IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 17 August 2023, accepted 4 September 2023, date of publication 7 September 2023,
date of current version 12 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3312711

RESEARCH ARTICLE

Short-Term Traffic Flow Prediction Based on VMD and IDBO-LSTM

KE ZHAO[1], DUDU GUO[2], MIAO SUN[1], CHENAO ZHAO[1], AND HONGBO SHUAI[1]
[1]School of Intelligent Manufacturing Modern Industry, Xinjiang University, Urumqi 830017, China
[2]School of Transportation Engineering, Xinjiang University, Urumqi 830017, China

Corresponding author: Dudu Guo (guodd@xju.edu.cn)

This work was supported in part by the Autonomous Region Key Research and Development Program Project 2022B01015, and in part by the Key Laboratory Open Project 2023ZDSYSKFKT06.

ABSTRACT To improve the accuracy of short term traffic flow prediction and to solve the problems of nonlinearity of short term traffic flow, more noise in the data, and more difficult to determine the parametes of long short term memory networks, a combined traffic flow prediction model based on variational modal decomposition (VMD) and improved dung beetle optimization-long short term memory network (IDBO-LSTM) is proposed. First, to extract various modal components, the historical traffic flow data are smoothed using variational modal decomposition (VMD). Second, the LSTM prediction model is built for each individual subsequence, and the parameters of the LSTM are optimized using the IDBO algorithm which combines Singer chaos mapping, variable spiral search strategy, and Levy flight strategy. Finally, to acquire the final prediction results, the predicted values of various subsequences are added up and reassembled. Experiments were conducted using data collected from eight sensors along an interstate highway in California, and taking the straight road morning peak (S-M) data as an example, compared with LSTM and VMD-LSTM, the MAE of VMD-IDBO-LSTM is reduced by 26.69 and 7.5108, MAPE is reduced by 8.08059% and 2.27569%, and RMSE is reduced by 33.6912 and 8.7657. According to the findings, the VMD-IDBO-LSTM model that was proposed is capable of significantly improving the accuracy of short-term traffic flow prediction while also effectively addressing nonlinearity, data noise, and the difficulty of identifying the LSTM parameters.

INDEX TERMS Short-time traffic flow prediction, variational modal decomposition, dung beetle optimization algorithm, long short term memory.

I. INTRODUCTION

The quantity of cars on the road is growing as a result of urbanization, which is occurring quickly, and the growth in people's living conditions, which is making the problem of traffic congestion worse [1]. Intelligent Transportation System (ITS) can solve the problem of road traffic congestion, and predictions of traffic flow over short periods of time can give it the necessary data on traffic flow for the time period that lies ahead. Therefore, it is essential to increase the stability and accuracy of short-term road traffic flow prediction in order to relieve congestion [2], [3]. Among them, short-term traffic flow prediction generally refers to the historical traffic flow data to predict the traffic flow in the next 5 to 30 minutes [4], depending on the application scenario of the prediction and the timeliness of the decision to be met. In general, short-term traffic flow forecasting focuses more on upcoming traffic conditions to provide timely and accurate information. This kind of short-term prediction can help traffic managers and traffic users make real-time decisions to optimize traffic operation and travel experience, which is of great significance for traffic management, traffic users, and urban traffic planning.

To increase the short-term traffic flow's predictability, stability, and accuracy, researchers have proposed many prediction models divided into three main categories: parametric models (statistical theory models), nonparametric models (machine learning models), and combinatorial

The associate editor coordinating the review of this manuscript and approving it for publication was Emanuele Crisostomi.

models, respectively. Parametric models mainly include Kalman filter methods [5], Grey prediction models [6], autoregressive integrated moving average models (ARIMA) [7], and so on. Non-parametric models mainly include decision trees [8], [9], support vector machine [10], [11], neural network [12], [13], [14], etc. Parametric models are particularly good at predicting smooth data, but they are not suitable for predicting traffic flows. As a result of the powerful nonlinear fitting abilities of nonparametric and combinatorial models, they are both excellent at dealing with nonlinear traffic flow data. Among them, due to their fast learning speed and powerful data-fitting capabilities, long short term memory (LSTM) networks are often used in addressing short-term traffic flow problems.

Due to the nonlinearity and high noise level of traffic flow data, it is challenging to make an accurate short-term prediction of traffic flow with a single LSTM. Therefore, many researchers at home and abroad solve the problems of nonlinear and noise-laden data by combining decomposition algorithms with LSTM. By utilizing singular spectrum analysis (SSA), Shuai et al. [15] were able to divide the flow of traffic into one main portion and three random portions. Based on these components' other properties, they predicted the various components using LSTM and support vector regression (SVR), respectively. Still, this method is prone to frequency confusion. Li et al. [16] proposed using wavelet decomposition (WD) to decompose the original traffic flow data into high and low frequencies and input CNN-LSTM for prediction. Still, the larger the number of layers decomposed by this method, the larger the distortion of the reconstructed signal will be, which has a poor effect on the effect of signal denoising to some extent. A combination EMD-PSO-LSTM prediction model was proposed by Zhao et al. [17]. This model makes use of empirical modal decomposition (EMD) to reduce the amount of noise in the data and the particle swarm optimization (PSO) method to optimize the parameters of the LSTM model. Although this method achieves data noise reduction processing, it is easy to fall into modal confounding using EMD. Although above the decomposition methods have achieved good results, the techniques' shortcomings will affect the final prediction accuracy. For this reason, Huang et al. [18] confirmed that variational modal decomposition (VMD) is the most widely adapted algorithm by comparing the performance of different decomposition algorithms combined with BiLSTM in short-time traffic flow prediction. Although different decomposition algorithms have advantages, they are generally inferior to VMD. Therefore, VMD is chosen as the decomposition algorithm in this paper.

Although LSTM has excellent advantages in short-time traffic flow prediction, improper parameter selection will lead it to fall into local optimum and poor generalization ability, thus affecting its prediction accuracy. With the emergence of emerging metaheuristics algorithm, more and more scholars have started combining metaheuristics algorithm with LSTM to solve the problems of low accuracy and difficulty tuning the LSTM prediction model. For instance, Wen et al. [19] optimized the LSTM's hidden layer count, training number, and dropout using a genetic algorithm (GA). However, the approach tends to converge slowly and enter the local optimum. On the basis of this observation, Zhang et al. [20] modified the genetic algorithm in order to accelerate the convergence of the optimized LSTM. The attention mechanism was added to the LSTM by Lan et al. [21]. The Gray Wolf Optimization (GWO) method was used to change the initial weight values of the attention mechanism in order to increase the prediction model's attention to important information. Their experiments show that combining the metaheuristic algorithm with the LSTM neural network can improve its prediction accuracy.

Using the aforementioned metaheuristic algorithm optimization method, it is possible to increase the LSTM prediction accuracy by substantially reducing the probability that the algorithm will enter a local optimum. However, due to the algorithm's lack of stochasticity, the issue of easily falling into local optimum may still occur, resulting in improper model parameter selection and low model prediction accuracy. Similarly, despite the fact that the original dung beetle optimization algorithms [22], [23] can effectively handle practical applications and are characterized by faster convergence and relatively higher solution accuracy than most algorithms, they still suffer from the issue of easily falling into local optima, which impacts their prediction accuracy in practical applications. It is logically demonstrated by the No Free Lunch (NFL) theorem [24] that no single metaheuristic method is effective for handling all optimization issues. This motivates us to keep developing new metaheuristics to solve different problems. Therefore, Zhang and Zhu [25] proposed using segmented linear chaotic mapping to improve the search capabilities of the original DBO algorithm, adaptive parameter adjustment strategy, and dimensional learning enhanced foraging search strategy and applied it with BP neural network in the field of heat-treated wood mechanical property prediction, and proved that the proposed IDBO-BP model has excellent prediction capability. Still, the improved strategy should have considered the problem that the dung beetle is prone to choose the optimal local solution and ignore the better global solution when the dung beetle stealing position is updated. There is still room for improvement in terms of the accuracy of its predictions.

In conclusion, the findings of this study present a short-term traffic flow prediction model that is based on VMD and includes an improved dung beetle algorithm to optimize LSTM for the difficulties posed by nonlinear traffic flow data, comprising high noise and difficulty in establishing the parameters of the LSTM. The original traffic flow sequence is initially broken down by the VMD algorithm into several intrinsic modal components. Then, individual LSTM prediction models are created for every component. The parameters of the LSTM models are optimized using

the IDBO algorithm that introduces Singer chaos mapping, variable spiral search strategy, and Levy flight before each model is trained. The final traffic flow prediction is created by adding and reconstructing the projected values for each component.

The rest of the paper is structured as follows: Section II delves into the model's concepts as well as the mechanics of the IDBO algorithm's improvement. Section III assesses the effectiveness of the suggested IDBO algorithm. Section IV describes the specific flow of the combined VMD-IDBO-LSTM model. Section V presents the experimental outcomes. Finally, Section VI outlines the research's significant contributions.

II. METHODOLOGY

A. VARIATIONAL MODEL DECOMPOSITION

A brand-new adaptive and entirely non-recursive signal decomposition and preprocessing technique was called Variational mode decomposition (VMD) [26]. It is possible to achieve the practical separation of intrinsic modal function (IMF), which is the most major advantage that this method has over standard modal decomposition methods such as EMD and EEMD. Additionally, the original signal can be decomposed into IMFs reflecting varying levels of randomness and volatility in the flow of traffic, which is another advantage. These IMFs reflect the regularity of the timing on different frequency bands, reducing the complexity of the traffic flow signals.

Due to the fact that the sequence of traffic flow data represents a nonlinear time series, and the decomposition algorithm of classic EMD series may have different degrees of modal confounding, the model in this paper selects VMD as the sequence decomposition algorithm when decomposing the traffic flow sequence. The specific decomposition steps are as follows:

Step1: Construct the constrained variational model.

$$\min_{\{\mu_k,\omega k\}}\left\{\sum_{k=1}^{k}\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)\times u_k(t)\right]e^{-j\omega_k^t}\right\|_2^2\right\}$$
$$\text{s.t.}\sum_{k=1}^{k}\mu_k=f \tag{1}$$

where ∂_t is the partial derivative of t, $\delta(t)$ is the impulse function, μ_k is the k-th mode function, ω_k is the center frequency of each mode, and f is the original signal.

Step2: Change the constrained variational model to the unconstrained variational model.

Equation (2) displays the altered Lagrange expression and introduces a quadratic penalty factor α and Lagrange multiplicative operator $\lambda(t)$ based on equation (1):

$$\begin{aligned}&L\left(\{\mu_k\},\{\omega_k\},\lambda\right)\\&=\alpha\sum_k\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)\mu_k(t)\right]e^{-j\omega_k^t}\right\|_2^2\\&+\left\|f(t)-\sum_{k=1}^{k}\mu_k(t)\right\|_2^2\kappa_2^2+\lambda(t),f(t)-\sum_{k=1}^{k}\mu_k(t)\end{aligned} \tag{2}$$

where $\lambda(t)$ is the Lagrangian multiplier and α is the penalty factor.

Step3: The saddle points of Eq. (2) are searched using the alternating multiplier method, and the expressions of $\{\mu_k\}$, $\{\omega_k\}$ and λ, after the alternating merit-seeking iterations are Eq. (3)-Eq. (5):

$$\hat{\mu}_k^{n+1}=\frac{\left(\hat{f}(\omega)-\sum_{i\neq k}\hat{\mu}_i(\omega)+\frac{\hat{\lambda}(\omega)}{2}\right)}{\left(1+2\alpha(\omega-\omega_k)^2\right)} \tag{3}$$

$$\omega_k^{n+1}=1\frac{\left(\int_0^\infty\omega\left|\hat{\mu}_k^{n+1}(\omega)\right|^2d\omega\right)}{\left(\int_0^\infty\left|\hat{\mu}_k^{n+1}(\omega)\right|^2d\omega\right)} \tag{4}$$

$$\hat{\lambda}^{n+1}(\omega)=\hat{\lambda}^n(\omega)+\gamma\left(\hat{f}(\omega)-\sum_k\hat{\mu}_k^{n+1}(\omega)\right) \tag{5}$$

where γ denotes the noise tolerance, and $\hat{\mu}_k^{n+1}(\omega)$, $\hat{\mu}_i(\omega)$, $\hat{f}(\omega)$ and $\hat{\lambda}(\omega)$ correspond to the Fourier transforms of μ_k^{n+1}, $\mu_i(t)$, f(t) and λ(t), respectively.

B. DUNG BEETLE OPTIMIZATION ALGORITHM

The rolling, dancing, foraging, stealing, and reproducing behaviors of the dung beetle served as the foundation for the Dung Beetle Optimization (DBO) algorithm, which is a one-of-a-kind meta-heuristic that was developed by Xue and Shen [22]. In particular, the DBO algorithm mimics a dung beetle's activity to traverse the search space and locate the best answer. Rolling, breeding, foraging, and stealing are the four primary behavioral categories of the DBO, which correlate to the four different types of dung beetles.

1) ROLLING BALL DUNG BEETLE

Dung beetles can roll in two different ways: when there are obstacles in their path and when there aren't.

As the dung beetle moves forward and does not hit an obstacle, the dung beetle must use celestial cues (sun position) to maintain the dung ball rolling in a straight line. During this process, the dung beetle's position changes with light intensity, and the place is updated as follows:

$$x_i^{t+1}=x_i^t+\alpha\times k\times x_i^{t-1}+b\times\left|x_i^t-x^w\right| \tag{6}$$

No obstacle mode. Where t denotes the number of current iterations, and $xt\ i$ is in terms of the position of the i-th dung beetle in the population at the t-th permutation. $k\in(0,0.2]$ denotes a constant of the deflection coefficient, b represents a constant belonging to (0,1), and α is a natural coefficient assigned to -1 or 1, where 1 means no deviation and -1 implies variation from the original direction. x^w the worst position in the current population, and $\left|x_i^t-x^w\right|$ is used to simulate the change in light intensity.

With obstacle mode. The dung beetle must dance to reposition itself to a different path when it reaches a barrier and is unable to proceed. By simulating the dung beetle's dancing behavior, the tangent function creates a new rolling direction that is only thought to lie between [0, π]. When the dung beetle finds a new path, the dung ball will continue to roll. As a result, the following position defines the dung beetle's dancing behavior:

$$x_i^{t+1} = x_i^t + tan(\theta)\left|x_i^t - x_i^{t-1}\right| \tag{7}$$

When $\theta = 0$, $\frac{\pi}{2}$ or π, tan(θ) is 0 or meaningless, so the dung beetle position does not change.

2) BREEDING DUNG BEETLE

In nature, dung balls are rolled to places suitable for spawning and hidden by dung beetles. Dung beetles must select the ideal location to lay their eggs in order to offer a secure environment for their young. As a result of the conversation that took place above, the authors offer the following boundary selection technique as a way to mimic the area in which female dung beetles spawn:

$$\begin{cases} Lb^* = max\left\{x^* \times (1-R), Lb\right\} \\ Ub^* = min\left\{x^* \times (1+R), Ub\right\} \end{cases} \tag{8}$$

where R = $1-\frac{t}{T_{max}}$, and $T_{\max}$ is the maximum number of iterations. Lb and Ub are the lower and upper bounds of the optimization problem, respectively, and x^* is the optimal position of the current population. The lower bound and higher terms of the optimization problem are shown by the letters Lb^* and Ub^*, respectively.

If a female dung beetle finds a suitable area to use for egg laying, she will lay her eggs in that area. It is important to keep in mind that during each iteration of the DBO algorithm, only one brood ball is laid by each female dung beetle. In addition, it is clear from Eq. (8) that the spawning area's boundary range is dynamically changing and that the R-value is the primary factor that determines this variability. The brood ball's position is therefore dynamic throughout the iterative process, which is characterized as follows:

$$B_i^{t+1} = x^* + b_1 \times \left(B_i^t - Lb^*\right) + b_2 \times \left(B_i^t - Ub^*\right) \tag{9}$$

where B_i^t is the position information of the i-th brood ball at the t-th iteration, b_1 and b_2 represent two independent random vectors of size 1 × D, and D denotes the dimensionality of the optimization problem.

3) FORAGING DUNG BEETLE

The term "baby dung beetles" refers to adult dung beetles that venture above ground in search of food and are given this moniker by certain people. In order to replicate how these dung beetles forage in nature, we also need to develop the best foraging zones to direct the dung beetles. The following are specific definitions of the borders of the best foraging area:

$$\begin{cases} Lb^l = max\left\{x^l \times (1-R), Lb\right\} \\ Ub^l = min\left\{x^l \times (1+R), Ub\right\} \end{cases} \tag{10}$$

where x^l is the ideal location globally, Lb^l and Ub^l stand for the ideal foraging area's lower and upper boundaries, respectively, and the remaining parameters are defined in Equation (8). Thus, the following information about the little dung beetle's location has been updated:

$$x_i^{t+1} = x_i^t + C_1 \times \left(x_i^t - Lb^l\right) + C_2 \times \left(x_i^t - Ub^l\right) \tag{11}$$

where C_1 signifies a random number adhering to the positive-terrestrial distribution, and C_2 means a random vector falling within the range (0,1), and x_i^t denotes position information for the i-th little dung beetle at the t-th iteration.

4) STEALING DUNG BEETLE

Dung balls are taken from other dung beetles by some dung beetles known as thieves. Equation (10) also demonstrates that x^l is the best source of food. As a result, it is advisable to compete for food close to x^l. The stolen dung beetle's location information is updated during the iterative process and is as follows:

$$x_i^{t+1} = x^l + S \times g \times \left(\left|x_i^t - x^*\right| + \left|x_i^t - x^l\right|\right) \tag{12}$$

where x_i^t denotes the position information of the i-th stealing dung beetle at the t-th iteration, g is a random vector of size 1 × D that obeys a normal distribution, and S denotes a constant.

C. IMPROVING THE DUNG BEETLE OPTIMIZATION ALGORITHM

The DBO algorithm has drawbacks, including low global searchability and premature convergence to a local optimum, despite its good performance and effective application to some engineering design problems. This study proposes an enhanced DBO algorithm with particular augmentation strategies to address these problems.

1) SINGER CHAOS MAPPING-BASED POPULATION INITIALIZATION

The original DBO algorithm relies on a randomly generated beginning population, which is prone to an uneven distribution of population, which causes the process to converge slowly and makes it simple to reach a local optimum.

Chaotic mappings are employed in the population initialization stage of the DBO to produce extremely different beginning populations, hence increasing the diversity of initial population solutions. Numerous chaotic mappings are currently available [27], primarily Singer mapping, Chebyshev mapping, Bernoulli mapping, Gaussian mapping, PWLCM mapping, etc. Singer chaotic mapping is a method for generating chaotic sequences. It is a typical representative of chaotic mapping and is widely used for the initial population generation of metaheuristic algorithms because of its traversal and non-repetitive characteristics. Additionally, Qu and Du [28] showed that Singer mappings' traversal uniformity and convergence speed are ideal for chaotic population initialization, and they experimentally proved that Singer

mappings can balance the capabilities of local and global search. To increase the initial solutions with high-quality uniform distribution, Singer chaotic mapping is presented in this research to replace the random search approach in DBO for population initialization. It is iteratively formulated as follows:

$$\varphi_{n+1} = \mu\left(7.86\varphi_n - 23.31\varphi_n^2\right) + \mu\left(28.75\varphi_n - 13.302875\varphi_n^4\right) \quad (13)$$

where φ_n, φ_{n+1} are the n-th and n+1-th chaotic values, respectively, and μ is a constant, $\mu \in (0.9,1.08)$.

2) VARIABLE SPIRAL SEARCH STRATEGY

In this paper, we add a variable spiral exploration factor [29] to update the nestling ball position and the best foraging area of the dung beetle so that the brood ball and the best foraging place of the dung beetle have multiple search paths to better adjust their positions, and balance the global search and local search ability of the algorithm.

In the process of position updating, a fixed value of the spiral parameter z will lead to a monotonic search method that may fall into a local optimum, thus weakening the search capability of the algorithm. To increase the ability of the optimal foraging area to search for brood balls and small dung beetles to explore uncharted territory, the parameter z is designed as an adaptive variable for dynamically adjusting the spiral shape of the algorithm. This increases search effectiveness and global search performance. The best foraging site approach for brood balls and small dung beetles can be calculated using the equation below:

$$B_i^{t+1} = x^* + e^{zl} \times cos(2\pi l) \times b_1 \times \left(B_i^t - Lb^*\right) + e^{zl} \times cos(2\pi l) \times b_2 \times \left(B_i^t - Ub^*\right) \quad (14)$$

$$x_i^{t+1} = e^{zl} \times cos(2\pi l) \times x_i^t + C_1 \times \left(x_i^t - Lb^l\right) + C_2 \times \left(x_i^t - Ub^l\right) \quad (15)$$

where z is the spiral parameter, k is the coefficient of variation, and in this paper, we take $k = 5$, l is the number of shapes, $l = 2\times\tau$-1, and τ is a random number within [0,1). The following equation gives the spiral parameter z:

$$z = \mathrm{e}^{k\times\cos\left(\frac{\pi t}{maxt}\right)} \quad (16)$$

3) LEVY FLIGHT

This alternating long and short-distance flight characteristic of the Levy flight [30] plays a vital role in balancing the optimization algorithm's local and global search capabilities: smaller steps of random wandering are beneficial to the algorithm for local search. In comparison, occasional jumps of more significant measures are advantageous to the algorithm for jumping out of the local optimum and improving the global search capability.

Due to the insufficiency of its stochastic strategy and the absence of efficient evasion techniques, the DBO algorithm may become local optimal in the situation of stealing dung beetle location updates and neglect the more optimal global solution. To address this issue, we introduce the Levy flight search approach to improve the ability of dung beetle stealing to find the global optimal solution while eliminating the local optimal. The updated dung beetle stealing formula is:

$$x_i^{t+1} = Levy(s) \times x^l + S \times g \times \left(\left|x_i^t - x^*\right| + \left|x_i^t - w \times x^l\right|\right) \quad (17)$$

The formula for the step size of the Levy flight is shown below:

$$\text{Levy}(s) = \frac{0.01 \times x \times \eta}{|y|^{\frac{1}{s}}} \quad (18)$$

where η is a random number between [0,1) and x and y obey a normal distribution, as follows:

$$x \sim N(0, \sigma_x), y \sim N(0, \sigma_y)$$
$$\sigma_x = \left[\frac{\Gamma(1+s)\, sin\left(\frac{\pi s}{2}\right)}{\Gamma\left(\frac{1+s}{2}\right) \times s \times 2^{\frac{s-1}{2}}}\right]^{\frac{1}{s}}, \sigma_y = 1 \quad (19)$$

W is a weighting factor with the following equation:

$$w = \frac{e^{2\times(1-\frac{t}{maxt})} - e^{-2\times(1-\frac{t}{maxt})}}{e^{2\times(1-\frac{t}{maxt})} + e^{-2\times(1-\frac{t}{maxt})}} \quad (20)$$

D. LONG SHORT TERM MEMORY NETWORK

The LSTM was an improved RNN structure that Hochreiter and Schmidhuber [31] jointly proposed. It can handle time series problems because it can approximation complex nonlinear relations and also perform associative memory. High precision, distributed storage, powerful learning capabilities, excellent robustness, and fault tolerance to noisy nerves are all benefits of this technology. Therefore, LSTM is an ideal model for traffic flow prediction.

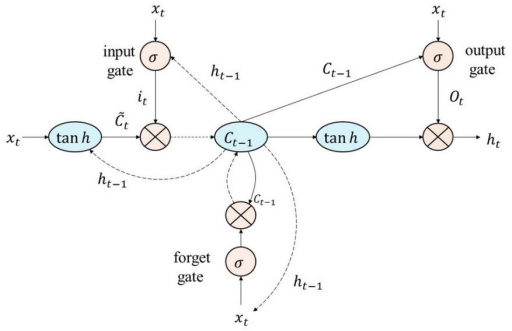

FIGURE 1. LSTM structure.

The LSTM structure is shown in Fig. 1. The normal sigmoid function and matrix multiplication are represented by σ and $\otimes$ wrapped in circles, respectively, while the tanh function is represented by tanh covered with ellipses. The three dashed lines indicated by C_{t-1} indicate hidden state h_{t-1} transitions, while the other three solid lines shown by C_{t-1} indicate normal state transitions. In addition, two dashed

lines pointed out by ⊗ complete the state update process of the neuron.

Three gating units with different functions determine when the information flow in the LSTM module flows in, out, and is forgotten, respectively. As equation (21) shows, the forget gate takes x_t and h_{t-1} as inputs and discards the information using the sigmoid function.

$$f_t = \sigma\left(\omega_{ft} x_t + \omega_{fi} h_{t-1} + b_f\right) \tag{21}$$

The input gate is completed by Equation (22) and Equation (23). The sigmoid function first determines which value to update. Afterwards the candidate vector C'' created with the tanh function and adds the state.

$$i_t = \sigma\left(\omega_{it} x_t + \omega_{ih} h_{t-1} + b_i\right) \tag{22}$$

$$\tilde{C}_t = tanh(\omega_{ct} ht - 1 + \omega_{cx} x_t + b_c) \tag{23}$$

By multiplying the old state C_{t-1} by f_t and adding it to $i_t * \tilde{C}_t$, as illustrated in equation (24), we update the state of the cell.

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{24}$$

The output gate eventually generates h_t's ultimate output. Equations (25) and (26), which divide the process into two parts, characterize these phases.

$$O_t = \sigma\left(\omega_{ot} x_t + \omega_{ot} C_{t-1} + b_o\right) \tag{25}$$

$$h_t = O_t * tanh\left(C_t\right) \tag{26}$$

where σ determines the fraction of the output C_{t-1}. Finally, equation (26) combines the new states C_t and O_t to calculate he final output h_t.

III. IDBO PERFORMANCE TESTING

This section tests the IDBO algorithm suggested in this research using eight common test functions to determine its efficacy. The experimental environment is shown in Table 1.

A. STANDARD TEST FUNCTIONS

We selected eight standard test functions from the literature [32] to evaluate the IDBO algorithm in this paper. The f_1 to f_3 are unimodal test functions, and the f_4 to f_8 are multimodal test functions.

A unimodal test function is a function that has a globally optimal solution in its domain. Specifically, a unimodal function has a unique minimum value at a certain position, and the value of the function at the rest of the position is higher than that minimum. The unimodal test function is usually used to test the performance of the optimization algorithm on global search ability. By using the unimodal test function, we can more intuitively observe the speed and accuracy of the algorithm approaching the global optimal solution in the search process. If the algorithm can find the global optimal solution in a short time and achieve high precision, it shows that the algorithm has good convergence and accuracy.

A multimodal test function is a function with multiple locally optimal solutions on the domain. Specifically, a multimodal function has a local minimum at multiple locations, of which only one is a global optimal solution, and the rest are local optimal solutions. The multimodal test function is usually used to test the performance of the optimization algorithm in local search ability. By using the multimodal function, we can observe the search behavior of the algorithm around various local optimal solutions, including whether it can jump out of the local optimal solution and whether it can carry out extensive exploration in the search space. These observations can help us judge the local search ability, convergence, robustness, and other performance indicators of the algorithm, and improve and optimize the algorithm.

Table 3 gives basic information about the standard test functions, including their function names, dimensions, expressions, search ranges, and optimal values.

B. PERFORMANCE COMPARISON

The IDBO algorithm proposed in this paper is used to compute the eight standard test functions mentioned above. The results are compared with the Gray Wolf Optimization (GWO) algorithm [33], the Sparrow Optimization (SSA) algorithm [34], the Whale Optimization (WOA) algorithm [35], the Nighthawk Optimization (NGO) algorithm [36], and the original DBO algorithm in a comprehensive manner.

To avoid the chance of experimental results, the initial population size was set to 30, the maximum number of iterations was set to 500, and 30 independent experiments were conducted on 8 standard test functions respectively. To evaluate the search performance of the algorithm, the best value (BV), the worst value (WV), the mean value (MV), and the standard deviation (SD) of the search results are recorded.

The parameter Settings of each algorithm are shown in Table 2. It should be noted that the values of these parameters are set according to the suggestions of their reference papers.

The experimental results are shown in Table 4. From Table 4, the original DBO algorithm cannot better solve the complex standard test functions. In contrast, the IDBO algorithm has zero BV, WV, MV, and SD on the test functions except for f_6, i.e., it has solved to the global optimum stably. For the multi-peaked function f_6 with valley shape, the global optimum is located at the bottom of the valley-seeking difficulty, and the DBO algorithm has better seeking accuracy for this function. Hence, the IDBO algorithm fails to improve the seeking accuracy of this function. Although the f_6 function fails to find the global optimum, it is still evident that all three improvement strategies have different degrees of improvement compared to the DBO algorithm.

The convergence curves of the six metaheuristic algorithms are plotted according to the experimental data, which can more intuitively observe the effectiveness of the IDBO algorithm and reflect the convergence speed, stability, and ability of the algorithm to escape from the local optimum.

TABLE 1. Experimental environment setting.

Name	Setting
CPU	Intel(R) Core(TM) i7-7700HQ processor
RAM	16 GB
Software	MATLAB R2018a
Operating system	Windows 10

TABLE 2. Parameter setting of the various algorithms.

Algorithm	Parameter	Setting
GWO	a	[2,0]
SSA	Proportion of leaders	20%
	The safety threshold	0.8
WOA	a	[2,0]
NGO	r	[0,1]
	i	1 or 2
DBO	k	0.1
	s	0.5
IDBO	k	0.1
	s	0.5

Fig. 2 depicts the convergence trend with the horizontal axis representing the number of iterations and the vertical axis representing the order of magnitude of the fitness value, which is expressed as a logarithm with a base of 10 to better demonstrate the convergence trend.

The six methods' convergence curves for the single-peak function are shown in Fig. 2(a) through 2(c). The degree of curve drop among them impacts how closely the algorithm's output resembles the ideal value. The IDBO method converges the fastest, as shown in Fig. 2. In contrast, the other five algorithms have a slowly decreasing trend, and the curve reduces little in 500 iteration cycles, making it difficult to approach the optimal value. This demonstrates that the population initialization by Singer chaotic mapping makes the overall population location distribution more uniform and increases the population diversity, enabling the IDBO algorithm to update the location quickly and accelerate the convergence in the early stage. At the same time, improving dung beetle reproduction and small dung beetle foraging by variable spiral search strategy speeds up the algorithm search efficiency and global search performance. It avoids duplicate values in the process of updating the position, further accelerates the convergence speed of the IDBO algorithm, dramatically improves the ability of the IDBO algorithm global search, and also has a vital help to balance the last local search.

Fig. 2(d) through 2(h) display the convergence curves for the six approaches on the multi-peak function. When the multi-peak function is used, it is clear that the IDBO algorithm has the fastest convergence speed among the six methods, demonstrating once more the effectiveness of the improved population initialization, dung beetle reproduction, and small dung beetle foraging strategies, all of which significantly speed up convergence. This also significantly speeds up the convergence speed. Also, the solution accuracy of the IDBO algorithm is optimal due to the introduction of Levy flight in the dung beetle stealing phase (local search), which significantly improves the drawback that the standard DBO algorithm tends to fall into local optimal solutions later. By further judging the dung beetle stealing location update, it can help to jump out of the local optimum, which makes the algorithm's local optimization finding the ability to be improved considerably.

The convergence of various algorithms for single-peak and multi-peak functions is depicted in Fig. 2. The IDBO algorithm performs exceptionally well in both single-peak and multi-peak functions, searching for convergence fast and deepening development, demonstrating a fair balance between the global search ability and the local development capacity of the IDBO algorithm.

Under the same practical condition limitations, the IDBO method has a faster convergence time and a greater solution accuracy, demonstrating the efficacy of the improvement technique suggested in this research and significantly improving the performance of the dung beetle algorithm. This is demonstrated by observing and analyzing experimental data and convergence curves.

IV. SHORT-TIME TRAFFIC FLOW PREDICTION BASED ON A COMBINED VMD-IDBO-LSTM MODEL

A. LSTM MODEL BASED ON IDBO OPTIMIZATION

The number of implied layer neurons, learning rate, and training times are all crucially important choices that might affect how accurately the LSTM model predicts short-term traffic flow in this study. Although these parameters can be selected by optimizing the LSTM using the DBO algorithm, the DBO algorithm suffers from problems such as uneven initial population distribution, weak local exploitation ability, and the tendency to fall into local optimality. The DBO algorithm has some drawbacks, so this work suggests a better DBO algorithm to optimize the number of hidden layer neurons, learning rate, and training durations of LSTM in order to increase the predictability of the LSTM model for short-term traffic flow.

In the beginning Singer chaotic mapping is used to initialize the population, increasing the homogeneity and diversity of dung beetle populations while also making the search space more homogeneous, improving the ability to conduct global searches. what's more, In order to improve global exploring capability, a variable spiral search strategy is employed to dynamically modify the position of the brood ball and the location of the optimal foraging region for small dung beetles. In the end, the Levy flying approach strikes a balance between convergence precision and search diversity.

The process of the IDBO-LSTM model algorithm is shown in Fig. 3, and the main steps are as follows:

Step 1: Initialize the IDBO algorithm parameters, which contain the population size of dung beetles (pop), the fraction

TABLE 3. Standard test functions.

Name	Dim	Function	Range	Optimal value
Sphere	30	$f_1(x)=\sum_{i=1}^{n}x_i^2$	[-100,100]	0
SumSquares	30	$f_2(x)=\sum_{i=1}^{n}ix_i^2$	[-10,10]	0
Schwefel2.22	30	$f_3(x)=\sum_{i=1}^{n}\lvert x_i\rvert+\prod_{i=1}^{n}\lvert x_i\rvert$	[-10,10]	0
SumPower	30	$f_4(x)=\sum_{i=1}^{n}\lvert x_i\rvert^{i+1}$	[-10,10]	0
Rastrigin	30	$f_5(x)=\sum_{i=1}^{n}\left[x_i^2-10\cos(2\pi x_i)+10\right]$	[-5.12,5.12]	0
Ackley	30	$f_6(x)=-20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i}^{n}x_i^2}\right)-exp\left(\frac{1}{n}\sum_{i=1}^{n}cos(2\pi x_i)\right)+20+e$	[-32,32]	0
Griewank	30	$f_7(x)=\frac{1}{4000}\sum_{i=1}^{n}x_i^2-\prod_{i=1}^{n}cos\left(\frac{x_i}{\sqrt{i}}\right)+1$	[-600,600]	0
SchafferN.2	2	$f_8(x)=0.5+\frac{sin^2\left(x_1^2-x_2^2\right)-0.5}{\left[1+0.001\left(x_1^2+x_2^2\right)\right]^2}$	[-100,100]	0

of distinct species of dung beetles in the population, variable parameter dimension (dim), upper and lower bounds, and other parameters.

Step 2: Determine the hyperparameters of the optimization search in the LSTM model and the optimization search's range.

Step 3: Initialize the dung beetle population location using Singer chaos mapping, as shown in equation (13).

Step 4: Calculate each dung beetle's adaptation ratings and record the ideal places.

Step5: All dung beetles should have their positions updated; If they are rolling, the rolling action in equation (6) in the unobstructed mode or the dancing action in equation (7) in the obstacle mode should update their positions; If it is a breeding or foraging dung beetle, it changes its position through breeding or foraging using equations (14) or (15), which involve a variable spiral search technique. If it's a stealing dung beetle, its position is updated by stealing actions, as described in equation (17), which introduces Levy flight strategy.

Step 6: Determine whether the individual dung beetle's location exceeds the lower bound *Lb* and upper bound *Ub* after updating.

Step 7: Update the current optimal solution and its fitness value.

Step 8: Steps 4 through 7 should be repeated as needed to attain the maximum number of iterations and output the best parameters to the LSTM model.

B. VMD-IDBO-LSTM MODEL PREDICTION PROCESS

This paper's short-term traffic flow forecasting model consists of four steps:

Step1: Using VMD decomposition, the original traffic flow sequence is divided into intrinsic model functions (IMF) with finite bandwidth and various center frequencies.

Step2: The IDBO-LSTM model of each IMF is built separately to obtain the predicted values of each IMF.

Step3: Reconstruct the predicted values of each IMF. The predicted values $P_1, P_2, \ldots$, and P_n of each IMF are summed to get the final prediction.

Fig. 4 depicts the flow chart of the combined VMD-IDBO-LSTM based prediction model.

V. EXPERIMENTAL ANALYSIS

Two typical road layouts (straight road and crossroads) are examined in this study in order to assess

TABLE 4. Function test results.

Function	Index	GWO	SSA	WOA	NGO	DBO	IDBO
f_1	BV	1.7348e-29	0	5.2218e-87	5.6797e-90	5.022e-152	0
	WV	5.9903e-27	1.8782e-57	1.834e-72	5.3574e-87	7.1639e-112	0
	MV	1.0541e-27	6.2805e-59	6.3691e-74	6.6176e-88	2.5442e-113	0
	SD	1.3803e-27	3.4288e-58	3.3459e-73	1.2604e-87	1.3078e-112	0
f_2	BV	9.3765e-30	2.2061e-56	5.4738e-88	1.3712e-91	4.1327e-166	0
	WV	1.9702e-27	0	1.1889e-73	8.4084e-88	1.4584e-104	0
	MV	1.628e-28	8.2669e-58	6.4615e-75	1.3777e-88	5.0713e-106	0
	SD	3.648e-28	4.0257e-57	2.4902e-74	2.116e-88	2.6613e-105	0
f_3	BV	2.343e-17	1.0629e-107	2.3299e-56	9.5814e-47	1.4972e-83	0
	WV	2.4452e-16	1.3856e-28	5.8735e-50	8.0128e-45	3.6254e-54	0
	MV	1.0058e-16	4.6849e-30	3.5544e-51	1.1912e-45	1.2086e-55	0
	SD	6.3196e-17	2.5287e-29	1.0879e-50	1.5772e-45	6.6191e-55	0
f_4	BV	2.6723e-99	6.2132e-202	1.984e-127	9.5097e-183	1.0159e-187	0
	WV	5.498e-85	1.7317e-50	3.5827e-100	8.4709e-174	4.6142e-112	0
	MV	1.8424e-86	5.7724e-52	1.2022e-101	8.6362e-175	1.543e-113	0
	SD	1.0036e-85	3.1616e-51	6.5397e-101	0	8.4235e-113	0
f_5	BV	5.6843e-14	0	0	0	0	0
	WV	8.2608	0	0	0	9.1724	0
	MV	1.7812	0	0	0	0.30575	0
	SD	2.5556	0	0	0	1.6746	0
f_6	BV	6.4837e-14	8.8818e-16	8.8818e-16	4.4409e-15	8.8818e-16	8.8818e-16
	WV	1.4655e-13	8.8818e-16	7.9936e-15	7.9936e-15	8.8818e-16	8.8818e-16
	MV	1.0096e-13	8.8818e-16	4.6777e-15	6.0988e-15	8.8818e-16	8.8818e-16
	SD	1.8045e-14	0	2.2726e-15	1.8027e-15	0	0
f_7	BV	0	0	0	0	0	0
	WV	0.037525	0	0.26304	0	0.0098618	0
	MV	0.0063977	0	0.014427	0	0.00032873	0
	SD	0.01026	0	0.056253	0	0.0018005	0
f_8	BV	0	0	0	0	0	0
	WV	0.0097159	0	0.0097159	0.0025191	0.0097159	0
	MV	0.0047763	0	0.0055057	0.00010168	0.00032386	0
	SD	0.0048778	0	0.0048969	0.00046675	0.0017739	0

the performance of the proposed VMD-IDBO-LSTM model.

A. EXPERIMENTAL DATA

The data used in this research were gathered from the California Department of Transportation's PeMS system at eight observation stations along interstate roads in the state of California.

The data collected by the PeMS system is characterized by large data volume, real-time, dynamic, as well as data accuracy and reliability. This high-quality data is crucial for traffic flow prediction and can provide credible basic data to support the modeling and prediction accuracy, making the models based on these data for traffic flow prediction have higher accuracy and reliability.

The PeMS system is a collection and management system for traffic data such as traffic flow, speed, and vehicle occupancy. It is primarily used to collect and analyze traffic data on the California freeway network and provides real-time traffic status information and historical traffic data analysis. Therefore, PeMS can be considered as a roadway travel dyna-mics data collection system.

In this study, observation stations near the intersections for both road layouts were chosen to assess model performance. In the straight layout (Fig. 5), three sites were chosen that had traffic volumes travelling from north to south. The data obtained at sites 1 and 2 were then utilized to anticipate the traffic flow at site 3 for the following five minutes. In the intersection layout (Fig. 6), five observation sites were selected that have positive traffic connections to each other, with the specific traffic flow directions shown by the arrows in Fig. 6 and the data collected at the four sites (1)-(4) around site five were used to predict the traffic flow at site 5, located in the center of the intersection, for the next five minutes.

Due to variances in weekend traffic patterns, the data utilized for the experiment only spans 10 weekdays (March 4, 2019 to March 15, 2019) at each site. On top of that, considering that few vehicles pass by during part of the day, the experiment only used data from peak hours of 6 a.m. to 10 a.m. and 4 p.m. to 8 p.m. Data were recorded at 5 min intervals, resulting in 96 sampling points per site per day.

Two traffic flow data groups were created to build prediction models and assess their effectiveness. The 2019 training dataset included the first seven working days (March 4 to March 8 and March 11 to March 12). The test dataset was created using data from the last three working days (March 13 to 15, 2019). subsequently acquired four experimental datasets for the two distinct road layouts, namely the straight road morning peak time (S-M), straight road evening peak time (S-E), crossroad morning peak time (C-M), and crossroad evening peak time (C-E).

The traffic flow at the target station at time t+i is predicted in this study using data obtained from all stations at times t, t-i, t-2i, t-3i, and t-4i, where i is a five minute sampling period.

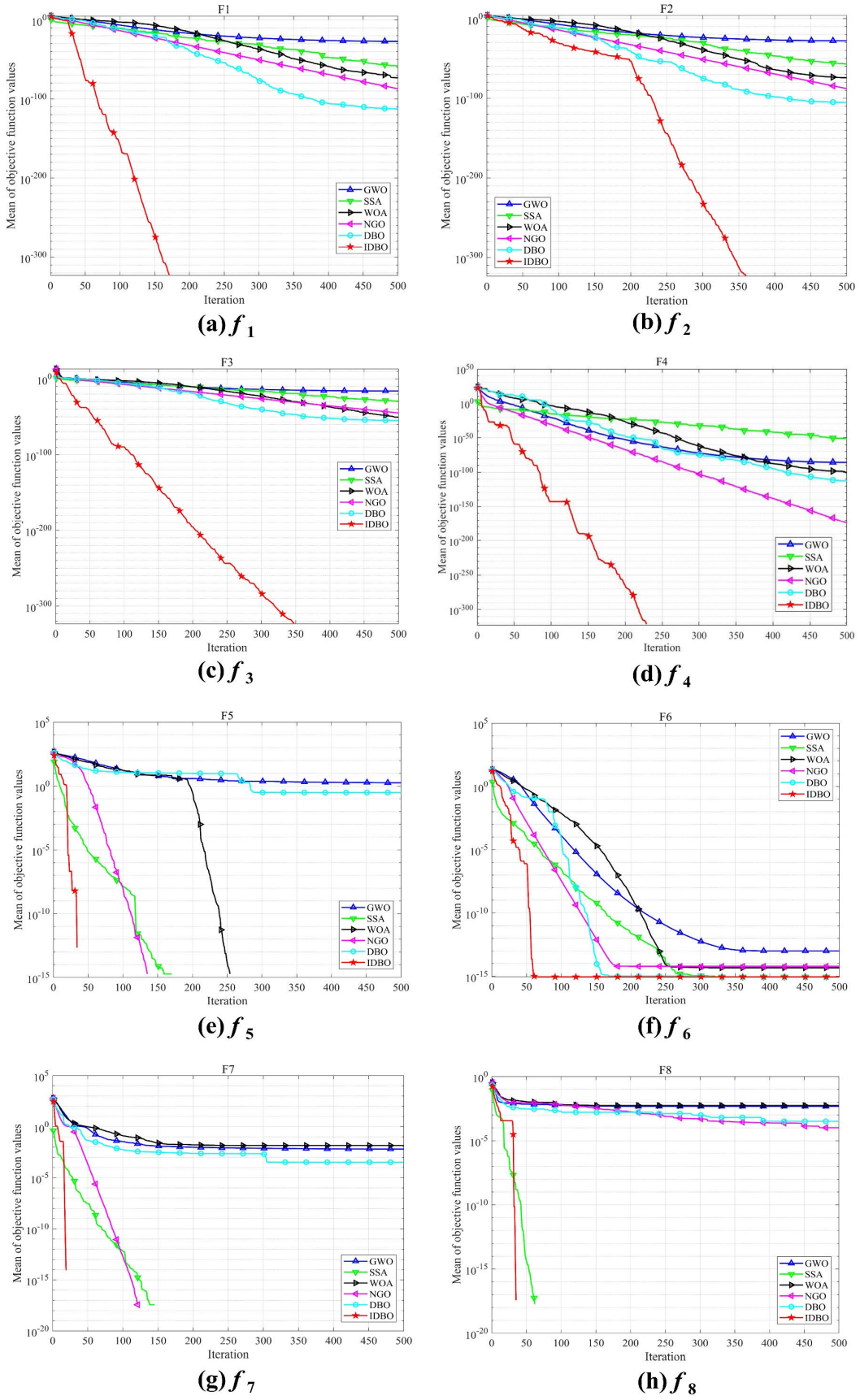

FIGURE 2. Test function convergence curve.

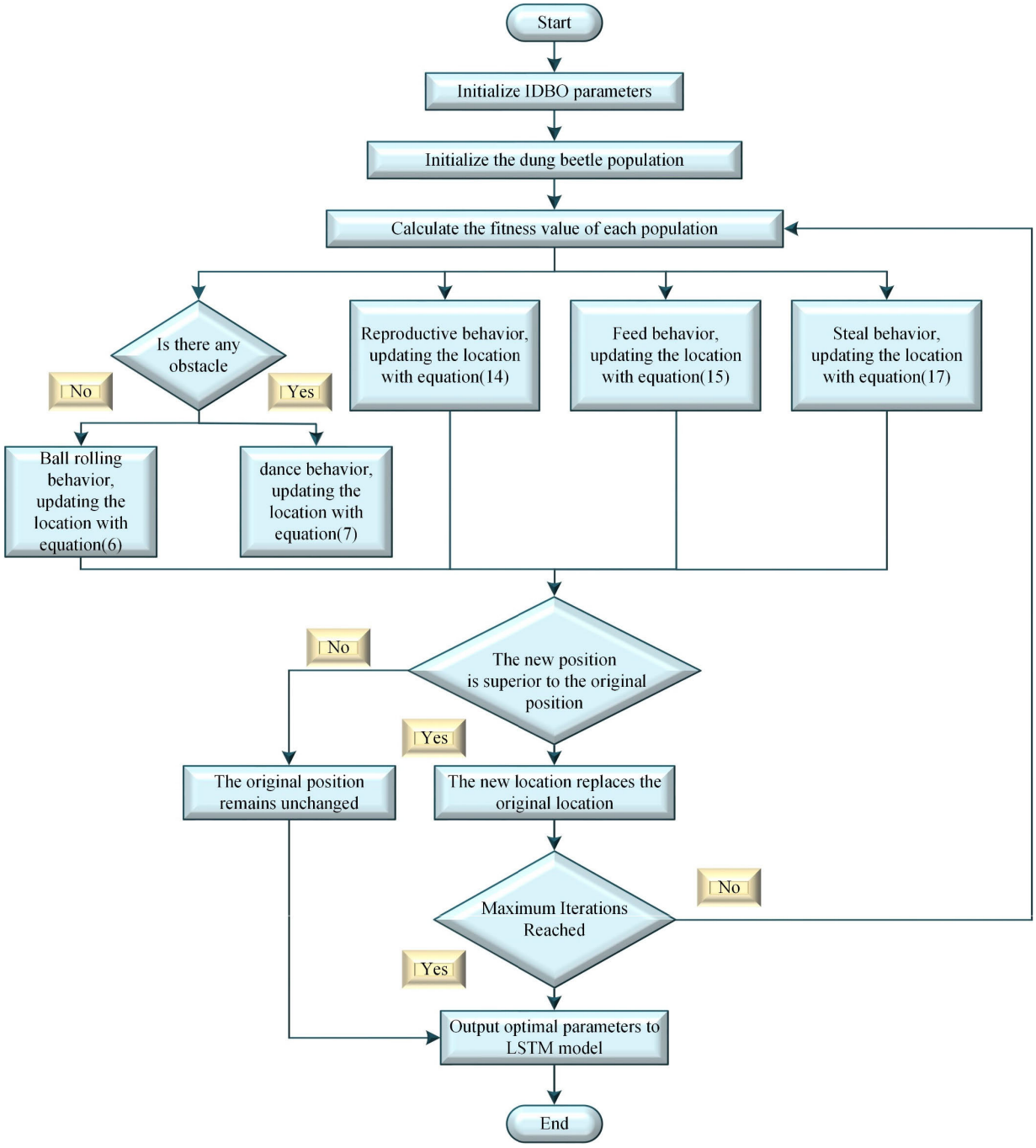

FIGURE 3. LSTM parameter optimization process based on IDBO algorithm.

B. VMD DECOMPOSITION

In Fig. 7, the raw traffic flow data shows apparent volatility, which makes direct forecasting difficult. To begin, VMD is utilized to breakdown the original traffic flow signal. Still, when performing the decomposition, the value of the decomposition number K needs to be determined first to avoid the useless components generated by the modal under-decomposition or over-decomposition due to the wrong choice of K. The final K value in this study is calculated by tracking how each IMF's center frequency changes [37] when K is steadily raised. Table 5 shows the center frequencies of each mode for different K values of S-M.

As seen in Table 5, the center frequencies of IMF5 and IMF6 are closer to each other when K=6, which makes them susceptible to modal mixing, and similarly for K=7. And when K=5, the effect of frequency separation is relatively good; that is, the center frequency interval of adjacent modes is more significant, effectively avoiding the issue of mode mixing, and the frequency information inside the sequence is relatively well discovered. To balance the number of K

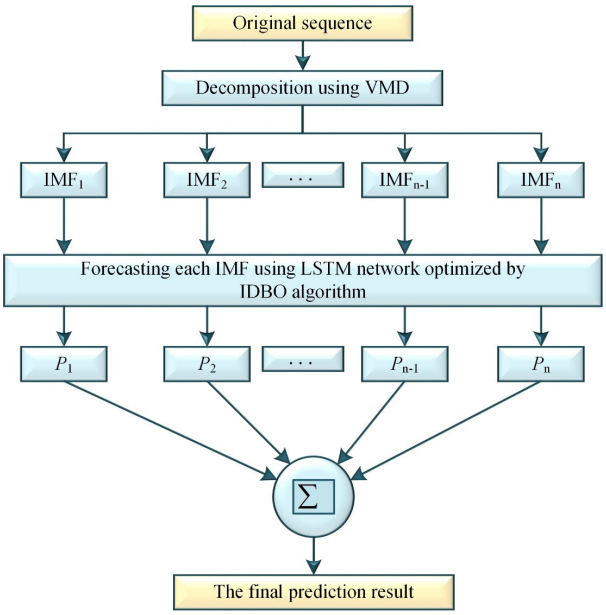

FIGURE 4. VMD-IDBO-LSTM prediction model process.

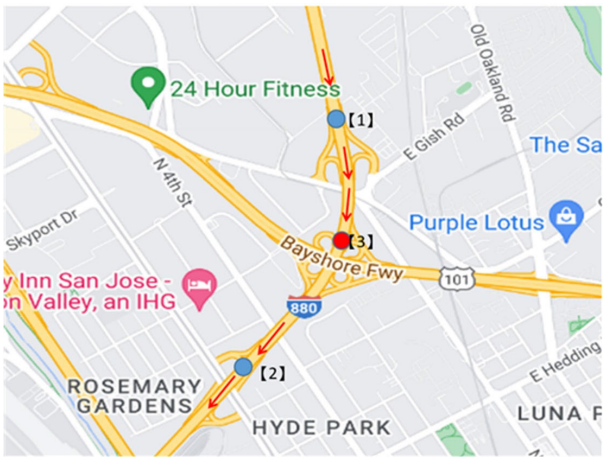

FIGURE 5. Straight road.

values and the accuracy of the decomposition, a K value of 5 is chosen. Meanwhile, to ensure the fidelity of the original sequence decomposition, the penalty parameter α, the initial central frequency ω_0, and the convergence criterion tolerance τ are set to $\alpha = 2000$, $\omega_0 = 1$ and $\tau = 1 \times 10.7$, respectively. Fig. 7 illustrates the subsequence that is obtained by VMD decomposition with K=5 when using S-M as an example.

In Fig. 7, the first sequence represents the original signal of the traffic flow, and the remaining sequences represent the intrinsic modes IMF1 to IMF5 obtained from the traffic flow after VMD decomposition. Modes IMF1 to IMF2 are weakly volatile and roughly reflect the primary trend of traffic flow over time. The frequency of modal IMF3 volatility is relatively low, but the periodicity is apparent, which can reflect the periodicity of traffic flow to a certain extent. Modes IMF4 to IMF5 have a higher frequency and relatively drastic changes, which can reflect the randomness of traffic flow to a certain extent. Each IMF can somewhat reflect the

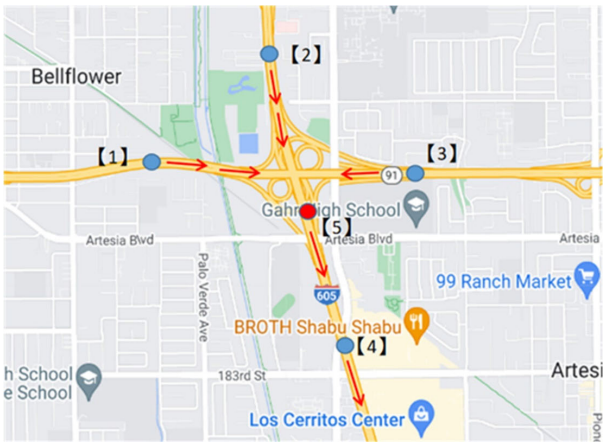

FIGURE 6. Crossroads.

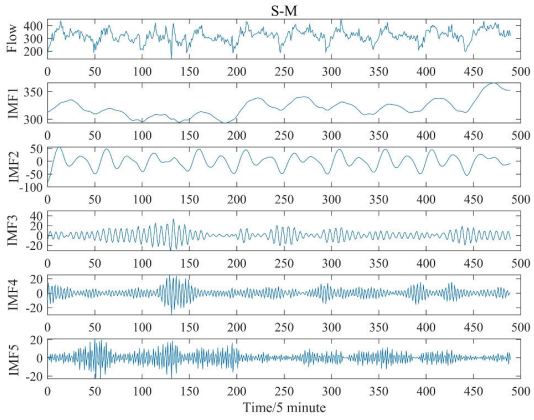

FIGURE 7. VMD decomposition of the S-M.

traffic flow characteristics, and the low-frequency IMF has weak randomness and high prediction accuracy. Therefore, the high-frequency modal prediction error determines the total traffic flow prediction error. Compared with the direct use of traffic flow raw series for prediction, VMD can lessen the influence of nonlinearity in data from traffic flows on prediction accuracy and reduce the complexity of the flow of traffic data to be forecasted, thus enhancing LSTM accuracy in forecasting.

C. EVALUATION CRITERION

The evaluation indices root mean square error (RMSE), mean absolute error (MAE), and mean fundamental percentage error (MAPE) are all used in this work.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{27}$$

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \tag{28}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{29}$$

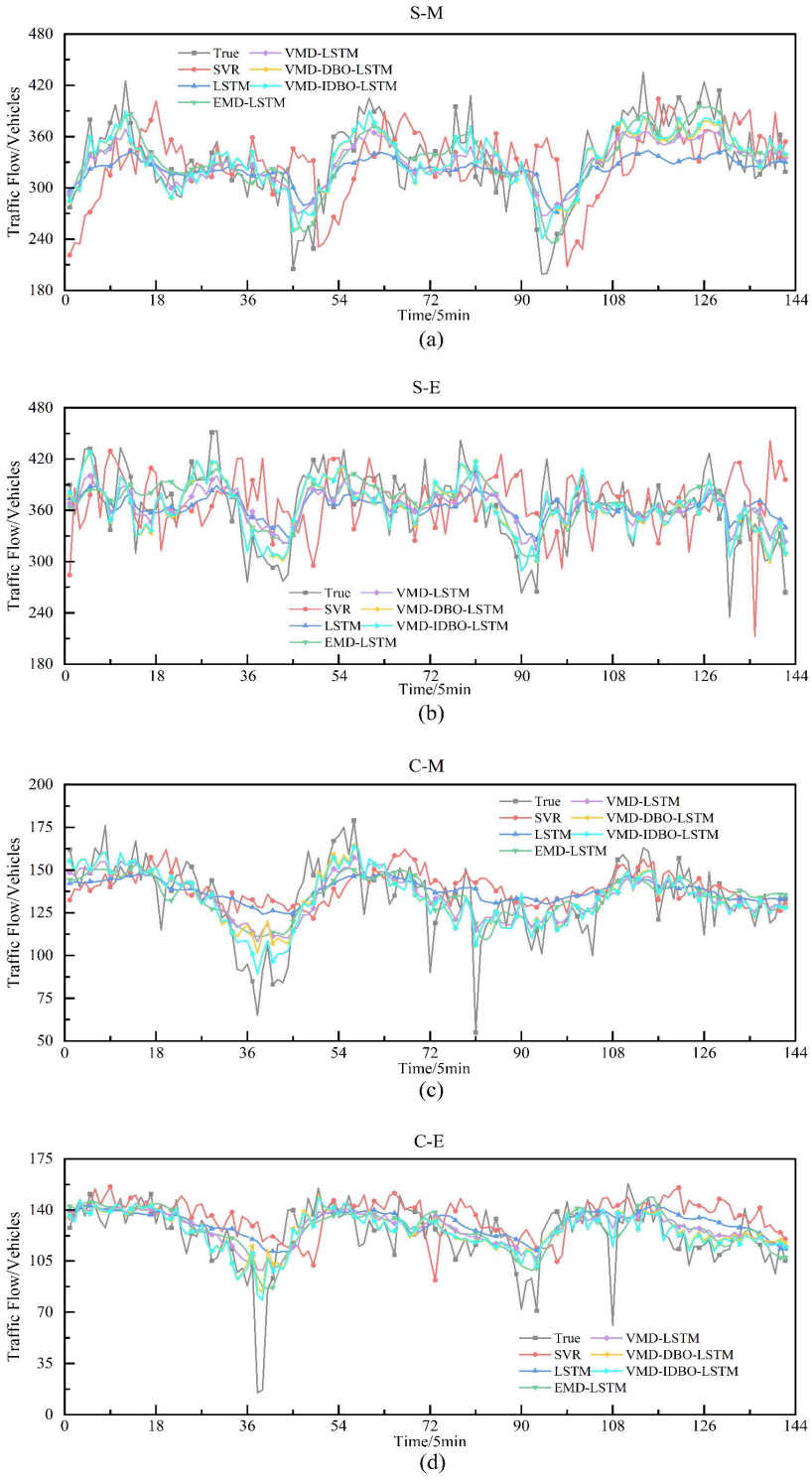

FIGURE 8. Prediction results of each model.

where the actual traffic flow (y_i) and the forecast traffic flow ($\hat{y}_i$) are the two variables, respectively. The error of the model's prediction outcomes will be reduced and the generalization ability will be stronger as MAE, MAPE, and RMSE values decrease.

D. ANALYSIS OF EXPERIMENTAL RESULTS

To validate the predictive model's effectiveness and superiority, a variety of typical prediction models are selected for comparison with it, including SVR, LSTM, EMD-LSTM, VMD-LSTM, VMD-DBO-LSTM, and the model parameters

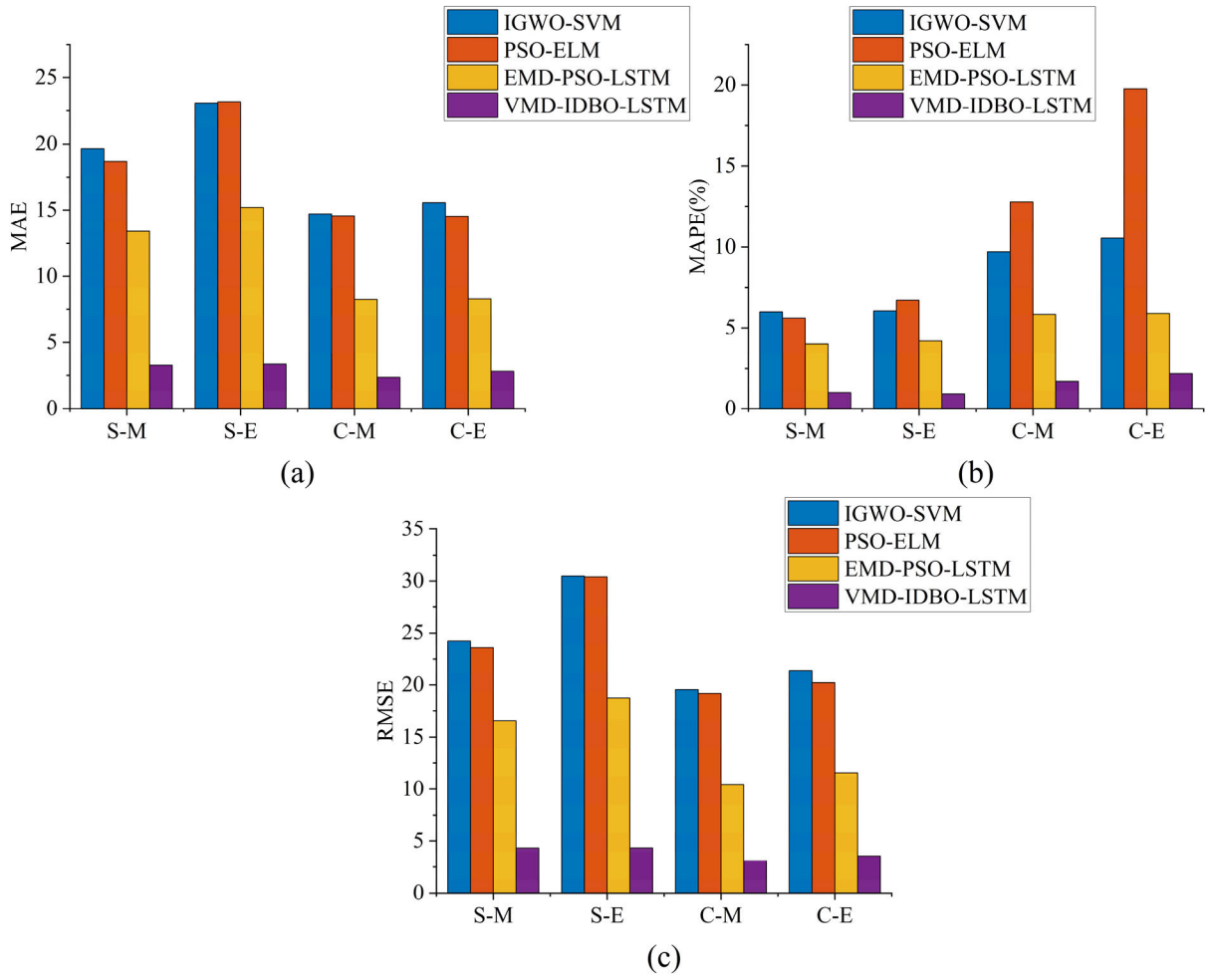

FIGURE 9. Histogram of evaluation metrics of VMD-IDBO-LSTM and other methods in the literature.

are set as follows: Population size pop=30, and dung beetle roles were divided in the following ratio 6:6:7:11. This means that out of 30 individuals, six dung beetles are used for ball rolling behavior, six dung beetles are used for reproductive behavior, seven dung beetles are defined as small dung beetles that perform foraging behavior, and the remaining 11 dung beetles are positioned as thieves and are used to perform stealing behavior. The dimension D=3, the number of neurons in the hidden layer, the maximum training period, and the initial learning rate are set to (500,300), (0,150), and (0.001, 0.01), respectively, in the range of finding the best.

Table 6 compares the errors of the six models, despite the fact that all six models made more accurate predictions for the four data sets. Still, in comparison, the prediction accuracy obtained by the combined prediction model is better than that of the single SVR and LSTM models. For the two combined prediction models of VMD-LSTM and EMD-LSTM, the prediction error obtained by VMD-LSTM is smaller, thus indicating that smoothing traffic flow data based on VMD is more advantageous and can get higher prediction accuracy. Compared with the VMD-LSTM model, the MAE of the VMD-DBO-LSTM model decreased by 7.0839, 9.3278, 1.9347, and 2.5072 for the four data sets S-M, S-E, C-M, and C-E, respectively, and MAPE decreased by 2.1341%, 2.51563%, 1.4163%, and 1.8282%. The RMSE decreased by 8.1107, 11.3763, 2.1816, and 3.6343, respectively, thus indicating that the LSTM optimized using the DBO algorithm has a more robust prediction performance after the VMD smoothing process. While the VMD-IDBO-LSTM model is more adaptive compared to the VMD-DBO-LSTM model, the MAE under S-M, S-E, C-M, and C-E decreased by 0.4269, 0.2593, 0.4683 and 0.4533, respectively, MAPE decreased by 0.14159%, 0.07573%, 0.1977%, 0.3156%, and RMSE decreased by 0.6549, 0.3256, 1.5082, 0.7765, respectively. This demonstrates that the IDBO algorithm has

TABLE 5. S-M center frequency at different *K* values.

K	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7
3	2.5674e-05	0.159 5	0.324 3				
4	9.6657e-06	0.037 1	0.250 7	0.36			
5	9.4518e-06	0.037	0.161 9	0.300 9	0.402 8		
6	9.4077e-06	0.037	0.161 3	0.250 6	0.325 5	0.404 4	
7	8.1471e-06	0.035 4	0.099 4	0.164 9	0.294 9	0.362 8	0.444 6

TABLE 6. Evaluation indicators of each model.

Data	Model	MAE	MAPE	RMSE
S-M	SVR	19.9408	6.0144%	24.2136
	LSTM	29.9937	9.0803%	38.0183
	EMD-LSTM	18.821	5.5203%	23.5085
	VMD-LSTM	10.8145	3.2754%	13.0928
	VMD-DBO-LSTM	3.7306	1.1413%	4.982
	VMD-IDBO-LSTM	**3.3037**	**0.99971%**	**4.3271**
S-E	SVR	21.7971	5.7541%	27.8949
	LSTM	26.6785	7.0978%	33.7834
	EMD-LSTM	22.2391	5.9321%	28.1328
	VMD-LSTM	12.979	3.5079%	16.0303
	VMD-DBO-LSTM	3.6512	0.99227%	4.654
	VMD-IDBO-LSTM	**3.3919**	**0.91654%**	**4.3284**
C-M	SVR	14.9492	9.9414%	19.9679
	LSTM	13.393	8.8737%	18.8398
	EMD-LSTM	10.9059	7.4941%	14.5096
	VMD-LSTM	4.7514	3.2875%	6.7945
	VMD-DBO-LSTM	2.8167	1.8712%	4.6129
	VMD-IDBO-LSTM	**2.3484**	**1.6735%**	**3.1047**
C-E	SVR	16.0431	11.059%	21.9012
	LSTM	13.8582	9.3124%	20.046
	EMD-LSTM	11.0984	7.7737%	15.7675
	VMD-LSTM	5.8019	4.297%	7.979
	VMD-DBO-LSTM	3.2947	2.4688%	4.3447
	VMD-IDBO-LSTM	**2.8414**	**2.1532%**	**3.5682**

TABLE 7. Evaluation metrics of VMD-IDBO-LSTM versus other methods in the literature.

Data	Model	MAE	MAPE	RMSE
S-M	IGWO-SVM[11]	19.6214	5.9821%	24.1968
	PSO-ELM[14]	18.6679	5.5939%	23.5865
	EMD-PSO-LSTM[17]	13.3976	3.994%	16.5411
	VMD-IDBO-LSTM	**3.3037**	**0.99971%**	**4.3271**
S-E	IGWO-SVM[11]	23.0979	6.0394%	30.4703
	PSO-ELM[14]	23.1884	6.6896%	30.3899
	EMD-PSO-LSTM[17]	15.2033	4.1806%	18.753
	VMD-IDBO-LSTM	**3.3919**	**0.91654%**	**4.3284**
C-M	IGWO-SVM[11]	14.7211	9.7123%	19.5463
	PSO-ELM[14]	14.5712	12.7732%	19.189
	EMD-PSO-LSTM[17]	8.2354	5.819%	10.4414
	VMD-IDBO-LSTM	**2.3484**	**1.6735%**	**3.1047**
C-E	IGWO-SVM[11]	15.5543	10.544%	21.3997
	PSO-ELM[14]	14.5396	19.7643%	20.2173
	EMD-PSOLSTM[17]	8.2736	5.8794%	11.5568
	VMD-IDBO-LSTM	**2.8414**	**2.1532%**	**3.5682**

greater stability and superiority-seeking ability than the DBO algorithm, which can enhance the prediction performance of the LSTM model more effectively, and also demonstrates the efficacy and superiority of the VMD-IDBO-LSTM.

Fig. 8 depicts the traffic flow prediction results from the six models. From Fig. 8, we can see that the predicted values of SVR and LSTM models deviate too much from the actual value curve and fluctuate more; VMD-IDBO-LSTM is closer to the real deal than EMD-LSTM, VMD-LSTM, VMD-DBO-LSTM, VMD-IDBO-LSTM model prediction curve is more relative to the actual value, has a higher fit. The forecasting of the outcome is much better than the other five models.

To confirm even more the benefits of the suggested model, the traffic flow error evaluation metrics of VMD-IDBO-LSTM under four different road types are compared with other models proposed in the literature, and the results are shown in Table 7 and Fig. 9. From these results, we can conclude that the VMD-IDBO-LSTM significantly outperforms other models proposed in the literature, and the MAE, MAPE, and RMSE of the VMD-IDBO-LSTM are the lowest among the predictions under the four different road types. Although the literature [11] used an enhanced GWO algorithm for optimizing SVM parameters and the literature [14] used PSO to optimize the parameters of the extreme learning machine (ELM), they ignored the nonlinearity of the original traffic flow and the fact that the data contained a lot of noise, resulting in low prediction accuracy. Meanwhile, compared with EMD-PSO-LSTM, the MAE of VMD-IDBO-LSTM decreased by 10.0939, 11.8114, 5.887, and 5.4322 for S-M, S-E, C-M, and C-E datasets respectively, and MAPE decreased by 2.99429%, 3.26406%, 4.1455%, and

3.7262%, and RMSE decreased by 12.214, 14.4246, 7.3367, and 7.9886, respectively. This is because although the EMD used in the literature [17] reduces the effect of data noise, modal aliasing is unavoidable. And both in terms of how quickly it converges and how accurately it finds a solution, the IDBO method that was detailed in this paper was superior to PSO.

VI. CONCLUSION

Improving the prediction accuracy of short-term traffic flow helps to optimize traffic planning and management, improve travel experience and efficiency, reduce energy consumption and environmental pollution, and improve traffic safety, thus contributing to the sustainable development of the urban transportation system and the improvement of residents' quality of life. Therefore, a short-time traffic flow prediction method based on VMD-IDBO-LSTM is proposed in this paper. The analysis of the arithmetic examples leads to the following conclusions:

(1) In this study, we suggest the IDBO method to address the DBO algorithm's drawbacks. The four key steps in the DBO algorithm, namely population initialization, brood ball location and optimal foraging area for small dung beetles, and dung beetle stealing location update, are optimized to improve the global and local search capability of the DBO algorithm, respectively. This is accomplished by the introduction of Singer chaotic mapping, variable spiral search strategy, and Levy flight strategy. Meanwhile, the IDBO algorithm's effectiveness was examined for eight standard test functions, and five metaheuristics, including the DBO algorithm, were selected. IDBO algorithm outperforms other metaheuristic algorithms in solving single-peak and multi-peak functions, according to the results., demonstrating the efficacy of the revised method.

(2) For the traffic flow sequences that are nonlinear and contain much noise in the data, the VMD method is used to smooth them and obtain several sub-series with more robust regularity, which can not only avoid the modal mixing phenomenon but also reduce the error in decomposition prediction reconstruction, with better adaptability and decomposition effect.

(3) LSTM model accuracy depends directly on hidden layer neurons, learning rate, and training times. This problem can be solved by using the IDBO algorithm to provide the LSTM model with optimal parameters. Combining these two methods can improve short-time traffic flow prediction accuracy and stability.

(4) The experimental findings demonstrate that, whether compared to other single models or the combination of EMD and VMD models, the combined VMD-IDBO-LSTM model suggested in this work greatly increased prediction accuracy. Compared with LSTM, the MAE of the VMD-IDBO-LSTM model decreased by 26.69, 23.2866, 11.0446, and 11.0168 for S-M, S-E, C-M, and C-E datasets, respectively, and MAPE decreased by 8.08059%, 6.18126%, 7.2002%, and 7.1592%, and RMSE decreased by 33.6912, 29.455, 15.7351, and 16.4778, respectively. Compared with the VMD-LSTM, the MAE of the VMD-IDBO-LSTM model decreased by 7.5108, 9.5871, 2.403, and 2.9605 for the S-M, S-E, C-M, and C-E datasets, respectively, and the MAPE decreased by 2.27569%, 2.59136%, 1.614%, and 2.1438%, and RMSE decreased by 8.7657, 11.7019, 3.6898, and 4.4108, respectively. The effectiveness and superiority of VMD-IDBO-LSTM in dealing with the problems of short-time traffic flow nonlinearity, high noise content in the data, and the difficulty of determining the parameters of the LSTM neural network are demonstrated.

REFERENCES

[1] K. Jin, S. Sun, H. Li, and F. Zhang, "A novel multi-modal analysis model with Baidu search index for subway passenger flow forecasting," *Eng. Appl. Artif. Intell.*, vol. 107, Jan. 2022, Art. no. 104518.

[2] L. Cai, M. Lei, S. Zhang, Y. Yu, T. Zhou, and J. Qin, "A noise-immune LSTM network for short-term traffic flow forecasting," *Chaos, Interdiscipl. J. Nonlinear Sci.*, vol. 30, no. 2, Feb. 2020, Art. no. 023135.

[3] H. Zheng, F. Lin, X. Feng, and Y. Chen, "A hybrid deep learning model with attention-based conv-LSTM networks for short-term traffic flow prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 11, pp. 6910–6920, Nov. 2021.

[4] M. A. Mondal and Z. Rehena, "Stacked LSTM for short-term traffic flow prediction using multivariate time series dataset," *Arabian J. Sci. Eng.*, vol. 47, no. 8, pp. 10515–10529, Aug. 2022.

[5] L. Cai, Z. Zhang, J. Yang, Y. Yu, T. Zhou, and J. Qin, "A noise-immune Kalman filter for short-term traffic flow forecasting," *Phys. A, Stat. Mech. Appl.*, vol. 536, Dec. 2019, Art. no. 122601.

[6] B. Zeng and C. Li, "Improved multi-variable grey forecasting model with a dynamic background-value coefficient and its application," *Comput. Ind. Eng.*, vol. 118, pp. 278–290, Apr. 2018.

[7] J. Wang, W. Deng, and Y. Guo, "New Bayesian combination method for short-term traffic flow forecasting," *Transp. Res. C, Emerg. Technol.*, vol. 43, pp. 79–94, Jun. 2014.

[8] S. S. Yassin and Pooja, "Road accident prediction and model interpretation using a hybrid K-means and random forest algorithm approach," *Social Netw. Appl. Sci.*, vol. 2, no. 9, pp. 1–13, Sep. 2020.

[9] S. Yang, J. Wu, Y. Du, Y. He, and X. Chen, "Ensemble learning for short-term traffic prediction based on gradient boosting machine," *J. Sensors*, vol. 2017, pp. 1–15, May 2017.

[10] C. Luo, C. Huang, J. Cao, J. Lu, W. Huang, J. Guo, and Y. Wei, "Short-term traffic flow prediction based on least square support vector machine with hybrid optimization algorithm," *Neural Process. Lett.*, vol. 50, no. 3, pp. 2305–2322, Dec. 2019.

[11] Z. He, X. Wu, and Z. Liu, "Optimized SVM model for short-term traffic flow prediction based on improved gray wolf optimize," *Xiamen Univ. Nat. Sci.*, vol. 61, no. 2, pp. 288–297, Mar. 2022.

[12] W. Zheng, D.-H. Lee, and Q. Shi, "Short-term freeway traffic flow prediction: Bayesian combined neural network approach," *J. Transp. Eng.*, vol. 132, no. 2, pp. 114–121, Feb. 2006.

[13] Z. Zhao, W. Chen, X. Wu, P. C. Y. Chen, and J. Liu, "LSTM network: A deep learning approach for short-term traffic forecast," *IET Intell. Transp. Syst.*, vol. 11, no. 2, pp. 68–75, Feb. 2017.

[14] W. Cai, J. Yang, Y. Yu, Y. Song, T. Zhou, and J. Qin, "PSO-ELM: A hybrid learning model for short-term traffic flow forecasting," *IEEE Access*, vol. 8, pp. 6505–6514, 2020.

[15] C. Shuai, Z. Pan, L. Gao, and H. Zuo, "Short-term traffic flow prediction of expressway: A hybrid method based on singular spectrum analysis decomposition," *Adv. Civil Eng.*, vol. 2021, pp. 1–10, Sep. 2021.

[16] Y. Li, S. Chai, Z. Ma, and G. Wang, "A hybrid deep learning framework for long-term traffic flow prediction," *IEEE Access*, vol. 9, pp. 11264–11271, 2021.

[17] M. Zhao, W. Zhang, K. Wang, and H. Li, "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.

[18] 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.

[19] H. Wen, D. Zhang, and S. Lu, "Application of GA-LSTM model in highway traffic flow prediction," *J. Harbin Inst. Technol.*, vol. 51, no. 9, pp. 81–87 and 95, 2019.
[20] J. Zhang, S. Qu, Z. Zhang, and S. Cheng, "Improved genetic algorithm optimized LSTM model and its application in short-term traffic flow prediction," *PeerJ Comput. Sci.*, vol. 8, p. e1048, Jul. 2022.
[21] T. Lan, X. Zhang, D. Qu, Y. Yang, and Y. Chen, "Short-term traffic flow prediction based on the optimization study of initial weights of the attention mechanism," *Sustainability*, vol. 15, no. 2, p. 1374, Jan. 2023.
[22] J. Xue and B. Shen, "Dung beetle optimizer: A new meta-heuristic algorithm for global optimization," *J. Supercomput.*, vol. 79, no. 7, pp. 7305–7336, May 2023.
[23] C. Wu, J. Fu, X. Huang, X. Xu, and J. Meng, "Lithium-ion battery health state prediction based on VMD and DBO-SVR," *Energies*, vol. 16, no. 10, p. 3993, May 2023.
[24] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 67–82, Apr. 1997.
[25] R. Zhang and Y. Zhu, "Predicting the mechanical properties of heat-treated woods using optimization-algorithm-based BPNN," *Forests*, vol. 14, no. 5, p. 935, May 2023.
[26] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014.
[27] F. B. Demir, T. Tuncer, and A. F. Kocamaz, "A chaotic optimization method based on logistic-sine map for numerical function optimization," *Neural Comput. Appl.*, vol. 32, no. 17, pp. 14227–14239, Sep. 2020.
[28] P. Qu and F. Du, "Improved particle swarm optimization for laser cutting path planning," *IEEE Access*, vol. 11, pp. 4574–4588, 2023.
[29] C. Ouyang, Y. Qiu, and D. Zhu, "Adaptive spiral flying sparrow search algorithm," *Sci. Program.*, vol. 2021, pp. 1–16, Aug. 2021.
[30] R. Jensi and G. W. Jiji, "An enhanced particle swarm optimization with Levy flight for global optimization," *Appl. Soft Comput.*, vol. 43, pp. 248–261, Jun. 2016.
[31] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.
[32] X. Song, M. Zhao, Q. Yan, and S. Xing, "A high-efficiency adaptive artificial bee colony algorithm using two strategies for continuous optimization," *Swarm Evol. Comput.*, vol. 50, Nov. 2019, Art. no. 100549.
[33] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014.
[34] J. Xue and B. Shen, "A novel swarm intelligence optimization approach: Sparrow search algorithm," *Syst. Sci. Control Eng.*, vol. 8, no. 1, pp. 22–34, Jan. 2020.
[35] S. Mirjalili and A. Lewis, "The whale optimization algorithm," *Adv. Eng. Softw.*, vol. 95, pp. 51–67, May 2016.
[36] M. Dehghani, Š. Hubálovský, and P. Trojovský, "Northern goshawk optimization: A new swarm-based algorithm for solving optimization problems," *IEEE Access*, vol. 9, pp. 162059–162080, 2021.
[37] N. Huang, Y. Wu, G. Cai, H. Zhu, C. Yu, L. Jiang, Y. Zhang, J. Zhang, and E. Xing, "Short-term wind speed forecast with low loss of information based on feature generation of OSVD," *IEEE Access*, vol. 7, pp. 81027–81046, 2019.

KE ZHAO was born in 2000. He is currently pursuing the master's degree with Xinjiang University under the supervision of Assoc. Prof. Guo Dudu. His research interest includes transportation big data.

DUDU GUO was born in 1985. She received the master's degree from Beijing Jiaotong University, in 2011. She is currently an Associate Professor with the School of Transportation Engineering, Xinjiang University. Her research interest includes information acquisition for intelligent transportation systems.

MIAO SUN was born in 2000. She is currently pursuing the master's degree with Xinjiang University under the supervision of Assoc. Prof. Dudu Guo. Her research interest includes transportation big data.

CHENAO ZHAO was born in 1999. He is currently pursuing the master's degree with Xinjiang University under the supervision of Assoc. Prof. Dudu Guo. His research interest includes computer vision.

HONGBO SHUAI was born in 2001. He is currently pursuing the master's degree with Xinjiang University under the supervision of Assoc. Prof. Dudu Guo. His research interest includes computer vision.