A Novel Combined Prediction Model for Monthly Mean Precipitation With Error Correction Strategy

GUOHUI LI[ID], WANNI CHANG[ID], AND HONG YANG[ID]

School of Electronic Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China

Corresponding authors: Guohui Li (lghcd@163.com) and Hong Yang (uestcyhong@163.com)

ABSTRACT Precipitation is an important parameter of water resource management, flood warning and hydrological analysis, so it is important to predict rainfall accurately. However, many previous studies did not extract the information of error series and only used a single model to predict rainfall data, ignoring the importance of model stability. Therefore, based on the idea of combination prediction and error correction strategy, this paper proposes a novel combined prediction model for monthly mean precipitation. It combines the variational mode decomposition (VMD), the improved butterfly optimization algorithm (IBOA), the least squares support vector machine model (LSSVM), the adaptive Volterra and autoregressive moving average (ARMA) model. Firstly, in order to find the best parameters of LSSVM, an improved butterfly optimization algorithm is proposed. The simulation results show the performance of IBOA is better than that of other algorithms, such as PSO, DE and BOA. Then the IBOA-LSSVM model and Volterra model are established for the mode components of the VMD, named VMD-IBOA-LSSVM and VMD-Volterra. Secondly, to solve the problem that the uncertainty of the hydrological prediction model, a combined precipitation prediction method based on the induced ordered weighted average (IOWA) operator of VMD-IBOA-LSSVM and VMD-Volterra is proposed. Finally, the ARMA model is established to correct the error sequence of the combined forecasting model. The precipitation data of two stations in Shaanxi Province are predicted. Experiment 1 is taken as an example, the maximum error of the proposed prediction model for rainfall is less than 9 mm, and the performance of the proposed model is improved by at least 43%. It shows that the proposed model can effectively reduce the prediction error of precipitation, and provide a new idea for precipitation prediction.

INDEX TERMS Variational mode decomposition, least squares support vector machine, improved butterfly optimization algorithm, combined forecasting, precipitation prediction.

I. INTRODUCTION

Precipitation is an important parameter of water resource management, flood warning and hydrological analysis [1], [2], and the change of precipitation has a direct impact on the runoff of surface river [3]. It is a complex nonlinear system, which is affected by many factors such as sea and land location, terrain, air pressure, ocean currents and human activities [4]. Blöschl *et al.* [5] put forward 23 problems in hydrological research, and how to solve the instability of the model has become the focus of many researchers. Therefore, it is of great significance to establish an effective prediction model to accurately predict rainfall [6], [7].

Now, a lot of the traditional forecasting methods of precipitation time series are based on statistics. Papacharalampous *et al.* [8], [9] compared the application of multiple statistical models and artificial intelligence models in hydrological time series forecasting. Dastorani *et al.* [10] used ARMA, ARIMA and SARIMA to predict the monthly precipitation under the semi-arid condition in Iran. The statistical model has the advantage of small computation, but the forecasting ability of complex series is weak. With the development of artificial intelligence prediction model, it has been widely used in precipitation forecasting research in recent years. Chang *et al.* [11] established a wavelet analysis combined with artificial neural network (ANN) model to predict 30-year non-linear and unsteady precipitation signals in a certain area. The results show that this method has higher forecasting precision than the traditional method. In the field

of machine learning, Support vector machine [12] (SVM) is a typical machine learning model. Compared with the ANN method, SVM uses the minimum structural risk to optimization, which avoids falling into the local minimum, and its prediction result is more stable. Yuan et al. [13] established a WA-SVM model to forecast the change of precipitation in Shanxi Province. The prediction results provide a useful reference for controlling the drought situation in the North China Plain. Du et al. [14] established the SVM model of particle swarm optimization (PSO) to predict precipitation instead of the traditional linear threshold precipitation algorithm. SVM algorithm is simple in structure, robust and not easy to fall into a local minimum. However, the convergence time of the SVM model is long, and it is difficult to implement for large-scale samples [15]. The least squares support vector machine (LSSVM) inherits the advantages of SVM, and improves the disadvantages of SVM. At the same time, it reduces the complexity of the solution and improves the efficiency [16], [17]. However, the regularization parameters and kernel function parameters of LSSVM have a great influence on the prediction ability of the model. So some scholars introduce intelligent optimization algorithms into LSSVM to find the best model parameters. Zhao et al. [18] optimized the LSSVM model with the bat algorithm (BA) to predict the electric charge. Xiang et al. [19] optimized the LSSVM with the bird swarm algorithm (BSA) to predict the short-term wind speed. Meng [20] used PSO to optimize the LSSVM for the optimization of supporting parameters of underground caverns. Butterfly optimization algorithm (BOA) is a new intelligent optimization algorithm proposed by Arora and Singh [21] in 2018. Compared with some existing meta-heuristic algorithms, the BOA algorithm is simple in operation, with few parameters to adjust and good in robustness. Du et al. [22] used BOA to optimize the extreme learning machine (ELM) model for port throughput prediction. However, the accuracy of BOA algorithm is low. Besides, the convergence speed of algorithm is slow and easy to fall into local optimum. Inspired by [23], [24], this paper introduces nonlinear time-varying adaptive weight to improve local search and global search. Besides, Cauchy disturbance is added to mutate the butterfly position information of the global search to improve the global search ability of butterfly.

Due to the strong instability and nonlinearity of precipitation data, multi-scale decomposition of the data can decrease the modeling complexity of systems and increase the precision of the forecasting model [25], [26]. Bokde et al. [27] reviewed the application of empirical mode decomposition (EMD) and ensemble empirical mode decomposition (EEMD) methods in wind speed power prediction, and pointed out that data preprocessing plays an important role in improving the accuracy of the prediction model. Ramana et al. [28] used the wavelet neural network (WNN) to predict the precipitation in the Darjeeling mountain in eastern India. Hu et al. [29] used the combined model of EEMD and general regression neural network (GRNN) to forecast the annual precipitation. Bokde et al. [30] used

EEMD-PSF model to predict wind speed series. Experiments showed that EEMD decomposition method can reduce the influence of trend, seasonal and irregular components to a certain extent. The variational mode decomposition (VMD) is a new mode decomposition method [31]. Compared with the traditional decomposition methods, VMD has stronger decomposition ability and anti noise interference ability, and its operation speed is faster [32], [33]. Wu and Lin [34] combined VMD with prediction method to forecast AQI. Li et al. [35] proposed a forecasting model of sunspot number based on the combination of VMD and BP neural network. At present, most scholars use the combination of the mode decomposition and the single forecasting model to predict the rainfall data. Compared with the single prediction method, the prediction accuracy has been greatly improved. But each model has its advantage and disadvantage, and the prediction ability is also different. So it can not achieve the best prediction effect. Hybrid prediction model combines different single prediction methods and integrates the advantages of multiple models to improve the performance of the prediction model [36]. The key problem of the hybrid prediction model is to determine the appropriate weight coefficient to improve the forecasting precision of the hybrid forecasting model. IOWA operator [37] weights the fitting precision of each time point in the sample interval by each single prediction method. It solves the problem that the weight coefficient of the same single prediction method in the traditional combination weight calculation remains unchanged at each time point. It is an effective method to calculate the weight of combined forecasting model [38]. The hybrid forecasting model based on the idea of decomposition integration and weight distribution can combine the advantage of each model and improve the forecasting precision effectively. However, due to the strong randomness of some time points, the prediction model can't achieve good prediction results. Huang et al. [39] adopted the error correction strategy based on the EEMD-LSTM prediction model to further improve the accuracy of wind speed prediction. If the idea of error correction can be introduced into the prediction of precipitation signal, the prediction accuracy of precipitation will be further improved. By analyzing that the error sequence is stable, this paper proposes the use of ARMA to fit the error sequence of the hybrid model to achieve error correction.

In recent years, some forecasting methods for rainfall series have been put forward. For example, Li et al. [40] adopted VMD-ELM method. Hu et al. [29] established EEMD-GRNN. Farajzadeh and Alizadeh [41] established ARIMAX and LSSVM model based on wavelet transform. However, with the development of forecasting technology, there are still some limitations: (i) each prediction model has its advantage and disadvantage, and only a single prediction model is established without considering the stability of the model; (ii) the prediction ability at each time point is also different; (iii) the influence of rainfall error series on the prediction results is not considered; (iiii) it is difficult to select model parameters of LSSVM. In order to obtain

higher precision and certainty of prediction model, a new hybrid prediction model based on the weight combination prediction model and error correction strategy proposed in this paper. The model consists of four modules: data preprocessing module, improved butterfly optimization algorithm (IBOA) module, weight combination module, error correction module. Firstly, the original precipitation data is decomposed by VMD. And then the LSSVM model optimized by IBOA and Volterra model is established for each mode components. The prediction results of VMD-IBOA-LSSVM and VMD-Volterra are combined by IOWA operator to get the prediction values of the combined prediction model. Third, the ARMA model is established to forecast the error sequence of the combined prediction results. Finally, the ultimate prediction results are acquired by summing the prediction results of the error series and combined model. The final prediction results are verified by the monthly mean precipitation of the Yan'an and Xianyang stations. The prediction model proposed in this paper provides a new idea for solving the instability of hydrological forecasting model, and is an effective precipitation forecasting model.

II. BASIC THEORY

A. VARIATIONAL MODE DECOMPOSITION

It is assumed that each subsequence has different frequency center value and bandwidth, so that the total bandwidth of each component is minimum [42]. Hilbert transform is applied to $\{m_k\}$ to acquire the single side spectrum of the mode function. Then, the frequency spectrum of each component is modulated to the corresponding fundamental band based on the center frequency $e^{-jw_k t}$ of the analytical signal of each mode component. The norm L_2 of the demodulated signal is calculated and the bandwidth of the component is estimated. The constrained variational problems are as follows:

$$\begin{cases} \min_{\{m_k\},\{\omega_k\}} \left\{ \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) m_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ s.t. \sum_{k=1}^{K} m_k(t) = x(t) \end{cases} \quad (1)$$

where, $\{m_k\} = \{m_1, m_2, \cdots, m_k\}$, $\{w_k\} = \{w_1, w_2, \cdots, w_k\}$; K is the number of IMF decompositions; δ (t) is the Dirac distribution.

α and θ (t) is the quadratic penalty factor and Lagrange multiplier respectivily. They are introduced to solve the variational problem, which transforms the constrained problem into the unconstrained problem. $\{m_k\}$ and $\{\omega_k\}$ is the set of sub signals and their center frequencies, f (t) is the original signal. The extended Lagrangian function is as follows:

$$L(\{m_k\}, \{\omega_k\}, \theta)$$
$$= C \sum_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) m_k(t) \right] e^{-j\omega_k t} \right\|_2^2$$
$$+ \left\| f(t) - \sum_{k=1}^{K} m_k(t) \right\|_2^2 + \left\langle \theta(t), f(t) - \sum_{k=1}^{K} m_k(t) \right\rangle \quad (2)$$

VMD uses alternating direction multiplier method to solve the above problems. By alternating updating $\hat{m}_k^{n+1}$, ω_k^{n+1} and $\hat{\theta}^{n+1}$ (n representing the number of iterations), VMD finds the saddle point of the extended Lagrangian expression. The specific formula is as follows:

$$\hat{m}_k^{n+1}(\omega) = \frac{\hat{f}(\omega) - \sum_{k=1}^{K} \hat{m}_k(\omega) + \frac{\hat{\theta}(\omega)}{2}}{1 + 2C(\omega - \omega_k)^2} \quad (3)$$

$$\omega_k^{n+1} = \frac{\int_0^{\infty} \omega |\hat{m}_k(\omega)|^2 d\omega}{\int_0^{\infty} |\hat{m}_k(\omega)|^2 d\omega} \quad (4)$$

$$\hat{\theta}^{n+1}(\omega) = \hat{\theta}^n(\omega) + \tau \left[\hat{f}(\omega) - \sum_{k=1}^{K} \hat{m}_k^{n+1}(\omega) \right] \quad (5)$$

When e > 0, VMD stops iteration.

$$\sum_{k=1}^{K} \frac{\left\| \hat{m}_k^{n+1} - \hat{m}_k^n \right\|_2^2}{\left\| \hat{m}_k^n \right\|_2^2} < e \quad (6)$$

B. BUTTERFLY OPTIMIZATION ALGORITHM AND ITS IMPROVEMENT

1) BUTTERFLY OPTIMIZATION ALGORITHM

BOA is a new intelligent optimization algorithm proposed by Arora and Singh [21] in 2018. This optimization algorithm is generated by imitating the butterfly's foraging and courtship behavior. Each butterfly can produce fragrance with different intensity, and the intensity of the fragrance is related to the adaptability of a butterfly to seek optimization. The butterfly will move towards the direction with the strongest fragrance, which is called global search in the algorithm. This stage is called the global search in the algorithm. In another case, the butterfly will move randomly when it cannot perceive the fragrance from its surroundings. This stage is a local search stage.

When the butterflies are looking for food, all the butterflies will send out fragrance to attract each other. The intensity of the fragrance is calculated as follows:

$$A = cI^a \quad (7)$$

where, A is expressed according to the physical intensity, I is the stimulation intensity related to the optimization adaptability, a is 0.1, C is the sensory factor, usually 0.01.

According to Equation (7), the algorithm enters the stage of global search and local search. In the global search, the iterative formula is shown as follows:

$$c_i^{t+1} = c_i^t + (r_1^2 \times g^* - c_i^t) \times A_i \quad (8)$$

where c_i^t is the position vector of the i-th butterfly in the t-th iteration, r_1 is the random number between [0-1], and A_i is the fragrance of the i-th butterfly. The local position update formula is as follows:

$$c_i^{t+1} = c_i^t + (r^2 \times c_j^t - c_k^t) \times A_i \quad (9)$$

where c_j^t and c_k^t is the position vectors of the k-th butterfly and j-th butterfly, and r is the random number between [0-1], indicating local random walk. Global search and local search are transformed by switching probability. In this paper, the random number and switching probability are used to determine the global search or local search. The switching probability is set to 0.8 following the reference [21].

2) IMPROVED BUTTERFLY OPTIMIZATION ALGORITHM

Because of the low accuracy of the butterfly optimization algorithm, it is easy to stuck at local optimization and slow convergence. In this paper, nonlinear time-varying adaptive weights are introduced to improve local search and global search. Besides, Cauchy disturbance is added to mutation the butterfly position information of global search, which improves the performance of the BOA. In the basic BOA, a weight factor varying with the number of iterations is introduced to control. Influenced by reference [23], the nonlinear time-varying adaptive weight is used to update the butterfly position.

$$\begin{cases} w_1 = \frac{1}{2} \cdot \left[1 + \cos\left(\frac{\pi t}{Max\_iter}\right)\right]^{\frac{1}{k}}, & t \le \frac{Max\_iter}{2} \\ w_2 = \frac{1}{2} \cdot \left[1 - \cos\left(\frac{\pi t}{Max\_iter}\right)\right]^{\frac{1}{k}}, & t > \frac{Max\_iter}{2} \end{cases} \quad (10)$$

where, Max_iter represents the maximum iteration, t represents the current number of iterations, k is the adjustment coefficient of the weight factor, $k = 2$.

The nonlinear time-varying adaptive weight decreases slowly at the beginning, and the algorithm can maintain a good global exploration ability. After a certain number of iterations, the weight decreases rapidly, which makes the BOA get the optimal position more precisely in the local search. Therefore, the Equation (8) and (9) are changed as follows:

$$\begin{cases} c_i^{t+1} = w_1 \times c_i^t + (r_1^2 \times g^* - c_i^t) \times A_i, t \le \frac{Max\_iter}{2} \\ c_i^{t+1} = w_2 \times c_i^t + (r_1^2 \times c_j^t - c_k^t) \times A_i, t > \frac{Max\_iter}{2} \end{cases} \quad (11)$$

In this paper, the Cauchy variation is added to the global search. Using Cauchy mutation can generate more perturbation near the current variant individuals. Cauchy mutation is easier to make the BOA algorithm jump out of the local optimal value and improve the ability of global search. The probability density function of Cauchy distribution is as follows:

$$f(h, h_0, \gamma) = \frac{1}{\pi \gamma [1 + (\frac{h - h_0}{\gamma})^2]} = \frac{1}{\pi} [\frac{\gamma}{(h - h_0)^2 + \gamma^2}] \quad (12)$$

where, h_0 is the location parameter defining the peak value of the perturbation and γ is the scale parameter at half of the maximum value. When $h_0 = 0$, $\gamma = 1$, it is called the standard Cauchy distribution formula as follows:

$$f(x) = \frac{1}{\pi} \left(\frac{1}{x^2 + 1}\right) \quad (13)$$

In this paper, the Cauchy distribution random variable is used to generate function $Cauchy(0, 1) = tan[(\eta - 0.5) \cdot \pi]$, where η is a random variable on [0,1]. So according to Cauchy variation, the current optimal solution formula is improved as follows:

$$c_{best} = c_i^{t+1} + c_i^{t+1} \cdot Cauchy(0, 1) \quad (14)$$

C. LEAST SQUARES SUPPORT VECTOR MACHINE

The least squares support vector machine (LSSVM) is an improved method of SVM. The quadratic programming problem is transformed into the solution of linear equations, which improves the convergence speed of the algorithm [42]. Compared with SVM, LSSVM keeps the characteristics of structural risk minimization and small samples, which greatly reduces the computational complexity. The principle is as follows:

According to the principle of minimizing the structural risk of Vapnick, the objective function of LSSVM can be expressed as follows:

$$\begin{cases} \min J(\omega^T \xi) = \frac{1}{2} \omega^T \omega + C \sum_{i=1}^{l} \xi^2, & i = 1, 2, \cdots l \\ s.t. \ y_i = \omega^T \varphi(x_i) + b + \xi_i \end{cases} \quad (15)$$

where, $\varphi(\cdot)$ represents the nonlinear mapping from the original space to the high-dimensional space, ω represents the weight vector, b represents the offset, ξ represents the relaxation factor, C is the penalty parameter. Equation (15) can be solved by Lagrange function and Karush-Kuhn-Tucker (KKT) condition. The results of the Lagrange function are as follows:

$$L(\omega, b, \xi, a) = J(\omega^T \xi) - \sum_{i=1}^{l} a_i[\omega^T \varphi(x_i) + b + \xi - y_i] \quad (16)$$

According to KKT condition, calculate $\frac{\partial L}{\partial w} = 0$, $\frac{\partial L}{\partial c} = 0$, $\frac{\partial L}{\partial e_i} = 0$, $\frac{\partial L}{\partial \alpha_i} = 0$, and get the following linear equations:

$$\begin{bmatrix} 0 & Q^T \\ Q & K + C^{-1}I \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ Y \end{bmatrix} \quad (17)$$

where, a_i is the Lagrange multiplier, I is the unit matrix, $Q = [1, \cdots, 1]$.

The kernel function satisfying Mercer condition is $K(x_i, x_j) = \varphi(x_i) * \varphi(x_j)$, and the radial basis function is used in this paper. The LSSVM function obtained by the least square method is Equation (18):

$$f(x) = \sum_{i=1}^{n} \alpha_i K(x, x_i) + b \quad (18)$$

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \quad (19)$$

where, σ is the width of the kernel function.

D. VOLTERRA MODEL

The adaptive volterra algorithm can automatically adjust its parameters according to the input signal [43], [44]. $X(n) = [x(n), x(n-1), \ldots, x(n-N+1)]$ is the input of the system, and $\hat{x}(n+1)$ is the output of the system as follows:

$$\hat{x}(n+1) = h_0 + \sum_{i_1=0}^{M-1} h_k(i_1)x(n-m_1)$$
$$+ \sum_{i_1,i_2=0}^{M-1} h_k(i_1, i_2)x(n-m_1)x(n-m_2) \quad (20)$$

where, $h_k = (i_1, i_2, \cdots, i_k)$ is the Volterra kernel of k-order, M is the length of the filter, m is the best embedding dimension. According to Equation (21), (22), Equation (20) can be rewritten as shown in Equation (23).

$$U_{(n)} = [1, x(n), x(n-1), \ldots, x(n-m+1),$$
$$x^2(n), x(n)x(n-1), \ldots, x^2(n-m+1)]^T \quad (21)$$
$$H_{(n)} = [h_0, h_1(0), h_1(1), \ldots, h_1(m-1), h_2(0, 0), h_2(0, 1),$$
$$\ldots, h_2(m-1, m-1)]^T \quad (22)$$
$$\hat{x}(n+1) = H^T(n)U(n), \quad (23)$$

where, $U(n)$ and $H(n)$ are the input vectors and filter coefficients of the filter respectively. By using the least mean square adaptive algorithm to solve Equation (23), the Volterra model of time series can be obtained and the nonlinear approximation of the model can be realized.

E. ARMA MODEL

The autoregressive moving average model (ARMA) is created by Box and Jenkins. It is used to deal with the random event sequence. The model parameters are few and the application is simple [45]. There are three basic types of ARMA: AR model, MA model and ARMA model. Equation (24) is the basic structure:

$$X_t = \xi_1 X_{t-1} + \xi_2 X_{t-2} + \cdots + \xi_p X_{t-p} + \delta + h_t$$
$$+ \theta_1 h_{t-1} + \theta_2 h_{t-2} + \cdots + \theta_q h_{t-q} \quad (24)$$

where, $\xi_1, \xi_2, \cdots, \xi_p$ is the autoregressive coefficient, $\theta_1, \theta_2, \cdots, \theta_p$ is the sliding average coefficient. In Equation (24), p and q are the order of coefficients, $h_t, h_{t-1}, \cdots, h_{t-p}$ are independent white noise sequences, which are recorded as ARMA (p, q). If the original sequence is non-stationary and needs to be stable after the d-order difference, the original sequence can be expressed as ARIMA (p, d, q) sequence.

F. IOWA OPERATOR

IOWA operator is called an induced ordered weighted average operator. The observation value of the series is $\{x_t, t = 1, 2, \cdots, N\}$, and the prediction accuracy of m single prediction methods is:

$$a_{it} = \begin{cases} 1 - |(x_t - x_{it})/x_t|, & |(x_t - x_{it})/x_t| < 1 \\ 0, & |(x_t - x_{it})/x_t| \geq 1 \end{cases} \quad (25)$$

where, a_{it} represents the prediction accuracy of the i-th prediction method at t time. $i = 1, 2, \cdots, m$, $t = 1, 2, \cdots, N$, $a_{it} \in [0, 1]$. x_t is the actual value, x_{it} is the predicted value. The a_{it} is regarded as the induced value of the x_{it}, then the forecasting precision at the t-th time of the m-th single prediction methods and their predictive value in the sample interval constitute m two-dimensional arrays [46]. $W = (w_1, w_2, \cdots, w_n)^T$ is the weight vector, where $\sum_{i=1}^{m} w_i = 1, w_i \geq 0, i = 1, 2, \cdots, m$. Arrange $a_{1t}, a_{2t}, \cdots, a_{mt}$ from large to small, and mark the i-th forecasting precision as $a - index(it)$. The combined prediction value of the series is:

$$f_L[(a_{1t}, x_{1t}), (a_{2t}, x_{2t}), \cdots, (a_{mt}, x_{mt})] = \sum_{i=1}^{m} w_i x_{a-index(it)} \quad (26)$$

It can be seen from Equation (26) that the weighting coefficient of the combined forecasting model is not related to the single forecasting method, but closely related to the forecasting precision of the single prediction model at each time. Therefore, the combination prediction model is expressed as the following optimization model.

$$\min S(L) = \sum_{i=1}^{m} \sum_{i=1}^{m} w_i w_j \left(\sum_{i=1}^{N} e_{a-index} e_{a-index(jt)} \right)$$

$$s.t. \begin{cases} \sum_{i=1}^{m} w_i = 1 \\ w_i \geq 0, \quad i = 1, 2, \cdots, m \end{cases} \quad (27)$$

Each single prediction method is weighted by the order of fitting precision of each time point in the sample interval, and a new combined prediction method is established. The combination forecast can reduce the forecast error effectively by the weighted average of the forecast value at each time point of a single forecast.

III. THE PREDICTION MODEL FOR MONTHLY MEAN PRECIPITATION

A. THE PROPOSED HYBRID MODEL

To improve the precision of precipitation prediction, a new hybrid precipitation prediction model is proposed by combining VMD-IBOA-LSSVM, Volterra model and ARMA. Fig. 1 is prediction the model block diagram.

Step 1: VMD process

VMD decomposes precipitation signal into modes. Each mode have a central frequency, for example, IMF1, IMF2,..., IMFn. The function of this technique is to reduce the non-stationarity of the sequence and improve the accuracy of prediction.

Step 2: Weight combination forecast

The IBOA-LSSVM model and Volterra model are established for each component. The final single model prediction result is obtained by superposing the prediction results of each component. Finally, this paper uses the IOWA operator to combine the prediction results of the VMD-IBOA-LSSVM and VMD-Volterra model.
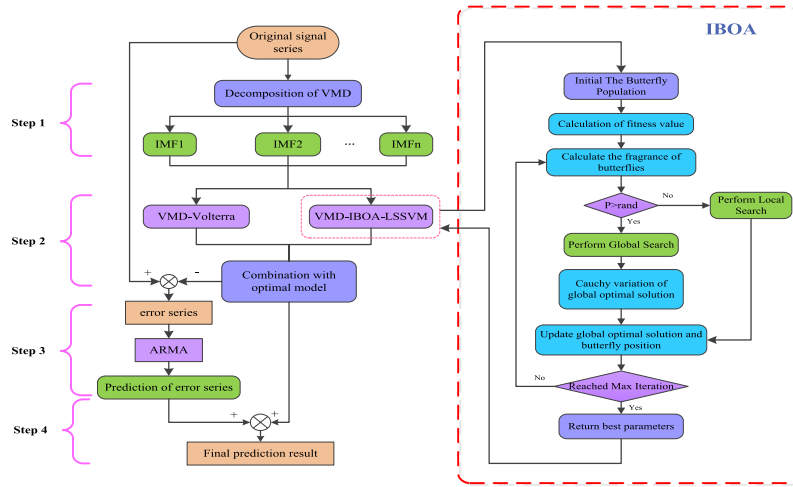
FIGURE 1. The framework of the proposed hybrid model.

Step 3: Error correction of the hybrid forecasting model

Because the prediction result of the model is the estimation of the overall development trend of the data, the prediction value will be too high or too low. To solve this problem, the ARMA model is used to predict and fit the prediction error sequence in step 2, and the corrected error value is obtained.

Step 4: Integration process

Add the prediction error result corrected by step 3 and the prediction value of step 2 to produce the final prediction value.

B. EVALUATION INDEX OF PREDICTION PERFORMANCE

In order to evaluate the prediction performance of the model, three different prediction evaluation criteria will be used: (i) the index of accuracy evaluation; (ii) the index of directional evaluation; (iii) statistical test.

In terms of accuracy evaluation, three commonly used error evaluation methods are used to evaluate the model prediction accuracy of the model such as mean absolute error (MAE), root mean squared error ($RMSE$) and correlation coefficient (R). $\hat{x}(t)$ and $\hat{x}(t)'$ represent the predicted data and the predicted average data, respectively. $x(t)$ and $x(t)'$ represent the actual data and the actual average data, respectively. N is the number of samples.

$$MAE = \frac{1}{N} \sum_{t=1}^{N} \left| \hat{x}(t) - x(t) \right| \tag{28}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^{N} \left[\hat{x}(t) - x(t) \right]^2} \tag{29}$$

$$R = \frac{\sum_{t=1}^{N} (x(t) - x(t)')(\hat{x}(t) - \hat{x}(t)')}{\sqrt{\sum_{t=1}^{N} (x(t) - x(t)')^2} - \sqrt{\sum_{n=1}^{N} (\hat{x}(t) - \hat{x}(t)')^2}} \tag{30}$$

Based on the above three evaluation indexes, the improvement rate (IR) [26] of the model can be established
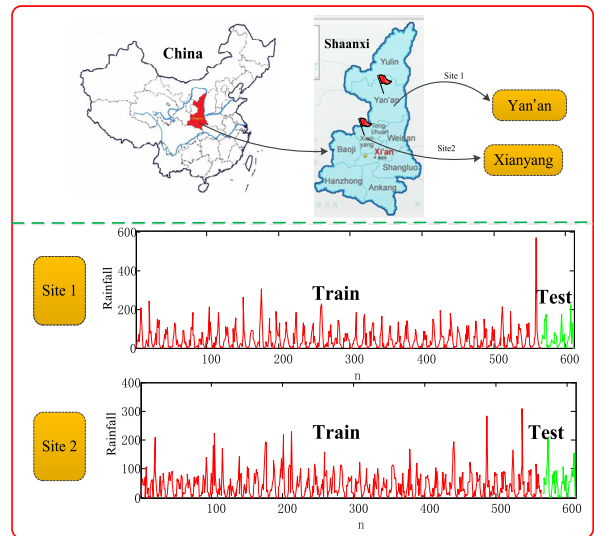

FIGURE 2. The precipitation data series.

in Equation (31) and (32). IR_{MAPE} and IR_{RMSE} are the improvement rates of model 1 to benchmark model 2 on the evaluation criteria RMSE and MAPE, respectively.

$$IR_{RMSE} = \left| \frac{RMSE_1 - RMSE_2}{RMSE_1} \right| \times 100\% \tag{31}$$

$$IR_{MAPE} = \left| \frac{MAPE_1 - MAPE_2}{MAPE_1} \right| \times 100\% \tag{32}$$

In addition to horizontal accuracy, directional accuracy of prediction is also very important [47].

$$D_{stat} = \frac{1}{N} \sum_{t=1}^{N} a_t \times 100\% \tag{33}$$

If $\left(\hat{x}_{t+1} - x_t \right)(x_{t+1} - x_t) \geq 0$, then $a_t = 1$, indicating that the model's direction prediction in $t + 1$ period is correct. Otherwise $a_t = 0$, indicating that the model's direction prediction in $t + 1$ period is wrong.

TABLE 1. The central frequency corresponding to different *K* values in site 1.

K	central frequency/HZ					
2	0.0841	0.296				
3	0.000868	0.092	0.469			
4	0.000547	0.086	0.271	0.474		
5	0.000452	0.085	0.180	0.296	0.421	
6	0.000407	0.084	0.170	0.280	0.387	0.479
7	0.000395	0.083	0.165	0.248	0.335	0.412	0.480
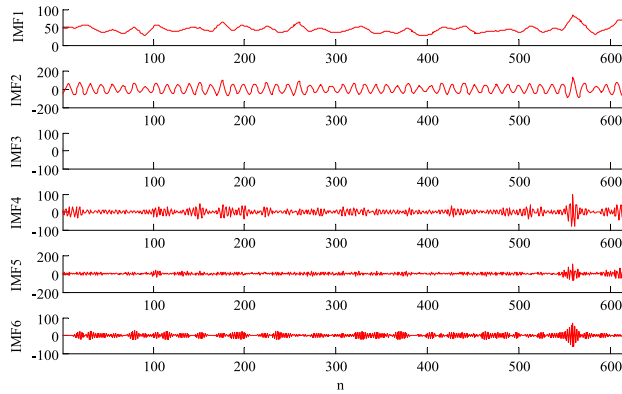
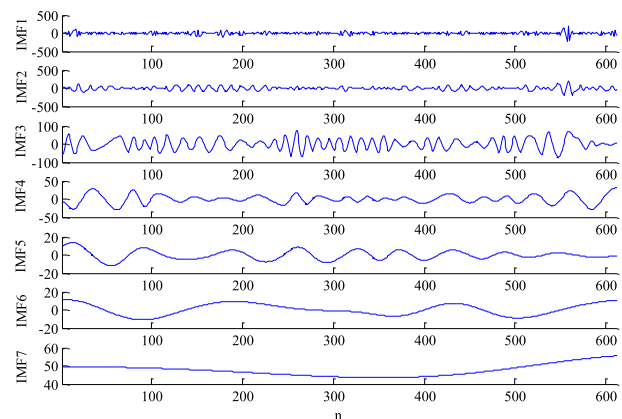**FIGURE 3.** Decomposition results of VMD in site 1.

FIGURE 4. Decomposition results of EMD in site 1.

In order to test the difference in prediction performance in a statistical sense, a statistical test is further introduced to determine whether the prediction accuracy of model 1 is significantly better than model 2. Therefore, this paper introduces the Diebold-Mariano (DM) statistics [48]. The target model is the proposed model. The original assumption of DM is that the prediction accuracy of target model A is not higher than that of benchmark model B, that is, the prediction error of model A($e_{A,t} = x_t - \hat{x}_{A,t}$) is greater than or equal to the prediction error of model B($e_{B,t} = x_t - \hat{x}_{B,t}$).

$$S = \frac{\bar{g}}{\left(\widehat{V}_{\bar{g}}/M\right)^{1/2}} \sim N(0, 1) \tag{34}$$

where, $\bar{g} = \frac{1}{M}\sum_{t=1}^{M} g_t$ is the average loss function value, $\widehat{V}_{\bar{g}} = \gamma_0 + 2\sum_{l=1}^{\infty} \gamma_l$. By testing the S statistic and P-value,
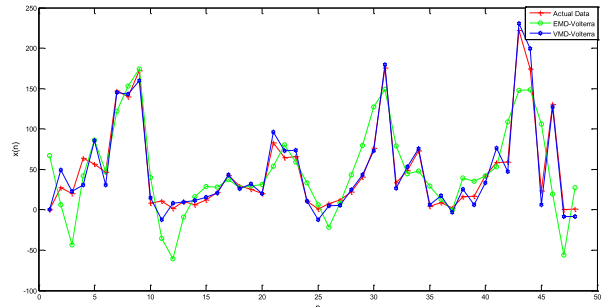

FIGURE 5. The prediction results of EMD-Volterra and VMD-Volterra.

TABLE 2. The error exponents of EMD-Volterra and VMD-Volterra.

Model	$RMSE$	MAE	R
EMD-Volterra	36.1901	28.8228	0.7784
VMD-Volterra	11.3899	8.3870	0.9810

TABLE 3. Test function.

Equation	Variable range	Global optimum	Function expression
Sphere	[-5.12,5.12]	0	$f_1 = \sum_{i=1}^{n} x_i^2$
Quartic	[-1.28,1.28]	0	$f_2 = \sum_{i=1}^{n} i x_i^4 + rand(0,1)$
Rastrigin	[-5.12,5.12]	0	$f_3 = \sum_{i=1}^{n}(x_i^2 - 10\cos(2\pi x_i) + 10)$
Griewangk	[-600,600]	0	$f_4 = \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n}\cos(\frac{x_i}{\sqrt{i}}) + 1$

TABLE 4. The result comparison of four test functions by using PSO, DE, BOA, and IBOA.

Fun	index	PSO	DE	BOA	IBOA
F1	MV	65.1749	88.9601	1.6234e-09	0
	SR(%)	90	100	100	100
F2	MV	0.0966	0.461	7.496e-04	4.428e-04
	SR(%)	50	50	58	67
F3	MV	1.158e-04	8.126e-05	1.280e-11	2.703e-35
	SR(%)	36	55	64	90
F4	MV	0.0172	0.0121	3.4270e-12	0
	SR(%)	50	60	100	100

the superiority of model A over benchmark model B can be effectively identified.

IV. MODEL PREDICTION AND RESULT ANALYSIS
A. STUDY AREA AND DATASET
In this paper, the monthly mean precipitation observation data of two stations in Shaanxi Province from 1967 to 2017 is used as the test data. The number of sample points in each data set is 612, and the data is from China Meteorological Data Network (http://data.cma.cn/). The data of site 1 comes from Yan'an City. The precipitation in Yan'an City is mainly concentrated in summer, with heavy precipitation and great intensity. From October to May of the next year, it only accounts for 29% of the total annual precipitation. The data of site 2 is from Xianyang city. The annual average precipitation
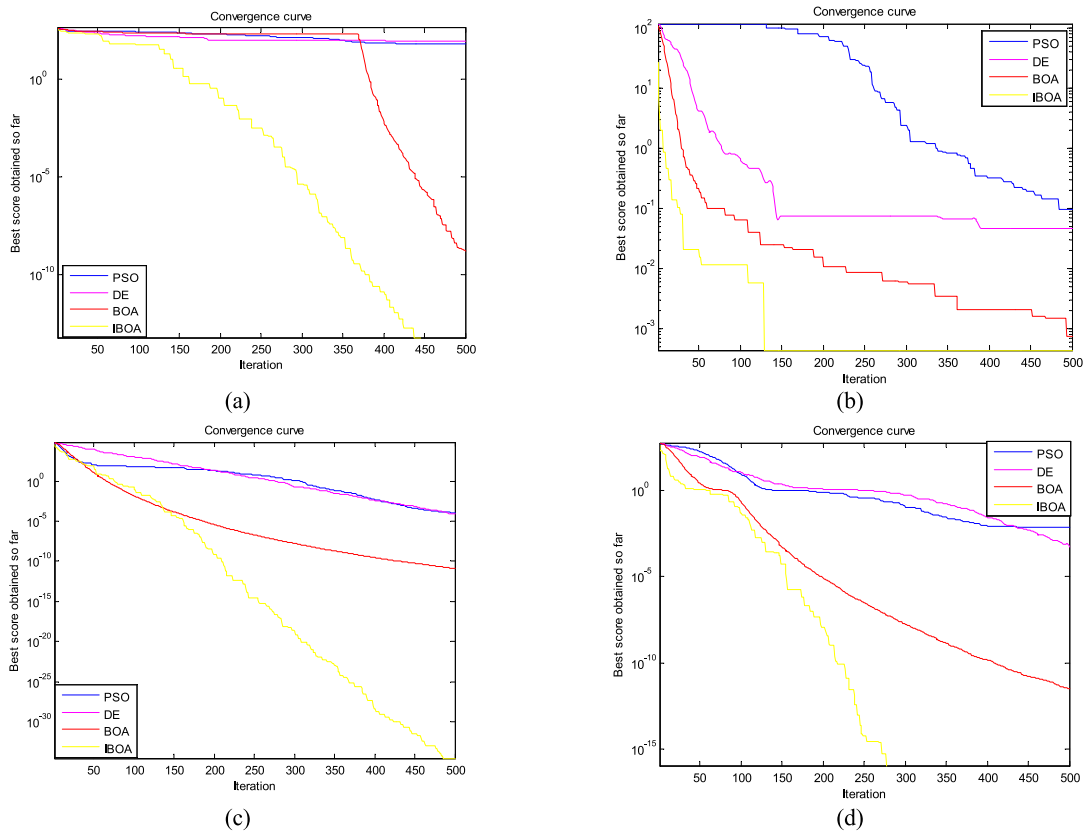
FIGURE 6. Iterative curve of optimization algorithm. (a) F1; (b) F2; (c) F3; (d) F4.

of Xianyang city is 500-600mm, and the precipitation is concentrated in summer.

The number of each samples data is 612. The first 564 observations are used as training sets, and the remaining 48 observations are used as test sets. The experimental environment is as follows: the program running platform is the PC of Windows 8 system, the processor is Intel Core i7, the main frequency is 3.6GHz, ROM storage is 4GB, memory is 32GB, and the experimental software is MATLAB R2014a. The time series diagram of the monthly average precipitation of the two sites is shown in Fig. 2.

B. RESULTS OF VARIATIONAL MODE DECOMPOSITION
The number of modes K of the VMD needs to be set manually. When the value of K is too small, the sequence cannot be completely decomposed. When the value of K is too large, the problem of over decomposition will occur. At the same time, too many mode numbers will increase the complexity of operation. It is the most common and intuitive method to determine K [49] by observing the center frequency. The mode center frequency corresponding to each K value is checked. If the center frequency value is close, it is regarded as over decomposition, and the optimal decomposition layer is K-1. Take the precipitation data of the site 1 as an example to calculate the center frequency values corresponding to different decomposition numbers, as shown in Table 1.

Table 1 shows that when K is 7, the center frequencies of IMF6 and IMF7 are similar. It is considered that over

decomposition occurs when $K = 7$. The number of decomposition $K = 6$. The VMD decomposition results of precipitation are shown in Fig. 3.

VMD can overcome the phenomenon of mode mixing and end effect of EMD, and solve the problem that the components with similar frequency can not be separated. The results of EMD in Site 1 are shown in Fig. 4.

As shown in Fig. 4, the IMF decomposed by EMD has different characteristics. Fig. 4 shows that the amplitude difference of each component is large, and the frequency of IMF 1 is the highest, which indicates that EMD cannot effectively separate the components with similar frequency. Besides, Fig. 3 and Fig. 4 show that VMD decomposes the high-frequency part more thoroughly than EMD, and the amplitude fluctuation of the high-frequency component is smaller. Fig. 5 and Table 2 show the prediction results of precipitation in site 1 by EMD and VMD respectively applied to the Volterra model. Fig. 5 shows that VMD-Volterra has higher prediction accuracy, which proves that VMD decomposition has higher prediction accuracy. In addition, the RMSE index of VMD-Volterra and EMD-Volterra are substituted into the Equation (31), which shows that the model performance after VMD decomposition increases by 68.52% compared with EMD.

C. THE IMPROVED OPTIMIZATION BUTTERFLY ALGORITHM TEST
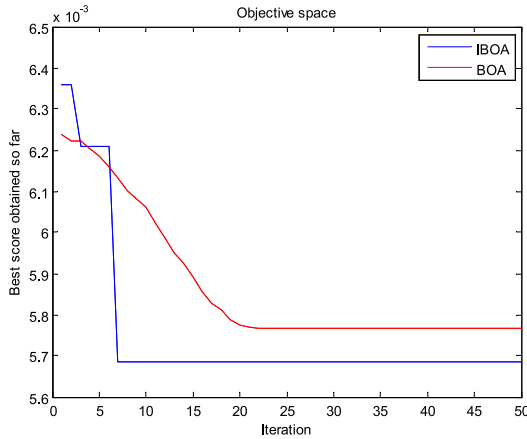Four different test functions are used to test the IBOA. Table 3 shows the function expressions of the four

FIGURE 7. Iterative curve of IBOA and BOA for LSSVM model optimization.

TABLE 5. The error exponents of BOA-LSSVM and IBOA-LSSVM.

Model	$RMSE$	MAE	(Gam, Sig2)
BOA-LSSVM	43.1343	32.0776	(81.6082,11.6007)
IBOA-LSSVM	42.8237	31.6532	(800, 24.0070)

test functions, where F1, F2 are unimodal functions, F3, F4 are multimodal functions. The test results are compared with the other three optimization algorithms, including particle swarm optimization (PSO), differential evolution algorithm (DE), butterfly optimization algorithm (BOA). In order to make the experiment objective and fair, the population size of the test function selection optimization algorithm is 30, and the number of iterations is 500.

The test function sphere and quartic selected in this paper are single-mode test functions, which are used to test the convergence performance of the algorithm. Rastrigin and Griewangk are multimodal test functions. Generally, there is a lot of local extremum to test the capacity of global mining. In Table 4, where MV is the minimum value. It can be seen that the optimized effect of the IBOA on the four test functions is closest to the global optimal value, among which the optimized effect of F3 and F4 can reach 100%, indicating that the optimization ability of IBOA is better than that of other optimization algorithms. In addition, Fig. 6 is the iterative curves of benchmark algorithms on the test function. From Fig. 6, we can find that the optimization ability of IBOA is better than that of BOA, and the convergence speed is faster.

The global optimal solution convergence success rate is an important indicator to measure the performance of the algorithm [50]. The success rate is defined as $SR = \frac{N_b}{N} \times 100\%$, where SR is the convergence success rate, N is the total number of trials, and N_b is the number of tests converging to the optimal solution. Four algorithms are used to optimize the four test functions, and each algorithm is repeated for 20 times. The optimal convergence value and success rate of the algorithm are shown in Table 4.

Through the analysis, it can be seen that compared with the other three optimization algorithms, the solution obtained by IBOA is the closest to the optimal solution, and the
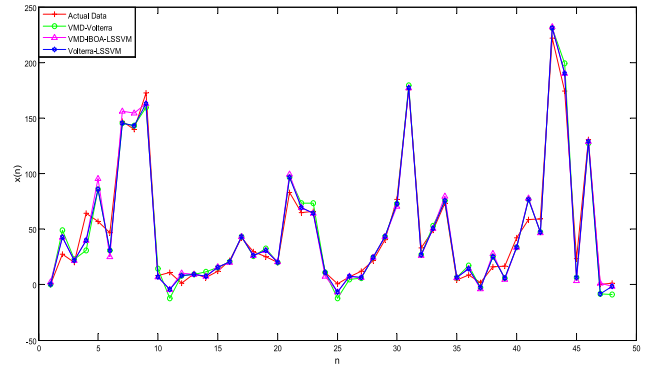

FIGURE 8. The Forecast results of VMD-Volterra-LSSVM.

TABLE 6. The error-index of VMD-Volterra-LSSVM.

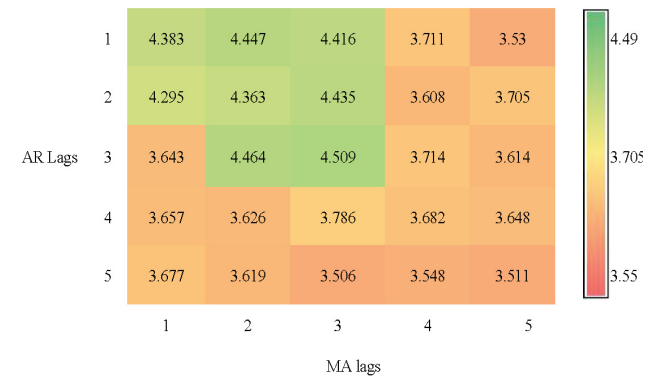| Model	$RMSE$	MAE	R
VMD-Volterra	11.3899	8.3870	0.9810
VMD-IBOA-LSSVM	10.9941	7.8976	0.9827
VMD-Volterra-LSSVM	9.3909	6.6782	0.9867

FIGURE 9. AIC criterion order distribution.

convergence success rate is the highest. To further verify the practicability of the IBOA, the actual precipitation data in the site 1 is used to verify the optimization effect of IBOA on LSSVM model parameters. The population size of IBOA and BOA is 10, and the number of iterations is 50. Fig. 7 shows that the convergence precision of the IBOA is higher than that of BOA, and the convergence speed is faster. Besides, Table 5 is the optimization results of IBOA and BOA for LSSVM. From Table 5, we can see that the IBOA-LSSVM model is better than BOA-LSSVM in precipitation prediction. The results show that the IBOA can effectively find the optimal regularization parameter gam and kernel parameter sig2 for LSSVM.

D. EXPERIMENT I: THE ANALYSIS OF SITE 1 PRECIPITATION PREDICTION MODEL RESULTS

Because the information source of single model prediction is not extensive and the model setting form is arbitrary, the prediction ability of the single model is different at different times. In this paper, the prediction results of LSSVM and Volterra model are combined by IOWA operator, and each single prediction method is weighted by the order of fitting
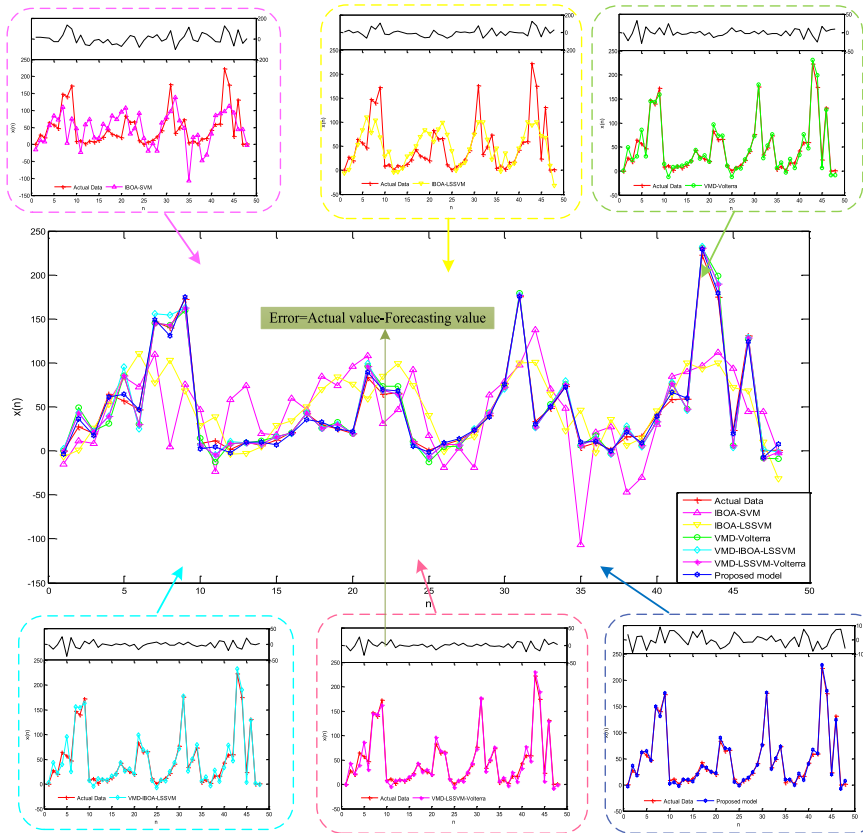
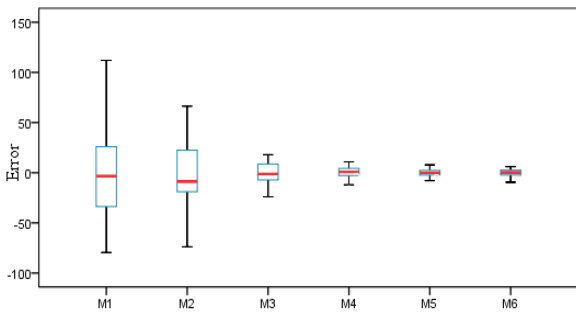**FIGURE 10.** Prediction results of various models in Site 1.

FIGURE 11. Prediction error box diagram of each model.

TABLE 7. Error analysis and comparison of forecasting model.

	Model	$RMSE$	MAE	R	D_{stat}
M1	IBOA-SVM	53.8634	42.3874	0.4578	0.5625
M2	IBOA-LSSVM	42.8280	31.8696	0.6207	0.6458
M3	VMD-Volterra	11.3899	8.3870	0.9810	0.8125
M4	VMD-IBOA-LSSVM	10.9941	7.8976	0.9827	0.7917
M5	VMD-Volterra-LSSVM	9.3909	6.6782	0.9867	0.8125
M6	Proposed model	4.5580	3.7762	0.9967	0.8958

precision of each time point in the sample interval, and a new combined prediction method is established based on the sum of error squares. According to Equation (27), the IOWA

TABLE 8. Performance improvements by proposed model for Site 1.

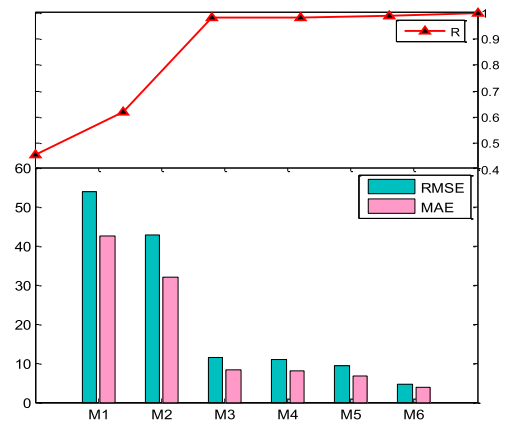| Performance metrics	M6 Vs M1	M6 Vs M2	M6 Vs M3	M6 Vs M4	M6 Vs M5
P_{RMSE} (%)	91.5378	89.3574	59.9820	58.5413	51.4636
P_{MAE} (%)	91.0912	88.1510	54.9779	52.1854	43.54548

FIGURE 12. Prediction error index of each model.

combined prediction model is as follows:

$$\min S = \min(4233.06w_1^2 + 7795.73w_2^2 + 10931w_1w_2)$$

$$s.t. \begin{cases} w_1 + w_2 = 1 \\ w_i \geq 0, \quad i = 1, 2 \end{cases} \quad (35)$$
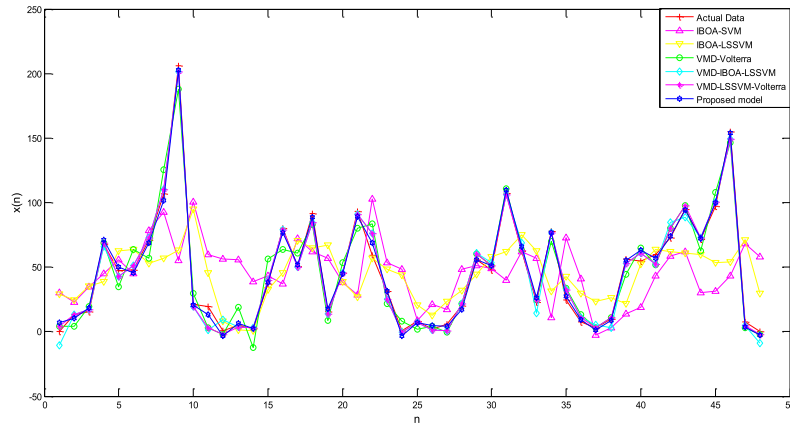
FIGURE 13. The prediction results of various models in site 2.

TABLE 9. DM test results across different models of site1.

Target Model		Benchmark Model				
		M1	M2	M3	M4	M5
M6	DM statistic	4.7565*	4.1068*	3.3816*	2.9439*	2.9896*
	P-value	0.0000	0.0000	0.0000	0.0016	0.0014

*A confidence interval of 1%.

TABLE 10. Error analysis and comparison of forecasting model.

	Model	$RMSE$	MAE	R	D_{stat}
M1	IBOA-SVM	45.4510	34.6474	0.1615	0.5833
M2	IBOA-LSSVM	38.7078	28.0887	0.4442	0.7083
M3	VMD-Volterra	10.1456	8.4334	0.9721	0.9167
M4	VMD-IBOA-LSSVM	5.9928	4.4240	0.9906	0.9167
M5	VMD-Volterra-LSSVM	4.8155	3.3130	0.9938	0.9167
M6	Proposed model	3.3187	2.6461	0.9970	0.9375

The above formula is solved by Matlab optimization toolbox to get $w_1 = 1, w_2 = 0$, and the prediction value of the hybrid prediction model can be obtained by substituting the weight into Equation (26). The prediction results of each point obtained by combining the LSSVM model and Volterra model with IOWA operator is shown in Fig. 8, and the error-index of the hybrid forecasting model is shown in Table 6.

Fig. 8 and Table 6 show that the VMD-Volterra-LSSVM model considers the forecasting precision of the Volterra model and LSSVM model at different time points. The single prediction value with high prediction accuracy is given a large weight. After the combination, the prediction performance of the model has been improved. However, there are still some points with large errors. To make the forecasting value of the model more accurate, the error correction is carried out for the prediction result of the combined model. The ARMA model is established to predict and fit the prediction error of VMD-Volterra-LSSVM. Finally, the modified prediction error is combined with the prediction value of VMD Volterra LSSVM to get the final prediction result. In the ARMA model, the order *p* and *q* of the model should be calculate firstly. In this paper, the distribution of the AIC criterion is drawn into a heat map, the vertical axis is the end of AR, and the horizontal axis is the order MA. Fig. 9 can intuitively select the model with the minimum AIC criterion, establish ARMA (5,3) model, and get the corrected error value. The modified prediction error results are combined with the combined model prediction values to produce the final prediction values as shown in Fig. 10.

To further measure the forecasting performance of the proposed model, the combined model is compared with the other five benchmark models, and the forecasting effect of

each model is quantitatively analyzed by MAE, RMSE, R and D_{stat}. The fitting curve of forecasting data and the original data of different prediction models are shown in Fig. 10, where the predicted data of the proposed model is represented by the blue curve, and the blue curve is relatively close to the red original data curve. In Fig. 10, the prediction error of each model is represented by black lines. By comparing the error curve, it can be seen that the error amplitude of the prediction model proposed in this paper is smaller. Corresponding to the prediction error box diagram in Fig. 11, the distribution of prediction error of each model can be seen more intuitively. The height of the box reflects the degree of data fluctuation to a certain extent, and the upper and lower edges represent the maximum and minimum values of the group of data. The flatter the box, the more concentrated the data, and the shorter the end line, the more concentrated the data. In Fig. 11, M6 is the model proposed in this paper, the forecasting error of the M6 is distributed around 0, and the change range is the smallest compared with other models. It shows that the combined forecasting model proposed in this paper has a good prediction effect.

We can see from the quantitative analysis results of forecasting results of each model in Table 7.

As an improved method of SVM, the LSSVM model can effectively improve the forecasting performance of the SVM model by optimizing the key parameters. The smaller the error index, the better the effect. According to Table 7, the RMSE of VMD-IBOA-LSSVM is 10.9941, but the RMSE value of IBOA-LSSVM model is 42.8280. VMD is used to decompose the complex data sequence into several
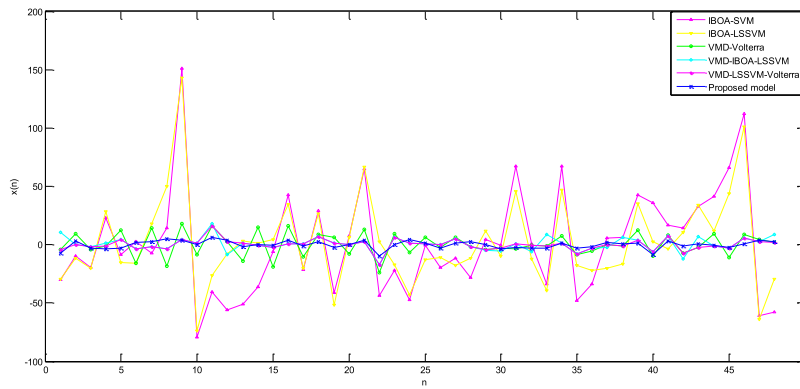
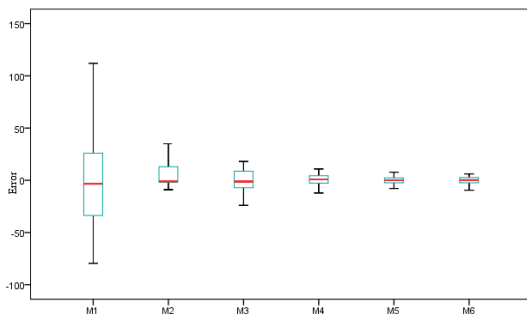**FIGURE 14.** The prediction error of various models in site 2.

FIGURE 15. Prediction error box diagram of each model.

FIGURE 16. Prediction error index of each model.

TABLE 11. Performance improvements by proposed model for site 2.

Performance metrics	M6 Vs M1	M6 Vs M2	M6 Vs M3	M6 Vs M4	M6 Vs M5
P_{RMSE} (%)	92.6982	91.4262	67.2892	44.6218	31.0829
P_{MAE} (%)	92.3627	90.5794	68.6235	40.1876	20.1297

TABLE 12. DM test results across different models of site2.

Target Model		Benchmark Model				
		M1	M2	M3	M4	M5
M6	DM statistic	3.7681*	3.0924*	5.5609*	3.0542*	1.8214**
	P-value	0.0000	0.0000	0.0000	0.0011	0.0343

*A confidence interval of 1%; **A confidence interval of 5%.

simple sequences, which can effectively reduce the forecasting difficulty and the forecasting error. Besides, the error-index of VMD-Volterra-LSSVM is smaller than that of VMD- Volterra and VMD-IBOA-LSSVM. Table 7 shows that the accuracy index and directional index of the proposed prediction model are optimal. Table 8 shows compared with other models, the performance of the proposed model is improved by at least 43%. In addition, Table 9 uses DM test for statistical proof. Through the DM test, general speaking, the proposed hybrid forecasting model outperforms other benchmark models at a 1% level of statistical significance, which proves that the model is an effective rainfall prediction model.

E. EXPERIMENT II: THE ANALYSIS OF SITE 2 PRECIPITATION PREDICTION MODEL RESULTS

Experiment I shows that the combination prediction model proposed in this paper has a better prediction performance. To illustrate the general applicability of this method in precipitation prediction, the precipitation data of another station in Shaanxi Province is used to test. There are 612 samples in total. The first 564 are trained and the last 48 are predicted.

The precipitation prediction results of site 2 are shown in Fig. 13, Table 10 quantitatively analyzes the error index of each prediction model, and the histogram of error index is shown in Fig. 16. When the numerical value of RMSE and MAE is higher, the prediction accuracy is better. The closer R is to 1, the better the prediction effect is. Fig. 14 and Fig. 15 show the prediction error curve and error box diagram of several models respectively. In Fig. 14, the final error
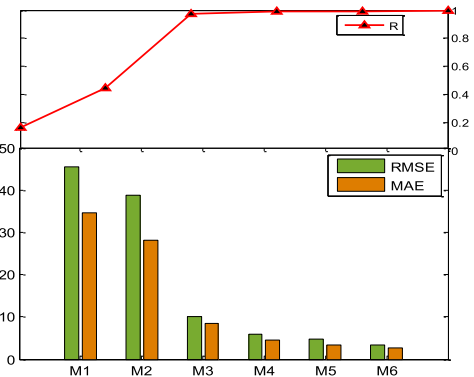
curve of the proposed model is represented by a blue line. It can be seen that the error fluctuation of the model proposed in this paper is the minimum. According to Table 10, the prediction indexes of M6 are 3.3187, 2.6461, 0.9970 and 0.9375 respectively, which are better than other models. From the improvement rate of the model, the prediction performance of the proposed model is greatly improved compared with other benchmark models. Through the analysis of site 2 and site 1, it shows that the model proposed in this paper has good general adaptability and stability, and can be applied in precipitation prediction research.

In addition, Table 12 uses DM test for statistical proof. Through the DM test, general speaking, the proposed hybrid forecasting model outperforms the VMD-Volterra-LSSVM

model at a 5% level of statistical significance, and outperforms other benchmark models at a 1% level of statistical significance, which proves that the model is an effective rainfall prediction model.

V. CONCLUSION

To improve the stability and accuracy of precipitation prediction, a novel combined prediction model for monthly mean precipitation with error correction strategy is proposed in this paper. The main conclusions are as follows:

(1) Compared with EMD, VMD algorithm decomposes the precipitation series more thoroughly and the prediction accuracy is higher. The simulation results show that the prediction model after VMD decomposition is enhanced in the overall RMSE index, which is increased by 68.52% compared with EMD. Therefore, VMD decomposition is more suitable for precipitation prediction, and the prediction accuracy of the model is effectively improved.

(2) Aiming at the problem of low convergence precision and easy to fall into the local optimum of the BOA algorithm, an improved BOA algorithm is proposed. Four test functions are compared with DE, PSO and BOA. The results in Section 4.3 show that the improved BOA algorithm can obtain the optimal value and the highest convergence success rate, indicating that IBOA has better global convergence than the benchmark algorithm.

(3) Because each prediction model has its advantage and disadvantage, this paper proposes using the IOWA operator to combine VMD-IBOA-LSSVM and VMD-Volterra. From the analysis of the prediction results of Experiment I and Experiment II, it can be seen that the combined model improves the prediction performance.

(4) By analyzing that the error sequence of VMD-IBOA-LSSVM-Volterra is stable, this paper uses ARMA with a simple model and fewer parameters to correct the error. The prediction results from the site1 data show that the prediction performance of the combined prediction model with error correction is improved by at least 43% compared with the other models.

(5) Compared with the method proposed in reference [40], this paper uses IBOA optimized LSSVM to predict the precipitation data in Xianyang and Yan'an areas. The experiment shows that the forecasting precision of VMD-IBOA-LSSVM is much higher than that of VMD-ELM proposed in reference [40]. Besides, in this paper the idea of combined prediction and the idea of error correction are introduced to improve the precipitation method VMD-IBOA-LSSVM, which is used to increase the performance of precipitation prediction model.

REFERENCES

[1] T. Jiang, B. D. Su, and H. Hartmann, "Temporal and spatial trends of precipitation and river flow in the Yangtze River Basin, 1961–2000," *Geomorphology*, vol. 85, nos. 3–4, pp. 143–154, 2007.

[2] X. C. Xu, X. Z. Zhang, E. F. Dai, and W. Song, "Research of trend variability of precipitation intensity and their contribution to precipitation in China from 1961 to 2010," *Geographical Res.*, vol. 33, no. 7, pp. 1335–1347, 2014.

[3] T. Jiang, B. D. Su, Y. J. Wang, Q. Zhang, N. X. Qin, and Y. F. Shi, "Trends of temperature, precipitation and runoff in the Yangtze river basin from 1961 to 2000," *Adv. Climate Change Res.*, vol. 1, no. 2, pp. 65–68, 2015.

[4] R. Maheswaran and R. Khosa, "A wavelet-based second order nonlinear model for forecasting monthly rainfall," *Water Resour. Manage.*, vol. 28, no. 15, pp. 5411–5431, Dec. 2014.

[5] G. Blöschl, M. F. P. Bierkens, A. Chambel, "Twenty-three unsolved problems in hydrology (UPH)—A community perspective," *Hydrol. Sci. J.*, vol. 64, no. 10, pp. 1141–1158, 2019.

[6] J. Qiu, Z. Shen, G. Wei, G. Wang, H. Xie, and G. Lv, "A systematic assessment of watershed-scale nonpoint source pollution during rainfall-runoff events in the Miyun reservoir watershed," *Environ. Sci. Pollut. Res.*, vol. 25, no. 1, pp. 1–18, 2017.

[7] K. W. Chau and C. L. Wu, "A hybrid model coupled with singular spectrum analysis for daily rainfall prediction," *J. Hydroinform.*, vol. 12, no. 4, pp. 458–473, Oct. 2010.

[8] G. Papacharalampous, H. Tyralis, and D. Koutsoyiannis, "One-step ahead forecasting of geophysical processes within a purely statistical framework," *Geosci. Lett.*, vol. 5, no. 12, pp. 1–19, Dec. 2018.

[9] G. Papacharalampous, H. Tyralis, and D. Koutsoyiannis, "Comparison of stochastic and machine learning methods for multi-step ahead forecasting of hydrological processes," *Stochastic Environ. Res. Risk Assessment*, vol. 33, no. 2, pp. 481–514, Feb. 2019.

[10] M. Dastorani, M. Mirzavand, M. T. Dastorani, and S. J. Sadatinejad, "Comparative study among different time series models applied to monthly rainfall forecasting in semi-arid climate condition," *Natural Hazards*, vol. 81, no. 3, pp. 1811–1827, Apr. 2016.

[11] N.-B. Chang, Y. J. Yang, S. Imen, and L. Mullon, "Multi-scale quantitative precipitation forecasting using nonlinear and nonstationary teleconnection signals and artificial neural network models," *J. Hydrol.*, vol. 548, pp. 305–321, May 2017.

[12] C. Cortes and V. Vapnik, "Support vector machine," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[13] R. Q. Yuan, X. T. Long, P. Wang, and X. F. Song, "Tempo-spatial variation and forecast of precipitation in Shanxi province," *J. Natural Resour.*, vol. 30, no. 4, pp. 651–663, 2015.

[14] J. L. Du, Y. Y. Liu, Y. A. Yu, and W. L. Yan, "A prediction of precipitation data based on support vector machine and particle swarm optimization (PSO-SVM) algorithms," *Algorithms*, vol. 10, no. 2, pp. 1–15, 2017.

[15] W. Qiao, K. Huang, M. Azimi, and S. Han, "A novel hybrid prediction model for hourly gas consumption in supply side based on improved whale optimization algorithm and relevance vector machine," *IEEE Access*, vol. 7, pp. 88218–88230, 2019.

[16] G. Sun, P. Jiang, H. Xu, S. Yu, D. Guo, G. Lin, and H. Wu, "Outlier detection and correction for monitoring data of water quality based on improved VMD and LSSVM," *Complexity*, vol. 2019, pp. 9643921-1–9643921-12, Jan. 2019.

[17] Q. Wu and H. Li, "Short-term wind speed forecasting based on hybrid variational mode decomposition and least squares support vector machine optimized by bat algorithm model," *Sustainability*, vol. 11, no. 3, pp. 1–18, 2019.

[18] F. Z. Zhao, S. Hao, Y. Zhang, S. H. Du, X. G. Dan, J. Su, T. J. Jing, and T. T. Zhao, "Short-term load forecasting for distribution transformer based on VMD-BA-LSSVM algorithm," *Trans. Chin. Soc. Agricult. Eng.*, vol. 35, no. 14, pp. 190–197, 2019.

[19] L. Xiang, Z. Deng, and A. Hu, "Forecasting short-term wind speed based on IEWT-LSSVM model optimized by bird swarm algorithm," *IEEE Access*, vol. 7, pp. 59333–59345, 2019.

[20] J. G. Meng, "Model of medium-long-term precipitation forecasting in arid areas based on PSO and LSSVM methods," *J. Yangtze River Sci. Res. Inst.*, vol. 33, no. 10, pp. 36–40, 2016.

[21] S. Arora and S. Singh, "Butterfly optimization algorithm: A novel approach for global optimization," *Soft Comput.*, vol. 23, no. 3, pp. 715–734, Feb. 2019.

[22] P. Du, J. Wang, W. Yang, and T. Niu, "Container throughput forecasting using a novel hybrid learning method with error correction strategy," *Knowl.-Based Syst.*, vol. 182, pp. 104853-1–104853-12, Oct. 2019.

[23] T. T. Dang and D. Lin, "Spider monkey optimization algorithm with dynamic self-adaptive inertia weight," *Comput. Eng. Appl.*, vol. 55, no. 14, pp. 40–47, 2019.

[24] X. Zhang, G. C. Chen, and W. X. Gong, "Artificial bee colony algorithm based on adaptive cauchy mutation and its convergence analysis," *Comput. Syst. Appl.*, vol. 24, no. 12, pp. 118–124, 2015.

[25] H. Yang, L. P. Gao, and G. H. Li, "Underwater acoustic signal prediction based on MVMD and optimized kernel extreme learning machine," *Complexity*, vol. 2020, pp. 6947059-1–6947059-17, Apr. 2020.

[26] M. Santhosh, C. Venkaiah, and D. M. V. Kumar, "Short-term wind speed forecasting approach using ensemble empirical mode decomposition and deep Boltzmann machine," *Sustain. Energy, Grids Netw.*, vol. 19, pp. 100242-1–100242-16, Sep. 2019.

[27] N. Bokde, A. Feijóo, D. Villanueva, and K. Kulat, "A review on hybrid empirical mode decomposition models for wind speed and wind power prediction," *Energies*, vol. 12, no. 2, pp. 254-1–254-42, 2019.

[28] R. Venkata Ramana, B. Krishna, S. R. Kumar, and N. G. Pandey, "Monthly rainfall prediction using wavelet neural network analysis," *Water Resour. Manage.*, vol. 27, no. 10, pp. 3697–3711, Aug. 2013.

[29] H. Hu, K. Yang, D. W. Zhu, X. J. Shen, and Z. W. Tian, "Prediction of precipitation based on EEMD and GRNN," *Water Resour. Power*, vol. 35, no. 4, pp. 10–14, 2017.

[30] N. Bokde, A. Feijóo, and K. Kulat, "Analysis of differencing and decomposition preprocessing methods for wind speed prediction," *Appl. Soft Comput.*, vol. 71, pp. 926–938, Oct. 2018.

[31] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 531–544, Feb. 2014.

[32] Y. Wang, R. Markert, J. Xiang, and W. Zheng, "Research on variational mode decomposition and its application in detecting rub-impact fault of the rotor system," *Mech. Syst. Signal Process.*, vols. 60–61, pp. 243–251, Aug. 2015.

[33] C. Aneesh, S. Kumar, P. M. Hisham, and K. P. Soman, "Performance comparison of variational mode decomposition over empirical wavelet transform for the classification of power quality disturbances using support vector machine," *Procedia Comput. Sci.*, vol. 46, pp. 372–380, 2015.

[34] Q. Wu and H. Lin, "A novel optimal-hybrid model for daily air quality index prediction considering air pollutant factors," *Sci. Total Environ.*, vol. 683, pp. 808–821, Sep. 2019.

[35] G. H. Li, X. Ma, and H. Yang, "A hybrid model for forecasting sunspots time series based on variational mode decomposition and back propagation neural network improved by firefly algorithm," *Comput. Intell. Neurosci.*, vol. 2018, pp. 3713410-1–3713410-17, Jan. 2018.

[36] Y. Ren, P. N. Suganthan, and N. Srikanth, "Ensemble methods for wind and solar power forecasting—A state-of-the-art review," *Renew. Sustain. Energy Rev.*, vol. 50, pp. 82–91, Oct. 2015.

[37] R. R. Yager, "Induced aggregation operators," *Fuzzy Sets Syst.*, vol. 137, no. 1, pp. 59–69, Jul. 2003.

[38] J. M. Zhu, P. Wu, and H. Y. Chen, "A hybrid forecasting approach to air quality time series based on endpoint condition and combined forecasting model," *Int. J. Environ. Res. Public Health*, vol. 15, no. 9, pp. 1–19, 2018.

[39] Y. Huang, L. Yang, S. Liu, and G. Wang, "Multi-step wind speed forecasting based on ensemble empirical mode decomposition, long short term memory network and error correction strategy," *Energies*, vol. 12, no. 10, pp. 1822-1–1822-22, May 2019.

[40] G. Li, X. Ma, and H. Yang, "A hybrid model for monthly precipitation time series forecasting based on variational mode decomposition with extreme learning machine," *Information*, vol. 9, no. 7, pp. 177-1–177-13, Jul. 2018.

[41] J. Farajzadeh and F. Alizadeh, "A hybrid linear–nonlinear approach to predict the monthly rainfall over the Urmia lake watershed using wavelet-SARIMAX-LSSVM conjugated model," *J. Hydroinform.*, vol. 20, no. 1, pp. 246–262, Jan. 2018.

[42] Y. Zhang, P. Han, D. F. Wang, and S. Wang, "Short term prediction of wind speed for wind farm based on variational mode decomposition and LSSVM model," *Acta Energiae Solaris Sinica*, vol. 39, no. 1, pp. 195–202, 2018.

[43] R. D. Nowak and B. D. Van Veen, "Random and pseudorandom inputs for Volterra filter identification," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 2124–2135, Aug. 1994.

[44] G. H. Li, W. N. Chang, and H. Yang, "Monthly mean meteorological temperature prediction based on VMD-DSE and Volterra adaptive model," *Adv. Meteorol.*, vol. 2020, pp. 8385416-1–8385416-17, Mar. 2020.

[45] Y. Wang, D. Wang, and Y. Tang, "Clustered hybrid wind power prediction model based on ARMA, PSO-SVM, and clustering methods," *IEEE Access*, vol. 8, pp. 17071–17079, 2020.

[46] Y. Sun, K. Li, Y. Yan, X. Wei, and C. Zhang, "Combination load forecasting method for CCHP system based on IOWA operator," in *Proc. Chin. Autom. Congr.*, Oct. 2017, pp. 4193–4197.

[47] L. Yu, S. Wang, and K. K. Lai, "Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm," *Energy Econ.*, vol. 30, no. 5, pp. 2623–2635, Sep. 2008.

[48] F. X. Diebold and R. S. Mariano, "Comparing predictive accuracy," *J. Bus. Econ. Stat.*, vol. 13, no. 3, pp. 253–263, 1995.

[49] C. Liu, Y. Wu, and C. Zhen, "Rolling bearing fault diagnosis based on variational mode decomposition and fuzzy C means clustering," *Proc. CSEE*, vol. 35, no. 13, pp. 3358–3365, Jul. 2015.

[50] H. Zhou, C. Xu, S. H. Shao, and D. M. Li, "Adaptive free search algorithm," *Comput. Sci.*, vol. 35, no. 10, pp. 188–191, 2008.

GUOHUI LI received the bachelor's degree in electronic information engineering from the Chongqing University of Technology, Chongqing, China, in 2001, the master's degree in circuits and systems from the University of Electronic Science and Technology of China, Chengdu, China, in 2004, and the Ph.D. degree in acoustics from Northwestern Polytechnical University, Xi'an, China, in 2015. He is currently an Associate Professor with the School of Electronic Engineering, Xi'an University of Posts and Telecommunications, Shaanxi, China. His research interests include underwater acoustic signal processing and chaotic signal processing.

WANNI CHANG received the bachelor's degree in communication engineering from Xi'an Polytechnic University, Xi'an, China, in 2018. She is currently pursuing the master's degree in electronics and communication engineering with the Xi'an University of Posts and Telecommunications, Xi'an. Her research interests include underwater acoustic signal processing and chaotic signal prediction.

HONG YANG received the bachelor's and master's degrees in mechanical and electronic engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2003 and 2006, respectively, and the Ph.D. degree in acoustics from Northwestern Polytechnical University, Xi'an, China, in 2015. She is currently an Associate Professor with the School of Electronic Engineering, Xi'an University of Posts and Telecommunications, Shaanxi, China. Her research interests include underwater acoustic signal processing and chaotic signal processing.

• • •