Research Article

Short-Term Traffic Flow Prediction of Expressway: A Hybrid Method Based on Singular Spectrum Analysis Decomposition

Chunyan Shuai,¹ Zhengyang Pan,¹ Lun Gao,¹ and HongWu Zuo ²

¹Faculty of Transportation Engineering, Kunming University of Science and Technology, Kunming 650093, China

²School of Continuing Education, Kunming University of Science and Technology, Kunming, Yunnan 650093, China

Correspondence should be addressed to HongWu Zuo; earth0806@kust.edu.cn

Received 30 June 2021; Revised 8 September 2021; Accepted 14 September 2021; Published 8 October 2021

Academic Editor: Xiuling Wang

Copyright © 2021 Chunyan Shuai 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.

Real-time expressway traffic flow prediction is always an important research field of intelligent transportation, which is conducive to inducing and managing traffic flow in case of congestion. According to the characteristics of the traffic flow, this paper proposes a hybrid model, SSA-LSTM-SVR, to improve forecasting accuracy of the short-term traffic flow. Singular Spectrum Analysis (SSA) decomposes the traffic flow into one principle component and three random components, and then in terms of different characteristics of these components, Long Short-Term Memory (LSTM) and Support Vector Regression (SVR) are applied to make prediction of different components, respectively. By fusing respective forecast results, SSA-LSTM-SVR obtains the final short-term predictive value. Experiments on the traffic flows of Guizhou expressway in January 2016 show that the proposed SSA-LSTM-SVR model has lower predictive errors and a higher accuracy and fitting goodness than other baselines. This illustrates that a hybrid model for traffic flow prediction based on components decomposition is more effective than a single model, since it can capture the main regularity and random variations of traffic flow.

1. Introduction

Expressway traffic flow prediction and management is always an important research field of intelligent transportation [1]. Mastering the change laws of traffic flow and accurately predicting the real-time traffic flow are conducive to studying the traffic capacity of expressway and managing and inducing traffic flow, so as to relieve traffic congestion, reduce fuel consumption, and save the total social cost of traffic travel [2]. Expressway traffic flow is affected by human travel behaviour, in sequence with strong regularity, suddenness, and randomness. Currently, different prediction models, such as classical statistical models, machine learning models, and Artificial Neural Network (ANN), have predicted the short-term traffic flow with different accuracies by obtaining different characteristics and changing laws of traffic flow [1].

The classical statistical models mainly include Automatic Regressive Integrated Moving Average (ARIMA), Seasonal ARIMA (SARIMA) [3], and ARIMA hybrid models. To solve

the problem of SARIMA requiring a sound data for model building, Kumar and Vanajakshi [3] made the traffic flow a stationary one by differencing and then used the autocorrelation function, partial autocorrelation function, and maximum likelihood method to identify the suitable parameters of the SARIMA model and made traffic flow forecast with limited data. Considering their limitations of the ARIMA needing the stationarity and autocorrelation of the time series and the variability of traffic flow, the ARIMA family is usually combined with other nonlinear models to make prediction of flow. Ding et al. [4] incorporated the dynamic volatility into the subway short-term ridership forecasting process, and they construct four kinds of the integrated ARIMA and GARCH models to model the mean part and volatility part of the short-term ridership in subway. However, this method needs identifying the volatility characteristic and parameters of different subway stations one by one. Hou et al. [5] proposed an adaptive hybrid model, in which ARIMA captures the linear laws and the Wavelet Neural Network method obtains the nonlinear

changes. Then, outputs of the two individual models are combined by fuzzy logic and their weighted result is regarded as the final predicted values of the traffic flow. This hybrid model performs well in the data processing stage, but vehicles which are hard to identify will be ignored. Support Vector Regression (SVR) has been employed in linear and nonlinear time series regression widely; however, reasonable hyperparameters in SVR are a difficult problem. To overcome this shortage, Cai et al. [6] employed gravitational search algorithm (GSA) to search optimal SVR parameters and achieved a more accurate short-term traffic flow forecasting. SVR is generally combined with other models to achieve a short-term traffic flow forecasting. According to the different categories of the passenger flows entered in the subway station, Li et al. [7] determined the optimal time granularity and interval by utilizing Pearson's correlation coefficient and then used Empirical Mode Decomposition (EMD) and SVR (EMD-SVR) to predict the passenger flow for each station. This framework obtains a significant improvement in forecasting passenger flow. However, it is difficult for the GSA and EMD in [6, 7] to obtain a uniform hyperparameter of SVR due to volatility characteristics of traffic or passenger flow on different stations. Feng et al. [8] explored both the randomness and nonlinearity of the traffic flow and combine polynomial kernel and Gaussian kernel to constitute an adaptive multi-kernel SVM (AMSVM), in which an adaptive particle swarm optimization algorithm is employed, that optimizes the parameters of AMSVM. The spatial-temporal correlation information is incorporated with AMSVM to predict the short-term traffic flow. Multiple kernels and hybrid methods make AMSVM can better adapt to the dynamic characteristic of traffic flow on urban roads, thus providing a more accurate predicted result, but the predicted values have a certain delay compared to the real values. K -nearest neighbors (KNN-based) models, as another common choice, implement traffic flow forecast by matching similarities in the historical data. Habtemichael and Cetin [9] gave more weight to the near neighbors, used a rank exponent to aggregate the candidates found by the enhanced KNN algorithm, and implemented short-term traffic forecasting. To explore the the inherent spatial heterogeneity of city traffic and unclear spatiotemporal dependency relationships, Cheng et al. [10] proposed an adaptive spatiotemporal KNN model (adaptive STKNN) for short-term traffic forecasting. Adaptive STKNN determines the sizes of spatial neighbors and the lengths of time windows for traffic influence using cross-correlation and autocorrelation functions and then introduces adaptive spatiotemporal weights into the distance functions to optimize the candidate neighbor search mechanism. The results demonstrated that the adaptive STKNN model outperforms other models during all time periods and especially the peak period. However, once the query does not appear in the historical data, the KNN-based models will suffer a big forecasting error.

Compared with the traditional intelligent models, ANN and Deep Neural Networks (DNN-based) models can provide a more accurate traffic prediction due to their complex structure and nonlinear functional

approximation. In which Long Short-Term Memory network (LSTM), Gated Recurrent Unit (GRU) performs well as it can learn about the inherent regularity of time series data, and the linear and nonlinear correlations with short-long distances. Tian et al. [11] presented multiscale temporal smoothing model to infer the missing data and used a revised LSTM approach to learn the prediction residual and forecast traffic flow. To exploit the impact of temporal features on prediction, Mou et al. [12] proposed a T-LSTM model that inputs the addition of temporal information into T-LSTM to improve the accuracy of short-term traffic flow prediction effectively. Furthermore, space-time hybrid models expand traffic flow forecast for a single station to the whole road network by fully considering the influences of the surrounding stations. Wu et al. [13] proposed a DNN-based traffic flow prediction model, which makes full use of weekly/daily periodicity and automatically learns the importance of past traffic flow. Then, a convolutional neural network and GRU are used to mine the spatial and the temporal features of traffic flow. Bogaerts et al. [14] utilized Graph Convolution Network [14] and Graph Attention mechanism [15] to capture the spatial features of traffic, and LSTM is used to obtain both short-term and long-term laws of traffic flows. This combination improves the accuracy of the traffic flow. However, the DNN-based models perform well in traffic flow prediction; however, they are difficult to interpret and require a large amount of data and a great time and space to train [3].

To distinguish the regularities of traffic flow, Empirical Mode Decomposition (EMD) and Singular Spectrum Analysis (SSA) are applied to reduce the noise of traffic flow sequence. Chen et al. [16] introduced Ensemble EMD (EEMD) and Wavelet to suppress the potential data outliers and then provided LSTM to fulfill the traffic flow prediction. Kolidakis et al. [17] combined SSA with ANN to analysis time series and forecast road traffic volume. However, these algorithms of component decomposition probably eliminate some random elements in the traffic flow, which leads to a distortion of the raw traffic flow.

Considering the complexity characteristics of the traffic flow, this paper proposes a hybrid model (SSA-LSTM-SVR), which decomposes the flow into different components and, according to their respective features, employs different models to make forecasting. Firstly, SSA is used to decompose the traffic flow into one principle component and multiple random components. LSTM is used to make a short-term prediction of the principle component due to its regularity and periodicity. For the random components, the SVR model is introduced to predict them. The corresponding prediction results of all components are superimposed to form the final predictive value. At the same time, Particle Swarm Optimization (PSO) algorithm [18] is applied to optimize the decomposition parameters of the SSA-LSTM-SVR model.

The rest of the paper is organized as follows. Section 2 illustrates the methodology. Section 3 contains the experiment results. Section 4 makes deep discussion about the experiment results. Section 5 contains the conclusions and future work.

2. Materials and Methods

2.1. Singular Spectrum Analysis (SSA). Singular Spectrum Analysis (SSA) [19] is usually applied to the component decomposition and noise reduction of time series signals. In order to find out the regularities and random factors of the expressway traffic flow, this paper introduces the SSA algorithm to decompose the trajectory matrix of short-term traffic flow into principle and random components. SSA mainly includes five steps: embedding, singular value decomposition (SVD), grouping, diagonal average, and component decomposition (subcomponent extraction).

Step 1. Embedding. Given a traffic flow series of length T , $\mathbf{X}(t) = [x_1, x_2, x_3, \dots, x_T]$, and the embedding dimension $L (2 < L < T/2)$, $\mathbf{X}(t)$ is arranged with time delay to obtain an $L \times K$ track matrix:

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_K \\ x_2 & x_3 & x_4 & \dots & x_{K+1} \\ \vdots & \vdots & \vdots & & \vdots \\ x_L & x_{L+1} & x_{L+2} & \dots & x_T \end{bmatrix}, \quad (1)$$

where $K = T - L + 1$, and the embedded dimension L is also called window's length of $\mathbf{X}(t)$, usually being an integer multiple of the period of $\mathbf{X}(t)$.

Step 2. Singular Value Decomposition (SVD). SVD of $\mathbf{X}\mathbf{X}^T$ is performed to obtain L eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \dots \geq \lambda_L \geq 0$ and corresponding orthogonal eigenvectors $\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3, \dots, \mathbf{U}_L$. Let $\mathbf{V}_i = \mathbf{X}^T \mathbf{U}_i / \sqrt{\lambda_i} (i = 1, 2, 3, \dots, L)$ and $d = \text{rank}(\mathbf{X}) = \max\{i: \lambda_i > 0\}$ and

$$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 + \mathbf{X}_3 + \dots + \mathbf{X}_d, \quad (2)$$

where $\mathbf{X}_i = \sqrt{\lambda_i} \mathbf{U}_i \mathbf{V}_i^T (i = 1, 2, 3, \dots, d)$. $\sqrt{\lambda_i}$ is the singular value of $\mathbf{X}$, $\mathbf{U}_i$ is the left eigenvector, $\mathbf{V}_i$ is the right eigenvector, and $(\sqrt{\lambda_i}, \mathbf{U}_i, \mathbf{V}_i)$ is a trip of $\mathbf{X}_i$.

Step 3. Grouping. Decomposing each $\mathbf{X}_i (i = 1, 2, 3, \dots, d)$ into p groups, then

$$\mathbf{X}_i = \mathbf{X}_i^1 + \mathbf{X}_i^2 + \mathbf{X}_i^3 + \dots + \mathbf{X}_i^p. \quad (3)$$

Given the contribution rate of $\mathbf{X}_i$ to $\mathbf{X}$ as $\lambda_i / \sum_{i=1}^d \lambda_i$, the contribution rate of the top r $\mathbf{X}_i$ to $\mathbf{X}$ will be

$$C_r = \frac{\sum_{i=1}^r \lambda_i}{\sum_{i=1}^d \lambda_i}. \quad (4)$$

Step 4. Diagonal Averaging (Reconstruction Component). Obtain the diagonal average values of $\mathbf{X}_i = (z_{ij})_{L \times K}$, and convert $\mathbf{X}$ into the time series $\mathbf{R}_i(t) = [r_1, r_2, r_3, \dots, r_T]$ of length T . This process is called reconstruction component (RC) and the sum of RCs is equal to the original sequence; namely,

$$\mathbf{X}(t) = \sum_{i=1}^d \mathbf{R}_i(t). \quad (5)$$

The k th member of $\mathbf{R}_i(t)$ is the average value of all members of $\mathbf{X}_i = (z_{ij})_{L \times K}$ with respect to $i + j = k + 1$, and the details are shown as follows:

$$r_k = \begin{cases} \frac{1}{k} \sum_{m=1}^k z_{m,k-m+1}, & 1 \leq k \leq L, \\ \frac{1}{L} \sum_{m=1}^L z_{m,k-m+1}, & L \leq k \leq K, \\ \frac{1}{T-k+1} \sum_{m=k-K+1}^L z_{m,k-m+1}, & K \leq k \leq T. \end{cases} \quad (6)$$

Step 5. Component Decomposition. By obtaining the diagonal average values of $\mathbf{X}_i = (z_{ij})_{L \times K}$, we obtain the time series $\mathbf{R}_i(t)$, $i = 1, 2, 3, \dots, d$. According to the subscript order $\{1, 2, 3, \dots, d\}$, it is divided into Γ parts, and the sum of the k th part $\{\mathbf{R}_i(t), \mathbf{R}_{i+1}(t), \dots, \mathbf{R}_{i+n-1}(t)\}$ constitutes a new time series with certain characteristics:

$$\mathbf{X}_k(t) = \sum_{j=i+0}^{j=i+n} \mathbf{R}_j(t), \quad (7)$$

where $i + n \leq d$ and $1 \leq k \leq \Gamma$. $\mathbf{X}_k(t)$ is the feature series of the original time series $\mathbf{X}(t)$, and each one has its own special features and they vary from each other. The sum of $\mathbf{X}_k(t)$ is equal to the original series; namely,

$$\mathbf{X}(t) = \sum_{k=1}^{\Gamma} \mathbf{X}_k(t). \quad (8)$$

In order to make the prediction result of short-term traffic flow as close as the ground truth, we firstly decompose the traffic flows into principle components and some random components by using SSA, where Particle Swarm Optimization (PSO) [19] is used to optimize the decomposition parameters of SSA. Then, according to their respective characteristics, the next section will discuss the corresponding prediction models suitable to respective characters.

2.2. Long Short-Term Memory Neural Network (LSTM). Since the principle component of expressway traffic flow decomposed by SSA can well reflect the periodicity and regularity of traffic flow, LSTM is introduced to make prediction. LSTM is an improvement of Recurrent Neural network (RNN) [20], which effectively solves the problems of gradient disappearance and gradient explosion of RNN. As LSTM can capture long short-term dependencies of a time series, it has been widely applied in the short-term traffic flow prediction. The hidden layer of LSTM is composed of several memory blocks, and each memory block contains a cell and three gates, namely, the input gate, the forget gate, and the output gate. The memory unit of LSTM is shown in Figure 1.

FIGURE 1: Memory cell of LSTM.

As Figure 1 shows, x_t and h_t represent the input and output at the interval t . f_t , g_t , and o_t are the outputs of the forget gate, input gate, and output gate, respectively, and their mapping functions are as follows:

$$\begin{aligned} f_t &= \sigma(W_f[h_{t-1}, x_t] + b_f), \\ g_t &= \sigma(W_g[h_{t-1}, x_t] + b_g), \\ J_t &= \text{ReLU}(W_J[h_{t-1}, x_t] + b_J), \\ o_t &= \sigma(W_o[h_{t-1}, x_t] + b_o), \\ s_t &= f_t * s_{t-1} + g_t * J_t, \\ h_t &= o_t * \text{ReLU}(s_t), \end{aligned} \quad (9)$$

where $\sigma(x) = (1 + e^{-x})^{-1}$ and $\text{ReLU}(x) = \max(0, x)$.

2.3. Support Vector Regression (SVR). As the stochastic component of expressway traffic flow has complex features and strong nonlinearity, the Support Vector Regression (SVR) algorithm is introduced to make prediction of these components. As a widely used traffic flow prediction model, SVR maps data points to an optimal hyperplane in the high-dimensional state space as far as possible to realize regression and prediction. For a sample set $\{(\mathbf{x}_i, y_i), i = 1, 2, 3, \dots, T\}$ of size T , in which x_i and y_i are the input and output values of the i th sample, the regression function of SVR is

$$f(\mathbf{x}) = \langle \mathbf{w} \cdot \mathbf{x}_i \rangle + b, \quad (10)$$

where $\langle \mathbf{w} \cdot \mathbf{x}_i \rangle$ represents inner product of vectors $\mathbf{w}$ and $\mathbf{x}_i$, and b is the offset. The loss function will be

$$L_\varepsilon(y, f(\mathbf{x})) = \begin{cases} 0, & |y - f(\mathbf{x})| \leq \varepsilon, \\ |y - f(\mathbf{x})|, & |y - f(\mathbf{x})| > \varepsilon, \end{cases} \quad (11)$$

where the error $\varepsilon > 0$. $\mathbf{w}$ and b can be obtained by solving the minimized objective function

$$\min_{\mathbf{w}} \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \frac{1}{n} \sum_{i=1}^n L_\varepsilon(y_i, f(\mathbf{x}_i)), \quad (12)$$

where $C > 0$ is the penalty factor. Among different kernel functions, the radial basis kernel function (RBF) is widely used due to its excellent nonlinearity, and the RBF is defined as

$$K_{\text{RBF}}(\mathbf{x}_i, \mathbf{x}) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_i\|^2}{2\gamma^2}\right), \quad (13)$$

where $\gamma > 0$. At the same time, Lagrange multipliers τ_i and τ_i^* are introduced to the duality problem of equation (12). By solving the duality problem under the KKT restriction, τ_i , τ_i^* , and b are obtained, and the regression function of SVR is described as

$$f(\mathbf{x}) = \sum_{i=1}^n (\tau_i - \tau_i^*) K_{\text{RBF}}(\mathbf{x}_i, \mathbf{x}) + b. \quad (14)$$

2.4. SSA-LSTM-SVR Structure. As shown in Figure 2, SSA decomposes the input traffic flow $\mathbf{X}(t)$ into Γ subcomponents with parameters, $\mu_1, \mu_2, \dots, \mu_\Gamma$, which are optimized by the PSO algorithm. According to formula (8), the sum of all components is equal to $\mathbf{X}(t)$ and the sum of μ_i is equal to 1; namely,

$$\sum_{i=1}^{\Gamma} \mu_i = 1. \quad (15)$$

The first principle component $\mathbf{X}_1(t)$ reflects the traffic volumes at different moments, the long-term regularities, and periodicity of the traffic flow. As discussed in Section 2.2, LSTM can capture these main and long-term law of the flow by iterating the inputs, so LSTM is used to forecast $\mathbf{X}_1(t)$. On the other hand, the left components, $\mathbf{X}_2(t), \dots, \mathbf{X}_\Gamma(t)$, mainly contain the short-term changes and fluctuations of the traffic flow, and SVR is more sensitive to short-term changes; therefore, SVR is applied to predict the left random flows, respectively. Finally, the corresponding prediction results, $\mathbf{Y}_1(t), \mathbf{Y}_2(t), \dots, \mathbf{Y}_\Gamma(t)$, of different components are superimposed to form the predictive value $\mathbf{Y}(t)$ of the original time series $\mathbf{X}(t)$ and expressed as

$$\mathbf{Y}(t) = \sum_{i=1}^{\Gamma} \mathbf{Y}_i(t). \quad (16)$$

3. Case Study

3.1. Dataset Description and Performance Indicators. The traffic flows of 347 toll stations of Guizhou province in China in January 2016 are obtained. Figure 3 shows the positions of 347 toll stations. The maximum daily average traffic volume reaches 14120, the minimum value is 5, and the traffic volumes varies greatly from site to site. According to the daily average traffic volumes, the toll stations are divided into three types, namely, less than 2000, 2000 to 6000, and above 6000. The traffic flows with different volumes from 46 stations (about 13.5% of toll stations) in one month are selected as experimental data to evaluate the performance of our algorithm. The intervals are set to 5, 10, 15, 30, 45, and 60 minutes (min).

Usually, mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE), accuracy, and goodness of fit (R^2) are used to evaluate the performances of predictive models, and they are defined as follows:

FIGURE 2: SSA-LSTM-SVR.

FIGURE 3: Positions of 347 highway toll stations of Guizhou province in China.

$$\begin{aligned}
 \text{MAE} &= \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|, \\
 \text{RMSE} &= \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}, \\
 \text{MAPE} &= \frac{1}{n} \sum_{i=1}^n \frac{|\hat{y}_i - y_i|}{y_i}, \\
 \text{accuracy} &= 1 - \frac{\sqrt{\sum_{i=1}^n (\hat{y}_i - y_i)^2}}{\sqrt{\sum_{i=1}^n y_i^2}},
 \end{aligned} \tag{17}$$

and

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}, \tag{18}$$

where $\hat{y}_i$ and y_i are predictive value and ground truth, and $\bar{y}$ is the mean of y_i . Low prediction errors (MAE, RMSE, and MAPE) and a high accuracy and R^2 imply a good prediction performance.

3.2. Characteristics of Traffic Flows and SSA Decomposition. Setting one day as the statistical period and 15 minutes (min) as the time interval, as an example, the traffic flows in the same historical periods in January 2016 of Guiyang North Toll Station are shown in Figure 4. It can be seen the traffic

FIGURE 4: Characteristics of traffic flows during 15 minutes at Guiyang North Station.

flows vary from day to day, with relatively fixed change trends and laws and strong randomness.

Set the population as 20, the maximum number of iterations as 100, the initial and termination weight factor as 0.9 and 0.2, both learning factors (ζ_1 and ζ_2) as 2, minimum and maximum moving speed as -1 and 1 , and the objective function as the goodness of fitting (R^2). PSO and SSA are applied on the traffic flows in 15 min of Guiyang North Toll Station from January 25 to 31, 2016. Then, the number of the components and their weights are obtained, and they are $\mu_1 = 0.3$, $\mu_2 = 0.3$, $\mu_3 = 0.3$, and $\mu_4 = 0.1$. Figure 5 demonstrates the original traffic flow and the decomposed 4 components. The principle component (the first one in Figure 5(b)) is basically consistent with the magnitude and change trends of the original traffic flow, and the traffic curve with obvious change rules is relatively smooth. The left three random components (Figures 5(c)–5(e)) have small magnitudes, large

FIGURE 5: Traffic flow decomposed by SSA. (a) Original traffic flow of a week, (b) the first component—principle component, (c) the second component—random component 1, (d) the third component—random component 2, and (e) the fourth component—random component 3.

random disturbances, and no obvious change rules. These characters will greatly increase the difficulty of prediction; however, their influences on the original traffic flow cannot be ignored and discarded as noise in traffic flow prediction.

3.3. Predictive Performance. With these decomposition parameters, the traffic flow of toll stations of Guiyang North Station in 5 min, 10 min, 15 min, 30 min, 45 min, and 60 min intervals are selected to verify the prediction performance of the SSA-LSTM-SVR model. Five prediction models, ARIMA, SVR, BP, KNN, and LSTM, are introduced for comparative analysis. These 5 comparative models have different implementation mechanisms, and they obtain the characteristics and variation laws of time series data and implement the traffic flow forecasting in different means. In which, ARIMA, BP, and LSTM directly obtain the temporal dependencies of the time serial, ARIMA is more sensitive to the short-term linear changes, and BP and LSTM can both capture the linear and nonlinear variations of the time series with long short-term dependencies, which are more robust and stable than the ARIMA. On the other hand, SVR and KNN adopt the hyperplane and similar pattern matching to achieve the traffic flow prediction.

As shown in Table 1, the MAE, RMSE, and MAPE of SSA-LSTM-SVR model are all minimum, and the fitting degree R^2 and prediction accuracy are higher than those of others. They all indicate that the SSA-LSTM-SVR model can better recognize the time series short-term change rules of the traffic flow and has a better prediction performance.

3.4. Predictive Performance on More Stations. To verify the generalization of the SSA-LSTM-SSR, with the same parameters, we carry out more experiments on other 45 toll

stations with different traffic volumes in whole month. Due to space limitation, only the prediction results in the 15 min, 30 min, and 60 min interval are shown in Table 2. The MAE, MAPE, and RMSE of the SSA-LSTM-SVR model are still the minimum, and the fitting degree R^2 and prediction accuracy are also higher than those of ARIMA, SVR, KNN, BP, and LSTM. This indicates that the SSA-LSTM-SVR model has a good generalization and better prediction performance than others. MAPE of the SSA-LSTM-SSR model is reduced by at least 5.00% and 6.33% on average, and the prediction accuracy is improved by at least 5.0% and 6% on average.

Let time interval be 15 min. SVR, LSTM, and SSA-LSTM-SVR are selected to further observe the differences between the predictive values and the ground truth on three stations with different traffic volumes. As Figure 6 shows, it can be seen that the predicted values of SSA-LSTM-SVR fit the ground truth well both in off-peak and peak hours, and it can capture the subtle changes of time series including the downward and upward tendencies. Overall prediction errors of SVR and LSTM are relatively large, especially in case of sudden changes of the traffic flow. As the traffic flow declines and ascends, LSTM and SVR cannot well capture the change tendencies in time.

4. Discussion

Any prediction model has its applicable scope, advantages, and disadvantages. For example, ARIMA is suitable for stable time series, while it is difficult to obtain the nonlinear relationship in the series. When the query mode does not appear in the history database, KNN will produce a large prediction error, and when the input time step of LSTM exceeds 32, the prediction performance will decline. Therefore, reasonable model selection is the basis for

TABLE 1: Performance comparison of prediction models (Guiyang North Station).

Time interval (min)	Models	MAE	MAPE	RMSE	Accuracy	R^2
5	ARIMA	6.4677	0.2394	9.0327	0.8574	0.9399
	SVR	6.6324	0.1778	9.273	0.8569	0.9369
	BP	7.6623	0.3671	10.1135	0.844	0.9250
	KNN	8.8914	0.2836	12.2265	0.8116	0.8905
	LSTM	6.8288	0.2261	9.364	0.8547	0.9348
	SSA-LSTM-SVR	4.7361	0.1284	6.821	0.9106	0.9421
10	ARIMA	11.1200	0.19999	15.6300	0.8763	0.9541
	SVR	10.7635	0.1765	15.388	0.8812	0.9559
	BP	11.2143	0.2161	15.6944	0.8788	0.9541
	KNN	14.6499	0.2261	20.4401	0.8422	0.9222
	LSTM	11.2955	0.1866	15.9469	0.8768	0.9525
	SSA-LSTM-SVR	8.8346	0.1206	13.2418	0.9385	0.9623
15	ARIMA	15.4965	0.1973	21.6791	0.8853	0.9603
	SVR	15.9333	0.1588	22.16	0.8857	0.9587
	BP	21.4699	0.3244	27.8068	0.8566	0.9351
	KNN	21.4068	0.2179	30.2227	0.8441	0.9238
	LSTM	16.1834	0.1913	22.2071	0.8855	0.9588
	SSA-LSTM-SVR	10.1942	0.1376	16.9009	0.9320	0.9758
30	ARIMA	30.6037	0.2118	42.1165	0.8885	0.9622
	SVR	31.1061	0.1402	43.5484	0.8877	0.9599
	BP	48.4023	0.5103	59.7906	0.8457	0.9246
	KNN	39.3415	0.2543	54.7551	0.8577	0.9367
	LSTM	38.1273	0.3719	49.2125	0.8724	0.9489
	SSA-LSTM-SVR	25.2886	0.1273	33.985	0.9324	0.9502
45	ARIMA	47.3747	0.1911	66.9107	0.8818	0.9573
	SVR	47.8773	0.1496	68.8685	0.8812	0.9552
	BP	73.554	0.4118	94.3918	0.837	0.9159
	KNN	57.8264	0.2016	78.2123	0.8634	0.9421
	LSTM	69.918	0.2824	97.7382	0.8297	0.9093
	SSA-LSTM-SVR	42.4357	0.0864	56.9149	0.9309	0.9528
60	ARIMA	74.2852	0.2414	104.810	0.8612	0.9410
	SVR	77.6223	0.1775	113.9067	0.8525	0.9309
	BP	125.1846	0.5853	157.7238	0.7952	0.8675
	KNN	85.3978	0.2353	114.6119	0.8499	0.9305
	LSTM	117.9721	0.3366	169.38	0.778	0.847
	SSA-LSTM-SVR	68.2209	0.1246	93.5078	0.9107	0.9493

accurate prediction. For the expressway traffic flow with periodic, sudden, and random changes, it is difficult for a single model to obtain all characteristics and accordingly make accurate short-time prediction.

In our SSA-LSTM-SVR model, SVD firstly decomposes the track matrix constructed by the traffic flow serials, which is equivalent to finding and extracting the intrinsic autocorrelation relationships of the traffic flow sequence. And then according to these correlations, SSA finds out the inherent change regularities and random features in the sequence through grouping, reconstruction, and component decomposition. These operations decompose the traffic flow serials into different components and are conducive to highlighting their respective regularities and characteristics and weakening the mutual influence caused by component hybridity. In order to fully take advantage of the acquired features of different components, SSA-LSTM-SVR puts forward the idea of separately forecasting. LSTM is more suitable for the principle component with great regularity and smoothness, while SVR is more suitable for

the left components with great random and sudden changes. Finally, in terms of respective predictive results, the SSA-LSTM-SVR model achieves a more accurate prediction performance, which actually reflects the short-term changes of the expressway traffic flow. In SSA-LSTM-SVR, SSA implements components decomposition according to the linear autocorrelation of the traffic flow serial, and the LSTM and SVR capture nonlinear relationships of different components. Therefore, compared with the ARIMA, KNN, BP, SVR, and LSTM models, the proposed SSA-LSTM-SVR model can catch more features and accordingly achieves a better prediction performance. In the SSA-LSTM-SVR, the differences between the predicted value and the real value of are the least, and the fitting effects are the best.

The above experiments and discussion illustrate that it is difficult for a single model to fully capture the characteristics of the complicated traffic flow sequence of the expressway, while a mixed model will be a more reasonable choice.

TABLE 2: The performance of different methods on Guiyang City Road Toll Stations.

Traffic volume (TV)	Models	Time interval being 15 min/30 min/60 min				
		MAE	MAPE	RMSE	Accuracy	R^2
15 stations ($TV \geq 6000$)	ARIMA	12.33/25.51/61.83	0.29/0.31/0.43	17.37/35.46/86.77	0.84/0.84/0.81	0.92/0.92/0.89
	SVR	12.68/25.74/61.53	0.19/0.20/0.23	17.96/36.89/93.12	0.84/0.84/0.80	0.92/0.91/0.87
	KNN	15.83/30.30/56.46	0.27/0.25/0.24	22.65/43.35/78.75	0.80/0.81/0.81	0.87/0.87/0.90
	BP	16.15/37.43/92.90	0.57/0.67/0.88	20.86/47.83/118.68	0.82/0.80/0.75	0.90/0.87/0.82
	LSTM	13.71/27.16/84.17	0.39/0.40/0.50	18.54/37.87/119.20	0.84/0.83/0.75	0.92/0.91/0.82
	SSA-LSTM-SVR	9.47/18.64/45.51	0.15/0.14/0.18	15.92/29.05/56.83	0.89/0.89/0.86	0.94/0.94/0.91
15 stations ($2000 \leq TV < 6000$)	ARIMA	6.86/12.33/24.49	0.84/0.38/0.41	9.58/17.13/33.55	0.80/0.82/0.82	0.89/0.90/0.90
	SVR	7.02/12.44/24.50	0.23/0.23/0.25	9.83/17.63/34.79	0.80/0.82/0.82	0.88/0.90/0.89
	KNN	8.41/14.44/26.95	0.35/0.31/0.36	12.01/20.56/38.92	0.75/0.79/0.80	0.82/0.86/0.88
	BP	8.12/17.05/35.58	0.64/0.74/0.82	10.49/21.06/43.30	0.78/0.78/0.77	0.86/0.86/0.84
	LSTM	6.95/13.31/30.57	0.38/0.44/0.56	9.58/17.59/41.23	0.80/0.82/0.79	0.88/0.90/0.87
	SSA-LSTM-SVR	5.22/10.01/20.51	0.19/0.18/0.20	7.27/15.36/28.07	0.85/0.88/0.88	0.90/0.93/0.93
15 stations ($TV < 2000$)	ARIMA	4.51/8.03/16.57	0.43/0.50/0.62	6.33/11.41/23.45	0.74/0.76/0.75	0.82/0.85/0.83
	SVR	4.55/8.08/16.68	0.26/0.26/0.31	6.46/11.73/24.66	0.74/0.76/0.75	0.82/0.84/0.82
	KNN	5.79/10.18/18.45	0.44/0.41/0.69	8.38/15.14/27.76	0.66/0.69/0.71	0.69/0.73/0.76
	BP	5.15/9.83/21.65	0.64/0.81/1.09	6.85/12.86/27.49	0.72/0.74/0.72	0.79/0.81/0.78
	LSTM	4.56/8.81/19.00	0.44/0.62/0.90	6.49/12.15/26.33	0.74/0.75/0.72	0.81/0.83/0.79
	SSA-LSTM-SVR	3.57/7.58/12.21	0.21/0.18/0.19	5.49/10.54/18.46	0.80/0.83/0.84	0.85/0.88/0.86

FIGURE 6: Continued.

FIGURE 6: The prediction results of SVR, LSTM, and SSA-LSTM-SVR.

5. Conclusion and Future Work

The real-time prediction of expressway traffic flow is always an important topic in traffic management. Due to the inherent complexity of expressway traffic flow, the prediction accuracy of a single model is limited; thus, the idea of combined prediction model is proposed in this paper. First, in terms of the characteristics of the traffic flow, it is decomposed into 1 principle component and 3 random components by SSA. As the magnitude and change regularities of the principle component are basically consistent with the original traffic flow, with smoother changes, LSTM is applied to capture these features and make short-term forecast. Since the left three random components have small magnitude, large random disturbance, and no obvious change rule, SVR is introduced to forecast, respectively. The estimates of each component are then added together to obtain the final forecast values. Experiments on the actual traffic flows of Guizhou expressway show that compared with the ARIMA, LSTM, KNN, BP, and SVR models, the SSA-LSTM-SVR model has a better prediction performance, whose differences between the predicted value and the real value are the least, and the fitting goodness is the best. Meanwhile, MAPE of SSA-LSTM-SSR is reduced by at least 5.0% and 6.33% on average, and the prediction accuracy is improved by at least 5.0% and 6% on average.

From the findings, we would like to suggest that if the time series has complex characteristics, such as regularity, suddenness, and randomness, it may be an optimal choice to use the mixed model for short-term prediction. This paper only considers the traffic flows of single toll station. In the future research, the interaction affection of traffic flows between toll stations in the expressway network will be further discussed, so as to achieve a more accurate short-term traffic flow prediction.

Data Availability

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

Conflicts of Interest

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

References

- [1] L. N. N. Do, N. Taherifar, and H. L. Vu, "Survey of neural network-based models for short-term traffic state prediction," *Wiley Interdisciplinary Reviews-Data Mining and Knowledge Discovery*, vol. 9, no. 1, 2019.
- [2] P. Xie, "Urban flow prediction from spatiotemporal data using machine learning: a survey," *Information Fusion*, vol. 59, pp. 1–12, 2020.
- [3] 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, 2015.
- [4] C. Ding, J. Duan, Y. Zhang, X. Wu, and G. Yu, "Using an ARIMA-GARCH modeling approach to improve subway short-term ridership forecasting accounting for dynamic volatility," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 4, pp. 1054–1064, 2018.
- [5] Q. Z. Hou, J. Leng, G. Ma, W. Liu, and Y. Cheng, "An adaptive hybrid model for short-term urban traffic flow prediction," *Physica A-Statistical Mechanics and Its Applications*, vol. 527, 2019.
- [6] L. R. Cai, Q. Chen, W. Cai, and X. Xu, "SVRGSA: a hybrid learning based model for short-term traffic flow forecasting," *IET Intelligent Transport Systems*, vol. 13, no. 9, pp. 1348–1355, 2019.
- [7] P. K. Li, C. Ma, J. Ning, Y. Wang, and C. Zhu, "Analysis of prediction accuracy under the selection of optimum time granularity in different metro stations," *Sustainability*, vol. 11, no. 19, 2019.
- [8] X. Feng, X. Ling, H. Zheng, Z. Chen, and Y. Xu, "Adaptive multi-kernel SVM with spatial-temporal correlation for short-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 6, pp. 2001–2013, 2019.
- [9] F. G. Habtemichael and M. Cetin, "Short-term traffic flow rate forecasting based on identifying similar traffic patterns," *Transportation Research Part C: Emerging Technologies*, vol. 66, pp. 61–78, 2016.
- [10] S. F. Cheng, F. Lu, P. Peng, and S. Wu, "Short-term traffic forecasting: an adaptive ST-KNN model that considers spatial

- heterogeneity,” *Computers, Environment and Urban Systems*, vol. 71, pp. 186–198, 2018.
- [11] Y. Tian, K. Zhang, J. Li, X. Lin, and B. Yang, “LSTM-based traffic flow prediction with missing data,” *Neurocomputing*, vol. 318, pp. 297–305, 2018.
 - [12] L. T. Mou, P. Zho, H. Xie, and Y. Chen, “T.-Lstm: A long short-term memory neural network enhanced by temporal information for traffic flow prediction,” *IEEE Access*, vol. 7, Article ID 98053, 2019.
 - [13] Y. K. Wu, H. Tan, L. Qin, B. Ran, and Z. Jhang, “A hybrid deep learning based traffic flow prediction method and its understanding,” *Transportation Research Part C: Emerging Technologies*, vol. 90, pp. 166–180, 2018.
 - [14] T. Bogaerts, A. D. Masegosa, S. A. Z. Juan, E. Onieva, and P. Hellinckx, “A graph CNN-LSTM neural network for short and long-term traffic forecasting based on trajectory data,” *Transportation Research Part C: Emerging Technologies*, vol. 112, pp. 62–77, 2020.
 - [15] T. Zhang and G. Guo, “Graph attention LSTM: a spatio-temporal approach for traffic flow forecasting,” *IEEE Intelligent Transportation Systems Magazine*, 2020.
 - [16] X. Chen, H. Chen, Y. Yang et al., “Traffic flow prediction by an ensemble framework with data denoising and deep learning model,” *Physica A: Statistical Mechanics and Its Applications*, vol. 565, Article ID 125574, 2021.
 - [17] S. Kolidakis, G. Botzoris, V. Profillidis, and P. Lemonakis, “Road traffic forecasting-a hybrid approach combining artificial neural network with singular spectrum analysis,” *Economic Analysis and Policy*, vol. 64, pp. 159–171, 2019.
 - [18] P. Adam and A. Kotsialos, “Second order macroscopic traffic flow model validation using automatic differentiation with resilient backpropagation and particle swarm optimisation algorithms,” *Transportation Research Part C: Emerging Technologies*, vol. 71, pp. 356–381, 2016.
 - [19] H. Hassani and A. Zhigljavsky, “Singular spectrum analysis: methodology and application to economics data,” *Journal of Systems Science and Complexity*, vol. 22, no. 3, pp. 372–394, 2009.
 - [20] H. Sak, A. Senior, and F. Beaufays, “Long short-term memory recurrent neural network architectures for large scale acoustic modeling,” in *Proceedings of the 15th Annual Conference of the International Speech Communication Association (Interspeech 2014)*, vol. 1–4, pp. 338–342, Singapore, September 2014.