Urban Mobility Analytics using Large-Scale Vehicle Trajectory Data

Lead Guest Editor: Jinjun Tang
Guest Editors: Xiaolei Ma and Lijun Sun

Urban Mobility Analytics using Large-Scale Vehicle Trajectory Data

Urban Mobility Analytics using Large-Scale Vehicle Trajectory Data

Lead Guest Editor: Jinjun Tang
Guest Editors: Xiaolei Ma and Lijun Sun

Hongtai Yang , China
Vincent F. Yu , Taiwan
Mustafa Zeybek, Turkey
Jing Zhao, China
Ming Zhong , China
Yajie Zou , China

Contents

WILEY | Hindawi

Research Article

A Time-Varying Coupling Analysis of Expressway Traffic Volume and Manufacturing PMI

Shuo Sun [1,2] **Mingchen Gu,** [1,2] **Yingping Wang,** [1,2] **Rongjie Lin,** [1] **Lifeng Xing,** [1] **and Zhiyuan Xu** [1,2]

[1]*Transport Planning and Research Institute, Ministry of Transport, Beijing 100028, China*
[2]*Laboratory for Traffic & Transport Planning Digitalization, Beijing 100028, China*

Correspondence should be addressed to Shuo Sun; sunshuo@tpri.org.cn

This study investigates the time-varying coupling relationship between expressway traffic volume and manufacturing purchasing manager index (PMI). First, for the traffic volume and manufacturing PMI time-series data, unit root stability test and Johansen cointegration test are applied to determine the stability of single sequence and the long-term stable correlation between variables, respectively. Then, a time-varying vector autoregressive model (TVP-VAR) is developed to quantify the time-varying correlation between variables. The time-varying parameters of TVP-VAR are estimated using the Markov chain Monte Carlo (MCMC) theory. Finally, the model is validated using examples from China. In the numeric example, three variables, i.e., expressway car traffic volume, expressway truck traffic volume, and manufacturing PMI, are selected for analysis. Results show that there is a positive interaction between expressway traffic volume (both car and truck) and manufacturing PMI. Express traffic volume slowly promotes the development of manufacturing industry. However, with the reform policy of road freight structure in China, the promotion effect of truck traffic on manufacturing PMI in the past two years has decreased significantly. Moreover, as affected by the China demand-led economic development model in recent years, the stimulus effect of manufacturing PMI on expressway passenger traffic volume has increased year by year. And, while the expressway freight structure remains stable, truck traffic volume is hardly affected by fluctuations in manufacturing PMI. These research results are helpful for policy makers to understand the time-varying coupling relationship between expressway traffic volume and manufacturing development and finally to improve the expressway management level.

1. Introduction

Expressway is the critical infrastructure for modern economic and social development and is also the important foundation for the construction of transportation modernization [1–3]. By the end of 2019, the expressway in China reached 149,600 km in mileage, ranking first in the world. Traffic volume of the section reached 27,000 passenger car unit per day (pcu/d). The fast-developing process of the expressway has played a positive role in promoting the sustained and rapid growth of manufacturing economy [4]. During the coronavirus pandemic (COVID-19) period in 2020, China's expressway vehicle tolls were waived until May 5, and under such condition, traffic volume has increased by

20% compared with the same period in the last year. Clearly, the increase in traffic has helped the resumption of work and production in economic recovery period. However, existing methods cannot dynamically measure the promotion effect of expressway traffic on manufacturing economy. The mutual relationship between expressway traffic and manufacturing economy needs in-depth study.

In recent years, many researchers have analyzed the impact of expressway on economic development of manufacturing industries. Their results showed a positive impact of expressway capital stock on gross domestic product (GDP) [5]. For example, areas with high traffic mileage have higher manufacturing and employment levels, and there is a positive lagging correlation between road

attributes and manufacturing [6]. From the perspective of spatial dependence and heterogeneity, the impact of improving transportation infrastructure on manufacturing agglomeration is also positive [7]. Moreover, transportation infrastructure has a typical spatial nonlinear diminishing effect on manufacturing agglomeration [8].

Expressway traffic volume is one of the most important indicators for operation analysis. Many researchers have proved that expressway traffic volume is closely related to GDP growth. Within years, this relationship gradually strengthens [9, 10]. Results of the cointegration test showed that there is a long-term equilibrium relationship between the freight volume and the output value of primary, secondary, and tertiary industries [11]. These results indicate that expressway traffic volume is strongly related to the development of national economy [12].

Current investigations generally have the following problems. ① Traffic volumes of private car and truck are not analyzed separately. Current studies only analyze the correlation between overall traffic volume and manufacturing economy. ② Time-varying characteristics of this relationship are ignored in current studies, despite the relationship between expressway traffic and manufacturing economy changes month by month. ③ Long-term and stable relationships are usually not considered. Therefore, there is an urgent need to investigate the time-varying and long-term effect of expressway traffic on manufacturing economy.

This article applies the time-varying analysis method in econometrics to examine the long-term inter-relationship between time-varying economic index of manufacturing industry and expressway traffic (including private car and truck). The promotion effect of manufacturing activity on expressway traffic, as well as the promotion effect of expressway traffic manufacturing activity, is quantitatively analyzed. Therefore, their long-term mutual promotion effect can be observed. This research is conducive for policy makers to deepen the understanding of social and economic promotion functionality of expressway development.

2. Methods

Over the course of modern macroeconomic analysis, the time-varying vector autoregressive model (TVP-VAR) is often used to explore the time-varying interaction of economic variables on macroeconomic issues [13]. The TVP-VAR model was originated from the vector autoregressive model (VAR), which is a large-scale simultaneous equation model pioneered by Sims in 1980 [14]. The VAR model can simulate the feedback mechanism between variables when analyzing the relationship between variables. However, the model did not distinguish the specific meaning and variable types of each equation. The VAR model is also constructed based on the lag value of the explained variable and the lag value of other variables. Such facts cause the model to ignore the influence of contemporaneous value of different variables.

Contemporaneous correlation refers to the correlation between the current disturbance term and the explanatory variable. In order to analyze the contemporaneous correlation between variables, Blanchard and Watson [15] constructed a structural vector autoregressive model (SVAR) with structural characteristics. While, in SVAR, the coefficients of equation variables are fixed over time, suggesting that the action mechanism between variables is relatively fixed. This model setting lets the model violate the fact that the actual interaction between variables will fluctuate over time. In recent years, Primiceri [16] proposed a TVP-VAR model with random fluctuations. The model adds time-varying coefficients and variance-covariance matrix to the SVAR model. Thus, the interaction intensity and the time-varying transmission mechanism between variables can be observed. Based on the work of Primicer, Nakajima [17] proposed a set of estimation methods for the TVP-VAR model.

The basic VAR model can be written as

$$y_t = B_0 + B_1 y_{t-1} + B_2 y_{t-2} + \cdots + B_p y_{t-p} + u_t, \qquad (1)$$

where p represents the lag order and $t = p + 1, p + 2, \ldots, n$. $y_t, y_{t-1}, \ldots, y_{t-p}$ refer to a set of $k \times 1$ endogenous variables, which are composed of manufacturing PMI and traffic volume of expressway private car and truck. B_0 is a $k \times 1$ vector of constants, $B_1, B_2, \ldots, B_p$ are coefficient matrices and represent the intertemporal correlation, and u_t is the structural shock at time step t and satisfies the following criteria [18]:

$$E(u_t) = 0,$$

$$E(u_t u_t') = \sum_u = \begin{bmatrix} \sigma_{u1}^2 & 0 & \ldots & 0 \\ 0 & \sigma_{u2}^2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \sigma_{uk}^2 \end{bmatrix}, \qquad (2)$$

$$E(u_t u_\tau) = 0, t \neq \tau.$$

Originated from the VAR model, the SVAR model considers the contemporaneous effect between variables in order to capture the variables' instantaneous structural relationship. The SVAR model can be structured as

$$A y_t = B_0 + B_1 y_{t-1} + B_2 y_{t-2} + \cdots + B_p y_{t-p} + u_t, \qquad (3)$$

where $A, B_1, \ldots, B_p$ are $k \times k$ time-invariant coefficient matrices and represent the intertemporal correlation. A represents the contemporaneous correlation, and $B_1, \ldots, B_p$ represent the intertemporal correlation. B_0 is a $k \times 1$ vector of constants.

Generally, A is supposed to be a lower triangular matrix in order to reduce the computational complexity without affecting the inference results:

$$A = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ a_{2,1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ a_{k,1} & \ldots & a_{k,k-1} & 1 \end{bmatrix}. \qquad (4)$$

SVAR can be converted into the following induced form:

$$y_t = F_0 + F_1 y_{t-1} + F_2 y_{t-2} + \cdots + F_p y_{t-p} + A^{-1} \sum_u \varepsilon_t, \quad \varepsilon \sim N(0, I_k),$$
(5)

where $F_i = A^{-1} B_i$, for $i = 0, 1, \ldots, p$. $\sum_u$ is the disturbance term in the diagonal form. By defining $X_t = I_k \otimes (1, y'_{t-1}, \ldots, y'_{t-p})$ and $\beta = (F_0, F_1, \ldots, F_p)'$, where $\otimes$ represents Kronecker product, equation (5) can be transformed into

$$y_t = X_t \beta + A^{-1} \Sigma_u \varepsilon_t.$$
(6)

In the SVAR model, coefficients in formula (6) remain unchanged over time. By allowing coefficients and error terms to change over time, SVAR can be extended to the TVP-VAR model:

$$y_t = X_t \beta_t + A_t^{-1} \Sigma_t \varepsilon_t,$$
(7)

where β_t, A_t, and Σ_t are time-varying coefficients. The $k \times (k-1)/2$ elements in A_t can be further stacked into a vector $a_t = (a_{21,t}, a_{31,t}, a_{32,t}, \ldots, a_{k(nk-1),t})'$. Similarly, the diagonal matrix Σ_t is transformed into a vector $\sigma_t = (\sigma_{1t}, \sigma_{2t}, \ldots, \sigma_{kt})'$. Therefore, $h_t = (h_{1t}, \ldots, h_{kt})'$ and $h_{jt} = \ln \sigma_{jt}^2$ for $j = 1, \ldots, k$. According to the work of Primieri [16], parameters in equation (7) obey the first-order random walk process:

$$\beta_{t+1} = \beta_t + v_{t+1}, a_{t+1} = a_t + \varsigma_{t+1}, h_{t+1} = h_t + \xi_{t+1},$$
(8)

where $v_t \sim N(0, \Omega_\beta)$, $\varsigma_t \sim N(0, \Omega_a)$, and $\xi_t \sim N(0, \Omega_h)$ are new disturbance items. Under the assumption that coefficients are uncorrelated, the variance-covariance matrix of all perturbations in equation (7) can be written as

$$V = \text{Var}\left(\begin{bmatrix} \varepsilon_t \\ v_t \\ \varsigma_t \\ \xi_t \end{bmatrix}\right) = \begin{pmatrix} I_k & 0 & 0 & 0 \\ 0 & \Omega_\beta & 0 & 0 \\ 0 & 0 & \Omega_a & 0 \\ 0 & 0 & 0 & \Omega_h \end{pmatrix}.$$
(9)

The above calculation process assumes that the time-varying coefficient follows a first-order random walk process. According to the research of Engle and Watson [19], when an individual receives new information and adjusts the estimation process of equation state, this fluctuation should follow the random Wandering process. Secondly, the random walk process allows the greatest degree of structural parameter changing. This feature allows the TVP-VAR model to capture subtle disturbances. Moreover, the random walk process can reduce the calculation difficulty and the number of calculations, thereby preventing the model from overfitting. Many studies have also proved that the traffic flow variability can be described as a Markov random process with random fluctuations. At the same time, according to the estimation of Primiceri [16], results of the TVP-VAR model are insensitive to assumptions.

For the TVP-VAR model, we use the Markov Chain Monte Carlo (MCMC) method under the Bayesian inference framework to estimate the unknown parameters.

Parameters are first assumed to be constant as in the SVAR model. Then, the least square method is applied to calculate prior initial values of β_{ols}, a_{ols}, h_{ols}, $V(\beta_{ols})$, and $V(a_{ols})$. Replace the variance of h_{ols} with I_k. Initial values can be obtained as

$$\begin{aligned} \beta_0 &\sim N(\beta_{ols}, 4V(\beta_{ols})), \\ a_0 &\sim N(a_{ols}, 4V(a_{ols})), \\ h_0 &\sim N(h_{ols}, 4I_k). \end{aligned}$$
(10)

The variance part is multiplied by 4 in order to capture more uncertain information and to prevent prior parameter constraints from being too tight.

The hyperparameters are independent of each other and follow inverse-Wishart distributions. Then, according to Primiceri [16], set its prior distribution as follows:

$$\begin{aligned} \Omega_\beta^{-1} &\sim W\left(1 + k_\beta, \left((r_\beta)^2 \cdot (1 + k_\beta) \cdot V(\beta_{ols})\right)^{-1}\right), \\ \Omega_a^{-1} &\sim W\left(1 + k_a, \left((r_a)^2 \cdot (1 + k_a) \cdot V(a_{ols})\right)^{-1}\right), \\ \Omega_h^{-1} &\sim W\left(1 + k_h, \left((r_h)^2 \cdot (1 + k_h) \cdot I_k\right)^{-1}\right), \end{aligned}$$
(11)

where k_β, k_a, and k_h represent the latitude of β_t, a_t, and h_t, respectively. For the n-dimensional p-order lag SVAR model, compute $k_\beta = pk^2$, $k_a = k(k-1)/2$, and $k_h = k(k-1)/2$. In addition, r_β, r_a, and r_h are set as 0.01, 0.1, and 0.01, respectively, based on the parameter setting defined by Koop and Korobilis [20].

After obtaining the initial value, divide the parameters of different categories into different parameter blocks so that different parameter blocks have greater independence. Then, sample the conditional posterior of each part, and finally, form the conditional posterior distribution of all parameters. Finally, dynamic parameters of each simultaneous equation in the TVP-VAR model is obtained based on the Monte Carlo simulation technology, in which a balanced distribution of Markov process is constructed.

In order to explain the contemporaneous correlation between variables and to investigate their feedback mechanism, this study uses impulse response to quantify the magnitude of influence between explanatory variables. The impulse response is performed by first fixing perturbations of all other explanatory variables and then allowing the value or the standard deviation of one explanatory variable to deviate one unit from its original value. The newly obtained value of the explanatory variable is called the impulse response [21]. Through impulse response analysis, dynamic influences between unclear variables can be well explained and quantified. The analysis also facilitates the comparison of shock magnitude between variables. The impulse response explains the synchronous correlation between variables and the dynamic influence of lag. Through impulse response function, mutual influences between variables can be quantitatively captured. In the paper, impulse response function is applied to analyze the influence degree between expressway traffic and manufacturing PMI.

3. Data Description

Expressway traffic can be classified into passenger cars (small and medium passenger cars and coaches) and trucks (small trucks, medium trucks, large trucks, extralarge trucks, and containers) according to data collection samples of national expressway. While sensors on some road sections will also collect traffic data about motorcycles and tractors, we consider traffic volume of passenger cars and trucks only since these two types of traffic can intuitively reflect the development of expressways in recent years. Thus, three types of variables, i.e., traffic volume of private cars, traffic volume of trucks, and manufacturing PMI, are analyzed in this paper.

Expressway traffic volumes of passenger cars and trucks are collected by continuously spaced traffic survey stations in the China National Traffic Survey Data Collection System. At present, about 2,200 continuous traffic survey stations are deployed on expressways throughout the country. These traffic survey stations transmit data to the traffic survey system in real time at 5-minute intervals. The data accuracy is over 90%. This paper uses the China National Traffic Survey Data Collection System to extract expressway passenger car and truck traffic volume.

Monthly manufacturing PMI data are obtained from questionnaire surveys of sample companies. The index covers multiple aspects of manufacturing economic activities and has shortest lag period among all macroeconomic series data. Overall, manufacturing PMI reveals the development trend of manufacturing industries in detail. Thus, the index provides an important basis for the country's macroeconomic regulation and guidance of business operations. Generally, the higher the manufacturing PMI, the better the economic trend of manufacturing industry. The department of China National Bureau of Statistics regularly releases the PMI of China's manufacturing industry on the first working day of each month.

In order to analyze the stable and long-term contemporaneous relationship between expressway traffic and PMI, this paper selects expressway passenger car and truck traffic volume and manufacturing PMI data from January 2010 to December 2019 on a monthly basis. It should be noted that the TVP-VAR model explains the interaction between variables by capturing the fluctuations of explanatory variables. Therefore, the dimensionality reduction of data will not affect the calculation results under the condition that the time-series trend remains unchanged. For expressway passenger car and truck traffic volume, divide them by 1000 and take the arithmetic square root. For manufacturing PMI, divide them by 10 and take the arithmetic square root. Figure 1 shows the over-time trend of expressway passenger car and truck traffic volume and manufacturing PMI.

3.1. Data Test.

In order to test whether there is a long-term stable correlation between passenger car and truck traffic volume and manufacturing PMI and to ensure that the causal relationship between variables is not a pseudoregression, it is necessary to conduct data stability tests and cointegration relationship tests on the three variables [22].

3.2. Stationarity Test.

Stationarity test is to verify whether time series of one variable is stationary. In this paper, the augmented Dickey-Fowler test (ADF) and the unit root test are used to determine the stability of a single variable. Table 1 presents the test results.

The results show that, at 5% significant level, the original ADF test values of passenger car traffic, truck traffic, and manufacturing PMI are all greater than the critical value, indicating that the time series of three variables are not stable at the original level. Under the condition of first-order difference, the ADF test values of passenger car traffic, truck traffic, and manufacturing PMI are all less than the critical value at 5% significant level, suggesting that the time series of three variables are stable. Such result further supports the cointegration relationship test.

3.3. Cointegration Relationship Test.

The cointegration relationship test is to verify the long-term correlation between traffic volume (i.e., passenger car and truck) and manufacturing PMI. A VAR model is established for the three variables. Lag order of the three variables is determined to be 1 through the Akaike information criterion (AIC). Table 2 shows the cointegration test results of the three variables under the condition of first-order difference.

Trace and maximum eigenvalue tests in Table 2 shows that the three variables have at most 2 long-term stable correlations when 5% critical value is considered. Therefore, there is a long-term and stable correlation between passenger car, truck traffic volume, and manufacturing PMI. The test data can be used for TVP-VAR analysis.

4. Results and Analysis

The three variables, i.e., traffic volume of expressway passenger car, traffic volume of truck, and manufacturing PMI, are substituted into the TVP-VAR model. C represents expressway passenger car traffic volume. T represents expressway truck traffic volume. P represents manufacturing PMI index. The lag order is set as $p = 1$ according to AIC test results. In impulse response analysis, time lags of 1 month, 2 months, 3 months, and 6 months are considered in order to measure the short-term, midterm, and long-term interactions between variables. In model estimation, least squares algorithm is conducted 4000 times for prior sampling and parameter initialization. Then, 40,000 iterations of Markov chain Monte Carlo simulation are carried out. Figure 2 shows the final calculation result.

Figure 2 shows the time-varying impulse response between the three variables in China from 2010 to 2019 under time lags of 1 month, 2 months, 3 months, and 6 months. In the figure, x- and y-axis represent month and the magnitude of impulse response, respectively. A positive impulse response indicates a positive influence between the two variables, while a negative one represents a negative influence relationship. Moreover, the greater the absolute value of impulse response, the more significant the impact. Taking response of T to C in Figure 2(b) as an example, the result represents the time-varying influence of expressway passenger car traffic volume
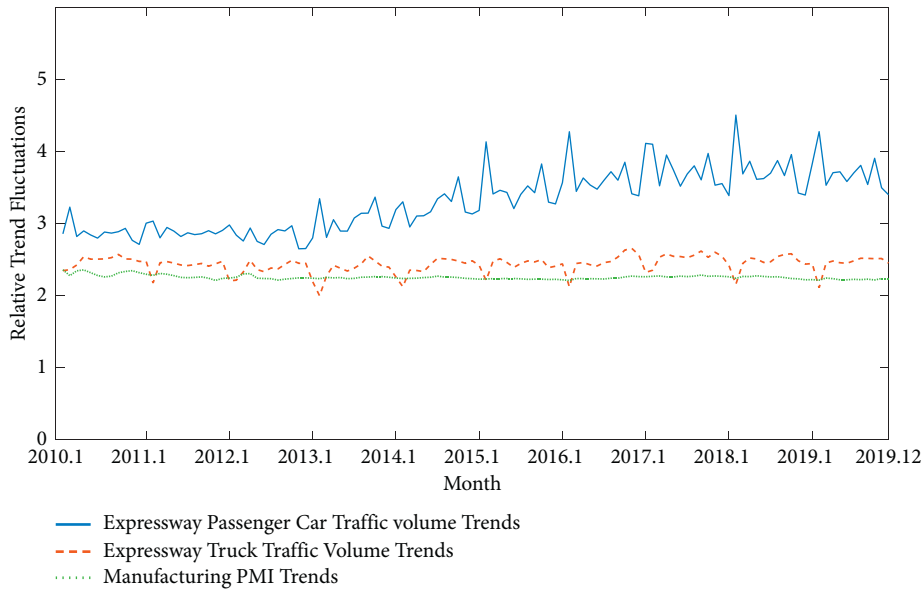
Figure 1: Variation tendency of expressway passenger car and truck traffic volume and manufacturing PMI.

Table 1: Unit root test of expressway passenger car, freight traffic, and manufacturing PMI.

Difference order	5% level	ADF test results		
		Private car	Truck	Manufacturing PMI
Level	−1.94	2.5944	−0.3569	−0.6736
1st difference	−2.89	−5.9552	−3.9243	−12.4554

Table 2: Cointegration test of expressway passenger car, freight traffic, and manufacturing PMI.

Hypothesized no. of CE (s)	Eigenvalue	Trace			Maximum eigenvalue		
		Trace statistic	5% critical value	Prob.	Maxeigen statistic	5% critical value	Prob.
None	0.3104	63.0907	29.7971	0.0000	43.4819	21.1316	0.0000
At most 1	0.1312	19.6088	15.4947	0.0113	16.4514	14.2646	0.0222
At most 2	0.0266	3.1574	3.8415	0.0756	3.1574	3.8415	0.0756

(a)

(b)

Figure 2: Continued.

(c)

(d)

(e)

(f)
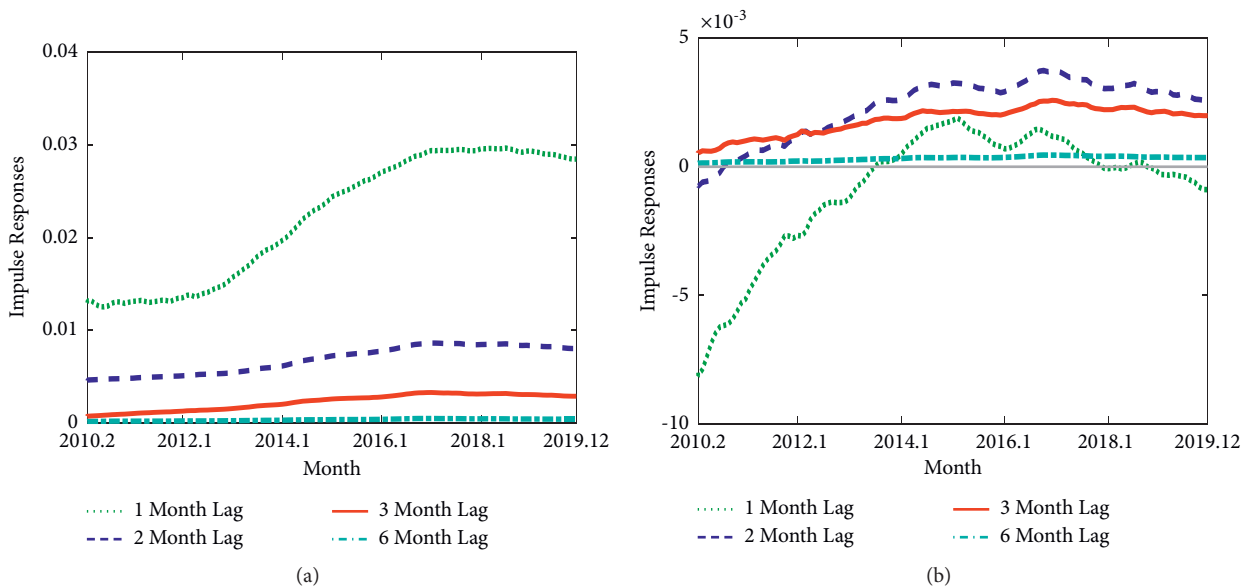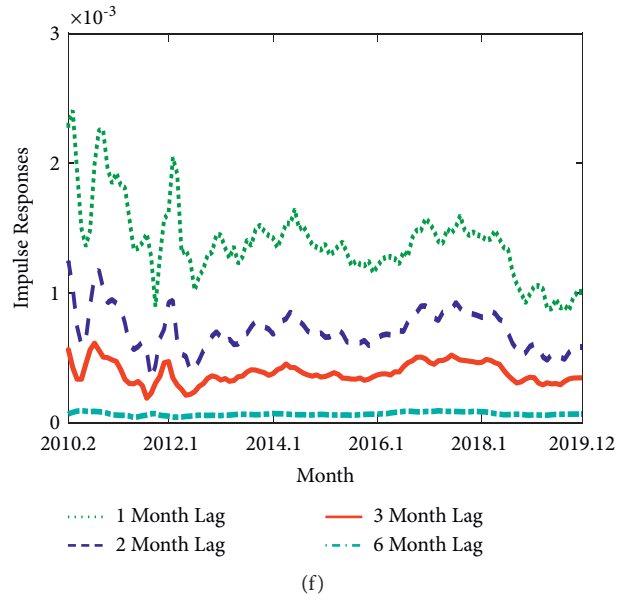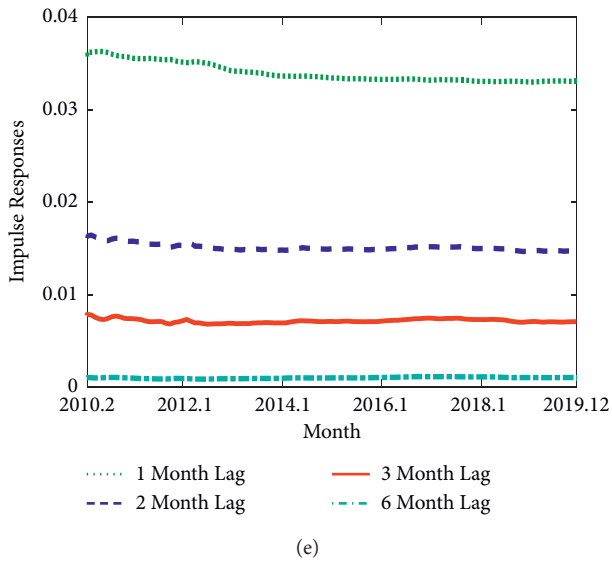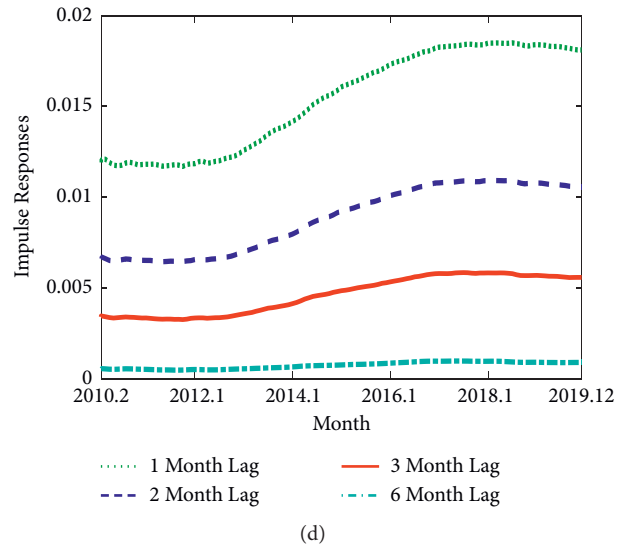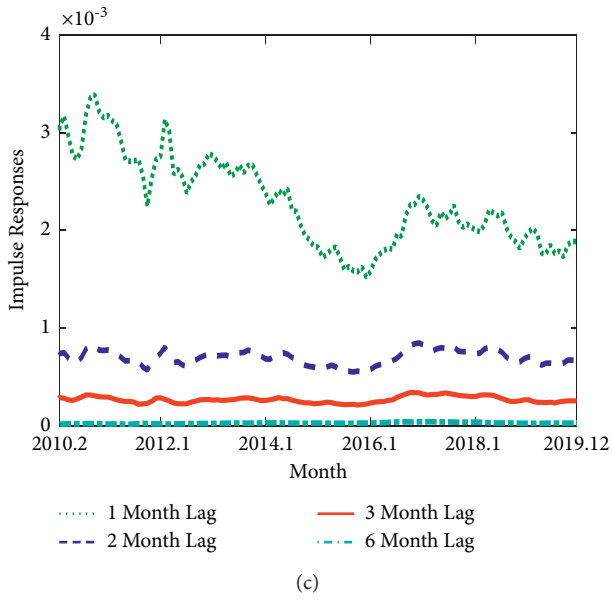
Figure 2: Continued.
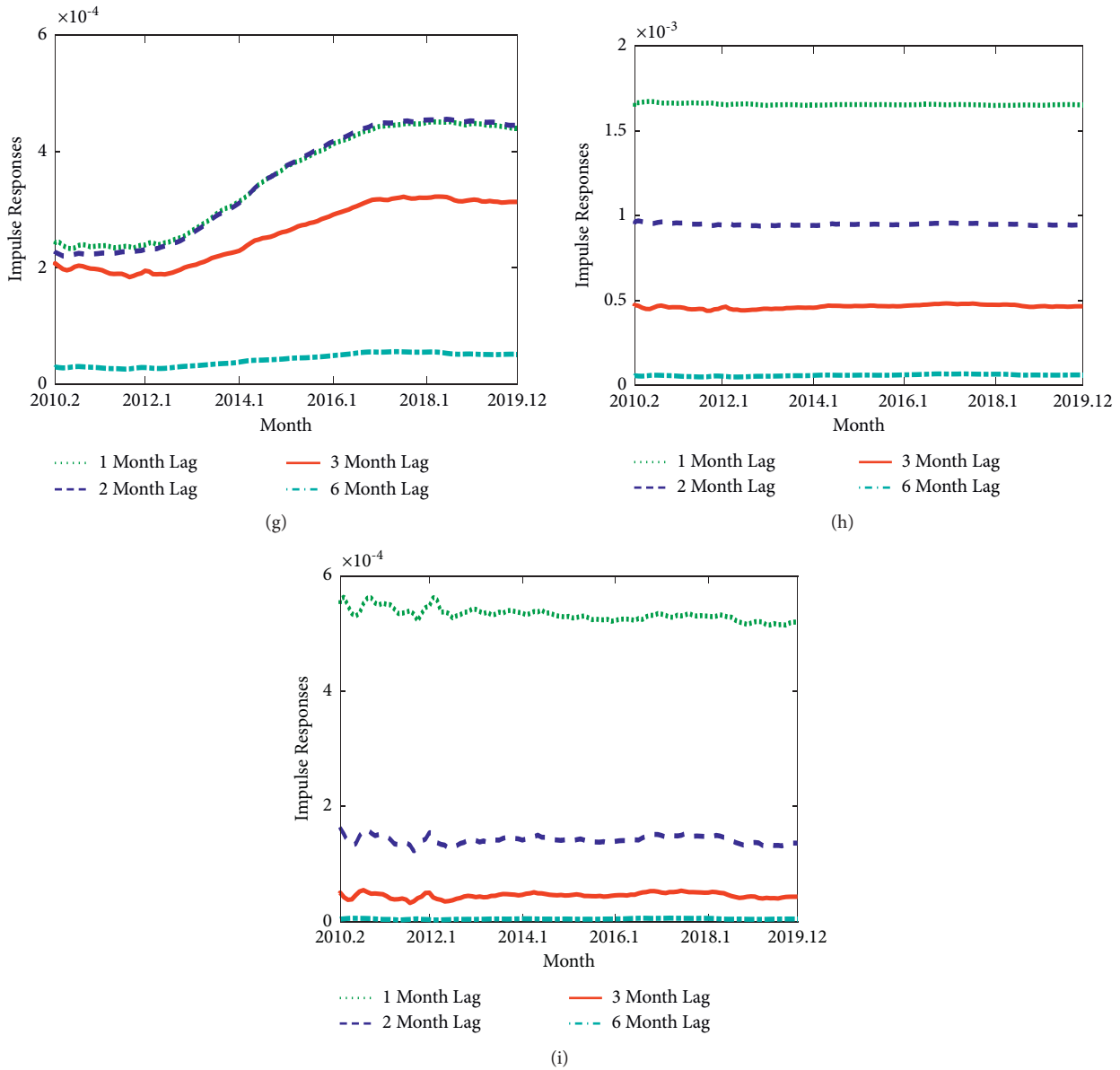
FIGURE 2: Time-varying impulse response between traffic volume of expressway passenger cars, traffic volume of trucks, and manufacturing PMI for different lag times. (a) Response C to C. (b) Response T to C. (c) Response P to C. (d) Response C to T. (e) Response T to T. (f) Response P to T. (g) Response C to P. (h) Response T to P. (i) Response P to P.

on truck traffic volume. The mutual relationship between the three variables is discussed below.

The response of P to C in Figure 2(c) has been positive for ten years, suggesting that expressway passenger car traffic volume has a positive impact on manufacturing PMI for a long time. For example, an increase in passenger car traffic volume on expressways will promote the development of manufacturing. Longitudinal analysis shows that the impact is highest when the lag is 1 month. As the time lags become larger, the impact gradually decreases and stabilizes. The impact reaches almost zero when the time lag becomes 6 months. The result indicates that the impact of expressway passenger car traffic on manufacturing PMI will gradually weaken as the lag time increases. Horizontal analysis of the

impulse response curve shows that the impact of expressway passenger traffic volume on the manufacturing PMI gradually weakened from 2010 to 2015. In 2016, the impact began to rebound to a higher and stable position. Overall, the positive promotion effect is still obvious although the increase in expressway passenger car traffic has a decline promotion effect on the development of manufacturing industry.

Response of P to T in Figure 2(f) has also been positive for ten years, suggesting that the increase in truck traffic volume on expressways will promote the development of manufacturing. Longitudinal analysis shows that the impact of expressway truck traffic volume on manufacturing PMI is basically similar to the passenger car traffic volume on

manufacturing PMI. The impact is highest when the time lag is 1 month, and then, the impact gradually decreases. When the lag enlarged to 6 months, the impulse response curve has been steadily dropped to around 0, which indicates the situation of no promotion effect. Horizontal analysis shows that fluctuations of impulse response curves for lags of 1 month, 2 months, and 3 months are quite similar. Different from passenger car traffic volume, the impact of truck traffic volume on manufacturing PMI has dropped significantly after 2018 due to the impact of transportation structure reform policies, e.g., "revolution to rail and water to revolution."

Response of C to P in Figure 2(g) shows a positive impact of manufacturing PMI on expressway passenger car traffic volume. That is, the increase in manufacturing PMI will reversely stimulate the increase in expressway passenger car traffic volume. For lags of 1 month and 2 months, sizes of the impulse response curve are basically the same. The reduction in impulse response is smaller when the lag is 3 months. The impulse response of manufacturing PMI to expressway passenger car traffic volume has gradually increased since 2012. The figure further shows the following. ① The development of manufacturing industry will reversely stimulate the increase of expressway passenger car traffic, and the stimulus lasts for a long time. The stable impact is about 2 months. ② In recent years, the development of manufacturing industry has gradually stimulated expressway passenger car traffic volume. The convenient and fast travel conditions provided by expressways make the role of "traffic first" stand out.

Response of T to P in Figure 2(h) shows a stable and an almost constant impulse response. As the lag increases, the increase in manufacturing PMI has significantly reduced the stimulus effect on expressway truck traffic volume. Such result means the development of manufacturing industry will promote the increase of expressway truck traffic volume. However, this promotion will not be affected by the degree of manufacturing development.

5. Conclusions

In this paper, the TVP-VAR model is applied to explore the time-varying interaction relationship among the three variables of expressway passenger car and truck traffic volume and manufacturing PMI. The results can help policy makers to better understand the mutual reinforcing relationship between China's Expressway traffic volume and manufacturing economy. Based on the mutual relationship, in-depth suggestions can be given to the development of China's transportation and new infrastructure construction. The results can also allow government to make corresponding policy measures for different development opportunities. Furthermore, this research is helpful for the understanding the role of free expressway pass policy in promoting the economic recovery of manufacturing industry during COVID-19.

Specifically, this article starts with China's expressway passenger car, truck traffic volume, and manufacturing PMI index data from 2010 to 2019. The TVP-VAR model is applied to analyze the time-varying interaction between the

three variables. The result shows that there is a long-term and stable interaction between the traffic volume of passenger cars and trucks on expressway and the manufacturing PMI in China. And, the interaction relationship is asymmetric. The increased traffic volume during COVID-19 freeway period is of great significance to the resumption of work and production in China. Other conclusions are briefly discussed as follows:

(1) There is a long-term positive interaction between expressway passenger car and truck traffic volume and manufacturing PMI. The increase in passenger car and truck traffic volume on expressways will promote the development of manufacturing. In turn, the development of manufacturing industry will also increase expressway traffic volume. This mutual promotion has a strong influence within 1 month. However, no interaction impact is shown for half-year lag.

(2) The increase in expressway passenger car traffic volume has continued to stimulate the development of manufacturing industry at a high level. The increase in the number of tourist trips is still significant in promoting the manufacturing economy. At the same time, under the influence of freight structure adjustment policies, the promotion of the development of manufacturing industry by expressway truck traffic volume has declined after 2018. The implementation effect of road transport structure adjustment policy is remarkable.

(3) In recent years, the stimulus effect of the development of manufacturing industry on expressway passenger car traffic volume has gradually increased. The economic development model led by domestic demand has gradually been ineffective. Domestic demand has gradually become the main driving force, development orientation, and support for economic growth. However, the stimulus effect of the development of manufacturing on expressway truck traffic has remained basically unchanged in past decade. That is, the structure of road freight transportation in China is relatively stable and reliable, and it is less affected by policy factors.

This article has some limitations. This article only considers the interaction between expressway passenger car and truck traffic volume and manufacturing PMI. Other relevant factors in the operation of expressways, including road network mileage, road congestion, and adaptive traffic volume, are ignored. However, considering these factors can give a much more complete and detailed understanding of the promotion effect of various indicators in the development of expressways on manufacturing economy, in future research, we will try to improve it.

Data Availability

The data on the traffic volume of passenger cars and trucks were collected by China National Traffic Survey Data Collection System, so they are not free. Requests for access to

these data should be made to the corresponding author (sunshuo@tpri.org.cn). The data on manufacturing PMI were collected by China National Bureau of Statistics. People can get these data from the official website.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] P. Li, "The influence of expressway on social and economic development," *China's Foreign Trade*, vol. 6, pp. 59–61, 2011.

[2] S. Kim, Z. Zafari, M. Bellanger, and P. A. Muennig, "Cost-effectiveness of capping freeways for use as parks: the New York cross-bronx expressway case study," *American Journal of Public Health*, vol. 108, no. 3, pp. 379–384, 2018.

[3] R. M. Xiao, B. Li, and Y. S. Chen, "Transportation status of Chinese expressway network in 2013," *Journal of Traffic and Transportation Engineering*, vol. 14, pp. 67–73, 2014.

[4] X. N. Long and X. Gao, "Transportation infrastructure and productivity of manufacturing firmseevidence from county-level expressway data and Chinese industrial enterprises database," *Journal of Huazhong Normal University(Humanities and Social Sciences)*, vol. 53, pp. 43–52, 2014.

[5] O. Kalan and I. Gokasar, "A dynamic panel data approach for the analysis of the growth impact of highway infrastructures on economic development," *Modern Economy*, vol. 11, no. 03, pp. 726–739, 2020.

[6] J. B. Mills and J. D. Fricker, "Spatial panel econometric analysis of economic impacts of bypasses," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2242, no. 1, pp. 122–133, 2011.

[7] Y. H. Ji and J. Wen, "Research on the influence of transportation infrastructure on manufacturing agglomeration in China," *Development and Research*, vol. 10, pp. 39–43, 2016.

[8] W. P. Bai, Z. Lu, and L. P. Liu, "Foreign direct Investment, Transportation infrastructure improvement and manufacturing agglomeration -- empirical research on panel data of 285 cities in China from 2003 to 2016," *Journal of Guizhou University of Finance and Economics*, vol. 199, pp. 16–28, 2019.

[9] P. N. Patatoukas and A. Skabardonis, "Traffic volume and aggregate economic activity: implications for taking the pulse of the U.S. Economy. Transportation research record," *Journal of the Transportation Research Board*, 2016.

[10] N. Caceres, L. M. Romero, F. J. Morales, A. Reyes, and F. G. Benitez, "Estimating traffic volumes on intercity road locations using roadway attributes, socioeconomic features and other work-related activity characteristics," *Transportation*, vol. 45, no. 5, pp. 1449–1473, 2018.

[11] H. Xiong and Y. W. Sun, "Analysis of Co-integration and causality between freight traffic volume and economic growth in China," *Systems Engineering*, vol. 28, pp. 50–54, 2010.

[12] A. M. Sha, "Discussion on countermeasures and intelligent operation mode of division construction for the expressway passenger and freight transport system," *China Journal of Highway and Transport*, vol. 33, pp. 1–7, 2020.

[13] B. B. Yan, *Application of Sign Restriction on TVP-VAR Model in Loan Supply and Demand Shocks in China*, Huazhong University, Huazhong, China, 2013.

[14] C. A. Sims, "Macroeconomics and reality," *Econometrica*, vol. 48, no. 1, pp. 1–48, 1980.

[15] O. Blanchard and M. Watson, *Are Business Cycles All Alike?*, University of Chicago Press, Chicago, IL, USA, 1986.

[16] G. E. Primiceri and E. Giorgio, "Time varying structural vector autoregressions and monetary policy," *The Review of Economic Studies*, vol. 72, no. 3, pp. 821–852, 2005.

[17] J. Nakajima, M. Kasuya, and T. Watanabe, "Bayesian analysis of time-varying parameter vector autoregressive model for the Japanese economy and monetary policy," *Journal of the Japanese and International Economies*, vol. 25, no. 3, pp. 225–245, 2011.

[18] B. Pfaff, "VAR, SVAR and SVEC models: implementation within R package vars," *Stat. Softw*, vol. 27, pp. 1–32, 2008.

[19] R. F. Engle and M. W. Watson, "The kalman filter model: applications to forecasting and rational expectations," *Advances in Econometrics: Fifth World Congress of Econometric Society*, vol. 1, pp. 245–283, 1987.

[20] G. Koop and D. Korobilis, "Bayesian multivariate time series methods for empirical macroeconomics," *Foundations and Trends(R) in Econometrics*, vol. 3, 2010.

[21] W. Y. Yap and J. S. L. Lam, "Competition dynamics between container ports in east asia," *Transportation Research Part A: Policy and Practice*, vol. 40, no. 1, pp. 35–51, 2006.

[22] S. Johansen, "Statistical analysis of cointegration vectors," *Journal of Economic Dynamics and Control*, vol. 12, no. 2-3, pp. 231–254, 1998.

WILEY | Hindawi

Research Article

Method of Evaluating and Predicting Traffic State of Highway Network Based on Deep Learning

Jiayu Liu [iD],[1] **Xingju Wang** [iD],[1,2] **Yanting Li** [iD],[1] **Xuejian Kang** [iD],[1,2] **and Lu Gao** [iD][3]

[1]*School of Traffic and Transportation, Shijiazhuang Tiedao University, Shijiazhuang, China*
[2]*State Key Laboratory of Mechanical Behavior and System Safety of Traffic Engineering Structures,*
 Shijiazhuang Tiedao University, Shijiazhuang, China
[3]*Department of Construction Management, University of Houston, 4730 Calhoun Road No. 300 Houston, Houston,*
 TX 77204-4021, USA

Correspondence should be addressed to Xingju Wang; wangxingju@stdu.edu.cn

The accurate evaluation and prediction of highway network traffic state can provide effective information for travelers and traffic managers. Based on the deep learning theory, this paper proposes an evaluation and prediction model of highway network traffic state, which consists of a Fuzzy C-means (FCM) algorithm-based traffic state partition model, a Long Short-Term Memory (LSTM) algorithm-based traffic state prediction model, and a K-Means algorithm-based traffic state discriminant model. The highway network in Hebei Province is employed as a case study to validate the model, where the traffic state of highway network is analyzed using both predicted data and real data. The dataset contains 536,823 pieces of data collected by 233 continuous observation stations in Hebei Province from September 5, 2016, to September 12, 2016. The analysis results show that the model proposed in this paper has a good performance on the evaluation and prediction of the traffic state of the highway network, which is consistent with the discriminant result using the real data.

1. Introduction

With the increasing of vehicle ownership, the demand for transportation has also expanded rapidly, which not only brings great pressure to the highway traffic system, but also causes a series of congestion problems and safety issues. The 2019 Hebei Statistical bulletin on national economic and social development indicates that the mileage open to traffic increased by 1.9% in the province, but the vehicle ownership increased by 7.4% year-over-year [1]. Therefore, it is particularly essential to be able to evaluate and predict the traffic state of the highway network accurately in real time. Traffic managers can develop prevention strategies in advance to effectively avoid the occurrence of traffic congestion and prevent the further expansion of the impact of congestion. How to realize the accurate evaluation and prediction of traffic state is the most important task faced by traffic managers and traffic problem researchers.

Appropriate indicators need to be selected for the evaluation of traffic state, mainly including service level [2], traffic flow parameters [3–5], and travel time [6]. In addition, a large number of scholars have studied the evaluation method of traffic state. The existing researches on the traffic state evaluation of road primarily focus on the intersection [7], road [8], and road network [9, 10]. For transportation node, most studies focus on urban roads and use major intersections as the research object. Xu et al. have conducted in-depth research on the evaluation of road traffic state. They used various theories and methods to estimate and predict the road traffic state [11–14]. In addition, D.W. Xu et al. proposed a novel deep learning framework for estimating road traffic state [15]. As the development of big data [16] and machine learning [17] advances, more methodologies will be applied to the evaluation of highway network traffic state. B. Hillel used cellular phone service provider information to measure traffic speed and travel time [18]. Based

on a large amount of floating car data, G. Fusco et al. made short-term forecasts of traffic speed and carried out comparative analysis using a variety of methods [19]. A. Wu et al. adopted floating car data to identify the traffic state of signalized intersections [20]. M. Wang et al. used a large cell phone network data to estimate traffic flow [21]. A large number of loop detectors are installed in the highway network to collect traffic information. These data can also be used to analyze traffic state [22, 23]. In addition, based on machine learning theory, neural networks have been widely applied to traffic flow prediction [24, 25]. J. Yu et al. presented a graph convolutional generative autoencoder model to estimate traffic speed in real time [26]. I. Laña et al. presented a method with an adaptation mechanism for long-term urban traffic volume forecasting [27].

The previously related research is mainly divided into two aspects including predicting traffic flow parameters, such as traffic volume, density, speed, and saturation, and evaluating the road traffic state. Little attention is paid to the evaluation and prediction of road network traffic state. Taking highway network as the research object, this study aims to conduct the evaluation and prediction model of highway network traffic state based on deep learning. The evaluating and predicting traffic state of highway network can comprehensively reflect the traffic condition of the entire highway network. This study could be useful for future policy making.

2. Methods

The traffic state of the highway network is divided into four categories, namely, severe congestion state, congestion state, general state, and smooth state. The speed and density of traffic flow are selected as evaluation indicators to develop the proposed evaluation and prediction model of highway network traffic state based on deep learning. The model mainly includes the partition model of traffic state based on the Fuzzy C-means (FCM) algorithm, the prediction model of traffic state based on the Long Short-Term Memory (LSTM) algorithm, and the discriminant model of traffic state based on K-Means algorithm.

2.1. The Partition Model of Highway Network Traffic State Based on FCM Algorithm. The traffic state of the highway network has certain fuzziness. The speed and density of traffic flow are used as evaluation indicators, and the FCM algorithm is used to evaluate the traffic state of the highway network. The FCM algorithm can well analyze the attribute characteristics of big data in practical problems, and it uses the membership function to cluster [28].

The FCM algorithm divides the n vectors $\mathbf{x}_i (i = 1, 2, \ldots, n)$ into c classes and calculates the cluster centers for each class. For $\forall \mathbf{x}_i = (v_i, k_i), v_i, k_i$ represent the speed and density of the *ith* sample, respectively, and $1 < c < n$. The objective function defined by the membership function is as follows:

$$\min\{J(\mathbf{U}, \mathbf{A})\} = \min \left\{ \sum_{j=1}^{c} \sum_{i=1}^{n} (u_{ij})^m (d_{ij})^2 \right\},$$

$$s.t. \begin{cases} 0 \le u_{ij} \le 1, \\ \sum_{j=1}^{c} u_{ij} = 1, \end{cases} \quad (1)$$

where $J(\mathbf{U}, \mathbf{A})$ is the weighted value of the sample to each cluster center of the highway network traffic state; u_{ij} is the membership function of the *ith* sample for the j class, and $1 \le j \le c, 1 \le i \le n$; $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_c]$ is the $2 \times c$ order cluster center matrix; $\mathbf{a}_j$ is the eigenvector of the *jth* cluster center of the highway network traffic state and $\mathbf{a}_j = (v_j, k_j)$; m is the fuzzy weighted index; $(d_{ij})^2 = \|\mathbf{x}_i - \mathbf{a}_j\|$ is the Euclidean distance from the *ith* sample to $\mathbf{a}_j$; and $\mathbf{U} = \{u_{ij}\}$ is the $c \times n$ order fuzzy classification matrix.

Construct the following formula by using Lagrange multiplication:

$$F = \sum_{j=1}^{c} (u_{ij})^m (d_{ij})^2 + \lambda \left(\sum_{i=1}^{c} u_{ij} - 1 \right). \quad (2)$$

Taking the derivatives for all input variables, we can get the following formula:

$$u_{ij} = \frac{1}{\sum_{r=1}^{c} (d_{ij}/d_{kj})^{2/m-1}},$$

$$\mathbf{a}_j = \frac{\sum_{k=1}^{n} (u_{jk})^m x_k}{\sum_{k=1}^{n} (u_{jk})^m}. \quad (3)$$

2.2. Algorithm Implementation

(1) Initialization. Give the iterative standards $\varepsilon > 0$, set b as the number of iterations, and let $b = 0$. The membership matrix is calculated as

$$u_{ij}^{(b)} = \frac{1}{\sum_{r=1}^{c} (d_{ij}^{(b)}/d_{rj}^{(b)})^2}. \quad (4)$$

If i, r exist, so that $d_{ir}^{(b)} = 0$, then there is $u_{ir}^{(b)} = 1$, and $j \ne r, u_{ir}^{(b)} = 0$.

(2) The iterative operation of the cluster center is

$$\mathbf{a}_j^{(b+1)} = \frac{\sum_{i=1}^{n} (u_{ij}^{(b)})^2 \mathbf{x}_i}{\sum_{i=1}^{n} (u_{ij}^{(b)})^2}, \quad j = 1, 2, \ldots, c. \quad (5)$$

(3) Use a matrix norm $\| \cdot \|$ to compare $\mathbf{a}^{(b+1)}$ and $\mathbf{a}^{(b)}$; if $\| \mathbf{a}^{(b+1)} - \mathbf{a}^{(b)} \| \le \varepsilon$, stop iteration; otherwise, let $b = b + 1$, and then turn to the second step.

(4) The cluster center matrix of the highway network traffic state is finally obtained as $\mathbf{A} = [\mathbf{a}_1, \ldots, \mathbf{a}_j, \ldots, \mathbf{a}_c]$.

The driving speed and density of the vehicle on the highway network can well reflect the traffic state of the highway network. The faster the driving speed of the vehicle, the smaller the density, the better the highway condition, and the smoother the traffic state of the highway network. According to the above algorithm, the cluster center matrix and the classification result of the traffic state of the highway network can be obtained.

2.3. The Prediction Model of Highway Network Traffic State Based on LSTM Algorithm.

LSTM is a long-term and short-term memory neural network. The model can be trusted for a long time, which is suitable for processing and predicting time series data and important events with relatively long delays. It has been used in many fields of science and technology. The model training methods proposed in this paper include forward calculation using activation function and backward estimation using the backpropagation through time (BPTT) according to the loss function [29].

The data are preprocessed by Z-score normalization method before training. The processed function has a mean value of 0 and a standard deviation of 1 as shown in (6). After outputting the predicted result, the data need to be anti-normalized to obtain the true predicted values of the traffic state evaluation indicators. The antinormalization formula is shown in (7):

$$x^* = \frac{x - \mu}{\sigma}, \tag{6}$$

$$x = x^* * \sigma + \mu, \tag{7}$$

where μ is the mean of all sample data, and σ is the standard deviation of all sample data.

In order to verify the accuracy of the prediction result and the feasibility of the prediction model, it is necessary to perform error analysis on this result. The relative error formula of each predicted point is shown in (8). The Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE) of the highway or the highway network are calculated using (9) and (10), respectively.

$$\varepsilon_i = \frac{y_i^* - y_i}{y_i}, \tag{8}$$

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^{N} \frac{|y_i^* - y_i|}{y_i} \times 100\%, \tag{9}$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i^* - y_i)^2}, \tag{10}$$

where ε_i is the relative error between the predicted value and the real value of the *ith* sample data, y_i is the real value of the *ith* sample data, y_i^* is the predicted value of the *ith* sample data, and N is the number of sample data of the traffic state evaluation indicators.

2.4. The Discriminant Model of Highway Network Traffic State Based on K-Means Algorithm.

The K-Means algorithm is an unsupervised clustering algorithm, and the basic idea of the algorithm is that, for a given sample set, it is divided into multiple clusters according to the distance between the samples. Using the Euclidean distance as the measure of similarity, the shortest distance between each sample and each cluster center can be found. Then, we can know the traffic state of each sample. The specific formula is as follows:

$$G(x_p) = \min_{k=1}^{c} \left\{ D\left((v_p, k_p), (v_{a_j}, k_{a_j}) \right) \right\}, \tag{11}$$

where

$$D\left((v_p, k_p), (v_{a_j}, k_{a_j}) \right) = \sqrt{\beta_1 \left(v_p - v_{a_j} \right)^2 + \beta_2 \left(k_p - k_{a_j} \right)^2}, \tag{12}$$

where x_p is the *pth* highway traffic state prediction sample; v_p and k_p are, respectively, the speed and density parameters of the *pth* highway traffic state prediction sample; a_j is the cluster center of the *jth* class of the highway traffic state; v_{a_j} and k_{a_j} are the speed and density of a_j, respectively; β_1 and β_2 are the weight of the influence of speed and density, respectively, on the classification of the traffic state, and this paper considers the influence as the same, namely, $0 \leq \beta_1 \leq 1, 0 \leq \beta_2 \leq 1, \beta_1 + \beta_2 = 1$.

In order to evaluate the traffic state of the highway network, the average speed and density of all highways need to be calculated. Based on the K-means algorithm, the average traffic state of the highway network can be obtained.

3. Case Study

3.1. Dataset.

The dataset is collected by 233 continuous observation stations from September 5, 2016, to September 12, 2016. The original dataset has certain quality problems, so the dataset is cleaned and processed. The dataset is the comprehensive traffic flow information by direction, lane, and vehicle mode. The observation interval is set to be five minutes, and each observation station collects 288 pieces of data per day. Thus, 536,832 pieces of data are collected from 233 observation stations. The layout of the observation stations in the highway network is shown in Figure 1. The main attribute information of the observation station is shown in Table 1. The data for one day the at Babaihu observation station are listed in Table 2. There is an observation station on each highway, and the observation data of the observation station represent the traffic information of the highway on which it is located.

3.2. Partition of Highway Network Traffic State.

The highway network is composed of different highways, and each highway has its specific design speed. Therefore, before using the FCM algorithm for clustering, the traffic flow speed and density of the highway network should be standardized. The standardized formula is as follows:
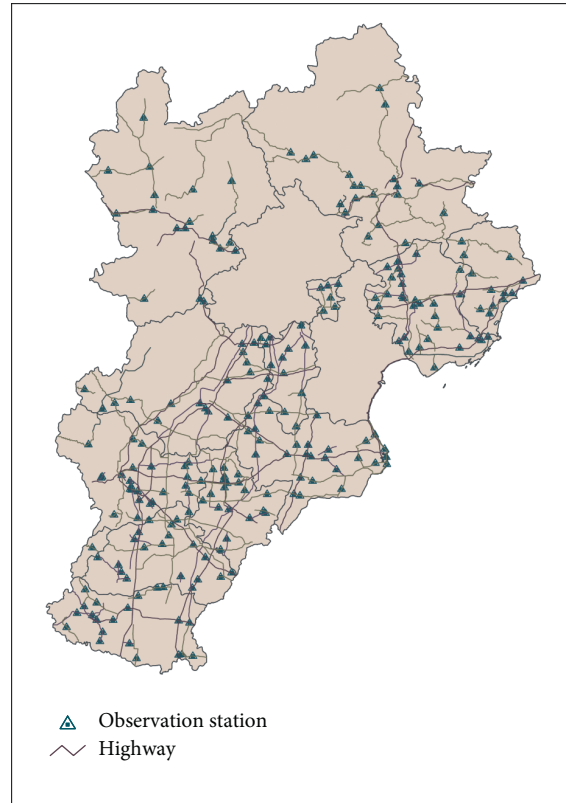
FIGURE 1: The layout of the observation stations in the highway network.

TABLE 1: The main attribute information of the observation station.

Index	Year	Observation station name	Observation station type	Investigation method	Design speed (km/h)
1	2016	Qingxianzhuxianzhan	Continuous	Equipment observation	120
2	2016	Cangzhoubei	Continuous	Equipment observation	120
3	2016	Cangzhouxi	Continuous	Equipment observation	120
4	2016	Xiadian	Continuous	Equipment observation	80
5	2016	Dadingfu	Continuous	Equipment observation	80
6	2016	Liangjiadian	Continuous	Equipment observation	80
. . .	. . .	. . .	. . .	. . .	. . .
228	2016	Pianqiaozi	Continuous	Equipment observation	80
229	2016	Dongying	Continuous	Equipment observation	80
230	2016	Taipingzhuang	Continuous	Equipment observation	60
231	2016	Mujiangkou	Continuous	Equipment observation	40
232	2016	Yeheqiaoyanghuzhongxin	Continuous	Equipment observation	40
233	2016	Erbaozi	Continuous	Equipment observation	30

TABLE 2: Babaihu observation station data.

Index	Observation station	Hour	Minute	Flow volume [pcu]	Speed (km/h)	Density[pcu/km]
1	Babaihu	1	0	8	25.0000	10.4800
2	Babaihu	1	5	41	24.8750	10.5327
3	Babaihu	1	10	21	24.8333	10.5503
4	Babaihu	1	15	25	22.3333	11.7313
5	Babaihu	1	20	22	22.5000	11.6444
6	Babaihu	1	25	16	22.3750	11.7095
. . .	. . .	. . .	. . .	. . .	. . .	. . .
283	Babaihu	24	30	17	24.4000	11.5984
284	Babaihu	24	35	7	22.7500	12.4396
285	Babaihu	24	40	14	22.0000	12.8636
286	Babaihu	24	45	38	23.5000	12.0426
287	Babaihu	24	50	12	26.0000	10.8846
288	Babaihu	24	55	39	27.2500	10.3853
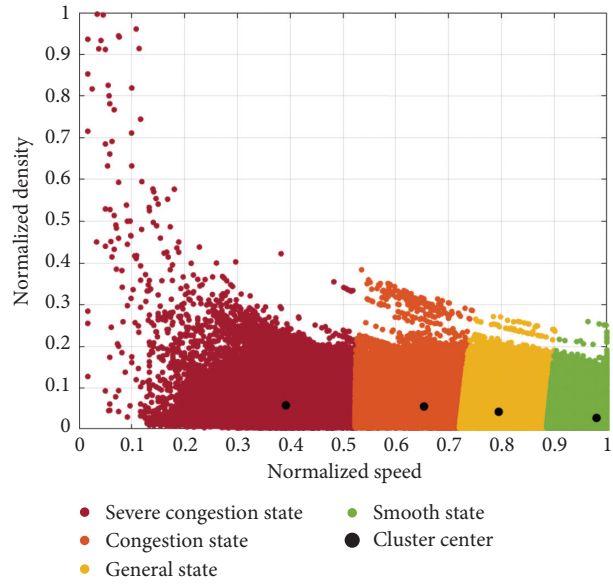
FIGURE 2: Highway network classification result.
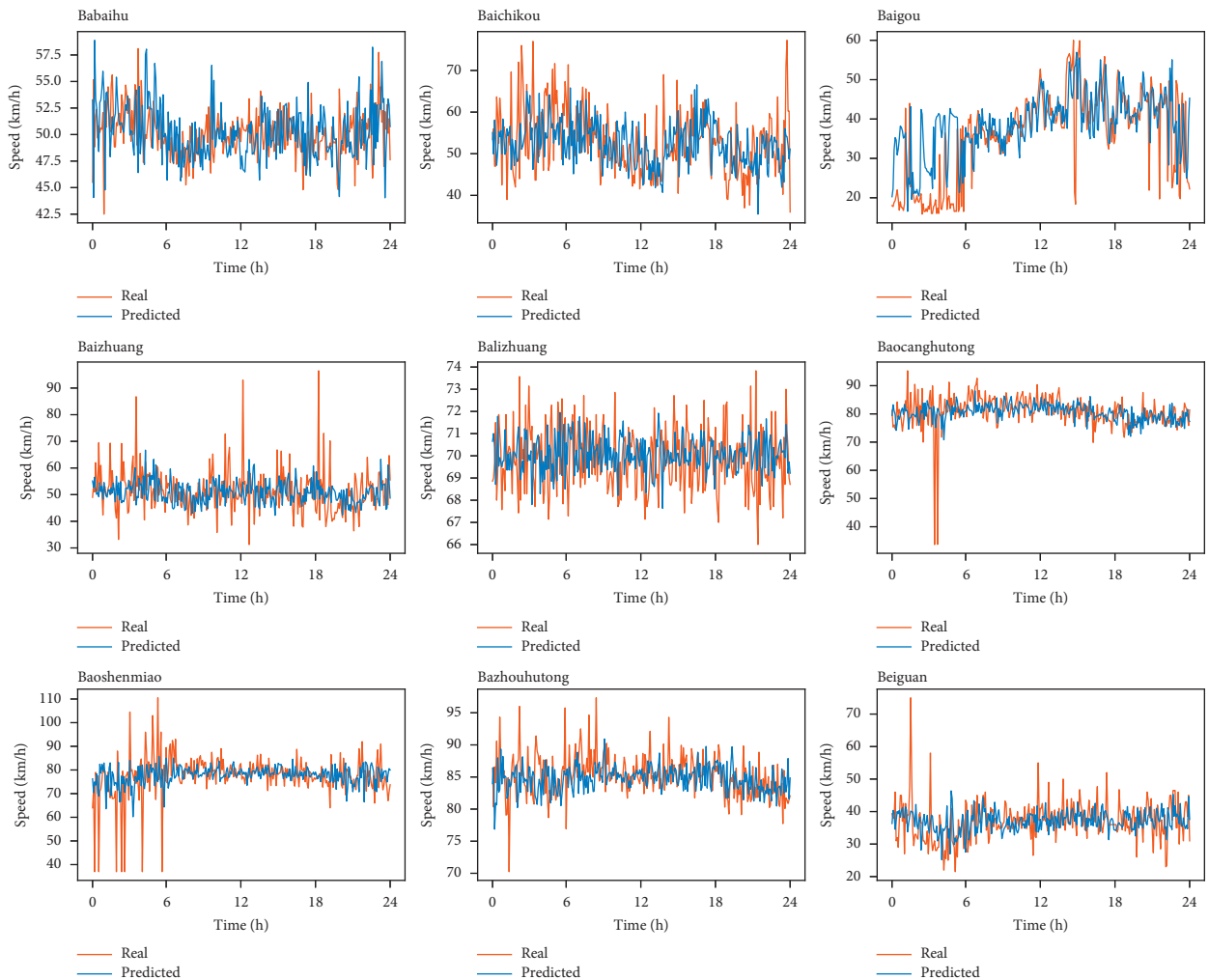
FIGURE 3: Comparison of predicted speed and true speed in 9 observation stations.
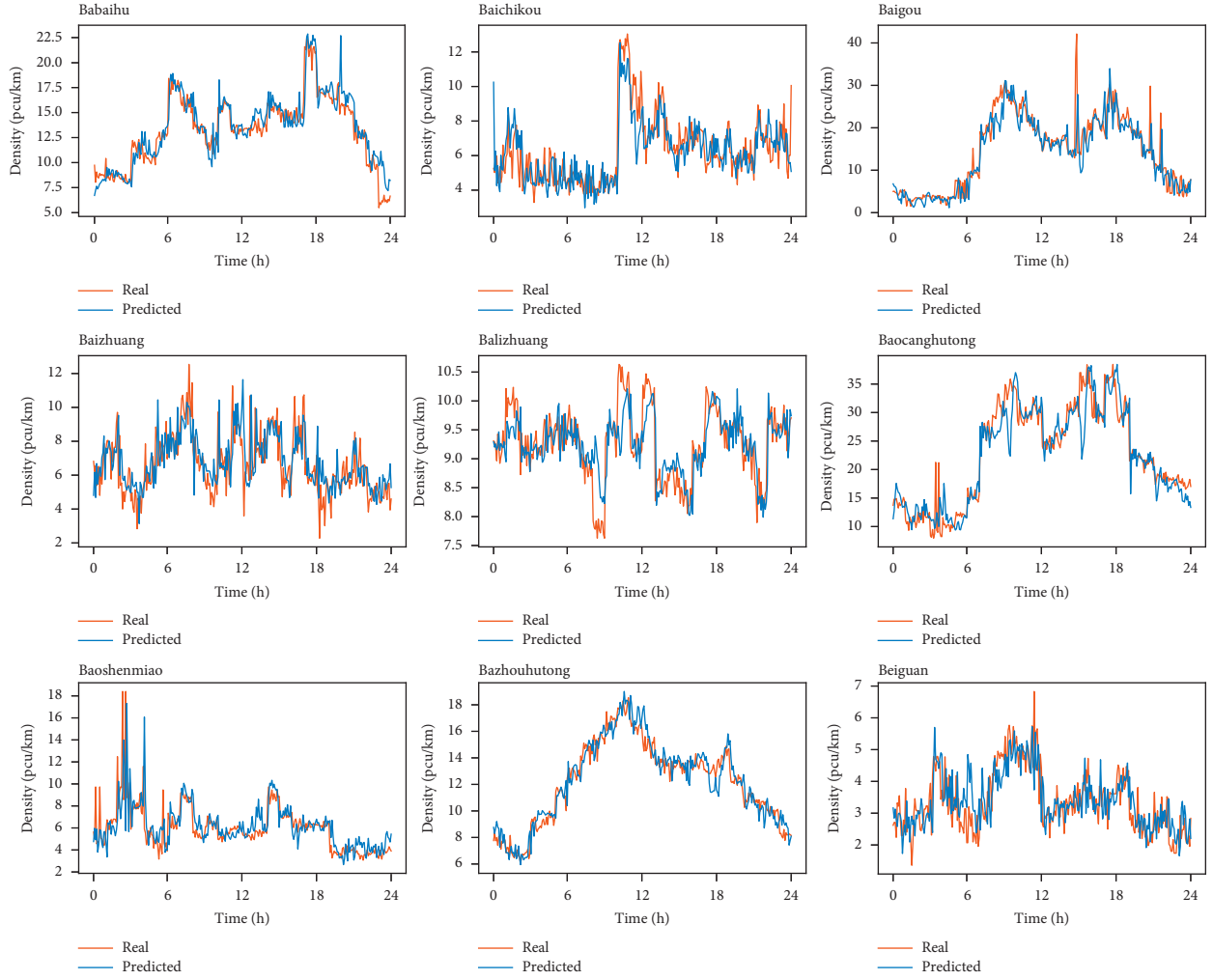
FIGURE 4: Comparison of predicted density and true density in 9 observation stations.

$$v_i' = \frac{v_i - v_{min}}{v_{max} - v_{min}}, \tag{13}$$

$$k_i' = \frac{k_i - k_{min}}{k_{max} - k_{min}}. \tag{14}$$

When using FCM algorithm to classify traffic state of the highway network, the model parameters need to be selected first. As discussed previously, the number of clusters $c = 4$. For fuzzy weighted index m, numerous studies have been conducted, and it is widely accepted that, for general classification, $m = 2$ is rational [30]. Speed and density data collected by 233 observation stations from September 5, 2016, to September 11, 2016, are studied. The collected dataset contains 469,728 pieces of data. The calculated cluster center is shown below.

$$A = \begin{bmatrix} 0.3919 & 0.0584 \\ 0.6535 & 0.0558 \\ 0.7951 & 0.0430 \\ 0.9804 & 0.0282 \end{bmatrix}. \tag{15}$$

The first column of matrix A shows the normalized speed value of the cluster centers. The second column shows the normalized density value of the cluster centers. Since the faster the speed, the smaller the density, and the better the traffic, the four rows in matrix A show cluster centers with severe congestion state, congestion state, general state, and smooth state from top to bottom, respectively. The clusters found by FCM are shown in Figure 2.

3.3. Prediction of Highway Network Traffic State. The LSTM algorithm is utilized to develop the short-term prediction. The data from September 5, 2016, to September 11, 2016, are the training set and the data from September 12, 2016, are the testing set. The speed and density of each observation station are predicted separately, and the prediction errors are calculated. The comparison of predicted speed and true speed in some observation stations is shown in Figure 3. And the comparison of predicted density and true density in some observation stations is shown in Figure 4. According to (9) and (10), the MAPE and RMSE of the speed and density of 233 observation stations are obtained. The resulting errors
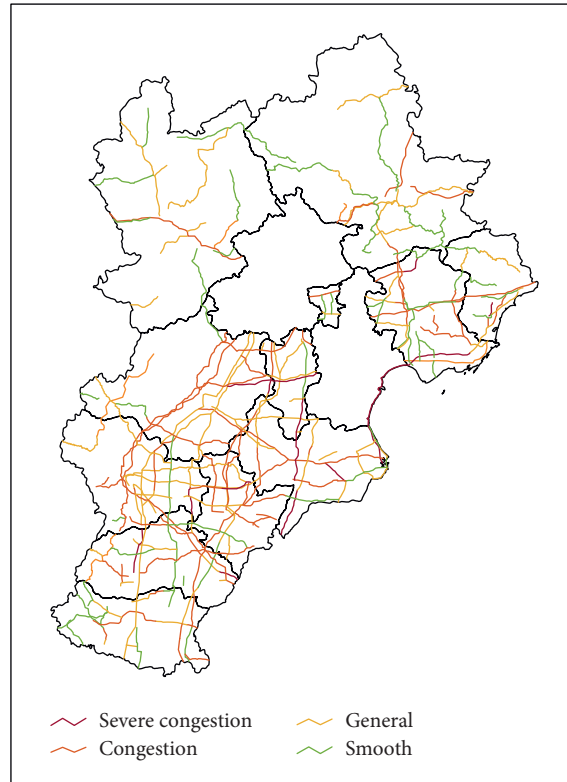
FIGURE 5: The predicted average traffic state of each highway.

are extremely small, indicating that the predicted value is relatively accurate.

$$
\begin{aligned}
\text{MAPE (speed)} &= 8.7892, \\
\text{RMSE (speed)} &= 0.1264, \\
\text{MAPE (density)} &= 2.0948, \\
RMSE\text{ (density)} &= 0.1824.
\end{aligned}
\tag{16}
$$

3.4. Discriminant of Highway Network Traffic State. The speed and density of each highway are determined by the observations of the observation station where it is located. Before judging the traffic state of each highway, the predicted speed and density values should be normalized according to (13) and (14). The traffic state of each highway is discriminated based on the K-means algorithm. Then, calculating the average speed and density of all highways can obtain the average traffic state of the highway network. The discriminant result of the average traffic state of each highway is shown in Figure 5. And the normalized average predicted speed and density of the highway network are 0.7719 and 0.0467, respectively. Therefore, according to the proposed evaluation and prediction model of highway network traffic state based on deep learning, on September 12, 2016, the average traffic state of the highway network constituted by 233 highways is in general state.
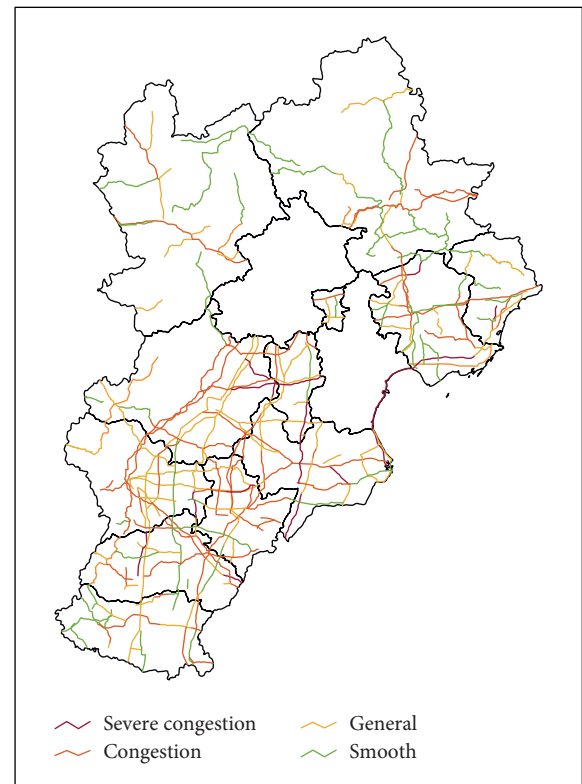
FIGURE 6: The real average traffic state of each highway.

TABLE 3: The proportion of each traffic state category.

Traffic state percentage	Severe congestion state (%)	Congestion state (%)	General state (%)	Smooth state (%)
The proportion of each predicted traffic state	5.58	30.47	38.32	25.75
The proportion of each real traffic state	6.01	30.90	37.77	25.32

3.5. Analysis of Evaluation Results. To prove the accuracy of the proposed model, the true average traffic state of the highway network on September 12, 2016, needs to be analyzed. First, the true average speed and density of each highway on September 12, 2016, are calculated and normalized, and then the average traffic state of each highway is discriminated, which is shown in Figure 6. The true normalized average speed and density of the highway network are 0.7710 and 0.0471, respectively. And the predicted and true traffic state are consistent, and both are in general state. The predicted and real average traffic states of each highway are compared, and the proportion of each traffic state in the highway network is shown in Table 3. The results show that congestion state and general state account for the largest proportion. About 6% of highways are in severe congestion state, indicating that the speed of vehicles is restricted to a certain extent. The operating conditions are not good, and the highway is in high pressure. Such problem could be caused by the imbalance between supply and demand of existing highway infrastructure and traffic demand caused. The traffic management departments should pay attention to the reconstruction and development of these highways. Approximately 30% of the highways are in congestion state, and these highways should also be focused on. Reasonable traffic organization plans and control measures should be formulated in advance to alleviate traffic congestion. When the highways are in general state and smooth state, it means that the vehicles can travel at a faster speed and are in good operating condition. And in the highway network, more than 50% of the highways are not congested. It shows that the overall layout of the highway network is relatively reasonable, and the traffic condition is good. The evaluation and prediction of the traffic state of the highway network can play a role in prevention and planning in advance.

4. Conclusions

This paper proposes an evaluation and prediction model of highway network traffic state based on deep learning. The model consists of a FCM algorithm-based traffic state partition model, a LSTM algorithm-based traffic state prediction model, and a K-Means algorithm-based traffic state discriminant model. In this study, the traffic state of the highway network is divided into four categories: severe congestion state, congestion state, general state, and smooth state. The speed and density of traffic flow are selected as evaluation indicators. To validate the effectiveness of the proposed model, the traffic flow data of 233 observation stations in Hebei Province from September 5, 2016, to September 12, 2016, are collected. The observation interval is

5 minutes, and the total of data is 536,823 pieces. The main findings are concluded as follows:

On September 12, 2016, the predicted average traffic state of the highway network constituted by 233 highways is in general state, which is consistent with the real traffic state. The overall layout of the highway network is relatively reasonable, and the supply and demand of transportation are basically balanced.

The highway network traffic state is an objective reflection of the traffic condition. The evaluation and prediction of highway network traffic state provide the dynamic characteristic of traffic change. Based on this information, travelers can make their travel plans better and improve their travel efficiency. Traffic managers can also work out better traffic organization plans in advance to reduce traffic congestion.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] Hebei Bureau of Statistics, http://www.hetj.gov.cn/hetj/app/tjgb/101591084423716.html Hebei 2019 Provincial Economy and Social Development Statistical Report, 2019.

[2] A. C. Tay and H. H. Lee, "Traffic condition with road upgrading during construction and operation stages based on level-of-service (LOS)," *IOP Conference Series Materials Science and Engineering*, vol. 344, no. 1, Article ID 012018, 2018.

[3] S. Toru, M. B. Alexandre, K. Takahio et al., "Traffic state estimation on highway: a comprehensive survey," *Annual Reviews in Control*, vol. 43, pp. 128–151, 2017.

[4] H.-Y. Yao, J. Wang, Y. Chen, Q. Tu, and B. Cheng, "Evaluation method for traffic operation status on ordinary highway roads," *Journal of Highway and Transportation Research and Development (English Edition)*, vol. 10, no. 2, pp. 73–78, 2016.

[5] A. Zeroual, F. Harrou, Y. Sun et al., "Road traffic density estimation and congestion detection with a hybrid observer-based strategy," *Sustainable Cities and Society*, vol. 46, 2019.

[6] E. Ugnenko, E. Uzhvieva, Y. Voronova et al., "Simulation of traffic flows on the road network of urban area," *Procedia Engineering*, vol. 134, pp. 153–156, 2016.

[7] H.-T. Zhao, X.-R. Liu, X.-X. Chen, and J.-C. Lu, "Cellular automata model for traffic flow at intersections in internet of vehicles," *Physica A: Statistical Mechanics and Its Applications*, vol. 494, pp. 40–51, 2018.

[8] Y. Kawasaki, Y. Hara, M. Kuwahara et al., "Traffic state estimation on a two-dimensional network by a state-space model," *Transportation Research Part C: Emerging Technologies*, vol. 113, pp. 176–192, 2020.

[9] A. Jayasinghe, K. Sano, C. C. Abenayake, and P. K. S. Mahanama, "A novel approach to model traffic on road segments of large-scale urban road networks," *MethodsX*, vol. 6, pp. 1147–1163, 2019.

[10] G. Zhu, K. Song, P. Zhang, and L. Wang, "A traffic flow state transition model for urban road network based on Hidden Markov Model," *Neurocomputing*, vol. 214, no. 19, pp. 567–574, 2016.

[11] D.-W. Xu, H.-H. Dong, L.-M. Jia, and Y. Tian, "Road traffic states estimation algorithm based on matching of regional traffic attracters," *Journal of Central South University*, vol. 21, no. 5, pp. 2100–2107, 2014.

[12] D. W. Xu, H. H. Dong, H. J. Li, L. M. Jia, and Y. J. Feng, "The estimation of road traffic states based on compressive sensing," *Transportmetrica B: Transport Dynamics*, vol. 3, no. 2, pp. 131–152, 2015.

[13] D.-W. Xu, Y.-D. Wang, L.-M. Jia, Y. Qin, and H.-H. Dong, "Real-time road traffic state prediction based on ARIMA and Kalman filter," *Frontiers of Information Technology & Electronic Engineering*, vol. 18, no. 2, pp. 287–302, 2017.

[14] D. Xu, Y. Wang, P. Peng, S. Beilun, Z. Deng, and H. Guo, "Real-time road traffic state prediction based on kernel-KNN," *Transportmetrica A: Transport Science*, vol. 16, no. 1, pp. 104–118, 2018.

[15] D. Xu and P. Peng, "A novel deep learning framework for road traffic state estimation," *Transportation Research Part C: Emerging Technologies*, vol. 117, 2020.

[16] S. Amini, I. Gerostathopoulos, C. Prehofer et al., "Big data analytics architecture for real-time traffic control," in *Proceedings of the IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pp. 710–715, Naples, Italy, June 2017.

[17] J. Wang, R. Chen, and Z. He, "Traffic speed prediction for urban transportation network: a path based deep learning approach," *Transportation Research Part C: Emerging Technologies*, vol. 100, pp. 372–385, 2019.

[18] H. Bar-Gera, "Evaluation of a cellular phone-based system for measurements of traffic speeds and travel times: a case study from Israel," *Transportation Research Part C: Emerging Technologies*, vol. 15, no. 6, pp. 380–391, 2007.

[19] G. Fusco, C. Colombaroni, N. Isaenko et al., "Short-term speed predictions exploiting big data on large urban road networks," *Transportation Research Part C: Emerging Technologies*, vol. 73, pp. 183–201, 2016.

[20] A. Wu, W. Yin, X. Yang et al., "Research on the real-time traffic status identification of signalized intersections based on floating car data," *Procedia - Social and Behavioral Sciences*, vol. 96, pp. 1578–1584, 2013.

[21] M.-H. Wang, S. D. Schrock, and T. Mulinazzi, "Estimating dynamic origin-destination data and travel demand using cell phone network data," *International Journal of Intelligent Transportation Systems Research*, vol. 11, no. 2, pp. 76–86, 2013.

[22] B. Li, "Recursive estimation of average vehicle time headway using single inductive loop detector data," *Transportation Research Part B: Methodological*, vol. 46, no. 1, pp. 85–99, 2012.

[23] J. Li, H. J. Van Zuylen, G. Wei et al., "Diagnosing and interpolating loop detector data errors with probe vehicle data," *Transportation Research Record*, vol. 2423, pp. 61–67, 2014.

[24] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 187–197, 2015.

[25] Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, and J. Liu, "LSTM network: a deep learning approach for short-term traffic forecast," *Iet Intelligent Transport Systems*, vol. 11, no. 2, pp. 68–75, 2017.

[26] J. J. Q. Yu and J. Gu, "Real-time traffic speed estimation with graph convolutional generative autoencoder," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 10, pp. 3940–3951, 2019.

[27] I. Laña, J. L. Lobo, E. Capecci, J. Del Ser, and N. Kasabov, "Adaptive long-term traffic state estimation with evolving spiking neural networks," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 126–144, 2019.

[28] N. R. Pal and J. C. Bezdek, "On cluster validity for the fuzzy c-means model," *IEEE Transactions on Fuzzy Systems*, vol. 3, no. 3, pp. 370–379, 1995.

[29] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: continual prediction with LSTM," *Neural Computation*, vol. 12, no. 10, pp. 2451–2471, 2000.

[30] Z. Cheng, W. Wang, J. Lu, and X. Xing, "Classifying the traffic state of urban expressways: a machine-learning approach," *Transportation Research Part A: Policy and Practice*, vol. 137, pp. 411–428, 2020.

WILEY | Hindawi

Research Article

Research on Travel Behavior with Car Sharing under Smart City Conditions

Zhimin Tao,[1] Quan Nie [ID],[2] and Weibin Zhang [ID][2]

[1]School of Transportation Science and Engineering, Beihang University, Beijing 100191, China
[2]School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

Correspondence should be addressed to Weibin Zhang; weibin.zhang@njust.edu.cn

As a sustainable transportation system, car-sharing schemes have been attracting increasing attention. A large amount of research and practice has proved that the application and promotion of car sharing can help reduce the number of private cars purchased, increase the utilization rate of automobiles, effectively alleviate traffic congestion, save energy, and reduce emissions. Therefore, research on car sharing is imperative. The logit model is widely used in studies on car sharing and is an effective tool for analyzing traffic problems. This study first introduces the status of research into car sharing and analyzes the potential users and market prospects for shared cars. The study then provides the results from a questionnaire survey in Nanjing, China, to obtain sample data. Finally, a mixed logit model is established to analyze the influencing factors of car-sharing selection behavior. The results show that factors such as an individual's housing situation and income significantly affect car-sharing decisions and that respondents who choose to use shared cars are relatively similar to commuters. The main contribution of this study is to use empirical analysis to determine the key influencing factors of car-sharing behavior in China and to provide practical insights for commercial practitioners and traffic planners.

1. Introduction

In the past decade, the rapid development of urban modernization has brought many social problems to China's transportation systems. Since 2010, China has become one of the largest car manufacturing regions in the world, producing tens of millions of cars [1]. The increase in private cars has had a huge negative impact on China's urban development and caused environmental issues, such as traffic congestion, air pollution, and overloaded parking areas, which have meant that policy makers must find effective solutions to control the level of car ownership. Car sharing is considered a possible innovative approach to solving these problems and is growing rapidly.

In recent years, as a new and more sustainable mode of transportation, car sharing in many countries is causing a shift in car ownership from private transportation to shared services. The first car-sharing organization can be traced back to Zurich, Switzerland, in 1948. However, for economic

reasons, the development of car-sharing systems was not smooth in the years that followed. It was not until the 1980s that car sharing first entered the market, and in the early 1990s, there was an increase in car sharing due to the popularity of ICT and mobile services [2]. In recent years, with the development of electric vehicles, it has become possible to use these in car sharing. Electric vehicle sharing avoids the high cost and range issues of these vehicles and has great potential in reducing operating costs and greenhouse gas emissions [3].

Car sharing can improve vehicle utilization, reduce air pollution, and relieve urban traffic pressure while also raising awareness among citizens of private car use and environmental protection. From the perspective of building sustainable cities, vehicles used for car sharing are usually new energy vehicles such as electric car, and they play a vital role in reducing urban emissions and congestion [4]. Some researchers have found that analyzing the characteristics of user behavior is very important for the development of car-

sharing services because behavior can fundamentally explain some system operating problems. Research in Europe and the United States has analyzed the behavior patterns of users and concluded that behavior is a key factor in determining the successful operation of car-sharing services [2, 5].

Based on data from shared-car travel questionnaires, this study establishes a mixed logit model to analyze the significant influencing factors for residents in Nanjing, China, in choosing shared-car travel. This study explores the pattern rules of residents' choice behavior in car sharing and quantifies the interactions and influence mechanisms of various factor. Finally, based on the research results, practical suggestions are made for policy makers and commercial practitioners seeking to promote the development of car sharing in China as a means of alleviating the problems of urban transportation and environmental impacts. The main contribution of this study is to conduct empirical analysis of the key influencing factors of China's car-sharing choice behavior and to propose effective options that will help with the implementation and development of car sharing in China.

The second chapter of this study reviews the literature, outlines the current status of car sharing and the potential user market, and conducts a comprehensive assessment of the modeling methods used in previous research. The third chapter presents the research methods adopted in this study, including data acquisition methods, modeling analysis methods, and their parameter tests. The fourth chapter describes the questionnaire design, summarizes the questionnaire data, and predicts the model simulation results based on the intuitive results from the sample data. Next, the fifth chapter evaluates the simulation results, focusing on the empirical results and discussing the significant influencing factors for residents in Nanjing, China, to choose car-sharing travel. Finally, we summarize our findings and propose a future research direction.

2. Literature Review

2.1. Research Status.
A number of studies in Europe and the United States have analyzed car-sharing users' behavior patterns. Some studies have shown that, after joining a car-sharing club, members used shared cars less than three times per month, and this was mainly for shopping, visiting friends, leisure, and vacations. A small proportion of members used shared cars for regular commuting, and most members used them to carry heavy items or to travel to many destinations [6]. The development of China's car-sharing system is still in the initial exploratory stage. Due to a lack of government and social support, there is relatively little research on car sharing in China.

Current research has illustrated the potential benefits of car-sharing systems, including benefits to society and users. A study by Nijland and Meerkerk [7] showed that car sharing can effectively reduce car ownership and can remove the need for a second or third car. Car sharing has been identified as a new model that can reduce costs, improve transportation efficiency, reduce car ownership, reduce traffic congestion, reduce parking demand, and improve the

environment [8]. In terms of the users' choice of car-sharing behavior, previous studies indicate that individuals' personal situations have a stronger influence on their willingness to use shared cars than the properties of the cars [9]. Other studies have shown that the service cost gap and service characteristics play a vital role in the choice of car sharing [10–12]. These factors should be taken into account when developing car sharing in emerging markets such as China.

Car sharing was introduced into mainland China in 2015, and many auto companies have since started to operate shared cars in large cities. Currently, the main shared car service companies in Nanjing are EvCard, GoFun, and GreenGo. Most of these companies use new energy vehicles as shared vehicles. Several studies have shown that China is a potential development and application market and have proposed some potential market segmentation systems [13]. Although some studies on the behavioral patterns of shared-car users provide valuable information, most studies only analyze the users' intentions. Recently, many researchers have concluded that analyzing the potential demand of users' travel behavior is important in determining the successful operation of a car-sharing system [4].

2.2. Potential Market and Users.
The potential market for shared cars has been analyzed in a number of previous studies, and these have found that the main groups using shared cars are community residents, business circles, college students, low-income groups, and commuters [14]. Studies have shown the market potential of car sharing in university campuses, finding that the university market has the ability to develop and expand the scope of a community or commercial market [12]. Younger people prefer car sharing because their attitudes toward cars are different from older generations; they consider that car sharing is a new form of car culture [15]. Shared cars are more convenient in urban centers than in suburbs because of the higher population densities; therefore, people in downtown areas are more likely to use car sharing. In addition, the frequency of car use is an active driving force for car sharing, and high-frequency users are more likely to join shared car services to meet their transportation needs [16]. Kim et al. [17] found that the availability of shared cars plays a crucial role in choice behavior, based on a hybrid selection model. Research by Gheorghiu and Delhomme [18] analyzed the use of carpooling for different types of daily travel and found that the commonest use was leisure travel, followed by shopping, then work, and children-related travel. This provides a reference for the market analysis of shared car services.

Wang et al. [13] considered that the potential for car sharing in the Chinese market is very large and that shared-car sites located near public transport facilities will attract more potential users. This study also pointed out that groups with the most potential to join car-sharing systems were the most educated citizens, young people, and middle-income people [16]. Many researchers believe that people who are familiar with the concepts of car sharing are more likely to join a system. Shaheen [14] conducted a survey of 840 respondents in Beijing to explore the impact of familiarity on

participation in car-sharing systems and found a correlation between the degree of familiarity of respondents and their choice of interests. This means that the deepening of familiarity will increase the likelihood of participation.

2.3. The Main Modeling Method. The discrete choice model is an effective tool for the study of shared-car travel behavior. A major current research direction is the study of the characteristics of shared-car users' travel behavior. In some studies, population characteristics were modeled using car-sharing user samples and analyzed using logit models. To understand the driving factors in the choice of urban car-sharing systems, Luca and Pace [11] studied multiple logit, hierarchical logit, cross-nested logit, and mixed logit models and determined which was most effective. The logit model is a type of the discrete selection model with various forms, such as binomial log [19], multiple logit [20], hierarchical logit, cross logit, nested logit, and hybrid logit [21], which are widely used to model travel decisions for car-sharing users.

The mixed logit model provides insights into travel behavior, decision-making, and demand forecasting and provides a wealth of theoretical support. In particular, it contributes greatly to travel demand theory, which helps transport policy analysis [22]. The characteristics of heterogeneity, the less restrictive substitution patterns, and the correlation between unobserved variables mean that the mixed logit model has a good effect in analyzing traffic behavior. Furthermore, because it enhances the behavioral connection between individual travel demand and the attributes of the transportation system, the mixed logit model can better analyze the factors that influence the behavior of car-sharing choices.

3. Research Methods

3.1. Research Design. As a megacity, Nanjing has undergone rapid economic growth, motorization, and urbanization in the past few decades. The focus of this study is to use logit models to analyze the factors that affect the choice of shared-car modes, to explore the patterns of residents' choices of shared-car behavior, and to quantify the interaction and influence mechanism of each factor. This study uses quantitative analysis methods including the design of stated preference- (SP-) based survey questionnaires, the collection of questionnaire data, and the establishment of a logit model to simulate data analysis.

There are three main approaches to this study. First, in order to better understand the concept of car-sharing systems, a large number of previous studies on car sharing in China and other countries are reviewed. Based on this previous research, we introduce the modeling method of this study and the main data to be collected. Second, we identify important data collection samples and design a questionnaire to enable analysis of the factors influencing Nanjing citizens' choice of car sharing, based on the SP method, and we check the validity and reliability of the important variables of the data. Finally, sample data from logit model analysis are used to estimate multiple unknown parameters,

and the research objectives of this study are quantified and explained according to simulation results.

3.2. Data Collection Method. This section illustrates the advantages and disadvantages of the SP survey method used in the questionnaire design. The SP survey method can analyze the behavior of the respondents under various hypothetical scenarios, reflecting the impact of unobserved variables such as comfort, convenience, and safety on the respondents. It is also possible to test the basic attributes of a research area that does not exist in real life.

In recent years, the SP survey [23] has been widely used in research topics such as travel destination selection, transportation mode selection, and travel time value analysis. The advantages of questionnaire data based on the SP method are that

(1) The data have extremely high operability
(2) The data error can be adjusted
(3) The set of selection options is determined for the stated preference survey

However, the SP survey is not perfect. It can easily lead to various biases, such as survey bias in which the respondents tend to reinforce their first choice based on other incentives. Content effect bias occurs when respondents respond to questions based on situations they have not experienced before, and when interviewees receive too much information from the investigator, they are bound by circumstances that make them confused about the decision. Nevertheless, these biases can be mitigated through proper design and management [20].

China is an emerging car-sharing market, and the concept of innovative car sharing is not common in Nanjing. The preference survey asks interviewees to respond to things that exist in reality, and it relies on the actual choices of the respondents. Therefore, using the revealed preference (RP) survey method to design the questionnaire would cause some limitations and make the method unsuitable for the study. Compared with RP surveys, SP surveys are more conducive to collecting data on scenarios that cannot be directly observed. Furthermore, the SP method can bring more flexibility and possibility, and the data obtained are more conducive to the survey participants' perception and satisfaction. Therefore, for the survey of car-sharing systems in China, the SP method is considered the best choice, and we designed a questionnaire using this method.

3.3. Modeling Method Description. The analysis and modeling of travelers' choice behavior is one of the most important research directions in traffic research [2]. The main model used in this research field is the discrete choice model. Discrete choice models can be divided into two major types of models: those with closed-form expressions and those that approximate a numerical solution based on a formula with integral values using simulation.

Discrete choice models are an effective method for analyzing the choice of the traffic mode and are used by

many scholars. For example, Ran [8] used Nanjing residents' travel data to establish a mixed logit model to determine sixteen variables in four ways to fully explain and quantitatively apply the characteristics of the residents' travel mode selection. The mixed logit model can effectively reflect the heterogeneity of individuals and explain the behavior of residents' choice of the travel mode. However, due to the high computational cost of the simulation process for the mixed logit model, its initial use is only in a very limited estimation and application. However, in recent years, the remarkable improvement in computer performance has increased people's interest in simulation models, and the mixed logit model has gradually become one of the most widely used models in transportation research [24].

Based on the above discussion, we chose the logit model to analyze the travel behavior of car-sharing users. According to the theory of utility maximization, it is assumed that the total utility of traveler n to choose travel scheme i is U_{jn}. Assume that the utility is a random variable:

$$U_{in} = V_{in} + \varepsilon_{in}, \tag{1}$$

where V_{in} is the determination or observable part of the utility of traveler n choosing scheme I, and ε_{in} is the uncertainty or unobserved utility portion of the utility function of the traveler n choosing scheme i.

According to the utility maximization theory, the probability that traveler n chooses the scheme i is

$$P_{in} = \text{Prob}\left(U_{in} > U_{jn}; \quad i \neq j, \ j \in A_n\right), \tag{2}$$

where $0 \leq P_{in} \leq 1$. Assume that ε obeys a double exponential distribution, and the expression for the general logit model probability is

$$P_{in} = \frac{eV_{in}}{\sum_{j \in A_n} eV_{jn}}. \tag{3}$$

When the fixed unknown parameters are replaced by random numbers subject to a distribution, the probability that scheme i is chosen should be the expected value of the probability of β traversing all possible values, and the probability function of the mixed logit model can be regarded as the integral of the binomial logit probability function on the probability density function of β. The selection probability can be expressed as

$$P_{in} = \int L_{in}(\beta) f\left(\frac{\beta}{\theta}\right) d\beta. \tag{4}$$

In this formula is the probability of selection of the unknown parameter β as deduced by the above logit:

$$L_{in} = \frac{eV_{in}}{\sum_{j \in A_n} eV_{jn}}, \tag{5}$$

where $V_{in}(\beta) = \beta' X_{in}$. It can be seen that the standard expression of the probability of the mixed logit model is

$$P_{in} = \int \frac{eV_{in}}{\sum_{j \in A_n} eV_{jn}} f\left(\frac{\beta}{\theta}\right) d\beta. \tag{6}$$

The unknown parameter β obeys a certain probability distribution, such as normal, uniform, or lognormal, and θ is the parameter of the probability distribution, such as the upper and lower limits of a uniform distribution. The unknown parameter β of the mixed logit model obeys a certain distribution form, which reflects the randomness characteristics of different travelers' choices of the travel mode and overcomes the irrelevance of the ratio caused by the independence of irrelevant alternatives (IIA) of the general logit model.

Because the symbol function of Matlab software cannot find the analytical solution, we consider the approximate solution of Monte Carlo simulation [25]. The following is an approximate algorithm using a Monte Carlo simulation method [24, 26].

Step 1: seeking simulation probability $\breve{P}_{in}$

(1) First determine the initial value of θ and then randomly extract a vector β from the probability density function $f(\beta/\theta)$, denoted as β_t. Remember the first time as $t = 1$.

(2) Calculate the value of $L_{in}(\beta_t)$ according to equation (5)

(3) Repeat steps (1) and (2) T times, using $T = 500$ in this study. Calculate the value of $L_{in}(\beta_t)$ each time and record it. Then, take the average value as the simulation value of the selection probability:

$$\breve{P}_{in} = \frac{1}{R} \sum_{r=1}^{R} L_{in}(\beta_r). \tag{7}$$

Previous studies have shown that increasing the number of sampling times T can effectively reduce the deviation and variance of the simulation process [27, 28].

Step 2: construct a logarithmic maximum likelihood function

Remember that the total number of samples is N and the number of selected limbs is J. Define the auxiliary variables

$$y_{in} = \begin{cases} 1, & \text{traveler } n \text{ has chosen scheme } i, \\ 0, & \text{others.} \end{cases} \tag{8}$$

So, the logarithmic maximum simulation likelihood function is

$$SLL(\beta) = \sum_{n=1}^{N} \sum_{i=1}^{J} y_{in} \ln \breve{P}_{in}. \tag{9}$$

Step 3: change θ to solve the likelihood function

Change the value of θ until the maximum simulated likelihood function achieves the maximum value to obtain the parameter estimation value. The function solution can use the Newton–Rapson method and the gradient method.

In the above method, we apply the maximum likelihood estimation method, which expresses the

probability of the respondents' choice behavior as a function of the unknown parameter and which can obtain the parameter estimate that maximizes the logarithmic maximum likelihood function value. This set of parameters is an estimate that best explains the sample data.

This study uses Matlab software programming to establish a model to construct the log-likelihood function of the above method. To realize the maximum likelihood of the log-likelihood function, the routine fminunc in Matlab is used to implement the unconstrained optimization descent algorithm proposed by Kalouptsidis and Psaraki [26]. Then, the parameter estimates of the characteristic variables are derived, and the factors that affect the travel behavior of car sharing are analyzed based on the simulation results.

3.4. Parameter Test. After modeling the data and performing maximum likelihood estimation to obtain the estimated values of the parameters, we must test the fitting effect of the model and study how well the estimated values of the parameters match the observed values. If the match is good, we can consider the model's fitting effect to be good (i.e., the model can effectively reflect the influencing factors of the travelers' selection behavior). The indicators for model parameter testing are

(1) McFaddenor goodness ratio is with a value between 0 and 1; the closer the value is to 1, the better the fitting effect. The formula for calculating this value is

$$\rho^2 = 1 - \frac{SLL(\widehat{\beta})}{SLL(0)}, \tag{10}$$

where $\widehat{\beta}$ is the parameter estimate value. Generally, in practical situations, when the value of ρ^2 reaches 0.2–0.4, the fitting accuracy of the model can be considered to be quite high [19].

(2) $\rho^2 = ((N - K)/N)\rho^2$ is also an indicator of model fitting accuracy and is equivalent to the degree of freedom correction correlation coefficient in regression analysis

(3) Hit ratio

This is a test indicator that shows whether the actual selection result of a traveler matches the selection result predicted by the model. The intermediate rate calculation steps are as follows.

(1) Step 1: substitute the parameter estimation vector $\widehat{\beta}$ and the characteristic variable data X_{ink} into equation (5) to find the selection probability $\widehat{P}_{in}$ of traveler n.

(2) Step 2: for all n, assume that the selection scheme with the selection probability greater than 50% is the selected scheme, and find the selection result:

$$\widehat{y}_{in} = \begin{cases} 1, & \widehat{P}_{in} \geq 50\%, \\ 0, & \widehat{P}_{in} < 50\%. \end{cases} \tag{11}$$

(3) Step 3: assume that S is 1 when the actual selection result y_{in} is equal to the prediction result $\widehat{y}_{in}$, and 0 otherwise. Hit ratio can be deduced from Hit

$$R = (1/N)\sum\nolimits^{n=1 N} S_{in}, \quad \text{where} \quad S_{in} = \begin{cases} 1, & y_{in} = \widehat{y}_{in} \\ 0, & y_{in} \neq \widehat{y}_{in} \end{cases}.$$

Generally, hit ratio is unlikely to reach 100%, and if it is greater than 80%, then the model's fitting accuracy is quite high.

4. Questionnaire Design and Summary

4.1. Questionnaire Design. Considering that car-sharing services were only recently introduced to Nanjing and the actual coverage is not widespread, the SP method is the most suitable questionnaire design method for this study. In this study, we divide the SP-based questionnaire into three parts.

In the first part of the questionnaire, we asked the demographic attributes of respondents (age, housing status, and income level). The subject of the study was the person with a driver's license. Age was divided into four groups. We defined "young people" as 18–30 years old. The typical characteristics are that they are still single, most of them are students, and their working conditions are not stable. Middle-aged people are those between the ages of 31 and 50. They are typically married, own cars, and have stable income. The middle and old age people are defined between the ages of 51 and 60. They have some savings and more free time. Older persons are those over the age of 60. They have a large amount of free time and prefer to use public transport or other means of transportation when their driving ability declines due to age restrictions.

The income level of respondents is divided into four options, based on data on the average monthly income of Nanjing from Nanjing Human Resource Bureau. The four groups are listed in Table 1. In general, the high-income group is highly educated and has private cars.

The relationship between housing and family situation with likelihood of using cars and car-sharing service is considered in this study. This section also discusses issues related to traffic habits, such as car ownership, travel preferences, parking, and the use of mobile apps. Respondents were also asked to describe their preferred mode of transportation, which helps researchers to propose suggestions for mobility. Family car ownership was used to analyze the impact of car ownership on car-sharing acceptance. Questions were also included regarding the frequency of use of taxi apps to assess the relationship between the frequency of use of these apps and the likelihood of participating in car-sharing systems.

Respondents were asked whether they wished to buy a car. We assumed that people who do not own a car or who do not wish to buy a car are more likely to participate in car sharing. To assess driving experience, respondents were asked when they obtained their driving license; we assume that the longer the time since receiving a license, the richer the driving experience of respondents.

The second part of the questionnaire produced a picture of the car-sharing system. Based on an understanding of the

TABLE 1: Monthly income category in Nanjing.

Category	Income (RMB)
Low	<5000
Middle	5000–10000
Middle-high	10000–15000
High	>15000

SP survey methodology, we used software to design a picture of a shared car profile that provided respondents with general concepts of car-sharing services, including visual and textual forms. The image briefly introduces the main processes of using shared cars, the characteristics of shared car services, and the charging standards of several car-sharing operators in Nanjing; the aim was to increase the level of understanding of car sharing by respondents and attract more environmental protection. At the same time, it helped respondents to better understand the concept of the shared car program they choose before completing the questionnaire. This helps respondents to make better choices and improves the validity of the questionnaire. This improves the credibility of the data and reduces the error, so that the results of the simulation analysis are more reliable.

The third part of the questionnaire addresses the properties of shared cars and asks respondents whether they choose to participate in car sharing. Each interviewee was presented with three of the nine plans from Nanjing's shared-car companies, each showing multiple shared-car properties, including

(1) "Deposit fee," paid by each user before using the shared car

(2) "Fee/minute," based on the time of use

(3) "Fee/kilometer," based on the distance traveled

(4) "Distribution distance," the distance from the interviewee to where the shared car is parked

This section aims to determine the impact of these different properties on choice behavior. Our research is based on the billing methods of Nanjing's car-sharing operators; companies such as GoFun and EvCard provide point-to-point car-sharing services. Each questionnaire included three options for sharing cars: respondents were asked to choose between the three schemes; therefore, each had only the two options of "yes" or "no." This can reflect the likelihood of respondents choosing to join each shared-car system (Figure 1). Table 2 lists the shared-car schemes used for the survey.

4.2. Summary. Based on the SP-based questionnaire, this study investigates Nanjing residents' car-sharing choices. The main factors affecting shared-car travel behavior are demographic, socioeconomic, and sharing program attributes. Based on these three aspects, we set relevant questions to conduct our survey.

The questionnaire used an online survey method. The investigation period was from April 2 to April 25, 2018. A total of 189 questionnaires were completed. Based on the

needs for constructing quantitative models and on the principle of complete and effective data, the quality of the questionnaire data was controlled; a number of invalid questionnaires were removed, and the questionnaire variables were discretized. There were 160 valid questionnaires, and the effective rate was 84.7%.

Age and education are important characteristics of an individual's basic attributes and have certain influences on the willingness of a traveler to share a car. Women comprised 46.25% of respondents and men 53.75%, which matches the gender ratio of Nanjing residents. Young people comprised 65.00% of respondents, followed by middle-aged respondents (18.75%).

In terms of the educational level, 55.00% of respondents were undergraduates, followed by graduate students (18.13%). In terms of family status, 28.13% of respondents were single, 16.25% was unmarried but had male (female) friends, and 55.30% was married. Most respondents (66.87%) lived in college dormitories or rental housing. 41.25% of respondents had a low income (monthly average of <5,000 yuan), and 27.50% had a middle-high income (10,000–15,000 yuan).

66.25% of respondents owned at least one car. Of these, 83.12% had a driver's license, and 26.26% had more than 5 years' driving experience. 68.75% of respondents said they did not wish to buy a car.

In terms of use of transportation modes, 36.25% of respondents chose public transport such as buses and subways, 26.25% chose private cars, and 18.75% preferred taxis. The remaining 18.75% of respondents chose to use a battery car or bicycle to travel. When asked about their familiarity with car sharing, the respondents were asked to rate their level of knowledge on a scale from "not familiar" to "very familiar." 63.13% of respondents said they had an understanding of the car-sharing system, 22.5% was very familiar with the concept, and 14.38% said they were not familiar with it. 91.25% of respondents indicated that they had to use software such as DiDi, and most indicated that they often use mobile applications.

When asked about the reason for using taxi software, 60.00% of respondents chose it for convenience, 54.38% chose it to save time, and 53.13% chose it because it had a cheaper price. Sharing cars, network taxis, and taxis had many similarities; therefore, the degree of convenience and the parking situation have a great influence on the choice of car sharing. Among the problems contributing to the choice of car sharing, 80.63% of respondents regarded the parking situation as an important factor. This shows that people are becoming less inclined to spend time looking for parking spaces as the pace of life gradually accelerates. 67.50% of respondents said that they use shared cars more for daily use and commuting.

Figure 2 shows that 72.09% of men chose to use a shared car program, compared to 64.86% of women, indicating that men are more willing to participate in car sharing. Young people were more likely to use car sharing (Figure 2). Some studies point out that young Chinese people's emotional identification with cars is weakening and that they now think of them simply as a mode of transport; consequently, they are more likely to accept the sharing model [29].

14.Car sharing scheme 1 in Nanjing , are you willing to use? *

Deposit: 699 yuan
Car models: New energy electric vehicle
Cost: 0.1 yuan / minute + 1.00 yuan / km, 400 yuan daily cap
Pickup distance: 1.2km

○ Yes   ○ No

FIGURE 1: Questionnaire of car-sharing plan one.

TABLE 2: Car-sharing scheme.

Plan	Deposit fee (yuan)	Fee/ minute (yuan)	Fee/ kilometer (yuan)	Distribution distance (km)
1	699	0.1	1.0	1.2
2	699	0.2	1.2	0.5
3	1000	1.2	0	1.5
4	699	0.2	1.0	1.5
5	699	0.4	1.2	1.0
6	1000	1.4	0	0.5
7	699	0.3	0.5	1.0
8	699	0.6	1.2	0.5
9	1000	1.0	0	1.2

Figure 3 shows that people with higher education are more likely to share vehicles and have a certain interest in car sharing. In addition, respondents who have fewer cars are more willing to use car-sharing service, which is consistent with the observed situation; people who do not have cars at home will want access to more convenient car travel.

Figure 4 shows that, in both the married group and the "unmarried but not single" group, approximately 73% of respondents would prefer to share cars. This is because they have more travel needs than single people. Approximately 79% of respondents who were renting chose to use shared cars. Generally, people who have no fixed property have relatively low incomes, so they have a greater desire to share cars.

Figure 5 shows that people who are more familiar with car sharing are more likely to choose to share cars, which is consistent with past research results and indicates that shared-car promotion is necessary to promote its development. In addition, there is greater interest in car sharing among people who use a taxi app more often; this suggests that a car-sharing system with mobile apps is more likely to attract potential customers.

When asked the reasons for using shared cars, 60% of respondents stated that convenience was a reason, and 41.25% of respondents chose them to protect the environment (Figure 6). Price is also an important factor (53.13% of respondents) and always plays a negative role in people's choice. 54.38% of respondents chose car sharing to save time, which indicates that this is also a key factor for people accepting car sharing in fast-paced cities.

When asked about the impact of parking conditions on the choice of shared vehicles, 58.13% of respondents considered it to be the determining factor, and 22.5% of respondents thought that it had a big impact (Figure 7). This

indicates that, with the high number of cars and the general shortage of parking spaces in the city, shared car operators can attract potential users by adding infrastructure such as shared car pick-up and drop-off networks.

5. Model Evaluation and Discussion

5.1. Model Evaluation. As discussed in the previous section, demographic attributes and socioeconomic characteristics had a significant impact on individuals' choice behavior. In terms of demographics, men are expected to have a higher interest in car sharing than women, and young people, highly educated people, and low- and middle-income people may be more willing to participate in car sharing. In addition, we believe that the respondents' choices will be influenced by other factors, such as familiarity with car sharing and usage of mobile phone applications. In this section, we use Matlab software to create a mixed logit model and examine the effect of each characteristic variable on individual selection behavior. The model's characteristic variables and their meanings are provided in Table 3.

The maximum log likelihood method is used to estimate the model, analyzing the variables in Table 3, and the Matlab software is used to solve the model. The model parameter estimation results are shown in Table 4.

We set each characteristic variable in the mixed logit model to obey the normal distribution and changed the expectation and variance of the normal distribution corresponding to each variable until the maximum simulation likelihood function achieved the maximum value. This value evaluates the parameter of the characteristic variable. The simulation results of the two models are in Figures 8 and 9; the model explained well users' travel selection behavior. The McFadden value of the model is 0.2283, the corrected McFadden value is 0.2100, and the medium rate Hit $R = 73.12\%$, which show a good fit.

The simulation results show that the fitted precision of the mixed logit model is high, which better explains the influence mechanism of shared-car travel behavior. According to the model parameter estimation results, the expression for selecting the shared car's power function can be obtained as

$$
\begin{aligned}
V_{in} = {} & 0.4379\text{Gen} + 0.3335\text{FS} - 0.8066\text{CarO} \\
& + 0.5164\text{Income} - 0.8954\text{HS} - 0.2119\text{Dl} \\
& + 0.5483\text{Tp} + 0.6030\text{Carb} + 0.2947\text{CsF} \\
& + 0.7789\text{UseP} - 0.1960\text{Fee/h} - 0.0326\text{Fee/km} \\
& - 0.0326\text{CT}.
\end{aligned}
$$

$$(12)$$

6. Discussion

As expected, demographic characteristics have a significant impact on car-sharing acceptance. The simulation results show that the "Gender" parameter is estimated to be positive and shows positive significance, indicating that men are more interested in driving than women. This result is

(a)

(b)
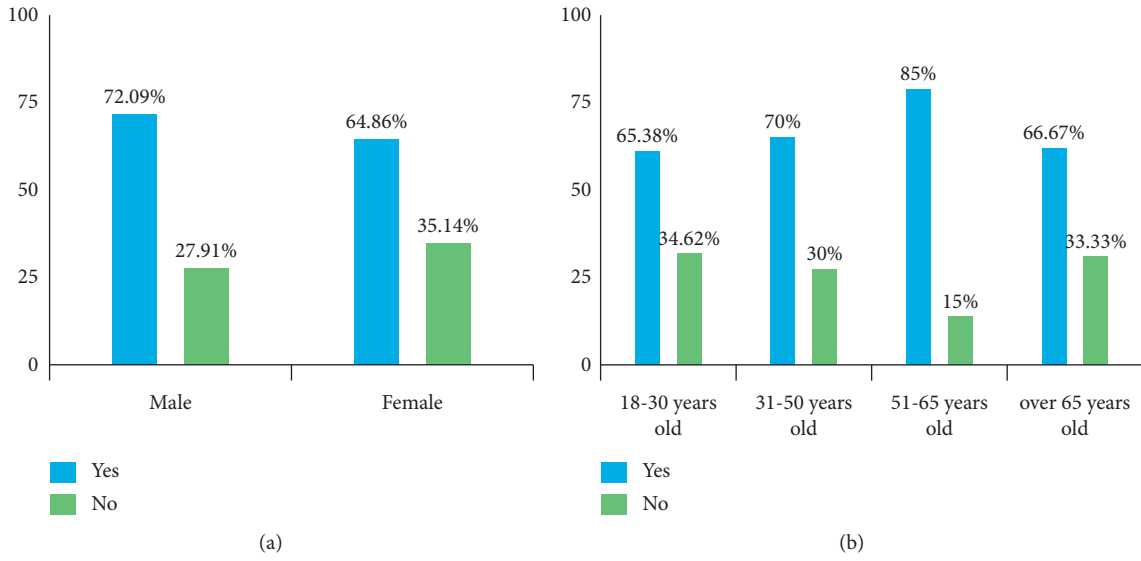
FIGURE 2: Breakdown of choice behavior according to sex and age.

(a)

(b)
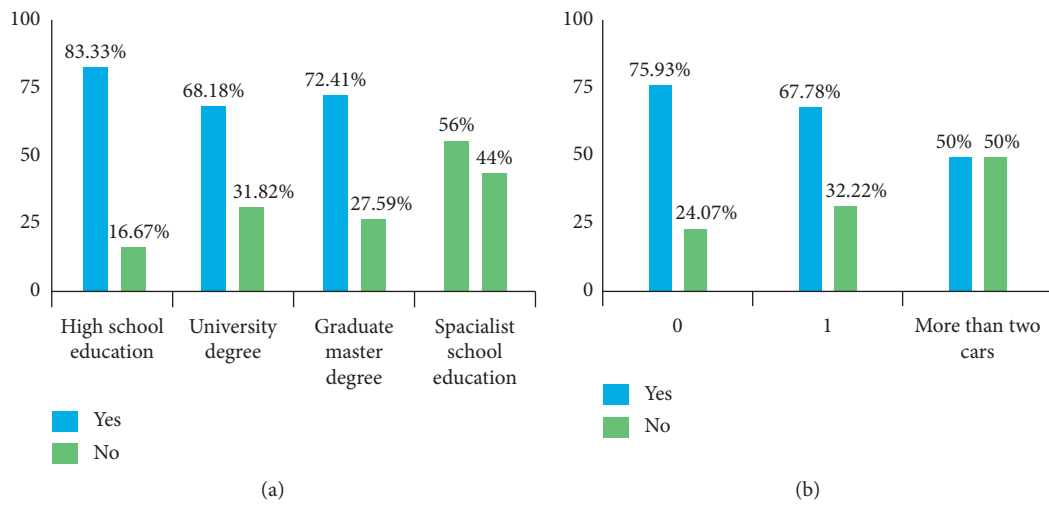
FIGURE 3: Breakdown of choice behavior according to education and private car ownership.

(a)

(b)
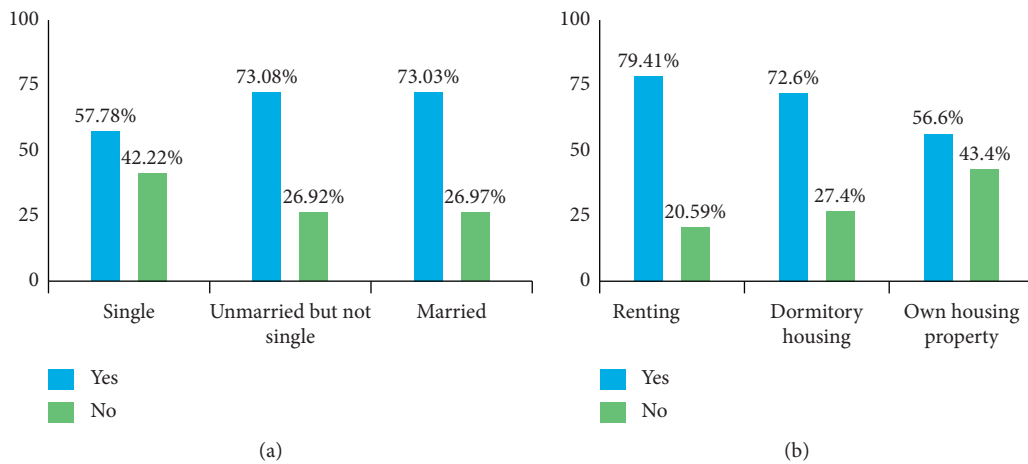
FIGURE 4: Breakdown of choice behavior according to family status and housing.
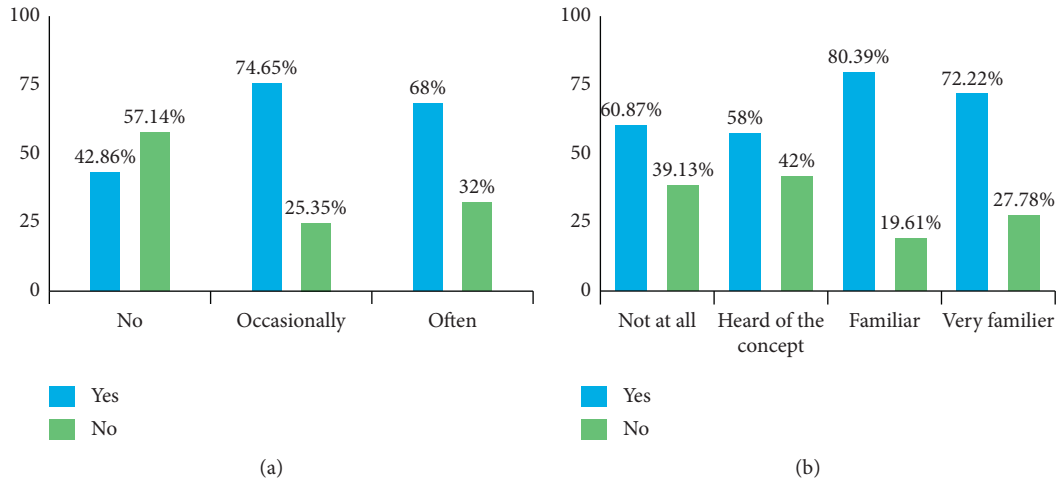
FIGURE 5: Breakdown of choice behavior according to taxi app usage and familiarity with car sharing.
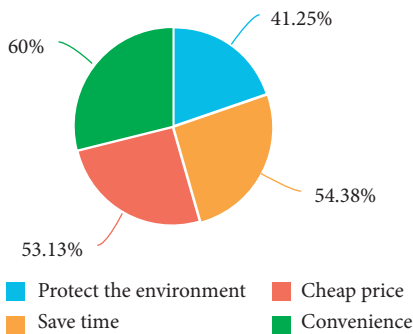
FIGURE 6: Reasons for using shared cars.
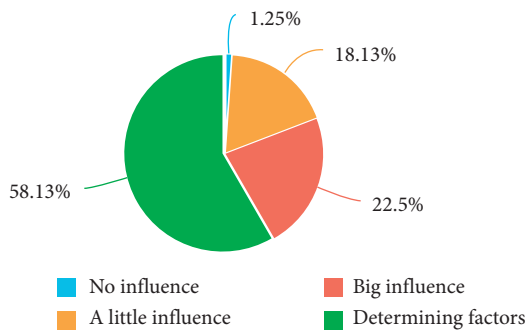
FIGURE 7: Effect of parking situation.

contrary to some studies on car sharing in North America and Europe, which found that women are more likely to join the car-sharing market. The utility of the "Age" variable is negative, indicating that younger people are more likely to choose car sharing. The "Edu" effect is positive, which means that those with higher education are more likely to use car sharing. This may be because highly educated people are more likely to be familiar with new concepts such as car sharing. Therefore, it is recommended that policy makers implement shared car systems in areas with highly educated populations, such as shopping malls and university towns.

Family status (FS) also has a positive effect, which means that married people are more likely to choose to use a shared car. The results accord with the reality that married people and those who are unmarried but not single usually have more driving needs than those who are single.

As expected, the "HS" and "CarO" variables are negative. This means that those who own a property or car are usually less interested in other transportation services such as rental cars and shared cars. The variable "Income" has a positive effect, indicating that those with higher incomes are still willing to join a car-sharing system. This is slightly different from the previous study, which may be because the age of most people in this survey were under 30 years old and thoughts of young people are changing compared with elder persons.

In terms of social and economic attributes, we provide evidence that if people are more familiar with the system, then they are more likely to join car sharing. The simulation results show that the variable "CsF" has a positive and significant effect, indicating that the likelihood of participating in car sharing increases as the respondents' familiarity with the concept increases. As expected, the variable "Mobile" has positive effects, which means that car-sharing systems that develop mobile apps are more likely to attract potential customers. The positive correlation between variables related to travel mode preferences indicates that people using private cars, especially electric car, are more interested in car sharing than those who usually travel in other ways, such as via taxis, buses, and subways. The significant positive effect of "Carb" indicates that respondents who wish to buy cars have greater interest in sharing cars, which suggests that shared cars could effectively reduce the purchase of cars in China; this is consistent with many previous studies. The variable "Dl" was estimated to be negative, indicating that as driving experience increases, people's interest in joining a car-sharing system decreases.

Simulation results show that car-sharing attributes have little effect on respondents' selection behavior. The cost-related characteristic variable estimate is negative, meaning that the

TABLE 3: Categorical variable coding table.

Characteristic variable	Symbol	Meaning	Amount	Frequency (%)
Age	Age	1, male	86	53.75
		0, female	74	46.25
Gender	Gen	1, 18–30 years old	104	65.00
		2, 31–50 years old	30	18.75
		3, 51–65 years old	20	12.50
		4, over 65 years old	6	3.75
Education level	Edu	1, high school education	18	11.25
		2, specialist school education	25	15.63
		3, university degree	88	55.00
		4, graduate degree	29	18.13
Family status	FS	0, single	45	28.13
		1, unmarried but not single	26	16.25
		2, married	89	55.63
Private car ownership	CarO	0, no car at home	54	33.75
		1, one car at home	90	56.25
		2, two cars at home	16	10.00
Housing situation	Hs	0, renting or dormitory housing	107	66.87
		1, own house/property	53	33.13
Income level	Income	1, monthly income below 5,000	66	41.25
		2, income in the range 5,000–10,000	26	16.25
		3, income in the range 10,000–15,000	44	27.50
		4, income above 15,000	24	15.00
Driving license	Dl	0, no driver's license	27	16.88
		1, driving less than 1 year	51	31.88
		2, driving 1–5 years	40	25.00
		3, driving 5–10 years	21	13.13
		4, driving for more than 10 years	21	13.13
Travel preferences	Tp	0, bus or subway travel	93	58.00
		1, private car travel	67	42.00
Car buying desire	Carb	0, do not wish to buy a car	110	68.75
		1, wish to buy a car	50	31.25
Shared car familiarity	CsF	0, not at all familiar	23	14.38
		1, heard of the concept	50	31.25
		2, familiar	51	31.88
		3, very familiar	36	22.50
Taxi software usage	Mobile	0, do not use taxi software	14	8.75
		1, occasionally use taxi software	71	44.38
		2, often use taxi software	75	46.88
Influence of parking situation	Park	0, no effect	2	1.25
		1, a little effect	29	18.13
		2, greater effect	36	22.50
		3, the decisive factor	93	58.13
Using shared cars' purpose	UseP	0, temporary use such as moving and short trips.	52	32.50
		1, daily use, work shift, and life shopping.	108	67.50

TABLE 4: Mixed logit simulation results.

Variable	Symbol	B
Gender	Gen	0.4379
Family situation	FS	0.3335
Car ownership	CarO	−0.8066
Income level	Income	0.5164
Housing situation	HS	−0.8954
Driving license	Dl	−0.2119
Travel preferences	Tp	0.5483
Car buying desire	Carb	0.6030
Shared car familiarity	CsF	0.2947
Purpose of usage	UseP	0.7789
Time fee	Fee/h	−0.2012
Distance fee	Fee/km	−0.0335
Pick up distance	CT	−0.0335
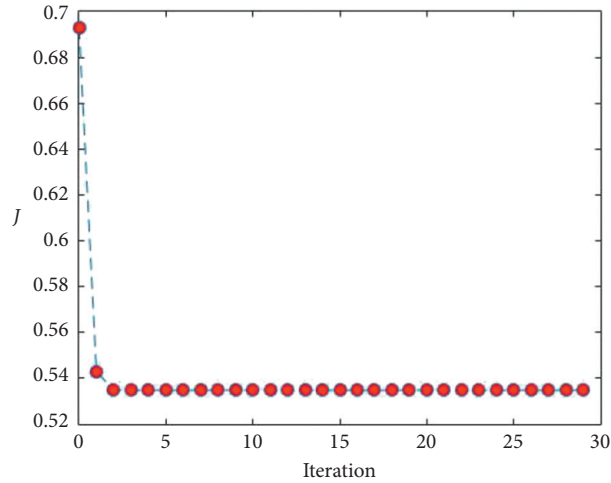
FIGURE 8: BL (binary logit) model likelihood function optimization.

(a)

(b)
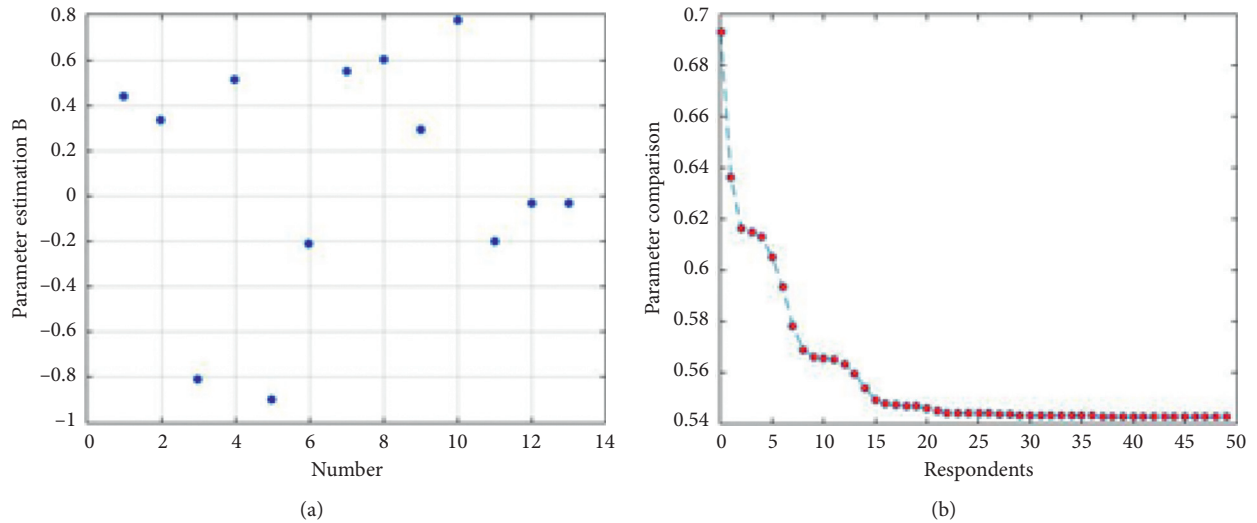
FIGURE 9: Likelihood function optimization diagram of the mixed logit model.

utility of sharing a car decreases as the cost increases. However, the estimated value of the cost per kilometer and the distance from the vehicle is small, and the impact is not significant. The following two assumptions can be used to explain the situation. First, respondents may mostly regard shared cars as short-distance travel tools, so they are sensitive to distance costs. Note that the degree is not high; therefore, the impact of distance costs is not significant. Second, it may be that people who are willing to accept car sharing often use shared bicycles, which are increasingly popular in China; therefore, the distance to pick up cars may not be a problem.

7. Conclusion

This study uses the mixed logit model to investigate the factors that have an impact on the travel behavior of people using shared cars in Nanjing, China. A questionnaire was designed based on the SP method, using selected characteristic variables covering demographics, socioeconomics, and shared vehicle attributes. The basic data and estimates of model parameters analyze the influencing factors for people choosing shared car travel, identify the significant factors, and provide corresponding insights on the development of shared car schemes.

The study found that young people in Nanjing have a higher acceptance of shared cars than people in other age categories and that travel preferences, family status, private car ownership, and housing conditions have significant impacts on the selection of shared cars. Although previous studies have shown that women are more attracted to car sharing, our research shows that men are more inclined to use shared cars. More highly educated people are more willing to join a car-sharing system, indicating that people with a higher level of education are more willing to try new things. Therefore, it is suggested that the number of shared cars be increased near schools and office buildings.

Housing condition had the greatest impact on the likelihood of a traveler choosing to participate in car sharing;

its utility coefficient is negative, indicating that people who own real estate are less likely to participate in car sharing. People who are married and people who are unmarried but single were more concerned about car sharing, which may be due to an increased demand for travel in the family and an interest in cost savings.

Familiarity with the concepts of shared cars also had a significant impact on selection results. We found that most of the respondents who were familiar with shared cars expressed their willingness to join a scheme. Therefore, it is necessary to introduce and promote car sharing to travelers, which can promote the development of car sharing in the Chinese market. The utility of "taxi software use" was also positive, indicating that the convenience of mobile Internet technology is also an important influence on the uptake of shared cars; therefore, increasing the development of mobile phone apps for shared cars can promote their popularity.

The simulation results also show that the properties of shared cars did not have a substantial impact on respondents' choices. The deposit and time costs had a more significant negative impact on the choice of respondents. Therefore, it is suggested that shared car companies consider adopting low-cost concessions to capture the market. The utility coefficient of the distance between vehicles is small, which indicates that the distance from the vehicle is not a problem for the traveler. This may be because the ready availability of shared bicycles now allows people to reach the shared car network quickly and easily. This also indicates that the cars will be used in conjunction with other public vehicles and that combining with transportation systems such as shared bicycles will attract more car-sharing customers.

In future research, we will consider more fully the influencing factors of car sharing and expand the scope of investigation and the number of samples. Because the questionnaire based on the SP method only surveyed the willingness of respondents, the choices made by respondents in the actual situation may be subjected to many other limitations or may be quite different. Therefore, it may be beneficial to combine SP and RP methods to produce survey questionnaires that provide more accurate results.

Data Availability

The data are collected and applied in Nanjing by an online questionnaire. The original data cannot be provided for privacy reasons.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

References

[1] S. Feng and Q. Li, *Car Ownership Control in Chinese Mega Cities: Shanghai, Beijing and Guangzhou*, Social Science Electronic Publishing, NY, USA, 2018.

[2] F. Ferrero, G. Perboli, M. Rosano, and A. Vesco, "Car-sharing services: an annotated review," *Sustainable Cities & Society*, vol. 37, pp. 501–518, 2018.

[3] P. Wu, "Which battery-charging technology and insurance contract is preferred in the electric vehicle sharing business?" *Transportation Research Part A: Policy and Practice*, vol. 124, pp. 537–548, 2019.

[4] E. W. Martin and S. A. Shaheen, "Greenhouse gas emission impacts of car sharing in North America," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 4, pp. 1074–1086, 2011.

[5] Y. Hui, W. Wang, M. Ding, and Y. Liu, "Behavior patterns of long-term car-sharing users in China," *Transportation Research Procedia*, vol. 25, pp. 4666–4682, 2017.

[6] S. A. Shaheen, A. P. Cohen, and M. S. Chung, "North American car sharing," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2110, no. 1, pp. 35–44, 2009.

[7] H. Nijland and J. V. Meerkerk, "Mobility and environmental impacts of car sharing in The Netherlands," *Environmental Innovation & Societal Transitions*, vol. 23, pp. 84–91, 2017.

[8] X. Ran, *Research on Characteristics of Residents' Travel Mode Selection Behavior in Big Cities: Taking Nanjing as an Example*, Southeast University, Nanjing, China, 2016.

[9] J. Zeng, M. Scott, M. Podriguez et al., "Car sharing in a university community: assessing potential demand and distant market characteristics," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2110, pp. 18–26, 2009.

[10] M. Catalano, B. L. Casto, and M. Migliore, "Car sharing demand estimation and urban transport demand modelling using stated preference techniques," *European Transport Trasporti Europei*, vol. 40, pp. 33–50, 2008.

[11] S. D. Luca and R. D. Pace, "Modelling users' behaviour in inter-urban car sharing program: a stated preference approach," *Transportation Research Part A: Policy and Practice*, vol. 71, no. C, pp. 59–76, 2015.

[12] J. Zhou, "Car sharing on university campus: subsidies, commuter benefits, and their impacts on car sharing," *Transportation Research Part D: Transport and Environment*, vol. 32, no. 32, pp. 316–319, 2014.

[13] M. Wang, E. Martin, and S. Shaheen, "Car sharing in Shanghai, China: analysis of behavioral response to a local survey and potential competition," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2319, no. 1, pp. 86–95, 2012.

[14] S. Shaheen, "Innovative mobility car sharing outlook: car sharing market overview, analysis, and Trends," *Spring*, vol. 2, no. 1, 2013.

[15] K. Xia and M. He, "Theory and practice of foreign car sharing services," *Urban Development Issues*, vol. 4, pp. 87–92, 2006.

[16] H. Becker, F. Ciari, and K. W. Axhausen, "Comparing car-sharing schemes in Switzerland: user groups and usage patterns," *Transportation Research Part A: Policy and Practice*, vol. 97, pp. 17–29, 2017.

[17] J. Kim, S. Rasouli, and H. Timmermans, "Satisfaction and uncertainty in car-sharing decisions: an integration of hybrid choice and random regret-based models," *Transportation Research Part A: Policy and Practice*, vol. 95, pp. 13–33, 2017.

[18] A. Gheorghiu and P. Delhomme, "For which types of trips do French drivers carpool? motivations underlying carpooling for different types of trips," *Transportation Research Part A: Policy and Practice*, vol. 113, pp. 460–475, 2018.

[19] H. Guan, *Disaggregate Model-A Tool for Traffic Behavior Analysis*, China Communications Press, Beijing, China, 2004.

[20] W. Zhang, Y. Qi, Y. Yan, J. Tang, and Y. Wang, "A method of emission and traveller behavior analysis under multimodal traffic condition," *Transportation Research Part D: Transport and Environment*, vol. 52, pp. 139–155, 2017.

[21] A. Lorimier and A. El-Geneidy, "Understanding the factors affecting vehicle usage and availability in car sharing networks: a case study of communauto car sharing system from Montréal, Canada," *International Journal of Sustainable Transportation*, vol. 7, no. 1, pp. 35–51, 2013.

[22] T. A. Domencich and D. L. McFadden, "Urban travel demand: a behavioral analysis," *Canadian Journal of Economics/revue Canadienne D`economique*, vol. 10, no. 4, 1975.

[23] P. Carroll, B. Caulfield, and A. Ahern, "Examining the potential for car-shedding in the greater Dublin area," *Transportation Research Part A: Policy and Practice*, vol. 106, pp. 440–452, 2017.

[24] S. Hess, K. E. Train, and J. W. Polak, "On the use of a modified Latin hypercube sampling (MLHS) method in the estimation of a mixed logit model for vehicle choice," *Transportation Research Part B: Methodological*, vol. 40, no. 2, pp. 147–163, 2006.

[25] G. Perboli, F. Ferrero, S. Musso, and A. Vesco, "Business models and tariff simulation in car-sharing services," *Transportation Research Part A: Policy and Practice*, vol. 115, pp. 32–48, 2018.

[26] N. Kalouptsidis and V. Psaraki, "Approximations of choice probabilities in mixed logit models," *European Journal of Operational Research*, vol. 200, no. 2, pp. 529–535, 2010.

[27] A. E. Watkins, A. Bargagliotti, and C. Franklin, "Simulation of the sampling distribution of the mean can mislead," *Journal of Statistics Education*, vol. 22, no. 3, p. 3, 2014.

[28] W. Zhang, F. Goerlandt, P. Kujala, and Y. Wang, "An advanced method for detecting possible near miss ship collisions from AIS data," *Ocean Engineering*, vol. 124, pp. 141–156, 2016.

[29] T. Yoon and R. C. Cherry, *EV (Electric Vehicle) Sharing Demand Estimation: A Case Study of Beijing, China*, KINTEX, Korea, 2015.

WILEY | Hindawi

Research Article

Examining the Impact of Adverse Weather on Travel Time Reliability of Urban Corridors in Shanghai

Yajie Zou [ID],[1] **Ting Zhu** [ID],[1] **Yifan Xie** [ID],[1] **Linbo Li** [ID],[1] **and Ying Chen** [ID][2]

[1]*Key Laboratory of Road and Traffic Engineering of Ministry of Education, Tongji University, Shanghai 201804, China*
[2]*School of Traffic and Transportation Engineering, Changsha University of Science & Technology, Changsha 410114, China*

Correspondence should be addressed to Linbo Li; llinbo@tongji.edu.cn

Travel time reliability (TTR) is widely used to evaluate transportation system performance. Adverse weather condition is an important factor for affecting TTR, which can cause traffic congestions and crashes. Considering the traffic characteristics under different traffic conditions, it is necessary to explore the impact of adverse weather on TTR under different conditions. This study conducted an empirical travel time analysis using traffic data and weather data collected on Yanan corridor in Shanghai. The travel time distributions were analysed under different roadway types, weather, and time of day. Four typical scenarios (i.e., peak hours and off-peak hours on elevated expressway, peak hours and off-peak hours on arterial road) were considered in the TTR analysis. Four measures were calculated to evaluate the impact of adverse weather on TTR. The results indicated that the lognormal distribution is preferred for describing the travel time data. Compared with off-peak hours, the impact of adverse weather is more significant for peak hours. The travel time variability, buffer time index, misery index, and frequency of congestion increased by an average of 29%, 19%, 22%, and 63%, respectively, under the adverse weather condition. The findings in this study are useful for transportation management agencies to design traffic control strategies when adverse weather occurs.

1. Introduction

Travel time reliability (TTR) is an important measure of traffic condition and is widely used to evaluate transportation system performance. The Federal Highway Administration (FHWA) [1] formally defined the TTR as "the consistency or dependability in travel times, as measured from day-to-day and/or across different times of the day." The TTR can be used to represent the probability of on-time arrival; thus, it is a key factor for roadway users to make decisions on travel routes and departure time. A survey conducted by Abdel-Aty et al. [2] shows that about 54% respondents consider TTR as an important factor for choosing commute routes. Unreliable travel times can cause significant inconvenience to roadway users and result in high time and monetary losses. Thus, some studies analysed the monetary value of TTR [3–5]. For instance, Lam [5] conducted a study in California and found that the value of travel time is $22.47 per hour (as 68 percentile of average

salary at that time), while the value of TTR is $31.16 per hour (as 95 percentile of average salary at that time). Meanwhile, there are several studies focusing on the relationship between TTR and traffic safety, and it is confirmed that the reduction of travel time reliability would lead to severe traffic collisions and crashes [6–10]. In recent years, some studies have found that improving travel time reliability can help reduce fuel consumption and emissions by avoiding stop-and-go movements of vehicles [11, 12]. Thus, improving the TTR on the entire road network can be essential and valuable.

To accurately evaluate the travel time reliability on transportation facilities, substantial studies focus on developing appropriate TTR measures. Lomax et al. [13] comprehensively summarized the common TTR measures and categorized them into four types: statistical range methods, buffer time methods, tardy-trips measures, and probabilistic measures. Table 1 provides a summary of the widely used TTR measures in recent years. Gao et al. [18]

TABLE 1: Summary of widely used TTR measures.

Category	Measure	Equation
Statistical range methods	Coefficient of deviation [14]	Standard deviation/average travel time
	Travel time variability [15]	$TT90 - TT10$
	Skew of travel time distribution [16]	$TT90 - TT50/TT50 - TT10$
	Width of travel time distribution [16]	$TT90 - TT10/TT50$
Buffer time methods	Planning time index [1]	95th percentile travel time/free flow travel time
	Buffer time index [13]	(95th percentile travel rate/free flow travel rate) $- 1$
	Travel time index [17]	Average travel time/free flow travel time
	Misery index [17]	(Average travel rate (top 20% trips)/average travel rate) $- 1$
Tardy-trip measures	On-time arrival [17]	Percent of trips with travel time less than 110% or 125% of the median travel time
Probabilistic measures	Frequency of congestion [1]	Frequency of trips experiencing traffic congestion

investigated the value of several TTR measures under different traffic conditions, and they found that the volatility trends of them were similar, while the magnitude of change values of Frequency of Congestion was significant. However, TTR measures demonstrate various characteristics. For example, the probabilistic measures can be used to calculate the probability of arrival before a certain time, which are user-friendly and can help travelers to plan trips. In recent years, TTR has been widely used to evaluate traffic conditions and help optimize transportation management based on these TTR measures [19–23]. For instance, Cedillo-Campos et al. [23] conducted a study to assess freight fluidity of transportation systems with the help of TTR measures (i.e., percentiles of travel time, planning time index, buffer time index, and skew and range of travel).

Knowledge about which factors affect TTR and how they influence TTR can help improve the travel time reliability. Based on previous studies, the factors affecting TTR can be classified into two categories [24], factors causing demand variation (such as season, time of day, and weather), and factors causing supply variation (such as traffic incident, road geometry, road work, and weather). It can be seen that weather influences the TTR on two ways. Several studies evaluated the impact of adverse weather on travel time reliability, and it is demonstrated that the adverse weather has negative effects on TTR above a certain critical inflow [15, 25–27]. However, most studies among them focus on freeways, and very few studies comprehensively evaluate the impact of adverse weather on urban corridor (especially arterial roadway) travel time reliability. Travel time and travel time reliability of different highway types are supposed to exhibit diverse patterns due to the difference of traffic characteristics. Thus, there is a need to explore the difference of adverse weather's impacts on TTR of urban corridor. In addition, the analysis of the impact of adverse weather on TTR of urban corridor under different traffic conditions can be useful for traffic management. The study has two primary objectives. The first objective is to explore the characteristics of urban corridor travel time. The second objective is to quantitatively estimate the impact of adverse weather on TTR under different conditions.

The rest of this paper is organized as follows. The second section introduces the single-mode distributions and travel time reliability measures used in this study to explore the

travel time characteristics of the urban corridor. In Section 3, the data preparation procedures are introduced. Section 4 illustrates the results and compares the impact of adverse weather on TTR under different conditions. Finally, Section 5 provides conclusions and identifies future directions for this research.

2. Methodology

2.1. Travel Time Distribution. Travel time distribution (TTD) describes the shape of travel time under different travel conditions. It provides a straightforward and visualized tool for modelling the average travel time and travel time variability. In this study, five commonly used single-mode distribution types are considered to fit travel time data, including Weibull distribution, Gamma distribution, Normal distribution, Lognormal distribution, and Log-logistic distribution. The details of five distributions are given in Table 2.

Based on the travel time data, the parameters of each distribution are estimated using software R. In order to compare the goodness-of-fit results of different distributions, the information criteria are adopted to select the better fit one. Information criteria mainly refer to Akaike information criterion (AIC) and Bayesian information criterion (BIC) [28].

AIC is a measure to compare relative quality of different models. Based on the principle of asymptotically unbiased estimation, Akaike proposed an exact estimation algorithm, as shown in

$$\text{AIC} = 2k - 2\ln(L), \tag{1}$$

where $k =$ the number of parameters and $L =$ the value of likelihood function.

Analogous to AIC, BIC introduces a penalty term associated with the parameter number to avoid the occurrence of over fitting problem. Thus, only AIC is calculated to evaluate the performance of five distributions. A smaller AIC value indicates better fitting performance.

2.2. Travel Time Reliability Measures. As aforementioned, there are several measures to describe travel time reliability. In this study, four widely used measures are adopted to evaluate the TTR, and they are travel time variability (TTV),

Table 2: Description of five distributions.

	Function	Parameters
Weibull distribution	$f(y\mid\gamma,\beta) = (\gamma/\beta)y^{\gamma-1}e^{-y^{\gamma}/\beta}$	γ, β
Gamma distribution	$f(x\mid\alpha,\beta) = (1/\Gamma(\alpha)\beta^{\alpha})x^{\alpha-1}e^{-x/\beta}$	α, β
Normal distribution	$f(x\mid\mu,\sigma^2) = (1/\sqrt{2\pi}\sigma)e^{-(x-\mu)^2/2\sigma^2}$	μ, σ^2
Lognormal distribution	$f(x\mid\mu,\sigma^2) = (1/\sqrt{2\pi}\sigma x)e^{-(\log x-\mu)^2/2\sigma^2}$	μ, σ^2
Log-logistic distribution	$f(x\mid\mu,\sigma) = (1/\sigma)e^{(x-\mu/\sigma)}/(1+e^{(x-\mu/\sigma)})^2$	μ, σ

buffer time index (BTI), misery index (MI), and frequency of congestion (FOC).

Travel time variability (TTV) is a statistical range method, representing the difference between the 90th and 10th percentile travel time, as shown in (2). Some previous studies found that the statistical properties of TTV can be more robust than moment-based measures (such as mean and deviation) [29, 30]:

$$TTV = TT90 - TT10, \tag{2}$$

where $TT90 = 90$th percentile travel time (min) and $TT10 = 10$th percentile travel time (min).

Buffer time index (BTI) is a typical buffer time method, proposed by Lomax et al. [13]. It is calculated by (3), which reflects the percentage of extra travel time that most travelers need to add to the mean travel time in order to arrive on time. This index depends on the travel rate, which is calculated as the travel time divided by the distance:

$$BTI = \frac{95\text{th percentile travel rate} - \text{average travel rate}}{\text{average travel rate}}, \tag{3}$$

where the 95th percentile travel rate and average travel rate are in minutes per km.

Misery index (MI) proposed by Lomax is a tardy-trip measure. Similar to BTI, it is related to travel rate as well as the worst trips. This TTR measure can represent the traffic conditions and identify the roadway segments experiencing congestion. MI is defined as the ratio of the difference between the average travel rate for the longest 20% of trips and average travel rate to the average travel rate, and the equation is provided as follows:

$$MI = \frac{\text{average travel rate of the top 20\% trips} - \text{average travel rate}}{\text{average travel rate}}, \tag{4}$$

where two average travel rates are in minutes per km.

Frequency of congestion (FOC) represents the frequency of trips exceeding a threshold value, which is a probabilistic measure. This research assumed that the traffic congestion occurs when the travel time is higher than twice the free-flow travel time. The free-flow speed of each route is obtained as the 85th percentile speed during overnight hours (10 p.m. to 5 a.m.) [9, 31, 32]. Note that the free-flow travel time of each route is inversely proportional to free-flow speed; thus, the free-flow travel time is defined as the 15th percentile travel time during overnight hours. The FOC can be calculated by using the following:

$$G(p) = 2G(0.15), \tag{5}$$

$$FOC = 1 - p, \tag{6}$$

where p is the frequency of travel time higher than twice the free-flow travel time.

3. Data Description

Two datasets are adopted in this study to investigate the impact of adverse weather on TTR of urban corridor. The traffic data collected from the Gaode Map website are used to generate travel times, while the weather data associated with the travel times are collected from China Meteorological Administration.

3.1. Description of Study Site.
Four routes along Yanan elevated expressway and arterial road are selected for this study. The Yanan elevated expressway has six lanes and connects the Hongqiao International Airport to the Bund of Shanghai, as shown in Figure 1. The selected study section of the elevated expressway is 14.32 km long and covers the eastbound and westbound directions. The Yanan elevated expressway is basically constructed along Yanan West Road, Yanan Middle Road, and Yanan East Road. Thus, the corresponding arterial road is around 14.32 km long as well, and traffic data for both directions are collected.

The main reasons for choosing these four routes are summarized as follows. Previous studies suggest that adverse weather has minor effect on TTR for free flow traffic condition but can significantly affect TTR at high traffic flow level [15, 30]. In addition, the Yanan elevated expressway is a main corridor in Shanghai and experiences heavy traffic during the peak hours.

3.2. Description of Travel Time Data.
The Gaode Map divides one route into fixed consecutive sections of the specified length and sets five congestion levels. From the end of the road, it records the congestion levels of road sections. Adjacent road sections with the same congestion level converge on one piece of data, recording starting point, ending point, direction, and average speed. The raw traffic data are collected every 4 minutes. Traffic data from October 2019 to February 2020 are used in this study, which contain the following information: road ID, driving direction, congestion level, speed, timestamp and so on. This study aims to analyse the differences of adverse weather's impact on TTR during peak and off-peak hours. Thus, the traffic data collected from 6 : 00 a.m. to 22 : 00 p.m. on weekdays are utilized.
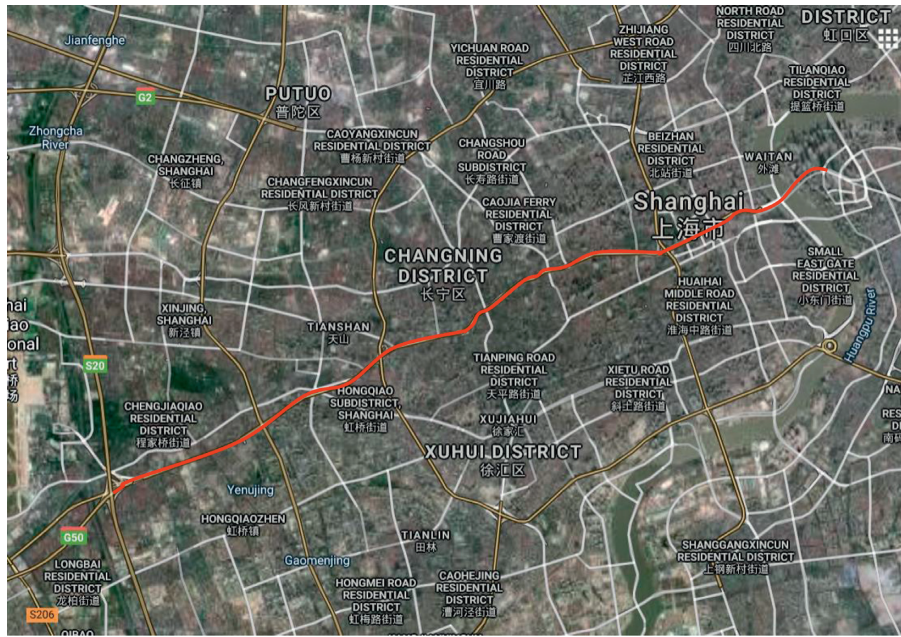
FIGURE 1: Map of Yanan corridor in Shanghai, China.

After excluding the traffic data with invalid or missing essential information, route-level travel times are obtained through several steps. First, calculate single section travel time based on each piece of data. Based on the starting and ending points, the length of section is determined. Then, the single section travel time is calculated by dividing the length of section by average speed. Second, route-level travel times are taken by summing all sections' travel time.

3.3. Description of Weather Data. The historical weather data of four districts (Huangpu, Changning, Jing'an, and Baoshan districts) within the study time period are obtained. The raw weather data classify weather conditions into four groups, including normal, overcast, mostly cloudy, and rain. According to previous studies, it is demonstrated that only adverse weather can cause a significant impact on traffic conditions [15]. Thus, "overcast" and "mostly cloudy" are classified into "normal" in this study due to no obvious influence on travel times. Meanwhile, this study tried to classify "rain" into more types (e.g., snow, ice, and fog) based on detailed weather information such as temperature, visibility, and humidity. In order to ensure the acceptable sample size, this study combined the adverse weather together as "rain" due to the rarity of other adverse weather types. Therefore, the weather variables are categorized as "normal weather" and "adverse weather."

Two datasets are matched based on timestamp. However, the weather data are collected every 20 minutes, while the travel time data are aggregated into 4-minute intervals. In order to merge the link travel time data with weather data, the weather data are expanded by filling the absence as the weather condition of the most recent timestamp. The Yanan elevated expressway covers four districts. To improve the

accuracy of analysis, this study only selects the data when the weather conditions of four districts are the same for research.

The average travel time (in 15-minute period) of each route under different weather conditions were calculated and shown in Figure 2. Note that in Figure 2, the horizontal axis denotes the day from 6:00 to 22:00 with 15 minutes as a time interval, and the vertical axis represents the mean of the travel times during desired time period. Previous studies [15, 33] concluded that time intervals between 10 and 15 minutes are suitable for analysing traffic conditions. For a better visual effect, this paper adopted a 15-minute time interval to draw scatter plots.

According to Figure 2, there are several points worth mentioning. First, the difference of the travel time between normal weather condition and adverse weather condition is remarkable during the peak time. This finding indicates that the adverse weather has significant impact on travel time only when traffic volume is above a certain value. Second, there is no a.m. peak hour characteristic of travel time in Figure 2(b). This can be probably explained by the influence of commuting resulting in the difference of traffic volume between driving directions at the same time period. Thus, it is also worth exploring the difference of TTR with respect to driving direction. Third, the travel times of elevated expressway are high at noon (13:00–15:00). The main reason for this could be that only vehicles with Shanghai license plate are permitted to enter Yanan elevated expressway after 15:00. Therefore, before that time, a majority of cars without a Shanghai license plate leave the city centre through the elevated expressway.

For analysing the impact of adverse weather on TTR under different traffic conditions, two typical time periods are selected as peak hours, that is, AM peak hours (7:30–9:00) and PM peak hours (16:30–18:30).
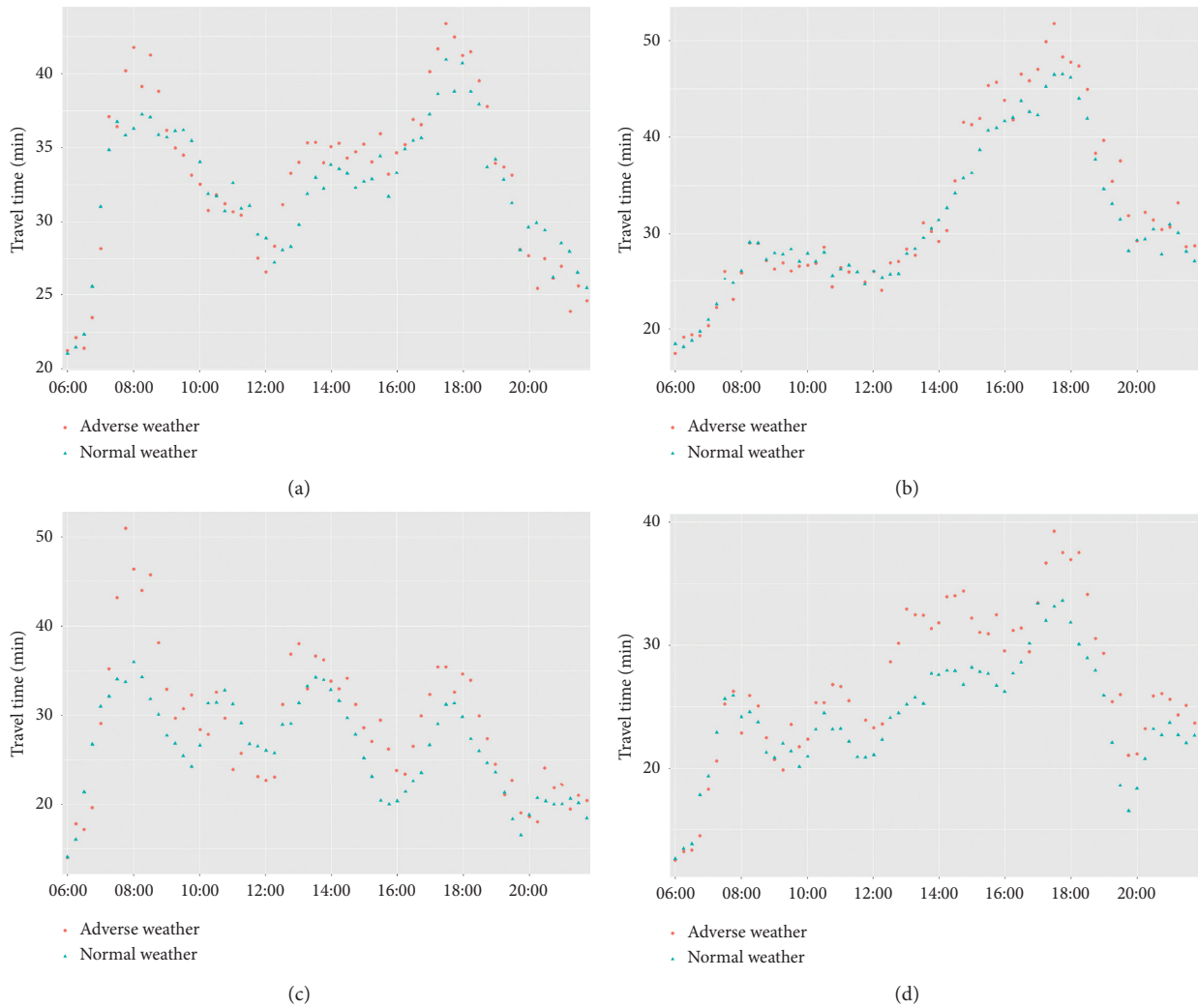
Figure 2: Travel time of each route under different weather conditions. (a) Yanan arterial road Eastbound. (b) Yanan arterial road Westbound. (c) Yanan elevated expressway Eastbound and (d) Yanan elevated expressway Westbound.

4. Results and Discussion

4.1. Travel Time Distribution Analysis. Five distribution types were fitted with the observed travel time data for each route. In this paper, the best-fit distribution is defined as the one with the minimum AIC value. Figure 3 shows the probability density of the best-fit distribution and the observed travel time data for each route. The legend displays the distribution type and parameters of the best-fit distribution. It can be found that the lognormal distribution provides the best model fit for all routes. According to the density curves of observed travel time data, the average travel time on arterial road is higher than that of expressway.

To illustrate the travel time distribution under different weather conditions, the travel time data are divided into four categories: elevated expressway data under adverse weather and normal weather conditions, arterial road data under adverse weather and normal weather conditions. Then, five distribution types are fitted with the travel time data.

Figure 4 shows the probability density of the best fitting distribution and the observed travel time data under different weather conditions. Based on Figure 4, it can be found that the lognormal distribution fits better than other distributions.

To illustrate the travel time distribution under different time periods, the travel time data are divided into four categories. Then, five distribution types are fitted with the travel time data for each dataset. Figure 5 shows the probability density of the best-fit distribution and the observed travel time data during different time periods. The results show that the lognormal distribution fits best for most scenarios, except the arterial road data observed during peak hours.

4.2. TTR Analysis Results for Different Traffic Directions. To investigate the travel time reliability of four studied routes, the TTR measures of each route are calculated, as shown in Figure 6. The travel time on westbound Yanan arterial road is less reliable. The BTI and MI of westbound
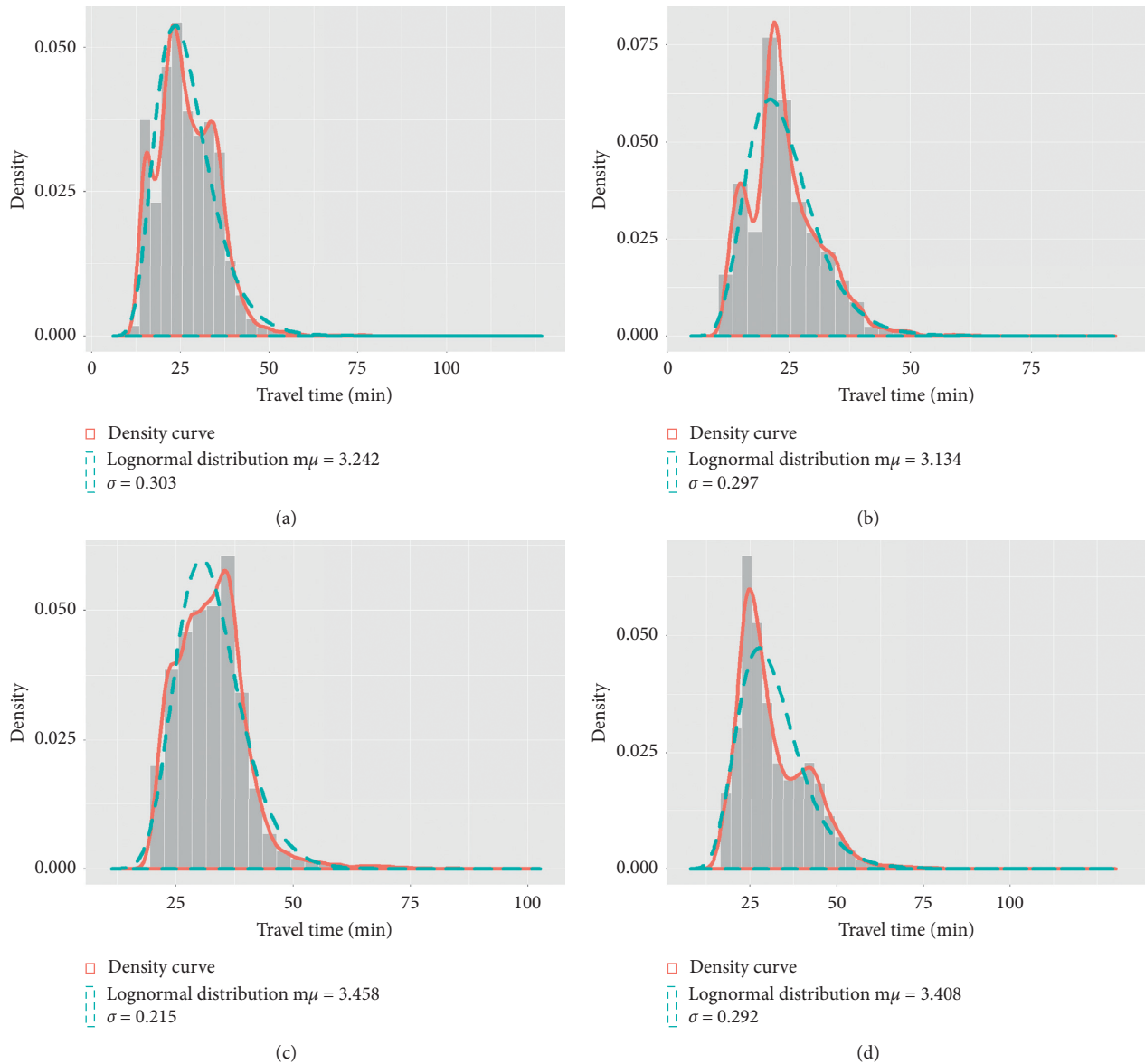
FIGURE 3: Selected distributions for the observed travel time data of four routes. (a) Yanan elevated expressway Eastbound, (b) Yanan elevated expressway Westbound. (c) Yanan arterial road Eastbound, and (d) Yanan arterial road Westbound.

Yanan arterial road have obviously higher values than those of eastbound. The BTIs (and MIs) of these two routes are 0.56 (0.50) and 0.34 (0.32), respectively. The possible reason for this is that tidal traffic phenomenon on Yanan arterial road westbound can result in more frequent traffic congestions during afternoon peak hours.

4.3. Impact of Adverse Weather on TTR. The analysis results in Section 4.1 indicate that the travel time distribution characteristics can be different with the consideration of time period and highway type. In order to investigate the impacts of adverse weather on TTR under different conditions, four scenarios are studied in this paper:

   (i) Scenario 1: the traffic on elevated expressway during off-peak hours

   (ii) Scenario 2: the traffic on elevated expressway during peak hours

   (iii) Scenario 3: the traffic on urban arterial road during off-peak hours

   (iv) Scenario 4: the traffic on urban arterial road during peak hours

For each scenario, the data are divided into two categories (adverse weather and normal weather) and four TTR measures are calculated. The results are shown in Figure 7, and there are three main findings. First, adverse weather has negative impacts on the TTR. For all scenarios, the values of four TTR measures all increase to some extent under adverse weather conditions. Second, travel times are less reliable during peak hours than that during off-peak hours. Third, the results suggest that the travel times are more reliable on elevated expressway.
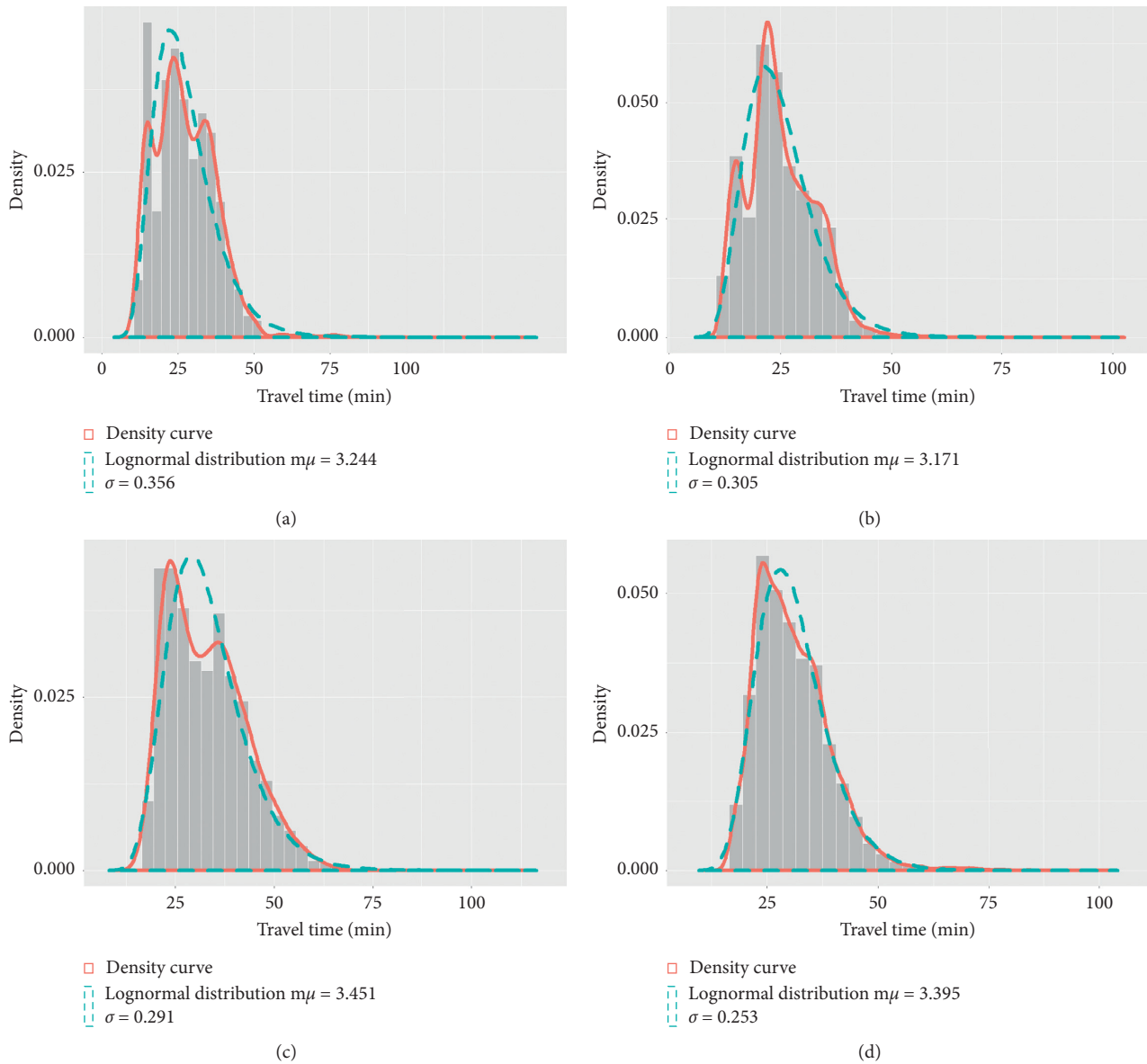
FIGURE 4: Selected distributions for the observed travel time data under different weather conditions. (a) Adverse weather of elevated expressway, (b) normal weather of elevated expressway, (c) adverse weather of urban arterial road, and (d) normal weather of urban arterial road.

Then, compared with the normal weather condition, the increase of TTR measures under adverse weather condition is calculated for each scenario. It is computed as the difference of TTR measure between two weather conditions divided by the TTR measure under normal weather conditions. The results on Table 3 show that, compared with off-peak hours, adverse weather has larger influences on TTR during peak hours. In the latter case, travel time variability under adverse conditions increased by an average of 29% when compared with normal conditions, while the buffer time index, misery index, and frequency of congestion increased by an average of 19%, 22%, and 63%, respectively. The possible reason is that the volume of traffic during peak hours is high, and, therefore, traffic collisions and congestions are more likely to occur during adverse weather.

4.4. Impact of Adverse Weather on Travel Delay. This section calculates the travel time delay under different conditions. The average delay was computed by subtracting the calculated travel time from the free-flow travel time for each four minutes period. As mentioned before, the free-flow travel time is defined as the 15th percentile of travel time during overnight hours. The results are displayed in Figure 8(a). It is apparent that the adverse weather causes an overall larger travel delay. The possible explanation is that the adverse weather can result in traffic congestion. Note that the difference of the average travel delay between adverse weather and normal weather during off-peak period is not very noticeable but is remarkable during peak period. This can be explained by the fact that adverse weather may result in higher rates of the traffic congestion at higher inflow levels.
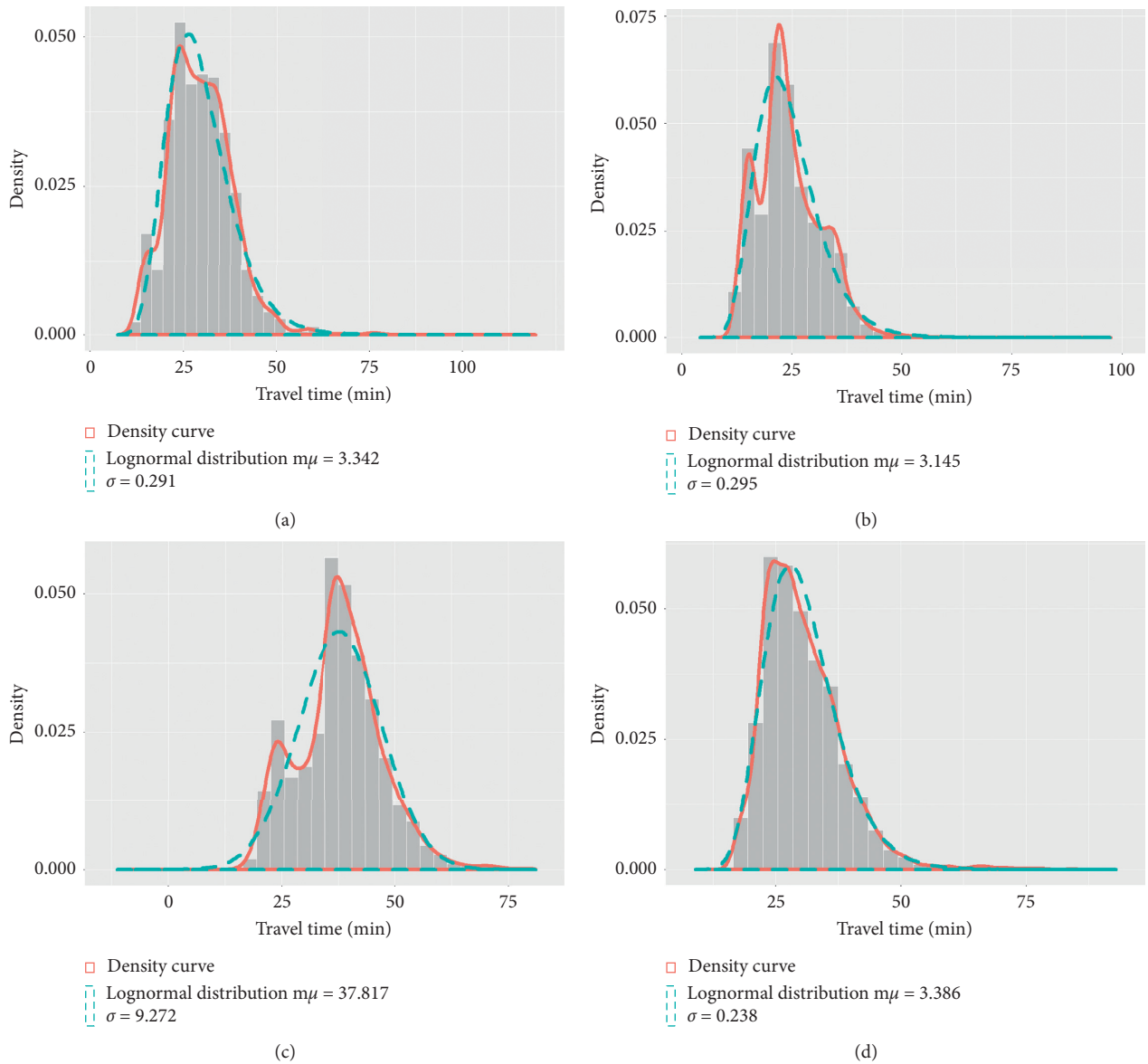
FIGURE 5: Selected distributions for the observed travel time data during different time periods. (a) Peak hours of elevated expressway, (b) off-peak hours of elevated expressway, (c) peak hours of urban arterial road, and (d) off-peak hours of urban arterial road.
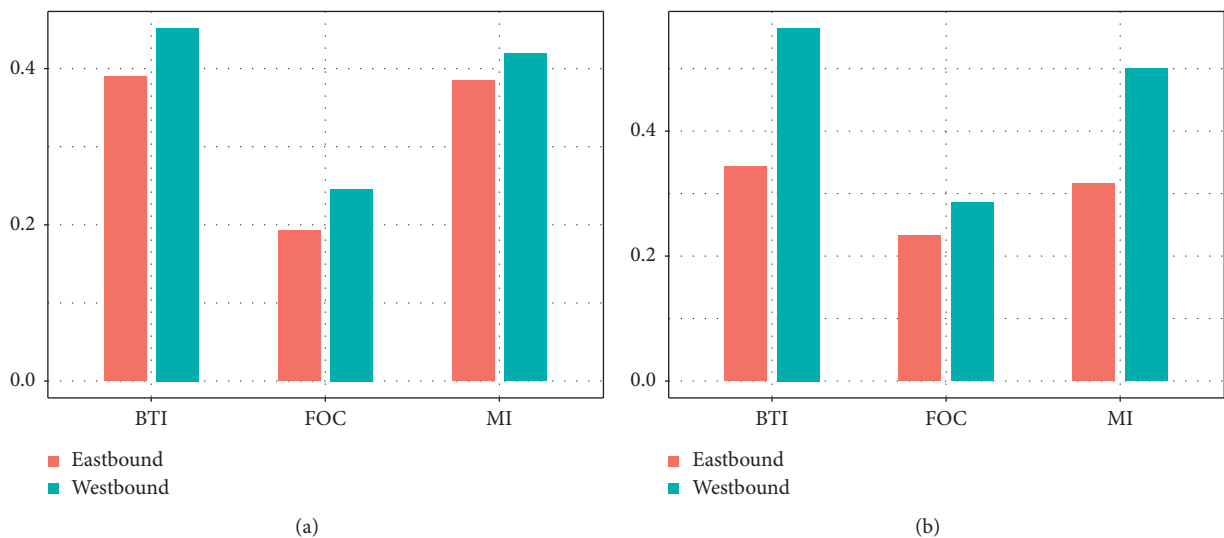
FIGURE 6: Buffer time index, misery index, and frequency of congestion of four routes. (a) Elevated expressway and (b) urban arterial road.

(a)

(b)

(c)

(d)
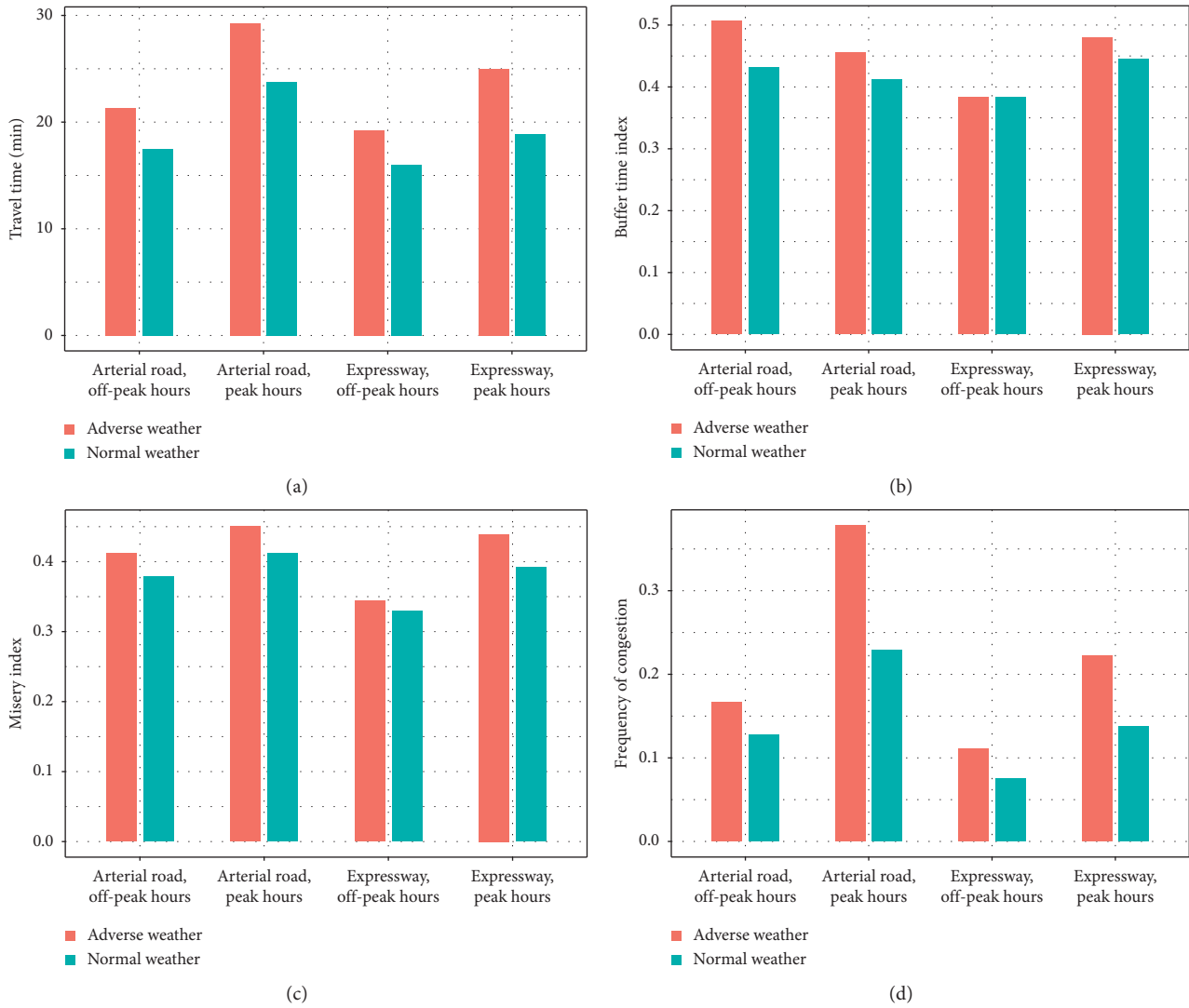
FIGURE 7: TTV, BTI, MI, and FOC under different scenarios. (a) TTV, (b) buffer time index, (c) misery index, and (d) frequency of congestion.

TABLE 3: The increase of TTR measures under adverse weather condition.

Roadway type	Time	TTV (%)	BTI (%)	MI (%)	FOC (%)
Elevated expressway	Peak hours	33	8	12	61
	Off-peak hours	21	8	4	49
Urban arterial road	Peak hours	24	30	31	65
	Off-peak hours	22	17	9	31

The average delay time rate is defined as the ratio of the average delay to free-flow travel time in this paper. The results are summarized in Figure 8(b) and show that, during off-peak hours, the average delay for adverse weather increases by 6–9% compared to normal weather. The average delay increases by 18–30% during peak hours. Meanwhile, the average delay time rate of elevated expressway is larger than that of the arterial road under the same traffic and weather condition. Due to the speed limit and intersections, the free-flow travel time of arterial road is higher.
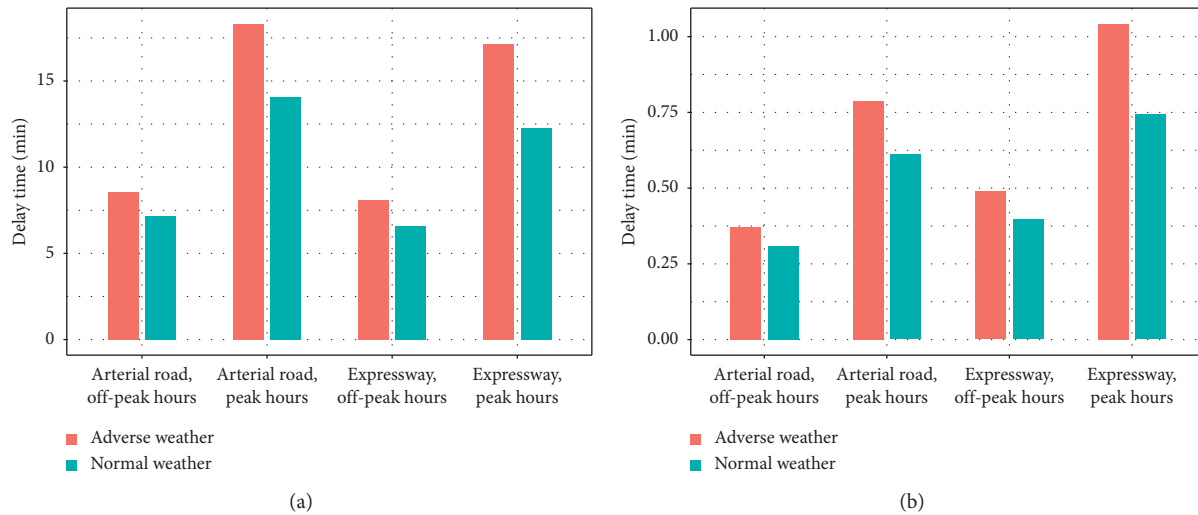
FIGURE 8: Travel delay under each condition. (a) Average delay and (b) average delay time rate.

5. Conclusions

In order to investigate the impact of adverse weather on travel time reliability of urban corridor under different conditions, this paper conducted an empirical study by using traffic data and weather data collected on Yanan corridor. Four typical scenarios with different highway types and time periods were studied in this paper, and the impact of adverse weather on urban corridor TTR under each scenario was evaluated against normal weather conditions. In addition, this study explored the distribution characteristics of travel time on urban roads with respect to the roadway type, time of day, and weather. Hence, the findings in this study are useful for determining traffic control strategies to address the adverse weather-related traffic congestions. And the distribution fitting results are useful for predicting travel time. The main findings can be summarized as follows.

(1) Among five widely used single-mode distribution types (i.e., Weibull, Gamma, Normal, Lognormal, and Log-logistic), the Lognormal distribution outperforms other models for most conditions, except the urban arterial road data observed during peak hours.

(2) Adverse weather clearly shows negative impacts on the travel time reliability of urban corridor, and the magnitude of the effects can be different under different scenarios. Compared to off-peak hours, adverse weather has larger influences on TTR during peak hours.

(3) During peak hours, travel time variability under adverse weather conditions increased by an average of 29% when compared to normal conditions, while the buffer time index, misery index and frequency of congestion increased by an average of 19%, 22%, and 63%, respectively.

For future work, the choice of distribution types can be expanded when analysing the travel time distribution, since the multimode distribution can provide a better fitting performance under some traffic conditions [34–37]. This paper mainly examined the difference of rainy weather's impact on TTR under different conditions; thus, the difference of other adverse weather should also be investigated if the sample size for each category is sufficient. In addition, the traffic composition and traffic flow rate should also be further considered in the future study.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] D. Fha, *Travel Time Reliability: Making it There on Time, All the Time*, US Department of Transportation, Federal Highway Administration, Washington, NJ, USA, 2010.

[2] M. Abdel-Aty, R. Kitamura, and P. P. Jovanis, "Investigating effect of travel time variability on route choice using repeated-measurement stated preference data," *Transportation Research Record*, vol. 1493, pp. 39–45, 1995.

[3] Z. Li, D. A. Hensher, and J. M. Rose, "Willingness to pay for travel time reliability in passenger transport: a review and some new empirical evidence," *Transportation Research Part E: Logistics and Transportation Review*, vol. 46, no. 3, pp. 384–403, 2010.

[4] C. Chen, A. Skabardonis, and P. Varaiya, "Travel-time reliability as a measure of service," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1855, no. 1, pp. 74–79, 2003.

[5] T. C. Lam, *The Effect of Variability of Travel Time on Route and Time-Of-Day Choice*, University of California, Oakland, CA, USA, 2000.

[6] Q. Shi and M. Abdel-Aty, "Evaluation of the impact of travel time reliability on urban expressway traffic safety," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2582, no. 1, pp. 8–17, 2016.

[7] E. Kidando, R. Moses, E. E. Ozguven, and T. Sando, "Evaluating traffic congestion using the traffic occupancy and speed distribution relationship: an application of Bayesian Dirichlet process mixtures of generalized linear model," *Journal of Transportation Technologies*, vol. 7, no. 3, p. 318, 2017.

[8] R. J. Javid and R. J. Javid, "A framework for travel time variability analysis using urban traffic incident data," *IATSS Research*, vol. 42, no. 1, pp. 30–38, 2018.

[9] W. Fan and L. Gong, *"Developing a Systematic Approach to Improving Bottleneck Analysis in North Carolina," FHWA/NC/2016-10*, North Carolina Dept. of Transportation, Raleigh, NC, USA, 2017.

[10] A. T. Hojati, L. Ferreira, S. Washington, P. Charles, and A. Shobeirinejad, "Reprint of: modelling the impact of traffic incidents on travel time reliability," *Transportation Research Part C: Emerging Technologies*, vol. 70, pp. 86–97, 2016.

[11] M. Figliozzi and A. Bigazzi, *Value of Travel-Time Reliability, Part II: A Study of Tradeoffs between Travel Reliability, Congestion-Mitigation Strategies and Emissions*, Oregon Transportation Research and Education Consortium, Portland, OR, USA, 2012.

[12] F. Yang, M.-P. Yun, and X.-G. Yang, "Travel time distribution under interrupted flow and application to travel time reliability," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2466, no. 1, pp. 114–124, 2014.

[13] T. Lomax, D. Schrank, S. Turner, and R. Margiotta, *Report of Selecting Travel Reliability Measures*, Federal Highway Administration, Washington, NJ, USA, 2003.

[14] W. Pu, "Analytic relationships between travel time reliability measures," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2254, no. 1, pp. 122–130, 2011.

[15] H. Tu, H. W. Van Lint, and H. J. Van Zuylen, "Impact of adverse weather on travel time variability of freeway corridors," in *Proceedings of the Transportation Research Board 86th Annual Meeting*, Washington, NJ, USA, January 2007.

[16] J. W. C. Van Lint and H. J. van Zuylen, "Monitoring and predicting freeway travel time reliability: using width and skew of day-to-day travel time distribution," *Transportation Research Record: Journal Transportation Research Record*, vol. 1917, no. 1, pp. 54–62, 2005.

[17] P. Chen, R. Tong, G. Lu, and Y. Wang, "Exploring travel time distribution and variability patterns using probe vehicle data: case study in Beijing," *Journal of Advanced Transportation*, vol. 2018, Article ID 3747632, 13 pages, 2018.

[18] L. Gao, *Research on Reliability of Path Travel Time Based on Dynamic Travel Time*, Hebei University of Technology, Tianjin, China, 2011.

[19] E. Milliken, R. Young, and M. P. Consortium, *Use of Travel Time, Travel Time Reliability, and Winter Condition Index Information for Improved Operation of Rural Interstates*, Mountain Plains Consortium, Fargo, ND, USA, 2015.

[20] B. Yao, B. Hu, X. Lu, J. Gao, and M. Zhang, "Transit network design based on travel time reliability," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 233–248, 2014.

[21] S. Mathew and S. S. Pulugurtha, "Assessing the effect of a light rail transit system on road traffic travel time reliability," *Public Transport*, vol. 12, pp. 1–21, 2020.

[22] K. Lyman and R. L. Bertini, "Using travel time reliability measures to improve regional transportation planning and operations," *Transportation Research Record*, vol. 2046, no. 1, pp. 1–10, 2008.

[23] M. G. Cedillo-Campos, C. M. Pérez-González, J. Piña-Barcena, and E. Moreno-Quintero, "Measurement of travel time reliability of road transportation using GPS data: a freight fluidity approach," *Transportation Research Part A: Policy and Practice*, vol. 130, pp. 240–288, 2019.

[24] M. A. P. Taylor, "Dense network traffic models, travel time reliability and traffic management. II: application to network reliability," *Journal of Advanced Transportation*, vol. 33, no. 2, pp. 235–251, 1999.

[25] Z. Li, L. Elefteriadou, and A. Kondyli, "Quantifying weather impacts on traffic operations for implementation into a travel time reliability model," *Transportation Letters*, vol. 8, no. 1, pp. 47–59, 2016.

[26] I. Tsapakis, T. Cheng, and A. Bolbol, "Impact of weather conditions on macroscopic urban travel times," *Journal of Transport Geography*, vol. 28, pp. 204–211, 2013.

[27] M. A. Yazici, C. N. Kamga, and A. Singhal, "Weather's impact on travel time and travel time variability in New York city," *Transportation Research*, vol. 40, p. 41, 2013.

[28] K. P. Burnham and D. R. Anderson, "Multimodel inference," *Sociological Methods & Research*, vol. 33, no. 2, pp. 261–304, 2004.

[29] J. W. C. Van Lint, H. J. Van Zuylen, and H. Tu, "Travel time unreliability on freeways: why measures based on variance tell only half the story," *Transportation Research Part A: Policy and Practice*, vol. 42, no. 1, pp. 258–277, 2008.

[30] H. Tu, H. Li, H. Van Lint, and H. Van Zuylen, "Modeling travel time reliability of freeways using risk assessment techniques," *Transportation Research Part A: Policy and Practice*, vol. 46, no. 10, pp. 1528–1540, 2012.

[31] Z. Chen and W. Fan, "Data analytics approach for travel time reliability pattern analysis and prediction," *Journal of Modern Transportation*, vol. 27, no. 4, pp. 250–265, 2019.

[32] D. Schrank, B. Eisele, T. Lomax, and J. Bak, *Urban Mobility Scorecard*, The Texas A&M Transportation Institute and The Texas A&M University System, College Station, TX, USA, 2015.

[33] B. L. Smith and J. M. Ulmer, "Freeway traffic flow rate measurement: investigation into impact of measurement time interval," *Journal of Transportation Engineering*, vol. 129, no. 3, pp. 223–229, 2003.

[34] S. Park, H. Rakha, and F. Guo, "Calibration issues for multistate model of travel time reliability," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2188, no. 1, pp. 74–84, 2010.

[35] H. Rakha, I. El-Shawarby, and M. Arafeh, "Trip travel-time reliability: issues and proposed solutions," *Journal of Intelligent Transportation Systems*, vol. 14, no. 4, pp. 232–250, 2010.

[36] X. Zhong, Y. Zou, Z. Dong, S. Yuan, and M. Ijaz, "Finite mixture survival model for examining the variability of urban arterial travel time for buses, passenger cars and taxis," *IET Intelligent Transport Systems*, vol. 45, pp. 1524–1533, 2020.

[37] L. Wang, H. Zhong, W. Ma, M. Abdel-Aty, and J. Park, "How many crashes can connected vehicle and automated vehicle technologies prevent: a meta-analysis," *Accident Analysis & Prevention*, vol. 136, Article ID 105299, 2020.

WILEY | Hindawi

Research Article

Multilane Spatiotemporal Trajectory Optimization Method (MSTTOM) for Connected Vehicles

Pangwei Wang [ID],[1] **Yunfeng Wang** [ID],[1] **Hui Deng** [ID],[1] **Mingfang Zhang** [ID],[1] **and Juan Zhang** [ID][2]

[1]*Beijing Key Lab of Urban Intelligent Traffic Control Technology, North China University of Technology, Beijing 100144, China*
[2]*College of Engineering, Mathematics and Physical Sciences, University of Exeter, EX4 4QF, Exeter, UK*

Correspondence should be addressed to Juan Zhang; jz397@exeter.ac.uk

It is agreed that connected vehicle technologies have broad implications to traffic management systems. In order to alleviate urban congestion and improve road capacity, this paper proposes a multilane spatiotemporal trajectory optimization method (MSTTOM) to reach full potential of connected vehicles by considering vehicular safety, traffic capacity, fuel efficiency, and driver comfort. In this MSTTOM, the dynamic characteristics of connected vehicles, the vehicular state vector, the optimized objective function, and the constraints are formulated. The method for solving the trajectory problem is optimized based on Pontryagin's maximum principle and reinforcement learning (RL). A typical scenario of intersection with a one-way 4-lane section is measured, and the data within 24 hours are collected for tests. The results demonstrate that the proposed method can optimize the traffic flow by enhancing vehicle fuel efficiency by 32% and reducing pollutants emissions by 17% compared with the advanced glidepath prototype application (GPPA) scheme.

1. Introduction

The traffic capacity of urban road network is restricted by the time delay, potential safety hazards, and environmental pollution caused by traffic congestion. According to the US Department of Transportation Statistics issued in 2019, 27% of highways in the United States are blocked at the peak of urban traffic and 54% of vehicles are in the state of congestion [1]. Intersections, as important nodes of urban road network, cause 80% of these urban traffic congestions [2]. Therefore, alleviating urban traffic congestion by taking the intersection as the optimization object is a challenging work.

The realization of spatiotemporal trajectory optimization mainly includes two parts: (1) realizing the dynamic speed control of multiple vehicles in the longitudinal direction; (2) achieving the cooperative lane-changing control of multiple vehicles in the horizontal direction. In the intelligent traffic system (ITS) strategic plan published by the US Department of Transportation in 2010, dynamic speed coordination based on spatiotemporal trajectory was denoted as one of the important methods for traffic flow optimization of road network [3]. Grumert and Tapani [4] established a variable speed limit (VSL) algorithm based on the traffic occupancy. The results show that VSL can increase traffic flow by 16%. Jo et al. [5] updated a VSL algorithm by detecting vehicles at multiple stations from the perspective of road safety, which solved the guidance delay problem. With the development of cooperative vehicle infrastructure system (CVIS) and connected vehicles (CVs), the problem of discontinuous dynamic speed limit can be solved by continuous dynamic speed control [6]. Considering the random traffic conditions, Zhu and Ukkusuri [7] proposed a method to realize dynamic speed limit control through RL. In this method, a link based dynamic network load model was established in the environment of CVs, which can reduce the total travel time by 18%. Aldana-Muñoz et al. [8] proposed a method by combining environmental driving and speed guidance to optimize the continuous linear spatiotemporal trajectory. The experimental results show that the linear spatiotemporal trajectory optimization can stabilize the change of

disturbance on traffic flow. Wang et al. [9] proposed a speed guidance model based on model predictive control (MPC), by which the spatiotemporal trajectory can be constructed in advance. Wei et al. [10] proposed a cooperative optimization method for multivehicle spatiotemporal trajectories, which is established based on the simplified Newell linear car-following model. Considering different road environments and weather conditions, De Mello and Chiodi [11] proposed a fuzzy logic control model of spatiotemporal trajectory, which increases the effectiveness of decision-making process by 13%. Liang et al. [12] presented a method to optimize the common spatiotemporal trajectories for both autonomous vehicles and manual driving vehicles. With the proportion of autonomous vehicles increasing, average delay time is reduced by 25% and average stopping time is reduced by 47%. Wang et al. [13, 14] designed a joint control model for CVs and traffic signal, which aims to optimize vehicular platoon and signal timing to improve traffic efficiency at intersections. The simulation results show that the joint control model can reduce the stopping time by 54%. Mirheli et al. [15] proposed a distributed cooperative control logic based on spatiotemporal trajectory, which can be applied to unsignalized intersections. In this study, vehicles with this logic application scheme can pass through the unsignalized intersection in optimal trajectories without collision. Wang et al. [16] developed a cooperative optimization method of spatiotemporal trajectory with multiple intersection phases. In this method, a lower bound estimator based on dynamic load network was constructed to improve robustness and stability of vehicle dynamic network.

Urban roads or expressways with single lane or dual lane were mainly taken in the above studies; however, the multilane spatiotemporal trajectory optimization is still lacked in the research. The main challenge of vehicles in the multilane spatiotemporal trajectory lies in how to solve its complex lane-changing problem. Moreover, accidents caused by lane-changing behaviors make the multiple lanes occupied, which makes the road congestion more serious [17]. Considering the risk of lane-changing produced by vehicle-to-vehicle (V2V) communication delay, Hongil and Jae-Il [18] proposed an oriented bounding box (OBB) method which reduces communication delay by defining the lateral and longitudinal gradient controllers. Li et al. [19] proposed a new lane-changing model based on CVs to evaluate the critical vehicle distance in lane-changing process. The safety potential field theory was integrated, and the vehicle movement state under different speed and acceleration conditions is dynamically presented. Finally, the lane-changing route can be determined through the selection of spatiotemporal trajectory. The majority of studies provide ideas and theoretical support for multivehicle cooperative lane-changing based on spatiotemporal trajectory. However, due to the massive amount of data calculation and high system complexity in multilane spatiotemporal trajectories, the common multilane road section traffic efficiency problem has not been effectively solved and improved through the optimization of multiple single lanes method.

To solve the above problems, MSTTOM for connected vehicles is established in this paper. The spatial temporal trajectories are generated and combined with vehicular current state information and the signal timing information. In addition, during the driving process, the trajectories can be optimized by RL. Moreover, the vehicle can change the lane and update its trajectory through the cooperative lane-changing strategy designed in this model. The significant contribution of this research is that vehicles can pass through intersections in optimal spatiotemporal trajectory and further solve the congestion problem in current multilane road section.

The rest of this paper is organized as follows: The typical application scenario and the structure of the method are introduced in Section 2. The specific content of MSTTOM is described in Section 3, where the trajectory is generated and optimized. The simulation experiment results through SUMO/Python are presented and analyzed in Section 4, and the stability of headway, fuel consumption rate, pollutant emission, and calculation time are also selected as the evaluation indexes. Finally, Section 5 concludes the paper.

2. Typical Application Scenario for Connected Vehicles

2.1. Application Scenario. The upstream section of the intersection, as the application of multilane scenario, is covered by vehicle-to-everything (V2X) wireless communication and shown in Figure 1. The vehicular trajectory from upstream to intersection process will be dynamically calculated based on MSTTOM.

2.2. System Structure. The flow chart of MSTTOM is shown in Figure 2. The procedure can be described as follows: the traffic signal and vehicular driving information is collected firstly by road-side unit (RSU) and on-board unit (OBU), so that the target lane and arrival time at current intersection can be calculated; secondly, the optimal trajectory is generated through connected vehicular state and constraints, where traffic capacity of current intersection can be optimized online through the cooperative lane-changing strategy; and finally, every connected vehicle can pass through the intersection with least travel time through the optimization result of RL. The spatiotemporal trajectory and the optimized state are clarified in Figure 3.

3. Multilane Spatiotemporal Trajectory Optimization Method (MSTTOM)

Firstly, state vectors defined in this paper contain the information of vehicular driving, such as position, speed, and acceleration, and the information of traffic signal timing. Secondly, the cost function and constraints of vehicular trajectory are formulated. Finally, the optimal vehicular trajectory is dynamically calculated through maximum principle and RL.

3.1. State Vector. In the typical scenario of urban intersection, we define the vehicle entering the road section as C_{mn}, where m and n represent the m-th lane of the road section
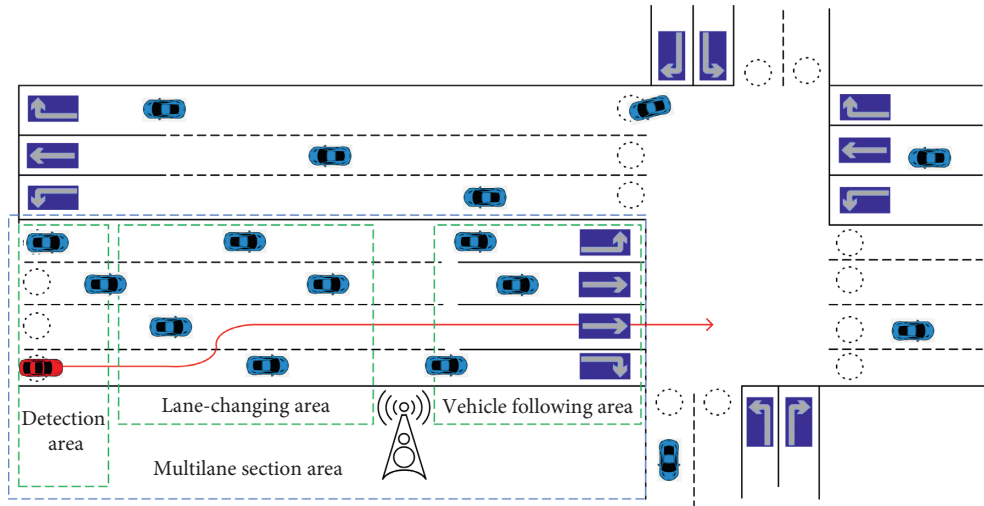
Figure 1: The scenario of multilane section based on connected vehicles.
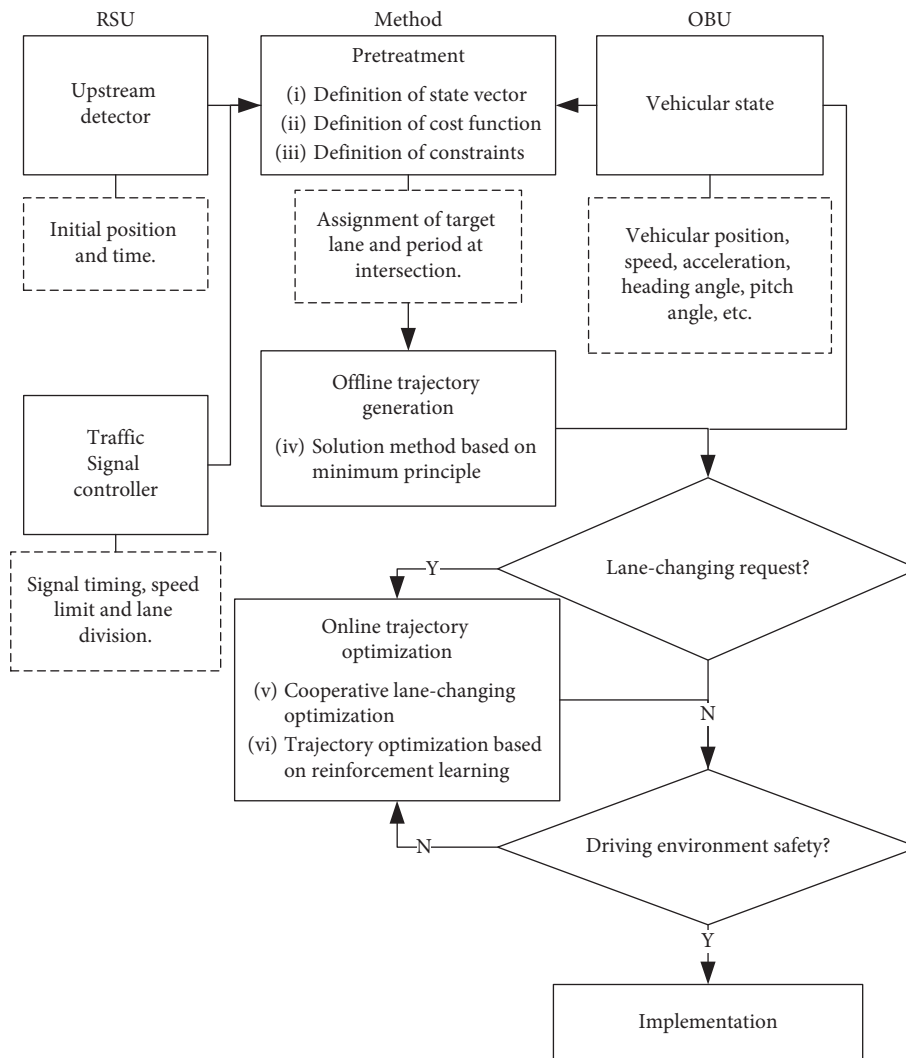
Figure 2: The flow chart of spatiotemporal trajectory optimization process for connected vehicles.
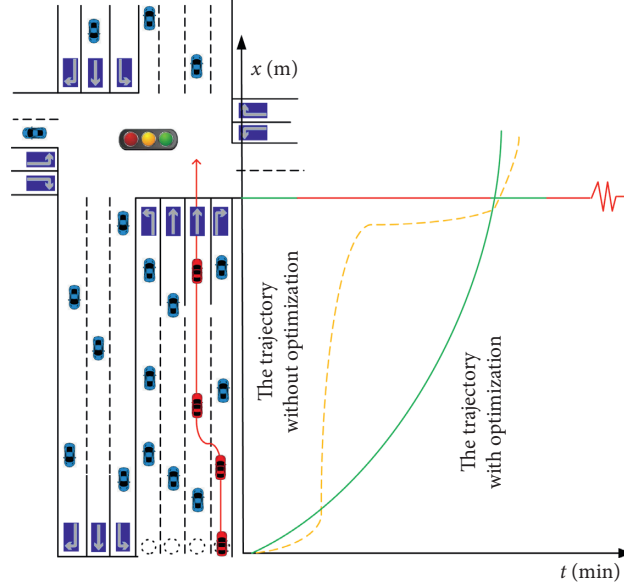
FIGURE 3: The spatiotemporal trajectory with and without optimization.

from the right side to the left side (the target vehicle is selected as the reference) and the n-th vehicle from the downstream to the upstream, respectively. Therefore, we define the state vector of vehicle C_{mn} as $\mathbf{x}_{mn}(t)$:

$$\mathbf{x}_{mn}(t) = \left[x_{mn}(t), y_{mn}(t), v_{mn}(t), v'_{mn}(t)\right]^T, \qquad (1)$$

where $(x_{mn}(t), y_{mn}(t))$ is the position of vehicle C_{mn} at time t, $v_{mn}(t)$ is the driving speed of vehicle C_{mn} at time t, and $v'_{mn}(t)$ is the lateral speed of vehicle C_{mn} at time t.

The method input is the vehicle acceleration $\mathbf{u}_{mn}(t)$, which can be expressed as follows:

$$\mathbf{u}_{mn}(t) = \left[u_{mn}(t), u'_{mn}(t)\right]^T, \qquad (2)$$

where $u_{mn}(t)$ is the longitudinal acceleration of vehicle C_{mn} at time t and $u'_{mn}(t)$ is the current lateral acceleration of vehicle C_{mn} at time t.

Meanwhile, the signal timing information of vehicle target lane $\varphi_{mn}(t)$ and the traffic flow information I_{m-1} and $I_{m-1}(t)$ of adjacent lanes $m+1$ and $m-1$ are defined as

$$\begin{aligned} \boldsymbol{\varphi}_{mn}(t) &= \left[\varphi_{mn}(t), t_{\varphi_{mn}(t)}, R_m, G_m\right], \\ \mathbf{I}_{m+1}(t) &= \left[k_{m+1}(t)\right], \\ \mathbf{I}_{m-1}(t) &= \left[\overline{v}_{m+1}(t), k_{m-1}(t)\right], \end{aligned} \qquad (3)$$

where $\varphi_{mn}(t)$ is the target phase state of vehicle C_{mn}, $t_{\varphi_{mn}(t)}$ is the remaining time of phase $\varphi_{mn}(t)$, R_m is the red light duration of lane m, G_m is the green light duration of lane m, $\overline{v}_{m+1}(t)$ and $\overline{v}_{m-1}(t)$ stand for the average traffic flow velocity of left lane $m+1$ and right lane m-1, and k_{m+1} and k_{m-1} are the average traffic flow density of left lane $m+1$ and right lane $m-1$. The average traffic flow speed $\overline{v}_m(t)$ and average traffic flow density $k_m(t)$ can be expressed as

$$\overline{v}_m(t) = \frac{\sum_{N=N_m^{\min}}^{N_m^{\max}} mN}{N_m^{\max} - N_m^{\min} + 1},$$

$$k_m(t) = \frac{N_m^{\max} - N_m^{\min} + 1}{L}, \qquad (4)$$

where N is the total number of vehicles, $N_m^{\max}$ is the maximum number of vehicles in lane m, $N_m^{\min}$ is the minimum number of vehicles in lane m, and L is the length of road section.

The system state equation is described as

$$\dot{x}_{mn}(t) = f\left(\mathbf{x}_{mn}(t), \mathbf{u}_{mn}(t), \boldsymbol{\varphi}_{mn}(t), \mathbf{I}_{m+1}(t), \mathbf{I}_{m-1}(t)\right). \qquad (5)$$

3.2. Cost Function. In order to ensure that the vehicle can accurately follow the designed trajectory of the system, fixed and variable costs in the process from upstream to downstream are considered and formulated, which can be represented by

$$J = g\left(\mathbf{x}_{mn}\left(t_{mn}^{\text{out}}\right)\right) + \int_{t_{mn}^0}^{t_{mn}^{\text{out}}} h\left(\mathbf{x}_{mn}(t), \mathbf{u}_{mn}(t)\right)dt, \qquad (6)$$

where t_{mn}^{out} is the time when vehicle C_{mn} leaves the road section, t_{mn}^0 is the time when the vehicle C_{mn} enters the road section, $g(\mathbf{x}_{mn}(t_{mn}^{\text{out}}))$ is the fixed cost of the process, and $\int_{t_{mn}^0}^{t_{mn}^{\text{out}}} h(\mathbf{x}_{mn}(t), \mathbf{u}_{mn}(t))dt$ is the variable cost of the process. The fixed cost includes the cost of fixed items (such as the driving distance and expected speed at the intersection) and the travel time of the road section. The process can be described as

$$g\left(x_{mn}\left(t_{mn}^{out}\right)\right) = w_1\left(t_{mn}^{out} - \bar{t}_{mn}^{out}\right)^2 + w_2\left(x_{mn}\left(t_{mn}^{out}\right) - L\right)^2 + w_3\left(v_{mn}\left(t_{mn}^{out}\right) - \bar{v}_{mn}^{out}\right)^2 + w_4\left(y_{mn}\left(t_{mn}^{out}\right) - d\left(\bar{t}_{mn}^{out}\bar{m} - \frac{1}{2}\right)\right)^2, \quad (7)$$

where $\bar{t}_{mn}^{out}$ is the target time of vehicle C_{mn} driving out of the road section, $\bar{v}_{mn}^{out}$ is the expected speed of vehicle C_{mn} at the intersection, $\bar{m}$ is the target lane of vehicle, and d is the lane width. w_1, w_2, w_3, w_4 are the weights of the travel time of the road section, driving length of the road section, expected speed at the intersection, and the target lane at the intersection, respectively. $w_1, w_2, w_3, w_4 \in R+$.

Variable cost involves acceleration and deceleration of the vehicle driving in longitudinal and lateral lanes. Its function can be expressed as

$$h\left(\mathbf{x}_{mn}(t), \mathbf{u}_{mn}(t)\right) = w_5\left(u_{mn}(t)^2 + 2u_{mn}(t)v_{mn}(t)\right)\chi\left(u_{mn}(t)\right) + w_6\left(u_{mn}'(t)^2 + 2u_{mn}'(t)v_{mn}'(t)\right), \quad (8)$$

where w_5 and w_6 are the weights of energy changes caused by longitudinal and lateral acceleration of the vehicle, receptively. $w_5, w_6 \in R+$. $\chi\left(u_{mn}(t)\right)$ is the Heaviside function of longitudinal acceleration of networked vehicles and can be used to separate the variable cost induced by acceleration in longitudinal deceleration process:

$$\chi\left(u_{mn}(t)\right) = \begin{cases} 0, u_{mn}(t) \leq 0, \\ 1, u_{mn}(t) > 0. \end{cases} \quad (9)$$

$\bar{t}_{mn}^{free}$ denotes the time of vehicles passing through the intersection under the condition of low-density traffic flow, also known as the time for vehicles C_{mn} leaving the road without restriction. The time consists of four parts: initial time t_{mn}^0, time t_{mn}^{acc} at acceleration state, time $t_{mn}^{c.s.}$ at constant speed state, and time t_{mn}^{dec} at deceleration state, which can be described as follows:

$$\bar{t}_{mn}^{free} = t_{mn}^0 + t_{mn}^{acc} + t_{mn}^{c.s.} + t_{mn}^{dec}, \quad (10)$$

where t_{mn}^{acc}, $t_{mn}^{c.s.}$, and t_{mn}^{dec} can be expressed as

$$t_{mn}^{acc} = \frac{v_0^{lim} - v_{mn}\left(t_{mn}^0\right)}{u_{mn}^{+max}},$$

$$t_{mn}^{c.s.} = \frac{L - \left(\left(\left(v_0^{lim}\right)^2 - v_{mn}\left(t_{mn}^0\right)^2\right)/2u_{mn}^{+max}\right) - \left(\left(\left(v_0^{lim}\right)^2 - \left(v_m^{lim}\right)^2\right)/2u_{mn}^{-max}\right)}{v_0^{lim}}, \quad (11)$$

$$t_{mn}^{dec} = \frac{v_0^{lim} - v_m^{lim}}{u_{mn}^{-max}},$$

where v_0^{lim} is the speed limit of the road section, v_m^{lim} is the speed limit at the downstream exit of the lane m, u_{mn}^{+max} is the maximum acceleration of vehicle C_{mn}, and u_{mn}^{-max} is the maximum deceleration of vehicle C_{mn}.

However, it is impossible for the urban road section to maintain a low-density state. Therefore, connected vehicles have to pass through current intersection as a temporary platoon. The candidate time $\bar{t}_{mn}^{temp}$ of vehicle C_{mn} can be calculated as

$$\bar{t}_{mn}^{temp} = \max\left(\bar{t}_{m(n-1)}^{out} + t_m^{h2h}, \bar{t}_{mn}^{free}\right), \quad (12)$$

where t_m^{h2h} is the minimum headway between two adjacent vehicles in lane m at the downstream exit of the section.

In order to improve the traffic capacity of intersections, the target time $\bar{t}_{mn}^{out}$ of vehicle C_{mn} at current intersection is controlled within the passable green signal, which can be selected according to

$$\bar{t}_{mn}^{out} = \begin{cases} \bar{t}_{mn}^{temp}, \bar{t}_{mn}^{temp} \in \xi_m, \\ \\ floor\left(\frac{\bar{t}_{mn}^{temp}}{R_m + G_m}\right)(R_m + G_m) + R_m, \bar{t}_{mn}^{temp} \notin \xi_m, \end{cases} \quad (13)$$

where ξ_m is the set of green light periods of lane m, and $floor$ (t) is the function of rounding down.

The expected speed of vehicles is defined as the speed limit at the downstream exit to ensure the maximum efficiency of vehicles and can be described as

$$\bar{v}_{mn}^{out} = v_m^{lim}. \quad (14)$$

In conclusion, the cost function can be integrated as

$$J = w_1\left(t_{mn}^{\text{out}} - \bar{t}_{mn}^{\text{out}}\right)^2 + w_2\left(x_{mn}\left(t_{mn}^{\text{out}}\right) - L\right)^2 + w_3\left(v_{mn}\left(t_{mn}^{\text{out}}\right) - \bar{v}_{mn}^{\text{out}}\right)^2$$

$$+ w_4\left(y_{mn}\left(t_{mn}^{\text{out}}\right) - d\left(\overline{m} - \frac{1}{2}\right)\right)^2 + \int_{t_{mn}^0}^{t_{mn}^{\text{out}}}\left(w_5\left(u_{mn}(t)^2 + 2u_{mn}(t)v_{mn}(t)\right)\chi\left(u_{mn}(t)\right)\right) \tag{15}$$

$$+ w_6\left(u_{mn}'(t)^2 + 2u_{mn}'(t)v_{mn}'(t)\right)\right)\mathrm{d}t.$$

3.3. Constraints.

In order to minimize the above cost function J, seven constraints, namely, initial vehicular state, vehicle spacing, speed, acceleration, jerk, signal timing, and wireless communication, should be satisfied in the optimization problem.

3.3.1. Initial Vehicular State.

Suppose that a connected vehicle C_{mn} enters the road section with detectors; when $n \geq n_{\max}$, the counter will reset. The initial state of vehicle C_{mn} is defined as

$$\mathbf{x}_{mn}\left(t_{mn}^0\right) = \begin{pmatrix} x_{mn}\left(t_{mn}^0\right) \\ y_{mn}\left(t_{mn}^0\right) \\ v_{mn}\left(t_{mn}^0\right) \\ v_{mn}'\left(t_{mn}^0\right) \end{pmatrix} = \begin{pmatrix} 0 \\ d\left(m - \dfrac{1}{2}\right) \\ v_{mn}^0 \\ v_{mn}^{0'} \end{pmatrix}. \tag{16}$$

3.3.2. Vehicle Spacing.

The vehicle will inevitably be in the vehicle-following state in the driving process. Therefore, vehicle C_{mn} on the lane m and the front vehicle $C_{m\,(n-1)}$ should ensure a certain displacement in space and time. The safety constraint can be formulated as

$$x_{mn}(t + \tau_{mn}) \leq x_{m(n-1)}(t) - d_{mn}, \tag{17}$$

where τ_{mn} is the time displacement of vehicle C_{mn} and the front vehicle $C_{m\,(n-1)}$, and d_{mn} is the spatial displacement of vehicle C_{mn} and the front vehicle $C_{m\,(n-1)}$.

3.3.3. Speed.

In order to ensure the safety of the vehicles in the road section, the speed constraint is applied to them. The longitudinal speed constraint embraces the speed information involving speed limit within the section and the minimum speed 0. Thus, we have

$$0 \leq v_{mn}(t) \leq v_0^{\lim}. \tag{18}$$

For the lateral speed constraint, the vehicle deflection angle is mainly constrained by vehicle dynamics, which can be expressed as

$$\alpha_{-\max} \leq \alpha_{mn}(t) \leq \alpha_{\max}, \tag{19}$$

where $\alpha_{\max}$ is the maximum angle through which the front wheel of the vehicle can turn to the left, $\alpha_{mn}(t)$ represents the steering angle of the front wheel of the current vehicle, and

$\alpha_{\max}$ represents the maximum angle through which the front wheel of the vehicle can turn to the right.

Therefore, the lateral restraint condition of the vehicle is shown as

$$v_0^{\lim}\alpha_{-\max} \leq v_{mn}'(t) \leq v_0^{\lim}\alpha_{\max}. \tag{20}$$

3.3.4. Acceleration.

In order to guarantee that the engine can provide enough power and the brake pads can provide the vehicle enough power limit, the acceleration or deceleration of the vehicle should be specified as a constraint. The acceleration constraints can be expressed as

$$\begin{cases} u_{-\max} \leq u_{mn}(t) \leq u_{\max}, \\ u_{-\max}' \leq u_{mn}'(t) \leq u_{\max}', \end{cases} \tag{21}$$

where $u_{\max}$ is the acceleration of the maximum longitudinal braking deceleration of the vehicle, $u_{\max}$ is the acceleration of the maximum longitudinal acceleration of the vehicle, $u'_{\max}$ is the maximum transverse acceleration of the vehicle to the left, and $u'_{\max}$ is the maximum acceleration of the vehicle to the right.

3.3.5. Jerk.

The jerk constraint is the change rate constraint of vehicle acceleration, also known as the impact constraint or comfort constraint. The purpose of jerk constraint is to eliminate the negative impact of the acceleration change in the driving process. The jerk constraint is shown as

$$\begin{cases} j_{-\max} \leq \dfrac{\partial u_{mn}(t)}{\partial t} \leq j_{\max}, \\[2mm] j_{-\max}' \leq \dfrac{\partial u_{mn}'(t)}{\partial t} \leq j_{\max}', \end{cases} \tag{22}$$

where $j_{\max}$ represents the maximum longitudinal backward jerk, $j_{\max}$ represents the maximum longitudinal forward jerk, $j'_{\max}$ represents the maximum lateral left acceleration, and $j'_{\max}$ represents the maximum lateral right acceleration.

3.3.6. Signal Timing.

The signal timing constraint of vehicles can ensure that vehicles avoid violations when passing through the intersection, which can be represented as

$$v_{mn}\left(t_{mn}^{\text{out}}\right) = 0, \quad t_{mn}^{\text{out}} \notin \xi_m. \tag{23}$$

3.3.7. Wireless Communication.

V2X is a communication system through which a vehicle communicates with any entity that may affect the vehicle. Thus, the essence of V2X

technology is wireless communication technology. In the practical application of wireless communication technology, time delay and packet loss are inevitable, which may affect the stability and security of networked vehicles in the system. The relevant parameters of wireless communication are constrained as follows:

$$\tilde{\tau}_{mn}(t) < 0.1,$$
$$P_{lp|mn}(t) < 15\%, \tag{24}$$

where $\tilde{\tau}_{mn}(t)$ and $P_{lp|mn}(t)$ represent the time delay and the packet loss probability of the vehicle C_{mn} at time t.

3.4. Solution Method Based on Maximum Principle.
MSTTOM is solved by Pontryagin's maximum principle, and the Hamilton function of the problem is established as

$$H(\mathbf{x}, \mathbf{u}, \boldsymbol{\lambda}, t) = \boldsymbol{\lambda}^T f(\mathbf{x}, \mathbf{u}, t) + h(\mathbf{x}, \mathbf{u}, t), \tag{25}$$

where $\boldsymbol{\lambda}$ is defined as the costate vector of vector $\boldsymbol{x}$, which represents the additional cost of the change of J caused by the small change ∂x of vector $\boldsymbol{x}$.

In the admissible set U, the minimum input $\mathbf{u}^*$ of the cost must satisfy the minimum state of the Hamilton function, which can be expressed as follows:

$$H(\mathbf{x}*, \mathbf{u}*, \boldsymbol{\lambda}*, t) \le H(\mathbf{x}*, \mathbf{u}, \boldsymbol{\lambda}*, t), \forall \mathbf{u} \in U, t \in \left[t_{mn}^0, t_{mn}^{out}\right],$$

$$\begin{cases} \dfrac{\partial H}{\partial u} = 0, \\\\ \dfrac{\partial H}{\partial x} = -\dot{\lambda}, \\\\ \dfrac{\partial H}{\partial \lambda} = \dot{x}. \end{cases} \tag{26}$$

Therefore, the Hamilton function of MSTTOM is shown as

$$H_{mn} = \lambda_1 v_{mn}(t) + \lambda_2 u_{mn}(t) + w_5$$
$$\left(u_{mn}(t)^2 + 2u_{mn}(t)v_{mn}(t)\right)\chi\left(u_{mn}(t)\right), \tag{27}$$

where the variation of the lateral movement trajectory is optimized in the process of vehicular cooperative lane-changing.

In the process of solving the problem, the state vector λ should meet the condition of the fixed cost, which is shown as

$$\lambda(t_{mn}^{out}) = \dfrac{\partial g(x_{mn}(t_{mn}^{out}))}{\partial x_{mn}} \Rightarrow \begin{cases} \lambda_1(t_{mn}^{out}) = 2w_2(x_{mn}(t_{mn}^{out}) - L) \\\\ \lambda_2(t_{mn}^{out}) = 2w_3(v_{mn}(t_{mn}^{out}) - \overline{v}_{mn}^{out}) \end{cases} \tag{28}$$

3.5. Cooperative Lane-Changing Optimization.
In order to accurately describe the process of cooperative lane-changing strategy, the request of the vehicular lane-changing is displayed in Figure 4. The vehicle C_{mn}, selected as the subject vehicle, is a red one in the middle of lane m and ready to change lanes from the current lane m to the target lane $m + 1$. Besides, yellow vehicles are defined as the primary threat vehicles of C_{mn}, orange vehicles are defined as the secondary threat vehicles of C_{mn}, and blue vehicles are defined as the non-threat vehicles of C_{mn}.

If target lane is not the adjacent lane, it can be decomposed into multiple lane-changing processes. The optimal trajectory is loaded into the traffic environment, and the potential lane-changing conflict safety inspection can be carried out. The flow chart of lane-changing safety inspection is shown in Figure 5.

If two adjacent vehicles receive the same requests concurrently, there is a priority for the vehicle with a higher speed. If adjacent vehicles based on daily driving habits have similar speed, there is a priority for the vehicle in the left lane.

Different traffic saturation should be satisfied in the cooperative lane-changing strategy. The higher the saturation involved in the case is, the higher the complexity of the lane-changing strategy will be supposed. Therefore, the case of higher saturation is discussed as follows:

(i) The subject vehicle C_{mn} initiates a lane-changing request to the target vehicle $C_{(m+1)n}$ after the target gap selected.

(ii) The target vehicle $C_{(m+1)n}$ and its rear vehicle $C_{(m+1)(n+1)}$ adjust their speeds to confirm a safe gap for vehicle C_{mn}.

(iii) Vehicle C_{mn} drives into the target lane $m + 1$.

(iv) The spatiotemporal trajectories of vehicles in the original lane and target lane are updated, which is shown in Figure 6.

3.6. Trajectory Optimization Based on RL.
In order to improve the efficiency of MSTTOM, this paper designs an optimization algorithm based on the RL paradigm, which can quickly match the optimal trajectory. The optimization algorithm is formulated with the current position and speed of vehicles. The target lane and time period are modeled as inputs, and the sets of vehicle acceleration are formulated as outputs. When the lane-changing request is initiated, the vehicle can be connected to the trajectory of the cooperative lane-changing strategy through the RL matching network. Once the lane-changing is completed, the spatiotemporal trajectory will be matched by the RL to achieve the multilane trajectory optimization.

The state vector of MSTTOM satisfies Markov property; that is, the next state s_{t+1} of the system is only related to the current state s_t, but not directly related to the preceding state. Thus, the state vector can be shown as follows:

$$s_t = \left[\mathbf{x}_{mn}(t), \boldsymbol{\varphi}_{mn}(t), \mathbf{I}_{m+1}(t), \mathbf{I}_{m-1}(t)\right],$$
$$P\left[\mathbf{s}_{t+1}|\mathbf{s}_t\right] = P\left[\mathbf{s}_{t+1}|\mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \ldots, \mathbf{s}_t\right], \tag{29}$$

where P is the probability matrix of transition between state.
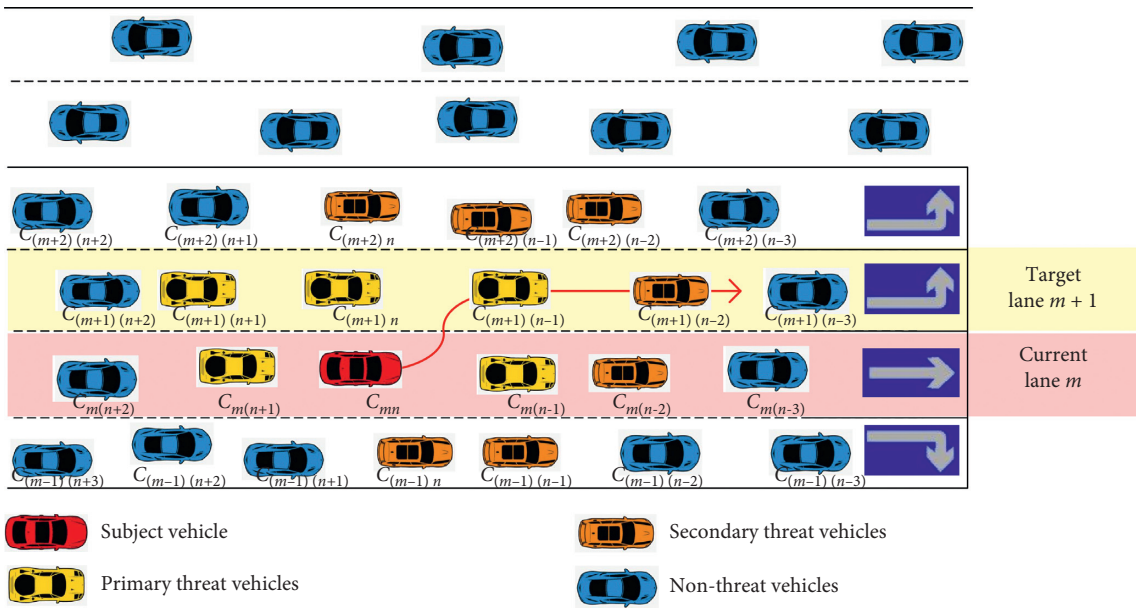
FIGURE 4: Schematic chart of lane-changing environment for vehicles C_{mn}.
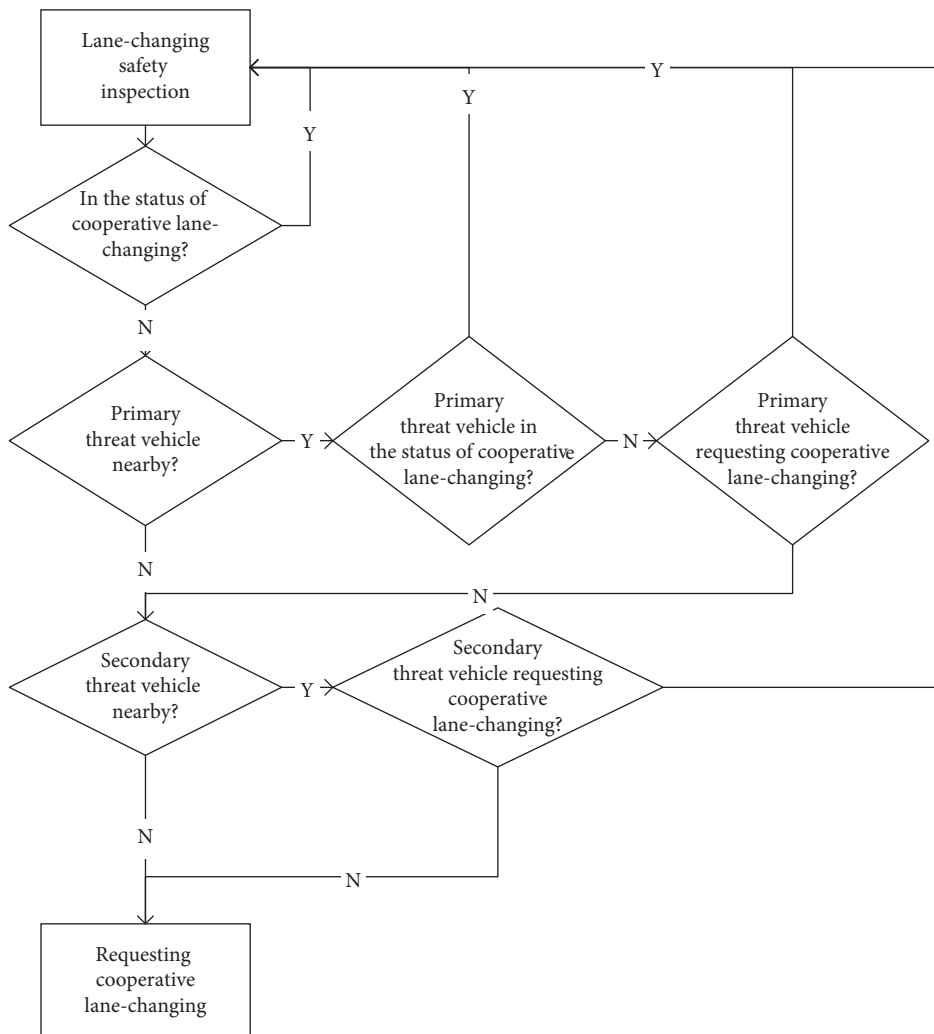
FIGURE 5: The flow chart of lane-changing safety inspection.
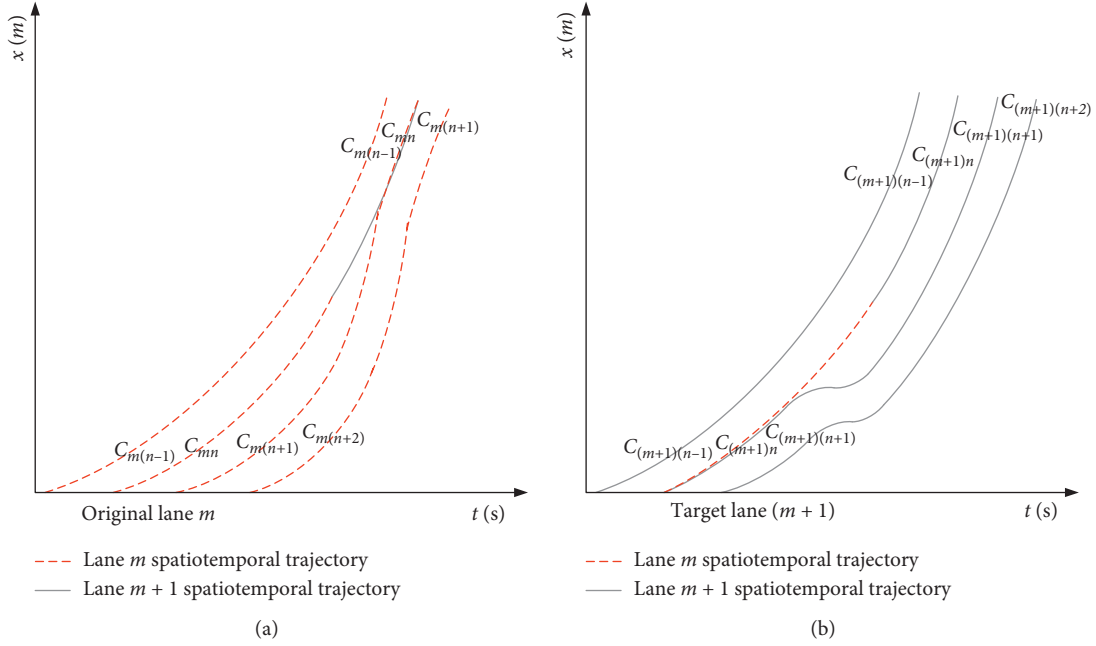
FIGURE 6: Vehicular spatiotemporal trajectories in lane-changing process.

Through integration, the set of action vectors will be presented in the form of trajectory data. The action vector $\boldsymbol{a}_t$ at current state $\boldsymbol{s}_t$ is shown as

$$\mathbf{a}_t = [\mathbf{u}_{mn}(t)]. \tag{30}$$

Five tuples (S, A, P, R, γ) are defined to describe the process of MSTTOM. Among them, S is the state set, which includes the current state of vehicles and the traffic flow state in the road section; A is the set of executing actions, that is, the set of lateral and longitudinal acceleration output by optimization; R is the reward function of the process, which is negatively linear with the cost function J; and γ is the discount factor when calculating the value function. For a fixed policy π, the value function $v_\pi(\mathbf{s})$ can be calculated as

$$v_\pi(\mathbf{s}) = E_\pi\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | S_t = s\right]. \tag{31}$$

An action value function $q_\pi(\mathbf{s}, \mathbf{a})$ can be calculated by defining the value of each action a, which is described as

$$q_\pi(\mathbf{s}, \mathbf{a}) = E_\pi\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | S_t = \mathbf{s}, A_t = \mathbf{a}\right]. \tag{32}$$

Bellman optimal recursive equations of the optimal state value function $v * \mathbf{s}$ and the optimal action value function $q * \mathbf{s}$ can be calculated by introducing Markov property into (31) and (32), which can be shown as follows:

$$v * (s) = \max_a R_s^a + \gamma \sum_{s_t \in S} p_{ss'}^a v * (s'),$$

$$q_\pi(\mathbf{s}, \mathbf{a}) = R_s^a + \gamma \sum_{s' \in S} p_{ss'}^a \max_a q * (s', a'), \tag{33}$$

where s' and $\mathbf{a}'$ are the state and action of the next moment, respectively.

The optimal strategy derived by maximizing the above functions can be represented as

$$\pi * (\mathbf{a}|\mathbf{s}) = \begin{cases} 1, & if\left(\mathbf{a} = \arg\max_{a \in A} q * (\mathbf{s}, \mathbf{a})\right), \\ 0, & else. \end{cases} \tag{34}$$

The function can be calculated by Q-learning algorithm, and the algorithm process is shown in Algorithm 1.

In Algorithm 1, $\mathbf{s}_T$ is the termination state in the fifth row. The current value function of the subsequent state estimation is used to update the optimal action function in the seventh row. In order to enhance the diversity of the exploration ability of the algorithm, the specific formula is shown as

$$\pi(\mathbf{a}|\mathbf{s}) = \begin{cases} 1 - \varepsilon + \dfrac{\varepsilon}{N(a)}, & \mathbf{a} = \arg\max_a q(\mathbf{s}, \mathbf{a}), \\ \dfrac{\varepsilon}{N(a)}, & \mathbf{a} \neq \arg\max_a q(\mathbf{s}, \mathbf{a}), \end{cases} \tag{35}$$

where $N(a)$ is the total number of actions, ε-greedy is the optimal action strategy taken based on 1-ε probability, and random action is the policy of ensuring the possibility of being selected for each action according to the probability of ε.

4. SUMO/Python Simulation and Analysis

In this section, we will simulate and analyze the proposed MSTTOM through the SUMO/Python-based platform to verify the feasibility and effectiveness of the scheme. Meanwhile, the optimal level of mobility at the signal intersection will be guaranteed in the experiment.

```
    Q-learning algorithm
(1) Initialize q (s, a);
(2) While (s_t = s_T)
(3) {Select the initial state s_0 and action a_0 according to the ε-greedy strategy;
(4) While (s_t = s_T)
(5) {Select the action a_t the state s_t according to the ε-greedy strategy, get reward r_t and the next state s_{t+1};
(6) q(s_t, a_t) ← q(s_t, a_t) + a(r_{t+1} + γ max_a q(s_{t+1}, a) − q(s_t, a_t));
(7) s_t = s_{t+1}; } }
(8) Get the optimal strategy π(s) = arg max_a q(s, a)
```

ALGORITHM 1: Q-learning algorithm.

4.1. Experiment Platform Based on SUMO. SUMO is microscopic traffic simulation software [20]. The simulation of SUMO is discrete in time and continuous in space, and the location of each vehicle is described internally. In SUMO, the vehicle model is collision free, so that the variation caused by incomplete model is not allowed to appear in the simulation. SUMO allocates appropriate routes to vehicles considering different traffic demands through Dijkstra algorithm. The structure of SUMO map is shown in Figure 7.

4.2. Experimental Scheme. In the paper, the Netedit of SUMO is used to carry out the multilane experimental scenario based on Figure 1, and the simulation scenario is shown in Figure 8. The intersection is a typical cross intersection, and the traffic light is set as a typical four-phase signal light. Due to the similar spatial characteristics of each section, the 4-lane section from west to east is selected for the experiment.

In the simulation process, the benchmark scheme in the software, the glidepath prototype application (GPPA) scheme [21], and the MSTTOM scheme proposed in this paper are used to simulate and compare the performance of traffic flow conditions with various vehicle saturation 0.6, 0.8, and 1.0. Vehicles simulation diagram is shown in Figure 9.

The three schemes compared in this paper are defined as follows:

(i) The benchmark scheme: in this scheme, the traditional human driving vehicle model in SUMO software is adopted, which means that there is no controlled vehicle and no network communication. Therefore, the driving situation of vehicles under the current traditional driving habits can be simulated in this scenario.

(ii) GPPA scheme: GPPA driving optimization scheme developed by the Federal Highway Administration (FHWA) has been tested and verified. In this scheme, all vehicles in the traffic flow are CVs and are controlled by the optimized system of GPPA.

(iii) MSTTOM scheme: all vehicles in this scheme are CVs, and MSTTOM was applied to all vehicles in this paper. Compared with GPPA scheme, there are the advantages of strong stability and fast calculation efficiency in MSTTOM. Meanwhile, this

scheme is adaptive in the complex multilane and multivehicle scenario.

Simulation conditions and parameter settings for these three schemes are shown in Table 1. Meanwhile, car-following state is simulated by Krauss car-following model [22].

In order to eliminate the influence of unrelated factors in the scenario, the simulation settings in SUMO are simplified by considering the following conditions:

(i) In the simulation environment, there are the same model, size, and kinematic characteristics with all vehicles.

(ii) There is no special weather effect in the simulation environment, and the road adhesion coefficient remains constant.

(iii) The roads are straight in the simulation environment. There are no ramps, nonmotorized vehicular lanes, or parking spaces, which lead to vehicles entering or exiting.

(iv) There is no inclination change in the road section, which means that the vehicle will not travel uphill or downhill.

(v) The arrival probability of vehicles in the simulation environment is stochastic and obeys the Poisson distribution.

The simulation process based on SUMO is shown in Figure 10. It can be seen that the red vehicle from the west to the east attempts to go straight through the intersection within the green duration from the rightmost lane. In this process, the red vehicle completes cooperative lane-changing with vehicles in the left lane through MSTTOM.

4.3. Result Analysis. Through the simulation of SUMO software, the output data of the vehicle can be exported and analyzed. The vehicular spatiotemporal trajectories on the left lane of the test road section are shown in Figure 11, where the spatiotemporal trajectory of the vehicle running on the right lane is denoted by the blue dotted line, and the spatiotemporal trajectory of the vehicle running on the left lane is represented by the red solid line. It is worth noting that the lane-changing vehicles can complete lane-changing quickly and safely
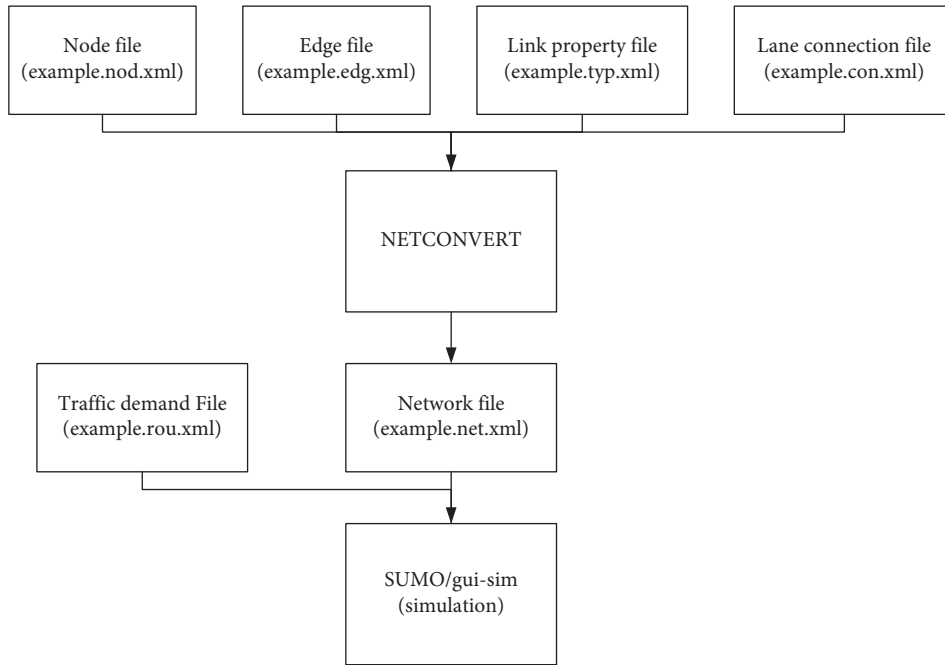
FIGURE 7: The structure of SUMO map.
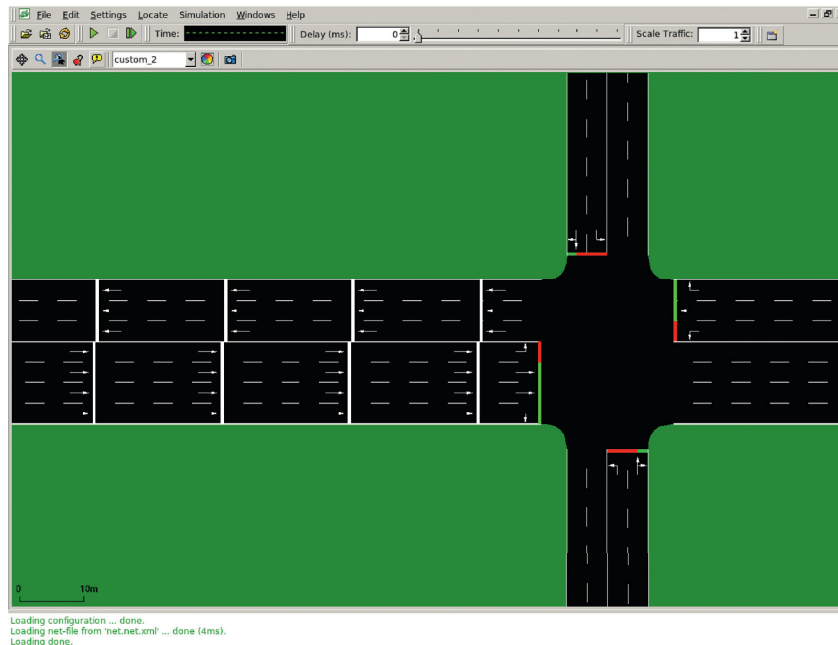
FIGURE 8: Diagrammatic sketch of simulation scenario.

under the optimized method. Once a vehicle requests the change of its lane, vehicles in the target lane can adjust the vehicle position and reserve the optimal safety gap for the vehicle requesting lane-changing. Under the environment of multilane signalized intersection, there is no sharp speed change when vehicles approach the intersection. Therefore, the efficiency of achieving the expected optimization of vehicular lane-changing can be proved for the vehicular cooperative optimization.

The headway and lost time at the stop line of intersection in the road network are analyzed in Figure 12, where the blue curve is the test statistical result of the left lane in the benchmark scheme, the green curve is the test statistical result of the left lane in GPPA scheme, and the red curve is the test statistical result of the left lane in MSTTOM scheme. The meaningless data of the first vehicle in the green and red phases in the experimental results are not included in the above statistics. The jumping part of the curve is the loss time
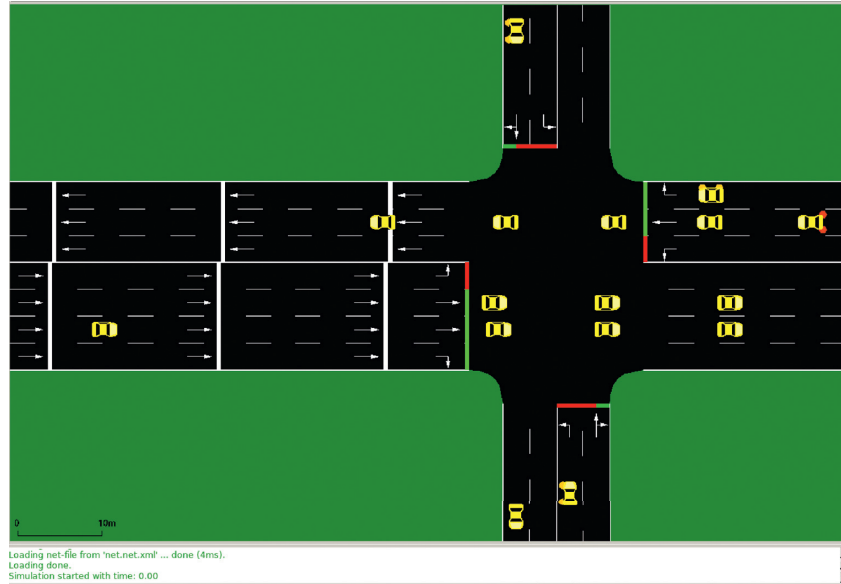
FIGURE 9: Diagrammatic sketch of vehicles simulation.

TABLE 1: Simulation conditions and parameter settings.

Parameters	Unit	Value
Distance from detector to intersection	(m)	200
Minimum headway at rest	(m)	2
Vehicle length	(m)	4.5
Vehicle width	(m)	1.8
Vehicle height	(m)	1.5
Lane width	(m)	3.5
Signal cycle duration	(s)	60
Green light duration	(s)	12
Red light duration	(s)	45
Yellow light duration	(s)	3
Saturated traffic flow rate	(Veh/h/lane)	1800
Maximum vehicle speed	(km/h)	60
Minimum vehicle speed	(km/h)	0
Maximum vehicle acceleration	(m/s2)	2.6
Maximum vehicle deceleration	(m/s2)	−4
Maximum vehicle lateral acceleration	(m/s2)	2
Maximum vehicle lateral deceleration	(m/s2)	−2
Maximum vehicle jerk	(m/s3)	10
Weight coefficient w_1	—	100
Weight coefficient w_2	—	50
Weight coefficient w_3	—	10
Weight coefficient w_4	—	200
Weight coefficient w_5	—	1
Weight coefficient w_6	—	5

of the first vehicle at the beginning of the green light or the lost time of the headway of the following vehicles. Through observation, it can be found that the headway of vehicles can be effectively reduced, the number of vehicles passing through the intersection stop line in a unit time is increased, and the saturation flow rate at the intersection is improved in MSTTOM scheme. In the process of GPPA scheme, some vehicles are affected by vehicular lane-changing, which results in a significant increase in the headway and raise in the

green light loss time. Under MSTTOM scheme, the headway of vehicles at the intersection is significantly reduced, and the saturated flow rate at the intersection is increased. According to the traffic control characteristics of signalized intersections, the increase of effective green time and saturation flow rate will directly lead to an increase in the number of vehicles passing through the stop line at the intersection, which means that the throughput of the intersection will be improved.

(a)

(b)

(c)

(d)
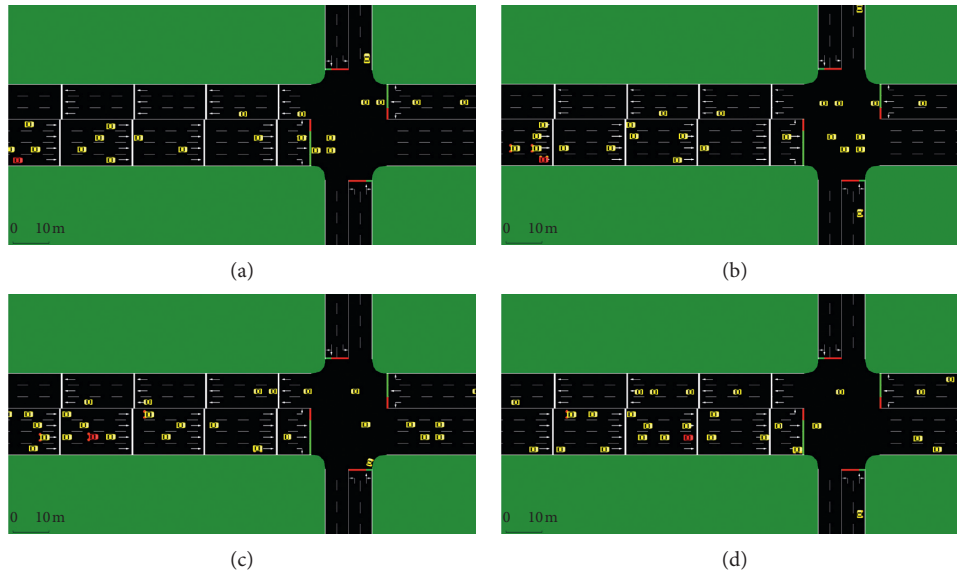
Figure 10: Part of the simulation process based on SUMO.
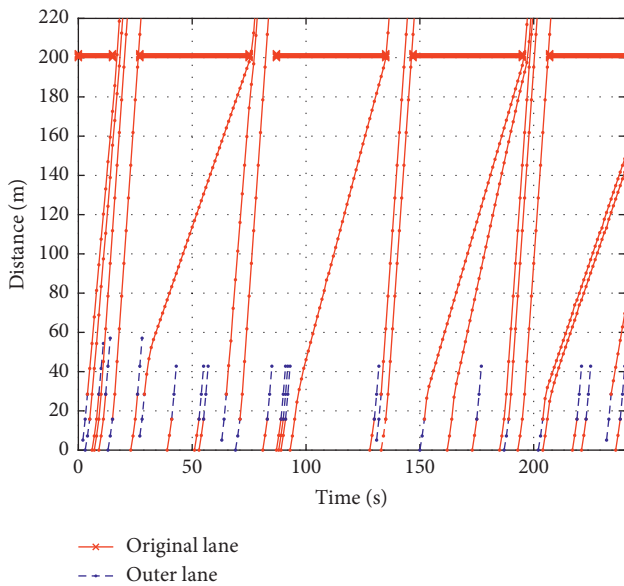
Figure 11: Spatiotemporal trajectories of left lane.

Original lane
Outer lane

Figure 12: Headway variation under different schemes.

Benchmark scheme
GPPA scheme
MSTTOM scheme

The vehicle fuel efficiency and corresponding profit of the whole system under MSTTOM are analyzed in Figure 13. The results of T-test show that there is a significant statistical difference in the vehicle fuel efficiency data between the MSTTOM scheme and the benchmark scheme. Through observation, it can be seen that the use of MSTTOM can effectively improve the vehicle fuel efficiency and reduce the vehicle fuel consumption.

After calculation, the vehicle fuel consumption profit is shown in Figure 14. In the case of different vehicle saturation, there is a significant fuel efficiency benefit in GPPA scheme and MSTTOM scheme. Comparing GPPA scheme

with the scheme of optimization method proposed in this paper, the advantages of MSTTOM scheme in improving traffic mobility, enhancing vehicle fuel efficiency, and reducing pollutants emissions are obvious.

In MSTTOM scheme, the system calculation time analysis under different iteration time steps and optimization time spans is demonstrated in Figure 15. The maximum calculation time is up to 0.93 when the optimization time span is 50 seconds, and the iteration time step is 0.2 seconds. In the simulation process, the optimization time spans of all CVs are less than 50 seconds, and each iteration time step is greater than 0.2 seconds. Therefore, the calculation time of
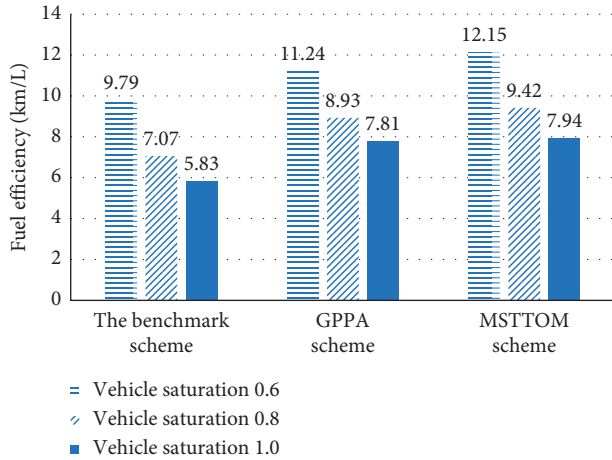
FIGURE 13: Fuel efficiency under different schemes.
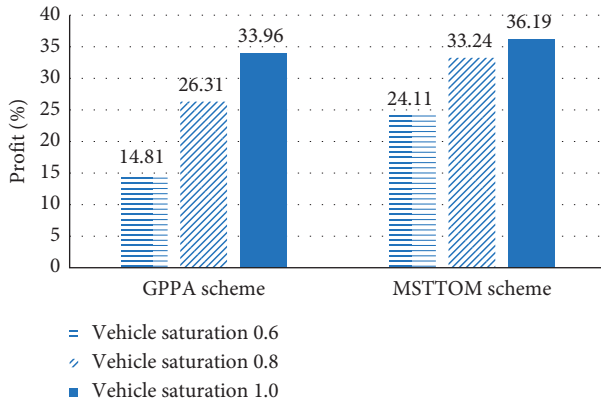
FIGURE 14: Vehicle fuel consumption profit under different schemes.
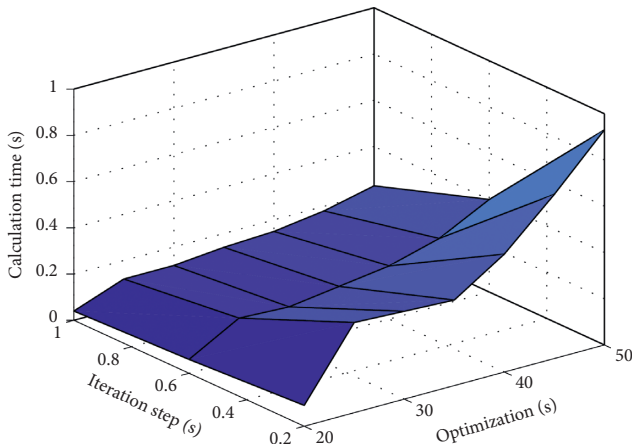
FIGURE 15: System calculation time under different optimization parameters.

the system is less than 1 second; that is, the effective performance of the system time meets the requirements of real-time optimization.

The objectives of ensuring the optimal mobility of vehicles at signalized intersections, enhancing vehicle fuel efficiency, and reducing pollutants emissions can be achieved in MSTTOM scheme. Compared with the existing GPPA algorithm, the strengths of MSTTOM are as follows:

(i) The function of multilane cooperative lane-changing strategy under the condition of random traffic flow is well realized in MSTTOM scheme.

(ii) Compared with the benchmark scheme, the mobility of vehicles within the road section is improved, the vehicle fuel efficiency is enhanced by 32%, and the emissions of pollutants are reduced by 17% in MSTTOM scheme.

(iii) Compared with the advanced GPPA scheme, there is a 24% reduction in the vehicle fuel consumption in MSTTOM scheme.

(iv) The traffic fluctuation caused by the intersection signal control is smooth and the traffic flow in the road section is more reasonable in MSTTOM scheme.

(v) For the randomness in the traffic process, the MSTTOM scheme shows its superiority in robustness than other existing schemes.

(vi) The calculation time in MSTTOM scheme is less than 1 second, and the distribution is mainly concentrated in 0.3 seconds to 0.7 seconds, which satisfies the requirements of real-time optimization.

5. Conclusions

This paper proposes a new scheme, MSTTOM, to ease traffic pressure and reduce traffic congestion in multilane road sections. The problem of low traffic efficiency at the multilane intersections is studied and analyzed in this paper. Four stages of work, namely, literature summary, scenario and structure introduction, method optimization design and experimental validation, and optimization objects determination, are conducted. The research results and innovations are as follows:

(i) A vehicular lane-changing method based on V2X is proposed. Firstly, the lane-changing environment information is analyzed to classify and define the potential threat vehicles in the driving environment. Secondly, the lane-changing behaviors of potential threat vehicles are analyzed, and conflict detection is carried out to determine the priority order in the lane-changing process. Then, lane-changing gap and target cooperative vehicles are determined and declared. Finally, the spatiotemporal trajectories of related vehicles are updated. The vehicular cooperative lane-changing strategy lays a foundation for the reasonable planning of MSTTOM.

(ii) This paper designs a multilane spatiotemporal trajectory optimization method. Based on the optimal control theory, the state control models of CVs driven by Pontryagin's maximum principle are constructed. The cost function is formulated by continuous functional target completion and

optimal profit. The constraints in line with the actual situation of the vehicle are also established. Finally, the trajectories data are optimized by Q-learning algorithm to achieve the real-time optimization.

(iii) SUMO software is used to verify the feasibility and effectiveness of MSTTOM by comparing it with GPPA scheme of FHWA in the United States. The experimental results demonstrate that MSTTOM can effectively optimize the traffic flow, improve the vehicle fuel efficiency, reduce pollutants emissions, and improve the traffic efficiency of vehicles at intersections.

Data Availability

All data generated or analyzed during this study are included in this article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] P. Reina and S. Ahn, "Lane flow distribution of congested traffic on three-lane freeways," *International Journal of Transportation Science and Technology*, vol. 9, no. 1, pp. 1–13, 2020.

[2] E. Suryani, R. A. Hendrawan, P. F. EAdipraja, A. Wibisono, and L. P. Dewi, "Modelling reliability of transportation systems to reduce traffic congestion," in *Proccedings of the Journal of Physics: Conference Series*, vol. 1196, no. 1, Article ID 12029, Palembang, Indonesia, November 2019.

[3] M. Tajalli and A. Hajbabaie, "Distributed optimization and coordination algorithms for dynamic speed optimization of connected and autonomous vehicles in urban street networks," *Transportation Research Part C: Emerging Technologies*, vol. 95, pp. 497–515, 2018.

[4] E. Grumert and A. Tapani, "Impacts of a cooperative variable speed limit system," *Procedia-Social and Behavioral Sciences*, vol. 43, pp. 595–606, 2012.

[5] Y. Jo, H. Choi, S. Jeon, and I. Jung, "Variable speed limit to improve safety near traffic congestion on urban freeways," *International Journal of Fuzzy Systems*, vol. 14, no. 2, 2012.

[6] Y. Fang and Z. Y. Guo, "Research on the method of dynamic speed limit on expressway under complex climate based on pavement skid-resistant performance," *Advanced Materials Research*, vol. 2478, 2013.

[7] F. Zhu and S. V. Ukkusuri, "Accounting for dynamic speed limit control in a stochastic traffic environment: a reinforcement learning approach," *Transportation Research Part C Emerging Technologies*, vol. 41, pp. 30–47, 2014.

[8] M. Aldana-Muñoz, E. Maeso-González, and A. García-Rodríguez, "Contributions of eco-driving on traffic safety," *Securitas Vialis*, vol. 7, pp. 1–3, 2016.

[9] J. F. Wang, C. C. Li, J. R. Lv, and X. D. Yan, "Speed guidance model during the green phase based on a connected vehicle," *Simulation*, vol. 92, no. 10, pp. 899–905, 2016.

[10] Y. Wei, C. Avcı, J. Liu et al., "Dynamic programming-based multi-vehicle longitudinal trajectory optimization with simplified car following models," *Transportation Research Part B Methodological*, vol. 106, pp. 102–129, 2017.

[11] R. De Mello and R. D. Chiodi, "A safe speed guidance model for highways," *International Journal of Injury Control and Safety Promotion*, vol. 54, no. 4, pp. 408–415, 2018.

[12] X. Liang, S. I. Guler, and V. V. Gayah, "Joint optimization of signal phasing and timing and vehicle speed guidance in a connected and autonomous vehicle environment," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2674, no. 4, pp. 70–83, 2019.

[13] P. Wang, Y. Jiang, L. Xiao, Y. Zhao, and Y. Li, "A joint control model for connected vehicle platoon and arterial signal coordination," *Journal of Intelligent Transportation Systems*, vol. 24, no. 1, pp. 81–92, 2020.

[14] P. Wang, H. Deng, J. Zhang, and M. Zhang, "Real-time urban regional route planning model for connected vehicles based on V2X communication," *Journal of Transport and Land Use*, vol. 13, no. 1, pp. 517–538, 2020.

[15] A. Mirheli, M. Tajalli, L. Hajibabai, and A. Hajbabaie, "A consensus-based distributed trajectory control in a signal-free intersection," *Transportation Research Part C: Emerging Technologies*, vol. 100, pp. 161–176, 2019.

[16] P. Wang, P. Li, F. R. Chowdhury, L. Zhang, and X. Zhou, "A mixed integer programming formulation and scalable solution algorithms for traffic control coordination across multiple intersections based on vehicle trajectories," *Transportation Research Part B: Methodological*.vol. 134, pp. 266–304, 2020.

[17] M. A. Benloucif, J. Popieul, and C. Sentouh, "F for multi-level cooperation and dynamic authority management in an Automated Driving System-a case study on lane change cooperation," *IFAC PapersOnLine*, vol. 49, no. 19, pp. 615–620, 2016.

[18] A. Hongil and J. Jae-Il, "Design of a cooperative lane change protocol for a connected and automated vehicle based on an estimation of the communication delay," *Sensors*, vol. 18, no. 10, p. 3499, 2018.

[19] L. Li, J. Gan, K. Zhou, X. Qu, and B. Ran, "A novel lane-changing model of connected and automated vehicles: using the safety potential field theory," *Physica A: Statistical Mechanics and Its Applications*, vol. 559, Article ID 125039, 2020.

[20] J. Zambrano-Martinez, C. Calafate, D. Soler, J. C. Cano, and P. Manzoni, "Modeling and characterization of traffic flows in urban environments," *Sensors*, vol. 18, no. 7, 2018.

[21] O. D. Altan, G. Wu, M. J. Barth, K. Boriboonsomsin, and J. A. Stark, "GlidePath: eco-friendly automated approach and departure at signalized intersections," *IEEE Trans. Intelligent Vehicles*, vol. 2, no. 4, pp. 266–277, 2017.

[22] V. Kanagaraj, G. Asaithambi, C. N. Kumar, K. K. Srinivasan, and R. Sivanandan, "Evaluation of different vehicle following models under mixed traffic conditions," *Procedia-Social and Behavioral Sciences*, vol. 104, pp. 390–401, 2013.

Research Article

Restriction Analysis of Transport Policy for Bridges Using the Trajectory Data

Zhenghua Hu,[1,2] Jibiao Zhou ⬤,[3,4] Shuichao Zhang,[3] Songhan He,[1] and Bo'an Yu[1]

[1]*School of Electronic and Information Engineering, Ningbo University of Technology, Fenghua Rd. No. 201, Ningbo 315211, Zhejiang Province, China*
[2]*College of Information Science & Electronic Engineering, Zhejiang University, Zheda Rd. No. 38, Hangzhou 310027, Zhejiang Province, China*
[3]*School of Civil and Transportation Engineering, Ningbo University of Technology, Fenghua Rd. No. 201, Ningbo 315211, Zhejiang Province, China*
[4]*Department of Transportation Engineering, Tongji University, Caoan Rd. No. 4800, Shanghai 201804, China*

Correspondence should be addressed to Jibiao Zhou; zhoujibiao@tongji.edu.cn

Roads are becoming increasingly congested with continuous rise in the number of vehicles. Restriction policies are selected to alleviate congestion in many cities. However, conclusions regarding the substantial effects of restriction policies have not been fully demonstrated. This study primarily aims to demonstrate whether traffic restrictions can control the driving habits of people to alleviate traffic pressure. Furthermore, this study investigates the effect on the traffic on the premise of a normalized restriction policy. Data were collected by bayonet systems in Ningbo. Results showed that vehicles restricted by the restriction policy only accounted for approximately 13%. Most drivers bypass restricted roads to avoid restrictions. The method proposed can effectively amend the trajectory deviation caused by the inaccuracy from the bayonet. Based on the results, some suggestions about the policy of restrictions were discussed.

1. Introduction

With the continuous development of urbanization and the acceleration of motorization, traffic congestion has become a bottleneck in urban development. The congested roads in urban areas not only affect the traveling environment and physical health of residents but also seriously hinder the sustainable development of social economy and daily life [1–5]. According to the internationally accepted sayings, holding more than 1 million cars is a sign that a city has entered the "auto society." After the cities of Beijing, Shanghai, Chengdu, Guangzhou, Shenzhen, Tianjin, Hangzhou, and Suzhou, Ningbo has become the ninth city in China with more than one million cars. At the end of 2018, the total number of cars and drivers in Ningbo had reached 2.54 and 3.23 million, respectively. A total of 249,000 new cars and 213,000 new drivers were recorded in 2018 alone. The ownership of cars in the city of Ningbo

continued to grow rapidly in 2019. Ningbo took only eight years to become a city with more than three million cars in 2019 from that with one million in 2011. The rapid growth in the total number of cars has caused congestion during rush hours in the core areas of Ningbo. The problem of "driving difficulty on the road" has become a livelihood issue that affects urban development and the lives of residents.

However, with the gradual improvement in road networks in the central area of Ningbo, the room for growth in infrastructure continues to decrease. Considering the local and international experience of development, the growth rate of transportation facilities has not sustained the demand from transportation.

In order to fundamentally solve the imbalances between supply and demand, it is necessary to treat the management of traffic demand as the fundamental means to solve transportation problems, fully tap the supply

capacity of transportation facilities, and adjust the spatial–temporal distribution of traffic flows [6–8]. This finding has become a vital issue for the development of the transportation system and is also an essential direction for the strategies of transportation development in big cities in China.

Many researchers have comprehensively analyzed the causes of congestion and proposed corresponding strategies to solve the problem [9–12]. A popular method is to curb traffic demand by controlling the usage intensity of vehicles (purchase restriction policies [13], parking policies [14], schemes to encourage carpooling [15], etc.), thereby reducing the amount of traffic. Restrictions by license plate number are the most economical and effective way to reduce traffic pressure on the roads [16–20]. Traffic demand regulation through policies and economical means has been recently adopted in major cities. The city of Ningbo has also implemented traffic restrictions based on even- and odd-numbered license plates on bridges (Jiangxia and Ling Bridges) across the river in the core area of the city. However, some controversies over such strategies still exist. Some people believe that restrictions based on even- and odd-numbered license plates are unnecessary because these restrictions objectively break the original structure of urban traffic patterns, which causes traffic flow oscillation between the two bridges and other adjacent passages [21]. And, the effect of traffic demand management is slightly observed. However, others view that the restriction policy must be continuously implemented and further expanded to form restriction zones so as to achieve demand management in the core areas. Therefore, it is necessary to study this restriction policy, evaluate its effect and deficiencies in real-time, and propose suitable policies regarding traffic demand management in Ningbo.

2. Literature Review

2.1. Different Strategies on Traffic Restriction. The restriction policy has been implemented in some first-tier cities in China, such as Beijing, Shanghai, Guangzhou, and Hangzhou. Based on the rules, two main types of strategies are available: according to the tail number of the license plate [22, 23] and according to the exhaust emissions [24, 25]. Among these strategies, traffic restriction according to plate number can be further divided into two subtypes, namely, restriction rules by even- and odd-numbered license plates and restriction of several (N) numbers on each day of a week. Theoretically, the former can reduce traffic flow by 50% daily. In comparison, the latter can reduce the traffic flow by $10\% \times N$. Based on the restricted time, the traffic restriction policies can be further divided into two subtypes: traffic restrictions during the special or nonspecial period. The former refer to restrictions during particular periods, such as the Olympic Games, Asian Games, or holidays, while the latter refer to restrictions during rush hours in the morning and evening on working days.

2.2. Evaluation of Restriction Policies. Researchers have obtained different evaluation results from different perspectives in studies. At present, the relatively acceptable evaluation conclusions are as follows:

(a) Policies of restrictions have changed the travel patterns of some residents; thus, these residents turn to public transportation, such as buses and subways [26]. However, these policies do not remarkably affect the travel of full-time personnel [27].

(b) Restricted rules have short-term effectiveness, but long-term restrictions will offset by various evasion behaviors [28, 29].

(c) Restriction policies can improve air quality to some extent [30]. Still, the effect is not evident [31, 32].

The advantages of implementing restriction policies are as follows.

(a) The cost of implementing restrictions is low

(b) In the early stage, the policies can reduce the traffic flow and alleviate traffic congestion [25, 33]

(c) The policies control the exhaust emissions and improve the air quality to a certain extent [34, 35]

The disadvantages of these policies are as follows [36, 37]:

(a) The traffic pressure can only be relieved in the short term and cannot fundamentally solve the problem of traffic congestion in urban areas

(b) These policies stimulate some people to purchase a second car to avoid traffic restrictions

The response from the public to restriction rules is various. Supportive and wait-and-see attitudes, as well as questioning voices, are observed [18, 20, 38]. Most citizens support the policy of traffic restriction. They believe that restriction rules can effectively stimulate citizens to take public transportation, such as buses and subways, thereby reducing exhaust emissions. Although the restriction policies have inconvenienced some car owners, they still support these restriction rules. However, some citizens believe that the restriction rules limit the legitimate rights and interests of citizens; and whether or not such restrictions provide corresponding compensation, the policy is considered a violation of private rights [39]. Many residents believe that with the gradual increase in motor vehicles, even with the adoption of restrictions rules, the willingness for people to drive out will not be curbed.

Overall, although restriction rules have limited the motorized vehicles out to a certain extent, long-term restrictions for travelers with complex needs may adopt evasion measures. Besides, the inconvenience caused by restriction may prompt the purchase of a second car, thereby inducing a new increase in vehicle ownership.

No researcher has implemented a detailed analysis from a quantitative perspective due to the different opinions on the restriction policy. Therefore, taking the restriction policy

in Ningbo as an example, this study took the license plate data from bayonet systems to extract the trajectory of vehicles and then conducted a detailed evaluation of the restriction policy. The method used in this study refers to trajectory acquisition technology based on the license plate. Although the recognition of license plate has successful achievements at home and abroad, the conventional trajectory extraction technology cannot be directly applied in Ningbo, due to the particularity of the bayonets system (aging and inconsistency of the equipment or the inaccuracy of basic data, etc.). Therefore, introducing a series of data preprocessing and analysis methods is necessary.

3. Data and Method

3.1. Data Collection. Through real-time video monitoring in the Big Data Lab of Ningbo University of Technology, the license plate data are captured when cars pass through each bayonet. The data format is shown in Table 1 and Figure 1.

Furthermore, the information about traffic flow toward each direction can be calculated. Simultaneously, the erroneous data are eliminated, and the missing data are interpolated and fitted to obtain the complete trajectory of each vehicle. Thus, the flow, the degree of congestion (saturation), and the spatial distribution characteristics of traffic flow could be calculated.

3.2. Extraction Method of Traveling Trajectory. The original data must be preprocessed initially to remove errors because the recognition data of license plates captured by the bayonet system contain numerous gross errors due to objective reasons. Then, the unreasonable records are eliminated using positional and topological relationships between road network and bayonet stations. In the sample data, the set of bayonets passing by each vehicle for a single trip is queried after the grosses are removed, and the Dijkstra algorithm for the restricted area is called to calculate the path between two adjacent points. However, an absolute deviation is also found in the position of the bayonet; that is, the position of the bayonet and the road network data do not entirely match. Only the coverage tree of the travel path can be obtained. Finally, the actual trajectory is computed by finding the longest path in that covering tree.

3.3. Basic Data Analysis. The trajectory data from November 3, 2019, to November 30, 2019 (four entire weeks), were used to ensure the reliability and stability of statistical results. These data include more than 1.65 billion records with a volume of up to 100 G. Therefore, the statistical results are convincing. Finally, saturation and service level analyses for the bridges (Jiangxia and Ling Bridges) and the surrounding intersections were conducted using the aforementioned method. The daily average traffic volume of the road in the city of Ningbo reached 4.54 million vehicles, and the average flow on cross section was 162 vehicles/h. The changes in the traffic flow and saturation on the main roads are shown in Table 2.

4. Case Study

4.1. Basic Profiles. Ningbo is a subprovincial and a separate city under the jurisdiction of Zhejiang Province. It is an important port city approved by the State Council and is an economic center on the south side of the Yangtze River Delta. Ningbo is located in the middle of China's coastline, with the southern wing of the Yangtze River Delta and the Zhoushan Islands being a natural barrier in the east. The city of Ningbo is one of the eight major water systems in Zhejiang Province. The rivers include Yuyao, Fenghua, and Yongjiang Rivers. The Yuyao and Fenghua Rivers merge into the Yongjiang River at Sanjiangkou in the urban area. The Yongjiang River flows northeast and enters the East China Sea through Zhaobao Mountain. More than 10 bridges span over the central part of the Yuyao, Fenghua, and Yongjiang Rivers. Among these bridges, Ling and Jiangxia Bridges are the core bridges in Ningbo and are both the central artery of the city. The ends of the two bridges connect the Old Town of Haishu and Jiangdong Districts, as shown in Figure 2.

According to statistics, the number of private cars in Ningbo has reached 2.76 million as of May 2019, and the average daily number of trips by car in the central city reached 200,000. The restriction policy has been implemented in Ningbo for over 15 years (2004–2019). The effect of the policy indicates that despite traffic congestion alleviation or traveling structure adjustment, the actual role of the restriction rules is not as evident as expected. Considerable uncertainty is found in its actual effect.

Jiangxia and Ling Bridges are two east-west bridges located near Sanjiangkou in the center of Ningbo. They are both bridges that span over Fenghua River and connect Haishu with Yinzhou District. Both bridges are also an essential gateway to the new eastern town and the business district in the south. They are adjacent to the old district of Jiangdong and connect bustling Tianyi Square to the west. Jiangxia and Ling Bridges are two two-way and two-lane bridges. As essential traffic arteries in the center of Ningbo, Jiangxia and Ling Bridges bear the main traffic functions and traffic flow between Haishu and Yinzhou Districts in the downtown area of Ningbo.

Ling and Jiangxia Bridges can only accommodate limited traffic due to historical conditions. However, due to their essential location, these two bridges have the highest traffic flow daily among all main roads in Ningbo, which can accommodate nearly 30,000 cars per day. Therefore, the two bridges are always in a congested state before the restriction policy was carried out, and the rush hours lasted for more than four hours a day.

4.2. Analysis of Trips under Traffic Restrictions. The management department should take economic or administrative measures, such as restriction rule, not only to control the travel demand and guide people to switch to public transportation (such as buses and subways), but also to reduce the actual traffic flow of road networks. This will undoubtedly optimize the relationship between traffic demand and supply, ease the traffic pressure on the road, maintain

TABLE 1: Data description of license plate captured by bayonet.

Name	Definition	Description
DeviceID	Number of each device	An 18-digit number that uniquely identifies a bayonet
SnapShottime	Shooting time	Time of each shot
LaneNo	Sequence of the lane	Sequence of the lane
CarType	Vehicle type	1 for car, 2 for truck, 3 for bus, and 4 for minibus
PlateNum	Plate number	Number of the recognized license plate

● Survey site

FIGURE 1: Road network and bayonet distribution on the road.

TABLE 2: Statistics of traffic flow on restricted bridges and intersections nearby.

Statistical section	Statistical period													
	7:00–8:00		8:00–9:00		10:00–11:00		13:00–14:00		17:00–18:00		18:00–19:00		20:00–21:00	
	F	S	F	S	F	S	F	S	F	S	F	S	F	S
Jiangxia Bridge	1013	0.96	820	0.78	795	0.76	829	0.79	1078	1.03	989	0.94	1308	1.25
Ling Bridge	739	0.70	623	0.59	609	0.58	628	0.60	778	0.74	772	0.74	903	0.86
Intersections around Jiangxia Bridge	1103	1.05	1108	1.06	1119	1.07	1197	1.14	1327	1.26	1235	1.18	1483	1.41
Intersections around Ling Bridge	749	0.71	704	0.67	694	0.66	700	0.67	838	0.80	746	0.71	772	0.74

F: flow (vehicles/hour) and S: saturation.

relatively smooth state of traffic flow, and improve the traveling environment. Vehicles running at a relatively smooth speed on the road can effectively reduce vehicle exhaust emissions, thereby further improving the air quality of the city.

4.2.1. Analysis of the Effect of Restriction Rules. This study uses the data of license plate from bayonet system in November 2019 to extract trajectory from vehicles and analyze whether the restriction policy has achieved the intended purpose of restricting traffic in Ningbo.

According to the restriction rules in Ningbo, the trajectory of vehicles that pass through the restricted bridges on the first day was calculated the next day. Then, these trajectories are used to verify whether the restriction policy effectively restrained people's habit of driving out and shifted the traveling mode to public transportation.

The statistical data show that the vehicle restriction policy in Ningbo is not as effective as expected. The number

of vehicles that are restricted only accounts for approximately 13%. That is, only 13% of vehicle owners give up driving out when facing the policy and select to travel by bus, subway, or nonmotorized vehicles. The restriction policy has changed and adjusted people's traveling mode to some extent. The policy has also alleviated the city's traffic pressure to a certain extent. Therefore, the restriction policy has played a role in the city of Ningbo, but the effect is not evident. Most car owners do not change their traveling modes but respond to restriction rules by bypassing corresponding roads.

4.2.2. Analysis of the Effect of Limiting Traffic. Jiangxia and Ling Bridges have become the most critical bridges to the new town in the east due to their unique locations. Long before the implementation of the traffic restriction, the two bridges were well known to the public. The average traffic flow of the two bridges on working days was computed to

FIGURE 2: Overview of the research area.

verify the effect of restriction policies on the bridge, as shown in Figure 3.

The chart shows that, during the restricted time, the traffic volume on Jiangxia and Ling Bridges did not show a significant peak and remained within a relatively stable range. The effect of traffic control on the corresponding road is evident. As the most important arterial road in the city center, the restriction policy has reduced the overall traffic volume on Jiangxia and Ling Bridges by 20%–30%. The restriction policy also reduced the traffic volume of the relevant roads and increased the average running speed. Thus, the roads were in a state of "basically unobstructed." However, according to the regulations of restrictions based on even- and odd-numbered license plates, nearly half of the traffic flow should be restricted from passing through the restricted roads in theory, but the ideal state was never reached. This finding may be due to the fact that the reduced traffic volume on the restricted roads has attracted traffic elsewhere or that some families have purchased a second car to deal with the restriction policy.

Outside the restricted period, the traffic volume of the restricted roads may be slightly higher than that during the restricted period. Significantly, the traffic volume of the restricted roads shows a maximum value within half an hour before and after the restricted period. This finding indicates that many travelers have changed the original travel time and passed the restricted road during nonrestricted periods (before or after the restricted time). Many drivers even pass through the restricted roads in the restricted time.

The traffic restriction has played a role in limiting the traffic volume for the two bridges, and the effect on traffic is evident. This finding also shows that the restriction policy diverts traffic from rush hours to flat hours. The policy improves the passability of the road and alleviates the current situation of traffic congestion in the morning and evening peaks.

4.2.3. Analysis of the Shunt Effect. Furthermore, the detour trajectories (trajectory of vehicles with odd number on even days and trajectory of vehicles with even number on odd days) during workdays are computed within four complete weeks, as shown in Figure 4.

The statistical data show that, due to the restriction policy, vehicles that originally traveled from Jiangxia and Ling Bridges were detoured from other bridges nearby. Moreover, the detour routes were mainly concentrated in Jiefang and Bund Bridges that locate at the north side of Jiangxia Bridge, Qin and Xingning Bridges that locate at the south side of the Ling Bridge, and the Zhilan Bridge that locates at the southwest of the city.

To the north of Jiangxia Bridge, nine detours comprising six bridges, including the Bund, Xinjiang, Yongjiang, and Jiefang Bridges, are available for vehicles to bypass. Among these bridges, the Jiefang and Bund Bridges have become the most welcomed detour routes for drivers, accounting for 12% of all detoured traffic because they are the most convenient alternative passages. The Xinjiang and Yongjiang Bridges are the closest to the Jiangxia Bridge; thus, many vehicles (approximately 8%) choose to pass through the two bridges. The reason is mainly that some drivers forget the regulations of that day when traveling. Moreover, many
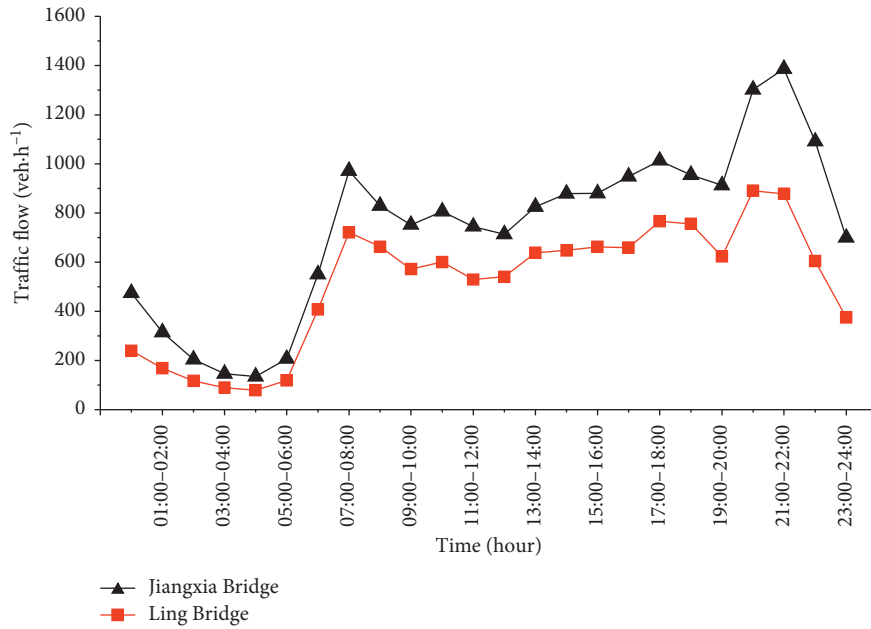
FIGURE 3: Daily average flow of the two restricted bridges.
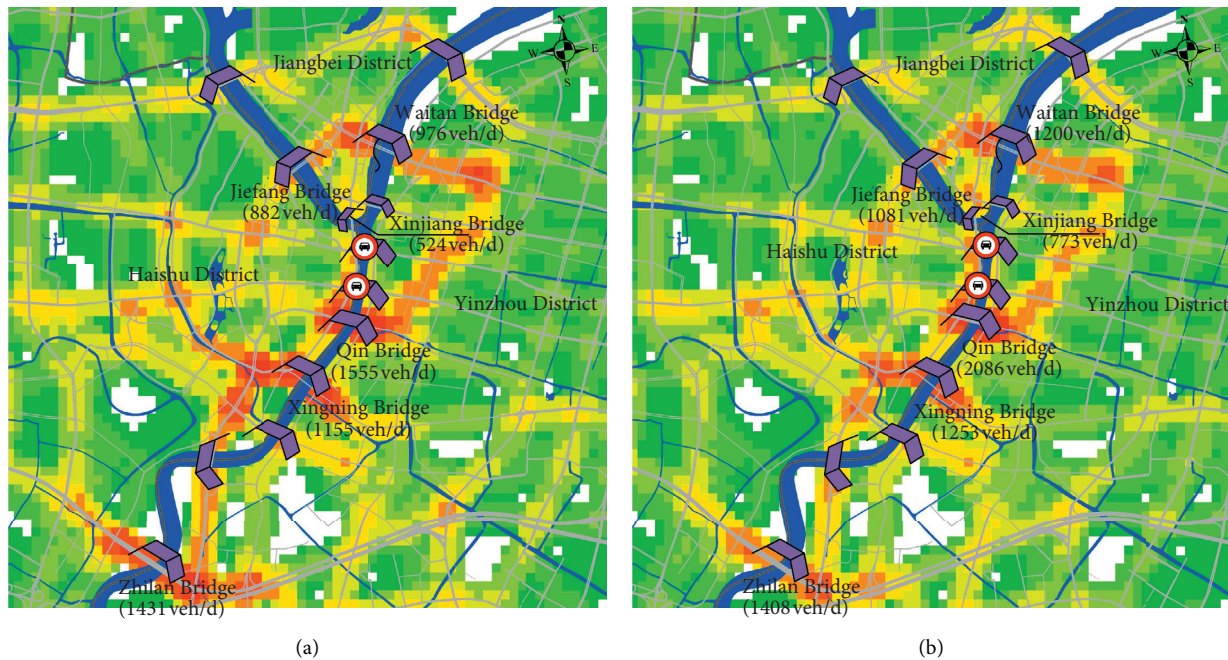
FIGURE 4: Heat maps of the detours. (a) Trajectory of vehicles with odd number on even days. (b) Trajectory of vehicles with even number on odd days.

foreign drivers do not know the local traffic restrictions in Ningbo; thus, these drivers change their paths temporarily after seeing a warning sign in front of the bridge. Other detour routes have either certain inconveniences or poor road conditions, and relatively few detour vehicles are available.

The detour route to the south mainly includes the Qin and Xingning Bridges. Although the Qin Bridge is geographically closest to Ling Bridge, this bridge has a long

history and has limited traffic capacity. Therefore, the traffic conditions of the bridge are relatively bad. Most drivers detour from Xingning Bridge, which is not far from the Qin Bridge. Therefore, the Qin and Xingning Bridges have become the most crucial bypass bridges in the south, bearing 15% and 18% of the detoured traffic, respectively. The statistical results also show that the South Ring Road, which is the farthest from the restricted bridge, is a significant detour. The reason is that this road combines viaduct in the

mainline and auxiliary on the ground, so the overall traffic condition is good. Moreover, the west side of the road is connected to the airport viaduct; therefore, the road has a considerable traffic capacity. Hence, a large proportion (16%) of drivers is willing to sacrifice the running length to ensure the punctuality of the travel.

4.2.4. Analysis of the Impact of Restriction Rules on Traffic Flow. Although the restriction policy does not effectively curb private cars running on the roads, this policy has controlled the traffic flow on restricted roads to a certain extent, exhibiting a diversion effect on the traffic flow. This finding also has a specific impact on the roads around the restricted sections. This study used the license plate data captured by the bayonet to extract the travel trajectory and calculate the changes in indicators on the detours around the two restricted bridges (Jiangxia and Ling Bridges). The impact of the restriction policy on unrestricted roads is then studied. The results are shown in Table 3.

Table 3 shows that the restriction policy affected the flow of bypass-roads around the restricted bridges. The policy has caused many travelers to detour from other bridges around restricted ones. Thus, the traffic flow on the bypass-roads was significantly improved, especially for Qingfeng and Xinjiang Bridges. However, these roads have not experienced explosive growth, and the traffic flow on the entire urban road network can be maintained in a smooth state. This finding is also consistent with the conclusions from the previous analysis. The restricted vehicles have experienced additional miles due to detours. The total mileage has increased for each trip, but the increment is within an acceptable range. The average length of the detour indicates that the distance due to the restriction accounts for 10% of the total mileage. The average driving speed denotes that as the traffic volume on the relevant road increases, the number of vehicles passing through increases and the average speed of the corresponding road decreases due to the increased traffic flow. These findings are consistent with the actual situation.

Based on the length of a single trip and the required travel time, the delayed time is further estimated, as shown in Figure 5. The figure indicates that although the traffic flow on the road and the total length of the path have increased to varying degrees, no explosive growth in the final travel time was observed. The time with and without detour is almost equal. For many detours, the floating range of the delay is within the acceptable level. Because of the restriction policy, the travel period has been extended. Hence, the distribution of travel flow per hour has become more balanced. Significantly, the proportion of trips in the morning and evening rush hours has reduced. Simultaneously, the proportion of trips before morning rush and after evening rush has increased. This indicates that the restriction policy effectively realizes the regulation of traveling time and optimizes the supply and demand relationship of traffic during rush hours. Therefore, increasing the travel distance in exchange for an improved driving environment is relatively unacceptable. As a whole,

if a comfortable driving environment is ensured, then most travelers are willing to travel from detours.

4.2.5. Scenario Simulation. We analyzed the different effects of restriction policy based on even- and odd-numbered license plates upon cross-river passages in Ningbo. Based on the existing travel trajectory, we explore the following three traffic scenarios with different restriction policies:

> Scenario I: the status of all cross-river passages remains unchanged; that is, traffic restriction policy is continued to be carried out on Jiangxia and Ling Bridges
>
> Scenario II: all cross-river passages adopt traffic restriction policy based on even- and odd-numbered license plates
>
> Scenario III: all cross-river passages do not adopt traffic restriction policy based on even- and odd-numbered license plates

(1) Scenario I. The traveling trajectories are counted using extraction method proposed in this paper. The traffic flow of the two restricted bridges is about 1137 veh/h, the saturation is 1.03, the number of vehicles that run across the river is 14219, and the number of detouring vehicles is 8866.

(2) Scenario II. Combined with the optimization of the traffic structure and development trend, a low-carbon traveling structure is proposed, and all cross-river passages are restricted by even- and odd-tail numbers. The traffic flow of the two restricted bridges is 1006 veh/h, the saturation is 0.91, the number of vehicles that run across the river is 12578, and the number of detouring vehicles is 6805.

(3) Scenario III. Since all cross-river passages do not adopt restriction policy, this suppresses the tendency of citizens to buy a second car to a certain extent. In the beginning, the traffic volume of the cross-river passages that cancel the restriction rules will increase, but as time goes on, the traffic will gradually and automatically tend to a balanced state. The traffic flow of the two restricted bridges is 1847 veh/h, the saturation is 1.67, the number of vehicles that run across the river is 23088, and the number of detouring vehicles is 5353.

Table 4 summarizes the information (such as traffic volume) on Jiangxia and Ling Bridges under three different traffic control methods.

On the whole, due to the increasing travel demand from residents, the growth trend of car ownership is obvious. In Scenario I, if the existing restriction policy remained unchanged, then most drivers would choose to detour to avoid restrictions that are imposed by the policy. In Scenario II, if all cross-river passages adopt the restriction policy, the proportion of public transport trips will undoubtedly increase, and saturation of roads will decrease. But there would be a tendency in saturation to increase year by year. Due to wealth of household and other factors, the number of private cars in a family may be more than one, which would cause more serious congestion. Within a few years, the traffic volume of the two bridges will return to the

TABLE 3: Indicators of traffic flow on detours before and after adopting the restriction rule.

Bridges	Average flow on detours (v/h)		The average length of the detours (km)		Average speed (km/h)	
	Before	After	Before	After	Before	After
QFB–YFB	2978	3091	12.77	13.13	25.6	25.2
QFB–JFB	475	487	12.06	14.21	17.9	22.2
QFB–XJB	644	759	11.65	11.94	23.1	23.9
WTB–YFB	1996	2115	12.35	12.58	20.9	20.7
WTB–JFB	2386	2656	9.19	10.18	18.4	19.3
WTB–XJB	392	408	14.19	14.14	11.1	15.7
YJB–YFB	663	665	9.84	9.86	19.5	19.5
YJB–JFB	1075	1084	8.05	7.86	19.2	19.2
YJB–XJB	869	909	8.52	8.21	17.5	18.9
Qin Bridge	10066	11290	7.98	8.05	17.3	17.5
XNB	7140	7630	10.39	10.49	19.8	19
CFB	11395	11637	11.26	11.9	20.3	19.3
CLB	10706	10986	9.46	11.17	18	17.2
Zhilan Bridge	11115	11556	12.12	13.7	26.8	25.2

QFB is short for Qingfeng Bridge, YFB is short for Yongfeng Bridge, JFB is short for Jiefang Bridge, XJB is short for Xinjiang Bridge, WTB is short for Waitan Bridge, YJB is short for Yongjiang Bridge, XNB is short for Xingning Bridge, CFB is short for Changfeng Bridge, and CLB is short for Chenglang Bridge.
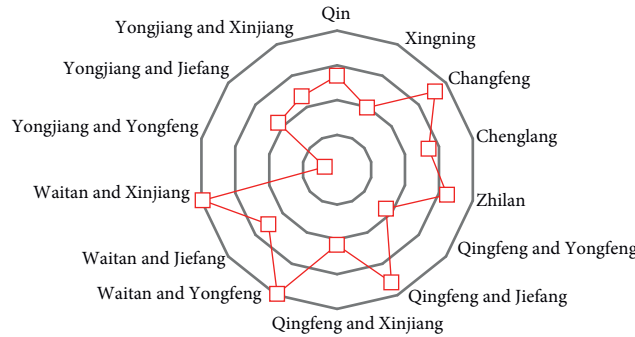
FIGURE 5: Delays from driving around the detoured bridges.

TABLE 4: Traffic information on Jiangxia and Ling Bridges under three different scenarios.

Scenarios	Traffic flow	Saturation	Number of vehicles run across the river	Number of detouring vehicles
Scenario I	1137	1.03	14219	8866
Scenario II	1006	0.91	12578	6805
Scenario III	1847	1.67	23088	5353

level before the restriction policy was carried out. In Scenario III, the restriction policy was not carried out on all passages across the river. Although the traffic volume on the previously restricted bridges has increased, due to the self-adaptability, the traffic volume of either bridge will gradually stabilize without sharp fluctuations. In contrast, Scenario III is more effective in alleviating traffic congestion in Ningbo.

4.3. Evaluation of Restricted Policies.
From the previous analysis, the conclusions are drawn as follows:

(a) Traffic restriction policy can effectively control the traffic flow on restricted roads, thereby improving the speed and substantially reducing the traffic pressure on restricted roads. Meanwhile, the restriction policy can restrain residents from using private cars to a certain extent and guide them to shift their travel ways to public transportation, such as buses or subways. This can alleviate exhaust pollution on the road and improve air quality in cities. The government has achieved the effect of guiding users' options of travel modes by implementing restrictions rules.

(b) In the face of restrictions, only a small number of citizens travel by bus or subway during the restricted period, and most citizens still use private cars to travel. Among these citizens, some travel out of the restricted period to avoid restriction rules. An increasing number of citizens will opt to detour roads around the restricted roads, thereby increasing the traffic pressure on the surrounding detours. The restriction policy plays a role in diverting traffic on restricted roads.

(c) In the long run, high-income people will purchase a second car to cope with the restriction policy. The increment in the number of vehicles will not fundamentally improve the situation of congestion on the road. Therefore, the long-term implementation of the restriction policy cannot control the overall traffic volume, and the effect of the restriction policy will be significantly reduced.

Overall, the policy of traffic restriction has limited people from using cars to a certain extent. In particular, this policy significantly affects traffic on restricted roads, ensures driving efficiency, and improves the traffic order. Restriction policy has reasonably promoted people's dependence on public transportation. However, these measures have not achieved the expected effect. Most car owners either travel during the nonrestriction period or bypass the restricted road. The occupancy rate of vehicles on the detours has significantly increased during the restriction period. In other words, traffic restriction has only played a role in diverting traffic on restricted roads.

5. Conclusion

We believe that the traffic flow on Ling Bridge and Jiangxia Bridge has experienced a historical decline, indicating that the restriction policy controls the traffic flow on relevant roads to a certain extent. Therefore, the traffic restriction policy has played a role in regulating traffic demand. However, most drivers circumvent restrictions by bypassing from either side of restricted bridges. If the traffic restriction policy was canceled, the traffic flow of Jiangxia Bridge and Ling Bridge, as well as the nearby roads, would inevitably rebound. However, due to the limited capacity itself, the traffic flow would be self-adaptive and return to a stable state. In particular, we believe that the traffic problem in Ningbo is the problem of commuter traffic. The traffic pressure during rush hours is relatively small (in core areas, trips by car in the four rush hours in the morning and evening account for 69.7% of the entire day and do not exceed 6% during flat hours). If the restriction policy was canceled, then no unprecedented pressure would be found from traffic.

From the perspective of the traffic control functions, the effect of restriction policy includes traffic flow regulation on certain roads. This effect is limited for the adjustment of regional traffic demand and may still be affected by the overall capacity of roads. The separate effect of restriction policy is small. So the restriction rules should be complemented by other transportation strategies (such as increasing parking fees and reducing public transportation fees) to improve their effects.

Further research on the reoptimization of road spaces to realize the continuity of motor vehicles, public transportation, nonmotorized vehicles, and sidewalks will be conducted. The focus is on improving the quality of public transportation, nonmotorized vehicles, walking, and maximizing the role of slow-moving traffic. Relevant departments should gradually increase the cost of cars and promote the transformation from short-distance travel by car to nonmotorized vehicles or walking. Meanwhile, additional crowdsourced data, such as trajectory and mobile phone signaling data, must be collected to further verify the effectiveness of the proposed method, especially in cities (such as Wuhan and Nanjing) with similar terrain in Ningbo where the entire city is divided by rivers or mountains into different regions.

Data Availability

The codes that support the findings of this study are available at https://figshare.com/articles/software/Restriction_Analysis_of_Transport_Policy_for_Bridges_Using_the_Trajectory_Data/11783649.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] F. Chen, M. Song, and X. Ma, "Investigation on the injury severity of drivers in rear-end collisions between cars using a random parameters bivariate ordered probit model," *International Journal of Environmental Research and Public Health*, vol. 16, no. 14, Article ID 2632, 2019.

[2] F. Chen and S. Chen, "Injury severities of truck drivers in single- and multi-vehicle accidents on rural highways," *Accident Analysis & Prevention*, vol. 43, no. 5, pp. 1677–1688, 2011.

[3] F. Chen, J. Wang, and Y. Deng, "Road safety risk evaluation by means of improved entropy TOPSIS-RSR," *Safety Science*, vol. 79, pp. 39–54, 2015.

[4] C. Ma, W. Hao, W. Xiang, and W. Yan, "The impact of aggressive driving behavior on driver-injury severity at highway-rail grade crossings accidents," *Journal of Advanced Transportation*, vol. 2018, Article ID 9841498, 10 pages, 2018.

[5] Y. Guo, T. Sayed, and M. Essa, "Real-time conflict-based Bayesian tobit models for safety evaluation of signalized intersections," *Accident Analysis & Prevention*, vol. 144, Article ID 105660, 2020.

[6] F. Chen, S. Chen, and X. Ma, "Analysis of hourly crash likelihood using unbalanced panel data mixed logit model and real-time driving environmental big data," *Journal of Safety Research*, vol. 65, pp. 153–159, 2018.

[7] Y. Guo, M. Essa, T. Sayed, M. M. Haque, and S. Washington, "A comparison between simulated and field-measured conflicts for safety assessment of signalized intersections in

Australia," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 96–110, 2019.

[8] Y. Guo, Z. Li, P. Liu, and Y. Wu, "Modeling correlation and heterogeneity in crash rates by collision types using full Bayesian random parameters multivariate tobit model," *Accident Analysis & Prevention*, vol. 128, pp. 164–174, 2019.

[9] X. Li, W. Wang, C. Xu, Z. Li, and B. Wang, "Multi-objective optimization of urban bus network using cumulative prospect theory," *Journal of Systems Science and Complexity*, vol. 28, no. 3, pp. 661–678, 2015.

[10] X. Li, J. Tang, X. Hu, and W. Wang, "Assessing intercity multimodal choice behavior in a touristy city: a factor analysis," *Journal of Transport Geography*, vol. 86, Article ID 102776, 2020.

[11] Y. Guo, T. Sayed, L. Zheng, and M. Essa, "An extreme value theory based approach for calibration of microsimulation models for safety analysis," *Simulation Modelling Practice and Theory*, vol. 106, Article ID 102172, 2021.

[12] Y. Guo, T. Sayed, and L. Zheng, "A hierarchical Bayesian peak over threshold approach for conflict-based before-after safety evaluation of leading pedestrian intervals," *Accident Analysis & Prevention*, vol. 147, Article ID 105772, 2020.

[13] F. Liu, F. Zhao, Z. Liu, and H. Hao, "The impact of purchase restriction policy on car ownership in China's four major cities," *Journal of Advanced Transportation*, vol. 2020, no. 5, 14 pages, Article ID 7454307, 2020.

[14] A. Ibeas, L. dell'Olio, and J. L. Moura, "Parking behavior and policy," *Journal of Advanced Transportation*, vol. 20182 pages, Article ID 1075946, 2018.

[15] D. Ding and B. Shuai, "A traffic restriction scheme for enhancing carpooling," *Discrete Dynamics in Nature and Society*, vol. 2017, Article ID 9626938, 9 pages, 2017.

[16] J. A. Acero, A. Simon, A. Padro, and O. S. Coloma, "Impact of local urban design and traffic restrictions on air quality in a medium-sized town," *Environmental Technology*, vol. 33, no. 21, pp. 2467–2477, 2012.

[17] X. Liu, "Analysis of the policy of "peak shifting the limit line" in Hangzhou," *Advances in Social Sciences*, vol. 8, no. 4, pp. 556–559, 2019.

[18] S. Lucrezi, M. Saayman, and P. Van der Merwe, "Impact of off-road vehicles (ORVs) on ghost crabs of sandy beaches with traffic restrictions: a case study of Sodwana Bay, South Africa," *Environmental Management*, vol. 53, no. 3, pp. 520–533, 2014.

[19] K. Ostrowski and M. Tracz, "Availability and reliability of a signalised lane," *Transportmetrica B: Transport Dynamics*, vol. 7, no. 1, pp. 1044–1061, 2019.

[20] A. Szarata, K. Nosal, U. Duda-Wiertel, and L. Franek, "The impact of the car restrictions implemented in the city centre on the public space quality," *Transportation Research Procedia*, vol. 27, pp. 752–759, 2017.

[21] X. Cheng, K. Huang, L. Qu, T. Zhang, and L. Li, "Effects of vehicle restriction policies on urban travel demand change from a built environment perspective," *Journal of Advanced Transportation*, vol. 2020, no. 3, 13 pages, Article ID 9848095, 2020.

[22] X. Dong, *The Impact of Travel Habits for Private Car on Travel Choice under the One-Day-a-Week Limiting Policy in Beijing*, Beijing Jiaotong University, Beijing, China, 2019.

[23] Z. Sun, *The Influence of Beijing Vehicle Restrictions on the Travel Mode Choice of Residents-an Empirical Study*, Beijing Jiaotong University, Beijing, China, 2018.

[24] M. Liu, *The Influence of Traffic Restriction Measures on Air Quality in Lanzhou City Based on CMAQ*, Lanzhou University, Lanzhou, China, 2018.

[25] Z. Liu, R. Li, X. Wang, and P. Shang, "Effects of vehicle restriction policies: analysis using license plate recognition data in Langfang, China," *Transportation Research Part A: Policy and Practice*, vol. 118, pp. 89–103, 2018.

[26] Y. Gu, E. Deakin, and Y. Long, "The effects of driving restrictions on travel behavior evidence from Beijing," *Journal of Urban Economics*, vol. 102, pp. 106–122, 2017.

[27] L. Wang, J. Xu, and P. Qin, "Will a driving restriction policy reduce car trips?-the case study of Beijing, China," *Transportation Research Part A: Policy and Practice*, vol. 67, pp. 279–290, 2014.

[28] V. Cantillo and J. De Dios Ortúzar, "Restricting the use of cars by license plate numbers: a misguided urban transport policy," *Dyna*, vol. 81, no. 188, pp. 75–82, 2014.

[29] L. de Grange and R. Troncoso, "Impacts of vehicle restrictions on urban transport flows: the case of Santiago, Chile," *Transport Policy*, vol. 18, no. 6, pp. 862–869, 2011.

[30] H. M. Worden, Y. Cheng, G. Pfister et al., "Satellite-based estimates of reduced CO and CO_2 emissions due to traffic restrictions during the 2008 Beijing olympics," *Geophysical Research Letters*, vol. 39, no. 14, 2012.

[31] S. Chowdhury, S. Dey, S. N. Tripathi, G. Beig, A. K. Mishra, and S. Sharma, ""Traffic intervention" policy fails to mitigate air pollution in megacity Delhi," *Environmental Science & Policy*, vol. 74, pp. 8–13, 2017.

[32] J. Ye, "Better safe than sorry? evidence from Lanzhou's driving restriction policy," *China Economic Review*, vol. 45, pp. 1–21, 2017.

[33] C. Sun, S. Zheng, and R. Wang, "Restricting driving for better traffic and clearer skies: did it work in Beijing?" *Transport Policy*, vol. 32, pp. 34–41, 2014.

[34] D. Han, H. Yang, and X. Wang, "Efficiency of the plate-number-based traffic rationing in general networks," *Transportation Research Part E: Logistics and Transportation Review*, vol. 46, no. 6, pp. 1095–1110, 2010.

[35] S. Wang, C. Shao, F. Wang, and J. Sun, "Evaluation on vehicle restriction measure in Beijing," in *Proceedings of the Seventh International Conference on Traffic and Transportation Studies 2010*, pp. 433–443, Kunming, China, August 2010.

[36] F. Gallego, J.-P. Montero, and C. Salas, "The effect of transport policies on car use: evidence from Latin American cities," *Journal of Public Economics*, vol. 107, pp. 47–62, 2013.

[37] H. Ma and G. He, "Effects of the post-olympics driving restrictions on air quality in Beijing," *Sustainability*, vol. 8, no. 9, p. 902, 2016.

[38] R. San Jose, J. L. Pérez, L. Pérez, and R. M. González, "A health impact assessment of traffic restrictions during Madrid NO_2 episode," in *Proceedings of the IOP Conference Series: 2018 9th International Conference on Environmental Science and Technology*, IOP Publishing, Prague, Czech Republic, June 2018.

[39] D.-x. Wu, R. Li, and J.-f. Wu, "The economics analysis of the vehicle restriction rule," *Technology & Economy in Areas of Communications*, vol. 2, 2012.

WILEY | Hindawi

Research Article

A New Model for Locating Plate Recognition Devices to Minimize the Impact of the Uncertain Knowledge of the Routes on Traffic Estimation Results

Santos Sánchez-Cambronero ⓘ **, Fernando Álvarez-Bazo** ⓘ **, Ana Rivas** ⓘ **, and Inmaculada Gallego** ⓘ

Department of Civil and Building Engineering, University of Castilla-La Mancha, 13071 Ciudad Real, Spain

Correspondence should be addressed to Santos Sánchez-Cambronero; santos.sanchez@uclm.es

A number of research papers have recently shown that the use of techniques based on the installation of vehicle identification devices allows us to address the observability problem of a traffic network in a much more efficient way than if it were done with traditional techniques. The use of such devices can lead to a better data set in terms of flows and therefore to a better definition of traffic flows, which is essential for traffic management in cities and regions. However, the current methodologies aimed at network modeling and data processing which are not fully adapted to the use of these devices in obtaining the necessary data for analyzing traffic and making network forecasts. This is because the essential variable in models which used data from plate scanning (as a particular case of AVI sensors) is composed of the route flows, while traditional methods are based on the observation of link and/or origin-destination flows. In this context, this paper proposes several practical contributions, in particular: (1) a traffic network design method aimed to use the plate scanning data to estimate traffic flows and (2) an algorithm for locating plate reader devices to reduce the effect of the uncertain knowledge of route enumeration. Next, using the well-known Nguyen-Dupuis network, a sensitivity analysis has been carried out to evaluate the influence of different parameters of the model on the final solution. These parameters are the considered routes, the degree of network simplification, and the available budget to install devices. Finally, the method has been applied to a real network.

1. Introduction

As is well known, estimating the origin-destination trip matrix, route flows, and link flows is essential to achieving efficient traffic management. Many authors have dealt with this problem, trying to estimate these traffic flows using either information from traditional sources such as traffic counts (see, among others, Castillo et al. [1, 2] and Perrakis et al. [3]) or information from more innovative sources such as mobile phones and GPS data (Huang et al. [4], Ibarra-Espinosa et al. [5], and Moreira-Matias et al. [6]), Big Data (Toole et al. [7] and Zin et al. [8]), or automatic vehicle identification (AVI) data (Castillo et al. [9], Fu et al. [10], and Fedorov et al. [11], among others).

Yang et al. pointed out [12], to deal with traffic flow estimation (not only for short-term predictions but for more

generalized studies too), classical statistical methods are widely applied, but also, machine learning methods are shown very useful due to their many advantages as, for example, problem adaptability, generalization, and also the learning ability, which is very important to estimate traffic flows using field data. For example, Sánchez-Cambronero et al. [13] used Bayesian networks; Bai and Chen [14] used neural networks; and Lui et al. [15] used deep learning. In any case, without going into detail concerning the models used to predict traffic flows, to conduct such traffic analysis both for static and dynamic studies, technicians need two things: a good representation of the traffic network and a good data set with which to simulate routes of the network and to predict the flows (Nigro et al. [16]). This means that the optimal number and locations of the sensors that can collect such data must be determined. The next sections of

this introduction deal with these problems in order to clarify concepts and describe problems that this paper faces.

1.1. The Network Representation.

A traffic network is a pair $(\mathcal{N}, \mathcal{A})$, where $\mathcal{N}$ is a set of nodes and $\mathcal{A}$ is a set of directed links connecting these nodes. The links represent the streets of a city, and the nodes typically represent the intersections of these streets. To build a transportation network model, two aspects of the network must be considered (see Sheffi [17]):

(i) The quantitative information associated with each link: this is the uncongested travel time, the travel cost, the parking places, the number of street lanes, the number of residents, etc. Each network link is associated with some impedance function derived from this information that, for a given flow, provides the "generalized cost" of using this link (see, for example, Spies [18], Huntsinger and Rouphail [19], or Mtoi and Moses [20]).

(ii) The graph representation: the role of the graph representation is to translate the physical structure of a city into a model of nodes and links. Some simplifications are needed:

  (a) The division into traffic zones: the beginning of a commute is, for example, a person's house, and the end is his or her workplace. To model this situation for all city inhabitants, thousands of origins and destinations would be needed. Thus, the transportation planning process is typically based on a partition of the city into traffic zones that are represented by nodes known as "centroids" from which all traffic routes are assumed to start and/or finish. They represent an aggregation of all the actual origins and all the actual destinations in each zone. Once the centroids are defined (and thus the set $\mathcal{O}$ of origins and $\mathcal{D}$ of destinations), the movement over an urban network can be expressed in terms of an origin-destination (OD) matrix $\mathbf{T}$, where t_{ks} is the number of trips originating from zone k and ending at zone s.

  (b) The connection between zones and links: each zone (modeled by its centroid) is joined to the road network by special links called "connectors." These links are fictitious links that do not represent any street of the city. The number of connectors depends on the level of detail with which an urban area is represented, but sometimes are chosen arbitrarily.

Indeed, the decision of how many centroids and then how many connectors must be used is closely related to the flow estimation error. For example, the links directly connected to the centroids may lead to incorrect flow and artificial congestion, or most importantly, for the purpose of this work, the resulting routes may be unrealistic.

Although the distribution of centroids and connectors seems crucial to obtain a coherent traffic flow estimation, it has received limited attention. Among the few studies found, Mann [21] presented a model in which every zone was divided into subareas with the aim of reducing assignment error; Friedrich and Galster [22] suggested methods for generating connectors based on geometric features in a microscopic reference scenario; Quian and Zhang [23] proposed a connector optimization algorithm to decide the number and location of connectors in order to minimize the maximum volume/capacity ratio in a given subset of network links by changing the connector travel time; and Jafari et al. [24] used a bilevel method to distribute each centroid demand both to its nearby nodes and to its peripheral nodes. Other methods for traffic network modeling are presented as in Hao and Yang [25] where they introduced the theory of granular computing to model the elements of the multilayer traffic network.

In summary and following some of the conclusions drawn by Quian and Zhang [23], building a transportation model with a good distribution of centroids and connectors is both a difficult and an important task because some problems may arise when any traffic assignment model is used:

  (i) The estimated link flows change significantly depending on the connector configurations

  (ii) If the network model is not well designed and the connectors' travel time is not well defined, the final demand assignment can lead to a solution in which connectors are used to bypass congested links that would otherwise have to be used

  (iii) Too few connectors often lead to artificial congestion in those links adjacent to the connector

  (iv) Since the routes begin with the connectors, if few connectors are designed, the set of routes may be unrealistic or uncompleted

Up to this point, we have not discussed using anything but traditional methods to predict traffic flows, which are usually updated using observed representative link flows, in the network modeling problem. However, the model proposed by Castillo et al. [26] (and then extended by Mínguez et al. [27] or Sánchez-Cambronero et al. [13]) suggests vehicle plate recognition as an alternative way to collect traffic data since this method is much more informative than traditional ones and can therefore be used more efficiently for traffic flow estimation. Other authors, such as Zhou and Mahmassani [28] or Liu et al. [29] or Li et al. [30], also used the information based on the automatic number of plate recognition in their models.

If little attention has been paid to build an appropriate traffic network using traditional methods, even less attention (to our knowledge) has had the development of models to build an appropriate network for traffic flow estimation using plate-recognition-based data, where the route flow is the key variable to be estimated (note that this procedure is just one of the possible techniques of automatic vehicle identification which can also be applied with the method

proposed in this paper). Sánchez-Cambronero et al. [31] addressed this problem by assuming that every node of the network can be the origin and destination of trips and built a node-based OD matrix used as a reference. With this, the authors proposed to use a route enumeration algorithm to build an "exhaustive set of routes" ($\mathfrak{R}$) between the nodes of the network. Then, a route simplification algorithm is proposed based on transferring to adjacent nodes the generated or attracted (reference) demand of those nodes that generate or attract fewer trips than a given threshold (F_{thres}). After the simplification process, a new set of routes $\mathbb{Q}$ is obtained to be used in the plate scanning device location model. However, the authors did not mention the criteria to determine those sets of origin-destination nodes or the method to build the exhaustive set of routes, which, in practice, are two key points of network model design. As it will be shown, this paper deals with these problems.

1.2. The Device Location for Traffic Flow Prediction. Given a traffic network model ($\mathcal{N}, \mathcal{A}$), the device location for the traffic flow prediction problem consists of determining which subset of $\mathcal{A}$ should be observed in order to estimate the traffic flow in the most reliable way. Due to the importance of device locations to achieving trustworthy flow predictions, many authors have addressed the issue of determining the optimal number and locations of traffic counts (see, for example, Yang and Zhou [32], Ehlert et al. [33], Gentili and Mirchandani [34], or Salari et al. [35]). Most of their models are formulated with the assumption that a set $\mathfrak{R}$ of routes is given and fixed. However, we have found a lack of analysis of how the network model design (traffic zones, centroids, connectors, links, etc.) affects this route enumeration.

Location models for plate recognition devices have been studied by Mínguez et al. [27], Castillo et al. [36], Yang and Sun [37], Sánchez-Cambronero et al. [38], Fu et al. [10], and Gentili and Mirchandani [39], among others. Again, all of these authors assume a given and fixed set of routes for the location model but do not analyze how to determine those routes and, even more important, how these location models (and the resulting flow estimates) are affected by uncertain knowledge of the routes. In addition, note that the device location problems do not have a unique solution, and the implications of such selection for later flow estimation are very relevant and important and hence deserve a deep analysis.

1.3. Contributions of This Paper. In the view of the above, this paper proposes a two-step methodology which leads to the following contributions compared with some of the previous studies in the same topic:

(i) We propose (in the first step) a traffic network design method (based on the one proposed by Sánchez-Cambronero et al. [31]) to be used in traffic flow estimation models that use data collected by AVI sensors. This method is an alternative to the classic modeling of the network using centroids and

connectors as those proposed by Quian and Zhang [23] or Jafari et al. [24]. Note that since the estimation models that use this field information from AVI sensors try to reconstruct users' routes, the precision to define network routes must be very high. This is one of the main advantages of the proposed method.

(ii) We propose (in the second step) a new heuristic algorithm to obtain the plate scanning devices' location aimed to obtain the best possible result in terms of traffic flow estimation. The main step forward of this contribution is that it deals with the uncertain knowledge of the network routes. Up to now, the existing models (proposed, for example, by Mínguez et al. [27] or Cerrone et al. [40]) assume the set $\mathbb{Q}$ of modeled routes known and fixed. However, due to traffic conditions, the proposed set of scanned links ($\mathcal{SL}$) may give a set of observed combinations of scanned links ($\mathcal{OSCSL}$) different from the expected set ($\mathcal{ESCSL}$). Therefore, since the routes actually used by vehicles may not have been included in $\mathbb{Q}$, it may be possible that some of these $\mathcal{OSCSL}$ have no intersection with the routes in $\mathbb{Q}$, and hence, the observed flow cannot be assigned. For this, the algorithm expands this set of modeled routes to a bigger set $\mathscr{C}$ of routes so that all of the combinations in $\mathcal{OSCSL}$ have intersections with at least one route.

(iii) We propose a sensibility analysis to evaluate the influence of different parameters of the model on the final solution. This contribution is very important since it may give tools to the transportation planner to decide which is the best value for each parameter depending on the particular case to study.

The rest of the paper is organized as follows: in Section 2, the problem of the uncertain knowledge of network routes is explained; in Section 3, the model proposed by Sánchez-Cambronero et al. [31] is discussed and improved; in Section 4, the proposed algorithm is presented, described, and analyzed; Section 5 is devoted to performing the sensibility analysis of the model parameters in the solution using the well-known Nguyen-Dupuis network to next apply the method to a real network; and finally, some conclusions are provided in Section 6.

2. The Impact of the Uncertain Knowledge of Routes on the Traffic Flow Estimation Results

Traffic estimation models based on vehicle plate recognition (as a particular case of AVI sensors) are based on identifying the circulating vehicles on some subsets of links to reconstruct vehicle routes or partial routes, from which route, OD, and link flows can be derived. As mentioned before, plate recognition has become a useful technique because of the great amount of information it provides compared with that provided by other very common standard methods (see, among others, Castillo et al. [26]).

To illustrate the concepts involved, the simple network with 6 nodes and 18 links, shown in Figure 1, is going to be used. Let us consider the set of reference routes $\mathfrak{R}$, shown in Figure 2(a). Sánchez-Cambronero et al. [31] proposed to obtain this set from a k-shortest path enumeration algorithm using a node-based OD matrix (i.e., assuming that every single node is able to be an origin and destination). In fact, with this, they tried to cover all possible feasible routes in the network although its route flow may be negligible (this paper extends this procedure in order to give more tools to obtain them). From this set, the method proposed by the authors is applied, and a simplified set $\mathcal{Q}$ (see Figure 2(b)) of routes is obtained. This set of routes $\mathcal{Q}$ is assumed to be good enough to perform a traffic flow analysis, and we also assume that this set is the one that the existing methods would use both for locating the sensors and after for the flow estimation. Although the set of routes $\mathcal{Q}$ may be considered very exhaustive after the simplification process, there is a great uncertainty on whether these routes represent reliably the routes actually used by the users.

Let us illustrate this with an example. Consider that the plate recognition devices are installed in the set SL = {1, 2, 3, 4, 5, 7, 8, 9, 12, 13, 14, 15, 16, 17}. With this information and using existing methods, Table 1 is developed which shows each set of expected combinations (s) of scanned links ($\mathcal{ESCSL}$) as the intersection of each route in $\mathcal{Q}$ with the set $\mathcal{SL}$. Let us now consider that we develop a field test, i.e., we install plate reader devices (or AVI devices), so we can obtain the associated observed flow $\widehat{w}_s$ (see the last column in the table). This means, for example, that those vehicles identified in links 1 and 5 (60.31 in this case) belong to route 1. Also, vehicles identified in links 9 and 17 (5.62) belong to route 19. In this example, (which is not the general case), each set s is associated with just one route (q) in set $\mathcal{Q}$, i.e., it is expected that the observed flow $\widehat{w}_s$ would be able to derive all the route flows by using the following relation:

$$\widehat{w}_s = \sum_{q \in \mathcal{Q}} \lambda_s^q f_q, \quad \forall s \in \mathcal{ESCSL}, \tag{1}$$

where $\widehat{w}_s$ is the observed flow in each set $\mathcal{ESCSL}$, f_q is the estimated flow in routes in set $\mathcal{Q}$, and λ_s^q is the element of the route-scanned combination incidence matrix for route q, which equals one if route q contains the subset s of scanned links and only those, and zero, otherwise. Therefore, the link flows can be calculated as follows:

$$v_a = \sum_{q \in \mathcal{Q}} \delta_r^a f_q; \quad \forall a \in \mathcal{A}, \tag{2}$$

where v_a is the flow for link a and δ_r^a is the relation incidence matrix between link and route flows.

However, suppose that the related field test reveals that new sets of combinations of scanned links $\mathcal{NSCSL}$ appear in addition to those shown in Table 1, but in this case, we do not find any intersection with routes of set $\mathcal{Q}$. This new set together with its corresponding observed flow is shown in Table 2. For example, we have found that there were vehicles scanned only in link 16 (18.86) which do not match with any
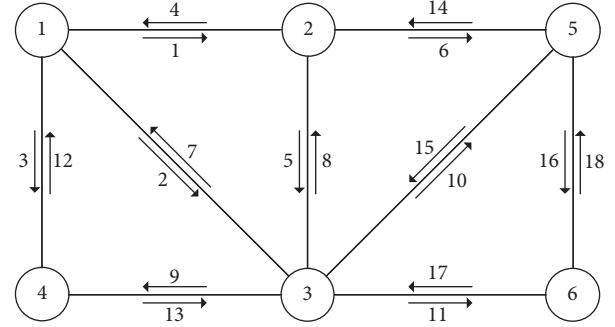
Figure 1: Elementary example network.

route in $\mathcal{Q}$. Note that $\mathcal{ESCSL} \cup \mathcal{NSCSL} = \mathcal{OSCSL}$, i.e., the observed set of combinations of scanned links.

To assign these new combinations, it is necessary to look for compatible routes with the new set of scanned links, for example, within set $\mathfrak{R}$. Taking the simplified sets in $\mathcal{Q}$ plus the new ones gives us a new, larger set $\mathcal{C}$ of routes (see Figure 2(c)). In other words, it is necessary to complete the set of routes once the field data were collected. Doing this, the uncertain knowledge of routes will be reduced. This paper proposes to include this procedure in the location model to improve the expected traffic flow estimation results.

With this last set of routes, it is already possible to solve a flow estimation problem using one of the methods proposed in the literature (see Mínguez et al. [27], which used a generalized least square method and Sánchez-Cambronero et al. [13], which used a Bayesian network model).

To compare estimation results using set $\mathcal{Q}$ and using set $\mathcal{C}$, let us define the link relative absolute error (LRAE) as

$$\mathrm{LRAE}(a) = \frac{|v_a - v_a^{\mathrm{real}}|}{v_a^{\mathrm{real}}}, \tag{3}$$

where v_a and v_a^{real} are the estimated flow and (assumed) real flow for link a. Such measure of the solution quality should be calculated over the link flows because the set of links $\mathcal{A}$ will remain constant regardless of the studied simplification or the number of links in set $\mathcal{SL}$ (note that the number of routes in $\mathcal{C}$ may vary depending on the field data collected). Therefore, once the estimation of flows has been made, it is possible to calculate the LRAE using equation (3). Table 3 compares the LARE after estimating the link flows using the set of routes in R versus using the set of routes in $\mathcal{C}$, proving the value of performing this calculation.

To check if the estimation of the link flows in the whole network is adequate, we can use the root mean absolute value relative error (RMARE), defined as

$$\mathrm{RMARE} = \frac{1}{n} \sum_a \frac{|v_a - v_a^{\mathrm{real}}|}{v_a^{\mathrm{real}}}, \tag{4}$$

where n is the number of links in the network. The value obtained using equation (4) for the example given in this section is shown in the last row of Table 3. It indicates that although both sets of routes yield good flow estimations,

Route $\mathcal{R}$	Route links
1	1
2	2
3	3 13
4	1 5
5	3
6	1 6
7	2 11
8	3 13 11
9	4
10	5
11	4 2
12	4 3 13
13	4 3
14	5 9
15	4 2 9
16	6
17	6 16
18	5 11
19	4 2 11
20	7
21	8 4
22	9 12
23	8
24	11 18 14
25	9
26	10
27	11 18
28	8 6
29	11
30	12
31	12 1
32	13 8
33	13 10 14
34	13
35	13 10
36	12 1 6
37	13 11 18
38	13 11
39	14 4
40	14
41	15
42	16 17
43	14 5
44	14 4 3
45	15 9
46	16 17 9
47	16
48	15 11
49	17 8 4
50	17 7
51	17 8
52	17
53	17 9
54	18

(a)

Sánchez-Cambronero et al. [31] →

Route $\mathcal{Q}$	Route links
1	1 5
2	3
3	2
4	3 13
5	4
6	4 3 13
7	4 3
8	5 9
9	4 2 9
10	5
11	4 2
12	9 12
13	14 4
14	16 17
15	14 5
16	14 4 3
17	15 9
18	16 17 9
19	15
20	8 4
21	7
22	9

(b)

Set {$\mathcal{SL}$} →

Route C	Route links
1	1 5
2	3
3	2
4	3 13
5	4
6	4 3 13
7	4 3
8	5 9
9	4 2 9
10	5
11	4 2
12	9 12
13	14 4
14	16 17
15	14 5
16	14 4 3
17	15 9
18	16 17 9
19	15
20	8 4
21	7
22	9
23	1
24	1 6
25	3 13 11
26	6 16
27	8
28	11 18 14
29	8 6
30	12
31	12 1
32	13 8
33	13 10 14
34	13
35	13 10
36	12 1 6
37	13 11 18
38	13 11
39	14
40	16
41	17 8 4
42	17 7
43	17 8
44	17
45	17 9

(c)

FIGURE 2: Operation scheme and sets of (a) reference routes {$\mathcal{R}$}, (b) simplified routes {$\mathcal{Q}$}, and (c) combination routes {$\mathcal{C}$}.

when the set of routes $\mathcal{C}$ is used, the traffic estimation model performs better.

Note that the solution for this problem is not unique in terms of links included in $\mathcal{SL}$ but constitutes a particular solution obtained through an optimization problem (see, for example, Castillo et al. [36]). Taking advantage of this fact, this paper proposes a heuristic algorithm to find set $\mathcal{SL}$ that minimizes the RMARE obtained using equation (4) and that provides the best set of routes which are able to represent the traffic flow in the entire network.

3. Discussion and Improvement of the Sánchez-Cambronero et al. [31] Model

Let us now change the network used to the well-known Nguyen-Dupuis network shown in Figure 3 which will be used. It consists of 13 nodes and 38 links. Let us suppose that, to have a reference level for demand, we have used the data from a study where traditional methods were used. Figure 3 shows the network divided into 4 zones together with its associated origin-destination (OD) matrix and the traffic

TABLE 1: Set of expected combinations of scanned links ($\mathscr{ESCSL}$) associated with simplified routes in set $\mathcal{Q}$ and their values of observed flow.

Routes, $\mathcal{Q}$	Route links	Set $\mathscr{ESCSL}$	Info obtained from set $\mathscr{SL}$														Observed flow
			1	2	3	4	5	7	8	9	12	13	14	15	16	17	
1	1 5	1	X				X										60.31
2	3	2			X												6.79
3	2	3		X													83.80
4	3 13	4			X							X					31.85
5	4	5				X											8.75
6	4 3 13	6			X	X						X					86.42
7	4 3	7			X	X											1.04
8	5 9	8					X			X							1.00
9	4 2 9	9		X		X				X							1.09
10	5	10					X										138.32
11	4 2	11		X		X											49.39
12	9 12	12								X	X						24.73
13	14 4	13				X							X				8.70
14	16 17	14													X	X	45.68
15	14 5	15					X						X				44.35
16	14 4 3	16			X	X							X				3.15
17	15 9	17								X				X			1.00
18	16 17 9	18								X					X	X	5.62
19	15	19												X			63.21
20	8 4	20				X			X								49.65
21	7	21						X									88.85
22	9	22								X							117.87

TABLE 2: New sets of combinations of scanned links ($\mathscr{NSCSL}$) beside their values of observed flow.

Route, $\mathcal{Q}$	Route links	Set $\mathscr{NSCSL}$	Info obtained from set $\mathscr{SL}$														Observed flow
			1	2	3	4	5	7	8	9	12	13	14	15	16	17	
		23	X														3.82
		24													X		18.86
		25							X								2.43
		26										X					4.56
		27								X							1.34
		28	X							X							2.00
		29							X			X					1.00
		30									X	X					1.33
		31										X					6.91
		32				X			X							X	1.22
		33						X								X	1.24
		34							X							X	1.55
		35														X	2.36
		36								X						X	4.62

link parameters (BPR function parameters and attraction and generation capacities) that will be used and explained in the following.

As discussed in Section 1, the use of a centroid as an aggregation of all origins and destinations within a determined zone implies a flow estimation error. The same occurs with the use of connectors (see Quian and Zhang [23]). Therefore, the number and location of connectors and centroids should be determined carefully due to the errors they may cause during traffic flow estimation.

Trying to solve some of these problems, Sánchez-Cambronero et al. [31] proposed a model that allows to design a traffic network that minimizes the negative effects of the use of centroids and connectors by replacing them with

"origin nodes" and "destination nodes" in such a way that all trip origins and destinations are assigned to these nodes of the network in accordance with the vehicle paths and the network shape. An application of this can be observed in Figure 4. Suppose a vehicle actually performs the trip indicated in Figure 4(a), whose true origin is somewhere in link 34 and whose true destination is somewhere in link 5. This method assumes that every vehicle has its origin in the first node of its trip. In this example, node 1 is the first node, so it is the origin node (see Figure 4(b)). Similarly, the destination node is taken to be the last node, where the vehicle passes; in this example, the last node of the route is node 3, so it would be the destination node. This resulting route is finally included in the traffic model. Note that this

TABLE 3: Link flow in estimated and real situations and relative absolute error value.

Link	Results using set $\mathcal{Q}$			Results using set $\mathcal{C}$		
	v_a^{est}	v_a^{real}	LRAE	v_a^{est}	v_a^{real}	LRAE
1	60.31	66.13	0.09	66.13	66.13	0.00
2	134.29	134.28	0.00	134.29	134.28	0.00
3	129.26	129.26	0.00	129.26	129.26	0.00
4	208.19	209.41	0.01	209.41	209.41	0.00
5	243.97	243.97	0.00	243.97	243.97	0.00
6	0.00	23.62	1.00	23.62	23.62	0.00
7	88.85	90.09	0.01	90.09	90.09	0.00
8	49.65	55.84	0.11	55.84	55.84	0.00
9	151.32	151.32	0.00	151.32	151.32	0.00
10	0.00	3.80	1.00	1.99	3.80	0.48
11	0.00	39.38	1.00	2.50	39.38	0.94
12	24.73	28.08	0.12	28.07	28.08	0.00
13	118.28	127.52	0.07	127.52	127.52	0.00
14	56.20	62.09	0.09	62.09	62.09	0.00
15	64.21	64.21	0.00	64.21	64.21	0.00
16	51.30	63.18	0.19	69.71	63.18	0.10
17	51.30	55.32	0.07	55.32	55.32	0.00
18	0.00	4.68	1.00	0.77	4.68	0.83
	RMARE = 0.27			RMARE = 0.13		

trip is one of the 470 trips that go from zone 1 to zone 3 (see Figures 3 and 4(b)), and depending on the choice of connectors, the path of this trip may be wrongly defined if the traditional method based on centroids and connectors is used.

3.1. Characterization of Origin-Destination Nodes.
According to the assumption described above, every node of the network will generate or attract trips depending on the characteristics of the adjacent links, i.e., depending on the capability of the adjacent links to attract and generate trips (i.e., number of on-street parking spaces, number of private parking spaces, etc.). This can be quantified in terms of the capacity of link a to attract trips, CA_{la}, and to generate trips, CG_{la} (similar assumption was made in Levy and Benenson [41]). Then, the capacities of each node are calculated as follows:

$$CA_{ni} = \sum_{a \in \mathscr{A}_{out}^i} CA_{la}, \quad (5)$$

$$CG_{ni} = \sum_{a \in \mathscr{A}_{in}^i} CG_{la}, \quad (6)$$

where equation (5) expresses the capacity of node i to attract trips, depending on the capacities of the adjacent links leaving the node ($\mathscr{A}_{out}$), and equation (6) expresses the capacity of node i to generate trips according to the capacities of the adjacent links arriving at the node ($\mathscr{A}_{in}$). Such capacity values for each link in the network are shown in Figure 3.

Then, according to these capacity values, one can obtain the proportion of the total trips attracted and generated by zone Z that begins or ends at node i. To do this, we propose the following expressions:

$$PA_{ni} = \frac{CA_{ni}}{\sum_{j \in \mathbf{Z}} CA_{nj}}, \quad (7)$$

$$PG_{ni} = \frac{CG_{ni}}{\sum_{j \in \mathbf{Z}} CG_{nj}}, \quad (8)$$

where PA_{ni} and PG_{ni} are these proportions which are shown in Figure 5 for the case of the example with the Nguyen-Dupuis network.

With the values obtained using equations (7) and (8), the relationship between the number of trips made between the origin and destination nodes and the number of trips made between the zones to which these nodes belong can be established as follows:

$$T_{ij} = \widehat{T}_{ZiZj} PA_{ni} PG_{nj}, \quad (9)$$

where T_{ij} is the number of trips from node i to node j and $\widehat{T}_{ZiZj}$ is the number of trips obtained through an out-of-date OD matrix (see Figure 5).

The completion of this step entails the definition of a new OD matrix defined by trips between nodes rather than a matrix defined by trips between zones. This means that, for the Nguyen-Dupuis network used in this example, the matrix with 8 OD flows is transformed in this step into a matrix with 84 OD flows (note that the simplification process proposed in this paper will reduce this number of OD flows. The process will be explained in the following section). This may be seen as a drawback, but note that, by doing this together with the simplification process proposed in Sánchez-Cambronero et al. [31], the problems associated with the use of connectors can be faced better. In addition, the key variables for the plate scanning technique models are the route flows (not the OD flows), whose number should be almost the same for both OD matrices.

3.2. The Definition of the Exhaustive Set of Routes $\mathfrak{R}$.
Network path enumeration is a requirement for developing a model based on plate scanning data because it is needed for both traffic estimation and device location. This implies that, at this stage, we need an exhaustive set of routes $\mathfrak{R}$. To construct this set, we propose to find the k-shortest paths of the extended OD matrix. Although transportation planners usually recommend to use $k = 3$ and discard routes with more than 1.5 times the shortest path (see, for example, Sheffi [17]), in Section 4.1, a sensitivity analysis of various suitable values for k will be performed.

3.3. The Network Simplification and the Set of Routes $\mathcal{Q}$.
The aim of the simplification process proposed by the authors is to transfer to adjacent nodes the generated or attracted demand of those nodes that generate or attract fewer trips than a given threshold (F_{thres}). This is a good way to avoid problems derived from the use of connectors since the start point of the simplification process is a set of routes built based on the physical characteristics of the real network and not based on the definition of artificial links as are the

Zone	1	2	3	4	Generated trips
1	0	640	470	0	1110
2	640	0	0	530	1170
3	470	0	0	110	580
4	0	530	110	0	640
Attracted trips	1110	1170	580	640	

(b)

Link	c_a	q_a	α_a	γ_a	CA_{la}	CG_{la}	Link	c_a	q_a	α_a	γ_a	CA_{la}	CG_{la}
1	7	700	1	4	250	2	20	9	700	1	4	250	300
2	9	560	1	4	250	100	21	9	700	1	4	200	7
3	9	700	1	4	300	250	22	14	560	1	4	250	10
4	9	280	1	4	250	2	23	5	375	1	4	2	250
5	8	560	1	4	250	2	24	9	420	1	4	5	3
6	11	560	1	4	200	5	25	5	280	1	4	3	5
7	12	560	1	4	300	1	26	9	280	1	4	5	2
8	5	375	1	4	250	5	27	4	280	1	4	2	2
9	7	700	1	4	5	250	28	5	280	1	4	3	2
10	12	560	1	4	2	200	29	4	700	1	4	3	3
11	12	420	1	4	6	5	30	9	280	1	4	5	250
12	9	420	1	4	2	2	31	8	560	1	4	2	250
13	12	420	1	4	5	4	32	4	700	1	4	2	7
14	5	700	1	4	10	1	33	4	700	1	4	2	4
15	4	280	1	4	2	2	34	9	560	1	4	10	250
16	7	140	1	4	5	50	35	7	140	1	4	12	5
17	5	700	1	4	5	5	36	14	560	1	4	15	200
18	9	700	1	4	10	150	37	11	560	1	4	5	200
19	4	700	1	4	2	5	38	9	280	1	4	2	5

(c)

FIGURE 3: Nguyen-Dupuis network and its division into 4 traffic zones (Z) and the associated OD matrix.

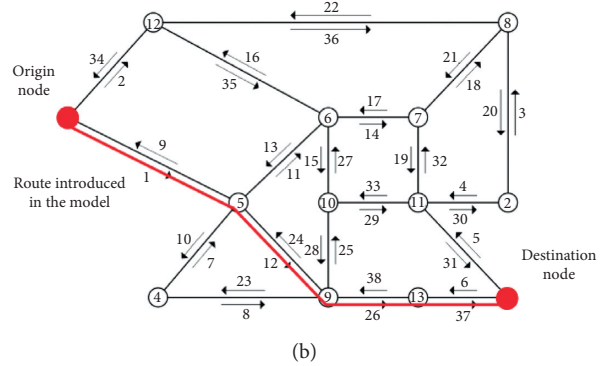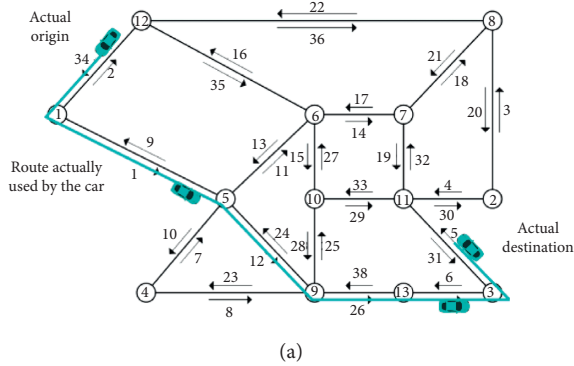(a)   (b)

FIGURE 4: Origin and destination nodes' definition: (a) route actually used by a certain vehicle; (b) route to be included in the traffic model.
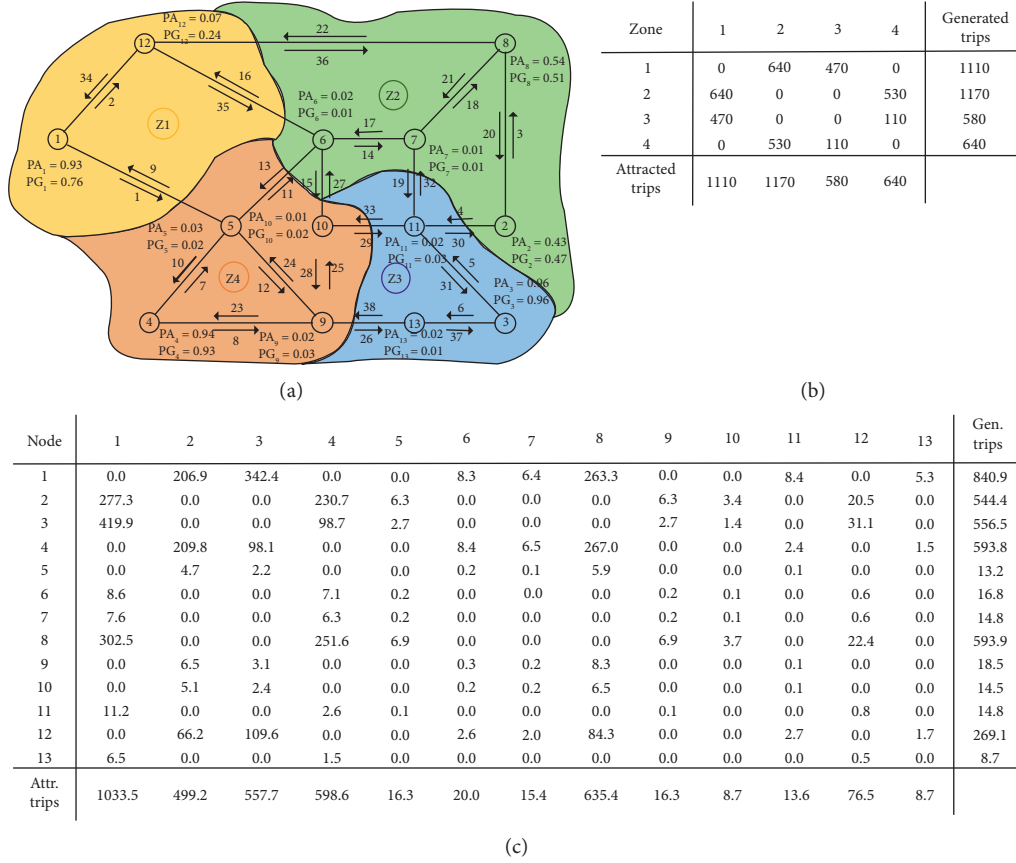
(a)

Zone	1	2	3	4	Generated trips
1	0	640	470	0	1110
2	640	0	0	530	1170
3	470	0	0	110	580
4	0	530	110	0	640
Attracted trips	1110	1170	580	640	

(b)

Node	1	2	3	4	5	6	7	8	9	10	11	12	13	Gen. trips
1	0.0	206.9	342.4	0.0	0.0	8.3	6.4	263.3	0.0	0.0	8.4	0.0	5.3	840.9
2	277.3	0.0	0.0	230.7	6.3	0.0	0.0	0.0	6.3	3.4	0.0	20.5	0.0	544.4
3	419.9	0.0	0.0	98.7	2.7	0.0	0.0	0.0	2.7	1.4	0.0	31.1	0.0	556.5
4	0.0	209.8	98.1	0.0	0.0	8.4	6.5	267.0	0.0	0.0	2.4	0.0	1.5	593.8
5	0.0	4.7	2.2	0.0	0.0	0.2	0.1	5.9	0.0	0.0	0.1	0.0	0.0	13.2
6	8.6	0.0	0.0	7.1	0.2	0.0	0.0	0.0	0.2	0.1	0.0	0.6	0.0	16.8
7	7.6	0.0	0.0	6.3	0.2	0.0	0.0	0.0	0.2	0.1	0.0	0.6	0.0	14.8
8	302.5	0.0	0.0	251.6	6.9	0.0	0.0	0.0	6.9	3.7	0.0	22.4	0.0	593.9
9	0.0	6.5	3.1	0.0	0.0	0.3	0.2	8.3	0.0	0.0	0.1	0.0	0.0	18.5
10	0.0	5.1	2.4	0.0	0.0	0.2	0.2	6.5	0.0	0.0	0.1	0.0	0.0	14.5
11	11.2	0.0	0.0	2.6	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.8	0.0	14.8
12	0.0	66.2	109.6	0.0	0.0	2.6	2.0	84.3	0.0	0.0	2.7	0.0	1.7	269.1
13	6.5	0.0	0.0	1.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.5	0.0	8.7
Attr. trips	1033.5	499.2	557.7	598.6	16.3	20.0	15.4	635.4	16.3	8.7	13.6	76.5	8.7	

(c)

FIGURE 5: PA_{ni} and PG_{ni} proportions and the extended OD matrix for the example with the Nguyen-Dupuis network.

connectors. The next section deals with a detailed description of the process to obtain the simplified set of routes $\mathcal{Q}$.

4. The Proposed Algorithm

In this section, we present the proposed two-step algorithm that allows us to determine the traffic flow network to be used and the location of AVI devices to perform the traffic flow estimation.

Inputs:

(i) Network model: sets $\mathcal{A}$ of links and $\mathcal{N}$ of nodes and links parameters C_a and Q_a

(ii) Cost and flow thresholds for the simplification process (C_{thres}, F_{thres})

(iii) Capacities of links to attract and generate trips (CA_{la}, CG_{la})

(iv) Number of trips $\widehat{T}_{ZiZj}$ based on network zones obtained through an out-of-date OD matrix

(v) RMARE$^{\text{best}}$ set to an initial value of 10 and iter$^{\text{max}}$ set to a maximum of 1000 iterations to carry out

Step 0: obtaining reference routes:

(i) Obtaining a new extended matrix T_{ij} of the number of trips from node i to node j using the procedure described in Section 2.1.1.

(ii) Find the k-shortest paths using the extended matrix to obtain the set of reference routes ($\mathfrak{R}$) with their respective reference route flows f_r^0 that can be obtained by using, for example, MNL stochastic user equilibrium (see, for example, Sheffi [17] and Sánchez-Cambronero et al. [42]).

Step 1: network simplification:

The simplified network is obtained using the green part of the algorithm shown in Figure 6. It allows us to decide what nodes can be origin/destination or not, based on a demand threshold flow F_{thres} established by the transportation planner. The simplification process finishes when it does not exist any node which complies simplification criterion.

Step 1.0: initialization:

Initialize the set of simplified routes $\mathcal{Q}$ as the exhaustive set $\mathfrak{R}$.

Step 1.1: search the least demand node:

The algorithm searches the node with the lower demand for both cases: origins and destination, i.e., least $G_i = \sum_j T_{ij}$ or $A_j = \sum_i T_{ij}$.

Step 1.2: checking the simplification criterion:

Once the candidate node has been identified to lose its origin/destination condition, it is checked if it meets the simplification criteria established, i.e., if $G_i \leq F_{\text{thres}}$ or $A_j \leq F_{\text{thres}}$, go to Step 1.3. Else, continue with Step 1.5.

Step 1.3: demand transmission:

If the candidate node is able to lose its OD condition, i.e., its OD demand is below the threshold flow F_{thres}, then the possibility that this demand has its origin or destination in some other node of the network is evaluated. Transmission will be made to the closer node of each route if and only if the node that could receive or emit demand is at a shorter distance than C_{thres}, and the involved routes will have to be modified accordingly. Otherwise, the candidate node and the demand will disappear, and the set of routes will be remade (see Sánchez-Cambronero et al. [31] for more details).

Step 1.4: set of routes' update:

Update the set of routes $\mathcal{Q}$ and their corresponding flow values f_q^0. Go to Step 1.1.

Step 1.5: output of Step 1 and inputs to Step 2:

As a result of the application of this simplification algorithm, we will obtain a new set of routes $\mathcal{Q}$ from the original set $\mathfrak{R}$ due to removal and re-enumeration of reference routes. Simultaneously, the route flows f_q^0 for this new set will be updated. Proceed to Step 2.

Step 2: location and estimation problem:

The AVI device location problem is a complex problem that does not have a unique solution. In this step, we assume simulated "real flows" in order to obtain the values for the observed flow $\hat{w}_s$ depending on the device location. The main objective of this step of the algorithm is to obtain the subset of the set of scanned links $\mathcal{SL}$ that gives the best possible flow estimation.

To achieve this objective, an optimization problem is incorporated into the algorithm. This problem is based on previous problems studied by Mínguez et al. [27] or Cerrone et al. [40], but as an improvement, we have included a new restriction that examines different options for the device location in order to assess which of them lead to the best solution in terms of flow estimation.

Step 2.1: scan device location problem:

The scan device location problem is formulated as

$$\max_{z_a; y_q} M = \sum_{q \in \mathcal{Q}} f_q^0 y_q, \tag{10}$$

subject to

$$\sum_{a \in \mathcal{A}} P_a z_a \leq B, \quad \forall a \in \mathcal{A}, \tag{11}$$

$$\sum_{a \in \mathcal{A}} \delta_a^q z_a \geq y_q, \quad \forall q \in \mathcal{Q}, \tag{12}$$

$$\sum_{a \in \mathcal{A}} \gamma_a^{qq_1} z_a + \sum_{a \neq b \in \mathcal{A}} \sigma_{ab}^{qq_1} x_{ab} \geq y_q, \quad \forall (q, q1) \in \mathcal{Q}^2 | q > q1;$$
$$\sum_{a \in \mathcal{A}} \delta_q^a \delta_{q_1}^a > 0, \tag{13}$$

$$2x_{ab} - z_a - z_b \leq 0, \quad \forall a \neq b \in \mathcal{A}, \tag{14}$$

$$\sum_{a \in \mathcal{A}} S_a^{\text{iter}} z_a \leq \sum_{a \in \mathcal{A}} S_a^{\text{iter}}, \quad \forall \text{iter} \in \mathbf{I}. \tag{15}$$

Objective function (10) maximizes the observed route flow in terms of f_q^0, which is the reference flow through route q; y_q, theoretically, is a binary variable which equals 1 if a route can be distinguished from others and 0, otherwise; however, to speed up the model, it is set as a continuous variable $y_q \in [0, 1]$ (Mínguez et al. [27]).

Constraint (11) satisfies the budget requirement, where z_a is a binary variable that equals 1 if link a is scanned and 0, otherwise. This constraint guarantees that we will have a number of scanned links with cost P_a for link a that does not exceed the established limited budget B.

Constraint (12) ensures that any distinguished route contains at least one scanned link (for this reason, they are usually known as covering constraints). This constraint is indicated by the parameter δ_a^q, which is the element of the link-route incidence matrix.

Constraints (13) are the diversification constraints. They indicate that route q must be distinguished from the other routes in at least one scanned link a. This happens if

(i) $\sum_a \gamma_a^{qq_1} z_a \geq 1$ because $\gamma_a^{qq_1}$ is 1 if link a is contained either in q or in $q1$ (not in both) and 0, otherwise.

(ii) $\sum_a \sigma_{ab}^{qq_1} x_{ab} \geq 1$ because vehicles using q and $q1$ use the same links but in different order since $\sigma_{ab}^{qq_1}$ is 1 if links a and b are both in routes q and $q1$, but they appear in a different order.

Note that if $y_q = 0$, constraints (13) always hold. The definition of these constraints taking into account that $q > q1$ and $\sum_{a \in A} \delta_q^a \delta_{q_1}^a > 0$ avoids the comparison of a great amount of routes without common links, resulting an important reduction of the computational time. This is important for the analysis of real-size networks and is usually forgotten in many papers, for example, in Castillo et al. [9, 43] and Cerrone et al. [40], among others.
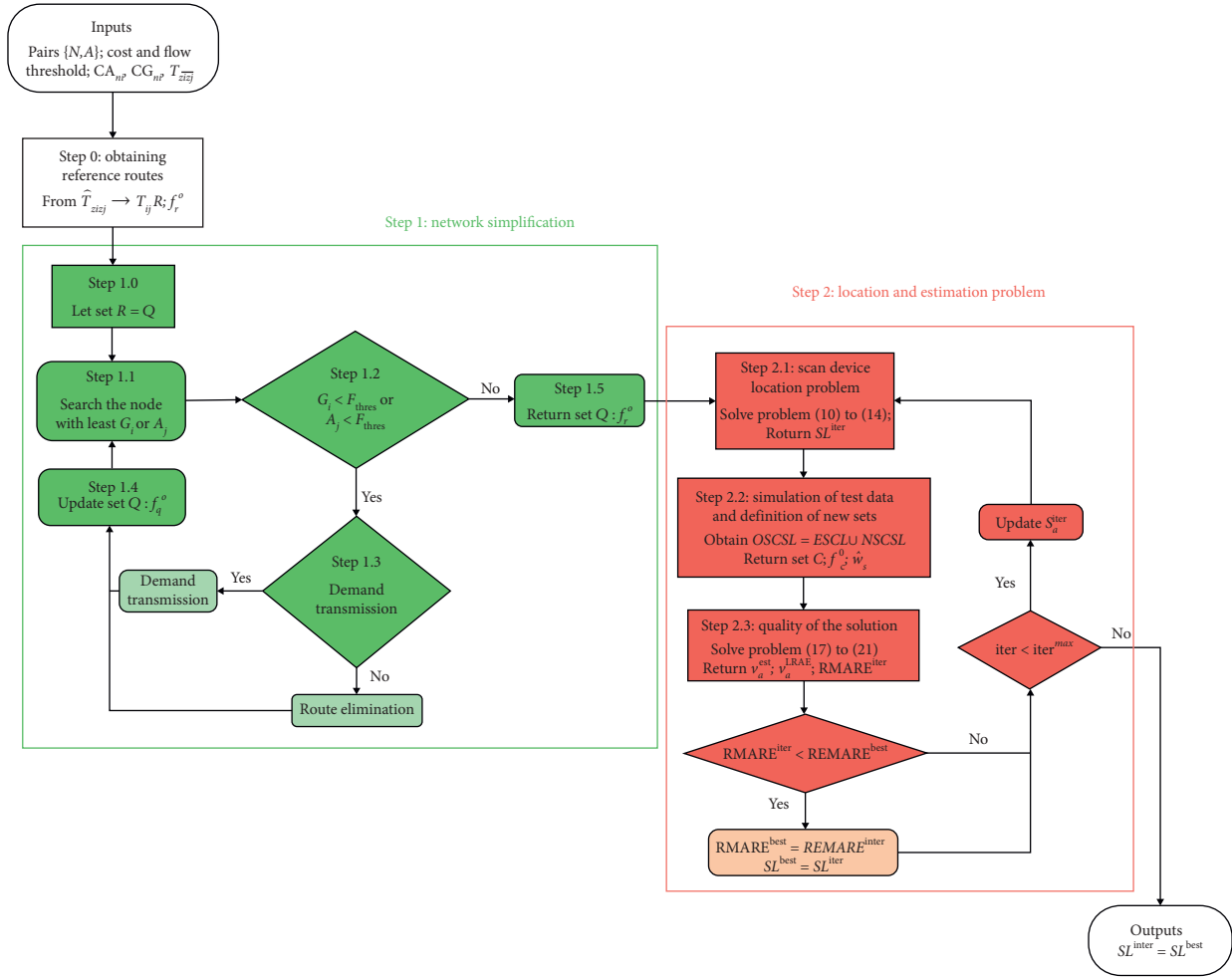
Figure 6: Flowchart of the algorithm.

Constraint (14) is the logical constraint linking the binary variable x_{ab} to z_a and z_b (see Cerrone et al. [40]).

Finally, since this model is a part of an iteration process, we propose additional constraints (15). S_a^{iter} is a matrix that grows with the number of iterations, in which each row reflects the set $\mathscr{SL}$ resulting from each iteration carried out up to then by the model in such a way that S_a^{iter} is 1 if link a was proposed to be scanned in the solution provided on iteration iter and 0, otherwise. This ensures that each iteration keeps the previous solutions and does not permit the process to repeat a solution in future iterations. That is, each iteration carried out by the algorithm is forced to search for a different solution with the same value of objective function (10).

Step 2.2: simulation of test ("real") data and definition of the new set of routes $\mathscr{C}$:

To test the quality of the solution, we have simulated "real" route data to carry out the process

described in Section 1.3, which through the new set of combinations of scanned links $\mathscr{NSCSL}$, observed by subset $\mathscr{SL}$ (obtained in Step 2.1), searches compatible combinations for these new observed routes and forms the observed set of combinations of scanned links ($\mathscr{ESCSL} \cup \mathscr{NSCSL} = \mathscr{OSCSL}$) and finally obtains set $\mathscr{C}$ of all available routes.

The real data were obtained multiplying each T_{ij} of the extended matrix in Step 0 by a random number $U(0.8\text{–}1.2)$ and executing a MNL SUE assignment to a k-shortest path enumeration set by using $k = 7$. From this set of "real" routes, we obtain $\hat{w}_s$, where $s \in \mathscr{OSCSL}$.

Step 2.3: measuring the quality and updating the solution:

Once we have the $\hat{w}_s$ and $\hat{v}_a$ "observed flow" (note that plate scanning devices also allow us to observe link flows on links of $\mathscr{SL}$), a traffic flow estimation can be carried out. As we have shown in Section 1.3,

this can be done by several mathematical methods. In this paper, we have used a generalized least square method as follows:

$$\min_{V, Fz} = \sum_{c \in \mathscr{C}} U_c^{-1} \left(\frac{f_c - f_c^0}{f_c^0} \right)^2 + \sum_{a \in \mathscr{SL}} Y_a^{-1} \left(\frac{v_a - \widehat{v}_a}{\widehat{v}_a} \right)^2, \quad (16)$$

subject to

$$\widehat{w}_s = \sum_{c \in \mathscr{C}} \beta_s^c f_c, \quad \forall s \in \mathscr{OSCSL}, \quad (17)$$

$$v_a = \sum_{c \in \mathscr{C}} \delta_a^c f_c, \quad \forall a \in \mathscr{A}, \quad (18)$$

$$f_c \geq 0, \quad \forall c \in \mathscr{C}, \quad (19)$$

$$v_a \geq 0, \quad \forall a \in \mathscr{A}, \quad (20)$$

where U_c^{-1} and Y_a^{-1} are the inverses of the variance-covariance matrices corresponding to the flow in routes in $\mathscr{C}$ and the observed flow in a, respectively, $\widehat{w}_s$ is the observed flow in each set $\mathscr{OSCSL}$, f_c is the estimated flow of routes in set $\mathscr{C}$, and β_s^c is the element of the route-scanned combination incidence matrix corresponding to route c, which is 1 if route c contains subset s of scanned links and only those, and 0 otherwise.

With this traffic estimation, one can compare the quality of the results obtained for the link flows using equation (3). As shown in Section 1.3, the quantification of the error is made using the results of the link flows since the set of these remains constant in all the iterations of the proposed algorithm. Additionally, the use of equation (4) allows us to check the global quality of the modeled network. For each iteration carried out by the algorithm, the value of RMARE[iter] is evaluated with respect to the best value RMARE[best]. If RMARE[iter] < RMARE[best], then the algorithm updates RMARE[iter] = RMARE[best] and considers the set of scanned links $\mathscr{SL}^{iter}$ to be the best solution up to that iteration. If iter < iter[max], then the algorithm updates S_a^{iter} and goes to Step 2.1; otherwise, return the solution $\mathscr{SL}^{iter}$ as the best solution.

The complete process of this step can be observed on the right side of Figure 6.

5. A Sensitivity Analysis of the Model Results

5.1. The Nguyen-Dupuis Network. In this section, a sensitivity analysis of the model results depending on the value of some parameters is presented. In particular, we have used the Nguyen-Dupuis network shown in Figure 4, and we have analyzed the influence of (i) the partial knowledge of the routes in terms of the k value for obtaining the reference set of routes, (ii) the degree of network simplification in terms of F_{thres}, and (iii) the available budget B for locating AVI devices.

For comparison purposes, a set of initial values was considered, and then the algorithm was applied. As a base situation, it was assumed that $P_a = 1$, $B = 16$ (hence, the number of installed cameras is 16), and the threshold flow F_{thres} required for the simplification network method (Step 1 in Figure 6) is 50. In all studied cases, the parameters on different links in the network, shown in Figure 3, remain constant throughout the iterations carried out by the model.

5.1.1. Influence of the Partial Knowledge of the Routes. To perform this analysis, a route enumeration algorithm was used to check the effect of considering k routes on each OD pair for the reference set $\mathfrak{R}$. The algorithm used here is based on Yen's k-shortest path algorithm (Yen [44]). This algorithm introduces into the model an initial reference route matrix $\mathfrak{R}$ that varies in size according to the value of k. Then, in Step 1, shown in Figure 6, the reference route matrix is simplified to $\mathcal{Q}$, taking into account $F_{thres} = 40$, i.e., those nodes which attract or generate flow which is below 40 lose their OD condition, its demand is transferred to other adjacent nodes, and the corresponding routes are grouped and hence simplified. This reduces the set of routes Q used in the model, (10)–(15), as shown in the third column of Table 4.
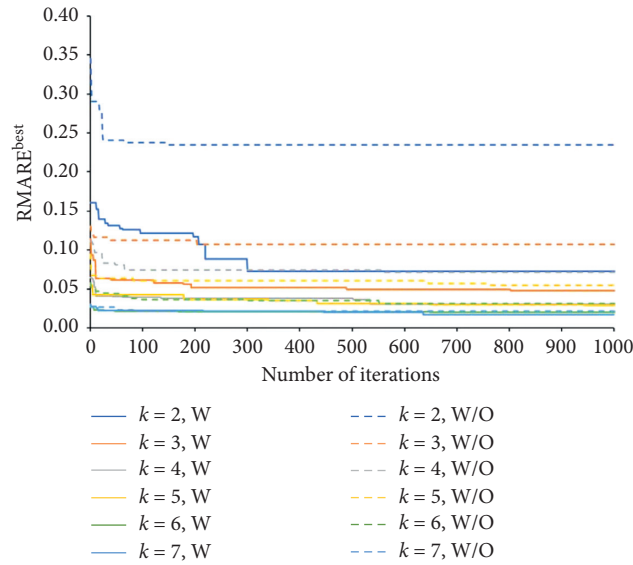
It is well known that the enumeration of routes is essential to solve this location problem. In this analysis, the k-shortest path is used to generate routes between each pair of nodes, but it may exist other routes (actually used by the vehicles) which they do not have been taken into account in the model (this is the reason why the "real situation" has been simulated using $k = 7$) (see Section 1). In this sensitivity analysis, a range of values of k, which is usually used in this kind of transportation analysis, has been selected, with values of $k = 3$ or 4 (Owais et al. [45]). Although values such as $k = 5$, 6 or 7 are rarely used to enumerate routes (Bonsall et al., [46]; Hazelton, [47]), because it is not normal that it exists such amount of routes for each O-D pair and because it increases the computation time and complexity of the problem, we have checked its influence in the simple example network. According to some authors, a value of k less than 2-3 would be unreasonable when working with this type of model (Sheffi [17]).

Figure 7 shows the evolution of RMARE[best] for all iterations performed, considering different k routes by each OD pair to construct the set of routes $\mathfrak{R}$. The continuous line represents the results obtained with the proposed algorithm. Each step in the lines indicates that a new solution has been found that is better than the previous best solution, i.e., with lower RMARE. As expected, better global solutions are found for higher number of k (higher number of reference routes in $\mathfrak{R}$).

Note also that the first solution in all cases provides the worst results, and it is improved with the iteration process, corresponding to a step forward compared with the previous models. Note that models of Mínguez et al. [27] or Cerrone et al. [40] used a fixed set of routes (let us say set $\mathcal{Q}$), so their estimation would be worse than the one provided in the first iteration since they do not work with the improved set $\mathscr{C}$. In any case and in order to explicitly probe this, we have applied our proposed Step 2 but with the set of routes in $\mathcal{Q}$ assumed

TABLE 4: Number of routes that appear with the application of the proposed algorithm according to the k value in the best solution in each case study.

k value	Number of routes in $\mathcal{R}$	Number of routes in $\mathcal{Q}$	Added routes compatible with $\mathcal{SL}$	Number of routes in $\mathcal{C}$
2	117	44	46	90
3	150	66	25	91
4	162	72	20	92
5	168	76	20	96
6	172	80	26	106
7	174	82	27	109

FIGURE 7: Evolution of the error RMARE^{best} with the number of iterations made, considering different values of k with (W) or without (W/O) the proposed model.

TABLE 5: Number of routes that appear with the application of the proposed algorithm according to the F_{thres} value.

Threshold flow, F_{thres}	Number of routes in set $\mathcal{R}$	Number of routes in set $\mathcal{Q}$	Added routes with $\mathcal{SL}$	Number of routes in $\mathcal{C}$
0	162	162	9	171
10	162	152	9	161
20	162	114	13	127
30	162	80	18	98
40 and 50	162	72	20	92
60	162	63	20	93

fixed, i.e., assuming that $\mathcal{C} = \mathcal{Q}$ in any case. The results, presented in dashed lines in Figure 7, show that the proposed model clearly outperforms the estimations given by other methods. This difference is more evident for lower values of k because difference in the number of routes in $\mathcal{Q}$ and in $\mathcal{C}$ is bigger.

In addition, higher improvements in terms of lowering the RMARE occur in the first 100–200 iterations and remain almost constant (with lower steps) till the end of the experiment.

The effect of assigning different values of k is observed in the second column of Table 4. When the value of k increases, the number of routes in $\mathcal{R}$ also increases. The third column collects the number of routes that appear in $\mathcal{Q}$ after the reference set has been simplified by means of the simplification algorithm. For the best solution after applying the

algorithm, the fourth column shows the number of added routes compatible with the $\mathcal{OSCSL}$ set. The fifth column shows the total number of routes finally included in $\mathcal{C}$.

Indeed, the last column of Table 4 indicates that we are able to reinclude many routes once $\mathcal{OSCSL}$ is obtained from $\mathcal{SL}$.

In summary, a high value for k yields a low error due to the existence of more routes, and this implies that more information is available for the estimation problem; however, the difference in terms of the quality of the final solution is not very large. This means that, for large networks, lower values of k can be chosen to avoid problems related to computational costs since the algorithm would operate with a smaller number of routes in each iteration.

Finally, from the results shown in Figure 7, we confirm that, as other authors have noted, a value of $k = 3$ or 4 is a

TABLE 6: Sets of routes resulting after the application of Step 1 according to $F_{thres} = 30$.

OD pair	Path ID	Links	f_q
1-2	1	1 11 14 19 30	37.82
	2	2 35 14 19 30	22.68
	3	1 11 15 29 30	39.78
	4	1 11 14 18 20	106.60
1-3	5	1 11 14 19 31	57.75
	6	2 35 14 19 31	34.63
	7	1 12 26 37	189.25
	8	1 11 15 29 31	60.74
1-6	9	1 11	14.40
	10	2 35	8.64
1-8	11	1 11 14 18	12.40
	12	2 35 14 18	7.44
	13	2 36	243.47
2-1	14	3 22 34	190.42
	15	3 21 17 13 9	53.66
	16	4 33 27 16 34	17.49
	17	3 21 17 16 34	15.71
2-4	18	4 33 28 23	230.68
2-6	19	4 33 27	1.95
	20	3 21 17	1.75
	21	4 32 17	1.60
2-12	22	3 22	16.42
	23	4 33 27 16	1.51
	24	3 21 17 16	1.35
	25	4 32 17 16	1.24
3-1	26	5 32 17 13 9	131.62
	27	6 38 24 9	202.75
	28	5 33 27 16 34	46.99
	29	5 32 17 16 34	38.53
3-4	30	6 38 23	70.11
	31	5 33 28 23	28.61
3-6	32	5 32 17	0.61
	33	5 33 27	0.75
3-12	34	5 32 17 16	14.22
	35	5 33 27 16	17.33
4-2	36	8 25 29 30	209.82
4-3	37	8 26 37	75.31
	38	8 25 29 31	22.83
4-6	39	7 11	8.32
	40	8 25 27	4.56
4-8	41	8 25 29 32 18	82.13
	42	7 11 14 18	119.51
	43	8 25 27 14 18	65.40
6-2	44	14 18 20	2.24
	45	15 29 30	0.84
	46	14 19 30	0.79
6-3	47	14 19 31	0.34
	48	15 29 31	0.36
6-8	49	14 18	8.82
6-1	50	16 34	6.85
	51	13 9	17.40
6-4	52	13 10	8.71
	53	15 28 23	3.00
6-12	54	16	2.02

TABLE 6: Continued.

OD pair	Path ID	Links	f_q
8-1	55	22 34	221.73
	56	21 17 13 9	62.48
	57	21 17 16 34	18.29
8-4	58	21 19 33 28 23	68.35
	59	21 17 13 10	136.37
	60	21 17 15 28 23	46.93
8-6	61	21 17	11.14
8-12	62	22	20.67
	63	21 17 16	1.71
9-1	79	24 9	9.62
9-2	64	25 29 30	7.32
9-3	65	26 37	3.47
	66	25 29 31	1.05
9-4	80	23	5.88
9-6	67	25 27	0.35
9-8	68	25 29 32 18	4.63
	69	25 27 14 18	3.68
12-2	70	35 15 29 30	0.72
	71	35 14 19 30	0.68
	72	35 14 18 20	1.92
	73	36 20	62.88
12-3	74	35 15 29 31	57.03
	75	35 14 19 31	54.23
12-6	76	35	7.38
12-8	77	35 14 18	2.50
	78	36	81.76

reasonable choice since, in terms of error in the estimation of flows, it offers similar results to models with a greater number of routes per OD pair, such as $k = 6$ or 7.

5.1.2. Influence of Threshold Flow (F_{thres}).

The degree of network simplification implemented in Step 1 of the proposed algorithm depends on the demand and cost threshold values. The definition of these values will determine the total number of routes in the set $\mathcal{Q}$, obtained from the simplification of the reference set $\mathfrak{R}$.

In the case in which the reference route set is not simplified, i.e., with a null threshold flow, the algorithm operates with no simplified network and, therefore, with the same number of routes as that in the reference set. When the threshold flow increases, however, the network will be simplified according to the defined flow value. For the network model studied in this paper, different values of threshold flow have been considered. Following the conclusions drawn in the previous section, this analysis was carried out using $k = 4$ (i.e., the number of routes in $\mathfrak{R}$ is 162) and $B = 16$ again.

The effects of the values of F_{thres} considered in this paper on the size of the route set are reflected in the third column of Table 5. If higher values of F_{thres} are considered, the number of routes with respect to the reference is reduced; hence, the number of routes in set $\mathcal{Q}$ is smaller, and once the

TABLE 7: OD matrix resulting after the application of Step 1 according to $F_{thres} = 30$.

Node	1	2	3	4	6	8	12	Gen. trips
1	0.00	206.88	342.37	0.00	23.04	263.31	0.00	835.60
2	277.28	0.00	0.00	230.68	5.30	0.00	20.52	533.78
3	419.89	0.00	0.00	98.72	1.36	0.00	31.55	551.52
4	0.00	209.82	98.14	0.00	12.88	267.04	0.00	587.88
6	24.25	3.87	0.70	11.71	0.00	8.82	2.02	51.37
8	302.50	0.00	0.00	251.65	11.14	0.00	22.38	587.67
9	9.62	7.32	4.52	5.88	0.35	8.31	0.00	36.00
12	0.00	66.20	111.26	0.00	7.38	84.26	0.00	269.10
Attr. trips	1033.54	494.09	556.99	598.64	61.45	631.74	76.47	

cameras are installed on the network, a bigger number of new routes in $\mathscr{C}$ will appear (from 9 added routes with $F_{thres} = 0$ to 20 with $F_{thres} = 60$). Having a lower number of routes in set $\mathscr{Q}$ has important advantages in terms of computational cost solving location problems (10)–(15). Note also that due to the OD matrix configuration (see Figure 5), the simplification process leads to the same final network for $F_{thres} = 40$ than 50. For illustration purposes, Tables 6 and 7 show the simplified set of routes and the simplified OD matrix for $F_{thres} = 30$.

Figure 8 shows RMARE^{best} obtained for different cases shown in Table 5. The smallest quantified error in link flows corresponds to a case without simplification, as expected. The value of the error for the remaining scenarios increases with the value of the threshold flow used in the simplification step, and again, higher improvements in terms of error occur in the first iterations of the algorithm.

It is important to point out that, despite all the facts exposed above, the results of all the cases are similar in terms of error. This is because the simplification algorithm always keeps the routes with higher reference flow which are used for problems (10)–(15) to locate devices. Therefore, this leads to lower estimation errors and hence a better performance.

For the sake of comparison with existing models, Figure 8 also shows the evolution of the RMARE^{best}, keeping the set of routes in the estimation process as constant (i.e., setting $\mathscr{C} = \mathscr{Q}$ for all the cases). In this case also, the proposed model outperforms the estimates given by the existing models.

5.1.3. Influence of the Available Budget B.
In terms of traffic flow estimation, the number of installed devices on a network may be the most important factor to consider, even more if one evaluates the unusual possibility of obtaining full observability of the network.

Several authors have presented two versions of the scan device location problem: the full flow-observability problem and the partial flow-observability problem. In the first version, given a set of scanned links $\mathscr{SL}$, the equation system is fully observable, and the coefficient matrix has full rank, and it is not necessary to estimate the flows of interest. In the second version, given a set of scanned links s, the equation system is not observable for all flow variables (Gentili and Mirchandani [34]), either due to limitations of the number of devices to be installed and hence not to have a
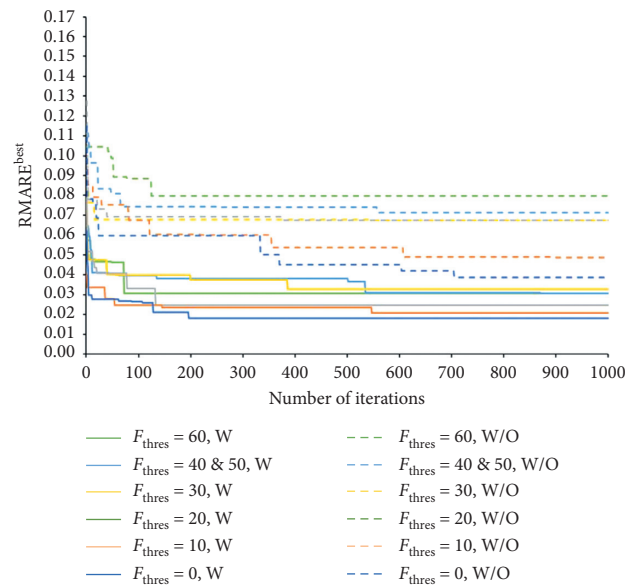

FIGURE 8: Evolution of the error RMARE^{best} with the number of iterations made, considering different values of F_{thres} with (W) or without (W/O) the proposed model.

full-rank coefficient matrix. In any case, the effects of partial knowledge of the routes analyzed in this paper have received very limited attention in the literature.

For the analysis, several values of B have been introduced into the location model so that the influence of the number of devices to be installed can be analyzed in terms of the quantification of the error in the flow estimation. In this case, we have developed the analysis assuming $k = 4$ and $F_{thres} = 40$.

Figure 9 compares the evolution of RMARE^{best} for different values of B. As expected, the cases with a higher budget value B show less error than the other cases, i.e., the error increases, while the value of the budget decreases. Here, a comparison with the results of the existing methods is also provided to show the improvements achieved using the proposed model.

Table 8 shows the number of added routes to set $\mathscr{Q}$ depending on the available budget for constant simplification F_{thres} and k. It is interesting that, for higher number of installed devices, the number of added routes is also higher. This means that the transportation planner can perform severe simplification of the network (higher F_{thres}) if the
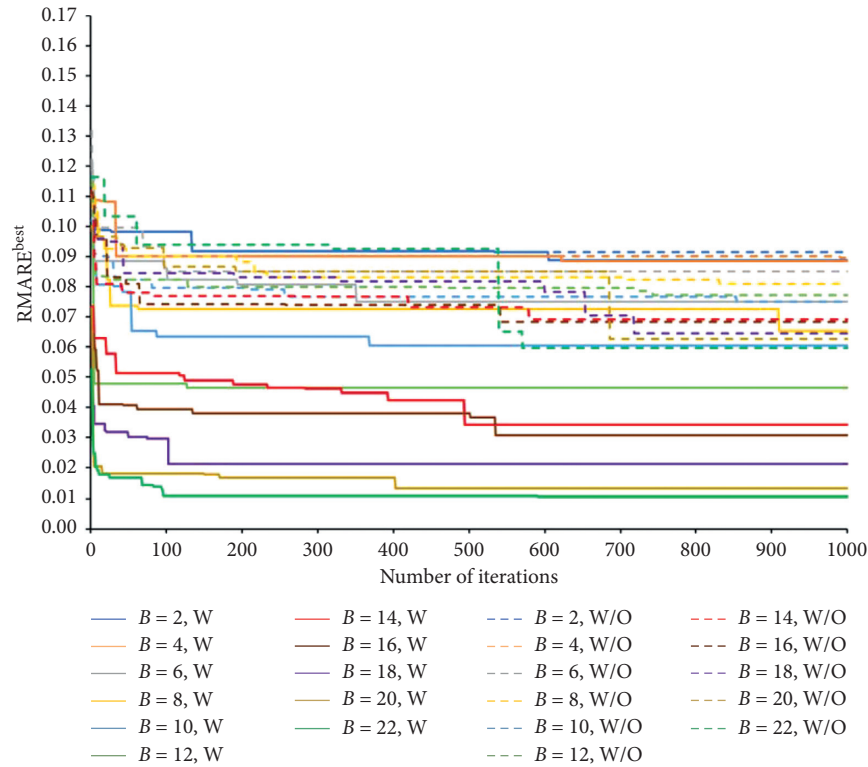
FIGURE 9: Evolution of RMARE$^{\text{best}}$ with the number of iterations considering different values of B with (W) or without (W/O) the proposed model.

TABLE 8: Number of routes that appear with the application of the proposed algorithm according to the B value.

B	Number of routes in set $\mathfrak{R}$	Number of routes in set $\mathcal{Q}$	Added routes with $\mathcal{SL}$	Number of routes in $\mathcal{C}$
2	162	72	—	72
4	162	72	2	73
6	162	72	3	75
8	162	72	3	75
10	162	72	5	77
12	162	72	6	78
14	162	72	18	90
16	162	72	20	92
18	162	72	37	109
20	162	72	70	142
22	162	72	75	147

number of devices to install is higher. This leads to a better performance of the algorithm while the expected estimation results will also be satisfactory.

5.2. The Ciudad Real Network. In this section, we illustrate the application of the proposed method to a real-size network. In particular, we have adapted the Ciudad Real network used in Castillo et al. [1] and in Owais et al. [45]. This network has 218 links and 105 nodes and has been divided into a total of 20 traffic zones, originating a matrix of 380 OD pairs (see Figure 10). After applying the method exposed in

Section 3, the extended OD matrix is composed by 10,374 node-based OD pairs. After that, the k-shortest path was carried out, assuming $k = 3$, discarding routes with more than 1.1 times the shortest path, resulting in a total of 18,630 routes in set $\mathfrak{R}$. Then, the proposed algorithm has been applied assuming $B = 50$ and $F_{\text{thres}} = 100, 150, 200,$ and 250, and the number of resulting routes in $\mathcal{Q}$ is shown in Table 9. Note that the number of resulting routes in $\mathcal{C}$ after applying the algorithm is on the same order of magnitude which again leads to good results in all cases. Figure 11 shows the evolution of RMARE$^{\text{best}}$ showing the same trend as in the Nguyen-Dupuis network.
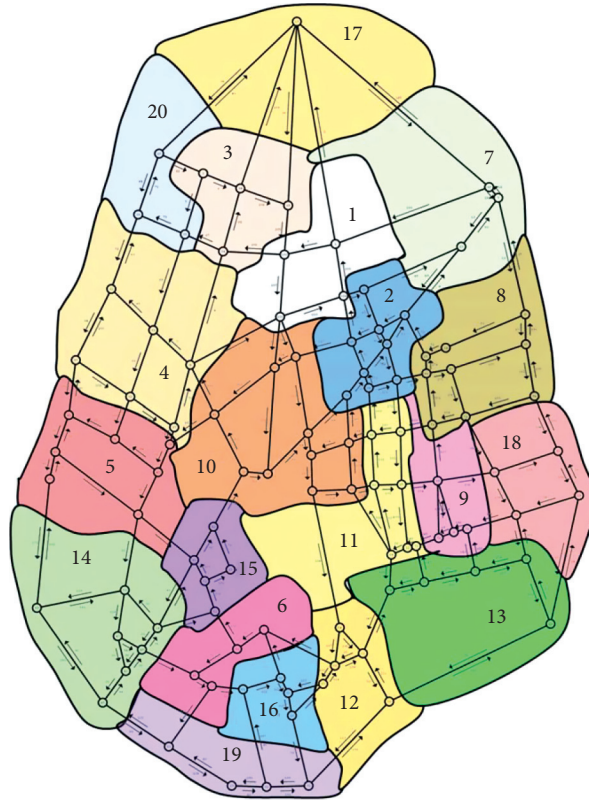
FIGURE 10: The Ciudad Real network (adapted from Castillo et al. [1]).

TABLE 9: Sets of routes that appear with the application of the proposed algorithm for the Ciudad Real network case.

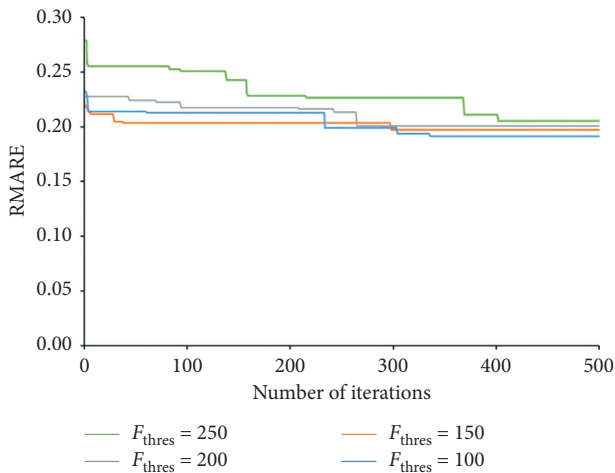| Threshold flow, F_{thres}	Number of routes in set $\mathfrak{R}$	Number of routes in set $\mathcal{Q}$	Added routes with $\mathcal{SL}$	Number of routes in $\mathcal{C}$
100	18,630	2987	6151	9138
150	18,630	1389	9122	10,511
200	18,630	361	11,417	11,778
250	18,630	138	11,687	11,825

FIGURE 11: Evolution of RMARE$^{\text{best}}$ for the Ciudad Real case.

6. Conclusions

This paper proposes a two-step methodology that may have important advantages from a practical point of view compared with the existing methodologies that estimate traffic flows using plate scanning data.

In the first step, a new methodology for traffic network modeling that does not use centroids and connectors is presented. The problems derived from the use of these tools are well known and have been analyzed in several studies and projects. Instead, the proposed methodology is an important step forward from a practical point of view since it defines a network with more detail so that the displacements between nodes (i.e., the network routes) have a better definition and also do not lead to artificial congestion in links. Both considerations make this method compatible with the data obtained from plate scanning as a particular case of AVI readers.

In the second step, a heuristic algorithm is proposed to face the uncertain or partial knowledge of routes which is essential for the correct application of the plate scanning technique or other AVI-based methods. For this, an iterative process has been formulated to obtain the device location to expect the best possible result in terms of link flow estimation. The proposed algorithm improves the expected flow estimation quality for the same value of the objective function used in other papers found in the literature. In addition, this reduction is achieved in the first 300–500 iterations of the algorithm. One of the reasons for this improvement is the incorporation of new routes to the estimation model once the field data have been collected. This fills a gap of the existing methods since they did not mention what to do with those scanned vehicles whose scanned patterns do not match with the modeled routes.

Finally, to evaluate the influence of different parameters of the algorithm on the final solution, we have performed a sensibility analysis using the well-known Nguyen-Dupuis network. In particular,

(i) We have performed an analysis by varying the parameter k in the process of the enumeration of the k-shortest path between the nodes of the designed network. It has been observed that a high value of k allows for better estimates of flows in terms of a smaller RMARE. However, a high value for k would entail working with a network with a high number of routes, which would have some computational cost mainly in the location optimization problem. Small values of this parameter would avoid this problem, and we have proved that the quality of the solution is similar to the solutions obtained with a higher number of routes. This is because of the incorporation of routes to the estimation model once the field data are collected. The sensitivity analysis performed confirms that, for the plate scanning technique, a value of k between 3 and 4 is an acceptable value (endorsed by several authors) for this type of model.

(ii) The analysis carried out to test the implications of the simplification of the network has been done by means of the elimination of those OD-pairs whose demand is lower than a defined threshold flow value F_{thres}. Again, the cases studied with different values of F_{thres} have obtained almost the same solution in terms of RMARE. This means that a medium-low degree of simplification leads to a good network model in terms of the final estimation process as well as the performance of the algorithm.

(iii) It has been verified that the budget B or number of AVI devices to be installed on the network has a great influence on the estimation results. For all the cases studied, the number of devices has a substantial effect on the number of new routes and sets $\mathcal{OSCSL}$ and therefore on the quality of the estimation errors. As expected, increasing the number of installed scanning devices on the network will yield better observability and better estimation of flows, and hence more information about routes (both included in set $\mathcal{Q}$ or not), which does not occur when working with a limited number of devices.

Finally, the methodology has also been applied to a real-size network. Despite, it has been observed that higher improvements on the solutions occur during the first 300–500 iterations, leading to few improvements when the algorithm has carried out a high number of iterations. It is of interest to find a method that allows to obtain the best solution for the scanned link set $\mathcal{SL}$ in earlier instances or iterations to avoid the computation costs arising from iterations that may be unnecessary. This observation deserves to be investigated and worked out in greater detail in future research, for which the use of advanced tools for heuristic optimization or machine learning is proposed to increase the efficiency of finding solutions in a shorter operating time.

Data Availability

All the data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] E. Castillo, J. M. Menéndez, and S. Sánchez-Cambronero, "Traffic estimation and optimal counting location without path enumeration using Bayesian networks," *Computer-Aided Civil and Infrastructure Engineering*, vol. 23, no. 3, pp. 189–207, 2008.

[2] E. Castillo, J. M. Menéndez, S. Sánchez-Cambronero, A. Calviño, and J. M. Sarabia, "A hierarchical optimization problem: estimating traffic flow using gamma random variables in a Bayesian context," *Computers & Operations Research*, vol. 41, pp. 240–251, 2014.

[3] K. Perrakis, D. Karlis, M. Cools, D. Janssens, K. Vanhoof, and G. Wets, "A Bayesian approach for modeling origin-destination matrices," *Transportation Research Part A: Policy and Practice*, vol. 46, no. 1, pp. 200–212, 2012.

[4] Z. Huang, X. Ling, P. Wang et al., "Modeling real-time human mobility based on mobile phone and transportation data fusion," *Transportation Research Part C: Emerging Technologies*, vol. 96, pp. 254–269, 2018.

[5] S. Ibarra-Espinosa, R. Ynoue, M. Giannotti, K. Ropkins, and E. D. de Freitas, "Generating traffic flow and speed regional

model data using internet GPS vehicle records," *MethodsX*, vol. 6, pp. 2065–2075, 2019.

[6] L. Moreira-Matias, J. Gama, M. Ferreira, J. Mendes-Moreira, and L. Damas, "Time-evolving O-D matrix estimation using high-speed GPS data streams," *Expert Systems with Applications*, vol. 44, pp. 275–288, 2016.

[7] J. L. Toole, S. Colak, B. Sturt, L. P. Alexander, A. Evsukoff, and M. C. González, "The path most traveled: travel demand estimation using big data resources," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 162–177, 2015.

[8] T. A. Zin, K. Kyaing, K. K. Lwin, and Y. Sekimoto, "Estimation of originating-destination trips in Yangon by using big data source," *Journal of Disaster Research*, vol. 13, no. 1, pp. 6–13, 2018.

[9] E. Castillo, I. Gallego, J. M. Menéndez, and A. Rivas, "Optimal use of plate-scanning resources for route flow estimation in traffic networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 2, pp. 380–391, 2010.

[10] C. Fu, N. Zhu, and S. Ma, "A stochastic program approach for path reconstruction oriented sensor location model," *Transportation Research Part B: Methodological*, vol. 102, pp. 210–237, 2017.

[11] A. Federov, K. Nikolskaia, S. Ivanov, V. Shepelev, and A. Minbaleev, "Traffic flow estimation with data from a video surveillance camera," *Journal of Big Data*, vol. 6, no. 1, p. 73, 2019.

[12] X. Yang, Y. Zou, J. Tang, J. Liang, and M. Ljaz, "Evaluation of short-term freeway speed prediction based on periodic analysis using statistical models and machine learning models," *Journal of Advanced Transportation*, vol. 2020, Article ID 9628957, 16 pages, 2020.

[13] S. Sánchez-Cambronero, E. Castillo, J. M. Menéndez, and P. Jiménez, "Dealing with error recovery in traffic flow prediction using Bayesian networks based on license plate scanning data," *Journal of Transportation Engineering*, vol. 137, no. 9, pp. 615–629, 2010.

[14] J. Bai and Y. Chen, "A deep neural network based on classification of traffic volume for short-term forecasting," *Mathematical Problems in Engineering*, vol. 2019, Article ID 6318094, 10 pages, 2019.

[15] Z. Liu, Z. Li, K. Wu, and M. Li, "Urban traffic prediction from mobility data using deep learning," *IEEE Network*, vol. 32, no. 4, pp. 40–46, 2018.

[16] M. Nigro, A. Abdelfatah, E. Cipriani, C. Colombaroni, G. Fusco, and A. Gemma, "Dynamic O-D demand estimation: application of SPSA AD-PI method in conjunction with different assignment strategies," *Journal of Advanced Transportation*, vol. 2018, Article ID 2085625, 18 pages, 2018.

[17] Y. Sheffi, *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, Prentice Hall, Engelwood Cliffs, NJ, USA, 1985.

[18] H. Spies, "Conical volume-delay functions," *Transportation Science*, vol. 24, no. 2, pp. 153–158, 1990.

[19] L. Huntsinger and N. Rouphail, "Bottleneck and queuing analysis: calibrating volume-delay functions of travel demand models," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2255, no. 1, pp. 117–124, 2011.

[20] E. T. Mtoi and R. Moses, "Calibration and evaluation of link congestion functions: applying intrinsic sensitivity of link speed as a practical consideration to heterogeneous facility types within urban network," *Journal of Transportation Technologies*, vol. 4, no. 2, pp. 141–149, 2014.

[21] W. W. Mann, "B-node model: new subarea traffic assignment model & application," in *Proceedings of the Eighth TRB Conference on the Application of Transportation Planning Methods Transportation Research Board*, Corpus Christi, Washington, DC, USA, July 2002.

[22] M. Friedrich and M. Galster, "Methods for generating connectors in transport planning models," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2132, no. 1, pp. 133–142, 2009.

[23] Z. S. Quian and H. M. Zhang, "On centroid connectors in static traffic assignment: their effects on flow patterns and how to optimize their selections," *Transportation Research Part B: Methodological*, vol. 46, no. 10, pp. 1489–1503, 2012.

[24] E. Jafari, M. D. Gemar, N. R. Juri, and J. Duthie, "Investigation of centroid connector placement for advanced traffic assignment models with added network detail," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2498, no. 1, pp. 19–26, 2015.

[25] S. Hao and L. Yang, "Raffic network modeling and extended max-pressure traffic control strategy based on granular computing theory," *Mathematical Problems in Engineering*, vol. 2019, Article ID 2752763, 11 pages, 2019.

[26] E. Castillo, J. M. Menéndez, and P. Jiménez, "Trip matrix and path flow reconstruction and estimation based on plate scanning and link observations," *Transportation Research Part B: Methodological*, vol. 42, no. 5, pp. 455–481, 2008.

[27] R. Mínguez, S. Sánchez-Cambronero, E. Castillo, and P. Jiménez, "Optimal traffic plate scanning location for OD trip matrix and route estimation in road networks," *Transportation Research Part B: Methodological*, vol. 44, no. 2, pp. 282–298, 2010.

[28] X. Zhou and H. S. Mahmassani, "Dynamic origin-destination demand estimation using automatic vehicle identification data," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 1, pp. 105–114, 2006.

[29] J. Liu, F. Zheng, H. J. Van Zuylen, and J. Li, "A Dynamic OD prediction approach for urban networks based on automatic number plate recognition data," *Transportation Research Procedia*, vol. 47, pp. 601–608, 2020.

[30] J. Li, H. v. Zuylen, Y. Deng, and Y. Zhou, "Urban travel time data cleaning and analysis for automatic number plate recognition," *Transportation Research Procedia*, vol. 47, pp. 712–719, 2020.

[31] S. Sánchez-Cambronero, A. Rivas, R. M. Barba, L. Ruiz-Ripoll, I. Gallego, and J. M. Menéndez, "A methodology to model a traffic network which have their field data obtained by plate scanning technique," *Transportation Research Procedia*, vol. 18, pp. 341–348, 2016.

[32] H. Yang and J. Zhou, "Optimal traffic counting locations for origin–destination matrix estimation," *Transportation Research Part B: Methodological*, vol. 32, no. 2, pp. 209–126, 1998.

[33] A. Ehlert, M. G. H. Bell, and S. Grosso, "The optimisation of traffic count locations in road networks," *Transportation Research Part B: Methodological*, vol. 40, no. 6, pp. 460–479, 2006.

[34] M. Gentili and P. B. Mirchandani, "Locating sensors on traffic networks: models, challenges and research opportunities," *Transportation Research Part C: Emerging Technologies*, vol. 24, pp. 227–255, 2012.

[35] M. Salari, L. Kattan, W. H. K. Lam, H. P. Lo, and M. A. Esfeh, "Optimization of traffic sensor location for complete link flow observability in traffic network considering sensor failure," *Transportation Research Part B: Methodological*, vol. 121, pp. 216–251, 2019.

[36] E. Castillo, M. Nogal, A. Rivas, and S. Sánchez-Cambronero, "Observability of traffic networks. Optimal location of

counting and scanning devices," *Transportmetrica B: Transport Dynamics*, vol. 1, no. 1, pp. 68–102, 2013.

[37] J. Yang and J. Sun, "Vehicle path reconstruction using automatic vehicle identification data: an integrated particle filter and path flow estimator," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 107–126, 2015.

[38] S. Sánchez-Cambronero, P. Jiménez, A. Rivas, and I. Gallego, "Plate scanning tools to obtain travel times in traffic networks," *Journal of Intelligent Transportation Systems*, vol. 21, no. 5, pp. 390–408, 2017.

[39] M. Gentili and P. B. Mirchandani, "Review of optimal sensor location models for travel time estimation," *Transportation Research Part C: Emerging Technologies*, vol. 90, pp. 74–96, 2018.

[40] C. Cerrone, R. Cerulli, and M. Gentili, "Vehicle-id sensor location for route flow recognition: models and algorithms," *European Journal of Operational Research*, vol. 247, no. 2, pp. 618–629, 2015.

[41] N. Levy and I. Benenson, "GIS-based method for assessing city parking patterns," *Journal of Transport Geography*, vol. 46, pp. 220–231, 2015.

[42] S. Sánchez-Cambronero, E. Castillo, and J. Menéndez, *Traffic Prediction Models Using Bayesian Networks and Other Tools*, LAP Lambert Academic Publishing, Riga, Latvia, 2011.

[43] E. Castillo, Z. Grande, A. Calviño, W. Szeto, and H. Lo, "A state-of-the-art review of the sensor location, flow observability, estimation, and prediction problems in traffic networks," *Journal of Sensors*, vol. 2015, Article ID 903563, 26 pages, 2015.

[44] J. Yen, "Finding the *K* shortest loopless paths in a network," *Management Science*, vol. 17, no. 11, pp. 661–786, 1971.

[45] M. Owais, G. Moussa, and K. Hussain, "Sensor location model for O/D estimation: multi-criteria meta-heuristics," *Operations Research Perspectives*, vol. 6, pp. 1–12, 2019.

[46] P. Bonsall, P. Firmin, M. Anderson, I. Palmer, and P. Balmforth, "Validating the results of a route choice simulator," *Transportation Research Part C: Emerging Technologies*, vol. 5, no. 6, pp. 371–387, 1997.

[47] M. L. Hazelton, "Estimation of origin-destination matrices from link flows on uncongested networks," *Transportation Research Part B: Methodological*, vol. 34, no. 7, pp. 549–566, 2000.