

## Research Article

# Traffic Speed Forecast in Adjacent Region between Highway and Urban Expressway: Based on MFD and GRU Model

Yuan Gao <sup>1</sup>, Jiandong Zhao <sup>2,3</sup>, Ziyang Qin,<sup>2</sup> Yingzi Feng <sup>2</sup>, Zhenzhen Yang <sup>2</sup>,  
and Bin Jia <sup>2,3</sup>

<sup>1</sup>School of Mechanical and Electronic Control Engineering, Beijing Jiaotong University, Beijing 100044, China

<sup>2</sup>School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

<sup>3</sup>Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Ministry of Transport, Beijing Jiaotong University, Beijing, 100044, China

Correspondence should be addressed to Jiandong Zhao; zhaojd@bjtu.edu.cn

Received 26 September 2020; Revised 15 November 2020; Accepted 20 November 2020; Published 3 December 2020

Academic Editor: Xiaoming Chen

Copyright © 2020 Yuan Gao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Traffic congestion in the adjacent region between the highway and urban expressway is becoming more and more serious. This paper proposes a traffic speed forecast method based on the Macroscopic Fundamental Diagram (MFD) and Gated Recurrent Unit (GRU) model to provide the necessary traffic guidance information for travelers in this region. Firstly, considering that the road traffic speed is affected by the macroscopic traffic state, the adjacent region between the highway and expressway is divided into subareas based on the MFD. Secondly, the spatial-temporal correlation coefficient is proposed to measure the correlation between subareas. Then, the matrix of regional traffic speed data is constructed. Thirdly, the matrix is input into the GRU prediction model to get the predicted traffic speed. The proposed algorithm's prediction performance is verified based on the GPS data collected from the adjacent region between Beijing Highways and Expressway.

## 1. Introduction

With the continuous growth of the scale of China's highway network and traffic volume, the traffic load of the intercity highway in some developed cities is increasing. Many adjacent regions (network of in- and out-of-town roads) between highway and expressway have become part of the urban commuting road. Besides, due to the restrictions of traffic management measures, trucks are not allowed to enter the urban area in fixed hours. They can only drive on adjacent regions between the highway and urban expressway, which caused the trucks to accumulate, resulting in low road capacity and service level during peak hours, severe traffic congestion, and frequent traffic accidents. What is more, the abnormal weather also often leads to local traffic congestion, which gradually evolves into spread congestion in the regional network. To sum up, it is essential to conduct research on traffic states predicting at the intersection of highways and urban

expressways and publish accurate traffic guidance information to travelers to alleviate traffic congestion.

The evolution of road traffic flow has complex nonlinear characteristics [1], which makes it challenging to realize accurate traffic flow predictions. Many machine learning algorithms have been used to research transportation [2], especially traffic flow prediction. Liu et al. proposed a hybrid road network traffic speed prediction model based on the state-space neural network and extended Kalman filter [3]. Zhang et al. predicted the traffic speed considering the heterogeneity of different roads [4]. Dhivyabharathi et al. proposed a method for predicting river traffic time using the particle filtering method [5]. Zhao et al. integrated the charging data and microwave detection data to predict traffic speed [6]. Zhao et al. proposed a prediction algorithm combining equal spacing interpolation and Sage-Husa adaptive Kalman filtering [7]. Wang et al. improved the reliability method of driving time prediction based on GPS point velocity distribution by calculating the

variable velocity distribution coefficient [8]. Zhou et al. proposed a recurrent neural network based microscopic car-following model on predicting traffic oscillation [9]. Wang and Goodchild developed a logit model to determine the truck route's influencing factors and estimate the driving time [10]. Jula et al. developed a hybrid method composed of dynamic programming and genetic algorithm to find trucks' shortest path [11]. Dong et al. proposed a traffic crash prediction method based on the support vector regression (SVR) model [12]. Yu et al. proposed the random forests based on the near neighbor (RFNN) method to predict bus travel time [13]. Xie and Wei proposed Elman neural network to predict truck speed [14]. Wang and Xu constructed the short-term traffic flow prediction model of urban expressway based on the Long Short-Term Memory (LSTM) network under deep learning [15]. Luo et al. proposed a short-term traffic flow prediction model based on deep learning combining the features of convolutional neural network and support vector regression classifier [16]. Yao et al. discussed the application of Support Vector Machine theory to predict road travel time [17]. Wu et al. [18] proposed a traffic flow prediction model based on Deep Neural Networks (DNN) by utilizing the weekly/daily periodicity and space-time characteristics of traffic flow [18]. Jia et al. proposed a deep learning method for short-term traffic speed information prediction-deep belief network (DBN) model [19]. Due to the problem that the commonly used weight optimization algorithm could not adjust the learning rate adaptably, Zhao et al. adopted Adam, Adadelta, and Rmsprop to optimize the weight in the GRU model of the deep learning algorithm [20]. Wang et al. established a travel time prediction model based on the LSTM (Long Short-Term Memory) considering the precipitation data [21]. Zang et al. proposed an all-day traffic speed prediction method for elevated highways based on deep learning [22]. Peng et al. proposed a 3D Convolutional Neural Network-Deep Neural Network method to recognize and predict traffic status from aerial videos [23]. The author has also conducted many studies on how to improve prediction accuracy [24–26]. The existing machine learning algorithm cannot fully dig out the essence of traffic flow characteristics. Deep learning model, such as GRU, can help us learn and seize the inherent complex features effectively and predict traffic flow without prior knowledge [27]. Using the deep learning algorithm to mine traffic flow rules becomes the direction of traffic state prediction.

The prediction of road segment traffic speed belongs to the study of microscopic traffic flow characteristics. It is also affected by the traffic states at the macrolevel, such as road network and neighboring region. At present, it is rare to combine macro- and microcharacteristics to predict traffic speed. The Macroscopic Fundamental Diagram (MFD) is a model reflecting the road network's macroscopic traffic state. According to specific indicators, MFD divides a complex and large road network into several independent subregions and implements appropriate control optimization strategies according to the subareas' characteristics. Based on the subregion division results, several road sections that are most similar to the predicted target road can be selected. Meanwhile, the spatial and temporal correlation between traffic flows in each subregion can be analyzed. The

subregions with strong correlation can be selected to construct the traffic flow sequence dataset and input the prediction model, which is conducive to improving the prediction accuracy. Many scholars have studied the subregional division method of MFD. Ji and Geroliminis proposed a static normalized cut (Ncut) based subregions division method based on traffic congestion's spatial characteristics and minimizing the subarea's vehicle density [28]. Ji et al. also proposed a dynamic subarea delineation method based on GPS data targeting the maximum connected element [29]. Haddad and Geroliminis proposed a division method based on the operational stability of the subregions [30]. Ma et al. used a spectral approach to divide traffic zones based on neighboring intersections [31]. Ncut is a graph theory-based partitioning algorithm derived from the image partitioning domain. This algorithm not only considers the similarities within regions but also normalizes the similarities within regions using the similarities between regions. And then, the cutting scheme that minimizes the similarities between regions after normalization is found. In this paper, the Ncut algorithm is used to divide the traffic subregions. Then, the stability of the MFD for each subdivided subregion is calculated and analyzed to justify the division results.

GPS data has a wide coverage area and can better reflect the characteristics of urban road traffic flow [32]. What is more, in recent years, China has strengthened the supervision of freight vehicles. It is required that the GPS devices be installed on the large heavy-haul trucks to monitor the trucks' running status. This produces a large amount of trajectory data, especially in the adjacent region between highway and expressway. In this paper, based on the average roadway speed and flow data extracted from the truck GPS data, a short-time traffic flow prediction method combining MFD and GRU is proposed. Using the characteristics of MFD, the road network area is divided into subregions, and the microtraffic flow characteristics and macrotraffic conditions are combined to develop a traffic forecasting method. The test results of real traffic flow data show that the method proposed in this paper has lower prediction errors and higher accuracy than the existing prediction models. It is a reasonable and effective method to predict short-time traffic flow. The technical framework of this paper is shown in Figure 1.

## 2. Subdivision Method of Road Network Based on MFD

*2.1. Construction of Road Network Weighted Graph Based on Traffic Operation Similarity.* A stable MFD exists in a network of roads with operational homogeneity. A large area can be divided into subregions based on the operational homogeneity of traffic. The starting point of the MFD theory is to study the relationship between traffic demand and traffic supply in the road network, the maximum traffic volume can directly reflect the traffic supply of each road section and the overall road network, and the traffic volume data can be easily obtained through traffic flow detection. Thus, the traffic volume is taken as the fundamental traffic characteristic of the road section in this paper. The road

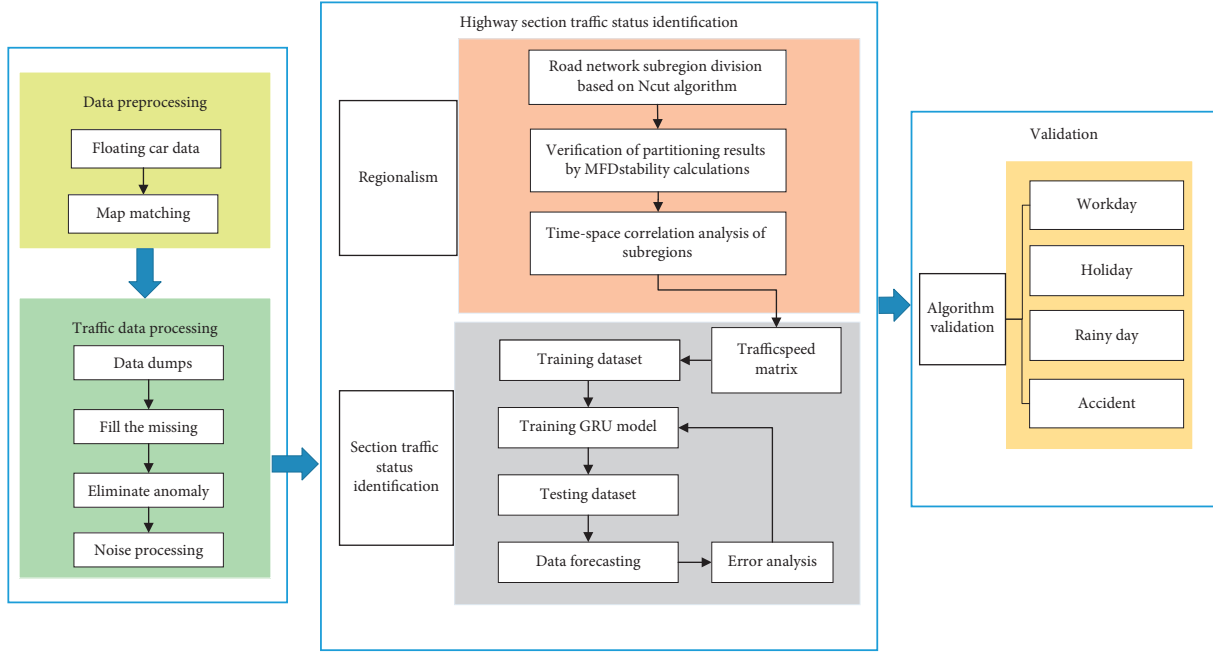


FIGURE 1: The technical framework of the paper.

section's maximum traffic volume is used to define the traffic operation similarity between adjacent connected sections.

Let the similarity degree of traffic operation between adjacent connected sections  $i$  and  $j$  in road network  $G$  be:

$$w(i, j) = \exp\left[-(q_i^m - q_j^m)^2\right], \quad (1)$$

where  $q_i^m$  is the maximum traffic volume of section  $i$  in road network  $G$  and  $q_j^m$  is the maximum traffic volume of section  $j$  in road network  $G$ . Using the natural constant transformation, the difference value of the maximum traffic volume between adjacent connected sections is squared mapped to the interval of 0-1. If the similarity is 0, the traffic operation similarity between the road sections is the least. If the similarity is 1, the traffic operation similarity between the road sections is the biggest.

Based on graph theory, the road network is first "node-arc transformation," so that the similarity degree of traffic operation between adjacent road sections is expressed as the weight of the arc section in the graph. The Laplace matrix is constructed based on the similarity degree of traffic operation, which is the basis for the subgraph division in graph theory. The details are as follows.

When road segment  $i$  in road network  $G$  is connected with road segment  $j$ ,  $a_{ij} = 1$ . When segment  $i$  is disconnected from segment  $j$ ,  $a_{ij} = 0$ . When  $i$  is equal to  $j$ ,  $a_{ij} = 0$ . The weighted adjacency matrix of road network  $G$  is  $W$ , and the element in  $W$  is

$$w_{ij} = \begin{cases} w'(i, j), & a_{ij} = 1, \\ 0, & a_{ij} = 0. \end{cases} \quad (2)$$

This paper adopts road sections as node  $V$  in the undirected graph  $G$ , so that  $w'(i, j) = -w(i, j)$ . The matrix  $D$  is the diagonal matrix:

$$D = \text{diag}\{d_i\}, \quad (3)$$

$$d_i = \sum_j w'(i, j).$$

$D-W$  is the Laplace matrix of road network  $G$ , in which the sum of all the rows and columns is zero. Based on the transformations and calculations, it is possible to obtain a road network weighted by traffic operations similarity between adjacent connecting sections.

**2.2. Road Network Subarea Division Method Based on Normalized Cut.** Ncut is one of the neutron graph partition methods in graph theory, which is a subgraph division method at the macrolevel. The focus is not on the graph's details, but on the overall characteristics of the graph. The optimal normalized cut problem of the graph can be expressed as

$$\begin{aligned} \min_x & \text{Ncut}(x) \\ \text{s.t.} & x \in V. \end{aligned} \quad (4)$$

It is an NP-hard problem to solve the minimum value of the normalized cut. The spectral clustering method is a widely used method that can solve NP-hard problem approximately by solving eigenvalue and eigenvalue vector. Therefore, the Fiedler method is used to calculate the

eigenvalues and eigenvectors of the matrix and divide the subareas of the road network. The point set  $V$  of figure  $G$  is divided into two subsets, and the transformation can express the optimal normalized cut problem of  $A$  and  $B$ :

$$\begin{aligned} \min_x \text{Ncut}(x) &= \min_y \frac{y^T(D-W)y}{y^T D y} \\ x_i &= 1, \quad \text{if } x_i \in A \\ x_i &= -1, \quad \text{if } x_i \in B \\ \text{s.t.} \quad y_i &\in \{-1, 1\} \\ y^T D 1 &= 0, \end{aligned} \quad (5)$$

where  $y = (1 + x) - b(1 - x)$ ,  $b = ((\sum_{x_i > 0} d_i) / (\sum_{x_i < 0} d_i))$ ,  $d_i = \sum_j w_{ij}$ , and  $x$  is a vector of columns consisting of 1 and -1. When the  $i$ -th node is subdivided into subregion  $A$ ,  $x = 1$ ; when it is subdivided into subregion  $B$ ,  $x = -1$ .

Since all rows and columns in a Laplace matrix have a sum of 0, the matrix always has an eigenvalue of 0. If graph  $G$  is connected, then the second small eigenvalue is positive. The corresponding eigenvector is called the Fiedler vector, which contains important information about the graph; that is, the numerical size of the elements in the Fiedler vector reflects the correlation of their corresponding vertices. When the road network is divided according to the Fiedler vector, the vertices corresponding to the Fiedler vector can be divided according to different critical value  $S$ . There are many methods to select the  $S$  value, among which the 0-point method is practical and straightforward.

### 3. Stability Calculation of MFD

By calculating the stability of MFD, the rationality of the subregion division can be proved. MFD stability depends on stability in the critical state. In the critical state, if the average traffic volume fluctuates less under the same road network density, the road network traffic operation will be more controllable. This article refers to the method (Fuzzy  $c$ -means algorithm) in our research [33] to divide the test data sets into three categories: unblocked, critical, and congested. Two indicators, the regional traffic volume and regional density, are used to classify the three traffic states. Firstly, FCM (Fuzzy  $c$ -means algorithm) is used to divide the data points of the spatial distribution of multidimensional data into specific classes. Each data point belongs to a certain class to some extent, and the membership degree is used to indicate the degree to which each data point belongs to a certain clustering. FCM divides  $n$  vectors into  $c$  fuzzy groups and calculates each group's clustering center to minimize the objective function of nonsimilarity indexes. Then, the traffic state is divided into three stages: unblocked, critical, and congested.

The dispersion of road network traffic operation, that is, the dispersion of weighted average traffic volume of road network in the critical state, represents MFD's stability in the critical state. The lower the dispersion of road network traffic operation is, the more stable the road network operation is,

and the higher the MFD stability is. The higher the dispersion of road network traffic operation, the more unstable the road network operation will be and the lower the MFD stability. The dispersion of road network traffic operation is

$$s = \frac{\sqrt{(1/n) \sum_{e=1}^n (q_e^w - q_c^w)^2}}{q_c^w}, \quad (6)$$

$$q_c^w = \alpha q_m^w,$$

where  $q_e^w$  is the average traffic volume in the test data.  $q_c^w$  is the critical average traffic volume of road network,  $q_m^w$  is the maximum average traffic volume of road network, and  $\alpha$  is undetermined parameters,  $0 < \alpha < 1$ .

The whole road network's dispersion degree is calculated by the weighted average of each subarea's dispersion degree. The calculation results can be used as the judgment index of the subarea division to characterize the whole network's MFD stability. If the entire road network is divided into  $N$  subareas, and the dispersion of road network traffic operation in subarea  $i$  is  $s_i$ , then the dispersion of the whole road network traffic operation is  $S_N$

$$S_N = \frac{\sum_{i=1}^N q_{ci}^w s_i}{\sum_{i=1}^N q_{ci}^w}, \quad (7)$$

where  $q_{ci}^w$  is the critical average traffic volume in subarea  $i$ .

## 4. Traffic Speed Prediction Method Based on the Spatial-Temporal Correlation of Subareas

The evolution of traffic speed on a road section in a specific subarea is affected not only by the temporal evolution law of the traffic flow on the road sections but also by the spatial influence of the road sections in other subareas. This paper proposes a traffic speed prediction algorithm that considers the spatial-temporal correlation of subareas.

### 4.1. Correlation Analysis of Subareas

**4.1.1. Spatial Correlation Analysis.** Firstly, the spatial correlation between each subarea is analyzed.

In spatial correlation analysis, it is necessary to measure the adjacency relationship of the neighboring subregions. This requires quantitatively describing the adjacency relationship of adjacent regions to perform the calculation of spatial correlation statistics.

In this paper, the spatial adjacency matrix is used to express the spatial relationship between subregions.

Suppose that there are  $m$  subregions in the study area, and the spatial weight matrix  $W_{sp} = [w_{ij}^{sp}]_{m \times m}$  is used to express the spatial correlation between the subregions:

$$w_{ij}^{sp} = \begin{cases} 1, & \text{when subregions } i \text{ and } j \text{ are adjacent,} \\ 0, & \text{when subregions } i \text{ and } j \text{ are not adjacent,} \end{cases} \quad (8)$$

where  $W_{sp}$  is  $m \times m$  dimensional spatial weight matrix and  $w_{ij}^{sp}$  is the spatial weight between the regional units  $i$  and  $j$ .

Besides, to ensure that the subregions cannot be adjacent to themselves specifies that when  $j = i$ ,  $w_{ij}^{\text{sp}} = 0$ . When two subregions are sharing one or more nodes, it is adjacent.

The standardized formula of the spatial weights is

$$\omega_{ij}^{\text{sp}} = \frac{w_{ij}^{\text{sp}}}{\sum_j w_{ij}^{\text{sp}}} \quad (9)$$

**4.1.2. Temporal Correlation Analysis.** The Pearson correlation coefficient formula is improved to measure the time correlation of the two regions. If two subregions in the study area have spatial adjacencies, the time correlation can be calculated by the following formula over a certain period:

$$\gamma_{ij} = \frac{E(q_i(t) - \mu_i)(q_j(t) - \mu_j)}{\sigma_i \sigma_j}, \quad (10)$$

where  $q_i(t)$  and  $q_j(t)$  are the traffic volume of subregions  $i$  and  $j$  at time  $t$ ;  $\mu_i$  and  $\mu_j$  are the mean traffic volume for  $i$  and  $j$ ; and  $\sigma_i$  and  $\sigma_j$  are the variances.

According to the road network area studied in this paper, the regional correlation is calculated as  $\omega_{ij}^{\text{sp}} \gamma_{ij}(\text{ac})$ , where  $\omega_{ij}$  is the spatial correlation of areas  $i$  and  $j$  and  $\gamma_{ij}(\text{ac})$  is the correlation of two areas in the study period.

In this paper, the data of the  $K$  most relevant regions to the region where the predicted target segment is located are selected for constructing the input matrix of the prediction model. In this paper, the  $K$  values are determined as follows:  $K$  ( $K=0, 1, \dots, 4$ ) are used to input the data from the most relevant regions into the GRU model, and the  $K$ -value with the smallest prediction error is taken.

One input sample of the deep learning algorithm is

$$\begin{bmatrix} V_{11}(t-1) & \cdots & V_{1(o-1)}(t-1) & V_{21}(t-1) & \cdots & V_{2o}(t-1) & \cdots & V_{ko}(t-1) & V_a(t-1) \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ V_{11}(t-L) & \cdots & V_{1(o-1)}(t-L) & V_{21}(t-L) & \cdots & V_{2o}(t-L) & \cdots & V_{ko}(t-L) & V_a(t-L) \end{bmatrix}, \quad (11)$$

where  $L$  is timestep,  $o$  is the number of sections of road included in Area  $k$ , and  $a$  represents the target section.

**4.2. GRU-Based Traffic Speed Prediction Algorithm.** RNN (recurrent neural network) is a kind of deep neural network designed to process sequence data, which plays an important role in the field of sequence mining.

The GRU model is an improvement of recurrent neural network, which is one of the hot technologies of deep learning in recent years. Different from the traditional recurrent neural network, the internal structure of the GRU's hidden layer nodes does not use a single activation function.

The specific calculation steps of GRU are as follows: firstly, the current state input  $z_t$  and the previous time output  $h_{t-1}$  are input into the update gate, and then a value between 0 and 1 can be output, where 0 represents the complete discarding information and 1 represents the complete reserving information, and the calculation formula is as shown in formula (12). Secondly,  $z_t$  and  $h_{t-1}$  entering the reset gate of the sigmoid layer output the value between 0 and 1. Meanwhile,  $\tanh$  layer will create a new candidate value vector  $\tilde{h}_t$ , and the calculation formulas are shown in equations (13) and (14). Thirdly, the update gate is used as

the weight vector, and the candidate vector and the output vector at the last moment are weighted averages to obtain the output  $h_t$  of GRU cells. The calculation formula is shown in equation (15):

$$\eta_t = \sigma(p_\eta + U_\eta z_t + Q_\eta h_{t-1}), \quad (12)$$

$$r_t = \sigma(p_r + U_r z_t + Q_r h_{t-1}), \quad (13)$$

$$\tilde{h}_t = \tanh(p_s + U_s z_t + Q_s r_t h_{t-1}), \quad (14)$$

$$h_t = (1 - \eta_t) h_{t-1} + \eta_t \tilde{h}_t, \quad (15)$$

where  $\eta$  represents the update gate vector;  $r$  represents the reset gate vector;  $p$  represents the bias vector;  $U$  represents the input weight;  $Q$  represents the cyclic weight;  $z_t$  represents the input vector at  $t$  time; and  $h_t$  represents the output vector at  $t$  time.

Regularization is generally defined as the modification of the learning algorithm, and the goal is to reduce generalization error rather than training error. Common regularization methods include  $L1$  and  $L2$  parameter paradigm penalty, Dropout, multitask learning, and early termination, which are common, where the penalty terms  $L2$  and  $L1$  of parameter normal form can be expressed as

$$L2(\theta) = \frac{1}{2} \sum_i |\theta_i^2|, L1(\theta) = \sum_i |\theta_i|, \quad (16)$$

where  $\theta_i$  can be expressed as the reciprocal of weight  $Q$  of each layer, indicating that, for the layer with too high weight learned, its updating degree should be reduced. On the contrary, for the node with too low weight learned in the layer, its updating degree should be improved to achieve the goal of amortizing the ownership value in the layer.

To sum up, the flow of the stroke speed prediction algorithm proposed in this paper is shown in Figure 2. The input is a three-dimensional vector composed of features, time step, and samples. This 3D tensor is input to the GRU model with a dropout layer and fully connected layer to get the travel time's predicted result. One column of the matrix in equation (11) corresponds to the input of one time step of the GRU model.

## 5. Algorithm Validation

**5.1. Road Network Subarea Division and Stability Calculation of MFD.** This paper uses truck GPS data as the basis for algorithm validation. As shown in Figure 3, the experimental area selected in this paper is located in Beijing's southeast. Beijing's expressway and the main road, including the 5<sup>th</sup> and 6<sup>th</sup> ring roads, Jingtai, Jinghu, and Jingha highway, are selected to verify algorithm accuracy. The area is approximately 110 square kilometers. The total length of roads in the road network is about 131 km. According to the analysis, the selected area has more accidents and more GPS data of truck.

The time range of the data used for validation is May 1, 2018, to July 31, 2018.

The methods of map matching, anomaly data processing, and traffic speed time series extraction of truck GPS data in this paper are from literature [20]. In this paper, the collected truck GPS data is organized into the form of a time series of the traffic speed of the road section with a period of 5 minutes. Then, according to the chosen  $K$ -value, it is organized into equation (11).  $L$  takes 12; that is, the prediction is made using the previous hour's data.

Sample size is a critical concern when using probe vehicles to collect real-time traffic information, and it is necessary to determine the number of probe vehicles needed for traffic state estimation. In this paper, the required sample size for different combinations of confidence levels of the study area is determined with reference to the method in [34].

May and June's average speed data are used as the training data set for GRU model training. The rest of the data serves as a test set for the algorithm. In this paper, the study

area's road network is abstracted into the road network diagram, as shown in Figure 4, and there are 32 road sections and 21 nodes.

The regional division of the cases of the parties is shown in Figure 5. The dispersion of traffic operation in subareas and the whole road network is shown in Table 1. When the network is divided into 5 subareas, the whole road network's traffic discrete degree is the smallest, 0.05673. The network has been divided into smaller regions. If continuing, the change of discrete degree of the whole road network is not too big. However, for speed prediction, the dimensionality of the data input to the model will increase. Thus, the prediction difficulty will increase. So, take five as the optimal scheme of areas division.

The MFD of each subarea is shown in Figure 6. The traffic state classification results based on the FCM algorithm for subarea 1 are shown in Figure 7. The clustering centers are shown in Table 2.

**5.2. Traffic Speed Prediction.** Two indexes, MAPE and RMSE, are selected to evaluate the prediction accuracy of the model. The calculation method of MAPE and RMSE is shown in the following formulas:

$$\text{MAPE} = \frac{1}{L} \sum_L \frac{|V_Y(t) - V(t)|}{V(t)}, \quad (17)$$

$$\text{RMSE} = \sqrt{\frac{1}{L} \sum_L (V_Y(t) - V(t))^2}, \quad (18)$$

where  $V_Y(t)$  is the predicted traffic speed at time  $t$ ,  $V(t)$  is the actual traffic speed at time  $t$ , and  $L$  is the total number of predicted cycles.

This paper selects nodes 17 to 18 (Section 1) and 9 to 10 (Section 2) belonging to different regions as the experimental verification sections. The accuracy of the algorithm was verified in four different scenarios. The prediction results were compared with the GRU prediction algorithm based on a single time series of the road segment. This GRU model has the same parameter settings as the model presented in this paper.

The first step is to determine the number of regions  $K$  that are input to the GRU model, so the relationship between the number of inputs and the prediction accuracy of the model is analyzed. Table 3 shows the prediction accuracy of Section 1 for different  $K$  values. Table 4 shows the prediction accuracy of Section 2 for different  $K$  values. From Table 3, it can be seen that the  $K$  of road Section 1 take 1. From Table 4, it can be seen that the  $K$  of road Section 2 take 1.

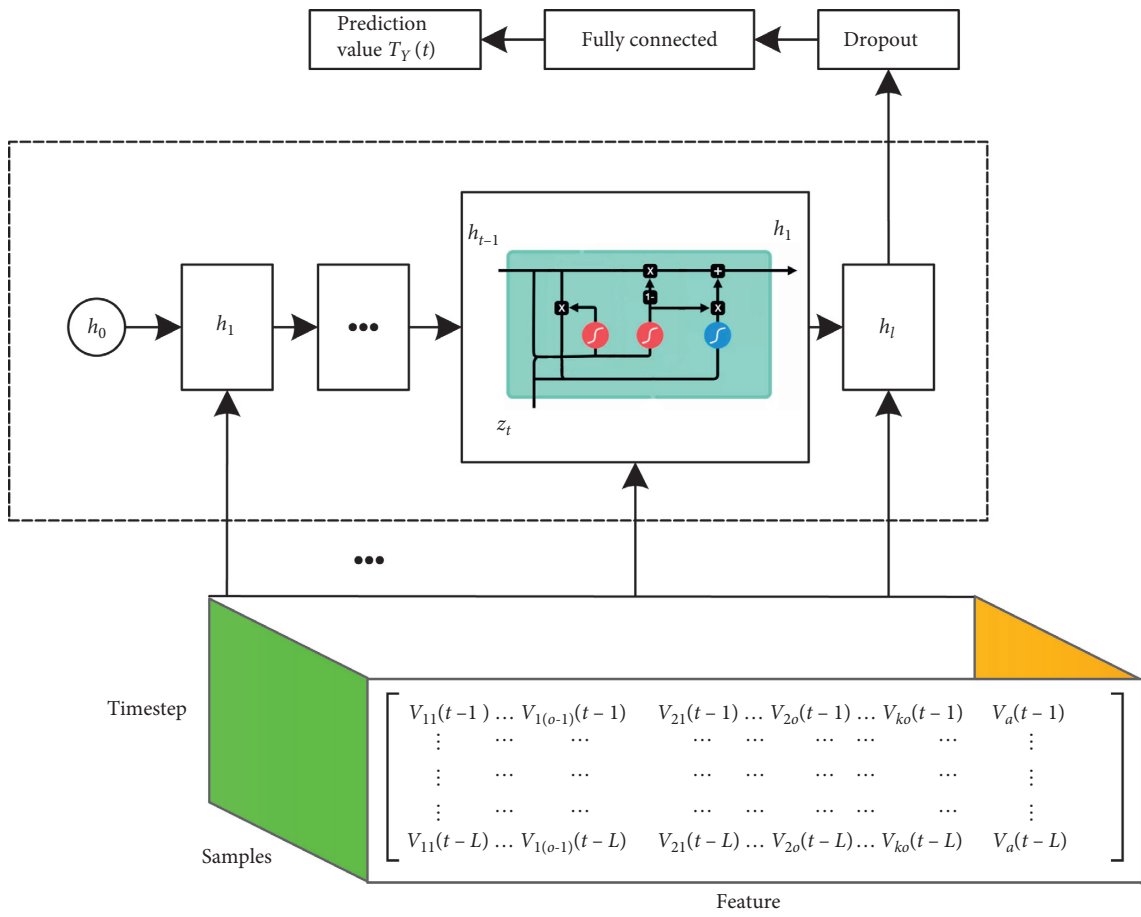


FIGURE 2: GRU algorithm framework.

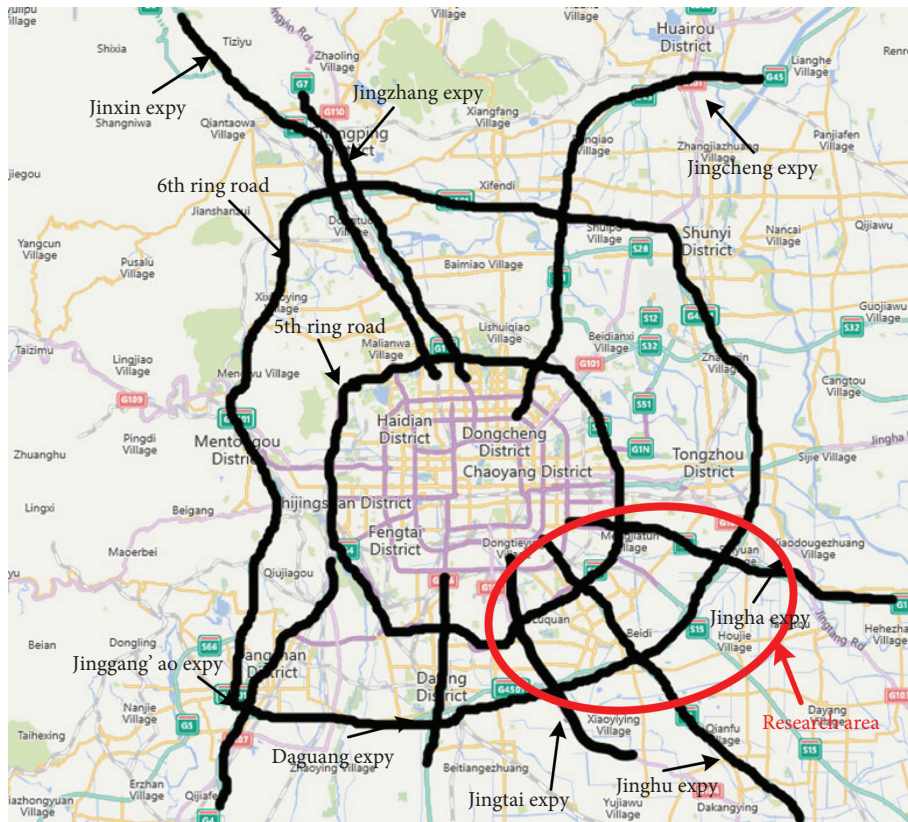


FIGURE 3: Adjacent region between the highway and urban expressway.

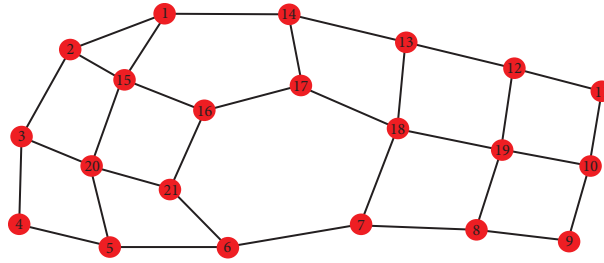


FIGURE 4: Node number of the road network.

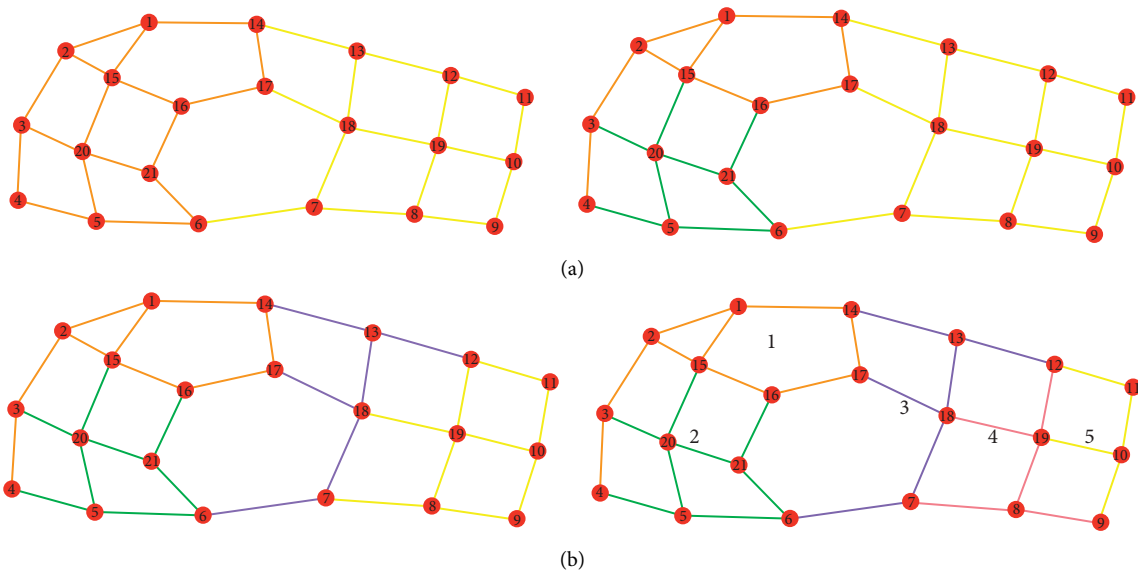


FIGURE 5: Subregional division of each scheme.

TABLE 1: Dispersion of traffic flow of each scheme.

Subarea number	Transportation dispersion in subareas	Dispersion of traffic flow in the whole road network
2	0.08350	0.06334
	0.07194	
3	0.05627	0.06130
	0.05957	
	0.06334	
4	0.05627	0.06499
	0.05957	
	0.03872	
	0.11177	
5	0.05627	0.05673
	0.05957	
	0.03872	
	0.07729	
	0.06224	



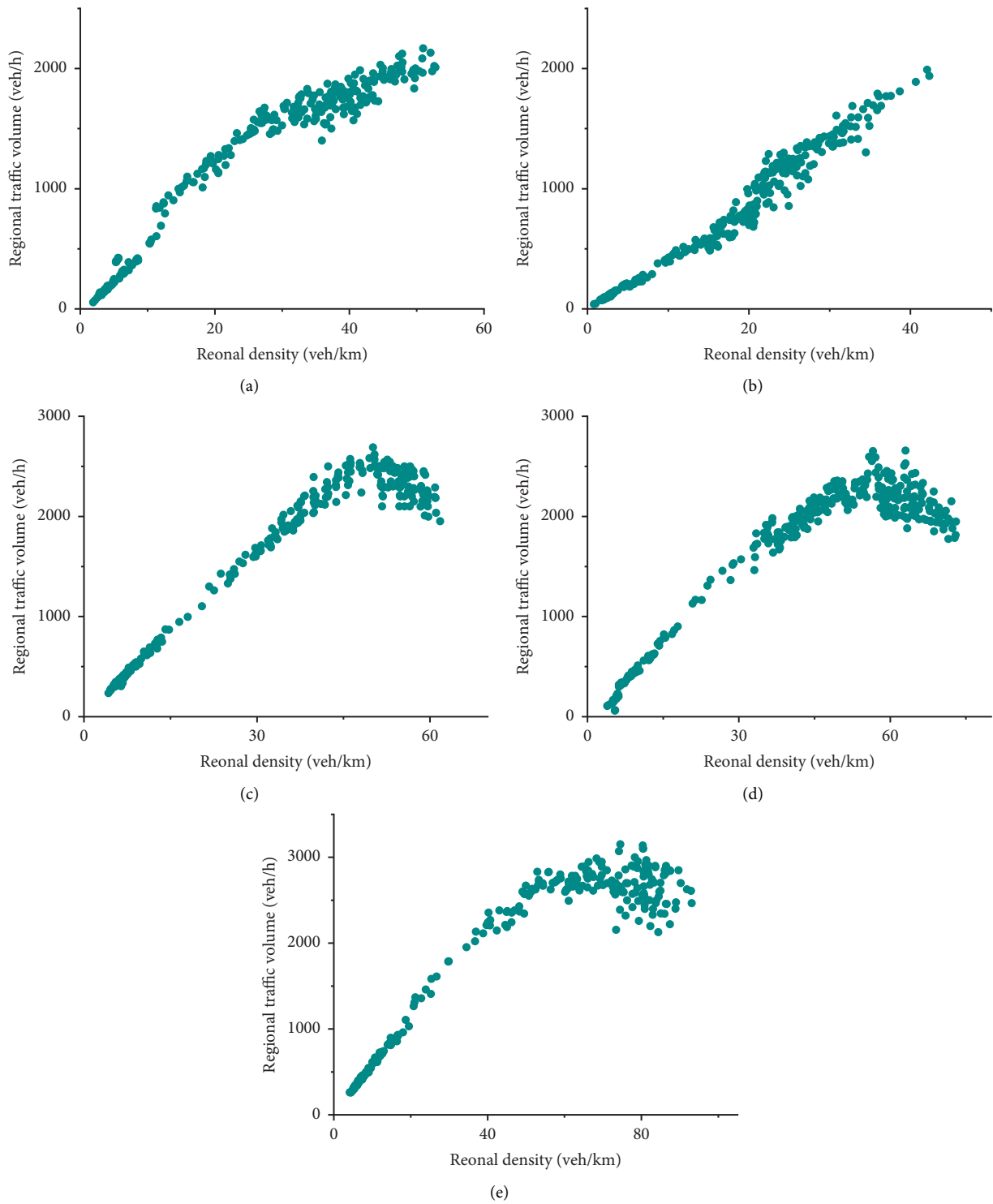


FIGURE 6: MFD of each subregional. (a) Subregional one. (b) Subregional two. (c) Subregional three. (d) Subregional four. (e) Subregional five.

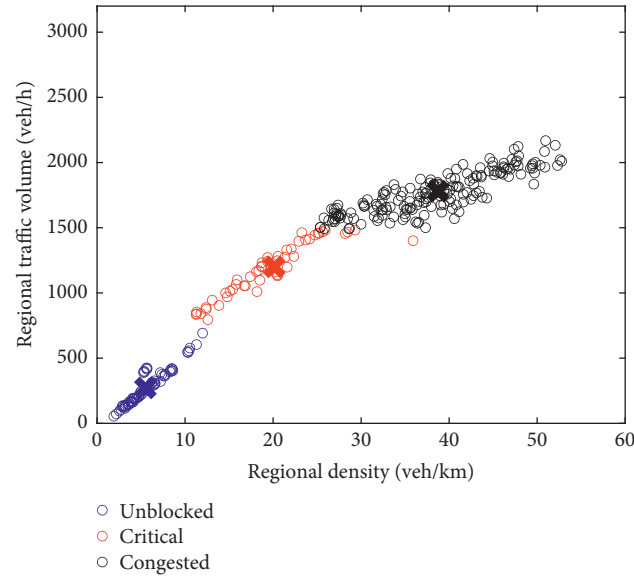


FIGURE 7: The traffic state classification results for subarea 1.

TABLE 2: The clustering centers.

Traffic state	Regional density	Regional traffic volume
Unblocked	5.49	272
Critical	20.13	1200
Congested	38.74	1780

TABLE 3: Prediction accuracy of road Section 1 for different K values.

	GRU	$K=0$	$K=1$	$K=2$	$K=3$	$K=4$
MAPE (%)	4.15	4.275	4.018	4.120	4.49	4.667
RMSE	2.331	2.251	2.217	2.294	2.402	2.42

TABLE 4: Prediction accuracy of road Section 2 for different K values.

	GRU	$K=0$	$K=1$	$K=2$	$K=3$	$K=4$
MAPE (%)	4.003	4.016	3.903	4.009	4.16%	3.955
RMSE	3.136	3.122	2.97	3.051	3.142	3.145

### 5.2.1. Working Days

(1) *Section 1.* The predicted results of Section 1 on July 2 (working day) are shown in Figure 8. The errors are shown in Figure 9. It can be seen that on July 2, the average speed of the morning rush hour and noon rush hour sections is low, in a congested state, and the road section was in a state of

smooth flow at night. The algorithm proposed in this paper can achieve acceptable prediction results. The results of error evaluation indicators MAPE and RMSE are shown in Table 5. MAPE was 2.30%, 3.05%, and the RMSE was 1.34 and 1.68, respectively.

(2) *Section 2.* The predicted results of Section 2 on July 2 (working day) are shown in Figure 10. The errors are shown

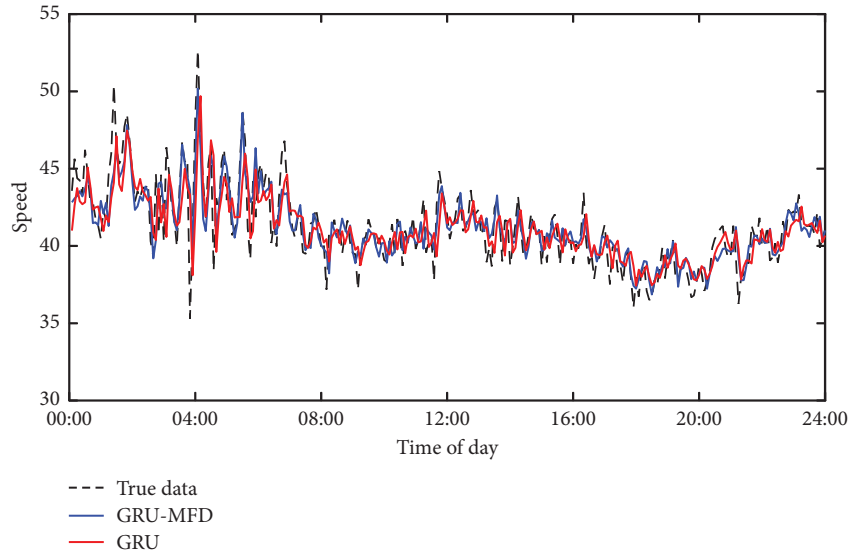


FIGURE 8: Prediction results on July 2, 2018, Section 1.

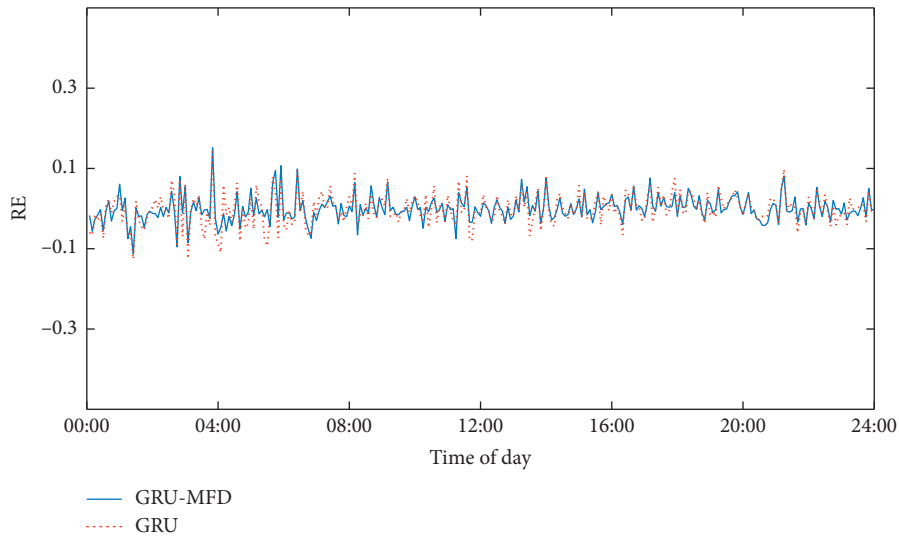


FIGURE 9: Errors on July 2, 2018, Section 1.

TABLE 5: Section 1 prediction error.

Evaluation index	Working day 7.2		Weekend 7.1		Rainy day 7.5	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
GRU-MFD	2.30	1.34	3.93	2.17	2.86	1.65
GRU	3.05	1.68	4.25	2.31	3.23	1.86

in Figure 11. The algorithm proposed in this paper can achieve acceptable prediction results. The results of error evaluation indicators MAPE and RMSE are shown in Table 6. MAPE was 1.42%, 2.53%, and the RMSE was 1.05 and 1.79, respectively.

5.2.2. *Weekend. (1) Section 1.* The predicted results of Section 1 on July 1 (weekend) are shown in Figure 12. The errors are shown in Figure 13. The algorithm proposed in this paper can achieve acceptable prediction results. The results of error evaluation indicators MAPE and RMSE are

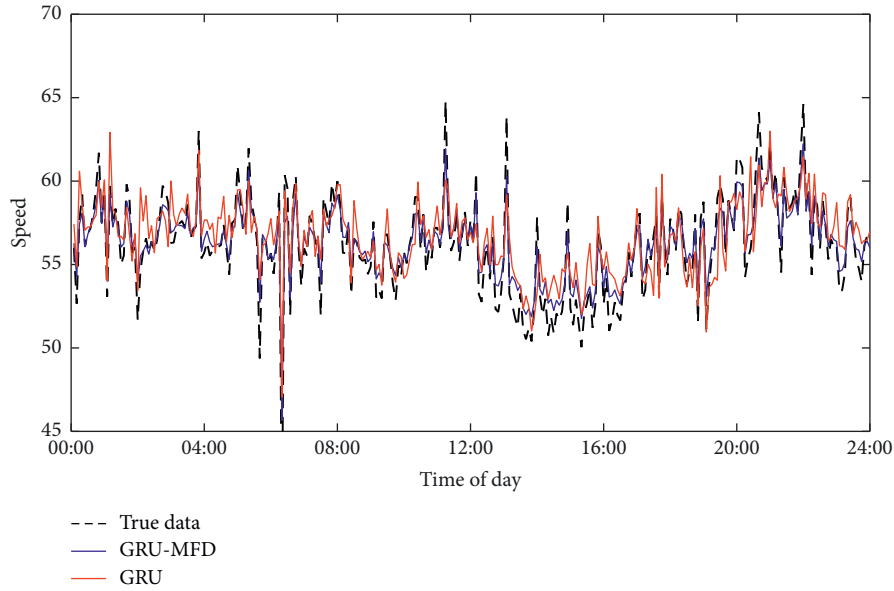


FIGURE 10: Prediction results on July 2, 2018, Section 2.

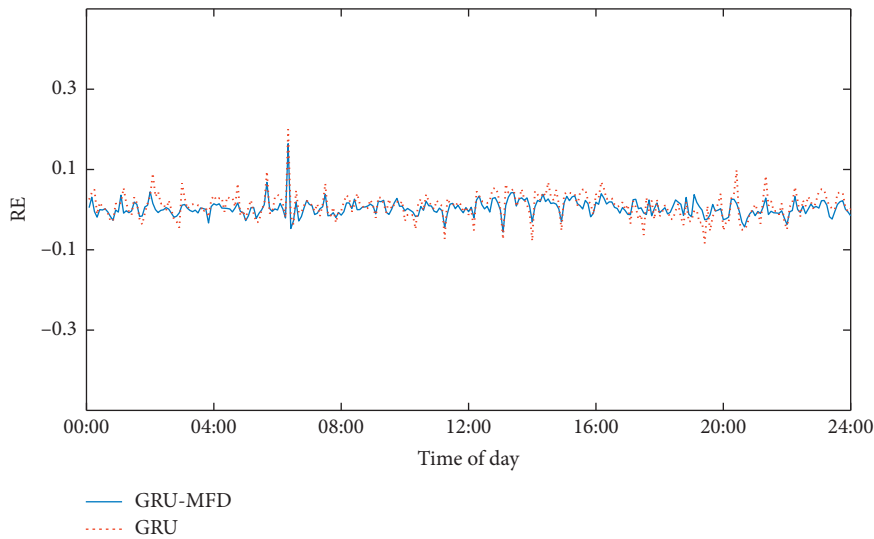


FIGURE 11: Errors on July 2, 2018, Section 2.

shown in Table 5. MAPE was 3.93%, 4.25%, and the RMSE was 2.17 and 2.31, respectively.

(2) *Section 2.* The predicted results of Section 2 on July 1 (weekend) are shown in Figure 14. The errors are shown in Figure 15. The algorithm proposed in this paper can achieve acceptable prediction results. The results of error evaluation indicators MAPE and RMSE are shown in Table 6. MAPE was 2.46%, 3.93%, and the RMSE was 1.89 and 2.18, respectively.

5.2.3. *Rainy Day.* (1) *Section 1.* The predicted results of Section 1 on July 5 (rainy day) are shown in Figure 16. The errors are shown in Figure 17. The algorithm proposed in this paper can achieve acceptable prediction results. The results of error evaluation indicators MAPE and RMSE are shown in Table 5. MAPE was 2.86%, 3.23%, and RMSE were 1.65 and 1.86, respectively.

(2) *Section 2.* The predicted results of Section 2 on July 5 (rainy day) are shown in Figure 18. The errors are shown in

TABLE 6: Section 2 prediction error.

Evaluation index	Working day 7.2		Holiday 7.1		Rainy day 7.5		Accident 7.11	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE (%)	RMSE
GRU-MFD	1.42	1.05	2.46	1.89	1.49	1.03	7.35	5.05
GRU	2.53	1.79	3.93	2.18	2.55	1.82	8.94	5.83

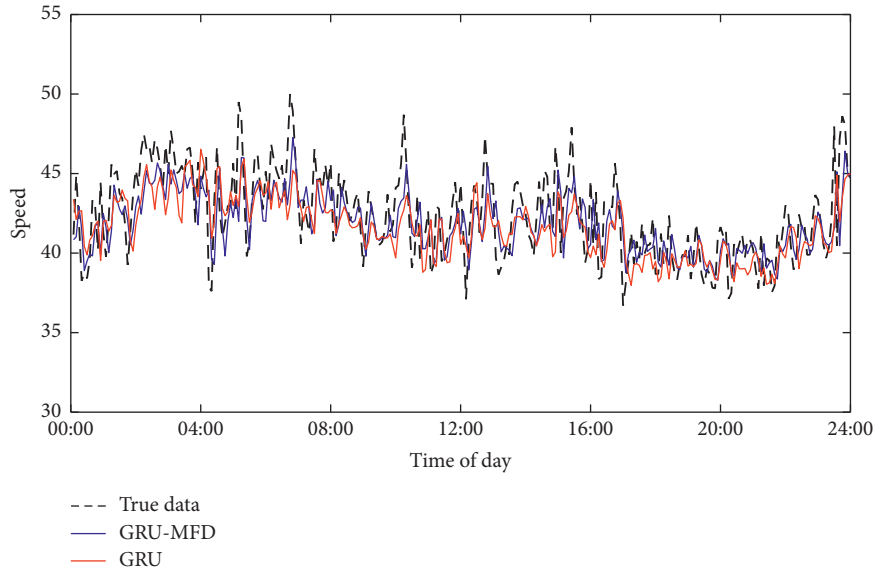


FIGURE 12: Prediction results on July 1, 2018, Section 1.

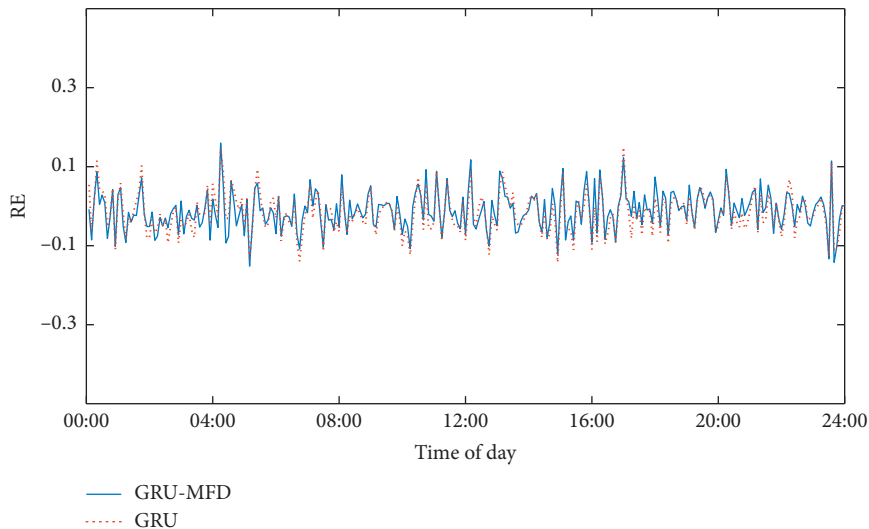


FIGURE 13: Errors on July 1, 2018, Section 1.

Figure 19. The algorithm proposed in this paper can achieve acceptable prediction results. The results of error evaluation indicators MAPE and RMSE are shown in Table 6. MAPE was 1.49% and 2.55%, and the RMSE was 1.03 and 1.82, respectively.

5.2.4. Accident. The predicted results of Section 2 on July 11 are shown in Figure 20. The errors are shown in Figure 21.

There was a traffic accident in the early hours of the morning. The results of MAPE and RMSE are shown in Table 6. MAPE was 7.35% and 8.94%, and the RMSE was 5.05 and 5.83, respectively.

It can be seen from the above figures and tables that, compared with the traditional prediction method based on single segment time series, the prediction accuracy of the proposed algorithm is improved. When no accident happened,

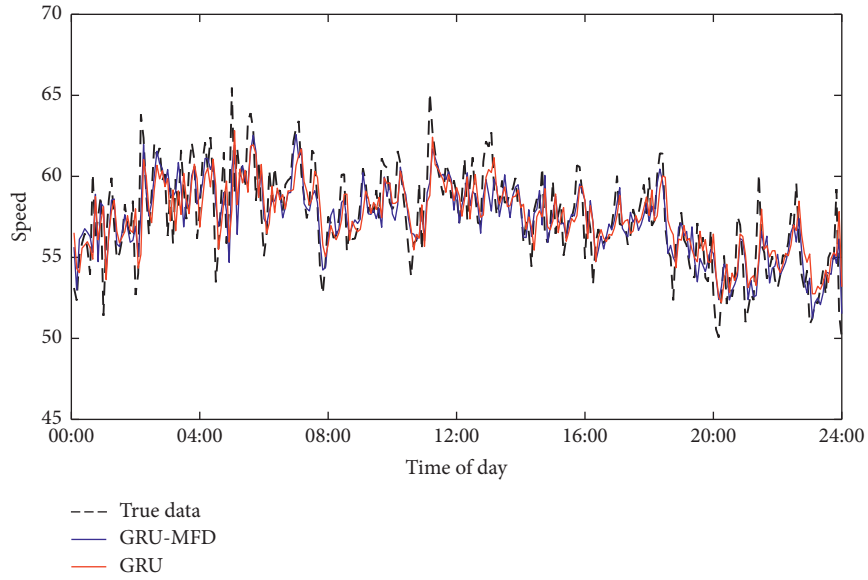


FIGURE 14: Prediction results on July 1, 2018, Section 2.

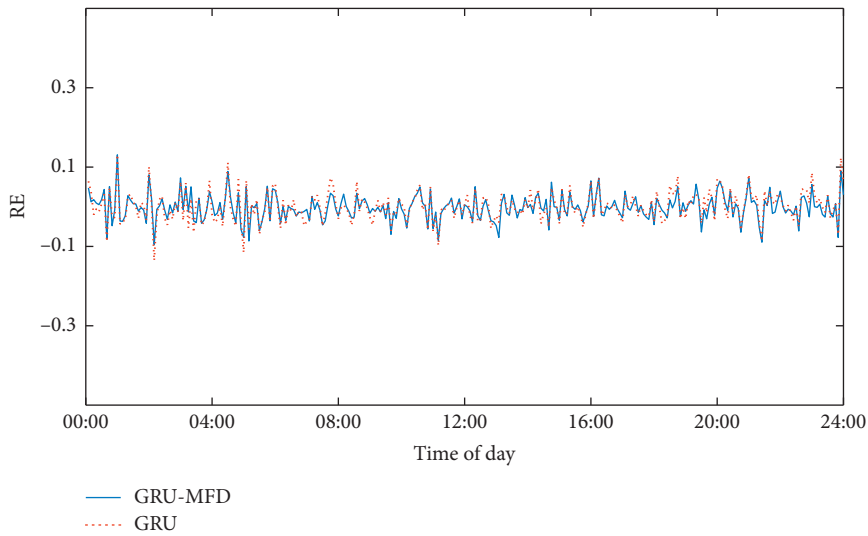


FIGURE 15: Errors on July 1, 2018, Section 2.

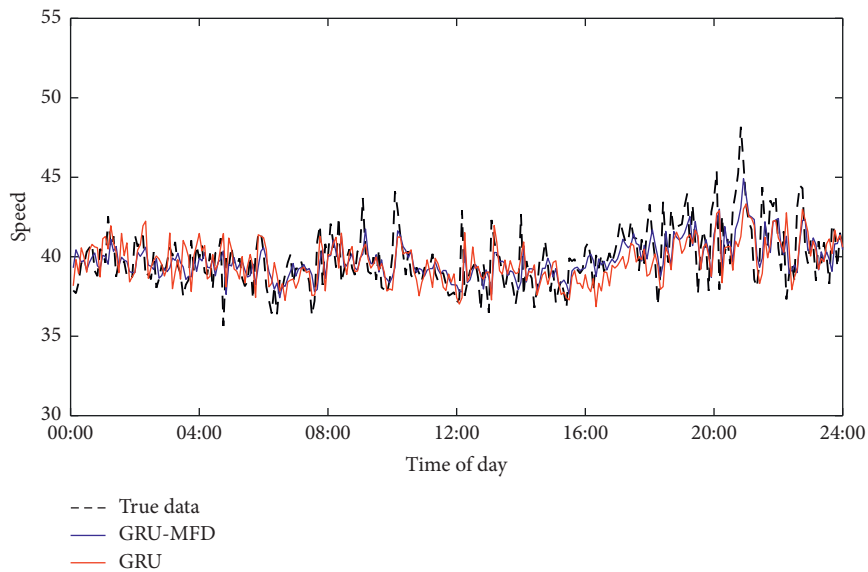


FIGURE 16: Prediction results on July 5, 2018, Section 1.

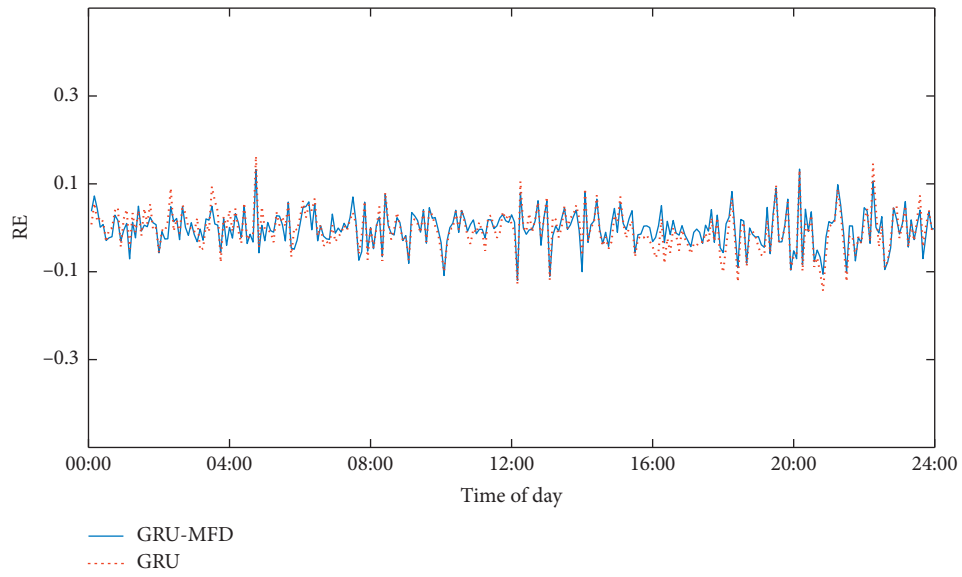


FIGURE 17: Errors on July 5, 2018, Section 1.

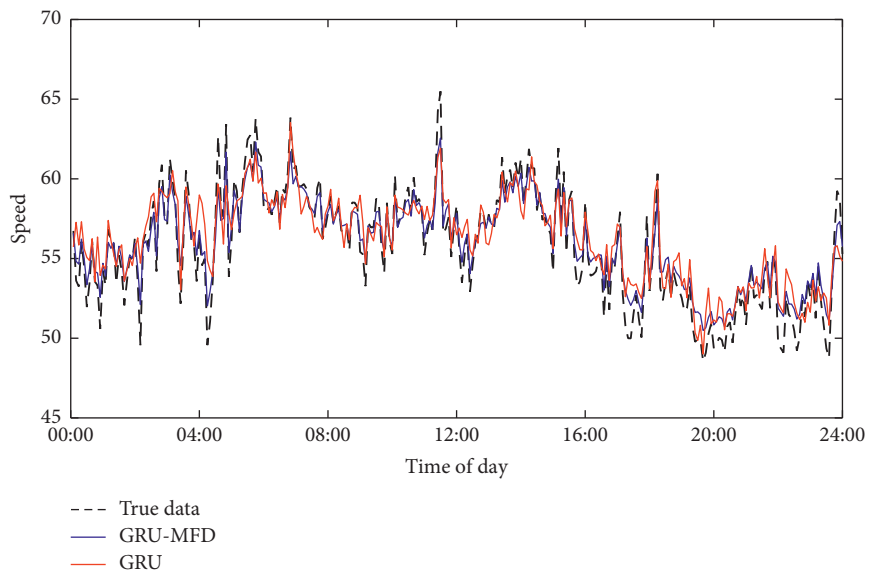


FIGURE 18: Prediction results on July 5, 2018, Section 2.

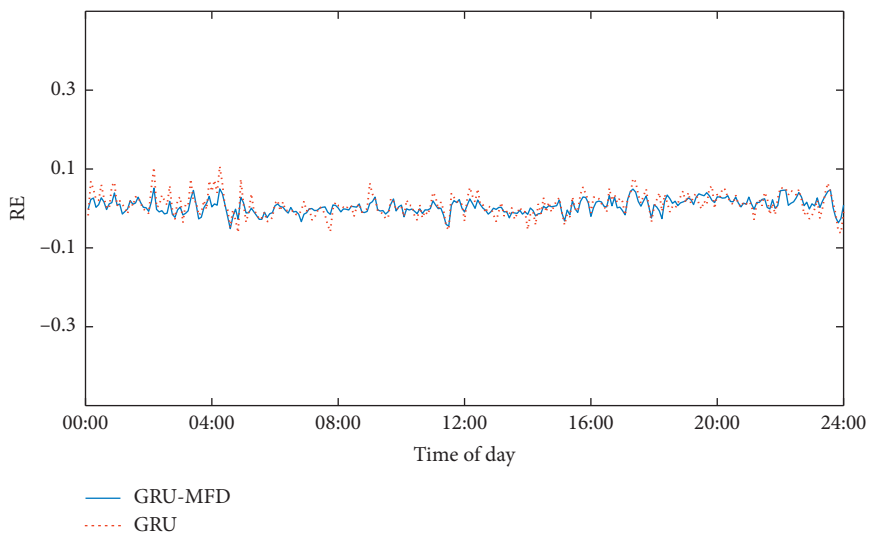


FIGURE 19: Errors on July 5, 2018, Section 2.

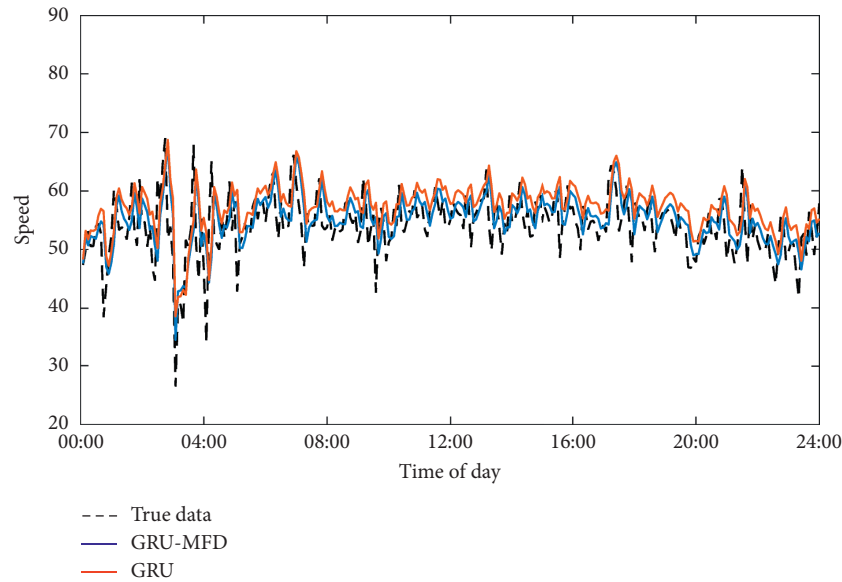


FIGURE 20: Prediction results on July 11, 2018, Section 2.

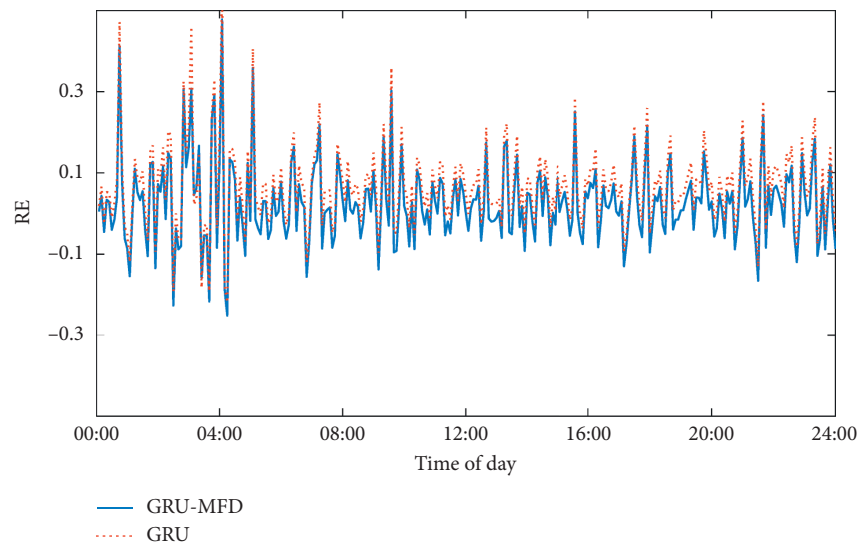


FIGURE 21: Errors on July 11, 2018, Section 2.

MAPE increased by about 0.5%, and on the day of the accident, MAPE increased by about 1.5%. The RMSE also improved more on accident day, compared with the day of no accident.

## 6. Conclusions

The traffic speed in adjacent regions between highway and expressway has gradually become important information concerned by highway managers and travelers. This paper proposes a prediction method of road traffic speed that considers microscopic traffic flow characteristics and macroscopic traffic status based on the road section average speed and flow data extracted from the GPS data.

Based on MFD, road network subareas are divided and evaluated. Firstly, the Ncut algorithm is used for the division of the road network. Secondly, to ensure the stability of the

divided subarea's MFD, the definition of the road network's discrete degree is proposed. The traffic state is divided combined with FCM to get the best scheme for dividing the subregions after the calculation of the discrete degree of the whole network.

The spatial-temporal correlation coefficient is proposed to measure the correlation between subareas. Then, the traffic speed time sequence of the study subarea and the related area is used to build a matrix of traffic speed. The regional matrix of traffic speed data is input into the GRU model, and the output result is the predicted traffic speed of the studied region.

This paper takes the adjacent region between the highway and expressway of Beijing as an example to verify the algorithm. The southeast corner of the Beijing road network is selected as the research area. The area consists of two ring expressways and



three highways with a total area of approximately 110 square kilometers. Truck GPS data from this region is the basis of this study. The proposed algorithm's accuracy is verified under the working days, weekend, rainy days, and accident scenarios. The result shows that, compared with the traditional prediction method based on single segment time series, the prediction accuracy of the proposed algorithm is improved. This will enhance the level of traffic information services in the adjacent region between the highway and urban expressway and ease traffic congestion.

## Data Availability

The data used to support the findings of this study have not been made available because the authors have signed the confidentiality agreement with the data providers.

## Conflicts of Interest

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

## Acknowledgments

This work was supported by the National Key Research and Development Program of China (no. 2018YFB1600900) and National Natural Science Foundation of China (71871011, 71890972/71890970, and 71621001).

## References

- [1] L. Peng, Z. Li, C. Wang, and T. Sarkodie-Gyan, "Evaluation of roadway spatial-temporal travel speed estimation using mapped low-frequency AVL probe data," *Measurement*, vol. 165, Article ID 108150, 2020.
- [2] X. Qu, Y. Yu, M. Zhou, C.-T. Lin, and X. Wang, "Jointly dampening traffic oscillations and improving energy consumption with electric, connected and automated vehicles: a reinforcement learning based approach," *Applied Energy*, vol. 257, Article ID 114030, 2020.
- [3] H. Liu, K. Zhang, and J. Z. Henk, "Short-term travel time prediction for urban networks," *Journal of Transportation Systems Engineering & Information Technology*, vol. 7, no. 3, pp. 118–124, 2007.
- [4] D. Zhang, C. OuYang, and W. Wu, "N-order neighbor road network based vehicle travel time estimation model," *Chinese Journal of Electronics*, vol. 43, no. 12, pp. 2491–2496, 2015.
- [5] B. Dhivyabharathi, E. S. Hima, and L. Vanajakshi, "Stream travel time prediction using particle filtering approach," *Transportation Letters*, vol. 10, no. 9, pp. 1–8, 2016.
- [6] J. Zhao, F. Xu, K. Zhang, and J. Bai, "Highway travel time prediction based on multi-source data fusion," *Journal of Transportation Systems Engineering & Information Technology*, vol. 16, no. 1, pp. 52–57, 2016.
- [7] J. Zhao, H. Wang, and W. Liu, "Prediction of highway travel time by adaptive kalman filtering," *Journal of South China University of Technology (Natural Science Edition)*, vol. 2014, no. 2, pp. 109–115, 2014.
- [8] Z. Wang, A. Goodchild, and E. McCormack, "Measuring truck travel time reliability using truck probe GPS data," *Journal of Intelligent Transportation Systems*, vol. 20, no. 2, pp. 103–112, 2016.
- [9] M. Zhou, X. Qu, and X. Li, "A recurrent neural network based microscopic car following model to predict traffic oscillation," *Transportation Research Part C: Emerging Technologies*, vol. 84, pp. 245–264, 2017.
- [10] Z. Wang and A. V. Goodchild, "GPS data analysis of the impact of tolling on truck speed and routing," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2411, no. 1, pp. 112–119, 2014.
- [11] H. Jula, M. Dessouky, P. Ioannou, and A. Chassiakos, "Container movement by trucks in metropolitan networks: modeling and optimization," *Transportation Research Part E*, vol. 41, no. 3, pp. 235–259, 2011.
- [12] C. Dong, K. Xie, X. Sun, M. Lyu, and H. Yue, "Roadway traffic crash prediction using a state-space model based support vector regression approach," *PLoS One*, vol. 14, no. 4, Article ID e0214866, 2019.
- [13] B. Yu, H. Wang, W. Shan, and B. Yao, "Prediction of bus travel time using random forests based on near neighbors," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 4, pp. 333–350, 2018.
- [14] S. Xie and L. Wei, "Application of elman neural networks to predict truck's operating speed," *Advanced Materials Research*, vol. 2013, no. 4, pp. 694–697, 2013.
- [15] X. Wang and L. Xu, "Study on short-term traffic flow prediction based on deep learning," *Journal of Transportation Systems Engineering & Information Technology*, vol. 18, no. 1, pp. 81–88, 2018.
- [16] W. Luo, B. Dong, and Z. Wang, "Short-term traffic flow prediction based on CNN-SVR hybrid deep learning model," *Journal of Transportation Systems Engineering & Information Technology*, vol. 2017, no. 5, pp. 68–74, 2017.
- [17] Z. Yao, C. Shao, and Z. Xiong, "Research on use of support vector machine for forecasting link travel time," *Journal of Highway and Transportation Research and Development*, vol. 24, no. 9, pp. 96–99, 2007.
- [18] Y. Wu, H. Tan, L. Qin, B. Ran, and Z. Jiang, "A hybrid deep learning based traffic flow prediction method and its understanding," *Transportation Research Part C: Emerging Technologies*, vol. 90, no. 90, pp. 166–180, 2018.
- [19] Y. Jia, J. Wu, and Y. Du, "Traffic speed prediction using deep learning method," *IEEE International Conference on Intelligent Transportation Systems*, vol. 2016, no. 11, pp. 1–4, 2016.
- [20] J. Zhao, Y. Gao, Z. Yang et al., "Truck traffic speed prediction under non-recurrent congestion: based on optimized deep learning algorithms and GPS data," *IEEE Access*, vol. 7, no. 7, pp. 9116–9127, 2019.
- [21] Z. Wang, D. Li, and X. Cui, "Travel time prediction based on LSTM neural network in precipitation," *Journal of Transportation Systems Engineering & Information Technology*, vol. 20, no. 1, pp. 137–144, 2020.
- [22] D. Zang, J. Ling, J. Cheng, K. Tang, and X. Li, "Using convolutional neural network with asymmetrical kernels to predict speed of elevated highway," in *Proceedings of the International Conference on Ion Sources*, pp. 212–221, Shanghai, China, October 2017.
- [23] B. Peng, Ju. Tang, X. Cai, J. Xie, Y. Zhang, and Y. Wang, "3DCNN-DNN based traffic status prediction from aerial videos," *Journal of Transportation Systems Engineering & Information Technology*, vol. 20, no. 3, pp. 39–46, 2020.
- [24] J. Zhao, Y. Gao, J. Tang, L. Zhu, and J. Ma, "Highway travel time prediction using sparse tensor completion tactics and K-nearest neighbor pattern matching method," *Journal of Advanced Transportation*, vol. 2018, Article ID 5721058, 2018.

- [25] J. Zhao, Y. Gao, Y. Guo, and Z. Bai, "Travel time prediction of expressway based on multi-dimensional data and the particle swarm optimization–autoregressive moving average with exogenous input model," *Advances in Mechanical Engineering*, vol. 10, no. 2, pp. 1–16, 2018.
- [26] J. Zhao, Y. Gao, Y. Qu, H. Yin, Y. Liu, and H. Sun, "Travel time prediction: based on gated recurrent unit method and data fusion," *IEEE Access*, vol. 6, no. 6, pp. 70463–70472, 2018.
- [27] M. Zhou, Y. Yu, and X. Qu, "Development of an efficient driving strategy for connected and automated vehicles at signalized intersections: a reinforcement learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 1, pp. 433–443, 2020.
- [28] Y. Ji and N. Geroliminis, "On the spatial partitioning of urban transportation networks," *Transportation Research Part B: Methodological*, vol. 46, no. 10, pp. 1639–1656, 2012.
- [29] Y. Ji, J. Luo, and N. Geroliminis, "Empirical observations of congestion propagation and dynamic partitioning with probe data for large-scale systems," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2422, no. 2422, pp. 1–11, 2014.
- [30] J. Haddad and N. Geroliminis, "On the stability of traffic perimeter control in two-region urban cities," *Transportation Research Part B: Methodological*, vol. 46, no. 9, pp. 1159–1176, 2012.
- [31] Y. Ma, X. Yang, and Y. Zeng, "Urban traffic signal control network partitioning using spectral method," *Systems Engineering-Theory & Practice*, vol. 30, no. 12, pp. 2290–2296, 2010.
- [32] F. Yang, L. Peng, A. Teng, and C. Wang, "Measuring temporal and spatial travel efficiency for transit route system using low-frequency bus automatic vehicle location data," *Advances in Mechanical Engineering*, vol. 10, no. 10, pp. 1–11, 2018.
- [33] H. Su, Z. Bai, L. Deng, S. She, Y. Qi, and J. Zhao, "Research on expressway state recognition based on Beidou positioning and navigation data," in *Proceedings of International Conference on Smart Grid and Electrical Automation*, pp. 354–357, Changsha, China, June 2018.
- [34] H. Han, L. Peng, A. Teng, C. Wang, and T. Z. Qiu, "Evaluation of freeway travel speed estimation using anonymous cell-phones as probes: a field study in," *Canadian Journal of Civil Engineering*, 2020.