

Research Article

Evaluation of Short-Term Freeway Speed Prediction Based on Periodic Analysis Using Statistical Models and Machine Learning Models

Xiaoxue Yang,¹ Yajie Zou ,¹ Jinjun Tang ,² Jian Liang,² and Muhammad Ijaz¹

¹Key Laboratory of Road and Traffic Engineering of Ministry of Education, Tongji University, Shanghai 201804, China

²School of Traffic and Transportation Engineering, Smart Transport Key Laboratory of Hunan Province, Central South University, Changsha, China

Correspondence should be addressed to Yajie Zou; yajiezou@hotmail.com and Jinjun Tang; jinjuntang@csu.edu.cn

Received 7 June 2019; Revised 6 August 2019; Accepted 30 August 2019; Published 20 January 2020

Academic Editor: Filomena Mauriello

Copyright © 2020 Xiaoxue Yang 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.

Accurate prediction of traffic information (i.e., traffic flow, travel time, traffic speed, etc.) is a key component of Intelligent Transportation System (ITS). Traffic speed is an important indicator to evaluate traffic efficiency. Up to date, although a few studies have considered the periodic feature in traffic prediction, very few studies comprehensively evaluate the impact of periodic component on statistical and machine learning prediction models. This paper selects several representative statistical models and machine learning models to analyze the influence of periodic component on short-term speed prediction under different scenarios: (1) multi-horizon ahead prediction (5, 15, 30, 60 minutes ahead predictions), (2) with and without periodic component, (3) two data aggregation levels (5-minute and 15-minute), (4) peak hours and off-peak hours. Specifically, three statistical models (i.e., space time (ST) model, vector autoregressive (VAR) model, autoregressive integrated moving average (ARIMA) model) and three machine learning approaches (i.e., support vector machines (SVM) model, multi-layer perceptron (MLP) model, recurrent neural network (RNN) model) are developed and examined. Furthermore, the periodic features of the speed data are considered via a hybrid prediction method, which assumes that the data consist of two components: a periodic component and a residual component. The periodic component is described by a trigonometric regression function, and the residual component is modeled by the statistical models or the machine learning approaches. The important conclusions can be summarized as follows: (1) the multi-step ahead prediction accuracy improves when considering the periodic component of speed data for both three statistical models and three machine learning models, especially in the peak hours; (2) considering the impact of periodic component for all models, the prediction performance improvement gradually becomes larger as the time step increases; (3) under the same prediction horizon, the prediction performance of all models for 15-minute speed data is generally better than that for 5-minute speed data. Overall, the findings in this paper suggest that the proposed hybrid prediction approach is effective for both statistical and machine learning models in short-term speed prediction.

1. Introduction

To alleviate the traffic congestion in large cities, it is particularly important to make full use of existing infrastructure resources such as the application of Intelligent Transportation System (ITS) [1–10]. Real-time and accurate prediction of traffic parameter, such as traffic flow, travel time, and travel speed, is an important input of ITS. Advanced Traveller Information System (ATIS) and Advanced Traffic Management System (ATMS) are essential parts of ITS, while dynamic traffic

assignment (DTA) is a significant task for the operation of ATIS and ATMS. For the purpose of DTA, traffic flow and travel time were estimated and predicted to describe the traffic conditions in DynaMIT (Dynamic Network Assignment for the Management of Information to Travelers) system [10]. When traffic incidents occurred, the predicted travel time was used to evaluate the performance of the application of ITS based on DTA [11]. In particular, evacuation time was predicted to analyze the effectiveness of ITS for evacuation purpose [12]. Using the predicted information, road users can re-plan reasonable

travel modes and routes before traveling, and further adjust the route to improve travel efficiency. With rigorous structure and strong computational ability, a good prediction algorithm can usually capture all kinds of characteristics in traffic data. Up to date, lots of short-term traffic prediction methods have been proposed. Vlahogianni et al. [2, 3] provided a summary of existing short-term traffic methodological approaches until 2013 and divided them into two types: parametric methods and non-parametric methods. Van Lint and Van Hinsbergen [4] summarized the application of Neural Network in short-term traffic prediction. So far, traffic prediction models can be generally categorized into following several types: statistical methods, machine learning methods and hybrid methods.

Due to good theoretical interpretation ability and clear computational structure, statistical methods are widely applied to short-term traffic prediction. The conventional vector autoregressive (VAR) model and the autoregressive integrated moving average (ARIMA) model have been widely utilized in traffic prediction [13–16]. To improve the applicability of the conventional ARIMA model, Kumar and Vanajakshi [17] developed a seasonal ARIMA (SARIMA) model. Zou et al. [16] used space time (ST) model to describe both time and space correlations of traffic. A spatio-temporal autoregressive moving average (STARIMA) model was proposed to utilize upstream volumes in the current moment to predict traffic condition [18]. Yang et al. [19] found that the spatial traffic information from upstream and downstream road segments can improve the prediction performance. Recently, some new models have also been developed. For example, Zhang et al. [20] proposed Granger causality to predict travel time and obtained better prediction performance. Agafonov and Yumaganov [21] presented a distributed model for short-term traffic flow prediction based on the k -nearest neighbors method.

In addition to statistical methods, the machine learning methods have been widely used in traffic prediction due to its strong generalization ability, learning ability and adaptability. Neural network is commonly used in traffic flow prediction [22]. Recurrent neural network (RNN) has been widely adopted for modeling nonlinear time series data because of its short term memory [23]. Recently, some researchers found that state-space neural network model [24] and long short-term memory neural network (LSTM) [25], which were improved based on the neural network, showed better prediction performance with high computational efficiency. Tang et al. [26] introduced an improved fuzzy-neural networks (FNN) to enhance traffic flow prediction accuracy. Dimitriou et al. [27] optimized the parameters of the adaptive fuzzy rule system using genetic algorithm. Lv et al. [28] proposed a deep-learning-based method using autoencoders to predict traffic flow. To capture the nonlinear spatial and temporal effect of traffic flow, Polson and Sokolov [29] introduced a deep learning approach combining a linear method. Based on mathematical methods and optimization techniques, support vector machine (SVM) transforms variables into a high dimensional space, and creates a hyperplane with maximum spacing. SVM contains two branches: support vector regression (SVR) and support vector classification (SVC). Furthermore, SVR model has a common application in traffic prediction [30, 31]. To speed up the parameter optimizing, the least square support vector machine (LS-SVMs) was

used to predict traffic [32]. Jeong et al. [33] proposed an online learning weighted support-vector regression (OLWSVR) model to implement real-time traffic flow prediction. The Kaman filter theory (KFT), an algorithm for optimal estimation of system state using state equation of linear system, was used to predict traffic flow and travel time [34, 35]. Wang et al. [36] proposed an improved extended Kalman filter (EKF) for travel time prediction. Guo et al. [37] introduced an adaptive Kalman filter (AKF) to predict unstable traffic flow.

Hybrid models combining the advantages of different methods are proposed to improve prediction accuracy. For example, the conditional probability theory and Bayesian rule were combined with ANN [38], and the statistical methods and heuristic models were combined with SVM [39]. Yanchong et al. [40] proposed a short-term traffic flow prediction model combining Mallat wavelet and BP neural network. Tang et al. [41] combined fuzzy c -means and Genetic algorithm to predict missing traffic volume data. Huang et al. [42] proposed a traffic flow prediction model based on fuzzy mean clustering (FCM) and advanced neural network (NN). To further improve the prediction performance, Huang et al. [43] developed a deep learning method incorporating a deep belief network (DBN) and a multitask regression layer. Tang et al. [44] introduced a hybrid model by combining double exponential smoothing (DES) and a support vector machine to predict traffic flow. It has been shown that hybrid models have better prediction performance than single method in traffic prediction.

Traffic data usually represent periodic features which help understanding the variation patterns of traffic flow and improving the prediction performance. To demonstrate the periodic pattern, researchers have introduced many prediction models considering cyclicity. To capture the weekly patterns of data, Williams and Hoel [45] applied Seasonal ARIMA model to traffic flow prediction. However, the outlier detection and the parameter estimation of Seasonal ARIMA model are time-consuming. Thus, Hong [46] applied seasonal support vector regression model with chaotic immune algorithm (SSVRCIA), Seasonal ARIMA model and seasonal Holt-Winters model to traffic flow prediction, and concluded that SSVRCIA model performed better than other models. Moreover, Lippi et al. [47] proposed two improved support vector regression models and compared the prediction performance in terms of accuracy and efficiency. Overall, Seasonal ARIMA model was proved to be more accurate; additionally, the new seasonal support vector regressor model performed better in peak hours. In addition, Li et al. [7] compared the prediction performance of hybrid models considering the periodicity of traffic time series data. And a frequently used strategy is the combination of prediction models and detrending methods [48–50]. For example, Dendrinis [51] divided the traffic-flow time series into two parts: periodic part and nonperiodic part, and took the nonperiodic part as the research focus. Some researchers modeled periodic components using spectral analysis technique [52, 53]. Zhang et al. [49] developed a hybrid traffic prediction method, which supposed that the traffic data contain three parts: periodic part, deterministic part, and volatility part. Furthermore, Zhang et al. [49] found that multi-step ahead prediction can provide more accurate prediction results.

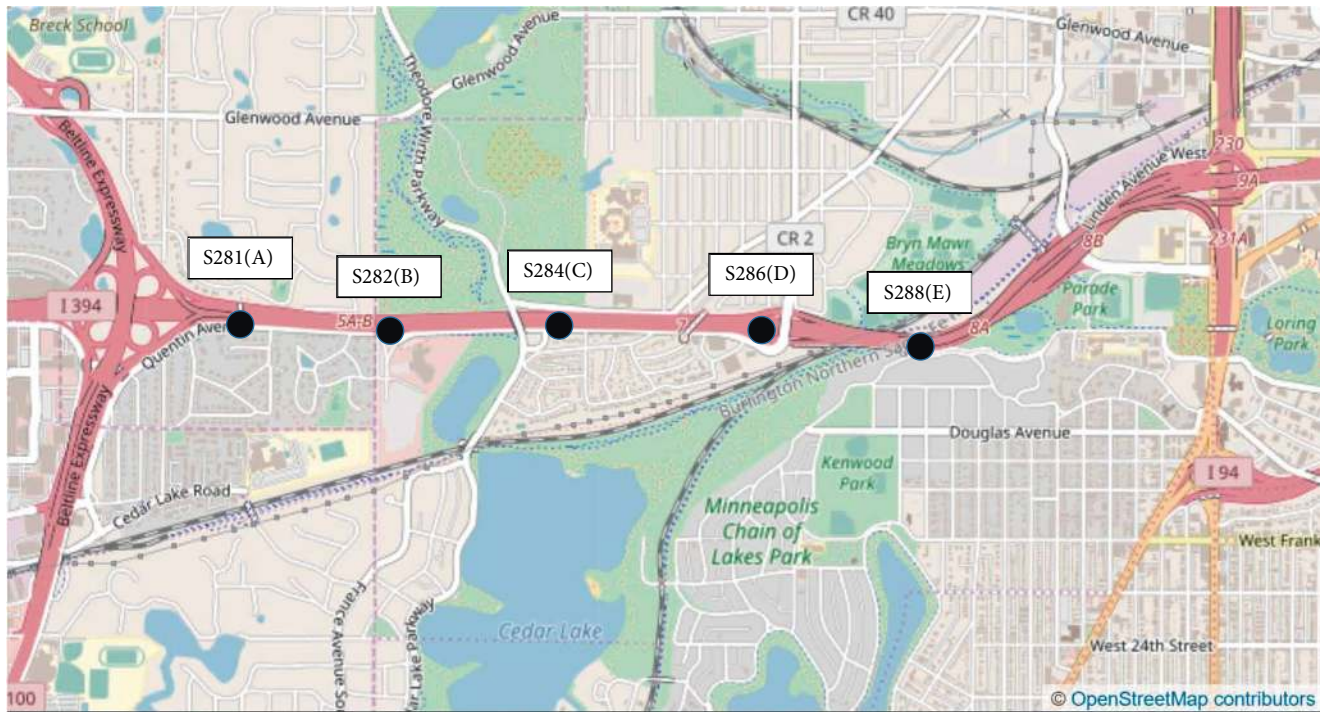


FIGURE 1: Location of I-394 road, from station A to station E.

Although a few studies have considered the periodic feature in traffic prediction, very few studies comprehensively evaluate the impact of periodic component on statistical methods and machine learning methods. Focusing on speed prediction, the specific research objectives are: (1) to evaluate the effectiveness of hybrid methods based on three statistical models (i.e., ST, VAR, ARIMA) and three machine learning models (i.e., SVM, Multi-layer Perceptron (MLP), RNN) in multi-step ahead prediction (5, 15, 30, 60 minutes ahead predictions) considering peak hours and off-peak hours, separately, (2) to compare the prediction performance improvement considering the impact of periodic component for all models, (3) to examine the difference in freeway speed prediction under two different data aggregation levels (5-minute and 15-minute).

The following sections of this paper are described below. The second part provides the data description, which shows the data gathered from five loop detectors of an eastbound road of Interstate 394 freeway stretch, Minnesota, from November 2017 to April 2018. The third part introduces two main methodologies used in this study: statistical models and machine learning approaches. The next section provides the modeling results. The short-term traffic speed prediction accuracy is evaluated under different scenarios. Finally, the model results are summarized and discussed.

2. Data Description and Preliminary Data Analysis

This study is carried out on an eastbound road of Interstate 394 freeway stretch, Minnesota, when suffering from heavy congestion during the rush hours of morning and afternoon.

In this study, the road equipped with 5 neighboring stations is selected (see Figure 1). The length of selected road is about 1.7 miles. And the distance between two neighboring stations is approximately 0.5 miles. There are 3 lanes for the eastbound direction.

Speed data can be downloaded using the publicly available data tool developed by the Minnesota Department of Transportation. The speed data are collected for 5-minute and 15-minute aggregation level, respectively, 24 hours a day from loop detectors, from November 2017 to April 2018. The data missing rate is less than 0.01%, thus, the data repair method based on historical average is adopted to repair missing data. Specifically, traffic at night or on weekends is usually smooth and free of congestion. Therefore, this study selects the speed data from 6:00 AM to 8:00 PM of weekdays, which contain the morning and afternoon peak hours. Figure 1 provides the location of the selected I-394 road.

Figure 2 provides the distribution of median value of historical speeds on weekdays (Nov. 2017–April 2018) at all stations. From Figure 2, we can see that there are two peak hours for all stations: one is from about 7:00 AM to 9:00 AM and the other is from 3:30 PM to 7:30 PM.

It has been demonstrated that the traffic data of adjacent stations are spatially and temporally correlated [54, 55]. The speed values of stations in the downstream are influenced by the speed values of stations in the upstream. Furthermore, the traffic condition in the downstream also affects the traffic in the upstream because of traffic jam. In this study, cross-correlation function (CCF) is utilized to evaluate the time and spatial correlation of speed.

In this study, we choose station C as the target station. In Figure 3, the cross-correlation consequences between

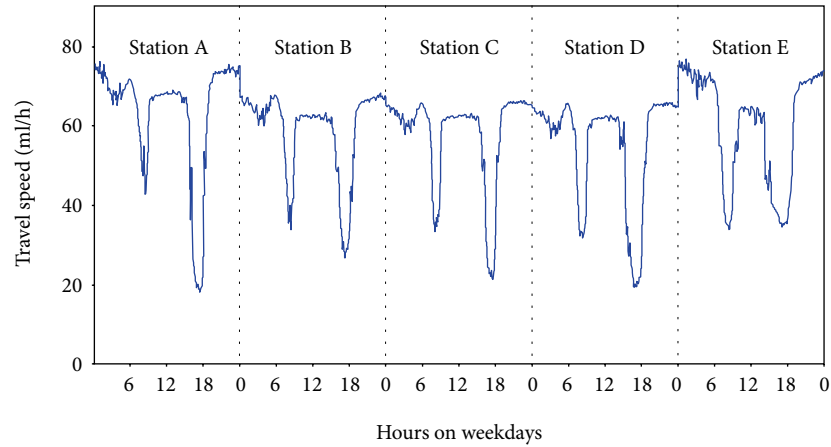


FIGURE 2: Median value of historical speeds at all stations.

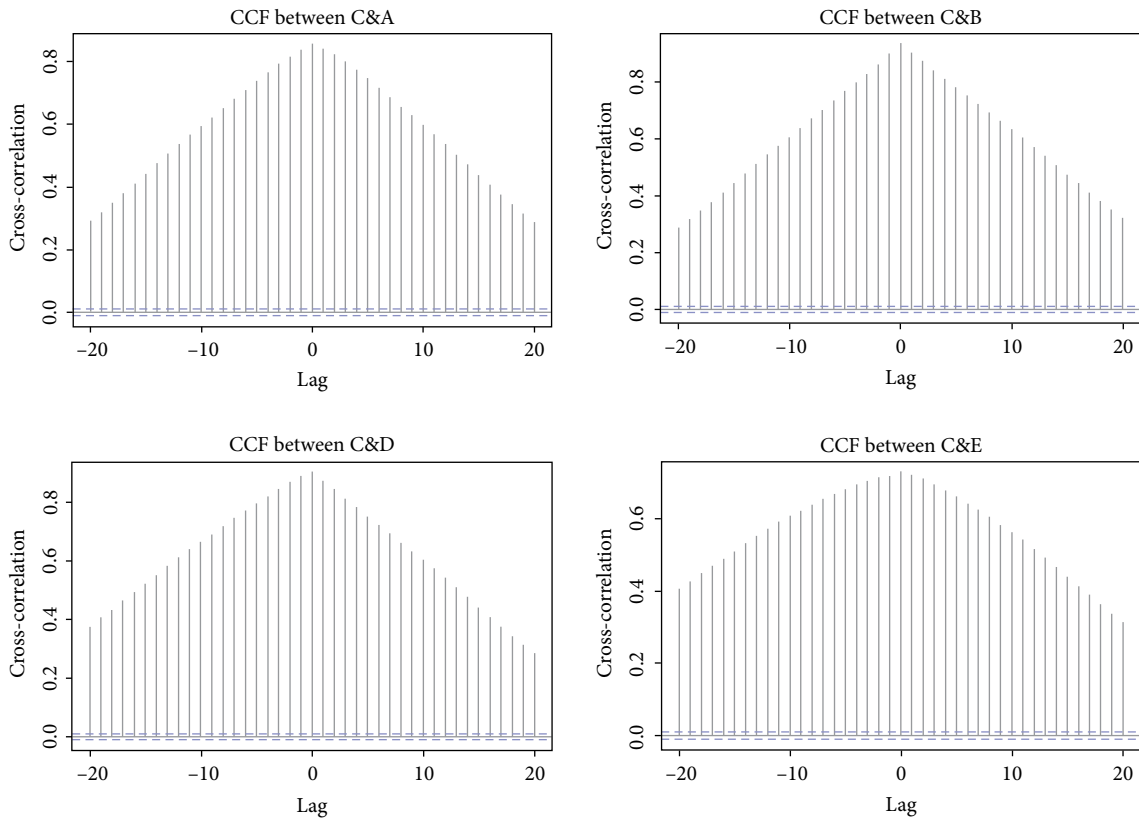


FIGURE 3: Cross-correlation values of speed between station C and other adjacent stations.

station C and other adjacent stations are shown. As provided in Figure 3, absolute value of lag has a steady decreasing effect on the CCF values of speed. More specifically, the cross correlation value between station C and other stations is largest when the lag is equal to 0. Furthermore, when the absolute value of lag equals to 20, the CCF value is as low as 0.3. Based on Figure 3, as the distance between two stations increases, it can be observed that the correlation between two stations decreases, that is, the value presents a downward trend. As shown in Figure 3, 0.8543, 0.9342, 0.9000, and 0.7281 are the maximum CCF values between

station C and stations A, B, D, and E, respectively. The result is reasonable, because the farther away the station is, the less impact it has.

Periodic patterns are another significant feature of traffic speed, except the temporal and spatial correlation above. In Figure 4, the speed distribution of station C during five continuous representative weekdays is shown. It has been shown that the period pattern exists in the traffic parameters [49, 56, 57]. As shown in Figure 4, the speed data show a periodic pattern every 24 hours. Obviously, during the peak hours in the morning and afternoon, the speed values suffer

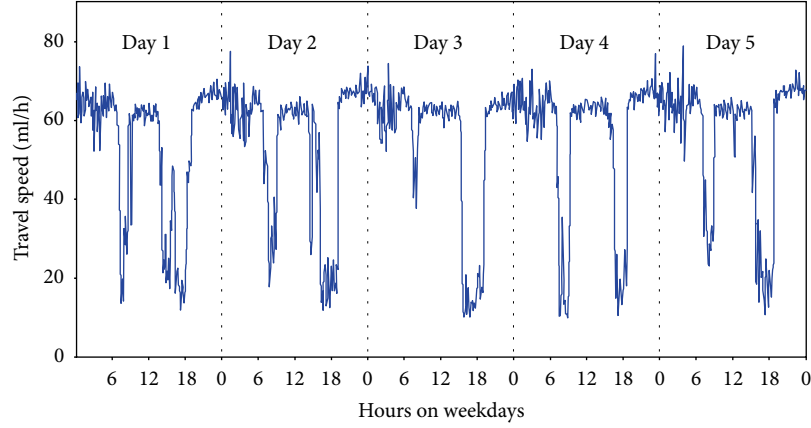


FIGURE 4: Periodic pattern of speed at station C during representative weekdays.

from a significant decline. In the off-peak hours, the speed values suffer from the random fluctuation during the free flow traffic condition. Furthermore, similar periodic patterns have been demonstrated at four other stations. To capture this cyclical pattern, a hybrid prediction method is adopted in the following section.

3. Methodology

In this part, statistical models (i.e., ST, VAR, ARIMA), machine learning models (i.e., SVM, MLP, RNN) and hybrid models are described, respectively. In this study, due to its marked drop in speed over the peak hours, and the accessibility of speed data of loop detectors from the upstream and downstream, station C is chosen as the target station. The target station C is strongly correlated with the adjacent station, thus, the historical speed data of the adjacent station can be used to predict speed of the target station. In the section below, two aggregation levels (5-minute and 15-minute) are considered. $A_{t,p}$, $B_{t,p}$, $C_{t,p}$, $D_{t,p}$ and $E_{t,i}$ are used to represent the speed at stations A, B, C, D, and E. The period, 6:00 AM –8:00 PM on weekdays (Nov. 2017–April 2018), is selected. Multi-horizon ahead prediction is calculated by using the proposed methods, and the prediction horizons are 5 minutes, 15 minutes, 30 minutes, and 1 hour. $C_{t+p,i}$ is the predicted speed at station C, where p denotes the time step, and i represents the aggregation level (for example, $C_{t+1,5}$ represents the 5 minutes ahead prediction (or one-step ahead prediction under the aggregation level of 5-minute) and $C_{t+4,15}$ represents the 60 minutes ahead prediction (or 4-step ahead prediction under the aggregation level of 15-minute)).

3.1. Statistical Models. In this part, three statistical models are described as follows: ST model, VAR model, and ARIMA model.

3.1.1. ST Model. ST model, modeling via the proper probability distribution of data, is introduced to traffic flow and traffic speed prediction [16, 57]. The model contains two kinds of prediction values: point prediction and interval prediction.

In this study, it is supposed that the speed at time $t + p$ under the aggregation level i at station C, $C_{t+p,i}$ is assumed to follow a normal distribution [58], that is, $C_{t+p,i} \sim N(\mu_{t+p,i}, \sigma_{t+p,i}^2)$. $\mu_{t+p,i}$, the mean of normal distribution, is the point prediction of speed. And the quantile, which is used to calculate the prediction interval, is defined below:

$$z_{\alpha,i} = \mu_{t+p,i} + \sigma_{t+p,i} \times \Phi^{-1}[\alpha], \quad (1)$$

where, $p = 1, 2, \dots, 12$; $i = 5, 15$; and, Φ represents the cumulative density function (cdf) of a standard normal distribution.

A linear regression combining the current value and the past value of speed at all stations is used to model $\mu_{t+p,i}$. For instance, when $p = 1$, $i = 5$ (i.e., 5 minutes ahead prediction),

$$\begin{aligned} \mu_{t+1,5} = & \alpha_0 + \alpha_1 C_{t,5} + \alpha_2 C_{t-1,5} + \alpha_3 B_{t,5} + \alpha_4 D_{t,5} + \alpha_5 D_{t-1,5} \\ & + \alpha_6 A_{t,5} + \alpha_7 A_{t-1,5} + \alpha_8 E_{t-1,5}. \end{aligned} \quad (2)$$

Analyzing the speed data from November to December 2017, independent variables for $\mu_{t+p,i}$ (for example, in (2)) are selected. The construction process of $\mu_{t+p,i}$ is described as follows: starting from the simplest model, independent variables are added stepwise until Bayesian information criterion is not further improved [59].

$\sigma_{t+p,p}$, the predictive spread, is fitted by a linear function of the fluctuating value $v_{t,p}$

$$\sigma_{t+p,i} = b_0 + b_1 v_{t,i}, \quad (3)$$

where, coefficients b_0 and b_1 are nonnegative; and, the fluctuating value, which reflects the range of the recent fluctuations of speed, is fitted as follows:

$$\begin{aligned} v_{t,i} = & \left(\frac{1}{10} \sum_{j=0}^1 \left((A_{t-j,i} - A_{t-j-1,i})^2 + (B_{t-j,i} - B_{t-j-1,i})^2 \right. \right. \\ & + (C_{t-j,i} - C_{t-j-1,i})^2 + (D_{t-j,i} - D_{t-j-1,i})^2 \\ & \left. \left. + (E_{t-j,i} - E_{t-j-1,i})^2 \right) \right)^{1/2}. \end{aligned} \quad (4)$$

The range of the recent fluctuations of speed can be reflected by the fluctuating value.

3.1.2. VAR Model. Focusing on interrelated time series, VAR model can capture the effect of the upstream and downstream stations. In this study, a 5-equation VAR(m) model is defined as follows:

$$\mathbf{X}_{t+1,i} = \boldsymbol{\varphi}_0 + \boldsymbol{\varphi}_1 \mathbf{X}_{t,i} + \boldsymbol{\varphi}_2 \mathbf{X}_{t-1,i} + \cdots + \boldsymbol{\varphi}_m \mathbf{X}_{t-m+1,i} + \mathbf{u}_{t+1}, \quad (5)$$

$$\begin{aligned} \mathbf{X}_{t,i} &= \begin{pmatrix} A_{t+1,i} \\ B_{t+1,i} \\ C_{t+1,i} \\ D_{t+1,i} \\ E_{t+1,i} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\varphi}_{10} \\ \boldsymbol{\varphi}_{20} \\ \boldsymbol{\varphi}_{30} \\ \boldsymbol{\varphi}_{40} \\ \boldsymbol{\varphi}_{50} \end{pmatrix} + \begin{pmatrix} \boldsymbol{\varphi}_{11,1} & \boldsymbol{\varphi}_{12,1} & \boldsymbol{\varphi}_{13,1} & \boldsymbol{\varphi}_{14,1} & \boldsymbol{\varphi}_{15,1} \\ \boldsymbol{\varphi}_{21,1} & \boldsymbol{\varphi}_{22,1} & \boldsymbol{\varphi}_{23,1} & \boldsymbol{\varphi}_{24,1} & \boldsymbol{\varphi}_{25,1} \\ \boldsymbol{\varphi}_{31,1} & \boldsymbol{\varphi}_{32,1} & \boldsymbol{\varphi}_{33,1} & \boldsymbol{\varphi}_{34,1} & \boldsymbol{\varphi}_{35,1} \\ \boldsymbol{\varphi}_{41,1} & \boldsymbol{\varphi}_{42,1} & \boldsymbol{\varphi}_{43,1} & \boldsymbol{\varphi}_{44,1} & \boldsymbol{\varphi}_{45,1} \\ \boldsymbol{\varphi}_{51,1} & \boldsymbol{\varphi}_{52,1} & \boldsymbol{\varphi}_{53,1} & \boldsymbol{\varphi}_{54,1} & \boldsymbol{\varphi}_{55,1} \end{pmatrix} \begin{pmatrix} A_{t,i} \\ B_{t,i} \\ C_{t,i} \\ D_{t,i} \\ E_{t,i} \end{pmatrix} \\ &+ \cdots + \begin{pmatrix} \boldsymbol{\varphi}_{11,m} & \boldsymbol{\varphi}_{12,m} & \boldsymbol{\varphi}_{13,m} & \boldsymbol{\varphi}_{14,m} & \boldsymbol{\varphi}_{15,m} \\ \boldsymbol{\varphi}_{21,m} & \boldsymbol{\varphi}_{22,m} & \boldsymbol{\varphi}_{23,m} & \boldsymbol{\varphi}_{24,m} & \boldsymbol{\varphi}_{25,m} \\ \boldsymbol{\varphi}_{31,m} & \boldsymbol{\varphi}_{32,m} & \boldsymbol{\varphi}_{33,m} & \boldsymbol{\varphi}_{34,m} & \boldsymbol{\varphi}_{35,m} \\ \boldsymbol{\varphi}_{41,m} & \boldsymbol{\varphi}_{42,m} & \boldsymbol{\varphi}_{43,m} & \boldsymbol{\varphi}_{44,m} & \boldsymbol{\varphi}_{45,m} \\ \boldsymbol{\varphi}_{51,m} & \boldsymbol{\varphi}_{52,m} & \boldsymbol{\varphi}_{53,m} & \boldsymbol{\varphi}_{54,m} & \boldsymbol{\varphi}_{55,m} \end{pmatrix} \begin{pmatrix} A_{t-m+1,i} \\ B_{t-m+1,i} \\ C_{t-m+1,i} \\ D_{t-m+1,i} \\ E_{t-m+1,i} \end{pmatrix} + \begin{pmatrix} u_{1,t+1} \\ u_{2,t+1} \\ u_{3,t+1} \\ u_{4,t+1} \\ u_{5,t+1} \end{pmatrix}. \end{aligned} \quad (6)$$

Through the evaluation of characteristic polynomial, the stabilization of VAR(m) model can be guaranteed:

$$\det(I_5 - \boldsymbol{\varphi}_1 z - \cdots - \boldsymbol{\varphi}_m z^m) \neq 0 \text{ for } |z| \leq 1. \quad (7)$$

No characteristic roots locate inside the unit circle, which is sufficient and necessary for stability.

3.1.3. ARIMA Model. Unlike VAR model, ARIMA model only takes the effect of time series into account, and has been used in various traffic data analysis [60–62]. A nonseasonal ARIMA model can be defined as follows: ARIMA (p, d, q), where, p is the number of autoregressive terms, d is the number of non-seasonal differences and q is the number of lagged prediction errors.

The autoregressive moving average (ARMA) model can be extended to ARIMA. The mathematical expression of ARMA (p, q) process is shown below:

$$\begin{aligned} (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p) X_t \\ = (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q) Z_t, \end{aligned} \quad (8)$$

The data series are required to be stable in the ARMA model. The time series which are nonstationary should be converted into a stationary series. The data, which can not be modeled by ARMA, can be transformed by ARIMA model to fit an ARMA model. $I(d)$ in ARIMA model means the d th difference of the unstable data.

The mathematical expression of ARIMA (p, d, q) process is shown below:

$$\begin{aligned} (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)(1 - B)^d X_t \\ = (1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q) Z_t, \end{aligned} \quad (9)$$

where, $B^k x_t = x_{t-k}$; Z_t is a Gaussian white noise series whose mean is zero and variance is σ^2 ; and, $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$ and $\theta(z) = 1 + \theta_1 z - \cdots - \theta_q z^q$ are polynomials of p and q , and $\phi(z) \neq 0$ for $|z| \leq 1$. $\boldsymbol{\varphi}_1$ to $\boldsymbol{\varphi}_m$ are matrices of coefficients; and, \mathbf{u}_{t+1} is a corresponding 5×1 independently and identically

where, $\mathbf{X}_{t+1,i} = (A_{t+1,i}, B_{t+1,i}, C_{t+1,i}, D_{t+1,i}, E_{t+1,i})^T$ is a 5×1 vector of variables; $\boldsymbol{\varphi}_0$ is a 5×1 absolute term; $\boldsymbol{\varphi}_1$ to $\boldsymbol{\varphi}_m$ are matrices of coefficients; and, \mathbf{u}_{t+1} is a corresponding 5×1 independently and identically distributed random vector with $E(\mathbf{u}_{t+1}) = 0$ and time invariant positive definite covariance matrix $E(\mathbf{u}_{t+1} \mathbf{u}_{t+1}^T) = \Sigma_{\mathbf{u}}$.

Equation (5) can also be written as follows:

3.2. Machine Learning Models. In this section, we describe three machine learning models: SVM model, MLP model, and RNN model.

3.2.1. SVM Model. SVM transforms vector variables into a higher dimensional space, and creates a hyperplane with maximum spacing. SVM contains two branches: support vector regression (SVR) and support vector classification (SVC). SVC, addressing the classification problems, calculates a decision boundary and maximizes the distance between the boundary and the nearest sample data. Like SVC, SVR uses a similar approach for regression problems and ignores the error which is less than ε between the observed value and the estimated value [60, 63, 64]. More specifically, given a group of training data, the objective is to seek a function $f(x)$ that the maximum deviation between actual values y_i and predicted values is at most ε . For instance, a linear function $f(x) = \omega^T x + b$ is flat if it has a small ω —this can be achieved by minimizing $\|\omega\|^2$. Due to the function that satisfies all the necessary constraints C may not exist, some slack variables (ξ, ξ^*) are introduced to allow for some errors.

The formulation of SVR is defined as follows:

$$\begin{aligned} \text{minimize} \quad & \left\{ \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l (\xi + \xi^*) \right\} \\ \text{subject to} \quad & y_i - \omega^T x_i - b \leq \varepsilon + \xi_i \\ & \omega^T x_i + b - y_i \leq \varepsilon + \xi_i^* + \xi_i^* \\ & \xi_i, \xi_i^* \geq 0, \end{aligned} \quad (10)$$

where, $f(x) = \omega^T x + b$ is the objective function; y_i is actual value; ε is the deviation between actual values and predicted values; and, ξ, ξ^* are slack variables allowing for some errors.

We can also extend SVR to nonlinear problems by combining nonlinear kernel functions. Common kernels include linear kernel and the Radial Basis Function (RBF) kernel, which convert the input sample into a higher dimensional space that results in better separation (for classification) or estimation of values (for regression). In this study, we experimentally choose to use a RBF kernel for SVR because it generally provides better results.

3.2.2. MLP Model. Multilayer perception is the most common Artificial Neural Network (ANN) model [61, 65, 66]. Jiang et al. [67] selected MLP model as one of the candidates to compare the prediction accuracy of different models. A neural network consists of some layers, each of which has one or more neurons. Every neuron is linked with all the neurons in the adjacent layer, while the neurons in the layer are not. Each neuron takes a linear weighted combination of all its input x (from the layer in front of it) and generates output y through a nonlinear activation function:

$$y = \sigma \left(\sum_{i=1}^N (\omega_i x_i) \right). \quad (11)$$

Each of these outputs y is used as an input to the next layer of neurons until the last (i.e., output) layer is reached. The weights correlated with each neuron can be randomly initialized so that each neuron can potentially learn a different function of its input.

And the loss function measures the difference between the estimated output value of the network and the real value of the training data. For regression problems, more specifically, the squared error between the estimate and the actual value is often used as a loss function. The backpropagation algorithm is then utilized to compute the gradient of this error, and the gradient is propagated back through the network (towards the input layer), and the weight of each neuron is updated by gradient descent. The weights associated with each neuron are the parameters that define the neural network model, which are estimated by minimizing a loss function. A random gradient-based optimizer called Adam, which is computationally efficient and extends well to larger data sets, is adopted [68]. Set all parameters of this optimizer to their default values. The rectified linear unit (ReLU) activation function is used in the MLP network.

3.2.3. RNN Model. As shown in Figure 5, RNN is a particular neural network, which consists of at least one feed-back link that serves as the internal state from the neuron's outputs to the inputs. The structure has the ability of time processing and sequence learning. RNN is widely used to process nonlinear time series data because of its short-term memory [25, 69–71]. The calculation formula of RNN is shown as follows:

$$s_t = H(Ux_t + Ws_{t-1} + b_s), \quad (12)$$

$$o_t = Vs_t + b_o, \quad (13)$$

where, $x = (x_1, \dots, x_T)$ is the input vector; $s = (s_1, \dots, s_T)$ is the hidden vector; $o = (o_1, \dots, o_T)$ is the output vector; W , U , and V are the weight matrices; b_o and b_s are the bias vectors; and, H represents the hidden layer function and is usually a sigmoid function.

3.3. Hybrid Models. As shown in Figure 4, the traffic speed usually demonstrates a daily cyclical pattern. Therefore, it is feasible to decompose the speed data into two components: periodic component which demonstrates the cyclical trend in

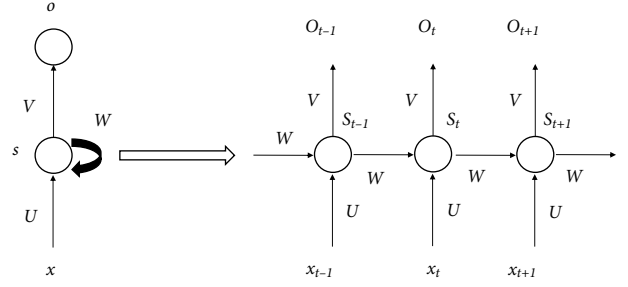


FIGURE 5: RNN network structure.

the weekdays, and a residual component. Therefore, the hybrid prediction methods can be defined as follows:

$$C_t = M_t + C_t^r, \quad (14)$$

where, C_t is the original speed at station C ; M_t is the cyclical trend at station C ; and, C_t^r is the residual component at station C .

A trigonometric regression function, combining sinusoids and cosinusoids, is used to fit the cyclical pattern observed in Figure 4. Adorf [72] applied the trigonometric regression function to capture the periodic patterns when analyzing time series data. Furthermore, to improve the prediction performance for the wind speed, Gneiting et al. [59] adopted the trigonometric functions in the ST models and achieved accurate prediction results.

Taking the aggregation level of 5-minute as an example, the average speed of every station is computed as follows:

$$S_t = \frac{1}{30} \sum_{d=1}^{30} s_t^d, \quad (15)$$

where, S_t is the average speed at time t ; s_t^d is 5-minute average speed at time t on day d ; $t = 1, 2, \dots, 288$; and, $d = 1, 2, \dots, 30$ is the number of days.

The periodic component can be fitted as below:

$$M_u = m_0 + m_1 \sin\left(\frac{2\pi u}{288}\right) + m_2 \cos\left(\frac{2\pi u}{288}\right) + m_3 \sin\left(\frac{4\pi u}{288}\right) + m_4 \cos\left(\frac{4\pi u}{288}\right) + \dots + m_{2n-1} \sin\left(\frac{2n\pi u}{288}\right) + m_{2n} \cos\left(\frac{2n\pi u}{288}\right), \quad (16)$$

where, $u = 1, 2, \dots, 288$; n is the number of trigonometric polynomials.

Previous study [16] has examined different number of trigonometric functions on speed prediction and found that the prediction accuracy improves slightly after n reaches 15. Thus, we set n to 15 in this study.

The residual component is fitted by the ST, VAR, ARIMA, SVM, MLP, and RNN models. Statistical prediction models (i.e., ST, HST (hybrid ST), VAR, HVAR (hybrid VAR), ARIMA, and HARIMA (hybrid ARIMA)) are estimated in the Software R, and machine learning models (i.e., SVM, HSVM (hybrid SVM), MLP, HMLP (hybrid MLP), RNN, and HRNN (hybrid RNN)) are estimated in the Software Python.

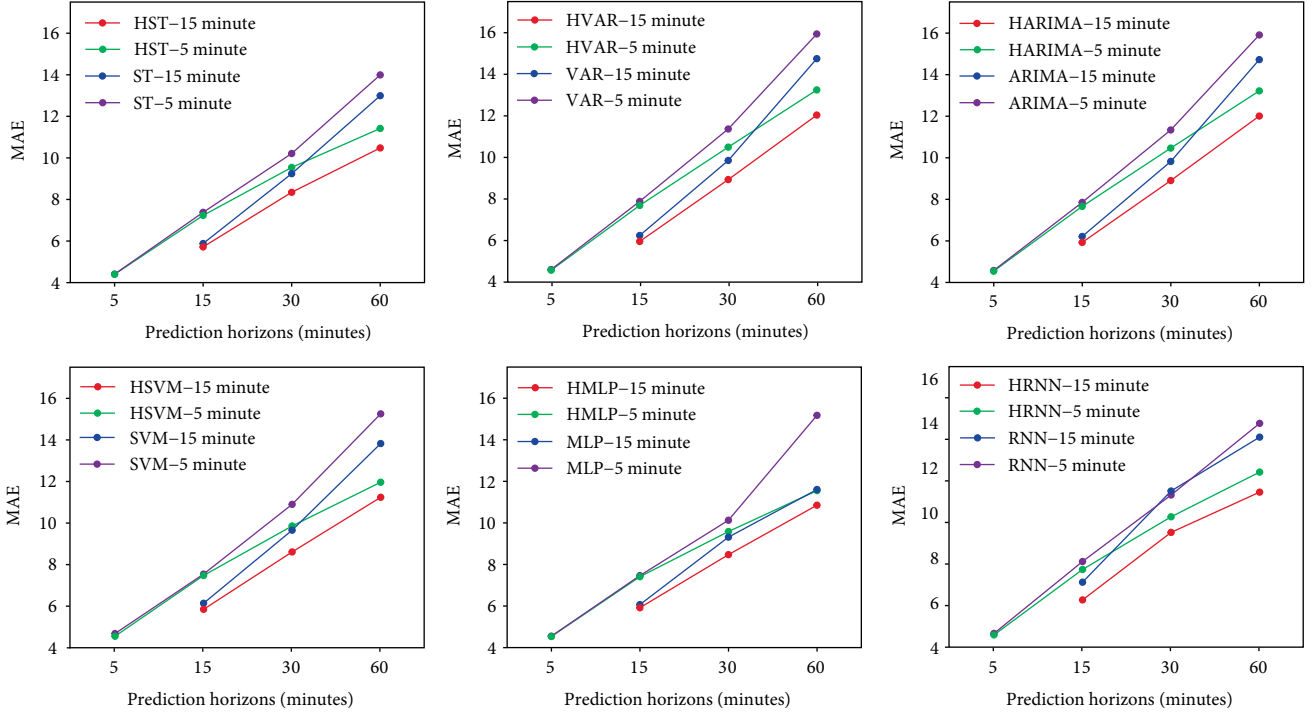


FIGURE 6: The MAE values of models for different horizons and aggregation levels in peak hours on station C.

4. Results

4.1. Evaluation Indicators. To calculate the multi-step ahead prediction performance of different models, three performance evaluation indexes (i.e., the mean absolute error (MAE), the mean absolute percentage error (MAPE), and the root mean square error (RMSE)) are adopted. The calculation formulas of these three indicators are shown as follows:

$$\text{MAE} = \frac{\sum_{i=1}^N |s'_i - s_i|}{N}, \quad (17)$$

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \frac{|s'_i - s_i|}{s_i} \times 100\%, \quad (18)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (s'_i - s_i)^2}{N - 1}}, \quad (19)$$

where, N is the number of observations; s_i is the actual speed at time i at station C; and, s'_i is the predicted speed at time i at station C.

To evaluate the performance of all models, both one-step and multi-step ahead prediction are considered.

4.2. Comparison of Prediction Results. In this study, the prediction performance of ST, VAR, ARIMA, SVM, MLP, RNN, and hybrid models is compared using the speed data at station C. The data samples collected during 6:00 AM–8:00 PM from April 2nd to April 30th (21 weekdays) are selected as the

testing period. Specifically, we utilize a sliding training period that consists of the last 65 days before the prediction point for ST, VAR, and ARIMA models. For instance, the parameters of models for speed prediction on April 10, 2018 are calculated using the last 65 days (i.e., April (6 days), March (22 days), February (20 days), and January (17 days)). To be fair, the data samples from January 1st to March 30th (65 weekdays) are chosen as the training period to optimize parameters in machine learning models.

In the prediction, the optimal parameters of statistical models are calculated in R . More specifically, ARIMA and VAR models use forecast and vars packages. The optimal order in ARIMA model is chosen by the Akaike Information Criterion (AIC) values with the last 65 days of data. A maximal order of 10 is used in VAR model. And the optimal order of VAR model is also determined via the AIC values. All machine learning algorithms are implemented using scikit-learn and keras software packages in the Python programming language. Radial Basis Function (RBF) is set as the kernel functions in SVM. All models consider multi-horizon ahead prediction (5, 15, 30, 60 minutes ahead predictions).

From Figure 2, we can see that there are two peak hours for all stations: one is about from 7:00 AM to 9:00 AM and the other is from 3:30 PM to 7:30 PM. Thus, we divide the target period into two parts: peak hours and off-peak hours.

Figure 6 shows prediction results of MAE values of from 5 minutes to 60 minutes ahead prediction over all models with different aggregation levels in peak hours. In terms of 5 minutes ahead prediction, the performance of hybrid models is slightly better than that of single models. The average MAE

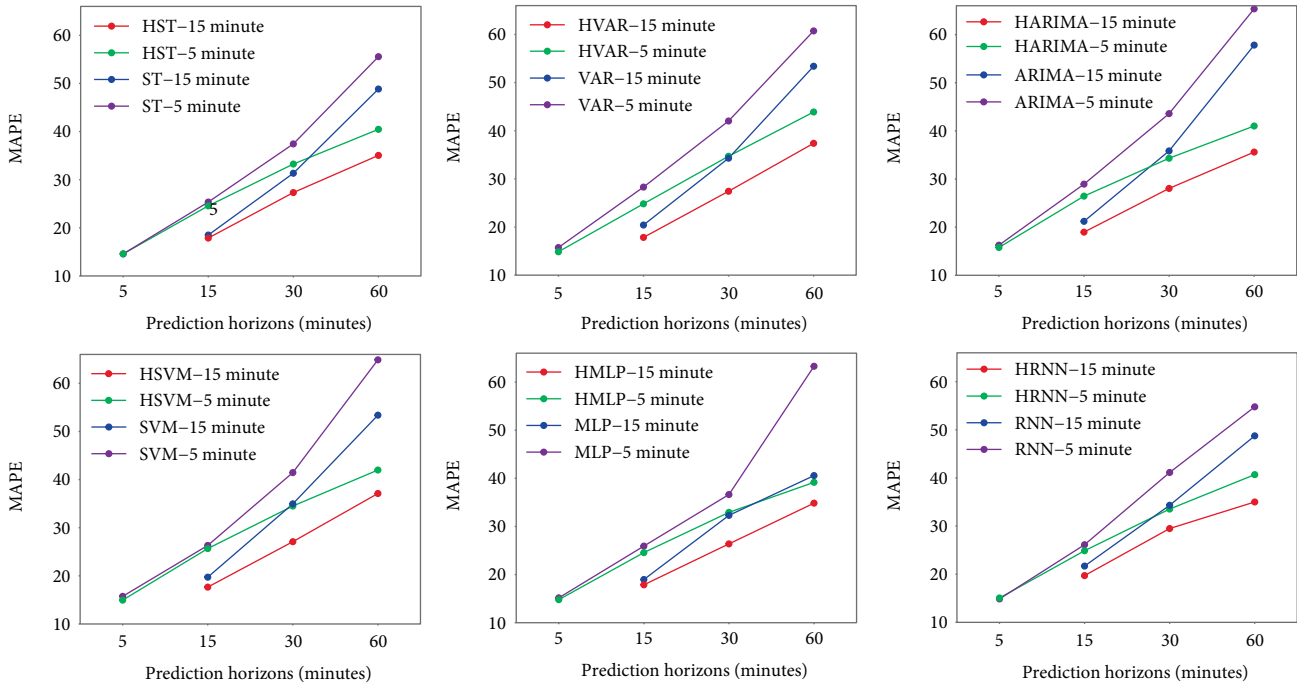


FIGURE 7: The MAPE values of models for different horizons and aggregation levels in peak hours on station C.

value of HARIMA model is 4.7999 while the value of ARIMA model is 4.7683 for 5 minutes ahead prediction. Specifically, the improvement of prediction accuracy is 0.66%, that is, both the hybrid model and the single model perform equally well for 5 minutes ahead prediction. However, for 5-minute data, the MAE value for HMLP model is 11.5593 and the value for MLP model is 15.1838 in terms of 60 minutes ahead prediction. That demonstrates the improvement is 23.87% in prediction performance. Overall, hybrid models perform better than single models in terms of MAE values. Hybrid models can provide explicit description of the basic structure of data and give better insights into potential characteristics of time series data. In addition, Figures 7 and 8 summarize the results of MAPE values and RMSE values. Moreover, similar findings with MAE values can be drawn: for 15-minute speed data, hybrid models show slight improvement in terms of MAPE values and RMSE values of one-step ahead prediction. Moreover, for 60 minutes ahead prediction, hybrid model improves accuracy by 38.43% of MAPE value for ARIMA model, and 26.46% is the improvement of the RMSE value for SVM model. In summary, hybrid models outperform the ST, VAR, ARIMA, SVM, MLP, and RNN models. This is because hybrid models treat residual components and periodic components separately, and the variation pattern of traffic time series data can be amplified focusing on the residual component and thus improves the prediction accuracy.

Further, from Figures 6, 7, and 8, it can be drawn that 15-minute speed data perform better than 5-minute speed data on both hybrid models and single models. For example, for 15-minute ahead prediction, the MAE value for HRNN model of 5-minute speed data, 11.8321, is greater than that of

15-minute data, 10.9388. Additionally, similar conclusion can be drawn from MAPE value and RMSE value. The possible reason is that with smaller aggregation level, more prediction steps are required under the same prediction horizon. Specifically, for 60 minutes ahead prediction, the step of 5-minute speed data is 12 while that of 15-minute aggregation level is 4. As Zou et al. [16] demonstrated, the prediction performance becomes worse as the prediction time step increases. Another reason is that the speed values of 15-minute aggregation level often show more significant fluctuations than those of 5-minute aggregation level.

In summary, two interesting findings can be summarized as follow: (1) the hybrid models can outperform the ST, VAR, ARIMA, SVM, MLP, and RNN models in terms of the prediction accuracy. This is because the hybrid models, considering the trigonometric polynomials, can better capture the periodic features of speed data; (2) under the same prediction horizon, the prediction performance of hybrid models for 15-minute speed data is generally better than that for 5-minute speed data.

Figures 9, 10, and 11 provide MAE, MAPE, and RMSE values of different prediction horizons for all models in off-peak hours. Compared with peak-hours, the prediction performance in off-peak hours is better on both single models and hybrid models. For example, for 5-minute speed data, the MAE value for HVAR model of 30 minutes ahead prediction in peak hours, 10.4796, is much larger than that in off-peak hours, 4.4149. In addition, similar conclusions can be found in terms of MAPE values and RMSE values. The possible reason is described as follows: the speed in off-peak hours is close to free-flow speed and fluctuates slightly. To be specific, the

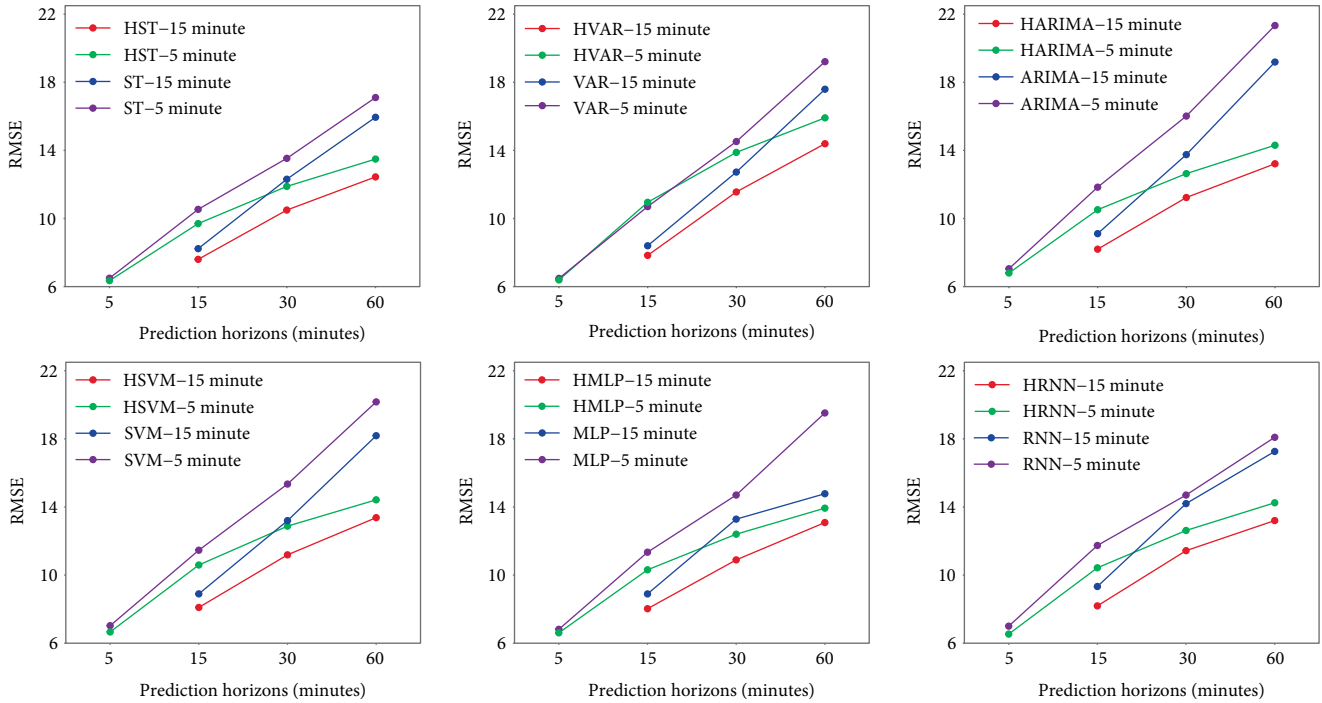


FIGURE 8: The RMSE values of models for different horizons and aggregation levels in peak hours on station C.

feature of speed in off-peak hours is much easier to characterize than that in peak hours. Moreover, similar findings with peak hours can be drawn: (1) the prediction performance of the hybrid model is better than the conventional models (ST, VAR, ARIMA, SVM, MLP, and RNN); (2) under the same prediction horizon, the prediction accuracy for 15-minute speed data is normally more accurate than that for 5-minute speed data.

Composing the residual component from original data, Tables 1 and 2 provide the improvement percentage of different prediction horizons in peak hours. From tables below, we can find that as the time step increases, the proportion of performance improvement gradually increases in sequence over all methods. That is, the hybrid model demonstrates its advantages with larger time steps. Taking 5-minute speed data as an example, the proportion of performance improvement of MLP model for 60 minutes ahead prediction, 23.8708, is larger than that for 30 minutes ahead prediction, 5.3738. This is because trigonometric polynomials successfully describe the long-term periodicity of speed time series data.

5. Discussion

5.1. Effects of Hybrid Models. Decomposing original data into two components (residual component and periodic component), hybrid models can capture the dynamic fluctuations of traffic data accurately. The periodic component, the main trend of time series, generally represents the time-independent part of original data; while the residual component reflects the time-dependent part of time series. In this study, a trigonometric regression function

is selected to construct the periodic component of daily similarity, while three statistical models (space time model, ARIMA model, and VAR model) and three machine learning models (SVM model, MLP model, and RNN model) are used to describe the residual component. Although statistical models and machine learning models both aim to predict traffic speed, these two types of approaches differ greatly in model structure and result interpretation. Statistical models, based on a rule-based mechanism, are presented in mathematical formula form. Additionally, each variable in the formula has a specific or practical meaning. However, the application of statistical models is generally limited to the assumption embedded in the model. For example, data are commonly required to vary linearly with time in ARIMA model and VAR model. In the contrast, machine learning models directly learn from data without any explicit model structure, and operate as a “black box”. Thus, the application of these models is more flexible, especially when handling big data. Nevertheless, lacking the capability of result interpretation is generally the limitation of these models. Given the significant difference between these two types of prediction, this study explored whether periodic components have the same impact on statistical model or machine learning model. Consequently, based on the above results, it can be inferred that the periodic component improves prediction performance both for statistical model and machine learning model. In summary, in terms of the unique characteristics of two kinds of models, researchers need to select the corresponding model according to specific requirements and assumptions. However, it is beneficial to introduce periodic components for both statistical model and machine learning model.

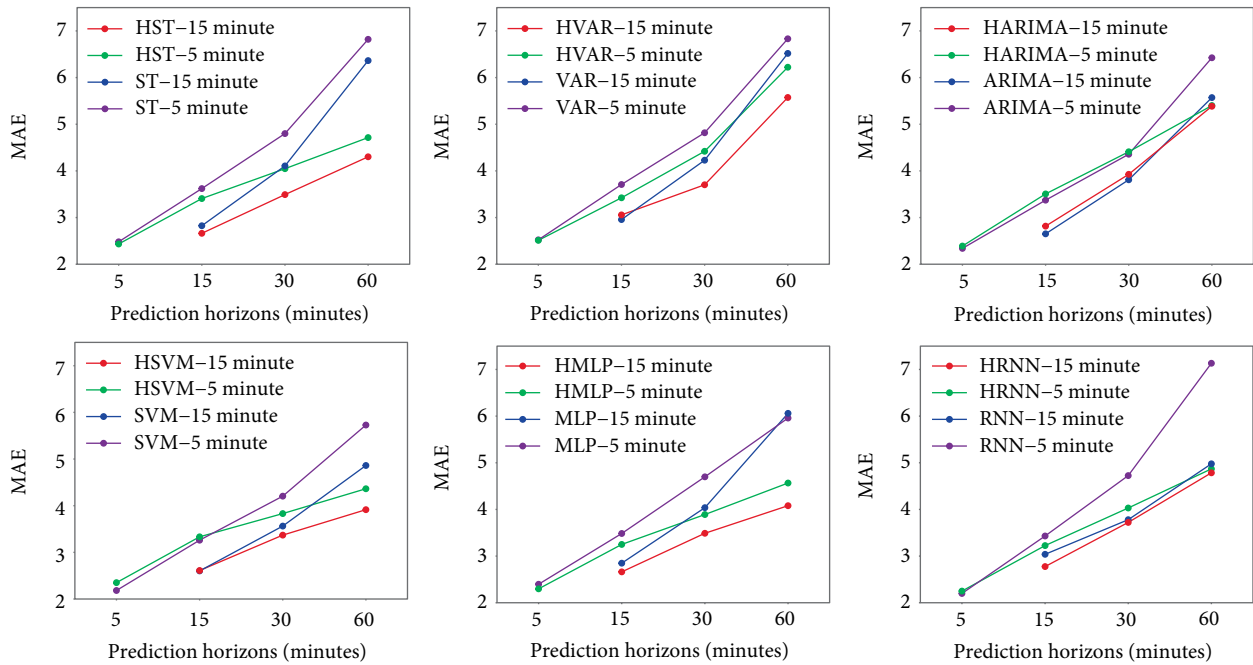


FIGURE 9: The MAE values of models for different horizons and aggregation levels in off-peak hours on station C.

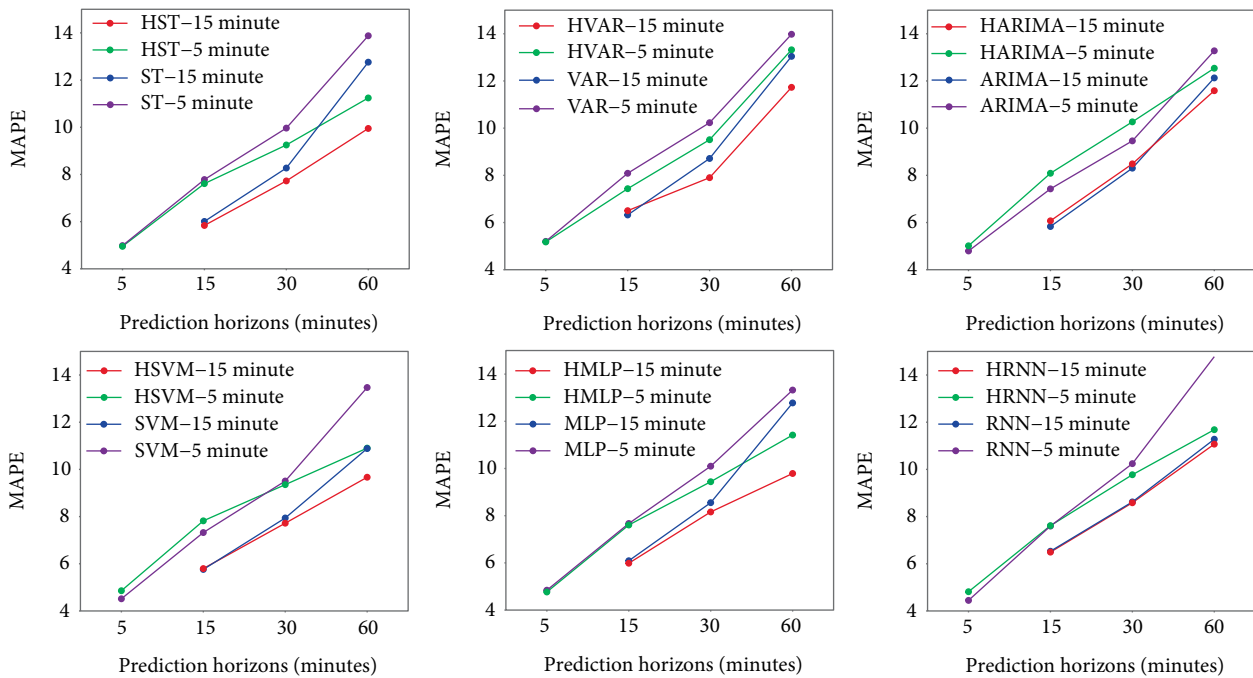


FIGURE 10: The MAPE values of models for different horizons and aggregation levels in off-peak hours on station C.

5.2. *Improvements of Hybrid Models over Single Models.* Improvement percentage, the ratio of prediction accuracy improvement from single models to hybrid models, is utilized to measure the effect considering periodic component. Compared with off-peak hours, the prediction performance in peak hours is generally more critical.

Specifically, the impact of periodic component in peak hours on both statistical model and machine learning model is examined. From one-step ahead prediction to 12-step ahead prediction, the improvement of accuracy ranges from 0.65% to 28.37% of ARIMA model in terms of 5-minute data. Similar conclusions can be drawn from other models (VAR

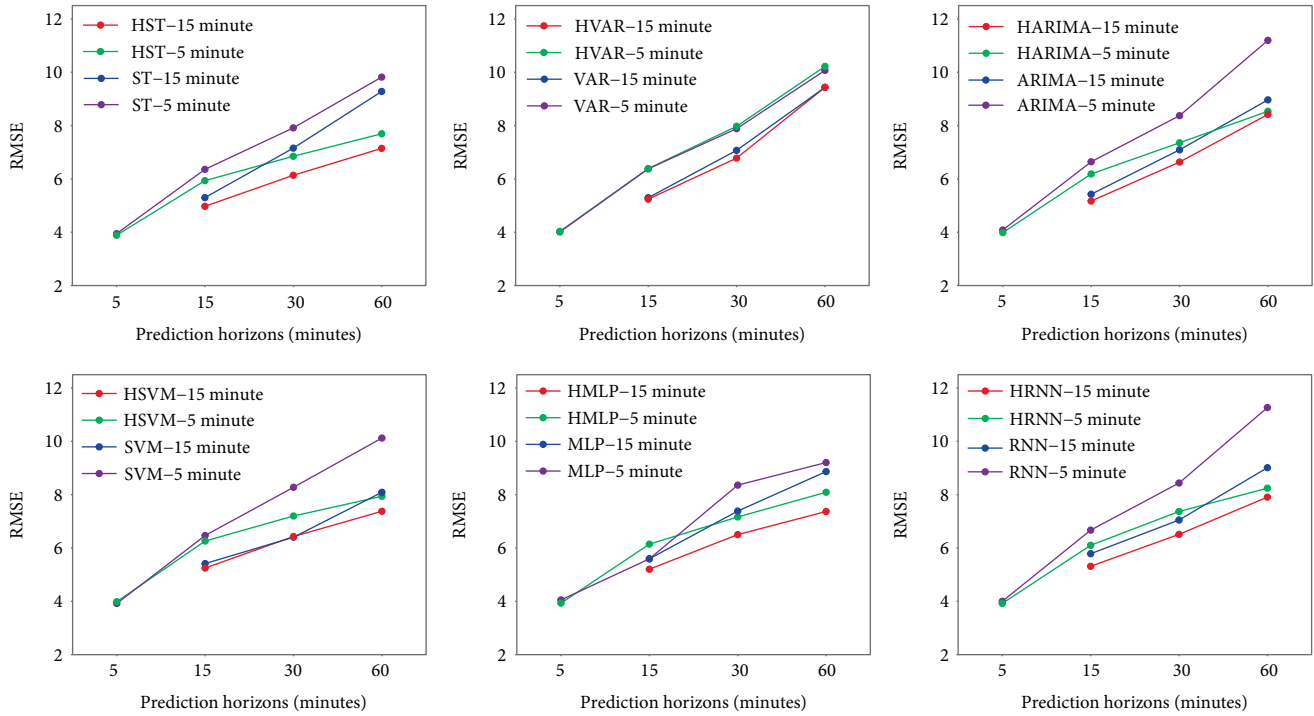


FIGURE 11: The RMSE values of models for different horizons and aggregation levels in off-peak hours on station C.

model, ST model, SVM model, MLP model, RNN model). It can be found that longer prediction step is associated with larger improvement, which explains that the decomposed periodic component of original data has its advantages over long-term trend.

5.3. Effects of Data Aggregation Level. Further, the impact of two different data aggregation levels on prediction precision can be examined from the prediction results above. For traffic data, if the time interval is too small, dynamic fluctuations of data are too complex and data generally contain redundant information. Moreover, if the time interval is too large, the variation trend of data is smooth which leads to loss of useful information. That is, 15-minute data contain less noise than 5-minute data. Based on the prediction results shown above, it can be concluded that 15-minute aggregation level is superior to 5-minute aggregation level in terms of the three evaluation indicators (MAE, MAPE, and RMSE). In other words, it is easier to accurately predict average speed in 15 minutes than to separately predict three consecutive 5-minute periods. After extracting periodic components from the original data, the dynamic variation of 15-minute aggregation level is slighter than that of 5-minute aggregation level, and thus it is easier to capture the underlining characteristic of remaining part of original data. In the meanwhile, larger prediction steps generally result in lower prediction accuracy. Thus, with the same prediction horizon, such as 30 minutes ahead prediction, 5-minute data requires a 6-step ahead prediction while a 2-step ahead prediction is adequate for 15-minute data. Overall, it is worthwhile to make a trade-off between prediction accuracy and data fluctuation in practical prediction. Specifically, if accurate high-aggregation

predictions are required, then it is worthy dedicating further study into effectively utilizing high-aggregation data while minimizing the negative effects of noise. Additionally, in terms of low-aggregation data, it is required to reduce information loss while ensuring high prediction accuracy.

6. Conclusions

In this study, we examine the impact of periodic component on three statistical models and three machine learning models when predicting freeway traffic speed using the data collected from five loop detectors of an eastbound road of Interstate 394 freeway stretch, Minnesota. In addition, multi-horizon ahead prediction and two aggregation levels of data are also considered. The prediction performance is measured by three evaluation indicators: MAE, RMSE, and MAPE. The important conclusions can be summarized as follows: (1) the multi-step prediction accuracy improves considering the periodic component of speed data for both statistical models and machine learning models, especially in the peak hours; (2) as the time step increases, the prediction performance improvement gradually becomes larger considering the impact of periodic component over all models; (3) under the same prediction horizon, the prediction performance of all models for 15-minute speed data is generally better than that for 5-minute speed data. Thus, the selection of prediction models can significantly affect the prediction accuracy in terms of MAE, MAPE and RMSE, and further influences the performance regarding traffic management tools such as ATIS and ATMS of ITS system. Inaccurate prediction of traffic parameters may have a significant impact

TABLE 1: Improvement percentage of hybrid models for different prediction horizons in peak hours at station C (time scale: 5 min).

MAE	Prediction horizons (minutes)			
	5	15	30	60
ST	0.1678	2.0201	6.5398	18.4136
VAR	0.7073	2.4666	7.7073	16.9290
ARIMA	0.6583	2.7986	13.5338	28.3774
SVM	2.7703	0.8935	9.5560	21.5672
MLP	0.3708	0.6867	5.3738	23.8708
RNN	1.4256	4.5202	8.9773	15.5103
MAPE (%)	Prediction horizons (minutes)			
	5	15	30	60
ST	-0.0816	3.1284	11.1746	27.1694
VAR	5.5966	12.2921	17.4324	27.7299
ARIMA	3.1773	10.5383	23.3195	38.2678
SVM	5.0060	2.4855	16.7698	35.2814
MLP	2.4710	5.2158	10.1209	38.1002
RNN	-1.3680	4.7849	18.4671	25.7139
RMSE	Prediction horizons (minutes)			
	5	15	30	60
ST	2.2066	7.9375	12.1864	21.1225
VAR	1.4003	-2.3589	4.3848	17.1806
ARIMA	5.0394	11.0708	20.5272	32.1107
SVM	5.2385	7.5790	16.1069	28.5130
MLP	3.0359	9.1199	15.6351	28.6400
RNN	6.6514	11.1524	14.1581	21.2511

TABLE 2: Improvement percentage of hybrid models for different prediction horizons in peak hours at station C (time scale: 15 min).

MAE	Prediction horizons (minutes)			
	5	15	30	60
ST	—	2.8043	9.6993	19.3439
VAR	—	4.6387	9.3649	18.4457
ARIMA	—	3.3953	10.6782	26.6451
SVM	—	4.7455	10.8446	18.7158
MLP	—	2.4939	9.0440	6.4904
RNN	—	11.3976	16.7804	18.3007
MAPE (%)	Prediction horizons (minutes)			
	5	15	30	60
ST	—	3.3989	12.8081	28.2290
VAR	—	12.6058	20.0241	29.9824
ARIMA	—	10.6009	21.7511	38.4383
SVM	—	10.4714	22.4997	30.4860
MLP	—	5.8083	18.3380	14.1194
RNN	—	9.1287	14.0741	28.1865
RMSE	Prediction horizons (minutes)			
	5	15	30	60
ST	—	7.6198	14.7111	21.9747
VAR	—	6.7191	9.1703	18.1812
ARIMA	—	10.0176	18.3135	31.1368
SVM	—	8.9609	15.2175	26.4659
MLP	—	9.7862	18.0127	11.4506
RNN	—	12.2067	19.4621	23.5209

on the operation of ITS system, which aims to improve the efficiency and capacity of traffic and transportation system. Given the accurate real-time information about traffic condition (traffic parameters such as traffic speed, traffic flow, and travel time), drivers can replan or adjust travel routes to avoid the heavy traffic of original routes. Additionally, either ramp controlling or variable speed limiting can effectively alleviate traffic congestion while the traffic speed of mainline closes to the critical speed. In summary, transportation management agencies should be cautious further when predicting traffic condition to maintain the regular operation of ITS [45].

In the future, several in-depth research can be conducted. First, nonrecurring congestion events, such as incidents and special events, will influence the distribution of traffic speed. Thus, it is necessary to examine the prediction performance of these models for nonrecurring congestion condition. Second, to address the heterogeneous speed data, some prediction methods based on finite mixture model and copula model may be developed [73, 74]. Third, it is useful to compare the traffic prediction performance between the traffic flow models [75–77] and the proposed prediction models in this study. Fourth, the possible impact of periodic component on interval prediction should also be considered [37, 78–80].

Data Availability

The data used to support the findings of this study are available from the publicly available data tool developed by the Minnesota Department of Transportation. Please visit: <http://data.dot.state.mn.us/datatools/>.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was funded jointly by the National Key Research and Development Program of China, grant number 2018YFE0102800; the National Natural Science Foundation of China, grant number 71971160, 71701215; Shanghai Science and Technology Committee, grant number 18510745400; Foundation of Central South University, grant number 502045002; Science and Innovation Foundation of the Transportation Department in Hunan Province, grant number 201725; Postdoctoral Science Foundation of China, grant number 140050005.

References

- [1] X. Yu and P. D. Prevedouros, "Performance and challenges in utilizing non-intrusive sensors for traffic data collection," *Advances in Remote Sensing*, vol. 2, no. 2, pp. 45–50, 2013.
- [2] E. I. Vlahogianni, J. C. Golias, and M. G. Karlaftis, "Short-term traffic forecasting: overview of objectives and methods," *Transport reviews*, vol. 24, no. 5, pp. 533–557, 2004.

- [3] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Short-term traffic forecasting: where we are and where we're going," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 3–19, 2014.
- [4] J. Van Lint and C. Van Hinsbergen, "Short-term traffic and travel time prediction models," *Artificial Intelligence Applications to Critical Transportation Issues*, vol. 22, pp. 22–41, 2012.
- [5] S. Das, A. Dutta, G. Medina, L. Minjares-Kyle, and Z. Elgart, "Extracting patterns from twitter to promote biking," *IATSS Research*, vol. 43, no. 1, pp. 51–59, 2018.
- [6] C. Ma, R. He, and W. Zhang, "Path optimization of taxi carpooling," *PLoS One*, vol. 13, no. 8, Article ID e0203221, 2018.
- [7] L. Li, X. Su, Y. Zhang, Y. Lin, and Z. Li, "Trend modeling for traffic time series analysis: an integrated study," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, pp. 3430–3439, 2015.
- [8] R. Zhang, X. Ye, K. Wang, D. Li, and J. Zhu, "Development of commute mode choice model by integrating actively and passively collected travel data," *Sustainability*, vol. 11, no. 10, p. 2730, 2019.
- [9] J. Tang, J. Liang, C. Han, Z. Li, and H. Huang, "Crash injury severity analysis using a two-layer stacking framework," *Accident Analysis & Prevention*, vol. 122, pp. 226–238, 2019.
- [10] M. Ben-Akiva, M. Bierlaire, D. Burton, H. N. Koutsopoulos, and R. Mishalani, "Network state estimation and prediction for real-time traffic management," *Networks and Spatial Economics*, vol. 1, no. 3–4, pp. 293–318, 2001.
- [11] H. S. Mahmassani, "Dynamic network traffic assignment and simulation methodology for advanced system management applications," *Networks and Spatial Economics*, vol. 1, no. 3–4, pp. 267–292, 2001.
- [12] C. Giovanna, M. Giuseppe, P. Antonio, R. Corrado, R. Francesco, and V. Antonino, "Transport models and intelligent transportation system to support urban evacuation planning process," *IET Intelligent Transport Systems*, vol. 10, no. 4, pp. 279–286, 2016.
- [13] S. R. Chandra and H. Al-Deek, "Predictions of freeway traffic speeds and volumes using vector autoregressive models," *Journal of Intelligent Transportation Systems*, vol. 13, no. 2, pp. 53–72, 2009.
- [14] P. Shang, X. Li, and S. Kamae, "Chaotic analysis of traffic time series," *Chaos, Solitons & Fractals*, vol. 25, no. 1, pp. 121–128, 2005.
- [15] B. M. Williams, "Multivariate vehicular traffic flow prediction: evaluation of ARIMAX modeling," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1776, no. 1, pp. 194–200, 2001.
- [16] Y. Zou, X. Hua, Y. Zhang, and Y. Wang, "Hybrid short-term freeway speed prediction methods based on periodic analysis," *Canadian Journal of Civil Engineering*, vol. 42, no. 8, pp. 570–582, 2015.
- [17] S. V. Kumar and L. Vanajakshi, "Short-term traffic flow prediction using seasonal ARIMA model with limited input data," *European Transport Research Review*, vol. 7, no. 3, p. 21, 2015.
- [18] Y. Kamarianakis and V. Vouton, "Forecasting traffic flow conditions in an urban network: comparison of multivariate and univariate approaches," *Transportation Research Record*, vol. 1857, no. 1, pp. 74–84, 2003.
- [19] S. Yang, S. Shi, X. Hu, and M. Wang, "Spatiotemporal context awareness for urban traffic modeling and prediction: sparse representation based variable selection," *PLoS One*, vol. 10, no. 10, Article ID e0141223, 2015.
- [20] Y. Zhang, A. Haghani, and X. Zeng, "Component GARCH models to account for seasonal patterns and uncertainties in travel-time prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 719–729, 2015.
- [21] A. Agafonov and A. Yumaganov, "Spatial-temporal K nearest neighbors model on mapreduce for traffic flow prediction," *International Conference on Intelligent Data Engineering and Automated Learning*, pp. 253–260, 2018.
- [22] Q. Ye, W. Y. Szeto, and S. C. Wong, "Short-term traffic speed forecasting based on data recorded at irregular intervals," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 4, pp. 1727–1737, 2012.
- [23] X. Ma, H. Yu, Y. Wang, and Y. Wang, "Large-scale transportation network congestion evolution prediction using deep learning theory," *PLoS One*, vol. 10, no. 3, Article ID e0119044, 2015.
- [24] J. W. C. van Lint, S. P. van Hoogendoorn, and H. J. Zuylen, "Accurate freeway travel time prediction with state-space neural networks under missing data," *Transportation Research Part C: Emerging Technologies*, vol. 13, no. 5–6, pp. 347–369, 2005.
- [25] X. Ma, Z. Tao, Y. Wang, H. Yu, and Y. Wang, "Long short-term memory neural network for traffic speed prediction using remote microwave sensor data," *Transportation Research Part C: Emerging Technologies*, vol. 54, pp. 187–197, 2015.
- [26] 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.
- [27] L. Dimitriou, T. Tsekeris, and A. Stathopoulos, "Adaptive hybrid fuzzy rule-based system approach for modeling and predicting urban traffic flow," *Transportation Research Part C: Emerging Technologies*, vol. 16, no. 5, pp. 554–573, 2008.
- [28] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wanag, "Traffic flow prediction with big data: a deep learning approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 2, pp. 865–873, 2015.
- [29] N. G. Polson and V. O. Sokolov, "Deep learning for short-term traffic flow prediction," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 1–17, 2017.
- [30] C.-H. Wu, J.-M. Ho, and D.-T. Lee, "Travel-time prediction with support vector regression," *IEEE Transactions on Intelligent Transportation Systems*, vol. 5, no. 4, pp. 276–281, 2004.
- [31] Y. Zhang and Y. Xie, "Forecasting of short-term freeway volume with v support vector machines," *Transportation Research Record Journal of the Transportation Research Board*, vol. 2024, no. 1, pp. 92–99, 2007.
- [32] Y. Zhang and Y. Liu, "Traffic forecasting using least squares support vector machines," *Transportmetrica*, vol. 5, no. 3, pp. 193–213, 2009.
- [33] 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.
- [34] Y.-C. Zhao, Y.-C. Liu, L.-Q. Shan, and B. Zhou, "Dynamic analysis of kalman filter for traffic flow forecasting in sensor networks," *Information Technology Journal*, vol. 11, no. 10, pp. 1508–1512, 2012.
- [35] S. I.-J. Chien and C. M. Kuchipudi, "Dynamic travel time prediction with real-time and historic data," *Journal of Transportation Engineering*, vol. 129, no. 6, pp. 608–616, 2003.

- [36] Y. Wang, M. Papageorgiou, and A. Messmer, "RENAISSANCE – a unified macroscopic model-based approach to real-time freeway network traffic surveillance," *Transportation Research Part C: Emerging Technologies*, vol. 14, no. 3, pp. 190–212, 2006.
- [37] J. Guo, W. Huang, and B. M. Williams, "Adaptive kalman filter approach for stochastic short-term traffic flow rate prediction and uncertainty quantification," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 50–64, 2014.
- [38] W. Zheng and D. H. Lee, "Short-term freeway traffic flow prediction: bayesian combined neural network approach," *Journal of Transportation Engineering*, vol. 132132, no. 2, pp. 114–121, 2006.
- [39] C. Dong, S. H. Richards, Q. Yang, and C. Shao, "Combining the statistical model and heuristic model to predict flow rate," *Journal of Transportation Engineering*, vol. 140, no. 7, p. 04014023, 2014.
- [40] C. Yanchong, H. Darong, and Z. Ling, "A short-term traffic flow prediction method based on wavelet analysis and neural network," in *2016 Chinese Control and Decision Conference (CCDC)*, pp. 7030–7034, IEEE, Yinchuan, China, 2016.
- [41] J. Tang, G. Zhang, Y. Wang, H. Wang, and F. Liu, "A hybrid approach to integrate fuzzy C-means based imputation method with genetic algorithm for missing traffic volume data estimation," *Transportation Research Part C: Emerging Technologies*, vol. 51, pp. 29–40, 2015.
- [42] H. Huang, Q. Tang, and Z. Liu, "Adaptive correction forecasting approach for urban traffic flow based on fuzzy-mean clustering and advanced neural network," *Journal of Applied Mathematics*, vol. 2013, Article ID 195824, 7 pages, 2013.
- [43] W. Huang, G. Song, H. Hong, and K. Xie, "Deep architecture for traffic flow prediction: deep belief networks with multitask learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 5, pp. 2191–2201, 2014.
- [44] J. Tang, H. Wang, Y. Wang, X. Liu, and F. Liu, "Hybrid prediction approach based on weekly similarities of traffic flow for different temporal scales," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2443, no. 1, pp. 21–31, 2014.
- [45] B. M. Williams and L. A. Hoel, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results," *Journal of Transportation Engineering*, vol. 129, no. 6, pp. 664–672, 2003.
- [46] W.-C. Hong, "Application of seasonal SVR with chaotic immune algorithm in traffic flow forecasting," *Neural Computing and Applications*, vol. 21, no. 3, pp. 583–593, 2012.
- [47] M. Lippi, M. Bertini, and P. Frasconi, "Short-term traffic flow forecasting: an experimental comparison of time-series analysis and supervised learning," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 2, pp. 871–882, 2013.
- [48] X. Jiang and H. Adeli, "Dynamic wavelet neural network model for traffic flow forecasting," *Journal of transportation engineering*, vol. 131, no. 10, pp. 771–779, 2005.
- [49] Y. Zhang, Y. Zhang, and A. Haghani, "A hybrid short-term traffic flow forecasting method based on spectral analysis and statistical volatility model," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 65–78, 2014.
- [50] M.-C. Tan, S. C. Wong, J.-M. Xu, Z.-R. Guan, and P. Zhang, "An aggregation approach to short-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 1, pp. 60–69, 2009.
- [51] D. S. Dendrinos, "Traffic-flow dynamics: a search for chaos," *Chaos, Solitons & Fractals*, vol. 4, no. 4, pp. 605–617, 1994.
- [52] T. T. Tchraikian, B. Basu, and M. O'Mahony, "Real-time traffic flow forecasting using spectral analysis," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 519–526, 2012.
- [53] A. Stathopoulos and M. G. Karlaftis, "Spectral and cross-spectral analysis of urban traffic flows," in *2001 IEEE Intelligent Transportation Systems Proceedings*, pp. 820–825, IEEE, Oakland, CA, USA, 2001.
- [54] S. R. Chandra and H. Al-Deek, "Cross-correlation analysis and multivariate prediction of spatial time series of freeway traffic speeds," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2061, no. 1, pp. 64–76, 2008.
- [55] J. Yang, L. D. Han, P. B. Freeze, S.-M. Chin, and H.-L. Hwang, "Short-term freeway speed profiling based on longitudinal spatiotemporal dynamics," *Transportation Research Record*, vol. 2467, no. 1, pp. 62–72, 2014.
- [56] J. Xia, M. Chen, and W. Huang, "A multistep corridor travel-time prediction method using presence-type vehicle detector data," *Journal of Intelligent Transportation Systems*, vol. 15, no. 2, pp. 104–113, 2011.
- [57] Y. Zou, X. Zhu, Y. Zhang, and X. Zeng, "A space-time diurnal method for short-term freeway travel time prediction," *Transportation Research Part C: Emerging Technologies*, vol. 43, pp. 33–49, 2014.
- [58] Y. Zou, Y. Zhang, and X. Zhu, "Constructing a bivariate distribution for freeway speed and headway data," *Transportmetrica*, vol. 10, no. 3, pp. 255–272, 2014.
- [59] T. Gneiting, K. Larson, K. Westrick, M. G. Genton, and E. Aldrich, "Calibrated probabilistic forecasting at the stateline wind energy center: The regime-switching space-time method," *Journal of the American Statistical Association*, vol. 101, no. 475, pp. 968–979, 2006.
- [60] X. Luo, D. Li, Y. Yang, and S. Zhang, "Spatiotemporal traffic flow prediction with KNN and LSTM," *Journal of Advanced Transportation*, vol. 2019, Article ID 4145353, 10 pages, 2019.
- [61] M. Chen, X. Yu, and Y. Liu, "PCNN: deep convolutional networks for short-term traffic congestion prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 11, pp. 3550–3559, 2018.
- [62] C. Xu, Z. Li, and W. Wang, "Short-term traffic flow prediction using a methodology based on autoregressive integrated moving average and genetic programming," *Transport*, vol. 31, no. 3, pp. 343–358, 2016.
- [63] A. J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199–222, 2004.
- [64] Z. Zhao, W. Chen, X. Wu, P. C. Chen, and J. Liu, "LSTM network: a deep learning approach for short-term traffic forecast," *IET Intelligent Transport Systems*, vol. 11, no. 2, pp. 68–75, 2017.
- [65] X. Luo, L. Niu, and S. Zhang, "An algorithm for traffic flow prediction based on improved SARIMA and GA," *KSCCE Journal of Civil Engineering*, vol. 22, no. 10, pp. 4107–4115, 2018.
- [66] J. Tang, F. Liu, W. Zhang, R. Ke, and Y. Zou, "Lane-changes prediction based on adaptive fuzzy neural network," *Expert Systems with Applications*, vol. 91, pp. 452–463, 2018.
- [67] H. Jiang, Y. Zou, S. Zhang, J. Tang, and Y. Wang, "Short-term speed prediction using remote microwave sensor data: machine learning versus statistical model," *Mathematical Problems in Engineering*, vol. 2016, Article ID 9236156, 13 pages, 2016.
- [68] D. Kinga and J. B. Adam, "A method for stochastic optimization," in *International Conference on Learning Representations, ICLR, San Diego*, 2015.

- [69] M. Hüsken and P. Stagge, "Recurrent neural networks for time series classification," *Neurocomputing*, vol. 50, pp. 223–235, 2003.
- [70] Z. Khan, S. M. Khan, K. Dey, and M. Chowdhury, "Development and evaluation of recurrent neural network-based models for hourly traffic volume and annual average daily traffic prediction," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2673, no. 7, pp. 489–503, 2019.
- [71] X. Cheng, R. Zhang, J. Zhou, and W. Xu, "Deep Transport: Learning spatial-temporal dependency for traffic condition forecasting," in *2018 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8, IEEE, Rio de Janeiro, Brazil, 2018.
- [72] H. M. Adorf, "Interpolation of irregularly sampled data series - a survey," *Astronomical Data Analysis Software & Systems IV*, vol. 77, 1995.
- [73] Y. Zou, J. E. Ash, B.-J. Park, D. Lord, and L. Wu, "Empirical bayes estimates of finite mixture of negative binomial regression models and its application to highway safety," *Journal of Applied Statistics*, vol. 45, no. 9, pp. 1652–1669, 2018.
- [74] Y. Zou, X. Zhong, J. Tang et al., "A copula-based approach for accommodating the underreporting effect in wildlife-vehicle crash analysis," *Sustainability*, vol. 11, no. 2, p. 418, 2019.
- [75] C. F. Daganzo, "The cell transmission model: a dynamic representation of highway traffic consistent with the hydrodynamic theory," *Transportation Research Part B: Methodological*, vol. 28, no. 4, pp. 269–287, 1994.
- [76] D. C. Festa, G. Longo, G. Mazzulla, and G. Musolino, "Experimental analysis of different simulation models for motorway traffic flow," in *ITSC 2001. 2001 IEEE Intelligent Transportation Systems. Proceedings (Cat. No. 01TH8585)*, pp. 675–680, IEEE, Oakland, CA, USA, 2001.
- [77] J. R. D. Frejo, I. Papamichail, M. Papageorgiou, and B. De Schutter, "Macroscopic modeling of variable speed limits on freeways," *Transportation research part C: emerging technologies*, vol. 100, pp. 15–33, 2019.
- [78] G. Shi, J. Guo, W. Huang, and B. M. Williams, "Modeling seasonal heteroscedasticity in vehicular traffic condition series using a seasonal adjustment approach," *Journal of Transportation Engineering*, vol. 140, p. 04014012, 2014.
- [79] J. Guo, W. Huang, and B. M. Williams, "Integrated heteroscedasticity test for vehicular traffic condition series," *Journal of Transportation Engineering*, vol. 138, no. 9, pp. 1161–1170, 2012.
- [80] W. Huang, W. Jia, J. Guo et al., "Real-time prediction of seasonal heteroscedasticity in vehicular traffic flow series," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, pp. 3170–3180, 2017.



Hindawi

Submit your manuscripts at
www.hindawi.com

