

## Research Article

# Underwater Acoustic Signal Prediction Based on MVMD and Optimized Kernel Extreme Learning Machine

Hong Yang , Lipeng Gao, and Guohui Li 

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

Correspondence should be addressed to Hong Yang; [uestcyhong@163.com](mailto:uestcyhong@163.com) and Guohui Li; [lghcd@163.com](mailto:lghcd@163.com)

Received 13 November 2019; Revised 2 March 2020; Accepted 27 March 2020; Published 24 April 2020

Academic Editor: Bernhard C. Geiger

Copyright © 2020 Hong Yang 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.

Aiming at the chaotic characteristics of underwater acoustic signal, a prediction model of grey wolf-optimized kernel extreme learning machine (OKELM) based on MVMD is proposed in this paper for short-term prediction of underwater acoustic signals. To solve the problem of  $K$  value selection in variational mode decomposition, a new  $K$  value selection method MVMD is proposed from the perspective of mutual information, which avoids the blindness of variational mode decomposition (VMD) in the preset modal number. Based on the prediction model of kernel extreme learning machine (KELM), this paper uses grey wolf optimization (GWO) algorithm to optimize and select its regularization parameters and kernel parameters and proposes an optimized kernel extreme learning machine OKELM. To further improve the prediction performance of the model, combined with MVMD, an underwater acoustic signal prediction model based on MVMD-OKELM is established. MVMD-OKELM prediction model is applied to Mackey–Glass chaotic time series prediction and underwater acoustic signal prediction and is compared with ARIMA, EMD-OKELM, and other prediction models. The experimental results show that the proposed MVMD-OKELM prediction model has a higher prediction accuracy and can be effectively applied to the prediction of underwater acoustic signal series.

## 1. Introduction

Underwater acoustic signal prediction is the basis of underwater acoustic signal processing, which can be applied to many aspects such as underwater target signal reduction, detection, and feature extraction [1–3]. The underwater acoustic signal has not only nonlinear, non-Gaussian, and nonstationary characteristics but also typical chaotic, fractal, and other characteristics [4, 5], which provide a basis for short-term prediction of underwater acoustic signal.

In recent years, with continuous development of ship noise reduction technology, the detection of underwater target signal is more and more difficult. If a more accurate prediction model is adopted, the lower signal-to-noise ratio signal can be detected [6]. Therefore, the study of the prediction problem of underwater acoustic signal is of great significance in the processing of underwater acoustic signal. In traditional forecasting models [7, 8], autoregressive and moving average model (ARMA) and autoregressive integrated moving average model (ARIMA) models have been

widely used in time series prediction. However, ARMA and ARIMA can only capture the linear relationship of the signal in nature, but not the nonlinear relationship of the signal, which will have certain limitations on the time series prediction. Artificial intelligence algorithms developed in recent years, such as BP, RBF, WNN, and ESN, are suitable for capturing the nonlinear relationship of signals and are gradually beginning to be applied in time series prediction. Zhou et al. [9] used the combination of particle swarm optimization (PSO) and RBF neural network to predict the underwater acoustic signal and verified the short-term predictability of chaotic time series. He and Zhang et al. [10] applied the BP neural network and RBF neural network to the prediction of underwater acoustic signal and achieved good prediction results. Yang et al. [11] used the fruit fly algorithm to optimize the wavelet neural network, which further improved the prediction accuracy of underwater acoustic signal. Chitsazan et al. [12] improved the classical echo state network and applied it to the prediction of wind speed and direction, with good prediction results. Although

this prediction method relying solely on the neural network model has achieved good prediction results, it does not consider importance of data preprocessing. In recent years, the decomposition technology in data preprocessing has attracted the attention of researchers, and some achievements have been made in time series prediction [13–19]. Li et al. [13] proposed a chaotic time series prediction model of monthly precipitation based on the combination of variational mode decomposition and extreme learning machine (ELM). The model can predict the precipitation trend and improve the prediction accuracy. Xiong et al. [14] used a combination of wavelet decomposition and LSSVM to achieve short-term prediction of wind speed. Büyükşahin and Ertekin [15] proposed a hybrid prediction method combining ARIMA and artificial neural network (ANN) and added empirical mode decomposition technology to further improve the prediction accuracy of time series. Li et al. [16] proposed a deep learning prediction model based on extreme-point symmetric mode decomposition and cluster analysis to predict monthly mean value of sunspots and have a good prediction effect. Cheng et al. [17] used ensemble empirical mode decomposition and LSSVM to achieve short-term prediction of wind power and verified that this prediction method has better prediction accuracy than EMD and LSSVM methods. Nazir et al. [18] proposed an improved CEEMDAN decomposition method and applied it to the prediction of river flow, and achieved good prediction results. Li et al. [19] proposed an underwater acoustic signal prediction method based on ESMD and ELM, which further improved the prediction accuracy based on ELM. To sum up, the combination of decomposition methods and various prediction models can effectively improve the prediction accuracy of time series. However, this kind of report on the prediction of underwater acoustic signal using decomposition prediction idea is rare, so it is necessary to carry out further research in this field.

In time series prediction, if an appropriate decomposition method can be selected, the original complex signal can be decomposed into multiple stationary component signals, so as to optimize the prediction model and improve the prediction accuracy. The VMD decomposition method proposed by Dragomiretskiy and Zosso [20] is an effective decomposition method. The method is suitable for decomposing nonlinear and nonstationary signals, and to a certain extent, it eliminates the modal aliasing phenomenon of decomposition methods such as EMD and EEMD and has been applied in various time series predictions [13, 21, 22]. However, the decomposition mode number  $K$  needs to be set in advance before VMD decomposition signal. Too large value will lead to over decomposition, while too small value will lead to insufficient decomposition. Therefore, the selection of  $K$  value plays an important role in the decomposition of the VMD. The number of VMD decomposition layers is usually determined by the observation center frequency [23, 24] or by the EMD decomposition layer [25, 26]. The observation of the center frequency is subject to subjective factors, and the EMD decomposition principle is far

from the decomposition principle of VMD, so it is difficult to be persuasive. In recent years, Sun et al. [27] proposed a  $K$  value determination method based on the correlation coefficient, which determines the final modal number by judging the correlation coefficient between the decomposition mode and the original signal. Wang et al. [28] proposed a method for determining  $K$  value based on the center frequency, which determines the optimal number of modes by judging the ratio of the center frequencies between before and after modes. Wang et al. [29] proposed a method for determining  $K$  value based on energy conservation. The method determines the optimal  $K$  value by judging the energy magnitude between sum of each component and original signal energy. In this paper, a  $K$  value selection method based on mutual information is proposed, which is to determine the optimal  $K$  value by judging the change in mutual information between the reconstructed series and the original series.

Based on the idea of decomposition prediction, this paper proposes an MVMD-OKELM prediction model. Firstly, the underwater acoustic signal is decomposed into a series of relatively stable components by MVMD. Then, a kernel extreme learning machine (KELM) prediction model is established for each component. The parameters of the KELM model are selected by grey wolf optimization (GWO). Finally, the prediction results of all components are reconstructed to obtain the final prediction results. To fully illustrate the effectiveness of MVMD-OKELM in time series prediction, this paper applies the proposed MVMD-OKELM prediction model to Mackey–Glass chaotic time series and underwater acoustic signal, respectively, and compares it with ARIMA, EMD-OKELM, and other prediction models.

## 2. Basic Theory

**2.1. Variational Mode Decomposition.** VMD can decompose the input signal into subsequences with different center frequencies and limited bandwidth  $\{u_k(t)\}$  ( $k = 1, 2, \dots, K$ ). The decomposition process is a process of solving the variational problem. The constrained variational model is as follows [20]:

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

where  $\omega_k$  represents the center frequency of the  $k$ th modal function.

In order to solve the constrained variational problem (1), a penalty factor  $C$  and a Lagrangian multiplication operator  $\theta(t)$  are introduced to transform the constrained variational problem into an unconstrained variational problem:

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

where  $\|f(t) - \sum_{k=1}^K u_k(t)\|_2^2$  is the second penalty term and  $\langle \cdot \rangle$  is the inner product operation.

The unconstrained variational problem (2) is solved by the multiplication operator alternating direction method and  $u_k(t)$ ,  $\omega_k$ , and  $\theta(t)$  are alternately updated as follows:

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

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

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

performing an inverse Fourier transform on  $\hat{u}_k(\omega)$ , and the actual part  $\{u_k(t)\}$  is the solution  $\{u_k(t)\}$  ( $k = 1, 2, \dots, K$ ).

The VMD algorithm introduces the original signal into the variational model and then uses the process of finding the optimal solution of the constrained variational model to obtain the component. In this process, each component is alternately iteratively updated in the frequency domain, adaptively decomposing the frequency band of the signal, and finally gets  $K$  narrowband components.

**2.2. Modal Number Selection Method Based on Mutual Information.** The  $K$  value of decomposition mode number should be determined first when VMD is used for decomposition. The selection of the  $K$  value has a great influence on the decomposition effect. When  $K$  is too small, signal decomposition may be insufficient, and some signals may be treated as noise. When  $K$  is too large, the signal may be overdecomposed, generating excessive noise or causing modal aliasing. In order to overcome the influence of  $K$  value selection on the decomposition effect, this paper proposes a method of selecting the number of modes based on mutual information, called the MVMD algorithm.

Mutual information is a criterion to measure the degree of interdependence between two random variables [30, 31], and the greater the mutual information between two random variables, the stronger the correlation between them.

The mutual information between two variables can be expressed as

$$M(X; Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log \frac{p(x, y)}{p(x)p(y)}, \quad (6)$$

where  $p(x)$  and  $p(y)$  are the probability of edge distribution of two random variables  $X$  and  $Y$ , respectively, and  $p(x, y)$  is the probability of joint distribution.

In this paper, the mutual information relationship between the reconstructed series after VMD decomposition and the original series is used to select the  $K$  value. When the signal is decomposed by VMD, the decomposition signal will slowly transition from the state of insufficient decomposition to the state of full decomposition with the increase in the number  $K$  of decomposition layers. In this process, the mutual information between the reconstructed series and the original series is increasing continuously. When the number of decomposition layers  $K$  is greater than the optimal  $K$  value, the VMD decomposition will enter the overdecomposition state, and the mutual information of the reconstructed series and the original series between adjacent  $K$  values will tend to be stable. In addition, the overdecomposition state may also bring redundant noise signal, which will reduce the mutual information between the reconstructed series and the original series. Therefore, this paper selects a threshold value  $\sigma$  for the determination of the optimal  $K$  value. If the mutual information difference between the adjacent decomposition layers  $K_i$  and  $K_{i+1}$  is less than the threshold, the optimal value of VMD is considered as  $K_i$ .

The algorithm flow to solve the optimal  $K$  value is shown in Figure 1.

**2.3. KELM.** The extreme learning machine is a single-hidden layer feedforward neural networks (SLFNs) proposed by Huang et al. [32], and KELM is an extension of ELM by combining regularization theory and kernel method [33]. KELM has better stability and generalization ability compared with ELM and has been widely used in regression fitting problems.

The algorithm model of KELM is expressed as follows:

$$f(x) = h(x)\beta, \quad (7)$$

where  $h(x)$  is the output matrix of the hidden layer and  $\beta$  is the connection weight of the hidden layer and the output layer.

In order to improve the stability and generalization ability of the model, the penalty coefficient  $C$  is introduced, and then, the least squares solution of the output weight can be obtained according to the generalized inverse matrix theory:

$$\hat{\beta} = H^T \left( \frac{I}{C} + HH^T \right)^{-1} y, \quad (8)$$

where  $H$  is the hidden layer output matrix and  $y$  is the vector of the desired output. So the output function of the KELM can be written as

$$f(x) = \left( H^T \left( \frac{I}{C} + HH^T \right)^{-1} y \right) h(x). \quad (9)$$

Considering the case where the feature mapping function  $h(x)$  is unknown, the kernel function needs to be introduced into the KELM. To define the kernel matrix

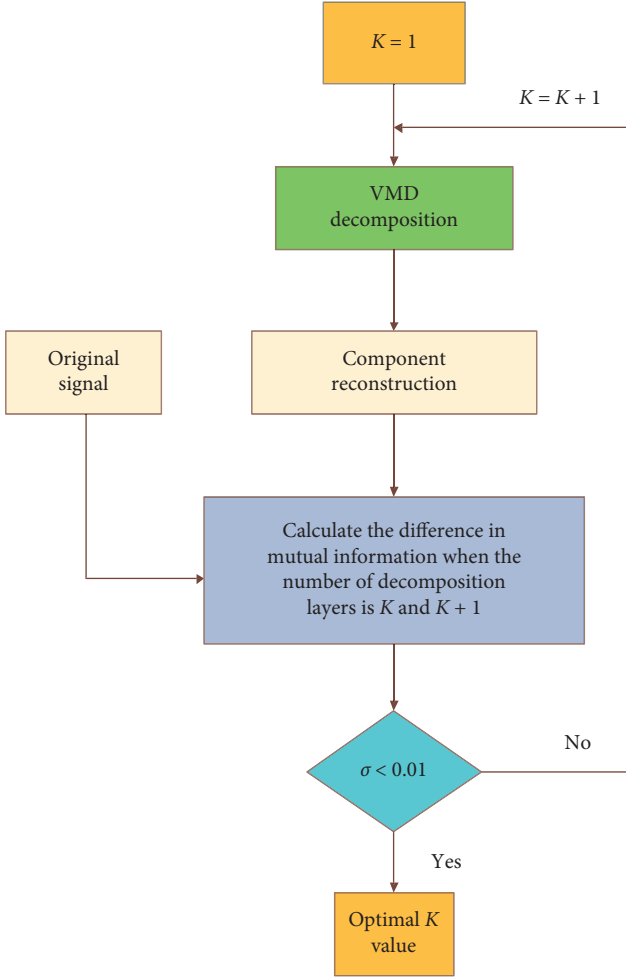


FIGURE 1: MVMD solves the optimal  $K$  value.

$\Omega_{\text{KELM}} = HH^T$ , the element of the matrix is  $\Omega_{\text{KELM}}(i, j) = h(x_i)h(x_j) = K(x_i, x_j)$ , and then, the network output of KELM can be expressed as follows:

$$f(x) = h(x)H^T \left( \frac{I}{C} + HH^T \right)^{-1} y = \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}^T \left( \frac{I}{C} + \Omega_{\text{KELM}} \right)^{-1} y, \quad (10)$$

where  $K(x_i, x_j)$  is a kernel function; this paper selects a radial basis function kernel with a wide convergence domain and good generalization ability:

$$K(x_i, x_j) = \exp\left(-\frac{1}{\lambda^2} \|x_i - x_j\|^2\right). \quad (11)$$

**2.4. GWO.** The grey wolf optimization (GWO) algorithm is a novel swarm intelligence optimization algorithm proposed by Mirjalili et al. [34]. Compared with the particle swarm optimization (PSO) algorithm, GWO has high optimization precision, fast convergence, and good robustness.

The algorithm searches for the optimal solution by simulating the social class and predation behavior of the wolves. The grey wolf relies on the strict hierarchical system within the wolves to complete the hunting and predation process. The pyramid model is built with the social rank of the wolves. It is divided into four grades  $\alpha$ ,  $\beta$ ,  $\delta$ , and  $\omega$  from top to bottom, as shown in Figure 2. In the wolves, the  $\alpha$  wolf dominates and the wolves are rounded up from all directions with the prey as the center. Under the leadership of the  $\alpha$  wolf, the attack is carried out by the  $\beta$  and  $\delta$  closest to the prey. Group  $\omega$  is used to fill vacancies to prevent prey from escaping. The wolves continue to attack to narrow the encirclement and eventually capture the prey.

The siege model of wolves against the prey is as follows:

$$\begin{aligned} D &= |CX_p(t) - X(t)|, \\ X(t+1) &= X_p(t) - AD, \end{aligned} \quad (12)$$

where  $X$  is the current position of the grey wolf individual,  $X_p$  is the position of the prey,  $D$  is the distance between the prey and the grey wolf,  $A$  is the convergence factor,  $C$  is the swing factor, and  $A$  and  $C$  can be expressed as follows:

$$\begin{aligned} A &= 2a \cdot r_1 - a, \\ C &= 2r_2, \end{aligned} \quad (13)$$

where  $a$  is linearly reduced and has  $a \in [0, 2]$  and  $r_1$  and  $r_2$  are random vectors in  $[0, 1]$ .

In GWO,  $\alpha$ ,  $\beta$ , and  $\delta$  are the optimal solutions in the whole wolf group. Other wolves can update their hunting positions according to the positions of  $\alpha$ ,  $\beta$ , and  $\delta$ . When the grey wolf updates the position, it should first calculate the distance from the three wolves  $\alpha$ ,  $\beta$ , and  $\delta$ , and the formula is as follows:

$$\begin{cases} D_\alpha = |C_1 X_\alpha(t) - X(t)|, \\ D_\beta = |C_2 X_\beta(t) - X(t)|, \\ D_\delta = |C_3 X_\delta(t) - X(t)|, \end{cases} \quad (14)$$

where  $D_\alpha$ ,  $D_\beta$ , and  $D_\delta$  represent the distance of the grey wolf from  $\alpha$ ,  $\beta$ , and  $\delta$  and  $X_\alpha$ ,  $X_\beta$ , and  $X_\delta$  represent the position of the grey wolf  $\alpha$ ,  $\beta$ , and  $\delta$ , respectively.

After calculating  $D_\alpha$ ,  $D_\beta$ , and  $D_\delta$ , the other wolves will approach the three wolves of  $\alpha$ ,  $\beta$ , and  $\delta$  and update their position. The formula is as follows:

$$\begin{cases} X_1 = X_\alpha(t) - A_1 D_\alpha, \\ X_2 = X_\beta(t) - A_2 D_\beta, \\ X_3 = X_\delta(t) - A_3 D_\delta, \end{cases} \quad (15)$$

where  $X_1$ ,  $X_2$ , and  $X_3$  are vectors of other wolves moving to  $\alpha$ ,  $\beta$ , and  $\delta$ , respectively.

According to the following formula, the new position of the grey wolf individual in the population after iteration can be obtained:

$$X(t+1) = \frac{(X_1 + X_2 + X_3)_3}{3}. \quad (16)$$

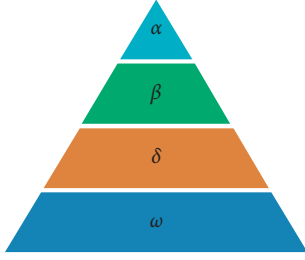


FIGURE 2: Wolves classification.

**2.5. MVMD-OKELM Prediction Model.** In the implementation of KELM learning algorithm, if the selection of nuclear parameters and regularization coefficients is not appropriate, the performance of the model will be affected. If the intelligent optimization algorithm is combined with KELM and the kernel parameters and regularization coefficients are optimized, the performance of the model can be improved to some extent. This paper proposes an OKELM method, which uses the grey wolf group intelligent optimization algorithm to optimize the parameters of KELM.

In implementation of the OKELM method, the root mean square error is used as the fitness evaluation function:

$$\text{fitness} = \sqrt{\frac{1}{N} \sum_{i=1}^N [y(i) - y_d(i)]^2}, \quad (17)$$

where  $y(i)$  is the predicted output of the model and  $y_d(i)$  is the desired output.

This paper proposes an underwater acoustic signal prediction model based on MVMD-OKELM. The modeling process of the prediction model is shown in Figure 3. Specific steps are as follows:

- (1) Decompose the underwater acoustic signal using MVMD to obtain several IMF components
- (2) Establish a KELM prediction model for each IMF component obtained by MVMD decomposition
- (3) Optimize the selection of each KELM parameter
- (4) The prediction result of each component is reconstructed to obtain a prediction result
- (5) Perform error analysis on the prediction results

### 3. Application Examples

In order to verify the validity of the MVMD-OKELM prediction model, the MVMD-OKELM prediction model is applied to chaotic time series prediction and underwater acoustic signal time series prediction, respectively.

In this paper, GWO and PSO are used to optimize KELM parameters, and the parameters are set as shown in Table 1.

The evaluation index uses the root mean square error  $E_{\text{RMSE}}$ , the average absolute error  $E_{\text{MAE}}$ , and the determination coefficient  $R^2$ , respectively:

$$E_{\text{RMSE}} = \sqrt{\frac{\sum_{i=1}^N (y(i) - y_d(i))^2}{N}},$$

$$E_{\text{MAE}} = \frac{1}{N} \sum_{i=1}^N |y(i) - y_d(i)|, \quad (18)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y(i) - y_d(i))^2}{\sum_{i=1}^N (y(i) - \bar{y})^2},$$

where  $N$  is the number of predicted sample points,  $y_d(i)$  is the actual output,  $y(i)$  is the predictive output, and  $\bar{y}$  is the average of the time series. The closer the value of  $R^2$  is to 1, the better the fit of the predicted output data to the original data.

**3.1. Mackey–Glass Chaotic Time Series Prediction.** The Mackey–Glass equation is a classical chaotic dynamic system. It is often used to test the predictive performance of nonlinear time series. The chaotic time series is generated by the following equation [35]:

$$\frac{dy(t)}{dt} = \frac{\alpha x(t - \tau)}{1 + x^\gamma(t - \tau)} - \beta x(t), \quad (19)$$

where  $\alpha = 0.2$ ,  $\beta = 0.1$ ,  $\gamma = 10$ , and  $\tau$  are time-delay parameters.

The Mackey–Glass time series exhibits chaotic characteristics when  $\tau > 16.8$ , and the larger the  $\tau$  value, the higher the degree of chaos, and the greater the prediction difficulty. The Mackey–Glass chaotic time series generated by (19) can be solved by the dde23 function in Matlab. Figure 4 shows a chaotic time series of 700 generated when  $\tau = 30$ . It can be seen from the figure that the time series has more complex nonlinear chaotic characteristics. In the sample division, 700 data points are divided into 694 samples, and the first 80% of the samples are selected as training samples and the last 20% are used as test samples.

Firstly, the original chaotic time series is decomposed by using VMD, and the  $K$  value is determined by the mutual information method proposed in this paper before decomposition. With the change in the  $K$  value, the mutual information trend of Mackey–Glass chaotic time series is shown in Figure 5.

As we can see from Figure 5, when  $K = 8$ , value of mutual information is 1.886; when  $K = 9$ , value of mutual information is 1.874. That is, when  $K > 8$ , the VMD decomposition appears overdecomposed, so we set the  $K$  value to 8.

The MVMD decomposition results of the Mackey–Glass chaotic series are shown in Figure 6.

To compare the prediction effect of the MVMD-OKELM model, this paper will use EMD to decompose it and establish the EMD-OKELM prediction model. The EMD decomposition result is shown in Figure 7.

As shown in Figures 6 and 7, the Mackey–Glass chaotic time series is decomposed into 8 IMF components by VMD and 7 IMF components by EMD decomposition. It can be seen from the decomposition results that the high-frequency component fluctuation of EMD is not stable, and the

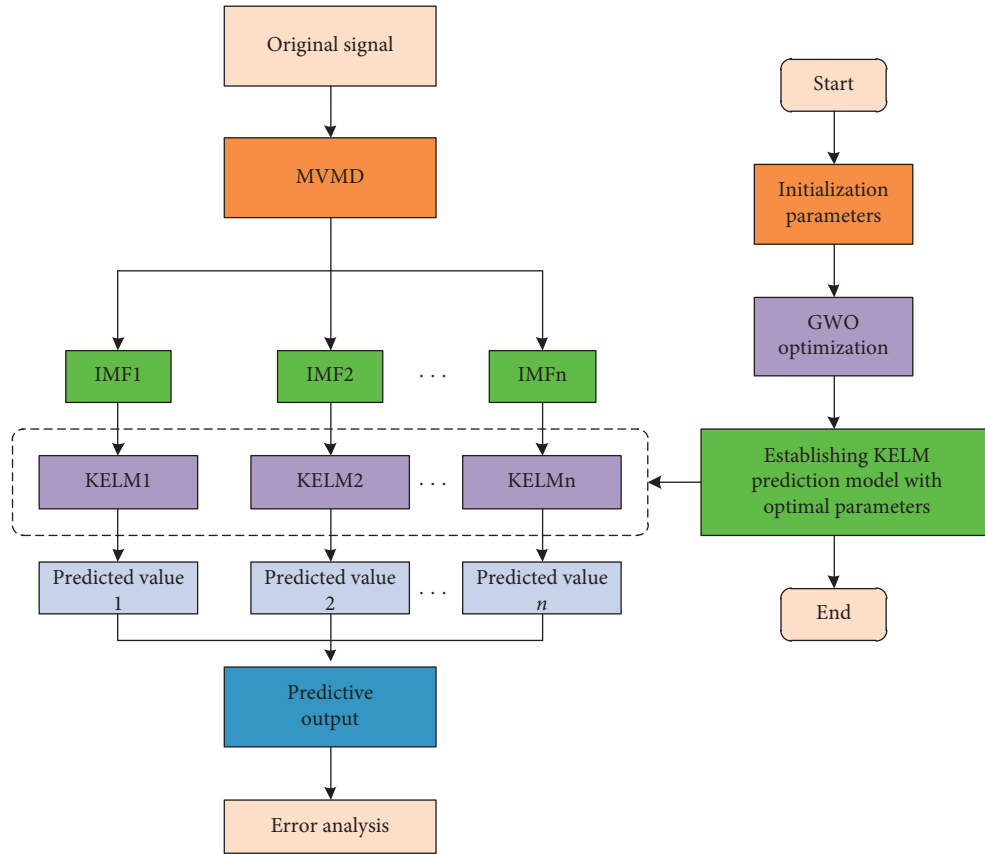


FIGURE 3: MVMD-OKELM prediction model.

TABLE 1: Comparison of MVMD-OKELM and performance indicators of each model.

The main parameters	Parameter value
Population size	10
The maximum number of iteration	50
Regularization parameter optimization range	[1, 800]
Kernel function width optimization range	[1, 70]

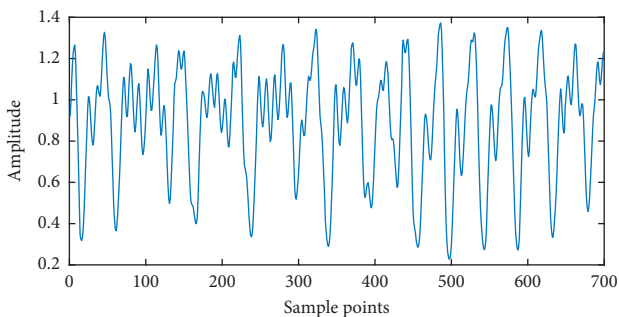


FIGURE 4: Mackey-Glass chaotic time series.

endpoint effect appears, which greatly affects the performance of the subsequent prediction model. To compare the prediction effect of MVMD-OKELM prediction model, SVM, RBF, KELM, ARIMA, PSO-KELM, OKELM, and EMD-OKELM will be used to predict Mackey-Glass chaotic time series. When using SVM to predict, we set the upper

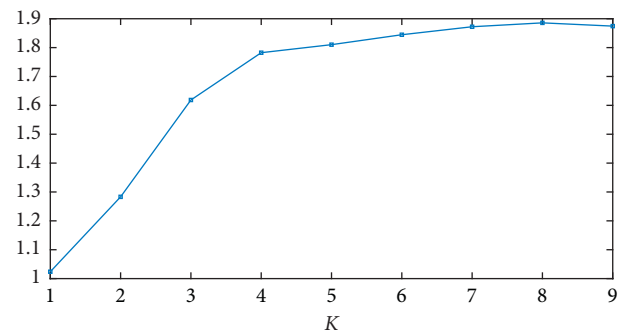


FIGURE 5: Mackey-Glass mutual information trend chart.

bound of Lagrange multiplier to 2 and the parameter of insensitive loss function to 10. When forecasting with RBF, we use newrb function in Matlab and set the radial basis expansion parameter to 0.07; when KELM prediction is used alone, we set the regularization parameter to 2 and the kernel parameter to 20. When ARIMA is used for prediction, auto.arima function in R language is used to obtain parameters, and then ARIMA (4, 0, 2) prediction model is obtained, where 4 represents the lag number of data itself, 0 represents the difference in order, and 2 represents the lag number of prediction error. The optimization algorithm is adopted to determine the parameter setting in the combined prediction model. The prediction results of SVM, RBF, KELM, and ARIMA are shown in Figure 8; the combined

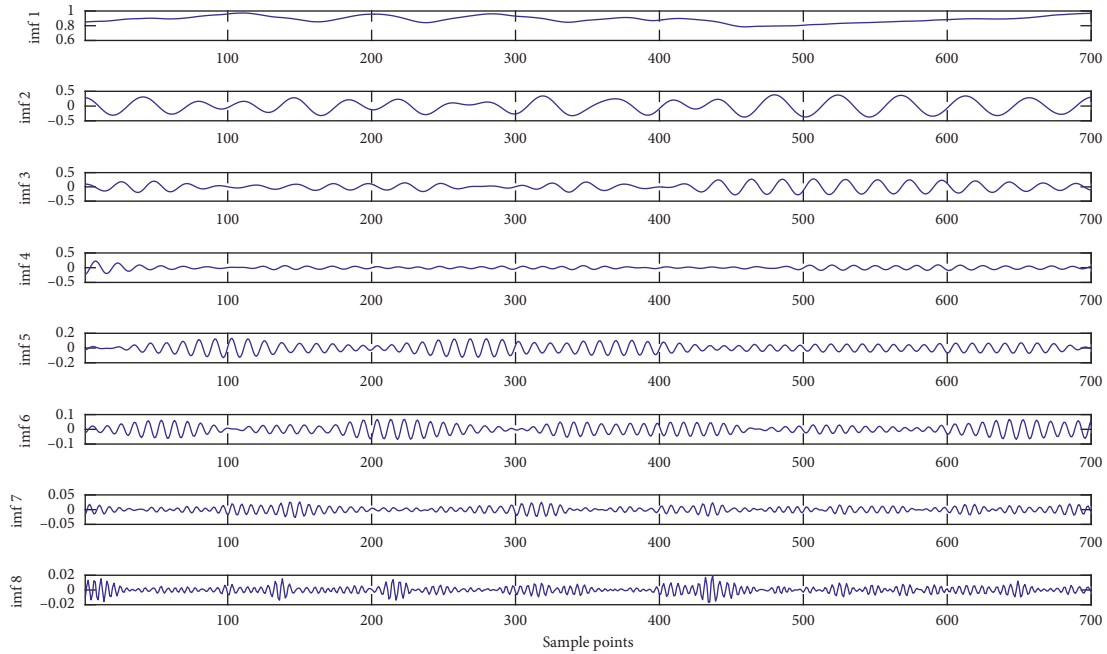


FIGURE 6: MVMD decomposition of Mackey-Glass chaotic sequence.

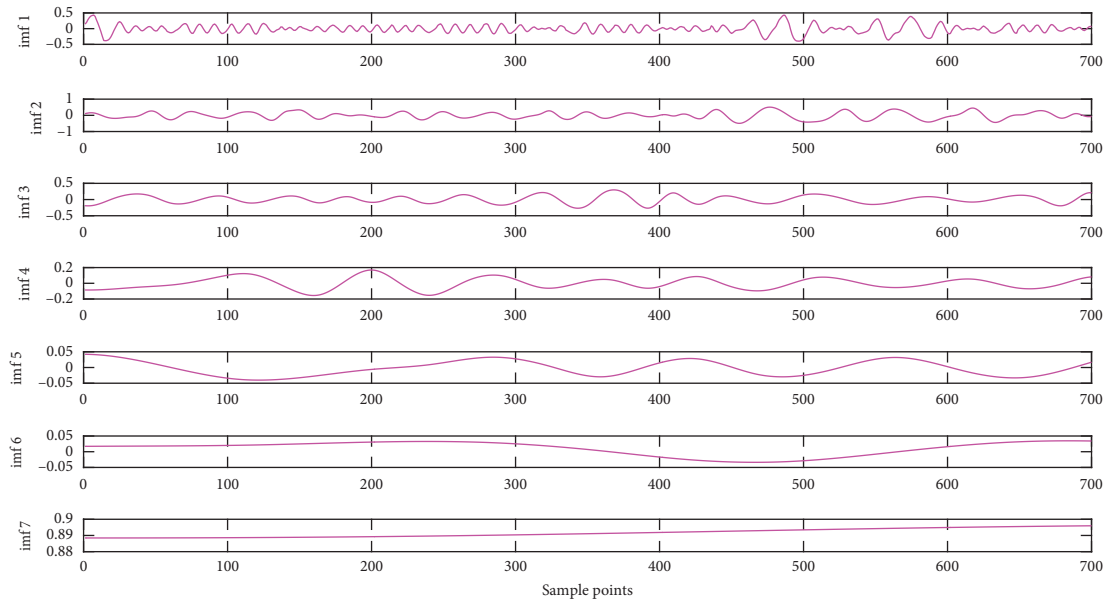


FIGURE 7: EMD decomposition of Mackey-Glass chaotic series.

prediction models of PSO-KELM, OKELM, and EMD-OKE are shown in Figure 9. The error distribution between the predicted output and real value of each model is shown in Figure 10, and Figure 11 further shows a boxplot of the prediction error.

It can be seen from Figures 8 and 9 that the ARIMA prediction model has a better prediction effect in a single model, and the prediction result of combined prediction model is far better than that of single prediction model. As can be seen from Figures 10 and 11, MVMD-OKELM model has the smallest fluctuation of prediction error and

better prediction performance. However, the difference between OKELM method and PSO-KELM method is small, and the performance of the two algorithms cannot be clearly distinguished. To verify the superiority of OKELM, Figure 12 gives the optimal value iteration curves of OKELM and PSO-KELM. It can be seen that OKELM has strong optimization ability and tends to be stable after 25 iterations. To compare the iteration speeds of the two algorithms, this paper performed 10 iterations and averaged times of the 10 algorithm iterations. The average iteration time of OKELM is 16.365267 seconds,

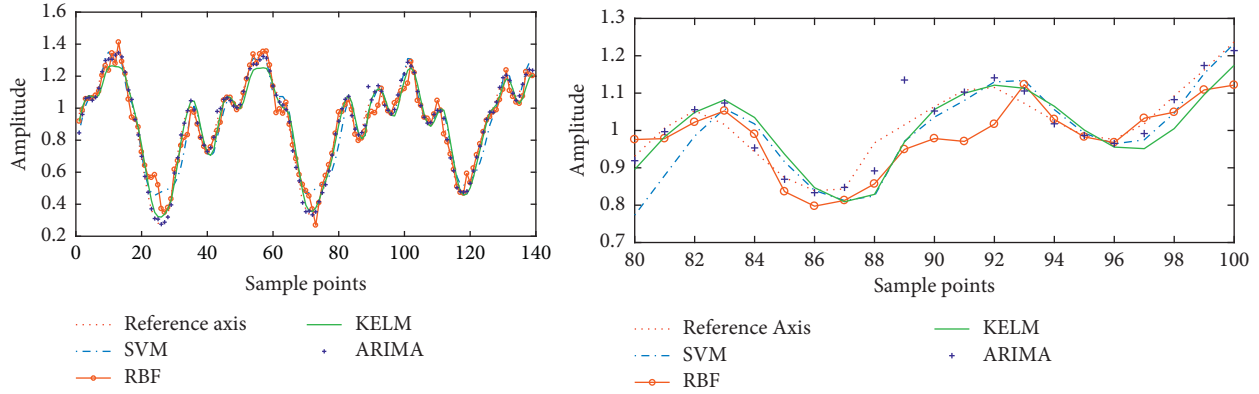


FIGURE 8: Prediction results of single model.

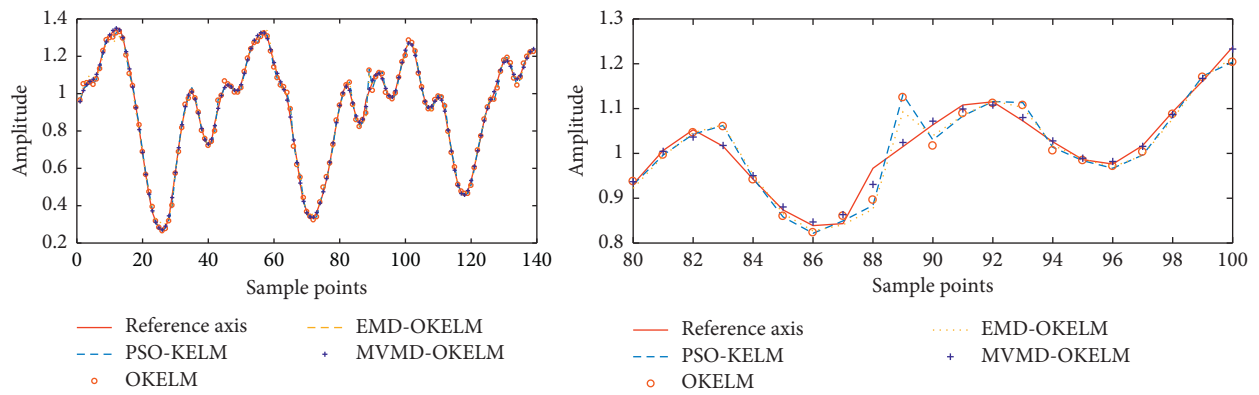


FIGURE 9: Prediction results of combined model.

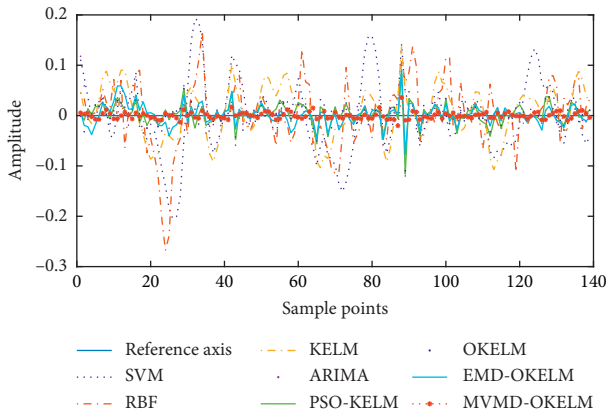


FIGURE 10: Prediction error distribution.

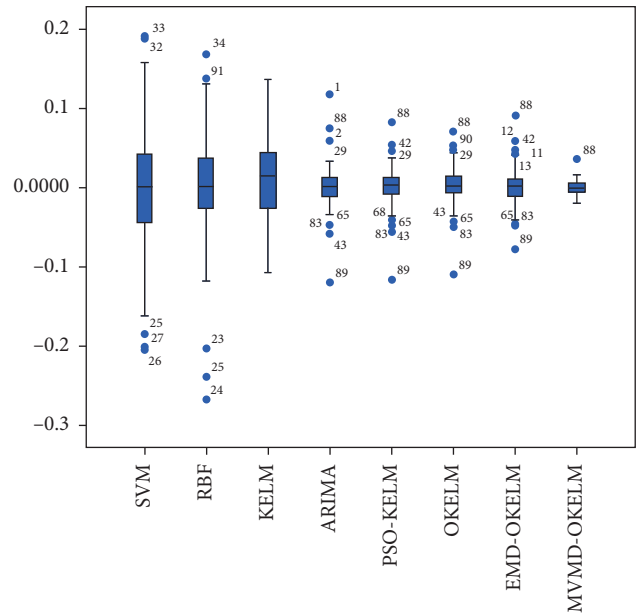


FIGURE 11: Boxplot of prediction error.

and the average iteration time of PSO-KELM is 16.080680 seconds. In summary, we can conclude that iteration speed of OKELM and PSO-KELM is not much different, and optimization ability of OKELM is stronger. Therefore, the OKELM model proposed in this paper has better prediction performance and is more suitable for KELM parameter optimization.

To quantitatively analyze the predictive performance of each predictive model, Table 2 gives a comparison of performance indicators for each predictive model, and

Figure 13 shows a comparison of RMSE and MAE for each prediction model.

As can be seen from Table 2 and Figure 13, a single ARIMA model is better than SVM, RBF, and KELM, and



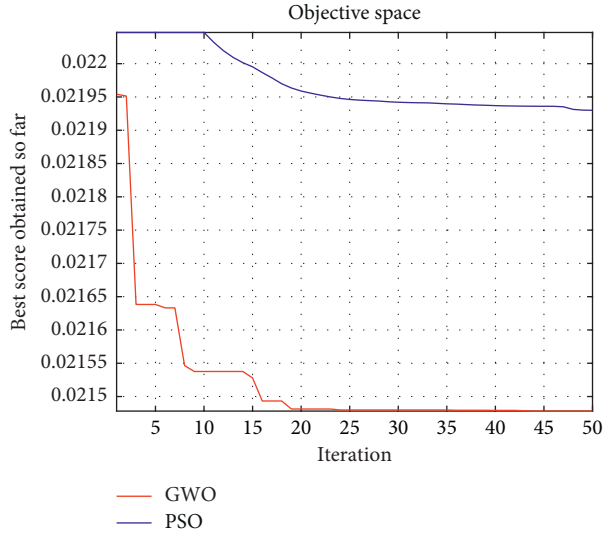


FIGURE 12: GWO-KELM and PSO-KELM optimal value convergence curve.

TABLE 2: Comparison of MVMD-OKELM and performance indicators of each model.

Predictive model	RMSE	MAE	$R^2$
SVM	0.07495	0.05798	0.9451
RBF	0.06446	0.04604	0.9573
KELM	0.05128	0.04280	0.9668
ARIMA	0.02464	0.01672	0.9924
PSO-KELM	0.02194	0.01525	0.9941
OKELM	0.02147	0.01520	0.9943
EMD-OKELM	0.02127	0.01516	0.9945
MVMD-OKELM	0.00701	0.00544	0.9993

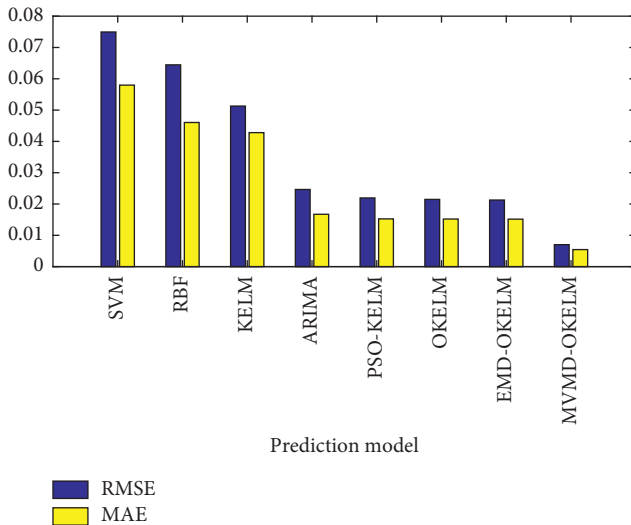


FIGURE 13: RMSE and MAE histograms for each prediction model.

the predicted accuracy of the optimized KELM model is improved compared to ARIMA. The prediction results of OKELM and PSO-KELM are very close and have similar prediction accuracy, but the search ability of GWO is stronger than that of PSO. Compared with PSO, it is not

easy to fall into the local optimal solution. Therefore, we chose the OKELM prediction model based on the GWO optimization algorithm for subsequent decomposition prediction. The EMD-OKELM prediction model based on EMD decomposition only improves a small amount of prediction accuracy based on OKELM. The prediction accuracy of MVMD-OKELM prediction model based on MVMD decomposition is about one order of magnitude higher than that of other models, and its prediction performance is significantly higher than other predictive models.

**3.2. Underwater Acoustic Signal Prediction Experiment.** The underwater acoustic signal selected in this paper is the ship radiated noise signal. The underwater acoustic signal has typical chaotic characteristics and can be used for short-term prediction. In order to contrast with Mackey–Glass chaotic time series, this section also divides 694 sets of data and selects the first 80% sets of data as training samples and the latter 20% sets of data as test samples. The underwater acoustic signal is shown in Figure 14.

As can be seen from Figure 14, the data complexity is much higher than the Mackey–Glass chaotic time series, which means that the prediction of underwater acoustic signal is more difficult than the Mackey–Glass chaotic time series.

Before the prediction, the underwater acoustic signal is first decomposed by MVMD, and  $K$  value is selected by the mutual information method. The trend of mutual information is shown in Figure 15.

As can be seen from Figure 15, when  $K=8$ , the value of mutual information is 1.709; when  $K=9$ , value of mutual information is 1.682. That is, when  $K>8$ , the VMD decomposition appears overdecomposed, so we set the  $K$  value to 8.

The MVMD decomposition result of the underwater acoustic signal and the decomposition result of the EMD is shown in Figures 16 and 17.

The chaotic component is mainly concentrated in the high-frequency component [36], and the stronger the chaos, the higher the prediction difficulty. To compare the prediction performance of MVMD-OKELM and EMD-OKELM in more detail, this paper selects the high-frequency component  $\text{imf1}$  of EMD decomposition and the high-frequency component  $\text{imf8}$  of MVMD decomposition for comparison. The results are shown in Figures 18 and 19, respectively.

The root mean square error of EMD-OKELM is 0.0309, the root mean square error of MVMD-OKELM is 0.0017, and the prediction accuracy of MVMD-OKELM is one order of magnitude lower than that of EMD-OKELM. As can be seen from Figure 16, the prediction effect of some points in the high-frequency component decomposed by EMD is poor, which is caused by the large fluctuation of the high-frequency component decomposed by EMD, while the  $\text{imf8}$  component decomposed by MVMD fluctuates closely around a central frequency and the fluctuation is relatively stable, so the prediction accuracy of MVMD-OKELM is

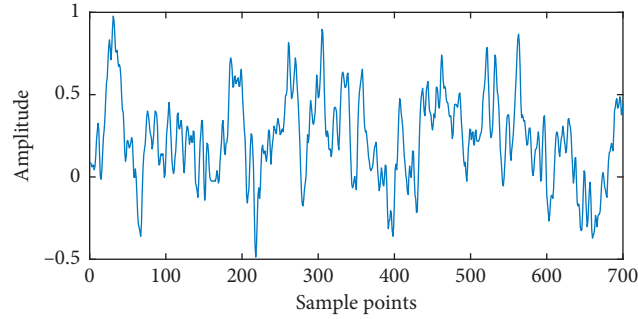


FIGURE 14: Underwater acoustic signal series I.

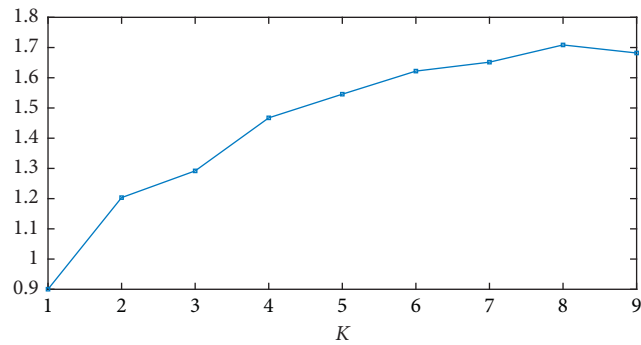


FIGURE 15: Underwater acoustic signal I mutual information trend.

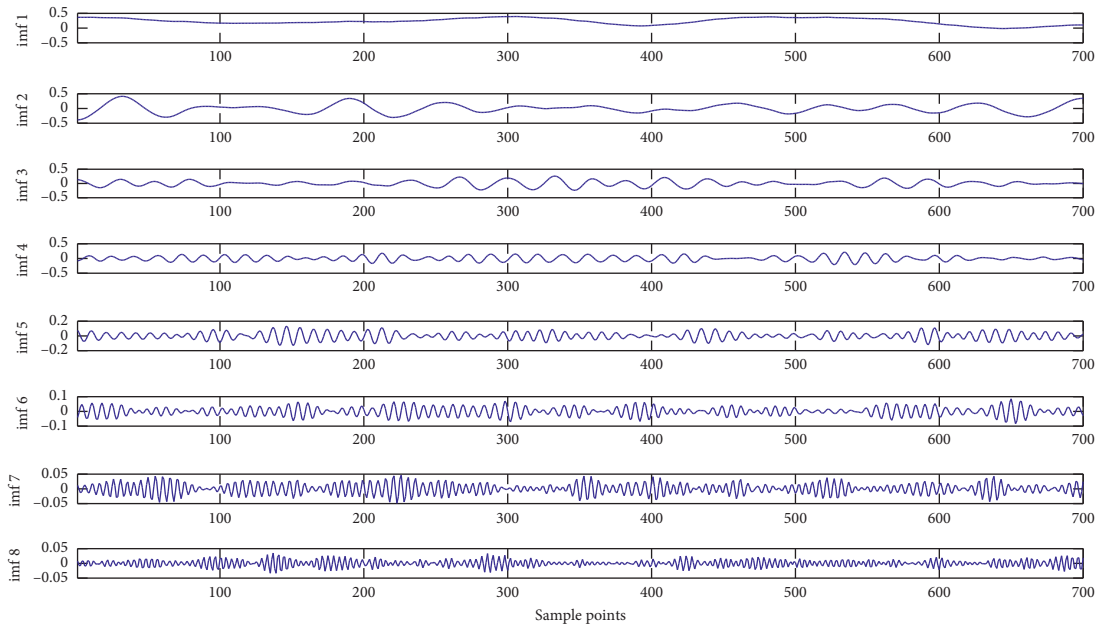


FIGURE 16: Underwater acoustic signal I MVMD decomposition.

higher. Since the prediction results are predicted and reconstructed from all components separately, the MVMD-OKELM reconstruction should produce better results than EMD-OKELM. In this section, except that the parameters of the ARIMA model are changed to ARIMA (1, 0, 3), the parameter settings of other comparative models refer to the

previous section. The prediction results of single model of underwater acoustic signal are shown in Figure 20, and the prediction results of combined prediction models are shown in Figure 21. The error distribution of the underwater acoustic signal is shown in Figure 22, and the corresponding error boxplot is shown in Figure 23.

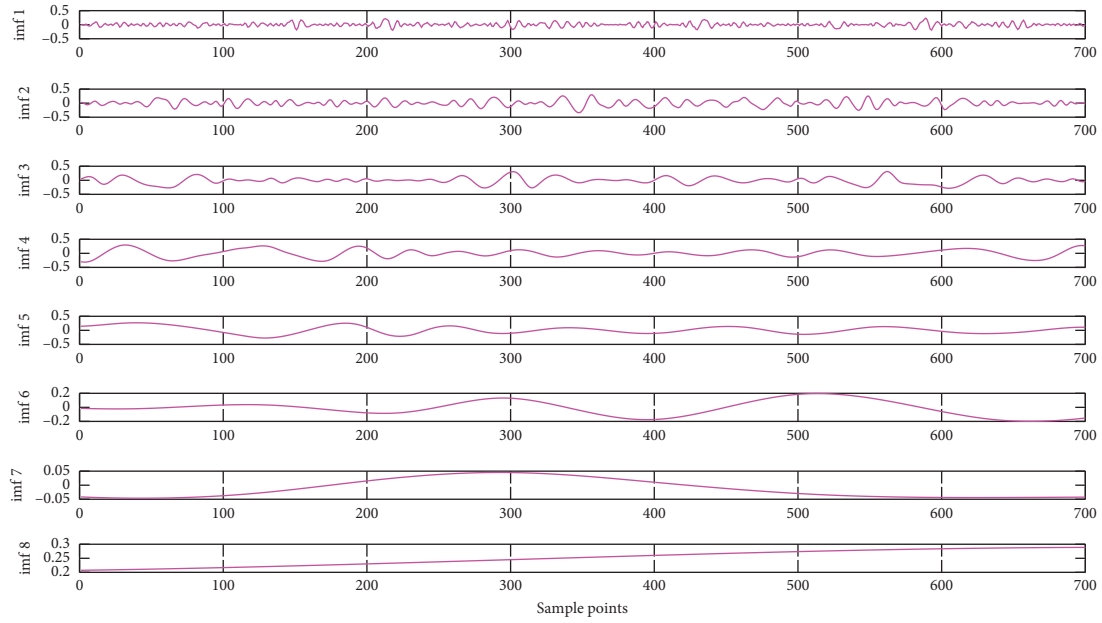


FIGURE 17: Underwater acoustic signal I EMD decomposition.

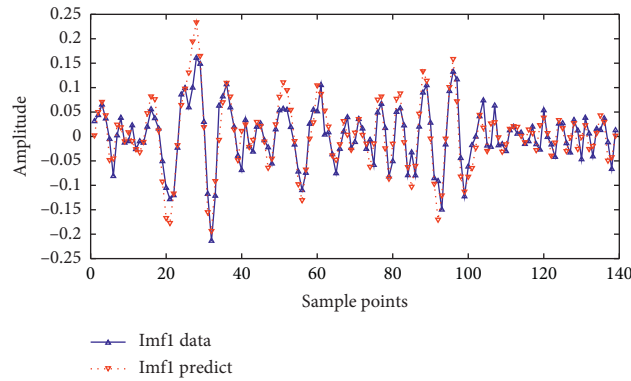


FIGURE 18: EMD imf1 component prediction result.

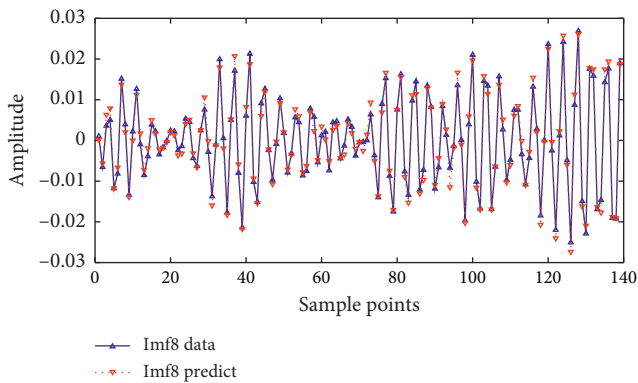


FIGURE 19: MVMD imf8 component prediction result.

As can be seen from Figures 20 and 21, in underwater acoustic signal prediction, the prediction accuracy of the combined prediction model is still higher than the single prediction model. As can be seen from the error distribution

diagram in Figure 22, the MVMD-OKELM prediction error distribution proposed in this paper is relatively uniform and fluctuates near the baseline, and its prediction performance is significantly better than other models. It can be seen in the boxplot of Figure 23 that the prediction error of ARIMA prediction model at some points is large, and it cannot fit all the points. In Figure 23, the fluctuation of MVMD-OKELM is the smallest, and the error distribution is relatively concentrated, which is consistent with the results obtained in Figure 22, which further illustrates that MVMD-OKELM has better prediction performance. Table 3 gives the specific performance indicators, and Figure 24 shows the RMSE and MAE of the underwater acoustic signal prediction model.

As can be seen from Table 3 and Figure 24, MVMD-OKELM has smallest RMSE value and MAE value and highest  $R^2$  value, which indicates that the MVMD-OKELM prediction model can be effectively applied in underwater acoustic signal prediction. However, the overall prediction performance index of the underwater acoustic signal is slightly worse than the prediction result of the

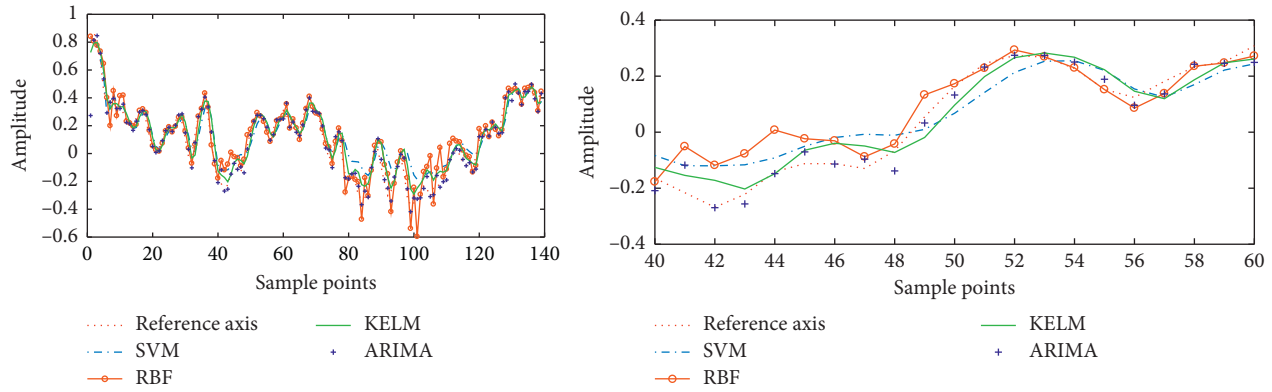


FIGURE 20: Underwater acoustic signal I prediction result of single model.

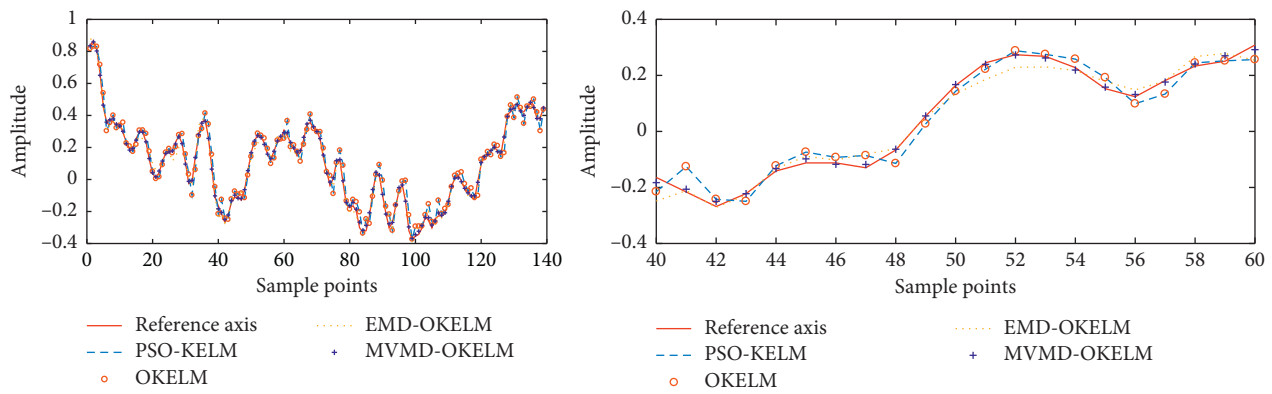


FIGURE 21: Underwater acoustic signal I prediction result of combined model.

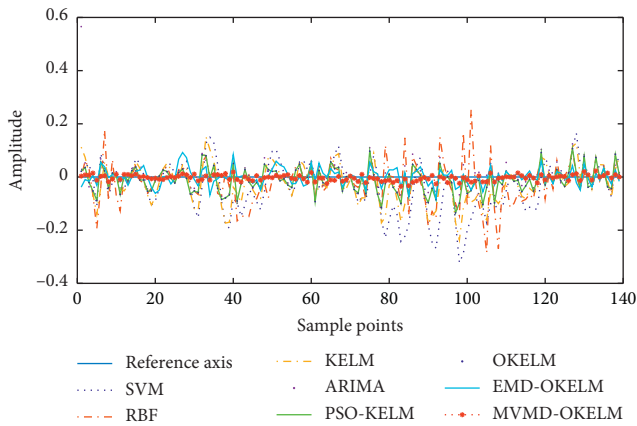


FIGURE 22: Underwater acoustic signal I prediction error distribution.

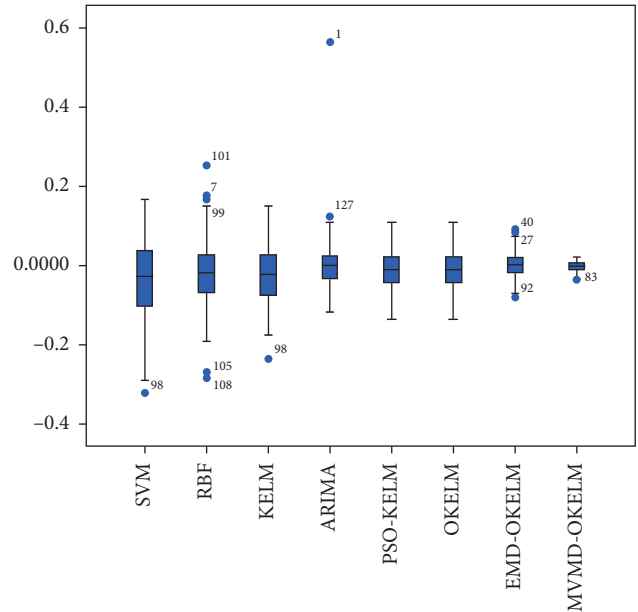


FIGURE 23: Boxplot of underwater acoustic signal I prediction error.

Mackey–Glass chaotic time series. This also shows that the chaotic degree of the actual underwater acoustic signal is much higher than that of the Mackey–Glass chaotic time series.

To avoid the accidental occurrence of underwater acoustic signal, this paper extracts 700 data points in the same data set as underwater acoustic signal II. Underwater acoustic signal II is shown in Figure 25.

In the prediction of underwater acoustic signal II, SVM, RBF, KELM, ARIMA, PSO-KELM, OKELM, and EMD-OKELM is also used as comparison models to verify the

TABLE 3: Predictive model performance indicator comparison.

Predictive model	RMSE	MAE	$R^2$
SVM	0.10216	0.07931	0.8938
RBF	0.08401	0.06219	0.9108
KELM	0.07630	0.06045	0.9288
ARIMA	0.04686	0.03666	0.9672
PSO-KELM	0.04869	0.03877	0.9667
GWO-KELM	0.04868	0.03877	0.9667
EMD-OKELM	0.03166	0.02449	0.9853
MVMD-OKELM	0.01171	0.00938	0.9984

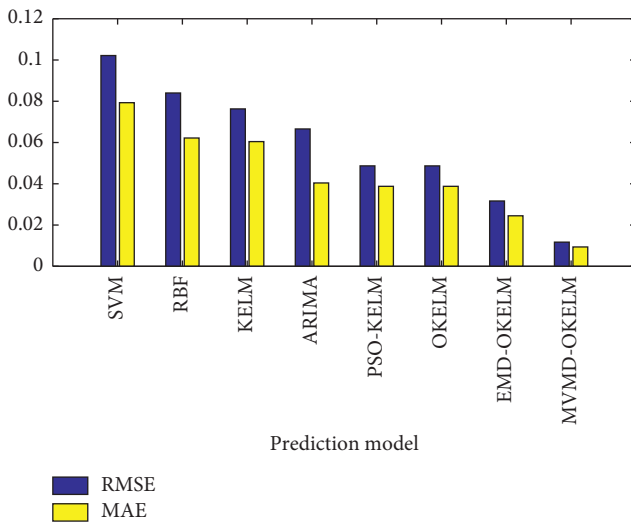


FIGURE 24: RMSE and MAE histograms for each prediction model.

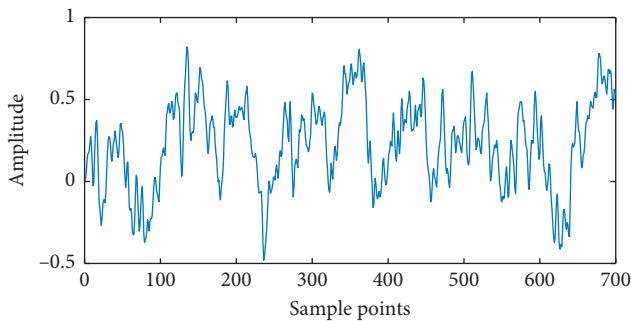


FIGURE 25: Underwater acoustic signal series II.

prediction performance of MVMD-OKELM. When MVMD decomposition is used, the mutual information method is also used to determine the  $K$  value, which can be determined as 5 in the mutual information trend in Figure 26. In the parameter settings, ARIMA model parameters of the underwater acoustic signal II are ARIMA (2, 1, 2). The single model prediction result of underwater acoustic signal II is shown in Figure 27, and the combined prediction model result is shown in Figure 28.

From Figures 27 and 28, it can be seen that, in single-step prediction, although the fitting degree of each model is

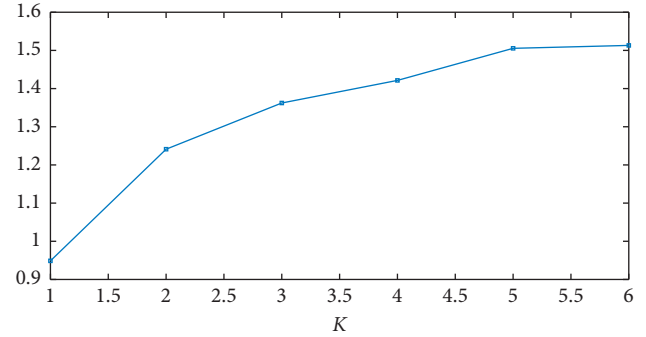


FIGURE 26: Underwater acoustic signal II mutual information trend.

already very high, the prediction accuracy of single prediction model is far lower than that the combined model. Although the prediction accuracy of combined prediction models is relatively high, we can still see that MVMD-OKELM has the best prediction effect. To judge the model better, this paper further gives the error boxplot of underwater acoustic signal II and the specific prediction indexes of the model.

In the boxplot of Figure 29, it can be seen that the MVMD-OKELM has the smallest error fluctuation. In Table 4, it can be seen that MVMD-OKELM has the lowest RMSE and MAE values and the highest coefficient of certainty. This fully shows that the prediction effect of MVMD-OKELM is the best in these models. Through the analysis of underwater acoustic signal II, we can conclude that MVMD-OKELM has strong applicability and can be effectively used in underwater acoustic signal prediction.

**3.3. Diebold–Mariano Test.** In order to find out the optimal prediction model statistically, this paper uses the Diebold–Mariano test [37] to evaluate whether there is significant difference between MVMD-OKELM and other comparative models in prediction ability. The Diebold–Mariano test results of Mackey–Glass time series are shown in Table 5. From Table 5, it can be seen that the  $P$  values tested by Diebold–Mariano are all less than 0.01, which indicates that the MVMD-OKELM model has significant differences compared with other 7 comparative models at a significance level of 1%. In the Diebold–Mariano test of underwater acoustic signal, except ARIMA has significant difference with the MVMD-OKELM model at 5% significance level, and other models have significant differences with the MVMD-OKELM model at 1% significance level. In the underwater acoustic signal II, the  $P$  values tested by Diebold–Mariano are all less than 0.01, which indicates that the MVMD-OKELM model has significant differences compared with other 7 comparative models, which is consistent with the conclusion obtained in the underwater acoustic signal I. Tables 6 and 7 further show that the MVMD-OKELM prediction model has significant advantages in prediction performance and can be effectively applied to underwater acoustic signal prediction.

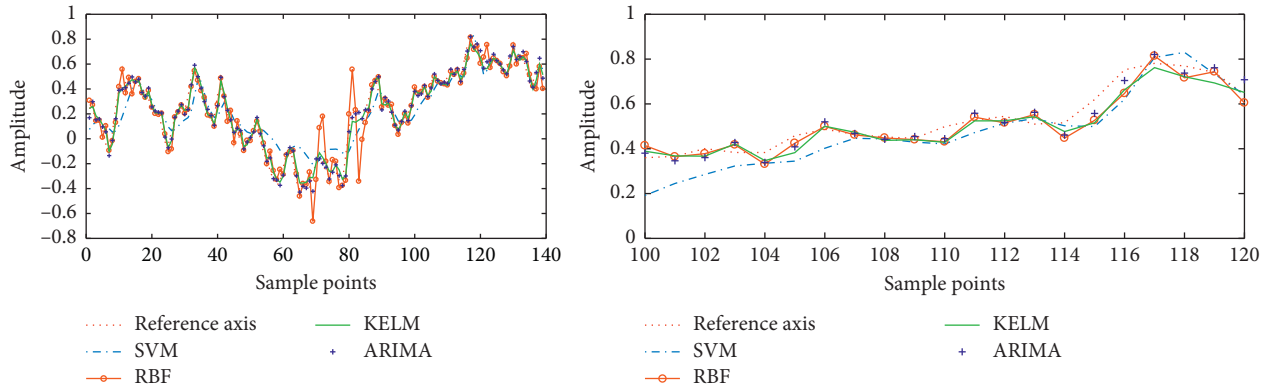


FIGURE 27: Underwater acoustic signal II prediction results of single model.

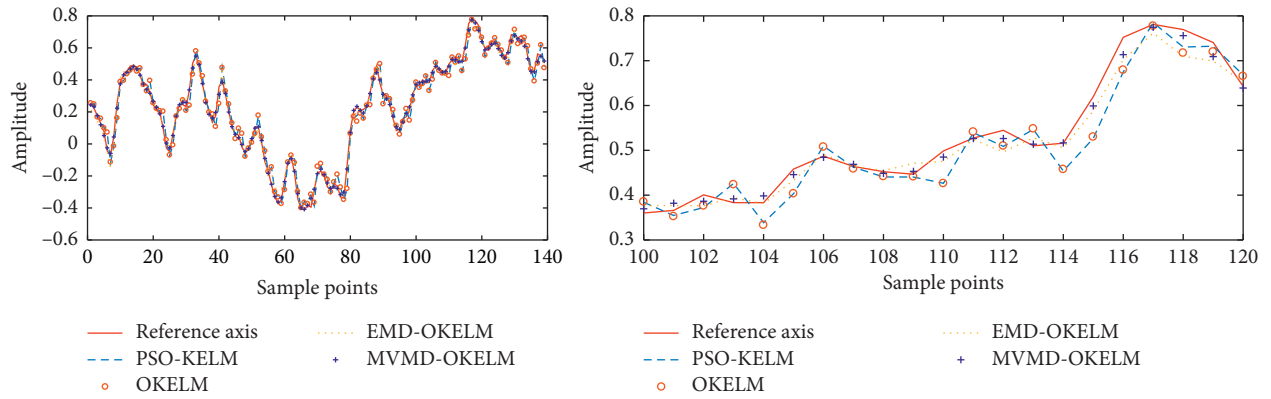


FIGURE 28: Underwater acoustic signal II prediction results of combined model.

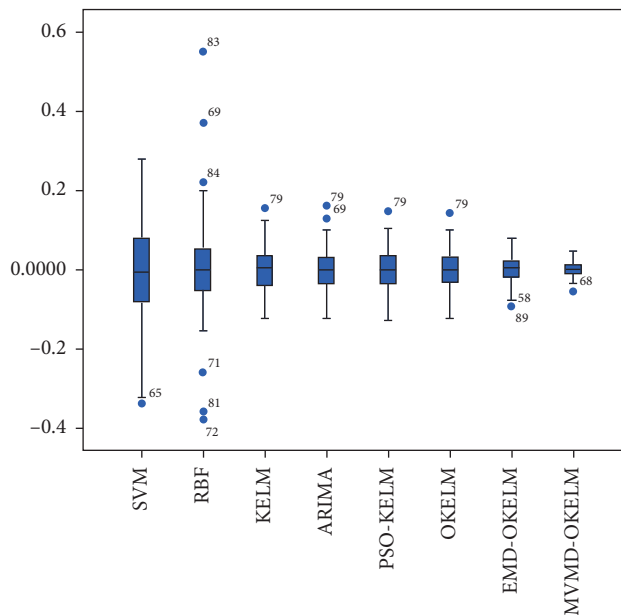


FIGURE 29: Boxplot of underwater acoustic signal II prediction error.

TABLE 4: Predictive model performance indicator comparison.

Predictive model	RMSE	MAE	$R^2$
SVM	0.13189	0.10251	0.8518
RBF	0.10370	0.06974	0.8946
KELM	0.05441	0.04422	0.9699
ARIMA	0.05024	0.04088	0.9738
PSO-KELM	0.04947	0.03965	0.9746
GWO-KELM	0.04897	0.03922	0.9751
EMD-OKELM	0.03228	0.02580	0.9901
MVMD-OKELM	0.00944	0.01392	0.9973

TABLE 5: Mackey–Glass time series Diebold–Mariano test.

Contrast model	SVM	RBF	KELM	ARIMA	PSO-KELM	OKELM	EMD-OKELM
DM statistics	7.5331*	5.2689*	9.9808*	3.6787*	3.8636*	4.2316*	4.9802*
$P$ value	0.0000	0.0000	0	0.0001	0.0000	0.0000	0.0000

\* A confidence interval of 1%.

TABLE 6: Underwater acoustic signal I Diebold–Mariano test.

Contrast model	SVM	RBF	KELM	ARIMA	PSO-KELM	OKELM	EMD-OKELM
DM statistics	7.5331*	6.4660*	7.8707*	1.8723**	8.3717*	8.3753*	7.0447*
$P$ value	0.0000	0.0000	0.0000	0.0306	0	0	0.0000

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

TABLE 7: Underwater acoustic signal II Diebold–Mariano test.

Contrast model	SVM	RBF	KELM	ARIMA	PSO-KELM	OKELM	EMD-OKELM
DM statistics	8.2559*	3.7431*	8.1767*	7.0529*	7.5690*	7.5290*	5.7260*
$P$ value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

\* A confidence interval of 1%.

## 4. Conclusions

From the perspective of mutual information, this paper uses the trend of mutual information between the reconstructed series and the original series to select the best  $K$  value and proposes a new  $K$  value selection method. To further improve the performance of the KELM method, an OKELM method is proposed to optimize the kernel parameters and regularization coefficients of KELM. Combined with MVMD decomposition, an underwater acoustic signal prediction model based on MVMD-OKELM is established. Through the comparative study of the experiment, it can be concluded that

- (1) OKELM has similar iterative speed compared with PSO-KELM, but GWO has a stronger optimization ability; it is not easy to fall into a local optimal solution compared to PSO
- (2) The MVMD-OKELM prediction model is significantly better than the contrast model in Mackey–Glass chaotic time series and underwater acoustic signal prediction, which shows the potential of MVMD-OKELM in underwater acoustic signal prediction
- (3) The prediction result of Mackey–Glass chaotic time series is better than that of underwater acoustic

signal, indicating that the chaotic of underwater acoustic signal is stronger than that of Mackey–Glass chaotic time series when  $\tau = 30$

## 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. X. Meng, S. E. Yang, S. C. Piao, and B. H. Shao, "Prediction and recognition method for signal of targets on the water

- surface,” *Journal of Harbin Engineering University*, vol. 37, no. 1, pp. 1–6, 2016.
- [2] 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.
  - [3] H. Y. Xing, Q. Q. Zhu, and W. Xu, “A method of weak target detection based on the sea clutter,” *Acta Physica Sinica*, vol. 63, no. 10, Article ID 100505, 2014.
  - [4] X. H. Zhang and X. M. Zhang, “Researches on chaotic phenomena of noises radiated from ships,” *Acta Acustica*, vol. 23, no. 2, pp. 134–140, 1998.
  - [5] Y. A. Li, D. M. Xu, and X. M. Zhang, “A new way for detecting and recognizing underwater ships using takens theorem,” *Journal of North Western Polytechnical University*, vol. 19, no. 2, pp. 266–269, 2001.
  - [6] H. P. Wang, Y. F. Jin, H. J. Cao, and X. H. Chen, “Research on prediction model of underwater acoustic signal based on radial basis function neural network,” *Ship Electronic Engineering*, vol. 28, no. 10, pp. 170–173, 2008.
  - [7] Aasim, S. N. Singh, and A. Mohapatra, “Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting,” *Renewable Energy*, vol. 136, pp. 758–768, 2019.
  - [8] S. Noureen, S. Atique, V. Roy, and S. Bayne, “Analysis and application of seasonal ARIMA model in Energy Demand Forecasting: a case study of small scale agricultural load,” in *Proceedings of the 2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 521–524, Dallas, TX, USA, August 2019.
  - [9] 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.
  - [10] 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.
  - [11] 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.
  - [12] M. A. Chitsazan, M. Sami Fadali, and A. M. Trzynałowski, “Wind speed and wind direction forecasting using echo state network with nonlinear functions,” *Renewable Energy*, vol. 131, pp. 879–889, 2019.
  - [13] G. H. 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, p. 177, 2018.
  - [14] J. H. Xiong, T. L. Wang, and R. S. Li, “Research on a hybrid LSSVM intelligent algorithm in short term load forecasting,” *Cluster Computing*, vol. 22, no. 4, pp. 8271–8278, 2019.
  - [15] Ü. Ç. Büyüksahin and Ş. Ertekin, “Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition,” *Neurocomputing*, vol. 361, pp. 151–163, 2019.
  - [16] G. H. Li, S. L. Zhang, and H. Yang, “A deep learning prediction model based on extreme-point symmetric mode decomposition and cluster analysis,” *Mathematical Problems in Engineering*, vol. 2017, Article ID 8513652, 6 pages, 2017.
  - [17] Q. M. Cheng, L. Chen, Y. M. Cheng, Q. Zhang, and J. Gao, “Short-term wind power forecasting method based on EEMD and LS-SVM model,” *Electric Power Automation Equipment*, vol. 38, no. 5, pp. 27–35, 2018.
  - [18] H. M. Nazir, I. Hussain, I. Ahmad, M. Faisal, and I. M. Almanjahie, “An improved framework to predict river flow time series data,” *PeerJ*, vol. 7, Article ID e7183, 2019.
  - [19] 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.
  - [20] K. Dragomiretskiy and D. Zosso, “Variational mode decomposition,” *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 531–544, 2014.
  - [21] F. X. Meng and Y. J. Zhang, “Ultra-short term load forecasting method based on VMD and ELM,” *Foreign Electronic Measurement Technology*, vol. 38, no. 1, pp. 24–28, 2019.
  - [22] S. Q. Zhang, X. S. Su, R. F. Chen et al., “Short-term load forecasting based on the VMD and FABB,” *Chinese Journal of Scientific Instrument*, vol. 39, no. 4, pp. 67–73, 2018.
  - [23] D. J. Chen, Y. D. Zhang, C. Y. Yao, B. W. Lai, and S. J. Lü, “Fault diagnosis method based on variational mode decomposition and multi-scale permutation entropy,” *Computer Integrated Manufacturing Systems*, vol. 23, no. 12, pp. 2604–2612, 2017.
  - [24] X. X. Zheng, G. W. Zhou, H. H. Ren, and Y. Fu, “A rolling bearing fault diagnosis method based on variational mode decomposition and permutation entropy,” *Journal of Vibration and Shock*, vol. 36, no. 22, pp. 22–28, 2017.
  - [25] Y. X. Li, Y. A. Li, X. Chen, and J. Yu, “Research on ship-radiated noise denoising using secondary variational mode decomposition and correlation coefficient,” *Sensors*, vol. 18, no. 1, p. 48, 2018.
  - [26] Y. Jiang, G. Huang, Q. Yang, Z. Yan, and C. Zhang, “A novel probabilistic wind speed prediction approach using real time refined variational mode decomposition and conditional kernel density estimation,” *Energy Conversion and Management*, vol. 185, pp. 758–773, 2019.
  - [27] G. P. Sun, P. Jiang, H. Xu et al., “Outlier detection and correction for monitoring data of water quality based on improved VMD and LSSVM,” *Complexity*, vol. 2019, Article ID 9643921, 12 pages, 2019.
  - [28] C. Wang, H. Li, G. Huang, and J. Ou, “Early fault diagnosis for planetary gearbox based on adaptive parameter optimized VMD and singular kurtosis difference spectrum,” *IEEE Access*, vol. 7, pp. 31501–31516, 2019.
  - [29] F. T. Wang, C. X. Liu, T. Zhang et al., “Fault diagnosis of rolling bearing based on k-optimized VMD,” *Journal of Vibration, Measurement & Diagnosis*, vol. 38, no. 3, pp. 540–547, 2018.
  - [30] R. Moddemeijer, “A statistic to estimate the variance of the histogram-based mutual information estimator based on dependent pairs of observations,” *Signal Processing*, vol. 75, no. 1, pp. 51–63, 1999.
  - [31] R. Moddemeijer, “On estimation of entropy and mutual information of continuous distributions,” *Signal Processing*, vol. 16, no. 3, pp. 233–248, 1989.
  - [32] 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.
  - [33] G. B. Huang and C. K. Siew, “Extreme learning machine with randomly assigned RBF kernels,” *International Journal of Information Technology*, vol. 11, no. 1, pp. 16–24, 2005.
  - [34] S. Mirjalili, S. M. Mirjalili, and A. Lewis, “Grey wolf optimizer,” *Advances in Engineering Software*, vol. 69, pp. 46–61, 2014.
  - [35] M. Mackey and L. Glass, “Oscillation and chaos in physiological control systems,” *Science*, vol. 197, no. 4300, pp. 287–289, 1977.



- [36] G. H. Li, M. T. Zheng, and H. Yang, "Cycle analysis method of tree ring and solar activity based on variational mode decomposition and hilbert transform," *Advances in Meteorology*, vol. 2019, Article ID 1715673, 8 pages, 2019.
- [37] F. X. Diebold and R. Mariano, "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, vol. 20, no. 1, pp. 134–144, 1995.