Mobley et al., 2008). MTTR is calculated from the ratio of total labour working hours required to carry out repairs and interventions to the total number of failures in the reporting period (Parida & Kumar, 2009). However, for the fleet maintenance process, a special interest should be given to the indicator of *Mean Vehicle Downtime* (hereinafter referred as M2), i.e. vehicle in the state "unready for operation". This indicator is obtained by the ratio of total hours when vehicle was unavailable for operation and number of failures in the observed period. Unlike indicator MTTR, when calculating indicator M2 it is taken into account the required repair or maintenance intervention preparation time (e.g. waiting for spare parts, free (available) workers, free dedicated workspace in the workshop and so on), as well as eventual time between (to or from) the maintenance workshop and transport company.

**Table 1:**
*Indicators for vehicle fleet maintenance management*
*source: (Vujanović et al, 2012)*

| **Interdependent groups** | **Indicators (factors)** | **Indicator definition** |
|---|---|---|
| **1.Transport process - $T_P$** | Operation Plan (OP) Realisation Percentage - ***T1*** | ***T1*** = (amount of realised tonne-kilometres / amount of planned tonne-kilometres, in OP) × 100 |
| | Vehicle Payload Utilisation - ***T2*** | ***T2*** = consignment mass (volume) / cargo (compartment volume) capacity |
| | Vehicle Fleet Utilisation Rate - ***T3*** | ***T3*** = number of required vehicles for operation / total number of vehicles |
| **2.Maintenance process - $M_P$** | Mean Time Between Failures - ***M1*** | ***M1*** = vehicle realised working hours (or kilometres) / number of failures |
| | Mean Vehicle Downtime - ***M2*** | ***M2*** = vehicle total hours "unready for operation" / number of failures |
| | Maintenance Plan (MP) Realisation - ***M3*** | ***M3*** = number of realised work orders from MP / total number of planned work orders in MP |
| | Planned Maintenance percentage - ***M4*** | ***M4*** = (labour working hours on planned maintenance work orders / total labour working hours) × 100 |
| **3. Environment - E** | Percentage of Fleet Roadworthiness - ***E1*** | ***E1*** = (number of vehicles complying with minimum requirements / total number of vehicles controlled on technical inspections) × 100 |
| | Percentage of Vehicle Roadworthiness in Accidents - ***E2*** | ***E2*** = (number of vehicles in accidents that complied with minimum safety requirements / total number of vehicles in accidents) × 100 |

Maintenance Plan (MP) is a useful tool in the maintenance management (Momčilović et al., 2007). MP is a document used by managers while making decisions on planned workspaces and timeframe to carry out necessary maintenance interventions on vehicles.

As appropriate indicator for measuring the MP realisation efficiency within the maintenance process appears *Realisation of MP* (hereinafter referred as M3). According to (Arts et al., 1998) indicator M3 is the ratio of number of work orders realised according to the MP and total number of planned work orders from the MP in analysed period.
The *Planned Maintenance Percentage* (hereinafter referred as M4) is another suitable indicator of the maintenance process, which represents the percentage of labour working hours on planned maintenance work orders from total labour working hours in a period (Arts et al., 1998). According to (Mobley et al., 2008), while the percentage of planned maintenance increases, it reduces the percentage of unplanned maintenance, which equally affects the maintenance costs reduction.
Regarding the environmental impact on the fleet maintenance management the indicator *Percentage of Fleet Roadworthiness* (hereinafter referred as E1). Indicator E1 represents the percentage of vehicles complying with minimum requirements regarding technical condition checked on technical inspection, i.e. vehicle roadworthiness test, in terms of vehicle safety and emissions, compared to the total number of vehicles controlled on technical inspection in certain period. In the paper (Bin, 2003) is shown the percentage of road vehicle manufacturers not meeting the legal requirements (thresholds) on emissions inspection ranging about 4-10% of the number of tested vehicles.
According to (Rechnitzer et al., 2000) it has been stated that in 3% of analysed accidents vehicle condition had certain impact on traffic accidents, i.e. technical malfunction or failure were the principal cause of accident. Therefore an appropriate indicator is the *Percentage of Vehicle Roadworthiness in Accidents* (hereafter referred as E2). Indicator E2 is the percentage of vehicles that participated in traffic accident and complied with minimum requirements regarding critical vehicle safety systems (e.g. braking, steering, suspension, lighting system and so on) relative to the total number of vehicles that participated in accidents in the observed period.

## 3. Methodology for evaluation of manager's efficiency

In the paper (Vujanović et al., 2012), by combining two methods: DEMATEL and ANP interdependencies have been calculated and rankings of defined indicators have been determined in relation to the achievement of the defined company objectives. As a result a model containing three interdependent areas (transport process, maintenance process and environment) and in each of the areas relative weights of defined indicators have been determined in relation to the defined company objective (Table 2). According to Table 2 it is obvious that while managing maintenance with the objective of increasing vehicle fleet energy efficiency the most relevant indicators are: *Maintenance Plan realisation* (M3), *Operation Plan Realisation Percentage* (T1), and *Vehicle Payload Utilisation* (T2).

**Table 2:**
*Model with relative indicator weights*
*source: (Vujanović et al, 2012)*

| Indicators | Relative weights | Rank of importance | Relative weights of fields |
|---|---|---|---|
| Operation Plan (OP) Realisation Percentage - ***T1*** | 0.195 | 2 | $T_P$=0.472 |
| Vehicle Payload Utilisation - ***T2*** | 0.169 | 3 | |
| Vehicle Fleet Utilisation Rate - ***T3*** | 0.108 | 5 | |
| Mean Time Between Failures - ***M1*** | 0.044 | 8 | $M_P$=0.431 |
| Mean Vehicle Downtime - ***M2*** | 0.122 | 4 | |
| Maintenance Plan (MP) Realisation - ***M3*** | 0.233 | 1 | |
| Planned Maintenance percentage - ***M4*** | 0.032 | 9 | |
| Percentage of Fleet Roadworthiness - ***E1*** | 0.051 | 6 | E=0.097 |
| Percentage of Vehicle Roadworthiness in Accidents - ***E2*** | 0.046 | 7 | |

Based on the developed model (Table 2) for a period of observation an "Overall management score" (S) could be calculated. Every single one of the defined indicators is evaluated with grades from 1 to 5 (where 1 is a grade for indicators with low value, and grade 5 for indicators with high realised value), according to expert determined grading scale (e.g. grade 1 could be associated to the indicator $T_1$ if its realised value is from 0 to 0.2; grade 2 if the attained value of the indicator $T_1$ falls into the interval from 0.2 to 0.4 and so on).
In such a way evaluated indicators are subject to relative weights from the Table 2 and the value an "Overall management score" (S) is obtained as follows (Vujanovic, 2013):

$$S = O_1 x g_{T1} + O_2 x g_{T2} + O_3 x g_{T3} + O_4 x g_{M1} + O_5 x g_{M2} + O_6 x g_{M3} + O_7 x g_{M4} + O_8 x g_{E1} + O_9 x g_{E2} \quad (1)$$

where: $O_1, O_2, \ldots O_8$ – grade of the management indicator (from 1 to 5);
$g_{T1}, g_{T2}, g_{T3}, g_{M1}, g_{M2}, g_{M3}, g_{M4}, g_{E1}, g_{E2}$ – relative weights of management indicators.
According to the paper (Vujanović et al., 2012), an "Overall management score" (S) determines the efficiency rating of the manager in view of vehicle fleet maintenance management.

**4. Results**

Methodology for evaluating manager's efficiency when managing vehicle fleet maintenance has been implemented in the company "Delmax" d.o.o. from Stara Pazova. The primary (core) activity of the company is distribution of automobile spare parts. The distribution of spare parts is realised daily on defined routes on the entire territory of the Republic of Serbia, with their own road vehicle fleet. The objective is, among others, to deliver to clients certain amount of spare parts in a defined time frame, with as lowest as possible transport and maintenance costs.
As a reference observation period of chosen company operation it has been selected the period from January to February 2011. Based on realised values of indicators in the reference period, to every indicator has been assigned a grade between 1 and 5 (grade 1 is the worst, while grade 5 is the best), according to expert determined limit values for evaluation. By using the developed model (Table 2) it has been calculated according to (1), value of the "Overall management score" (S), which in the observed reference period was 2.885 (Table 3).

**Table 3:**
*Realised values and grades of indicators according to model with relative weights in the company „Delmax" d.o.o., in the reference period*

| Management indicator | Realised value | Indicator grade | Value of the indicator factor as per model | Grade of the indicator according to model |
|---|---|---|---|---|
| $T_1$ | 90.96 | 3.20 | 0.195 | 0.624 |
| $T_2$ | 27.79 | 2.55 | 0.169 | 0.431 |
| $T_3$ | 83.33 | 3.66 | 0.108 | 0.395 |
| $M_1$ | 7445 | 1.00 | 0.044 | 0.044 |
| $M_2$ | 3.3 h | 2.70 | 0.122 | 0.329 |
| $M_3$ | 55.71 | 2.07 | 0.233 | 0.482 |
| $M_4$ | 64.29 | 2.93 | 0.032 | 0.094 |
| $E_1$ | 100 | 5.00 | 0.051 | 0.255 |
| $E_2$ | 100 | 5.00 | 0.046 | 0.230 |
| **Overall management score (S)** | | | | **2.885** |

Based on implementation of particular enhancement measures while managing the maintenance process (better realisation of agreed time frames for repairs of vehicles in specialized workshops in the surroundings, introduction of additional control interventions, decrease of the number of reserve/backup vehicles, etc.) during the period from March to August 2011, which have brought some enhancement of the realised values upon majority of defined indicators. For the period from July to August 2011the value of the "Overall management score" (S) has been enhanced compared to the reference period and amounts to 3.91 (see Table 4). Therefore, managers in the observed company have become more efficient while managing their maintenance.

**Table 4:**
*Realised values and grades of indicators according to model with relative weights in the company „Delmax“ d.o.o., in the period July-August, 2011*

| Management indicator | Realised value | Indicator grade | Value of the indicator factor as per model | Grade of the indicator according to model |
|---|---|---|---|---|
| $T_1$ | 98.21 | 5.00 | 0.195 | 0.975 |
| $T_2$ | 29.20 | 2.84 | 0.169 | 0.480 |
| $T_3$ | 88.89 | 4.77 | 0.108 | 0.515 |
| $M_1$ | 11 | 3.64 | 0.044 | 0.160 |
| $M_2$ | 3.03 | 2.97 | 0.122 | 0.362 |
| $M_3$ | 69.35 | 3.44 | 0.233 | 0.802 |
| $M_4$ | 75.81 | 4.08 | 0.032 | 0.131 |
| $E_1$ | 100 | 5.00 | 0.051 | 0.255 |
| $E_2$ | 100 | 5.00 | 0.046 | 0.230 |
| **Overall management score (S)** | | | | **3.910** |

By realisation of an integrated maintenance management has led to realisation of greater transport volume, although with lower number of vehicles compared to the reference period (Table 5). The decrease in the number of vehicles has led to decreasing capacity costs as well as vehicle fleet maintenance costs per transport volume. Since in the integrated approach only vehicles from appropriate CO vehicle groups were allocated according to the vehicle payload utilization, realised in the period from July to August 2011, the value of the indicator T2 has been enhanced (i.e. fuel consumption was lowered per unit of transport volume) for 14.68% compared to the reference period and amounts 15.39 *l*/100 *tkm*. By this value an increase in energy efficiency is expected.

**Table 5:**
*Vehicle fleet operation characteristics within „Delmax“ d.o.o., in the reference period and between July and August, 2011*

| Observation period | Number of vehicles | Distance travelled (*km*) | Realised transport volume (*tkm*) | Specific fuel consumption per transport volume – $q_t$ (*l* / 100 *tkm*) |
|---|---|---|---|---|
| January-February, 2011 | 16 | 186 122 | 111 249 | 18.04 |
| July-August, 2011 | 15 | 176 226 | 125 216 | 15.39 |

Therefore, by implementation of an integrated maintenance management and applying the methodology for evaluation of manager’s efficiency in the observed period has led to more efficient and rational vehicle fleet operation in the observed company, wherein all objectives regarding OP and environment have been fulfilled. It will influence a decrease in total transport and maintenance costs which will allow attaining of company objectives.

## 5. Conclusions

In this paper, the importance of measuring the effects of the implementation of specific measures in the maintenance management in order to increase vehicle fleet energy efficiency. For efficient fleet maintenance management it is necessary to observe integrally the transport process, the vehicle maintenance process and the environment. Thus, for measuring managers' efficiency in the maintenance management it was necessary to define indicators in these areas. Based on the literature review and authors' experience nine indicators were proposed. The model with the relative weights of maintenance management indicators with the objective of increasing vehicle fleet energy efficiency. A methodology for measuring managers' efficiency in the vehicle fleet maintenance management.

The developed methodology for measuring managers' efficiency has been implemented in a company with own vehicle fleet. Actually it has been attained an improvement in the "Overall management score" (S) in this period. In the same period, the improvement in specific fuel consumption per unit transport volume has been attained by order of 14.68%. In this sense, it can be concluded that the evaluation of the managers' efficiency in the maintenance management is significant for energy efficient and effective operation of the vehicle fleet, taking into account the requirements of the OP and the environment.

Further research will focus on applying the developed methodology for evaluating managers in public utility companies with large vehicle fleets where significant improvements in the efficiency of maintenance management is expected.

## Acknowledgments

The research in this paper has been realised within the project TR36027 named: "Software development and national database for strategic management and development of transportation means and infrastructure in road, rail, air and inland waterways transport using the European transport network models" supported by the Ministry of Science and Technological Development of the Republic of Serbia.

**References:**

Arts, R.H.P.M., Knapp, G. M., Mann L. J. (1998). Some aspects of measuring maintenance performance in the process industry. *Journal of Quality in Maintenance Engineering,* 4(1), 6-11.

Ashayeri J., Teelen A., Selen W. (1996). A production and maintenance planning model for the process industry. *International Journal of Production Research,* 34 (12), pp.3311-3326.

Bin, O. (2003). A logit analysis of vehicle emissions using inspection and maintenance testing data. *Transportation Research, Part D*, 8, 215–227.

Brah S., Chong W. (2004). Relationship between total productive maintenance and performance. *International Journal of Production Research*, 42 (12), pp.2383-2401.

Brandolese M., Franci M., Pozzetti A. (1996). Production and maintenance integrated planning. *International Journal of Production Research*, 34 (7), pp.2059-2075.

Haghani, A., Shafahi, Y. (2002). Bus maintenance systems and maintenance scheduling: model formulations and solutions. *Transportation Research Part A*, 36, 453–482.

Kamakaté, F., Schipper, L. (2009). Trends in truck freight energy use and carbon emissions in selected OECD countries from 1973 to 2005. *Energy Policy*, 37, 3743–3751.

Li, J.-Q., Mirchandani, P. B., Borenstein, D. (2009). Real-time vehicle rerouting problems with time windows. *European Journal of Operational Research*, 194, 711–727.

Marquez A. (2007). The Maintenance Management Framework; Models and Methods for Complex Systems Maintenance. *SpringlerVerlag*, London.

Marquez A., Gupta J. (2006). Contemporary maintenance management: process, framework and supporting pillars. *Omega-International Journal of Management Science,* 34, pp.313-326.

McKinnon A., (1999). A Logistical Perspective on the Fuel Efficiency of Road Freight Transport. Report on the Workshop "Improving Fuel Efficiency in Road Freight: The Role of Information Technologies. *International Energy Agency*, *European Conference of Ministers of Transport*, Paris,

McKone K., Schroeder R., Cua K. (2001). The impact of total productive maintenance practices on manufacturing performance. *Journal of Operations Management, 19*, pp.39-58.

Milosavljević N., Teodorović D., Papić V., Pavković G. (1996). A fuzzy sets theory approach to the vehicle assignment problem. *Journal of Transportation Planning and Technology, 20 (1)*, pp.33-47.

Mobley, R. K. Higgins, L. R. Wikoff, D. J. (2008). Maintenance Engineering Handbook, Seventh Edition. ISBN 0-07-164101-7, *The McGraw-Hill Companies, Inc.*, New York.

Momčilović V., Papić V., Vujanović D. (2007). Quality of the Fleet Maintenance Management: A Factor Influencing Sustainable Development. *Proceedings of the International Congress "Transport Science & Technology Congress"*, Prague.

Papić V., Medar O., Pejčić Tarle S. (1999). Management of vehicle fleet maintenance in conditions of sustainable development, *Mobility & Vehicles Mechanics*, 25, (2&3), 37-46.

Parida, A., Kumar, U. (2009). Maintenance Productivity and Performance Measurement. Handbook of Maintenance Management and Engineering. (Eds.) Ben-Daya, M., Duffuaa, S.O., Raouf, A., Knezevic, J., Ait-Kadi, D. ISBN 978-1-84882-471-3, *Springer-Verlag*, London.

Rechnitzer, G., Haworth, N., Kowadlo N. (2000). The effect of vehicle roadworthiness on crash incidence and severity, *Monash University Accident Research Centre, Report No. 164*, Clayton, Australia.

Ruzzenenti, F., Basosi, R. (2009). Evaluation of the energy efficiency evolution in the European road freight transport sector. *Energy Policy*, 37, 4079–4085.

Vujanovic, D. (2013). Contribution to the Development of the Vehicle Fleet Maintenance Management. Doctoral Dissertation. *University of Belgrade, Faculty of Transport and Traffic Engineering*. Original on Serbian.

Vujanović D., Momčilović V., Bojović N., Papić V. (2012). Evaluation of vehicle fleet maintenance management indicators by application of DEMATEL and ANP. *Expert Systems with Applications, 39 (12)*, pp.10552-10563.

Vujanović, D., Mijailović, R., Momčilović, V., Papić, V. (2010). Energy efficiency as a criterion in the vehicle fleet management process, *Thermal Science*, 14(4), 865-878

Waeyenbergh G., Pintelon L. (2002). A framework for maintenance concept development. *International Journal of Production Economics, 77 (3)*, pp.299-313.

Yamamoto, T., Madre, J.-L., Kitamura, R. (2004). An analysis of the effects of French vehicle inspection program and grant for scrappage on household vehicle transaction. *Transportation Research, Part B*, 38, 905–926.

Zhu G., Gelders L., Pintelon L. (2002). Object/objective-oriented maintenance management. *Journal of Quality in Maintenance Engineering,* 8 (4), pp.306-318.

Zhu G., Pintelon L. (2001). Integrated production maintenance management (IPMM) as an enterprise approach to maintenance management. *Proceedings of IIE Annual Conference*, *Institute of Industrial Engineers*, Dallas

*Sustainable Transport*

# DIAL–A–RIDE TRANSPORT SERVICE OF THE PROVINCE OF REGGIO CALABRIA: "CHI-AMA BUS"

**Domenica Catalfamo[1], Giuseppe Amante[2], Giovanna Chilà[3]**
*[1,2,3] Provincia di Reggio Calabria, Reggio Calabria, Italy*

**Abstract:** The "CHI-AMA BUS" is an integrated local development project, proposed by the Province of Reggio Calabria with the partnerships of local municipalities, and financed with European funds. The aim is to implement a Dial-a-Ride transport service in the entire area of the Province. This unconventional transport service, the first type of flexible transport service in the Calabria Region, will be organized in macro-zones of traffic, defined in relation to mobility demand as well as with socio-economic characteristics. The project includes the implementation of an operation centre, located at the building of the Province of Reggio Calabria that will be the smart centre dedicated to the management and planning of the transport service. Based on ITS technology, it will handle bookings, allocation of vehicles and any routing problem. Transport service will be performed with efficient vehicles suitable for the transport of all users and connected with the operation centre. A totem for booking and an information service will be installed on the main roadways of the Province.

**Keywords**: Accessibility, Dial-a-ride.

## 1. Introduction

The Province of Reggio Calabria is a peninsula that is touched both by the Tyrrhenian Sea and the Ionian Sea which stretches out towards Sicily forming a bridge for links between Europe and Sicily itself. In recent times urban settlements have developed along the coastal corridor, influencing the dynamics of the development of road infrastructure and traffic. The entire transport system and regional planning on a provincial scale have been influenced by this socio-economic development. The territory is characterized by the presence, in the space of a few kilometers, of different environments: in particular, the hinterland is characterized by a morphological structure that is connected with the Aspromonte mountain system, while the coastline is characterized by coastal settlements which have main development between the shoreline and the coastal roads.

A peculiar place which stands out on its own is the City of Reggio Calabria, the largest city of Calabria and the largest in terms of municipal territory (236 sq km), that is characterized by a considerable extension along the coast (over 10 km) and by a territory with altitudes ranging from sea level to over 1800 metres high. The city, which alone accounts for 70% of the population of the entire area covered by this study, is indeed predominant in relation to other nearby towns and is a real pole of attraction of mobility, both for work quality reasons and for the number of services offered. In the city you can find the following: the University with four faculties; schools of every level; an Airport with national connections; a port with regular ferry and hydrofoil services; the railway station, a long-distance railway line terminal; Large-capacity roads (coastal, Ionic and Tyrrhenian SA-RC motorway); the National Museum of Magna Grecia and other cultural heritage; the Office of the Province; offices of the region; hospital centres. All of this, together with an improved quality of life, an increasing economic activity and the allocation of services offered make the city a centre for mobility.

At the northern edge of the area you can find the town of Villa San Giovanni, whose importance is based on the connections with Sicily; further north, the Port of Gioia Tauro, 50 km from Reggio Calabria that is one of the most important ports in the Mediterranean for container handling.

Almost all components of the social and economic system interacting with different levels of intensity are covered on the entire territory. If we take a look at other main towns of the Province, inland areas of the Province are characterized by low accessibility in terms of infrastructures and services: the privately owned car represents the only modal alternative, especially for specific groups of citizens and in the peak-off time period. The high percentage of those who use their own private car produces very low levels of environmental quality and energy efficiency, a low level of both accessibility and attractiveness of areas. Accessibility has always been a decisive factor in economic development: improving accessibility allows a change of economic activities, increasing the economic surplus and the levels of local employment. The "elimination of distance" may offer new opportunities for territorial development in the medium and long term, such as development of trade at different spatial scales, relocation of businesses and consumers, expansion of the market and globalization of trade, employment growth and an increase in tourism.

The aim of this work is to implement a flexible transport service in order to improve the accessibility of the specific area of the Province of Reggio Calabria and to promote a modal shift towards more sustainable alternative transport.

[1] Corresponding author: domenica.catalfamo@provincia.rc.it

During the last, several studies have been developed on flexible transport service simulation, such as Deflorio et al. (2002), Fu (2002), Noda et al. (2003), Quadrifoglio et al. (2008) and so on. In the following, we propose a general vision of demand mobility in the Province of Reggio Calabria and the main characteristics of the proposed transport service in this area.

## 2. Demand Mobility in the Province of Reggio Calabria: a general vision

The mobility demand derives from the need to carry out activities in different places than where a user or a commodity is. Each move is the result of a series of choices made by the user of the transport service: the traveler in the mobility of persons or the operators (producers, shippers, transporters) in the transport of goods. The choices that determine the demand for transport with its many features are: to or not to move for some reason in a certain time slot, to which destination it will be done, how to do it and following which path. There are many issues which can determine by which means the demand for transport can be analyzed and classified; amongst these services, the territorial scale of reference may be included.

In this study we have analyzed the demand mobility on a provincial territorial scale.

Demand mobility in the Province of Reggio Calabria is related to a study area including 576,693 residents and 189,946 families; a percentage of about 47% of these moves for work or study reasons.

In particular, we have estimated the transport demand on a provincial scale in a generic day, for the following reasons: home – work, home - school, home-personal services, beginning from data available in transport literature and from the application of demand model of transport (Cascetta, 1998; Train, 2003; Cascetta, 2009). The applied demand model could be classified according to a statistical approach and includes: a generation model, finalized to define main generator poles of transport demand; a distribution model, finalized to define main attractor poles of demand; a modal choice model, finalized to evaluate choice percentage of transport mode in the Province of Reggio Calabria.

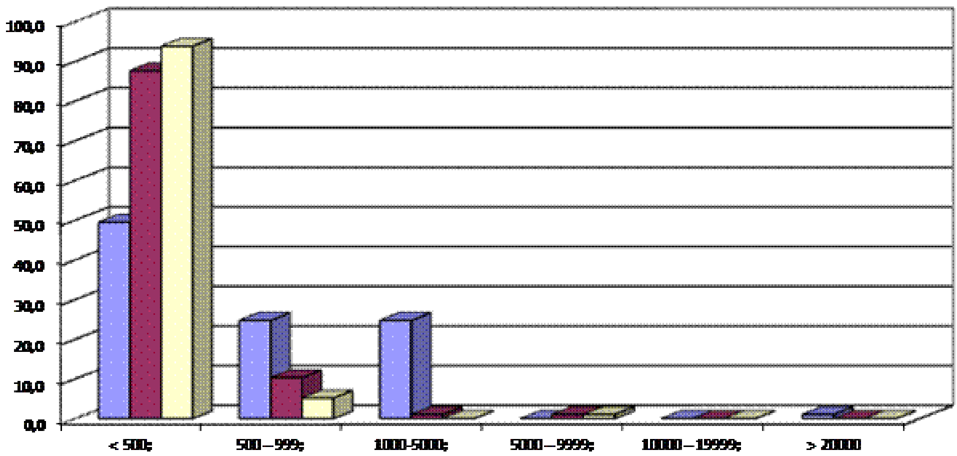
Figure 1 shows the percentage of municipalities in the province, classified according to the demand generated and for which reason, distinguishing between:

- percentage of municipalities with less than 500 emission users/day;
- percentage of municipalities with emission between 500 and 999 users/day;
- percentage of municipalities with emission between 1,000 and 4,999 users/day;
- percentage of municipalities with emission between 5,000 and 9,999 users/day;
- percentage of municipalities with emission between 10,000 and 19,999 users/day;
- percentage of municipalities with issuing more than 20,000 users/day;

and, with regards to the reasons of the movement:

- home-work (CL);
- home-school (CS);
- home-personal services (CA).

It is important to emphasize that the above analysis was performed including, values, total emission, also traveling with origin and destination within the same municipality.




*Legend*

*Home - work*

*Home - school*

*Home – personal services*

**Fig. 1.**

*Percentage of municipalities in the province per average daily demand generated and reason*

*Source: Province of Reggio Calabria*

Table 1 shows the main generator poles of demand for the various reasons. In particular, for the reason home-work the poles with emission by over 2,000 users/day are considered; for home-school poles with emission by over 500 users/day; for home-personal service poles with demand exceeding 350 users/day. In the last column you can find in descending order, the 10-pin characterized by higher average daily demand issued, for all the reasons.
With regards to the main attractors, figure 2 shows main municipalities for the home – work reason. The municipality of Reggio Calabria is in 42% of cases, the main town of the destination home - work. You can also identify two other poles of reference: Locri, in the Ionian basin, with rates that are around 16%, and the area of Gioia Tauro - Polistena - Palmi, in the Tyrrhenian basin, with rates that globally amount to around 18%.

**Table 1**
*Main generator poles of transport demand*

| **Home - work** | **Home -school** | **Home – personal services** | **All** |
|---|---|---|---|
| Reggio Calabria | Reggio Calabria | Reggio Calabria | Reggio Calabria |
| Palmi | Siderno | Palmi | Palmi |
| Gioia Tauro | Palmi | Gioia Tauro | Gioia Tauro |
| Taurianova | Taurianova | Siderno | Taurianova |
| Siderno | Locri | Taurianova | Siderno |
| Rosarno | Villa San Giovanni | Rosarno | Locri |
| Locri | Cittanova | Villa San Giovanni | Rosarno |
| Villa San Giovanni | Gioia Tauro | Locri | Villa San Giovanni |
| Polistena | Bagnara Calabra | Polistena | Polistena |
| Cittanova and Melito di Porto Salvo | Melito di Porto Salvo | Cittanova | Cittanova |

*Source: Data elaboration by Province of Reggio Calabria*

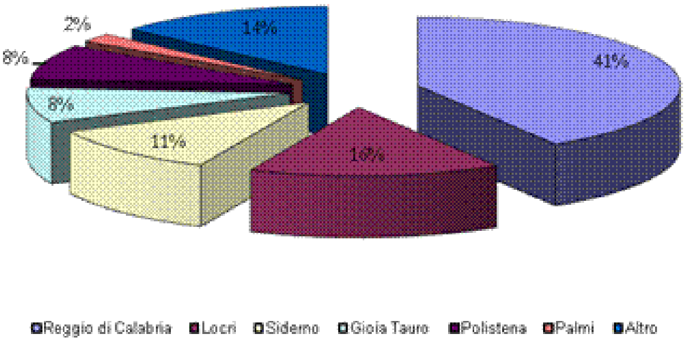



**Fig. 2.**
*Main pole attractors for the home – work reason*
*Source: Province of Reggio Calabria*

Figure 3 shows main municipalities for the home – school reason. The municipality of Reggio Calabria is, in 13% of cases, the main town of destination of trips for school. You can also identify other poles of reference: Locri, in the Ionian basin, with rates that are around 12%, with Bovalino and Roccella Ionica; Polistena, Palmi and Oppido Mamertina in the Tyrrhenian basin, with percentages that, globally, are around 26%.
Figure 4 shows the main attractors for home – personal services trips: the city of Reggio Calabria is, in 31% of cases, the main destination; then, Locri, Siderno and Roccella Ionica, in the Ionian basin; Polistena, Gioia Tauro and Palmi in the Tyrrhenian basin.
From the analysis carried-out, the result is that daily inter-zonal trips which are inside the Province have approximately 89,000 users/day. With respect to exchange movements, it is assumed that they have a total of 52,000 users/day, of which 26,000 interior - exterior, 26,000 external - internal; for trips with origin and destination outside the study area, it is assumed that they are about 15,000 users /day (*Source: Basin Plans of Local Public Transport, Province of Reggio Calabria, 2004*).

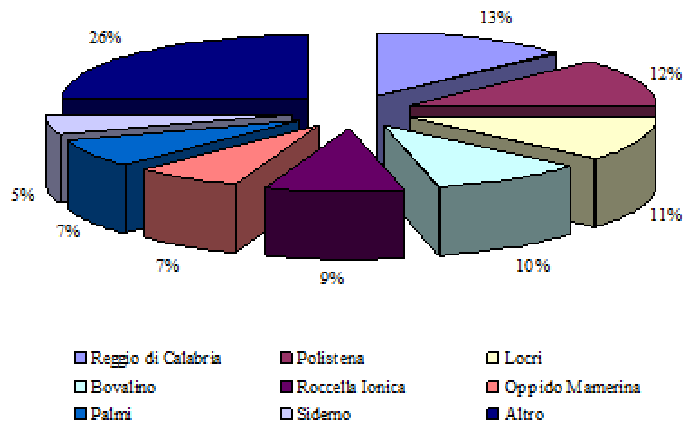



**Fig. 3.**
*Main pole attractors for the home – school reason*
*Source: Province of Reggio Calabria*

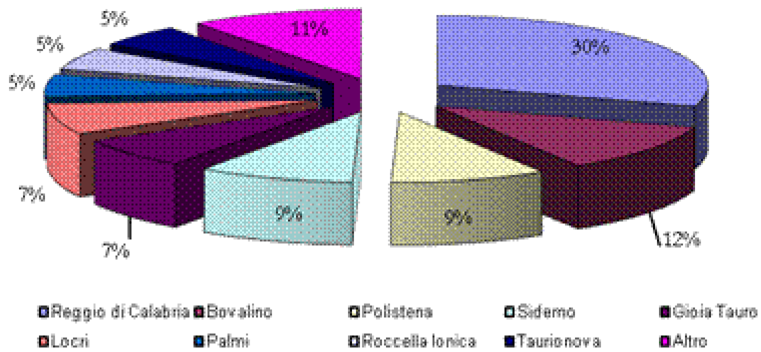



**Fig. 4.**
*Main pole attractors for the home – personal services reason*
*Source: Province of Reggio Calabria*

### 3. *CHI – AMA* Bus transport service: study area and demand

In order to maximize the efficiency and effectiveness of the proposed transport service, the study area has been subdivided in basin of mobility, within which to implement the proposed operations, starting from:

- data provided by the Calabria Region in the Framework of the Integrated Planning (2011);
- basins of traffic introduced in the Transport Plan of Calabria Region (1997) and defined in Guidelines of Regional Transport Planning (2013);
- Local Units of Work, that is the territorial units consisting of multiple contiguous municipalities in which the majority of the population lives and works, bordered by commuting flows for work (Istat, 2001);
- data included in the Provincial and Territorial Coordination Plan of the Province of Reggio Calabria.

By analyzing the data mentioned above, 8 traffic basins have been defined; each basin has been identified by the main attractor included in the basin, in terms of demand and supply transport. Some data of transport mobility related to the traffic basins are reported in table 2.

Given the characteristics of the dial-a-ride transport service (Hall, 2011), it should refer not only to commute to work or study, currently served by traditional transport services, but also to the demand mobility for home-personal, much more variable and difficult to predict and manage.

Similar experiences made on a national and international scale, indicate that the main beneficiaries of dial-a-ride services are the users who move to recreational activities (38%), work (27%), medical visits in hospitals (16%) and shopping (13%). It is estimated that the service can also be used by students in peak-off period who are not covered with the traditional transport services.

Starting from this status quo, it is clear that there is a need to implement the proposed transport services, which should be a bridge between the private and the public transport mode, giving additional transport services in the *soft* periods and promoting a modal shift of displacements.

From the above data, the set of potential users of the proposed transport service has been defined; it takes into account the categories of users considered in the transport literature for the implementation of dial-a-ride services (e.g. the SAMPO Project, the TWIST Project).

In order to define the demand mobility available to use the proposed transport service, in this study the following user categories have been considered:

- users moving for home – personal services;
- weak users;
- users who are directed towards main transport nodes.

**Table 2**
*Transport data for traffic basin*

| | Ionian Area | | | Reggio Calabria Area | | Tyrrhenian Area | | |
|---|---|---|---|---|---|---|---|---|
| | **Roccella Ionica** | **Locri** | **Bovalino** | **Bova Marina** | **Reggio di Calabria** | **Palmi** | **Gioia Tauro** | **Polistena** |
| Residents | 46.677 | 42.580 | 35.485 | 29.082 | 228.196 | 53.994 | 71.829 | 56.380 |
| Families | 17.123 | 14.937 | 12.709 | 10.733 | 80.287 | 18.898 | 24.086 | 19.921 |
| Workers | 14.739 | 14.076 | 11.945 | 9.235 | 82.530 | 18.522 | 26.046 | 19.823 |
| Commuters | 16.105 | 15.504 | 12.964 | 10.156 | 95.873 | 21.458 | 27.691 | 20.197 |

*Source: Data elaboration by Province of Reggio Calabria*

### 3.1. Users moving for the home – personal services

The transport demand for the home – personal services has been estimated according to a statistical literature model for the entire Province of Reggio Calabria and is distributed amongst the basins, as reported in Table 3. This specific mobility demand is generally poorly served by the current public transport service, because of its variability both in time and in space; indeed, the modal choice percentage of public transport for the mobility demand for shopping, personal services and so on is even lower than the provincial average.
It is possible to estimate an additional set of potential users, moving for different reasons within associated urban systems (e.g. Locri - Siderno, Reggio Calabria - Villa San Giovanni, Gioia Tauro - Palmi), where the most important public services for the traffic basin (e.g. schools, hospitals, public offices, etc.) are located. Starting from available data, this demand is of approximately 4.400 users/day.

**Table 3**
*Emission of trips for the home – personal services aggregate for mobility basin*

| Area | Basin | | |
|---|---|---|---|
| Ionian Area | Roccella Ionica<br>1’500-2’000 | Locri<br>1’500-2’000 | Bovalino<br>1’000-1’500 |
| Reggio Calabria Area | Melito di Porto Salvo<br>1’000-1’500 | Reggio di Calabria<br>>3’000 | / |
| Tyrrhenian Area | Palmi<br>2’000-2’500 | Gioia Tauro<br>2’500-3’000 | Polistena<br>2’000-2’500 |

*Source: Data elaboration by Province of Reggio Calabria*

### 3.2. Weak users

Given the social value of this transport service, finalized to improve not only the internal territorial accessibility, but also the quality of life of residents, especially of weak user categories, we have focused our work particularly upon: seniors (over years), teenagers (14-18 years) the unemployed and special users. These categories, in most cases, do not have their own car; therefore, the shortage of supply of TPL has an influence on the quality of life, limiting activities and initiatives. Data on the weaker groups identified by basin within the province of Reggio Calabria are reported in table 4.

**Table 4**
*Data on vulnerable users in the Province of Reggio Calabria*

| Basin | Seniors | Teenager | Unemployed | Special users* | Tot. | Residents | Families |
|---|---|---|---|---|---|---|---|
| Reggio Calabria | 36.684 | 11.546 | 10.392 | 80 | 58.622 | 213.093 | 75.030 |
| Melito di Porto Salvo | 8.671 | 2.571 | 1.235 | 20 | 12.477 | 43.679 | 15.990 |
| Bovalino | 6.814 | 2.067 | 1.113 | | 9.994 | 35.485 | 12.709 |
| Locri-Siderno | 7.487 | 2.423 | 2.017 | | 11.927 | 42.580 | 14.937 |
| Roccella Ionica | 9.227 | 2.778 | 2.067 | | 14.072 | 46.677 | 17.123 |
| Palmi | 9.696 | 3.246 | 2.472 | 30 | 15.414 | 53.994 | 18.898 |
| Gioia Tauro | 10.713 | 4.479 | 4.966 | | 20.158 | 71.829 | 24.086 |
| Polistena | 9.843 | 3.491 | 3.597 | | 16.931 | 56.380 | 19.921 |

*Source: Data elaboration by Province of Reggio Calabria*

### 3.3. Users directed towards main transport nodes

In this study we have mainly considered demand mobility directed towards the "Tito Minniti" airport.
We have carried out RP (Revealed Preferences) and SP (Stated Preferences) surveys (Ortuzar and Willumsen, 2006); RP surveys are aimed at the analysis of the main characteristics of user mobility departing from and arriving at the airport, while SP surveys are finalized to evaluate the willingness, on the part of the current users of the airport, to use the proposed transport service.
With regards to RP surveys, we have carried out 188 investigations, amongst which 99 are about departing users and 89 are about arriving users. Surveys are carried out on a generic weekday, in a temporal period between 6 am and 12 pm. In comparison with the traffic basin of the CHI-AMA Bus, the distribution of departing and arriving users is reported respectively in figures 5 and 6.
Concerning SP surveys, a percentage of users of around 84% states to be willing to use the proposed transport service, a percentage around 7% state that they are not willing to use it, while a rate of around 9% did not answer. With regard to the knowledge of the service, it is useful to point out that, compared to the total users interviewed, a percentage of around 35% said they did not know dial-a-ride transport services, a percentage of 51% said they did and the remaining percentage did not answer.
Starting from RP and data demand mobility related to the Tito Minniti airport, an estimation has been made, following a statistical approach. In figures 7 and 8 estimated demand mobility of users of Reggio Calabria airport is reported, considering users available to choose the *CHI-AMA* Bus service and distinguishing departing (fig. 7) and arriving (fig. 8) users.

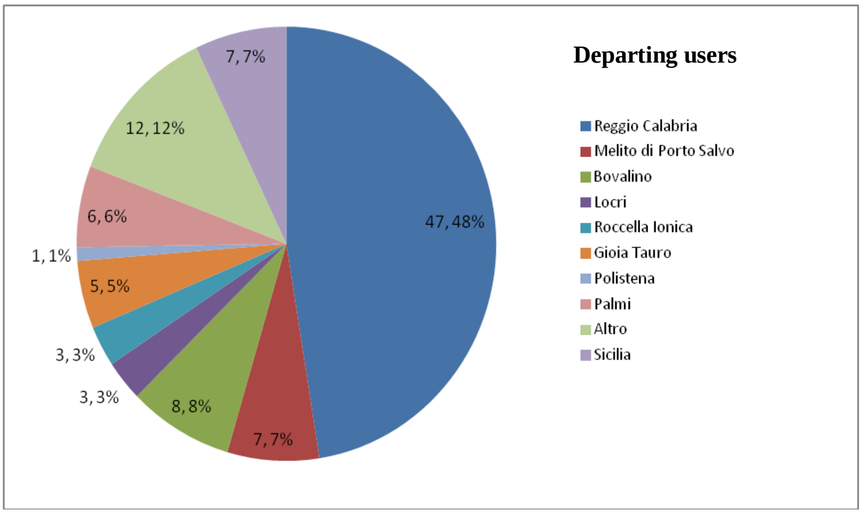



**Fig. 5.**
*Distribution of departing users vs. origin basin*
*Source: Data elaboration by Province of Reggio Calabria*

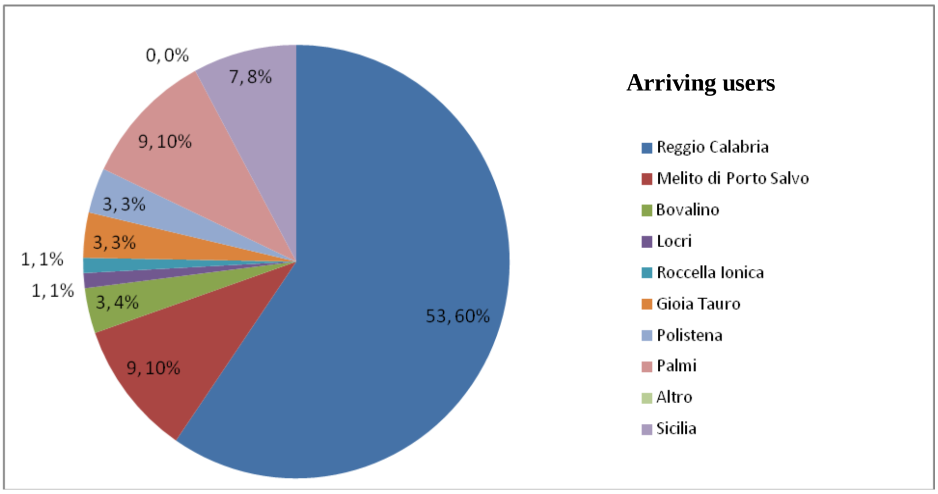



**Fig. 6.**
*Distribution of arriving users vs. destination basin*
*Source: Data elaboration by Province of Reggio Calabria*

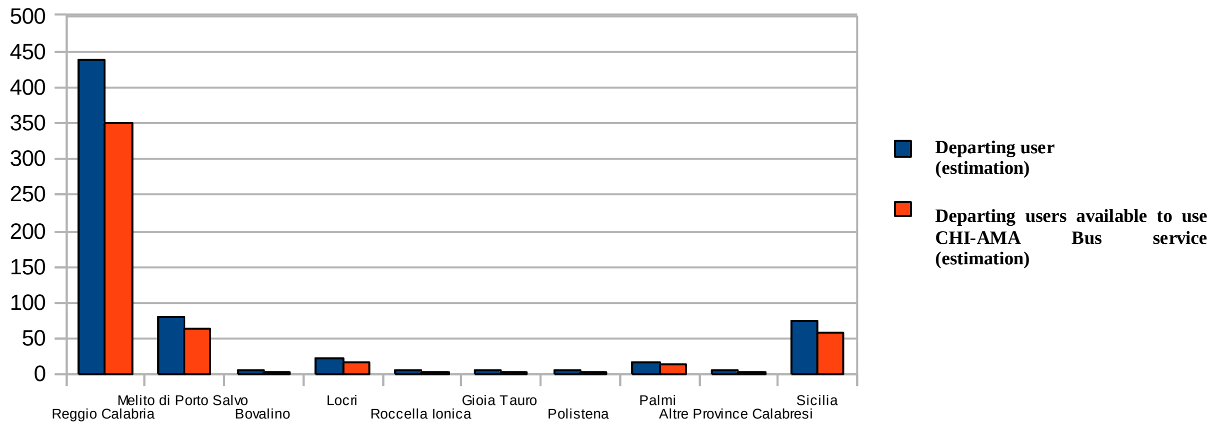



**Fig. 7.**
*Estimated demand of departing users vs. origin basin*
*Source: Data elaboration by Province of Reggio Calabria*

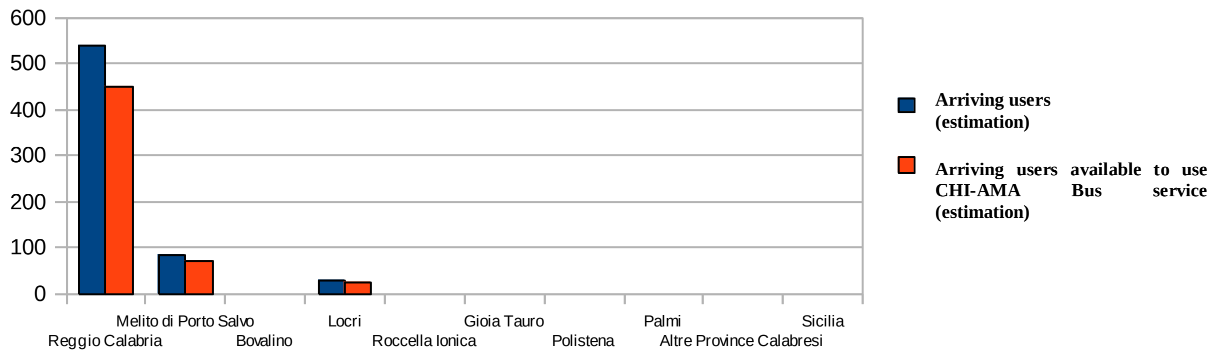



**Fig. 8.**
*Estimated demand of arriving users vs. destination basin*
*Source: Data elaboration by Province of Reggio Calabria*

## 4. *CHI – AMA* Bus transport service: main characteristics

The project includes the implementation of an operation centre, located at the Province of Reggio Calabria building which will be the smart centre whose objective is to manage and plan the transport service. Based on ITS technology, it will handle bookings, allocation of vehicles and any routing problems.

The operations centre will be linked in real time with the buses used for carrying out the service, the latter being equipped with hardware and software to ensure tracking and data transmission.

The operations centre will also have the function of giving user information and being a booking service. Specifically, the booking of the service will also be done through:

- The purchase of client packages for smartphones and the internet;
- An intelligent totem available in some municipalities of the province (e.g. in the main attractor of each basin) and in the main transport and services hub.

The operations centre will provide:

- The location of the vehicle;
- The location / representation of the fleet of vehicles (based on cartography and in relation to journeys made) on the system of land that serves as a control centre for real-time monitoring of the progress of races;
- The availability and accessibility of all information through "Web based" technology.

The operations centre will be divided into two main systems:

- System Earth (SDT), i.e. the function block aimed at managing communication with the media and with any mobile operation centres present in the companies that provide the transportation, processing and storage of data relating to the service;
- The transmission systems operating data, such as the systems board (SDB), installed on vehicles, equipped with wireless connectivity UMTS / GPRS to the central system of Earth and GPS devices for locating; System Earth will also be able to receive information directly from the driver of the vehicle if the system board is equipped with a console driver.

Transport service will be performed through efficient vehicles suitable for the transport of vulnerable users and connected with the operation centre. For the realization of the service the purchase of minibuses with 9 places and of a bus with 40 places is provided, which will be equipped with an on-board computer constantly able to locate the vehicle and the route. The board terminals, installed on each vehicle, will ensure the transmission of data between the operation centre and vehicles. Some of these minibuses will be equipped with electrical power, according to a sustainable and *smart* approach.

## 3. Conclusions and future perspectives

In this study we have proposed a flexible transport service finalized to improve the accessibility and the quality of life in the Province of Reggio Calabria, promoting a modal shift towards more sustainable transport alternatives. This is the first example of a dial-a-ride transport service in this territorial context and could represent the input for the launching of a *smart* user-oriented transport management system.

Future objectives will be to extend the management system of the *CHI-AMA* Bus service to the general system of transport which is being implemented in the Metropolitan City of Reggio Calabria.

**References**

Cascetta, E. 1998. Teoria e metodi dell'ingegneria dei sistemi di trasporto. UTET, Torino.

Cascetta, E. 2009. Transportation Systems Analysis. Models and Application. Springer, New York

Deflorio, F.; Chiara, B.D.; Murro, A. 2002. Simulation and performance of DRTS in a realistic environment. In Proceedings of the 13th Mini-Euro Conference Handing uncertainty in the analysis of traffic and Transportation systems and the 9th Meeting of the Euro Working Group on Transportation Intermodality, Sustanaibility and Intelligent Transport systems, 622-628.

Fu, L. 2002. A simulation model for evaluating advanced dial-a-ride paratransit systems. Transportation Research Part A, 36: 485-506.

Hall, C.H. 2011. Modelling and Simulation of Dial-a-Ride and Integrated Public Transport Sevices. Liu-Tryck, Linkoping.

Noda, I.; Ohta, M.; Shinoda, K.; Kumada, Y.; Nakashima, H. 2003. Evaluation of usability of dial-a-ride systems by social simulation. In Proc. of Fourth International Workshop on Multi-Agent-Based Simulation, 139–152.

Ortuzar, J.; de D. Willumsen, L.G. 2004. Modelling Transport, John Wiley & Sons Ltd, Chichester.

Progetto TWIST, Trasport With a Social Target, INTERREG III B – CADSES. 2007. Available from internet: <http://www.twistproject.org/IT/aspbite/categories/index.asp?inteCatID=80>

SAMPO Project, Systems for Advanced Management of Public Transport Operations. Available from internet: <http://cordis.europa.eu / telematics /tap_transport / research / projects / sampo.html>

Quadrifoglio, L.; Dessouky, M.; Ordonez, F. 2008. A simulation study of demand responsive transit system design. Transportation Research Part A 42: 718–737.

Train, K. 2003. Discrete choice methods with simulation. MIT Press, Cambridge, MA.

# THEORETICAL AND REAL PUBLIC TRANSPORT FLOWS – DAILY COMMUTING IN THE CZECH REPUBLIC

**Igor Ivan[1], Jiří Horák[2], Lenka Zajíčková[3]**
[1,2] *Institute of Geoinformatics, VSB-Technical University of Ostrava, 17. listopadu 15, 70800, Ostrava, Czechia*
[3] *Department of Geoinformatics, Faculty of Science, Palacky University in Olomouc, 17. listopadu 50, 771 46, Olomouc, Czechia*

**Abstract:** The effort to increase usage of public transport for daily mobility belongs to main strategies of national transport governments in many developed countries and applied strategies differ from country to country. Nevertheless in the Czech Republic, the only national-wide source of data being able to describe the result of such effort provides the census. It contains data about modal split of commuting flows on the municipality level. Even it does not cover all types of daily mobility, commuting and travels to school belong to the most frequent ones. New approach how to regularly analyse traffic flows on detailed geographical level provides valid time tables which are utilized in this paper. Anyway, both of these data sources describe transport situation on the municipality level from different perspective - real flows vs. theoretical flows which should however partly respond to real transport demand. Basic portraying of such data in a map reveals the existence of larger areas with municipalities indicating higher or lower level of public transport use and accessibility. To locate such hot/cold spots, univariate local indicator of spatial autocorrelation has been used. These methods enable to distinguish municipalities with a high level of both the public transport use and the accessibility and conversely; but many municipalities imply also combination of values from opposite portions of data distribution. Mutual comparison of results using methods of spatial clustering methods classifies municipalities into groups according to similar level of public transport use and public transport accessibility.

**Keywords**: public transport, time table, commuting, accessibility, census.

## 1. Introduction

Public transport has a long history in the Czech Republic and compared to another similar in the central Europe countries (e.g. Poland or Slovenia), high offer of public transport connections still remains also due to financial subsidies by government. This fact caused sufficient level of public transport accessibility to all municipalities what belongs to basic idea of law 194/2010 about public passenger transport services that generally ensures basic transport serviceability of a region but without specific number of bus or train connections. Anyway due to rationalization and short cuts to public transport offer, the level is getting lower. Even so development of modal split in the Czech Republic copies similar development in European countries. The main aspect is stagnant or decreasing share of public transport use and increasing number of journeys by cars (Burian et al., 2014). The graph below depicts the situation in Czechia where share of public transport use has decreased form 17.5% in 1995 to 10.9% in 2013 (without city public transport). This decrease is caused mainly due to smaller number of passengers using buses (from 644 mil. in 1995 to 338 mil. in 2013). Number of passengers using trains tends to increase in last three years (about 3 mil. more), nevertheless the total decrease is more than 52 million of passengers from 1995 to 2013.

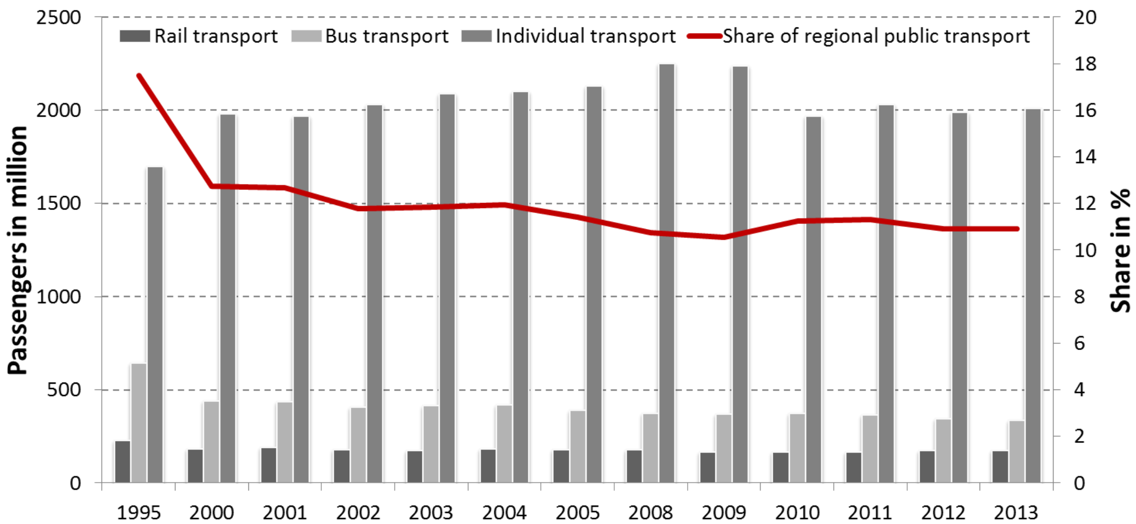



**Fig. 1.**
*Development of modal split in Czechia*
*Source: created by authors according to data from MoT (2014)*

[1] Corresponding author: igor.ivan@vsb.cz

Final mode choice is still often studied factor in transport studies. Current Czech national transport policy 2014-2020 has not been accepted yet but the draft tends to follow European trend with activities to increase public transport use. However some countries have been successful in shifting car users onto public transport, others are struggling despite their effort to make public transport more attractive. From all objective factors that have more or less significant impact on final travel mode choice, travel time and the level of variability of travel time are considered as the most important (van Vugt, van Lange, Meertens, 1996). Particularly in case of regular short-distance trips, e.g. commuting, the commuter's behaviour minimises the economical and time costs needed for the trip.
This paper deals with theoretical and reals commuting flows using public or individual travel modes in the Czech Republic. Even it is not so easy to quantify real commuting flows, census data can be used to describe commuting flows at municipality level under the conditions of Czechia. Also other data sources are possibly useable such as ticket sales or floating car data, but while data about ticket sales are considered as trade secret, utilisation of floating car data is still quite rare (e. g. Novák et al., 2013). Theoretical transport flows are usually modelled using various modifications of gravity models but in this paper data from valid time tables is used. The goal of the paper is to compare spatial distribution of theoretical and real transport flows and to categorise municipalities into several categories based on level of real public/individual transport use for commuting (real flows), number of municipalities accessible by public transport or car ownership index (theoretical flows) together with other socio-demographic variables. The remainder of this paper is organised as follows: Section 2 describes the data used; Section 3 presents theoretical transport potential; Section 4 describes real transport flows; Section 5 presents the results of combining both data sources and finally; Section 6 summarises the main findings.

## 2. Data

Specific data sources play an important role in this paper because several different databases have been combined to get final results. Two main data sources can be distinguished. Census provides data about commuting flows in 2011 and is used for description of real commuting flows. It contains records about all commuting flows with the origin municipality within the Czech Republic. Unfortunately quality of this data is not perfect because more than 30% of residents did not answer the questions relating to their commuting behaviour. Nevertheless it can be assumed that the structure of commuting behaviour is correct. For needs of this paper, a commuter is defined as an employee who has the job out of the residential municipality. That means that commuting within the same municipality is not considered as commuting. Besides municipality of origin and destination and number of commuters, demographical, employment and transport mode structures are distinguished. Altogether, the table contains 178,171 records and 1,551,918 commuters are travelling between these pairs of municipalities. The original commuting data has been reduced based on two criterias. Firstly, both, origin and destination of commuting must be within the Czech Republic and secondly, only commuting flows shorter than 100 kilometres (Euclidean distance) were included. This distance restriction is important to make this data consistent with the second data source described below.
For estimation of theoretical commuting flows or transport potential, the Database of transport connections has been used. This database has been developing since 2007 by authors and annually updated based on valid time tables. Czechia has the advantage that all time tables are centralized in central information system maintained by CHAPS Ltd. This database contains all combination of municipalities within 100 kilometres (Euclidean distance) with information about existing public transport connections between each pair. Several variables are searched in time tables for each combination of municipalities (e. g. travel time, number of changes, price, existence of return connection) for five times (to 6, 7, 8, 14 and 22 o'clock). These times define the beginnings of three work shifts. Valid public transport connection must meets defined criterias. Travel time is smaller than 90 minutes, number of changes is smaller than 5, arrival time cannot be earlier than 60 minutes before, and departure time from origin cannot be earlier than 120 minutes before arrival to destination (more in Ivan et al., 2013 or Horák et al., 2014). For purposes of this paper, timetables for trains and buses (no urban transport) valid in March 2011 have been utilized.
Central register of motor vehicles maintained by the Ministry of Transport was used for estimation of individual transport potential. To keep time consistency between used data sources, database from 10th of April 2011 was downloaded. In particular, number of cars registered in municipalities was selected for calculation of car ownership index (number of cars per resident older than 17).

## 3. Theoretical transport potential

Theoretical transport potential is defined as number of municipalities from where is the analysed municipality accessible *NMA* to total number of municipalities within 100 kilometres *NMT*

$$RA_{h,i} = \frac{NMA_{h,i}}{NMT_{h,i}} * 100$$

where *h* is start of work shift and *i* is index of municipality. Potential differs from 0 to 100, where 100 means that there is a possibility to travel to analysed municipality from all municipalities within 100 kilometres, so the ideal destination for commuters from surrounding municipalities.

However, many times commuters are using a public transport stop in nearby municipality within walking distance. Therefore spatial filter has been used and final value is equal to the average of potentials of municipalities within 2 kilometres. This distance is considered as maximal walking distance to a public transport stop.

Average municipality in Czechia is accessible for any of analysed five arrival times from 6% of surrounding municipalities within 100 kilometres. Spatial distribution is depicted in the map below (figure 2) and spatial clustering (effect of second order) as well as spatial trend (effect of first order) from west to east are evident in this map. Spatial trend is confirmed by positive and statistically significant ($p = 0.01$) correlation between public transport potential and $x$ coordinate ($R = 0.284$). The more east is the municipality the better is its accessibility. Roads (highways, motorways and first class roads) and railways have influence on public transport potential too. Average of theoretical transport potential of municipalities within 2 kilometres from railway or road network is 6.9%. This is higher than national average. Difference in public transport potential between municipalities within and farther than 2 kilometres from roads and railways (4.9%) has been proved by ANOVA. Bigger influence on average public transport potential is caused by roads with an average equal to 7.5%. Average for railways is smaller 7%.

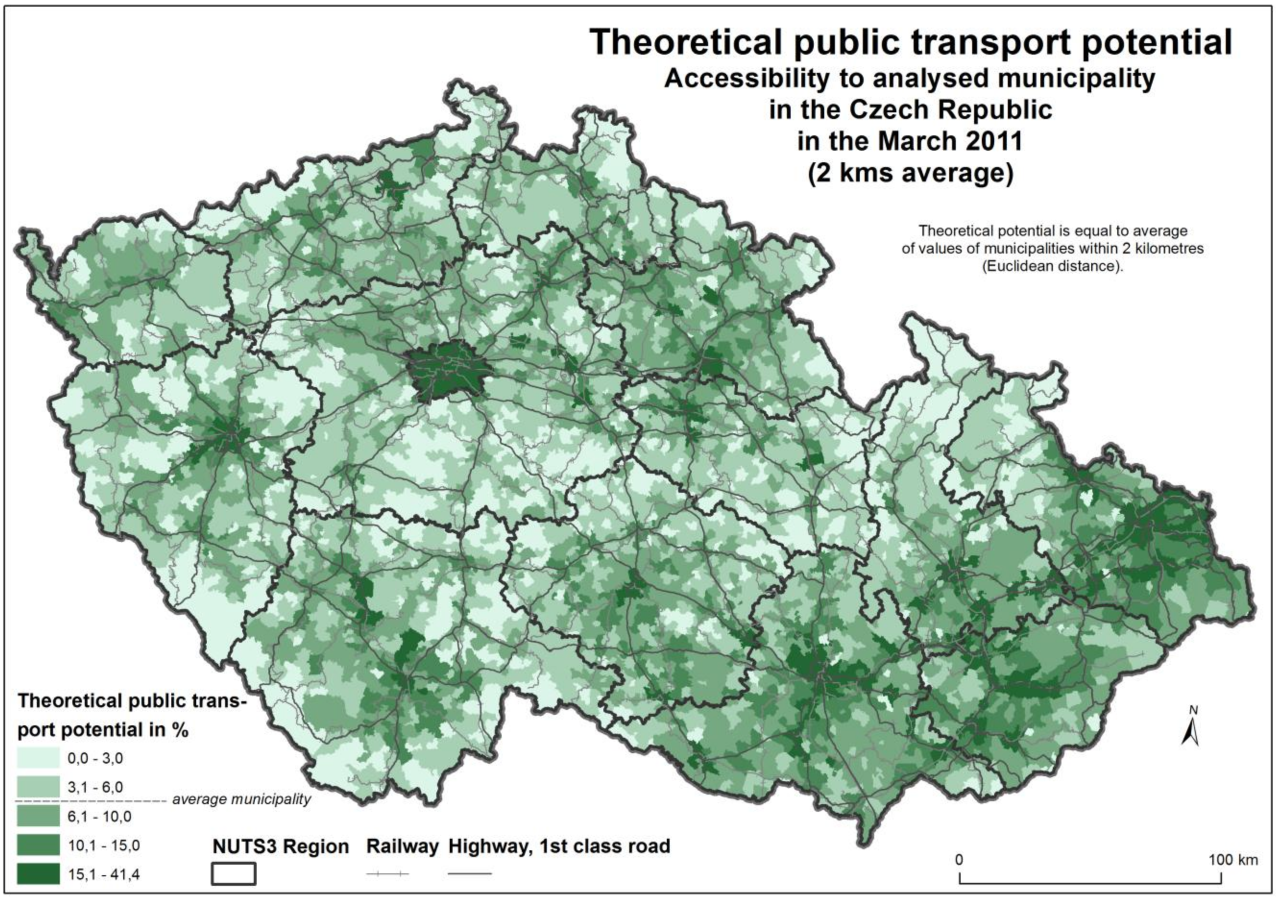



**Fig. 2.**
*Theoretical transport potential*
*Source: authors*

Obviously if a municipality has a good public transport potential, surrounding municipalities should have higher potential too. Particularly municipalities joined with this municipality by the same public transport link. Spatial clustering is analysed using local autocorrelation of public transport potential – Moran's $I$ (local indicator of spatial association) (more in Anselin, 1995). Moran's $I$ is 0.575 and because interpretation is the same as in case of correlation, it has been proved that public transport potential is clustered. The left map in figure 3 depicts clusters of municipalities. Red colour characterizes municipalities with higher public transport potential surrounded also by municipalities with higher potential. These municipalities are concentrated mainly in the eastern parts of Czechia what confirms previous hypothesis about western-eastern trend. On contrary blue colour describes municipalities with low potential surrounded by municipalities with low level. These municipalities create several clusters situated mainly in the western part of Czechia and in the northern part of Moravia (eastern part of Czechia).

So far only public transport potential has been analysed. For purposes of individual car transport potential, car ownership index has been used and Moran's $I$ is smaller than previously (0.304) but still statistically significant. From the cluster map (right map), it is evident that the whole eastern part (Moravia) creates one large cluster with low values of car ownership index. A few hot spots (high values surrounded by high values) are sparsely distributed in the western part and mainly around Prague, Pilsen and České Budějovice (regional capitals).

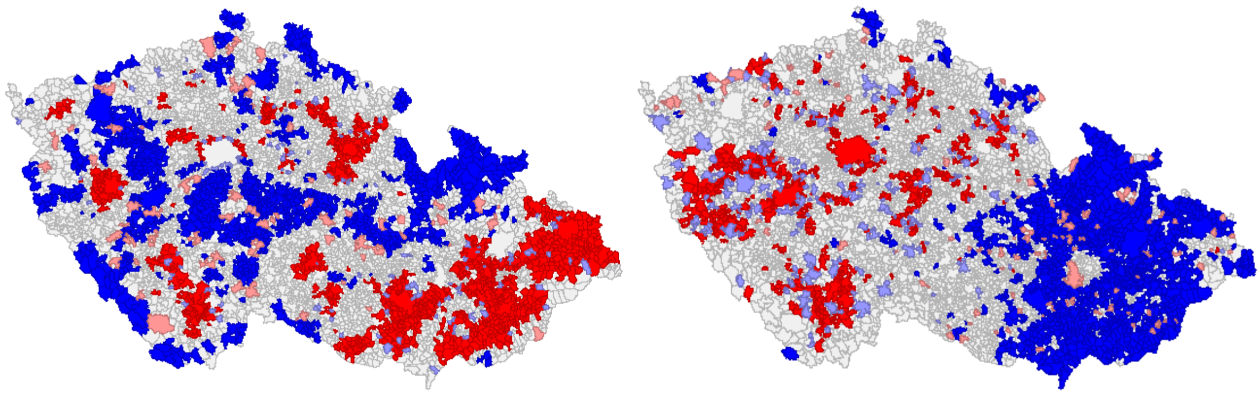


**Fig. 3.**
*Cluster map of public transport potential (left) and individual car transport (right)*
*Source: authors*

**4. Real transport flows**

Real transport flows are described by census data for both the public and the individual transport. The map below (figure 4) describes share of public transport use for commuting to analysed municipality from surrounding municipalities within 100 kilometres. To an average municipality, 16.9% of all commuters use public transport. But again significant spatial heterogeneity is present. The western-eastern trend is very similar as for public transport potential and again with statistically significant ($p$ = 0.01) correlation with $x$ coordinate (R = 0.204). Influence of transport infrastructure has been again proved by ANOVA. Compared to general average (16.9%), municipalities within 2 kilometres from a railway or a road (highway or first class road) network has the average slightly bigger 17.6% and significantly bigger than for municipalities farther than 2 kilometres from transport infrastructure (14.2%). Railways have a little bit bigger significant (18%) than roads (17.9%).

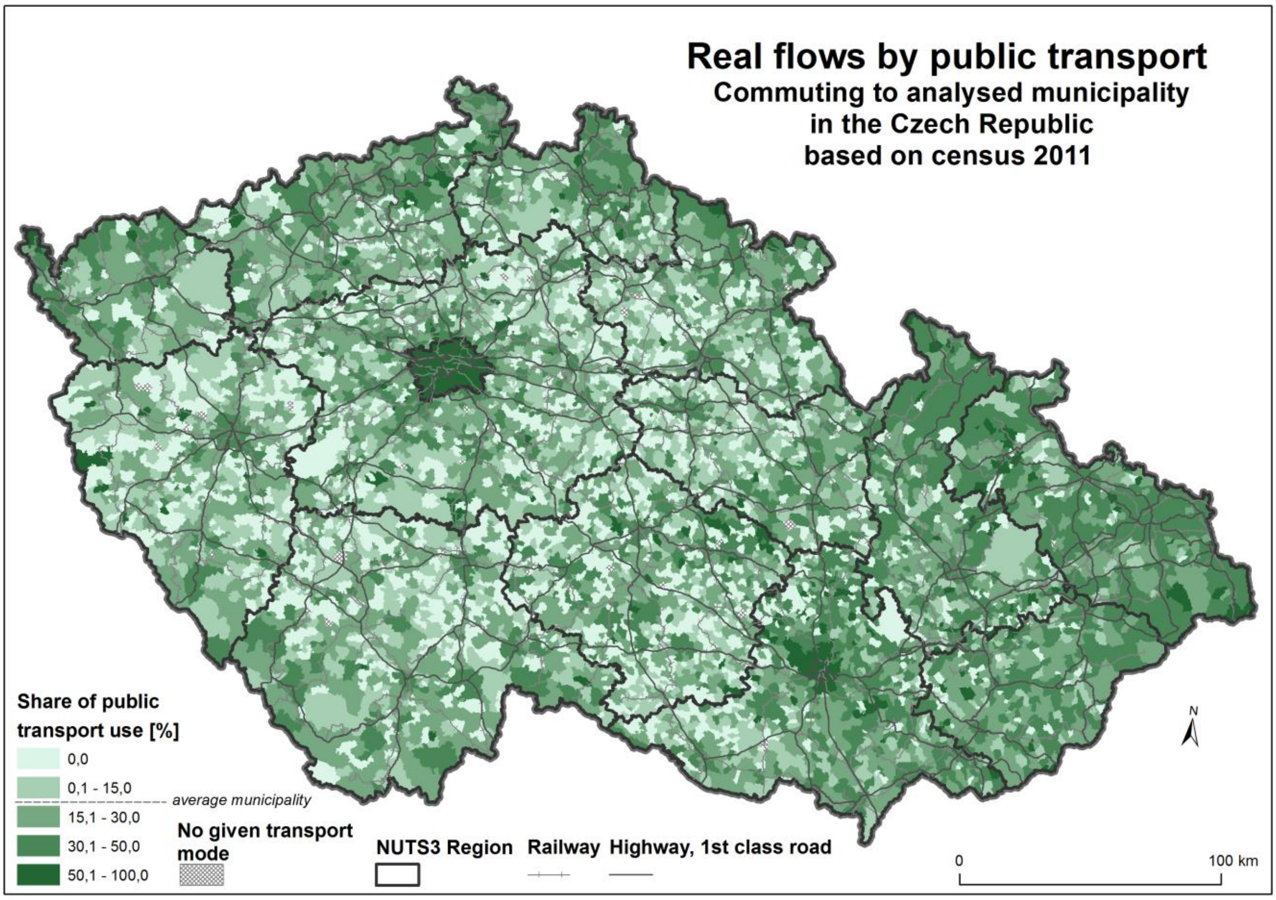



**Fig. 3.**
*Share of public transport use*
*Source: authors*

Spatial clusters were studied again using Moran's I, but the assumption did not predict some significant clustering of higher or lower values as was in theoretical flows. This has been proved by small values of Moran's I, that is equal to 0.173 for public transport use and 0.032 for individual transport use. But both are statistically significant. Hot spots of public transport use are concentrated in the eastern part of Czechia and correspond to public transport potential. Cold spots are very small and sparsely distributed in Bohemia (western part of Czechia). Individual transport does not create any large hot or cold spots and one bigger hot spot is located in the area to the east from Prague.

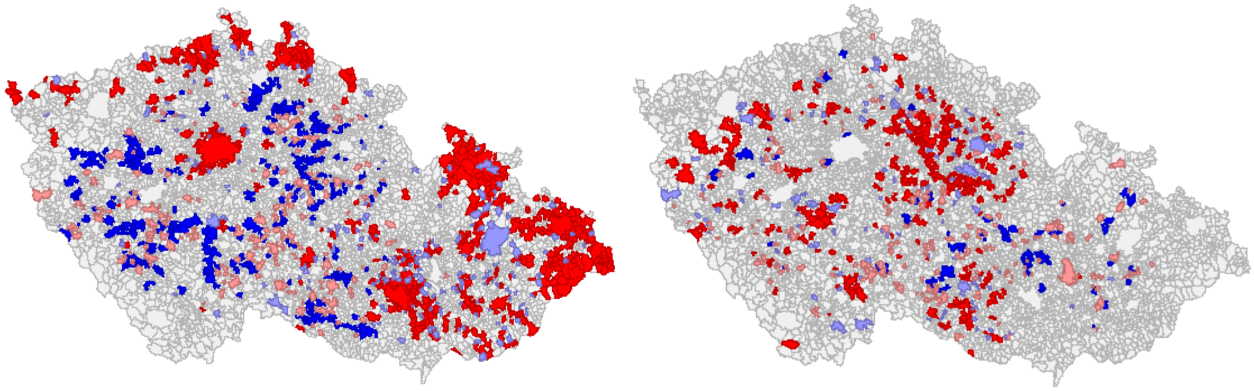


**Fig. 4.**
*Cluster map of public transport use (left) and individual car transport use (right)*
*Source: authors*

## 5. Comparison of transport potential and real commuting

As it has been noted above, the spatial distribution of municipalities with higher public transport potential and real use of public transport suggest similar spatial distribution. This chapter statistically proves this assumption using k-means clustering method (Smith, 2014). All municipalities have been classified into 4 different clusters based on their values of public transport potential and real use. The first cluster contains the biggest number of municipalities (43%) and could be named as "not knowing public transport" because these municipalities have very low both indicators, public transport potential as well as real use of public transport. These municipalities are located mainly in the western part of Czechia.

**Table 1**
*Final cluster centres for public transport and share of municipalities*

| Cluster | Real use in % | Potential use in % | Share of municipalities in % |
|---|---|---|---|
| 1 | 1,9 | 4,8 | 43,1 |
| 2 | 87,1 | 5,0 | 1,6 |
| 3 | 19,7 | 6,6 | 35,1 |
| 4 | 38,5 | 7,4 | 20,2 |
| Mean | 16,9 | 6,0 | - |

The second cluster consists of municipalities that "use the small they are provided" because with low level of public transport potential, these commuters use public transport on average in 87% of all journeys to work. Anyway this cluster is very small and contains less than 2% of all municipalities and does not created any larger cluster. Totally different situation is evident in case of third cluster. These municipalities could be named as "refusing public transport" because the average real use does not correspond with quite high potential. This cluster is the second biggest with 35% of all municipalities and is distributed all over the Czechia. More than 20% of municipalities "relying on public transport" have been classified to the fourth cluster. These municipalities have very high public transport potential and high public transport use for commuting. These are concentrated mainly in the eastern part of Czechia, around Prague and in areas close to national borders. All clusters are significantly different.

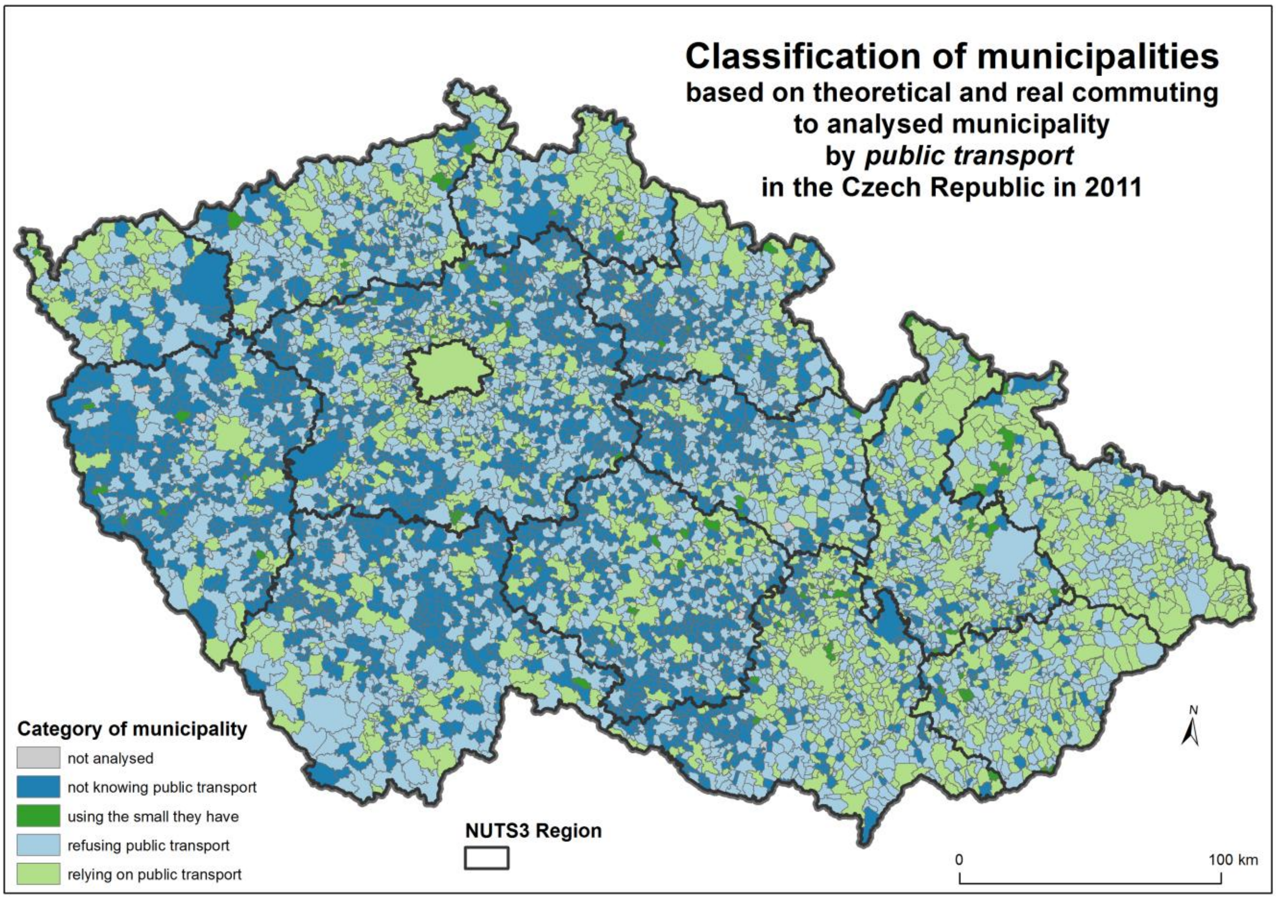



**Fig. 5.**
*Classification of municipalities based on theoretical and real commuting*
*Source: authors*

Additionally the same situation as previously extended about individual transport has been analysed too. K-mean clustering method was used and all municipalities have been divided into four clusters based on potential and real transport flows by individual and public transport. The first cluster contains only 7% of municipalities that could be named as public transport oriented with lower car ownership index but cars are also not so often used for commuting. On the other hand they have higher potential of public transport and buses and trains are very often (in almost 58%) used for commuting. Mainly big cities belong to this cluster e.g. Prague, Brno, Ostrava and other smaller cities in the eastern part of Czechia. More than 40% of all municipalities are between the first and third cluster belong to second cluster. These municipalities have lower car ownership index but higher use of cars for commuting. Regarding public transport, they are typical due to higher potential of public transport use but unfortunately commuters are not so often using it. They are concentrated mainly in Moravia (eastern part) and in north-western parts of Czechia.

**Table 2**
*Final cluster centres for public transport, individual transport and share of municipalities*

| Cluster | Public transport | | Individual transport | | Share of municipalities in % |
|---|---|---|---|---|---|
| | Real use in % | Potential use in % | Real use in % | Potential use in % | |
| 1 | 57.7 | 6.4 | 42.3 | 52.2 | 7.1 |
| 2 | 26.6 | 7.2 | 73.4 | 53.5 | 40.1 |
| 3 | 0.0 | 4.2 | 0.0 | 60.1 | 8.6 |
| 4 | 4.7 | 5.2 | 95.3 | 59.9 | 44.1 |
| Mean | 16.9 | 6.0 | 74.5 | 56.8 | - |

Third cluster consists of more than 8% of municipalities that could be considered as excluded from transport due to minimal use of public and individual transport. These municipalities have very low public transport use and very high individual transport potential due to high car ownership index. Due to small size of this cluster there is no evident spatial clustering but generally members are situated in central parts of Czechia and close to regional borders. The biggest cluster is oriented heavily on individual transport. They are typical by very high car ownership index and also by very high use of car transport for commuting. Public transport is not used very often (less than 5% of journeys) possibly due to lower theoretical potential of public transport. Members of this cluster are concentrated mainly in the western parts of Czechia as can be seen in map below (figure 6).

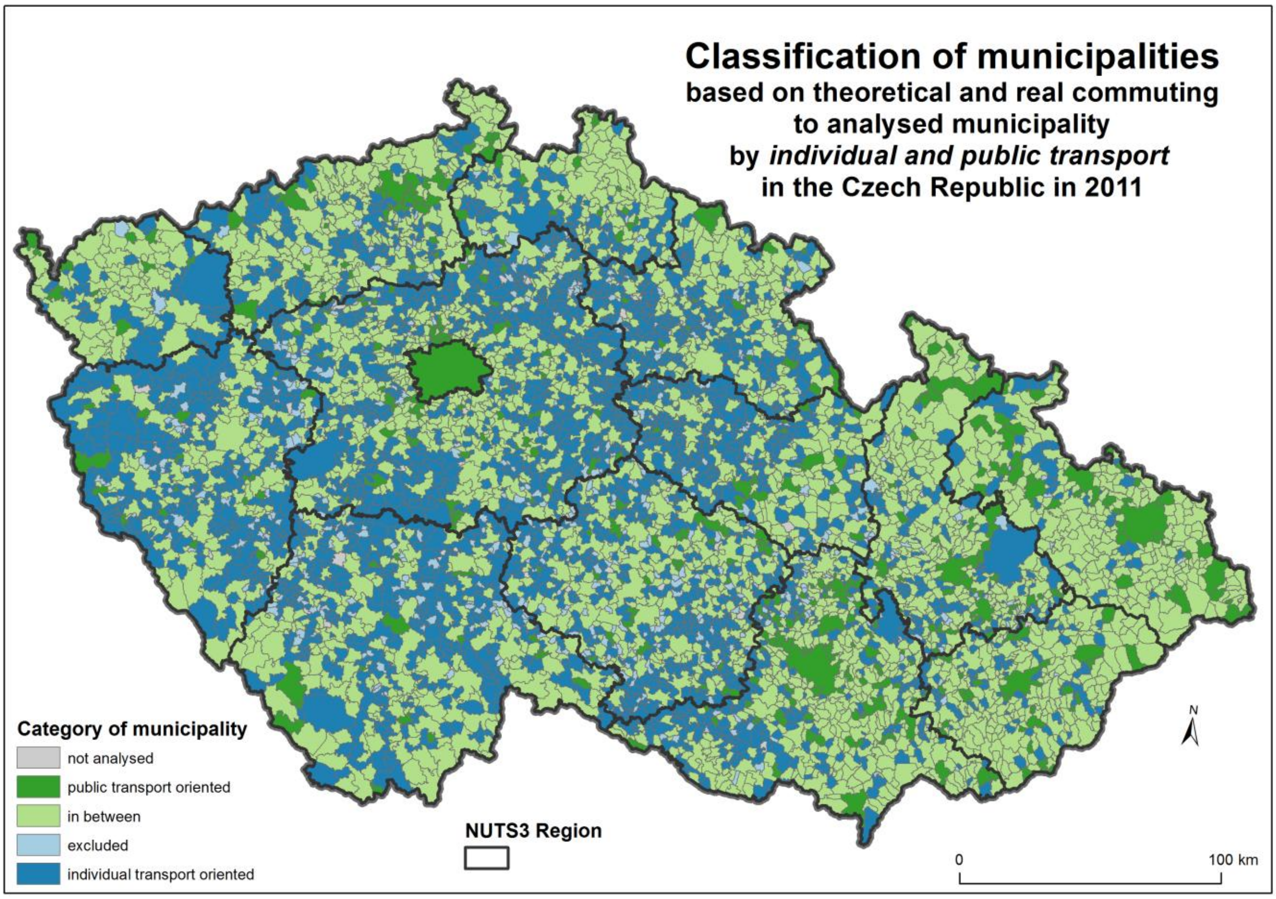



**Fig. 6.**
*Classification of municipalities based on theoretical and real commuting by public and individual transport*
*Source: authors*

## 6. Conclusion

Various data sources have been utilised for comparison of theoretical transport potential for commuting by public transport and individual transport and real commuting by both transport modes. Besides data form census from 2011 describing commuting that are very popular in transport (especially commuting) studies also time tables have been used to asses public transport potential of particular municipalities in Czechia. Spatial distribution of this public transport potential embodies spatial trend with increasing potential form west to east and are significantly spatially autocorrelated with several spatial clusters – hot and cold spots. While hot spots are located in Moravia (eastern part) cold spots are situated mainly in Bohemia (western part). Real usage of public transport for commuting has quite similar spatial trend as in case of theoretical potential. So higher rates in Moravia and sparsely distributed cold spots in Bohemia. Nevertheless, spatial concentration of higher or smaller rates of public transport use is smaller what has been proved by small (still significant) Moran's *I*. All municipalities have been classified using firstly only data describing potential and real use of public transport and secondly in addition to previous analysis potential and real use of individual transport have been added. As result of both analyses, municipalities have been classified into four clusters. Municipalities where commuters do not use public transport because they do not want to or they cannot because of its poor potential are mainly in Bohemia. While municipalities relying on public transport are in Moravia, in big cities, and close to national borders. Last analysis has divided municipalities in public transport oriented, individual transport oriented, somewhere in between and excluded from transport.

This paper provides innovative approach how combine various data sources describing transport potential and real transport behaviour. As the result it allows to find areas with lower level of potential for public transport use but also areas where regardless a good level of potential, real use remains at very low level.

## Acknowledgements

The research is supported by the Czech Science Foundation, project Spatial simulation modelling of accessibility, No. 14-26831S.

## References

Anselin, L. 1995. Local Indicators of Spatial Association–LISA. *Geographical Analysis*, 27:2, 93–115.

Burian, J.; Zajickova, L.; Tucek, P.; Vozenilek, V.; Langrova, B.; Boori, M. 2014. Traffic intensity changes and their influence on spatial distribution of suburbanization. *14th International Multidisciplinary Scientific GeoConference SGEM 2014.*

Horák, J.; Ivan, I.; Fojtík, D.; Burian, J. 2014. Large scale monitoring of public transport accessibility in the Czech Republic. In *Proceedings of ICCC 2014, Velké Karlovice, 28-30.5.2014,* pp. 1-7.

Ivan, I.; Horák, J.; Fojtík, D.; Inspektor, T. 2013. Evaluation of Public Transport Accessibility at Municipality Level in the Czech Republic. *13th International Multidisciplinary Scientific GeoConference SGEM 2013*, vol. 1, 1088 p.

MoT. 2014. *Transport Yearbook 2013*. Ministry of Transport. Available from Internet: <https://www.sydos.cz/cs/rocenka-2013/yearbook/htm_uk/index.html>.

Smith, M. J. 2014. *Statistical Analysis Handbook*. Available from Internet: <http://www.statsref.com>.

Novak, J.; Ahas, R.; Aasa, A.; Silm, S. 2013. Application of mobile phone location data in mapping of commuting patterns and functional regionalization: a pilot study of Estonia. *Journal of Maps*, 9:1, 10–15.

van VUGT, M.; van LANGE, P. A. M.; MEERTENS, R. M. 1996. Commuting by car or public transportation? A social dilemma analysis of travel mode judgements. *European Journal of Social Psychology*, 26:3, 373–395.

# CARPOOLING IN MUNICIPALITY OF GORNJA STUBICA

**Marko Slavulj**[1] **, Davor Brčić**[2] **, Dino Šojat**[3]
[1 ,2 ,3] *Faculty of Transport and Traffic Sciences, University of Zagreb, Croatia*

**Abstract:** Carpooling is an alternative approach to private car usage which encompasses various positive effects, such as reduction of traffic congestion, lower greenhouse gas emissions and lower operational costs of private cars. The goal of this paper is to obtain information regarding usage habits of carpooling in the Municipality of Gornja Stubica, which is influenced by the closeness of the City of Zagreb. For this purpose, a survey including 100 respondents was conducted. The survey covered the following: modal split of passenger trips, distance between residence and workplace, income of subjects, carpool usage, motives to use carpool and vehicle occupancy. Also, the cost-benefit analysis was conducted in order to compare carpooling opposite to public transportation (by train and bus) and private car (without carpooling). Based on the research conducted in this paper, solutions to increase quality of service as well as number of carpool users were proposed.

**Keywords**: carpooling, sustainable mobility, Gornja Stubica, cost-benefit analysis, vehicle occupancy

## 1. Introduction

Many European and worldwide countries face with large oscillations in transport demand and transport networks face high densities every day. Since the cities are characterized by increasing motorisation rate, the share of private car in the modal split is constantly increasing. Problems in the transport network such as congestion in the urban centres then become inevitable. There are many negative effects arising from traffic congestion - increased travel time, pollution, noise, higher traffic accident rate, higher external costs and inadequate car parks have a large ratio in an ordinary transportation system (Štefančić, 2008).

The solutions for such transportation problems in the scope of private car users are forms of trip sharing, in the world known as carpooling, car sharing, ride sharing, etc. Among them, carpooling is becoming an interesting and popular option for achieving many positive effects influencing traffic flow. Before the introduction of carpooling, it is necessary to define the elementary parameters of private cars affecting the proposed system. The introduction of a unique method for pairing city trips, citizens and reservations is also necessary. The process of trip pairing would have to be done in real-time, with an option to book the trip in advance through modern communication systems, to assign priorities in traffic (by special traffic lanes, parking priorities, etc.) and to determine starting points according to the user requirements (for commuters).

## 2. Carpooling user experience in the municipality of Gornja Stubica

Location of the municipality of Gornja Stubica is shown on Fig. 1. Located between Stubičke Toplice and Marija Bistrica on the northern slopes of the Medvednica Mountain, and along the southern edge of the Krapina-Zagorje County, the municipality of Gornja Stubica has a surface area of 50 km$^2$ and a population of 6000 inhabitants (SSOGS, 2014).


**Fig. 1.**
*The location of the municipality of Gornja Stubica*
*Source: Google Maps*

[1] Corresponding author: marko.slavulj@fpz.hr

The easiest way to approach Gornja Stubica is by car – there are 3 different possibilities:
- Using the A2 Highway (Zagreb – Macelj) to node Zabok, and the through Oroslavje, Stubičke Toplice, Donja Stubica and Gornja Stubica;
- Using the old road through Podsused and Gornja Bistra in the City of Zagreb followed by the Stubičke Toplice, or through Donja Bistra, Jakovlje and Stubičke Toplice;
- Through Stubički Laz, Kašina and Sesvete.

Gornja Stubica can also be accessed from the northern part of Hrvatsko Zagorje through Zlatar, Zlatar Bistrica and Marija Bistrica. The shortest path from Gornja Stubica to Zagreb in one direction has a length of 42 km, while the alternative routes can reach the length of 72 km.

### 2.1. The objectives of the carpooling survey

The objectives of the survey are the following:
- To get the insight about carpooling as a viable mode of transport and introducing carpooling to the residents of Gornja Stubica;
- To know what is the common opinion about carpooling and what are the most common habits and the mode of transport among commuters;
- To familiarize the residents with benefits of carpooling and how to get savings from it if implemented in private cars;
- To familiarize the residents with the advantages of carpooling as well as disadvantages that can occur while arranging a carpool, and how to avoid or eliminate them;
- To deal with stereotypes common for the observed area related to freedom and security of passengers travelling with strangers;

To present how to get in contact with strangers and how to carpool with them.

### 2.2. The results of the carpool survey

The survey was conducted on 100 participants from Gornja Stubica and its surrounding areas – 81% were male, and 19% female. Only 3% do not possess a driver's license. Participants are divided into age groups as follows: 48% between 25 and 34, 22% between 35 and 44, 13% between 18 and 24, 9% between 55 and 64 and 8% between 45 and 54. There were no participants younger than 18 and older than 65. The ratio of unemployed people is 8%. By comparing their monthly incomes, participants are divided into following groups: 49% between 5.001 HRK and 7.000 HRK, 32% between 2.801 HRK and 5.000 HRK, 7% has a monthly income less than 2.800 HRK, 3% have the income larger than 10.000 HRK and 1% have the income between 7.001 and 10.000 HRK. The mode of transport mostly used to commute (Fig. 2) is private car (55%), and the percentage becomes 79% if the participants use private car combined with train and streetcar (13%) or with train and bus (11%). There are also 10% of the participants who use only train and 3% of the participants who use only bus.

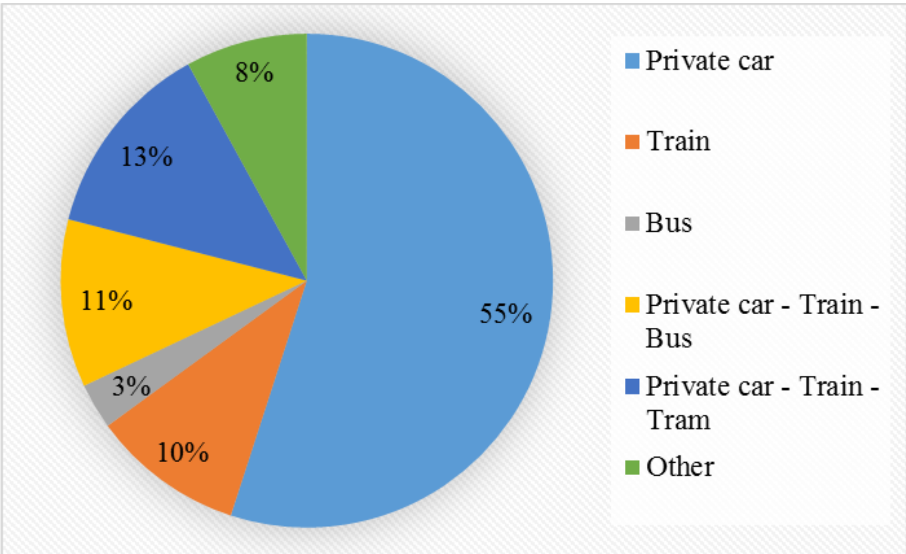



**Fig. 2.**
*Mode of transport*
*Source: created by authors*

Distance between place of residence and workplace is between 40 and 50 kilometres among the majority of the participants – 48%, which is common for the wider area of the City of Zagreb. It can be noted from further analysis that the majority of the residents have a workplace within the City of Zagreb or the Zagreb County. Furthermore, 17% travel less than 10 km, 5% between 10 km and 20 km, 11% between 20 km and 30 km, 9% between 30 km and 40 km, and only 2% of the participants travel more than 50 km.

Carpooling is currently used by 16% of the participants, 10% are former carpool users, and 74% have never used a carpool service (Fig. 3).

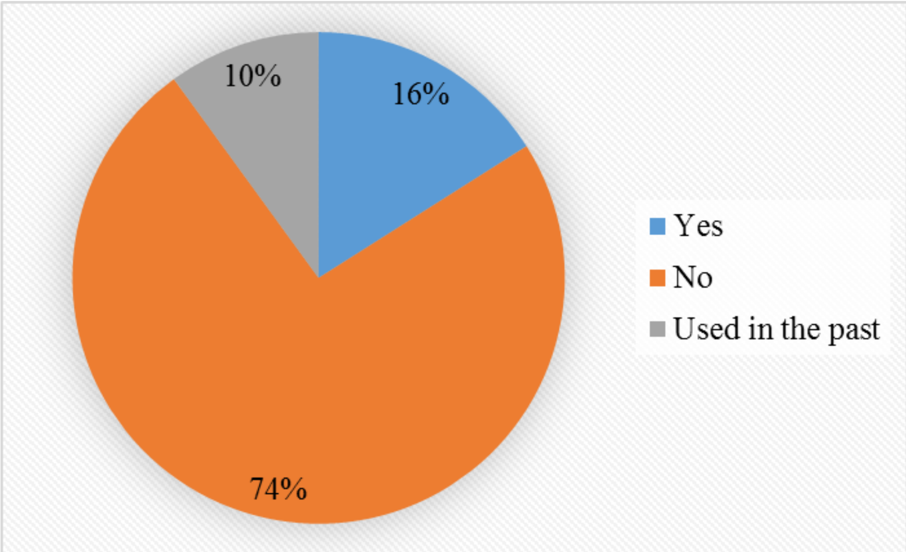



**Fig. 3.**
*Carpooling experience*
*Source: created by authors*

Due to the multiple answer option in the survey, the majority of the participants which use or have used carpooling state that the reasons for its usage were the following: fuel savings while commuting (65%), possibility to commute by car (people who don't own a private car – 19%) and the combination of commuting, fuel savings and socializing (16%).
Fig. 4 shows monthly expenses of each driver while commuting alone (blue), while using carpooling (red) and savings (green), which is the difference between values in the first two cases. The participants who have been commuting alone had monthly expenses between 400 HRK and 2.000 HRK. After the introduction of carpooling, their monthly expenses became between 150 HRK and 1.000 HRK, which are the savings between 250 HRK and 1.000 HRK.

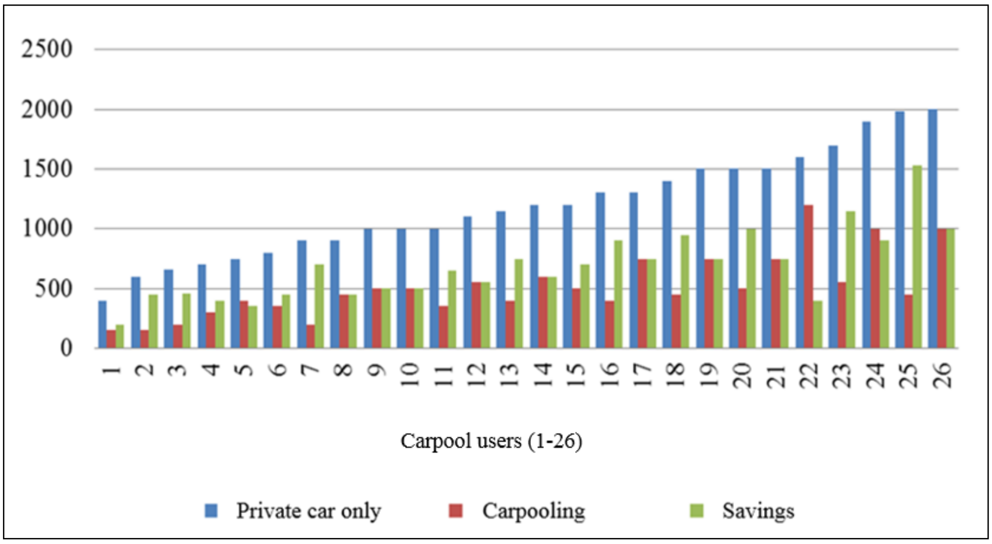



**Fig. 4.**
*Cost comparison between conventional trips, carpooling and savings among survey participants*
*Source: created by authors*

Vehicle occupancy in the survey was the following: two persons in 50%, three persons in 34% and four or five persons in 8% of the total number of carpools (Fig. 5).

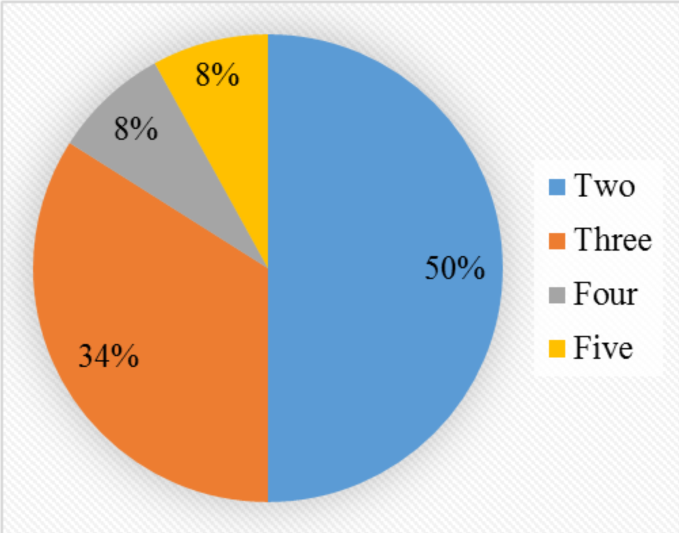



**Fig. 5.**
*Vehicle occupancy*
*Source: created by authors*

Among the participants, 81% are drivers and 19% don't drive. It can also be seen that the 54% of the participants exchange the driving order daily in their carpools, 46% exchange according to an agreement between them, and there were no cases of weekly exchanges. All the participants belong to the familiar category, i.e. category in which all the passengers know each other (mostly friends or neighbours) – there were no carpools of the unfamiliar category, and neither there were carpools of the employees working in the same company. The survey also gave information about the carpool arrangements – 81% of the participants make arrangements within a carpool, and 19% use a cell phone. Web portals weren't used among the participants in the survey. A free parking space at the destination can be offered to 89% of the participants (11% don't have the possibility to acquire a parking space).

### 3. Cost-benefit analysis of carpooling

The cost-benefit analysis of carpooling was done on a case-study carpool route Gornja Stubica – Zagreb, including only operating costs, shown on Fig. 6. In this case study route, users commuted daily in one direction crossing a distance of 45 km, or 90 km in both directions, with fuel consumption (Eurosuper BS 95) of 10 litres per 100 km and a price of 10 HRK per one litre of fuel. From the data above, fuel expenses were 90 HRK in both directions on daily basis, 450 HRK on weekly basis and 1.890 HRK on monthly basis.

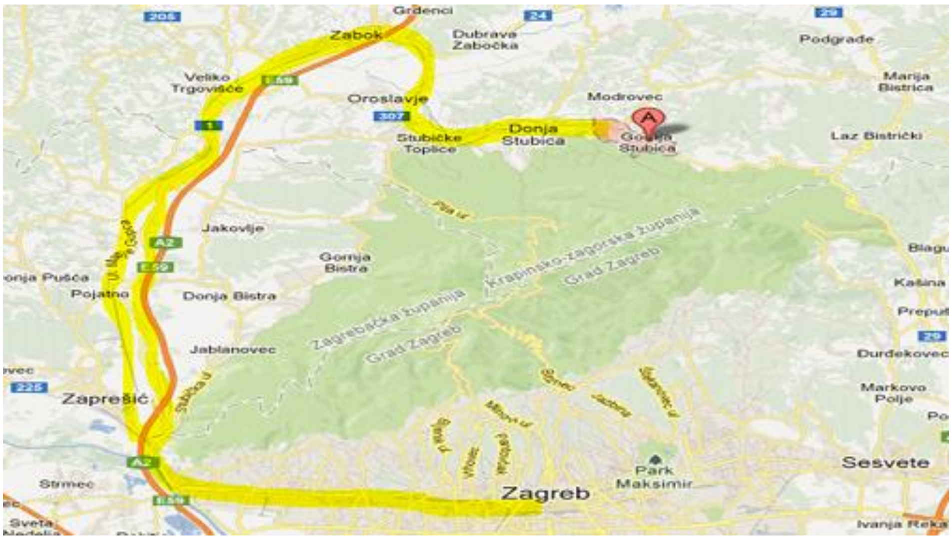



**Fig. 6.**
*Observed carpool route between Gornja Stubica and Zagreb*
*Source: Google Maps*

If carpooling was used, the expenses were significantly lower, so the savings that were achieved were 1.140 HRK on a monthly basis (360 HRK on a weekly basis).

The results of private car and carpooling from the financial aspect (based on operating costs only) can also be compared to the most common alternative modes of transport in the area – bus and train. The analysis is performed on weekly basis and monthly basis (20 workdays). It can also be seen from Fig. 6 that the roads and railways lie on the similar route and that all the major towns towards the City of Zagreb are located on their routes.

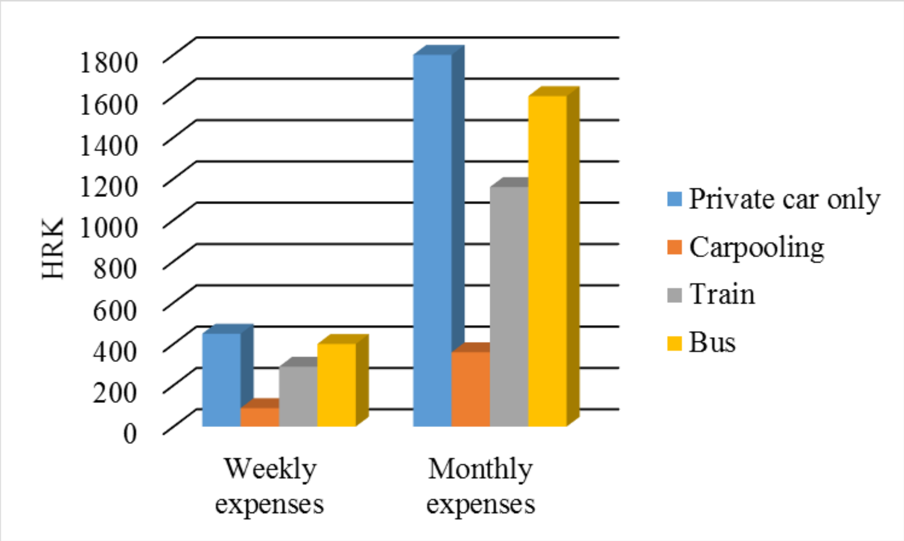



**Fig. 7.**
*Cost comparison for different modes of transport*
*Source: created by authors*

Bus transport between Zagreb and Gornja Stubica is operated by the Presečki Group four times a day during workdays, and once or twice on weekends. During workdays, the bus runs three times in the morning, and only once in the afternoon. The price of a ticket is 52 HRK in one direction, and 80 HRK in both directions. The bus stops at Podsused, Donja Bistra, Stubičke Toplice and Donja Stubica, and the trip lasts for 60 minutes. The expenses of travelling by bus become 1.600 HRK on monthly basis, which is 400 HRK on weekly basis.

Train transport from Zagreb to Gornja Stubica runs 8 times a day – the price of the ticket is 36 HRK in one direction, and 58 HRK in both directions. The train stops at Zaprešić, Veliko Trgovišće, Zabok, Oroslavje, Stubičke Toplice and Donja Stubica, and the trip has a duration of 90 minutes. The expenses of travelling by train are 1.160 HRK on monthly basis, which is 290 HRK on weekly basis.

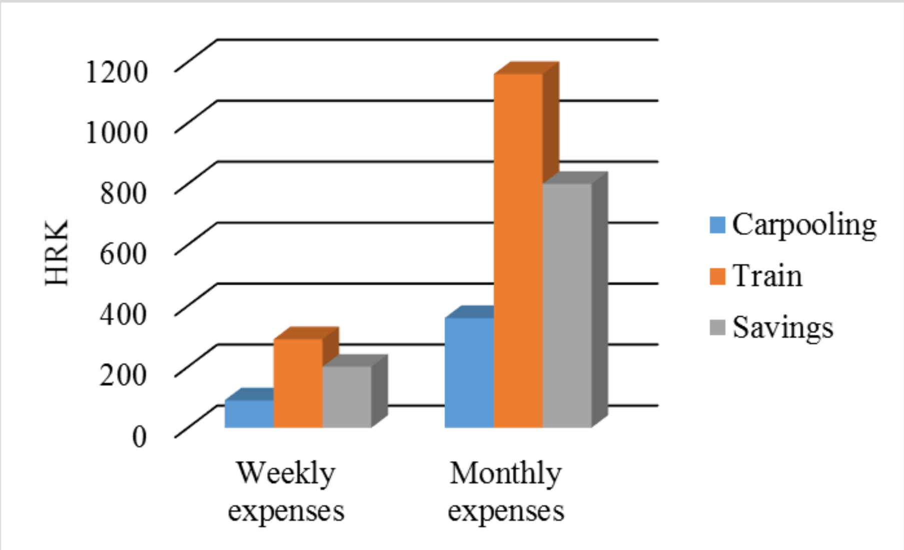



**Fig. 8.**
*Cost comparison between carpooling and commuting by train and savings*
*Source: created by authors*

The financial analysis of the private car, carpooling, bus and train is shown in Fig. 7.
From the Fig. 7 it can be observed that the carpooling has many financial benefits over the other modes of transport. Train transport is currently the most popular mode of transport in the Republic of Croatia (the data is obtained from the State Bureau of Statistics) because of its financial benefits compared to private car or bus, which are currently the most expensive modes of transport.

The difference in expenses between carpooling and train, i.e. savings which can be achieved by using carpooling compared to train is shown in Fig. 8.

## 4. Proposed carpooling solutions in Gornja Stubica

Carpooling as a mode of transport has a very low share in the Municipality of Gornja Stubica and neighbouring areas – it is almost never used. There is only a small portion of residents who use it on larger distances, mostly because they are introduced to it. The residents mostly do not use carpooling because they have a little knowledge of it and by that, they are not aware of its benefits.

The importance of carpooling promotion should be provided to Gornja Stubica and all major towns nearby through social media as radio, billboards in shopping malls, town and municipality centres, and local newspapers. Also, the companies and other related institutions (hospitals, factories, universities, schools, kindergartens, municipal and other authorities) should be encouraged to implement mobility plans for the employees in the form of pilot projects at the beginning. These mobility plans should be maintained for company personnel or individuals who have demands to carpool in the future.

The carpooling system can be promoted through local radio stations such as Radio Gornja Stubica and newspapers such as Gornjostubička Luč, which is the monthly newspaper related to the Municipality of Gornja Stubica and Gupčeva Lipa, which is the journal of the Gornja Stubica. Gornja Stubica and neighbouring areas do not have their specific traffic sign nor the billboard which would serve for carpool passengers. For this purpose, certain locations should be provided of such signs or billboards.

The survey found that the 48% of the participants are aged between 25 and 34 and that the 22% of them are aged between 35 and 44, from which can be concluded that they are capable of using computer and internet. Besides using computers, the internet can be accessed through smartphones and tablet computers which are available nowadays and by that, very suitable for carpool arrangements and finding travellers.

The promotion of carpooling and its improvements can be also done by the frequently visited official web pages of Gornja Stubica.

## 5. Conclusion

At present times, highly influenced by the financial crisis, and very low personal income, carpooling represents a convenience through many different reductions – such of those reductions are related to traffic congestion, air pollution, travel costs and fuel consumption. Also, the feasibility of the carpooling system is not intended for individuals only, but the entire transport for the purpose of commuting between Gornja Stubica and Zagreb and their surroundings in the scope of the sustainable urban mobility.
The goal of the paper was to explore the usage habits of the carpooling in the Municipality of Gornja Stubica and neighbouring areas. The survey conducted in the paper showed that the concept of carpooling is still underdeveloped in the municipality, because the local community and the City of Zagreb do not support this concept. The survey also showed that the 86% of the participants are familiar with the concept of carpooling, but only 26% actually use it. The majority of the participants state that the reason for using carpooling is commuting and fuel savings. Using private car only, monthly expenses of the participants can reach even 2.000 HRK, while using carpooling, those expenses become reduced by 50%, which is the main reason why the most participants see carpooling as a mode of transport with considerable savings. The maximum vehicle occupancy of 5 persons is only slightly accomplished – 50% travels in couples. Since the groups of carpool passengers are usually familiar people, the carpool is then usually formed by personal contact, and passengers exchange roles of drivers every day. From the 26 subjects that were carpool users, 23 of them has been provided a free parking space.

Despite all the advantages, disadvantages and rules associated with carpooling, this mode of transport is still used and has its financial benefits, and because of these facts, it needs to be better presented to the public so that could be used at its maximum capabilities. The presentation of the carpooling should be provided to both the Municipality of Gornja Stubica, the City of Zagreb and all potential areas.

**References**

Brčić, D.; Šimunović, Lj.; Slavulj, M. 2014. Sustainable Urban Mobility Plans in Europe. In *Proceedings of International Academic Conference on Places and Technologies*. Faculty of Architecture, Belgrade, the Republic of Serbia, 889-895

Google Maps. 2014. Availabe from Internet: http://maps.google.hr/

Grgurević, I; Slavulj, M.; Protega, V. 2012. Identification of Risks Regarding the Introduction of Measures Aimed at Improving Citizen Mobility in the City of Zagreb, *Results and lessons learned from the Civitas-Elan project, The Publishing House of the Faculty of Social Sciences, Ljubljana: Založba FDV*, 179-193

Slavulj, M.; Brčić, D.; Šimunović, Lj. 2012. Approach to dealing with the transport demand management in cities with review to the City of Zagreb. In *Proceedings of the 2th International Conference on Road and Rail Infrastructure*. Faculty of Civil Engineering, Dubrovnik, 929-935

Slavulj, M.; Grgurević, I..; Brabec, D. 2010. Vehicle occupancy measurement for the City of Zagreb. In *Proceedings of the 5th International Conference on Ports and Waterways*. Faculty of Transport and Traffic Sciences, Zagreb, 128-135

Slavulj, M.; Grgurević, I.; Golubić, J. 2012. Analysis of possible risks in introducing congestion charging and carpooling measures in Zagreb, *Wessex Institute of Technology, The Built Environment*, 128: 433-444

SSOGS. 2014. Službene stranice općine Gornja Stubica. Available from Internet: <http://www.gornjastubica.hr/o_gornjoj_stubici.htm>

Štefančić, G. 2008. Tehnologija gradskog prometa I (Urban Traffic Technology I). Fakultet prometnih znanosti Sveučilišta u Zagrebu. HR. 222-223 p.

# THE ECONOMIC PERFORMANCES OF SUPPLY CHAIN(S) SERVED BY THE MEGA FREIGHT TRANSPORT VEHICLES

**Milan Janic**[1]
*Department of Transport & Planning, Faculty of Civil Engineering and Geosciences and Department of Air Transport and Operations, Faculty of Aerospace Engineering, Delft University of Technology, Delft, the Netherlands*

**Abstract:** This paper deals with the economic performances of supply chain(s) served by different including the mega freight transport vehicles. These performances are considered as a dimension of the supply chain's sustainability together with the infrastructural, technical/technological, operational, environmental, and social performances. The supply chain(s) consists of the spoke and hub supplier(s) and the hub and spoke consumer(s) of goods/freight shipments. The economic performances include the inventory, handling, and transport cost of goods/freight shipments processed in the chain(s). The analytical model is developed for estimating the above-mentioned economic performances of the generic configuration of supply chain(s) operating according to the specified scenario(s) under given conditions. Then, the model is applied to the intercontinental supply chain exclusively served by the conventional and mega container ships aiming at investigating their effects/impacts on the chain's economic performances.

**Key words***:* supply chain(s), economic performances, analytical models, mega container ships

## 1 Introduction

A general definitions of supply chain states that it is: "the movement of materials as they flow from their source to the end customer including purchasing, manufacturing, warehousing, transportation, customer service, demand and supply planning, and Supply Chain management" (http://www.supplychaindefinitions.com/). In this paper, the supply chain is considered as the physical network, which produces, handles, transports, and receives goods/freight shipments consolidated into TEUs (Twenty Foot Equivalent Unit(s)) between their ultimate suppliers and customers. These can be the large production/consumption plants, distribution centers, sea-ports, airports, large surface modal (rail, road), and intermodal (rail/road/barge) terminals. They usually generate and attract rather substantive flows of these (consolidated) goods/freight shipments and as such operate as the hub nodes of the global (continental and intercontinental) freight transport network(s). In many cases, these substantive flows between particular hub nodes are transported by larger including the mega freight transport vehicles. These mega vehicles are easily recognizable within each transport mode: road – mega trucks, rail/intermodal - long freight trains, air - large cargo aircraft, and sea – large container ships. (http://www.investopedia.com/terms/s/supplychain.asp; http://en.wikipedia.org/wiki/Supply_chain.

This paper deals with the economic performances of supply chain(s), which are served by different including the mega freight transport vehicles. In addition to this, the paper consists of three other sections. Section 2 explains the concept and develops the analytical model for estimating the economic performances of a given supply chain(s). Section 3 presents an application of the proposed model. The last section summarises some conclusions.

## 2 The Concept and Model of Economic Performances of Supply Chain(s)

### 2.1. The concept

In general, the performances of supply chain(s) are considered as its inherent ability to deliver goods/freight shipments between their ultimate suppliers/senders and ultimate customers/receivers under given conditions, usually as planned, i.e., generally efficiently, effectively, and safely. Consequently, as at the similar systems, the performances of supply chains can be classified as infrastructural, technical/technological, operational, economic, environmental, and social (Janic, 2014). Specifically, the economic performances can generally be the chain's inventory, handling, and transportation cost of the goods/freight shipments.

### 2.2. Some previous research

The previous research on dealing directly or indirectly with particular performances of supply chains has been substantive. The research closely related to that presented in this paper addresses the role and influence of transport operations on the entire performances of supply chain(s).

[1] Corresponding author: M.Janic@tudelft.nl

This has mainly dealt with understanding the relationships between the transport and logistics operations and possible improvements through i) the goods/freight shipment(s) delivery speed, quality of service, operating costs, use of facilities and equipment, and savings energy (Tseng et al., 2005), ii) modelling performances of different spatial and operational configurations of the goods/freight collection/distribution networks (Janic 2005; 2014), and iii) understanding the potential interactions between location of European manufacturing industry, related services, and logistics and freight transport (EC, 1999). In addition, the research on an explicit investigation of the effects/impacts of the mega freight vehicles on the performances of supply chains has been scarce. An exception has been elaboration of the infrastructural, technical/technological, operational, economic, environmental, and social performances of these (mega) vehicles such as long intermodal freight trains, road mega trucks, large freight/cargo aircraft, and large container ships on the case-by-case, i.e., vehicle-by-vehicle, basis (Janic, 2014). Consequently, this paper intends to partially fill in this gap by considering the effects/impacts of mega vehicles on the economic performances of supply chain(s).   .

### 2.3. The Objectives and Assumptions

The objectives of the paper are to develop an analytical model for estimating the economic performances of a given supply chain(s) served by different including the mega freight transport vehicles. The model is based on the following assumptions (Daganzo, 2005; Janic, 2005):

- The hub supplier of a given supply chain(s) is ultimately the production location, i.e., origin, of the goods/freight shipments; the hub customer is ultimately their consumption location, i.e., destination;
- The chain(s)'s production/consumption cycle during the specified period of time satisfies the series of successive orders of goods/freight shipments to be transported between the hub supplier and the hub customer exclusively by the different vehicle fleets including the mega one;
- The size of a goods/freight shipment(s) is always less than or at most equal to the payload capacity of a given vehicle serving the chain(s);
- The fleet(s) serving the supply chain(s) consists of vehicles of the same size/payload capacity operating with the same load factor; and
- Exclusive use of a given vehicle fleet implies the "all-or-nothing principle" of serving goods/freight shipments within the chain.

### 2.5. Generic Configuration of a Supply Chain(s)

The generic configuration of a supply chain(s) served by any category of the freight transport vehicles is represented as the H-S (Hub-and-Spoke) transport network whose main nodes are the hub supplier(s) and the hub customer(s) connected by the transport link(s) between them as shown on Figure 1 (a, b). The spokes 'feeding' the hub supplier(s) and those 'fed' by the hub customer(s) are also shown. As can be seen, the inventories of goods/freight shipments take place at the hub supplier(s), the hub customer(s), and along the route between them. Figure 1a shows the case a) of exclusive and Figure 1b the case b) of simultaneous collecting and loading of goods/freight shipments at the hub supplier(s), and their exclusive unloading and distributing at the hub customer(s), respectively. 'Exclusivity' implies that the entire shipment is collected before starting its loading, and the entire shipment is unloaded before starting its distribution. 'Simultaneously' implies that both collecting and loading of goods/freight shipment(s) on the one end and its unloading and distribution on the other end of the chain can be partially or fully carried out at the same time. Is such way, it is possible to manage the inventories of goods/freight shipments and related costs.

### 2.6 Basic Structure of the Model

The economic performances of a given supply chain are considered to be the i) inventory, ii) handling, and iii) transport a) the total and b) the average cost of goods/freight shipment(s) served within the chain. If the size of goods/freight shipment corresponds to the vehicle payload capacity, these costs are determined as follows:

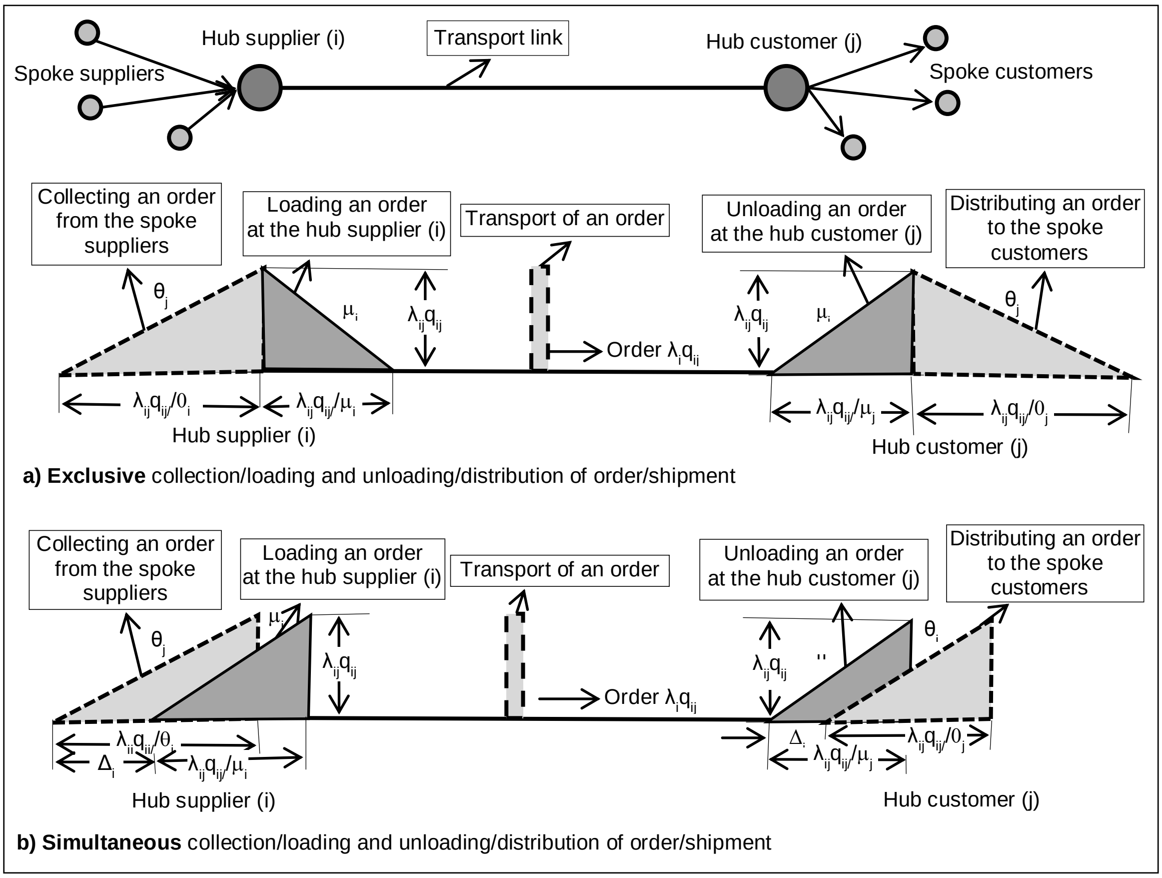



**Fig. 1**

*Simplified scheme of the generic configuration of supply chain(s) (Janic, 2014)*

***a) Inventory cost (€ or $US)***

$$C_{ij/INV}(\lambda_{ij}q_{ij}) = IT_i(\lambda_{ij}q_{ij})*\alpha_i + (\lambda_{ij}q_{ij})*\frac{d_{ij}}{s_{ij}*v_{ij}(d_{ij})}*\alpha_{ij} + IT_j(\lambda_{ij}q_{ij})*\alpha_j \tag{1}$$

The first and third term in Eq. 1 represent the inventory cost of a goods/freight shipment at the hub supplier *(i)* and at the hub customer *(j)*, respectively. The second term represents the inventory, i.e., the shipment's cost of time while in transportation between the hubs *(i)* and *(j)*. From Figure 1, the goods/freight shipment inventory time in Eq. 1 at the hubs *(i)* and *(j)*, respectively, is determined as follows:

$$IT_i(\lambda_{ij}q_{ij}) = \left\{ \begin{array}{l} \frac{1}{2}(\lambda_{ij}q_{ij})^2\left[\frac{1}{r_i\theta_i} + \frac{1}{p_i\mu_i}\right] \quad if \quad a) \\ \max\left\{0; (\lambda_{ij}q_{ij})^2\left(\frac{1}{p_i\mu_i} - \frac{1}{2r_i\theta_i}\right) + (\lambda_{ij}q_{ij})\Delta_i\right\} \quad if \quad b) \end{array} \right\} \tag{2a}$$

and analogously

$$IT_j(\lambda_{ij}q_{ij}) = \begin{cases} \frac{1}{2}(\lambda_{ij}q_{ij})^2\left[\frac{1}{r_j\theta_j}+\frac{1}{p_j\mu_j}\right] \quad if \quad a) \\ \max\left\{0;(\lambda_{ij}q_{ij})^2\left(\frac{1}{p_j\mu_j}-\frac{1}{2r_j\theta_j}\right)+(\lambda_{ij}q_{ij})\Delta_j\right\} \quad if \quad b) \end{cases} \tag{2b}$$

***b) Handling and transport cost (€ or $US)***

$$C_{ij/H-TRA}(\lambda_{ij}q_{ij}) = c_i*(\lambda_{ij}q_{ij}) + c_{ij}(\lambda_{ij}q_{ij})*(\lambda_{ij}q_{ij})*d_{ij} + c_j*(\lambda_{ij}q_{ij}) \tag{3}$$

***c) Total (inventory + handling + transport) cost (€ or $US)***

$$C_{ij}(\lambda_{ij}q_{ij}) = C_{ij/INV}(\lambda_{ij}q_{ij}) + C_{ij/H-TRA}(\lambda_{ij}q_{ij}) \tag{4}$$

***d) Average total cost (€ or $US/TEU-km or ton-km)***

$$\bar{c}_{ij}(\lambda_{ij}q_{ij}) = C_{ij}(\lambda_{ij}q_{ij})/[(\lambda_{ij}q_{ij})*d_{ij}] \tag{5}$$

where

$\lambda_{ij}$, $q_{ij}$ is the average load factor and the payload capacity, respectively, of a vehicles serving the chain *(ij)* (tons, m$^3$, or TEUs per vehicle);

$\Delta_i$, $\Delta_j$ is the time between starting vehicle's loading at the hub supplier *(i)* and unloading at the hub customer *(j)*, respectively (TU);

$d_{ij}$ is the length of transport route between the hubs *(i)* and *(j)* (km);

$v_{ij}$ *($d_{ij}$)* is the vehicle's average (planned) operating speed on the route *($d_{ij}$)* (km/TU or kts (knots); 1 kts = 1nm/h; nm – nautical mile = 1.852km));

$\mu_i$, $\mu_j$ is the loading and unloading rate of a vehicle at the hubs *(i)* and *(j)*, respectively (tons, m$^3$ or TEU/TU);

$p_i$, $p_j$ is the proportion of vehicle's loading and unloading rate used at the hubs *(i)* and *(j)*, respectively ($p_{i1}$, $p_{j1} \leq 1.0$);

$s_{ij}$ is the portion of maintained vehicle's average planned operating speed on the route *($d_{ij}$)* ($s_{ij} \leq 1.0$)

$\theta_i$, $\theta_j$ is the rate of collecting and distributing goods/freight shipments at the hubs *(i)* and *(j)*, respectively (tons, m$^3$ or TEU/TU);

$r_i$, $r_j$ is the proportion of rate of collecting and distributing goods/freight shipments used at the hubs *(i)* and *(j)*, respectively ($r_i$, $r_j \leq 1.0$);

$c_i$, $c_j$ is the handling (loading/unloading/transhipment) cost of a goods/freight shipment at the hubs *(i)* and *(j)*, respectively (€/(ton, m$^3$, or TEU)); and

$\alpha_i$, $\alpha_{ij}$, $\alpha_j$ is the cost of goods/freight shipment inventory time while at the hub *(i)*, in transportation, and at the hub *(j)*, respectively (€/(ton or m$^3$ or TEU)/h or day).

By replacing the shipment size *($\lambda_{ij}q_{ij}$)* by the quantity of goods/freight ($Q_{ij}$) generated during the chain's production/consumption cycle, the corresponding economic performances can be estimated similarly from Eqs. 1-5. In addition, these Eqs. indicate that the goods/freight shipment inventory time and related cost could be compromised in any handling phase in the chain, i.e., during collecting, loading, transporting, unloading, and distributing.

## 3. Application of the Model of Economic Performances of the Supply Chain(s)

### 3.1. The Case

The above mentioned model of economic performances is applied to the case of supply chain between North Europe and Far East Asia served by the liner container shipping. The hub supplier is assumed to be the port of Rotterdam – APM Terminals Rotterdam (The Netherlands) and the hub customer is assumed to be the port of Shanghai – Yangshan Deepwater Port Phases 1/2 or 3/4 (People Republic of China). Currently, this is one of the world's busiest chains (sea trading routes)[2] sharing about 70% of the total trading volumes between Europe and Asia. The container terminals at both ports of the given route enable access and operation of the large container ships including the currently largest Triple E Maersk. The collection and distribution of goods/freight shipments (TEUs) at both ports is carried out by rail/intermodal, road, inland waterway (barge), and feeder (including short-sea) vessel transport modes (Zhang et al, 2009).

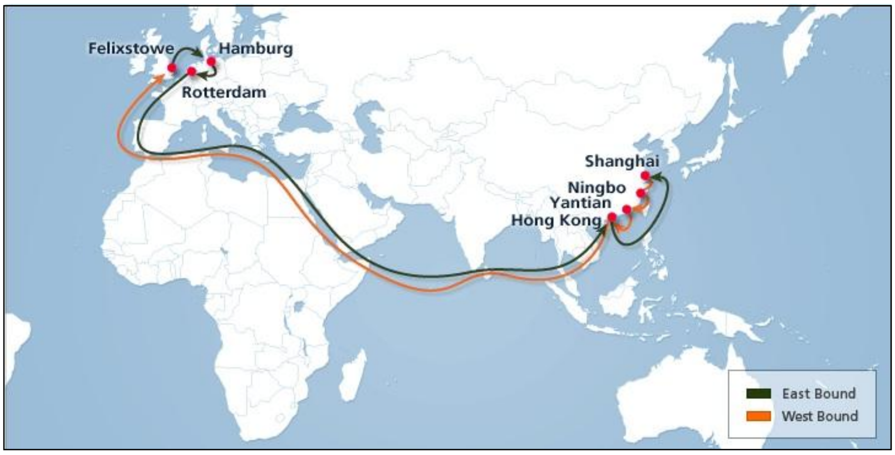



**Fig. 2**
*Simplified scheme of geography of the given supply chain – the liner shipping route Rotterdam – Shanghai (http://www.ship.gr/news6/hanjin28.htm)*

Two scenarios of operating the given chain (route) are considered: the first implies an exclusive use of container ships of the capacity of 4000 TEU (or the current Panamax) and the other an exclusive use of container ships of the capacity of 18000 TEU (i.e., Neo Panamax represented by Triple E class ship started operations by Maersk in the year 2013) The length and width (beam) of the container ships, similarly as their above-mentioned capacity, are specified by design as shown in on Figure 3 (http://www.worldslargestship.com/).

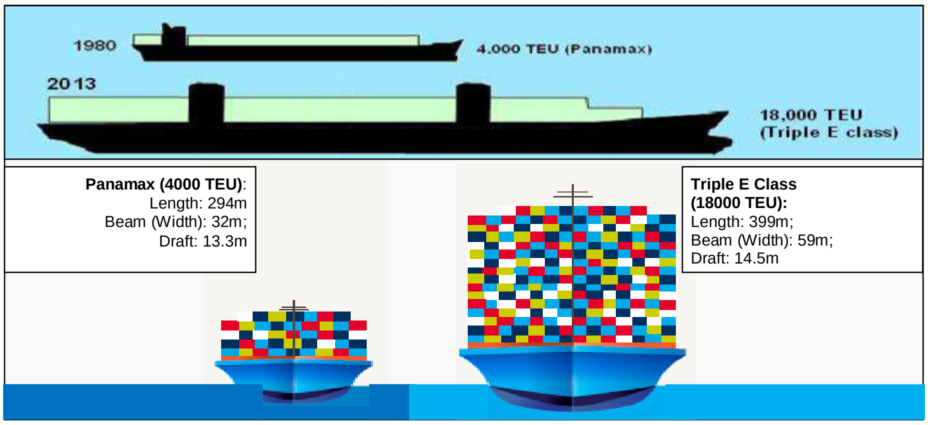



**Fig. 3**
*Scheme of the container ships used in the given supply chain (PR, 2011; http://en.wikipedia.org/wiki/Container_ship)*

[2] This chain (sea trading route) included in the WCI (World Container Index) together with other 10 most voluminous world's container chains (sea trading routes) shares about 35% of their total volumes (TEUs) (http://www.worldcontainerindex.com/).

In both scenarios, the container ships are assumed to operate at the typical slow steaming speed of 20 kts (knots) and the supper slow steaming speed of 15 kts (1kt = 1nm/h; nm – nautical mile) (SCG, 2013). In addition, only direct transportation of the containerized goods/freight shipments in the single direction of the chain is considered.

### 3.2. Input Data

The input data for application of the proposed models to the given supply chain are collected from the case itself and the other different sources and given in Table 1.

**Table 1**
*Input data for application of the models of performances to the given supply chain – liner shipping route Rotterdam (The Netherlands) – Shanghai (China)*

| Input variable | Notation/Unit | Value |
|---|---|---|
| • Container ship capacity | $q_{ij}$ (TEU/ship) | 4000; 18000 |
| • Container ship length | $L_{ij}$(m)/$q_{ij}$(TEU/ship) | 294 (4000); 399 (18000) |
| • Container ship beam (width) | $w_{ij}$(m)/$q_{ij}$ (TEU/ship) | 32 (4000); 59 (18000) |
| • Container ship load factor | $\lambda_{ij}$ | 0.80 (4000); 0.80 (18000) |
| • Collection rate of containers at the hub supplier port | $\theta_i$ (TEU/day) | 1100 |
| • Proportion of used collection rate of containers at the hub supplier port | $r_i$ | 1.0 |
| • Distribution rate of containers at the hub customer port | $\theta_j$ (TEU/day) | 1100 |
| • Proportion of used distribution rate of containers at the hub customer port | $r_j$ | 1.0 |
| • Loading rate of containers at the hub supplier port | $\mu_i$ (TEU/h) | 92 (3-4 cranes)/ 215 (7-8 cranes) |
| • Proportion of used loading rate of containers at the hub supplier port | $p_i$ | 1.0 |
| • Unloading rate of containers at the hub customer port | $\mu_j$ (TEU/h) | 94 (3-4 cranes)/ 215 (7-8 cranes) |
| • Proportion of used unloading rate of containers at the hub customer port | $p_j$ | 1.0 |
| • Time between starting colleting and loading containers at the hub supplier port | $\Delta_i$ (day(s)) | 1 |
| • Time between starting unloading and distributing containers at the hub consumer port | $\Delta_j$ (day(s)) | 1 |
| • Operating distance between the hub ports | $d_{ij}$ (nm) | 10525 |
| • Average operating speed of container ship | $v_{ij}$ (kts) | 20 (Slow steaming) 15 (Super slow steaming) |
| • Portion of the maintained average ship's operating speed | $s_{ij}$ | 1.0 |
| | | |
| • Container inventory cost at the hub ports | $\alpha_i, \alpha_j$ (€/TEU-day) | 124; 124 |
| • Container cost of time in transportation | $\alpha_{ij}$ (€/TEU-day) | 10.6 |
| • Container handling cost at the hub supplier port | $c_i$ (€/TEU) | 185 |
| • Container handling cost at the hub customer port | $c_j$ (€/TEU) | 58 |
| • Container ship operating cost | $c_{ij}$ (€cents/TEU-nm)/ $v_{ij}$(kts)/$q_{ij}$(TEU/ship) | 9.90/20; 5.49/15 (4000) 2.01/20; 1.13/15 (18000) |

| | | |
|---|---|---|

The number of containers (TEU) per chains' production/consumption cycle of duration of one year is determined by assuming the service frequency by the Triple E class ships of 1dep/week, the Panamax class ships of 5depts/week, and the average load factor of both ship classes of 0.80. These give the total annual number of 748800 TEUs to be transported within the chain according the specified scenarios implying using exclusively one class of ships under given conditions. This is, however, only about one sixth of the total annual number of TEUs transported within the chain (http://www.worldcontainerindex.com/).

The rates of collection and distribution of goods/freight shipments (TEUs) are set up regarding the service schedule of different inland transport modes serving the ports (terminals) at both ends of the chain (route) (Zhang et al., 2009). The container loading and unloading rates are set up based on the empirical evidence from both ports/terminals. Two cases are considered: first, both categories of ships are loaded/unloaded by using 3-4 cranes simultaneously (Mongelluzzo, 2013); second, the Triple E class ships are loaded/unloaded by up to seven to eight cranes simultaneously at both ends of the chain (route) (SCG, 2013). All selected crane rates are considered to be fully operational over the period of 24h/day. As well, the shipments are assumed to be completely collected at the hub supplier port before being loaded and completely unloaded at the hub consumer port before being distributed further.

The time between docking and starting loading and unloading of ships at the corresponding ports is chosen as an illustration (This could be reasonable regarding the administrative procedures to be carried out after the ship(s) docks at berths).

The ships are assumed to operate along the route at the constant (slow or super slow steaming) speed(s) without its substantive variations (http://www.sea-distances.org/). This implies that all transport services are assumed to be perfectly reliable, i.e., without delays along the route and consequently at the destination. .

The inventory cost of container(s) during collection and loading at the hub supplier port (Rotterdam) and unloading and distribution at the hub customer port (Shanghai) is estimated based on the average retail value of goods in containers and typical share of the inventory cost (25%) in that value (REM Associates, 2014; Rodrigue, 2013). The cost of container time during transportation is considered as an average for the goods/freight shipments carried out by the sea transport mode (VTI, 2013).

The handling cost of containers at both port terminals is based on the empirical evidence (EC, 2009). The cost of container ship(s) operating on open sea are estimated respecting the effects of cruising/operating speed(s) on the fuel consumption, fuel price (assumed constant), and the share of fuel cost in the total ship's operating costs (Cullinane and Khanna, 2000; Davidson, 2014; Sys et al., 2008; Stopford, 2003).

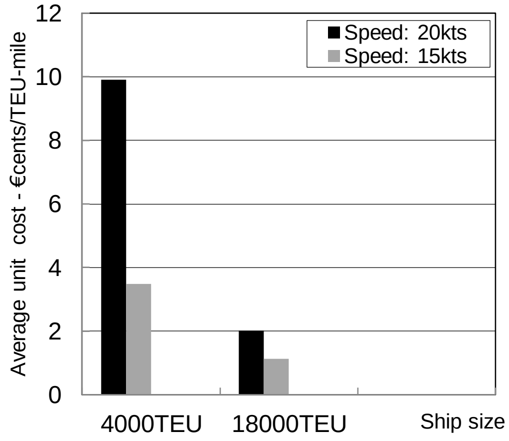
### 3.3. Analysis of Results

The results from application of the model of economic performances to the given case of supply chain based on the input data in Table 1 are shown in Figures 4. Figure 4a show that**,** the relative terms, if exclusively transport cost are considered, the mega ship(s) is for about 5 times more cost efficient than its smaller counterpart(s), while operating on open sea at either steaming speed (20kts or 15kts). This unit cost difference appears to be in line with differences in the ships' size/capacity, thus confirming existence of the substantive economies of scale of the mega container ship(s) under given conditions. Figure 4b shows the total chain's average cost consisting of the inventory and handling cost of collecting/loading and unloading/distributing containers (TEUs) at hub ports, their time cost in transportation, and transport cost. In such case, if the fleet of smaller ships serves the chain, it will be more cost efficient (for about 52% and 79%) than if being served by the fleet of mega ships at either the slow (20kts) and super slow (15kts) steaming speed, respectively. Speeding up of the loading and unloading of the fleet of mega ships at the hub ports decreases this still positive difference for the fleet of smaller ships to about 30% (at slow) and 52% (at super slow) steaming speed. In addition, reducing the steaming speed decreases the chain's average costs much more when served by the fleet of smaller than by the fleet of mega ships, i.e., for about 24% and 1-1.5%, respectively.

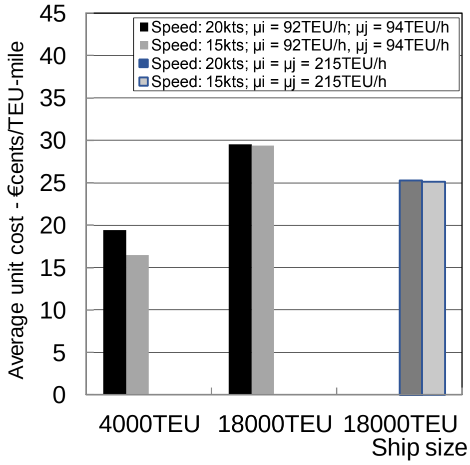
Figure 4c shows that the chain's total average cost decrease by excluding the inventory cost during collecting and distributing containers (TEUs) at the hub ports. This time the chain becomes more cost efficient when served by the fleet of mega ships operating at the slow steaming speed (20kts) (for about 14%). However, the chain becomes less cost efficient (for about 8%) if the fleet of mega ships serves it at the super slow steaming speed (15kts). In case of speeding up the loading and unloading of mega ships at the hub ports, the chain's inventory cost substantively decreases causing decreasing of the total average cost.

Consequently, in this of fast loading/unloading, if all other cost remain unchanged, the chain served by the fleet of mega ships operating at the slow and super slow steaming speed(s) becomes much more cost efficient (62% and 34%, respectively) than in the case when being served by the fleet of smaller counterparts.

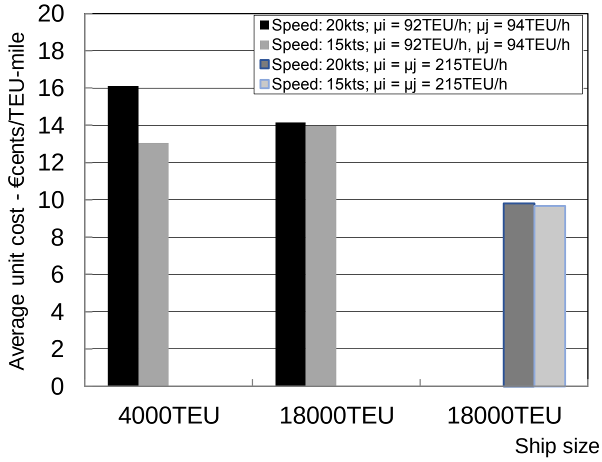
Table 2 gives the structure of the chain's average cost when the inventory cost during collecting/loading and unloading/distribution of containers (TEUs) at the hub ports are included. As can be seen, the share of this (inventory) cost is much lower and the share of transport cost is much higher in the total cost if the chain is served by the fleet of smaller than that of mega ships, independently on their operating speed(s). In any case, reducing the operating speed contributes to increasing of the share of inventory cost on the account of the share of transport cost. Speeding up loading/unloading of the mega ships at the hub ports reduces very little the share of inventory cost compared to that under common loading/unloading .rate(s). Table 3 gives the structure




*a) Average transport (ship) operating cost*




*b) Average chain's cost including the inventory cost during collecting/loading and unloading /distributing containers (TEUs)*




*c) Average chain's cost including only the inventory cost during loading and unloading containers (TEUs)*

**Fig. 4**
*Economic performances of the given supply chain*

**Table 2**
*Structure of the total cost of given supply chain: - The inventory cost during collecting/loading + unloading/distributing containers (TEUs) included*

| **Operating characteristics** | **Container ship capacity (TEU)** | | |
|---|---|---|---|
| | **40**<br>**00** | **18**<br>**000** | **1**<br>**8000** |
| Loading/Unloading rate (TEU/h) | 92<br>/94 | 92<br>/94 | 2<br>15/215 |
| Operating speed (kts) | 20<br>/15 | 20<br>/15 | 2<br>0/15 |
| **Cost component (%)** | | | |
| Inventory | 38<br>/49 | 85<br>/88 | 8<br>3/86 |
| Handling | 12<br>/14 | 8/<br>8 | 9<br>/9 |
| Transport | 50<br>/37 | 7/<br>4 | 8<br>/5\ |

of the chain's cost when only the inventory cost during loading and unloading of containers (TEUs) at the hub ports is taken into account.

**Table 3**
*Structure of the total cost of given supply chain: - The inventory cost During loading + unloading containers (TEUs) included*

| **Operating characteristics** | **Container ship capacity (TEU)** | | |
|---|---|---|---|
| | **400**<br>**0** | **1**<br>**8000** | **1**<br>**8000** |
| Loading/Unloading rate (TEU/h) | 92/<br>94 | 9<br>2/94 | 2<br>15/215 |
| Operating speed (kts) | 20/<br>15 | 2<br>0/15 | 2<br>0/15 |
| **Cost component (%)** | | | |
| Inventory | 24/<br>36 | 6<br>9/75 | 5<br>6/64 |
| Handling | 14/<br>18 | 1<br>6/17 | 2<br>4/23 |
| Transport | 62/<br>46 | 1<br>5/8 | 2<br>0/13 |

As can be seen, by excluding the inventory cost during collecting and distributing of containers (TEUs) at both ports, the share of this cost substantively decreases and the share of transport cost increases independently on the class of ship fleet serving the chain. However, the share of the former (inventory) cost remains much higher and the share of the latter (transport) cost remains much lower in case when the chain is served by the fleet of mega ships than in case when it is served by its smaller counterpart. In this case, reducing of the ships' operating speed also contributes to increasing of the share of inventory cost in the total chain's cost.

## 4. Conclusion

This paper has developed the analytical model for estimating the economic performances of supply chain(s) served by different classes of the freight transport vehicles including the mega ones. The model has been applied to the case of the intercontinental supply chain served by the liner shipping according to the specified scenarios of exclusively using: i) nominal container ships (i.e., the Panamax class of the capacity of 4000TEU (Twenty Foot Equivalent Units)), and ii) the mega container ships (i.e., the Triple E Class of the capacity of 18000TEU). The results from application of the model have shown the following effects of using the fleet of mega container ships on the chain's performances in the given case:

- Significantly lower transport (operational) cost of mega ship(s); and
- Substantively higher the average total cost of the chain served by mega ships due to dominance of the inventory cost, which otherwise can be reduced by speeding up collecting, loading, unloading, and distributing (i.e., handling) of goods/freight shipments (containers - TEUs) at the chain's hubs (This (inventory) cost tends to increase by reducing the ship operating speed of either class due to extending transport time).

These results clearly indicate that the liner shipping companies deploying mega ships should be prepared to pass a part of their benefits (profits) gained thanks to the much lower transport costs to the shippers and receivers in order to compensate their much higher inventory and handling costs of goods/freight shipments. Similar happens at the supply chains served by the mega vehicles operated by other transport modes – road, rail, and air, thus making effects/impacts of the mega vehicles on the economic performances of supply chain(s) they serve generous. .

**References**

Cullinane K., Khanna, M, 2000. Economies of Scale in Large Container Ships. *Journal of Transport Economics and Policy*. Vol. 33, Part. 2, pp. 185-208

Daganzo, C. 2005. *Logistics Systems Analysis*. 4th Edition, Springer Berlin Heilderberg, New York

Davidson, N. 2014. .Global Impacts of Ship Size Development and Liner Alliances on Port Planning and Productivity. *IPAH Mid-term Conference for Planning and Investment*, April 2014, Sidney, Australia

EC.1999. *Freight Logistics and Transport Systems in Europe*. Executive Summary Euro-CASE, European Council of Applied Sciences and Engineering, Paris, France

EC.2009. *Terminal Handling Charges During and After the Liner Conference Era*. European Commission, Publication Office of the European union, Luxembourg

Janic, M, .2005. .Modelling Performances of Intermodal Freight Transport Networks. *Logistics and Sustainable Transport*. Vol. 1, No. 1, pp. 19-26

Janic, M. 2014. *Advanced Transport Systems: Analysis, Modelling, and Evaluation of Performances*. Springer, UK

Mongelluzzo, B.2013. Bigger Ships and Tighter Supply Chains Shine a new Light on Port Productivity and Its Importance to Shippers. *The Journal of Commerce*. www.joc.com, pp. 11-17

PR.2011.*Port Vison 2030*. Prepared for the City Council of Rotterdam, Rotterdam, The Netherlands

REM Associates. 2014. *Methodology of Calculating Inventory Carrying Costs*. REM Associates Management Consultants, Princeton, New Jersey, USA

Rodrigue, P. 2013. Container Shipping Costs and Cargo Value. *Geography of Transport Systems*. https://people.hofstra.edu/geotrans/eng/ch3en/conc3en/table_containershippingcosts.html

SCG. 2013. Global Supply Chain News: Maersk Triple E Cost Advantages are Too Great to Ignore. *Supply Chain Digest*. http://www.scdigest.com/index.php

Sys, C., Blauwens, G., Omei, E., Van de Vooede, E., Witlox, F. 2008. In Search of the Link between Ship Size and Operations. *Transportation Planning and Technology*. Vol. 31, No. 2, pp. 435-463

Tseng, Y. Y., Yue, L. W., Taylor, P. M. 2005. The Role of Transportation in Logistics Chain. *Proceedings of Eastern Asia Society for Transportation Studies*. Vol. 5, pp. 1657 - 1672

VTI.2013. *Value of Freight Time Variability Reductions: Results from a Pilot Study for the Swedish Transport Administration*. VTI notat 39A–2013, The Swedish National Road and Transport Research Institute (VTI), Stockholm, Sweden

Zhang, M., Wiegmans, B., Tavasszy, L. A. 2009. A Comparative Study on Port Hinterland Intermodal Container Transport: Shanghai and Rotterdam. *The Fifth Advanced Forum on Transportation of China (AFTC).* 17th of October, Bejing, China

http://www.supplychaindefinitions.com/

http://www.investopedia.com/terms/s/supplychain.asp;

http://www.ship.gr/news6/hanjin28.htm

http://www.sea-distances.org/

http://www.worldcontainerindex.com/

http://www.worldslargestship.com/

# SUSTAINABLE TRANSPORT IN CITIES, LEARNING FROM BEST AND WORST PRACTICE?

**Anders Langeland**[1]

[1]*Department of Industrial Economics, Risk Management and Planning, Faculty of Science and Technology, University of Stavanger, Norway*

**Abstract:** It is generally recognized that a shift from car use to public transport, walking and cycling will reduce both local and global emissions. The goal of this paper was firstly to identify the European cities with the lowest and the highest car use, and secondly, to discuss *if* and *how* other cities can learn from the Best and Worst practice.

The Top Twenty of cycling cities in Europe is presented. Münster in Germany tops the list and is the European Bike City. The new TEMS EPOMM database makes it easy to compare modal split in cities and thus answer questions like: Which city is most transport sustainable? Which city is the most car dependent? It is not surprising that they cycle a lot in Dutch cities, nor in Copenhagen, possibly more surprising that the bike is used extensively in Berlin? Milton Keynes has indeed become a car city, 75 percent of all trips is with a car. A line can be drawn through Europe. North of the line one find "cycling countries" and south of the line "car countries". Belgium is divided in two. Cities in the Flemish part to the north cycle a lot, while in Valona they hardly cycle. The paper discusses some of the findings and gives tentative answers to these differences between countries and cities.

**Key words:** Top sustainable transport cities. Modal split. Comparison of cities. Learning from best and worst practice.

## Introduction

The Climate Challenge has led to an active interest for sustainable urban transport, from the EU level, through states and down to municipalities and businesses. Integrated *Land Use and Transport Planning* seeks to intensify land use, prevent sprawl and shift people from cars to public transport, walking and cycling. *"Cycling is good for you, your city and the environment"* – the new mantra says. Many cities are therefore trying to achieve a shift from the car towards cycling, walking and public transport. International studies indicate that health benefits of cycling are substantially larger than the risks of cycling relative to car driving. For the society as a whole, this can be even larger because there will be a reduction in air pollution emissions and eventually fewer traffic accidents as a result of reduced car traffic (de Hartog, J. J. et. al., 2010).

Some cities like Freiburg, Copenhagen and Amsterdam are presented as front runners in achieving the shift from car to more environmentally friendly modes. Cycling cities have become a hot topic in marketing and very often one can find unreliable information as in Virgin Flight Magazine where Sandnes, Norway, was placed 7$^{th}$ in the world, but in fact only 4% cycle. So what is the true picture of cycling in European cities? This paper has used the TEMS EPOMM database to rank cities. Which is the World Bike City? Which Capital is best on cycling? Which city is most car dependent? The different rankings are discussed and some tentative answers to why countries and cities differ so much are given. The paper concludes that there is much to learn, but implementing other cities success is not straightforward.

## Method

The comparisons of cities across country borders have been fraught with difficulties: lack of data, privately owned data, non-comparable data and not least "political data" (some glossed – some hidden) used for propaganda and as strategic misrepresentation. Strategic misrepresentation to support certain political goals, to "sell" the city, obtain money or favors from higher administrative levels, and/or obtain political goodwill is widespread, neither is deliberate lying to support a certain project or policy uncommon (Flyvbjerg 1998). There are difficulties in measuring modal shift due to lack of data and no common standard of measurement. Very often the data was produced for a specific purpose e.g. city branding, instead of monitoring the changes over time and evaluating the policy. (Langeland 2009)

The recently established database TEMS a huge step forward. TEMS – The EPOMM Modal Split Tool, is a database that was made with the support of intelligent Energy Europe in the project EPOMM-Plus. Since the start of TEMS in May 2011 initiatives have been taken to collect city data. There is still not a common standard or definition for: metropolitan area, urban area, city/municipality/district area, they differ both in size and population according to source. Trip purpose varies, some estimate all trips, some only business trips, while some only work journeys. Data collection methods are several: travel surveys (national or local), cordon counts, mode-counts. TEMS is a splendid tool which hopefully will promote more data from more cities and with standardized survey methods and quality controls.

## The Good, the Bad and the Capital Cities

### The Top Twenty Cycling Cities

Figure 1 shows the top twenty Bike Cities in the world. Münster in Germany is the cycling city of the world with 38% of all trips on the bike. Four of the top cities have a cycling share over 30 percent. The population data has come from the TEMS database and might or might not be the population that the mode split data covers.

The table is limited to cities in which the car share is below 50%, hence some cities with both a high bike share and high car might be exempted.

Seven of the twenty cities are German, five are Dutch, three Swedish and two from Denmark. One city is included in the table from Italy, Switzerland and Belgium.

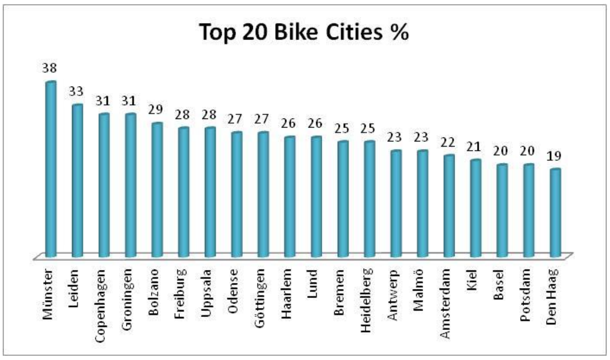



**Figure 1**
*Top 20 Bike Cities*

The data can be sorted and looked at from different angles, for example top walking cities, top public transport cities. The sum of walking and cycling is presented as top environmentally sustainable cities in this paper as shown in the following figure.

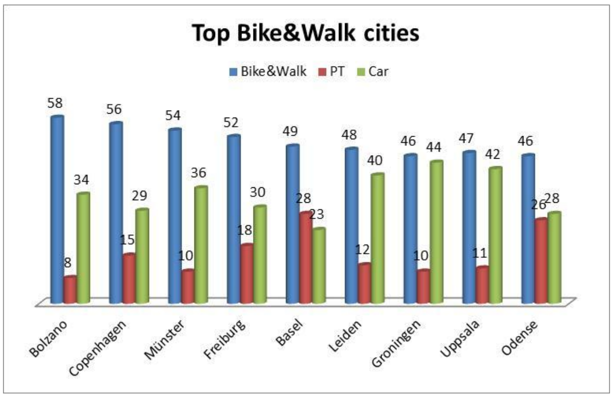



**Figure**
*2 Top Bike & Walk Cities*

The Figure 2 shows a ranking according to the highest sum of walking and cycling. All cities have more than 45% bike/walk share. In addition one should look at low car use and then Basel comes out as best with only 23% car share. Bolzano the Italian city which has shifted between nations after the first World War, is very untypical of the Italian cities coming to walking, cycling and bike use. They do cycle and walk a lot in Bolzano, as such the city resembles its neighbor to the north, Innsbruck.

**Münster – The World Bike City**

The mode split development in Münster 1982 - 2007 is shown in the following table and figure. In 2001 there was a swing back to higher car share, which was reversed in 2007. Hence, to maintain a level of cycling at 38% will require continual effort in the next years.

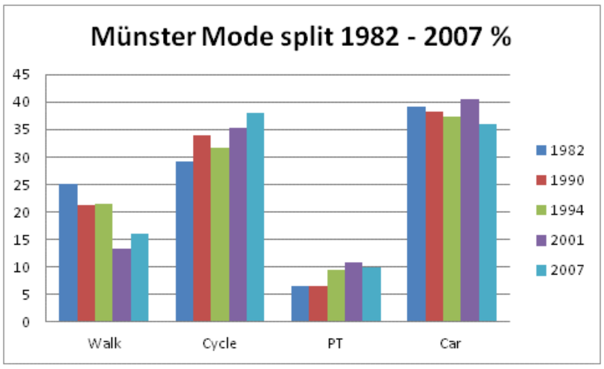



**Figure 3**
*Münster MS 1982 - 2007*

When cycling increases in Münster, the walking decreases. Less than 40% car share is very good, but it raises the question if it is possible to decrease the car share further in Münster? It might be that neither public transport nor cycling are alternatives to the car for many of the trips in the outer area of the city?

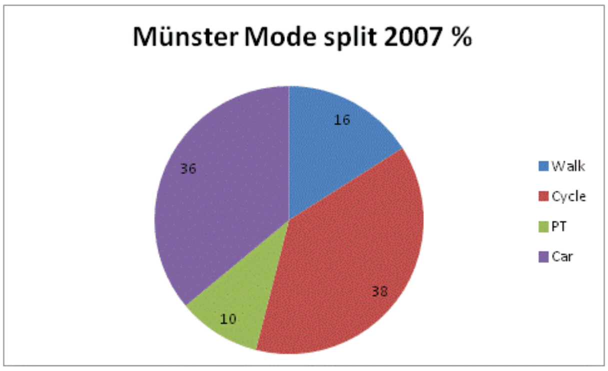



**Figure 4**
*Münster Mode Split 2007*

Münster is special when it comes to using the bicycle, 38 % of all trips. Other famous German cities for their transport policy and achievements like Freiburg and Karlsruhe has a smaller bike share.

**The Bad Ones – the car "only" cities**
In the following table the "bad cities" when it comes to car use are listed. The list is limited to three cities from each country. In the top ten of the worst car cities there are three cities from UK, three from Belgium and three from France! Among the top twenty the Netherlands, Italy and Norway has three cities on the list.

At the bottom of the list one finds the worst car city in Austria with only 49% car share, and the worst in Germany, Finland and Sweden with 57, 58 and 59 per cent respectively. Even better is the Swiss cities, all with car share less than 37%. It is obvious that these countries *control* the use of the car in cities far better than those on the top of the list.

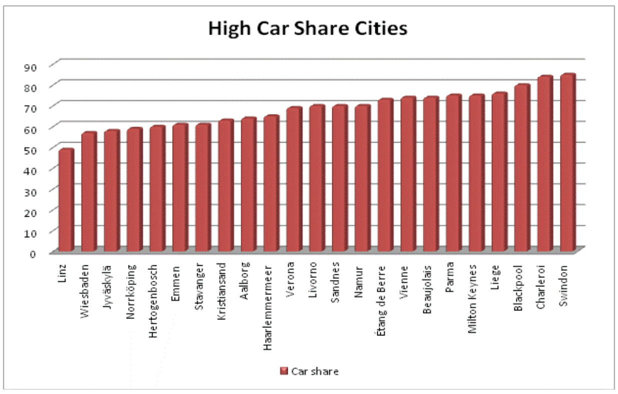



**Figure 5**
*High Care Share Cities*

**The Capital Cities**

The mode split in some of the European capitals are shown below, sorted by population high to low. Only three of the capital cities have bike share above 10%. Note that some of the data is for the metropolitan region e.g. Stockholm, while others are the city of municipality, e.g. Copenhagen. This give large differences in modal split.

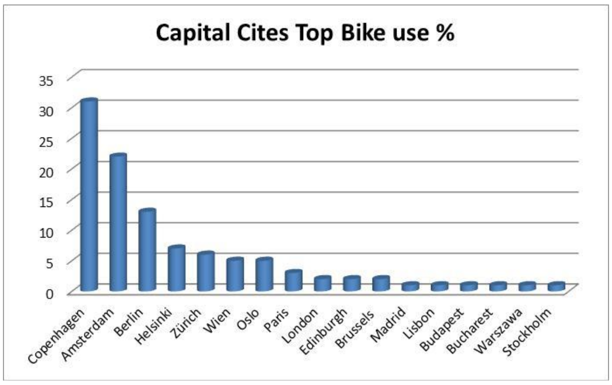



**Figure 6**

*Top Capital Bike Cities*

Only three of the capitals have more than ten per cent bike share. The capital city in Denmark, Copenhagen, is on top as a bike city, it is on top when walking and cycling are combined, and it has a car use of only 29%. Copenhagen has for several decades had increased cycling as a key strategy. From 1995 it has published a "Bicycle account" measuring progress towards goals. From 1995 to 2010 the number of serious cyclist casualties fell from 231 per year to 92. Still the 2015 goal of 59 serious cyclist casualties is a bit away. The kilometres cycled in Copenhagen increased in the same period from 0,80 to 1,21 (million km per day) in 2010. 35% cycled to work in 2010 slightly down from 2004, but far from *Copenhagen's very ambitious goal of 50% in 2015*!

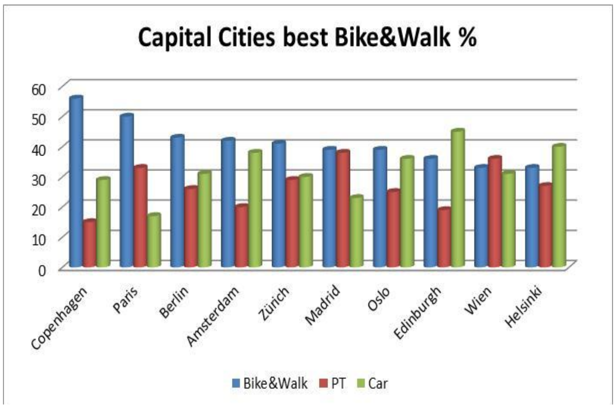



**Figure 7**

*Capital Cites, Best Bike and Walk*

Stockholm is one of the most interesting cities in this comparison, well known for the City Plan and Transit Oriented Development from the early fifties. "Stockholm is arguably the best example anywhere of coordinated planning of rail transport and urban development" and "Overall, experiences in greater Stockholm reveal that transit villages are not isolated islands within the larger metropolis, but rather are dependent upon each other as well as major urban centers. Clearly, jobs-housing balance and self-containment are not prerequisites to reducing automobile dependence." (Bernick and Cervero, 1996)

**National differences**

There are striking differences between cities across national borders and also within nations. There are some countries that hardly use the bike: Spain, France, Italy and UK., while others use the bike a lot. Netherlands, Denmark and Germany are the countries with several cities that are "Bike cities". In Table 2 are listed the number of cities in each country with a bike share above 12 percent cycling.

**Table 1**
*Nations with bike cities above 12%*

| Country | Bike cities |
|---|---|
| Netherlands | 26 cities |
| Germany | 21 cities |
| Sweden | 8 cities |
| Finland | 5 cities |
| Denmark | 4 cities |
| Italy | 4 cities |
| Austria | 3 cities |
| Belgium | 3 cities |
| Switzerland | 2 cities |
| United Kingdom | 1 city |

**Belgium**
Belgium is a special case. In the Flemish North, the cities of Brügge, Ghent and Antwerp are cycling cities with a high level of cycling resembling that in Dutch cities. In Valona to the south, hardly anyone cycle, just as in France further south, as shown on the following figure:

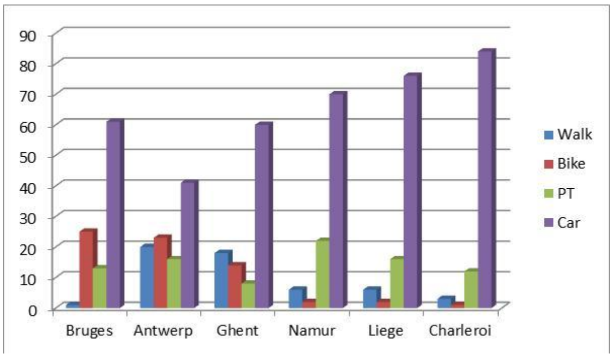



**Figure 8**
*Belgian Cities Mode Split*

**Great Britain**

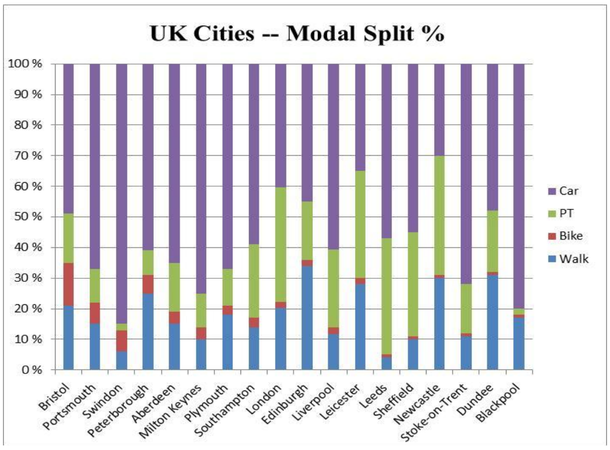



**Figure 9**
**British cities Mode Split**
Cycling has not a large modal share in Great Britain. Only Bristol resembles European cycling cities. It is the very high car share that single out Britain as being heavily car dependent. Newcastle has the lowest car share in Britain with only 30%.

## Discussion

### Does mode shift matter?

Does it matter to reduce car usage and transfer people from cars to public transport and bike and walking? The following calculation illustrate the effect of changing from the mode split similar to Milton Keynes to a mode split similar to Freiburg. The calculation is for work journeys only in an imaginary city with a population of about 250.000 and about *100.000 workers* that travel on average 11 km to work, 250 days every year. The imaginary model city with Freiburg MS has yearly $CO_2$ emissions of 156.000 tons, while with Milton Keynes MS the yearly $CO_2$ emissions are 360.000 tons, which is 204.000 tons or 230 per cent more than in the Freiburg MS case. With other words each worker produce $CO_2$ emissions per year respectively 1,56 tons with Freiburg MS, and 3,6 tons with Milton Keynes MS. There are substantial gains in a transport policy that reduce $CO_2$ emissions in “Car Cities” and there are a lot of such cities in Europe.

### Public attitude and image, history and culture

This paper shows a marked cultural difference between the protestant North Europe and the catholic South. This divide is clearly demonstrated in Belgium where bike use in the Valona south is negligible while in the Flemish North bike use in the cities is very high and comparable to the Netherlands. The arrival of the mountain bike in the eighties suddenly changed the image of the bicycle from an outdated transport mode to an artefact for the hip professional for training as a supplement to “work outs” in the training studios. In Oslo, Norway a traditional cross-country skiing competition has got its summer variant on bikes “Birken”, which has become a key status symbol among the finance brokers both to get into the race and complete the race below the set time limit. Thus the modern lifestyle mean cycling for pleasure/competition while the car is indispensable to sustain a “fast life”. That was in Oslo, Amsterdam and Copenhagen certainly have a different attitude to bicycling. There it still is common to cycle to work for all income groups and they cycle in their ordinary clothes. In Denmark and Netherlands it also seems that biking is embedded in the culture and a way of life people in the cities will hang on to?

### Income, car ownership and costs

Rising incomes makes car ownership and use relatively cheaper. According to the National transport plan it is expected that real incomes in Norway will increase by 70 % in thirty years. What will people do with increasing affluence? Past history shows that people use a share of it to buy more cars and use them. People will also travel more with air to major cities abroad, an increasingly unsustainable behaviour. The cost, travel time door to door and the convenience of alternative modes have a crucial impact on modal choice (Pucher et al, 1999). The vast majority of households have a car, many more than one. The car is by far the fastest door to door for most trips in most cities, is also the cheapest travel compared with public transport since most drivers only compare the petrol costs with the out of pocket costs for public transport. If time costs are included, the car also beat the bicycle on trips more than three to four kilometres. In addition it offers comfort, flexibility and a feeling of safety.

With very fast growth in China, India and many more nations, the number of cars may reach 2 billion cars in a few decades (Dargay 2007, Sperling and Gordon, 2009.). Does this trend leave a place and space for the bicycle? Yes, the evidence in this paper shows that it is possible, especially in cities. There are many examples of cities with both a very wealthy population and a high level of bike use.

### Infrastructure, Climate and Safety

Cities with high levels of bicycle use also have a good cycle network, but it is not clear if the infrastructure or the good network caused the amount of cyclists. The infrastructure may be a response to increased cycling instead of its cause. A network of separate bike paths and bike lanes as well as general streets with low speed seems to be a necessary, but not a sufficient element in a Bike City.

Does climate matter? Cycling is obviously affected by climate and topography, but there are questions to be raised. Why is cycling so prevalent in the wet Netherlands, while nearly non-existent in Southern France or Italy? Portland, which is both hilly and rainy, has the highest cycling rates in the USA (Pucher et al, 2009). Trondheim and Umeå both lies at the same latitude (64 degrees North), but they cycle twice as much in Umeå (19%) than in Trondheim (9%) (Langeland, 2013). Climate plays a role, but it is not the major barrier to increasing cycling. “Cycling is dangerous” is often used to explain why many persons are not cycling. It is obviously matters for some, but it is not the whole history. The Dutch culture where cycling is mixed with other traffic seem to show a more accommodating attitude between car drivers and bike users than in most other countries. If there exist a hostile attitude to cyclists on the roads, one may find that the cyclists feel more unsafe, perceive that cycling is dangerous and refrain from using the bike, irrespective of the actual risk involved.

**Planner's Vision or People's Vision?**

After the 2nd World War the car gradually became ubiquitous and a driving force in urban expansion. The Buchanan report *Traffic in Towns* asked the question: How to cope with the car? The answer was high capacity roads and "Traffic Architecture" to accommodate the car. Milton Keynes is one of the last English "New Towns" to be planned from these principles. The Buchanan report became "the bible" for generations of transportation engineers and planners, and car modernization their "leitbild". Rådberg states that to understand urban change one has to understand the planners' inner thoughts, ideas and visions (Rådberg, 1997). In the fifties and sixties, some 25 years behind USA in car ownership, the vision of the car society swept through the planners minds in Europe. The car would solve the transport problems and there was no need for trams, bikes or buses in the vision of the future. The bike disappeared as a transport mode in the planner's minds and in practise, except in Holland, Germany and to a certain extent in Denmark. Milton Keynes is a prime example of how the planners' "car vision" was implemented. 75 per cent of all trips in Milton Keynes are in a car. However, in many of the successful bike cities the planners' car modernisation ideas met opposition from the public. It is the cities where the protests won through that now are a successful Bike Cities. The *lesson* to be learnt from the successful bike cities is that cities should <u>not</u> be planned by a small group of planners and politicians, but be developed in a dialog with the public.

**More freedom to cities – less Government.**

The main *lesson* from the successful bike cities is that at a certain point the car modernisation path was broken and shifted towards a bike path or a public transport path, or an eco-path, often a combination of all these. The City of Davis in California for example, adopted a bike strategy in 1966. This was expanded into more environmentally designs for neighbourhoods and by 1980, the City was praised as an Eco-City, visited by celebrities as the wife of the then president Carter and the actress Jane Fonda. Freiburg, which retained the tramway in the seventies, became famous for developing "the environment card" a payment system for public transport. In the early nineties the design for the new urban area Vauban, became the example of how to design a sustainable city based on public transport as the main mode and with strong restrictions on cars (Langeland 2009).

Another *lesson* from the successful bike cities is that they managed to make a coordinated action across all levels and layers. Often one finds that the governing structure and the incentives for the city planners and politicians work against sustainable transport in cities. The car infrastructure is expanded and the cost of using the car is going down, while the public transport suffers from decreasing quality and increased fares, and the cyclists are literally forced off the roads. This has happened with the active support of the local politicians. The consequences of a fragmented and sector organised transport policy, is that the local politicians have strong incentives to acquire state grants for road building and disincentives to promote cycling. The German system of an "earmarked" tax to be used for city transport and the French tax on businesses according to the number of employees, are exceptions to the common picture. In most Western European countries the nation state strongly steer the city transport system through the development of the infrastructure and the level of subsidies of the public transport fares.

The shift away from the car modernisation path seems to have released creativity and innovation in the cities. It was very much a local focus that drove change. This underline that each city is unique and instruments and measures must be adapted to the local context. Most of the public transport systems in urban areas rely on subsidies in one way or another. Hardly any system is in operation where the revenue from the ticket sales makes profit for the transport company. Hence, the public transport system is continually a headache for the politicians both regarding the financing of the infrastructure and the services. Building bicycle networks in cities experience the same barriers. Who shall pay for the infrastructure? For roads the taxation of the car and petrol tax has paid for the roads, and many places toll roads are used to finance the infrastructure. Cyclists are however not a great source to be taxed to pay for the bike net. Establishing a separate bike net has therefore not been in the forefront of the politicians' agenda. The successful bike cities have against the odds, managed to develop the cycling system. This was possible, because the protests either won through in a ballot or forced the political majority to shift course away from the car modernization path. This is an important *lesson* to be learnt from the successful bike cities.

**Learning from other cities?**

Allan Jones, who lead the project turning Woking into a more Sustainable City, claim that for a city to shift to a sustainability vision, three things are necessary:

- Political support
- Chief officer support
- Insider in the organization able and qualified to deliver the vision

As Jones says "Many local authorities have one or two of these things but rarely all three as the case of Woking." " (Jones, 2009: 273) He continues: "--- this is not just about technology, it is also about politics and mindsets easily swayed by vested interests. Politics can be a very influential agent and lead the way in tackling climate change, but it can also be an agent for delay and procrastination in the fight to tackle climate change. " (Jones, 2009: 282)

Mode shift policies are deeply embedded in local context, the transport structure and the local culture. It is therefore not likely that learning from one city can be easily transferred to another. The successful sustainable cities have had hordes of delegations coming to learn and many papers are pointing to these successes. Freiburg and Groningen have had visits galore from foreign cities, and many planning documents advocate their solutions. Vauban in Freiburg has been the ideal model for how to plan a new neighbourhood, but hardly replicated. If one tries to find examples on how the learning has been implemented in other cities, there are few if any to find.

**Conclusions**

An intensive bicycle policy will not persist when it does not result in increased cycling, and a high degree of bicycle use will not persist if facilities are not upgraded and maintained (Fietsberaad, 2006). The bicycling country is still the Netherlands, but the best cycling city is Münster in Germany with 38% of all trips on the bike. The German cities Freiburg and Münster are the top environmentally sustainable cities, while the Swiss cities are foremost on integrated transport planning, low car use and high use of public transport. Another lesson from Switzerland is wide public participation and the use of referendums to decide important planning issues.

Practice in one country might encourage innovation in another country, it is after all the main purpose of comparison. However, different planning traditions and political, institutional and cultural circumstances might require adaption to different environments. Thus, it might not be possible to transfer the practice from one country to another, as: "the danger of proposing change in practice in the light of experience abroad is that practice may be dependent for its success upon a chain of circumstances which does not apply at home" (Booth, 1986 B).

The ranking of cities presented in this paper can give insights and inspiration, hence go to Münster, to Amsterdam, to Copenhagen and learn. But one should not only go to the best practice cities, there might be more relevant learning from a city striving to reach its goals and not succeeding, as the successful one.

**References**

Booth, P. (1986) The design and implementation of cross-national research projects - introduction. in: Masser, I and Williams, R (eds) *Learning from Other Countries,* Geo Books, Norwich

Booth, (1986 B), in Fischer, T.B. 2002. *Strategic Environmental Assessment in Transport and Land Use Planning.* EARTHSCAN

Fietsberaad, (2006). Continuous and integral: The cycling policies of Groningen and other European cycling cities. Publication no. 7. Rotterdam: Fiestberaad.

Flyvbjerg, B. 1998. Rationality and Power: Democracy in Practice. University of Chicago Press.

de Hartog, J. J. et.al., (2010). Do the health benefits of cycling outweigh the risks? In Environmental Health Perspectives, 118: 1109-1116.

Jones, A. (2009). *Moving Cities Towards a Sustainable Low Carbon Energy Future: Learning from Woking and London*. In Davoudi, S. et al., eds. Planning for Climate Change. Strategies for mitigation and adaptation for spatial planners. Earthscan, London.

Koetse, M.J. & Rietveld, P. (2009). The impact of climate change and weather on
transport: An overview of empirical findings. In Transportation Research Part D 14 (2009), pp. 205–221. Amsterdam: Elsevier Ltd.

Langeland, A. (2009). *The Quest for Environmentally Sustainable Transport Development. Land use and transport planning in 4 cities in 4 countries.* Aalborg University.

Langeland, A, (2013). *Sustainable bike cities and unsustainable car cities.* Paper presented: Joint AESOP/ASCP Conference Dublin July, 2013.

McClintock, H. 2002. Promoting cyling through "soft" (non-infrastructural) measures. In. H. McClintock (red.), Planning for cycing. Principles, practice and solutions for urban planners. Cambridge: Woodhead Publishing Ltd.

Pucher, J., Dill, J., Handy, S.(2009) *Infrastructure, programs and policies to increase bicycling: an international review.* Preventive Medicine, 2009, Volume 50, pages S106-S125

Pucher, J., Buehler, R., and Seinen, M. (2009). *Bicycle renaissance in North America? An update and reappraisal of cycling trends and policies.* Transportation Research Part A 45 (2011) 451-475

Rådberg, Johan, 1997: *Drømmen om Atlantångaren. Utopier og myter i 1900-tallets stadsbyggande.* Atlantis, Stockholm

Sperling, D., and Gordon, D. (2009). *Two Billion Cars*. Oxford University Press.

# KEY INDICATORS FOR THE FINANCIAL ASSESSMENT OF AVAILABILITY PAYMENT PPP PROJECTS

**Dejan Zlatkovic[1], Slaven Tica[2], Goran Mladenovic[3], Cesar Queiroz[4]**
*[1,2,3]Faculty of Transport and Traffic Engineering, University in Belgrade, Vojvode Stepe 305, 11000 Belgrade, Serbia.*
*[4]Roads and Transport Infrastructure Consultant, World Bank, 1818 H Street NW, Washington, DC, 20433, USA.*

**Abstract:** As a consequence of renewed interest in attracting private financing for big infrastructure investments, public-private partnership (PPP) arrangements are mostly seen as a suitable mechanism for ensuring sound and quicker delivery of transport infrastructure projects. Unlike conventional delivery methods, when correctly implemented, PPPs are expected to deliver projects more effectively with an efficient transfer of risk and better management of assets. However, a general concern is that expectations of mobilising private sector funds have been overestimated in a number of cases. Some pitfalls of PPP arrangements can be overcome by proper financial assessment of the project and by carrying out an appropriate value for money (VfM) analysis which is often a rationale for using PPPs.
Not all PPP models are suitable for all projects. Public sector agencies in different countries are increasingly using availability payment contract models of PPPs, for projects with uncertain revenue streams, to undertake critical infrastructure investments. Moreover, turning to availability payment projects where demand risk is retained by the government can significantly improve the attractiveness of transport investments to the private sector. An appropriate financial analysis and estimation of the minimum Annual Availability Payments that a potential concessionaire requires from the public sector to undertake the project is a key element of a successful availability payment PPP project. This paper reflects on critical indicators relevant for the financial assessment of availability payment PPP projects.

**Keywords**: PPPs, public private partnership, financial assessment, Annual Availability Payments

## 1. Introduction

An increasing tendency to implement public-private partnership (PPP) policies and projects in the transport sector can be noticed in the past few decades (Dewulf et al., 2011; Grimsey & Lewis, 2007; G. A. Hodge & Greve, 2007). As an illustration, between 1990 and 2013, 80% of European Investment Bank (EIB) loans directed to PPP schemes were absorbed by transport sector projects. In 2013, the aggregate PPP European market amounted to EUR 16.3 billion, with the UK being the most active market both by number of projects and its value (EPEC, 2014a, 2014b). Although there is no consensus on the definition of PPP, it can be defined as a “risk-sharing relationship based on a shared aspiration between the public sector and one or more partners from the private and/or voluntary sectors to deliver a publicly agreed outcome and/or public service” (Grimsey & Lewis, 2007). While PPP projects are based on the same principles, by their structure, complexness, uncertainty and the existence of numerous endogenous and exogenous variables, they can substantially differ from one another.

A quite large amount of research has been conducted in order to understand the nature of long lasting transport PPP projects. A lot of expertise and know-how have been capitalised during the implementation of PPPs in the past decades. Previous experience with successful PPP projects is a predominant factor encouraging public sector agencies to opt for PPP arrangements. Furthermore, the current public budgetary constraints make PPPs an essential mechanism for the governments to early and efficiently deliver costly transport infrastructure projects.

However, a general concern is that expectations of mobilizing private sector funds have been overestimated in a number of cases. Many authors have tried to define which elements play a key role on the successful PPPs. Cruz and Marques (2013) argue that there are several characteristics affecting the economic value of PPP projects which make them particularly sensitive to uncertainty, namely i) large sunk investments, meaning large construction costs and large debts (public and/or private), ii) high sensitivity to demand variations/estimations, iii) great exposure to financial markets (due to the large debts), and iv) vulnerability to political instability. Rockart (1978, 1982) defines critical success factors (CSF) as “the limited number of areas in which results, if they are satisfactory, will ensure successful competitive performance”. Mladenovic et al (2013) analysed the use of key performances indicators (KPI) based on the analysis of critical success factors (CSF) for monitoring of PPP transport projects from the different stakeholders perspective. The main findings and the most important CSF identified are the following:

- Public Sector: Political, social, and economic environment: Stable political and social environment, Transparent and predictable legal framework, Favorable investment climate, Stable macroeconomic environment, Transportation infrastructure needs
- Project-related CSFs: Detailed project planning and evaluation, Transparent, competitive and efficient procurement process, Appropriate risk allocation, Project economic efficiency, Capable public and private partners, Professional relationship between stakeholders
- Private sector: Partnership and communication between public and private partners, Implementation of innovative technologies, Appropriate risk allocation, Faster project completion, Transparency

[1] Corresponding author: d.zlatkovic@gmail.com

Transport PPP contracts are highly exposed and dependent to exogenous risks (Nikolaidis & Roumboutsos, 2013). A high degree of volatility of the key economic variables makes long term PPP projects vulnerable to various potential risks. A basic principle is that a risk should be borne by the party who can best manage it and hence minimize that risk. It is of utmost importance to adequately assess and address those risks in case they materialize. The following risks are of concern to both parties: legal and regulatory risk, design and construction risk, operational and maintenance risk, revenue risk, financial risk, environmental and social risks and political risk. The management and risk mitigation is a very complex task. Whereas some risks - such as construction, operational or legal risk - are more controllable by some stakeholders, traffic demand risk (revenue risk) cannot be controlled by any of them (Vassallo, 2006).

Uncertain revenue streams resulting in optimism bias on traffic demand forecasting is one of the most frequent reasons for project financial distortions. Such scenario could result in renegotiations which can produce change in the level of tolls, modification of contract length, compensation costs etc. It can also lead to possible opportunistic behaviour by the concessionaire, a vicious cycle of additional transaction costs and, in the worst case scenario, a project failure. In a survey of nearly 1.000 PPP concessions in Latin America, Guasch (2004) reports that 54% of transport sector PPPs were renegotiated at the request of the PPP Company and resulted in delays in investment or increases in tolls or availability payments.

If the traffic demand risk is to be retained by the public sector, a suitable repayment method mechanism is to be established, with a proper financial assessment of all key parameters, including sensitivity analysis. Experience suggests that possible pitfalls of PPP arrangements could be overcome by proper financial assessment of the project and value for money (VfM) analysis, which is typically the rationale for using PPPs.

The paper is organised as follows: Section 2 gives an overview of PPP repayment methods and its relation to demand risks. Section 3 reflects on principal financial indicators for the availability payment PPPs. Section 4 presents the concluding remarks and topics for further research.

## 2. Repayment methods and traffic risk mitigation

PPP models are not suitable for all project settings. The assessment of value for money has been widely used by public authorities as a tool to compare the viability of purchasing a project as a PPP or traditional procurement (Morallos et al., 2009). A PPP project is said to achieve value for money if it costs less than the best realistic alternative public sector project developed hypothetically, that is the Public Sector Comparator.

PPPs create different cash flows compared to traditional public procurement. They enable projects to be undertaken earlier than they would under public financing. In most PPP projects, no public expenditure is required at the construction phase of the project, and compensation to the private partners is made either through direct user charges or payments from the public budget spread in periodic instalments over the lifetime of the project (availability payments). Thus, the expected cash flows are highly dependent on the remuneration scheme applied.

On the other hand, traffic risk mitigation remains a challenging aspect of highway concessions. Bain et al.(2009) analysed the ratios of actual traffic to those forecast in more than 100 road, bridge and tunnel concessions. They found out that such average ratio for the first year of operation is 0.77 (i.e., the actual traffic is 23% below forecast), as indicated in Figure 1, with no significant improvement after year one (0.77 to 0.80 between years 1 and 5). Such optimism bias of the traffic forecast may cause the private sector to require a too high risk premium to take on the demand risk.

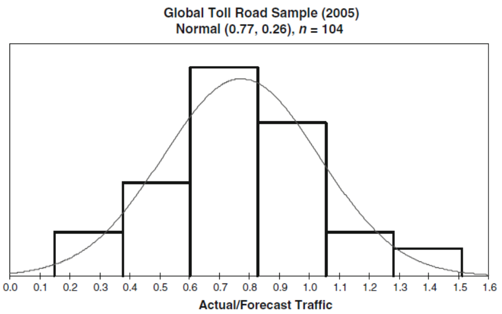



**Fig. 1.**
*Forecast performance distribution.*
*Source: Bain (2009)*

The payment mechanism is one of the principal means for allocating risks and providing incentives in the PPP contract. In practice, there are several types of payment mechanisms:

- **Revenue-based PPPs** are based on the user-pays principle and the remuneration is realised through user charges (e.g. conventional motorway tolls). Revenue-based PPPs are very specific and require examining the capacity and willingness of users to pay, in particular to ensure that tariffs are affordable during the lifecycle of the project. In several cases, toll funding can make the PPP project financially self-sustainable. Under revenue-based PPPs the financial burden is shifted from taxpayers to users. While in traditionally financed toll road projects, the government would collect the toll revenue, under a toll road PPP, usually the private sector is exposed to the traffic revenue risk.
- In contrast, under an **annuity-based PPP,** the public agency makes periodic payments to (i.e. availability payments) to the private partner throughout the operating period of the concession, thus postponing the public spending from the construction period to the operating period. The private party bears no traffic revenue risk, as it receives the fixed periodic payments as long as the road is available for use, under acceptable conditions.
- **Combination of the above** – such models combine relatively low user charges together with public support, such as: construction subsidies, minimum income guarantee models, revenue-sharing bands with thresholds for sharing traffic revenue risk (maximum and minimum targets), and flexible-term contracts with the "least present value of revenue" approach. Some of these forms of public support are widely used as a traffic risk mitigation mechanism. In EU countries, where EU funds are available for co-financing PPP projects, a number of, so called, hybrid PPP models have been developed. In recent practice, mixed payment mechanisms of an availability payment and user charge are very common.

For projects with uncertain revenue streams, the public sector is increasingly using availability payment contract models of PPP to undertake critical transport infrastructure investments. The main advantages associated with availability payments mechanisms are as follows:

- Payments should not commence until the full service is available at the agreed standard of service (availability-based principle).
- Projects are bankable – the private partner should be able to forecast cost and revenue flows with reasonable certainty.
- Payments are defined during a competitive bidding process.
- Payments are spread over the life of the contract.
- Lower return on equity (RoE) in comparison to usage payments, due to the fact that the demand risk is not borne by the private partner.
- No private sector control of toll rates (for tolled facility projects).

Moreover, turning to availability payment projects, where the demand risk is retained by the public sector, can significantly improve the attractiveness of transport investments. Proper financial analysis and estimation of the minimum Annual Availability Payments that a potential concessionaire requires from the public sector to undertake the project is a key element of a successful realisation of the project.

## 3. Key indicators in financial assessment of Availability Payment Projects

This section reflects on major indicators and critical variables relevant for the financial assessment of availability payment PPP projects and the estimation of the minimum Annual Availability Payments from the public sector with an aim to understand the likelihood of the project to be attractive to private capital.

One of the most important indicators is the Project Financial Internal Rate of Return (**Project IRR**). Basically, it represents the internal rate of return considering only the project cash flows without taking account of its financial structure (the proportion of subsidies, debt and equity). The project is considered to be financially viable when the Project Financial IRR is above a benchmark rate of return with respect to the country, sector and project characteristics. It usually should be 8% or more in real terms (Mladenovic and Queiroz, 2014). The Project IRR must be able to cover the weighted average cost of capital (WACC) of the project. WACC is equal to the average return expected from all sources of financing (including common stock, preferred stock, bonds and any other long-term debt). The WACC is higher for a higher equity-to-debt ratio.

Another important indicator is the Equity Internal Rate of Return (**Equity IRR**). Namely, stakeholders' equity represents the amount of the construction cost contributed by the owners in the project financial structure. It is typically 10% to 30% of a PPP's capital cost. The balance is funded from external debt finance which may be provided by IFIs, commercial banks or, in some cases, by the bond market. Equity IRR represents the yield of the project through the remuneration of shareholders´ investment with dividends. The project is profitable for the shareholders when Equity IRR is high. Generally, a minimum expected real rate of return (real return) is 10% (Shadow Toll) or 17% (Toll Roads) (Queiroz, 2010). This minimum Equity IRR is called Hurdle Rate.

The required equity IRR usually serves as a reference for bidders to estimate the required annual availability payment. While the Project IRR is independent from the project financial structure, the Equity IRR is related to it.

Before committing to a project, lenders will perform an in-depth review of the project viability to assure that the project is bankable. Any interruption of the project revenue stream could jeopardise the ability to make debt service payments. Next we describe ratios that are used to check the project capacity to repay debt in different scenarios, including if revenues are below forecasted levels.

To assess the project capacity to repay debt, the Annual Debt Service Cover Ratio (**ADSCR**) is normally calculated. This ratio is determined as follows:

$$ADSCRi = \frac{CBDSi}{DSi}$$

where:

- *CBDSi* - the net cash flow before debt service in year *i* (i.e., the amount of cash remaining in the project company after operating costs and taxes have been paid)
- *DSi* - the debt service to be paid in year i (principal and interests)

The project is considered viable for the lenders when the ADSCR is greater than 1. The higher the ADSCR, the more attractive the project will be to lenders because it leaves the reserve for undesirable circumstances that may occur during the project's life. Generally, for low-risk PPP projects involving availability payments the minimum required ADSCR is 1.15 to 1.2, while for high risk projects, such as toll-based concessions, it ranges between 1.5 and 2.0 (World Bank, 2009).

Another ratio that assesses the creditworthiness of the project is the Loan Life Cover Ratio (LLCR). The LLCR indicates the capacity for the project company to bear a sporadic shortfall of cash due to discrepancies in the assumptions in the model while maintaining its debt service to the end of the debt. This ratio is calculated as follows:

$$LLCRi = \frac{NPV(CBDSi \to end)}{DSi \to end}$$

where:

- *NPV (CBDSi → end)* is the present value of the net cash flow before debt service from year i to the end of the debt repayment period.
- *DSi → end* is the total of debt service remaining at year i (principal and interests).

The project is considered viable for the lenders when the LLCR is higher than 1 for every year of the project life (Mladenovic and Queiroz, 2014). The minimum initial LLCR requirement in lenders projections for PPP projects is typically about 10 percent higher than the required ADSCR.

The principal objective is to minimise the Annual Availability Payment that the concessionaire will expect from the government in the first year of project operation (the initial revenue year). The estimation of minimum Annual Availability Payment is subject to constraints that are to be met for the project to be able to attract private investors. As mentioned before, these constraints are related to minimum required thresholds for the three indicators considered critical for the project: Project IRR, Equity IRR and ADSCR. Annual Availability Payments in subsequent years are adjusted according to the specified inflation rate.

**4. Conclusions**

PPP financing structures are complex and often project-specific. These structures should ensure that financial and other related risks are well shared among all stakeholders. PPPs should be attractive enough to ensure the viability of project against downside scenarios. Key financial indicators presented in the paper represent the basic mechanism for the assessment of the investments, both by private and public parties. On the other hand, when designing the payment mechanism, comparable projects should be used as a useful benchmark considering features that may later give rise to disputes. Despite different pros and cons, availability payments projects, where demand risk is retained by the public sector, can significantly improve the attractiveness of the transport investments.

Assuming that previous studies have shown that one PPP transport project is economically justified, and socially and environmentally sound, the financial model estimate the minimum Annual Availability Payment that a potential concessionaire will require from the government to undertake the project. This estimated value should ensure that all targets are met for the project to be able to attract private investors.

**References:**


Bain, R. (2009), “Error and Optimism Bias in Toll Road Traffic Forecasts”, Transportation, Springer, 28 February 2009

Cruz, C. O., & Marques, R. C. (2013). Flexible contracts to cope with uncertainty in public–private partnerships. International Journal of Project Management, 31(3), 473-483. doi: 10.1016/j.ijproman.2012.09.006

Dewulf, G., Garvin, M., & Duffield, C. (2011). What are the benefits of standardised PPP procurement processes? Working paper proceedings. Paper presented at the Engineering Project Organizations Conference, August 9-11, Estes Park, CO.

EPEC. (2014a). Market Update: Review of the European PPP Market in 2013. Luxembourg: European PPP Expertise Centre.

EPEC. (2014b). PPPs Financed by the European Investment Bank since 1990. Luxembourg: European PPP Expertise Centre.

Grimsey, D., & Lewis, M. K. (2007). Public Private Partnerships: The Worlwide Revolution in Infrastructure Provision and Project Finance. Cheltenham: Edward Elgar.

Guasch, J. L., & Straub, S. (2006). Renegotiation of infrastructure concessions: An overview. Annals of Public and Cooperative Economics, 77(4), 479-493.

Hodge, G. A., & Greve, C. (2007). Public-private partnerships: An international performance review. Public Administration Review, 67(3), 545-558. doi: DOI 10.1111/j.1540-6210.2007.00736.x

Mladenovic, G., Queiroz, C. (2014), “Assessing the Financial Feasibility of Availability Payment PPP Projects” Transportation and Development Institute (T&DI) Second Congress, American Association of Civil Engineers, Orlando, FL, USA, June 8-11, 2014, Paper ID #260

Mladenovic, G., Vajdic, N., Wündsch, B., & Temeljotov-Salaj, A. (2013). Use of key performance indicators for PPP transport projects to meet stakeholders’ performance objectives. Built Environment Project and Asset Management, 3(2), 228-249. doi: 10.1108/BEPAM-05-2012-0026

Morallos, D., A. Amekudzi, C. Ross, and M. Meyer, (2009) “Value for Money Analysis in U.S. Transportation Public–Private Partnerships”, Journal of the Transportation Research Board, 2115, pp. 27-36

Nikolaidis, N., & Roumboutsos, A. (2013). A PPP renegotiation framework: a road concession in Greece. Built Environment Project and Asset Management, 3(2), 264-278. doi: 10.1108/BEPAM-05-2012-0031

Queiroz, C (2010) World Bank Workshop on the Toolkit for PPP in Roads and Highways, Brasilia, Brazil, June 8-9, 2010

Rockart, J. F. (1978). Chief executives define their own data needs. Harv Bus Rev, 57(2), 81-93.

Rockart, J. F. (1982). The changing role of the information systems executive: a critical success factors perspective: Massachusetts Institute of Technology.

Vassallo, J. M. (2006). Traffic risk mitigation in highway concession projects: the experience of Chile. Journal of Transport Economics and Policy, 359-381.

World Bank. (2009). Toolkit for Public-Private Partnerships in Roads and Highways, World Bank, Washington D.C.

# ASSESSING THE EFFECT OF TRAFFIC CONGESTION ON GREENHOUSE GAS EMISSIONS

**Eugen Roșca[1], Dorinela Costescu[2], Florin Ruscă[3], Ștefan Burciu [4]**
*[1,2,3,4]University Politehnica of Bucharest, Faculty of Transports, 313 Splaiul Independenței, 060042 Bucharest, Romania*

**Abstract:** Transport activities are among the greatest contributor to greenhouse gas (GHG) emissions and despite other economic sectors the trend is still positive. Using the latest European Environment Agency guidelines in estimating vehicles emissions, the paper presents a methodology for assessing GHG emissions integrating exhaust emission factors, geographic information system (GIS) data, traffic composition surveys, and road traffic simulation. According to EU recommendation for reporting national gas inventories, the estimation of GHG emission factors based on the average speed is adopted. The computer simulation of road traffic allows tracing individual car driving cycles and computing emissions based on average speed, mileage, fuel, and engine emission standards. The main advantage when using the computer traffic simulation is to reduce the amount of measurements which have to be performed in the field. Also the computer models have the ability to evaluate the effects of different traffic patterns (free flow, car following flow and traffic jam) and the future traffic scenarios. The carbon dioxide, methane, and nitrous oxide emissions are equated by their global warming potential. The case-study done in Carpathians outlines the way the superposition of the specific land use policy, transport network characteristics and road traffic patterns generated by the variability of human activities (the leisure-time and business tourism) can cause road congestion that increases vehicles emissions. The study presents aggregated emissions evolution by time periods and comparative values to European average levels and targets.

**Keywords**: traffic congestion, greenhouse gas emission, traffic simulation

## 1. Introduction

The most important source of greenhouse gases (GHG) emissions in the EU-28 was fuel combustion of energy industries, transport, manufacturing and construction industrial activities. Transport sector was the only source that presented an increase between 1990 and 2011 (19%), having a share of 20.2% from total emissions (EC, 2013). The GHG emissions of Romania are still below the target established in the Kyoto Protocol (-50%), allowing the trade of the emission rights on the European Trading Scheme. Despite the general favorable situation, the GHG emissions in the Romanian transport sector have been growing from an amount of 5.97 Mt in 1989 to 11.15 Mt in 2010, showing a high growing dynamic due to a constant increasing of the motorization rate and the dominant position of the road transport mode versus rail and inland water transport. Policy makers should take into consideration new ways to promote an efficient transport system that provides accessibility, satisfies the economic and social needs, and minimizes the negative external effects on the environment.

## 2. Assessment methodology for the road traffic GHG emissions

Across Europe, the most pertinent emission estimation tools that have been used are CORINAIR (Eggleston et al., 1992), DRIVE (Jost et al., 1992), COST 319 (Joumard, 2009) and HBEFA (Hausberg et al., 2009). The European Commission, through the European Environment Agency (EEA) issued the Emission Inventory Guidebook (EEA, 2013) that covers a high range of exhaust emissions (GHG, acidifying substances, ozone-precursors, particulate matters, heavy metals, toxic and carcinogenic substances). Also, the EEA methodology consists in the thoroughly classification of the emission factors according to a large variety of vehicles technical characteristics. The EEA guidelines have been also incorporated into the software tool COPERT 4 (Ntziachristos et al., 2009), which allows the compilation of the national emissions inventories on a yearly basis. Still, COPERT 4 uses traffic patterns and average traffic flows characteristics, which are not suitable for thorough details on local areas.

The following methodology for evaluating GHG emissions combines field recorded traffic data, computer traffic simulation using VISSIM software and emission factors estimated through Tier 3 approach of the EEA guidelines. The flow chart for the GHG emissions assessment is depicted in Fig. 1.

[4]Corresponding author: florin_rusca@yahoo.com

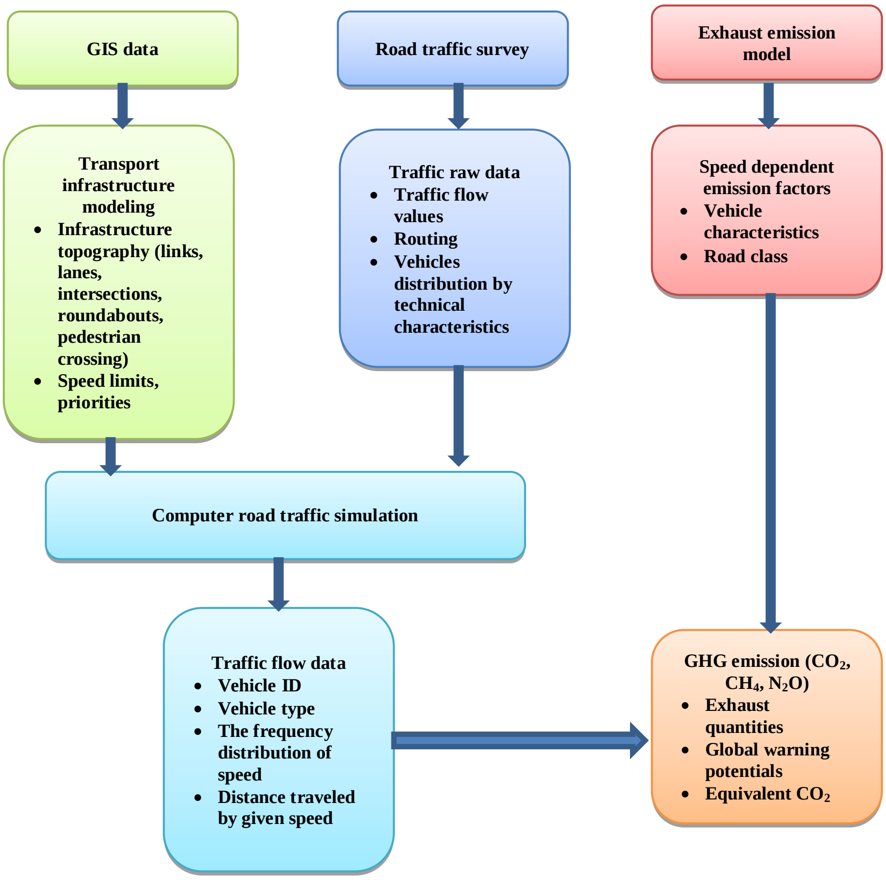



**Fig. 1.**
*GHG emission evaluation flow chart*

### 2.1. Transport infrastructure modeling

The Geographic Information System (GIS) base software provides scalable geographic maps and detailed transport network data used further into traffic simulation programs. Using real world coordinates, the transport network topology is described in terms of links, lanes, connections, roundabouts, public transport terminals/stations, speed limit areas, priority rules, signaling systems.

### 2.2. Traffic raw-data

This module considers survey data obtained in the field. The car traffic flow is recorded separately for each individual vehicle type (passenger car, LDH, HDV, bus/coach). Traffic flow variation has to be considered. The daily peak-hours, weekends or seasonal variations influence the exhaust emissions due to changes in the car driving cycles (e.g. speed, acceleration, and number of stops). The vehicles are classified according to their fuel type (gasoline, diesel, LPG, CNG, hybrids), technology/standards (pre ECE, Euro 1-5) and engine capacity. Other traffic participants (pedestrians, cyclists) have to be included as long as they interact with the car traffic flows.

### 2.3. Computer traffic simulation

The traffic raw-data and the transport infrastructure model represent inputs for the traffic simulation software. The main advantage when using the computer traffic simulation is to reduce the amount of measurements, which have to be performed in the field. Also the computer models have the ability to evaluate the future traffic scenarios based on various traffic management schemes. The traffic micro simulation models are able to describe a complete driving cycle including free-flow, car following, overtaking and interactions in junctions. Boulter and McCrae (2007), Tate et.al (2005), Panis et al. (2006), Jayaratne et al. (2009) provide relevant descriptions of the nowadays traffic simulation models and how they are implemented in assessing vehicle emissions using different emission factors databases.

### 2.4. Traffic flow characteristics

Traffic simulation allows getting the basic outputs used in GHG emission assessment:
- the vehicle ID;
- the vehicle type;
- the frequency distribution of the mean speeds of each vehicle along the route;
- the car mileage of each vehicle at a given mean speed.

The driving cycle contains free-flow, car following, acceleration/deceleration or queue waiting periods specific for each road segment. As flow increase, the traffic density increases too, while the average speed decreases. The flow is getting a maximum at a specific speed and density, according to traffic conditions, road topology, and cars interactions. Then the flow starts decreasing with increasing density and decreasing speed (congestion pattern), going towards zero at jam density.

### 2.5. Emission factors

The car traffic emissions with the highest greenhouse effect are $CO_2$, $CH_4$ and $N_2O$. $CO_2$ emissions are estimated based on the fuel consumption, while $CH_4$ and $N_2O$ are directly estimated, based on the specific emission factors covering different traffic situation (urban, rural, highway) and engine characteristics.

- **Carbon dioxide ($CO_2$) emission**

Considering an oxygenated fuel with the generic chemical formula $C_xH_yO_z$, the end of pipe mass of $CO_2$ emitted by vehicles with the engine technology *k*, combusting the fuel *m* is:

$$E_{CO_2,k,m} = 44.011 \times \left( \frac{FC_{k,m}}{12.011 + 1.008 r_{H:C,m} + 16 r_{O:C,m}} - \frac{E_{k,m}^{CO}}{28.011} - \frac{E_{k,m}^{VOC}}{13.85} - \frac{E_{k,m}^{EC}}{12.011} - \frac{E_{k,m}^{OC}}{13.85} \right) \quad (1)$$

where:

$E_{CO_2,k,m}$ - $CO_2$ exhaust emission [g/km];

$FC_{k,m}$ - the fuel consumption of the vehicles [g/km];

$EC_{k,m}^{CO}$ - carbon monoxide exhaust emission [g/km];

$EC_{k,m}^{VOC}$ - volatile organic components exhaust emission [g/km];

$EC_{k,m}^{EC}$ - elemental carbon exhaust emission [g/km];

$EC_{k,m}^{OC}$ - organic carbon exhaust emission [g/km];

$r_{H:C,m}$ - the ratio of hydrogen to carbon atoms in the fuel;

$r_{O:C,m}$ - the ratio of oxygen to carbon atoms in the fuel.

Eq. 1 takes into consideration the emission of carbon (C) atoms in the form of carbon monoxide (CO) and volatile organic components (VOC). EEA (2013) offers generic functions for estimating the fuel consumption (FC), CO and VOC emissions in relation to the vehicle type, fuel, engine capacity, engine emission standards and vehicle mean speed. These functions having quadratic, power, polynomial form or a combination of them, have been estimated through statistical analysis of the empirical data, providing for most of them coefficients of determination higher than 0.9.

- **Methane ($CH_4$) emission**

The $CH_4$ emission factors are directly estimated and not computed based on fuel consumption. The emission factors are discriminated according to vehicle type, fuel, engine technology and road category. The limits of the variation of $CH_4$ emission for different car/traffic characteristic are presented in Table 1.

**Table 1**
*Limits for CH4 emission factors*

| Vehicle type | Fuel | $CH_4$ emission factor [mg/km] | | |
|---|---|---|---|---|
| | | Urban | Rural | Highway |
| Passenger car and light-duty vehicle | Gasoline | 2 - 131 | 2 - 86 | 0 - 41 |
| | Diesel | 0 - 28 | 0 - 12 | 0 - 8 |
| | LPG | 80 | 35 | 25 |
| Heavy-duty vehicle, bus and coach | Diesel | 85 - 175 | 23 - 80 | 20 - 70 |
| | CNG | 980 - 6800 | | |
| Two-wheels vehicle | Gasoline | 150 - 219 | | |

*Source: (EEA, 2013)*

- **Nitrous oxide ($N_2O$) emission**

Based on the complementary studies done by Riemersma et al. (2003), Papathanasiou and Tzirgas (2005), EEA estimated the $N_2O$ emission factors by vehicle type, fuel, and engine technology.

**Table 2**
*Limits for $N_2O$ emission factors*

| Vehicle type | Fuel | $N_2O$ emission factor [mg/km] |
|---|---|---|
| Passenger car | Gasoline | 5 - 23 |
| | Diesel | 0 - 10 |
| | LPG | 0 - 24 |
| Light-duty vehicle | Gasoline | 10 - 28 |
| | Diesel | 0 - 9 |
| Heavy-duty vehicle | Diesel | 7 - 53 |
| Bus and coach | Diesel | 1 - 40 |
| | CNG | 0 |
| Two-wheels vehicle | Gasoline | 1 - 2 |

*Source: (EEA, 2013)*

The $N_2O$ emissions are not strongly dependent on vehicle average speed as $CO_2$ emissions are, but mainly on fuel and engine type. Diesel vehicles have lower $N_2O$ emissions than catalyst equipped cars.

**2.6. $CO_2$ equivalence**

The emissions of GHG can be aggregated into $CO_2$ equivalent taking into consideration the relative effect to the climate change determined by the global warming potentials (GWP). Table 3 presents GWP for methane and nitrous oxide.

**Table 3**
*Global Warming Potentials relative to $CO_2$*

| Gas | Radiation efficiency [$Wm^{-2}$/ppbv] | GWP (20 years time horizon) |
|---|---|---|
| Carbon dioxide $CO_2$ | $1.8\times10^{-5}$ | 1 |
| Methane $CH_4$ | $3.7\times10^{-4}$ | 62 |
| Nitrous oxide $N_2O$ | $3.1\times10^{-3}$ | 275 |

*Source: (IPCC, 2001)*

The equivalent $CO_2$ emissions of vehicles crossing a specific area during a given time interval should be computed as:

$$E_{eq.CO2} = \sum_{i=1}^{N} d_{i,k,m,r} \times (GWP_{CO2} \times e_{CO2,k,m,r} + GWP_{CH4} \times e_{CH4,k,m,r} + GWP_{N2O} \times e_{N2O,k,m,r}) \quad , \quad (2)$$

where:

$E_{eq.CO2}$ is the equivalent exhaust emission of $CO_2$;

$N$ - number of vehicles [veh] crossing the study area;

$d_{i,k,m,r}$ - mileage per vehicle [km] driven on road r by vehicle i of technology k and fuel m;

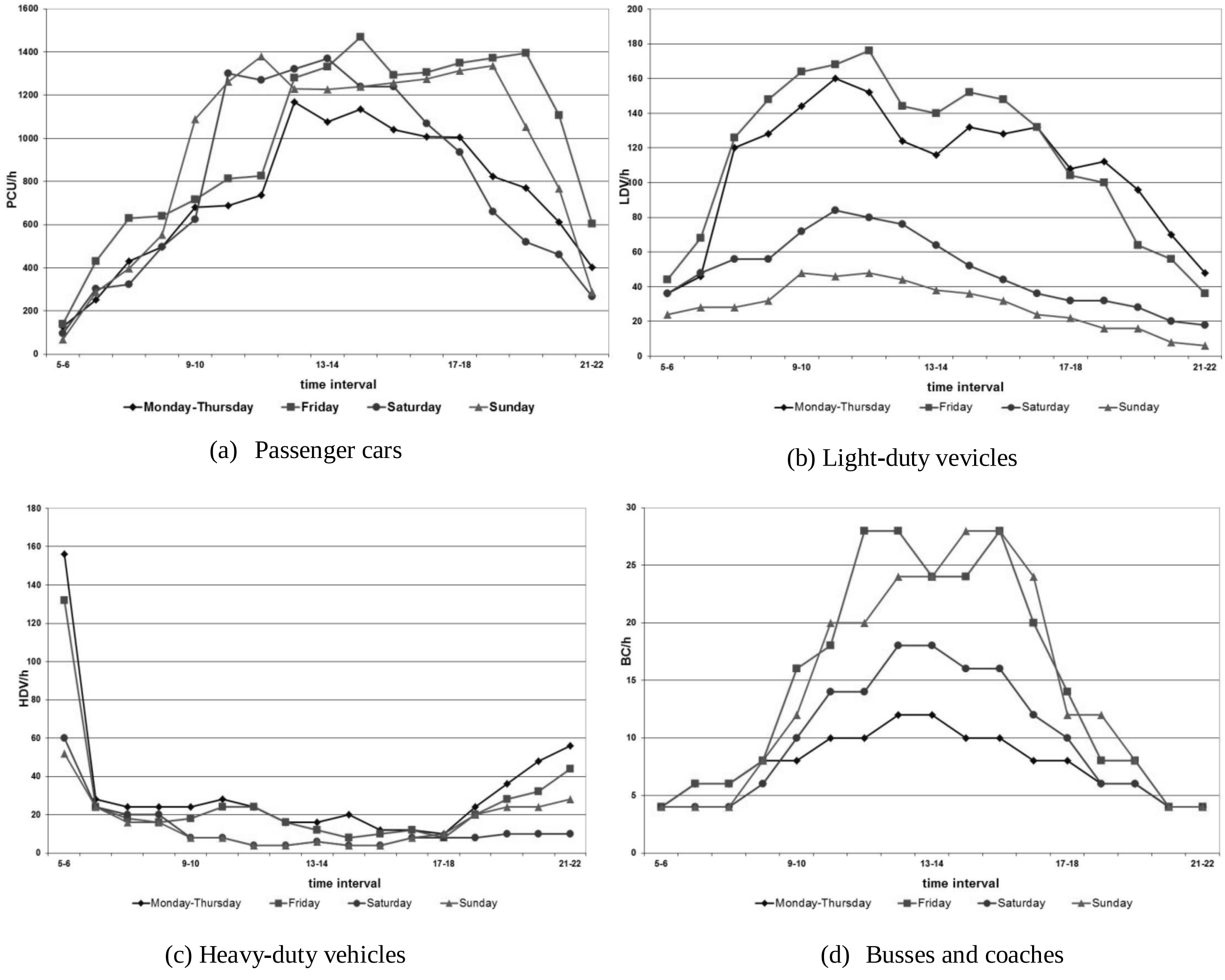
$e_{j,k,m,r}$ - emission factor [g/km] for pollutant j ($CO_2$, $CH_4$ or $N_2O$) by vehicle of technology k, fuel m on road type r;

$GWP_j$ - global warming potential for pollutant j.

### 3. Case study. Assessing the GHG emissions on Prahova Valley (Carpathians)

The route E60 along Prahova Valley, situated in the Meridional Carpathians, is one of the most transited routes connecting the southern and the central part of Romania. The transit route passes alongside the Bucegi Massif that is classified as a type B site (medium influence environment) by the European Network of Protected Areas – Natura 2000. Downhill of this massif there are a couple of mountain resorts transited by the route E60, among them Buşteni standing out as the main bottleneck of the road traffic. Due to the transport infrastructure characteristics (single lane, intersections, dense pedestrian crossing) and the traffic increasing during weekends and tourist seasons, the road traffic congestion is often presented in the area, reducing the level of performance of the network and increasing the car exhaust emissions.

For assessing the GHG emissions, the road traffic in Buşteni mountain resort is modeled using the VISSIM traffic simulation software, implementing the Wiedemann (1974) car-following model. The transport infrastructure layout is completed by the traffic rules (vehicles speed limits, priority rules in intersections and pedestrian crossing) and by the public transport elements (stations, stop times distribution). The traffic survey was conducted during May-June, period providing high traffic intensity and non-uniformities mainly during weekends. The traffic flows are recorded separately by vehicles type: passenger car (PCU), light-duty vehicle (LDV), heavy-duty vehicle (HDV), bus/coach (B/C). The data are differentiated by the days of the week (Monday-Thursday, Friday, Saturday and Sunday), taking into consideration the traffic pattern according to the socio-economic activities cycle and the leisure-time travels (Holden, 2007). The average traffic flows (cumulative both ways) are shown in Fig. 2.




(a) Passenger cars

(b) Light-duty vevicles

(c) Heavy-duty vehicles

(d) Busses and coaches

**Fig. 2**
*Daily average road traffic flows*

The survey also included the structure of the vehicles flow according to their fuel type, technology/emission standards and engine capacity. The data have been collected through a systematic sample scheme for each vehicle category, using a sample interval of 10 vehicles. The data structure based on vehicle type, fuel and emission standard is shown in Fig. 3.

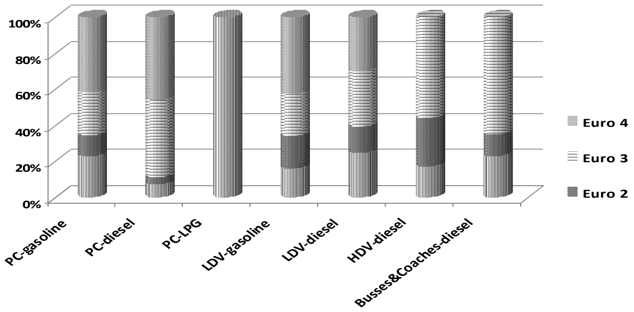



**Fig. 3**
*Statistical structure of the vehicles flow*

The traffic simulation outputs have been obtained for different daily time intervals and for working days/weekends. Each vehicle entering the simulation process has allocated individual ID number and characteristics concerning its type, fuel, engine capacity, emissions standard and maximum speed allowed.
The plot of the average speed to the traffic flow volume resulting from simulation is shown in Fig. 4. The total traffic flow volume is expressed in equivalent passenger car units (PCU) using the equivalent factors proposed by Hobbs (1979).

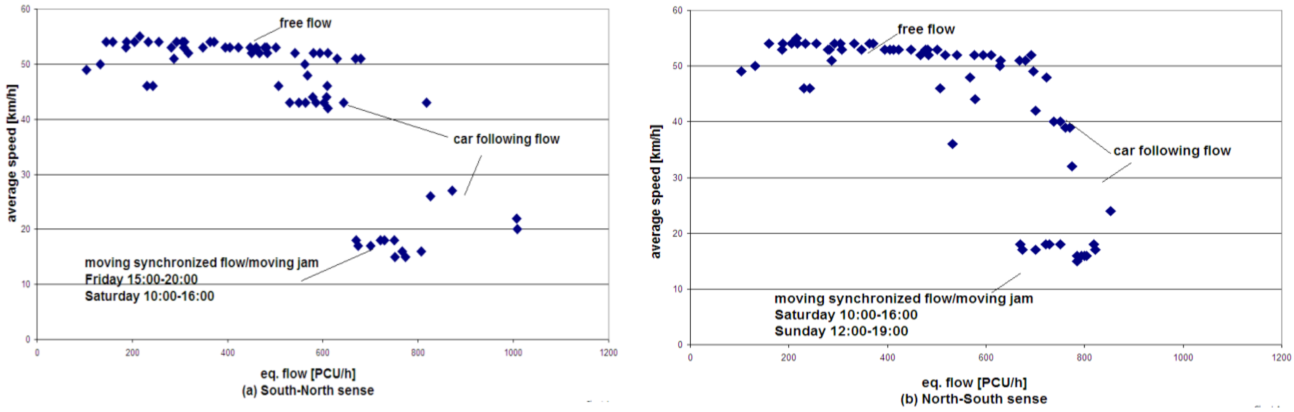



**Fig. 4**
*Traffic volume vs. average speed diagram*

The traffic situation creates specific congested phases characterized by a moving synchronized flow pattern that turns during some periods to a moving jam phase. The congestion appears on different traffic ways on Friday and Sunday and on both ways on Saturday.
Following the assessment methodology described in the paragraph 2, the average equivalent CO2 emission per unit of time and distance and by day of the week is depicted in Fig. 5.

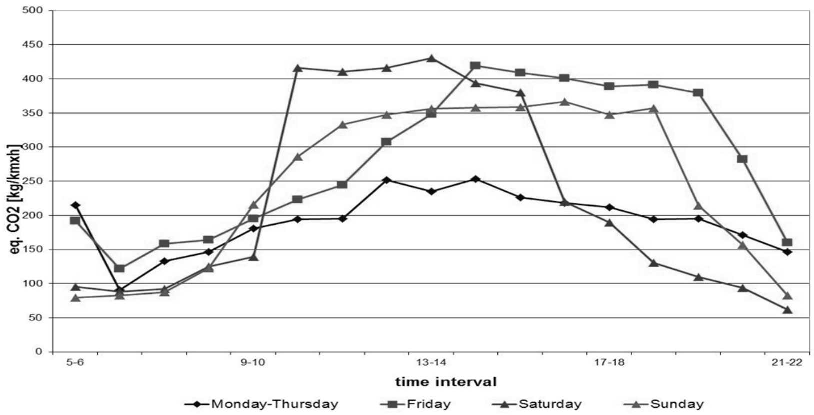



**Fig. 5**
*Equivalent $CO_2$ emission*

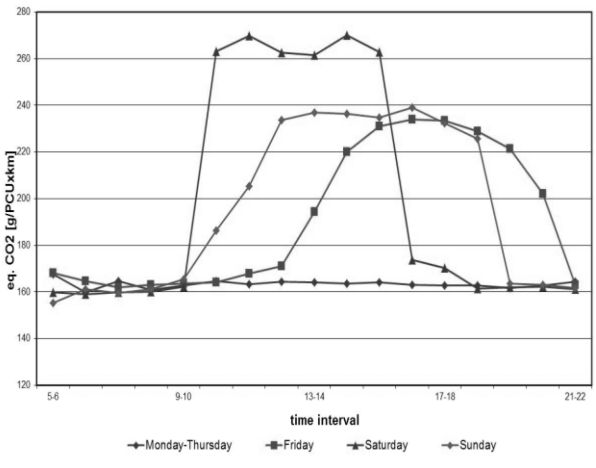
Also, the daily average emissions of GHG in the area are shown in table Table 4.

**Table 4.**
*Daily average GHG emissions from road traffic*

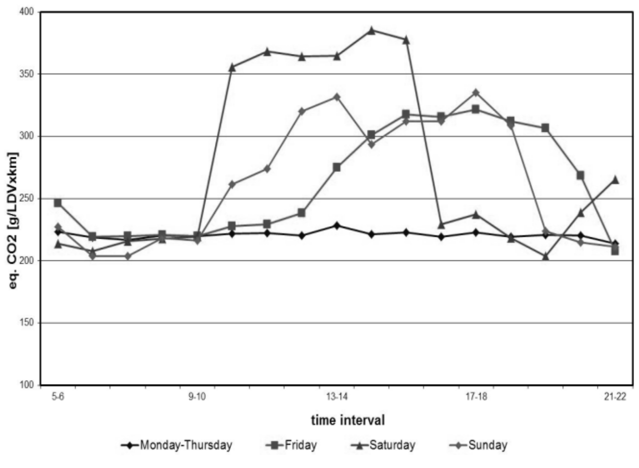
| Day of the week | Emisions [kg/day] | | | | | | |
|---|---|---|---|---|---|---|---|
| | **$CO_2$ full combustion process** | **CO** | **VOC** | **$CO_2$ end of pipeline** | **$CH_4$** | **$N_2O$** | **Equivalent $CO_2$** |
| Monday-Thursday | 3301.83 | 57.49 | 5.09 | 3195.33 | 0.49 | 0.12 | 3258.33 |
| Friday | 4898.50 | 103.86 | 9.48 | 4705.20 | 0.60 | 0.15 | 4784.67 |
| Saturday | 3908.96 | 93.83 | 8.61 | 3734.16 | 0.45 | 0.11 | 3792.87 |
| Sunday | 4272.65 | 101.19 | 9.23 | 4084.32 | 0.53 | 0.13 | 4152.37 |

Analyzing the equivalent $CO_2$ emission in Fig. 5, some remarks are issuing:

- The morning emissions are greater during the working days due to the heavy-duty vehicles circulation. These are restricted to circulate on Prahova Valley during 06:00-22:00, unless a special transport toll is paid.
- The mid-day and the afternoon emissions are highly raised during weekends, due to increase in car traffic volume and the resulting decrease of the average speed. As depicted also in Fig. 4., the congestion phenomena are presented during these time intervals of the weekends. The rising of the traffic flow and the decrease of the speed is turned in a higher increasing of the eq. $CO_2$ emission.




(a) Passenger cars




(b) Light-duty vehicles

**Fig. 6.**
*Average equivalent $CO_2$ emission per vehicle*

The average eq. $CO_2$ emissions for passenger cars and light-duty vehicles are depicted in Fig. 6. During the working days and the out-of-peak time intervals of the weekends, the eq. $CO_2$ average emissions for passenger cars are very closed to the average value recorded at European level – 150g/PCU×km, but also superior to the aim established – 140g/PCU×km (EEA, 2010). During congested periods, the average passenger car emission values increase between 40 – 60%. For light-duty vehicles, the eq. $CO_2$ average emissions overpass 200 g/LDV×km and are higher to those recorded at European level – 170g/LDV×km (EEA, 2010). This situation reflects the aged LDV fleet in Romania. The average emissions per LDV are up to 70% greater during the congested periods of the weekends compared to the working days.

## 4. Conclusions

The use of the traffic simulation is highly benefic for assessing vehicles emissions. It allows a detailed computing at individual vehicle level, and also could be used in evaluating the influence of the infrastructure modernization or traffic reorganization. The case study outlines that the weekend leisure-time and the business tourism are among the most responsible human activities for road traffic congestion and therefore for the GHG emission. The rising of the car ownership, the poor public transport services and the land use policies allowing the spread of the residential areas in the former restricted areas are translated into a high growth of the passenger car use with negative external effects on the air pollution.

## Acknowledgements

The work has been funded by the Sectoral Operational Programme Human Resources Development 2007-2013 of the Ministry of European Funds through the Financial Agreement POSDRU/159/1.5/S/132397.

**References**

Boulter P.G., McCrae I.S. 2007. The links between micro-scale traffic, emission and air pollution models, Transport Research Laboratory, UK.

EC 2013. Energy, transport and environment indicators, European Commission, Publications Office of the European Union, Luxembourg.

EEA 2013. EMEP/EEA Air Pollutant Emission Inventory Guidebook, European Environment Agency, Copenhagen, Denmark.

EEA 2010. Towards a resource-efficient transport system TERM 2009: indicators tracking transport and environment in the European Union, EEA Report 2/2010, European Environment Agency, Copenhagen, Denmark.

EEA 2012. Annual European Union greenhouse gas inventory 1990–2010 and inventory report 2012, Technical Report 3/2012, European Environment Agency, Copenhagen, Denmark.

Eggleston H.S., Gaudioso D., Gorisson N., Joumard R., Rijkeboer R.C., Samaras Z., Zierock K.H. 1992. CORINAIR working group on emission factors for calculating 1990 emissions from road traffic, Vol. 1: Methodology and Emission Factors, Final Report, Commission of the European Communities.

Hausberger S., Rexeis M., Zallinger M., Luz R. 2009. Emission Factors from the Model PHEM for the HBEFA, Version 3, Report Nr. I-20/2009, Graz University of Technology, Graz, Austria.

Holden E. 2007. Achieving Sustainable Mobility. Everyday and Leisure-time Travel in EU, Ashgate, UK.
Hobbs F.D. 1979. Traffic Planning and Engineering, Pergamon Press, Oxford, UK.

IPCC 2001. Third Assessment Report: Climate Change. The Scientific Basis, Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK.

Jayaratne E.R., Wang L., Heuff D., Morawska L., Ferreira L. 2009. Increase in particle number emissions from motor vehicles due to interruption of steady traffic flow, Transportation Research, Part D, 14, 521–526.

Jost P., Hassel D., Weber F.J. 1992. Emission and fuel consumption modeling based on continuous measurements, DRIVE project V 1053, TUV Rheinland, Koln, Germany.

Joumard R. 2009. Methods of estimation of atmospheric emissions from transport: European scientist network and scientific state-of-the art, Action COST 319 Final Report, INRETS, France.

Ntziachristos L., Gkatzoflias D., Kouridis Ch., Samaras Z. 2009. COPERT: a European road transport emission inventory model, In: Information Technologies in Environmental Engineering, Athanasiadis I., Mitkas P., Rizzoli A., Gomez J. (Eds.), Springer, Berlin-Heidelberg, 491-504.

Panis L.I., Broekx S., Liu R. 2006. Modelling instantaneous traffic emission and the influence of traffic speed limits, Science of the Total Environment, 371, 270 –285.

Papathanasiou L., Tzirgas S. 2005. $N_2O$ and $NH_3$ emission factors from road vehicles, Aristotle University of Thessaloniki, Technical Report 0507, Thessaloniki, Greece.

Riemersma I., Vermeulen R., Gense R., Smokers R. 2003. $N_2O$ emissions of LD and HD vehicles, TNO Automotive, 12th International Scientific Symposium on Transport and Air Pollution, Avignon, France, 16-18 June, 153-160.

Tate J.E., Bell M.C., Liu R. 2005. The application of an integrated micro-simulation and instantaneous emission model to study the temporal and spatial variation in vehicular emission at local scale, In Proceedings of the 14$^{th}$ International Conference Transport and Air Pollution, Graz, Heft 85/1, 138-147.

Wiedemann R. 1974. Simulation des Straßenverkehrsflusses, Schriftenreihe des Instituts für Verkehrswesen der Universität Karlsruhe, Heft 8, Karlsruhe, Germany.

# THE GOALS OF SUSTAINABLE TRANSPORT AND INTELLIGENT TRANSPORT SYSTEMS FOR URBAN BUSES

**Mojtaba Zoorbakhsh**[1]
[1]*Faculty member of department of civil engineering, Islamic Azad University, Estahban Branch, Estahban, Iran*

**Abstract:**

In the process of urban life and subsequent developments, transportation programming plays two important key roles in affecting and being affected by the whole system. Forms of new mega cities and their infrastructure have been specially influenced by road and rail transportation systems. On the other hand, the structure and formation of cities is now interrelated with employing Intelligent Transportation Systems (ITS), urban traffic management and their efficiency in the manner the tasks assigned within such systems is implemented.
In recent years augmentation and complications of challenges such as consumption of energy, deterioration of air and sound pollutions, weakness in planning and management, time wasted in traffic and delays due to traffic congestion and above all severity of increase in accidents and reduction of safety in urban transportation has all changed the conception of sustainable transportation viewpoint in most of countries worldwide. This article, therefore, focuses on sustainable conceptions from economic, social and environmental viewpoints for finding ways to tackle the existing hurdles with the primary definitions of these conceptions in addition to presenting some matrixes for evaluation.
In continuation, the article delineates some very important strategies for reaching these goals, namely, implementation of Intelligent Transportation Systems (ITS), efficient and opportune traffic management and practical policies in order to embody them.

**Keywords**: Urban Sustainable Development, Urban Traffic Management, Intelligent Transport Systems (ITS).

## 1. Introduction

Transportation sustainability is largely being measured by transportation system effectiveness and efficiency as well as the environmental and climate impacts of the system. (Jeon, C M; Amekudzi, 2004)Sustainable transport systems make a positive contribution to the environmental, social and economic sustainability of the communities they serve. Transport systems exist to provide social and economic connections, and people quickly take up the opportunities offered by increased mobility. (Schafer A., 1998)The advantages of increased mobility need to be weighed against the environmental, social and economic costs that transport systems pose. Transport systems have significant impacts on the environment, accounting for between 20% and 25% of world energy consumption and carbon dioxide emissions. (World Energy Council. Retrieved 2009-05-26, 2009) Greenhouse gas emissions from transport are increasing at a faster rate than any other energy using sector. (Intergovernmental Panel on Climate Change. Retrieved 2009-05-26, 2009) Road transport is also a major contributor to local air pollution and smog. (US Environmental Protection Agency. 2002. Retrieved 2009-03-18, 2009). The social costs of transport include road crashes, air pollution, physical inactivity, (Health effects of transport. Retrieved 2008-08-29, 2008) time taken away from the family while commuting and vulnerability to fuel price increases. Many of these negative impacts fall disproportionately on those social groups who are also least likely to own and drive cars. (Final report on transport and social exclusion (PDF). Retrieved 2003-02-01, 2003) Traffic congestion imposes economic costs by wasting people's time and by slowing the delivery of goods and services.
Alongside physical and regulatory measures for bus priority, there has been a significant advance in the use of new communications, information and control technologies in transport, under the general heading of intelligent transport systems (ITS). To date, ITS has been particularly prominent in the automobile sector, with various navigation and driver support systems, and in traveler information systems, including those managed by local authorities (variable message signs for congestions and/ or parking information, real-time displays at bus-stops, etc.). (Nick B.Hounsell, 2004)
Key applications of ITS to help keep buses moving are summarized in the following sections.

[1] Corresponding author: m_zourbakhsh@iauest.ac.ir

## 2. INTELLIGENT TRANSPORT SYSTEMS AND APPLICATIONS FOR BUSES

### 2.1. Automatic vehicle location

A number of AVL systems have also been installed in bus fleets over the past decade, as a platform to support ITS application such as those illustrated in Fig.1. (Nick B.Hounsell, 2004)

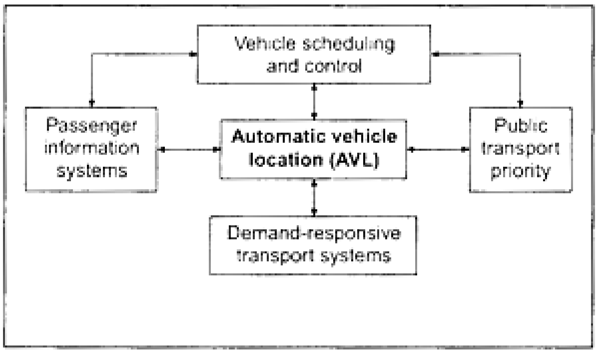



**Fig. 1.**
*AVL and application for buses*
*Source: Municipal Engineer 157 Issue MEI* (Nick B.Hounsell, 2004)

Automatic vehicle location (AVL) is a means for automatically determining and transmitting the geographic location of a vehicle. This data, from one or more vehicles, may then be collected by a vehicle tracking system for a picture of vehicle travel. Most commonly, the location is determined using GPS, and the transmission mechanism is SMS, GPRS, a satellite or terrestrial radio from the vehicle to a radio receiver. Vehicle location technologies can be used in the following scenarios:
***Fleet management:*** when managing a fleet of vehicles, knowing the real-time location of all drivers allows management to meet customer needs more efficiently. Vehicle location information can also be used to verify that legal requirements are being met: for example, that drivers are taking rest breaks and obeying speed limits.
***Passenger Information:*** Real-time Passenger information systems use predictions based on AVL input to show the expected arrival and departure times of Public Transport services.
***Asset tracking:*** companies needing to track valuable assets for insurance or other monitoring purposes can now plot the real-time asset location on a map and closely monitor movement and operating status. For example, haulage and logistics companies often operate trucks with detachable load carrying units. In this case, trailers can be tracked independently of the cabs used to drive them. Combining vehicle location with inventory management that can be used to reconcile which item is currently on which vehicle can be used to identify physical location down to the level of individual packages.
***Field worker management:*** companies with a field service or sales workforce can use information from vehicle tracking systems to plan field workers' time, schedule subsequent customer visits and be able to operate these departments efficiently.
***Covert surveillance:*** vehicle location devices attached covertly by law enforcement or espionage organizations can be used to track journeys made by individuals who are under surveillance. (Wikipedia)

### 2.2. Real-time passenger information

Real-time passenger information (RTPI) systems include internet- based facilities that provide valuable information and reassurance to passengers where there is service irregularity (e.g. due to congestion). The information may be perceived as a 'dynamic timetable' and it appears that some passengers feel buses are 'on time' if the bus arrives according to this timetable, even though adherence to the static timetable is the true yardstick of service quality.

Current operational information on service running is collected from automatic vehicle location systems and from control systems, including incident capture systems. This information can be compared by computers with the published service timetable to generate a prediction of how services will run in the next few minutes to hours. This may be informed by additional information: for instance, bus services will be affected by congestion on the road network, while all services may be affected by adverse weather conditions.

The capital and revenue costs for traveler information systems can be calculated with reasonable accuracy. However, the derivation of tangible financial benefits is far more difficult to establish and as a consequence, there is very little research. This directs the business model for information systems towards the "softer" merits such as traveler confidence etc. It is worth noting that there must be an actual value as individuals are willing to pay for systems that give them access to real time data relating to their journey. The difficulty is establishing what this is for each individual person and perhaps each individual piece of roadside hardware (Wikipedia)

### 2.3. Bus priority using AVL

Real-time knowledge of the progress of all buses on their routes, as provided by AVL, provides the opportunity for sophisticated algorithms for bus priority at traffic signals. Traditionally this form of priority has been provided equally to all buses using special bus detectors on the approaches to signals, detecting buses directly or through interaction with on-bus equipment (e.g. a transponder). While this reduces junction delays, it has little effect on punctuality (for low-frequency services) or regularity (for high-frequency services). AVL allows bus-specific priority to be allocated according to the needs of each bus at each junction. An example of this has been implemented in trials in London (Fig. 2), where a priority algorithm in the AVL center calculates priority levels for each bus according to the ratio of the actual headway to the bus in front to the scheduled headway. The higher this ratio, the higher the level of priority requested. No priority may be provided where this ratio is less than1.The priority level required is transmitted to each bus and then on to the traffic signal controller via the appropriate roadside beacon. The actual priority awarded is determined by the signal control system (SCOOT in this case). For example, high priority may allow non-priority traffic stages to operate up to a relatively high degree of saturation, a parameter which SCOOT is constantly calculating and monitoring on each link. (Bus Priority in SCOOT. Traff1c Advisory Unit, Traff1c Advisory Leaflet 8/00, 2000)

The advantages of this approach are that:
(a) Priority is targeted where it is most needed
(b) Regularity improvements can be achieved, as well as delays reduced
(c) Fewer priority actions occur (if they occur only for 'late' buses), with less consequent effect on non-priority traffic.

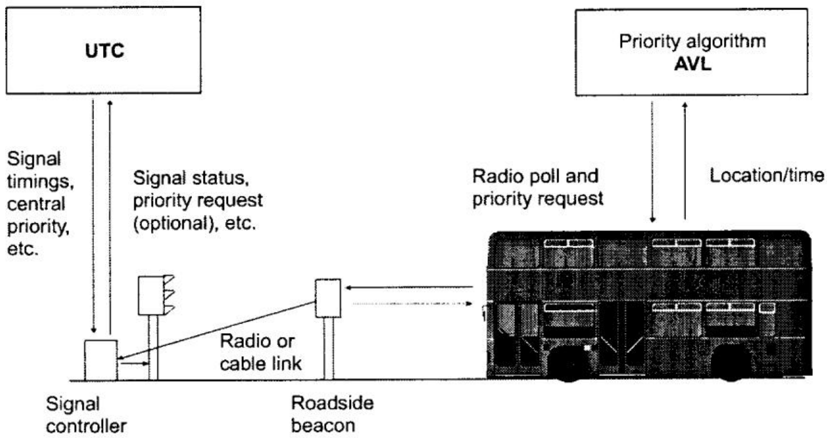



**Fig 2:**
*An Example Bus Priority Architecture*
*Source: Municipal Engineer 157 Issue MEI (Nick B.Hounsell, 2004)*

It is evident from recent research (McLeod F.N., HOUNSELL N.B. and PICCO A, 2003) that there are many systems such as this in cities across Europe, with a wide variety of system architectures and technologies. These depend to some extent on the characteristics of the urban traffic control (UTC) system in place, the available/preferred communications infrastructure and the organizational structure for public transport. It would appear that a more regulated structure for bus operations which exists in mainland Europe facilitates investment in technologies such as this. (Nick B.Hounsell, 2004)

### 2.4. Fleet management

With real-time knowledge of bus locations and performance, relative to schedules, bus operators can implement dynamic fleet management to improve situations which have deteriorated. These could include, for example, adding 'spare' buses to a route to regain required regularity, or removing (turning) buses from congested sections. AVL provides valuable information for what is predominantly operator-determined decisions based on experience. However, it appears that operators often have limited guidance on how to use AVL information to best effect and there would seem to be a good opportunity here to exploit AVL still further for fleet management by developing guidelines for operator support. (Nick B.Hounsell, 2004)The Fleet Management Systems Interface (FMS) is a standard interface to vehicle data of commercial vehicles.

The following data are broadcast at the FMS interface:

Vehicle improvement (all round)/ Vehicle speed (wheel based)/ Vehicle speed (from Tachograph)/Clutch switch (on/off)/Brake switch (on/off)/Cruise control (on/off)/PTO (Status/Mode)/Accelerator pedal position (0–100%)/Total fuel used (litres since lifetime)/Fuel level (0–100%)/Engine speed/Axle weight (kg)/Total engine hours (h)/FMS-Standard software version (supported modes)/Vehicle identification number (ASCII)/Tachograph information/High-resolution vehicle distance/Service distance/Engine coolant temperature. (Wikipedia)

### 2.5. Bus-stop dwell times

The time spent by buses at bus stops for passenger boarding an alighting is an important contributor to overall service speeds and, potentially, regularity /punctuality. Regularity is particularly susceptible to long and variable bus-stop dwell times; once headways become longer than scheduled, passenger boarding numbers often increase, causing increased dwell time. The converse occurs for following buses with shorter headways. In the extreme, this 'natural ' phenomenon can lead to severe irregularity and bus 'bunching'. Bus-stop dwell time can be related to four main factors:

(a) The number of boarding / alighting passengers
(b) The type of bus
(c) Ticketing / payment arrangements
(d) Possible congestion at the bus-stop(s) (e.g.  Due to other buses).

A potential contribution of ITS to reduced bus-stop dwell times can be in the use of (contactless) smart cards for ticketing; this can speed up passenger boarding, as well as providing other functions / data concerning passengers and their origins/ destinations. (Of course passenger boarding times can also be reduced using simpler techniques, such as off-bus ticketing as now introduced widely in London.) Multiple-door buses, such as articulated buses, can also speed up boarding and alighting, particularly where passenger volumes are high. So a key point to note here is that ITS in this case, as in many others can complement efficient design and operational practices rather than being a substitute for them. (Nick B.Hounsell, 2004). A public transport timetable is a document setting out information on service times, to assist passengers with planning a trip. Typically, the timetable will list the times when a service is scheduled to arrive at and depart from specified locations. It may show all movements at a particular location or all movements on a particular route or for a particular stop. It is now also often available in a variety of electronic formats (Wikipedia).

### 2.6. Congestion management

Traffic congestion is a condition on road networks that occurs as use increases, and is characterized by slower speeds, longer trip times, and increased vehicular queuing. As demand approaches the capacity of a road (or of the intersections along the road), extreme traffic congestion sets in. When vehicles are fully stopped for periods of time, this is colloquially known as a traffic jam or traffic snarl-up. Traffic congestion can lead to drivers becoming frustrated and engaging in road rage. (Wikipedia)
An important method for congestion management concerns queue relocation. To help buses, this can involve the use of traffic signals to relocate queues away from links in the network where buses are affected by congestion to links where buses are protected from congestion. This usually involves the integration of physical measures, such as bus lanes and/or bus-only streets with traffic signaling. Where this is undertaken locally at a junction or on a short route section, the application is often referred to as pre-signaling. (Wu J. and H0UNSELL N.B. A, 1998) the role of variable message signing (VMS) in managing congestion can also be noted. Such signs are used predominantly to inform road users of forthcoming events (road works, etc.) or current incidents/congestion. Route advice may also be offered. There is scope here to pay particular attention to sensitive bus routes, in an effort to keep buses moving as much as possible.

### 2.7. Travel Demand Management

Travel demand management (TDM) is the application of strategies and policies to reduce travel demand (specifically that of single-occupancy private vehicles), or to redistribute this demand in space or in time. (FHWA, Travel Demand Management) (Nelson, Donna C., 2000). Demand management measures can play an important role in attempting to manage traffic demand to levels sufficiently below capacity to keep all traffic freely flowing. Measures can include access control, parking management/pricing, park- and-ride and road-user charging. (Nick B.Hounsell, 2004)In transport, as in any network, managing demand can be a cost-effective alternative to increasing capacity. A demand management approach to transport also has the potential to deliver better environmental outcomes, improved public health, stronger communities, and more prosperous and livable cities. (Wikipedia)

### 2.8. Enforcement

A traffic enforcement camera system, consisting of a camera and a vehicle-monitoring device, is used to detect and identify vehicles disobeying a speed limit or some other road legal requirement and automatically ticket offenders based on the license plate number. Traffic tickets are sent by mail. Applications include:
***Speed cameras*** that identify vehicles traveling over the legal speed limit. Many such devices use radar to detect a vehicle's speed or electromagnetic loops buried in each lane of the road.
***Red light cameras*** that detect vehicles that cross a stop line or designated stopping place while a red traffic light is showing.
***Bus lane cameras*** that identify vehicles traveling in lanes reserved for buses. In some jurisdictions, bus lanes can also be used by taxis or vehicles engaged in carpooling.
***Level crossing cameras*** that identify vehicles crossing railways at grade illegally.
***Double white line cameras*** that identify vehicles crossing these lines.
***High-occupancy vehicle lane cameras*** that identify vehicles violating HOV requirements.
***Turn cameras at intersections*** where specific turns are prohibited on red. This type of camera is mostly used in cities or heavy populated areas. (Wikipedia)

### 2.9. Smart buses

Bus rapid transit (BRT) is a bus-based mass transit system. A true BRT system generally has specialized design, services and infrastructure to improve system quality and remove the typical causes of delay. Sometimes described as a "surface subway", BRT aims to combine the capacity and speed of light rail or metro with the flexibility, lower cost and simplicity of a bus system. Keeping Buses Moving. Local Transport Note 1/97, 1997) to be considered BRT, buses should operate for a significant part of their journey within a fully dedicated right of way (bus way) to avoid traffic congestion. In addition, a true BRT system will have most of the following elements:
Alignment in the center of the road (to avoid typical curb-side delays)
Stations with off-board fare collection (to reduce boarding and alighting delay related to paying the driver)
Station platforms level with the bus floor (to reduce boarding and alighting delay caused by steps)
Bus priority at intersections (to avoid intersection signal delay).
The concept of the 'smart bus' is continuously evolving, with current possibilities illustrated in the schematic in Fig.3 (Nick B.Hounsell, 2004)

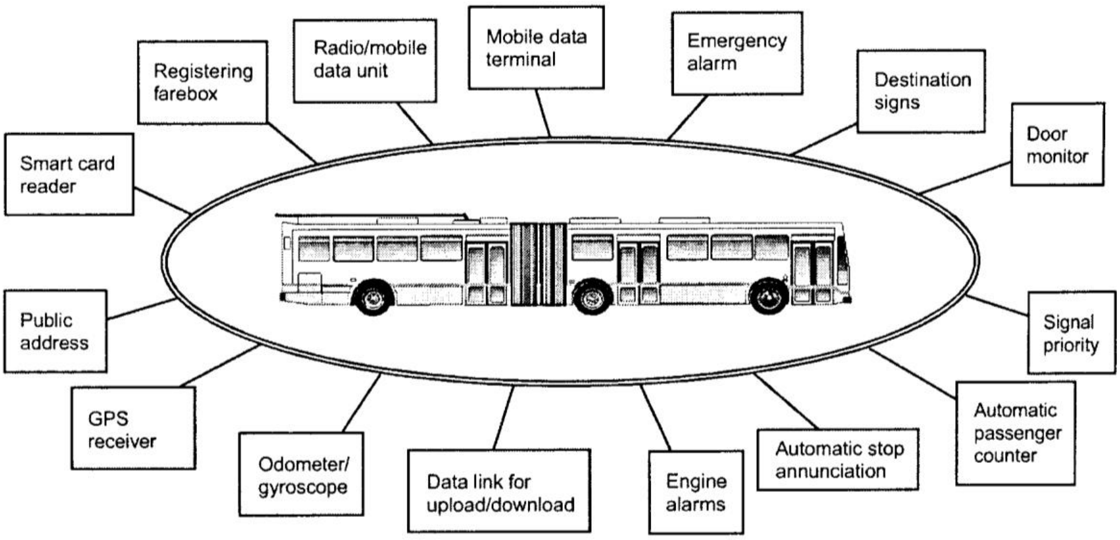



**Fig3:**
*Example of a smart bus with numerous ITS related function*
*Source: Municipal Engineer 157 Issue MEI (Nick B.Hounsell, 2004)*

### 3. Conclusion

This paper has illustrated a range of intelligent transportation systems (ITS) are advanced applications which, without embodying intelligence as such, aim to provide innovative services relating to different modes of transport and traffic management and enable various users to be better informed and make safer, more coordinated, and 'smarter' use of transport networks. The potential is dearly extensive, but the realization of this potential is, arguably, much more dependent on factors such as organizational structures, funding and commitment, rather than technology availability. With increasing sophistication and options for ITS, there is also a need for improved skills in systems specification and subsequent operations and management, if the benefits are to be achieved and long-lasting. (Nick B.Hounsell, 2004).

For most towns and cities in the world, buses are the main form of public transport, usually sharing road space with other traffic over significant parts of the network. This means that buses can then suffer from the same delays as other traffic when congestion builds up. This can have a serious impact on bus operations, where buses are needed for repeated round trips: delays can have significant knock-on effects on reliability/regularity downstream of the problem and on subsequent journeys, resulting in increased operating costs, increased waiting time for passengers and a poorer quality of service overall. A downward spiral of deteriorating service can then occur, with a loss of patronage and potential modal shift in the wrong direction buses moving, from which three main categories of bus priority can be identified:
(a) Priority through traffic management
(b) Physical / regulatory measures such as with-flow and contra-flow lanes, bus-only streets and bus stop treatments
(c) Priority at junctions, particularly at traffic signals. (Nick B.Hounsell, 2004)

Traditional transport planning aims to improve mobility, especially for vehicles, and may fail to adequately consider wider impacts. But the real purpose of transport is access - to work, education, goods and services, friends and family - and there are proven techniques to improve access while simultaneously reducing environmental and social impacts, and managing traffic congestion. (Todd Litman, 1998)Communities which are successfully improving the sustainability of their transport networks are doing so as part of a wider programmed of creating more vibrant, livable, sustainable cities.

**References**

DEPARTMENT OF THE ENVIRONMENT, TRANSPORT AND THE REGIONS. The Stationery Office, London, 1997. Keeping Buses Moving. Local Transport Note 1/97.

DEPARTMENT FOR TRANSPORT. Traffic Advisory Unit, London, 2003. Public Transport Priority, Traffic Advisory Leaflet ITS 5/ 03.

DEPARTMENT FOR TRANSPORT. , London, 2004. Bus Priority: the Way Ahead.
Bus Partnership Forum. Available from internet: www.buspriority.org.uk

DEPARTMENT OF THE ENVIRONMENT, TRANSPORT AND THE REGIONS. , London, 2000. Bus Priority in SCOOT. Traff1c Advisory Unit, Traff1c Advisory Leaflet 8/00.

DEPARTMENT OF THE ENVIRONMENT, TRANSPORT AND THE REGIONS, London, 2000.SCOOT Gating. Traffic Advisory Unit, Traffic Advisory Leaflet 7/ 00.

European Commission, Brussels, 2003 CODE. Guidelines for implementing real-time information
Projects for city-wide public transport. Available from Internet :< ftp: / / ftp.cordis.lu / pub/ telematics/ docs/ tap transport / real_time_guide.pdf>

FHWA, Travel Demand Management, Available from Internet :< http://ops.fhwa.dot.gov/tdm/index.htm>

Intergovernmental Panel on Climate Change 2007.IPCC Fourth Assessment Report: Mitigation of Climate Change, chapter 5, Transport and its Infrastructure (PDF). Intergovernmental Panel on Climate Change. Retrieved 2009-05-26

Jeon, C M; Amekudzi .March 2005, Addressing Sustainability in Transportation Systems: Definitions, Indicators, and Metrics, JOURNAL OF INFRASTRUCTURE SYSTEMS: 31–50

Joo s E, 1995, Economy and Ecology are no Contradictions: Three Messages from Zurich Concerning the New Transport Policy. Zurich Transport Authority unpublished paper.

London, 2002. CARTWRIGHT M, TILLEY G., RADIA B. and THANCANAMOOTOO B.UTMC demonstrators. Results so far. In Proceedings of the11th International Conference on Road Transport Information and Control, Institution of Electrical Engineers, Conference Publication No. 486.

Loughborough University. 14-18 September1998. HOUNSELL N.B., RILEY K. and Alexander M. Enhanced bus priority using signals-the INCOME project. In Proceedings of the 26th European Transport Conference.

March 2004.Nick B.Hounsell. Keeping buses moving: role of intelligent transport systems, In Proceedings of the institution of Civil Engineer 157 Issue MEI Pages 55-60.

National multipollutant emissions comparison by source sector in 2002. US Environmental Protection Agency. 2002. Retrieved 2009-03-18.

15. Nelson, Donna C., Editor 2000. Intelligent Transportation Primer. Institute of Transportation Engineers, Washington, D.C. pp. 10-1. ISBN 0-935403-45-0. 19. Schafer, A.1998.The global demand for motorized mobility. Transportation Research A 32(6), 455-477.

16. Social Exclusion Unit, Office of the Prime Minister (UK). Making the Connections - final report on transport and social exclusion (PDF). Retrieved 2003-02-01.

Todd Litman .1998.Measuring Transportation: Traffic, Mobility and Accessibility. Victoria Transport Policy Institute. Retrieved 2009-03-18.

Turin, Italy, November 2000. Real time passenger information and bus prioritySystem in Cardiff. . In Proceedings of the 7th World Congress on Intelligent Transport Systems.

TRANSPORT FOR LONDON. Congestion Charging: Six Months On2003, Available from Internet:
< www.tfl.gov.uk>

UK RTIG, 2003.UNITED KINGDDM REAL-TIME IMPLEMENTATION GROUP. Available from Internet: www.rtigimplementers.org.uk

UNIVERSITY OF SOUTHAMPTON.1973. Evaluation of the Bitteme Bus Priority Scheme. Transportation Research Group.

UNITED KINGDOM, UK RTIG, 2003. Outline Requirements Specification for Real Time Public Transport Information Systems. REAL-TIME IMPLEMENTATION GROUP. Real Time Information Group library reference RTIG02-2.1

Wu J. and H0UNSELL N.B. A, 1998, *Bus priority using pre-signals. Transportation Research* 32, No.8, 563-583.

World Energy Council 2007. Transport Technologies and Policy Scenarios. World Energy Council. Retrieved 2009-05-26.

World Health Organization, Europe. Health effects of transport. Retrieved 2008-08-29.

Wikipedia. Available from Internet :< www.wikipedia.org>

INTERNATIONAL CONFERENCE ON TRAFFIC AND TRANSPORT ENGINEERING

ictte

www.ijtte.com