Hybrid Model for Method for Short-Term Traffic Flow Prediction Based on Secondary Decomposition Technique and ELM

LeiNa Zhao, XinYu Wen, YiMing Shao, and ZhenYu Tang

1 Institute of Mathematics and Statistics, Chongqing Jiaotong University, Chongqing 400074, China
2 School of Traffic and Transportation, Chongqing Jiaotong University, Chongqing 400074, China

Correspondence should be addressed to XinYu Wen; 18982736892@163.com

Received 17 March 2022; Revised 7 May 2022; Accepted 13 June 2022; Published 21 August 2022

1. Introduction

With the increasing wealth of material life, vehicles have become the most popular transportation means for the most urban residents. This phenomenon inevitably leads to an increase of cars driving on the urban roads, thereby yielding serious traffic jams [1]. Meanwhile, the existing road network capacity cannot match the rapid growth of vehicles [2]. Although some attempts (e.g., the expansion and renewal of traffic facilities) have been made, the above problem still exists and has to resort to some new technical development. Fortunately, the rapid development of information technology (e.g., intelligent transportation system, ITS) makes it possible to address above the problems. The ITS has the ability to realize traffic control and guidance at the same time, which plays the important role in reducing traffic accidents and congestion [3]. Especially, it is beneficial to provide a reference for the management department to make strategies in time, such as realizing effective control by adjusting the time interval of intersection lights. In the ITS, short-term traffic flow prediction is one of the key technologies that attract more and more attention during the recent decades. The prediction accuracy is directly related to the performance of traffic management and control, and the higher one indicates the more reliable real-time road information [4]. Therefore, the realization of accurate short-term traffic prediction is urgently needed.

Traffic prediction models have been developed over the past few years, and they can be roughly divided into two categories including single models and hybrid models. In actual cases, single models normally only explain one characteristic. However, the true short-term traffic flow contains complex characteristics. To this end, the hybrid prediction models become mainstream. The hybrid models based on commonly methods include parametrical optimization, single model combination, and data decomposition. In contrast with parametrical optimization and single model combination, the data decomposition method has the
capability to explain a variety of complex data characteristics in short-term traffic flow, which attracts the growing attention of researchers.

However, traditional prediction models based on decomposition methods exist some problems in the actual application that may cause some deviation in the prediction result. For example, the empirical mode decomposition (EMD) method is sensitive to noise and limited by the mode mixing problem. To enhance the accuracy of the short-term traffic prediction, this paper proposes a new hybrid model based on the secondary decomposition technique and ELM. The contributions of this paper can be summarized as follows: (1) time-varying filtering-EMD (TVF-EMD) is used to decrease nonstationary and remove the high-frequency noise of the original traffic flow. Then, we obtain the all IMFs. (2) LMD can further handle the IMFs to obtain smoother subseries. To decrease the nonstationary and avoid extreme oscillations by LMD, we propose decomposition conditions to ensure that the IMFs can be decomposed or not. (3) Extreme learning machine (ELM) is employed to forecast each subsequence decomposed by the secondary self-adaptive decomposition method (TVF-EMD-LMD) and obtain the final prediction value. Then, two groups of actual traffic flow data with different characteristics are used to verify the effectiveness of the proposed method.

The organization in this paper is listed as follows: Section 2 exhibits the proposed models and methods in short-term traffic flow. Section 3 briefly presents the basic theory of corresponding methods. Section 4 gives the construction steps of the proposed hybrid model. In Section 5, two groups of data of short-term traffic flow are used to verify the proposed model, and the superiority of the proposed model is illustrated by comparing it with the hybrid model based on the traditional decomposition method. The last chapter makes a summary of the full article.

2. Literature Review

In the past decades, various methods have been developed to improve the forecasting performance. These single models are mainly involved in autoregressive integrated moving average (ARIMA), Kalman filter (KF), and so on. These models usually feature a simple structure and mainly aim at explaining the linear component hidden in the data [5, 6]. However, short-term traffic flow not only contains linear components but also has a significant nonlinearity. These features greatly affect the prediction accuracy of the model. To this end, many scholars turn their focus on the research of nonlinear prediction models. Fortunately, with the rapid development of computer technology, artificial intelligence models (AI-based models) have become a convenient tool to solve the nonlinear problems in short-term traffic flow prediction. By comparison, it can provide more accurate prediction due to its strong adaptive learning and nonlinear mapping abilities. As a typical intelligent model used in this field, artificial neural network (ANN) usually has a higher forecasting accuracy than traditional time series models [7]. For example, Huang et al. [8] proposed a type of method, i.e., ELM model. In this model, the output matrix of samples is obtained by randomly generating parameters, and the output connection weights are determined by solving the generalized inverse matrix. So it has the ability to realize a fast calculation. To effectively forecast the short-term traffic flow, the researchers start adopt deep learning models. For instance, the deep belief network (DBN) is used to extract the features of the traffic flow, then adopt support vector regression to predict. Compared with the traditional AI-based models, the application of the DBN model can realize satisfactory prediction effectiveness [9]. In addition, the deep learning model is also adopted as a prediction machine to learn the corresponding obtained features [10]. More details about the superiority of the AI-based models can be found in reference [11].

Although the single models have achieved well prediction results, they cannot explain multiple characteristics (e.g., nonstationary and nonlinearity) of short-term traffic flow at the same time [12]. Fortunately, the hybrid model can explain them. A large number of facts show that the prediction effect of the hybrid models is better than that of the single models. For instance, a hybrid model based on the long and short memory model is proposed to predict traffic flow, and the prediction effect is better than the single model [13]. In addition, nonlinear and nonstationary are obstacles in time series prediction, which increase the difficulty of prediction work. In order to separate the above important features hidden in short-term traffic flow and decrease the nonstationary, the hybrid models based on the data decomposition method are an optional scheme to address this issue. The construction steps of this kind of hybrid model are roughly divided into three parts: (1) decompose the original short-term traffic flow by using the data decomposition method. (2) The proper predictors are established to forecast the decomposed subseries, respectively. (3) The final prediction result is obtained by superimposing the sub-prediction results.

The data decomposition method has become one of the key technologies of the hybrid model in recent years [14–17]. Huang et al. [18] first proposed the EMD method, which can adaptively process nonstationary and nonlinear signals. Meanwhile, it is widely used in various fields and achieves satisfactory results [19, 20]. However, the EMD method cannot cope with the problem of signals affected by noise effectively. Therefore, ensemble empirical mode decomposition (EEMD) is proposed to further optimize EMD, which adds a white noise to the original signal for eliminating the negative influence of the noise, and the experiment shows that it can achieve good results in a certain range of applications [21]. Nevertheless, the applications of EEMD will be restricted due to an unreasonable assumption of white noise on the actual signal. In order to address the noise problem effectively, TVF-EMD method is proposed in the following period of time. This method adopts a B-spline instead of a cubic spline to interpolate the extreme points, which greatly improves the filtering effect of high-frequency noise. Moreover, TVF-EMD can eliminate the pseudo extreme points of “pollution” by noise and further realize accurate interpolation [22]. In the application of short-term traffic flow
prediction, the prediction models based on TVF-EMD are more effective than EMD [23, 24]. Although EMD-based models are widely applied in prediction fields, it is a decomposition method based on Hilbert-Huang transform theory [25], and the decomposed subsequence cannot well explain the negative frequency components in the frequency domain of signal. To this end, the local mean decomposition (LMD) method [26] was proposed to further improve EMD, and the combination of EMD and LMD model shows an excellent performance in application [27]. Besides the EMD-based decomposition methods, it is worth mentioning that a new signal processing method, i.e., variational mode decomposition (VMD), has been proposed in [28]. This method can not only remove the high-frequency noise adaptively but also decompose the original signal into several subseries whose central frequencies can be clearly separated. Its good performance has been verified in the field of prediction [29–32]. However, the choice of VMD decomposition level in the field of short-term traffic flow prediction lacks theoretical guidance, which may lead to a huge deviation in prediction accuracy [33].

3. Materials and Methods

3.1. TVF-EMD. TVF-EMD is an improved data decomposition method of EMD. Its two main processes are the calculation of the cut-off frequency of the signal and the filtering of the signal approximated by B-spline polynomials. The following is a brief introduction to their theory. Suppose that there is a signal \( x(t) \), the analytic form of the signal will be obtained by Hilbert transform, and its mathematical expression is as shown in the formula.

\[
Z(t) = x(t) + j\tilde{x}(t) = A(t) \cdot \exp(j\varphi(t)),
\]

where \( j \) is the imaginary unit; \( A(t) \) and \( \varphi(t) \) are the instantaneous amplitude and frequency of the analytic signal, respectively; and \( \tilde{x}(t) \) is the Hilbert transform of the original signal. More details about the Hilbert transform can be found in reference [34]. Then the analytic signal \( Z(t) \) can be further decomposed into two analytic signals by equation (2).

\[
Z(t) = a_1(t) \cdot \exp(j\varphi_1(t)) + a_2(t) \cdot \exp(j\varphi_2(t)).
\]

Similarly, \( a_1(t), a_2(t), \varphi_1(t), \) and \( \varphi_2(t) \) are the instantaneous amplitude and frequency of the two decomposed signals correspondingly. Denote the set of the local minimum and maximum values of the original signal amplitude \( A(t) \) as \( \{A(t_{\text{min}})\}, \{A(t_{\text{max}})\} \), respectively, and the corresponding timing node-sets are \( \{t_{\text{min}}\} \) and \( \{t_{\text{max}}\} \). Then, in terms of the theoretical derivation of the local minimum and maximum values of the signal (the detailed contents can be found in [22]), we can get formulas (3)–(6).

\[
\beta_1(t_{\text{min}}) = A(t_{\text{min}})[a_1(t_{\text{min}}) - a_2(t_{\text{min}})],
\]

\[
\beta_2(t_{\text{max}}) = A(t_{\text{max}})[a_1(t_{\text{max}}) + a_2(t_{\text{max}})],
\]

\[
\eta_1(t_{\text{min}}) = \varphi_1'(t_{\text{min}}) \cdot [\varphi_1(t_{\text{min}}) - a_1(t_{\text{min}}) \cdot a_2(t_{\text{min}})] +
\]

\[
\varphi_2'(t_{\text{min}}) \cdot [\varphi_2(t_{\text{min}}) - a_1(t_{\text{min}}) \cdot a_2(t_{\text{min}})],
\]

\[
\eta_1(t_{\text{max}}) = \varphi_1'(t_{\text{max}}) \cdot [\varphi_1(t_{\text{max}}) + a_1(t_{\text{max}}) \cdot a_2(t_{\text{max}})] +
\]

\[
\varphi_2'(t_{\text{max}}) \cdot [\varphi_2(t_{\text{max}}) + a_1(t_{\text{max}}) \cdot a_2(t_{\text{max}})].
\]

In the formula, \( a_1(t_{\text{min}}), a_2(t_{\text{min}}), a_1(t_{\text{max}}), a_2(t_{\text{max}}), \varphi_1'(t_{\text{min}}), \varphi_2'(t_{\text{min}}), \varphi_1'(t_{\text{max}}), \varphi_2'(t_{\text{max}}) \) can be obtained by relevant calculation. Then \( \beta_1(t) \) and \( \beta_2(t) \) will be expressed by interpolating between \( \beta_1(t_{\text{min}}) \) and \( \beta_2(t_{\text{max}}) \) of the original signal; \( \eta_1(t) \) and \( \eta_2(t) \) are also obtained in similar way. Naturally, \( a_1(t), a_2(t), \varphi_1(t), \varphi_2(t) \) can be obtained through formulas (7)–(10).

\[
a_1(t) = \frac{\beta_1(t) + \beta_2(t)}{2},
\]

\[
a_2(t) = \frac{\beta_1(t) - \beta_2(t)}{2},
\]

\[
\varphi_1'(t) = \frac{\eta_1(t)}{2a_1^2(t) - 2a_1(t) \cdot a_2(t)} + \frac{\eta_2(t)}{2a_2^2(t) + 2a_1(t) \cdot a_2(t)},
\]

\[
\varphi_2'(t) = \frac{\eta_1(t)}{2a_1^2(t) - 2a_1(t) \cdot a_2(t)} + \frac{\eta_2(t)}{2a_2^2(t) + 2a_1(t) \cdot a_2(t)},
\]

where equations (7) and (8) assume that \( a_1(t) \) is greater than \( a_2(t) \). Finally, we get the expression of the cut-off frequency (11).

\[
\psi(t) = \frac{\varphi_1'(t) + \varphi_2'(t)}{2}.
\]

B-spline is commonly applied in interpolation and exhibits a great approximation effect. It is a polynomial interpolation function so that any finite signal can be approximated by the B-spline basis function [35, 36], and the expression of approximating signal is shown as the following formula.

\[
h^n_m(t) = \sum_{k=-\infty}^{\infty} c(k) \cdot \psi^n(t - k),
\]

where \( \psi^n(\cdot) \) is the B-spline basis function of a polynomial with degree \( n \); \( m \) is the node step size, and \( c \) is the coefficients of the B-spline basis function. For a given signal \( y(t) \), the coefficients of the approximation signal are determined by the least square error between the real signal and the approximation signal, as shown in equation (13).

\[
c(k) = \arg\min_{c(t)} \left( y(t) - [c]_m \cdot \psi^n(t/m) \right)^2.
\]
where \( [\cdot]_m \) is the \( m \)-step up sampling of the \( c(t) \) and \( * \) is the convolution operation. The solution of the problem is shown in equation (14).

\[
c(k) = [p_m^* * y]_m(k), \quad (14)
\]

where \( [\cdot]_m \) is \( m \)-step down sampling operation and \( p_m^* \) is the filter term, and its expression is exhibited as follows:

\[
p_m^* = \left[ y_m^{\prime \prime} \left( \frac{t}{m} \right) \right]_{m+1}^{-1} \left[ y_m^{\prime \prime} \left( \frac{t}{m} \right) \right]_m, \quad (15)
\]

In this way, the original signal can be replaced by the approximation signal, and the approximation signal plays a very important role in filtering the noise of the original signal. According to the corresponding experiment in [22], the larger the value of \( m \) in the proper range, the better the suppression effect of high-frequency noise. The detailed steps about TVF-EMD are presented in Appendix A.

### 3.2. Extreme Learning Machine

The extreme learning machine is an artificial intelligence model, which can effectively solve the problem of nonlinear classification and prediction. Suppose that there is an observation sample vector \( O = \{o_1, o_2, \ldots, o_N\}, o_i \in \mathbb{R}^n \) whose corresponding observation value is \( T = \{t_1, t_2, \ldots, t_N\}, t_i \in \mathbb{R} \). Then the extreme learning machine can be used to learn the sample space adaptively. The structure is composed of the input layer, hidden layer, and output layer, and its learning process can be summarized as the following steps.

**Step 1.** Set the number of neurons in the hidden layer \( M \) and the neuron activation function \( f(\cdot) \).

**Step 2.** Generate input layer connection matrix \( \omega \) and bias vector \( b \) randomly, where \( \omega \) is the matrix of \( n \times M \), and the element of row \( i \) and column \( j \) is denoted as \( \omega_{ij} \); \( b \in \mathbb{R}^M \).

**Step 3.** From the randomly generated connection matrix and bias vector, the output matrix \( \psi \) can be obtained, as shown in formula (16).

\[
\psi = \begin{bmatrix}
f(\omega_1^T \cdot o_1 + b_1) & \cdots & f(\omega_M^T \cdot o_1 + b_M) \\
f(\omega_1^T \cdot o_2 + b_1) & \cdots & f(\omega_M^T \cdot o_2 + b_M) \\
\vdots & \cdots & \vdots \\
f(\omega_1^T \cdot o_N + b_1) & \cdots & f(\omega_M^T \cdot o_N + b_M)
\end{bmatrix}_{N \times M}, \quad (16)
\]

where \( \omega_{ij}, b_i \) are the \( i \)-th column of \( \omega \) and the \( i \)-th element of \( b \), respectively.

**Step 4.** Determine the output vector \( \pi, \pi \in \mathbb{R}^M \) so that formula (17) is established.

\[
\pi = \arg \min_{\pi_0} \left\| \psi \cdot \pi_0 - P \right\|,
\]

where \( P \) is the observed value vector, i.e., \( P = (t_1, t_2, \ldots, t_N)^T \). \( \| \cdot \| \) is the first norm of the vector.

According to the definition of the generalized inverse matrix, the solution of equation (17) can be obtained by equation (18).

\[
\pi = \psi^+ \cdot P, \quad (18)
\]

where \( \psi^+ \) is the generalized inverse of matrix \( \psi \).

From the above four steps, it can be seen that the learning process of the extreme learning machine does not need to learn the corresponding parameters through iteration, but only one calculation of a generalized inverse matrix is required.

### 4. The Hybrid Forecasting Model

The main research contents of this paper are how to realize reliable decomposition and obtain better prediction results. Based on the principles of TVF-EMD and LMD, we can
combine their advantages to establish a new method. TVF-EMD is a valid method to remove data noise and decrease the nonstationary. In addition, LMD has the capability to overcome the related frequency problems of EMD-based decomposition methods. The detailed steps of LMD can be found in Appendix B. Hence, this paper adopts the hybrid of TVF-EMD and LMD to decompose the original short-term traffic flow data.

Central to the decomposition method is to decrease the nonstationary of the original time series and then obtain the smoother subseries. To this end, the following conditions should be satisfied in the further decomposition of IMFs by LMD. First, the number of local extreme points of decomposed IMFs is not less than 3 according to the principle of the LMD method. Second, to avoid extreme oscillations in the decomposition sequence, the total energy of the subsequences obtained by decomposition cannot be higher than that of IMFs to be decomposed. The first condition means that not all IMFs need to be decomposed; the second one controls the amplitudes of the subseries, which reduces the nonstationary and obtains smoother subseries.

5. Case Study

To verify the performance of the proposed in a more thorough way, this work collected two groups of data with different characteristics at the intersection of main roads in the urban area of Chongqing, its schematic is shown in Figure 2.
5.1. Data Description. Data sets 1 and 2 were extracted in the A and B intersections, respectively, between December 15 and December 21, 2020, and each group of data is based on the traffic volume arriving at the intersection every 5 minutes.

From a longitudinal perspective, the maximum traffic flow of data set 1 does not exceed 50 (vec/5 min), which is a low-traffic traffic flow; from the other direction, the traffic flow has a certain periodicity, and each cycle is about one day, which is also consistent with social regularity.

Because the same experiment will be conducted on the two groups of data, we will discuss data set 1 in detail and the forecasting results of data set 2 are simply presented in Section 5.4. The statistical results of data set 1 are shown in Figure 3.

To make an effective analysis and evaluation to prediction results, this paper divides the data set into the training part and test part, in which two-thirds of the data set is considered as the training part and the remaining part is the test part [37].

5.2. Data Decomposition Results. According to the flowchart described in Section 4, the original short-term traffic flow is decomposed by TVF-EMD, and several IMFs are obtained. The results of TVF-EMD decomposition are displayed in Figure 4.

It can be intuitively observed from Figure 4 that the oscillation of IMFs is decreased gradually. Furthermore, compared with IMFs 1–5, the IMFs 6–10 show a smoother fluctuation. Because LMD has the capability of handling smooth and symmetric signals [27]. Besides, it can avoid a negative frequency trap. Then all IMFs are tested by LMD decomposition conditions in Subsection 4. Finally, the IMF 7 of the training set of data set 1 can be further decomposed into 5 different subseries, respectively. Their LMD decomposition results are presented in Figure 5.

| Table 1: Alternative names of three hybrid models. |
|--------------------------------------------------|
| EMD-ELM                                         | Model 1 |
| VMD-ELM                                         | Model 2 |
| TVF-EMD-ELM                                     | Model 3 |
5.3. Prediction Results and Comparative Analysis. This work adopts the ELM model to forecast subsequence decomposed by the proposed approach, then superimpose the prediction result of all subseries, by which the corresponding prediction results of the test set can be obtained. In order to highlight the advantages of the hybrid model, several commonly used single models (ARIMA, ELM, DBN) are used for comparison. In addition, this paper will prove the superiority of the proposed decomposition method by comparing it with the conventional methods (EMD, TVF-EMD, and VMD). Finally, it will make a comparative analysis in terms of the prediction results of all involved models. For the convenience of description, three contrastive hybrid models are renamed in Table 1.

To intuitively exhibit the comparative results among the above-involved models, the predicted results of single models and the proposed model are shown in Figure 6. Similarly, the forecasting results of the hybrid models based on the traditional decomposition method and the proposed model are shown in Figure 7.

Some information can be captured from Figures 8 and 9. First, the prediction trend of the proposed model is closer to the true one. Second, the proposed model achieves the best performance in local intervals [453, 457].

The above results can only reflect the advantages of the proposed hybrid model in the local part. Therefore, we use four popular indicators including mean absolute error

\[
\begin{array}{|c|c|c|}
\hline
\text{Model} & \text{ARIMA} & \text{Proposed} \\
\hline
\text{Mean Absolute Error} & 1.23 & 0.98 \\
\hline
\end{array}
\]

Figure 6: The prediction results of the single models and the proposed model.

Figure 7: The prediction results of the hybrid models and the proposed model.
mean absolute percentage error (MAPE), root mean square error (RMSE), and root means square relative error (RMSRE) to further evaluate the prediction performance. Their mathematical expressions are shown in (19)–(22).

\[
\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |\sigma_i - \hat{\sigma}_i|, \tag{19}
\]

\[
\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{\sigma_i - \hat{\sigma}_i}{\sigma_i} \right|, \tag{20}
\]

\[
\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\sigma_i - \hat{\sigma}_i)^2}, \tag{21}
\]

\[
\text{RMSRE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left( \frac{\sigma_i - \hat{\sigma}_i}{\sigma_i} \right)^2}, \tag{22}
\]

where \(\sigma_i\) is the \(i\)th true value in the test set, \(\hat{\sigma}_i\) is the corresponding predicted value, and \(n\) is the number of samples.
in the test set. It can be seen from equations (19) ~ (22) that the smaller the value of MAE, MAPE, RMSE, and RMSRE indicates, the higher the accuracy of prediction.

The evaluation indicators about these seven models are presented in Table 2 (data set 1).

On the one hand, the smaller average amount of the original traffic flow may lead to a larger MAPE. In addition, the prediction performance of the deep learning model is influenced by the hyperparameter, which will produce a poor prediction effect. On the other hand, Table 2 shows that the proposed hybrid model achieves the best performance in prediction accuracy. Furthermore, the comparison with the ELM model, models 1–3 indicate that the application of the decomposition method has the capability of decreasing nonstationary. Similarly, the prediction accuracy of the proposed model is higher than the model 3, which reveals the superiority of the LMD method.

To avoid the biased results by chance and prove the robustness of the proposed model [38], this paper selects the cross-validation strategy to verify the involved prediction models. The MAE value of the involved models by cross-validation strategy is showed as follows:

From Table 3 and Figure 8, the proposed method exhibits the stable results under the cross-validation strategy and owns the best performance in the involved models.

5.4. Additional Case. Another group of data is further used to verify the effectiveness of the proposed method. The corresponding statistical result is displayed in Figure 9. It is easy to find that the traffic flow data of data set 2 also feature a periodicity, while the peak value in the interval [250, 350] is smaller than other peak values. In addition, the average value of data set 2 is larger than that of data set 1.
Similarly, TVF-EMD is employed to decompose original traffic flow into several IMFs that the fluctuation could become weaker gradually. Then the IMFs that meet the decomposition condition of LMD are further decomposed. Finally, ELM is utilized to forecast using the same way.

With the update of the training set step by step, all the forecasting tasks are completed. Then, the corresponding evaluation indicators of all involved models are summarized in Table 4.

From Table 4, the MAE and RMSE values of all involved models are higher than that of Table 2, which shows the data set 1 and data set 2 own different characteristics. Further, the MAPE and RMSRE value of the proposed method is relatively low, there are perhaps two reasons: 1. the average traffic flow is larger than data set 1; 2. the decomposition model proposed in this paper can not only reduce the nonstationarity but also make the decomposed subsequence smoother, both of which are beneficial to handle the strong nonstationarity series (data set 2). But the proposed model both achieve the best performance.

6. Conclusions

In view of the shortcoming of the traditional decomposition methods, this paper proposes a new decomposition method based on the combination of TVF-EMD and LMD. To the best of our knowledge, this method is first applied to the field of short-term traffic flow prediction. Compared with the traditional TVF-EMD method, it not only decreases the nonstationary time series but also further decomposes the IMFs to avoid the negative frequency trap by the LMD method and obtain the smoother series. From the perspective of prediction accuracy, a case study based on two groups of the measured traffic flow data verifies that the TVF-EMD-LMD method can achieve the highest prediction accuracy. Therefore, this developed self-adaptive decomposition method (i.e., TVF-EMD-LMD) may have great potential in the prediction of short-term traffic flow.

Appendix

A. The complete steps of TVF-EMD

Step 1: Find the position of extreme maximum in input signal \( x(t) \) and denote them as \( u_i \) \((i = 1, 2, \ldots L)\);
Step 2: Seek all intermittences of \( u_i \) that satisfies the following formula and mark them as \( e_j \) \((j = 1, 2, \ldots K)\);
\[
\max\left(\varphi^t(u_l \leq t \leq u_{l+1})\right) - \min\left(\varphi^t(u_l \leq t \leq u_{l+1})\right) > \rho
\]
(A.1)

where \( \varphi^t(\cdot) \) is bisecting frequency introduced in Subsection 3.1 and \( \rho \) is the threshold. If vd, it means \( e_j \) locates on a rising edge of \( \varphi^t(\cdot) \) and vice versa;
Step 3: Ensure the true peak and bottom parts of \( \varphi^t(\cdot) \) in terms of rising and floor edge;
Step 4: Readjust the local cut-off frequency by interpolating extreme points in new peak parts and bottom parts;
Step 5: Take the extreme timing of \( h(t) \) as knots, and the expression of \( h(t) \) is as follows. Then utilize the B-spline technique to interpolate between knots, and the corresponding approximate results are marked as \( r(t) \)
\[
h(t) = \cos\left[\int \varphi^t(t)dt\right]
\]
(A.2)
Step 6: Select the stop criterion to examine whether the \( x(t) \) satisfies the condition. The criterion is shown as follows:
\[
\theta(t) = \frac{B_{\text{Loughlin}}(t)}{\varphi_{\text{avg}}(t)},
\]
(A.3)

where \( B_{\text{Loughlin}} \) and \( \varphi_{\text{avg}}(t) \) are Loughlin instantaneous bandwidth and the weighted average of the instantaneous frequency, respectively; \( \xi \) is given bandwidth threshold; the calculation method of \( B_{\text{Loughlin}} \) and \( \varphi_{\text{avg}}(t) \) is detailed in [39, 40];
Step 7: if \( x(t) \) satisfies the stop criterion, then it will be regarded as IMF; otherwise, \( x(t) = x(t) - r(t) \), repeat steps 1–6 until obtaining all IMFs.

B. The complete steps of TVF-EMD

Step 1: Initial \((p, q) = 0\), take \( IMF_1 \) as an input signal and redefine as \( u_{pq,i} \), i.e., \( u_{pq,i} = IMF_i \);
Step 2: Check whether \( u_{pq,i} \) satisfies the first decomposition condition introduced in Section 4, then \( p = p + 1 \);
Step 3: If the condition is satisfied, then calculate the local means and magnitude of the input signal by moving average and denoted them as \( m_{pq,i}(t) \) and \( a_{pq,i}(t) \), respectively;
Step 4: Calculate \( b_{pq,i}(t) = u_{pq,i}(t) - m_{pq,i}(t) \) and \( s_{pq,i}(t) = b_{pq,i}(t)/a_{pq,i}(t) \), then \((q = q + 1)\);
Step 5: Judge \( s_{pq,i}(t) \) is a purely frequency modulated signal or not. If \( s_{pq,i}(t) \) is not a purely frequency modulated signal, back to step 3; otherwise, obtain subseries of \( IMF_i \) by following formulas;
\[
a_{pq,i}(t) = \sum_{j=1}^{q} a_{pq,j}(t),
\]
(B.1)
\[
PF_{pq,i}(t) = a_{pq,i}(t) \cdot s_{pq,i}(t),
\]
where \( PF_{pq,i}(t) \) is one of subseries of \( IMF_i \);
Step 6: Regard \( u_{pq,i}(t) = u_{pq,i}(t) - PF_{pq,i}(t) \), then examine whether the \( u_{pq,i} \) satisfies the first decomposition condition. If it satisfies, back to step 2 and continue decomposing; otherwise, obtain all subseries \( PF_{j\cdot}(t) (j = 1, 2, \cdots p) \) and residual signal \( u_{pq,i} \);
Step 7: Check whether all decomposed subseries satisfy the second decomposition condition. If it is true, regard
them as eventual subseries of \( IMF_i \). Otherwise, it cannot further decompose.

**Data Availability**

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

**Conflicts of Interest**

The authors declare that they have no conflicts of interest.

**Acknowledgments**

This work was supported by the Program of Chongqing Municipal Education Commission (KJQN201900707), Natural Science Foundation of Chongqing (cstc2018jcyjA0028), and Joint Training Base Construction Project for Graduate Students in Chongqing (JDLHPYJD2021016).

**References**

[1] G. Yao, X. Liu, and Y. Chen, "Forecasting maximum autonomous ownership in urban areas," *Urban Transportation of China*, vol. 18, no. 5, pp. 110–119, 2020.

[2] Q. Chen, Y. Song, and J. Zhao, "Short-term traffic flow prediction based on improved wavelet neural network," *Neural Computing & Applications*, vol. 33, no. 14, pp. 8181–8190, 2020.

[3] L. Xu and X. Wang, "Application of SVD-based optimized robust UKF algorithm to estimation of short-term traffic flow state," *Journal of South China University of Technology*, vol. 44, no. 12, pp. 44–52, 2016.

[4] Q. Bing, B. Gong, Z. Yang, Q. Shang, and X. Zhou, "Short-term traffic flow local prediction based on combined kernel function relevance vector machine model," *Mathematical Problems in Engineering*, vol. 2015, Article ID 154703, 9 pages, 2015.

[5] C. Han, S. Song, and C. Wang, "A real-time short-term traffic flow adaptive forecasting method based on ARIMA model," *System Simulation*, vol. 16, no. 7, pp. 1530–1535, 2004.

[6] I. Okutani and Y. Stephaneides, "Dynamic prediction of traffic volume through Kalman filtering theory," *Transportation Research Part B: Methodological*, vol. 18, no. 1, pp. 1–11, 1984.

[7] K. Kumar, M. Parida, and V. Katiyar, "Short term traffic flow prediction in heterogeneous condition using artificial neural network," *Transport*, vol. 30, no. 4, pp. 397–405, 2013.

[8] G. Huang, Q. Zhu, and C. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1-3, pp. 489–501, 2006.

[9] H. Xu and C. Jiang, "Deep belief network-based support vector regression method for traffic flow forecasting," *Neural Computing & Applications*, vol. 32, no. 7, pp. 2027–2036, 2019.

[10] D. Ma, X. Song, and P. Li, "Daily traffic flow forecasting through a contextual convolutional recurrent neural network modeling inter- and intra-day traffic patterns," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 5, pp. 2627–2636, 2021.

[11] W. Zheng, X. Peng, D. Zhang et al., "Composite quantile regression extreme learning machine with feature selection for short-term wind speed forecasting: a new approach," *Energy Conversion and Management*, vol. 151, pp. 737–752, 2017.

[12] Y. Jiang, N. Zhao, L. Peng, J. Xin, and S. Liu, "Fast simulation of fully non-stationary wind fields using a new matrix factorization assisted interpolation method," *Mechanical Systems and Signal Processing*, vol. 172, pp. 1–18, 2022.

[13] F. Liu, Y. Xu, Y. Yang, and H. Ma, "A spatial-temporal hybrid model for short-term traffic prediction," *Mathematical Problems in Engineering*, vol. 2019, Article ID 4858546, 12 pages, 2019.

[14] J. Tang, X. Chen, Z. Hu, F. Zong, C. Han, and L. Li, "Traffic flow prediction based on combination of support vector machine and data denoising schemes," *Physica A Statistical Mechanics and Its Applications*, vol. 534, pp. 1–26, 2019.

[15] Y. Jiang, S. Liu, N. Zhao, J. Xin, and B. Wu, "Short-term wind speed prediction using time varying filter-based empirical mode decomposition and group method of data handling-based hybrid," *Energy Conversion and Management*, vol. 220, pp. 1–15, 2020.

[16] H. Huang, J. Chen, X. Huo, Y. Qiao, and L. Ma, "Effect of multi-Scale decomposition on performance of neural networks in short-term traffic flow prediction," *IEEE Access*, vol. 9, pp. 50994–51004, 2021.

[17] Y. Jiang, S. Liu, and L. Peng, "A novel wind speed prediction method based on robust local mean decomposition, group method of data handling and conditional kernel density estimation," *Energy Conversion and Management*, vol. 200, pp. 1–15, 2019.

[18] N. Huang, Z. Shen, S. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.

[19] H. Wang, L. Liu, Z. Qian, and H. Wei, "A novel work zone short-term vehicle-type specific traffic speed prediction model through the hybrid EMD-ARIMA framework," *Transportation Research Part B: Methodological*, vol. 4, no. 3, pp. 159–186, 2016.

[20] H. Wang, L. Liu, H. Wei, and S. Dong, "Empirical mode decomposition-autoregressive integrated moving average," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2460, no. 1, pp. 66–76, 2014.

[21] X. Chen, H. Chen, Y. Yang et al., "Traffic prediction by an ensemble framework with data denoising and deep learning model," *Physica A Statistical Mechanics and Its Applications*, vol. 565, pp. 1–11, 2021.

[22] H. Li, Z. Li, and W. Mo, "A time varying filter approach for empirical mode decomposition," *Signal Processing*, vol. 138, pp. 146–158, 2017.

[23] A. Longo, S. Bianchi, and W. Plastino, "TVF-EMD based time series analysis of 7Be sampled at the CTBTO-IMS network," *Physica A: Statistical Mechanics and Its Applications*, vol. 523, pp. 908–914, 2019.

[24] Y. Xu, Z. Cai, and K. Ding, "An enhanced bearing fault diagnosis method based on robust local mean decomposition, group method of data handling and conditional kernel density estimation," *Measurement Science and Technology*, vol. 29, no. 9, pp. 1–21, 2018.

[25] N. Huang and Z. Wu, "A review on Hilbert-Huang transform: method and its applications to geophysical studies," *Reviews of Geophysics*, vol. 46, no. 2, pp. 1–23, 2008.

[26] J. Smith, "The local mean decomposition and its application to EEG perception data," *Journal of the Royal Society Interface*, vol. 2, no. 5, pp. 443–454, 2005.

[27] S. Yue, Y. Wang, L. Wei, Z. Zhang, and H. Wang, "The joint empirical mode decomposition-local mean decomposition
method and its application to time series of compressor stall process,” *Aerospace Science and Technology*, vol. 105, pp. 1–10, 2020.

[28] K. Dragomiretskiy and D. Zosso, “Variational mode decomposition,” *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.

[29] W. Niu, Z. Feng, Y. Chen, and H. Zhang, “Annual streamflow time series prediction using extreme learning machine based on gravitational search algorithm and variational mode decomposition,” *Journal of Hydrologic Engineering*, vol. 25, no. 5, pp. 1–10, 2020.

[30] Q. Bing, F. Shen, X. Chen, W. Zhang, Y. Hu, and D. Qu, “A hybrid short-term traffic flow multistep prediction method based on variational mode decomposition and long short-term memory model,” *Discrete Dynamics in Nature and Society*, vol. 2021, pp. 1–13, 2021.

[31] W. Chen, H. Xu, and Z. Jiang, “A novel method for time series prediction based on error decomposition and nonlinear combination of forecasters,” *Neurocomputing*, vol. 426, pp. 85–103, 2021.

[32] J. Xin, Y. Jiang, J. Zhou, L. Peng, S. Liu, and Q. Tang, “Bridge deformation prediction based on SHM data using improved VMD and conditional KDE,” *Engineering Structures*, vol. 261, pp. 1–16, 2022.

[33] D. Wang, H. Luo, and O. Lin, “Multi-step ahead wind speed forecasting using an improved wavelet neural network combining variational mode decomposition and phase space reconstruction,” *Renewable Energy*, vol. 113, pp. 1345–1358, 2017.

[34] S. Hahn, *Hilbert Transforms in Signal Processing*, Artech House, Norwood, MA, USA, 1996.

[35] Q. Chen, N. Huang, S. Riemenschneider, and Y. Xu, “A B-spline approach for empirical mode decompositions,” *Advances in Computational Mathematics*, vol. 24, no. 1-4, pp. 171–195, 2006.

[36] Y. Yang, C. Miao, and J. Deng, “An analytical expression for empirical mode decomposition based on B-spline interpolation,” *Circuits, Systems, and Signal Processing*, vol. 32, pp. 2899–2914, 2006.

[37] Y. Wang, L. Zhao, S. Li, X. Wen, and Y. Xiong. “Short term traffic flow prediction of urban road using time varying filtering based empirical mode decomposition,” *Applied Sciences Basel*, vol. 10, pp. 1–15, 2020.

[38] N. Bokde, Z. Yaseen, and G. Andersen, “Forecast TB—an R package as a Test-Bench for time series forecasting-application of wind speed and solar radiation modeling,” *Energies*, vol. 13, pp. 1–24, 2020.

[39] P. Loughlin and K. Davidson, “Modified Cohen-Lee time-frequency distributions and instantaneous bandwidth of multicomponent signals,” *IEEE Transactions on Signal Processing*, vol. 49, no. 6, pp. 1153–1165, 2001.

[40] G. Jones and B. Boashash, “Instantaneous frequency, instantaneous bandwidth and the analysis of multicomponent signals,” in *Proceedings of the International Conference on Acoustic, Speech, and Signal Processing*, pp. 2467–2470, Albuquerque, NM, USA, April 1990.