IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 14 August 2023, accepted 28 August 2023, date of publication 30 August 2023, date of current version 5 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3310250

RESEARCH ARTICLE

A Method Based on Improved Ant Colony Algorithm Feature Selection Combined With GA-SVR Model for Predicting Chlorophyll-a Concentration in Ulansuhai Lake

CHENHAO WU[1], XUELIANG FU[1,2], HONGHUI LI[1], HUA HU[1,2], XUE LI[1], AND LIQIAN ZHANG[1]

[1]College of Computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot 010018, China
[2]Inner Mongolia Autonomous Region Key Laboratory of Big Data Research and Application of Agriculture and Animal Husbandry, Hohhot 010018, China

Corresponding author: Xueliang Fu (fuxl_imau@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 62041211 and Grant 61962047; in part by the National Key Research and Development Program of China under Grant 2019YFC049205; in part by the Natural Science Foundation of Inner Mongolia Autonomous Region of China under Grant 2020MS06011, Grant 2019MS06015, and Grant 2021MS06009; in part by the Basic Research Operation Funds for Universities under Inner Mongolia Autonomous Region under Grant BR22-14-05; in part by the Research Program of Science and Technology at Universities of Inner Mongolia Autonomous Region under Grant NJZZ23044; and in part by the Program for Innovative Research Team in Universities of Inner Mongolia Autonomous Region under Grant NMGIRT2313.

ABSTRACT Chlorophyll-a (Chl-a) is an important parameter of water bodies, but due to the complexity of optics in water bodies, it is currently difficult to accurately predict Chl-a concentration in water bodies by traditional methods. In this paper, Sentinel-2 remote sensing images is used as the data source combined with measured data, and Ulansuhai Lake is taken as the study area. An adaptive ant colony exhaustive optimization (A-ACEO) algorithm is proposed for feature selection and combined with a novel intelligent algorithm of optimizing support vector regression (SVR) by genetic algorithm (GA) for prediction of Chl-a concentration. The ant colony optimization (ACO) algorithm is improved to select remote sensing feature bands for Chl-a concentration by introducing relevant optimization strategies. The GA-SVR model is built by optimizing SVR using GA with the selected feature bands as input, and comparing with the traditional SVR model. The simulation results show that under the same conditions, using A-ACEO algorithm to select feature bands as inputs can effectively reduce the model complexity, and improve the model prediction performance, which provides a valuable reference for monitoring Chl-a concentration in lakes.

INDEX TERMS Chlorophyll-a (Chl-a), genetic algorithm, lake, machine learning algorithm, remote sensing, support vector regression.

I. INTRODUCTION

Lake wetland ecosystem is an important ecosystem in the world [1] and also the guarantee for the survival of human beings, animals, and other organisms [2]. However, due to the interference of human activities, various water resource problems have emerged, and the lakes have difficulties in maintaining normal ecosystem functions [3]. Chlorophyll-a (Chl-a) is an important carrier of algal organisms, and its concentration is also an important indicator for assessing the degree of algae and eutrophication in water bodies [4].

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

Ulansuhai Lake wetland ecosystem, not only has the characteristics of a typical cold and arid lake [5], but also carries the receding water of agricultural fields and the discharge of industrial pollution wastewater in the region [6]. Therefore, it is of great practical significance to carry out the monitoring of Chl-a concentration by taking Ulansuhai Lake as the study area to understand the lake water ecosystem and manage the water environment.

Traditional monitoring techniques are not capable of large-scale water quality monitoring due to high cost, time-consuming and susceptibility to external conditions [7]. Remote sensing as an important monitoring technology has become an effective method for monitoring

Chl-a concentration in lakes due to the advantages of fast speed, low cost and wide detection range [8]. In previous studies, most scholars predict Chl-a concentration based on the bands contained in remote sensing data. For the feature bands existing in remote sensing data, it can be roughly divided into single-band, multi-band combination and full-band combination predictive modeling. When a single band is used for modeling, the band data is less sensitive to Chl-a, and the performance of the model will be affected [9]. When full-band modeling is used, the feature bands will interfere with each other due to the influence of multicollinearity, which will also affect the accuracy of model prediction [10]. Therefore, when analyzing the feature bands of remote sensing satellites, determining the feature variables of Chl-a is the first prerequisite to improve the operation efficiency of the prediction model, simplify the model structure, and enhance the model stability, and it is more common to select the band combination most related to the concentration of Chl-a [11].

The feature band selection methods of remote sensing data can be divided into two types. One is the selection method based on mathematical statistical characteristics. Some commonly used methods are correlation coefficient analysis (CC), successive projections algorithm (SPA), competitive adaptive reweighted sampling method (CARS), etc. For example, Zhang et al. [12] used CC to analyze band data in Landsat 8 and measured Chl-a concentrations in Donghu Lake, China, where the feature band with a high correlation coefficient with Chl-a was selected to construct a prediction model. Zhang et al. [13] used SPA to analyze the sensitive band of the remote sensing feature data, and selected the most sensitive band to construct the Chl-a concentration prediction model of Qinghai Lake. Liu et al. [14] used CARS to select the optimal feature bands, and built a soil organic matter prediction model based on the selected feature bands. Such methods can eliminate part of the redundant information to some extent, but there are still a large number of variables after selection, and the stability of the model is also poor. The other is a new band feature extraction method based on group intelligent optimization algorithm. For example, Deng et al. [15] uses GA to select feature bands, which effectively reduces the data dimension and achieves good results. Yang et al. [16] used particle swarm optimization (PSO) and ACO algorithms for feature band selection and demonstrated that irrelevant feature bands can be filtered out, signal-to-noise ratio can be reduced, and the accuracy of the prediction model can be improved. Such algorithms solve problems by simulating the way groups behave in nature, and have a rigorous theoretical basis, providing feasible solutions to complex problems that cannot be addressed in conventional methods.

The traditional remote sensing prediction methods can be roughly divided into empirical method, semi-empirical method and analytical method. The empirical method [17] is relatively simple and has a good effect in linear relation, but the relationship between remote sensing spectral features and water body elements is difficult to be expressed by simple linear function, and the adaptability of empirical method is usually poor. Semi-empirical method [18] and analytical method [19] are prediction methods that mainly utilize the inherent optical characteristics in water bodies and combine with remote sensing reflectivity. The demand for optical data is large, the model is complex, and it is difficult to obtain a large amount of data in practical applications, so they are not suitable for wide application. At present, machine learning, as a new method, has been widely used in remote sensing prediction. For example, Xu et al. [20] built three machine learning models to study Chl-a concentration in Xingkai Lake, and the results show that combining machine learning algorithms with remote sensing technology can achieve large-scale monitoring more effectively. Zhao et al. [21] studied and analyzed the concentration of Chl-a in Taihu Lake and constructed multiple linear models and multiple machine learning models for prediction. The results showed that the prediction accuracy of the machine learning model was more accurate than that of the linear model. SVR, as an important machine learning algorithm, has shown strong usability in the field of remote sensing because of its good performance in solving small sample data and nonlinear problems. For example, Guo et al. [22] constructed a variety of machine learning methods to predict dissolved oxygen in Lake Huron, and the results show that SVR model has higher prediction accuracy and better model stability and generalization ability than other machine learning models. He et al. [7] analyzed the concentration of Chl-a in lakes and constructed SVR model and other methods for prediction. The results showed that SVR model was more effective and promising in predicting the concentration of Chl-a. In general, the traditional remote sensing prediction method is relatively simple, but it cannot be widely used in remote sensing prediction due to its existing characteristics, while machine learning, as a new method, has been widely used in remote sensing prediction. SVR also has strong availability in remote sensing prediction due to its existing advantages.

In summary, taking Ulansuhai Lake as the study area, combining Sentinel-2 remote sensing image and measured Chl-a concentration data, an adaptive ant colony exhaustive optimization (A-ACEO) algorithm is proposed by analyzing the correlation between Chl-a and remote sensing features. The selected feature bands are taken as input, and the advantages of artificial intelligence optimization algorithm are comprehensively considered. By introducing GA to optimize SVR parameters, the Chl-a concentration prediction model was constructed. Finally, the model was analyzed and the performance of the method was verified.

II. STUDY AREA AND DATA SETS

A. STUDY AREA

Ulansuhai Lake (40°36′ N-41°03′ N, 108°43′ E-108°57′ E), as shown in Fig. 1, located in Bayan Nur City, Inner Mongolia, is the biggest lake wetland at the same latitude in the world. The lake covers an area of 325.31km^2, in which 123.11km^2 is open water and the remaining is reed area,

with a reservoir capacity of about 2.5 to 3 × $10^8 m^3$. The wetland of Ulansuhai Lake is relatively rich in species, inhabited by a large number of fish and birds. However, due to the continuous exploitation of Ulansuhai Lake, the lake is decreasing in size, while the industrial and agricultural wastewater discharged into the lake is increasing year by year and a large number of fish are dying, resulting in serious ecological damage of the lake and gradual degradation of ecosystem function.

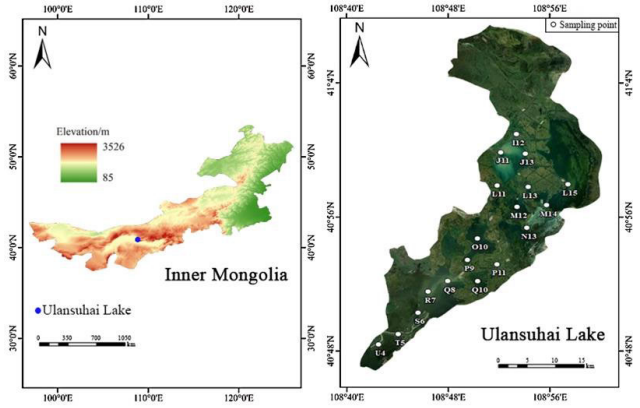

FIGURE 1. Location and sampling point distribution of study area.

B. MEASURED DATA ACQUISITION

The measured data is provided by the national research team of Inner Mongolia Agricultural University, ''River and Lake Wetland Water Environmental Protection and Restoration Technology Research and Innovation Talent Team'', which is dedicated to the research of lake wetlands in the northern cold and arid region for a long time, providing comprehensive big data management and sharing services for ecological protection in China. Ulansuhai Lake starts to freeze in early November every year, and the freezing and thawing period can be up to 5 months, so the sampling time is mainly concentrated in June to September, and the sampling is fixed to the middle and end of each month, and the sampling depth is 0.5m down from the water surface vertically. The measured data of Chl-a concentration were collected from 2015 to 2018 with 92 samples, and some measured data are shown in TABLE 1. The 92 samples were randomly divided into 64 samples as training set and 28 samples as test set.

C. MEASURED DATA ACQUISITION

Developed by the European Space Agency (ESA), the Sentinel-2 satellite can be used to monitor remote sensing images of terrestrial ecology, inland rivers and the coastal areas. It is equipped with a multispectral imager with 13 spectral bands and three spatial resolutions of 10m, 20m and 60m. Specific satellite data are shown in TABLE 2. The image data are available on the ESA website (https://scihub.copernicus.eu/dhus/#/home).

The quasi-synchronous Sentinel-2L1C class remote sensing image data synchronized with or one day removed from the date of the measured Chl-a concentration data were downloaded through ESA for pre-processing. The pre-processing of Sentinel-2 remote sensing image data includes cropping, resampling, geometric correction and band reflectance extraction to obtain remote sensing reflectance in each band. For atmospheric correction of the data, the Sentinel-2L2A level image data was acquired using Sen2Cor plug-in from ESA. When the data was resampled, the S2 Resampling Processor in SNAP software is used and the resolution is set to 20m using the nearest-neighbour method. The SNAP software is available from ESA website (https://step.esa.int/main/download/snap-download/).

TABLE 1. Selected measured data from 2015-2018.

Date of actual test	Remote Sensing Image Dates	Sampling point	Measured Chl-a concentration (μg/L)
2015/9/25	2015/9/25	I12	17.014
2016/6/20	2016/6/21	S6	4.704
2017/6/25	2017/6/26	P9	5.017
2017/8/29	2017/8/30	Q8	19.064
2017/9/25	2017/9/24	J13	30.289
2018/7/26	2018/7/26	M14	7.451

III. METHODS

A. REMOTE SENSING DATA PRE-PROCESSING METHOD

Fig. 2(a) shows the reflectance of the original band. Since the remote sensing image is greatly affected by weather, time, angle and other factors, the quality of the remote sensing image may be reduced, resulting in errors in the obtained band reflectance. Therefore, the second derivative (SD) method was used in this paper to preprocess the original bands in order to reduce the influence of external factors on the remote sensing data. The reflectivity of the waveband after pre-processing is shown in Fig. 2(b) After pretreatment, the feature band is selected.

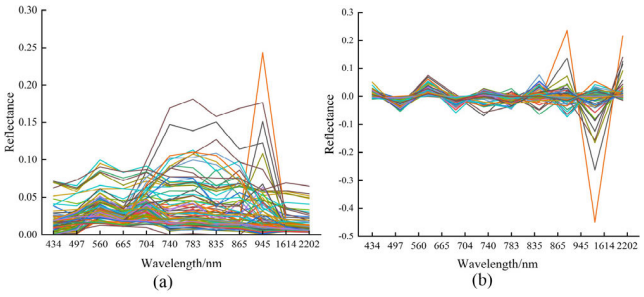

FIGURE 2. (a) Original and (b) pretreatment band reflectance.

B. FEATURE BAND SELECTION METHOD

1) COMPETITIVE ADAPTIVE REWEIGHTED SAMPLING

Competitive adaptive reweighted sampling (CARS) takes the Darwin's ''survival of the fittest'' as the theoretical idea [23], which treats each feature as an individual and selects the

TABLE 2. Sentinel-2 sensor spectral characteristic information.

Band Number	Sentinel-2A		Sentinel-2B		Spatial resolution(m)
	Central wavelength (nm)	Bandwidth(nm)	Central wavelength (nm)	Bandwidth(nm)	
1	433.9	27	442.3	45	60
2	496.6	98	492.1	98	10
3	560.0	45	559.0	46	10
4	664.5	38	665.0	39	10
5	703.9	19	703.8	20	20
6	740.2	18	739.1	18	20
7	782.5	28	779.7	28	20
8	835.1	145	833.0	133	10
8A	864.8	33	864.0	32	20
9	945.0	26	943.2	27	60
10	1373.5	75	1376.9	76	60
11	1613.7	143	1610.4	141	20
12	2202.4	242	2185.7	238	20

individual with strong adaptability. The specific steps are shown as follows:

1) N samples were randomly selected in the feature band by using the Monte Carlo algorithm, and the partial least squares regression (PLSR) model was constructed.
2) The absolute value of the regression coefficient of the model is calculated, and the variables are selected by the exponential decay function and the adaptive reweighted sampling algorithm, where the high regression coefficients are retained and the low regression ones are removed. The exponential decay function retains the proportion of variables as shown in equation (1):

$$\mathrm{R}_i = ae^{-ki} \tag{1}$$

where a and b are constants. Based on the first calculation with all variables and the N-th calculation with 2 variables, i.e. $r_1 = 1$ and $r_2 = \frac{2}{p}$, so a and b are shown in equations (2)-(3):

$$a = (\frac{p}{2})^{\frac{1}{N-1}} \tag{2}$$

$$k = \frac{In(\frac{p}{2})}{N-1} \tag{3}$$

where p is the variable number.

3) The retained variables were used as a new subset and the PLSR model was constructed while calculating the root mean square error of cross validation (RMSECV).
4) Step (1)-(3) are repeated. After N times of calculation, N samples will get N variable subsets, and N variable subsets will get N RMSECVs. Finally, the smallest variable subset in RMSECVs is selected as the optimal feature band combination.

2) STANDARD ANT COLONY OPTIMIZATION ALGORITHM

The ant colony optimization (ACO) algorithm is a swarm intelligence optimization algorithm that simulates the foraging behavior of ants [24]. ACO algorithm was originally designed to solve static combinatorial problems, but it also has good applicability in dynamic combinatorial optimization problems [25]. Therefore, ACO presents feasibility in the feature band selection method.

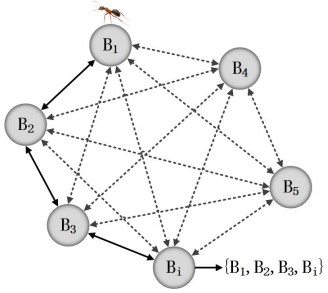

FIGURE 3. Visualization of ACO algorithm to select feature bands.

As shown in Fig. 3, each node in the figure corresponds to a feature band, where $\{B_1, B_2, \ldots, B_i\}$ is the set of original feature bands. An ant randomly starts from a node and selects another node according to the rules, after a period of traversal, a subset of original feature bands $\{B_1, B_2, B_3, B_i\}$ will be obtained, and if this subset satisfies the stopping condition, it is judged to be a feasible solution. The ACO algorithm is shown below:

1) Initialization of the pheromone concentration. The initial pheromone of all nodes is set to 1, and the initial nodes are randomly selected according to the roulette method.
2) Node selection probability. The ant selects the next node according to the pheromone concentration presented at the node, and the probability from node B_i to node B_j is shown in equation (4):

$$P_{ij}(t) = \begin{cases} \dfrac{\tau_{ij}(t)}{\sum_J \tau_{ij}(t)}, & j \in J \\ 0, & \text{Other} \end{cases} \tag{4}$$

where t is the iteration number, τ is the node pheromone concentration, and J is the collection of unselected nodes reachable by ants at node B_i.
3) Feature subset objective function selection. The root mean square error (RMSE) in the GA-BP model is used as the basis for the calculation of the feature subset function F in the ant colony algorithm as shown in equation (5):

$$F = \frac{C}{1 + RMSE} \tag{5}$$

where C is a constant term that is usually set to 1. The equation shows that the smaller the RMSE, the easier the subset is selected.
4) Pheromone concentrations updating. Once all ants have completed one iteration, the node pheromone concentration will be updated, the node pheromone concentration within the selected subset set increases, and the remaining node pheromone concentration volatilizes, as shown in equations (6)-(7):

$$\tau_{ij}(t+1) = (1-\rho)\,\tau_{ij}(t) + \tau'_{ij}(t) \tag{6}$$

$$\tau'_{ij}(t) = \begin{cases} F_{<i,j>}, & <i,j> \in B \\ 0, & Other \end{cases} \tag{7}$$

where ρ is the pheromone volatility factor between (0,1), B is the set of selected subsets of feature bands.

3) ADAPTIVE ANT COLONY EXHAUSTIVE OPTIMIZATION ALGORITHM

Although standard ACO algorithm has good advantages in solving dynamic combinatorial problems, the initial parameters tend to influence the algorithm, which can easily lead to too slow convergence and relatively low efficiency [26]. The adaptive ant colony exhaustive optimization algorithm (A-ACEO) introduces the corresponding adaptive strategy based on the standard ACO algorithm, so that it can find the optimal parameter value adaptive, shorten the running time of the algorithm and improve the efficiency. In addition, the combination of ACO algorithm and exhaustive method [27] is used to increase the diversity of subset features and avoid the defects of exhaustive search method, which greatly reduces the search time. The A-ACEO algorithm process is as follows:

1) Adaptive adjustment strategy of initialized pheromone concentration. In standard ACO algorithm, the initial pheromone concentration factor of the nodes is the same fixed value and the ant randomly selects a node as the initial node, which makes the algorithm inefficient. To solve the above problem, the reciprocal of RMSE corresponding to each node is used as the initial pheromone concentration based on GA-SVR model to guide ants in selecting the initial nodes, which avoids the shortage of the ants finding nodes randomly in the iterations.
2) Adaptive adjustment strategy of volatilization factor ρ. In standard ACO algorithm, the volatility factor is a fixed value, if the value is not set reasonably will affect the convergence speed, or even lose the global search ability. To solve the above problem, ρ is made adaptive adjustment strategy to enhance the algorithm search capability, as shown in equation (8):

$$\rho(t+1) = \frac{T}{t+T} e^{1-\rho(t)} \tag{8}$$

where t and T are the current and maximum iteration number respectively. It is clear that initially, ρ has a larger value, with more and more iterations, ρ decreases, the probability of the optimal node gradually increases.
3) Adaptive updating strategy of pheromone concentration. The pheromone concentration update strategy is improved with standard ACO algorithm as shown in equation (9):

$$\tau'_{ij}(t) = \begin{cases} \lambda_{ij} + Q \times \dfrac{1}{\mathrm{RMSE}_{ij}(t)}, & <i,j> \in B \\ 0, & Other \end{cases} \tag{9}$$

where λ_{ij} is the initialized pheromone concentration of the node selected by the ant, Q is the pheromone heuristic factor, B is the set of the selected subset of feature bands, $\mathrm{RMSE}_{ij}(t)$ is the RMSE of the selected subset of nodes at the t-th iteration.
4) Optimal threshold and optimal matrix strategy. Combining the standard ACO algorithm with the exhaustive enumeration method, an optimal threshold and optimal matrix strategy is proposed. The subset of nodes generated by the iteration is filtered by the optimal threshold, and only the subset of nodes better than the optimal threshold is deposited into the optimal matrix. Finally, the optimal node subset is filtered in the optimal matrix, which is the optimal feature subset. By filtering, this strategy significantly reduces the space of data storage and the consumption of computational resources. The optimal threshold is shown in equation (10):

$$O(t) = \frac{1}{RMSE_{xy}(t)}, <x,y> \in B_m \tag{10}$$

where B_m is the pheromone concentration set of nodes in order from highest to lowest, RMSE_{xy} is the RMSE of the set of subsets of nodes with the highest pheromone concentration.

C. PREDICTION METHOD FOR CHL-A CONCENTRATION

In order to improve the accuracy of model prediction. This paper proposes a GA-SVR prediction model by optimizing the SVR with GA, based on SVR.

1) SUPPORT VECTOR REGRESSION

Support vector regression (SVR), originally proposed by the scholar Vapnik [28]. SVR has fewer adjustable parameters and fewer overfitting problems [29]. Therefore, SVR has good feasibility in multidimensional function problems.

For a set of experimental data $\{(x_1, y_1), (x_2, y_2), \cdots, (x_i, y_i)\}$, where $x_i \in R^n$ is the input, $y_i \in R$ is the output, and n is the number of data, the input and output can be related by defining a linear regression function f(x) [30]. f(x) is shown in equation (11):

$$f(x) = \omega \cdot \Phi(x) + b \tag{11}$$

where ω and b are the weight vector and deviation, respectively. $\Phi(x)$ is a nonlinear function which is used to map the original input dataset to the high dimensional feature space. ω and b are calculated as shown in equation (12), based on the minimization of structural risk:

$$\begin{aligned} &minR(\omega) = \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}(\xi_i^- + \xi_i^+) \\ &s.t.\begin{cases} y_i - (\omega \cdot \Phi(x_i) + b_i) \le \varepsilon + \xi_i^+ \\ (\omega \cdot \Phi(x_i) + b_i) - y_i \le \varepsilon + \xi_i^- \\ \xi_i^-, \xi_i^+ \ge 0, i = 1, 2, \cdots, n \end{cases} \end{aligned} \tag{12}$$

where C is the penalty factor, ξ_i^- and ξ_i^+ are the relaxation variables, and ε is the insensitivity coefficient. By introducing Lagrange multipliers and constructing a Lagrange function, the function minimization problem is transformed into a pairwise problem whose dual is solved as shown in equation (13):

$$\begin{aligned} &\max \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)y_i - \sum_{i=1}^{n}(\alpha_i + \alpha_i^*)\varepsilon \\ &\quad - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)k(x_i, y_i) \\ &s.t.\begin{cases} \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0 \\ 0 \le \alpha_i, \alpha_i^* \le C, i = 1, 2, \cdots, n \end{cases} \end{aligned} \tag{13}$$

where α_i and α_i^* are Lagrangian multipliers and $k(x, x_i)$ is the kernel function. The final function is shown in equation (14):

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)k(x, x_i) + b \tag{14}$$

SVR contains a large number of kernel functions, and the selection of the appropriate kernel function will greatly affect the accuracy of the model prediction. Since the Chl-a concentration prediction model of Ulansuhai Lake belongs to a nonlinear relationship with multiple inputs and single output, and the radial basis function (RBF) has strong applicability in nonlinear problems, the RBF is chosen as the kernel function of the SVR for modeling, and the specific mathematical model is shown in equation (15):

$$k(x, x_i) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \tag{15}$$

where σ is the kernel parameter.

2) GENETIC ALGORITHM TO OPTIMIZE SUPPORT VECTOR REGRESSION

Although SVR can better solve the small sample and nonlinear problems and has been used in many fields, the values of its parameters directly affect the accuracy of SVR model prediction [31]. However, there is no clear method to determine the values of parameters, and most studies have been manually selecting parameter values relying on experience, making it difficult to obtain optimal values of parameters [32].

The genetic algorithm (GA), first proposed by John holland [33], is a stochastic search method that simulates natural selection in nature. GA is a gradient-free optimization and search technique that can randomly generate starting points from different directions and perform fast adaptive search in the solution space by the guidance of fitness function. Therefore, using the global search ability and learning speed of GA to find the optimal solution of the three parameters of penalty factor C, insensitivity coefficient ε and RBF kernel function σ parameters in SVR model can not only improve the shortcomings of SVR model, but also improve the learning speed of SVR model. The GA-SVR process is shown in Fig.4.

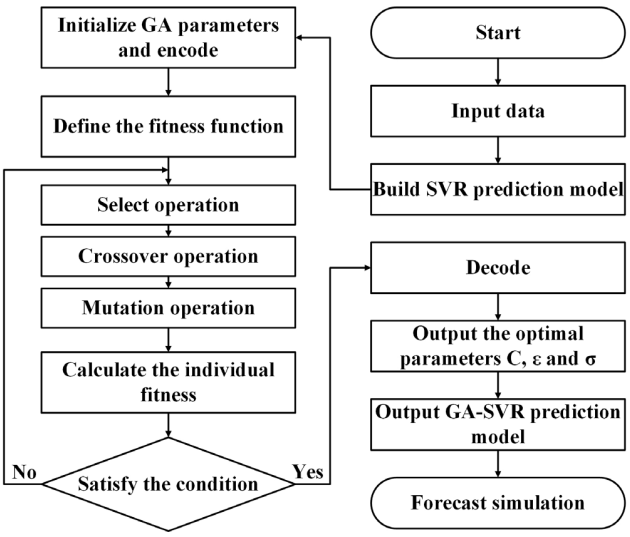

FIGURE 4. Flow chart of the GA-SVR model.

The specific steps are as follows:

1) Determination of network topology. The feature band selected by the A-ACEO algorithm is taken as the input layer, and the Chl-a concentration is taken as the output layer.

2) Initialization of parameters. Set the population size, the maximum number of iterations, the probabilities of selection, crossover and variation, the upper and lower bounds of the parameters C, ε and σ, and encode them as a set of chromosomes.
3) Definition of fitness function. The mean square error in the GA-SVR model is set as the fitness function F. The smaller the F, the higher the fitness. As shown in equation (16):

$$\mathrm{F} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 \tag{16}$$

where y_i and $\hat{y}_i$ are the measured and predicted values of Chl-a concentration.
4) Select operation. Roulette selection is used to determine the probability and select the best individuals to breed and enter the next generation.
5) Crossover operation. Two individuals in the population are selected for the cross operation to produce the more adaptive individuals. The mathematical model is shown in equation (17):

$$\begin{cases} X_{\mathrm{a}}' = pX_{\mathrm{a}} + (1-p)X_b \\ X_b' = (1-p)X_{\mathrm{a}} + pX_b \end{cases} \tag{17}$$

where $\mathrm{X_a'}$ and $\mathrm{X_b'}$ are child individuals, X_a and X_b are parent individuals, and p is a random factor between (0,1).
6) Mutation operation. An individual is selected in the population, and a portion of its genes in which are selected to exchange and vary with their alleles, so as to produce individuals with stronger fitness. The m gene variation of the $\mathrm{X_i}$ individual is shown in equations (18)-(19):

$$X_{im} = \begin{cases} X_{im} + (X_{im} - X_{\max}) \cdot f(m) r_1 > 0.5 \\ X_{im} + (X_{\min} - X_{im}) \cdot f(m) r_1 \le 0.5 \end{cases} \tag{18}$$

$$f(m) = r_2 \cdot \frac{1-m}{T} \tag{19}$$

where $\mathrm{X_{max}}$ and $\mathrm{X_{min}}$ are the upper and lower bounds of individual genes, m and T are the current and maximum iteration numbers, $\mathrm{r_1}$ and $\mathrm{r_2}$ are random numbers between [0, 1].
7) Calculation of individual fitness. Determine whether the stop condition is satisfied, if so, Step (8) is executed; otherwise, Step (4) is executed.
8) Decoding. The parameter values output by GA are used as the values of parameters C, ε and σ in SVR, and the final prediction results are output.

In conclusion, the prediction process of Chl-a concentration based on A-ACEO algorithm feature selection combined with GA-SVR is shown in Fig.5.

D. MODEL EVALUATION METRICS

In this paper, the coefficient of determination ($\mathrm{R^2}$) and the root mean square error (RMSE) are used as the evaluation indexes of the prediction model of Chl-a concentration. The details are shown in equations (20)-(21):

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}_i\right)^2} \tag{20}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2} \tag{21}$$

where y_i and $\hat{y}_i$ are the measured and predicted value of Chl-a concentration, $\bar{y}_i$ is the average value of Chl-a concentration. The closer the $\mathrm{R^2}$ is to 1, the more stable the model is; the smaller the RMSE, the more accurate the model is.

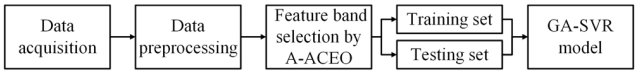

FIGURE 5. Overall process of Chl-a concentration prediction.

IV. RESULTS

A. MODEL PARAMETERS SETTING

The parameter values affect stability and efficiency of the algorithm. Whether A-ACEO or standard ACO algorithm is used to select feature bands, the parameters need to be set reasonably. TABLE 3 shows the relevant parameter settings for A-ACO-E and ACO algorithms.

TABLE 3. The relevant parameter settings for A-ACO-E algorithm and ACO algorithm.

Parameter	ACO	A-ACO-E
Maximum iterations number	50	50
Initial population size	20	20
Pheromone volatility factor	0.1	/
Pheromone inspiration factor	5	5

When the GA-SVR model was constructed, the population was selected by selection, crossover and variation behaviors in population, and the population was gradually screened to select the best adapted individuals, i.e., the optimal parameter value of SVR. Among them, the parameter values in the GA-SVR model are set reasonably to obtain the optimal individual more efficiently. In TABLE 4 the settings of the parameters related to the GA-SVR algorithm are displayed.

B. BAND REFLECTANCE PRE-PROCESSING RESULTS

The prediction results of the GA-SVR model before and after SD pre-processing are shown in TABLE 5. We can see that the $\mathrm{R^2}$ and RMSE of the GA-SVR model are improved after pre-processing the remote sensing data using SD algorithm. Therefore, SD algorithm can effectively eliminate the interference of some external factors on remote sensing images and improve the accuracy of the model.

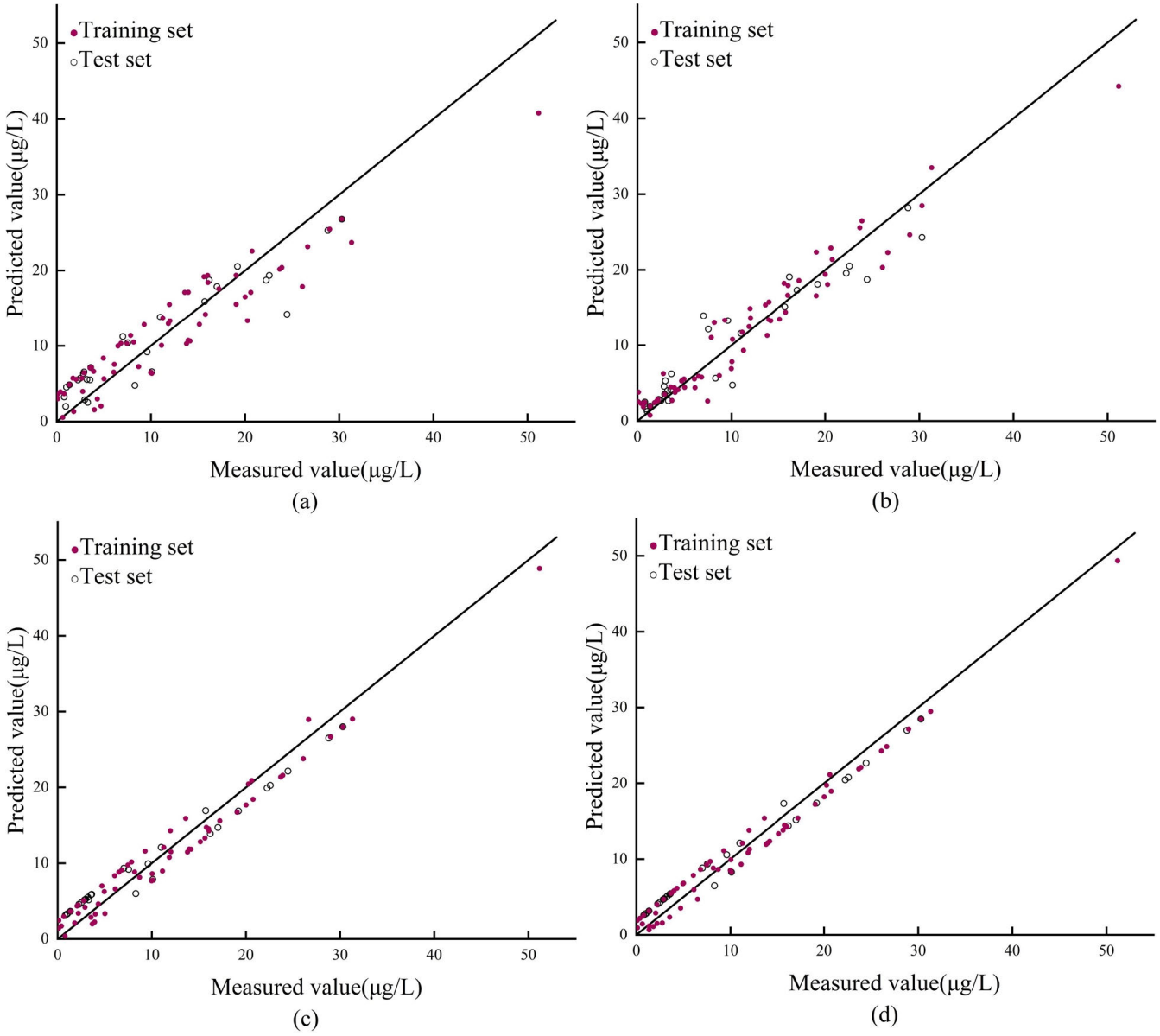

FIGURE 6. Prediction results of the GA-SVR model for (a) full-band, (b) CARS, (c) ACO and (d) A-ACEO feature band selection methods.

TABLE 4. GA-SVR parameter settings.

Parameter	Value (Probability)
Maximum iterations number	50
Initial population size	20
Selection	0.2
Crossover	0.2
Mutation	0.2
Upper boundary	0.0001
Lower boundary	100

TABLE 5. Prediction results of GA-SVR and SD-GA-SVR models.

Model	Training set		Test set	
	R^2-C	RMSE-C(μg/L)	R^2-P	RMSE-P(μg/L)
GA-SVR	0.8346	0.0712	0.8172	0.0774
SD-GA-SVR	0.8732	0.0646	0.8585	0.0663

C. GA-SVR MODEL BUILDING

The GA-SVR model is constructed to validate the effectiveness of the proposed A-ACEO algorithm and compared using full-band, CARS and ACO feature band selection methods, respectively. Fig. 6 presents the prediction results, where the diagonal line in the figure is the 1:1 line between predicted and measured values, and the prediction evaluation metrics are shown in TABLE 6.

The GA-SVR models constructed by four different feature band selection algorithms as input all have good prediction effects. From Fig. 6, most of the values are well-distributed along the 1:1 line, and only a few values are more discrete.

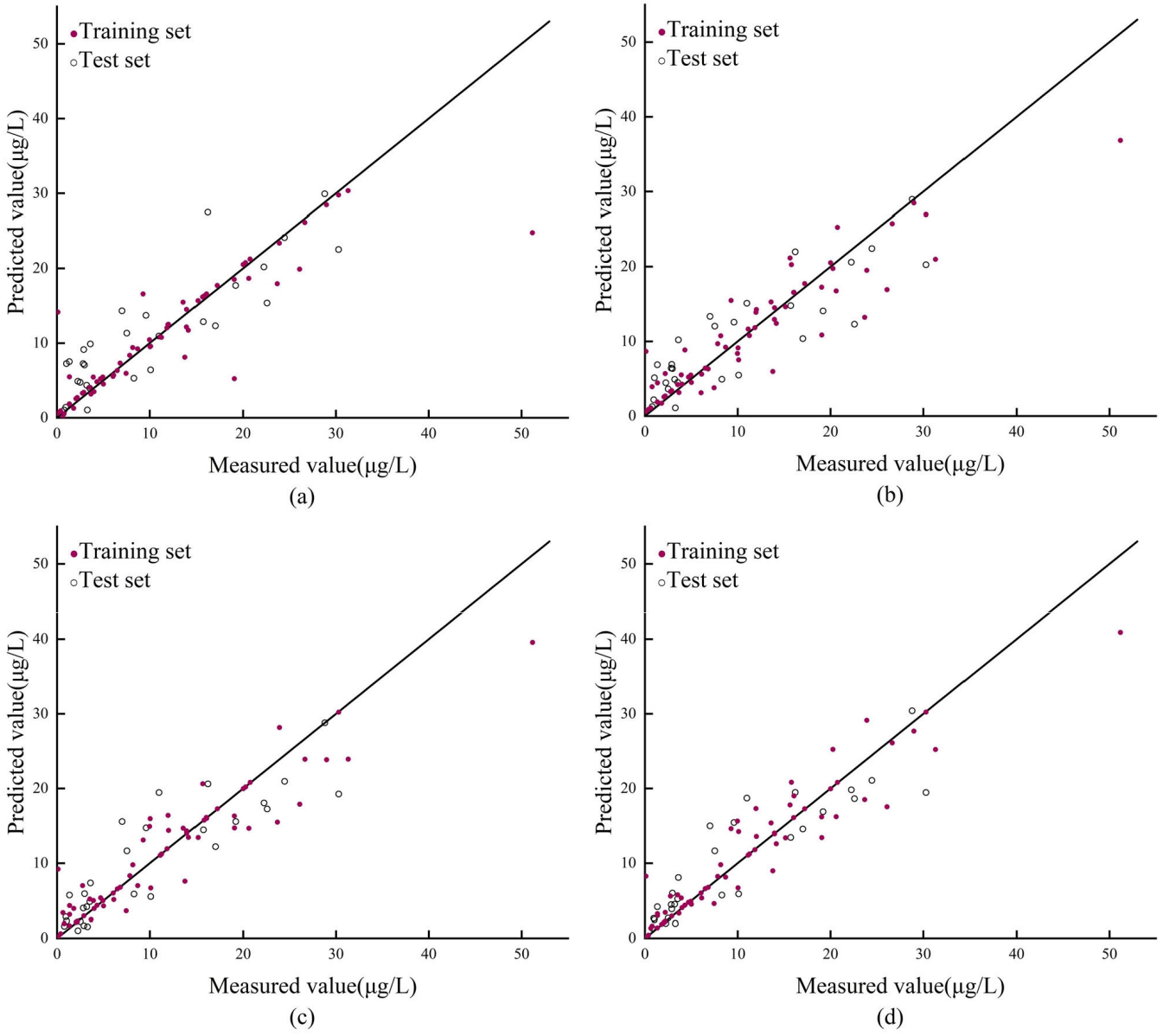

FIGURE 7. Prediction results of the SVR model for (a) full-band, (b) CARS, (c) ACO and (d) A-ACEO feature band selection methods.

From TABLE 6, the results using A-ACEO algorithm combined with GA-SVR model is the best, with R^2-C and R^2-P of 0.9696 and 0.9617, respectively, and RMSE-C and RMSE-P of 0.0330μg/L and 0.0345μg/L, respectively, which are better than the GA-SVR model built from feature bands selected by the full-band, CARS and ACO algorithms. Therefore, using A-ACEO algorithm for feature band selection simplifies the model and improves the model prediction performance.

D. SVR MODEL BUILDING

The traditional SVR model is built and compared with the GA-SVR model. Fig. 7 presents the prediction results of SVR model with different feature bands selection algorithms, and the prediction evaluation metrics are shown in TABLE 6.

From Fig. 7, the prediction accuracy of SVR model constructed with different feature band selection algorithms is not satisfactory. The prediction effect of the SVR based on the full-band is the worst compared with CARS, ACO and A-ACEO algorithms. The prediction results of the SVR model are generally distributed along 1:1 line, and there are some deviations.

From TABLE 6, compared with the GA-SVR models, all SVR models have a reduced prediction effect. The prediction accuracy of SVR model constructed using full-band is the lowest, while the best accuracy is achieved using the A-ACEO algorithm, with R^2-C and R^2-P are

TABLE 6. Prediction evaluation metrics of GA-SVR models with different feature band selection algorithms.

Model	Number of feature bands	Training set		Test set	
		R^2-C	RMSE-C(μg/L)	R^2-P	RMSE-P(μg/L)
GA-SVR	12	0.8732	0.0646	0.8585	0.0663
CARS-GA-SVR	9	0.9076	0.0557	0.8944	0.0573
ACO-GA-SVR	5	0.9524	0.0412	0.9422	0.0424
A-ACEO-GA-SVR	7	0.9696	0.0330	0.9617	0.0345

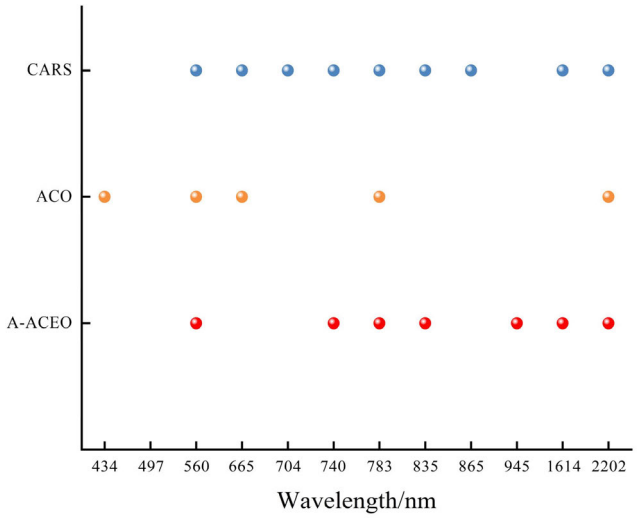

FIGURE 8. Visualization of three algorithms to select the feature band.

0.7982 and 0.7817, respectively, and RMSE-C and RMSE-P are 0.0801μg/L and 0.0823μg/L, respectively, and the number of selected feature bands was relatively small. This further demonstrates that A-ACEO algorithm has greater application potential in the selection of feature bands.

E. FEATURE BAND SELECTION

Fig. 8 shows the comparison of the optimal feature bands selected by CARS, ACO and A-ACEO algorithms. The results are visualized in Fig. 8, A-ACEO algorithm selected seven feature bands related to the Chl-a concentration, and CARS and ACO algorithms selected nine and five feature bands related to the Chl-a concentration, respectively. A-ACEO algorithm not only selects effective feature bands but also reduces the number of input variables to simplify the model.

V. DISCUSSION

Remote sensing technology has become an important tool for Chl-a concentration monitoring in lake wetlands, in which it is important to choose the appropriate feature band selection method and modeling method. This paper proposes an A-ACEO feature band selection algorithm and a GA-SVR modeling method based on the collected 92 samples.

A. CARS, ACO AND A-ACO-E FEATURE BAND SELECTION ALGORITHMS

By comparing the full band with CARS, ACO and A-ACEO algorithms, it can be found that all algorithms reduce the input of the feature band and improve the model prediction. This shows that the three algorithms are effective methods for the selection of Chl-a concentration feature bands in Ulansuhai Lake, among which the proposed A-ACEO algorithm has the best prediction effect.

The CARS algorithm is a relatively simple feature band selection method, but there is a part of redundant information in the selected feature bands, so it is difficult for the CARS algorithm to select the most relevant feature band with Chl-a concentration, which affects the accuracy of model prediction. As an intelligent algorithm, ACO algorithm can effectively avoid the interference of redundant information, but it is difficult to find the optimal solution when the problem faced is quite complex, therefore, ACO algorithm will lead to information loss, resulting in the reduction of model prediction accuracy. The A-ACEO algorithm improves on ACO algorithm by introducing the adaptive update strategy, optimal threshold and optimal matrix strategy, which improves the convergence speed and global search ability of the algorithm, and improves the diversity of feature bands while improving the information loss problem.

B. COMPARISON OF GA-SVR MODEL AND SVR MODEL

According to the model prediction results, the GA-SVR model established by different feature band selection methods has better prediction effect than the traditional SVR model, which indicates that the optimization of the parameters in the SVR model can effectively improve the model prediction accuracy by using the powerful optimization-seeking ability of the GA. This conclusion is consistent with the findings of the existing literature [34], [35]. This may be due to the fact that the prediction of Ulansuhai Lake Chl-a concentration is a nonlinear problem. Although the SVR model has strong advantages in small sample and nonlinear problems, the values of the parameters in SVR strongly affect the prediction effect of the SVR model itself. And the optimization of the SVR parameters by using the GA can avoid the shortcomings of manual search for the SVR parameters, but also find the optimal solution rapidly. Therefore, the GA-SVR

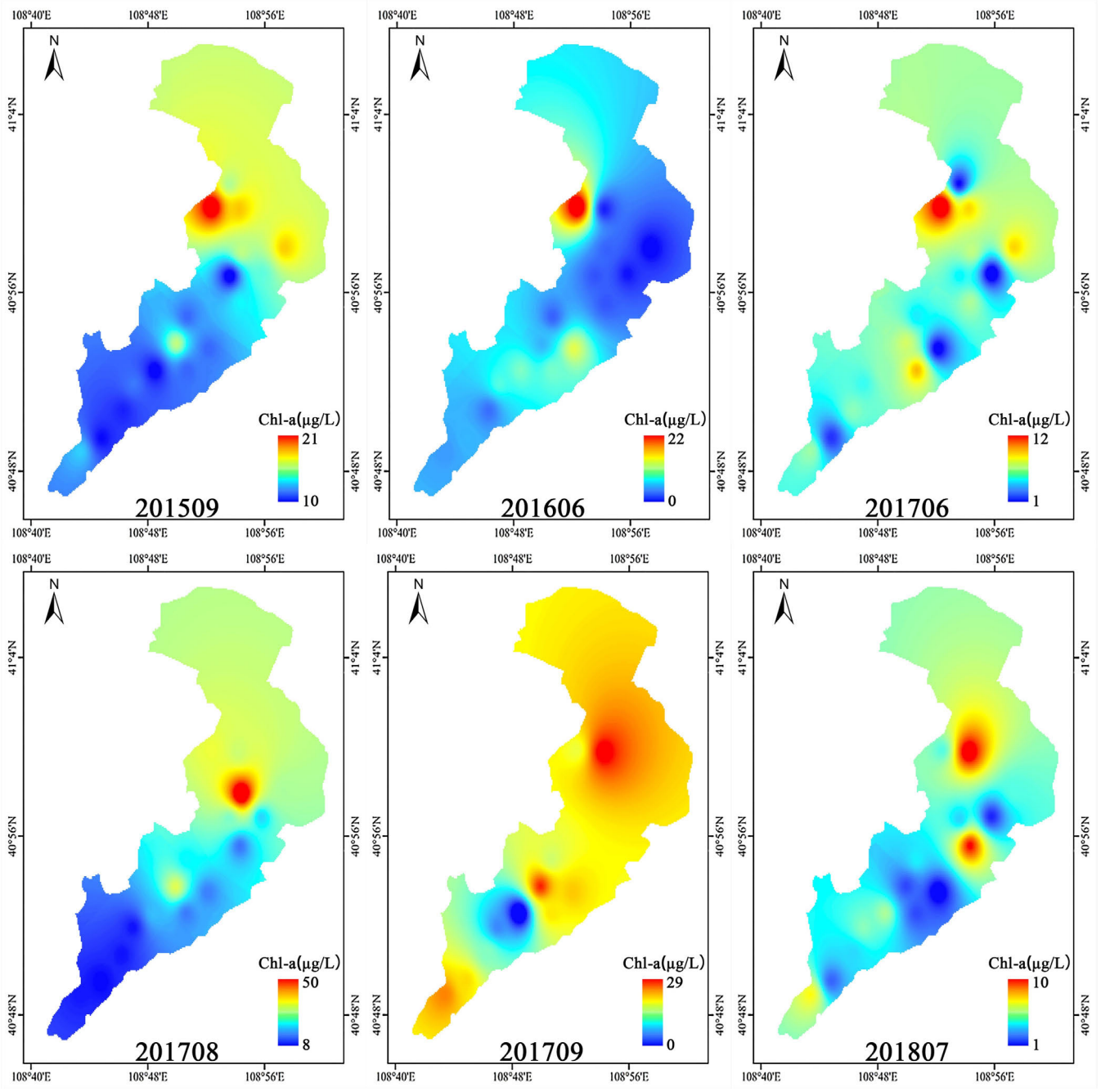

FIGURE 9. Characteristics of the spatial distribution of Chl-a in different times.

model can better predict the Chl-a concentration in Ulansuhai Lake.

C. SPATIAL AND TEMPORAL DISTRIBUTION OF CHL-A CONCENTRATION

Through analyzing different feature bands selection algorithms and different modeling methods, the GA-SVR model established with A-ACEO algorithm has the best prediction effect. Therefore, the Chl-a concentration distribution obtained based on this model is shown in Fig. 9. There are some differences in the concentration and spatial distribution of Chl-a in Ulansuhai Lake over time, which may be due to the inherent differences in the number of organisms with Chl-a concentrations in Ulansuhai Lake at different times. In addition, Ulansuhai Lake is an important component of the irrigation and drainage system of the Hetao Irrigation District in Inner Mongolia, carrying local domestic sewage, industrial wastewater and farmland drainage, which can influence the number of organisms with Chl-a concentration.

TABLE 7. Prediction evaluation metrics of SVR models with four feature band selection algorithms.

Model	Number of feature bands	Training set		Test set	
		R^2-C	RMSE-C(μg/L)	R^2-P	RMSE-P(μg/L)
SVR	12	0.7532	0.0882	0.7379	0.0902
CARS-SVR	11	0.7631	0.0867	0.7426	0.0894
ACO-SVR	7	0.7805	0.0829	0.7734	0.0839
A-ACEO-SVR	8	0.7982	0.0801	0.7817	0.0823

VI. CONCLUSION

Taking Ulansuhai Lake as the study area, combining with Sentinel-2 remote sensing data, proposing an A-ACEO feature band selection algorithm, and constructing a GA-SVR model to predict the Chl-a concentration, this paper draws the following conclusions:

When the SD algorithm is used to preprocess remote sensing data, the errors present in the remote sensing data can be reduced and the accuracy of the model prediction can be improved.

When the prediction model is the same, A-ACEO algorithm can select the feature bands related to Chl-a concentration more effectively, which can not only decrease the model complexity, but also improve the model prediction effect.

When the feature band selection method is the same, the GA-SVR model is more accurate and can better predict the Chl-a concentration. Among them, the prediction of the GA-SVR model based on A-ACEO algorithm is optimal.

Therefore, the proposed method can provide new ideas for monitoring Chl-a concentration in lakes. In future work, it is planned to accumulate additional data to more accurately monitor Chl-a concentration in lakes.

REFERENCES

[1] P. Wu, W. Zhan, N. Cheng, H. Yang, and Y. Wu, "A framework to calculate annual landscape ecological risk index based on land use/land cover changes: A case study on Shengjin Lake wetland," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 11926–11935, 2021.

[2] F. Wang, S. Zhang, H. Hou, Y. Yang, and Y. Gong, "Assessing the changes of ecosystem services in the Nansi Lake wetland, China," *Water*, vol. 11, no. 4, p. 788, Apr. 2019.

[3] M. C. Pendleton, S. Sedgwick, K. M. Kettenring, and T. B. Atwood, "Ecosystem functioning of Great Salt Lake wetlands," *Wetlands*, vol. 40, pp. 2163–2177, Jul. 2020.

[4] M. V. Nguyen, C.-H. Lin, M. A. Syariz, T. T. H. Le, and A. C. Blanco, "Multi-task convolution neural network for season-insensitive chlorophyll-a estimation in inland water," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 10439–10449, 2021.

[5] Z. Li, Q. Wang, M. Tang, P. Lu, G. Li, M. Leppäranta, J. Huotari, L. Arvola, and L. Shi, "Diurnal cycle model of lake ice surface Albedo: A case study of Wuliangsuhai Lake," *Remote Sens.*, vol. 13, no. 16, p. 3334, Aug. 2021.

[6] T. Yang, P. Hei, J. Song, J. Zhang, Z. Zhu, Y. Zhang, J. Yang, C. Liu, J. Jin, and J. Quan, "Nitrogen variations during the ice-on season in the eutrophic lakes," *Environ. Pollut.*, vol. 247, pp. 1089–1099, Apr. 2019.

[7] J. He, Y. Chen, J. Wu, D. A. Stow, and G. Christakos, "Space-time chlorophyll-a retrieval in optically complex waters that accounts for remote sensing and modeling uncertainties and improves remote estimation accuracy," *Water Res.*, vol. 171, Mar. 2020, Art. no. 115403.

[8] C. Wang, W. Jiang, Y. Deng, Z. Ling, and Y. Deng, "Long time series water extent analysis for SDG 6.6.1 based on the GEE platform: A case study of Dongting Lake," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 490–503, 2022.

[9] G. Hassan, M. E. Goher, M. E. Shaheen, and S. A. Taie, "Hybrid predictive model for water quality monitoring based on Sentinel-2A L1C data," *IEEE Access*, vol. 9, pp. 65730–65749, 2021.

[10] J. Zhang, P. Fu, F. Meng, X. Yang, J. Xu, and Y. Cui, "Estimation algorithm for chlorophyll-a concentrations in water from hyperspectral images based on feature derivation and ensemble learning," *Ecol. Informat.*, vol. 71, Nov. 2022, Art. no. 101783.

[11] T. S. Feng, Z. G. Pang, and W. Jiang, "Remote sensing retrieval of chlorophyll-a concentration in Lake Chaohu based on Zhuhai-1 hyperspectral satellite," *Spectrosc. Spectral Anal.*, vol. 42, no. 8, pp. 2642–2648, 2022.

[12] T. Zhang, M. Huang, and Z. Wang, "Estimation of chlorophyll-a concentration of lakes based on SVM algorithm and Landsat 8 OLI images," *Environ. Sci. Pollut. Res.*, vol. 27, pp. 14977–14990, Feb. 2020.

[13] D. Zhang, Q. Guo, L. Cao, G. Zhou, G. Zhang, and J. Zhan, "A multiband model with successive projections algorithm for bathymetry estimation based on remotely sensed hyperspectral data in Qinghai Lake," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 6871–6881, 2021.

[14] J. Liu, Z. Dong, J. Xia, H. Wang, T. Meng, R. Zhang, J. Han, N. Wang, and J. Xie, "Estimation of soil organic matter content based on CARS algorithm coupled with random forest," *Spectrochim. Acta A, Mol. Biomol. Spectrosc.*, vol. 258, Sep. 2021, Art. no. 119823.

[15] X. Deng, Z. Zhu, J. Yang, Z. Zheng, Z. Huang, X. Yin, S. Wei, and Y. Lan, "Detection of citrus huanglongbing based on multi-input neural network model of UAV hyperspectral remote sensing," *Remote Sens.*, vol. 12, no. 17, p. 2678, Aug. 2020.

[16] P. Yang, J. Hu, B. Hu, D. Luo, and J. Peng, "Estimating soil organic matter content in desert areas using in situ hyperspectral data and feature variable selection algorithms in Southern Xinjiang, China," *Remote Sens.*, vol. 14, no. 20, p. 5221, Oct. 2022.

[17] M. Xu, H. Liu, R. A. Beck, J. Lekki, B. Yang, Y. Liu, S. Shu, S. Wang, R. Tokars, R. Anderson, M. Reif, and E. Emery, "Implementation strategy and spatiotemporal extensibility of multipredictor ensemble model for water quality parameter retrieval with multispectral remote sensing data," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 4200616.

[18] S. Miao, Y. Li, Z. Wu, H. Lyu, Y. Li, S. Bi, J. Xu, S. Lei, M. Mu, and Q. Wang, "A semianalytical algorithm for mapping proportion of cyanobacterial biomass in eutrophic inland lakes based on OLCI data," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 5148–5161, Jul. 2020.

[19] M. Leppäranta and L. Wen, "Ice phenology in Eurasian Lakes over spatial location and altitude," *Water*, vol. 14, no. 7, p. 1037, Mar. 2022.

[20] S. Xu, S. Li, Z. Tao, K. Song, Z. Wen, Y. Li, and F. Chen, "Remote sensing of chlorophyll-a in Xinkai Lake using machine learning and GF-6 WFV images," *Remote Sens.*, vol. 14, no. 20, p. 5136, Oct. 2022.

[21] X. Zhao, H. Xu, Z. Ding, D. Wang, Z. Deng, Y. Wang, T. Wu, W. Li, Z. Lu, and G. Wang, "Comparing deep learning with several typical methods in prediction of assessing chlorophyll-a by remote sensing: A case study in Taihu Lake, China," *Water Supply.*, vol. 21, no. 7, pp. 3710–3724, 2021.

[22] H. Guo, J. J. Huang, X. Zhu, B. Wang, S. Tian, W. Xu, and Y. Mai, "A generalized machine learning approach for dissolved oxygen estimation at multiple spatiotemporal scales using remote sensing," *Environ. Pollut.*, vol. 288, Nov. 2021, Art. no. 117734.

[23] X. Xu, M. Ren, J. Cao, Q. Wu, P. Liu, and J. Lv, "Spectroscopic diagnosis of zinc contaminated soils based on competitive adaptive reweighted sampling algorithm and an improved support vector machine," *Spectrosc. Lett.*, vol. 53, no. 2, pp. 86–99, 2020.
[24] M. Dorigo and L. M. Gambardella, "Ant colony system: A cooperative learning approach to the traveling salesman problem," *IEEE Trans. Evol. Comput.*, vol. 1, no. 1, pp. 53–66, Apr. 1997.
[25] M. Mavrovouniotis, S. Yang, M. Van, C. Li, and M. Polycarpou, "Ant colony optimization algorithms for dynamic optimization: A case study of the dynamic travelling salesperson problem [research frontier]," *IEEE Comput. Intell. Mag.*, vol. 15, no. 1, pp. 52–63, Feb. 2020.
[26] Y. Tan, J. Ouyang, Z. Zhang, Y. Lao, and P. Wen, "Path planning for spot welding robots based on improved ant colony algorithm," *Robotica*, vol. 41, no. 3, pp. 926–938, Mar. 2023.
[27] K. Y. Wang, G. Tarr, J. Y. Yang, and S. Mueller, "Fast and approximate exhaustive variable selection for generalised linear models with APES," *Austral. New Zealand J. Statist.*, vol. 61, no. 4, pp. 445–465, Dec. 2019.
[28] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA: Springer Verlag, 2000.
[29] C. Xu, M. Amar, M. Ghriga, H. Ouaer, X. Zhang, and M. Hasanipanah, "Evolving support vector regression using Grey Wolf optimization; forecasting the geomechanical properties of rock," *Eng. Comput.*, vol. 38, pp. 1819–1833, Apr. 2022.
[30] Y. Meng, X. K. Zhang, and X. F. Zhang, "Identification modeling of ship nonlinear motion based on nonlinear innovation," *Ocean Eng.*, vol. 268, Jan. 2023, Art. no. 113471.
[31] X. Liu, W. Wang, Z. Guo, C. Wang, and C. Tu, "Research on adaptive SVR indoor location based on GA optimization," *Wireless Pers. Commun.*, vol. 109, pp. 1095–1120, May 2019.
[32] A. Zhan, F. Du, Z. Chen, G. Yin, M. Wang, and Y. Zhang, "A traffic flow forecasting method based on the GA-SVR," *J. High Speed Netw.*, vol. 28, no. 2, pp. 97–106, 2022.
[33] J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis With Applications to Biology, Control, and Artificial Intelligence*. Ann Arbor, MI, USA: University of Michigan Press, 1975.
[34] Q. Quan, Z. Hao, H. Xifeng, and L. Jingchun, "Research on water temperature prediction based on improved support vector regression," *Neural Comput. Appl.*, vol. 34, no. 11, pp. 8501–8510, Jun. 2022.
[35] Z. Luo, M. Hasanipanah, H. B. Amnieh, K. Brindhadevi, and M. M. Tahir, "GA-SVR: A novel hybrid data-driven model to simulate vertical load capacity of driven piles," *Eng. Comput.*, vol. 37, pp. 823–831, Apr. 2021.

CHENHAO WU was born in Hefei, Anhui, China, in 1998. He received the B.S. degree in computer science and technology from the College of Information Engineering, Suzhou University, Suzhou, China, in 2021. He is currently pursuing the M.S. degree in computer science and technology with Inner Mongolia Agricultural University, Hohhot, China.

His research interests include intelligent computing and data mining.

XUELIANG FU received the Ph.D. degree in computer science and technology from the Dalian University of Technology, Dalian, in 2008.

He is currently a Professor and the Dean of the College of Computer and Information Engineering, Inner Mongolia Agricultural University, where he is also the Director of the National Intellectual Property Information Service Center. His current research interests include intelligent computing and data mining.

HONGHUI LI received the B.S. degree in computer software from the University of Electronic Science and Technology of China, China, in 1992, and the Ph.D. degree in computer science from Concordia University, Canada, in 2013.

She is currently a Professor with the College of Computer and Information Engineering, Inner Mongolia Agricultural University, Inner Mongolia. Her research interests include optical network planning and optimization.

HUA HU was born in Bayannur, Inner Mongolia Autonomous Region, China, in 1981. He received the B.S. degree in computer science and technology from the College of Computer and Information Engineering, Inner Mongolia Agricultural University, Inner Mongolia, in 2005, where he is currently pursuing the Ph.D. degree in agricultural information technology.

His research interests include intelligent computing and data mining.

XUE LI was born in Yantai, Shandong, China, in 2000. She received the B.S. degree in computer science and technology from the College of Science and Information Science, Qingdao Agricultural University, Qingdao, China, in 2022. She is currently pursuing the M.S. degree in computer science and technology with Inner Mongolia Agricultural University, Hohhot, China.

Her research interests include intelligent computing and data mining.

LIQIAN ZHANG received the B.S. degree in communication engineering and the M.S. degree in signal and information processing from Inner Mongolia University, Inner Mongolia, China, in 2001 and 2004, respectively.

She is currently an Associate Professor with the College of Computer and Information Engineering, Inner Mongolia Agricultural University. Her current research interest includes computer networks.