Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2021.26059

222 (2021) 156–167
May

Prediction model of dissolved oxygen in marine pasture based on hybrid gray wolf algorithm optimized support vector regression

Baoan Yin[a], Rong Wang[a], Shengbo Qi[a,*], Jingdong Yu[b], Wenliang Jiang[b]

[a]*College of Engineering, Ocean University of China, Qingdao 266100, China, Tel. +86-186-6182-4998; email: qishengbo@ouc.edu.cn (S. Qi), Tel. +86-156-2115-6279; email: yinbaoan@stu.ouc.edu.cn (B. Yin), Tel. +86-178-6532-6597; email: wangrongrong312@163.com (R. Wang)*

[b]*Sencott Intelligent Instrument Co., Ltd., Qingdao 266100, China, Tel. +86-136-5648-7526; email: sencott@126.com (J. Yu), Tel. +86-156-8910-0086; email: 441887538@qq.com (W. Jiang)*

Received 26 September 2019; Accepted 2 May 2020

ABSTRACT

Water quality prediction plays a vital role in water pollution warning and control. However, traditional prediction models usually suffer from low efficiency and poor robustness. To predict accurately the dissolved oxygen concentration in the marine pasture, a dissolved oxygen prediction model, based on wavelet analysis and hybrid gray wolf algorithm (HGWO) optimized vector regression, was established. Because of the water quality data in the marine pasture is a stationary time series. To improve the accuracy of water quality data, wavelet analysis was applied for data pre-processing in this paper. Besides, after the gray wolf algorithm (GWO) was optimized by the differential evolution algorithm (DE), it was used to optimize the support vector regression (SVR). Hence, the SVR's disadvantages of optimization ability and prediction accuracy both were improved. Back propagation neural network (BPNN), SVR, GWO-SVR, DE-SVR, and this model were, respectively, used to predict the dissolved oxygen concentration of Beidaihe marine pasture. The experimental results show that the mean square error, mean absolute error, and average percentage error of the model are 0.1658, 0.359, and 0.0305, respectively, which are better than the traditional prediction model. So this model has higher prediction accuracy and stronger generalization ability, and it can provide a reference for the precise regulation of aquaculture.

Keywords: Wavelet analysis; Gray wolf algorithm; Differential evolution algorithm; Support vector regression; Dissolved oxygen

1. Introduction

Water resources play a vital role in the survival and development of human beings. With the development of industrialization, water pollution has become increasingly serious, which has caused close attention to water resources management departments [1]. Water quality prediction is an important means for water environmental planning, management, and control. It is also an important part of studying water pollution and basic work for water environmental protection and governance [2]. In aquaculture, the dissolved oxygen concentration is an important indicator of the growth status of aquaculture products and can reflect the water quality [3]. Therefore, to prevent the deterioration of water quality and control the outbreak of aquatic product diseases, it is necessary to make an accurate prediction of the dissolved oxygen concentration.

Achieving an accurate prediction of dissolved oxygen concentration has always been a scientific problem in the field of aquaculture [4]. Because the concentration of dissolved oxygen in water is easily affected by many factors, such as hydrometeorology, biology, physical chemistry, etc.

* Corresponding author.

Therefore, it has the characteristics of non-linearity, time lag, and fuzzy uncertainty [5]. Moreover, seawater is corrosive to water quality monitoring equipment, which may cause problems such as equipment aging and data transmission failure. These problems easily lead to data loss, errors, and unavoidable noise in the collected water quality data. If the raw water quality data is directly used for the prediction of dissolved oxygen concentration, the prediction accuracy of the model will be seriously affected [6].

At present, domestic and foreign scholars have conducted a lot of research on the prediction of dissolved oxygen. The most widely used model is the prediction model based on time-series, such as autoregressive integrated moving average model, linear discriminant analysis model, and logistic regression model [7]. These models are simple in structure and require that time-series data must be stable. Therefore, it can only capture linear relationships, and the predictions are not well for non-linear, high-dimensional, and small-sample water quality data [8]. The artificial neural network has good non-linear mapping characteristics and self-learning capabilities. Therefore, it has been widely used in water quality prediction [9]. For example, [10] proposed a water quality prediction model based on FS-RBFNN, which can dynamically change the structure of the algorithm to maintain prediction accuracy [11]. Combined principal component analysis (PCA), genetic algorithm (GA), and back propagation neural network (BPNN) to predict the water quality of rivers and achieved satisfactory results [12]. Established an ANFIS model and used it to predict biochemical oxygen demand, which performed quite well [13]. Combined ANN modeling technology with general regression neural network (GRNN) and multilayer perceptron (MLP), and successfully applied it to the prediction of seawater quality. Nevertheless, it also has some certain disadvantages in water quality prediction, like complex structure, fixed learning rate, slow convergence rate, the randomness of weights, etc. Hence, prediction results are difficult to meet the needs of precise regulation of aquaculture [14–16].

Support vector regression (SVR) follows the principle of structural risk minimization, which is good at solving some problems such as small sample, non-linear, and high dimensional pattern recognition [17]. Moreover, it is different from classical methods of machine learning which follow empirical risk minimization. The SVR model can avoid problems such as overfitting, difficulty in accommodating, and slow convergence [18–20]. Therefore, it has been widely used in many fields, particularly in environmental problems like solar radiation prediction [21], forest modeling [22], and air and water quality estimation, to give some examples [23–25]. However, the actual performance of the SVR depends on the parameters of the learning machine, such as the penalty factor, the nuclear parameter, etc. There is no uniform method in the world to select the parameters of the SVR [26]. At present, the meshing method is often used to find the best penalty factor and kernel parameter. Although the meshing method can find the global optimal solution in the sense of CV, it has low search efficiency and takes a long time. However, the heuristic optimization algorithm can be used to optimize the SVR parameters and improve the operating efficiency of the model [27].

The gray wolf optimization algorithm (GWO) is a new swarm intelligence optimization algorithm proposed by Mirialili et al. (2014) [28]. The GWO algorithm seeks the optimal solution by simulating the social rank and hunting behavior of the gray wolf. It has the advantages of not considering gradient information, simple structure, few parameter settings, and strong global search ability. Therefore, it is widely used in the optimization field of neural networks. However, it has the disadvantages of low accuracy and easy prematureness in solving high-dimensional and multimodal complex function optimization problems [29]. As a representative evolutionary algorithm, the differential evolution algorithm (DE) has strong global search ability and good robustness. Therefore, it works well when solving complex optimization problems such as discontinuous and multipeak. However, there are defects such as weak local search ability, low search efficiency, and the search performance has a certain dependence on parameters [30].

Aiming at the above problems, a dissolved oxygen prediction model, based on wavelet analysis and hybrid gray wolf algorithm (HGWO) optimized vector regression, was established. First, to improve the accuracy of the original data, the multi-scale decomposition characteristics of wavelet analysis were used to reduce the noise of the data. Meanwhile, the DE algorithm was used to optimize the GWO algorithm, which improved the global search ability and the optimal solution search speed of the GWO algorithm and avoids the premature phenomenon. Then, this HGWO was used to optimize the SVR model, which improved the problems of insufficient optimization ability and low prediction accuracy of the SVR model. Finally, BPNN, SVR, GWO-SVR, DE-SVR, and this model were respectively used to predict the dissolved oxygen concentration of Beidaihe marine pasture. The experimental results show that the model has higher prediction accuracy and stronger generalization ability, which can provide a reference for the precise regulation of aquaculture.

2. Experimental area and data source

Beidaihe District (39°47′N–39°53′N, 119°24′E–119°31′E) is located in Qinhuangdao City, Hebei Province, with a vast area and rich aquatic resources. To promote the sustainable development of fisheries, the Beidaihe marine pasture was established in the southwest waters of Jinshanzui. Beidaihe marine pasture is the best water area for the protection of fishery resources and ecological environment restoration in Qinhuangdao. It plans to build 650 ha and has already built 218 ha. It restores the ecological environment by artificially placing fish reefs, fry, and transplanting algae. After efforts, the aquatic species in Beidaihe marine pasture have increased significantly, the continued decline in fishery resources has been effectively curbed, and marine biodiversity has gradually been restored.

As shown in Fig. 1, to monitor water quality in real-time, Beidaihe marine pasture is equipped with a high-precision amlodipine (AML) sensor. It can effectively detect six parameters such as temperature, salinity, depth, pH, chlorophyll, and dissolved oxygen. The AML sensor supports setting acquisition frequency, and we set it to collect every minute. The AML sensor has been operating stably for 2 y, so, we

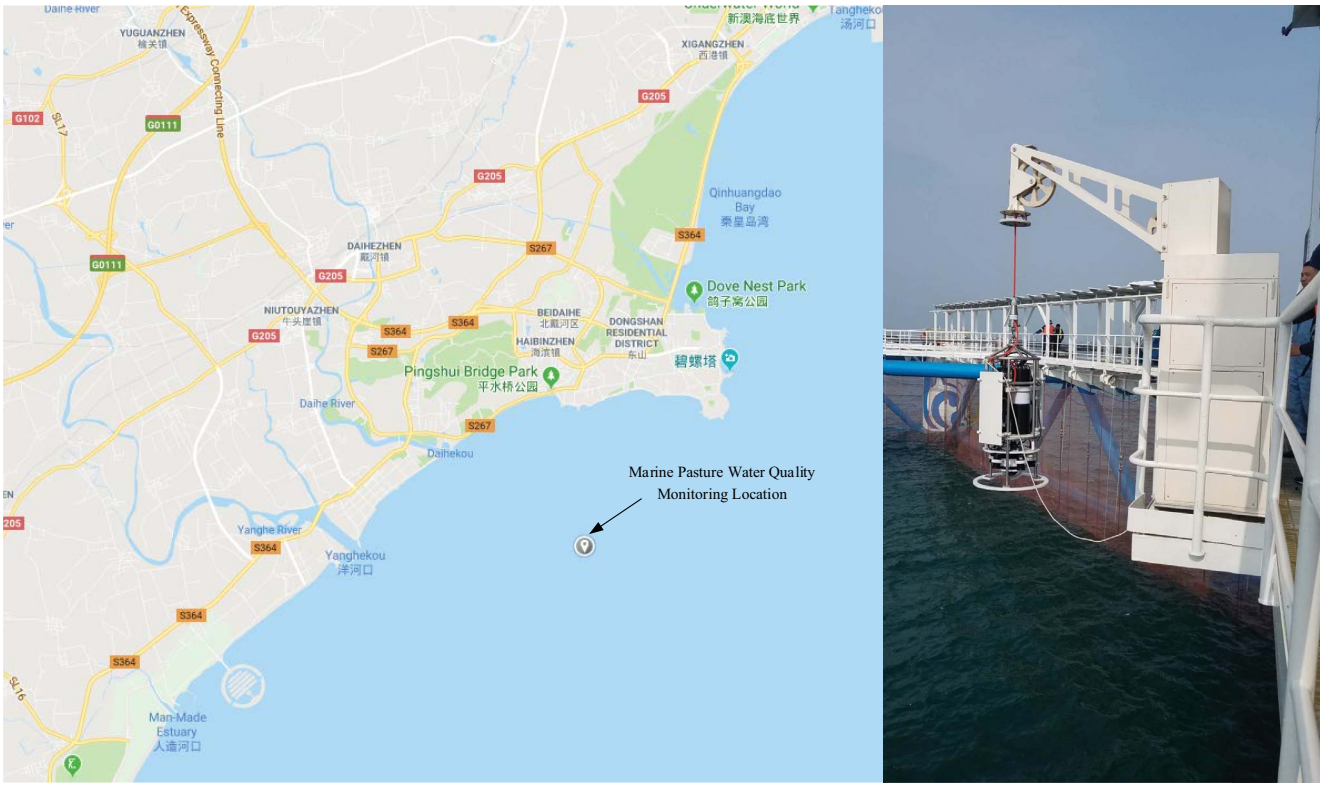

Fig. 1. Water quality monitoring location and monitoring equipment of marine pasture.

can use the continuous-time water quality data it records for experiments.

3. Data preprocessing

3.1. Data repair

Due to sensor failure or network interruption, data loss, and abnormal data may occur during the water quality data collection process. To reduce the data processing cost and improve the accuracy of the prediction model, the data collected by the sensor must be repaired [31].

3.1.1. Data missing repair

If the interval between the missing data is not large, the missing data can be repaired by simple linear interpolation of Eq. (1).

$$x_{k+i} = x_k + \frac{i \times \left(x_{k+j} - x_k\right)}{j}, 0 < i < j \tag{1}$$

where x_k is the water quality data collected at the known time k; x_{k+j} is the water quality data collected at the known time $k+j$; x_{k+i} is the missing water quality data at the time $k+i$.

If there are more missing data, it can be repaired by the weighted average method of Eq. (2).

$$x\left(d,t\right) = \omega_1 x\left(d_1,t\right) + \omega_2 x\left(d_2,t\right) \tag{2}$$

where $x(d,t)$ is the water quality data at the time t on the day d; $x(d_1,t)$ is the water quality data at the time t on the day d_1, d_1, and d have similar weather and they are closest; ω_i is the weight.

3.1.2. Abnormal data repair

The water quality data of marine pasture is a continuous-time series, and the change is relatively stable. Therefore, the moving average method can be used to solve the step-change problem of water quality data in a short time, and its formula is shown in Eq. (3).

$$x_k = \frac{x_{k-1} + x_{k+1}}{2}, \left|x_k - x_{k-1}\right| < \vartheta_1 \text{ or } \left|x_k - x_{k-1}\right| < \vartheta_2 \tag{3}$$

where ϑ_1 and ϑ_2 represent error thresholds between adjacent data.

When the meteorological conditions are similar, the water quality data of the adjacent dates will change little. Therefore, if the meteorological conditions are similar, the water quality data recorded at the same time on the second day is more than ±20% change from the previous day, the data is abnormal and can be processed by the average method of Eq. (4).

$$x_{(d,k)} = \begin{cases} \overline{x_k} + \vartheta_3 \\ \overline{x_k} - \vartheta_3 \end{cases}, \left|x_{(d,k)} - \overline{x_k}\right| > \vartheta_3 \tag{4}$$

where $x_{(d,k)}$ is the water quality data at the time k on day d; ϑ_3 is the error threshold between the acquired data; $\bar{x}_k$ is the average of the data in recent days.

3.2. Wavelet noise reduction

When collecting the water quality data, due to environmental interference or sensor failure, the collected water quality data may contain noise. To reduce noise interference and improve the accuracy of the dissolved oxygen prediction model, the collected data needs to be noise-reduced. Wavelet transform is an emerging signal analysis tool. Its development and ideas come from the Fourier transform. Wavelet analysis not only preserves the advantages of Fourier analysis but also resolves the contradiction between time resolution and frequency resolution [32]. According to the characteristics of water quality, the layered threshold wavelet noise reduction method was used to reduce the noise of water quality data.

- The water quality data collected by the sensor was used as the input sample, and the function *db*3 was selected as the wavelet base function. Then the water quality data was divided into three layers by the Mallat algorithm, and the lowest approximate component coefficients and discrete detail component coefficients of each layer were obtained.
- The soft threshold filtering rule was used to perform threshold quantization on discrete detail components of each layer. In this way, high-frequency interference in the signal was removed. The soft threshold filtering rule is shown in Eq. (5).

$$\text{soft}\left(\omega,\lambda\right)=\begin{cases}\omega+\lambda,\omega\le-\lambda\\0,\left|\omega\right|\le-\lambda\\\omega-\lambda,\omega\ge-\lambda\end{cases} \tag{5}$$

  where ω is the detail component coefficient and λ is the threshold.
- According to the detailed component coefficient and the approximate component coefficient, the Mallat algorithm was used to reconstruct the signal. The reconstructed signal was the water quality signal after noise reduction.

3.3. Data normalization

Because water quality parameters have different dimensions. To improve the prediction accuracy of the model, the water quality data needs to be normalized. In this paper, the water quality data was converted to a number between [0,1], Therefore, the order of magnitude difference between different water quality parameters was eliminated [33]. The specific formula is shown in Eq. (6).

$$x_k=\frac{x_k-x_{\min}}{x_{\max}-x_{\min}} \tag{6}$$

where $x_{\min}$ is the minimum value in the data sequence; $x_{\max}$ is the maximum value in the data sequence.

4. Hybrid gray wolf algorithm optimization SVR prediction model

4.1. SVR algorithm

SVR is based on the VC dimension theory and the principle of structural risk minimization in statistical learning. It uses the idea of kernel function to transform the non-linear problems in low-dimensional space into high-dimensional space. Then, seek a linear regression hyperplane in high-dimensional space to solve the non-linear problem. Even if the sample size is small, it can guarantee an accurate prediction effect and strong generalization ability [34].

Given the training data set $G = \{(x_1,y_1),\ldots,(x_2,y_2)\}$. Where $x_i \in R^n$ is the input eigenvector, $y_i \in R$ is the output target value, $n = 1,2,\ldots i$, and R^n is the n dimensional vector space. The goal of the SVR model is to find a regression function $g(x)$. The deviation of the regression function $g(x)$ from the training set needs to be less than the user-defined insensitive loss function ε. The calculation formula of the regression function $g(x)$ is shown in Eq. (7).

$$y=g(x)=\omega x+b \tag{7}$$

where y is the output; ω is the weight vector which determines the direction of the hyperplane; b is the displacement and determines the distance between the hyperplane and the origin. ω and b can be obtained by the structural risk minimization function of Eq. (8).

$$\min\left(\frac{\left\|\omega\right\|^2}{2}+c\sum_{i=1}^{N}\xi_i+\xi_i^*\right) \tag{8}$$

The constraint conditions are shown in Eq. (9):

$$\begin{cases}y_i-\left(\omega^T x_i+b\right)\le\varepsilon+\xi_i\\\left(\omega^T x_i+b\right)-y_i\le\varepsilon+\xi_i^*\\\xi_i,\xi_i^*\ge0\end{cases} \tag{9}$$

where c is the penalty coefficient, which is used to balance the weight between algorithm complexity and sample error; ε is the deviation between the training set and the actual observations; ξ_i and ξ_i^* are slack variables; α_i and α_i^* are Lagrange multipliers. According to the Eq. (9), the Lagrange function is established to solve the dual problem of the original problem, and finally, the regression function of the optimal hyperplane is obtained. The regression function is shown in Eq. (10).

$$f=\sum_{i}^{N}\left(\alpha_i+\alpha_i^*\right)K\left(x,x_i\right)+b \tag{10}$$

where $K(x,x_i)=\exp\left(-\frac{\left\|x_i-x\right\|^2}{\sigma^2}\right)$ is a kernel function that satisfies the Mercer condition. By comparison, this paper chooses the Gaussian radial basis function as the kernel function. Where σ is the Gaussian kernel width coefficient; x and x_i

are the feature vectors of the training set and the test set, respectively; α_i and α_i^* are Lagrange multipliers.

4.2. Gray wolf optimization algorithm

The gray wolf optimization algorithm is a new type of swarm intelligence optimization algorithm proposed by Mirialili et al. (2014). It obtains the optimal solution by simulating the gray wolf's predation process, which mainly includes tracking, encircling, hunting, and attacking. This algorithm has the advantages of ignoring gradient information, simple structure, and strong global search ability. The gray wolf population has a clear social hierarchy, which can be divided into α, β, δ, and ω according to their status in the population from high to low. Among them, α, β, and δ are responsible for hunting prey, and ω is responsible for tracking prey. Set α as the current best individual in the wolves, β and δ are the second and third best individuals, respectively, and the remaining individuals are ω. In the search space of wolves, α, β, and δ lead ω to search for the optimal region. Through continuous iteration, the prey position is finally found, that is, the global optimal solution [35].

During the predation process, the gray wolf first needs to determine the distance between itself and the prey. The distance formula is shown in Eq. (11).

$$D = \left| C \cdot X_p(t) - X(t) \right| \tag{11}$$

where D is the distance between the prey and the gray wolf; $X_p(t)$ is the position of the prey after t iteration; $X(t)$ is the position of the gray wolf after t iterations; $C = 2r_1$ is the swing factor; r_1 is a random number between [0,1].

Then the gray wolf adjusts its direction and position according to this distance. The location update formula is shown in Eq. (12).

$$X(t+1) = X_p(t) - A \cdot D \tag{12}$$

where $A = 2ar_2 - a$ is the convergence factor; r_2 is a random number between [0,1]; a decreases linearly from 2 to 0 as the number of iterations increases. When $|A| \geq 1$, the gray wolf shows better global search ability, and when $|A| < 1$, it shows better local search ability.

After the gray wolf estimates the position of the prey, α, β, and δ begin to hunt down the prey. The position update formula is shown in Eqs. (13)–(19).

$$D_\alpha = \left| C_\alpha \cdot X_\alpha(t) - X(t) \right| \tag{13}$$

$$D_\beta = \left| C_\beta \cdot X_\beta(t) - X(t) \right| \tag{14}$$

$$D_\delta = \left| C_\delta \cdot X_\delta(t) - X(t) \right| \tag{15}$$

$$X_1 = X_\alpha - A_1 \cdot D_\alpha \tag{16}$$

$$X_2 = X_\beta - A_2 \cdot D_\beta \tag{17}$$

$$X_3 = X_\delta - A_3 \cdot D_\delta \tag{18}$$

$$X_p(t+1) = \frac{X_1 + X_2 + X_3}{3} \tag{19}$$

According to Eqs. (13)–(18), the distance between the gray wolf and its prey can be calculated, and the updated position of the gray wolf can be determined. Besides, the position of the prey can be determined according to Eq. (19).

4.3. Differential evolution algorithm

Differential evolution algorithm (DE) is a heuristic global search algorithm proposed by Storm and Price in 1995. It seeks optimal solutions by simulating mutation, crossover, and selection in the biological evolution mechanism. The algorithm has the advantages of strong global search ability, simple structure, less adjustable parameters, and strong robustness. It is widely used to solve non-linear, high-dimensional complex functions, and linear system optimization problems [36].

First, the DE algorithm needs to initialize the population. Suppose the population consists of M vectors, and the dimension of the vector is D. Besides, the cross probability of the population is CR and the mutation probability is F. Initialize the population to $X(t) = (X_1(t), X_2(t), \ldots X_i(t), \ldots X_M(t))$, where $X_i(t) = (x_{i1}(t), x_{i2}(t), \ldots x_{ij}(t), \ldots x_{iD}(t))$ represents the individual i in the t generation, x_{ij} represents the component j of the individual i. The specific calculation steps are as follows.

4.3.1. Mutation operation

Three individuals are randomly selected from the population for mutation operation, and the mutation formula is shown in Eq. (20).

$$v_{ij}(t+1) = x_{cj}(t) + F \cdot (x_{aj}(t) - x_{bj}(t)) \tag{20}$$

where $v_{ij}(t+1)$ is the population obtained after mutation; t is the current number of iterations; F is the scaling factor, generally within [0,1]; $x_{aj}(g)$, $x_{bi}(g)$, and $x_{cj}(g)$, are three different individuals randomly selected from the population.

4.3.2. Cross operation

First, the cross operation needs to use the function randint(a,b,[]) to generate a random number R. Then, the mutant population is crossed with the original population to realize the diversity of the population. The operation method is shown in Eqs. (21) and (22).

$$R = r \text{ and } \operatorname{int}\left(1,1,\left[1,\mathrm{D}\right]\right) \tag{21}$$

$$V_{ij}(t+1) = \begin{cases} v_{ij}(t+1), r \text{ and} \leq \mathrm{CR} \text{ or } j = R \\ x_{ij}(t), r \text{ and} > \mathrm{CR} \text{ and } j = R \end{cases} \tag{22}$$

where $V_{ij}(t+1)$ is a new population obtained by crossing; r and is a random fraction between [0,1], $R \in [1,\mathrm{N}]$, and $\mathrm{CR} \in [0,1]$.

4.3.3. Select operation

By comparing the fitness values of $V_{ij}(t+1)$ and $X_i(t)$, individuals with better fitness are selected as the next generation. The selection formula is shown in Eq. (23).

$$X_i(t+1)=\begin{cases}V_{i,j}(t+1), f\left(V_{ij}(t+1)\right)<f\left(X_i(t)\right)\\ X_i(t), f\left(V_{ij}(t+1)\right)\geq f\left(X_i(t)\right)\end{cases} \tag{23}$$

where $X_i(t+1)$ is the selected offspring individual; $f(V_{ij}(t+1))$ and $f(X_i(t))$ are the fitness values of $V_{ij}(t+1)$ and $X_i(t)$, respectively. This selection mechanism can ensure that the offspring population is better than the parent population, which improves the average performance of the population.

After the population is mutated, crossed, and selected, a new population with better performance is obtained. The new population is shown in Eq. (24).

$$X(t+1)=\left(X_1(t+1),X_2(t+1),...X_i(t+1),...X_M(t+1)\right) \tag{24}$$

After the new population is generated, the differential evolution algorithm will continue to mutate, cross, and select. When the precision or number of iterations reaches the set value, the iteration stops.

4.4. Hybrid gray wolf algorithm optimization SVR

It can be seen from Section 3.1 (Data repair) that the SVR model needs to set the optimal penalty coefficient c and the radial basis parameter g to determine the optimal hyperplane. However, there is no internationally recognized best method to find the best parameters c and g. At present, the most widely used method is the grid method. Although the grid method can find the global optimal solution in the sense of CV, if you want to find the best parameters c and g in a wider range, it takes a long time and the precision is not high. However, heuristic optimization algorithms can be used to find the optimal parameters c and g to improve the operating efficiency of the algorithm.

The GWO algorithm has the advantages of fast convergence speed and high optimization precision. Its two random self-tuning parameters A and C guarantee that the algorithm can effectively search for the optimal value. However, as the number of iterations increases, A decreases linearly, when $|A| < 1$, it no longer emphasizes exploration. Therefore, the algorithm is easy to fall into local optimum when faced with multi-extreme global optimization problems. Besides, the GWO algorithm is a guided random search algorithm. It's α, β, and δ wolves play an absolute guiding role in the hunting process. If their position is not ideal, it is easy to make the whole wolf group fall into local optimum. However, the crossover, mutation, and selection of the DE algorithm are conducive to maintaining the diversity of the population. Hence, it has better optimization effects for non-continuous, non-differentiable, noisy, and multi-modal complex optimization functions. But its search efficiency is low [37].

To solve the above problems, a dissolved oxygen prediction model based on HGWO-SVR was proposed in this paper. First, the crossover and mutation of the DE algorithm were used to optimize the gray wolf algorithm. Therefore, the diversity of the gray wolf population was guaranteed. Then, the three values with the best fitness in the population were selected as α, β, and δ, respectively. Later, the crossover and selection of the DE algorithm were used to update the position of the gray wolf. Keep iterating until the final position of the prey was determined. Finally, the prey position output from the HGWO model was brought into the SVR model to predict the dissolved oxygen in the marine pasture. In this way, the global search capability of the model was improved, and the shortcoming of the algorithm easily falling into a local optimum was overcome. The algorithm flow chart is shown in Fig. 2. The specific implementation steps are as follows:

- To eliminate the influence of the dimension between the water quality factors, the original water quality data was normalized to the range of [0,1], which improved the prediction accuracy of the model.
- The normalized water quality data was divided into a training set and a test set by the 10-fold cross-validation method. That was, 90% of data was used as a training set and 10% of data was used as a test set.
- Initialized the population size $n$pop, the maximum iteration number MaxIt, the independent variable dimension nVar, the crossover probability pCR, the upper bound of the parameter value ub, the lower bound lb, the upper bound of the scaling factor Fu, and the lower bound of the scaling factor Fl.
- First, the parent and offspring populations of the gray wolf were initialized, and then, the position of each gray wolf was assigned to the parameters c and g of the SVR model, respectively. Finally, the training set was used for the dissolved oxygen regression prediction.
- The individual fitness of each wolf was calculated by Eq. (25). Then, the three wolves with the best fitness of the parent population were selected as α, β, and δ wolves.

$$f=\frac{\sum_{i=1}^{k}\mathrm{Acc}_i}{K} \tag{25}$$

where K = 10, Acc_i is the average accuracy of the i fold cross-validation.

- Set the current number of iterations t and its corresponding a value.
- Initialized the convergence factor A and the swing factor C. After that, the positions of the gray wolf parent population were updated according to Eqs. (13)–(19). Then the position of the parent gray wolf was assigned to the parameters c and g of the SVR model, respectively. Finally, the training set was used for regression prediction, and the individual fitness of each wolf was calculated by Eq. (25).
- First, the gray wolf parent population used the crossover and mutation of the DE algorithm to generate the gray wolf offspring population. Then, the position of the offspring gray wolf was assigned to the parameters c and g

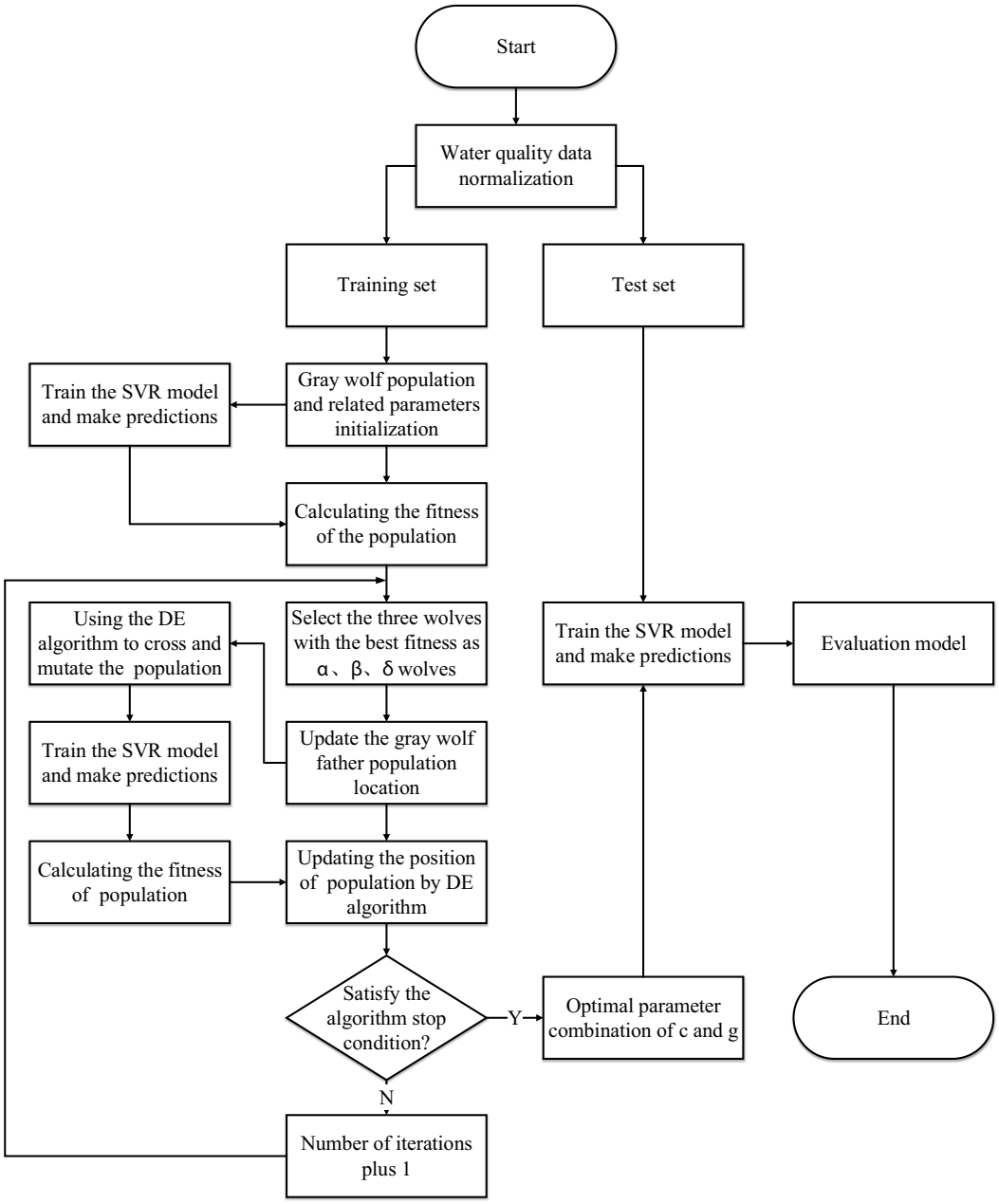

Fig. 2. HGWO-SVR dissolved oxygen prediction model flow chart.

of the SVR model, respectively. Finally, the training set was used for regression prediction, and the individual fitness of each wolf was calculated by Eq. (25).

- First, the individual fitness of the parent population and the offspring population were compared, and then the DE greedy selection algorithm was used to select individuals with better fitness to update the parent population.
- The three wolves with the best individual fitness in the updated parent population were selected as α, β, and δ wolves, and the number of iterations was increased by one.
- If the maximum number of iterations had not been reached, repeat steps (6)–(10). If the maximum number of iterations was reached, the position of the α wolf was assigned to the parameters c and g of the SVR model, respectively. Retrain the SVR model with the training set, and used the test set to predict the dissolved oxygen concentration in the marine pasture. Finally, the performance of the model was evaluated and analyzed.

4.5. Model evaluation method

To verify the performance of the HGWO-SVR model, the model was evaluated by the mean square error (MSE), average absolute error (MAE), average percentage error

(MAPE), and correlation coefficient (R^2). The calculation formula for each indicator is as follows.

$$\mathrm{MSE} = \frac{1}{N}\sum_{i}^{N}\left(y_i - \hat{y}_i\right)^2 \tag{26}$$

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

$$\mathrm{MAPE} = \frac{1}{N}\sum_{i}^{N}\left(y_i - \hat{y}_i\right)\times\frac{100}{y_i} \tag{28}$$

$$R^2 = \frac{\left(\sum_{i=1}^{N}\left(y_i - \bar{y}\right)\left(\hat{y}_i - \bar{y}\right)\right)^2}{\sum_{i=1}^{N}\left(y_i - \bar{y}\right)^2\left(\hat{y}_i - \bar{y}\right)^2} \tag{29}$$

where N is the total number of samples; $\hat{y}_i$ is the predicted value of the model; y_i is the observed value; $\bar{y}$ is the average of the predicted model output.

5. Experimental verification

5.1. Water quality data noise reduction experiment

This paper used the data of Beidaihe marine pasture from April 1 to 10, 2019 for experiments. Every day from 00:00, six kinds of water quality data such as temperature, salinity, depth, pH, chlorophyll, and dissolved oxygen were collected every 1 min. Due to external interference and sensor failure, the water quality data collected by AML sensors must have noise. First, to reduce the impact of erroneous data and improve the reliability of water quality data, the data was repaired by the formulas in Section 3.1 (Data repair). Then, the layered wavelet threshold noise reduction method was used to denoise the water quality data. Due to the limited space of the article, the wavelet noise reduction experiment was only taken as an example of temperature, salinity and depth. The wavelet noise reduction results are shown in Fig. 3.

According to Fig. 3, after using the layered wavelet threshold noise reduction processing, the signal-to-noise ratios of temperature, salinity, and depth are 23.3632, 21.9817, and 28.5008, respectively. The curve becomes smooth,

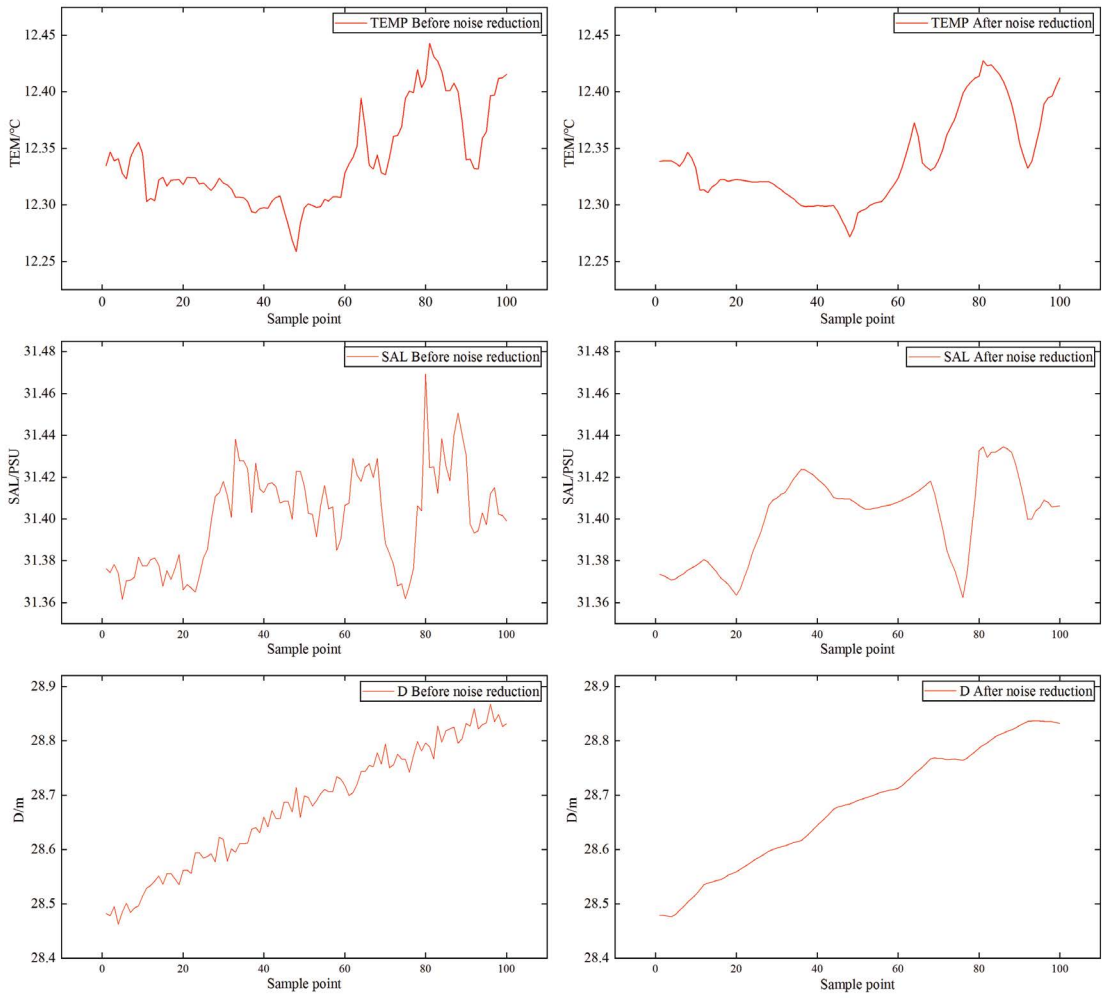

Fig. 3. Water quality data noise reduction experiment results.

besides the effects of noise and peaks are reduced, and the reliability of the data is improved.

5.2. Dissolved oxygen prediction experiment

The water quality data in the marine pasture will not be abrupt in a short time. To improve the efficiency of the model, the water quality data was selected every 10 min. Then, this data set was used for the dissolved oxygen prediction experiment in this paper. The data set was normalized by MATLAB, and then it was scrambled by random sorting. In this model, five kinds of water quality parameters such as temperature, salinity, depth, pH, and chlorophyll were used as input, and the dissolved oxygen was used as output. To ensure the accuracy of the HGWO-SVR dissolved oxygen prediction model, according to the existing research results, the model parameters were set according to Table 1.

First, set the model parameters according to Table 1. Then, the HGWO-SVR model was used to predict the dissolved oxygen of the Beidaihe marine pasture. The model prediction results are shown in Fig. 4, and the model prediction errors are shown in Fig. 5.

From Figs. 4 and 5, the MSE, MAE, MAPE, and R^2 of the HGWO-SVR model are 0.1658, 0.359, 0.0305, and 93.48%, respectively. Therefore, the model has a higher prediction accuracy. To verify the accuracy and superiority of the HGWO-SVR model in the prediction of dissolved oxygen, the HGWO-SVR model was compared with the BP neural network model, SVR model, GWO-SVR model, and DE-SVR model. Under the same conditions, the prediction results of each model are shown in Fig. 6, and the prediction errors are shown in Fig. 7.

As shown in Figs. 6 and 7, the HGWO-SVR dissolved oxygen prediction model has a better fitting degree, and higher prediction accuracy. Therefore, it is more suitable for the prediction of dissolved oxygen in the marine pasture. To comprehensively evaluate the HGWO-SVR model, MSE, MAE, MAPE and R^2 were used to evaluate each model. The evaluation results of each model are shown in Table 2.

As shown in Table 2, in all models, the MSE, MAE, and MAPE indicators of the HGWO-SVR model are the lowest and R^2 is the highest. Besides, the BP neural network model has the highest MSE, MAE, and MAPE indicators, and the lowest R^2. According to the evaluation results, compared with the BP neural network model, the SVR model is more suitable for small sample prediction. Intelligent optimization algorithms such as the GWO algorithm, DE algorithm, and HGWO algorithm can optimize the SVR model and improve the performance of the SVR model. Compared with the BP neural network model, the MSE, MAE, and MAPE indicators of the HGWO-SVR model are reduced by 0.7059, 0.3566, and 0.037, respectively, and the R^2 is increased by 17.04%. Compared with the SVR model, the MSE, MAE, and MAPE indicators of the HGWO-SVR model are reduced by 0.3877, 0.3566, and 0.037, respectively, and the R^2 is increased by 10.73%. Compared with the GWO-SVR model, the MSE, MAE, and MAPE indicators of the HGWO-SVR model are reduced by 0.2873, 0.1596, and 0.0187, respectively, and R^2 is increased by 3.09%. Compared with the DE-SVR model, the MSE, MAE, and MAPE indicators of the HGWO-SVR model are reduced by 0.3903, 0.2329, and 0.0256, respectively, and the R^2 is increased by 3.56%. In summary, compared with BP

Table 1
Dissolved oxygen prediction model parameters

Parameter meaning	Parameter symbol	Ranges
Population size	*n*pop	30
Number of iterations	MaxIt	500
Variable dimension	*n*Var	2
Cross probability	*p*CR	0.2
Parameter value	[lb,ub]	[0.01,100]
Scaling factor	[Fl,Fu]	[0.2,0.8]

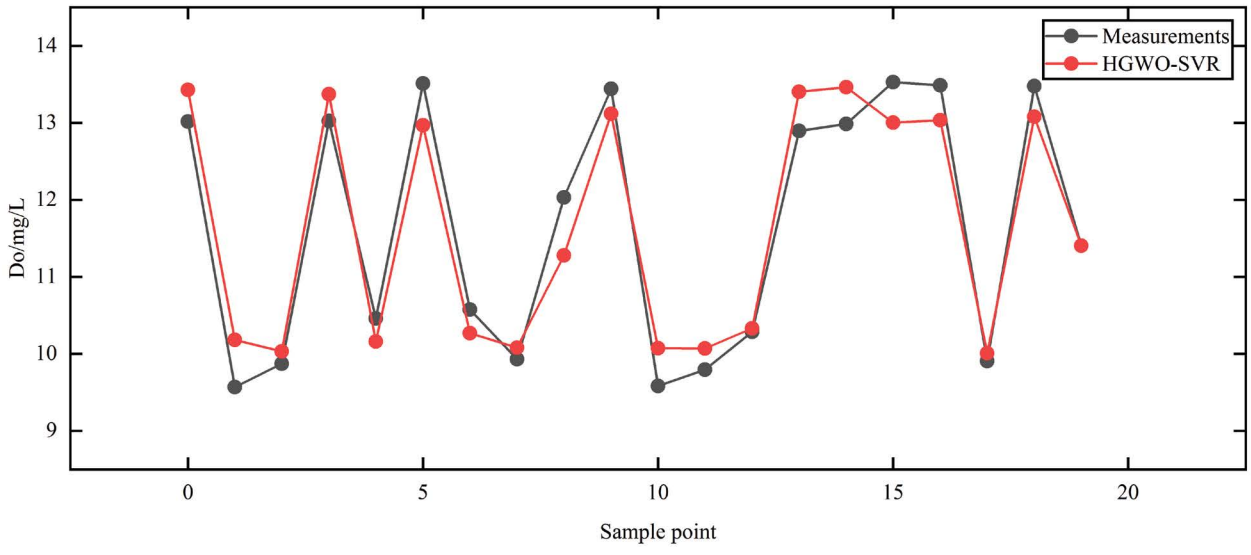

Fig. 4. Water quality data noise reduction experiment results.

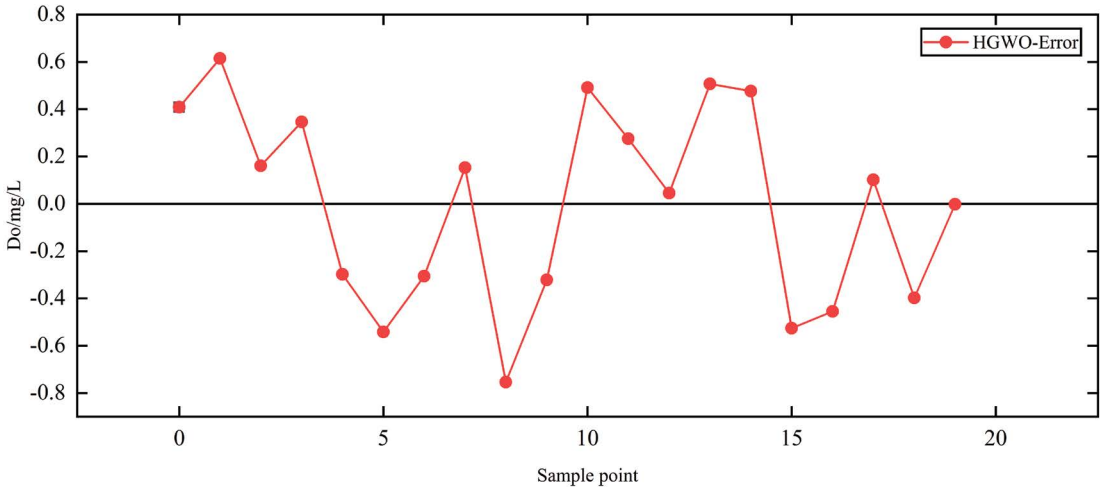

Fig. 5. HGWO-SVR prediction model prediction error.

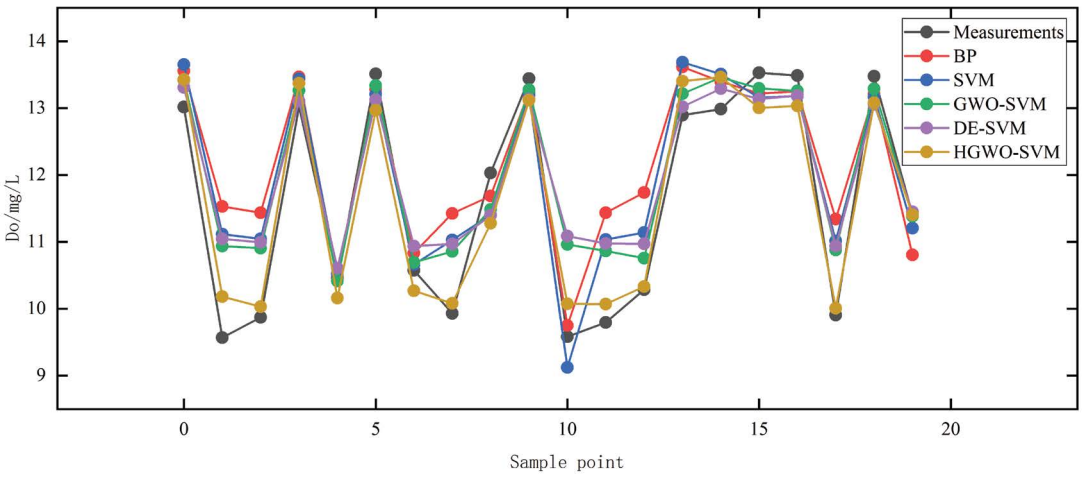

Fig. 6. Do prediction model prediction results.

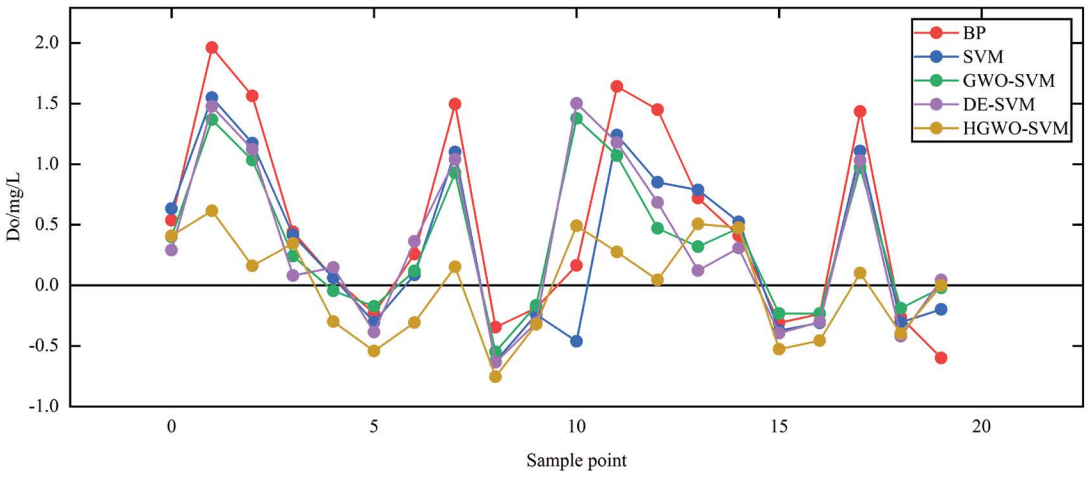

Fig. 7. Do predictive model prediction error.

Table 2
Dissolved oxygen prediction model evaluation index

Model	MSE	MAE	MAPE	R^2
BP	0.8717	0.7156	0.0675	76.44%
SVR	0.5535	0.6171	0.057	82.75%
DE-SVR	0.5561	0.5919	0.0561	89.92%
GWO-SVR	0.4531	0.5186	0.0492	90.39%
HGWO-SVR	0.1658	0.359	0.0305	93.48%

neural network model, SVR model, GWO-SVR model, and DE-SVR, the HGWO-SVR model has higher prediction accuracy. Therefore, it is more suitable for the prediction of dissolved oxygen concentration in the marine pasture.

6. Conclusion

Water quality prediction plays an important role in water quality monitoring, management, and planning. Through the prediction of the dissolved oxygen concentration in water, the changing trend of the dissolved oxygen concentration can be grasped in time. Therefore, it can lay a certain foundation for the precise regulation of aquaculture. Given the shortcomings of the traditional water quality prediction model, such as low prediction accuracy and poor generalization ability, a dissolved oxygen prediction model based on wavelet analysis and HGWO optimization SVR was proposed in this paper. First, the multi-scale decomposition characteristics of wavelet analysis were used to denoise the data collected by the sensor. Therefore, the accuracy of the data was improved. Then, the DE algorithm was used to improve the GWO algorithm, therefore the global search ability of GWO was improved. Besides, the defect that the algorithm was easily trapped into a local optimum was overcome. Subsequently, the HGWO was used to find the best penalty coefficient and kernel parameters of the SVR model. Hence the SVR's disadvantages of optimization ability and prediction accuracy both were improved. Finally, the HGWO-SVR model was used to predict the dissolved oxygen concentration of Beidaihe marine pasture. Besides, to fully evaluate the HGWO-SVR model, the BPNN model, SVR model, GWO-SVR model, and DE-SVR model were used for comparison. The experimental results show that the model has higher prediction accuracy and stronger generalization ability, which can provide a reference for aquaculture regulation.

Acknowledgments

The authors would like to acknowledge the financial support from the National Key Research and Development Project (2016YFC1400800) and the National Natural Science Foundation of China (51475197).

References

[1] L. Li, P. Jiang, H. Xu, G. Lin, D. Guo, H. Wu, Water quality prediction based on recurrent neural network and improved evidence theory: a case study of Qiantang River, Environ. Sci. Pollut. Res., 26 (2019) 879–896.
[2] M. Homami, S.A. Mirbagheri, S.M. Borghei, M. Abbaspour, Simulation modeling of nutrients, dissolved oxygen and total dissolved solids in Peer-Bazar River and Anzali wetland eutrophication prediction, Desal. Water Treat., 79 (2017) 108–124.
[3] O.T. Baki, E. Aras, U.O. Akdemir, B. Yilmaz, Biochemical oxygen demand prediction in wastewater treatment plant by using different regression analysis models, Desal. Water Treat., 157 (2019) 79–89.
[4] S.Y. Liu, H.J. Tai, Q.S. Ding, D.L. Li, L.Q. Xu, Y.G. Wei, A hybrid approach of support vector regression with genetic algorithm optimization for aquaculture water quality prediction, Math. Comput. Modell., 58 (2013) 458–465.
[5] J.J. Carbajal-Hernandez, L.P. Sanchez-Fernandez, L.A. Villa-Vargas, J.A. Carrasco-Ochoa, J.F. Martínez-Trinidad, Water quality assessment in shrimp culture using an analytical hierarchical process. Ecol. Indic., 29 (2013) 148–158.
[6] S.A. Dellana, D. West, Predictive modeling for wastewater applications: linear and nonlinear approaches, Environ. Modell. Softw., 24 (2009) 96–106.
[7] E.V. Hatzikos, G. Tsoumakas, G. Tzanis, An empirical study on sea water quality prediction, Knowledge-Based Syst., 21 (2008) 471–478.
[8] H. Amdevyren, N. Demyr, A. Kanik, Use of principal component scores in multiple linear regression models for prediction of Chlorophyll-a in reservoirs, Ecol. Modell., 181 (2005) 581–589.
[9] X. Ta, Y. Wei, Research on a dissolved oxygen prediction method for recirculating aquaculture systems based on a convolution neural network, Comput. Electron. Agric., 145 (2018) 302–310.
[10] S. Palani, S.Y. Liong, P. Tkalich, An ANN application for water quality forecasting, Mar. Pollut. Bull., 56 (2008) 1586–1597.
[11] J.H. Cho, S.K. Seok, H.S. Ryong, A river water quality management model for optimizing regional waster treatment using a genetic algorithm, J. Environ. Manage., 73 (2004) 229–242.
[12] H.G. Han, Q.L. Chen, J.F. Qiao, An efficient self-organizing RBF neural network for water quality prediction, Neural Networks, 24 (2001) 717–725.
[13] Y.R. Ding, Y.J. Cai, P.D. Sun, B. Chen, The use of combined neural networks and genetic algorithms for prediction of river water quality, J. Appl. Res. Technol., 12 (2004) 493–499.
[14] M. Mahmoodabadi, R.R. Arshad, Long-term evaluation of water quality parameters of the Karoun River using a regression approach and the adaptive neuro-fuzzy inference system, Mar. Pollut. Bull., 126 (2018) 372–380.
[15] H.M. Lee, C.M. Chen, T.C. Huang, Learning efficiency improvement of back-propagation algorithm by error saturation prevention method, Neurocomputing, 41 (2001) 125–143.
[16] H.C. Neiad, M. Farshad, F.N. Rahatabad, O. Khayat, Gradient-based back- propagation dynamical iterative learning scheme for the neuro-fuzzy inference system, Expert Syst., 33 (2016) 70–76.
[17] B. Scholkopf, A.J. Smola, R. Williamson, P. Bartlett, New support vector algorithms, Neural Comput., 12 (2000) 1207–1245.
[18] T. Hansen, C.J. Wang, Support vector based battery state of charge estimator, J. Power Sources, 141 (2005) 351–358.
[19] X. Li, D. Lord, Y. Zhang, Y. Xie, Predicting motor vehicle crashes using Support Vector Machine models, Accid. Anal. Prev., 40 (2008) 1611–1618.
[20] F. Nagata, K. Tokuno, K. Mitarai, Defect detection method using deep convolutional neural network, support vector machine and template matching techniques, Artif. Life Rob., 24 (2019) 512–519.
[21] V.H. Quej, J. Almorox, J.A. Arnaldo, L. Saito, ANFIS, SVM and ANN soft-computing techniques to estimate daily global solar radiation in a warm sub-humid environment, J. Atmos. Sol. Terr. Phys., 155 (2017) 62–70.
[22] P.J. Garcia Nieto, J. Martinez Torres, M. Araujo Fernandez, C. Ordonez Galan, Support vector machines and neural networks used to evaluate paper manufactured using *Eucalyptus globulus*, Appl. Math. Modell., 36 (2012) 6137–6145.
[23] A. Suarez Sanchez, P.J. Garcia Nieto, P. Riesgo Fernandez, F.J. Iglesias Rodriguez, Application of an SVM-based regression model to the air quality study at local scale in the Aviles urban area (Spain), Math. Comput. Modell., 54 (2011) 1453–1466.

[24] P.J. Garcia Nieto, E.F. Combarro, J.J. del Coz Diaz, E. Montanes, A SVM-based regression model to study the air quality at local scale in Oviedo urban area (Northern Spain): a case study, Appl. Math. Comput., 219 (2013) 8923–8937.
[25] P.J. García Nieto, E. Garcia Gonzalo, J.R. Alonso Fernandez, C. Diaz Muniz, Hybrid PSO–SVM-based method for long-term forecasting of turbidity in the Nalon river basin: a case study in Northern Spain, Ecol. Eng., 73 (2014) 192–200.
[26] Y.G. Oh, M. Busogi, K. Ransikarbum, D. Shin, D. Kwon, N. Kim, Real-time quality monitoring and control system using an integrated cost effective support vector machine, J. Mech. Sci Technol., 33 (2019) 6008–6020.
[27] S. Kulkarni, G. Harman, An Elementary Introduction to Statistical Learning Theory, Wiley, New York, NY, 2011.
[28] S. Mirjalili, S.N. Mirjalili, A. Lewis, Grey wolf optimizer, Adv. Eng. Softw., 69 (2014) 46–61.
[29] B. Xiao Qiang, Z. Lu, D. Zhi Min, C. Jing, Z. Jian Ye, Prediction of sulfur solubility in supercritical sour gases using grey wolf optimizer-based support vector machine, J. Mol. Liq., 261 (2018) 431–438.
[30] L.Z. Cui, G.H. Li, Z.X. Zhu, Z.K. Wen, N. Lu, J. Lu, A novel differential evolution algorithm with a self-adaptation parameter control method by differential evolution, Soft Comput., 22 (2018) 6171–6190.
[31] J.S. Chou, C.P. Yu, D.N. Truong, B. Susilo, A.Y. Hu, Q. Sun, Predicting microbial species in a river based on physicochemical properties by bio-inspired metaheuristic optimized machine learning, Sustainability, 11 (2019), doi: 10.3390/su11246889.
[32] S.H. Wang, Y. Li, Y. Shao, C. Cattani, Y.D. Zhang, S.D. Du, Detection of dendritic spines using wavelet packet entropy and fuzzy support vector machine, CNS Neurol. Disord. Drug Targets, 16 (2017) 116–121.
[33] H.W. Liu, K. Guo, Z.Y. Zhang, D.D. Yu, J.X. Zhang, F.P. Ning, High-power LED photoelectrothermal analysis based on backpropagation artificial neural networks, IEEE Trans. Electron Devices, 64 (2017) 2867–2873.
[34] W.Z. Lu, W.J. Wang, Potential assessment of the "support vector machine" method in forecasting ambient air pollutant trends, Chemosphere, 59 (2005) 693–701.
[35] X.P. Liao, G. Zhou, Z.K. Zhang, J. Lu, J.Y. Ma, Tool wear state recognition based on GWO-SVM with feature selection of genetic algorithm, Int. J. Adv. Manuf. Technol., 104 (2019) 1051–1063.
[36] N.M. Hatta, A.M. Zain, R. Sallehuddin, Z. Shayfull, Y. Yusoff, Recent studies on optimisation method of Grey Wolf Optimiser (GWO): a review (2014–2017), Artif. Intell. Rev., 52 (2019) 2651–2683.
[37] A. Korashy, S. Kamel, F. Jurado, A.R. Youssef, Hybrid whale optimization algorithm and grey wolf optimizer algorithm for optimal coordination of direction overcurrent relays, Electr. Power Compon. Syst., 47 (2019) 644–658.