AUTOMATIC MULTIREGIME FUNDAMENTAL DIAGRAM CALIBRATION USING LIKELIHOOD ESTIMATION

Saurav Barua[1], Nazmul Haque[2], Anik Das[3], Md Hadiuzzaman[4], Sanjana Hossain[5]

[1,3] *Department of Civil Engineering, University of Information Technology and Sciences, Dhaka, Bangladesh*
[2,4,5] *Department of Civil Engineering, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh*

Abstract: LIMB (Likelihood Identification for Multi-regime models' Breakpoint), a new tool, is developed to calibrate various Fundamental diagrams (FDs) under different road geometric and traffic operational conditions. This tool enables to estimate traffic state from real time data and is ready to implement into model based control strategy. Since efficient traffic management or control of transportation system remains big challenge due to the necessity of accurate traffic state estimation; model based control strategy incorporated with LIMB tool can ameliorate the complexity. This research found that the breakpoint of multi-regime FD models obtained from experience were not able to estimate traffic state precisely; therefore LIMB was used to calibrate those models. The investigation also endeavored to develop a guideline which was capable to calibrate suitable FD models for lane- wise traffic conditions. Our proposed technique is independent of speed limits and completely automatic without any threshold inputs. Furthermore, it is comparable with the well-recognized FD automatic calibration technique. The comparative study found a 5% to 8% variation in estimating FD parameters. Later, we investigated several novel single and multi- regime FD models utilizing field traffic data obtained from PeMS website. LIMB adopted likelihood estimation method to identify density at breakpoint in between free flow and congestion states for multi-regime models. It applies least square method to estimate critical density-free flow speed-capacity. The proposed interface is conducive and easily adaptable for transportation practitioners to select the best model based control strategy for smooth and efficient traffic operation.

Keywords: LIMB, FD structure, single regime model, multi-regime model, breakpoint.

1. Introduction

Fundamental Diagram (FD) describes the relationship among traffic parameters such as flow q, density k and speed v of an equilibrium traffic state. Several researchers conducted their investigation to formulate static relationship between flow-density-speed through theoretical and empirical modeling using field data. Among which, flow-density relation is concave, speed-density is monotonous decreasing and speed-flow is foliage shape branching into lower and upper limbs. FD is bounded to a specific location over a period of time because of road geometry and variations of traffic characteristics. Usually, those are developed to formulate intrinsic relationship of traffic stream within a range of observed data. Because of the deviation in assumptions and calculation techniques, even, similar models calibrated with same field data could

---

[1] Corresponding author: saurav.barua@uits.edu.bd

not produce results with same accuracy. Moreover, FD depends on geometric characteristics and operation of the site including number of lanes, existence of High Occupancy Vehicle (HOV) lane, position of off-ramp and on-ramp. Hence, calibration and parameter estimation are required to construct FD model for individual roadway environment separately—which is a tedious and laborious work. This study is focused on automatic calibration procedure by which one can select an appropriate FD structure and estimate its parameters. Our research developed an interface 'LIMB' which is capable to fit various FD structures and estimate corresponding traffic parameters using likelihood estimation method.

Control strategies such as—Variable Speed Limit (VSL) and Ramp Metering are implemented in highways to avoid flow breakdown. These strategies typically consider critical density (k_c) as the threshold value for their control parameter activation. k_c is the density at which traffic state is changed from free flow regime to congested regime. Critical density has been found in both single and multi- regime speed-density curve. In two regime models, k_c is the breakpoint in between free flow and congested traffic states. In contrast, single regime models do not have any breakpoint in speed-density curve. Control strategy designed on the basis of inaccurate k_c value, would activate control parameters either earlier or later (Lu *et al.*, 2009). If control strategy is applied into the highway system before critical density, traffic flow will reduce and the system will remain underutilized. However, the maximum benefit is achievable only through designing proper control strategy. Several macroscopic traffic modelling approaches use the concept of FD. LWR hydrodynamic model (Lighthill

and Whitham, 1995) is such a widely used model for traffic simulation and control. Cell Transmission Model (CTM) (Daganzo, 1994) and METANET Model (Messmer and Markos, 1990) utilize the parameters of FD. Model based control strategies, which are implemented for prediction purpose, also use FD structure. Besides, there is a widely known debate over the performance history of VSL. The performance of VSL depends on the suitability of FD structure. Hence, the selection of FD structure and accurate estimation of its parameter emerges equal importance. We investigated FD structure and its parameters under different traffic operating conditions and road geometry utilizing LIMB tool. This automatic calibration tool uses likelihood estimation method to search different FD parameters for single HOV lane and multilane roads.

The pivotal research on automatic calibration of FD was performed by (Dervisoglu *et al.* 2009). For the studied roadway section, they considered speed limit of 55 mph to draw left limb i.e. free flow portion of flow-density plot. However, in case of roads where speed limits are not posted properly—especially in developing countries, the implementation of the aforementioned calibration technique is difficult. The LIMB-based proposed automatic calibration technique is not speed limit dependent and hence it can be implemented for roadways of developed as well as developing countries.

The paper is structured as follows: Section 2 reviews previous works on FD models and their calibration procedure; Section 3 presents the adopted likelihood estimation method for FD calibration; Section 4 introduces the developed LIMB interface; Section 5 describes the procedure of traffic data extraction from Caltrans Performance

Measurement System (PeMS); Validation of the likelihood estimation method is presented in section 6; Section 7 applies LIMB for calibrating FD models using PeMS data; Finally, concluding remarks and future research scopes are given in Section 8.

2. Literature Review

The earlier researches on FD covered variety of mathematical forms, assuming single regime phenomenon over full range of flow condition. The seminal research was conducted by (Greenshields *et al.,* 1934). He derived speed as a linear function of density based on aerial photographic study. (Greenberg, 1959) proposed logarithmic speed-density function and mathematically bridges the gap between macroscopic models with microscopic car-following model. Underwood model (Underwood, 1961) gained popularity by exponential model and attempted to overcome the limitation of the Greenberg model. Further researches on single regime models were directed toward the generalized modeling approach by introducing new parameter in the formula. Drew proposed additional exponent parameter 'm' on Greenshields model. The Pipes-Munjal model (Pipes, 1967) was the same as Greenshields model by varying 'm' for a family of traffic models. However, single regime models were unable to estimate FD parameters properly when the field data were near capacity. Some field observations showed discontinuity in the flow-speed-density relationship. These led to develop multi-regime models for FD. Free flow and congested flow conditions were separated in two regime models attempting to distinguish traffic characteristics. (Edie, 1961) proposed two regime models considering Underwood and Greenberg model for the free flow and congested regime, respectively. In another

study (Jayakrishnan *et al.,* 1995), a modified Greenberg model was proposed considering a constant speed for the free flow regime. North-Western University research group (Drew, 1968) focused on Two Regime Linear model that adopted Greenshields model for the free flow and congested regime separately. Three Regime Linear model (Drake et al., 1967) was also developed to distinguish transitional flow from free flow and congested flow regime.

Generalized polynomial model proposed non-integer power for FD. (Zhang, 1999) described one parameter polynomial. (Hegyi *et al.,* 2002) used exponential in polynomial model, which generalizes Underwood model. BPR model (Skabardonis and Dowling, 1997) and Van Aerde model (Aerde, 1995) are type of FD models those are used for planning purpose. In another study, (Lu and Skabardonis, 2007) combined several existing models using Taylor expansion. They proposed variable FD structures with two limbs; especially generalized polynomial model with fractional coefficient for right limb.

The difficulty in calibrating multi-regime FD models is to estimate the breakpoint kc between regimes. Quandt's likelihood estimation technique (1958) was adopted by Northwestern researchers to identify breakpoints between regimes for freeway traffic and performed regression analysis to select the best models (May, 1990). It is generally accepted that the breakpoint or critical density k_c varies with the road geometry and traffic characteristics. Thus, breakpoint obtained for each road segment does not match with others. It may still varies based on traffic composition and traffic pattern over time. Accuracy and validity of the models depend on breakpoint significantly.

Hence, the positioning of breakpoint for multi-regime model is very crucial.

Traffic dynamics cannot be captured with limited number of measurement device. Nonetheless, the estimation of FD for each Vehicle Detector Station (VDS) is infeasible. Consequently, automatic calibration method for FD emerges along with the development of the macroscopic traffic flow modeling. (Dervisoglu *et al.,* 2009) proposed an automated technique to calibrate FD for freeways considering capacity drop. They considered piece-wise linear FD for obtaining required traffic parameters in order to predict traffic state using Cell Transmission Model (CTM). (Zhong *et al.,* 2016) proposed optimization based automatic calibration of FD. (Li, 2014) developed a tool to identify traffic state automatically and calibrate FD using historical data. (Knoop and Daamen, 2016) characterized FD by five parameters such as free flow speed, wave speed, free flow capacity, queue discharge rate and jam density. They utilized least square method for calibration of FD automatically. (Pompigna and Rupi, 2015) calibrated FD, adopting Van Aerde Model and Longitudinal Control Model for freeways. They tested suitability of level of service assessment as per HCM standard for freeways in Italy.

Finding appropriate FD structure under existing road condition is foremost required to estimate FD parameters accurately. In this regard, the developed tool- LIMB could play a vital role. LIMB toiled to develop a probabilistic approach using likelihood estimation method to calibrate various single and multi-regime FD models automatically. Initially, it generates multiple breakpoints related to each density value in the dataset.

After that, separate regression lines are established for the breakpoints to determine corresponding FD parameters. However, the optimal breakpoint and FD parameters were obtained by maximizing the likelihood function. The traffic data collected from PeMS were utilized to formulate and calibrate FD models to demonstrate the novelty of the proposed calibration technique. This research work could act as a guideline for transportation researchers and practitioners on modeling FD structures and optimizing traffic parameters.

3. Methodology

Speed-density (*v-k*) curve is the basis of many traffic models. Such curve was assumed linear, logarithmic and exponential in (Greenshields *et al.,* 1934), (Greenberg, 1959) and (Underwood, 1961) model, respectively. Since, Greenberg model is not suitable for sparse traffic, (Edie, 1961) considered two regime of which left and right limbs are logarithmic and exponential, respectively. In Two Regime Linear model, left and right limbs are linear with breakpoint at critical density k_c. LIMB incorporates the above models including 2nd order and 3rd order polynomials to calibrate FD. Furthermore, the tool searches breakpoint in the two regime FD using likelihood function as in (Quandt, 1958).

Let, v_i and k_i are the speed and density of traffic at *i*-th state respectively. The relationship in *v-k* is $v_i = f(k_i)$. The error ε_i is the deviation of fitted speed-density curve with the observed data at *i*-th state. The $\varepsilon_i = v_i - f(k_i)$. Assume, N is total number of observed data points. The least square fit uses Eq. (1):

$$Minimize \sum_{i=1}^{N} \varepsilon_i = Minimize \sum_{i=1}^{N} v_i - f(k_i) \qquad (1)$$

Speed-density relation can be generalized as regression model or fitted model as $v_i = f(k_i) + \varepsilon_i$ for single regime model. Moreover, for linear model, $f(k_i) = ak_i + c$; for logarithmic model $f(k) = a \ln k_i$, and for exponential model $f(k_i) = a \exp^{k}{}_i$. Polynomial models provide higher order of v-k relationship.

Whereas, for two regime models, the above equation could be split into two separate equations with one breakpoint. Those two equations can be written as follows Eq. (2) and Eq. (3):

$$v_i = f_1(k_i) + \varepsilon_1 \qquad (2)$$

$$v_i = f_2(k_i) + \varepsilon_2 \qquad (3)$$

ε_1 and ε_2 are normally and independently distributer error terms. Mean of errors are zero and standard deviations are σ_1 and σ_2 respectively. Since, total number of observations = N, Eq. (2) generates first n observations and Eq. (3) generates N-n observations.

Densities of ε_1 at i point, Eq. (4),

$$D_1 = \frac{1}{\sqrt{2\pi} \cdot \sigma_1} \exp\left[-\left(\frac{1}{2 \cdot \sigma_1^2}\right) \cdot (v_i - f_1(k))\right] \qquad (4)$$

Densities of ε_2 at i point, Eq. (5),

$$D_2 = \frac{1}{\sqrt{2\pi} \cdot \sigma_2} \exp\left[-\left(\frac{1}{2 \cdot \sigma_2^2}\right) \cdot (v_i - f_2(k))\right] \qquad (5)$$

Therefore, the likelihood of sample of n observations from Eq. (2):

$$L_1 = \left(\frac{1}{\sqrt{2\pi} \cdot \sigma_1}\right)^n \exp\left[-\left(\frac{1}{2 \cdot \sigma_1^2}\right) \cdot \sum_{i=1}^{n} (v_i - f_1(k))^2\right] \qquad (6)$$

And, likelihood of sample of N-n observations from Eq. (3):

$$L_2 = \left(\frac{1}{\sqrt{2\pi} \cdot \sigma_2}\right)^{N-n} \exp\left[-\left(\frac{1}{2 \cdot \sigma_2^2}\right) \cdot \sum_{i=n+1}^{N} (v_i - f_2(k))^2\right] \qquad (7)$$

The likelihood of entire sample is $l = L_1 \ x \ L_2$, Eq. (8):

$$l = \left(\frac{1}{\sqrt{2\pi} \cdot \sigma_1}\right)^n \left(\frac{1}{\sqrt{2\pi} \cdot \sigma_2}\right)^{N-n} \exp\left[-\left(\frac{1}{2 \cdot \sigma_1^2}\right) \cdot \sum_{i=1}^{n}(v_i - f_1(k))^2 - \left(\frac{1}{2 \cdot \sigma_2^2}\right) \cdot \sum_{i=n+1}^{N}(v_i - f_2(k))^2\right] \qquad (8)$$

Taking logarithm of likelihood function Eq. (9):

$$Log\, l = -N \cdot \log\sqrt{2\pi} - n \cdot \log\sigma_1 - (N-n)\cdot \log\sigma_2 -$$
$$- \frac{1}{2 \cdot \sigma_1^2} \sum_{i=1}^{n}(v_i - f_1(k))^2 - \frac{1}{2 \cdot \sigma_2^2}\sum_{i=n+1}^{N}(v_i - f_2(k))^2 \qquad (9)$$

Assuming, $L = \log l$, Eq. (10),

$$L = -N \cdot \log\sqrt{2\pi} - n \cdot \log\sigma_1 - (N-n)\cdot \log\sigma_2 -$$
$$- \frac{1}{2 \cdot \sigma_1^2} \sum_{i=1}^{n}(v_i - f_1(k))^2 - \frac{1}{2 \cdot \sigma_2^2}\sum_{i=n+1}^{N}(v_i - f_2(k))^2 \qquad (10)$$

Setting partial derivatives of (7) equal to zero and by substituting, we obtain, Eq. (11),

$$L(n) = -N \cdot \log\sqrt{2\pi} - n \cdot \log\sigma_1 - (N-n)\cdot \log\sigma_2 - \frac{N}{2} \qquad (11)$$

Eq. (11) provides the logarithm of maximum likelihood value for a given observation of N and it is a function of n only. The objective is to find the value of n at which Eq. (11) gives maximum value. LIMB iterates the value on n from 2 to N-1 for computing likelihood value using Eq. (11) and selects the maximum likelihood estimate for a particular n.

The breakpoint corresponding to critical density in the speed-density diagram was estimated by an iterative process using likelihood estimation method. The procedure was summarized as—firstly, the speed-density dataset (v_1,k_1), (v_2,k_2)... (v_N,k_N) were sorted as per traffic density value in ascending order using bubble sort algorithm. The initial breakpoint was assumed in the left-most portion of the speed-density plot. The dataset were divided into left limb and right limb considering breakpoint in the above mentioned plot. Separate regression lines were fitted for the two limbs using the least square method. There were first n number data points in the left limb and the rest of (N-n) number data points in the right limb. The variance of speed σ_1 and σ_2 were calculated for the both limbs. After that, the likelihood function $L(n)$ corresponding to the breakpoint was calculated using the Eq. (11). For the next iteration, the break point between two limbs was moved by one data point to the right in the speed-density plot. Then, separate regression lines for left and right limbs were estimated. LIMB evaluated the likelihood value corresponding to the new breakpoint. The point was moved further rightward. This iteration applied on the entire dataset to find the highest maxima in the likelihood-density plot. Corresponding traffic density is the critical density k_c. For the estimated k_c, other FD parameters such as capacity q_{max}, jam density k_j and free flow speed v_f were computed.

4. Developing Limb Tool

LIMB is an application tool developed using MATLAB GUI. The application is capable to work independently under .exe format as well. Fig. 1 shows the interface of the developed tool. In the middle portion of the interface, there is a 'Load Data File' button.

It is provided for uploading traffic data from MS Excel spread sheet. Column number with range for corresponding speed, flow and density value are required to mention here. In addition to the 'Load Data File' button, there is a radio button for user to choose various FD models. Several FD models can be chosen from the two groups of models— Single Regime Models (Greenshields, Greenberg, Underwood, Generalized Polynomial, 2nd Order Polynomial, 3rd Order Polynomial) and Multi-Regime Models (Edie and Two Regime Linear). Selecting the checkbox under Multi-Regime Model, user can estimate the breakpoint by likelihood estimation method. If the checkbox remains uncheck, LIMB uses default breakpoint value given in (20). After uploading the aggregated traffic data (measured speed v, flow q and density $k= q/v$) and selecting FD model, the detailed analysis starts automatically. Using v-k relationship LIMB generates corresponding fitted FD model. It provides estimated speed v' corresponding to each density k from the model. The estimated flow q' is computed using, $q' = v'k$ relationship. Later, Speed-density, flow-density and speed-flow scattered plot with fitted model are generated in the left most portion of the interface. R^2 and sum of square error are computed by LIMB for each selected FD model. In the right portion, tabulated value for measured flow q, density k, speed v are presented along with estimated flow q' and speed v' value. LIMB also generates estimated Likelihood vs density plot for Edie model and Two Regime Linear model as shown in Fig. 2. Edie model and Two Regime Linear Model utilize separate .m extension file to perform iteration for finding maximum likelihood value. The least square estimation method is encrypted in the PlotFD.m file to fit different FD structures and graphical interface is feed backed by MATLAB GUI.
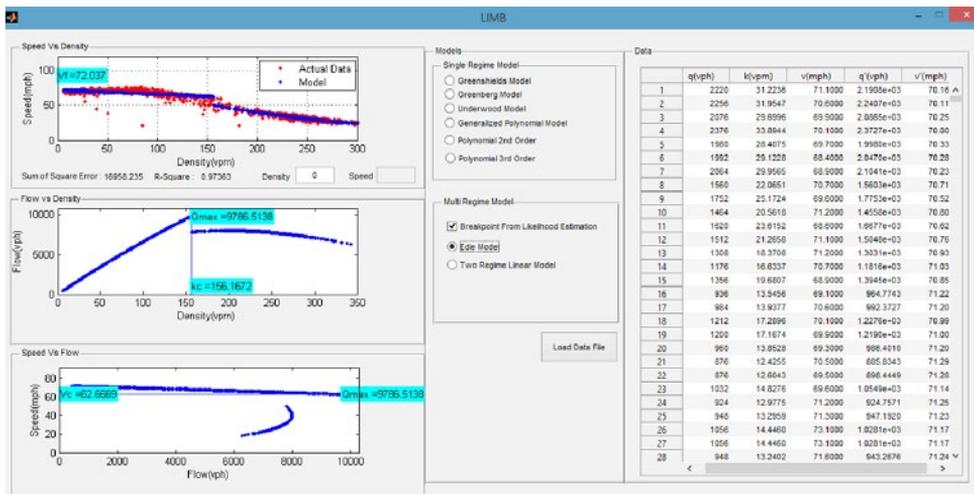
Fig. 1.
LIMB Interface Showing Analysis Results and Graphical Output for the Traffic Data Obtained at Mainline VDS 769346 - WHITSETT AVE on March 17 to March 23, 2016
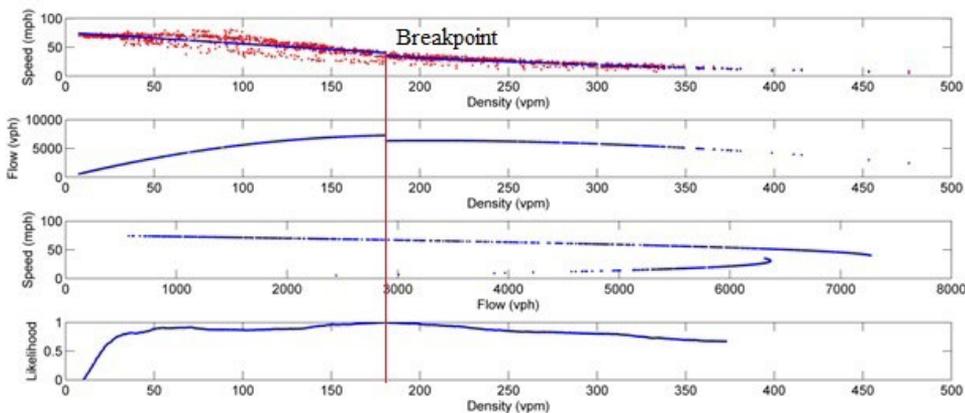
Fig. 2.
Maximum Likelihood Estimation Method; the Vertical Line Corresponding to Maxima of Likelihood vs. Density Plot Indicates the Breakpoint and Critical Density for Edie Model

5. Traffic Data

To check the accuracy of the proposed likelihood estimation based automatic calibration technique and to find the FD parameters for different traffic operations and road geometric conditions, traffic data was extracted from Caltrans Performance Measurement System (PeMS). It works as repository for real time traffic data of the California state, USA. PeMS allows the user to extract speed and flow data across

vehicle detector stations (VDS) on a time series basis. The traffic data are obtainable in different formats: PDF, .doc and .xlsx. For comparison and case studies, the data were aggregated for 5 minutes intervals following the optimal aggregation period proposed by Dervisoglu et al. (2009). The data were accumulated over 7 days from March 17 to March 23, 2016. Thus, for a 5 minutes aggregation, 7 days data provides 2016 data points for each VDS. Noted that, the coefficient of determination (R^2) value for speed-density relationship computed by LIMB considering 1-day data provides comparatively higher value than that of 7-day data. This is due to the fact that 7-day data is relatively scattered with adequate sampling. The value of R^2 decreases slightly for 365 days data; therefore, the 7-day traffic data was taken to avoid dealing with unnecessary huge dataset.

6. Accuracy of Likelihood Estimation Method

We have performed a comparative study to check the accuracy of the likelihood estimation based calibration technique with the well-established automatic calibration technique of (Dervisoglu *et al.,* 2009). We have featured same 7-day traffic data of Mainline VDS 769346 - WHITSETT AVE and analyzed the dataset for both of the techniques. The calibration performed by (Dervisoglu *et al.,* 2009) consists of three steps. Firstly, vehicle speed above 70 mph (corresponds to speed limit) are considered to perform the least square fit on flow-density data and estimate free flow speed v_f. The next step is to estimate capacity which is a deterministic and rather conservative method. Specifically, following (Dervisoglu *et al.,* 2009) capacity is calibrated by

averaging maximum flow over several ideal days where no significant incidents happened under good weather and road condition. This value is then horizontally projected to the free-flow line to establish the tip of the triangular FD. The intersection is defined as the critical density k_c for the section, above which the flow is congested. Lastly, congestion speed w is calibrated by quartile regression performed for the data points having $k > k_c$. The line corresponding to w also defines the jam density for the studied section. Ten ascending and successive data values are binned together to facilitate the quartile regression.

The likelihood estimation method accompanied by least square regression is utilized through LIMB tool on the same dataset. The FD parameters—Critical density k_c, capacity qmax and free flow speed v_f estimated by the proposed technique are 156.2 veh/mile, 9773.3 veh/hr and 72.1 mile/hr respectively. These parameters values were obtained by fitting Edie model to the PeMS data. Whereas, by applying the technique of (Dervisoglu *et al.,* 2009), the parameters were found to be 144.5 veh/mile, 10344 veh/hr and 71.6 mile/hr respectively. The comparison of these values shows a variation of 8.09%, 5.52% and 0.70% respectively. These small variations indicate that the proposed automatic calibration technique is quite comparable with that of (Dervisoglu *et al.,* 2009). Therefore, the proposed technique is verified with respect to the well-recognized automatic calibration technique as in (Dervisoglu *et al.,* 2009).

The establishment of the left limb of FD by (Dervisoglu *et al.,* 2009) depends on speed limit and requires data segregation at critical density. However, in developing

countries, often speed limits are not properly posted on highways; as such, the motorists drive vehicles as per their own judgment and convenience. This in turn makes the data segregation based on speed limit quite difficult. To this end, our proposed technique is global and not dependent on speed limit. Although the technique is not threshold based, it works well to calibrate FD for the roads where the corresponding speed limit is unknown. The probabilistic approach itself can estimate free flow speed and locate the breakpoint automatically at critical density by an iterative calculation and maximizing the likelihood function.

7. Calibrating FD for Different Road Conditions

The calibration was performed over road segments with different number of lanes, i.e. single lane, dual lane, three lane, four lane and five lane. For each case, 3 different VDS was chosen and overall 15 VDS dataset were considered to collect unbiased information. The flows and speed are extracted from PeMS and densities are calculated as— Density = flow/speed. The extracted data in .xlsx format are directly readable after uploaded the worksheet in LIMB interface via 'Load Data File' button as mentioned in the prior section 4. User can fit the observed dataset with the desired FD structures and can obtain R^2 and sum of squared errors (SSE). LIMB also provides estimated flow q' and speed v'.

The study found that two regime models work better compare to single regime

models. Considering default breakpoint in Edie model and Two Regime Linear model as presented in (May, 1990), they perform poorly. Since critical density varies over traffic operating conditions, incorporating the likelihood estimation method improve the estimation capability of Edie model and Two Regime Linear model significantly. For multilane roads, Edie Model and Two Regime Linear Model with optimized breakpoint perform better.

Our investigation based on LIMB showed that Edie model and Two Regime Linear model without optimization produce negative R^2 value for speed-density relationship in case of multilane freeway, which is infeasible. The R^2 value calculated for single HOV lane is much lower compared to that of multilane freeway. In case of multilane freeway, effects of individual vehicle type are minimized resulting in less scattered data points and higher R^2 value. The table 1 shows R^2 value of various FD models obtained under different scenarios according to our investigation.

In some cases, 2nd and 3rd order polynomial models fit speed-density data with better precision. Specifically, due to the data fitting flexibility, compared to the 2nd order model, 3rd order model performs better. However, these models may fail to find jam density. Thus, these single regime models may not always comply with the traffic engineering concepts. Root mean square error (RMSE) of estimated flow and Mean percent error (MPE) estimated speed are presented in the table 2 and 3 respectively.

Table 1

R^2 of Calibrated FD Models Under Different Scenarios

| Freeway	VDS	R^2									
		GS	GB	UW	GPM	P2O	P3O	ED	2RLM	ED_bp	2RLM_bp
Single Lane	772457	0.96	0.66	0.94	0.96	0.96	0.97	0.63	0.74	0.97	0.96
	775214	0.94	0.61	0.92	0.94	0.94	0.96	0.78	0.88	0.93	0.95
	771917	0.93	0.82	0.95	0.94	0.95	0.95	0.71	0.88	0.96	0.93
Dual Lane	314159	0.94	0.61	0.91	0.94	0.94	0.95	0.09	0.07	0.97	0.97
	716938	0.63	0.56	0.61	0.65	0.65	0.65	-0.62	-0.53	0.66	0.66
	766673	0.93	0.72	0.91	0.93	0.94	0.96	0.48	0.52	0.97	0.97
Three Lane	716797	0.89	0.74	0.87	0.89	0.89	0.92	-2.05	-2.22	0.92	0.92
	760236	0.94	0.76	0.89	0.96	0.96	0.98	-1.64	-1.86	0.99	0.98
	760196	0.92	0.69	0.85	0.95	0.94	0.98	-2.81	-3.00	0.98	0.97
Four Lane	717433	0.90	0.69	0.85	0.93	0.93	0.95	-4.42	-5.26	0.95	0.95
	805627	0.94	0.72	0.90	0.95	0.95	0.97	-2.44	-3.49	0.97	0.96
	763434	0.90	0.68	0.84	0.93	0.93	0.95	-2.47	-3.65	0.95	0.95
Five Lane	401698	0.84	0.71	0.82	0.84	0.84	0.86	-2.34	-3.10	0.85	0.85
	772501	0.89	0.77	0.89	0.90	0.90	0.91	-2.39	-4.04	0.91	0.91
	769346	0.93	0.72	0.89	0.96	0.95	0.97	-3.65	-5.60	0.97	0.97

N.B.: GS stands for Greenshields model, GB stands for Greenberg model, UW stands for Underwood model, GPM stands for Generalized polynomial model, P2O stands for Polynomial 2nd order model, P3O stands for Polynomial 3rd order model, ED stands for Edie model, 2RLM stands for Two Regime Linear Model, ED_bp stands for Edie model with breakpoint optimization and 2RLM_bp stands for Two Regime Linear Model with breakpoint optimization.

Table 2

RMSE of Estimated Traffic Flow

Freeway	Road/ VDS	%RMSE of Traffic Flow									
		GS	GB	UW	GPM	P2O	P3O	ED	2RLM	ED_bp	2RLM_bp
Single Lane	772457	3.14	30.83	3.90	2.29	1.44	0.88	0.88	9.45	1.41	3.33
	775214	4.55	26.23	4.14	5.13	3.74	2.50	4.35	3.42	2.70	5.00
	771917	2.57	4.57	0.89	1.28	0.79	0.84	2.62	1.35	1.35	2.43
Dual Lane	314159	13.98	48.01	7.03	13.41	6.35	5.68	38.90	80.60	2.78	3.90
	716938	4.80	6.03	5.30	5.30	4.67	4.68	16.34	15.51	4.24	4.36
	766673	3.73	12.83	2.39	3.48	2.12	1.97	19.49	23.10	1.07	0.90
Three Lane	716797	1.26	2.66	1.27	1.41	1.31	1.01	44.92	51.29	0.89	0.94
	760236	0.90	4.17	1.83	1.14	1.27	0.32	60.08	73.30	0.33	0.43
	760196	1.03	5.63	2.44	1.23	1.40	0.46	53.90	63.40	0.36	0.72
Four Lane	717433	2.08	4.93	2.76	2.14	2.25	1.74	76.21	99.60	1.63	1.83
	805627	1.22	7.26	2.22	1.51	1.58	0.94	95.49	183.06	0.62	0.68
	763434	3.83	12.09	6.91	3.88	4.20	2.73	108.87	195.83	2.69	3.03
Five Lane	401698	4.01	4.92	3.75	4.20	4.08	3.95	82.50	121.57	3.80	3.79
	772501	5.63	9.25	4.22	4.51	3.52	3.81	146.41	347.83	3.21	3.36
	769346	1.10	4.77	1.87	1.35	1.52	0.85	113.20	198.50	0.72	0.76

Table 3

MPE of Estimated Traffic Speed

Freeway	Road/VDS	MPE of Speed									
		GS	GB	UW	GPM	P2O	P3O	ED	2RLM	ED_bp	2RLM_bp
Single Lane	772457	-0.59	15.02	5.19	0.01	0.59	0.61	0.61	-23.61	1.84	-0.76
	775214	0.97	17.45	5.64	0.80	1.18	1.24	-6.61	-2.71	3.39	0.77
	771917	-0.03	4.40	1.31	0.44	0.67	0.66	-8.15	-4.87	-4.87	0.04
Dual Lane	314159	-0.01	19.05	6.75	0.07	0.90	1.29	-45.52	-51.69	0.29	0.35
	716938	3.95	5.19	4.76	4.76	3.59	3.59	-34.12	-33.36	3.31	3.33
	766673	0.63	11.85	4.99	0.73	1.13	0.66	-37.34	-37.62	-0.14	0.44
Three Lane	716797	0.92	3.09	1.83	0.80	0.88	0.66	-58.89	-60.86	0.66	0.67
	760236	1.19	5.44	3.84	0.50	0.53	0.23	-71.33	-76.71	0.15	0.24
	760196	1.40	5.61	3.86	0.46	0.46	0.33	-67.88	-71.49	0.23	0.18
Four Lane	717433	1.57	4.40	2.91	0.93	0.93	0.90	-80.03	-88.39	0.90	0.79
	805627	1.11	7.53	4.15	0.62	0.71	0.55	-83.68	-104.30	0.42	0.54
	763434	3.42	9.74	7.22	1.76	1.67	1.87	-91.27	-113.99	1.88	1.77
Five Lane	401698	2.45	5.32	3.30	2.48	2.45	2.64	-81.02	-93.54	2.44	2.44
	772501	1.28	8.55	5.26	1.77	2.20	2.23	-105.29	-141.96	2.39	2.33
	769346	1.29	5.56	3.24	0.52	0.52	0.49	-95.60	-119.57	0.47	0.48
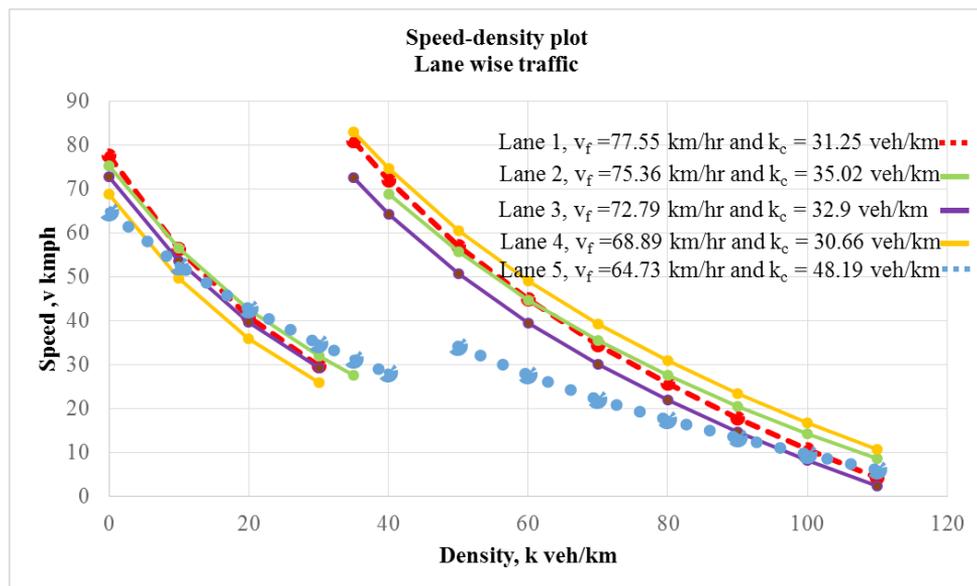
Fig. 3.

Speed vs Density Plot of Traffic for Various Lane of a Multilane Freeway, Detector ID, VDS-769346

Investigation was also done to find lane-wise FD models. The speed-density relationship obtained by Edie Model with breakpoint optimization for different lanes of a 5-lane freeway (WHITSETT AVE, VDS 769346) are presented in Fig. 3. It revealed that traffic

characteristics in the right-most lane are not similar to that of the left-most lane or middle lanes. Practically, the left-most lane has less friction compared to the right-most lane. Since, traffic in the right-most lane comprises of slower vehicles near shoulder; whereas, that in the left-most lane comprises of faster vehicle near median. These results were adequately captured by the developed LIMB tool, which shows the accuracy of likelihood estimation based calibration technique of multi-regime FD models.

8. Concluding Remarks

This paper develops a tool-LIMB to identify the breakpoint of two regime FD (speed-density curve). Here, the congestion regime is separated from the free flow regime at the breakpoint. Since this breakpoint varies for different traffic scenarios, identifying its position is crucial to calibrate FD. Specifically, LIMB adopted a probabilistic method which is capable of automatically finding the position of the breakpoint accurately. Utilizing the speed-density plot, the likelihood corresponding to each traffic state was computed iteratively and the maximum of likelihood function value was estimated. This value corresponds to the breakpoint of FD. After identifying the breakpoint, LIMB used the least square method to fit the right and left limbs with different types of FD structures.

The proposed calibration technique was checked with the well-recognized technique of (Dervisoglu *et al.,* 2009). The parameters of FD such as—critical density (abscissa of the breakpoint), free flow speed and maximum roadway capacity obtained by the likelihood estimation method varies 8.09%, 0.7% and 5.52% respectively corresponding to the aforementioned automatic calibration technique. Thus, the proposed technique is verified with respect to the well-recognized technique of (Dervisoglu *et al.,* 2009). Moreover, the proposed technique is independent of threshold value i.e. speed limit. This indicates the wider application of the proposed technique in developing countries where either the speed limit is not posted adequately or drivers are reluctant to follow it.

LIMB has a user friendly interface which provides flexibility to calibrate FD model. Aggregated speed-density dataset readable in .xlsx format can be uploaded and analyzed. User can choose various single regime (Greenshields, Greenberg, Underwood, Generalized Polynomial, 2nd Order Polynomial and 3rd Order Polynomial) models and multi-regime (Edie and Two Regime Linear model) models to fit with the observed data. The tool is capable to generate corresponding fitted FD plots for the models and estimates various FD parameters—critical density, free flow speed and maximum capacity. The R^2 value for the speed-density plot is directly computed and user can understand the fitness of the models. Besides, the estimated speed and estimated flow data for each traffic state are presented in the tabular form so that they can convey further in-depth statistical analysis. User can easily export these data to MS Excel format. They can evaluate various FD models corresponding to measured traffic data, estimates FD parameters and conduct comparative study to find the best fitted models for calibration.

PeMS data are used as an input to test the applicability of LIMB tool. The study covers to calibrate FD from single HOV lane to various multilane freeways. The study shows that, two regime models perform

better compared to single regime models. Edie and Two Regime Linear models with breakpoint optimization have better R^2 value and fit better with the measured data in most of the scenarios. However, Edie and Two Regime Linear models without optimizing breakpoint work poorly, which considers default breakpoint as presented in (May, 1990). Critical density varies over road geometry and traffic operating condition. Incorporating the likelihood estimation method to compute critical density improve the estimation capability of Edie model and Two Regime Linear model significantly. Though 2nd order Polynomial and 3rd order Polynomial perform well, those model may not comply with traffic engineering concept. Specifically, these models may fail to find jam density. Statistical analysis show that the single regime models are less capable to estimate the FD model for multilane freeways. Besides, the study tried to model FD for individual lane-wise traffic in a multilane freeway utilizing Edie model with breakpoint optimization. It shows that the left-most lane has higher free flow speed and lower critical density compared to the right-most lane. These results match with the real world traffic operation quite remarkably showing the novelty of LIMB tool.

The proposed automatic FD calibration is adaptive and robust with the variability of traffic data. The likelihood estimation method is capable to capture traffic dynamics in real time, as well as, predict abnormal traffic condition, such as, incident scenario. Our automatic FD calibration can be further implemented when traffic surveillance systems are incapable to collect comprehensive traffic data, roads are equipped with sparse number of detectors, and large number of detectors are out-of-order due to aging. The proposed calibration technique is global

and independent of speed limit. This tool is capable to calibrate critical density, free flow speed and capacity from real time traffic data. Theses traffic parameters are key throughputs for many macroscopic models. The tool can work with online database and fit with real time data as well. However, the models used in the tool are deterministic. This research could be extended to adjust stochastic models. Spatial and temporal probabilistic distribution of capacity and congestion traffic state can be investigated for several incidents and road conditions.

Acknowledgements

References

Aerde, V.M. 1995. Single regime speed-flow-density relationship for congested and uncongested highways. In *Proceedings of the Transportation Research Board 74th Annual Meeting*, 95080.

Daganzo, C.F. 1994. The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory, *Transportation Research Part B: Methodological* 28(4): 269-287.

Dervisoglu, G.; Gomes, G.; Kwon, J.; Horowitz, R.; Varaiya, P. 2009. Automatic calibration of the fundamental diagram and empirical observations on capacity. In *Proceedings of the Transportation Research Board 88th Annual Meeting,* Vol 15.

Drake, J.S.; Schofer, J.L.; May, A.D. 1967. A statistical analysis of speed–density hypotheses, *Highway Research Record* 154:112-117.

Drew, D.R. 1968. *Traffic Flow Theory and Control*. McGraw-Hill Inc., New York. USA. 467 p.

Edie, L.C. 1961. Car-following and steady-state theory for non-congested traffic, *Operations Research* 9(1): 66-76.

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

Greenshields, B.D.; Bibbins, J.R.; Channing, W.S.; Miller, H.H. 1934. A study of traffic capacity. In *Proceedings of the Fourteenth Annual Meeting of the Highway Research Board* 14(1): 448-477.

Hegyi, A.; Schutter, A. B.D.; Hellendoorn, H.; Boom, T.V.D. 2002. Optimal coordination of ramp metering and variable speed control-an MPC approach. In *Proceedings of the American Control Conference* 5: 3600-3605.

Jayakrishnan, R.; Wei K.T; Chen, A. 1995. A dynamic traffic assignment model with traffic flow relationships, *Transportation Research Part C: Emerging Technologies* 3(1): 51-72.

Knoop, V.L.; Daamen, W. 2016. Automatic fitting procedure for the fundamental diagram, *Transportmetrica B: Transport Dynamics*, 1-16.

Li, H. 2014. Automatically Generating Empirical Speed-flow Traffic Parameters from Archived Sensor Data. In *Proceedings of the 9th International Conference on Traffic and Transportation Studies (ICTTS)*, 138: 54–66.

Lighthill, M.J.; Whitham, G.B. 1955. On kinematic waves. II. A theory of traffic flow on long crowded roads. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences,* 229(1178): 317-345.

Lu, X.; Varaiya, P.; Horowitz, R. 2009. Fundamental diagram modelling and analysis based NGSIM data. In *Proceedings of the IFAC 12th Symposium on Control in Transportation Systems*, 308-315.

Lu, X.T.; Skabardonis, A. 2007. Freeway traffic shockwave analysis: exploring the NGSIM trajectory data. In *Proceedings of the Transportation Research Board 86th Annual Meeting*, 07-3016.

May, A.D. 1990. *Traffic Flow Fundamentals*. Pearson Education. New York. USA. 301 p.

Messmer, A.; Markos, P. 1990. METANET: A macroscopic simulation program for motorway networks, *Traffic Engineering & Control* 31(8): 466-470.

Pipes, L.A. 1967. Car-following models and the fundamental diagram of road traffic, *Transportation Research* 1(1): 21-29.

Pompigna, A.; Rupi, F. 2015. Differences between HCM Procedures and Fundamental Diagram Calibration for Operational LOS Assessment on Italian Freeways, *Transportation Research Procedia* 5(2015): 103-118.

Quandt, R.E. 1958. The estimation of the parameters of a linear regression system obeying two separate regimes, *Journal of the American Statistical Association* 53(284): 873-880.

Skabardonis, A.; Dowling, R. 1997. Improved speed flow relationships for planning applications, *Transportation Research Record* 1572(1997): 18-23.

Underwood, R.T. 1961. *Speed, Volume, and Density Relationship: Quality and Theory of Traffic Flow*, Yale Bureau Highway Traffic, New Haven, Connecticut. 47 p.

Zhang, H.M. 1999. A mathematical theory of traffic hysteresis, *Transportation Research Part B: Methodological* 33(1):1-23.

Zhong, R.; Chen, C.; Chow, A.H.F.; Pan, T.; Yuan, F.; He., Z. 2016. Automatic calibration of fundamental diagram for first-order macroscopic freeway traffic models, *Journal of Advanced Transportation* 50(3): 363-385.