

Research Article

A New Hybrid Model for Underwater Acoustic Signal Prediction

Guohui Li , Wannan Chang, and Hong Yang 

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

Correspondence should be addressed to Guohui Li; lghcd@163.com and Hong Yang; uestcyhong@163.com

Received 14 January 2020; Revised 16 June 2020; Accepted 26 June 2020; Published 25 July 2020

Academic Editor: Harish Garg

Copyright © 2020 Guohui Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The prediction of underwater acoustic signal is the basis of underwater acoustic signal processing, which can be applied to underwater target signal noise reduction, detection, and feature extraction. Therefore, it is of great significance to improve the prediction accuracy of underwater acoustic signal. Aiming at the difficulty in underwater acoustic signal sequence prediction, a new hybrid prediction model for underwater acoustic signal is proposed in this paper, which combines the advantages of variational mode decomposition (VMD), artificial intelligence method, and optimization algorithm. In order to reduce the complexity of underwater acoustic signal sequence and improve operation efficiency, the original signal is decomposed by VMD into intrinsic mode components (IMFs) according to the characteristics of the signal, and dispersion entropy (DE) is used to analyze the complexity of IMF. The subsequences (VMD-DE) are obtained by adding the IMF with similar complexity. Then, extreme learning machine (ELM) is used to predict the low-frequency subsequence obtained by VMD-DE. Support vector regression (SVR) is used to predict the high-frequency subsequence. In addition, an artificial bee colony (ABC) algorithm is used to optimize model performance by adjusting the parameters of SVR. The experimental results show that the proposed new hybrid model can provide enhanced accuracy with the reduction of prediction error compared with other existing prediction methods.

1. Introduction

Underwater acoustic signal processing is one of the most active disciplines in the field of ocean and information [1]. It was widely used in military and civil applications [2]. Underwater acoustic signal in the marine environment has been widely concerned by some scholars at home and abroad [3, 4]. Taroudakis et al. [5, 6] analyzed the statistical characteristics of underwater acoustic signal and studied the application of the statistical characteristics of signal in geoacoustic inversions and applications in ocean acoustic tomography, Li et al. [7] established a model to extract the characteristics of underwater acoustic signal. The research showed that underwater acoustic signal has not only the characteristics of nonstationary, non-Gaussian, and nonlinear but also the typical characteristics of chaos and fractal [8, 9]. This means that underwater acoustic signal with nonlinear characteristics has local predictability.

Liang et al. [10] stated that joint prediction results can be applied to the extraction of nonlinear features from underwater acoustic signal. In addition, the improvement of the prediction accuracy result for underwater acoustic signal

can help for the reduction of the signal-to-noise ratio [11, 12]. Therefore, it is very important to study the prediction of underwater acoustic signal in underwater acoustic signal processing. At present, artificial neural network and Volterra nonlinear model are mainly used to predict underwater acoustic signal of ships. Fang et al. [13] and Sun et al. [14] used the Volterra series theory to establish a nonlinear dynamic model of underwater acoustic signal and realized the suppression of background noise through local prediction of underwater acoustic signal. Liu [15] used the neural network and Volterra to predict sea clutter. Zhou et al. [16] used the theory of artificial neural network to predict underwater acoustic signal of ships. He et al. [17] proposed the automatic search algorithm of particle swarm optimization (PSO) for RBF neural network based on the phase space theory to predict underwater acoustic signals. Yang et al. [18] used wavelet neural network to predict underwater acoustic signal. Although the above methods have achieved good prediction results of underwater acoustic signal, the neural network is easy to fall into local optimum, long calculation time, and easy to oscillate. As a polynomial model, Volterra is difficult to obtain satisfactory

prediction results of strongly nonlinear series. Aiming at these problems, some artificial intelligence prediction models with good performance were used to improve the prediction performance of nonlinear signal. The extreme learning machine (ELM) was proposed by Huang et al. [19] which has the advantages of fast training speed and without falling into the local minimum value. It has been successfully applied in various fields of prediction [20, 21]. Support vector regression (SVR) is a machine learning algorithm. It maps the original data to high-dimensional space by nonlinear function to expand regression analysis. It has obvious advantages in solving nonlinear problems and can effectively improve the generalization ability and prediction accuracy of the model [22–24]. However, parameter optimization is an important problem in SVR research. The selection of penalty parameters and kernel parameters directly affects the prediction accuracy and generalization performance of SVR. An intelligent optimization algorithm was widely used in parameter optimization of SVR because of its good optimization performance [25, 26]. Wang et al. [27] used genetic algorithm (GA) to optimize the SVR model and applied it in the field of energy prediction. In order to improve the optimization ability of GA, Garg proposed GSA-GA and PSO-GA to solve the constraint optimization problem [28, 29]. The artificial bee colony algorithm was a new swarm intelligence optimization algorithm [30]. Garg [31] used an artificial bee colony algorithm to solve structural optimization design problems. The algorithm simulates the foraging behavior of the swarm. It solves the contradiction between expanding the new solution space and accurately searching in the old solution space through the cooperation between different solution spaces. It avoids falling into the local optimal solution problem and has better performance [32].

Many scholars apply signal features to various problems. Hossen et al. [33] used statistical signal features to extract parameters and classify different modulated signal. Taroudakis and Smaragdakis [6] used the characteristics of the signal for the inversion of underwater signal. Mode decomposition technology used the characteristics of signal frequency domain to divide the signal frequency band with the purpose of reduction modeling complexity. Empirical mode decomposition (EMD) proposed by Huang et al. [34] is a data-driven signal decomposition method, which has been proved as a good time-frequency analysis tool. However, mode mixing and end effects [35, 36] were suffered due to the sensitivity of EMD to noise. In 2014, Dragomiretskiy and Zosso [37] proposed the variational mode decomposition (VMD) to solve the problems of mode mixing and end effects in EMD. As VMD has strong decomposition ability, better noise robustness in signal decomposition, and fast processing speed, it has been successfully applied in many fields [38, 39]. Ali et al. [40] proposed the wind speed data decomposition using VMD to reduce the difficulty in prediction. Wu and Lin [41] combined VMD and wavelet decomposition to predict air quality index and improve the prediction accuracy of the AQI index. Li et al. [42] proposed a prediction model of sunspot number time series based on the combination of VMD and BP neural network. Yang et al.

[43] proposed a prediction model of underwater acoustic signal based on the combination of VMD and LSSVM.

As mentioned above, some prediction methods of underwater acoustic signal such as Volterra model [13], wavelet neural network [18], and mode decomposition technology combined prediction method [21, 43] were proposed for different prediction models. Although several prediction methods have been developed, they still have some limitations: (i) it is found that a single prediction model cannot fully capture the nonlinear data information and for the requirements of high prediction accuracy; (ii) the traditional decomposition integration model uses the decomposition method to predict each mode component, which take a long computation time. In order to overcome these issues, this paper proposes the combination of sequence decomposition technology and optimal prediction model for the underwater acoustic signal prediction. The dispersion entropy (DE) is used to calculate the entropy value of each mode component and analyze the complexity. Then, the mode components with approximate entropy are combined to reduce the calculation scale of each mode component with the improvement of the prediction performance. The results show that the proposed method is an effective prediction method for underwater acoustic signal.

The main contents of this paper are as follows: in Section 2, the basic theory for each part of the hybrid prediction model will be introduced; in Section 3, the overall framework of the model will be presented; in Section 4, results and discussion for the proposed prediction hybrid model for underwater acoustic signal will be discussed; and conclusions will be presented as the last section of this paper.

2. Basic Theory

2.1. Variational Mode Decomposition. VMD is a typical instantaneous frequency analysis method proposed by Dragomiretskiy and Zosso [37]. The main function of the VMD method is to stabilize the signal. The fluctuation of the original signal for different frequencies is decomposed into a series of sequences with different characteristics, and each sequence is called an intrinsic mode function (IMF). The accuracy of the prediction model is increased by decomposing the complex nonstationary signal into a series of simple stable signal with different frequencies using VMD. The principle of VMD decomposition is as follows.

In the following equation, each IMF obtained by VMD is defined as amplitude modulated frequency modulated (AM-FM) signal:

$$u_k(t) = A_k(t)\cos(\varphi_k(t)), \quad (1)$$

where $u_k(t)$ is the k th IMF component and $A_k(t)$ and $\varphi_k(t)$ are the instantaneous amplitude and phase of $u_k(t)$, respectively. The instantaneous frequency $u(t)$ is recorded as $\omega(t)$, which is obtained by $\varphi(t)$ differentiation.

VMD theory assumes that the input signal $x(t)$ is composed of a finite number of IMF with limited bandwidth and different frequencies. Under the constraint that the sum of each IMF component is equal to the input signal $x(t)$, the

variational model of signal decomposition is constructed with the goal of minimizing the sum of estimated bandwidth of each IMF. The process of establishing the variational model is as follows:

- (1) Hilbert transform is applied to each mode component, and the unilateral spectrum of the mode function is obtained by Hilbert transform and construction of analytic signal.
- (2) The analysis signal of each mode component is mixed with the corresponding center frequency, and the spectrum of each mode component to the baseband is moved.
- (3) The bandwidth of each mode component is estimated.
- (4) The constrained variational model is constructed by introducing constraints. The concrete structure is as follows:

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

In equation (2), $\{u_k\}$ and $\{\omega_k\}$ represent the set of subsignal and their center frequencies, respectively; K represents the total number of subsignal; and $\delta(t)$ represents the Dirac distribution.

For solving variational problems, the extended Lagrangian function $L(\{u_k\}, \{\omega_k\}, \{\lambda\})$ shown in equation (3) can transform the constrained problem into the unconstrained problem. In the following equation, α is the quadratic penalty factor, $\theta(t)$ is the Lagrange multiplier, $\{u_k\}$ and $\{\omega_k\}$ are the set of subsignal and their center frequencies, and $f(t)$ is the original signal:

$$L(\{u_k\}, \{\omega_k\}, \theta) = C \sum_{k=1}^K \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 + \left\| f(t) - \sum_{k=1}^K u_k(t) \right\|_2^2 + \left\langle \theta(t), f(t) - \sum_{k=1}^K u_k(t) \right\rangle. \quad (3)$$

The variational mode decomposition uses the alternate direction method of multipliers (ADMM) to solve equation (3); \hat{u}_k^{n+1} , $\hat{\omega}_k^{n+1}$, and $\hat{\theta}^{n+1}$ are updated alternately, where n denotes the number of iterations. The formulas are as follows:

$$\begin{aligned} \hat{u}_k^{n+1}(\omega) &= \frac{\hat{f}(\omega) - \sum_{k=1}^K \hat{u}_k(\omega) + (\hat{\theta}(\omega)/2)}{1 + 2C(\omega - \omega_k)^2}, \\ \hat{\omega}_k^{n+1} &= \frac{\int_0^\infty \omega |\hat{u}_k(\omega)|^2 d\omega}{\int_0^\infty |\hat{u}_k(\omega)|^2 d\omega}, \\ \hat{\theta}^{n+1}(\omega) &= \hat{\theta}^{n+1}(\omega) + \tau \left[\hat{f}(\omega) - \sum_{k=1}^K \hat{u}_k^{n+1}(\omega) \right]. \end{aligned} \quad (4)$$

Given discriminant accuracy $e > 0$, the convergence condition for stopping iteration is as follows:

$$\sum_{k=1}^K \frac{\|\hat{u}_k^{n+1} - \hat{u}_k^n\|_2^2}{\|\hat{u}_k^n\|_2^2} < e. \quad (5)$$

2.2. Dispersion Entropy. Dispersion entropy (DE) is a new method proposed by Rostaghi and Azami [44] in 2016 which is used to quantify the complexity of time series. It is faster than sample entropy in calculation and avoids the problem that permutation entropy ignores signal amplitude value. For a given time series $x = \{x_1, x_2, \dots, x_N\}$, the steps of calculating dispersion entropy are as follows:

Step 1: the normal cumulative distribution function is used to normalize the signal x to $y = \{y_1, y_2, \dots, y_N\}$, and $z_i^c = \text{round}(c \times y_i + 0.5)$ is used to assign $[1, 2, \dots, c]$ integers to y_i , where c is the number of categories.

Step 2: construct the embedding vector $z_i^{m,c}$, that is,

$$z_i^{m,c} = \{z_i^c, z_{i+d}^c, \dots, z_{i+(m-1)d}^c\}, \quad (6)$$

$$i = 1, 2, \dots, N - (m-1)d,$$

where m and d represent the embedding dimension and delay time, respectively. N is the number of sample points, and the final number of i is determined by the number of sample points n and m and d . Map $z_i^{m,c}$ to scatter mode $\pi_{v_0, v_1, \dots, v_{m-1}}$, where $z_i^c = v_0, z_{i+d}^c = v_1, \dots, z_{i+(m-1)d}^c = v_{m-1}$, and the number of scatter modes allocated to $z_i^{m,c}$ is c^m .

Step 3: calculate the relative frequency of c^m dispersion modes:

$$p(\pi_{v_0, v_1, \dots, v_{m-1}}) = \frac{N_b(\pi_{v_0, v_1, \dots, v_{m-1}})}{N - (m-1)d}, \quad (7)$$

where $N_b(\pi_{v_0, v_1, \dots, v_{m-1}})$ represents the number of $z_i^{m,c}$ mapped to $\pi_{v_0, v_1, \dots, v_{m-1}}$, so $p(\pi_{v_0, v_1, \dots, v_{m-1}})$ can represent the ratio of the number of elements in $N_b(\pi_{v_0, v_1, \dots, v_{m-1}})$ and $z_i^{m,c}$.

Step 4: the formula of DE is defined as follows:

$$\text{DE}(x, m, c, d) = - \sum_{\pi=1}^{c^m} p(\pi_{v_0, v_1, \dots, v_{m-1}}) \cdot \ln(p(\pi_{v_0, v_1, \dots, v_{m-1}})). \quad (8)$$

When calculating DE, m and d need to be set in advance. It is suggested in [43] that m usually takes 2 or 3, c is an integer between [4, 8], and the time delay d is generally 1. In this paper, m , d , and c are set to 3, 1, and 6, respectively.

2.3. ABC-SVR

2.3.1. Support Vector Regression. Vapnik [45] established support vector machine (SVM) based on the principle of structural risk minimization (SRM), which was applied to solve the classification problem. Then, the idea of SVM was

used to solve the regression problem. In SVR, data x are mapped into high-dimensional linear space by using nonlinear map in order to transform from nonlinear regression of low-dimensional feature space into high-dimensional feature space linear regression problem. Let the training set $T = \{(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)\}$, where the function is expressed as

$$f(x) = w \cdot \Phi(x) + b, \quad (9)$$

where $\Phi(x)$ is a nonlinear mapping from X to high-dimensional Hilbert space, w is the weight vector, b is the threshold, and w and b can be estimated by minimizing the following equation:

$$R(f) = \frac{1}{2}\|w\|^2 + C \frac{1}{l} \sum_{i=1}^l L_\varepsilon(y_i, f(x_i)), \quad (10)$$

where C is the penalty factor and L_ε is the loss function. By introducing relaxation variables ξ_i and ξ_i^* , the optimization objective can be expressed as follows:

$$\begin{aligned} \min \quad & \frac{1}{2}\|w\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \\ \text{s.t.} \quad & \begin{cases} y_i - w \cdot \Phi(x_i) - b \leq \varepsilon + \xi_i \\ w \cdot \Phi(x_i) + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0, i = 1, 2, \dots, l. \end{cases} \end{aligned} \quad (11)$$

Lagrange functions can be obtained by introducing Lagrange multipliers $\alpha_i, \alpha_i^*, u_i,$ and u_i^* :

$$\begin{aligned} L(w, b, \alpha_i, \alpha_i^*, u_i, u_i^*) \\ = \frac{1}{2}\|w\|^2 + C \sum_{i=1}^m (\xi_i + \xi_i^*) - \sum_{i=1}^m u_i \xi_i - \sum_{i=1}^m u_i^* \xi_i^* \\ + \sum_{i=1}^m \alpha_i (f(x_i) - y_i - \varepsilon - \xi_i) + \sum_{i=1}^m \alpha_i^* (y_i - f(x_i) - \varepsilon - \xi_i^*). \end{aligned} \quad (12)$$

By substituting equation (9) into equation (12) and making the partial derivative of $w, b, \xi_i,$ and ξ_i^* zero, we get

$$\begin{aligned} w &= \sum_{i=1}^m (\alpha_i^* - \alpha_i) x_i, \\ 0 &= \sum_{i=1}^m (\alpha_i^* - \alpha_i), \\ C &= \alpha_i + u_i, \\ C &= \alpha_i^* + u_i^*. \end{aligned} \quad (13)$$

Substituting equation (13) into equation (12) can be transformed into the corresponding dual problem:

$$\begin{aligned} \max \quad & -\frac{1}{2} \sum_{i,j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(x_i \cdot x_j) \\ & - \sum_{i=1}^l \alpha_i (\varepsilon - y_i) - \sum_{i=1}^l \alpha_i^* (\varepsilon + y_i) \\ \text{s.t.} \quad & \begin{cases} \sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0 \\ 0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, 2, \dots, l. \end{cases} \end{aligned} \quad (14)$$

By solving the above problems, the regression function of SVR can be obtained:

$$f(x) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) K(x_i, x_j) + b. \quad (15)$$

In the formula, $K(x_i, x_j)$ is the kernel function, which needs to meet Mercer condition. In this paper, Gaussian RBF kernel function [46] is selected as

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) = \exp(-g\|x_i - x_j\|^2), \quad g > 0. \quad (16)$$

It can be seen that penalty factors and kernel function parameters are involved in the calculation process of SVR, which are represented by c and g in this paper. c represents the tolerance of the model to errors, and the value of c affects the generalization ability of the model. g is the parameter of RBF function as kernel, which determines the distribution of data mapped to the new feature space and affects the speed of training and prediction. In this paper, the initial range of penalty parameter and kernel function parameter is set to [0.01–100], and the optimal model parameter value is selected through the iterative optimization of ABC algorithm.

2.3.2. Artificial Bee Colony Algorithm. Artificial bee colony (ABC) algorithm is an optimization algorithm to simulate bee colony behavior, which was proposed by Karaboga and Basturk [47] in the early 21st century. Karaboga et al. have successfully applied the artificial bee colony algorithm to the unconstrained numerical optimization function. During the process of honey gathering in nature, all bees are working together to search optimal food source by sharing the information among them. The artificial bee colony algorithm adopts heuristic search strategy for both local and global search for achieving optimal solution. The algorithm has strong adaptability and good universality. The foraging behavior of bee colony intelligent includes three parts: honey source, employed foragers (EF) called leader bees, and unemployed foragers (UF) including scout bees and follower bees. In addition, there are three basic behaviors: searching

for nectar sources (S), recruiting nectar sources (R), and giving up nectar sources. The working diagram of bee colony collecting food is shown in Figure 1.

The food source is defined as a location in the search space. The initial number of food sources is equal to the number of leader bees and follower bees. The main steps are as follows:

Step 1: after initialization, the bee begins to search all the initial solutions circularly. It includes the population number, the maximum iteration number, the control parameter, and the range of the solution.

Step 2: at the beginning of the search process, each leader bee generates a new food source by the following equation:

$$V_{ij} = X_{ij} + \Phi_{ij}(X_{ij} - X_{kj}), \quad (17)$$

where $k \in \{1, 2, \dots, SN\}$, $j \in \{1, 2, \dots, D\}$ and $k \neq i$ and Φ_{ij} is a random number in $[-1, 1]$. The fitness of new honey candidates is determined. If the fitness of the new candidate honey source v_i is larger than the solution x_i in memory, the new solution will replace the original solution.

Step 3: after all the leader bees have finished searching, they will share the nectar source information with the follower bees. Then, bees choose the location of the honey source with a certain probability of nectar source P_i . The calculation P_i is shown in the following equation:

$$P_i = \frac{\text{fit}_i}{\sum_{n=1}^{SN} \text{fit}_n}, \quad (18)$$

where fit_i is the fitness of the i th solution and SN is the number of food sources. After the honey source is determined, follower bees search in the neighborhood of the honey source to find a new honey source. If the fitness of the new honey source is larger than the solution in memory, the original solution is replaced by the new solution. On the contrary, the fitness of the new solution is preserved.

Step 4: although all the follower bees have finished the searching process, the follower bees and the leader bees still do not find a more adaptable honey source. Then, the solution falls into the local optimum, and the honey source will be abandoned. At the same time, the leader bee or the follower bee is transformed into a scout bee, and a new honey source is randomly generated according to the following equation:

$$\text{new_}X_{ij} = X_{ij} + \text{rand}(0, 1)(X_{\max j} - X_{\min j}), \quad (19)$$

where $j \in \{1, 2, \dots, D\}$ and then return to the leader bee search process and start to repeat the cycle.

Step 5: fitness is the objective function of optimization problem, and the calculation formula of fitness value is shown in equation (20), where fit is the fitness value and f_i is the objective function value of the i th solution. In this paper, the objective function represents the mean

square error (MSE) of SVR. The formula of MSE is shown in equation (21), where $\hat{x}(t)$ represents predictive data, $x(t)$ represents actual data, and n represents the number of sample points:

$$\text{fit}_i = \begin{cases} \frac{1}{1 + f_i}, & f_i \geq 0, \\ 1 + \text{abs}(f_i), & f_i < 0, \end{cases} \quad (20)$$

$$\text{MSE} = \frac{1}{n} \sum_{t=1}^n (\hat{x}(t) - x(t))^2. \quad (21)$$

2.4. Extreme Learning Machine. Extreme learning machine (ELM) is a new algorithm for single hidden layer feed-forward neural network (SLFNs) proposed by Huang et al. [19]. The ELM structure shown in Figure 2 consists of input layer, hidden layer, and output layer. Among them, there are n neurons in the input layer for n input variables, l neuron in the hidden layer, and m neurons in the output layer for m output variables.

In Figure 2, w_{nl} represents the connection weight between the n th neuron in the input layer and the l th neuron in the hidden layer. w_{lm} is the connection weight between the j th neuron in the hidden layer and the m th neuron in the output layer.

For n different random sample (x_i, t_i) , where $t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in R^N$, $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in R^N$, the ELM with the excitation function $g(x) = (g_1(x), g_2(x), \dots, g_l(x))$ can be expressed as follows:

$$\sum_{i=1}^l \beta_i g_i(w_i \cdot x_j + b_j) = y_j, \quad j = 1, 2, 3, \dots, N, \quad (22)$$

where $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ is the connection weight vector between the hidden layer and the input layer, the connection weight vector between the hidden layer neuron and the output layer neuron is $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$, b_j is the threshold value of the hidden layer, and y_j is the output value of the node. When the number of neurons in the hidden layer N is less than the number of samples m , the output function of the model can be expressed in the form of matrix as follows:

$$H\beta = T, \quad (23)$$

where H is the hidden layer output matrix of neural network and the specific form is as follows.

The least square solution of the connection weight $\hat{\beta}$ between the hidden layer and the output layer is obtained by solving the linear equations of the above formula:

$$\hat{\beta} = H^+ T, \quad (24)$$

where H^+ is the Moore generalized inverse matrix of the output matrix H of the hidden layer.

ELM model adopts three-layer structure, in which the number of input layer and output layer corresponds to the number of input and output variables. If the number of neurons in the hidden layer is too large, it will affect the accuracy of model prediction. So in this paper, the number

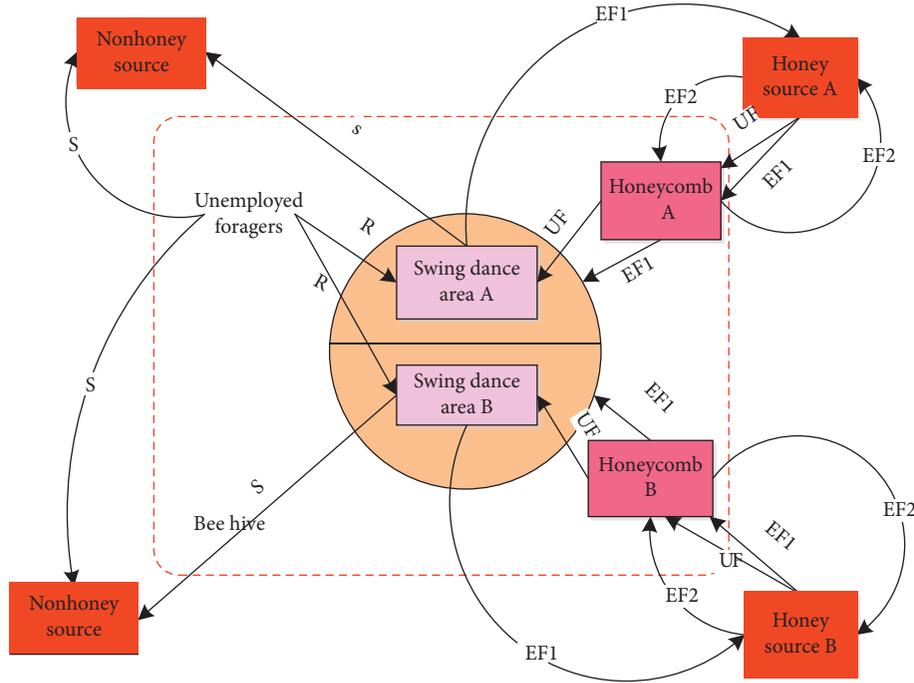


FIGURE 1: The working diagram of the bee colony collection.

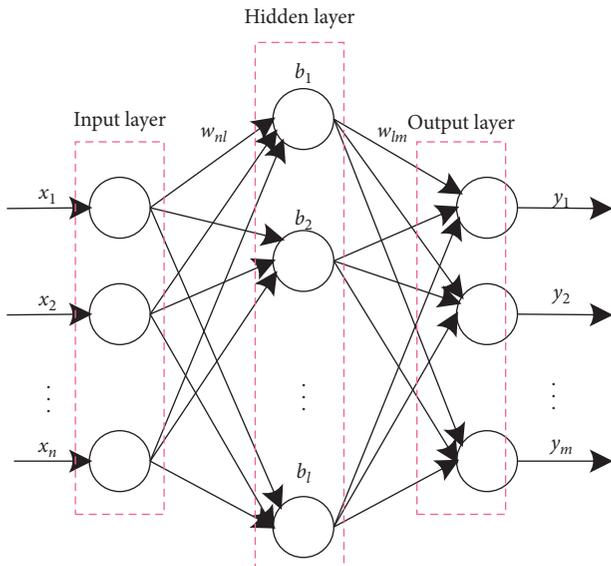


FIGURE 2: Structure diagram of extreme learning machine.

of neurons in the input layer is 5, the number of neurons in the hidden layer is 16, and the number of neurons in the input layer is 1. The activation function of hidden layer is sigmoid.

3. The Prediction Model for Underwater Acoustic Signal

3.1. The Proposed Hybrid Model. The flow chart based on the VMD-DE-ELM-ASVR hybrid prediction model for underwater acoustic signal is shown in Figure 3.

VMD decomposes underwater acoustic signal into modes, and each mode have a central frequency, for example, $IMF_1, IMF_2, \dots, IMF_n$. The function of this technique is to reduce the nonstationarity of the sequence and improve the accuracy of prediction.

In order to solve the problem of overdecomposition and computing burden, DE is used to calculate the entropy value of $IMF_1, IMF_2, \dots, IMF_n$ and analyze the complexity. Then, the IMF with similar entropy value is added to get the new subsequence $DE - IMF_{it} (i = 1, 2, \dots, s)$, where s is the number of mode components after merging.

In the combined model forecasting process, each new recombination sequence obtained in the DE process is divided into training set and test set. The training set is used for the training model, and the test set is used to verify the effect of prediction. According to the DE value and curve of each mode component, it can be divided into high-frequency sequence and low-frequency sequence. ELM is used to predict the low-frequency subsequence, SVR is used to predict the high-frequency subsequence, and ABC algorithm is used to optimize the model performance by adjusting the parameters of SVR. Figures 4 and 5 show that SVR has higher prediction accuracy in the high-frequency component of underwater acoustic signal, and ELM has higher accuracy in the low-frequency component. Therefore, the combination of the SVR and ELM model can improve the prediction accuracy and reduce the prediction error.

In the hybrid process, $DE - IMF_{it} (i = 1, 2, \dots, s)$ is the prediction result of $DE - IMF_{it} (i = 1, 2, \dots, s)$, which will be superposed to output the final prediction result $\hat{X}_t = \sum_{i=1}^s DE - IMF_{it} \cdot \hat{X}_t$ represents the predicted value of the original data.

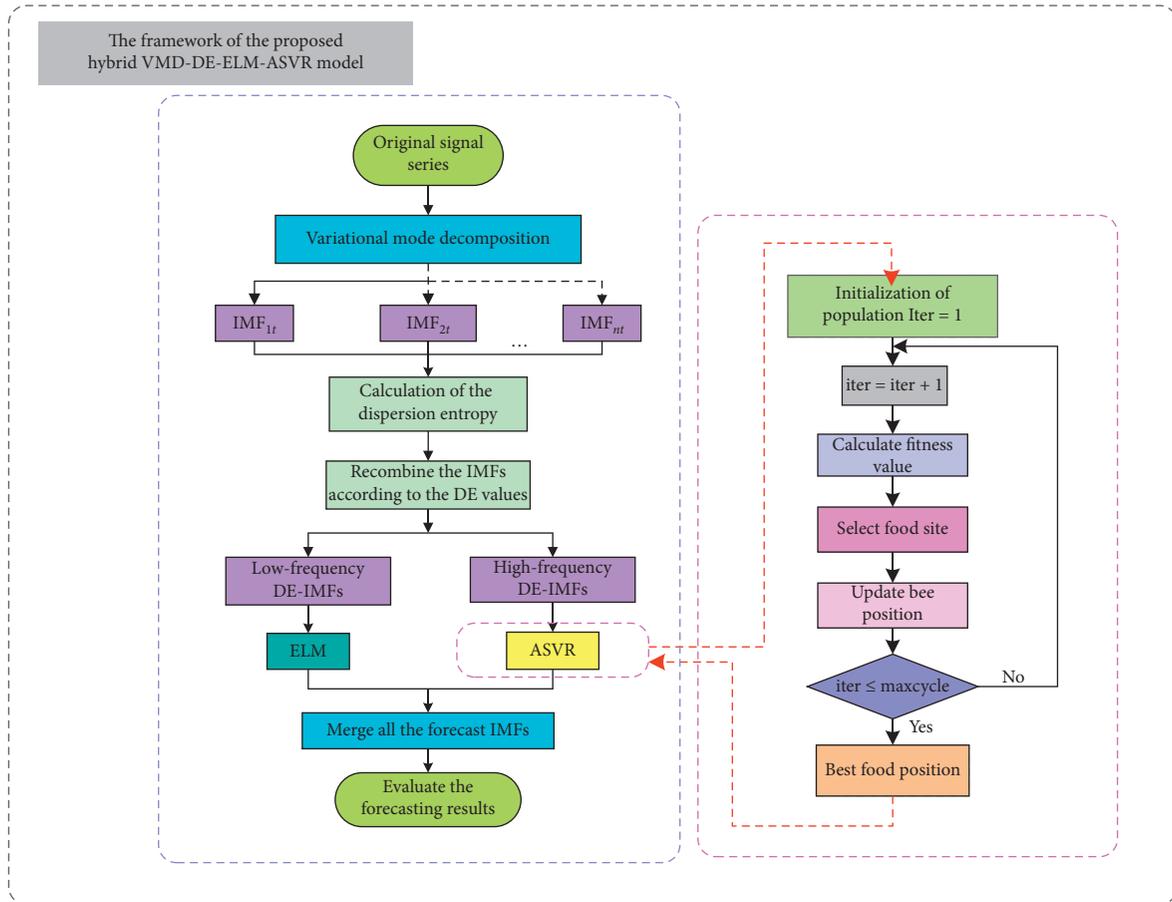


FIGURE 3: The framework of the proposed hybrid VMD-DE-ELM-ASVR model.

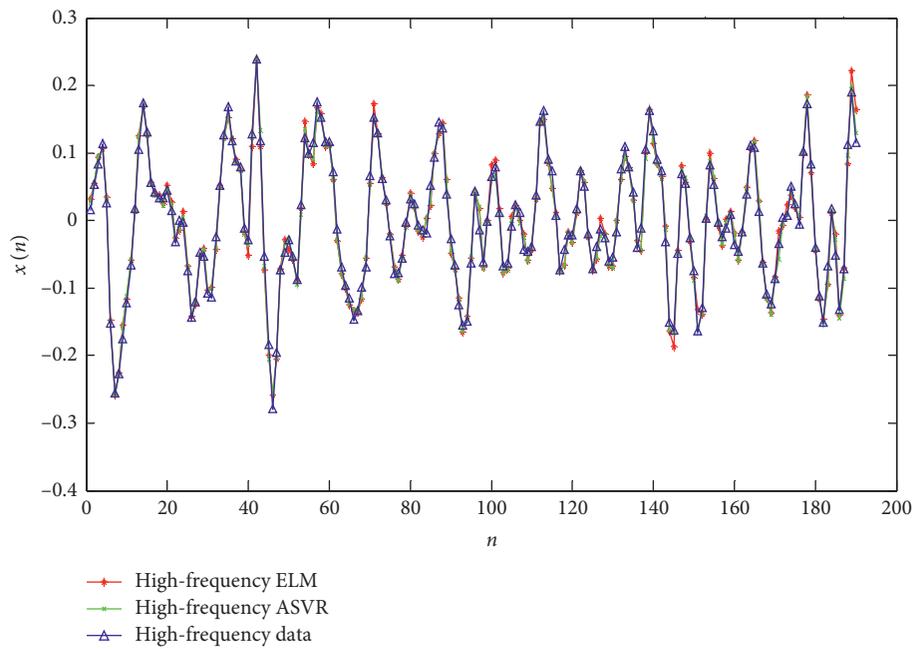


FIGURE 4: Forecasting results of high-frequency sequence of data set A by ELM and ASVR models.

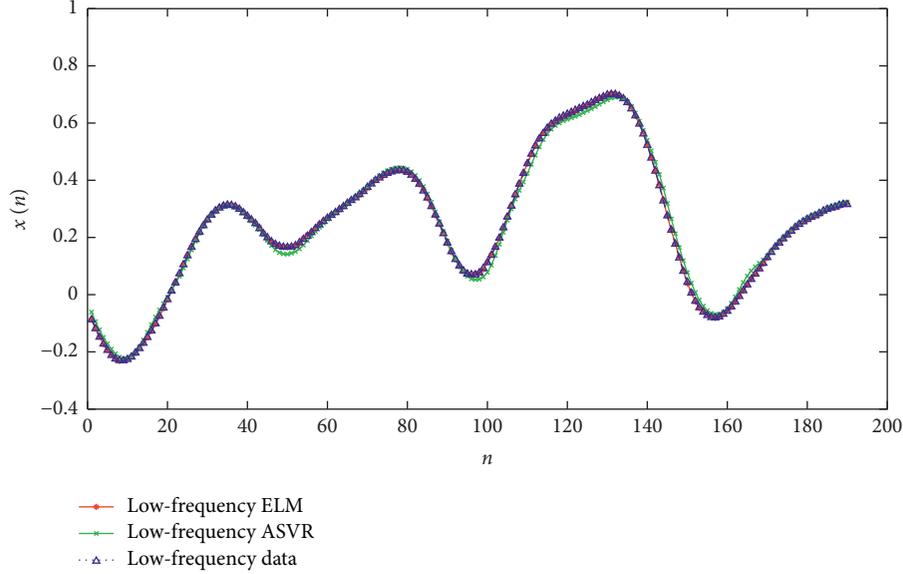


FIGURE 5: Prediction results of low-frequency sequence of data set A by ELM and ASVR models.

3.2. Performance Indicators of Prediction Accuracy. In this paper, the following three error indexes are selected to measure the prediction effect of the proposed prediction model: mean absolute error (MAE), root mean squared error (RMSE), and coefficient of determination (R^2). Using the performance index MAE and RMSE to quantify the error of prediction value, the smaller the value is, the better the prediction accuracy is. The closer the R^2 is to 1, the better the prediction performance is. The formulas are as follows:

$$\begin{aligned} \text{MAE} &= \frac{1}{n} \sum_{t=1}^n |\hat{x}(t) - x(t)|, \\ \text{RMSE} &= \sqrt{\frac{1}{n} \sum_{t=1}^n [\hat{x}(t) - x(t)]^2}, \\ R^2 &= 1 - \frac{\sum_{t=1}^n (\hat{x}(t) - x(t))^2}{\sum_{t=1}^n (\hat{x}(t) - \bar{x})^2}, \end{aligned} \quad (25)$$

where $\hat{x}(t)$ represents predictive data, $x(t)$ represents raw data, n represents the number of sample points, and \bar{x} is the average of the sample series.

4. Results and Discussion

4.1. Experimental Setup and Data Set. In order to verify the effectiveness of the proposed prediction model, experiments need to be carried out. A PC with Intel Core i7, 3.6 GHz RAM, 4 GB ROM, 32 GB memory, running the Microsoft Window 8 operating system was used as the platform on which to implement the proposed model. In addition, MATLAB R2014a software platform was also used for the implementation of the proposed model. In this paper, the normalized preprocessing method is applied to the experimental data set [48]. The sample frequency and length are

20 kHz and 2048 points, respectively. 1000 points are randomly selected as experimental data. In order to verify the prediction effect of different scale data sets, the prediction error (MSE) is taken as the index. The relationship between the prediction error (MSE) and the number of samples is shown in Figure 6.

It can be seen from Figure 6 that the size of data sets with different scales can achieve better prediction results, but when the sample data length is 1000, the prediction effect is the best. Therefore, this paper selects 1000 samples to test and uses two sample data sets A and B to verify the prediction model proposed in this paper. The length of each sample data is 1000 points, and the time-domain waveform is shown in Figure 7, where the ordinate $x(n)$ represents the amplitude of the normalized signal, and the abscissa n represents the number of sample points. For data sets A and B, 1000 data are divided into training samples and test samples. The last 190 data are test samples, and the rest are used as training samples.

4.2. VMD-DE Processing. Before decomposition, the mode number K in VMD needs to be set in advance. When the value of K is too large, additional noise will be generated or mode repetition will be caused. When the value of K is too small, the underdecomposition of components will be caused. Therefore, the selection of appropriate mode number K is very important for the result of decomposition. Because the center frequency of each mode number is different, the method of observing the center frequency is usually used to determine K [49] and check the mode center frequency corresponding to each K value. If the center frequency value is close, it is regarded as overdecomposition, and the optimal decomposition layer is $K - 1$. After repeated experiments, for data set A, when $K > 8$, the subsequent tends to be approximate, so $K = 8$ is chosen in this paper. For

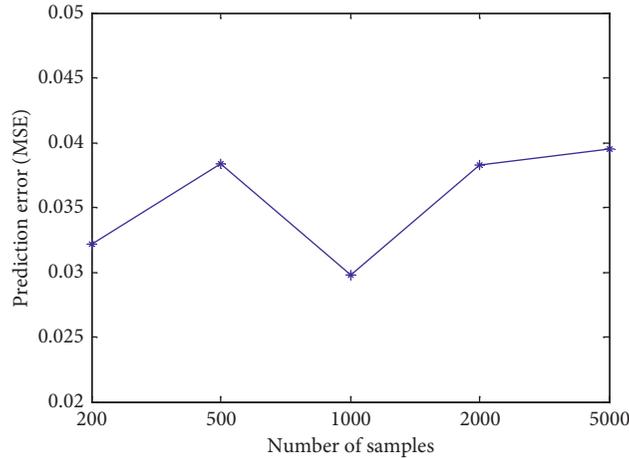


FIGURE 6: Prediction error (MSE) and sample number.

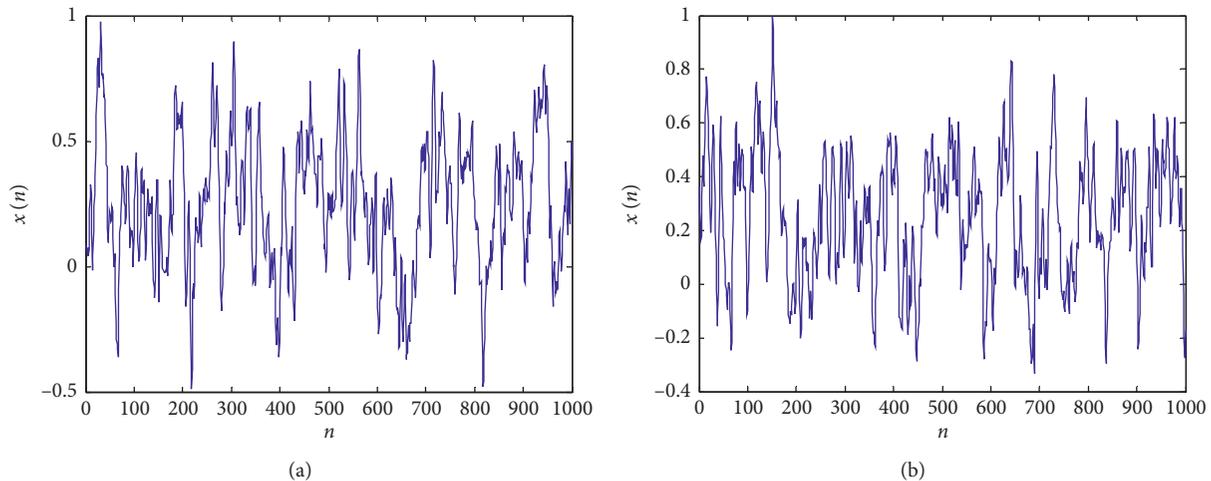


FIGURE 7: (a) Underwater acoustic signal data set A; (b) underwater acoustic signal data set B.

data set B, $K=9$ is chosen in this paper. The VMD decomposition results of two underwater acoustic signal data sets are shown in Figures 8(a) and 8(b).

As shown in Figures 8(c) and 8(d), the IMF decomposed by EMD has different characteristics. The IMF1 component is mainly random white noise. The IMF obtained by EMD decomposition is disturbed by noise, and there is a relatively large swing at both ends, which affects the whole component sequence. Therefore, the decomposition result is seriously distorted. Figure 8(a) shows that the underwater acoustic signal data set A is decomposed into eight IMFs by VMD. It can be seen that the IMF2 component decomposed by VMD is most similar to the original waveform, and the waveform distortion is small, so VMD has better noise robustness than EMD.

IMF with finite complexity approximation is obtained by VMD decomposition. If the prediction model is directly used to predict each IMF, the calculation scale will increase. In order to reduce the computation scale, the complexity of each IMF is analyzed by using the DE algorithm. The calculation results are shown in Figure 9.

It can be seen from Figure 9 that the DE value of each IMF sequence is increased for the underwater acoustic signal data sets A and B. It also shows that the complexity from low-frequency component to high-frequency component increases gradually, and the randomness of the sequence increases in turn. The results of entropy values of each IMF of the underwater acoustic signal data sets A and B are shown in Tables 1 and 2. Based on the similarity of entropy, the components are merged. For data set A, the complexity of IMF8 is the highest, and its DE value is 0.7227, which is significantly higher than other components, especially 0.3751 higher than DE value of IMF1. It further shows that IMF8 has strong randomness, which is difficult to predict. The DE values of IMF5 and IMF6 are similar to each other, and the difference is 0.0124, indicating that they can be combined into a new sequence. The difference between IMF3 and IMF4 is 0.0174, which can be combined into a new IMF component. The DE values of other modes are different from those of other modes and can be used as a new subsequence.

Table 3 shows the sample entropy (SE) of each IMF in data set A. It can be seen that the SE values of IMF3, IMF4,

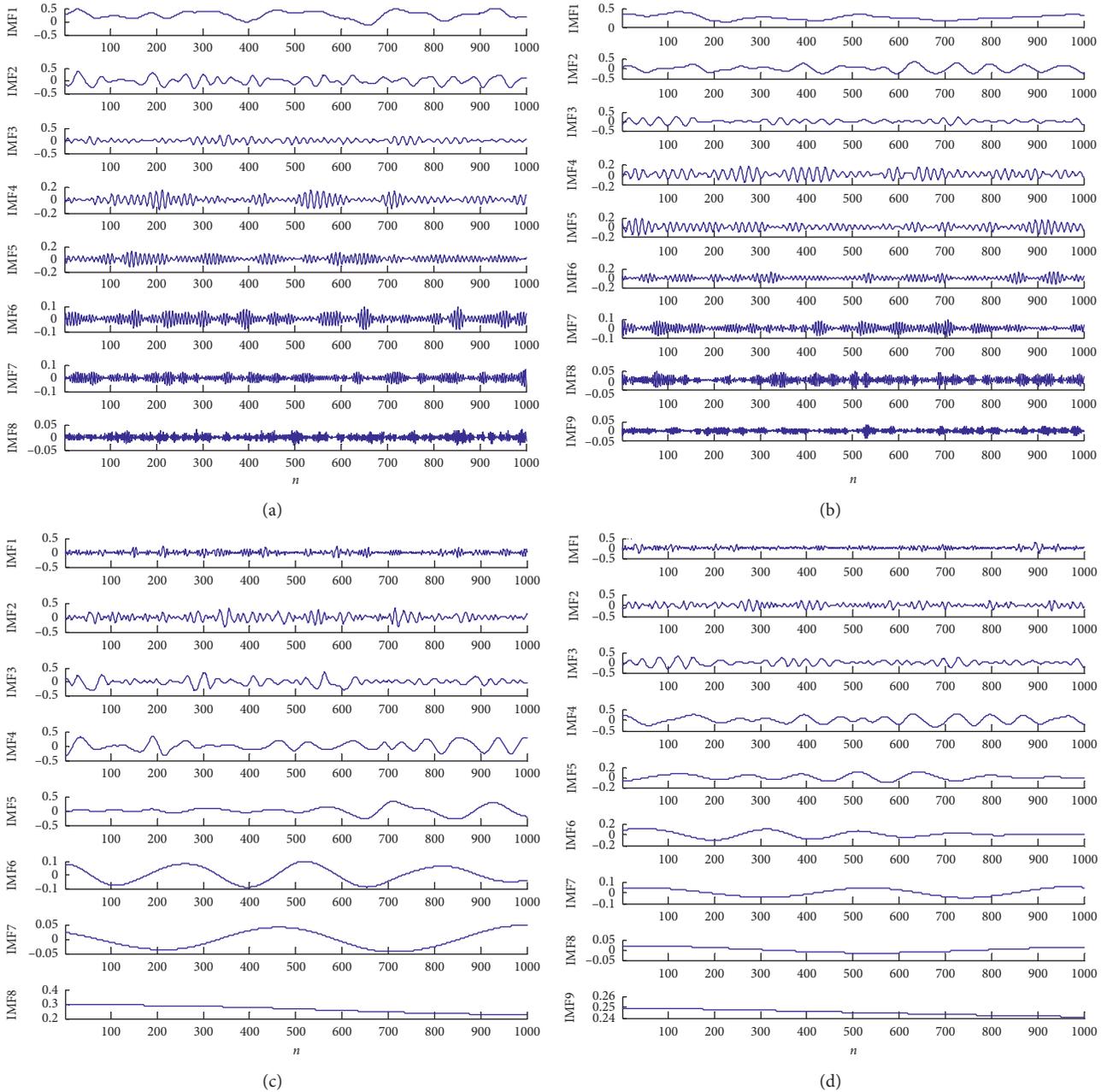


FIGURE 8: (a) VMD decomposition results of data set A; (b) VMD decomposition results of data set B; (c) EMD decomposition results of data set A; (d) EMD decomposition results of data set B.

and IMF6 are similar, so they can be combined. Compared with SE, DE avoids the problem of uncertain estimation of SE [50], so we choose DE to analyze the complexity of each sequence. In addition, in order to compare the effect of SE and DE, the index of orthogonality (IO) [51] of each IMF after reconstruction is calculated in this paper. IO describes the orthogonality of IMF. The larger the IO is, the more serious the mode mixing problem of the method is. The results show that the IO index of SE and DE is 0.0735 and 0.0718, respectively. It can be seen that the effect of DE is better.

For underwater acoustic signal data set B, after VMD, the entropy value of each component and the result of merging

are shown in Table 2. From Table 2, it can be seen that the entropy values of IMF4 and IMF5 are close, the entropy values of IMF8 and IMF9 are close, so the components with entropy values close are merged, and the rest components are treated as separate new sequences.

According to the combined results of mode components in Tables 1 and 2, the DE values of the first two mode components are less than 0.5, which is smaller than the other components. In addition, from the decomposition waveforms of VMD in Figures 8(a) and 8(b), it can be seen that the first two mode components have relatively small fluctuations and are in a gentle state as a whole. Therefore, the first two components can be regarded as low-

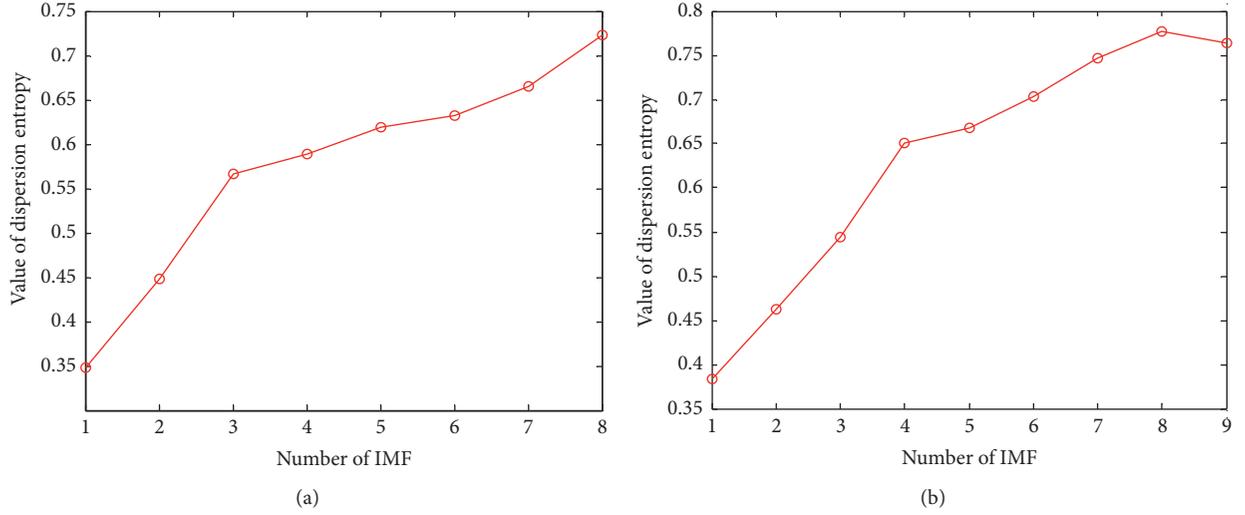


FIGURE 9: DE of each IMF sequences: (a) data set A; (b) data set B.

TABLE 1: Recombined by the DE value of VMD subsequence of data set A.

Sub	DE value	Recombination	New sub
IMF1	0.3476	IMF1	DE-IMF1
IMF2	0.4775	IMF2	DE-IMF2
IMF3	0.5668	IMF3&IMF4	DE-IMF3
IMF4	0.5842		
IMF5	0.6192	IMF5&IMF6	DE-IMF4
IMF6	0.6316		
IMF7	0.6648	IMF7	DE-IMF5
IMF8	0.7227	IMF8	DE-IMF6

TABLE 2: Recombined by the DE value of VMD subsequence of data set B.

Sub	DE value	Recombination	New sub
IMF1	0.3830	IMF1	DE-IMF1
IMF2	0.4627	IMF2	DE-IMF2
IMF3	0.5439	IMF3	DE-IMF3
IMF4	0.6508	IMF4&IMF5	DE-IMF4
IMF5	0.6681		
IMF6	0.7033	IMF6	DE-IMF5
IMF7	0.7471	IMF7	DE-IMF6
IMF8	0.7767	IMF8&IMF9	DE-IMF7
IMF9	0.7640		

TABLE 3: Recombined by the SE value of VMD subsequence of data set A.

Sub	SE value	Recombination	New sub
IMF1	0.1499	IMF1	SE-IMF1
IMF2	0.4962	IMF2	SE-IMF2
IMF3	0.6206	IMF3&IMF4&IMF6	SE-IMF3
IMF4	0.5921		
IMF5	0.5409	IMF5	SE-IMF4
IMF6	0.6054		
IMF7	0.8079	IMF7	SE-IMF5
IMF8	1.3062	IMF8	SE-IMF6

frequency components and the rest as high-frequency components. The high-frequency component has a large fluctuation, which often reveals the useful information hidden in the signal. The low-frequency component contains the characteristics of the signal itself, so accurate prediction is particularly important. The waveform of subsequence after the reorganization of underwater acoustic signal data sets A and B is shown in Figure 10.

4.3. Model Training and ABC Parameter Setting. The SVR model is a machine learning algorithm, which overcomes some shortcomings of traditional prediction methods. It has obvious advantages in solving nonlinear problems and strong generalization performance. In practical applications, the kernel function g and penalty parameter c directly affect the fitting accuracy and generalization performance of SVR. Therefore, in this paper, we choose the artificial with local and global search ability ABC algorithm to optimize the parameters of SVR. The fitness value is calculated by equation (20), and the MSE is used as the fitness evaluation function. The parameters of ABC algorithm are set as follows: the number of food sources (SN) is 20, the maximum cycle number (MCN) of food sources is set to 50, and the number of end cycles is 50.

After setting the initial parameters, the training data and test data are input into the ASVR model many times to achieve the best effect. The curve of fitness value and cycle times during training is shown in Figure 11. It can be seen from Figure 11 that the fitness value of ABC algorithm is decreasing with the increase in iteration times, which shows that the model gradually obtains the optimal parameters with the increase in iteration times. After several iterations, the fitness value gradually stabilizes to a fixed value. The results show that the model converges completely when the number of cycles reaches 50.

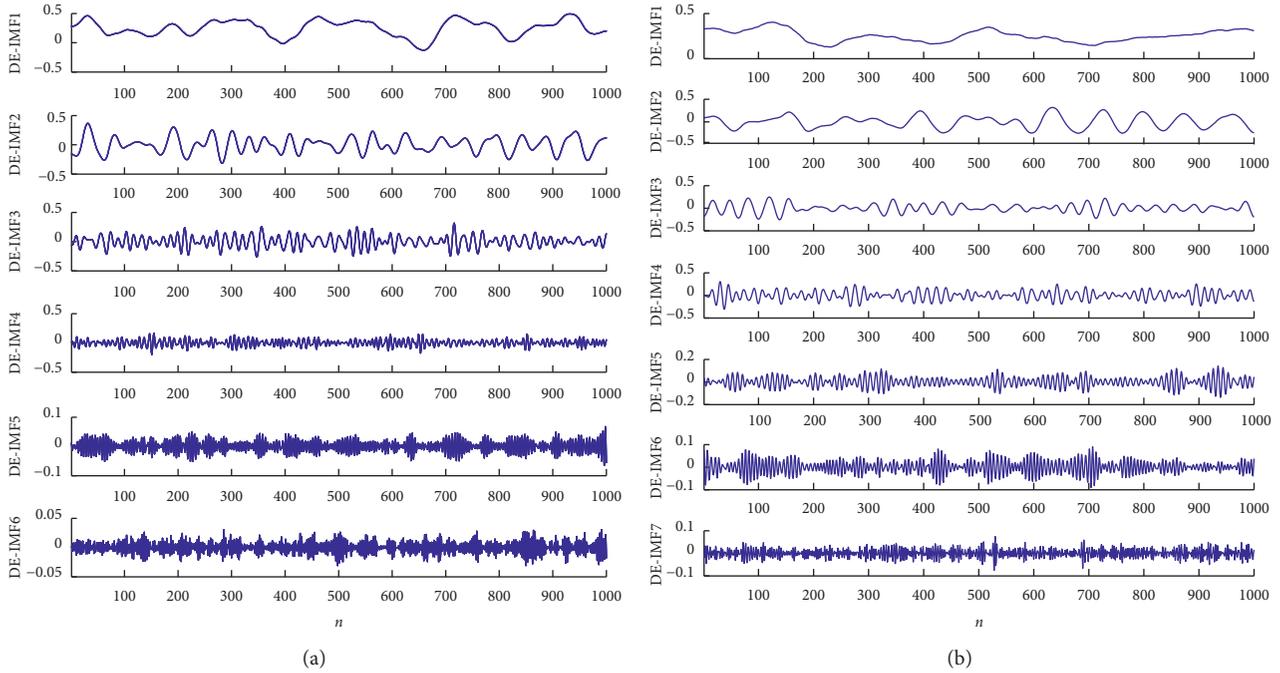


FIGURE 10: Recombination component processed by the VMD-DE method: (a) data set A; (b) data set B.

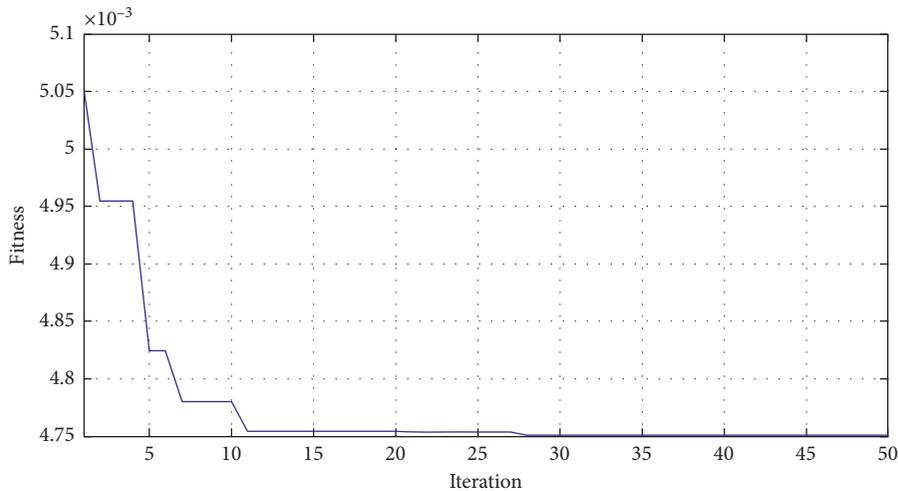


FIGURE 11: ABC-SVR optimal value convergence curve.

4.4. Comparison and Analysis of ELM and ASVR for Prediction of High- and Low-Frequency Series. In Section 4.2, according to the waveform of mode component and the corresponding DE value, it is divided into low-frequency component and high-frequency component. In order to better compare the prediction performance of ELM model and ASVR model corresponding to different frequency components, we take data set A as an example.

Each high-frequency mode component of data set A after VMD decomposition is superposed into a high-frequency sequence, and the prediction results of ELM model and ASVR model for high-frequency sequence are shown in Figure 4. It can be seen from the figure that the waveform of the high-frequency sequence changes greatly and is more

intensive. The red line in the figure represents the predicted value of the ELM model, the green line represents the predicted value of the ASVR model, and the blue line represents the original high-frequency sequence. It can be seen from Figure 4 that the green line and the blue line have a relatively high fitting degree. It also shows that the SVR model optimized by the ABC algorithm can effectively predict the highly nonlinear sequence with higher frequency by using the mapping of kernel function.

The low-frequency sequence often contains the characteristics of the original signal, and its waveform is closer to the real series. Similarly, we use the ELM model and ASVR model to predict the low-frequency series after the superposition of each low-frequency mode component. From

Figure 5, we can see that the prediction results of the two models for the low-frequency series are well. However, the fitting effect between the predicted value and the real data of the ELM model represented by the red line is better than that of the ASVR model, which also reflects that ELM has better prediction effect than ASVR for low-frequency series prediction of underwater acoustic signal data.

Through the above comparative analysis, this paper uses ELM model to predict the low-frequency sequence of underwater acoustic signal and ASVR model to predict the high-frequency sequence of underwater acoustic signal. The prediction results of ASVR for each component of data set A after VMD-DE decomposition are shown in Figure 12(a). The red line represents the real value of underwater acoustic signal, and the blue line represents the predicted value of ASVR. The prediction results of the combined prediction model proposed in this paper are shown in Figure 12(b). It can be clearly seen from Figure 12 that the prediction effect of the first two low-frequency components has been significantly improved, and the overall prediction effect is better.

4.5. Prediction Results of Each Model for Data Set A. In order to further measure the prediction effect of this method, the VMD-DE-ELM-ASVR model proposed in this paper is compared with the other six models, and the prediction effect of each model is quantitatively analyzed by MAE, RMSE, and R^2 , so as to verify the superiority of the combined model in the prediction performance of this paper.

Data A are divided into six relatively stable mode components after VMD-DE decomposition. The prediction model is established for each component, and the prediction results are shown in Figure 12. It can be seen that the predicted value of each mode component fits well with the original signal value. The prediction value of the model proposed in this paper is obtained by superposing the prediction results of each component, as shown in Figure 13, where the red line represents the real data and the blue line represents the prediction value. It can be seen that the prediction value of the model proposed in this paper and the original data fit well. Among them, RMSE, MAE, and R^2 are also listed in Figure 13. It can be seen that the error indexes are relatively small, and R^2 is close to 1, which further explains the effectiveness of the prediction model proposed in this paper.

The fitting curves of predicted values and original values of different prediction models are shown in Figure 14. In the figure, the predicted value of the model proposed in this paper is represented by blue curve, and the original data curve is represented by red line. It can be seen that the blue curve has the highest fitting degree compared with other curves. In order to further reveal the superiority of the model in this paper, the box diagram of prediction error of the seven models is shown in Figure 15(a). The prediction error of VMD-DE-ELM-ASVR is distributed around 0, and the change amplitude is the smallest compared with other models. It shows that the combined forecasting model proposed in this paper has a good prediction effect.

It can be seen from Figure 14 and Table 4 that the prediction accuracy of SVR is low because the best model parameters cannot be obtained. However, the parameters of SVR model optimized by ABC algorithm are all optimized, so the prediction performance has been greatly improved. By comparing EMD-ASVR and ASVR, it can be seen that the difficulty in prediction is reduced after mode decomposition. However, due to the mode mixing and endpoint effect of EMD, the prediction performance of the EMD-ASVR model is lower than that of the VMD-ASVR model.

The values of MAE, RMSE, and R^2 of all models are shown in Table 4. In addition, Figure 15(b) clearly shows the corresponding MAE and RMSE sizes of different models in the form of histogram. It can be seen from Table 4 that the prediction indexes of the prediction model proposed in this paper are minimum, and R^2 is the closest to 1, indicating that the proposed model has the best prediction accuracy.

By comparing the prediction results of SVR and ASVR model, it can be seen from Table 4 that the prediction accuracy of the SVR model is the lowest, and the R^2 of predicted value and real value is 0.9147, while the correlation coefficient of other models is above 0.96. However, the prediction accuracy of the SVR model optimized by the ABC algorithm has been significantly improved. MAE and RMSE are the mean absolute error and root mean square error of predicted value and actual value. Therefore, the smaller the index is, the better the result is. The RMSE value of the optimized SVR model is lower than that of the ASVR model. This shows that the ABC algorithm relies on its strong global and local search ability to find the optimal penalty parameter c and kernel function g and obtains satisfactory prediction results.

Both EMD-ASVR and VMD-ASVR models have improved the prediction accuracy compared with a single ASVR model. It shows that the prediction model with decomposition method can improve prediction accuracy and reduce the prediction error. However, by comparing the results in Table 4, it can be seen that the RMSE and MAE values of VMD-ASVR are 0.0140 and 0.0168, respectively, while the RMSE and MAE values of the prediction model added with EMD decomposition algorithm are 0.0298 and 0.0369, respectively, and the values of these two indexes and R^2 are also significantly worse than VMD decomposition. It also shows that the end effect of EMD decomposition and mode mixing will directly affect the prediction results. In addition, the complexity of each mode component is analyzed by DE to reconstruct the subsequence, which improves the prediction accuracy and shortens the training time of the model.

From Table 4, it can be seen that MAE and RMSE of VMD-DE-ELM-ASVR combination model are 0.0102 and 0.0129, respectively, which are lower than VMD-DE-ASVR and R^2 is 0.9977. It shows that the combination prediction model proposed in this paper has the highest fitting degree with the original data.

4.6. Prediction Results of Each Model for Data Set B. In order to further verify the prediction effect of the VMD-DE-ELM-ASVR model in different underwater acoustic signal

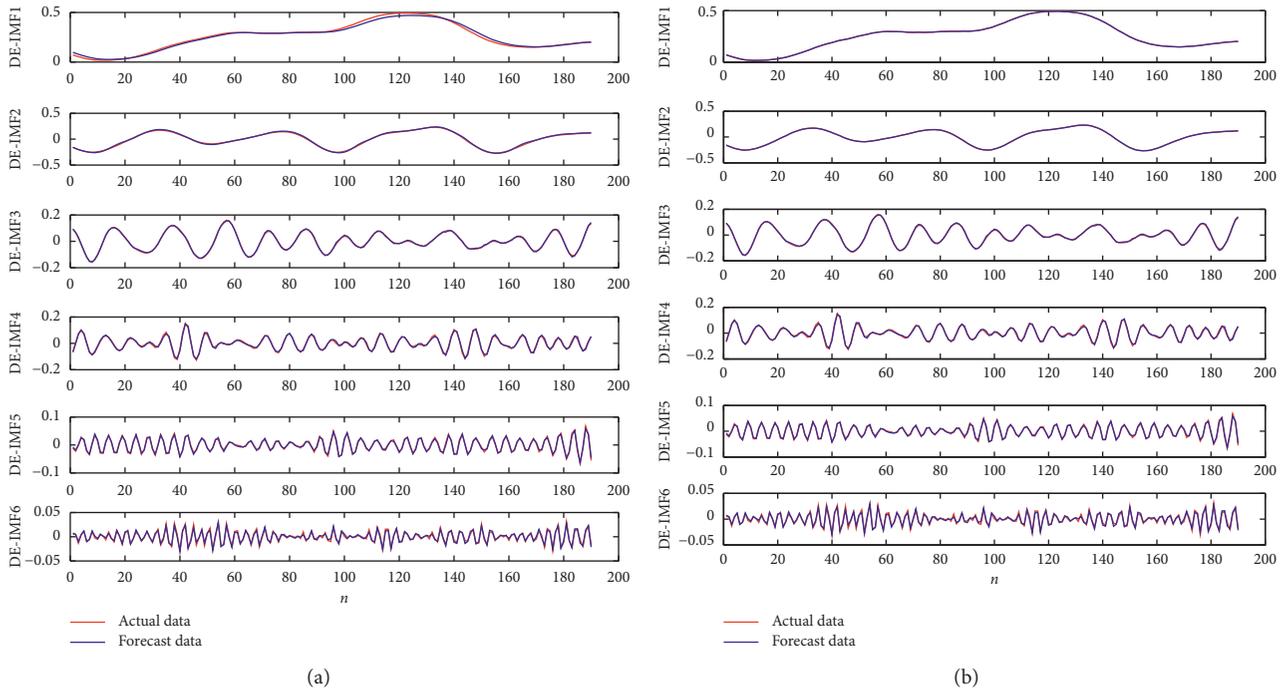


FIGURE 12: (a) VMD-DE-ASVR prediction results of each component; (b) VMD-DE-ELM-ASVR prediction results of each component.

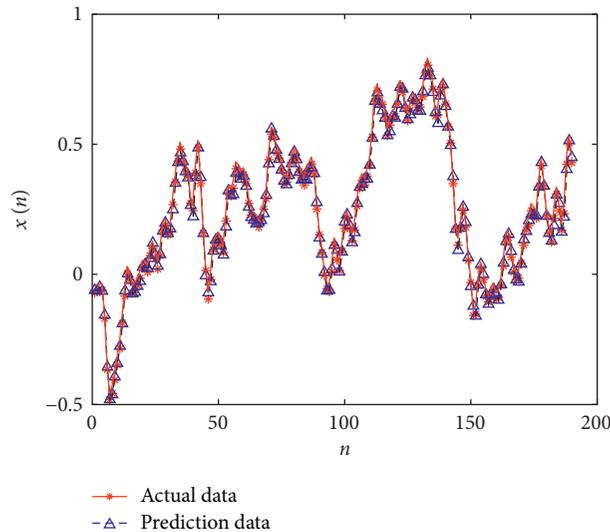


FIGURE 13: Prediction results of VMD-DE-ELM-ASVR. MAE = 0.010158; RMSE = 0.012886; $R^2 = 0.99773$.

sequences, the underwater acoustic signal data set B is selected for discussion and analysis in this section. The prediction results are shown in Figure 16. By comparing the fitting curves of the predicted values and the original values of different prediction models, we can see that the combined prediction model proposed in this paper has the best fitting effect with the real underwater acoustic signal data.

In addition, Figure 17(a) shows the error box of each model. For data B, the prediction error of the VMD-DE-ELM-ASVR model is distributed around 0, with the smallest change compared with other models. Furthermore, the

combination prediction model proposed in this paper has a better prediction effect. In order to quantitatively analyze the differences among the models, the values of MAE, RMSE, and R^2 are shown in Table 5. In addition, Figure 17(b) clearly shows the values of MAE and RMSE corresponding to different models in the form of histogram. From Table 5, it can be seen that the error index of the combined prediction proposed in this paper reaches the minimum, and R^2 reaches 0.9966, which is similar to the prediction performance of the underwater acoustic signal in data A.

Generally speaking, from the prediction effect of the underwater acoustic signal data sets A and B, the combined

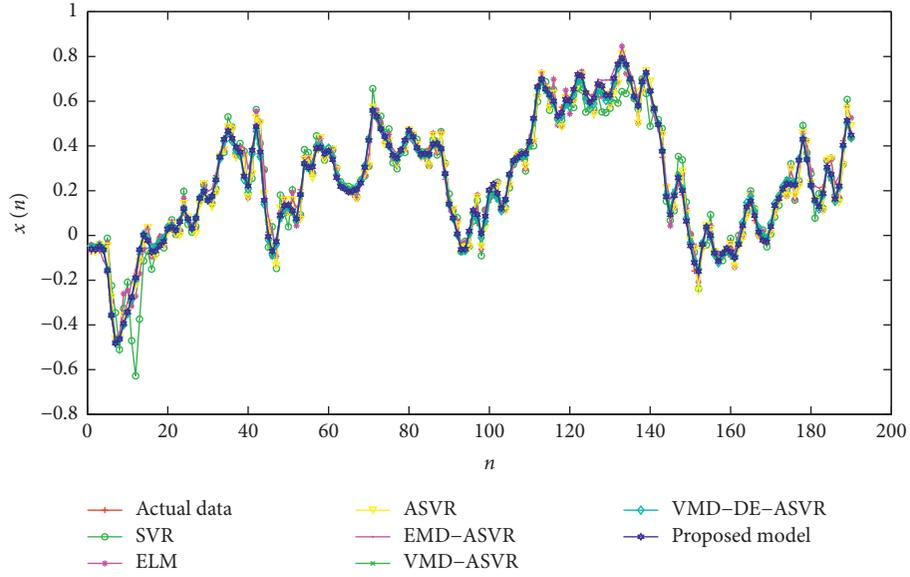


FIGURE 14: Prediction results for each model of data set A.

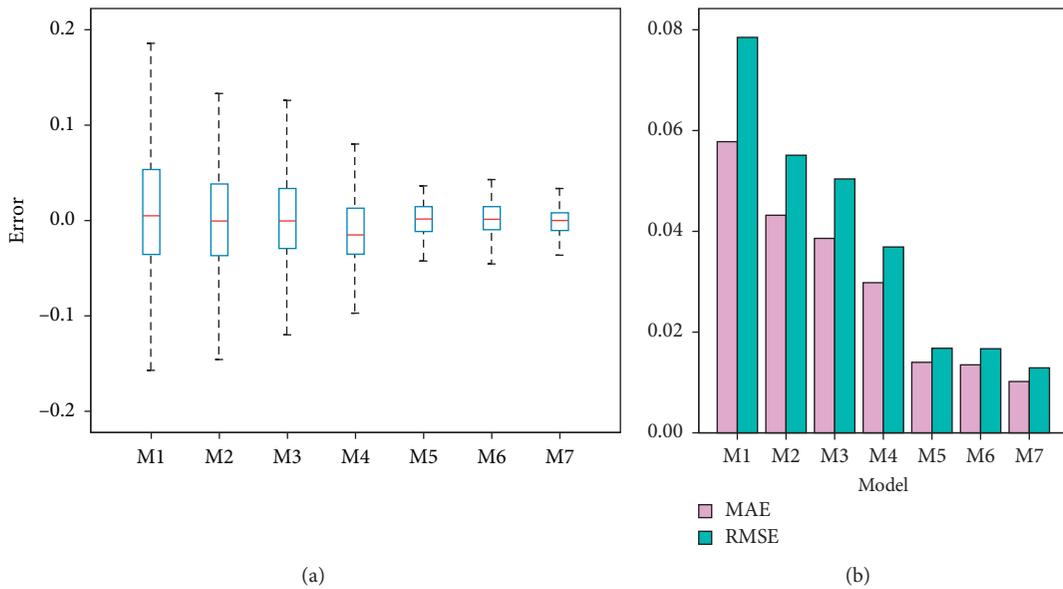


FIGURE 15: (a) Error box diagram of each model; (b) bar chart of model error index. M1: SVR; M2: ELM; M3: ASVR; M4: EMD-ASVR; M5: VMD-ASVR; M6: VMD-DE-ASVR; M7: VMD-DE-ELM-ASVR.

TABLE 4: The comparison result of different model performances for data set A.

Model	MAE	RMSE	R^2	T (s)
SVR	0.0578	0.0785	0.9147	7
ELM	0.0432	0.0551	0.9602	6
ASVR	0.0386	0.0504	0.9641	154
EMD-ASVR	0.0298	0.0369	0.9832	456
VMD-ASVR	0.0140	0.0168	0.9966	460
VMD-DE-ASVR	0.0135	0.0167	0.9967	374
VMD-DE-ELM-ASVR	0.0102	0.0129	0.9977	334

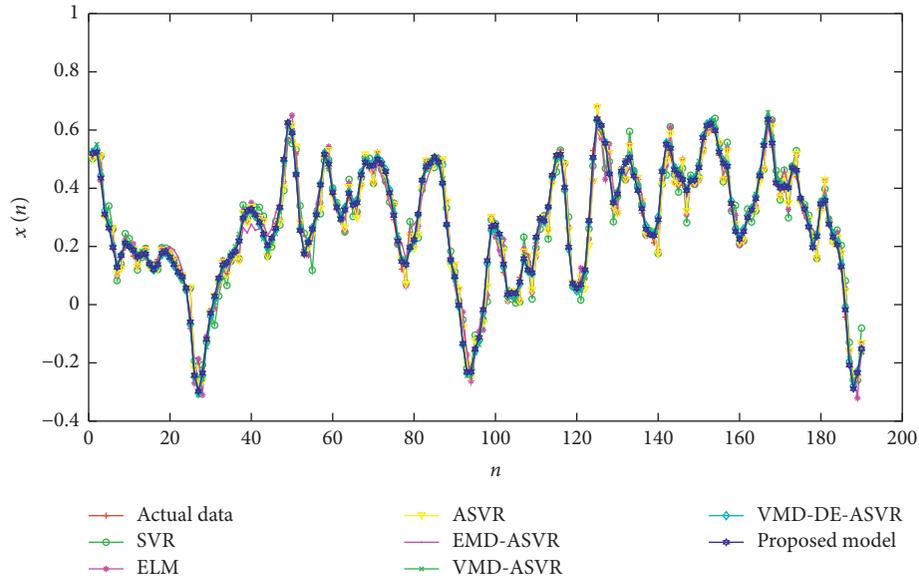


FIGURE 16: Prediction results for each model of data set B.

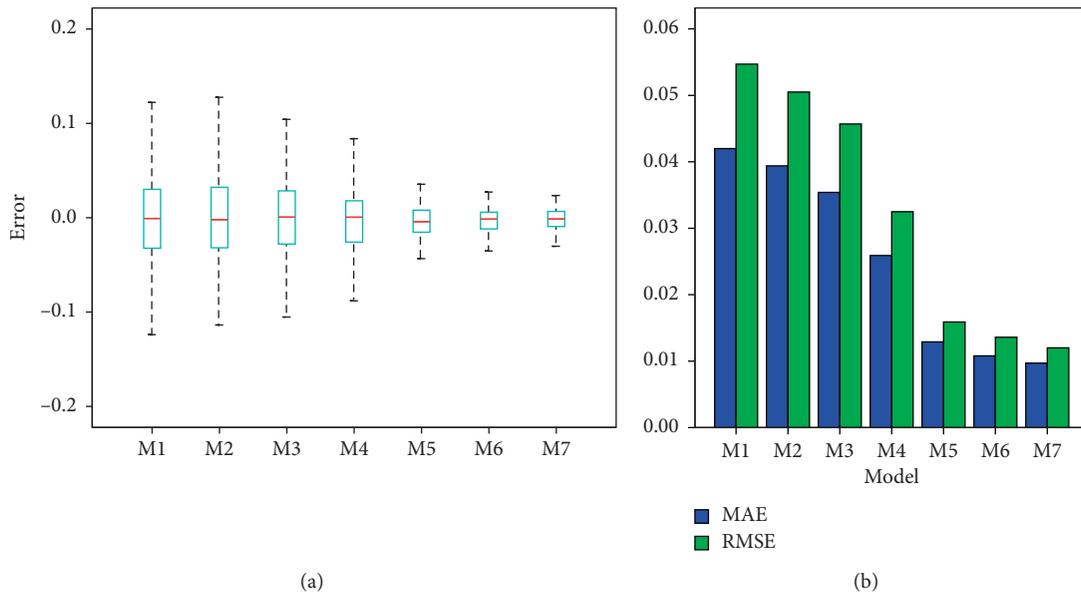


FIGURE 17: (a) Error box diagram of each model; (b) bar chart of model error index. M1: SVR; M2: ELM; M3: ASVR; M4: EMD-ASVR; M5: VMD-ASVR; M6: VMD-DE-ASVR; M7: VMD-DE-ELM-ASVR.

TABLE 5: The comparison result of different model performances for data set B.

Model	MAE	RMSE	R^2	T (s)
SVR	0.0420	0.0547	0.9285	10
ELM	0.0394	0.0505	0.9394	8
ASVR	0.0354	0.0457	0.9498	154
EMD-ASVR	0.0259	0.0325	0.9746	470
VMD-ASVR	0.0129	0.0159	0.9947	465
VMD-DE-ASVR	0.0108	0.0136	0.9961	396
VMD-DE-ELM-ASVR	0.0097	0.0120	0.9966	316

prediction model proposed in this paper shows better prediction performance, which can provide a reference for prediction of underwater acoustic signal.

5. Conclusions

In order to improve the prediction accuracy of underwater acoustic signal, a combined prediction model based on VMD-DE-ELM-ASVR is proposed and applied to the prediction of underwater acoustic signal. The main conclusions are as follows:

- (1) The VMD decomposition algorithm can effectively overcome the mode mixing of EMD. The simulation results show that the decomposition effect of VMD is clearer, and the prediction accuracy is higher.
- (2) In this paper, DE is used to calculate the entropy of the IMFs of VMD decomposition, and the components with DE approximation are merged and recombined. The simulation results show that eight IMF components obtained by VMD decomposition in data set A are combined into six IMF components, and nine mode components are combined into seven by DE in data set B. In this way, the complexity of calculation is effectively reduced, and the prediction accuracy is improved.
- (3) ABC algorithm has few parameter settings and can be used for global and local search. The results show that the optimal penalty parameter c and kernel function g can be found by finite iterations of SVR model after ABC optimization, which improves the prediction accuracy.
- (4) In this paper, ELM model is selected to predict the low-frequency component of underwater acoustic signal, and ASVR is used to predict the high-frequency component of underwater acoustic signal. It is proved by experiments that the combined prediction model can improve the prediction accuracy and reduce the prediction error compared with the single prediction model.
- (5) The prediction method proposed in this paper is tested by actual underwater acoustic signal data sets A and B and compares seven kinds of prediction models with three statistical indicators, including SVR, ELM, ASVR, EMD-ASVR, VMD-ASVR, VMD-DE-ASVR, and VMD-DE-ELM-ASVR. The experimental results show that VMD-DE-ELM-ASVR can effectively predict underwater acoustic signal. Compared with other models, the combined model improves the prediction accuracy, reduces the error, and has strong generalization ability and robustness.

Nomenclature

VMD:	Variational mode decomposition
EMD:	Empirical mode decomposition
ELM:	Extreme learning machine
SVR:	Support vector regression
SRM:	Structural risk minimization
ASVR:	Optimized SVR based on ABC algorithm
R^2 :	Coefficient of determination
GSA:	Gravitational search algorithm
PSO:	Particle swarm optimization
SE:	Sample entropy
GA:	Genetic algorithm
DE:	Dispersion entropy
IMF:	Intrinsic mode function
ABC:	Artificial bee colony
SVM:	Support vector machine

RMSE:	Root mean squared error
MSE:	Mean square error
MCN:	Maximum cycle number
IO:	Index of orthogonality.

Data Availability

The data used to support the findings of this study are currently under embargo, while the research findings are commercialized. Requests for data, 12 months after publication of this article, will be considered by the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (no. 51709228).

References

- [1] Q. H. Li, "Advances of research work in some areas of underwater acoustics signal processing," *Applied Acoustics*, vol. 20, no. 1, pp. 1–5, 2001.
- [2] R. J. Vaccaro, "The past, present, and the future of underwater acoustic signal processing," *IEEE Signal Processing Magazine*, vol. 15, no. 4, pp. 21–51, 1998.
- [3] M. Taroudakis, C. Smaragdakis, and N. R. Chapman, "Denoising underwater acoustic signal for applications in acoustical oceanography," *Journal of Computational Acoustics*, vol. 25, no. 2, Article ID 1750015, 2017.
- [4] G. H. Li, Q. R. Guan, and H. Yang, "Noise reduction method of underwater acoustic signal based on CEEMDAN, effort-to-compress complexity, refined composite multiscale dispersion entropy and wavelet threshold denoising," *Entropy*, vol. 21, no. 1, p. 11, 2018.
- [5] M. I. Taroudakis, "Statistical characterization of acoustic signal using 1D wavelet transforms with applications in acoustical oceanography," *Journal of Theoretical and Computational Acoustics*, vol. 26, no. 4, Article ID 1850047, 2018.
- [6] M. I. Taroudakis and C. Smaragdakis, "Tomographic and bottom geoacoustic inversions using genetic algorithms and a statistical characterization of the acoustic signal," *Acta Acustica United with Acustica*, vol. 95, no. 5, pp. 814–822, 2009.
- [7] Y.-X. Li, Y.-A. Li, Z. Chen, and X. Chen, "Feature extraction of ship-radiated noise based on permutation entropy of the intrinsic mode function with the highest energy," *Entropy*, vol. 18, no. 11, p. 393, 2016.
- [8] X. H. Zhang, X. M. Zhang, and L. J. Lin, "Researches on chaotic phenomena of noises radiated from ships," *Acta Acustica*, vol. 23, no. 2, pp. 134–140, 1998.
- [9] Y. A. Li, D. M. Xu, and X. M. Zhang, "A new way for detecting and recognizing underwater ships using takens theorem," *Journal of Northwestern Polytechnical University*, vol. 19, no. 2, pp. 266–269, 2001.
- [10] Y. Liang, Q. Meng, and J. R. Lu, "Extraction of nonlinear features from underwater acoustic signal based on joint

- prediction results," *Technical Acoustics*, vol. 26, no. 4, pp. 602–605, 2007.
- [11] H. C. Wang, "The research of detection of underwater acoustic signal based on the chaotic prediction theory," M.S. thesis, Northwestern Polytechnical University, Xi'an, China, 2007.
 - [12] J. Zhou, "Weak signal detection on underwater acoustic based on the prediction of chaos," M.S. thesis, National University of Defense Technology, Changsha, China, 2009.
 - [13] Y. Y. Fang, Y. A. Li, L. Cui, and J. Shang, "Research on prediction of underwater acoustic signal based on volterra adaptive filter," *Acta Armamentarii*, vol. 34, no. 9, pp. 1173–1179, 2013.
 - [14] Y. Sun, Y. A. Li, X. Chen, S. Dai, and W. J. Gao, "Nonlinear modeling and prediction of underwater acoustic signal," *Torpedo Technology*, vol. 22, no. 5, pp. 341–346, 2014.
 - [15] L. Liu, "Nonlinear prediction of sea clutter using radial neural network and volterra series filter," *Ship Science and Technology*, vol. 39, no. 2A, pp. 52–54, 2017.
 - [16] J. N. Zhou, Y. A. Li, and Y. S. Wu, "Prediction of underwater acoustic signal based on neural network," *Technical Acoustics*, vol. 25, no. 3, pp. 226–229, 2006.
 - [17] Y. Y. He and H. D. Zhang, "Modeling and prediction of underwater acoustic signal based on PSO and RBF neural network," *Computer Engineering*, vol. 34, no. 23, pp. 208–209, 2008.
 - [18] H. Yang, S. L. Wang, G. H. Li, and T. T. Mao, "A new hybrid model based on fruit fly optimization algorithm and wavelet neural network and its application to underwater acoustic signal prediction," *Mathematical Problems in Engineering*, vol. 2018, Article ID 3136267, 8 pages, 2018.
 - [19] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.
 - [20] J. Zhu, T. X. Tan, L. F. Wu, and H. M. Yuan, "RUL prediction of lithium-ion battery based on improved DGWO-ELM method in a random discharge rates environment," *IEEE Access*, vol. 7, pp. 125176–125187, 2019.
 - [21] G. H. Li, S. L. Wang, and H. Yang, "Prediction of underwater acoustic signal based on ESMD and ELM," *Indian Journal of Geo Marine Sciences*, vol. 48, no. 3, pp. 357–362, 2019.
 - [22] S. Mukherjee, E. Osuna, and F. Girosi, "Nonlinear prediction of chaotic time series using support vector machines," in *Proceedings of the 1997 IEEE Workshop on Neural Networks for Signal Processing VII*, pp. 511–520, Amelia Island, FL, USA, 1997.
 - [23] S. Samadianfard, S. Jarhan, E. Salwana, A. Mosavi, S. Shamshirband, and S. Akib, "Support vector regression integrated with fruit fly optimization algorithm for river flow forecasting in lake Urmia Basin," *Water*, vol. 11, no. 9, p. 1934, 2019.
 - [24] W. Yang, J. Wang, and R. Wang, "Research and application of a novel hybrid model based on data selection and artificial intelligence algorithm for short term load forecasting," *Entropy*, vol. 19, no. 2, p. 52, 2017.
 - [25] X. Y. Tang, L. Wang, J. R. Cheng, J. Chen, and V. S. Sheng, "Forecasting model based on information-granulated GA-SVR and ARIMA for producer price index," *Computers Materials & Continua*, vol. 58, no. 2, pp. 436–491, 2019.
 - [26] L. Zou, Y. Sun, and X. Yang, "An entropy-based neighborhood rough set and PSO-SVRM model for fatigue life prediction of titanium alloy welded joints," *Entropy*, vol. 21, no. 2, p. 117, 2019.
 - [27] S. Wang, H. Hae, and J. Kim, "Development of easily accessible electricity consumption model using open data and GA-SVR," *Energies*, vol. 11, no. 2, p. 373, 2018.
 - [28] H. Garg, "A hybrid GSA-GA algorithm for constrained optimization problems," *Information Sciences*, vol. 478, pp. 499–523, 2019.
 - [29] H. Garg, "A hybrid PSO-GA algorithm for constrained optimization problems," *Applied Mathematics and Computation*, vol. 274, pp. 292–305, 2016.
 - [30] D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm," *Journal of Global Optimization*, vol. 39, no. 3, pp. 459–471, 2007.
 - [31] H. Garg, "Solving structural engineering design optimization problems using an artificial bee colony algorithm," *Journal of Industrial & Management Optimization*, vol. 10, no. 3, pp. 777–794, 2014.
 - [32] D. Yang, Y. Liu, S. Li, X. Li, and L. Ma, "Gear fault diagnosis based on support vector machine optimized by artificial bee colony algorithm," *Mechanism and Machine Theory*, vol. 90, pp. 219–229, 2015.
 - [33] A. Hossen, F. Al-Wadahi, and J. A. Jervase, "Classification of modulation signal using statistical signal characterization and artificial neural network," *Engineering Applications of Artificial Intelligence*, vol. 20, no. 4, pp. 463–472, 2007.
 - [34] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
 - [35] Y. H. Jiang, W. D. Jiao, and R. Q. Li, "A study on the method for eliminating mode mixing in B-spline empirical mode decomposition based on adaptive bandwidth constrained signal," *Journal of Vibration and Shock*, vol. 37, no. 16, pp. 83–90, 2018.
 - [36] Z. P. He, Z. Q. Zhu, H. C. Xie et al., "Restrained boundary effect of EMD based on least square fitting," *Journal of System Simulation*, vol. 30, no. 9, pp. 3377–3385, 2018.
 - [37] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.
 - [38] 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," *Mechanical Systems and Signal Processing*, vol. 60–61, pp. 243–251, 2015.
 - [39] 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 Computer Science*, vol. 46, pp. 372–380, 2015.
 - [40] M. Ali, A. Khan, and N. Rehman, "Hybrid multiscale wind speed forecasting based on variational mode decomposition," *International Transactions on Electrical Energy Systems*, vol. 28, no. 1, Article ID e2466, 2017.
 - [41] Q. Wu and H. Lin, "A novel optimal-hybrid model for daily air quality index prediction considering air pollutant factors," *Science of the Total Environment*, vol. 683, pp. 808–821, 2019.
 - [42] G. 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," *Computational Intelligence and Neuroscience*, vol. 2018, Article ID 3713410, 9 pages, 2018.
 - [43] H. Yang, L. Gao, and G. Li, "Underwater acoustic signal prediction based on correlation variational mode decomposition and error compensation," *IEEE Access*, vol. 8, pp. 103941–103955, 2020.

- [44] M. Rostaghi and H. Azami, "Dispersion entropy: a measure for time-series analysis," *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 610–614, 2016.
- [45] V. N. Vapnik, "An overview of statistical learning theory," *IEEE Transactions on Neural Networks*, vol. 10, no. 5, pp. 988–999, 1999.
- [46] D. J. Armaghani, P. G. Asteris, B. Askarian, M. Hasanipanah, R. Tarinejad, and V. Van Huynh, "Examining hybrid and single SVM models with different kernels to predict rock brittleness," *Sustainability*, vol. 12, no. 6, p. 2229, 2020.
- [47] D. Karaboga and B. Basturk, "On the performance of artificial bee colony (ABC) algorithm," *Applied Soft Computing*, vol. 8, no. 1, pp. 687–697, 2008.
- [48] H. Yang, Y. A. Li, and G. H. Li, "Energy analysis of ship radiated noise based on ensemble empirical mode decomposition," *Journal of Vibration and Shock*, vol. 34, no. 16, pp. 55–59, 2015.
- [49] C. L. Liu, Y. J. Wu, and C. G. Zhen, "Rolling bearing fault diagnosis based on variational mode decomposition and fuzzy C means clustering," *Proceedings of the CSEE*, vol. 35, no. 13, pp. 3358–3365, 2015.
- [50] W. Cao, C. L. Liu, Z. Q. Wang, and H. J. Li, "Wind power ultra-short-term prediction based on MEEMD and permutation entropy," *Renewable Energy Resources*, vol. 37, no. 3, pp. 439–444, 2019.
- [51] H. Kuang, R. L. Wang, J. Zhang, and J. Yan, "Study on suppression of EMD end effect based on SVM," *Computer Engineering and Applications*, vol. 51, no. 11, pp. 196–200, 2015.