Research Article

A Hybrid Short-Term Traffic Flow Multistep Prediction Method Based on Variational Mode Decomposition and Long Short-Term Memory Model

Qichun Bing, Fuxin Shen, Xiufeng Chen, Weijian Zhang, Yanran Hu, and Dayi Qu

College of Mechanical and Automotive Engineering, Qingdao University of Technology, Qingdao 266520, China

Correspondence should be addressed to Xiufeng Chen; chenhill7765@163.com

Received 11 April 2021; Revised 30 July 2021; Accepted 13 September 2021; Published 12 October 2021

1. Introduction

High-precision traffic flow prediction information is definitely considered as one of the most important parts in intelligent transportation systems. The accurate and timely traffic flow prediction information can not only provide strong support for travel decisions but also transform traffic management mode. For travelers, it would enable them to plan their trips ahead of time and timely adjust their trip mode and trip route with the dynamic short-term traffic prediction information. For traffic managers, the short-term traffic flow prediction information would enable them to take traffic control measures early enough to avoid traffic congestion rather than to deal with the traffic problems after the traffic congestion has already occurred [1]. Prediction within 15 minutes ahead is typically called short-term prediction. However, owing to the randomness and volatility characteristics of the traffic flow data, it is difficult to develop the satisfactory short-term traffic flow prediction models.

2. Related Works

In the literature available, many scholars focus on the short-term traffic flow prediction models in the last few decades. These existing methods include regression method [2, 3], time series model [4–6], Kalman filtering methods [7, 8], support vector machine regression model [9–11], Bayesian model [12, 13], hidden Markov model [14], and neural network methods [15–19]. Considering the volatility of traffic flow and limitations of single prediction models, more...
and more hybrid prediction models have been proposed. For example, Hou et al. proposed an adaptive hybrid short-term traffic flow prediction method. Firstly, the linear autoregressive integrated moving average model and nonlinear wavelet neural network model were used to predict short-term traffic flow. Then, prediction results of the two single models were analyzed by fuzzy logic, and the weighted result was regarded as the final predicted results [20]. Zhang et al. proposed a traffic flow multistep forecasting method by decomposing the data into three modeling components: an intraday or periodic trend by introducing the spectral analysis technique, a deterministic part modeled by the ARIMA model, and the volatility estimated by the GJR-GARCH model [21]. Moretti et al. provided a one-hour forecast of urban traffic flow rates by combining artificial neural networks and a simple statistical approach [22].

In recent years, with the rapid development of data science, the data-driven models for short-term traffic flow prediction have attracted wide attention. Generally, these data-driven models mainly consist of the data processing algorithms and machine learning algorithms. The data processing algorithms are usually used for data denoising [23, 24], data decomposition, and data reconstruction [25, 26], which can effectively extract the significant features of the original traffic flow data [27]. Among the various data processing algorithms, the wavelet decomposition (WD) [28] and empirical mode decomposition (EMD) are the two mainstream algorithms for short-term traffic flow forecasting. Additionally, some improvement algorithms have also been extensively utilized, such as wavelet packet decomposition (WPD), ensemble empirical mode decomposition (EEMD), and complete ensemble empirical mode decomposition (CEEMDAN). Numerous research studies have proved the effectiveness of these algorithms. For example, Zhang et al. [29] presented a short-term traffic flow prediction method based on improved wavelet packet analysis and long short-term memory neural network model. Wei et al. [30] proposed a hybrid method by combining EMD and backpropagation neural networks (BPN) to predict the short-term passenger flow in metro systems. Yang [31] proposed a hybrid traffic flow multistep prediction method based on EMD and stacked autoencoder model. Li et al. [32] proposed a travel time prediction model based on EEMD and random vector functional link network. EEMD is employed to decompose the complex travel time data series into several simple functions. Tian et al. [33] presented a CEEMDAN-PE-OSELM forecasting model based on CEEMDAN, permutation entropy (PE), and online sequential extreme learning machine (OSELM). However, there are still many shortcomings in these algorithms. For instance, the decomposition effect of WD and WPD depends on the choice of the basis function and has poor adaptability, while EMD and its improved algorithms have the problems of endpoint effect and over envelope. In order to solve these drawbacks, a novel data decomposition algorithm called variational mode decomposition (VMD) is put forward. The VMD algorithm has been verified to have superior prediction ability in wind speed prediction [34], foreign exchange rates’ prediction [35], and carbon price prediction [36].

The machine learning algorithms are widely applied in the field of traffic flow forecasting because of their excellent performance in data mining and pattern recognition. Zhu et al. [37] put forward a new method based on RBF neural network to predict the short-term traffic volume. Vanajakshi et al. [38] compared the effect of the support vector machine model and the artificial neural network model in short-term traffic flow prediction. Wei et al. [39] integrated the time-varying deviation of the day-to-day traffic variation into support vector regression and proposed an adaptive support vector regression short-term traffic forecasting model. Recently, different types of deep learning algorithms, such as long short-term memory (LSTM), convolutional neural network (CNN), deep belief network (DBN), and recurrent neural network (RNN), have generated great interest among research studies due to their great ability to capture complex nonlinear relationships of traffic flow. Ma et al. [40] developed a long short-term memory neural network model for short-term travel speed prediction, and the proposed algorithm can determine the optimal time window for time series in an automatic manner. Polson [41] proposed an innovative deep learning architecture to predict traffic flows and showed that deep learning architectures can capture nonlinear spatio-temporal effects. Wu et al. [42] designed a novel traffic flow prediction model based on deep learning framework. Guo et al. [43] proposed a novel end-to-end deep learning framework to model the complex explanatory variables for traffic forecasting. Yu et al. [44] proposed graph-based learning models for traffic forecasting that mimicked real traffic propagation based on variable adjacency matrices.

Motivated by the successful applications of models with attention, this paper will put forward a hybrid short-term traffic flow multistep prediction method based on the VMD algorithm and the LSTM model. The main contributions of this study are summarized as follows. (1) The variational mode decomposition algorithm is employed to decompose the original traffic flow data into several intrinsic mode function (IMF) components. (2) Different LSTM models are established to predict different IMF components, and the component prediction results are integrated to obtain the final traffic flow prediction values. (3) Comprehensive performance comparisons are conducted using inductive loop data measured from the north-south viaduct in Shanghai.

The rest of this paper is organized as follows. Section 3 introduces the theoretical basis of relevant algorithms and the framework of the proposed hybrid method. Section 4 demonstrates case study where the prediction results of the proposed method and other involved methods are evaluated. Finally, Section 5 provides conclusion of this research.

3. Methodology

3.1. Variational Mode Decomposition Algorithm. The variational mode decomposition (VMD) algorithm is a novel process to examine nonstationary and nonlinear signals in signal processing, which can decompose a signal into a series of modes with specific bandwidth in spectral domain [45]. Each mode can be compacted around a center pulsation determined during the decomposition process. The whole framework of the variational mode decomposition is the
variational problem, which mainly includes the construction of the variational problem and its solution. The time-series data of traffic flow is regarded as nonstationary signal $f$, and the variational problem is described as seeking $K$ mode functions $u_k(t)$ ($k = 1, 2, \ldots, K$). In order to obtain the bandwidth of each mode, three steps should be fulfilled [46]:

(1) For each mode function $u_k(t)$, the Hilbert transform is adopted to obtain unilateral frequency spectrum:

$$
\delta(t) = \begin{cases} 
0 & t \neq 0 \\
\infty & t = 0 
\end{cases},
\int_{-\infty}^{\infty} \delta(t)dt = 1.
$$

(2) For each mode function $u_k(t)$, the frequency spectrum is modulated to the corresponding fundamental frequency band by applying an exponential tuned to the respective estimated center frequency:

$$
\left[\delta(t) + \frac{j}{\pi t}\right] \cdot u_k(t) e^{-j\omega_k t}.
$$

(3) The bandwidth of $u_k(t)$ could be estimated by utilizing the Gaussian smoothness of the modulation signal. Then, the constrained variational problem can be expressed as

$$
\min_{u_k, w_k} \left\{ \sum_{k=1}^{K} [\delta(t) + j/\pi t] \cdot u_k(t) e^{-j\omega_k t} \right\},
$$

s.t. \sum_{k=1}^{K} u_k = f(t),

where $f(t)$ indicates the original signal, $t$ indicates the time interval, $K$ denotes the number of modes, $u_k$ denotes the $k$th mode, $\delta(t)$ denotes the Dirac distribution, and $w_k$ is the center frequency.

In order to convert the above optimization problem into an unconstrained one, the penalty term and Lagrangian multipliers are employed, which can be shown as

$$
L(u_k, w_k, \lambda) = \alpha \sum_{k=1}^{K} \left[ [\delta(t) + j/\pi t] \cdot u_k(t) e^{-j\omega_k t} \right]^2 + \left\| f(t) - \sum_{k=1}^{K} u_k(t) \right\|^2_2 + \left\langle \lambda(t), f(t) - \sum_{k=1}^{K} u_k(t) \right\rangle,
$$

where $\alpha$ is the second penalty factor and $\lambda(t)$ is Lagrange multiplication operator.

The alternating multiplier direction method is used to solve the unconstrained problem in formula (4), which can find the saddle point of the augmented Lagrangian. Based on the alternating multiplier direction method, $u_k$ and $w_k$ can be continuously updated in two directions to complete the variational mode decomposition.

The solution of the optimization problem for $u_k$ can be explained as follows:

$$
\tilde{u}_{k}^{n+1}(w) = \frac{\tilde{f}(w) - \sum w \tilde{u}_i(w) + \bar{\lambda}(w)/2}{1 + 2\alpha(w - \tilde{u}_k)^2},
$$

where $n$ is the number of iterations $\tilde{f}(w), \tilde{u}_i(w), \bar{\lambda}(w),$ and $\tilde{u}_{k}^{n+1}(w)$ are the Fourier transforms of $f(t), u_i(t), \lambda(t),$ and $\tilde{u}_{k}^{n+1}(t),$ respectively.

The solution of the optimization problem in frequency domain for $w_k$ is as follows:

$$
\tilde{w}_k^{n+1} = \frac{\int_{-\infty}^{\infty} w|\tilde{u}_k(w)|^2 dw}{\int_{-\infty}^{\infty} |\tilde{u}_k(w)|^2 dw}.
$$

The specific process of VMD decomposition is as follows:

Step 1: initialize $\{u_k^1\}, \{w_k^1\}$, and $\lambda^1$, and set the number of iterations to 1.

Step 2: set $n = n + 1$. For $1 = 1$: $K$, update $\tilde{u}_k^{n+1} (w)$ and $\tilde{u}_k^{n+1}$ according to formulas (5) and (6).

Step 3: for all $w \geq 0$, update $\lambda^n$ according to the following formula:

$$
\lambda^{n+1} (w) = \lambda^n (w) + \tau \left( \tilde{u}(w) - \sum_{k=1}^{K} \tilde{u}_k^{n+1}(w) \right).
$$

Step 4: repeat Steps 2 to 3 until iteration constraints are met, that is,

$$
\sum_{k} \left[ \| u_k^{n+1} - \tilde{u}_k^{n+1} \|^2 \right]_2.
$$

3.2. Prediction Model Based on LSTM. Long short-term memory (LSTM) network is a special kind of recurrent neural network (RNN). Compared with the traditional RNN, LSTM adds a cell state in the hidden layer, which solves the problem that RNN is easy to fall into gradient disappearance or gradient explosion. LSTM network consists of the input layer, the hidden layers, and the output layer. The main characteristics of the LSTM are the memory cells in its hidden layers, which contain memory blocks rather than traditional neuron nodes. Each block has several self-connected memory cells and three multiplicative units, input, output, and forget gates. The network structure of LSTM is shown in Figure 1.

The LSTM generates a mapping from an input sequence vectors to an output probability vector by calculating the network unit activations using the following equations:

\begin{align*}
    i_t &= \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \\
    f_t &= \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \\
    o_t &= \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \\
    \bar{C}_t &= \tanh(W_c \cdot [h_{t-1}, x_t] + b_c), \\
    C_t &= \bar{C}_t + f_t \cdot C_{t-1}, \\
    h_t &= o_t \cdot \tanh(C_t)
\end{align*}
subsequence. The number of decomposition subsequences $K$ is determined by observing the central frequency. When the central frequency of the last layer is relatively stable, $K$ can be considered as the best value.

Step 2: different LSTM models are established to predict different IMF components obtained by VMD.

Step 3: for each prediction model, one-step, two-step, and three-step predictions are conducted. The multi-step predictions are carried out by introducing the prediction result of the previous step into the predicting model.

Step 4: the component prediction results of each step are accumulated to obtain the final traffic flow multi-step prediction values.

4. Case Study

4.1. Experiment Data and Evaluation Criteria. The north-south viaduct expressway of Shanghai, China, is selected as the study scenario; the satellite map of selected urban expressway is shown in Figure 3. This section of expressway runs from Yandong interchange to Gonghe interchange, which includes 24 main line detection sections and 30 ramp detection sections. A total of 88 main line coil detectors and 60 ramp coil detectors are arranged. The average distance between the main line detectors is about 500 meters. The data collection time period starts from August 27 to August 31 in 2018. The sampling interval of the data is set to 5 minutes. Figures 4 and 5 give the traffic flow data of NBDX 16(2) and NBXX 11(3) collecting from five consecutive days.

In this paper, three evaluation criteria are employed to judge the performances of proposed method: the root of mean square error (RMSE), the mean absolute error (MAE), and the mean absolute percentage error (MAPE). The less the values of these criteria are, the closer the prediction results are to the original traffic flow values. The evaluation criteria are defined as follows:

\[
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \tilde{y}_i)^2},
\]

\[
\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |y_i - \tilde{y}_i|,
\]

\[
\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \tilde{y}_i}{y_i} \right| \times 100%,
\]

where $y_i$ represents the actual data and $\tilde{y}_i$ represents the prediction data and $n$ is the number of samples.

4.2. VMD Results. Before VMD of traffic flow data, it is necessary to determine the number of decomposition subsequences $K$. If the value of $K$ is too small, the mode aliasing phenomenon will appear in the decomposed subsequence. If the value of $K$ is too large, the mode repetition or additional noise will be produced.
In this paper, the value of $K$ is determined by observing the central frequency. Firstly, the original traffic flow time-series data is decomposed by VMD at different $K$ values. Then, the center frequency in different layers is calculated. Next, the center frequency of IMF component in the last layer at different $K$ values is compared in order to obtain the best value of $K$. When the central frequency of the last layer is relatively stable, $K$ can be considered as the best value. The parameters of VMD algorithm are set as follows: $\alpha = 2000$, $\tau = 0.3$, and $\epsilon = 10^{-7}$. It can be seen from Tables 1 and 2,
when $K$ is greater than 4, the center frequency of IMF component in the last layer remains relatively stable, so $K = 5$ is selected as the decomposition number in the experiment.

The decomposition results of traffic flow data are illustrated in Figures 6 and 7.

In Figures 6 and 7, the display on the left is the five intrinsic mode functions obtained by the VMD, and the right side is the spectrums corresponding to the figures on the left. It is obvious that the signals on different scales exhibit completely different characteristics and the frequency of each component increases gradually from IMF1 to IMF5. The VMD algorithm decomposes the highly complex traffic flow data into multiple periodic and regular components, which is conducive to discover and learn more information contained in the traffic flow data and make the prediction results more accurate.

4.3. The Experimental Results of the Proposed Model. In order to evaluate the multistep prediction performance of the proposed method, one-step predictions to three-step predictions are carried out. The number of layers of the LSTM model is set to 4 which includes input layer, output layer, and two LSTM layers. The number of neurons is set to 128 and the input time window is 4. According to the different multistep prediction logic, the multistep prediction method can be divided into two types: direct multistep prediction and iterative multistep prediction. In the direct multistep prediction method, many different prediction models need to be established to realize multistep prediction. Moreover, the larger the prediction time span, the lower the correlation between the current input variable and the corresponding prediction object, which will lead to the increase of prediction error. In the iterative multistep
prediction method, introducing the prediction value into the input information will lead to certain error conduction, especially when there are too many prediction steps. According to the research results of [47], the prediction effect of iterative multistep prediction is better than that of direct multistep prediction. Therefore, iterative prediction is adopted in the multistep prediction process in this paper. As shown in Figure 8, when predicting $y_{t+2}$, the input variables are $\{y_{t-p+1}, \ldots, y_t, y_{t+1}\}$.

Figures 9–14 give the one-step to three-step prediction results of nbdx 16(2) and nbxx 11(3). The blue curve represents the actual traffic flow data and the red curve represents the prediction data.

As can be directly seen from Figures 9–14, the prediction results of the proposed method not only accurately describe the dynamic fluctuation trend of the original traffic flow sequence but also are similar to the real traffic flow data in the numerical value. Therefore, it can be concluded that the
Figure 8: Multistep prediction schematic diagram.

Figure 9: The one-step prediction results for nbdx16(2).

Figure 10: The two-step prediction results for nbdx16(2).

Figure 11: The three-step prediction results for nbdx16(2).
proposed hybrid method could effectively improve the prediction accuracy.

4.4. Experimental Results’ Comparison of Different Models.
In order to avoid randomness and receive reliable results, five-fold cross validation is used. As shown in Figure 15, the traffic flow dataset for five consecutive days is separated into five parts and five repeated experiments are conducted. For each experiment, four parts are taken as the training dataset to train the LSTM model, and the remainder is applied as testing dataset to test the trained model. The overall performance of the model was evaluated by the average criteria of the five repeated experiments.

In order to verify the superiority of the proposed algorithm in the paper, six methods are employed for comparing, which are included as the SVM model, LSTM model, WD-LSTM model, EMD-LSTM model, VMD-BP model, and VMD-SVM model. One-step predictions to three-step predictions are provided for all the involved models. The penalty parameter $C$ and kernel parameter $\sigma$ of the SVM model are set to 0.5 and 125, respectively. The number of layers of the BP neural network model is set to 3, which includes input layer, output layer, and a hidden layer. The
Table 3: Error results of different prediction methods for NBDX 16(2).

<table>
<thead>
<tr>
<th>Prediction method</th>
<th>Evaluation criterion</th>
<th>One-step</th>
<th>Two-step</th>
<th>Three-step</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>MAE</td>
<td>8.65</td>
<td>9.07</td>
<td>9.66</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>9.64</td>
<td>10.47</td>
<td>10.85</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>9.26%</td>
<td>9.82%</td>
<td>10.34%</td>
</tr>
<tr>
<td>LSTM</td>
<td>MAE</td>
<td>7.22</td>
<td>7.76</td>
<td>8.14</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>8.71</td>
<td>9.38</td>
<td>9.90</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>7.84%</td>
<td>8.32%</td>
<td>8.89%</td>
</tr>
<tr>
<td>WD-LSTM</td>
<td>MAE</td>
<td>5.06</td>
<td>5.59</td>
<td>6.10</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>6.02</td>
<td>6.48</td>
<td>6.77</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>5.43%</td>
<td>5.85%</td>
<td>6.20%</td>
</tr>
<tr>
<td>EMD-LSTM</td>
<td>MAE</td>
<td>4.65</td>
<td>4.90</td>
<td>5.13</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>5.66</td>
<td>6.10</td>
<td>6.38</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>4.97%</td>
<td>5.47%</td>
<td>5.84%</td>
</tr>
<tr>
<td>VMD-BP</td>
<td>MAE</td>
<td>4.75</td>
<td>4.96</td>
<td>5.24</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>5.82</td>
<td>6.17</td>
<td>6.51</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>5.38%</td>
<td>5.74%</td>
<td>5.97%</td>
</tr>
<tr>
<td>VMD-SVM</td>
<td>MAE</td>
<td>4.54</td>
<td>4.79</td>
<td>4.93</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>5.47</td>
<td>5.90</td>
<td>6.13</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>4.82%</td>
<td>5.23%</td>
<td>5.57%</td>
</tr>
<tr>
<td>VMD-LSTM</td>
<td>MAE</td>
<td>4.39</td>
<td>4.65</td>
<td>4.84</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>5.27</td>
<td>5.62</td>
<td>5.87</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>4.55%</td>
<td>4.76%</td>
<td>4.98%</td>
</tr>
</tbody>
</table>

Table 4: Error results of different prediction methods for Nbx11(3).

<table>
<thead>
<tr>
<th>Prediction method</th>
<th>Evaluation criterion</th>
<th>One-step</th>
<th>Two-step</th>
<th>Three-step</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>MAE</td>
<td>8.82</td>
<td>9.46</td>
<td>10.07</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>10.22</td>
<td>10.77</td>
<td>11.05</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>9.58%</td>
<td>10.26%</td>
<td>10.77%</td>
</tr>
<tr>
<td>LSTM</td>
<td>MAE</td>
<td>7.54</td>
<td>7.85</td>
<td>8.26</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>8.92</td>
<td>9.63</td>
<td>10.14</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>8.05%</td>
<td>8.48%</td>
<td>8.96%</td>
</tr>
<tr>
<td>WD-LSTM</td>
<td>MAE</td>
<td>5.25</td>
<td>5.94</td>
<td>6.37</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>6.34</td>
<td>6.79</td>
<td>7.02</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>5.78%</td>
<td>6.17%</td>
<td>6.43%</td>
</tr>
<tr>
<td>EMD-LSTM</td>
<td>MAE</td>
<td>5.01</td>
<td>5.56</td>
<td>5.94</td>
</tr>
<tr>
<td></td>
<td>RMSE</td>
<td>5.82</td>
<td>6.21</td>
<td>6.55</td>
</tr>
<tr>
<td></td>
<td>MAPE</td>
<td>5.23%</td>
<td>5.74%</td>
<td>6.07%</td>
</tr>
</tbody>
</table>
number of neurons and the input time window are set to 16 and 4, respectively. In addition, in order to verify the broad adaptability of the proposed method, all experiments are based on data from both east main lane and west main lane. In each experiment, optimal parameters are set in these models. The number of decomposition subsequences is all 5 in the proposed model and the other hybrid models. All of the prediction experiments are carried out in Matlab R2018a. Tables 3 and 4 show the prediction performance of different models using data source collected from NBDX16(2) and NBXX11(3).

From the observation of prediction results in Table 3 and 4, several important conclusions can be drawn as follows:

1. It is worth noting that all hybrid prediction methods obviously outperform the single models in terms of prediction accuracy, which indicates that decomposition of traffic flow time series prior to forecasting can significantly improve the prediction performance.

2. The multistep prediction accuracy of the VMD-LSTM hybrid method is generally much better than the EMD-LSTM method and the WD-LSTM method. This indicates that VMD algorithm is a more effective decomposition algorithm.

3. The prediction error of the VMD-LSTM hybrid method is much lower than that of the VMD-SVM method and the VMD-BP method, which indicates that the prediction performance of the LSTM model is better than that of the traditional prediction model.

From the above experiments, we can conclude that the VMD-LSTM hybrid model has superiority and wide applicability.

**5. Conclusion**

To achieve accuracy short-term traffic flow forecasting, a hybrid short-term traffic flow multistep prediction method based on VMD algorithm and LSTM model is proposed. The main idea consists of two parts: the data processing technique and the performance of the prediction model. For the volatility and randomness of traffic flow, VMD algorithm is employed to transform the unstable traffic flow sequence into a series of relatively stable intrinsic mode function (IMF) components. The periodic characteristics of the IMF components are more obvious than the original traffic flow sequence. Considering the nonlinear fitting ability and self-learning adaptive characteristic of the LSTM model, different LSTM models are established to predict different IMF components, and the results are accumulated to obtain the final prediction results. For each prediction model, 1-step to 3-step iterative predictions are carried out. The traffic flow data measured from the north-south viaduct expressway in Shanghai are used to carry out the experiment. According to the experiment results, it can be concluded that (1) compared with single models, the proposed hybrid model has a greater improvement in short-term traffic flow prediction and (2) compared with EMD algorithm and WD algorithm, VMD algorithm can effectively solve the problems of modal aliasing and endpoint effects, which could make the periodic characteristics of each IMF more obvious and improve the prediction performance.

Overall, this paper succeeds in proving the possibilities to forecasting short-term traffic flow based on VMD algorithm and LSTM model. The results of this paper can be used for the implementation of an advanced traffic management system and advanced traveler information system. However, there are still several limitations in the current study. Firstly, the hyperparameters of the LSTM model are determined by reference to the existing literature. In future research, different parameter optimization algorithms could be used to obtain the optimal hyperparameters. Secondly, different time-slice data such as 1 min and 2 min can be tested, which could check the generalization ability of the proposed method. Thirdly, short-term traffic flow data in different scenarios such as arterial road and expressway can be used to test the effectiveness of the proposed method.

**Data Availability**

The data used to support the findings of the study are included within the article.

**Conflicts of Interest**

The authors declare no conflicts of interest.

**Acknowledgments**

This research was supported by the Key Research and Development Program of Shandong Province (Grant no.
2019GGX101038) and the Shandong Provincial Natural Science Foundation (Grant no. ZR2019MG012.)

References


[47] Q. Li, Research on Monitoring and Forecasting Method of Traffic State Based on Multi-Source Data, Jilin University, Changchun, China, Doctor Degree, 2013.