Short-Term Traffic Flow Prediction Based on VMD-SMA-SVR

Zuchen Que *
Zhejiang Scientific Research Institute of Transport
Hangzhou, Zhejiang, China
zjlsqzc@126.com

Bo Jin
Zhejiang Scientific Research Institute of Transport
Hangzhou, Zhejiang, China
kimbojin@163.com

Ren Lou
Zhejiang Scientific Research Institute of Transport
Hangzhou, Zhejiang, China
179787711@qq.com

Zheting Si
Zhejiang Scientific Research Institute of Transport
Hangzhou, Zhejiang, China
596995631@qq.com

Hongliang Dai
Zhejiang Scientific Research Institute of Transport
Hangzhou, Zhejiang, China
dhl1972@163.com

2024 6th International Conference on Data-driven Optimization of Complex Systems (DOCS) | 979-8-3503-7784-2/24/$31.00 ©2024 IEEE | DOI: 10.1109/DOCS63458.2024.10704344

***Abstract*—Accurate prediction of short-term highway traffic flow is crucial for ensuring traffic safety and reducing the accident rate. This study proposes a traffic flow prediction model based on Variational Mode Decomposition-Slime Mould Algorithm-Support Vector Regression (VMD-SMA-SVR). The VMD algorithm is used to decompose traffic flow data, and the SMA algorithm optimizes the parameters of the SVR model. The SVR prediction models for each component are constructed and their predictions are aggregated. Comparative experiments with various models demonstrate that the VMD-SMA-SVR model excels in prediction accuracy and stability. The VMD algorithm shows good adaptability, while the SMA algorithm exhibits strong optimization capabilities. This model effectively improves the accuracy of short-term traffic flow prediction, featuring fewer model parameters, high computational efficiency, and low resource consumption, making it suitable for practical short-term traffic flow forecasting.**

Keywords-Traffic Flow Prediction; Support Vector Regression; Variational Mode Decomposition; Slime Mould Algorithm

I. INTRODUCTION

Statistics show that although the mileage of highways in China accounts for only 2.8% of the entire road traffic system, traffic accidents on highways are responsible for 10% of the annual traffic-related fatalities. This highlights the safety risks associated with highways. Therefore, accurately predicting short-term traffic flow on highways and understanding their risk conditions are crucial for taking timely measures to reduce traffic risks and decrease the occurrence of accidents.

To address the safety risks associated with highways, accurately predicting short-term traffic flow is crucial. This provides essential data for intelligent traffic systems to reduce accidents [1]. Extensive research has been conducted on various traffic flow prediction models, which can be categorized into four types: linear prediction methods, traditional nonlinear prediction methods, intelligent nonlinear prediction methods, and combined prediction methods.

Linear prediction models include the Historical Average (HA) model and the Autoregressive Integrated Moving Average (ARIMA) model. The HA model, which uses past traffic averages, struggles with dynamic changes and complex traffic conditions. Ahmed M S et al. [2] were the first to apply time series theory to traffic flow prediction using the ARIMA model for highways. Although accurate with continuous data, ARIMA faces challenges with limited or unstable data. Improvements like the Seasonal ARIMA model [3] and the Kalman filtering method have been proposed. Okutani and Stephanedes [4] introduced Kalman filtering to traffic flow prediction, validating it with Nagoya, Japan's road network data.

Nonlinear prediction methods offer greater flexibility and adaptability, better handling the variable states of traffic flow. Popular methods include wavelet analysis, Support Vector Regression (SVR), decision trees, and neural networks. Wavelet analysis, by performing multi-scale decomposition, effectively captures the nonlinear characteristics and sudden changes in traffic flow. Chai et al. [5] combined Mallat wavelets with BP (Back Propagation) neural networks to construct a short-term traffic flow prediction model, validated with real data from the Chengyu Expressway.

Neural networks, particularly BP neural networks, have been widely used in traffic prediction since the 1990s. Smith et al. [6] applied this type of neural network to short-term traffic flow prediction, proving its superior predictive power over the historical average and time series models. However, neural networks require high-quality data samples for training, which may affect performance with small datasets. To address this, some researchers have turned to Support Vector Machines (SVM) for traffic flow prediction. Ling et al. [7] proposed an adaptive particle swarm optimization algorithm to optimize a Multi-kernel Support Vector Machine model (MSVM), which performs well even during rapid traffic changes.

With the rapid development of artificial intelligence, deep learning has become a major trend in machine learning and has been widely applied across industries. Researchers have begun applying deep learning techniques to traffic flow prediction with notable success. Zhang et al. [8] developed a deep learning-based short-term traffic flow prediction model using Convolutional Neural Networks (CNN), significantly reducing prediction errors. Fang et al. [9] proposed a model combining

Kalman filtering and Long Short-Term Memory networks (LSTM), showing higher predictive accuracy than traditional LSTM models. However, deep learning models require substantial data, and their performance suffers with insufficient data.

Traffic systems often require more than a single model for accurate prediction, making combined models increasingly popular. These models integrate strengths of different algorithms for higher accuracy. For instance, Ding et al. [10] combined SVM with BP neural networks and optimized BP parameters using Particle Swarm Optimization (PSO), achieving better results than single models. Wei et al. [11] developed a model combining autoencoders with LSTM, showing good predictive accuracy and stability.

Due to the inherent nonlinearity and high noise of traffic flow data, researchers have combined decomposition algorithms with prediction algorithms. For example, Shuai et al. [12] used Singular Spectrum Analysis (SSA) to decompose traffic flow into components, predicting each with LSTM and SVR. However, this method is prone to frequency confusion. Zhao et al. [13] proposed an EMD-PSO-LSTM model to reduce noise and optimize LSTM parameters, but EMD suffers from modal confusion. Huang et al. [14] compared various decomposition algorithms with Bidirectional LSTM (BiLSTM) and found Variational Mode Decomposition (VMD) to be the most stable and adaptable. Therefore, VMD was chosen for this study.

Improper parameter selection can lead machine learning models to local optima, reducing their generalization and prediction accuracy. Heuristic algorithms have been used to optimize model parameters, enhancing accuracy and stability. Genetic Algorithms (GA) have been widely used, but require significant computational resources [15][16]. Recently, newer heuristic algorithms like the Grey Wolf Optimizer (GWO) [17], Sparrow Search Algorithm (SSA) [18], and Slime Mould Algorithm (SMA) [20] have been introduced. This study uses SMA for its fewer parameters and robust optimization capabilities.

This study introduces an innovative method for short-term traffic flow prediction using the Support Vector Regression (SVR) model. To enhance accuracy, the SVR model is optimized using the Slime Mould Algorithm (SMA). Additionally, the Variational Mode Decomposition (VMD) algorithm decomposes initial traffic flow data into multiple modal components. Separate SVR prediction models are constructed for each component, and the predictive values are aggregated for the final traffic flow prediction. This method's core advantage lies in effectively capturing complex patterns in traffic data, significantly enhancing prediction accuracy by combining SMA and VMD.

II. METHODOLOGY

This study introduces a highway traffic flow prediction model based on the VMD-SMA-SVR framework, whose detailed process flow is illustrated in the diagram below (Fig. 1). The model includes the following key steps:

1) Initially, the missing values in the dataset were imputed based on the historical average values corresponding to the specific time points where the data was missing;

2) Next, the completed data is processed using Variational Mode Decomposition (VMD), an algorithm that decomposes the data into several intrinsic mode functions (IMFs) of different frequencies;

3) Subsequently, these IMFs are normalized, and a separate Support Vector Regression (SVR) model is constructed for each component. The Slime Mould Algorithm (SMA) is then employed to fine-tune the hyperparameters of these SVR models, thereby training more accurate models;

4) The trained and optimized models are used to predict each IMF component individually, obtaining their forecasted values;

5) Finally, the predicted values for all IMFs are denormalized and aggregated to reconstruct the final prediction of highway traffic flow.

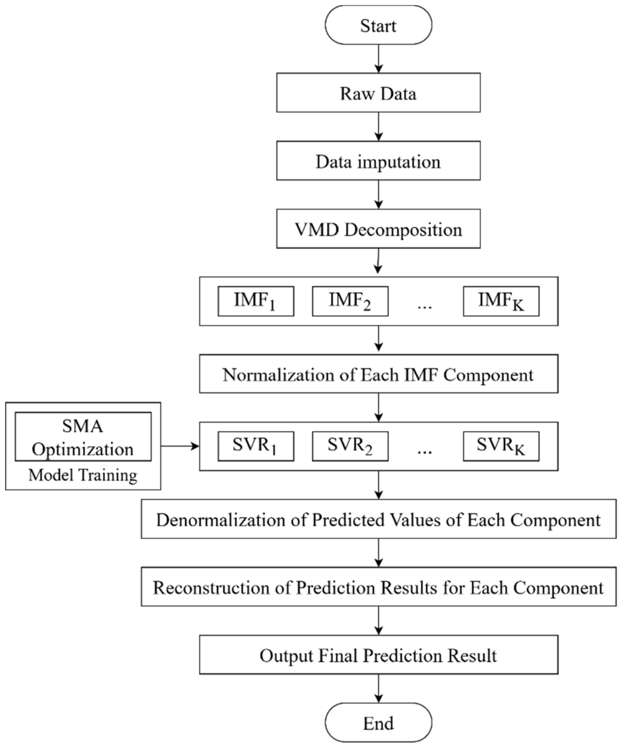

Figure 1. VMD-SMA-SVR model Flow Chart

A. Variational Mode Decomposition

Due to the inherent nonlinearity, instability, and randomness of highway traffic flow data, directly applying traditional regression models for prediction often fails to achieve high accuracy. To address this issue, this study introduces a multiscale decomposition algorithm to meticulously decompose the time series data of truck traffic flow. This process aims to break down the original traffic flow data into several more stable subsets, each of which can be predicted individually. By accurately forecasting these more stable data components, this research significantly enhances the overall accuracy of the model's predictions.

Variational Mode Decomposition (VMD) is a novel signal processing method for non-stationary signals, proposed by Dargomiretskiy et al. [19] in 2014. The VMD algorithm is adaptive, non-recursive, and capable of decomposing a signal into a sum of IMF components, each with a different bandwidth. It is particularly suitable for complex nonlinear and non-stationary signals. Based on Wiener filtering, VMD operates within a variational framework to search for the optimal solution of the input signal, automatically updating the central frequency, bandwidth, and corresponding sub-signals. This allows VMD to effectively separate the independent components of a signal in the frequency domain. Compared to the classical Empirical Mode Decomposition (EMD) method and Complete Ensemble Empirical Mode Decomposition with Adaptive Noise (CEEMDAN) method, VMD has a more robust theoretical foundation, offers higher decomposition precision, and effectively reduces the problem of mode mixing.

The decomposition process of the VMD algorithm is actually a process of solving a variational problem. Specifically, it decomposes a signal $f(t)$ into K modal functions $u_k(t)$ and minimizes the sum of the estimated bandwidths of each modal function $u_k(t)$. The steps for implementing the VMD algorithm are as follows:

Step 1: Perform a Hilbert transform on each modal function $u_k(t)$ to obtain the unilateral spectrum:

$$\left[\delta(t) + \frac{j}{\pi t}\right] * u_k(t) \tag{1}$$

Step 2: Multiply each modal function by the estimated center frequency $e^{-j\omega_k t}$ and then modulate the spectrum to the corresponding baseband:

$$\left\{\left[\delta(t) + \frac{j}{\pi t}\right] * u_k(t)\right\} \cdot e^{-j\omega_k t} \tag{2}$$

Step 3: the Gaussian smoothing method is used to estimate the bandwidths of the demodulated signal's modal functions $u_k(t)$, i.e., their gradient mean norms are calculated. Then, based on this bandwidth information, the constrained variational problem is solved. The expression of this constrained variational problem is as follows:

$$\begin{cases} \min\limits_{\{u_k\},\{\omega_k\}} \left\{\sum_k \left\|\partial_t \left[\delta(t) + \frac{j}{\pi t} * u_k(t)\right] \cdot e^{-j\omega_k t}\right\|_2^2\right\} \\ s.t. \sum_k u_k = f(t) \end{cases} \tag{3}$$

Where, $\{u_k\} = \{u_1, u_2, \ldots, u_K\}$ represents the K intrinsic mode function (IMF) components obtained after decomposition, $\{\omega_k\} = \{\omega_1, \omega_2, \ldots, \omega_K\}$ represents the center frequencies of the respective IMF components, ∂_t denotes the derivative with respect to time t, and $\delta(t)$ is the unit impulse function.

Step 4: Introduce the quadratic penalty factor α and the Lagrange multiplier λ to ensure the reconstruction accuracy of the original signal s(t), and convert the constrained variational formula into an unconstrained one.

$$\begin{aligned} L(\{u_k\},\{\omega_k\},\lambda) = \alpha \sum_k \left\|\partial_t \left[\left(\delta(t) + \frac{j}{\pi t}\right) * u_k(t)\right] \cdot e^{-j\omega_k t}\right\|_2^2 \\ + \left\|f(t) - \sum_k u_k(t)\right\|_2^2 + \langle \lambda(t), f(t) - \sum_k u_k(t) \rangle \end{aligned} \tag{4}$$

Step 5: Using the alternating direction method of multipliers, update $\hat{u}_k^{n+1}$, the corresponding center frequencies ω_k^{n+1}, and the Lagrange multiplier $\hat{\lambda}^{n+1}$ to find the minimum point of the augmented Lagrangian function expression. At this point, the solution to the variational problem is:

$$\hat{u}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{i \neq k} \hat{u}_i(\omega) + \frac{\hat{\lambda}(\omega)}{2}}{1 + 2\alpha(\omega - \omega_k)^2} \tag{5}$$

$$\omega_k^{n+1} = \frac{\int_0^\infty \omega |\hat{u}_i(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_i(\omega)|^2 d\omega} \tag{6}$$

$$\hat{\lambda}^{n+1}(\omega) = \hat{\lambda}^n(\omega) + \gamma \left[\hat{f}(\omega) - \sum_k \hat{u}_k^{n+1}(\omega)\right] \tag{7}$$

Step 6: Repeat step 5 until the iteration precision ϵ is satisfied, thereby obtaining k IMF components. The iteration termination condition can be expressed as:

$$\frac{\sum_k \|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2}{\sum_k \|\hat{u}_n^k\|_2^2} < \epsilon \tag{8}$$

B. Slime Mould Algorithm

The Slime Mould Algorithm (SMA) is an innovative metaheuristic optimization algorithm proposed by Li et al. [20] in 2020. The algorithm draws inspiration from the complex and sophisticated behaviors of the Physarum polycephalum, a slime mold, as it forages for food in the natural world. Researchers have developed a mathematical model that emulates the slime mold's search mechanism, based on a comprehensive study of these biological behaviors. Current research indicates that the SMA algorithm excels in finding optimal solutions, demonstrating rapid convergence capabilities and significantly improved precision and stability compared to conventional optimization algorithms. Owing to its outstanding optimization performance [21], this study employs the SMA algorithm to fine-tune the parameters of the SVR algorithm, aiming to achieve more accurate forecasting results. The following will outline the fundamental steps of this algorithm.

When the slime mold approaches food, the mathematical model of the slime mold algorithm can be represented by the following equation:

$$X(t+1) = \begin{cases} X_b(t) + v_b \cdot [W \cdot X_A(t) - X_B(t)], r < p \\ v_c \cdot X(t), r \geq p \end{cases} \tag{9}$$

Where, t is the current iteration number, $X_b(t)$ is the optimal position of the slime mold individual at the t-th iteration, $X_A(t)$ and $X_B(t)$ are the positions of two randomly selected slime mold individuals, v_b is the control parameter with a range of $[-a, a]$, v_c is a parameter that linearly decreases from 1 to 0, r is a random value between [0, 1], W is the mass of the slime mold, representing the fitness weight, p is the control variable, v_b is the control parameter with a range of $[-a, a]$, and the expression is as follows:

$$p = \tanh(|S(i) - DF|) \tag{10}$$

$$v_b = [-a, a] \tag{11}$$

$$a = \operatorname{arctanh}\left[-\left(\frac{t}{t_{max}}\right) + 1\right] \tag{12}$$

Where, $i \in 1,2,3 \cdots, n$, $S(i)$ is the fitness value of the i-th slime mold, DF is the best fitness value among all iterations, and the expression for the fitness weight W is as follows:

$$W(SmellIndex(i)) = \begin{cases} 1 + r \cdot \log\left[\frac{bF - S(i)}{bF - wF} + 1\right], condition \\ 1 - r \cdot \log\left[\frac{bF - S(i)}{bF - wF} + 1\right], other \end{cases} \quad (13)$$

$$SmellIndex = sort(S) \quad (14)$$

Where, bF is the best fitness value obtained in the current iteration process, wF is the worst fitness value obtained in the current iteration process, the $condition$ is for individuals whose $S(i)$ can be ranked in the front half of the slime mold population, $SmellIndex$ is the sequence of fitness values of the slime molds, but for solving minimization problems, it uses an ascending sort method.

When the slime mold envelops food, the mathematical model of the slime mold algorithm is as follows:

$$X = \begin{cases} rand \cdot (UB - LB) + LB, rand < z \\ X_b(t) + v_b \cdot [W \cdot X_A(t) - X_B(t)], r < p \\ v_c \cdot X(t), r \geq p \end{cases} \quad (15)$$

Where, $rand$ and r are random values between $[0,1]$, UB and LB are the upper and lower bounds of the search space, z represents the switching frequency, which determines whether the SMA will search for other food sources or search around the best individual.

When the slime mold grasps food, the vascular tissue and biological oscillator of the slime mold undergo changes. The higher the concentration of food in contact with the veins, the stronger the oscillations generated by the biological oscillator. Relying on these changes, the slime mold will grasp food with higher concentrations.

The changes in the width of the slime mold's veins are implemented using W, v_b, and v_c. W simulates the oscillation frequency of the slime mold in the vicinity under different food concentrations. v_b randomly varies within the range of $[-a, a]$, and gradually approaches zero as the number of iterations increases. The value of v_c oscillates between $[-1,1]$ and eventually tends to zero. When the slime mold chooses food, the synergistic relationship between v_b and v_c plays an important role.

C. Support Vector Regression Algorithm

Support Vector Regression (SVR) is a regression algorithm based on Support Vector Machines (SVM), used to solve regression problems. Its core idea is to handle regression issues while maintaining a maximum margin principle similar to that of Support Vector Machines. The optimization problem for the SVR model is:

$$\min_{w} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{n} (\xi_i + \xi_i^*) \quad (16)$$

$$s.t. \begin{cases} y_i - w \cdot \varphi(x_i) - b \leq \varepsilon + \xi_i^* \\ w \cdot \varphi(x_i) + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \ (i = 1,2,\cdots,n) \end{cases} \quad (17)$$

Where, w represents the weight vector, C is the penalty parameter, n is the number of samples in the SVR training set, ξ_i and ξ_i^* are the slack variables, $s.t.$ denotes the constraint conditions, ε is the tolerance deviation, x_i represents the i-th input data, y_i represents the response variable, and b is the bias term.

The final decision function of the SVR model can be represented as:

$$f(x) = \sum_{i=1}^{n} ((a_i - a_i^*) k(x_i, x) + b \quad (18)$$

Where, a_i and a_i^* are the introduced Lagrange multipliers, which are obtained through model optimization. $k(x_i, x)$ represents the kernel function, and here the Gaussian kernel function is used as the kernel function. When using the Gaussian kernel function, for any input sample x, its predicted output y_{pre} can be calculated as:

$$y_{pre} = \sum_{i=1}^{n} (a_i - a_i^*) exp\left\{\frac{-\|x_i - x\|}{2\sigma^2}\right\} + b \quad (19)$$

Where σ is the parameter of the Gaussian kernel function, which is used to control the width and smoothness of the kernel function. By adjusting the value of σ, the fitting capability and generalization performance of the model can be influenced.

III. Experiment

A. Data Processing

This study selected traffic flow data measured by a traffic monitoring device at a highway section in Zhejiang Province from March 1 to March 10, 2023. In this study, a method based on historical average values was employed to impute the missing data in the dataset, leveraging the trends observed in the historical data to estimate the missing information. Following data cleaning and organization, a comprehensive dataset was constructed, consisting of a total of 2880 records, with data collected at five-minute intervals. The first 60% of the dataset was designated as the training set, the middle 20% as the testing set, and the final 20% as the validation set. Furthermore, the Min-Max normalization method was employed to preprocess the data.

B. Baseline Models

To validate the effectiveness of the VMD-SMA-SVR model, this study conducted a comprehensive comparison of various models, including VMD-SVR, SMA-SVR, CEEMDAN-SMA-SVR, VMD-SMA-LSTM, VMD-SMA-RNN, and VMD-SSA-SVR, from both horizontal and vertical perspectives. In the process of hyperparameter optimization, the hidden size, learning rate, and dropout rate were selected for optimization in the LSTM model, while the regularization parameter, kernel coefficient, and tolerance were optimized in the SVR model.

C. Evaluation Metrics

In this study, mean absolute error (MAE), root mean square error (RMSE), mean absolute percentage error (MAPE) and explained variance (EV) were chosen as evaluation metrics to assess the accuracy of model predictions. The equation of these metrics is shown as follow:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \quad (20)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \quad (21)$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \quad (22)$$

$$EV = 1 - \frac{Var(y - \hat{y})}{Var(y)} \quad (23)$$

Where, y_i denotes the true value of the sample, $\hat{y}_i$ denotes the predicted value, and $\overline{y}_i$ denotes the sample mean of the true data. N denotes the full sample size. $Var()$ denotes variance, which quantifies the dispersion of values around the mean in a dataset.

D. Experiment Results and Analysis

Table 1 presents the various performance metrics of different models in traffic flow prediction. From Table 1, it is evident that the proposed VMD-SMA-SVR model exhibits the best performance across all metrics, including MSE, RMSE, MAPE, and EV, demonstrating its superior prediction accuracy, stability, and capacity to explain dataset variability. This also indicates that the application of the VMD algorithm for decomposing traffic flow data and predicting each component separately can significantly enhance model accuracy. Additionally, the applicability of VMD is shown to be superior to that of the CEEMDAN algorithm for traffic flow data. Moreover, the SMA method exhibits higher suitability in hyperparameter optimization compared to other optimization algorithms.

TABLE I. PERFORMANCE COMPARISON OF DIFFERENT MODELS

Model	MSE	RMSE	MAPE	EV
VMD-SVR	1.283	1.530	8.297	0.921
SMA-SVR	4.267	5.433	29.705	0.322
CEEMDAN-SMA-SVR	2.508	3.381	16.122	0.685
VMD-SMA-LSTM	0.629	0.985	4.430	0.965
VMD-SMA-RNN	0.218	0.612	42.709	0.987
VMD-SSA-SVR	0.281	0.342	24.675	0.991
VMD-SMA-SVR	0.228	0.292	1.292	0.996

Fig. 2 illustrates the results of the VMD algorithm, which decomposes the traffic flow time series data into five IMFs. The VMD algorithm breaks down the highly fluctuating traffic flow data into IMFs of different frequencies, enabling the prediction model to more accurately capture the trends and patterns of each component, thereby enhancing the model's performance.

Fig. 3 and Fig. 4 respectively compare the prediction results of the VMD-SMA-SVR model and the VMD-SMA-LSTM model against the actual data values. As shown, the prediction curve of the VMD-SMA-SVR model closely follows the actual values for most of the time, indicating a generally superior prediction performance. In contrast, the VMD-SMA-LSTM model exhibits larger deviations in areas with significant fluctuations. This suggests that, for single-source data with a smaller dataset, the SVR model offers better prediction performance. Furthermore, during model training, it was observed that due to the complex network structure of the LSTM model, it has higher computational complexity compared to the SVR model, requiring more computational resources and longer training time.

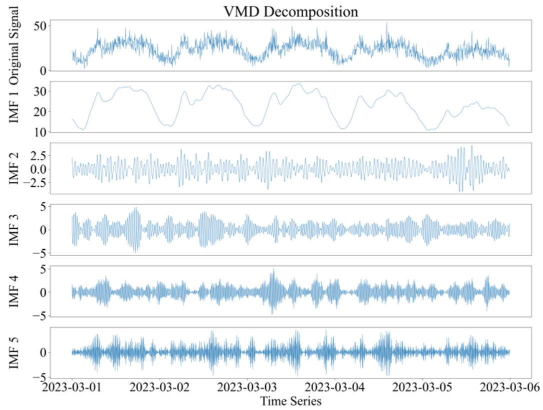

Figure 2. VMD decomposition results

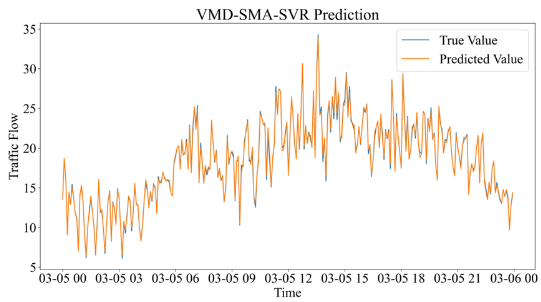

Figure 3. Comparison of VMD-SMA-SVR model predictions and true values

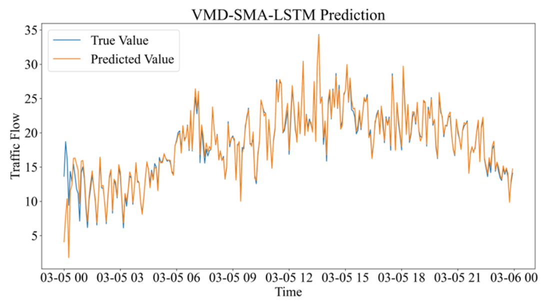

Figure 4. Comparison of VMD-SMA-LSTM model predictions and true values

IV. CONCLUSIONS

Enhancing the accuracy of short-term traffic flow forecasting is crucial for optimizing traffic planning and management comprehensively. This optimization not only significantly improves travel experience and efficiency but also effectively reduces energy consumption and environmental pollution, thereby greatly enhancing traffic safety. Accordingly, this paper proposes a short-term traffic flow forecasting model for highways based on VMD-SMA-SVR. Through comparative analysis of different models, the following conclusions were drawn:

(1) For nonlinear traffic flow time series data with substantial noise, employing the VMD algorithm for smoothing can effectively avoid mode mixing and yield components with different bandwidths, demonstrating good adaptability and decomposition performance.

(2) Comparative analysis with other heuristic algorithms indicates that the SMA algorithm, by leveraging the adaptive mechanism of slime molds, enhances global search capability. It requires fewer parameters and shows faster convergence and better performance in finding optimal hyperparameters.

(3) Experimental results show that the proposed VMD-SMA-SVR model significantly improves prediction accuracy in short-term traffic flow forecasting.

(4) The proposed VMD-SMA-SVR model exhibits high prediction accuracy, fewer model parameters, high computational efficiency, and limited resource consumption, making it well-suited for practical short-term traffic flow forecasting applications.

ACKNOWLEDGMENT

This work was supported by the Independent Research Project of Zhejiang Scientific Research Institute of Transport under grant no. ZK202411.

REFERENCES

[1] M. A. Mondal and Z. Rehena, "Stacked LSTM for short-term traffic flow prediction using multivariate time series dataset," Arabian Journal for Science and Engineering, vol. 47, no. 8, pp. 10515 - 10529, 2022.

[2] M. S. Ahmed and A. R. Cook, "Analysis of freeway traffic time - series data by using Box - Jenkins technique," Transportation Research Board, 1979, 722: 1 - 9.

[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, pp. 1 - 9, 2015.

[4] I. Okutani and Y. J. Stephanedes, "Dynamic prediction of traffic volume through Kalman filtering theory," Transportation Research Part B: Methodological, vol. 18, no. 1, pp. 1 - 11, 1984.

[5] Y. Chai, D. Huang, and L. Zhao, "A short - term traffic flow prediction method based on wavelet analysis and neural network," in 2016 Chinese Control and Decision Conference (CCDC), IEEE, 2016, pp. 7030 – 1034.

[6] B. L. Smith and M. J. Demetsky, "Short - term traffic flow prediction: neural network approach," Transportation Research Record, 1994 (1453).

[7] X. Ling, X. Feng, Z. Chen, et al., "Short - term traffic flow prediction with optimized multi - kernel support vector machine," in 2017 IEEE Congress on Evolutionary Computation (CEC), IEEE, 2017, pp. 294 – 300.

[8] W. Zhang, Y. Yu, Y. Qi, et al., "Short - term traffic flow prediction based on spatio - temporal analysis and CNN deep learning," Transportmetrica A: Transport Science, vol. 15, no. 2, pp. 1688 - 1711, 2019.

[9] W. Fang, W. Cai, B. Fan, et al., "Kalman - LSTM model for short - term traffic flow forecasting," in 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), IEEE, 2021.

[10] C. L. Ding and R. A. Juanatas, "A Short - term Traffic Flow Forecasting Method Based on SVM - PSO - BP Combined Model," in 2023 5th International Conference on Robotics, Intelligent Control and Artificial Intelligence (RICAI), IEEE, 2023, pp. 869 – 873.

[11] W. Wei, H. Wu, and M. H. Ma, "An autoencoder and LSTM - based traffic flow prediction method," Sensors, vol. 19, no. 13, p. 2946, 2019.

[12] C. Shuai, Z. Pan, L. Gao, et al., "Short - term traffic flow prediction of expressway: a hybrid method based on singular spectrum analysis decomposition," Advances in Civil Engineering, 2021, 2021: 1 – 10.

[13] M. Zhao, W. Zhang, K. Wang, et al., "Short - term passenger flow prediction of urban rail transit based on EMD - PSO - LSTM combined model," Railway Transp. Economy, vol. 44, no. 7, pp. 110 - 118, 2022.

[14] H. Huang, J. Chen, X. Huo, et al., "Effect of multi - scale decomposition on performance of neural networks in short - term traffic flow prediction," IEEE access, vol. 9, pp. 50994 - 51004, 2021.

[15] X. Li, "Forecasting Urban Traffic Flow Based on Support Vector Machine Optimized by Genetic Algorithm," Microelectronics & Computer, vol. 27, no. 10, pp. 186 - 188, 192, 2010.

[16] X. Yang, L. Zhang, and W. Xie, "Forecasting model for urban traffic flow with BP neural network based on genetic algorithm," in 2019 Chinese Control and Decision Conference (CCDC), IEEE, 2019, pp. 4395 – 4399.

[17] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," Advances in engineering software, vol. 69, pp. 46 - 61, 2014.

[18] J. Xue and B. Shen, "A novel swarm intelligence optimization approach: sparrow search algorithm," Systems science & control engineering, vol. 8, no. 1, pp. 22 - 34, 2020, DOI: 10.1080/21642583.2019.1708830.

[19] K. Dragomiretskiy and D. Zosso, "Variational Mode Decomposition," IEEE Trans. Signal Process., vol. 62, no. 3, pp. 531-544, 2013.

[20] S. Li, H. Chen, M. Wang, et al., "Slime mould algorithm: A new method for stochastic optimization," Future generation computer systems, vol. 111, pp. 300 - 323, 2020.

[21] D. Gürses, S. Bureerat, S. M. Sait, et al., "Comparison of the arithmetic optimization algorithm, the slime mold optimization algorithm, the marine predators algorithm, the salp swarm algorithm for real - world engineering applications," Materials Testing, vol. 63, no. 5, pp. 448 - 452, 2021.