

Research Article

Application of Extreme Learning Machine for Predicting Chlorophyll-a Concentration Inartificial Upwelling Processes

Yan Wei ¹, Haocai Huang ^{1,2}, Bin Chen,³ Bofu Zheng,¹ and Yihong Wang¹

¹Ocean College, Zhejiang University, Zhoushan 316021, China

²Laboratory for Marine Geology, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266061, China

³Zhoushan Agricultural and Forestry Institute of Zhejiang, Zhoushan 316021, China

Correspondence should be addressed to H. Huang; hchuang@zju.edu.cn

Received 1 January 2019; Revised 13 March 2019; Accepted 6 May 2019; Published 29 May 2019

Academic Editor: Zhaoxia Guo

Copyright © 2019 Y. Wei et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Artificial upwelling, artificially pumping up nutrient-rich ocean waters from deep to surface, is increasingly applied to stimulating phytoplankton activity. As a proxy for the amount of phytoplankton present in the ocean, the concentration of chlorophyll a (*chl-a*) may be influenced by water physical factors altered in artificial upwelling processes. However, the accuracy and convenience of measuring *chl-a* are limited by present technologies and equipment. Our research intends to study the correlations between *chl-a* concentration and five water physical factors, i.e., salinity, temperature, depth, dissolved oxygen (DO), and pH, possibly affected by artificial upwelling. In this paper, seven models are presented to predict *chl-a* concentration, respectively. Two of them are based on traditional regression algorithms, i.e., multiple linear regression (MLR) and multivariate quadratic regression (MQR), while five are based on intelligent algorithms, i.e., backpropagation-neural network (BP-NN), extreme learning machine (ELM), genetic algorithm-ELM (GA-ELM), particle swarm optimization-ELM (PSO-ELM), and ant colony optimization-ELM (ACO-ELM). These models provide a quick prediction to study the concentration of *chl-a*. With the experimental data collected from Xinanjiang Experiment Station in China, the results show that *chl-a* concentration has a strong correlation with salinity, temperature, DO, and pH in the process of artificial upwelling and PSO-ELM has the best overall prediction ability.

1. Introduction

In recent years, the ecological environment of the ocean has been deteriorated due to the excessive development and utilization of human beings [1, 2]. To ensure the sustainable development of the ocean and improve the marine environment, an artificial welling approach has been paid more attention worldwide [3–6]. It promotes the oceanic primary productivity and thus enriches phytoplankton in the euphotic layer by artificially bringing to surface the deep ocean water rich in nutrient salts such as Nitrogen and Phosphorus [7]. This process could affect the nutrient cycle, alter physical water characteristics, such as temperature, pH, and salinity, and indirectly influence sea creatures.

Chl-a concentration is a well-known indicator of ecological health of aquatic environment, and its distribution reflects richness and diversities of phytoplankton stocks [8]. It is significant to study the distribution mechanism *chl-a* and

its relationship with other water physical factors. Establishing mathematical models is an effective way to predict *chl-a* concentration with a purpose of improving marine productivity and preventing seawater eutrophication [9].

Extensive studies have been conducted in the current literature on predicting the effects of nutrients and the variations of other ecological parameters on *chl-a* concentration. They could be categorized into two mainstreams. One is measured spectrophotometrically over a long period of several years for preventing water eutrophication. In this time, the data are collected, preserved, and transferred to the laboratory regularly. The other is based on remote sensing of optical inversion and prediction. It allows researchers to map contemporaneous *chl-a* concentration on a large spatial scale. Malve developed Bayesian stratified linear model. It combined the advantages of both nonhierarchical lake-type-specific linear model and lake-type-specific linear models to predict the relation between *chl-a* concentration and total

nitrogen content and total phosphorus in Finnish Lake [10]. With the data of a tropical lake in Malaysia from 2001 to 2004, Malek established *chl-a* concentration prediction models with four computation methods, i.e., fuzzy logic, artificial neural networks, hybrid evolutionary algorithms, and multiple linear regression. The paper quantitatively evaluated and compared the performance of different models with an indicator of root mean square error (RMSE) [11]. Wang proposed a method for *chl-a* simulation in Baiyangdian Lake. It coupled the wavelet analysis and the artificial networks and took into account 14 variables in the developed and validated models, including hydrological, meteorological, and physicochemical data [12]. Based on the long-term monitoring of Yao Lake, Chen's research discovered that four environmental factors had the closest correlation with *chl-a*, i.e., transparency, nitrogen/phosphorus ratio, permanganate index, and total phosphorus. In this work, stepwise regression analysis method was adopted [13]. Mishra proposed a normalized difference chl index to estimate *chl-a* concentration from remote sensing data in estuarine and coastal turbid productive waters [14]. Wang suggested neural network technology to simulate the mathematical relationship between *chl-a* concentration and remote sensing reflectance in the Yellow Sea and East China Sea [15]. Hashim used satellite remote sensing images to study the *chl-a* distribution at Northern Region Coast of Peninsular Malaysia [16].

The *chl-a* concentration prediction in artificial upwelling processes has seldom been tackled. We have carried out the preliminary research to model the correlations between *chl-a* concentration and the associated water physical factors possibly affected by artificial upwelling with two novel neural network methods, i.e., genetic-algorithm-based neural network (GA-NN) and particle-swarm-optimization-based neural network (PSO-NN) [17]. Although the results show that these two methods can predict the correlations, they share a common disadvantage of requiring a relatively long training time. Particularly, with the increase of training data, the training time of neural network would rise accordingly. In addition, their reliability and accuracy need to be further studied. On the other hand, a novel learning algorithm for single-hidden-layer feedforward neural networks (SLFNs) named ELM was proposed in 2004 [18]. ELM has extremely faster learning speed, better generalization performance, and the least human intervention [19]. However, ELM needs a high number of hidden neurons and may lead to the ill-condition problem due to the random determination of the input weights and hidden biases. A hybrid learning algorithm later was proposed to overcome the drawbacks of ELM. It used an improved PSO algorithm to select the input weights and hidden biases and Moore-Penrose (MP) generalized inverse to analytically determine the output weights [20]. To clarify magnetic resonance image as healthy or pathological, a method combining modified particle swarm optimization and extreme learning machine (MPSO-ELM) was proposed. MPSO is applied to optimizing the hidden node parameters of SLFN, and the output weights are determined analytically. It shows the MPSO-ELM algorithm could achieve higher accuracy than BP-NN, SVMs, and conventional ELM [21]. With the development of ELM, ELM models have been

used in different fields. For example, in pathological brain detection, researchers discover that the modified differential evolution- (MDE-) ELM is able to optimize the input weights and hidden biases of ELM [22–24]. The improved ELM methods exhibit potential improvements in terms of classification accuracy and number of features. Therefore they are applied in our research.

The demand for frequent monitoring of artificial upwelling environmental primary productivity calls for an urgent need to find an efficient and accurate solution of predicting *chl-a* concentration. Meanwhile, the improvement of machine learning, in particular, ELM, inspired us to apply the machine learning to establish *chl-a* concentration prediction model. Here we try to predict the complex correlations between *chl-a* concentration and the water physical factors that are possibly affected in artificial upwelling processes, including salinity, temperature, depth, DO, and pH. Both traditional regression algorithm-based and intelligent algorithm-based prediction models were established. These models could reveal the complex relationship between water physical factors and assist researchers in analyzing the artificial upwelling's primary productivity based on *chl-a* concentration. With the experimental data collected in the Xinanjiang Experiment Station, the results of different models were analyzed and compared. The rest of the paper is organized as follows. Section 2 describes the artificial upwelling process. Section 3 explains each algorithm in detail. The experimental results and discussions for regression models and intelligent models are given to demonstrate the effectiveness of ELM-based methods in Section 4. The concluding remarks are drawn and future works are presented in Section 5.

2. Artificial Upwelling Process Description

The air-lift artificial upwelling device developed by Zhejiang University, China, utilizes an air-lift pump powered by compressed air for upwelling deep ocean water. The experiment was carried out in the Xinanjiang Experiment Station in China. Figure 1 is the schematic diagram of the experimental set-up. The main machine configuration and tools used for the experiment include research ship, monitor, air compressor, flow control valve, pressure control valve, air supply line, crane, air injection port, and upwelling pipe.

The upwelling pipe is 28.3m in length and 0.4m in internal diameter. It is composed of a suction pipe (from point B to point C, 20.8m) and a gas injection section pipe (from point A to point B, 7.5m) and is vertically deployed and completely submerged at the water depth of 2.1m. After the air is generated from the air compressor and passes through the air supply line, the pressure could be reduced by the pressure control valve to the working pressure with the range of 1.2–3.2 bar. Therefore, the pressure differences would cause the deep water first to be sucked in and then to climb up through the suction pipe. Owing to the air injection at point B, the gas injection section pipe is occupied by two-phase water-air flow. A more detailed description of this field test can be found in the report by Fan [25].

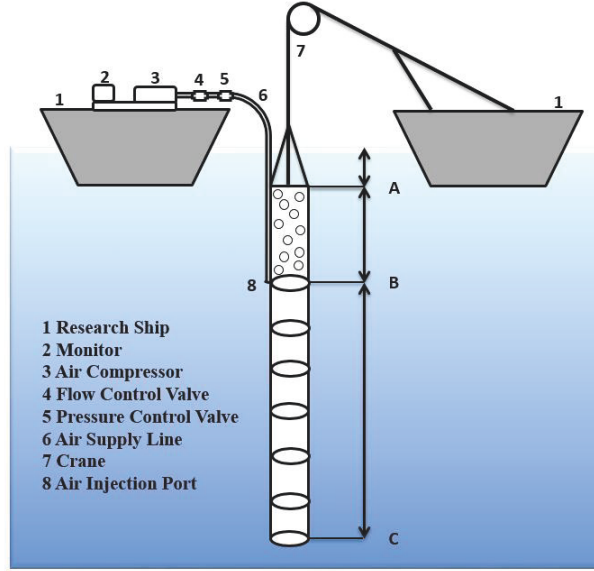


FIGURE 1: Schematic diagram of the air-lift artificial upwelling experiment set-up.

In this experiment, the deep, nutrient-rich water was sucked at point C, and the water physical factors could be regarded as inputs; the output is *chl-a* concentration at the surface water. It is predicted that while the artificial upwelling approach alters the inputs, the output varies accordingly. Five physical factors were chosen as the system input, i.e., depth, water temperature, salinity, pH, and DO. The data collection process was presented in detail by Zhou as well [17] [26].

3. Model Methods

In this section, two kinds of modeling are elaborated, i.e., linear regression model and machine learning model. The former is a traditional modeling method, simple to use and comparatively easy to analyze the relation between the model data types. Occasionally, it might ignore the nonlinear relation between the selected data types leading to inaccurate prediction result. The latter is frequently applied in big data modeling. BP-NN is one of machine learning modeling methods and is capable of self-correction. Given enough training data, it could yield accurate prediction result. ELM is a relatively new machine learning theory and exhibits fast modeling speed and is regarded applicable in quick prediction.

3.1. Linear Regression. The multiple linear regression (MLR) is a classical approach for modeling the relationship between a dependent variable and one or more independent variables [27, 28]. Every value of the independent variable is associated with the value of the dependent variable. Equation (1) is MLR function.

$$y = b_0 + \sum_{i=1}^n b_i x_i \quad (1)$$

where y is the dependent variable, x_i is the independent variable, b_i is the regression coefficient of the explanatory variable i , and b_0 is the value of the intercept in the linear fitting.

In the artificial upwelling process, MLR scheme models the relationship between every single input and output. Equation (1) can be expanded as Equation (2).

$$Chl\ a = b_0 + b_1 Sal + b_2 T + b_3 Dep + b_4 DO + b_5 pH \quad (2)$$

The training data estimate the unknown model parameters in Equation (2). Like all forms of regression analysis, MLR focuses on the conditional probability distribution of y for a given x_k value, rather than the union of x_k and y probability distribution. It depends on the model of its unknown parameters that is easier to fit than nonlinear analysis models.

Different from MLR, MQR takes into account the interaction feature between input variables. With quadratic terms, it depicts more precisely the correlation between output and input and thus is a better fit for nonlinear systems. Generated from Equation (1) by adding more terms, Equation (3) is the MQR function.

$$y = b_0 + \sum_{i=1}^k b_i x_i + \sum_{i=1}^k c_i x_i^2 + \sum_{i \neq j, i=1, j=2}^n d_{ij} x_i x_j \quad (3)$$

x_i represents Sal, T, Dep, DO, and pH here.

3.2. Backpropagation-Neural Network (BP-NN). To capture nonlinear relationships among water quality variables in a specific water system remains a technical challenge due to the complex physical, chemical, and biological processes involved [29]. Moreover, to accurately simulate the upwelling process requires a significant amount of field data to support the analysis. Superior to linear regression methods, BP-NN

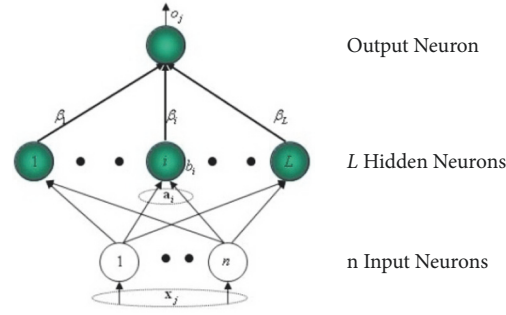


FIGURE 2: Topology of ELM.

is capable of simulating complex nonlinear systems because of its strong learning ability and self-adaption. Therefore, BP-NN is proposed to model the correlation between chl-a concentration and the five physical factors. BP-NN architecture consists of two or more layers of neurons connected by weights. The information is captured by the network when input data pass through the hidden layer of neurons to the output layer. In our research, we set the number of hidden layers and hidden nodes of every layer by the repeated tests. In order to ensure that the prediction results are stable and reliable, overmuch hidden nodes were avoided. This setting principle also helps to generate quick prediction results.

3.3. Extreme Learning Machine (ELM). ELM, a novel learning algorithm for SLFNs, was recently proposed as a unifying framework for different families of learning algorithms [30]. The optimization problem arising in learning the parameters of an ELM model can be solved analytically, resulting in a closed form involving only matrix multiplication and inversion. Hence, the learning process can be carried out efficiently without requiring an iterative algorithm, such as BP, or the solution to a quadratic programming problem as in the standard formulation of SVM. In ELM, the input weights (linking the input layer to the hidden layer) and hidden biases are randomly chosen, and the output weights (linking the hidden layer to the output layer) are analytically determined by using MP generalized inverse [31]. ELM not only has a simple structure but learns faster with better generalization performance. Therefore, it is regarded as suitable to apply ELM to modeling chl-a concentration as well.

For training SLFNs with L hidden neurons and activation function $g(x)$ shown in Figure 2, suppose there are N ($L \leq N$) arbitrary distinct samples (x_i, t_i) , where $x_i = [x_{i1}, x_{i2}, \dots, x_{in}]^T \in R^n$ and $t_i = [t_{i1}, t_{i2}, \dots, t_{im}]^T \in R^m$. In ELM, the input weights and hidden biases are randomly generated instead of being tuned. Thus the nonlinear system has been converted to a linear system. The output is expressed in Equation (4):

$$\sum_{i=1}^L \beta_i g(w_i \cdot x_j + b_i) = o_j, \quad j = 1, \dots, N \quad (4)$$

where $w_i = [w_{i1}, w_{i2}, \dots, w_{in}]^T$ is the weight vector connecting the i th hidden neuron and the input neurons, $\beta_i =$

$[\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ is the weight vector connecting the i th hidden neuron and the output neurons, $w_i \cdot x_j$ is the inner product of w_i and x_j , b_i denotes the bias of i th hidden neuron, and $o_j = [o_{j1}, o_{j2}, \dots, o_{jn}]^T$ is the matrix of the desired output.

The training goal of the neural network is that when the number of hidden neurons is equal to the number of distinct training samples, i.e., $L = Q$, standard SLFNs with L hidden neurons and activation function $g(x)$ can approximate N samples with zero error means:

$$\sum_{j=1}^N \|o_j - t_j\| = 0 \quad (5)$$

Then there should exist β_i , w_i , and b_i that satisfy this function:

$$\sum_{i=1}^L \beta_i g(w_i \cdot x_j + b_i) = t_j, \quad j = 1, \dots, N \quad (6)$$

Equation (6) can be rewritten in the matrix form:

$$H\beta = T \quad (7)$$

H is the hidden-layer output matrix, i.e.,

$$H(w_1, w_2, \dots, w_i, b_1, b_2, \dots, b_i, x_1, x_2, \dots, x_N) = \begin{bmatrix} g(w_1 x_1 + b_1) & g(w_2 x_1 + b_2) & \dots & g(w_i x_1 + b_i) \\ g(w_1 x_2 + b_1) & g(w_2 x_2 + b_2) & \dots & g(w_i x_2 + b_i) \\ \vdots & \vdots & \ddots & \vdots \\ g(w_1 x_N + b_1) & g(w_2 x_N + b_2) & \dots & g(w_i x_N + b_i) \end{bmatrix}_{N \times L} \quad (8)$$

When, in most cases, the number of hidden neurons is much less than the number of distinct training samples, i.e., $L \ll Q$, the training error of SLFNs could approach an arbitrary value $\varepsilon > 0$, i.e.,

$$\sum_{j=1}^N \|o_j - t_j\| < \varepsilon \quad (9)$$

Therefore, when activation function $g(x)$ is infinitely differentiable, there is no need for all the parameters of SLFNs to be

adjusted. The input w_i and the hidden layer biases b_i could be arbitrarily given and remain unchanged during the training process. To train an SLFN is simply equivalent to finding a least-squares solution $\hat{\beta}$ of $\min_{\beta} \|H\beta - T'\|$. The smallest norm least-squares solution is

$$\hat{\beta} = H^{\dagger} T' \quad (10)$$

where H^{\dagger} is the MP generalized inverse of the matrix H ; T' is the transposed matrix of T .

Based on the arithmetic formula discussed before, the main steps of the ELM learning algorithm are elaborated as follows:

Step 1. Determine the number of neurons in the hidden layer, and randomly set the weight vector w connecting the hidden neuron and the input neurons and the bias of the hidden layer neurons b . (The time of training model and prediction accuracy are two main factors in determining the hidden nodes.)

Step 2. Select an infinitely differentiable function as the activation function of the neurons of the hidden layer, and then calculate the hidden-layer output matrix H .

Step 3. Calculate the weight vector β connecting the hidden neuron and the output neurons.

3.4. GA-ELM & PSO-ELM & ACO-ELM. Although ELM is fast and presents good generalization performance, since the output weights β are computed based on randomly assigned input weights w and hidden biases b using MP generalized inverse, there may exist a set of nonoptimal input weights and hidden biases, and it might suffer from the overfitting as the learning model approximates all training samples well [32]. The model is capable of training samples but there is still space to improve in predicting samples.

Most work using an optimization algorithm to train the neural network is to fix the topology of the network in advance and then to use different algorithms to optimize the weighting matrix of the previous neural network. To overcome the disadvantages of ELM and to optimize the initial weight of the network, we propose three optimization algorithms to optimize ELM. The first is GA, a well-known random search and global optimization method based on the idea of natural selection and evolution [33]. The second is PSO technique, a population-based stochastic algorithm for optimization based on social-psychological principles [34]. The third is ACO, a metaheuristic inspired by the pheromone trail laying and following the behavior of some ant species, for solving hard combinatorial optimization problems [35]. Although the underlying theories of the three algorithms are different, the detailed running procedures of PSO and ACO algorithms are similar to GA, specifically in terms of initialization, fitness evaluation, the update of particle state, and goal test.

GA-ELM was taken as an example to explain the optimization procedure. Optimization of ELM network training methods can be divided into two main parts. The first is to

determine the coding scheme of network connection weights; the second is to apply the genetic algorithm to completing the evolution. To a fixed network structure, the process of evolving network connection weights typically follows the below steps:

Step 1. Initialize the original population. Randomly generate a population of coded individuals. The coding pattern is predesigned.

Step 2. Decode the population and calculate the fitness of each. For every individual, its fitness could be obtained from the designed fitness function:

$$F = \sum_{m=1}^M \sum_{k=1}^q (d_k^m - y_k^m)^2 \quad (11)$$

where d_k^m is the desired output of the node k in the set m ; y_k^m is the predicted output of the node k in the set m ; F is the fitness.

Step 3. Select the individual. The survival-of-the-fittest mechanism is imposed on the candidate individuals. The individuals with higher fitness values have a greater probability of giving offspring, while the low-fitness ones have the potential to be eliminated and their chance of survival is low. Many selection schemes have been proposed to accomplish this idea, including roulette-wheel selection, stochastic universal selection, ranking selection, and tournament selection. We choose the roulette-wheel selection here. It is based on proportionate fitness selection. The selection probability of each can be computed by

$$f_i = \frac{k}{F_i}, \quad (12)$$

$$p_i = \frac{f_i}{\sum_{i=1}^N f_i}$$

where F_i is the fitness for individual i . k is the coefficient; p_i is the probability of being selected of individual i ; N is the population size. The smaller fitness indicates the better individual.

Step 4. Get the offspring. After selection, the individuals from the mating pool are combined to generate hopefully better offspring. The crossover and mutation are involved in this stage. For specific problems, many crossover methods have been designed. Here, the real-parameter crossover operator was applied because the individual adopts real number coding. Let the individual k and individual i do a crossover at the position j .

$$a_{kj} = a_{kj}(1-b) + a_{ij}b$$

$$a_{ij} = a_{ij}(1-b) + a_{kj}b \quad (13)$$

where a refers to the individual; b is a random number between 0 and 1.

TABLE 1: Error comparison between different models.

	R-square	RMSE	ME
MLR	0.4528	0.2116	-0.0676
MLR-updated	0.4494	0.2101	-0.0619
MQR	0.6966	0.1951	0.1275
MQR-updated	0.7084	0.1892	0.1226
BP-NN	0.7060	0.1270	0.0186
ELM	0.5212	0.1251	0.0193
GA-ELM	0.5437	0.1241	0.0262
PSO-ELM	0.6557	0.1156	0.0357
ACO-ELM	0.6913	0.1195	0.0119

As for a mutation operator, select the gene j of individual i , namely, a_{ij} to mutate.

$$a_{ij} = \begin{cases} a_{ij} + (a_{ij} - a_{max}) * f(g) & r_1 \geq 0.5 \\ a_{ij} + (a_{min} - a_{ij}) * f(g) & r_1 < 0.5 \end{cases} \quad (14)$$

$$f(g) = r_2 \left(1 - \frac{g}{G_{max}} \right)$$

where a_{max} and a_{min} represent the upper and lower bounds of a_{ij} , respectively; g is the current iteration number; G_{max} is the maximum generation; r_1 and r_2 are random numbers between 0 and 1. Then the new generation of the population could be derived.

Step 5. Go back to Step 2 until the performance requirements are met.

After GA, assign the parameters of the best offspring to the input weights of ELM. Thus, GA-ELM has been optimized with GA and is ready for traditional training.

4. Experimental Results

4.1. Comparison of Analysis between Different Models. The data used in building the *chl-a* concentration prediction model were collected from the Xinanjiang Experiment Station, China. Five sets of input (depth, water temperature, salinity, pH, and DO) have been selected from the experiment data. According to our primary research [17], these five water parameters have a strong correlation with the *chl-a* concentration in the process of artificial upwelling. Based on the analysis of existing *chl-a* concentration prediction model, the five chosen water parameters could give effective prediction result. Meanwhile, these five data sets are comparatively easier to obtain than some other water parameters, like dissolved inorganic nitrogen concentration.

For the uniform measurement of model accuracy, 2100 sets of data were collected, and the first 2000 groups were taken for training and the remaining 100 groups were used for correlation prediction. For each model, 100 correlation-predicting output values were estimated and further compared with the corresponding real measurements. To facilitate the comparison between different models, three indicators were chosen and calculated to evaluate the models'

prediction effectiveness, i.e., the regression coefficient (R), the root mean square error (RMSE), and the mean error (ME). MLR-updated and MQR-updated are two improved models. They are the results of the reregression after shaking off the input variables insignificant to the output correlation. RMSE and ME are calculated by

$$RMSE = \sqrt{\frac{\sum d_i^2}{n}}, \quad i = 1, 2, 3, \dots, n \quad (15)$$

$$ME = \frac{\sum d_i}{n}, \quad i = 1, 2, 3, \dots, n \quad (16)$$

where d_i is the prediction error.

Table 1 shows that PSO-ELM and ACO-ELM have lower RMSE indicating that their prediction results are more reliable with the lowest prediction bias. With higher R-square value, MQR-updated and BP-NN result in a relatively high overestimation. Here, ME is the mean value of 100 errors, and the error values could be minus as well. Although the ME of PSO-ELM is not the lowest, largely it does not necessarily influence the overall prediction accuracy of PSO-ELM. In conclusion, PSO-ELM and ACO-ELM have the best prediction results.

4.2. MLR and MLR-Updated Results. Table 2 shows the results estimated by MLR model and their t-test (Student's t-test). With the estimated parameters obtained, the fitted equation is

$$Y = 6.0976 + 10.6243X_1 - 0.1026X_2 - 0.0023X_3 - 0.1433X_4 - 0.1564X_5 \quad (17)$$

Here Y is the *chl-a* concentration, X_1 is Salinity, X_2 is Temperature, X_3 is Depth, X_4 is Dissolved oxygen, and X_5 is pH.

In Table 2, the P value for X_3 in the t-test is 0.0116. Therefore, X_3 is removed from the independent variable, and the linear regression was performed again [17]. The following equation is the MLR-updated result:

$$Y = 6.1819 + 10.7405X_1 - 0.1046X_2 - 0.1721X_4 - 0.1451X_5 \quad (18)$$

TABLE 2: Correlation test for MLR.

Variable	Parameters Estimate	Standard Error	t Value	p-Value
intercept	6.0976	0.1104	55.2072	<0.0001
X1	10.6243	0.2743	38.7352	<0.0001
X2	-0.1026	0.0041	-25.0641	<0.0001
X3	-0.0023	0.0009	-2.5262	0.0116
X4	-0.1433	0.0203	-7.0658	<0.0001
X5	-0.1564	0.0065	-24.1301	<0.0001

4.3. MQR and MQR-Updated Results. As described in Section 3.1, MQR adds the quadratic term and the cross-term to MLR. The parameter estimates after first regression analysis are listed in Table 3. The coefficients of some cross items are larger than the first term, and the correlation between depth X3 and *chl-a* concentration is not strong. However, the coefficient of the cross term with depth X3 is larger. This demonstrates that the low linear correlation does not necessarily indicate low influences. Similar to the optimization in Section 4.2, although the p-value of the whole function is 0.0000, it is still found that the p values of X1, X3, X1 * X3, and X3 * X5 are relatively large. This means that these four correlations with *chl-a* concentration are not obvious enough. After removing the four items, the results of the MQR-updated model are shown in Table 4.

4.4. BP-NN and ELM Results. For BP-NN and ELM, the important parameters are the number of layers and hidden nodes. After repeating model testing, the suitable parameters of BP-NN and ELM are obtained. For BP-NN, it needs more than one layer to ensure its stability and accuracy. For ELM, one layer-structured model could yield ideal prediction results. The number of hidden nodes should be set at a reasonable range so the ELM prediction model could choose the number of most suitable hidden nodes. The detailed parameter setting is shown in Table 5.

The advantage of BP-NN is its small ME, indicating that BP-NN model has better prediction performance. Once the prediction error is higher or the deviation is larger caused by abnormal input values, BP-NN model has a feedback system and is able to correct itself.

The advantage of ELM model lies in two aspects. First is its lower RMSE value, demonstrating its better prediction accuracy for smaller points. Therefore, if the input error of the system is small, ELM models would yield better optimization results. The second is the shorter training time of ELM compared with BP-NN and its data prediction consumes less time as well. This is crucial especially when the total amount of data collected in the experiment becomes large. The long training time would weaken the timeliness of the data prediction. Therefore, it is considered that ELM model is more suitable for predicting *chl-a* concentration in artificial upwelling processes.

4.5. GA-ELM Results. Table 1 shows that the RMSE of GA-ELM model is reduced compared with that of ELM, but its ME is increased. To understand the reason, we compared

the GA-ELM prediction results based on 100 sets of test data with the corresponding measurement data, as displayed in Figure 3.

In general, the GA-ELM model can predict the dynamic change of *chl-a*. It performs well in some data segments, especially in the general trend at the turning point. Therefore, the RMSE is reduced compared with the ELM model. However, the GA-ELM model has a large error in the prediction of individual points. There are multiple points of which the error is more significant in Figure 3, resulting in an increase in the mean error. These large-error points appear in the area where the corresponding measurement results are stable. The errors could be incurred by a certain input value with sudden change. This may prove that the GA-ELM model is rather sensitive to the change of input values.

4.6. PSO-ELM Results. In Table 1, the RMSE of the PSO-ELM is the smallest, indicating that the overall fitting effect of PSO-ELM is the best among the seven models. Its prediction results were compared with the measurement data shown in Figure 4.

As the same with GA-ELM, PSO-ELM can also have a dynamic predictive ability of *chl-a*, especially in the partial data segment. This greatly reduces the RMSE of the model. However, similar to GA-ELM, the ME of PSO-ELM is also large and is mainly caused by the poor performance at the turning points of data in Figure 4. This shows that even though PSO-ELM has overall good forecast ability of the *chl-a* concentration, it may be not good at predicting the data with frequent fluctuation.

4.7. ACO-ELM Results. Table 1 presents that ACO-ELM has better performance with smaller RMSE and ME, which indicates that ACO-ELM is able to predict the dynamics of *chl-a* concentration. In the smooth data section and the transition data section, the ACO-ELM yields relatively accurate prediction result, and the numbers of extreme error points are less compared with that of GA-ELM or PSO-ELM, as is shown Figure 5.

Figure 5 illustrates that, for ACO-ELM model, large deviation only occurs at 2 points. It proves ACO-ELM significantly superior to the first two optimization algorithms. However, an obvious drawback of ACO-ELM is that a large number of evolutionary algebras and a huge number of groups are required in order to achieve Figure 5 prediction results. It indicates that a longer processing time is needed for model training.

TABLE 3: Correlation test for MQR.

Variable	ParametersEstimate	Standard Error	t Value	p-Value
Intercept	10.5306	1.4556	7.2346	<0.0001
X1	6.6003	5.3318	1.2379	0.2159
X2	0.6132	0.1217	5.0393	<0.0001
X3	0.0156	0.0261	0.5968	0.5507
X4	-4.0743	0.2622	-15.5412	<0.0001
X5	0.5394	0.1157	4.6605	<0.0001
X1*X2	0.7263	0.2010	3.6127	0.0003
X1*X3	0.1273	0.0834	1.5258	0.1272
X1*X4	-6.9839	0.7005	-9.9700	<0.0001
X1*X5	5.0367	0.2791	18.0447	<0.0001
X2*X3	-0.0086	0.0010	-8.6792	<0.0001
X2*X4	0.1025	0.0131	7.8027	<0.0001
X2*X5	-0.0506	0.0054	-9.3766	<0.0001
X3*X4	0.0207	0.0042	4.8882	<0.0001
X3*X5	0.0038	0.0013	2.9932	0.0028
X4*X5	-0.1162	0.0176	-6.6092	<0.0001
X1*X1	-57.7413	4.8081	-12.0091	<0.0001
X2*X2	-0.0357	0.0024	-14.7037	<0.0001
X3*X3	-0.0007	0.0002	-4.7051	<0.0001
X4*X4	0.3293	0.0209	15.7644	<0.0001
X5*X5	0.0251	0.0047	5.2821	<0.0001

TABLE 4: Correlation test for MQR-updated.

Variable	Parameters Estimate	Standard Error	TValue	p-Value
Intercept	10.8428	1.3143	8.2501	<0.0001
X2	0.6451	0.1085	5.9462	<0.0001
X4	-4.2175	0.2558	-16.4867	<0.0001
X5	0.5760	0.0896	6.4292	<0.0001
X1*X2	0.8776	0.1483	5.9170	<0.0001
X1*X4	-5.7966	0.4445	-13.0407	<0.0001
X1*X5	4.7375	0.2526	18.7538	<0.0001
X2*X3	-0.0073	0.0008	-9.6089	<0.0001
X2*X4	0.0983	0.0118	8.3297	<0.0001
X2*X5	-0.0494	0.0047	-10.3963	<0.0001
X3*X4	0.0265	0.0025	10.6561	<0.0001
X4*X5	-0.0869	0.0145	-6.0100	<0.0001
X1*X1	-56.3251	4.5516	-12.3748	<0.0001
X2*X2	-0.0369	0.0021	-17.3334	<0.0001
X3*X3	-0.0008	0.0001	-6.8581	<0.0001
X4*X4	0.3209	0.0205	15.6840	<0.0001
X5*X5	0.0113	0.0022	5.1592	<0.0001

TABLE 5: Parameters for BP-NN and ELM.

Model	Layer 1 hidden node	Layer 2 hidden node
BP-NN	10	1
ELM	[10, 15]	0

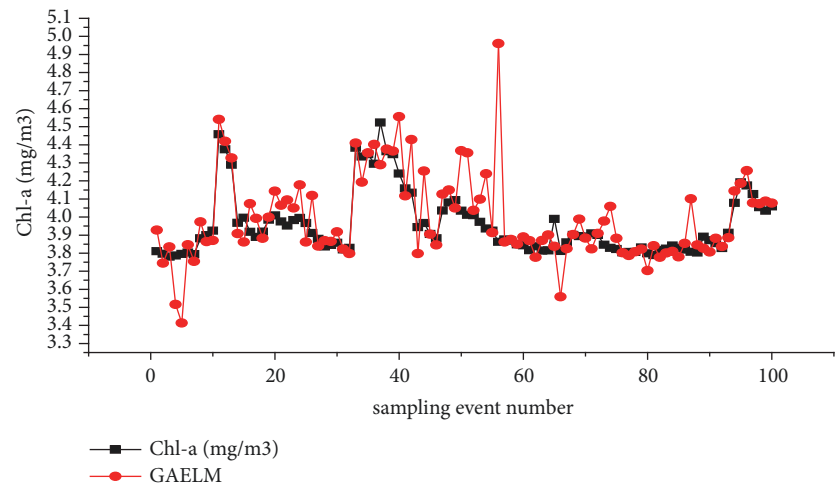


FIGURE 3: Prediction result of GA-ELM.

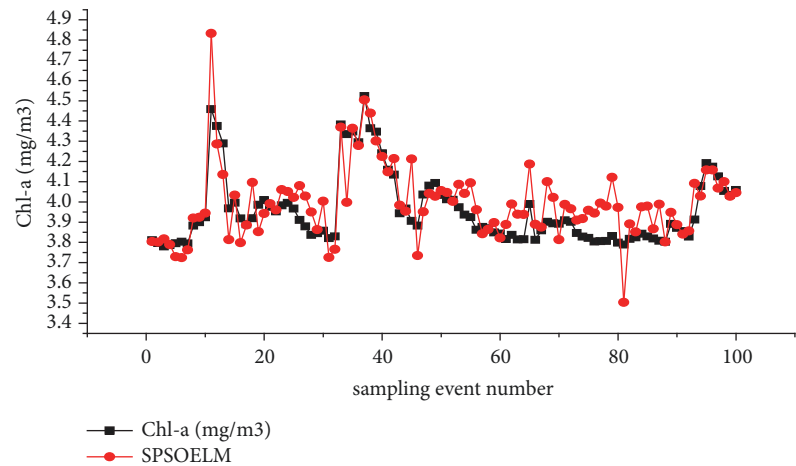


FIGURE 4: Prediction result of PSO-ELM.

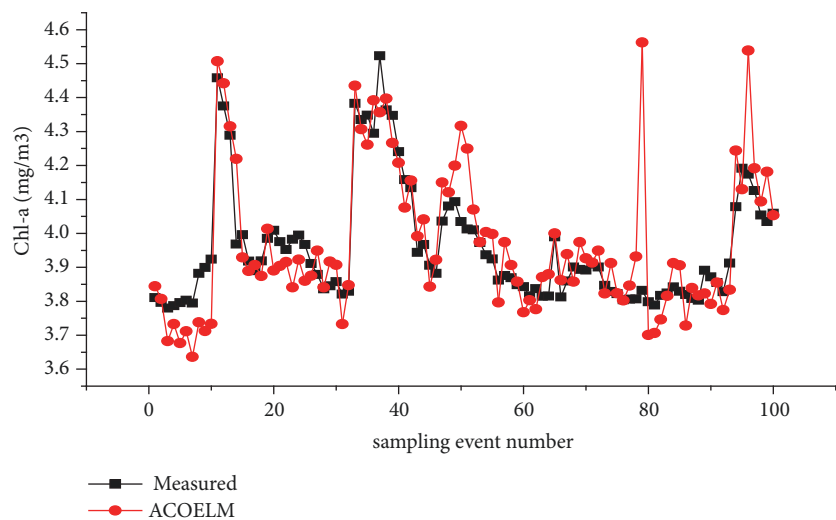


FIGURE 5: Prediction result of ACO-ELM.

5. Conclusions and Future Work

In this paper, the prediction models were established for predicting the correlation between the water physical parameters (salinity, temperature, depth, dissolved oxygen, and pH) and *chl-a* concentration in the artificial upwelling environment. Respectively, seven models, MLR, MQR, BP-NN, ELM, GA-ELM, PSO-ELM, and ACO-ELM, were presented. Furthermore, three indicators, i.e., R-square, RMSE, and ME, were applied as the evaluation standard for quantitative analysis of model prediction ability. The prediction results show that the first two traditional regression algorithms reveal not only the relationship between input and output parameters but also the effects of combined variables. The latter five models are collectively referred to as intelligent algorithms. Owing to the complex nonlinear relationship between the inputs and the output, the intelligent algorithm has higher prediction accuracy than the traditional regression algorithms.

Among these five intelligent algorithms, PSO-ELM has the best overall prediction ability after a comprehensive evaluation. However, each model has exhibited its own characteristics. Although the training speed of BP-NN is slower than ELM due to its feedback mechanism, its prediction accuracy is more acceptable, apart from a certain amount of input mutation. The apparent advantage of ELM is that its training speed is fast and the data could be forecasted efficiently. However, because of its hidden layer structure and the preset neuron node, it is indispensable to increase the nodes according to the increase of the training sample. The outstanding advantage of the GA-ELM model is that it has better prediction results at turning points, and conversely, PSO-ELM could predict the smooth segment data better. Their combination could be more preferable in practical application. For ACO-ELM, although no large deviation happens, the algorithm requires a large number of populations and more iterations to achieve better prediction results and would be slow in training. In practical applications, to obtain accurate prediction results, it is suggested to select more suitable models based on the available measurement data.

For the further work, three parts are considered to be incorporated to improve the prediction ability of proposed models. The first is that more data to be collected from different experiment stations would be used to further train and test the models. Furthermore, it is expected that turbidity will be added as the sixth input. In addition, since ELM with modified sine cosine algorithm has shown great ability in pattern classification, it might be that the application of this algorithm to PSO-ELM would help to increase the prediction accuracy of *chl-a*.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

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

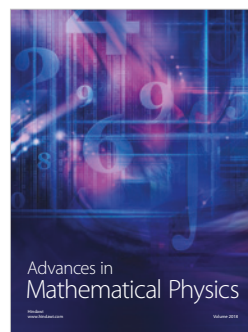
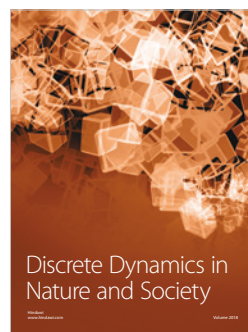
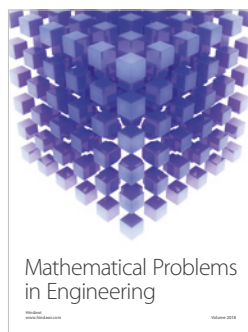
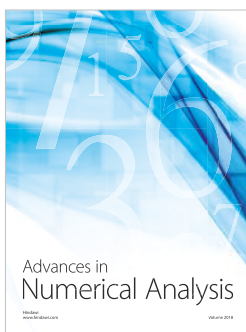
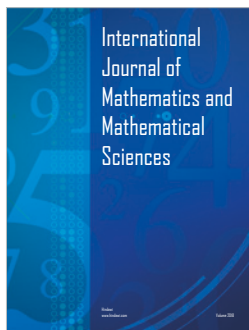
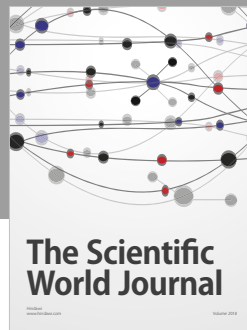
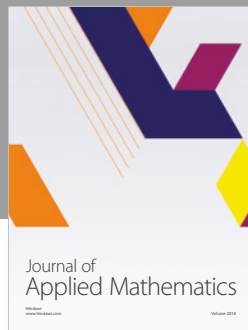
Acknowledgments

This research was funded by the NSFC Projects under Grants nos. 41576031 and 51120195001 and the Scientific Research Foundation for the Returned Overseas Chinese Scholars, State Education Ministry, with Grant no. [2014]1685. The authors would like to acknowledge Dr. Jiawang Chen, Dr. Han Ge, Miss Shan Lin, and Miss Jianying Leng who helped with the Qiandao Lake experiments.

References

- [1] J. D. Watson and F. H. C. Crick, "A structure for deoxyribose nucleic acid," *Nature*, vol. 171, no. 4356, pp. 737-738, 1953.
- [2] B. Kirke, "Enhancing fish stocks with wave-powered artificial upwelling," *Ocean & Coastal Management*, vol. 46, no. 9-10, pp. 901-915, 2003.
- [3] C. C. K. Liu and Q. Jin, "Artificial upwelling in regular and random waves," *Ocean Engineering*, vol. 22, no. 4, pp. 337-350, 1995.
- [4] K. Ouchi and H. Ohmura, "Design concept and experiment of ocean nutrient enhancer "TAKUMI"," in *Proceedings of the Ocean '04 - MTS/IEEE Techno-Ocean '04: Bridges across the Oceans - Conference Proceedings*, pp. 322-327, Japan, November 2004.
- [5] K. Ouchi, "Results from real sea experiment of ocean nutrient enhancer TAKUMI," in *Proceedings of the ASME 2009 28th International Conference on Ocean, Offshore and Arctic Engineering*, pp. 1473-1480, Honolulu, Hawaii, USA, 2009.
- [6] J. Aure, Ø. Strand, S. R. Erga, and T. Strohmeier, "Primary production enhancement by artificial upwelling in a western Norwegian fjord," *Marine Ecology Progress Series*, vol. 352, pp. 39-42, 2007.
- [7] A. Van Geen, R. K. Takesue, J. Goddard, T. Takahashi, J. A. Barth, and R. L. Smith, "Carbon and nutrient dynamics during coastal upwelling off Cape Blanco, Oregon," *Deep-Sea Research Part II: Topical Studies in Oceanography*, vol. 47, no. 5-6, pp. 975-1002, 2000.
- [8] J. Chen, M. Zhang, T. Cui, and Z. Wen, "A review of some important technical problems in respect of satellite remote sensing of chlorophyll - A concentration in coastal waters," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 5, pp. 2275-2289, 2013.
- [9] P. H. Dimberg and C. J. Olofsson, "A comparison between regression models and genetic programming for predictions of Chlorophyll-a concentrations in Northern Lakes," *Environ Model Assess*, vol. 21, pp. 221-232, 2015.
- [10] O. Malve and S. S. Qian, "Estimating nutrients and chlorophyll a relationships in Finnish lakes," *Environmental Science & Technology*, vol. 40, no. 24, pp. 7848-7853, 2006.
- [11] S. Malek, A. Salleh, and S. M. S. Ahmad, "A comparison between neural network based and fuzzy logic methods for Chlorophyll-a estimation," in *Proceedings of the ICCEA*, vol. 2, pp. 340-343.
- [12] F. Wang, X. Wang, B. Chen, Y. Zhao, and Z. Yang, "Chlorophyll a simulation in a lake ecosystem using a model with wavelet analysis and artificial neural network," *Journal of Environmental Management*, vol. 51, no. 5, pp. 1044-1054, 2013.
- [13] Y. Chen, *Study on relationship between water eutrophication and environmental factors in Yao Lake [MSc Thesis]*, East China Jiaotong University, Nanchang, China, 2012.

- [14] S. Mishra and D. R. Mishra, "Normalized difference chlorophyll index: A novel model for remote estimation of chlorophyll-a concentration in turbid productive waters," *Remote Sensing of Environment*, vol. 117, pp. 394–406, 2012.
- [15] D. Wang, Y. Li, and B. Gao, "Neural network technology and semi-analytical approach combined model for remote sensing chlorophyll-a concentration," in *Proceedings of the 36th IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2016*, pp. 5852–5855, China, July 2016.
- [16] S. Hashim, R. H. Narashid, and Z. M. Arof, "Studies of chlorophyll-a distribution in Northern Region Coast of Peninsular Malaysia using satellite remote sensing," in *Proceedings of the 2011 IEEE International Conference on Space Science and Communication: "Towards Exploring the Equatorial Phenomena"*, *IconSpace 2011*, pp. 261–264, Malaysia, July 2011.
- [17] Y. Zhou, H. Huang, J. Ding, H. Wang, Y. Chen, and J. Wang, "Modeling chlorophyll concentration affected by artificial upwelling in qiandao lake," *IEEE Journal of Oceanic Engineering*, vol. 41, no. 4, pp. 960–969, 2016.
- [18] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *Proceedings of the IEEE International Joint Conference on Neural Networks*, vol. 2, pp. 985–990, July 2004.
- [19] S. Ding, X. Xuand, and R. Nie, "Extreme learning machine," *Neural Computing and Applications*, vol. 25, pp. 549–556, 2014.
- [20] F. Han, H.-F. Yao, and Q.-H. Ling, "An improved evolutionary extreme learning machine based on particle swarm optimization," *Neurocomputing*, vol. 116, pp. 87–93, 2013.
- [21] D. R. Nayak, R. Dash, and B. Majhi, "Discrete ripplelet-II transform and modified PSO based improved evolutionary extreme learning machine for pathological brain detection," *Neurocomputing*, vol. 282, pp. 232–247, 2018.
- [22] D. R. Nayak, R. Dash, and B. Majhi, "An improved extreme learning machine for pathological brain detection," in *Proceedings of the TENCON 2017 - 2017 IEEE Region 10 Conference*, pp. 13–18, Penang, Malaysia, November 2017.
- [23] D. R. Nayak, R. Dash, and B. Majhi, "An improved pathological brain detection system based on two-dimensional PCA and evolutionary extreme learning machine," *Journal of Medical Systems*, vol. 42, no. 1, p. 1, 2018.
- [24] D. R. Nayak, R. Dash, B. Majhi, and S. Wang, "Combining extreme learning machine with modified sine cosine algorithm for detection of pathological brain," *Computers and Electrical Engineering*, vol. 68, pp. 366–380, 2018.
- [25] W. Fan, J. Chen, Y. Pan, H. Huang, C.-T. Arthur Chen, and Y. Chen, "Experimental study on the performance of an air-lift pump for artificial upwelling," *Ocean Engineering*, vol. 59, pp. 47–57, 2013.
- [26] American Water Association, *Standard Methods for The Examination of Water Wastewater*, part 4500, American Public Health Association, Washington, DC, USA, 1999.
- [27] J. Cohen, P. Cohen, S. G. West et al., *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*, Lawrence Erlbaum Associates, Hillsdale, NJ, USA, 2nd edition, 2003.
- [28] A. Velo, F. F. Pérez, T. Tanhua, M. Gilcoto, A. F. Ríos, and R. M. Key, "Total alkalinity estimation using MLR and neural network techniques," *Journal of Marine Systems*, vol. 111–112, pp. 11–18, 2013.
- [29] J.-T. Kuo, M.-H. Hsieh, W.-S. Lung, and N. She, "Using artificial neural network for reservoir eutrophication prediction," *Ecological Modelling*, vol. 200, no. 1–2, pp. 171–177, 2007.
- [30] A. R. Lima, A. J. Cannon, and W. W. Hsieh, "Nonlinear regression in environmental sciences using extreme learning machines: a comparative evaluation," *Environmental Modelling & Software*, vol. 73, pp. 175–188, 2015.
- [31] Q. Zhu, A. K. Qin, P. N. Suganthan, and G. Huang, "Evolutionary extreme learning machine," *Pattern Recognition*, vol. 38, no. 10, pp. 1759–1763, 2005.
- [32] H. Yang, J. Yi, J. Zhao, and Z. Dong, "Extreme learning machine based genetic algorithm and its application in power system economic dispatch," *Neurocomputing*, vol. 102, pp. 154–162, 2013.
- [33] A. Fraser, "Simulation of genetic systems by automatic digital computers. II: Effects of linkage on rates under selection," *Austral Journal of Biological Sciences*, vol. 10, no. 4, pp. 492–499, 1957.
- [34] J. Kennedy and R. Eberhart, *Swarm Intelligence*, Morgan Kaufmann Publishers, Inc, San Francisco, Calif, USA, 2001.
- [35] M. Dorigo and T. Stützle, *Ant Colony Optimization*, MIT Press, London, UK, 2004.



Submit your manuscripts at
www.hindawi.com