

THE CONNECTION BETWEEN SPATIAL AUTOCORRELATION AND BORDER CROSSING TRAFFIC

Zsombor Szabó^{1,2}, Tibor Sipos^{1,3}

¹ Budapest University of Technology and Economics, Faculty of Transportation Engineering and Vehicle Engineering, Department of Transport Technology and Economics – H-1111 Műegyetem rakpart 3. Budapest, Hungary

² KTI Institute for Transport Sciences Nonprofit Ltd., Directorate for Public Transport Services – H-1119 Than Károly utca 3-5. Budapest, Hungary

³ KTI Institute for Transport Sciences Nonprofit Ltd., Directorate for Research and Innovation – H-1119 Than Károly utca 3-5. Budapest, Hungary

Received 16 November 2021; accepted 27 December 2021

Abstract: Examining the traffic of border crossing points is a priority task due to the exploitation of the advantages of the national economy. An essential part in this process is the examination of the autocorrelation in the data. In this article, a theoretical approach was used: the geographically located physical parameters were removed, and random networks were generated and analysed to investigate the effect of autocorrelation. Spatial autocorrelation could explain up to nearly 50 percent of the effects with a well-chosen spatial weight matrix. This article can also be interpreted as the first element of a research series, thus defining future research directions and the steps of generalizability of the models is crucial.

Keywords: transport geography, spatial econometrics, network analysis, transport economics.

1. Introduction

Determining the traffic at border crossing points is a priority task, as it can be used to examine numerous parameters that influence economy. On the one hand, these infrastructure elements are essential parts of the international trade networks, so even a small change in parameters can harm international relations (Brown and Anderson, 2015). In addition, they also have an impact on regional relations; thus, infrastructure can improve the well-being of the borderlands through several effects, such as cross-border commuting (Anderson *et al.*, 2014; Szabó and Sipos, 2020), one-day trips (Bradbury, 2013) or shopping

tourism (Rietveld *et al.*, 2001). Moreover, these infrastructure elements can create jobs, and in addition to facilitating commuting conditions, the number of border guards also has a significant influence on the number of jobs (Avetisyan *et al.*, 2015).

Our research aims to prove under random conditions that the traffic data measured at cordon points on an object representing separation effect (for example, a country border) has spatial autocorrelation; thus, spatial econometric modelling techniques can be applied in their case.

As a transportation approach, the main characteristics between border sections

¹ Corresponding author: szabo.zsombor@kjk.bme.hu

can be given as the difference between the density and the administrative rules for crossing. In terms of the density of border crossings, there is an infrastructure element in every 25 kilometres in Hungary. Although this statistic is better than the average for Central and Eastern Europe (75 km), it lags far behind the average distance (5-6 km) for Western Europe (Illés, 2008). With joining the Schengen area within the European Union, the administrative barriers caused by borders have been significantly reduced, leading to an increase in the number of crossing points. Developments in this direction generally increase traffic at border crossing points, but the extent and manifestation of the interaction are far from clear.

The basis of the problem is that there are numerous methods for estimating border-crossing flows and analysing the coefficients of the explanatory variables used, but these do not consider the effects caused by the whole system. In most cases, the development of traffic at pre-selected border crossings is examined with respect to time. Opananon and Kitthamkesorn summarise the development opportunities for one of the busiest border crossings between Thailand and Malaysia (Opananon and Kitthamkesorn, 2016). In his article, Burt examines traffic data for twelve border crossings to see that the 9/11 terrorist attacks did not have a significant negative impact on the Canada-US international shipping and bilateral relations (Burt, 2009). Park and their co-authors' primary goal is to determine the cost of delays caused by crossing borders. This is based on two bridges on the Buffalo-Toronto axis, which are available for HGV traffic (Park et al., 2014). Border conditions in Southeast Europe were analysed by Miltiadou

and their co-authors for land transport chains (Miltiadou *et al.*, 2017). However, each year, border traffic is influenced by previous years' numbers and by traffic at other crossing points, analogous to Tobler's definition of spatial autocorrelation (Tobler, 1970).

In this article, a computer program is described which can be used for measuring the bare effect of spatial autocorrelation. The methodology is shown in detail, together with the first results. Finally, the suggestions for further development are presented.

2. Methodology

Since the goal is to set up a general model for the border crossing points, it is necessary to generate random networks. In general, a country's transport network can be of three types (Baráth and Rimóczi, 2011):

- Centralised: the main routes are in a radial direction from the centre (often capital), which are connected in circles by other main roads (for example, Hungary);
- Square lattice (Horiguchi and Sakakibara, 1998): major transport routes connect significant cities horizontally and vertically, the minor roads connect these infrastructures (for example, Germany, the United States, or the European International E-road network);
- Mixed: a combination of the two aforementioned types (for example, Poland).

Since these networks are quite different from each other on their own, this article will focus on centralised networks. As random networks are applied, the methodology

is based on Monte Carlo simulation (von Neumann and Ulam, 1949), which is also very often used in the engineering sciences – e.g., Mályusz and Varga (2018) or Babu *et al.* (2021). In random networks, the expected traffic at the border crossings is assigned using a gravity model (Odlyzko, 2015), for which a spatial econometric model is built only by using the spatial autocorrelation parameter as an explanatory variable (Anselin, 1988; Varga, 2002). Note that in this case, by random a randomly generated network is understood, which significantly differs from the mathematical definition – e.g., Barabási *et al.* (1999, 2000).

The study was performed in a MATLAB (MATLAB, 2019) environment, where a script was responsible for running it, calling five functions we defined. These were as follows:

- Network generation: generates a random graph that meets pre-set requirements. This determines the location of each settlement and connecting road, the country borders, and the border crossing points;
- Determining the population of cities: assigning the weights required for the gravity model as a population to each city in the model based on Auerbach's rule (Auerbach, 1913; Hagget, 2001);
- Border crossing traffic generation: the third step is to generate border crossing traffic using the gravity model.
- Spatial weight matrix generation: generates the spatial weight matrix required for spatial econometric models based on the position of border crossings.

2.1. Network Generation

During network generation, the goal was to create a random centralised model similar to a real network, which is illustrated by Figures 1 and 2 – Figure 1. is the real network, while Figure 2. is the modelled one. The network generation procedure is similar to the models constructed by Cats (2017). In terms of the structure of the network, it consists of a central city (this is the capital of the country being studied – later this will be necessary) surrounded by two concentric circles (in the distance of $r_1 = 1$, and $r_1 + r_2$) and the number of roads from the capital leading to the first circle and from the first circle towards the second one. Thus, a total of three types of terrain points can be identified: intersections, cities, and border crossings, which can be identified by polar coordinates due to their circular design. The terrain points describing the network and their location are generated from the number of roads starting from the capital and from the value of r_2 with random parameters.

Each terrain point is assigned to a country. Five countries were distinguished, one central (this is the country being studied) and four neighbouring countries. The total area of the central country is included in the model, while only that part of the neighbouring countries is included, which is close to the studied country. Numerous criteria were considered for the classification into countries, such as that the countries should be connected, that there should be no interference, and that the main country should be of an interpretable size; thus, half of the settlements in the inner ring should belong to the central country.

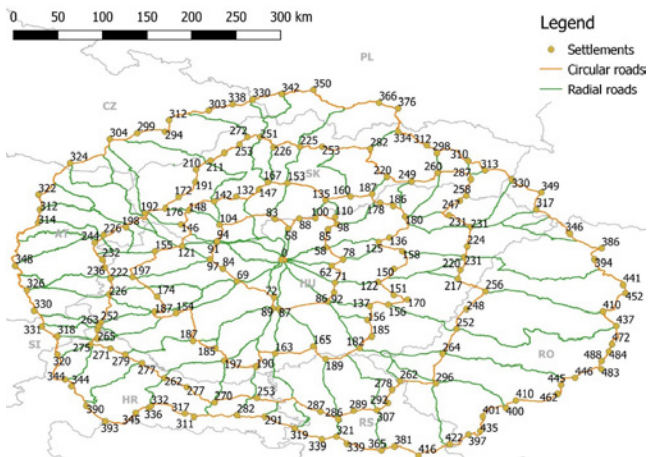


Fig. 1.
 Circular-radial Structure of the Road Network of the Carpathian Basin and Distances from Budapest
 Source: own work, based upon OSM (OpenStreetMap contributors, 2017) data

Based on all this, networks like Fig. 2. are created (different colours for different countries, squares for cities, and dots for border crossings). In the case of borders, the

colour means which country can be reached from the central country (green). Border crossing points that do not affect the central country were not considered.

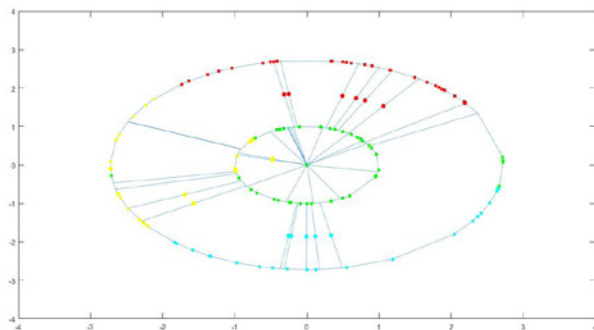


Fig. 2.
 Example of a Generated Network
 Source: own work, edited in MATLAB (2019)

Due to the application of the gravity model, weights have to be assigned to the settlements, which in this case will be the population. This was also randomly

generated based upon Auerbach (1913); as random networks are generated without real physical properties, we could assume ideal countries, where if we rank the settlements

according to their population, the population of the n^{th} settlement is $1/n$ part of its capital (Hagget, 2001).

Thus, we only need to generate two values; one is the capital population (this was set between one and two million, as this is the typical size in Central Europe), and the place occupied by the given city in the given country in terms of size. In the central country, we assume as many settlements as we generated in the first model, but in the case of the four neighbouring countries, we assume more cities than are shown in the model – because the analysis is on the central country; thus, from the neighbouring countries only the borderlands are in our interest.

2.2. Border Crossing Traffic Generation

To set up the gravity model, in addition to the weights of each starting and destination point, the distance between them is also required. There are numerous different methods for calculating this, the best known being perhaps the Dijkstra algorithm (Dijkstra, 1959), but it is worth examining how travellers choose their route. Several studies have addressed how travel cost developments affect those selecting a particular route (Bansal *et al.*, 2020; Prato, 2009). Thus, it is not enough to determine the shortest route, but it is also necessary to choose alternative routes. For this, we used Yen's k shortest path algorithm (Yen, 1976).

In general, there are two types of algorithms for finding the shortest k path, one allowing the use of loops and the other being loopless (Yen, 1976). Yen's algorithm belongs to the latter category, which is closer to reality in the sense that road users do not choose an alternative route by inserting a circle into their path.

After identifying the shortest routes, the traffic demand between each pair of settlements can be determined. Naturally, Yen's algorithm is also suitable for finding the shortest path, since the first shortest path is the same as the global one between two points. Thus, all the parameters required for the gravity model were produced. The essence is that the traffic demand between two settlements can be described by analogy with Newton's law of universal gravitation known from physics (1).

$$F = -\gamma \frac{m_1 m_2}{r^3} \mathbf{r} \quad (1)$$

where:

- F : attractive force;
- γ : gravitational constant ($6.67 \times 10^{-11} \frac{\text{m}^3}{\text{kg} \times \text{s}^2}$);
- m_1, m_2 : mass of the point-like bodies;
- r : distance vector between the point-like bodies.

This methodology is very often used in the transport sciences (Hummels, 1999; Jung *et al.*, 2008; Limão and Venables, 2001; Okubo, 2004; Tagai *et al.*, 2008) is an accepted traffic-generating procedure. In the present case, we will use the "classic" approach, i.e., we have chosen the population as weights and the distance from the result of Yen's algorithm. Thus, the travel demand (U) between two cities can be calculated by the following equation (2):

$$U_{ij} = \frac{p_i p_j}{d_{ij}^a} \quad \forall i \neq j \quad (2)$$

where:

- P_i : the population of the i^{th} settlement $\forall i$;
- d_{ij} : the length of the shortest path between the i^{th} and the j^{th} settlement $\forall i \neq j$;
- a : the exponent of the resistance.

In the model, it is possible to set the resistance exponent, which in this case was taken to be 1.94 according to Hagget (2001). Since the transportation demands in the present case have no physical equivalent – thus, they only represent a theoretical demand burden the given section – the value corresponding to γ can be chosen to be 1.

The needs determined in this way can be assigned to the defined shortest routes between the two settlements. The distance-decay property is used as the basis for this assignment (Hagget, 2001). Let $l \in [1..k]$ be the serial number of the given route between the settlement i and j onto which we assign part of the traffic. The volume of traffic on a given route can be calculated as follows (3):

$$N_{i,j}^l = U_{ij} \frac{d_{i,j}^{l-A}}{\sum_{l=1}^k (d_{i,j}^{l-A})} \quad (3)$$

where:

- $N_{i,j}^l$: traffic volume between settlements i and j on the l^{th} shortest path;
- $d_{i,j}^l$: distance between settlements i and j on the l^{th} shortest path;
- A : the parameter of distance-decay (currently also 1.94 (Hagget, 2001)).

Based on all this, if there is a border crossing point on the given route, the route's traffic will be assigned to the border crossing point. Thus, the traffic of all crossings can be determined.

2.3. Setting up the Spatial Econometric Model

The spatial econometric model was set up based on the work of Attila Varga (Varga, 2002) and Luc Anselin (Anselin, 1988). With the help of this, to verify our hypothesis, we seek answers to the following questions:

- Is there autocorrelation in the traffic data of border crossings?
- Does the spatial autocorrelation parameter have a significant effect on border crossing traffic?
- To what extent does spatial autocorrelation alone affect border crossing traffic?

However, based on all this, the formation of initial expectations is not clear, as the direction of autocorrelation is difficult to determine. In the case of Hungary, the border crossing points follow a unique pattern. On the one hand, the main crossing points can be considered competing (negative autocorrelation), so a diversion can be observed between them. The negative autocorrelation is also reinforced, because smaller crossings, which usually have less traffic than the larger ones, are located next to the main crossings. In contrast, however, it can be observed that smaller border crossings play a complementary role, supporting a positive autocorrelation. It may be equally important that in the case of a popular destination, the traffic at adjacent border crossings also increases, reinforcing the theory of positive autocorrelation.

There are several tests to prove the presence of autocorrelation. The most frequently used tool for this is Moran's I -test (Moran, 1948). The statistics can be calculated as follows (Varga, 2002).

$$I = \frac{N \sum_{i,j} (w_{ij}(x_i - \mu)(x_j - \mu))}{S_0 \sum_i (x_i - \mu)^2} \quad (4)$$

where:

- N : number of observations;
- x_i, x_j : data of two observed points;
- μ : the expected value of x ;
- w_{ij} : the elements of the spatial weight matrix;
- S_0 : normaliser – $S_0 = \sum_{i,j} w_{ij}$.

Our analysis aims to prove that spatial autocorrelation influences the development of cordon point traffic volumes; thus, we will build a spatial autoregression model (SAR) from the tools of spatial econometric modelling (5). However, since this is a theoretical model, we examine the network alone without area-specific properties; thus, the exogenous variables can be omitted (6).

$$y_{(N \times 1)} = \rho W_{(N \times N)} y_{(N \times 1)} + X_{(N \times K)} \beta_{(K \times 1)} + \varepsilon_{(N \times 1)} \quad (5)$$

$$y = a + \rho W y + \varepsilon \quad (6)$$

where:

- **Y**: the traffic flow vector of the border crossing points;
- **a**: constant term;
- **ρ** : spatial autocorrelation parameter;
- **X**: exogenous variable matrix;
- **β** : exogenous variable parameter vector;
- **Wy**: spatially lagged parameter vector;
- **W**: spatial weight matrix;
- **ε** : vector of the errors.

If we examine formula (5), we can see that everything is available except for the spatial weight matrix. Spatial weight matrices were determined based on our previous research (Szabó *et al.*, 2017; Szabó and Török, 2018a, 2018b). According to the literature (Getis, 1991), two types of weight matrices were used. One is when the weight is the power of the reciprocal of the distance with some exponent (in this case, 1.94). The other is when the elements of the weight matrix can take the value $w_{ij} = \begin{cases} 0 \\ 1 \end{cases}$, based on whether *i* and *j* are adjacent or not. The criterion of neighborhood can be defined as desired. Five weight matrices were used in the present analysis:

- Everything is connected to everything (inverse distance-based) – a) matrix;
- Only adjacent elements interact (inverse distance-based) – b) matrix;
- Only adjacent elements interact (binary) – c) matrix;
- Only those who go to one country interact (inverse distance-based) – d) matrix;
- Only those who go to one country interact (binary) – e) matrix.

For distance-based weight matrices, it is essential to choose the exponent of the distance. The program allows free parameter selection in this case; however, only 1.94 was examined (Hagget, 2001).

Three parameters were used for validation. The first parameter is R^2 , which describes the correlation. The second parameter is the significance level of the Moran's *I*-test, which can be calculated using the *z*-test. This requires the mean and the standard deviation. There are several approaches to determining these, depending on what expected value and standard deviation can be assumed; these are summarised by Cliff and Ord (1972) based on the work of Moran (1950). Based on all this, the applied calculation formulas are as follows (if we can assume a normal distribution, the expected value can be used instead of the average) (Anselin, 1988; LeSage, 1998; Varga, 2002):

$$E(I) = \frac{-1}{N-1} \quad (7)$$

$$V(I) = \left(\frac{N}{S_0}\right)^2 \frac{\text{tr}(MWMW') + \text{tr}(MW)^2 + (\text{tr}(MW))^2}{N(N+2)} - E(I)^2 \quad (8)$$

where:

- **W**: unstandardised weight matrix;
- **M**: projection matrix ($M = U - x(x'x)^T x'$);
- **x**: vector containing the observed data;
- **U**: unit matrix.

The third parameter is the significance level of the spatial autocorrelation parameter (ρ). Since we are talking about a classic linear regression model in the present case, the significance level can be determined by the *t*-test, where the value of the test statistics is the following (Nagy and Balogh, 2013):

$$t = \frac{\rho}{s_b} \quad (9)$$

$$s_b = \sqrt{\frac{s_e}{\sum(Wy_i - \bar{W}\bar{y})^2}} \quad (10)$$

$$s_e = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{n}} \quad (11)$$

where:

- *t*: the *t*-value;
- S_b : the standard error of the parameters;
- S_e : standard error of fit;
- \hat{y}_i : values estimated by the linear regression model.

3. Results

As a result of the program presented in detail in the methodological part, due to the capabilities of the Monte Carlo simulation, a large amount of data was generated, the

analysis of which requires the application of statistical methods. Three parameters need to be examined, which are the following:

- R^2 : the extent to which spatially lagged data affect traffic at border crossings;
- *I*: Moran's *I*-test value;
- ρ : spatial autocorrelation parameter.

It would be useful to estimate the expected value of the listed parameters from the data, but to do this, their distribution must first be determined. For this, we used the *R 3.4.0* software (R Core Team, 2017), including the following libraries to test the different distributions:

- normal distribution: *nortest* (Gross and Ligges, 2015);
- exponential distribution: *MASS* (Venables and Ripley, 2002);
- gamma distribution: *EnvStats* (Millard, 2013);
- universal distribution: *spgs* (Hart and Martínez, 2019);
- beta distribution: *EnvStats* (Millard, 2013) and *Sim.DiffProc* (Guidoum and Boukhetala, 2020);
- Gumbel distribution: *goft* (Gonzalez-Estrada and Villasenor-Alva, 2020).

A Kolmogorov–Smirnov-based test (Massey, 1951) was set up for the studied parameters in each case. Except for two issues, we found a suitable distribution for each weight matrix for each parameter estimate, so their expected value can also be determined. This is illustrated in Table 1.

Table 1*The Distribution and the Expected Value of the Important Parameters*

		a) matrix	b) matrix	c) matrix	d) matrix	e) matrix
R^2	Distribution	Exponential	Normal	Normal	Gamma	Gamma
	Statistics	0.0778	0.0635	0.0621	0.0813	0.0737
	Significance level	0.6964	0.5998	0.6345	$\geq 0.1^{**}$	$\geq 0.1^{**}$
	Expected value	0.1577	0.4229	0.4548	0.1884	0.2071
I	Distribution	Normal	Normal	Normal	Normal	Normal
	Statistics	0.0568	0.0664	0.0457	0.0831	0.0882
	Significance level	0.7635	0.5280	0.9508	0.2607	0.1877
	Expected value	-0.0673	0.5196	0.5050	0.2333	0.2445
ρ	Distribution	Gumbel	Normal	Normal*	Normal	Normal*
	Statistics	0.0841	0.0791	0.1198	0.0481	0.1810
	Significance level	0.6017	0.2563	0.0069	0.9227	9.43E-07
	Expected value	-0.0257	0.7348	0.8247	0.3451	0.3847

* Nonsignificant

** The peculiarity of the R library used, and the command containing it is that it does not give a more accurate result.

Source: authors

4. Discussion

4.1. Evaluation

Based on the results, it can be said in general that when each border crossing affects all of the other crossings, it gives worse results than those in which the connection between infrastructures selected according to a rule and close to each other can be assumed. Furthermore, in the case of matrices b)–e), the binary matrices (c) and e)) have a higher R^2 value, but their I value is lower, and no distribution could be fitted to the parameter ρ in these cases, so their expected value cannot be assessed beyond doubt.

Examining the R^2 values, the significantly higher value emerging from the assumption of the interaction of the adjacent border crossings is clear. On the one hand, a value around 0.4–0.5 assumes that the methodology is well applicable, as only spatially lagged parameters explain 40–50 percent of traffic. On the other hand, our previous results show (Szabó et al., 2017; Szabó and Török, 2018a, 2018b; Sipos et al., 2021) that if we do not decompose the physical parameters from

the border crossings, then neighbouring country-dependent assumptions represented by matrices d) and e) are proven to be more effective. The 20 percent explanatory power observed here may still be appropriate for multivariate regression analyses involving multiple parameters or spatial econometric analyses.

Moran's I values assume a positive autocorrelation in all cases except for matrix a). In case a), the expected value obtained by the Monte Carlo simulation is so close to the theoretically expected value of Moran's I that no clear conclusions can be drawn. Thus, the generally accepted positive autocorrelation assumes that the various border crossing development programs – for example (Szalóki, 2017a, 2017b) – have an overall positive effect on cross-border traffic. This is also in line with international research findings – for example (Cavallaro and Dianin, 2020a, 2020b).

This is confirmed by the values of ρ if, except in case a), positive values are taken. It can be said that the use of binary weight matrices assumes higher ρ values; however, in these

cases, we did not find a distribution that can be proven to describe the results.

4.2. Opportunities for Improvement

This article can also be interpreted as a starting element of a series, as further analyses can be performed by generalising the fixed parameters. The generalisation can cover the following parameters:

- network generation;
- the exponent of the spatial resistance;
- determining the settlement population;
- the number of the shortest routes taken into consideration.

In this article, only centralised networks similar to Hungary have been developed during network generation. The analysis of further regular (e.g., square lattice – Germany, or mixed – Poland) or even completely random (e.g., Romania) network structures will be the subject of future research. It is also part of network generation that the location of settlements follows a simplified scheme in this article, as they are always located on the concentric circles. However, this is far from the real-life situation in most cases, so it is advisable to incorporate it into the model in the future. Furthermore, the radial direction of the connecting roads is questionable in the case of centralised networks – such as the position and layout of the main roads 53 and 54 in Hungary – the integration of which into the model will also be a future research direction.

As an exponent of spatial resistances, based on Hagget (2001), we used a value of 1.94 in this article. However, this is only an empirical mean for the interval [0.4; 3.3] stated in the same research. By approaching

the extreme values of the interval, additional rules could be identified.

The population estimates of the settlements were also based on the book of Hagget (2001); although, the property formulated for the parameters is a qualified case of the Zipf distribution (Zipf, 1949). However, later tests revealed that the $1/n^{1.07}$ relation better describes the relationship between cities (Clauset *et al.*, 2009); thus, later implementation of this result may also be desirable into the methodology.

The last parameter to be mentioned is the number of shortest trips. Yen's (1976) algorithm is suitable for finding the k shortest loopless path; however, the determination of the optimal value of k is not yet clear. In the present article, for the sake of simplification, we have chosen $k = 10$, but it will be seen later that this can only be considered as a mean. We hypothesize that the number of the shortest roads to be considered depends on the distance of the settlements and their location. As an example, we examined four relations in Hungary, which differ in distance and location. The relations are as follows:

- 1) Budapest–Debrecen;
- 2) Debrecen–Nyíregyháza;
- 3) Budapest–Székesfehérvár;
- 4) Győr–Szolnok.

The identified relations are illustrated in Table 2. Only the main road network was considered in determining the routes. Thus, we determined the total number of possible routes between the two settlements (n), the length of the shortest road in kilometres and time (d) – by considering the expressway network – and then given the percentage increase of the k^{th} shortest route optimizing for both time and distance. Since we have no longer considered the expressway network

during the calculation of the k shortest path, $k = 1$ does not always correspond to the shortest road. The table's Opt. and Par. columns show that we examine the

shortest paths based on distance or time (Opt. – Optimisation) and which parameter's percentage increase can be seen (Par. – Parameter).

Table 2

The Examined Journeys

No.	n	d	Opt.	Par.	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 10$	$k = 15$	$k = 20$	$k = n$
1)	28	232 km	Dist	Dist	100.86%	100.86%	102.16%	106.90%	109.05%	117.24%	131.47%	146.98%	173.71%
				Time	140.00%	145.00%	146.43%	145.00%	155.00%	184.29%	185.71%	218.57%	272.14%
		140 min	Time	Dist	100.86%	100.86%	106.90%	102.16%	109.05%	116.38%	130.17%	146.98%	173.71%
				Time	140.00%	145.00%	145.00%	146.43%	155.00%	167.14%	188.57%	218.57%	272.14%
2)	4	50 km	Dist	Dist	100.00%	103.82%	219.32%	259.56%	-	-	-	-	259.56%
				Time	100.00%	102.22%	200.00%	271.11%	-	-	-	-	271.11%
		45 min	Time	Dist	100.00%	103.82%	219.32%	259.56%	-	-	-	-	259.56%
				Time	100.00%	102.22%	200.00%	271.11%	-	-	-	-	271.11%
3)	6	65 km	Dist	Dist	107.44%	107.91%	131.78%	175.19%	182.95%	-	-	-	182.95%
				Time	170.45%	165.91%	188.64%	263.64%	218.18%	-	-	-	222.73%
		44 min	Time	Dist	107.91%	107.44%	131.78%	182.95%	182.95%	-	-	-	175.19%
				Time	165.91%	170.45%	188.64%	218.18%	222.73%	-	-	-	263.64%
4)	66	226 km	Dist	Dist	100.44%	101.77%	107.52%	107.52%	110.18%	115.04%	123.01%	126.55%	171.68%
				Time	176.38%	171.65%	177.17%	185.04%	192.13%	215.75%	210.24%	215.75%	275.59%
		127 min	Time	Dist	101.77%	100.44%	107.52%	123.45%	107.52%	113.27%	116.81%	119.91%	146.02%
				Time	171.65%	176.38%	177.17%	183.46%	185.04%	194.49%	198.43%	204.72%	276.38%

Source: authors

Based on Table 2, in the case of very close settlements, the number of possible routes does not reach 10 (there are still other directions, but in that case, the increase in time is no longer viable), and in the periphery – 2) – the value is lower. In contrast, long distances often lead to more paths, especially if we do not start from the centre – 1) – but pass through it – 4). In these cases, the 10th shortest path causes an increase of about 15 and 20–40 percentage points in time compared to the $k = 1$ case, which can still be considered viable. Furthermore, it can be said that there are many alternate routes with roughly the same conditions in this range.

5. Conclusion

The aim of this paper was to create a basic theoretical model in which we generated

random networks to determine the effect of spatial autocorrelation alone – deprived of real geographical parameters – on the traffic of border crossings. As a result, during the analysis of such spatial econometric models, weight matrices built upon the principle of only the close units have an effect on each other can be better applied instead of the everything is related to everyone approach.

We have analysed numerous cases with different weight matrices – a) to e) cases. The best weight matrices are those where only adjacent infrastructure elements interact with each other. In this case, the spatial parameter explains nearly 50 percent of the data; although, in cases with the matrices used in our previous articles, there seems to be nearly 20 percent of explanatory force, which can be considered as acceptable

with taking into consideration that in such cases the physical properties are available. Overall, the analysis shows that in the case of infrastructures crossing a separation effect, it may be worthwhile to examine the traffic estimation models from a spatial econometric point of view.

This article can be interpreted as the beginning of a series of articles. The model was programmed so that several parameters – exponent of the distance of the gravity model, exponent of the spatial weight matrix resistance – could be freely changed; however, currently only one value, the expected value, was examined. In addition, there are several parameters – the population of settlements, the applied network structure – from which we also considered a special case, but the implementation of other cases is expected only at the end of another programming process.

References

- Anderson, W.P.; Maoh, H.F.; Burke, C.M. 2014. Passenger car flows across the Canada–US border: The effect of 9/11, *Transport Policy* 35: 50–56. <https://doi.org/10.1016/j.tranpol.2014.05.005>.
- Anselin, L. 1988. *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers, The Netherlands. 284 p.
- Auerbach, F. 1913. Das Gesetz der Bevölkerungskonzentration (The Law of Population Concentration), *Petermanns Geographische Mitteilungen* 59: 74–76.
- Avetisyan, M.; Heatwole, N.; Rose, A.; Roberts, B. 2015. Competitiveness and macroeconomic impacts of reduced wait times at U.S. land freight border crossings, *Transportation Research Part A: Policy and Practice* 78: 84–101. <https://doi.org/10.1016/j.tra.2015.04.027>.
- Babu, D.; Balan, S.; Anjaneyulu, M.V.L.R. 2021. Population Synthesis for a City in a Developing Country, *Periodica Polytechnica Transportation Engineering* 49(2): 156–163. <https://doi.org/10.3311/PPtr.14938>.
- Bansal, P.; Hörcher, D.; Graham, D.J. 2020. A Dynamic Choice Model with Heterogeneous Decision Rules: Application in Estimating the User Cost of Rail Crowding, arXiv preprint arXiv:2007.03682.
- Barabási, A.-L.; Albert, R.; Jeong, H. 1999. Mean-field theory for scale-free random networks, *Physica A: Statistical Mechanics and its Applications* 272(1–2): 173–187. [https://doi.org/10.1016/S0378-4371\(99\)00291-5](https://doi.org/10.1016/S0378-4371(99)00291-5).
- Barabási, A.-L.; Albert, R.; Jeong, H. 2000. Scale-free characteristics of random networks: the topology of the world-wide web, *Physica A: Statistical Mechanics and its Applications* 281(1–4): 69–77. [https://doi.org/10.1016/S0378-4371\(00\)00018-2](https://doi.org/10.1016/S0378-4371(00)00018-2).
- Baráth, F.; Rimóczi, M. 2011. *Közlekedési földrajz (Transport Geography)*. Műszaki Kiadó, Hungary. 292 p.
- Bradbury, S.L. 2013. The impact of security on travelers across the Canada–US border, *Journal of Transport Geography* 26: 139–146. <https://doi.org/10.1016/j.jtrangeo.2012.08.009>.
- Brown, M.W.; Anderson, W.P. 2015. How thick is the border: the relative cost of Canadian domestic and cross-border truck-borne trade, 2004–2009, *Journal of Transport Geography* 42: 10–21. <https://doi.org/10.1016/j.jtrangeo.2014.10.006>.
- Burt, M. 2009. Tighter Border Security and Its Effect on Canadian Exports, *Canadian Public Policy / Analyse de Politiques* 35(2): 149–169.
- Cats, O. 2017. Topological evolution of a metropolitan rail transport network: The case of Stockholm, *Journal of Transport Geography* 62: 172–183. <https://doi.org/10.1016/j.jtrangeo.2017.06.002>.

- Cavallaro, F.; Dianin, A. 2020a. Efficiency of public transport for cross-border commuting: An accessibility-based analysis in Central Europe, *Journal of Transport Geography* 89: 102876. <https://doi.org/10.1016/j.jtrangeo.2020.102876>.
- Cavallaro, F.; Dianin, A. 2020b. An innovative model to estimate the accessibility of a destination by public transport, *Transportation Research Part D: Transport and Environment* 80: 102256. <https://doi.org/10.1016/j.trd.2020.102256>.
- Clauset, A.; Shalizi, C.R.; Newman, M.E.J. 2009. Power-Law Distributions in Empirical Data, *SIAM Review* 51(4): 661–703. <https://doi.org/10.1137/070710111>.
- Cliff, A.; Ord, K. 1972. Testing for Spatial Autocorrelation among Regression Residuals, *Geographical Analysis* 4(3): 267–284. <https://doi.org/10.1111/j.1538-4632.1972.tb00475.x>.
- Dijkstra, E.W. 1959. A Note on Two Problems in Connexion with Graphs, *Numerische Mathematik* 1: 269–271. <https://doi.org/10.1007/BF01386390>.
- Getis, A.; 1991. Spatial interaction and spatial autocorrelation: a cross-product approach, *Environment and Planning A* 23: 1269–1277.
- Gonzalez-Estrada, E.; Villasenor-Alva, J.A. 2020. goft: Tests of Fit for some Probability Distributions. Available from Internet: <<https://cran.r-project.org/web/packages/goft/goft.pdf>>.
- Gross, J.; Ligges, U. 2015. nortest: Tests for Normality. Available from Internet: <<https://cran.r-project.org/web/packages/nortest/nortest.pdf>>.
- Guidoum, A.C.; Boukhetala, K. 2020. Performing Parallel Monte Carlo and Moment Equations Methods for Itô and Stratonovich Stochastic Differential Systems: R Package Sim.DiffProc, *Journal of Statistical Software* 96(2): 1–82. <https://doi.org/10.18637/jss.v096.i02>.
- Hagget, P. (Ed.). 2001. *Geography: a global synthesis*. Pearson Education Limited, United Kingdom. 627 p.
- Hart, A.; Martínez, S. 2019. spgs: Statistical Patterns in Genomic Sequences. Available from Internet: <<https://cran.r-project.org/web/packages/spgs/spgs.pdf>>.
- Horiguchi, T.; Sakakibara, T. 1998. Numerical simulations for traffic-flow models on a decorated square lattice, *Physica A: Statistical Mechanics and its Applications* 252: 388–404. [https://doi.org/10.1016/S0378-4371\(97\)00628-6](https://doi.org/10.1016/S0378-4371(97)00628-6).
- Hummels, D.L. 1999. Toward a Geography of Trade Costs. Available from Internet: <<https://ssrn.com/abstract=160533>> or <<https://dx.doi.org/10.2139/ssrn.160533>>.
- Illés, I. 2008. *Regionális gazdaságtan – területfejlesztés (Regional Economics - Regional Development)*. Typotex, Hungary. 262 p.
- Jung, W.-S.; Wang, F.; Stanley, H.E. 2008. Gravity model in the Korean highway, *EPL (Europhysics Letters)* 81(4): 48005. <https://doi.org/10.1209/0295-5075/81/48005>.
- LeSage, J.P. 1998. *Spatial Econometrics*. Department of Economics University of Toledo, Spain. 273 p.
- Limão, N.; Venables, A.J. 2001. Infrastructure, Geographical Disadvantage, Transport Costs, and Trade, *The World Bank Economic Review* 15(3): 451–479.
- Mályusz, L.; Varga, A. 2018. An Estimation of the Learning Curve Effect on Project Duration with Monte Carlo Simulation, *Periodica Polytechnica Architecture* 49(1): 66–71. <https://doi.org/10.3311/PPar.12759>.
- Massey, F.J. 1951. The Kolmogorov-Smirnov Test for Goodness of Fit, *Journal of the American Statistical Association* 46(253): 68–78. <https://doi.org/10.1080/01621459.1951.10500769>.

- MATLAB, 2019. 9.7.0.1296695 (R2019b). The MathWorks Inc. USA.
- Millard, S.P. 2013. *EnvStats: An R Package for Environmental Statistics*. Springer, USA. 291 p.
- Miltiadou, M.; Bouhours, E.; Basbas, S.; Mintsis, G.; Taxiltaris, C. 2017. Analysis of border crossings in South East Europe and measures for their improvement, *Transportation Research Procedia* 25: 603–615. <https://doi.org/10.1016/j.trpro.2017.05.445>.
- Moran, P.A.P. 1948. Some Theorems on Time Series: II The Significance of the Serial Correlation Coefficient, *Biometrika* 35(3/4): 255–260. <https://doi.org/10.2307/2332344>.
- Moran, P.A.P. 1950. Notes on Continuous Stochastic Phenomena, *Biometrika* 37(1/2): 17–23. <https://doi.org/10.2307/2332142>.
- Nagy, L.; Balogh, P. 2013. *Ökonometria (Econometrics)*. Debreceni Egyetem Gazdálkodástudományok Centruma, Hungary. 160 p.
- Odlyzko, A. 2015. The forgotten discovery of gravity models and the inefficiency of early railway networks, *Oeconomica* 5(1): 157–192.
- Okubo, T. 2004. The border effect in the Japanese market: A Gravity Model analysis, *Journal of the Japanese and International Economies* 18(1): 1–11. [https://doi.org/10.1016/S0889-1583\(03\)00047-9](https://doi.org/10.1016/S0889-1583(03)00047-9).
- Opasanon, S.; Kitthamkesorn, S. 2016. Border crossing design in light of the ASEAN Economic Community: Simulation based approach, *Transport Policy* 48: 1–12. <https://doi.org/10.1016/j.tranpol.2016.02.009>.
- OpenStreetMap contributors. 2017. Planet dump retrieved from <<http://download.geofabrik.de/europe.html>>.
- Park, J.; Kwon, C.; Son, M. 2014. Economic implications of the Canada–U.S. border bridges: Applying a binational local economic model for international freight movements, *Research in Transportation Business & Management* 11: 123–133. <https://doi.org/10.1016/j.rtbm.2014.06.003>.
- Prato, C.G. 2009. Route choice modeling: past, present and future research directions, *Journal of Choice Modelling* 2(1): 65–100. [https://doi.org/10.1016/S1755-5345\(13\)70005-8](https://doi.org/10.1016/S1755-5345(13)70005-8).
- RCore Team. 2017. *A language and environment for statistical computing*. R Foundation for Statistical Computing. Austria.
- Rietveld, P.; Bruinsma, F.R.; van Vuuren, D.J. 2001. Spatial graduation of fuel taxes; consequences for cross-border and domestic fuelling, *Transportation Research Part A: Policy and Practice* 35(5): 433–457. [https://doi.org/10.1016/S0965-8564\(00\)00002-1](https://doi.org/10.1016/S0965-8564(00)00002-1).
- Sipos, T.; Szabó, Z.; Török, Á. 2021. Spatial Econometric Cross-Border Traffic Analysis for Passenger Cars – Hungarian Experience, *Promet - Traffic & Transportation* 33(2): 233–246. <https://doi.org/10.7307/ptt.v33i2.3641>.
- Szabó, Z.; Sipos, T. 2020. Térstatistika a közlekedésben (Spatial statistics in transport), *Műszaki Szemle* 75: 1–7.
- Szabó, Z.; Sipos, T.; Török, Á. 2017. Spatial Econometric Analysis of the Hungarian Border Crossings. In: *MATEC Web of Conferences* 134, 00057. <https://doi.org/10.1051/mateconf/201713400057>.
- Szabó, Z.; Török, Á. 2018a. Tranzitforgalmak Magyarországon: Egy térökonometriai elemzés (Transit traffic in Hungary: A spatial econometric analysis). In: *Proceedings of the VIII. Közlekedéstudományi konferencia*, 201–212.

- Szabó, Z.; Török, Á. 2018b. Magyarország határátkelőinek térökonometriai elemzése (Evaluating the Hungarian Border Crossings from Spatial Econometric Point of View), *Közlekedéstudományi Szemle* 68(4): 46–60. <https://doi.org/10.24228/KTSZ.2018.4.4>.
- Szalóki, F. 2017a. M15 expressway – upgrade and widening to a 2+2 traffic lanes motorway between M1 Motorway and Rajka (HU-SK border). In: *Proceedings of the XV. European Transport Congress and X. Budapest International Road Congress*, 12–17.
- Szalóki, F. 2017b. M70 expressway – upgrade and widening to a 2+2 traffic lanes motorway between Letenye and Tornyiszentmiklós at HU-SL border. In: *Proceedings of the XV. European Transport Congress and X. Budapest International Road Congress*, 18–24.
- Tagai, G.; Péntes, J.; Molnár, E. 2008. Methods of the analysis of integration effect on border areas – the case of Hungary, *Eurolimes – Journal of the Institute for Euroregional Studies* 6: 150–160.
- Tobler, W.R. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region, *Economic Geography* 46(sup1): 234–240. <https://doi.org/10.2307/143141>.
- Varga, A. 2002. Térökonometria (Spatial econometrics), *Statisztikai Szemle* 80(4): 354–370.
- Venables, W.N.; Ripley, B.D. 2002. *Modern Applied Statistics with S*. Springer, USA. 495 p.
- von Neumann, J.; Ulam, S. 1949. The Monte Carlo method, *Journal of the American Statistical Association* 44(247): 335–341.
- Yen, J.Y. 1976. Finding the K Shortest Loopless Paths in a Network, *Management Science* 17(11): 712–716.
- Zipf, G.K. 1949. *Human behavior and the principle of least effort*. Addison-Wesley Press, USA. 573 p.