

A TRAFFIC FUNDAMENTAL DIAGRAM CALIBRATING METHODOLOGY TO AVOID UNBALANCED SPEED–DENSITY OBSERVATIONS

Chunbo ZHANG^{1*}, Zhaoguo HUANG², Yonggang WANG^{3#}

¹*School of Traffic and Transportation, Shijiazhuang Tiedao University, Shijiazhuang, China*

²*School of Civil Engineering, Lanzhou University of Technology, Lanzhou, China*

³*College of Transportation Engineering, Chang'an University, Xi'an, China*

Submitted 27 June 2019; resubmitted 30 April 2020, 29 July 2020; accepted 4 August 2020

Abstract. Traffic fundamental diagram is extremely important to analyse traffic flow and traffic capacity, and the central part of traffic fundamental diagram is to calibrate speed–density relationship. However, because of unbalanced speed–density observations, calibrating results using Least Square Method (LSM) with all speed–density points always lead to inaccurate effect, so this paper proposed a selecting data sample method and then LSM was used to calibrate four well-known single-regime models. Comparisons were made among the results using LSM with all speed–density points and the selecting data sample. Results indicated that the selecting data sample method proposed by this paper can calibrate the single-regime models well, and the method overcomes the inaccurate effect caused by unbalanced speed–density observations. Data from different highways validated the results. The contribution of this paper is that the proposed method can help researchers to determine more precise traffic fundamental diagram.

Keywords: traffic fundamental diagram, least square method, speed–density relationship, unbalanced speed–density observations, single-regime models, selecting data sample method.

Notations

Abbreviations

GA400 – Georgia state route 400 (US);
I-80 – interstate highway (US);
ITS – intelligent transportation system;
LSM – least square method;
MSE – mean squared error;
RE – relative error;
WLSM – weighted LSM.

\hat{v}_i – predicted speed;
 β – parameter/s;
 η – coefficient.

Variables

f – free-flow;
 k – density;
 k_i – corresponding density;
 k_j – jam density;
 k_0 – at-capacity density;
 m – density interval number;
 Q – sum of squared error;
 v – speed;
 v_f – free-flow speed;
 v_i – actual speed;
 v_0 – at-capacity speed;

Introduction

Traffic fundamental diagram is significant to analyse traffic flow and capacity (Zhang *et al.* 2019), and manage traffic operation (Alonso *et al.* 2019). It is also the basis to establish traffic flow models (Fiems *et al.* 2019) for traffic control (Zhu, Li 2019). The main part of traffic fundamental diagram is to determine the speed–density relationship. Therefore, it is very important to model the traffic fundamental diagram (Baer *et al.* 2019) and to calibrate the corresponding models precisely according to speed–density observations.

Researchers have done much work (Del Castillo, Benítez 1995a, 1995b; Jiang, Huang 2009; Lam *et al.* 2013; Wang *et al.* 2013) to analyse the speed–density relation-

*Corresponding author. E-mail: zhangchunbochn@yeah.net

#Editor of the TRANSPORT – the manuscript was handled by one of the Associate Editors, who made all decisions related to the manuscript (including the choice of referees and the ultimate decision on the revision and publishing).

ship, and many models were developed including single-regime models and multi-regimes. For single-regime models, Greenshields *et al.* (1935) according to the limited data thought the speed and density relationship was a straight line; Greenberg (1959) assumed traffic to behave like a continuous fluid and developed a speed–density relationship model, shown in Table 1; apparently, the Greenberg model fails to remain finite at zero density, Underwood thought that the infinity asymptote may be along the density scale (Drake *et al.* 1967) and the Underwood model was proposed as illustrated in Table 1; the speed–density data from the Eisenhower Expressway exhibited concavity at low densities and Drake *et al.* (1967) suggested a bell-shaped curve model, which is the Northwestern model as shown in Table 1; Newell (1961) considered the nonlinear car-following model and determined one speed–density model; Wang *et al.* (2011) established the logistic model of the equilibrium speed–density relationship motivated by the success of the logistic curves in modelling the growth phenomenon such as plant growth in agriculture, population dynamics, market growth in economics and epidemic growth in biology. For multi-regime models, Edie (1961) thought that the extreme of very light traffic and also the extreme of very dense traffic might show different behaviours and different models should be used; Sun and Zhou (2005) used cluster analysis to specify the number of regimes and it was shown that the k -means algorithm with original data worked well and could be conveniently used in practice. Although multi-regime models can accurately reflect the speed–density relationship to some extent, the mathematical elegance of multi-regime models is always not perfect, so the single-regime models are more often used. The four well-known single-regime models of Greenberg, Underwood, Northwestern and Newell are shown in Table 1, respectively.

Though almost all the researchers (Poole, Kotsialos 2016; Zhong *et al.* 2016; Knoop, Daamen 2017) calibrated the traffic fundamental diagram models using the LSM, it has been verified that the calibrating single-regime models

are not accurate under congested/jam conditions because of unbalanced observations of speed–density points (Qu *et al.* 2015). In other words, since most real speed–density observations are located in the uncongested condition (Maghrour Zefreh, Török 2020), the calibrated models using LSM can reflect the uncongested conditions precisely, but not under jam conditions, which can lead to significant errors for congested conditions. To overcome this problem, Qu *et al.* (2015, 2017) introduced WLSM to calibrate single-regime models, the weights were related to the density distance between adjacent data points and speed–density data points at congested conditions with little observations had large weights. Results showed better effect. However, the weight determination process of WLSM by Qu *et al.* (2015) is very complicated and the best weights of different models may be different (Zhang *et al.* 2017). Zhang *et al.* (2017) used five weight determination methods of WLSM to calibrate five one-regime speed–density models, and results showed that different models have different best calibrating models. Bhourri *et al.* (2019) proposed a data-driven approach for estimating the fundamental diagram, and unbiased fundamental diagram was obtained for both congested and uncongested observations. However, it is not suitable to single-regime models. To avoid using WLSM and to calibrate single-regime models precisely both under congested and uncongested conditions, this paper tries to propose a new method to calibrate single-regime models.

The structure of this paper is organized as follows. The methods and data used to calibrate single-regime speed–density models in the research are shown in Section 1. Results and discussion is presented in Section 2, followed by Results validation in Section 3. Then, conclusion is followed.

1. Methodologies and data

In this section, methodologies including the method to select data sample, LSM, and evaluation indicators are presented. The selecting data sample method is first designed to determine the data that are used to avoid unbalanced speed–density observations. Then, LSM is presented to calibrate the fundamental diagrams with the selected data. In addition, the calibration effectiveness is evaluated using RE and MSE, respectively. Besides, the original data information is presented at the end of this section.

1.1. Selecting data sample method

In order to avoid using the unbalanced speed–density observations, which leads to inaccurate calibrating, the selecting data sample method is proposed to balance the speed–density distribution to calibrate single-regime models. The specific steps of the method are as follows and are shown in Figure 1:

» *Step 1:* Rank the speed–density observations considering their densities (Qu *et al.* 2015) and speeds. Data points become:

Table 1. Four well-known single-regime speed–density models

Model	Function	Parameters
Greenberg	$v = v_0 \cdot \ln\left(\frac{k_j}{k}\right)$	v_0, k_j
Underwood	$v = v_f \cdot \exp\left(-\frac{k}{k_0}\right)$	v_f, k_0
North-western	$v = v_f \cdot \exp\left(-0.5 \cdot \left(\frac{k}{k_0}\right)^2\right)$	v_f, k_0
Newell	$v = v_f \cdot \left(1 - \exp\left(-\left(\frac{\eta}{v_f}\right) \cdot \left(\frac{1}{k} - \frac{1}{k_j}\right)\right)\right)$	v_0, k_j, η

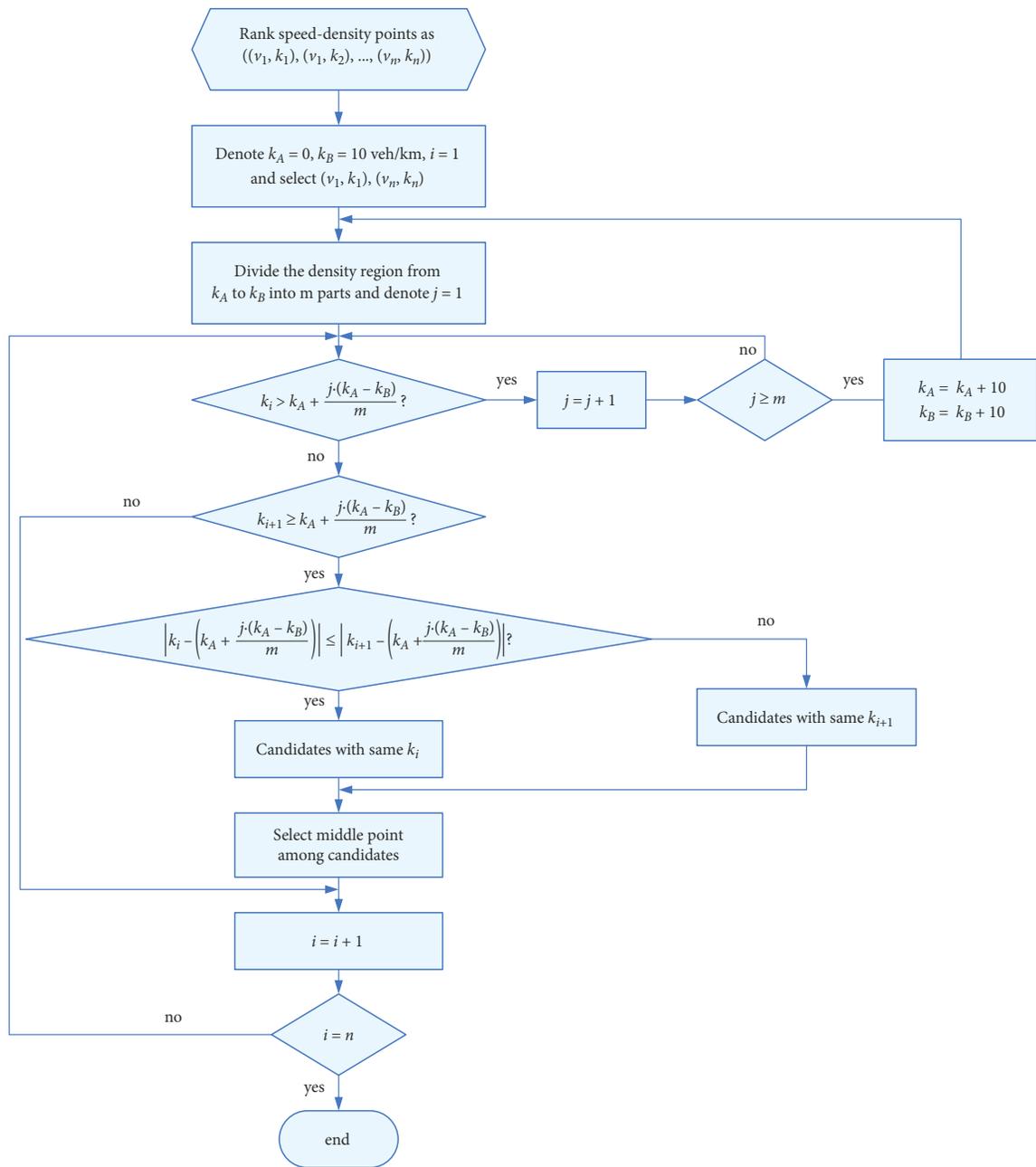


Figure 1. The flow chart of selecting data sample method

$$(v_1, k_1), (v_2, k_2), \dots, (v_i, k_i), \dots, (v_n, k_n), \quad (1)$$

where: $k_1 \leq k_2 \leq \dots \leq k_i \leq \dots \leq k_n$ and v_i is the corresponding speed, if $k_i = k_{i+1}$, then $v_i \geq v_{i+1}$;

»» Step 2: Denote $k_A = 0, k_B = 0$ veh/km, $i = 1$. Data points (v_1, k_1) and (v_n, k_n) are selected;

»» Step 3: Divide the density region from k_A to k_B into m parts and denote $j = 1$;

»» Step 4: If $k_i > k_A + \frac{j \cdot (k_B - k_A)}{m}$,

then

$$j = j + 1,$$

$$\text{if } j \geq m$$

then

$$k_A = k_A + 10, k_B = k_B + 10 \text{ and repeat Step 3}$$

otherwise

repeat Step 4

otherwise

$$\text{if } k_{i+1} \geq k_A + \frac{j \cdot (k_B - k_A)}{m},$$

then

if

$$\left| k_i - \left(k_A + \frac{j \cdot (k_B - k_A)}{m} \right) \right| \leq \left| k_{i+1} - \left(k_A + \frac{j \cdot (k_B - k_A)}{m} \right) \right|,$$

then

data points with same k_i are candidates as shown in Figure 2a and the middle point considering speed among these candidates is selected finally¹, go to Step 5

otherwise

data points with same k_{i+1} are candidates as shown in Figure 2b and the corresponding point is selected

otherwise

start Step 5;

» Step 5: Denote $i = i + 1$, if $i = n$, then stop otherwise repeat Step 4.

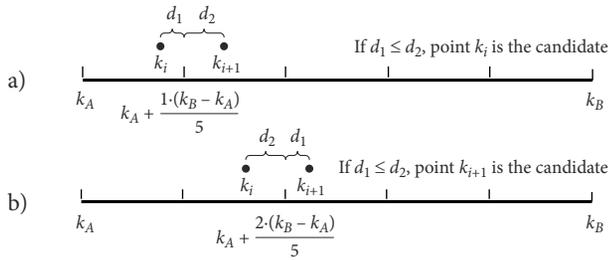


Figure 2. Selecting point process with two conditions: a – k_i is near $k_A + \frac{j \cdot (k_B - k_A)}{m}$; b – k_{i+1} is near $k_A + \frac{j \cdot (k_B - k_A)}{m}$

1.2. LSM

LSM is used in many fields to determine parameters, such as electric (Zheng et al. 2015), magnetic gradient tensor system (Yin et al. 2015), inductance estimation of electrically excited synchronous motor (Jeong et al. 2015), rating curves (Kim et al. 2014), electrochemical degradation of three reactive dyes (Djafarzadeh et al. 2014), laminar flow and heat transfer, thermal and flow analysis of microchannel heat sink (Hatami, Ganji 2014), electrohydrodynamic flow (Ghasemi et al. 2014) as well as transportation researches (Qu et al. 2015). The LSM seeks a solution that minimizes the function (Washington et al. 2020) as shown in Equation (2):

$$Q_{\min} = \sum_{i=1}^n (v_i - \hat{v}_i)^2 = \sum_{i=1}^n (v_i - f(\beta, k_i))^2. \quad (2)$$

By setting the partial derivative/s of Q with respect to β equal to zero, the least square estimated parameter/s β are/is obtained (Washington et al. 2020):

$$\frac{\partial Q}{\partial \beta} = 2 \cdot \sum_{i=1}^n (v_i - f(\beta, k_i)) \cdot \frac{\partial f(\beta, k_i)}{\partial \beta} = 0. \quad (3)$$

Solving Equation (3), the value/s of parameter/s β are/is obtained.

¹ If there is only one candidate, the candidate is as the middle point; if there are odd numbered candidates S , the $(S-1)/2$ th point of the candidates considering speed is the middle point; if there are even numbered candidates S , the $S/2$ th one is the middle point.

1.3. RE and MSE

To evaluate the effectiveness of the proposed methods, RE (Qu et al. 2015) and MSE (Washington et al. 2020) were used, as shown in Equations (4) and (5):

$$RE = \frac{1}{n} \cdot \sum_{i=1}^n \frac{|v_i - \hat{v}_i|}{v_i}, \quad (4)$$

$$MSE = \frac{1}{n} \cdot \sum_{i=1}^n (v_i - \hat{v}_i)^2, \quad (5)$$

where: the parameters are the same with above.

1.4. Data information

The original ITS data of GA400 were aggregated every 5 minutes, which were often used to calibrate the speed–density relationship models. One-year of 44787 continuous observations were obtained in 2003, of which the time interval is long enough to calibrate the speed–density relationships. The specific distribution of the data is shown in Figure 3a and in Table 2, and the unbalanced distribution of observation distribution can be seen easily.

2. Results and discussion

According to the method in Section 1.1, denote m equals 10, 20 and 50, respectively, and the three data samples are 125, 238 and 552 speed–density points respectively, as shown in Figure 3b and Table 3, Figure 3c and Table 4, Figure 3d and Table 5. And compared to Figure 3a, the speed–density points are balanced distribution.

The four well-known single-regime models in Table 1 with methods of LSM using all speed–density points, LSM using 125 sample points, LSM using 238 sample points and LSM using 552 sample points are calibrated in Figure 4 and Table 6.

From Figure 4, it is obvious that all single-regime models' calibrating results with methods LSM125, LSM238 and LSM552 are better than that with method of LSM except the Northwestern model, especially under congested/jam conditions. In order to analyse the specific calibrating efficiency, the REs and MSEs of these four well-known single-regime models with different methods are calculated, as shown in Figures 5 and 6.

From Figures 5 and 6, for Greenberg model, the REs of methods using LSM with selecting samples are less than 0.4 when the traffic density is more than 30 veh/km, which are less than that of method using LSM with all the speed–density points. The RE of using LSM with all speed–density points is even larger than 1 when the traffic density is more than 60 veh/km, and the REs of methods using LSM with selecting samples are much similar, which indicates that these methods all can calibrate the Greenberg model well. For REs of LSM with selecting samples, RE of LSM125 is the lowest one in a whole, RE of LSM552 comes second, and then is the RE of LSM238 and these indicates that the calibrating with method of LSM125 is the best one in the concern of RE. For MSEs, that is similar

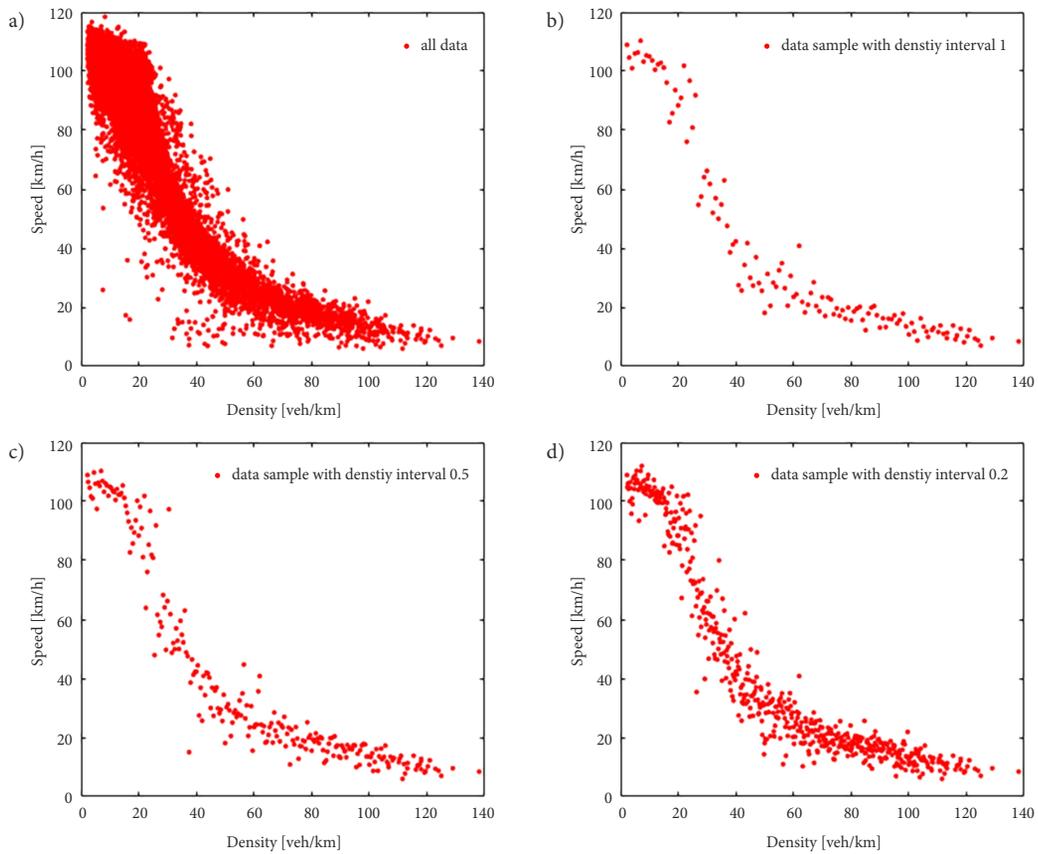


Figure 3. Speed–density data of GA400: a – all data; b – data sample of 125 points; c – data sample of 238 points; d – data sample of 552 points

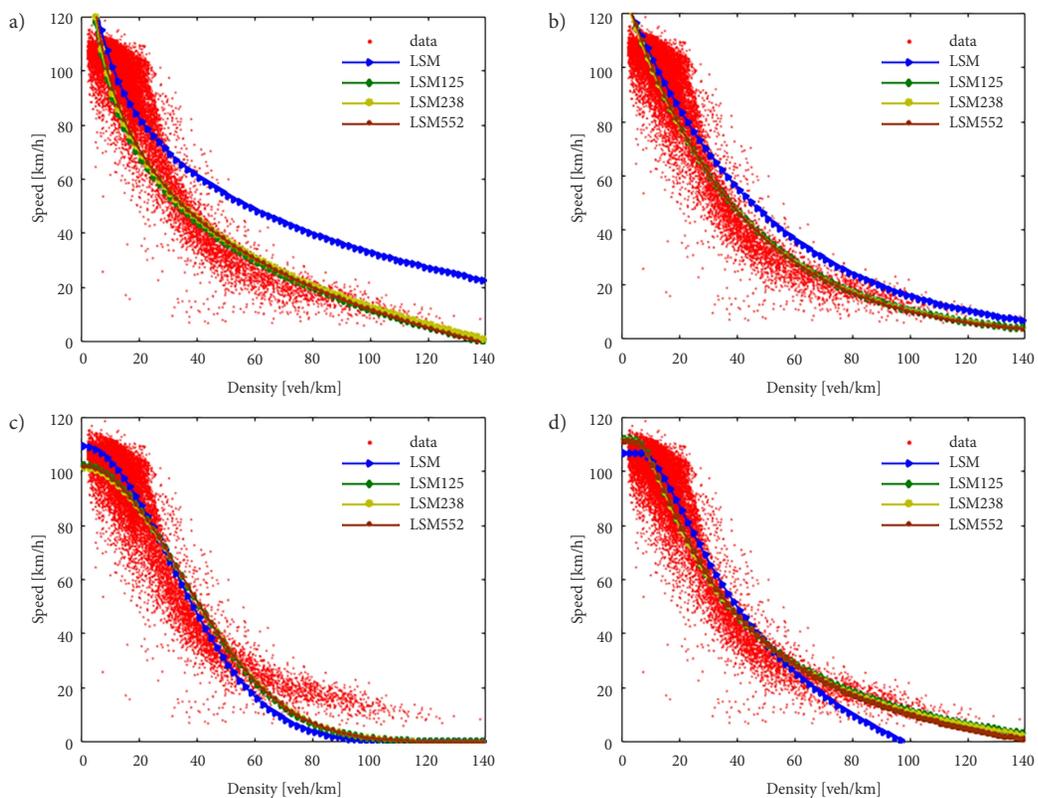


Figure 4. Four well-known single-regime models calibrated with different methods: a – Greenberg model; b – Underwood model; c – Northwestern model; d – Newell model (Notes: LSM means the LSM method using all speed–density points; LSM125 means the LSM using 125 speed–density sample points; LSM238 means the LSM using 238 speed–density sample points; LSM552 means the LSM using 552 speed–density sample points)

to REs of Greenberg model, MSEs of LSM with selecting speed–density samples are much lower than that of method using LSM with all points when the traffic density is more than 30 veh/km, and MSE of LSM125 is the lowest one in a whole, MSE of LSM552 comes second, and then is

the MSE of LSM238 and these indicates that the calibrating with method of LSM125 is the best one in the concern of MSE. In a word, the Greenberg model calibrating result of LSM selecting sample method is reasonable and much better than that of LSM using all speed–density points.

Table 2. Frequencies of GA400 speed–density data

Density [veh/km]	0...10	10...20	20...30	30...40	40...50	50...60	60...70
Frequencies	9333	29329	2665	1105	827	529	346
Density [veh/km]	70...80	80...90	90...100	100...110	110...120	120...130	130...140
Frequencies	268	173	136	48	21	6	1

Table 3. Frequencies of GA400 speed–density data sample with 125 points

Density [veh/km]	0...10	10...20	20...30	30...40	40...50	50...60	60...70
Frequencies	8	10	11	9	10	11	10
Density [veh/km]	70...80	80...90	90...100	100...110	110...120	120...130	130...140
Frequencies	9	11	9	10	10	6	1

Table 4. Frequencies of GA400 speed–density data sample with 238 points

Density [veh/km]	0...10	10...20	20...30	30...40	40...50	50...60	60...70
Frequencies	16	20	21	19	20	21	20
Density [veh/km]	70...80	80...90	90...100	100...110	110...120	120...130	130...140
Frequencies	19	21	19	18	17	6	1

Table 5. Frequencies of GA400 speed–density data sample with 552 points

Density [veh/km]	0...10	10...20	20...30	30...40	40...50	50...60	60...70
Frequencies	39	50	51	49	50	51	50
Density [veh/km]	70...80	80...90	90...100	100...110	110...120	120...130	130...140
Frequencies	49	51	48	36	21	6	1

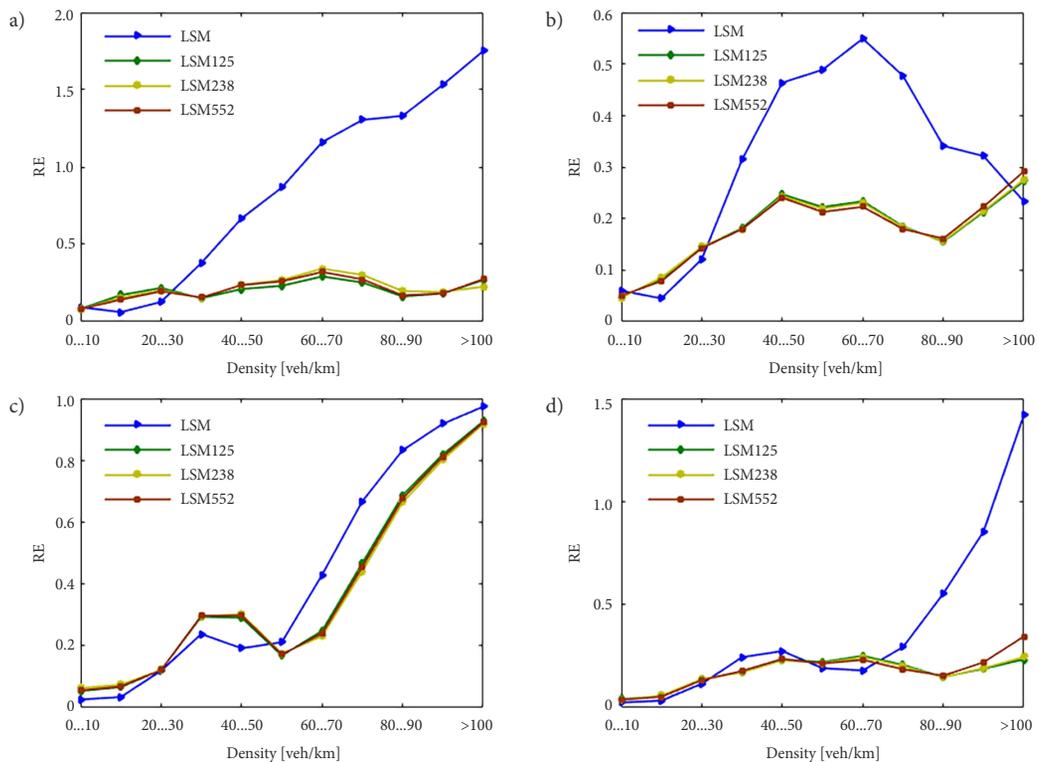


Figure 5. REs of four well-known single-regime models with different methods: a – REs of Greenberg model; b – REs of Underwoodmodel; c – REs of Northwestern model; d – REs of Newell model

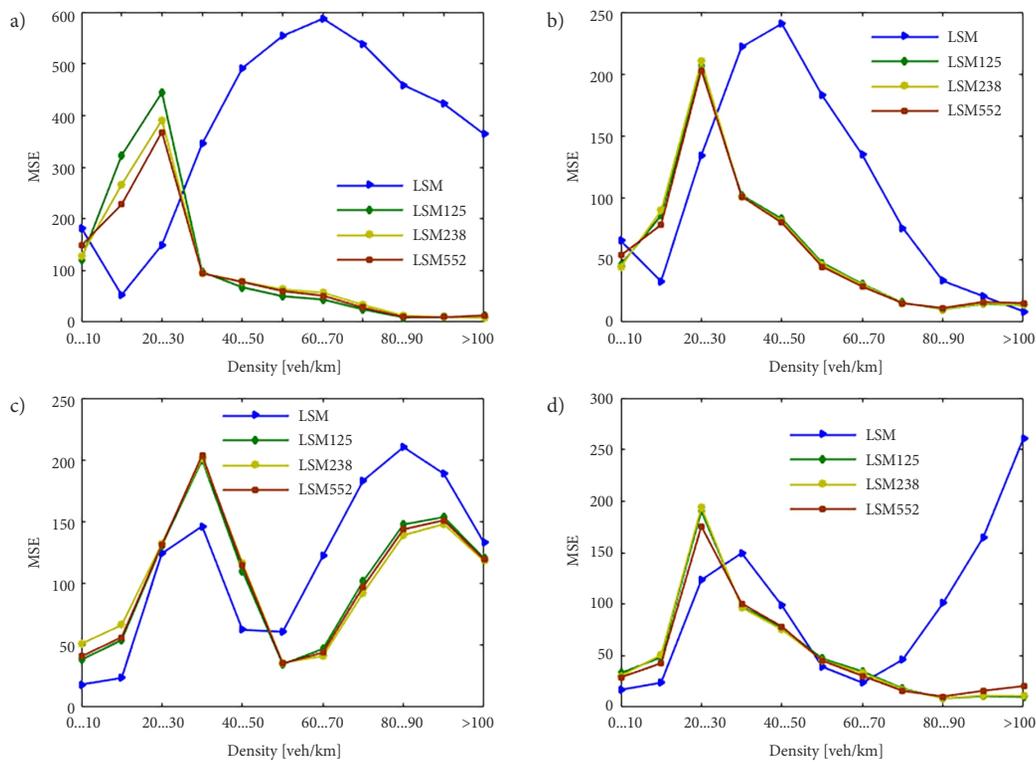


Figure 6. MSEs of four well-known single-regime models with different methods: a – MSEs of Greenberg model; b – MSEs of Underwood model; c – MSEs of Northwestern model; d – MSEs of Newell model

Table 6. Calibrating results of four well-known single-regime models with different methods

Model	Function	LSM	LSM125	LSM238	LSM552
Greenberg	$v = v_0 \cdot \ln\left(\frac{k_j}{k}\right)$	$v_0 = 30.88$ $k_j = 291.0$	$v_0 = 35.37$ $k_j = 140.1$	$v_0 = 35.70$ $k_j = 144.0$	$v_0 = 36.89$ $k_j = 138.4$
Underwood	$v = v_f \cdot \exp\left(-\frac{k}{k_0}\right)$	$v_f = 129.3$ $k_0 = 47.60$	$v_f = 129.8$ $k_0 = 39.82$	$v_f = 129.5$ $k_0 = 39.77$	$v_f = 131.4$ $k_0 = 39.16$
Northwestern	$v = v_f \cdot \exp\left(-0.5 \cdot \left(\frac{k}{k_0}\right)^2\right)$	$v_f = 109.5$ $k_0 = 31.06$	$v_f = 102.3$ $k_0 = 34.50$	$v_f = 101.0$ $k_0 = 34.99$	$v_f = 102.0$ $k_0 = 34.71$
Newell	$v = v_f \cdot \left(1 - \exp\left(-\left(\frac{\eta}{v_f}\right) \cdot \left(\frac{1}{k} - \frac{1}{k_j}\right)\right)\right)$	$v_f = 106.8$ $\eta = 4573$ $k_j = 98.36$	$v_f = 112.0$ $\eta = 3214$ $k_j = 161.8$	$v_f = 111.5$ $\eta = 3214$ $k_j = 159.8$	$v_f = 110.9$ $\eta = 3380$ $k_j = 145.4$

For Underwood model, the REs of methods using LSM with selecting samples are less than 0.3 when the traffic density is more than 30 veh/km. The REs of methods using LSM with selecting samples are much similar, which indicates that these methods all can calibrate the Underwood model well. For REs of LSM with selecting samples, RE of LSM552 is the lowest one in a whole, RE of LSM238 comes second, and then is the RE of LSM125 and these indicates that the calibrating with method of LSM552 is the best one in the concern of RE. For MSEs, that is similar to REs of Underwood model, MSEs of LSM with selecting speed–density samples are much lower than that of method using LSM with all points when the traffic density is more than 30 veh/km, and MSE of LSM552 is the lowest one in a whole, MSE of LSM238 comes sec-

ond, and then is the MSE of LSM125 and these indicates that the calibrating with method of LSM552 is the best one in the concern of MSE. In a word, the Underwood model calibrating result of LSM selecting sample method is reasonable and much better than that of LSM using all speed–density points.

For Northwestern model, the REs of methods using LSM with selecting samples are less than that of method using LSM with all the speed–density points when the traffic density is more than 50 veh/km. In addition, the REs of methods using LSM with selecting samples are much similar, which indicates that these methods all can calibrate the Northwestern model well. For REs of LSM with selecting samples, RE of LSM238 is the lowest one in a whole, RE of LSM552 comes second, and then is the RE

of LSM125. For MSEs, that is similar to REs of Northwestern model, MSEs of LSM of selecting speed–density samples are much lower than that of method using LSM with all points when the traffic density is more than 50 veh/km, and MSE of LSM238 is the lowest one in a whole, MSE of LSM552 comes second, and then is the MSE of LSM125. Though these indicated that Northwestern model calibrating result of LSM selecting sample method is better than that using all speed–density points, it is not recommended as the REs is even greater than 0.4 when the density is more than 70 veh/km.

For Newell model, the REs of methods using LSM with selecting samples are less than that of method using LSM with all the speed–density points when the traffic density is more than 70 and between 30 and 50veh/km. In addition,

the REs of methods using LSM with selecting samples are much similar, which indicates that these methods all can calibrate the Newell model well. For REs of LSM with selecting samples, RE of LSM552 is the lowest one in a whole, RE of LSM238 comes second, and then is the RE of LSM125, and these indicates that the calibrating with method of LSM552 is the best one in the concern of RE. For MSEs, that is similar to REs of Newell model, MSEs of LSM of selecting speed–density samples are much lower than that of method using LSM with all points when the traffic density is more than 70 and between 30 and 50 veh/km, and MSE of LSM552 is the lowest one in a whole, MSE of LSM238 comes second, and then is the MSE of LSM125 and these indicates that the calibrating with method of LSM552 is the best one in the concern

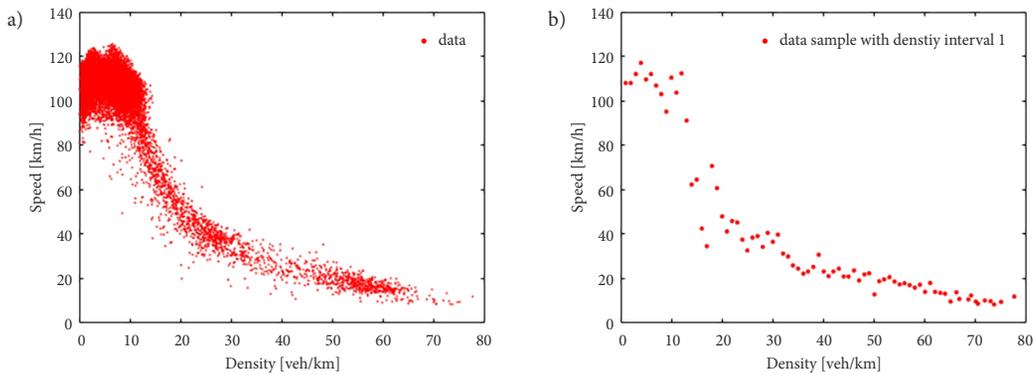


Figure 7. Speed–density observations of I-80: a – original data of I-80; b – data sample selected

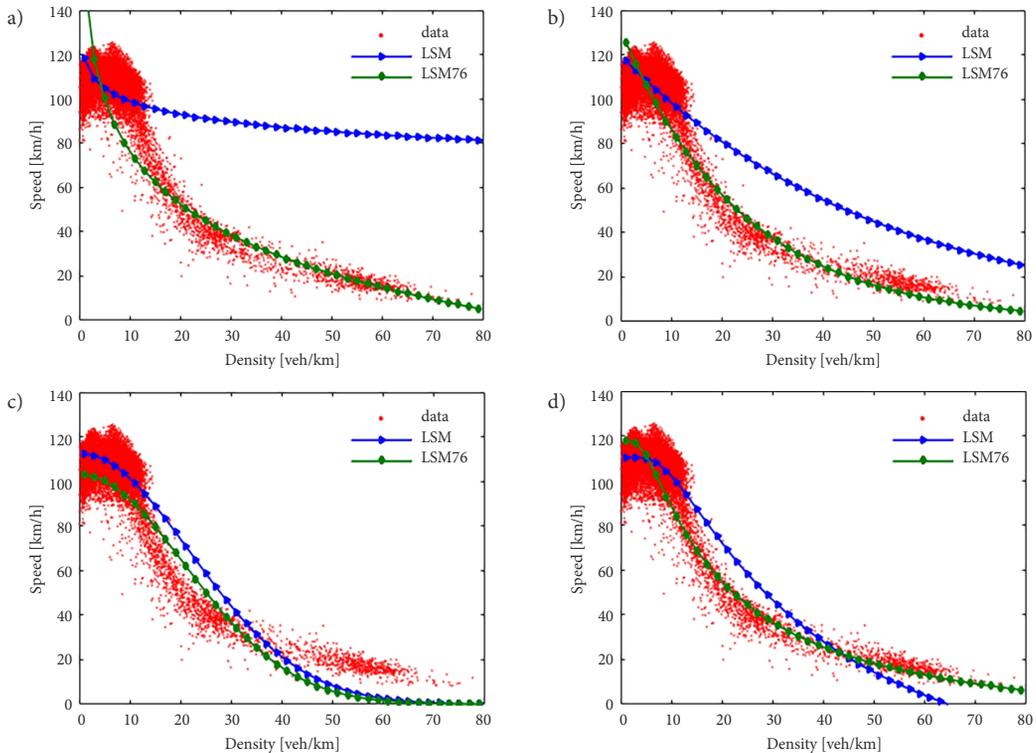


Figure 8. Fitting curves of four speed–density models for I-80: a – Greenberg model; b – Underwood model; c – Northwestern model; d – Newell model (Notes: LSM means the calibrating results using LSM with all speed–density data; LSM76 means those with the 76 selected speed–density points)

of MSE. In a word, the Newell model calibrating result of LSM selecting sample method is reasonable and much better than that of LSM using all speed–density points.

From above, the results of LSM125, LSM238 and LSM552 are different but very similar, so for convenience, we recommend the selecting data interval of density is 1 veh/km.

3. Results validation

To validate the effect of the proposed method in other highways, we used the data from I-80 in California to crosscheck the fitting results. The original data and the data sample selected with density interval 1 of I-80 are shown in Figure 7a and 7b, respectively. There are 76

speed–density points were selected as shown in Figure 7b. Unbalanced distribution of speed–density observation can be seen easily in Figure 7a and the data sample selected by the proposed method is almost evenly distributed shown in Figure 7b. The fitting curves using LSM with original data and data sample are shown in Figure 8 and Table 7. The results indicated that all calibrated models with data sample are better those with original data except North-western model.

To further analyse the calibration, RE and MSE were used to evaluated the effects and results are shown in Figures 9 and 10. If original data was used to calibrate speed–density models, the results with low density (0...20 veh/km) were perfect with no doubt, because most observations are located between 0 and 20 veh/km as shown in Figure 8.

Table 7. I-80 validation results of four well-known single-regime models with different methods

Model	Function	LSM	LSM76
Greenberg	$v = v_0 \cdot \ln\left(\frac{k_j}{k}\right)$	$v_0 = 8.450$ $k_j = 1.202 \cdot 10^6$	$v_0 = 34.35$ $k_j = 92.28$
Underwood	$v = v_f \cdot \exp\left(-\frac{k}{k_0}\right)$	$v_f = 119.6$ $k_0 = 51.12$	$v_f = 131.1$ $k_0 = 24.05$
Northwestern	$v = v_f \cdot \exp\left(-0.5 \cdot \left(\frac{k}{k_0}\right)^2\right)$	$v_f = 112.5$ $k_0 = 21.89$	$v_f = 103.1$ $k_0 = 20.87$
Newell	$v = v_f \cdot \left(1 - \exp\left(-\left(\frac{\eta}{v_f}\right) \cdot \left(\frac{1}{k} - \frac{1}{k_j}\right)\right)\right)$	$v_f = 110.3$ $\eta = 3383$ $k_j = 64.60$	$v_f = 117.7$ $\eta = 1794$ $k_j = 111.8$

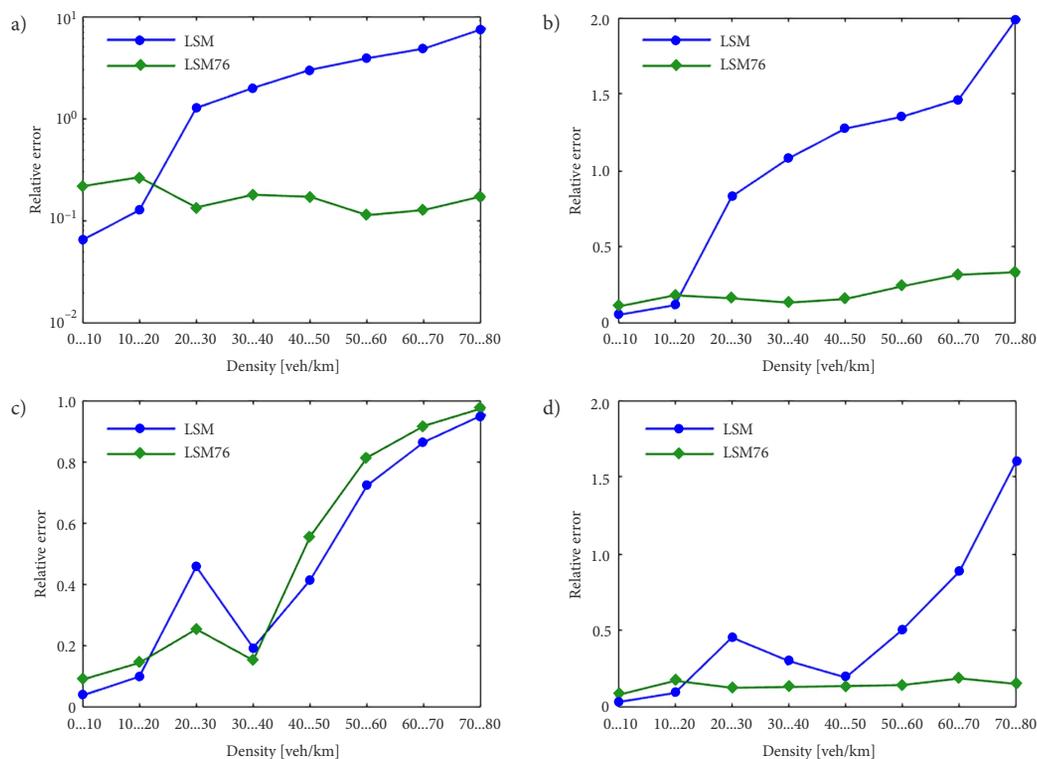


Figure 9. REs of four well-known single-regime models for I-80: a – REs of Greenberg model; b – REs of Underwood; c – REs of Northwestern model; d – REs of Newell model

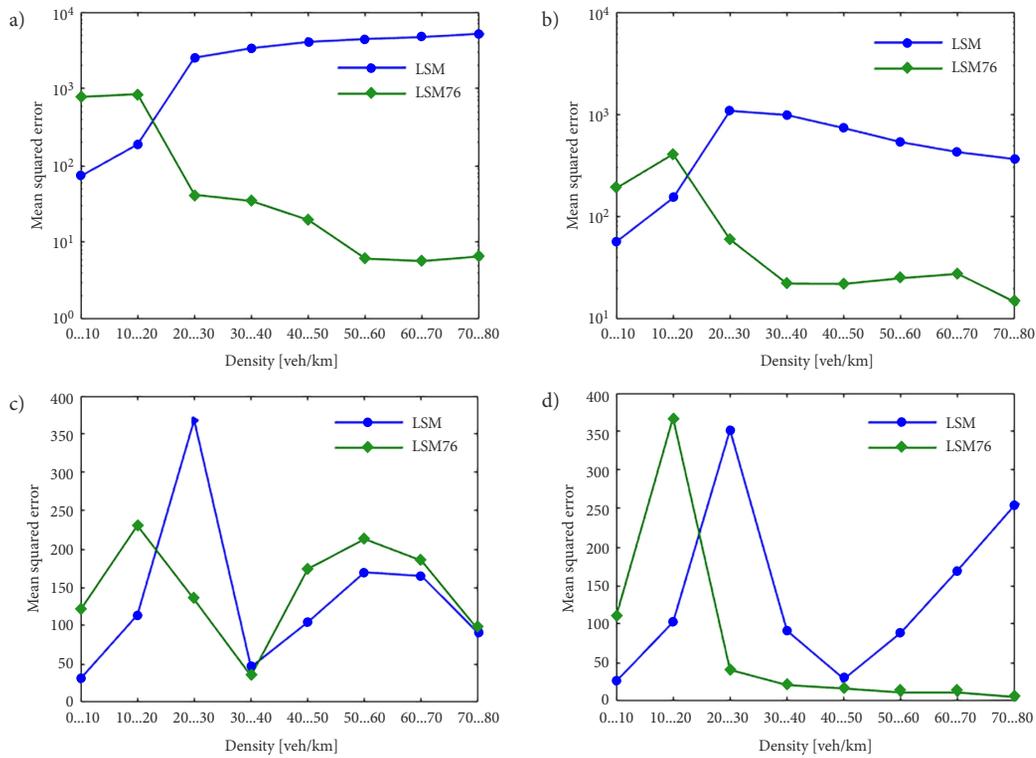


Figure 10. MSEs of four well-known single-regime models for I-80: a – MSEs of Greenberg model; b – MSEs of Underwood model; c – MSEs of Northwestern model; d – MSEs of Newell model

If density is high, the calibration with original data is not appropriate and REs are even greater than 0.5. However, the calibration with data sample selected using the proposed method is suitable whatever the density is, and REs are lower than about 0.3 except Northwestern model. MSEs also indicated the calibration with data sample is better than those with original data except Northwestern model. These verified that the proposed data section method is effective.

For Northwestern model, the calibration with either original data or data sample is not suitable, as the REs are even greater than 0.5 under high density as shown in Figure 9c. It is consistent with the result for GA400 as shown in Figure 5c. These showed that Northwestern model is not suitable for some specific speed–density data.

Conclusions

For highway and freeways, traffic flow operates at uncongested conditions for most time. Little speed–density data points measured are located in the congested area. If LSM is used to calibrate fundamental diagrams with the original speed–density data, it can lead to unprecise results. To overcome this problem, this study proposed a new method to calibrate fundamental diagrams.

The main contribution of this study is that a selecting data sample method was proposed to avoid the unbalanced speed–density observations, which can lead to inaccurate calibration of the single-regime speed–density models under congested/jam conditions calibrated using LSM. The specific steps of the method were explained and

illustrated in a figure. Real speed–density data from different highways were used to test the proposed method, and results indicated, in general, that the method can calibrate the single-regime speed–density relationship models precisely both under congested and uncongested conditions.

The method proposed by this paper can help to establish traffic flow models precisely and is useful for traffic control in practice. In the future, multi-regime speed–density models may be tested using the method proposed by this paper.

Acknowledgements

The authors would like to express gratitude to *Dr. Daiheng Ni*, University of Massachusetts Amherst (US), for generously supplying the data needed in the research.

This research was partially sponsored by Science and Technology Research Project of Colleges and Universities in Hebei province of the China (QN2020134).

Author contributions

Chunbo Zhang conceived the study and was responsible for the design and development of the data analysis.

Zhaoguo Huang and *Yonggang Wang* were responsible for data interpretation.

Disclosure statement

The authors declare that there is no conflict of interest regarding the publication of this paper.

References

- Alonso, B.; Ibeas, Á.; Musolino, G.; Rindone, C.; Vitetta, A. 2019. Effects of traffic control regulation on network macroscopic fundamental diagram: a statistical analysis of real data, *Transportation Research Part A: Policy and Practice* 126: 136–151. <https://doi.org/10.1016/j.tra.2019.05.012>
- Baer, N.; Boucherie, R. J.; Van Ommeren, J.-K. C. W. 2019. Threshold queueing to describe the fundamental diagram of uninterrupted traffic, *Transportation Science* 53(2): 585–596. <https://doi.org/10.1287/trsc.2018.0850>
- Bhouri, N.; Aron, M.; Hajsalem, H. 2019. A data-driven approach for estimating the fundamental diagram, *Promet – Traffic & Transportation* 31(2): 117–128. <https://doi.org/10.7307/ptt.v31i2.2849>
- Del Castillo, J. M.; Benítez, F. G. 1995a. On the functional form of the speed-density relationship – I: general theory, *Transportation Research Part B: Methodological* 29(5): 373–389. [https://doi.org/10.1016/0191-2615\(95\)00008-2](https://doi.org/10.1016/0191-2615(95)00008-2)
- Del Castillo, J. M.; Benítez, F. G. 1995b. On the functional form of the speed-density relationship – II: empirical investigation, *Transportation Research Part B: Methodological* 29(5): 391–406. [https://doi.org/10.1016/0191-2615\(95\)00009-3](https://doi.org/10.1016/0191-2615(95)00009-3)
- Djafarzadeh, N.; Safarpour, M.; Khataee, A. 2014. Electrochemical degradation of three reactive dyes using carbon paper cathode modified with carbon nanotubes and their simultaneous determination by partial least square method, *Korean Journal of Chemical Engineering* 31(5): 785–793. <https://doi.org/10.1007/s11814-013-0267-5>
- Drake, J. S.; Schofer, J. L.; May, A. D. 1967. A statistical analysis of speed–density hypotheses, *Highway Research Record* 154: 112–117.
- Edie, L. C. 1961. Car-following and steady-state theory for non-congested traffic, *Operations Research* 9(1): 66–76. <https://doi.org/10.1287/opre.9.1.66>
- Fiems, D.; Prabhu, B.; De Turck, K. 2019. Travel times, rational queueing and the macroscopic fundamental diagram of traffic flow, *Physica A: Statistical Mechanics and its Applications* 524: 412–421. <https://doi.org/10.1016/j.physa.2019.04.127>
- Ghasemi, S. E.; Hatami, M.; Mehdizadeh Ahangar, G. R.; Ganji, D. D. 2014. Electrohydrodynamic flow analysis in a circular cylindrical conduit using least square method, *Journal of Electrostatics* 72(1): 47–52. <https://doi.org/10.1016/j.elstat.2013.11.005>
- Greenberg, H. 1959. An analysis of traffic flow, *Operations Research* 7(1): 79–85. <https://doi.org/10.1287/opre.7.1.79>
- Greenshields, B. D.; Bibbins, J. R.; Channing, W. S.; Miller, H. H. 1935. A study of traffic capacity, *Highway Research Board Proceedings* 14: 448–477.
- Hatami, M.; Ganji, D. D. 2014. Thermal and flow analysis of microchannel heat sink (MCHS) cooled by Cu–water nanofluid using porous media approach and least square method, *Energy Conversion and Management* 78: 347–358. <https://doi.org/10.1016/j.enconman.2013.10.063>
- Jeong, I.; Gu, B.-G.; Kim, J.; Nam, K.; Kim, Y. 2015. Inductance estimation of electrically excited synchronous motor via polynomial approximations by least square method, *IEEE Transactions on Industry Applications* 51(2): 1526–1537. <https://doi.org/10.1109/TIA.2014.2339634>
- Jiang, Z.; Huang, Y.-X. 2009. Parametric calibration of speed–density relationships in mesoscopic traffic simulator with data mining, *Information Sciences* 179(12): 2002–2013. <https://doi.org/10.1016/j.ins.2009.02.005>
- Kim, B.; Lee, T.; Ouarda, T. B. M. J. 2014. Total least square method applied to rating curves, *Hydrological Processes* 28(13): 4057–4066. <https://doi.org/10.1002/hyp.9944>
- Knoop, V. L.; Daamen, W. 2017. Automatic fitting procedure for the fundamental diagram, *Transportmetrica B: Transport Dynamics* 5(2): 129–144. <https://doi.org/10.1080/21680566.2016.1256239>
- Lam, W. H. K.; Tam, M. L.; Cao, X.; Li, X. 2013. Modeling the effects of rainfall intensity on traffic speed, flow, and density relationships for urban roads, *Journal of Transportation Engineering* 139(7): 758–770. [https://doi.org/10.1061/\(ASCE\)TE.1943-5436.0000544](https://doi.org/10.1061/(ASCE)TE.1943-5436.0000544)
- Maghrour Zefreh, M.; Török, Á. 2020. Distribution of traffic speed in different traffic conditions: an empirical study in Budapest, *Transport* 35(1): 68–86. <https://doi.org/10.3846/transport.2019.11725>
- Newell, G. F. 1961. Nonlinear effects in the dynamics of car following, *Operations Research* 9(2): 209–229. <https://doi.org/10.1287/opre.9.2.209>
- Poole, A.; Kotsialos, A. 2016. Second order macroscopic traffic flow model validation using automatic differentiation with resilient backpropagation and particle swarm optimisation algorithms, *Transportation Research Part C: Emerging Technologies* 71: 356–381. <https://doi.org/10.1016/j.trc.2016.07.008>
- Qu, X.; Wang, S.; Zhang, J. 2015. On the fundamental diagram for freeway traffic: a novel calibration approach for single-regime models, *Transportation Research Part B: Methodological* 73: 91–102. <https://doi.org/10.1016/j.trb.2015.01.001>
- Qu, X.; Zhang, J.; Wang, S. 2017. On the stochastic fundamental diagram for freeway traffic: Model development, analytical properties, validation, and extensive applications, *Transportation Research Part B: Methodological* 104: 256–271. <https://doi.org/10.1016/j.trb.2017.07.003>
- Sun, L.; Zhou, J. 2005. Development of multiregime speed–density relationships by cluster analysis, *Transportation Research Record: Journal of the Transportation Research Board* 1934: 64–71. <https://doi.org/10.1177/0361198105193400107>
- Wang, H.; Li, J.; Chen, Q.-Y.; Ni, D. 2011. Logistic modeling of the equilibrium speed–density relationship, *Transportation Research Part A: Policy and Practice* 45(6): 554–566. <https://doi.org/10.1016/j.tra.2011.03.010>
- Wang, H.; Ni, D.; Chen, Q.-Y.; Li, J. 2013. Stochastic modeling of the equilibrium speed–density relationship, *Journal of Advanced Transportation* 47(1): 126–150. <https://doi.org/10.1002/atr.172>
- Washington, S.; Karlaftis, M.; Mannering, F.; Anastasopoulos, P. 2020. *Statistical and Econometric Methods for Transportation Data Analysis*. Chapman and Hall/CRC. 496 p.
- Yin, G.; Zhang, Y.; Fan, H.; Ren, G.; Li, Z. 2015. One-step calibration of magnetic gradient tensor system with nonlinear least square method, *Sensors and Actuators A: Physical* 229: 77–85. <https://doi.org/10.1016/j.sna.2015.03.026>
- Zhang, C.; Guo, X.; Xi, Z. 2017. Determination of observation weight to calibrate freeway traffic fundamental diagram using weighted least square method (WLSM), *Promet – Traffic & Transportation* 29(2): 203–212. <https://doi.org/10.7307/ptt.v29i2.2088>
- Zhang, P.; Yue, H.; Wang, P.; Shao C.; Zhang, X. 2019. Modeling the enveloping macroscopic fundamental diagram based on the traffic assignment with deterministic user equilibrium, *IEEE Access* 7: 69776–69794. <https://doi.org/10.1109/ACCESS.2019.2918551>

- Zheng, H.; Liang, Z.-F.; Li, M.-S.; Li, K. 2015. Optimization of parameters for LCL filter of least square method based three-phase PWM converter, *Journal of Electrical Engineering and Technology* 10(4): 1626–1634.
<https://doi.org/10.5370/JEET.2015.10.4.1626>
- 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.
<https://doi.org/10.1002/atr.1334>
- Zhu, W.-X.; Li, S. 2019. Study on discrete boundary-feedback-control strategy for traffic flow based on macroscopic fundamental diagram, *Physica A: Statistical Mechanics and its Applications* 523: 1237–1247.
<https://doi.org/10.1016/j.physa.2019.04.090>