



## Research Article

# Highway Traffic Speed Prediction in Rainy Environment Based on APSO-GRU

Dongqing Han,<sup>1</sup> Xin Yang,<sup>1</sup> Guang Li,<sup>2</sup> Shuangyin Wang,<sup>1</sup> Zhen Wang ,<sup>3</sup>  
and Jiandong Zhao ,<sup>3,4</sup>

<sup>1</sup>Zhong Dian Jian Ji Jiao Highway Investment Development Company Limited, Shijiazhuang, Hebei 050090, China

<sup>2</sup>Hebei Intelligent Transportation Technology Co., Ltd of HEBTIG, Shijiazhuang, Hebei 050090, China

<sup>3</sup>School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

<sup>4</sup>Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Ministry of Transport, Beijing 100044, China

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

Received 12 July 2021; Accepted 4 August 2021; Published 18 August 2021

Academic Editor: Ming Xu

Copyright © 2021 Dongqing Han 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.

In order to accurately analyse the impact of the rainy environment on the characteristics of highway traffic flow, a short-term traffic flow speed prediction model based on gate recurrent unit (GRU) and adaptive nonlinear inertia weight particle swarm optimization (APSO) was proposed. Firstly, the rainfall and highway traffic flow data were cleaned, and then they are matched according to the spatiotemporal relationship. Secondly, through the method of multivariate analysis of variance, the significance of the impact of potential factors on traffic flow speed was explored. Then, a GRU-based traffic flow speed prediction model in rainy environment is proposed, and the actual road sections under different rainfall scenarios were verified. After that, in view of the problem that the prediction accuracy of the GRU model was low in the continuous rainfall scenario, the APSO algorithm was used to optimize the parameters of the GRU network, and the APSO-GRU prediction model was constructed and verifications under the same road section and rain scene were carried out. The results show that the APSO-GRU model has significantly improved prediction stability than the GRU model and can better extract rainfall features during continuous rainfall, with an average prediction accuracy rate of 96.74%.

## 1. Introduction

Rainfall is the most frequently occurring severe weather, which brings serious impact to highway traffic safety. It is important to study the traffic flow characteristics of highways under rainy environment and grasp the regularities of rainfall on traffic flow, making stable prediction and analysis of traffic flow to implement effective traffic control [1–4].

In terms of the impact of adverse weather on traffic flow characteristics, with the improvement of the highway system and the continuous development of information observation and collection technology, scholars at home and abroad have conducted continuous research [5–8]. At present, the data analysis and modelling system for the

impact of weather factors on highway traffic flow is well established. In 1994, Ibrahim and Hall [9] studied the impact of adverse weather on the flow-occupancy and speed-flow relationships through regression analysis and showed that heavy rain and snow caused a 10–20% and 30%–48% reduction in maximum highway flow, respectively. In 2005, Agarwal et al. [10] used years of traffic flow data and contemporaneous weather data of the Northern United States to quantify the effects of adverse weather conditions and roadway conditions on highway traffic flow. The results showed that heavy rain, snow, and low visibility resulted in a 10%–17%, 19%–27%, and 12% of reduction in capacity, respectively, and a reduction of vehicle speed by 4%–7%, 11%–15%, and 10%–12%, respectively. In 2015, Li [11] derived the mean values of vehicle speeds of different

rainfall intensities on highways based on statistical analysis of data, used standard deviations to measure the dispersion of vehicle speeds, and investigated the variability of vehicle speed on different lane locations, vehicle types, and time periods during rainfall.

In terms of traffic flow prediction models, they are mainly divided into prediction models based on statistical theory analysis, nonlinear theory prediction models, machine learning prediction models, and combined prediction models [12]. However, nonlinear theoretical model related theories are very complex, especially in terms of mathematical processing. The model has a high degree of complexity and a large amount of calculation, which is suitable for more complicated emergency transportation systems. With the rise of artificial intelligence, machine learning has been more often applied to the field of traffic flow prediction, and related prediction algorithms have emerged. They are mainly divided into support vector machines, artificial neural networks, and deep learning [13]. In 2009, Castro-Neto et al. [14] proposed a supervised online SVR statistical learning model, which optimized the problem of limited applicability of general models in atypical cases. The developed model outperformed models such as Gaussian maximum likelihood, Holt exponential smoothing, and artificial neural networks in typical and atypical traffic flow prediction. In 2013, Jeong et al. [15] proposed an online learning weighted SVR model (OLWSVR) for short-term traffic flow prediction, which outperformed prediction models such as locally weighted regression, conventional SVR, and online learning SVR. Smith and Demetsky [16] analysed short-term traffic flow prediction models based on neural network and non-parametric regression. Cai et al. [17] proposed a neural network based on improved cuckoo algorithm with optimized radial basis function (CS-RBF) for highway traffic flow prediction under heavy rainfall, and the study showed that the algorithm has better prediction accuracy and convergence speed. In 2015, Lv et al. [18] proposed traffic flow feature learning using a stacked autoencoder model and trained it with greedy hierarchical unsupervised learning deep learning model. Zhang and Wang [19] built an urban trunk road travel time prediction model based on GRU network and simulated it using real road network data. In 2020, Wang et al. [20] proposed an LSTM travel time prediction model considering rainfall data, and the results showed that the prediction results with the inclusion of rainfall features were more accurate than when the rainfall features were not included. Meng proposed the LSTM-GRU combined model to predict the short-term traffic flow speed of highways in rainy days. This model is well adapted to the uncertainty and sudden change of traffic flow speed in rainy days [21].

Reviewing the above literature, we can find the following research trends regarding the influence of rainy weather environment on traffic flow characteristics and traffic flow prediction. (1) For the research on the influence of rainy weather environment on traffic flow characteristics, most domestic and foreign scholars divide the rainfall intensity into levels [22]. Using the traffic flow data and rainfall data of

the actual road section, the changes of the macro traffic flow parameter values of the road section under different levels of unfavorable weather are given. However, there is a lack of comprehensive consideration of the impact of multiple factors on traffic flow. (2) In terms of traffic flow prediction, various prediction models have different principles. Currently, machine learning and deep learning models have become the mainstream of research in the field of traffic flow prediction. Throughout the many previous studies on traffic flow prediction, there are fewer studies on traffic flow prediction under rainy environment, and more related studies only add the verification of rainy weather scenarios on the traditional prediction.

Therefore, in this paper, we consider the influence of various factors to carry out the research on traffic flow characteristics of highways under rainy environment. Also, we add rainfall features to the deep learning model to carry out the prediction of highway traffic flow speed under rainy environment. In view of the fact that the PSO algorithm can adjust the hyperparameters of the deep learning model and bring better prediction performance, this article will build the APSO-GRU model.

## 2. Data Preprocessing

*2.1. Preprocessing of Traffic Flow Data.* The data in this article come from the floating car data of Beijing-Harbin Highway (JingHa Highway), Beijing-Tianjin Highway (JingJin Highway), Beijing-Taipai Highway (JingTai Highway), and Beijing-Kaifeng Highway (JingKai Highway). By fusing multisource floating car data and combining the original data with relevant geographic information through the MapInfo interface, the space-time matching of traffic flow data is completed. After data preprocessing, the proportion of abnormal data accounts for 5% of the total original data.

The raw traffic flow data are recorded in 5-minute intervals, spanning a total of six months from June 1 to August 31, 2018, and June 1 to August 31, 2019. The raw data include information such as highway section ID, section direction, average vehicle speed, and traffic flow. The data format is shown in Table 1.

In the table, the first 13 digits of Section\_id indicate the number of a section of the highway, and the last digit indicates the direction of vehicle travel on the section, with 1 representing upward and 0 representing downward; Speed\_avg indicates the average speed of all vehicles passing the section during the collection time; volume indicates the traffic volume of the section during the collection time.

The data cleaning is divided into two parts: rejection of erroneous data and repair of missing data. For the rejection of erroneous data, a “rule rejection method” is used, which integrates the threshold method and the basic theory of traffic flow [23]. For the missing original data [24] of traffic flow, a simple nearest neighbor mean fill method is used to fill the data, which combines the mean filling method of replacing the missing data with the mean of the existing data and the nearest neighbor interpolation method using the observation values near the missing value to replace the missing value. The nearest neighbor interpolation method of

missing values is combined. Then, we take the average value of the valid data adjacent to the missing data as filling, as in the following equation:

$$H_t = \frac{(H_{t-1} + H_{t+1})}{2}, \quad (1)$$

where  $H_t$  represents the missing traffic flow data of the  $t^{\text{th}}$  cycle, including  $V$  and  $Q$ , and  $H_{t-1}, H_{t+1}$  are traffic flow data of the two adjacent cycles of the  $t^{\text{th}}$  cycle.

**2.2. Preprocessing of Rainfall Data.** The rainfall data were obtained from Beijing Nanjiao Observatory Station (No. 54511), Tongzhou Station (No. 54431), and Daxing Station (No. 54594). In this paper, only hourly rainfall data and their corresponding dates and times are extracted. A total of 4397 meteorological data from Daxing District and 4401 meteorological data from Tongzhou District were extracted, and the format of rainfall data is shown in Table 2. The time item indicates the end moment of the data collection time, and the rainfall amount is the accumulated rainfall amount within the data collection time.

Only a small amount of rainfall data was found to be missing through inspection. According to the method of traffic flow data filling, the average value of rainfall in the adjacent hours of the missing data was used to fill in the missing data.

**2.3. Spatial and Temporal Matching.** The recording period of traffic flow data is 5 minutes, while the recording period of rainfall data is 1 hour. It is necessary to match two data from time granularity. High-precision rainfall data are currently not available, and it is difficult to decompose long-period data into short-period data. In addition, the time accuracy of the floating car data acquisition system needs to be improved. Even if high-precision rainfall data are obtained, the time-space error of the two data matching is difficult to evaluate. So, it is more reasonable to combine the traffic flow data from 5-minute recording period to 1-hour recording period, and the combination rule is expressed as follows:

$$\begin{cases} \bar{C} = \sum_{i=1}^{12} C_i, \\ \bar{V} = \frac{\sum_{i=1}^{12} V_i C_i}{C}, \end{cases} \quad (2)$$

where  $\bar{C}$  represents traffic flow on the section in one hour (veh/h);  $C_i$  represents traffic flow on the section in 5 minutes (veh/5 min);  $\bar{V}$  represents the average speed of vehicles on the section in one hour (km/h); and  $V_i$  represents the average vehicle speed on the section in 5 minutes (km/h).

Based on the weighted average method, the average vehicle speed time series and traffic flow time series of each highway are calculated. The weight of the road section is the proportion of its length in the whole highway. The calculation method is shown in the following equation:

TABLE 1: Format of raw traffic flow data.

| Section_id     | Speed_avg (km/h) | Volume (veh/5 min) |
|----------------|------------------|--------------------|
| 59565500000081 | 86.52            | 267                |
| 59565500000090 | 85.57            | 238                |
| 59565500000101 | 86.34            | 314                |
| 59565500000221 | 85.61            | 284                |
| 59565500000231 | 85.82            | 329                |
| 59565600000010 | 85.66            | 283                |

TABLE 2: Rainfall data format.

| Station ID | Date     | Time  | Rainfall |
|------------|----------|-------|----------|
| 54594      | 20190615 | 22:00 | 0        |
| 54594      | 20190615 | 23:00 | 0.2      |
| 54594      | 20190615 | 24:00 | 1.0      |
| 54594      | 20190616 | 1:00  | 1.7      |

$$\begin{cases} V = \sum_{n=1}^K \frac{\bar{V}_n \cdot L_n}{L}, \\ C = \sum_{n=1}^K \frac{\bar{C}_n \cdot L_n}{L}, \end{cases} \quad (3)$$

where  $V$  represents the average speed of vehicles on the highway (km/h);  $\bar{V}_n$  represents the average vehicle speed of the road section (km/h);  $L_n$  represents the length of the road section (m);  $L$  represents the length of highway (m);  $K$  represents the total number of sections of the highway;  $C$  represents the overall traffic volume of the highway (veh/h); and  $C_n$  represents the traffic volume of the road section (veh/h).

The traffic flow time series and the rainfall time series in the region are integrated according to the corresponding time to complete the spatiotemporal matching, and the format of the matched data is shown in Table 3. In the table, Date\_hour indicates the date and time; precipitation indicates the rainfall amount in mm/h; Volume\_sum indicates the traffic flow in veh/h; and Speed\_avg indicates the average vehicle speed in km/h.

### 3. Analysis of the Influence of Rainy Weather on Traffic Flow Speed of Highway

#### 3.1. Analysis of the Factors Influencing the Traffic Flow Speed.

The traffic flow speed of highway is affected by many factors. Four potential factors, such as rainfall intensity, date category, time period, and number of lanes, are selected to explore whether these factors affect the traffic flow speed of highway by multivariate analysis of variance. According to the statistical analysis of the characteristics of highway traffic flow, it can be seen that the ‘‘morning peak hour’’ of highway is relatively lagging behind that of urban roads. Before the arrival of the peaking hour, it can be clearly seen that both traffic flow speed and traffic flow have experienced two processes of first decreasing and then increasing. Through observation, it is found that it is more reasonable to divide every four hours as a time period, as shown in Table 4.

TABLE 3: Data format after spatiotemporal matching.

| Highway                | Date_hour   | Precipitation | Volume_sum | Speed_avg |
|------------------------|-------------|---------------|------------|-----------|
| Beijing-Harbin Highway | 20190706_09 | 0             | 1106       | 89.61     |
| Beijing-Harbin Highway | 20190706_10 | 9             | 1051       | 85.34     |
| Beijing-Harbin Highway | 20190706_11 | 2.8           | 1137       | 87.74     |
| Beijing-Harbin Highway | 20190706_12 | 2.1           | 963        | 87.93     |
| Beijing-Harbin Highway | 20190706_13 | 1.5           | 897        | 88.57     |

TABLE 4: The division of different periods of the day.

| Period number | Time span   | Characteristic  |
|---------------|-------------|---|
| 1             | 0:00–4:00   | During the low period in the morning, the speed and flow decrease slowly                  |
| 2             | 4:00–8:00   | During the rising period, the speed and flow increase rapidly                             |
| 3             | 8:00–12:00  | In the morning peak, the speed and flow are high  |
| 4             | 12:00–16:00 | In the afternoon, the peak was flat, and the speed and flow decreased smoothly and slowly |
| 5             | 16:00–20:00 | In the evening peak, the flow reaches the peak again, and the speed decreases gradually   |
| 6             | 20:00–24:00 | In the low period of night, the speed and flow decrease rapidly                           |

SPSS software was used for multivariate analysis of variance, and the output of SPSS is shown in Figure 1. The results show that the four factors have significant influence on the traffic flow speed. In addition, the interaction of date category, time period, and number of lanes has a significant impact on traffic flow speed. The combination of other factors has no significant effect on traffic flow speed.

### 3.2. Influence of Rainfall on Traffic Flow Speed of Highway

**3.2.1. Distribution Characteristics of Traffic Flow Speed under Different Rainfall Intensities.** Considering the different levels of date categories and number of lanes, the distribution statistics of standard traffic flow speed under different rainfall intensities are carried out, which are divided into Tongzhou District and Daxing District, as shown in Figure 2.

It can be seen that the standard speed of highway vehicles in rainy days decreases with the increase of rainfall intensity. In terms of date category, weekend is more vulnerable to rainfall than working day. The slope of “rainfall intensity standard speed” of the four-lane highway in the two areas is greater than that of the three-lane highway, which indicates that the four-lane highway is more vulnerable to rainfall.

**3.2.2. Speed Distribution Characteristics of Traffic Flow in Different Periods.** Considering the different levels of rainfall intensity, date category, and number of lanes, the distribution statistics of the standard speed of vehicles in different periods of each highway are carried out, as shown in Figure 3.

It can be seen from Figure 3 that the standard speed of four-lane highway is generally lower than that of three-lane highway when other factors are the same, which indicates that its traffic flow speed is more easily affected. The same rainfall intensity has different influence on the traffic flow speed in different periods of the day, and the morning peak and evening peak are more easily affected by rainfall. Similarly, the speed of traffic flow in the first period, the second period, and the sixth period is relatively less

susceptible to rainfall. With the increase of rainfall intensity, the above differences will be more obvious.

## 4. The Establishment of the Prediction Model of the Traffic Flow Speed of APSO-GRU

**4.1. Design of Traffic Flow Speed Prediction Model Based on GRU.** The proposed GRU model is composed of three sections, i.e., input layer, hidden layer, and output layer. The output layer is a fully connected dense layer. Adam algorithm is selected as the weight optimizer to optimize the internal weight of the model. The structure of the prediction model is shown in Figure 4 [25].

The input of the model is a time series matrix composed of traffic flow speed, traffic flow, and rainfall, the output is traffic flow speed, and the loss function is MAE. MSE is more affected by outliers, while MAE is more stable. After actual data validation, the parameters of the GRU prediction model are set as follows: the number of hidden layer nodes is 15, dropout parameter is 0.3, batch size is 200, epoch is 180, and learning rate is 0.004. And the training set and test set are divided in a ratio of 4:1.

**4.2. PSO Algorithm.** Part of the parameters of the GRU model is automatically adjusted by the model, and the other part of the parameters needs to be set artificially, which are called superparameters, including the number of hidden layers, the number of hidden layer nodes, the number of iterations, etc., and the rationality of superparameter setting directly affects the convergence speed of model calculation and the accuracy of prediction. Therefore, this section uses PSO algorithm to optimize the GRU model.

Particle swarm optimization algorithm is described in the D-dimensional search space, and N different particles form a search population. The current position of the  $i^{\text{th}}$  particle is  $x_i = (x_{i1}, x_{i2}, \dots, x_{iD})$ , current speed is  $v_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ , and the best location currently searched by the individual is  $p_i = (p_{i1}, p_{i2}, \dots, p_{iD})$ ; it is called individual extremum [26]. The optimal position of the

| Intersubjective effect test                                     |                         |                   |             |             |              |
|---|-------------------------|-------------------|-------------|-------------|--------------|
| Dependent variable: traffic flow speed                          |                         |                   |             |             |              |
| category  | Type III sum of squares | Degree of freedom | Mean square | F           | Significance |
| Revised model   | 34335.592 <sup>a</sup>  | 95                | 361.427     | 434.476     | .000         |
| Bias  | 4594866.533             | 1                 | 4594866.535 | 5523544.621 | .000         |
| rainfall intensity  | 11347.833               | 3                 | 3782.611    | 4547.122    | .000         |
| date category   | 28.831                  | 1                 | 28.831      | 34.658      | .000         |
| number of lanes   | 2532.155                | 1                 | 2532.155    | 3043.934    | .000         |
| time period   | 13278.496               | 5                 | 2655.699    | 3192.448    | .000         |
| rainfall intensity* date category                               | 50.508                  | 3                 | 16.836      | 20.239      | .000         |
| rainfall intensity* number of lanes                             | 6.983                   | 3                 | 2.328       | 2.798       | .039         |
| rainfall intensity* time period                                 | 31.436                  | 15                | 2.096       | 2.519       | .001         |
| date category* number of lanes                                  | 22.878                  | 1                 | 22.878      | 27.502      | .000         |
| date category* time period                                      | 138.678                 | 5                 | 27.736      | 33.341      | .000         |
| number of lanes* time period                                    | 181.608                 | 5                 | 36.322      | 43.663      | .000         |
| rainfall intensity* date category* number of lanes              | 1.339                   | 3                 | .446        | .537        | .657         |
| rainfall intensity* date category* time period                  | 1.971                   | 15                | .131        | .158        | 1.000        |
| rainfall intensity* number of lanes* time period                | 2.142                   | 15                | .143        | .172        | 1.000        |
| date category* number of lanes* time period                     | 35.738                  | 5                 | 7.148       | 8.592       | .000         |
| rainfall intensity* date category* number of lanes* time period | 2.787                   | 15                | .186        | .223        | .999         |
| deviation   | 625.566                 | 752               |             |             |              |
| total   | 6023624.627             | 848               |             |             |              |
| Revised total   | 34961.157               | 847               |             |             |              |
| a. R square = .982 (revisedR square = .980)                     |                         |                   |             |             |              |

FIGURE 1: SPSS output.

whole population is called global extremum, and it is  $g = (g_{i1}, g_{i2}, \dots, g_{iD})$ .

The current position of the particle corresponds to a candidate solution of the optimization problem, and the flight process is the search process of the individual. Each particle iterates continuously to update its speed and position, which are determined by equations (4) and (5), respectively:

$$v_i(t+1) = v_i(t) + c_p \cdot r_1 \cdot (p_i(t) - x_i(t)) + c_g \cdot r_2 \cdot (g(t) - x_i(t)), \quad (4)$$

$$x_i(t+1) = x_i(t) + v_i(t+1), \quad (5)$$

where  $v_i(t)$  represents the velocity of the  $i^{\text{th}}$  particle at time  $t$ ;  $x_i(t)$  represents the position of the  $i^{\text{th}}$  particle at time  $t$ ;  $c_p, c_g$  represent the acceleration coefficients, where  $c_p$  is the cognitive learning factor and  $c_g$  is the social learning factor, respectively, representing the self-learning ability of particles

and the ability to learn from the optimal individual of the group,  $c_p, c_g > 0$ ;  $r_1, r_2$  represent random numbers with  $(0, 1)$  interval uniform distribution;  $p_i(t)$  represents the historical optimal position of the  $i^{\text{th}}$  particle at time  $t$ ; and  $g(t)$  represents global optimal position of particle swarm optimization at time  $t$ .

In order to further optimize the performance of the PSO algorithm, Shi introduced a new parameter inertia weight [27] into the particle velocity update formula of the original PSO algorithm, and equation (4) becomes

$$v_i(t+1) = w \cdot v_i(t) + c_p \cdot r_1 \cdot (p_i(t) - x_i(t)) + c_g \cdot r_2 \cdot (g(t) - x_i(t)). \quad (6)$$

Inertia weight determines the influence of particle velocity at the previous time on the current velocity, which can effectively balance the role of global search and local search. Equation (5) consists of three parts. The first part is inertial motion, which indicates the degree to which the particle

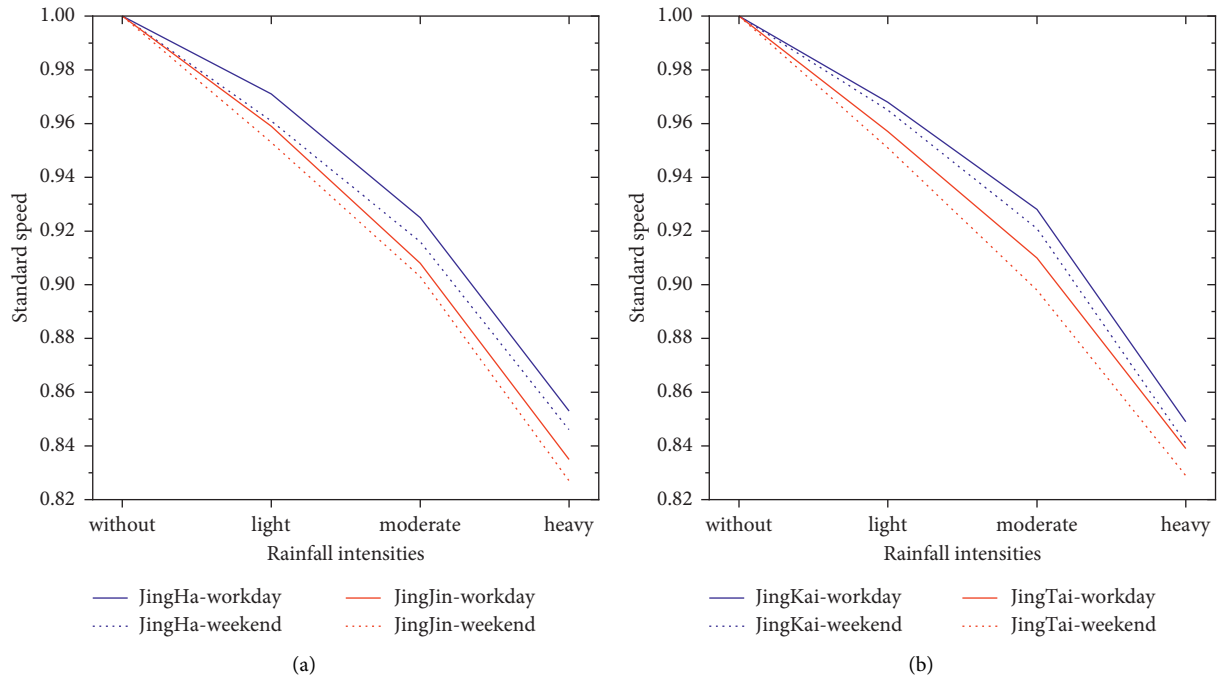


FIGURE 2: Standard speed distribution of highway under different rainfall intensities: (a) highway in Tongzhou District; (b) highway in Daxing District.

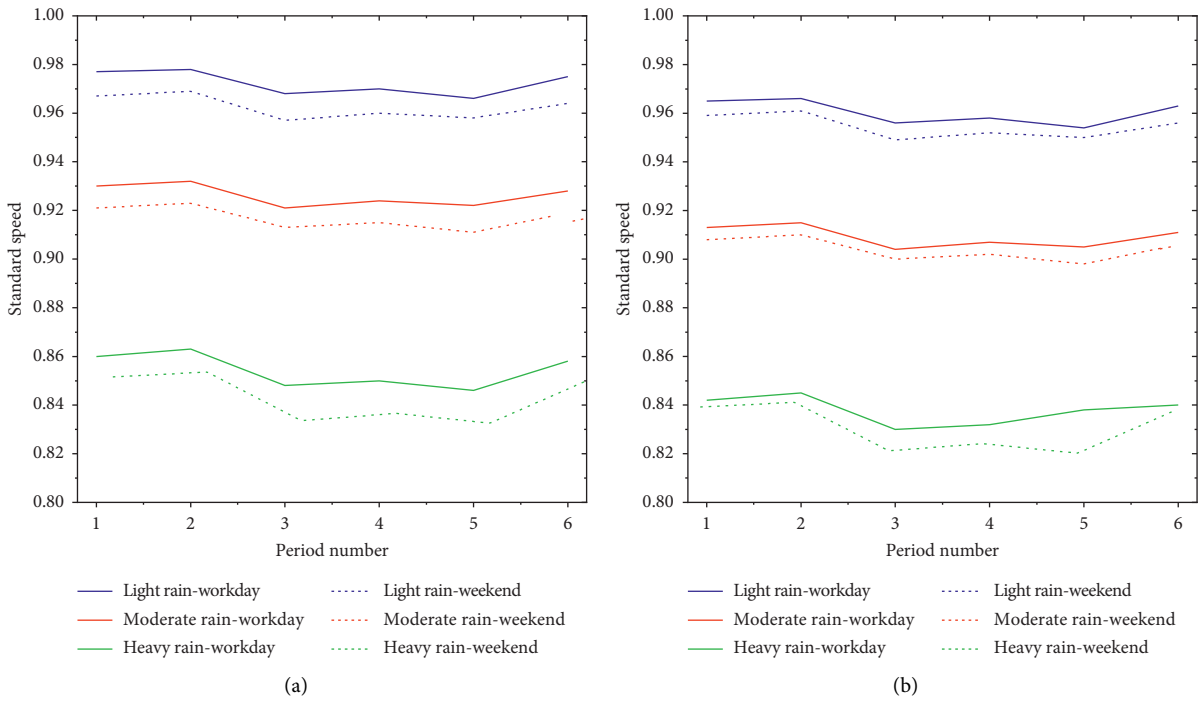


FIGURE 3: Continued.

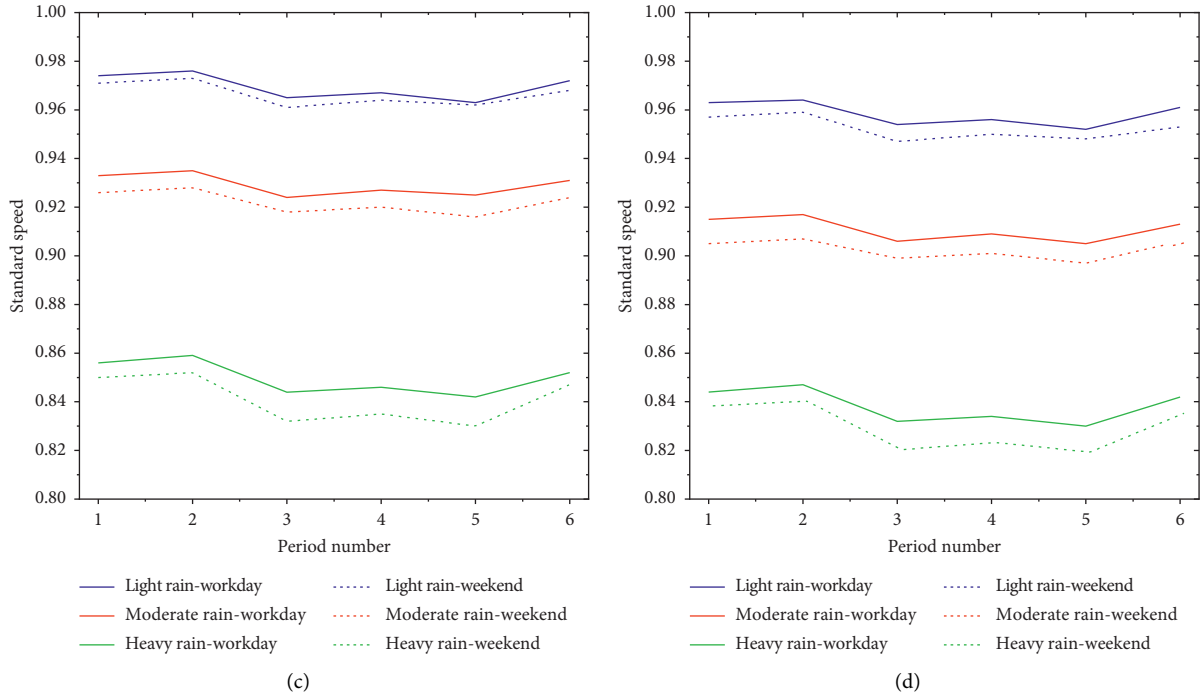


FIGURE 3: Time distribution characteristics of standard speed of highway in rainy environment: (a) Beijing-Harbin Highway; (b) Beijing-Tianjin Highway; (c) Beijing-Kaifeng Highway; (d) Beijing-Taipei Highway.

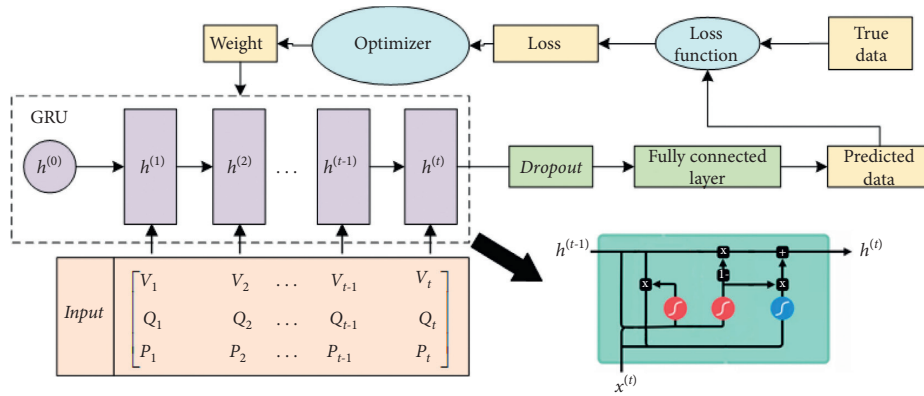


FIGURE 4: Structure of the GRU prediction model.

maintains its velocity at the previous moment; the second is cognitive learning, which means that the particles memorize their historical optimal position and make them close to the historical optimal position. Finally, social learning, which means the information exchange between particles, makes particles close to the historical optimal position of the population [28]. It can be seen from equation (4) that particles have memory, and they move towards the direction of the optimal particle combined with their own and group experience. Equations (5) and (6) constitute a new PSO algorithm called standard PSO algorithm (SPSO).

4.3. Adaptive Nonlinear Inertia Weight PSO. In order to further improve the problem that PSO algorithm is easy to

fall into local optimal solution and reasonably balance the ability of local search and global search of PSO algorithm, an adaptive nonlinear inertia weight method is used to adjust, as shown in the following equation:

$$w = \begin{cases} w_{\min} + \frac{(w_{\max} - w_{\min}) * (f_i - f_{\min})}{f_{\arg} - f_{\min}}, & f_i \leq f_{\arg}, \\ w_{\max}, & f_i > f_{\arg}, \end{cases} \quad (7)$$

where  $w_{\max}$ ,  $w_{\min}$  are the maximum and minimum values of inertia weight;  $f_i$  represents the adaptation value of particle  $i$ ;  $f_{\min}$  represents the minimum fitness of all current

particles;  $f_{\max}$  represents the maximum fitness of all current particles; and  $f_{\arg}$  represents the average fitness value of all current particles.

**4.4. Establishment of APSO-GRU Prediction Model.** Based on the adaptive nonlinear inertia weight PSO algorithm proposed in the previous section, an APSO-GRU prediction model is proposed, and the parameters of GRU network are optimized by PSO algorithm. It is finally determined that the number of hidden layers of the model is 2, the number of nodes in the first layer is 20, and the number of nodes in the second layer is 15 through cross combination test. And the relevant parameters of the APSO are set as follows: the population number is 20, the number of iterations is 30, the acceleration coefficients are both 2, and the maximum and minimum inertia weights are 0.9 and 0.1, respectively,  $[h_{1\min}, h_{1\max}]$  and  $[h_{2\min}, h_{2\max}]$  both are  $[1, 50]$ ,  $[\eta_{\min}, \eta_{\max}]$  is  $[0.001, 0.01]$ , and  $[n_{\min}, n_{\max}]$  is  $[1, 500]$ . The model flow is shown in Figure 5.

## 5. Instance Verification

**5.1. Experimental Verification Scenario and Evaluation Indexes.** In the field of traffic flow prediction, the most common loss functions include mean absolute error (MAE), root mean square error (RMSE), and mean absolute percentage error (MAPE) [29]. Their calculations are shown in equations (8)–(10). RMSE and MAPE are used as error functions to evaluate the performance of the prediction model.

$$\text{MAE} = \frac{1}{L} \sum_{t=1}^L |Y(t) - Y_p(t)|, \quad (8)$$

$$\text{RMSE} = \sqrt{\frac{1}{L} \sum_{t=1}^L (Y_p(t) - Y(t))^2}, \quad (9)$$

$$\text{MAPE} = \frac{1}{L} \sum_{t=1}^L \frac{|Y_p(t) - Y(t)|}{Y(t)} \times 100\%, \quad (10)$$

where  $L$  represents the total length of time series;  $Y_p(t)$  represents prediction value at time  $t$ ; and  $Y(t)$  represents true value at time  $t$ .

Based on the established model of the traffic flow speed prediction of GRU and APSO-GRU, the traffic flow speed of Beijing-Harbin Highway, Beijing-Tianjin Highway, Beijing-Taiwan Highway, and Beijing-Kaifeng Highway in Beijing is predicted, respectively. At the same time, the support vector regression (SVR) model is compared with the APSO-LSTM model on prediction performance.

In order to more comprehensively and accurately evaluate the performance of the model under different rainfall scenarios, the rainy environment is divided into two categories: noncontinuous rainfall and continuous rainfall. Noncontinuous rainfall is the situation that the rainfall process is relatively short and sparse in a specific period of

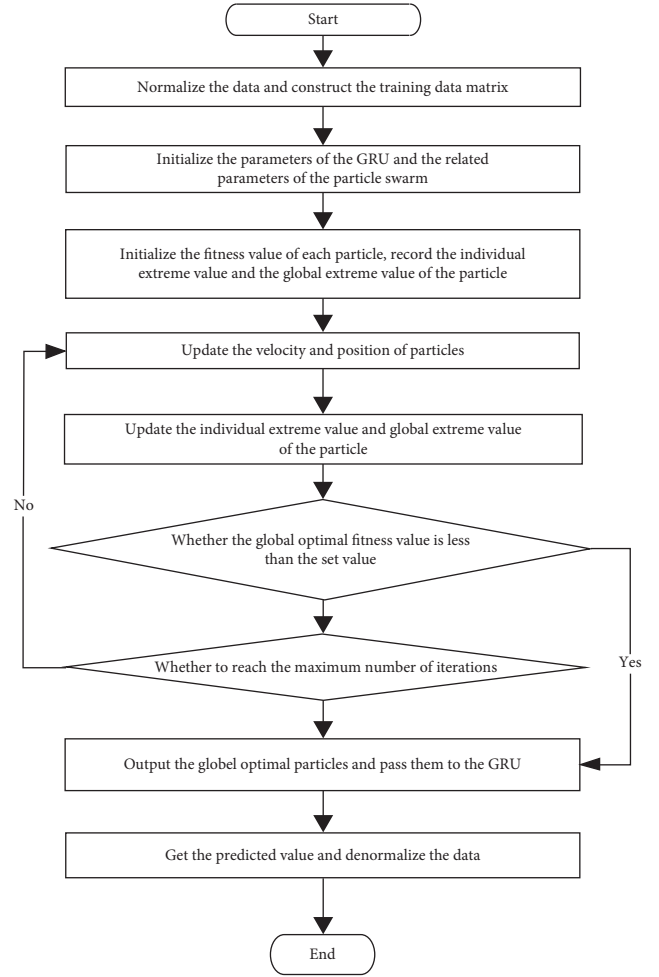


FIGURE 5: Process of the APSO-GRU prediction model.

time, while continuous rainfall is the situation that the rainfall process is relatively long and dense in a specific period of time.

### 5.2. Analysis of Prediction Results of APSO-GRU

**5.2.1. Prediction Results of Noncontinuous Rainfall.** The time span of noncontinuous rainfall in Tongzhou District is the interval between 0:00 on 9th August (working day) and 24:00 on 10th August (weekend) in 2019, and the time span of noncontinuous rainfall in Daxing District is from 0:00 on 11th August (weekend) to 24:00 on 12th August (working day) in 2019.

Due to article content limitation, this paper only selects to visualize prediction results of Beijing-Harbin Highway, as shown in Figure 6. The horizontal axis represents the time, and the left vertical axis represents the speed, corresponding to the broken line chart. The right vertical axis represents precipitation, corresponding to the histogram. This setting is used in the following.

Under the noncontinuous rainfall scenario, the traffic flow speed of highway is obviously disturbed during the rainfall (moderate rain and heavy rain), and the change



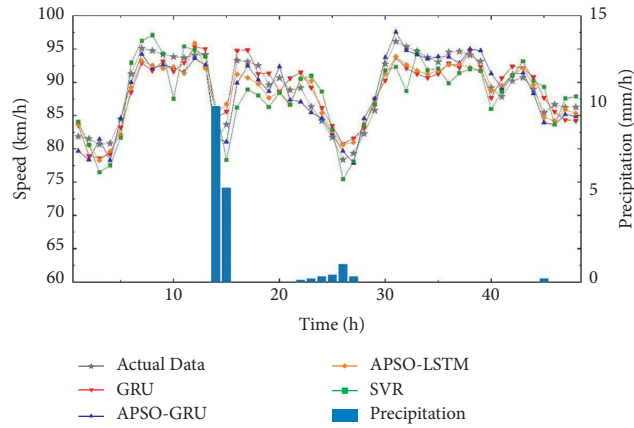


FIGURE 6: Prediction results of traffic flow speed during unsustained rainfall on Beijing-Harbin Highway.

TABLE 5: Errors in traffic flow speed prediction during unsustained rainfall.

| Model     | Beijing-Harbin Highway |      | Beijing-Tianjin Highway |      | Beijing-Kaifeng Highway |      | Beijing-Taiwan Highway |      |
|-----------|------------------------|------|-------------------------|------|-------------------------|------|------------------------|------|
|           | MAPE (%)               | RMSE | MAPE (%)                | RMSE | MAPE (%)                | RMSE | MAPE (%)               | RMSE |
| GRU       | 3.78                   | 2.91 | 4.15                    | 3.22 | 4.04                    | 3.07 | 4.20                   | 3.13 |
| APSO-GRU  | 2.84                   | 2.15 | 3.27                    | 2.76 | 2.53                    | 1.80 | 2.71                   | 2.09 |
| APSO-LSTM | 3.17                   | 2.68 | 3.77                    | 2.93 | 4.23                    | 3.26 | 3.13                   | 2.75 |
| SVR       | 6.02                   | 4.63 | 5.89                    | 4.39 | 6.21                    | 4.78 | 4.87                   | 3.89 |

trend of the predicted value of each model is basically consistent with the real value. The prediction error is shown in Table 5.

5.2.2. *Prediction Results of Continuous Rainfall.* The time span of continuous rainfall in Tongzhou District and Daxing District is from 0:00 on 28th July 2019 (weekend) to 24:00 on 29th July 2019 (working day).

The prediction results of traffic flow speed of Beijing-Harbin Highway during continuous rainfall are shown in Figure 7.

It can be seen from the figure that under the continuous rainfall scenario, the traffic flow speed of highway is greatly affected, and the operation state continuously fluctuates. In this unstable situation, the trend of the predicted values of each model is basically consistent with the real values, but the prediction accuracy is significantly lower than that of noncontinuous rainfall. The prediction error is shown in Table 6.

5.2.3. *Comparative Evaluation and Analysis of Prediction Results.* The statistics of the average prediction error of the traffic flow speed of each model under different rainfall scenarios on each highway can be seen in Table 7.

In the aspect of traffic flow speed prediction, the prediction accuracy of the APSO-GRU model is better than that of the GRU model and APSO-LSTM model under the two rainfall scenarios, and the accuracy of the SVR model is the lowest, which verifies the performance improvement of the

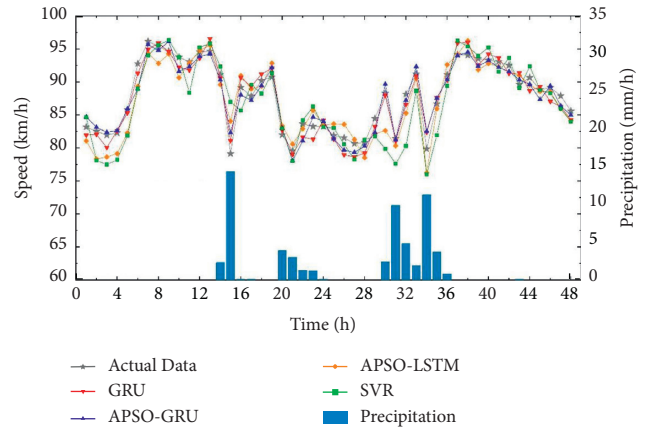


FIGURE 7: Prediction results of traffic flow speed during sustained rainfall on Beijing-Harbin Highway.

built deep learning model. The average prediction accuracy of the GRU model, APSO-GRU model, APSO-LSTM model, and SVR model is 95.96%, 97.16%, 96.42%, and 94.25%, respectively. The average prediction accuracy of the APSO-GRU model is 1.20% higher than that of the GRU model and 0.74% higher than that of the APSO-LSTM model.

The average prediction accuracy of the APSO-GRU model is 96.75% under the continuous rainfall scenario, which is 2.38% and 2.22% higher than that of the GRU model and APSO-LSTM model, respectively. The prediction accuracy of each model has declined, but the decline of the APSO-GRU model is not obvious, followed by the GRU

TABLE 6: Errors in traffic flow speed prediction during sustained rainfall.

| Model     | Beijing-Harbin Highway |      | Beijing-Tianjin Highway |      | Beijing-Kaifeng Highway |      | Beijing-Taiwan Highway |      |
|-----------|------------------------|------|-------------------------|------|-------------------------|------|------------------------|------|
|           | MAPE (%)               | RMSE | MAPE (%)                | RMSE | MAPE (%)                | RMSE | MAPE (%)               | RMSE |
| GRU       | 5.24                   | 3.77 | 6.19                    | 4.40 | 5.77                    | 4.12 | 5.31                   | 3.86 |
| APSO-GRU  | 3.52                   | 2.86 | 3.47                    | 2.83 | 2.88                    | 2.17 | 3.14                   | 2.31 |
| APSO-LSTM | 5.63                   | 4.56 | 4.69                    | 4.16 | 6.30                    | 5.10 | 5.24                   | 4.43 |
| SVR       | 10.37                  | 7.29 | 11.40                   | 7.71 | 12.63                   | 8.02 | 10.71                  | 7.55 |

TABLE 7: Error comparison of traffic flow speed prediction results.

| Model     | Noncontinuous |      | Continuous |      |
|-----------|---------------|------|------------|------|
|           | MAPE (%)      | RMSE | MAPE (%)   | RMSE |
| GRU       | 4.04          | 3.08 | 5.63       | 4.04 |
| APSO-GRU  | 2.84          | 2.20 | 3.25       | 2.54 |
| APSO-LSTM | 3.58          | 2.91 | 5.47       | 4.58 |
| SVR       | 5.75          | 4.42 | 11.28      | 7.64 |

model. The prediction accuracy of the APSO-LSTM model and SVR model is lower than that of the former two models.

## 6. Conclusions

The main conclusions obtained in this paper are as follows:

- (1) Based on the results of the multivariate analysis of variance, rainfall intensity, date category, time of day, and number of lanes have significant effects on traffic flow speed. The higher the intensity of rainfall is, the more the traffic flow is affected. Traffic flow is more likely to be affected by rainfall on weekends than weekdays, and it is more likely to be affected by rainfall during daytime (especially AM peak and PM peak) than at night.
- (2) An APSO-GRU traffic flow speed prediction model was built for the rainy environment. Under the noncontinuous rainfall scenario, the average prediction accuracy of the APSO-GRU model reaches 97.33%, which is 1.19% and 0.71% higher than that of the GRU model and the APSO-LSTM model, respectively. Under the continuous rainfall scenario, the average prediction accuracy of the APSO-GRU model reaches 96.74%, which is 2.69% and 2.39% higher than that of the GRU model and the APSO-LSTM model, respectively. The results show that the prediction accuracy and stability of the APSO-GRU model are significantly improved compared with the APSO-LSTM model under different rainfall scenarios.
- (3) Comparison of the traffic flow speed prediction results between the machine learning model SVR and the APSO-LSTM model in deep learning shows that the prediction accuracy of APSO-LSTM is higher than that of the SVR model by 2.18% and 5.55% under noncontinuous rainfall and continuous rainfall scenarios, respectively. It indicates that the prediction accuracy and stability of the model based on LSTM are better than those of the SVR model,

which fully proves that the prediction performance of the deep learning model is better than the traditional SVR model.

## Data Availability

The data used to support the findings of this study have not been made available because of data ownership issues.

## Conflicts of Interest

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

## Acknowledgments

This research was funded by the National Natural Science Foundation of China (71871011) and the Science and Technology Project of Hebei Provincial Department of Transportation, China (TH-201921).

## References

- [1] K. M. Saberi and R. Bertini, "Empirical analysis of the effects of rain on measured freeway traffic parameters," in *Proceedings of the 89th Annual Meeting of the Transportation Research Board*, Washington DC, USA, January 2010.
- [2] D. Akin, V. P. Sisiopiku, and A. Skabardonis, "Impacts of weather on traffic flow characteristics of urban freeways in Istanbul," *Procedia—Social and Behavioral Sciences*, vol. 16, pp. 89–99, 2011.
- [3] Y. J. Stephanedes, P. G. Michalopoulos, and R. A. Plum, "Improved estimation of traffic flow for real-time control (discussion and closure)," in *Proceedings of the 60th Annual Meeting of the Transportation Research Board*, pp. 28–39, Washington, DC, USA, January 1981.
- [4] 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.
- [5] Y. Tian, "Research on short-term traffic status prediction of freeway network," Shanghai Jiaotong University, Shanghai, China, 2016.

- [6] J. Tang, F. Liu, Y. Zou, W. Zhang, and Y. Wang, "An improved fuzzy neural network for traffic speed prediction considering periodic characteristic," *IEEE Transactions on Intelligent Transportation Systems*, vol. 18, no. 9, pp. 2340–2350, 2017.
- [7] J. Tang and Y. Zeng, "Spatiotemporal gated graph attention network for urban traffic flow prediction based on license plate recognition data," *Computer-Aided Civil and Infrastructure Engineering*, pp. 1–21, 2021.
- [8] X. Meng, H. Fu, L. Peng et al., "D-LSTM: short-term road traffic speed prediction model based on GPS positioning data," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–13, 2020.
- [9] A. T. Ibrahim and F. L. Hall, "Effect of adverse weather conditions on speed-flow-occupancy relationships," *Transportation Research Record Journal of the Transportation Research Board*, vol. 1457, pp. 184–191, 1994.
- [10] M. Agarwal, T. H. Maze, and R. R. Souleyrette, "Impacts of weather on urban freeway traffic flow characteristics and facility capacity," in *Proceedings of the 2005 Mid-Continent Transportation Research Symposium*, pp. 1121–1134, Ames, IA, USA, August 2005.
- [11] C. Li, "Research on traffic flow characteristics and guidance control of highway in bad weather," Beijing University of Technology, Beijing, China, 2015.
- [12] F. Xu, "Calculation and analysis of Urban road capacity under rainfall condition," Chang'an University, Xi'an, China, 2017.
- [13] C. Dong, C. Shao, Z. Xiong et al., "Short-term traffic flow forecasting of road network based on elman neural network," *Journal of Transportation Systems Engineering and Information Technology*, no. 1, pp. 149–155, 2010.
- [14] M. Castro-Neto, Y. Jeong, M. Jeong, and L. Han, "Online-SVR for short-term traffic flow prediction under typical and atypical traffic conditions," *Expert Systems with Applications*, vol. 36, no. 3, pp. 6164–6173, 2009.
- [15] Y.-S. Jeong, Y.-J. Byon, M. M. Castro-neto, and S. M. Easa, "Supervised weighting-online learning algorithm for short-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1700–1707, 2013.
- [16] B. L. Smith and M. Demetsky, "Short-term traffic flow prediction models—a comparison of neural network and nonparametric regression approaches," in *Proceedings of the Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, San Antonio, TX, USA, August 2002.
- [17] Y. Cai, B. Yue, E. Cai et al., "Short-term traffic flow forecast of highway under heavy rain," *Computer Engineering*, vol. 46, no. 6, pp. 34–39, 2020.
- [18] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic flow prediction with big data: a deep learning approach," *Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865–873, 2015.
- [19] M. Zhang and X. Wang, "Traffic time prediction of Urban main road based on GRU-RNN model," *Journal of Beijing Information Science & Technology University*, vol. 34, no. 4, pp. 30–35, 2019.
- [20] Z. Wang, D. Li, and X. Cui, "Travel time prediction based on LSTM neural network in precipitation," *Journal of Transportation Systems Engineering and Information Technology*, vol. 20, no. 1, pp. 137–144, 2020.
- [21] L. Meng, "Research on traffic safety assessment and early warning management of highway on rainy days based on speed prediction," Beijing Jiaotong University, Beijing, China, 2019.
- [22] E. Chung, "Does weather affect highway capacity," in *Proceedings of the 5th International Symposium on Highway Capacity & Quality of Service*, pp. 139–146, Yokohama, Japan, July 2006.
- [23] Z. Wang, "Highway traffic flow characteristic analysis and prediction in rainy environment," Beijing Jiaotong University, Beijing, China, 2020.
- [24] L. Zheng, H. Huang, C. Zhu, and K. Zhang, "A tensor-based K-nearest neighbors method for traffic speed prediction under data missing," *Transportation Business: Transport Dynamics*, vol. 8, no. 1, pp. 182–199, 2020.
- [25] J. Zhao, Y. Gao, Y. Guo, and Z. Bai, "Travel time prediction of expressway based on multi-dimensional data and PSO-ARMAX model," *Advances in Mechanical Engineering*, vol. 10, no. 2, pp. 1–16, 2018.
- [26] Y. Gao, J. Zhao, Z. Qin, Y. Feng, Z. Yang, and B. Jia, "Traffic speed forecast in adjacent region between highway and urban expressway: based on MFD and GRU model," *Journal of Advanced Transportation*, vol. 2020, Article ID 8897325, 18 pages, 2020.
- [27] Y. Shi, "A modified particle swarm optimizer," in *Proceedings of the 1998 IEEE International Conference on Evolutionary Computation*, pp. 67–73, IEEE, Anchorage, AK, USA, May 1998.
- [28] L. Kang, "Several improvements to particle swarm optimization and their theoretical fundamentals," Wuhan University, Wuhan, China, 2017.
- [29] G. Qi, A. Ceder, A. Huang, and W. Guan, "A methodology to attain public transit origin-destination mobility patterns using multi-layered mesoscopic analysis," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–19, 2020.