IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received August 19, 2019, accepted September 17, 2019, date of publication September 30, 2019, date of current version October 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2944422

Prediction of Significant Wave Heights Based on CS-BP Model in the South China Sea

SHAOBO YANG[1,3], TIANLIANG XIA[1], ZHENQUAN ZHANG[1], CHONGWEI ZHENG[2,4,5], XINGFEI LI[1,3], HONGYU LI[6], AND JIANJUN XU[7,8]

[1]State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China
[2]State Key Laboratory of Estuarine and Coastal Research, Shanghai 200062, China
[3]Qingdao Institute for Ocean Technology, Tianjin University, Qingdao 266200, China
[4]Shandong Provincial Key Laboratory of Ocean Engineering, Ocean University of China, Qingdao 266100, China
[5]Dalian Naval Academy, Dalian 116018, China
[6]School of Mechanical and Electronic Engineering, Shandong University of Science and Technology, Qingdao 266590, China
[7]South China Sea Institute of Marine Meteorology, Guangdong Ocean University, Zhanjiang 524000, China
[8]Southern Marine Science and Engineering Guangdong Laboratory (Zhanjiang), Zhanjiang 524000, China

Corresponding authors: Chongwei Zheng (chinaoceanzcw@sina.cn) and Jianjun Xu (jxu@gdou.edu.cn)

This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFC1506902, in part by the Fund of Southern Marine Science and Engineering Guangdong Laboratory (Zhanjiang) under Grant ZJW-2019-XX, in part by the State Key Laboratory of Estuarine and Coastal Research under GrantSKLEC-KF201707, in part by the State Key Laboratory of Precision Measuring Technology and Instruments of Tianjin University, in part by the National Natural Science Foundation of China, in part by the Innovation and Entrepreneurship Project of the Ocean Engineering Institute of Tianjin University under Grant 20190201-5, and in part by the Opening Fund of the Shandong Provincial Key Laboratory of Ocean Engineering.

ABSTRACT Forecasting the significant wave heights (Hs) is indispensable in H_S-related engineering studies and is exceedingly important in the assessment of wave energy in future. As a technique essential for the future of clean energy systems, reducing the forecasting errors related to Hs has always been a vital research subject. In this paper, an optimized hybrid method based on the back propagation neural network (BP) and the cuckoo search algorithm (CS) is proposed to forecast the Hs in the South China Sea. This approach employs the CS as an intelligent optimization algorithm to optimize the parameters of the BP model, which develop a hybrid model that is suit for the data set, reducing the forecasting errors. The proposed method is subsequently tested based on nine prediction points selected in the South China Sea, where the proposed hybrid model is proved to perform effectively and steadily.

INDEX TERMS CS-BP, significant wave heights, South China sea, predication performance.

I. INTRODUCTION

Ocean waves are complex phenomenon because their production depends on many atmospheric, meteorological and oceanographic factors [1]. The study of ocean waves is of great significance to marine engineering construction, marine development, transportation, marine fishery, aquaculture and so on. The rapid and accurate prediction of wave heights is also crucial for disaster warning and emergency prevention [2]. The complex and random characteristics of ocean waves make it more difficult to predict the height of ocean waves.

For effective prediction of wave heights, many experts and scholars put forward different methods such as empirical, numerical and soft computing approaches. The prediction of ocean waves is traditionally done by converting wind-related information to waves. The empirical models like SMB and Darbyshire were frequently used to predict ocean wave heights in the 1960s and 1970s [3]. Numerical models such as wave modeling (WAM) and nearshore wave modeling (SWAN) [4] are usually based on a form of spectral energy or action balance equation, which became popular in the 1980s and 1990s due to its mathematical rigor and the large temporal and spatial coverage with them [5]. However, this method is to build a physical model of wave heights, which requires not only basic knowledge of physical processes as a prerequisite, but also high performance computing infrastructure, as well as high cost and long time period [6]. In case of an ocean emergency, in order to take emergency measures in advance and reduce disasters, it is necessary to propose a faster and more accurate wave heights prediction method. Predicting the heights of wave from the knowledge of generating wind is

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

basically an uncertain and random process, so it is difficult to model with deterministic equations, which makes it ideally suit for neural network model [7].

Agrawal and Deo [8] studied the online prediction of wave heights by Artificial Neural Networks (ANNs), first-order autoregressive moving average (ARMA) and Autoregressive Integrated Moving Average (ARIMA) model. They reported that ANNs was more accurate for short-term predictions than either time series model. Mahjoobi and Adeli Mosabbeb [9] proposed a support vector machine (SVM) method for predicting wave heights. Compared with the traditional neural network, the pre diction error of SVM is small. Ozger [10] applied the hybrid method of wavelet and fuzzy to predict the significant wave heights and average wave period 48 hours ago, and the results were satisfactory compared with autoregressive, artificial neural network and fuzzy logic model. Nitsure *et al.* [11] used wind information and genetic programming (GP) for wave prediction and the results were satisfactory, especially for the peak of wave heights formed by the extreme events like hurricanes. In the same year, Kambekar and Deo [12] used GP and model trees to predict waves, and the results showed that GP model had a better tendency than the MT model. Truong and Ahn [13] used a modified gray model MGM (1,1) to predict wave for real-time control of wave energy converters in irregular waves. Fernández *et al.* [14] introduced classification techniques in marine energy prediction. Shahabi *et al.* [15] used a hybrid wavelet-genetic programming approach to predict the significant wave heights of two monitoring stations in the north Atlantic. The results showed that the model has higher prediction reliability. Cornejo-Bueno *et al.* [16] proposed a hybrid Grouping Genetic Algorithm - Extreme Learning Machine approach (GGA-ELM) to predict the significant wave height at the Western coast of the USA, obtaining good results. This approach can solve feature selection problems and may be applied to alternative regression approaches. Wang *et al.* [7] used a BP neural network model optimized by mind evolutionary algorithm (MEA-BP) to predict the ocean wave heights, and the results indicated that MEA-BP was superior to the genetic algorithm-BP neural network model (GA-BP) and standard BP neural network model (St-BP) with faster running time and higher prediction accuracy. Yang *et al.* [17] applied SARIMA model to predict the significant wave height of the South China Sea and adjacent waters in the long term, which has a good prediction performance.

BP neural network is the most widely used neural network, which can directly map input to output. Although BP neural networks can utilize relevant parameters in prediction of Hs, they also are easy to be trapped in local minimum. Therefore, it is very necessary to use powerful optimization algorithms to overcome this shortcoming. CS is a population-based optimization algorithm and a meta-heuristic algorithm, which can be implemented to adjust weights and biases of BP neural network by updating location of the nests in order to achieve the purpose of improving performance [18]. This paper presents a CS-BP prediction model, which combines the local optimization of BP neural network and the global optimization of cuckoo search algorithm. The model is used to predict the Hs in the South China Sea, transforming the uncertainty of Hs into predictability.

II. RESEARCH METHODOLOGY

According to the theory that wave is created by wind interacting with the ocean surface, Hs and wind data are combined with neural network to predict Hs. The Figure 1. shows the research method for predicting Hs, which can be divided into three steps.

1) The features selection. According to the theory that wave is driven by wind, the parameters, such as Hs and wind, are selected and processed.

2) The optimization of the BP neural network. The BP neural network is optimized by genetic algorithm and particle swarm optimization algorithm (PSO) and cuckoo search algorithm to GA-BP model, PSO-BP model and CS-BP model respectively.

3) The prediction of Hs using neural network model and analysis of prediction results. The Hs are predicted by the neural network model. It is compared with the Hs from WAVEWATCH-III (WW3) driven by wind data to analyze the generalization ability and prediction accuracy.

III. SELECTED PREDICTION POINTS

In this paper, we choose the South China Sea as the study area. The annual average Hs simulated by WW3 wave model in the South China Sea in 2017 are shown in the Figure 2. The blue points represent that, in the South China Sea, the nine prediction points were selected from 111.375°E-117.125°E and 9.125°N-21.125°N. Each research point has a latitude difference of 6° and a longitude difference of 3°. As can be seen from Figure 2, there are significant differences in Hs at the nine prediction points, which cover a wide range of geographical locations and different weather conditions. This could make research results more universal.

IV. AVAILABILITY DATA AND SETTING OF INPUT PARAMETERS

A. SIGNIFICANT WAVE HEIGHTS AND WIND DATA

The ocean surface is subjected to the action of wind to generate waves. As the wind gets stronger and stronger, the waves get bigger and bigger. Energy accumulates over time. The current wave is related not only to the current wind speed, but also the previous wind speed and wave. Therefore, Hs database and wind database were used as input parameters of the used model in this study.

The Hs data is from the WW3 model because of its high resolution and accuracy. The WW3 model is the third generation of full-spectrum space wave mode under the framework of the WAM, developed by a marine simulation team in NOAA/NCEP Environmental Simulation Center. In this study, WW3 was used to simulate wave field in the South China Sea from January 2012 to December 2017. Range selected for model calculation was 97.375°E-125.125°E,

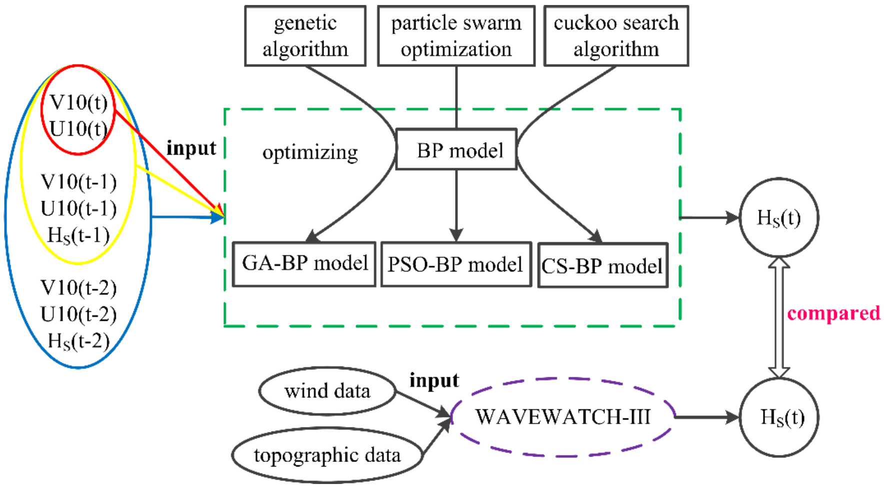


FIGURE 1. Research methodology for predicting significant wave heights.

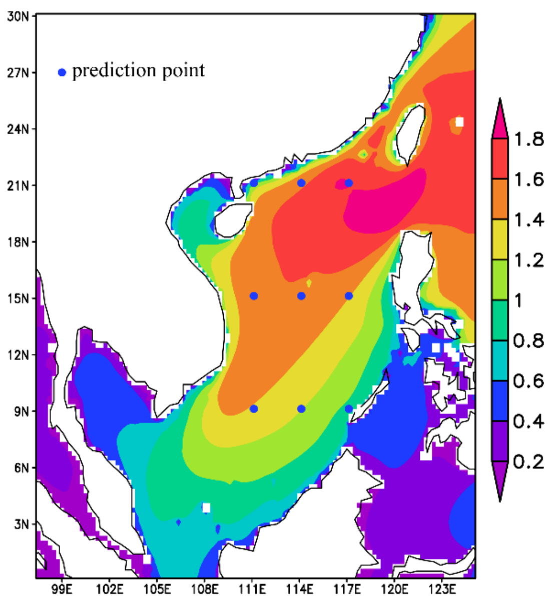


FIGURE 2. The annual average significant wave heights simulated by the WW3 model in the South China Sea in 2017.

0.125°N-30.125°N. Many studies have demonstrated that WW3 has a good ability to simulate the wave in the South China Sea and Simulated Hs accuracy is within 15% [19], [20]. So, the simulated Hs are used as input data for the used model and validation data for comparing with the predicted Hs.

The CCMP wind data comes from ESE (NASA earth science enterprise) [21], combining several kinds of data such as Advanced Earth Observing Satellite, 2nd Generation (ADEOS-II), QuikSCAT, Tropical Rainfall Measuring Mission Microwave Imager (TRMM TMI), Special Sensor Microwave Imager (SSM/I) and Advanced Microwave Scanning Radiometer Earth Observing System (AMSR-E), and is derived from variation method. CCMP wind field is chosen due to its high resolution and long-time sequence.

B. INPUT PARAMETER AND NETWORK STRUCTURE

According to the theory of wind generating wave, the wind data and Hs data were used as the input parameters of the used model. After daily-averaging the data, there are 1,826 sets of data at one prediction point, including five years' data from 2013 to 2017, so the nine prediction points contain 16434 sets of data. The default samples are divided into three categories: training samples, verification samples and test samples. The last 100 sets of data were selected as the test set and validation set and the rest as the training set.

There are three kinds of input parameters. The first kind of input parameters are V10(t) and U10(t), the second kind of input parameters are the first input parameters, V10(t-1), U10(t-1) and Hs(t-1), and the third kind of input parameters are the second input parameters, V10(t-2), U10(t-2) and Hs(t-2). The output parameter is Hs(t). While V10 and U10 are respectively the daily average wind speed in latitude-direction at 10 meters and in longitude-direction at 10 meters. Hs is daily average significant wave height. The time step is a day. To prevent prediction error, an array of all ones is added to each input data as training set.

A three-layer BP neural network was developed in this study due to its sufficient ability to approximate any continuous nonlinear functions [22]. The structure of three-layer BP neural network includes input layer, hidden layer and output layer. The number of three input layers corresponding to the above input data of the prediction model are 3, 6 and 9, respectively. The number of hidden layers was determined by Eq. (1).

$$q = 2p + 1 \tag{1}$$

where p is the number of input layer and q is the number of hidden layer. From the Eq. (1), the number of hidden layers are 7, 13 and 19. The output parameter is the Hs. Thus, three kinds of neural network structure are 3-7-1, 6-13-1 and 9-19-1 respectively. The network structure is shown in table 1.

TABLE 1. The network structure.

Parameters	Number of input layers	Number of hidden layers	Network structure
array of all ones, V10(t) and U10(t)	3	7	3-7-1
array of all ones, V10(t), U10(t), V10(t-1), U10(t-1) and Hs(t-1)	6	13	6-13-1
array of all ones, V10(t), U10(t), V10(t-1), U10(t-1), Hs(t-1), V10(t-2), U10(t-2) and Hs(t-2)	9	19	9-19-1

V. PREDICTION MODEL

Recently, the neural network has developed rapidly and achieved good results in economy, biology, ocean and other aspects. Zhang *et al.* [23] studied a hybrid algorithm combining PSO and BP together to train the weights and thresholds of neural network. After illustrating the result of different applications, they concluded that the hybrid algorithm integrating BP and PSO improved the performance, and exceeded the individual BP or PSO in terms of the convergence rate and error level. Adhikari *et al.* [24] investigated a PSO approach which improve the predicting ability of feed-forward artificial neural network optimized by particle swarm optimizations to predict uniaxial compressive strength and achieved good results. Wang *et al.* [18] used the BP model optimized by cuckoo search algorithm, which is superior to the traditional BP neural network in lightning prediction. In this paper, we try to apply the BP, GA-BP, PSO-BP and CS-BP model to the prediction of Hs in the South China Sea.

A. BP NEURAL NETWORK

In the mid-1980s, California PDP (Parallel Distributed Processing) applied Error Back Propagation Training to the research of neural network, which systematically solved the learning problem of connection weight of hidden layer in multi-layer neural network [25].

The back-propagation (BP) is the most well-known learning strategy among vast number learning algorithms [26], [27] and has arbitrary complex pattern classification ability and excellent multi-dimensional function mapping ability. Structurally, there are input layer, hidden layer and output layer. The calculation processes of BP neural network consists of forward calculation process and reverse calculation process. In the process of forward propagation, the input parameters are processed layer by layer from the input layer through the hidden layer and transferred to the output layer. Moreover, the state of neurons in each layer only affects the state of neurons in the next layer. Every neuron decides its net weighted contribution as:

$$X = \sum_{i=1}^{n} x_i \omega_i + \theta \tag{2}$$

where n is the number of inputs, and x_i and ω_i demonstrate respectively the value of ith input and weight. The thresholds applied to the neurons is denoted by θ.

If the desired output cannot be obtained in the output layer, it will be transferred to back propagation, and the error signal will be returned along the original connection path. By modifying the weights and thresholds of each neuron, the predicted output of BP neural network are gradually approaching the desired output [28].

B. OPTIMIZAED BP NEURAL NETWORK

The BP network is apt to plunge into local minimum, so the optimization algorithm is used to improve its performance.

In 1962, professor Holland from the university of Michigan proposed a genetic algorithm, which is a global optimization algorithm simulating the mechanism of biological evolution. However, it has highly sensitive to the initial population, which affects its global optimal search ability [29] and has been successfully combined with BP neural network [30]. According to the selected fitness function, individuals are screened through selection, cross and mutation in genetic algorithm. The weights and thresholds of BP neural network are optimized and updated to form a new algorithm, namely the BP neural network model optimized by genetic algorithm (GA-BP).

In this case, the roulette method based on fitness ratio was selected. The selection probability of each individual i is P_i , as follows.

$$f_i = \frac{k}{F_i} \tag{3}$$

$$p_i = \frac{f_i}{\sum_{j=1}^{n} f_j} \tag{4}$$

where F_i is the fitness value of individual i. Since the fitness value is getting better and better, the reciprocal of fitness value is performed before individual selection. The coefficient k is set as 10. n is the number of individuals in the population.

The encoding method of individual is real encoding, so the real crossover method is adopted. The crossover of the kth chromosome a_k and the lth chromosome a_l at position j is performed as follows.

$$\left.\begin{aligned} a_{kj} &= a_{kj}(1-b) + a_{lj}b \\ a_{lj} &= a_{lj}(1-b) + a_{kj}b \end{aligned}\right\} \tag{5}$$

where b is a random number between [0, 1].

The jth gene a_{ij} of the ith individual was selected to be mutated by the following methods.

$$a_{ij} = \begin{cases} a_{ij} + (a_{ij} - a_{\max}) * f(g) & r > 0.5 \\ a_{ij} + (a_{\min} - a_{ij}) * f(g) & r \le 0.5 \end{cases} \tag{6}$$

$$f(g) = r_2(1 - \frac{g}{G_{\max}})^2 \tag{7}$$

where a_{max} and a_{min} are upper and lower bounds of gene a_{ij} respectively. r_2 is a random number. g and G_{max} are respectively the number of current iteration and the maximum number of evolution. r is a random number between [0,1].

Originally proposed by Kennedy and Eberhart [31], the PSO algorithm is a global algorithm with strong global optimization capability [32]. In particle swarm optimization, all kinds of simple particles are put into the search space of n-dimensional problem or capacity [33]. Three indexes of position, velocity and fitness value are used to represent the particle characteristics. The position of every particle is determined by connecting some aspects of their own current and optimal position with those of other swarm particles. By comparing the fitness value of particles with the individual extreme and the global extreme, the individual extreme value and the global extreme value are updated. After the condition is met, the velocity and position of particles are updated, and then the weights and thresholds of BP neural network are updated [34]. Thus the BP neural network optimized by particle swarm optimization algorithm is formed.

In each iteration, the particle updates its speed and position through the individual extreme value and the global extreme value. The update formula is as follows:

$$X_i^{k+1} = X_i^k + V_i^{k+1} \tag{8}$$

$$V_i^{k+1} = \omega V_i^k + c_1 r_1 (P_{b,i}^k - X_i^k) + c_2 r_2 (G_{b,i}^k - X_{id}^k) \tag{9}$$

where X_i is the n-dimensional vector that denotes the position of particle i in the search space, and k is iteration. V_i means the speed of this particle. The best position of the individual and global extremum found by the swarm are respectively represented by $P_{b,i}$ and $G_{b,i}$, in Eq. (9). Besides, ω is inertia weight, r_1 and r_2 are random values in the scope of [0, 1], c_1 and c_2 are positive speeding up constants.

A large inertia weight is conducive to global search, while a small inertia weight is more conducive to local search. In order to balance the global search capability and local search capability of the algorithm, the weight of linear decline relation is proposed.

$$\omega(k) = \omega_{start} - (\omega_{start} - \omega_{end}) * \frac{k}{T_{\max}} \tag{10}$$

where ω_{start} is the initial inertia weigh. ω_{end} is the inertia weight at the maximum number of iterations. k and T_{max} are respectively the number of current iteration and the maximum number of iteration. In this paper, ω_{start} is 0.9 and ω_{end} is 0.4.

The cuckoo search algorithm (CS) was proposed in 2009 by Yang and Deb [35], Cambridge scholars. It is a meta-heuristic algorithm that simulates cuckoo parasitic reproduction and combines the cuckoo reproduction process with the Lévy flight search method [18], [36] of birds. In the algorithm, the host egg in the nest is regarded as a solution, and the cuckoo egg is regarded as a new solution. Thereafter, the nest with higher quality egg may be preserved to the next generation, the target is that the cuckoo's egg is incubated instead of the host's egg. In other words, the bad solution in the nest is replaced by the good cuckoo's solution.

In the CS algorithm, the initial location of the host nest is expressed as:

$$x_i = r \times (Ub - Lb) \tag{11}$$

where r is the uniform random numbers within the interval [0, 1], and the Ub and Lb respectively are the upper and lower bounds of the search space.

Yang et al. proposed that cuckoo's location is updated making use of the Lévy flight mechanism. The update formula is as follows:

$$x_i^{(t+1)} = x_i^{(t)} + \alpha \oplus L(\lambda) \tag{12}$$

where $x_i^{(t+1)}$ and $x_i^{(t)}$ denote the new solution and current location unit, respectively. α is the step size associated with the optimized problem, and the step size $\alpha = 1$ is selected. The notation $\oplus$ represents the entry-wise multiplication. $L(\lambda)$ the transition probability, and it carries out the random walk based on the Levy distribution, which is given by Eq. (13):

$$L(\lambda) = 0.01 \frac{\mu}{|v|^{\frac{1}{\beta}}} (x_i^{(t)} - x_b^{(t)}) \tag{13}$$

where $x_b^{(t)}$ is the best location, $\beta = 1.5$, μ and v are deducted from the normal distribution curves.

The next task in CS algorithm is to compare the discovery probability pa with a random number to determine whether a new solution can be generated. In case the cuckoo egg is discovered by the host bird and it would be abandoned from the nest, the event may occur based on the probability Pa. Thence, the Pa can be regarded as the probability of an individual to be retained. The new solution is calculated as

$$x_i^{(t+1)} = \begin{cases} x_i^{(t)} + \gamma \times (x_j^{(t)} - x_k^{(t)}) & r > p_a \\ x_i^{(t)} & otherwise \end{cases} \tag{14}$$

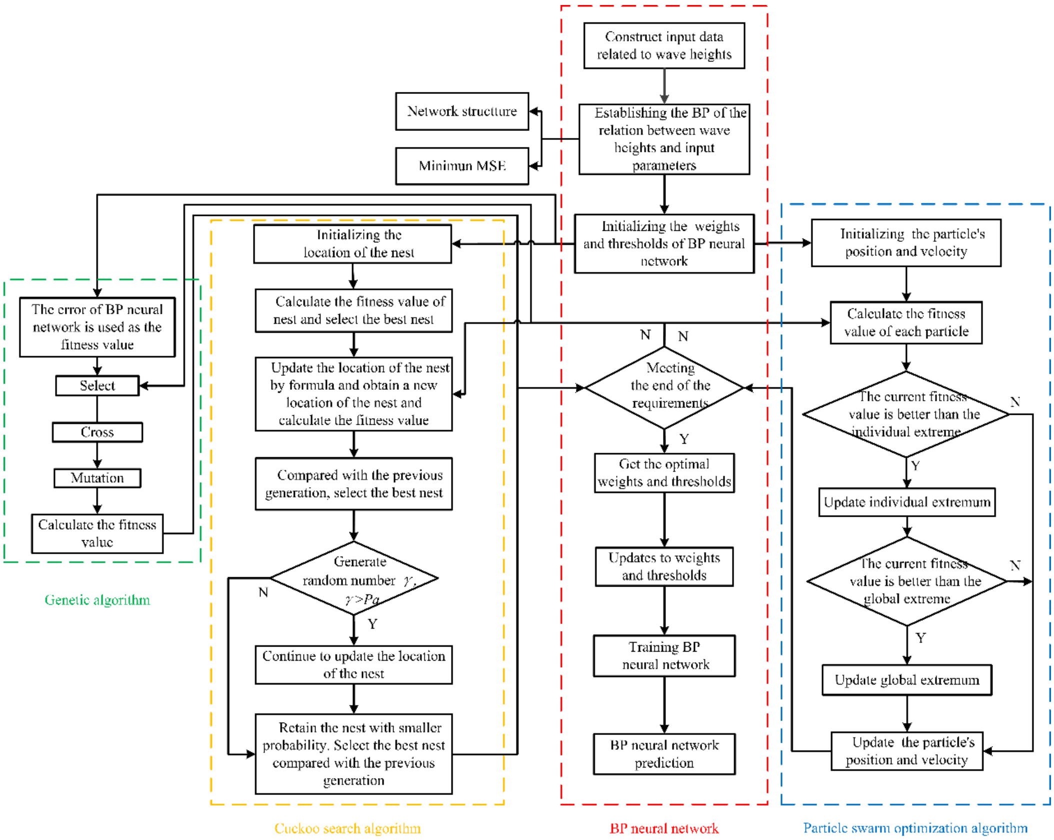


FIGURE 3. Algorithm flow of the BP, GA-BP, PSO-BP and CS-BP model.

where r and γ are the uniform random numbers within the interval [0, 1], $x_j^{(t)}$ and $x_k^{(t)}$ are two different solutions which are chosen at random.

C. ALGORITHM FLOW OF THE BP, GA-BP, PSO-BP AND CS-BP MODEL

In this paper, the mean square error (MSE) of simulated and predicted Hs have been selected to be minimized as an objective function defined in Eq. (15):

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(X_i - P_i)^2 \tag{15}$$

where X_i and P_i refer to simulated and predicted Hs, respectively, and n is the number of simulations.

As shown in Figure 3, the algorithm flow has four parts: BP neural network, genetic algorithm optimization, particle swarm optimization and cuckoo search algorithm.

1) Constructing input data related to Hs.

2) Establishing the BP neural network of the relation between Hs and input parameters: The minimum mean square error (MSE) and establishing the network structure, such as the number of hidden layers.

3) Initializing the weights and thresholds of BP neural network.

4) Calculating the fitness value to optimize the weights and thresholds of BP neural network. This step includes genetic algorithm, particle swarm optimization and cuckoo search algorithm.

The part of genetic algorithm: the error obtained by training BP neural network is taken as the fitness value. After selection, cross and variation, the fitness value is calculated. The population scale of the GA is set as 25, the crossover probability is set as 0.7, and the mutation probability is set as 0.1. If the condition is met, the step 5 should be conducted. Otherwise, the previous operation is repeated.

The part of particle swarm optimization: initializing the position and velocity of the particles. The population scale of the PSO is set as 25, c_1 and c_2 are set as 1.5. The fitness value of particles is calculated. If the current fitness value is better than individual extremum, the individual extremum should be updated. If the current fitness value is better than global extremum, the global extremum should be update. Otherwise, the particle's position and speed be updated directly. If the condition is met, the step 5 should be conducted. Otherwise, the previous operation is repeated.

The part of cuckoo search algorithm: initializing the location of the nest and setting relevant parameters. The number of nests is set as 25 and probability P_a is set as 0.25. The fitness value of nests is calculate to select the current best nest. According to Eq. (12), the location of the nest is updated

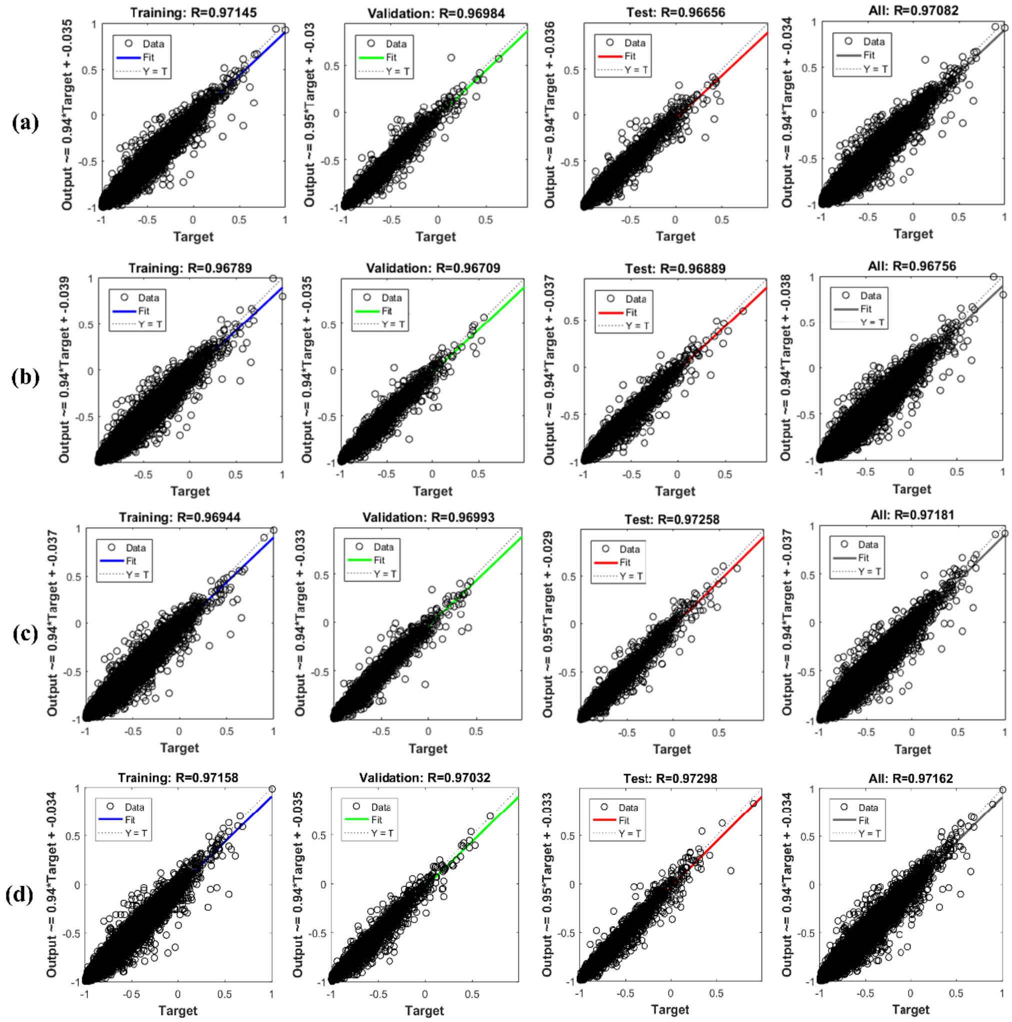


FIGURE 4. The correlation coefficients between the predicted output data and the simulated output data at the three stages of training, verification and testing based on (a) BP, (b) GA-BP, (c) PSO-BP and (d) CS-BP model.

to obtain a new location of nest and calculate its fitness value. By comparing the updated new nest with the previous generation one, the nest with the best fitness value is selected as the current one. Compared the random number γ generated by Eq. (14) with Pa, nests with smaller probability are retained, while nests with bigger probability continue to be updated and improved. If the condition is met, the step 5 should be conducted. Otherwise, the previous operation is repeated.

5) Getting the optimal weights and thresholds, and then updating them.

6) The BP neural network be trained to predict the Hs.

VI. SIMULATION RESULTS AND DISCUSSION

The Hs data simulated by WW3 in the South China Sea and wind data are used for the model applications. According to the theory that wave is created by wind interacting with the ocean surface, different combinations of data sets are imposed as input variable to predict Hs. In addition, different model structures are adopted. After the Hs are predicted based on CS-BP model, the network learning and prediction results will be checked to evaluate the prediction performance of the model. In order to achieve this goal, the data set is first divided into training, test and verification set. Then, correlation analysis is carried out on the prediction results and the models' generalization ability is compared. The prediction results are further evaluated and performance of the CS-BP model is also compared with the classical ARIMA, ELM, BP model, GA-BP and PSO-BP model. Through the comparison of error indicators, the influence of different models on the prediction results is analyzed when the input parameters and network structure are determined and the influence of different input parameters and network structure on the prediction results is analyzed when the prediction model is determined.

A. CORRELATION ANALYSIS OF PREDICTION RESULTS

In the South China Sea, we used the first 16434 sets of data as the training data set, and the latter 100 sets as the test

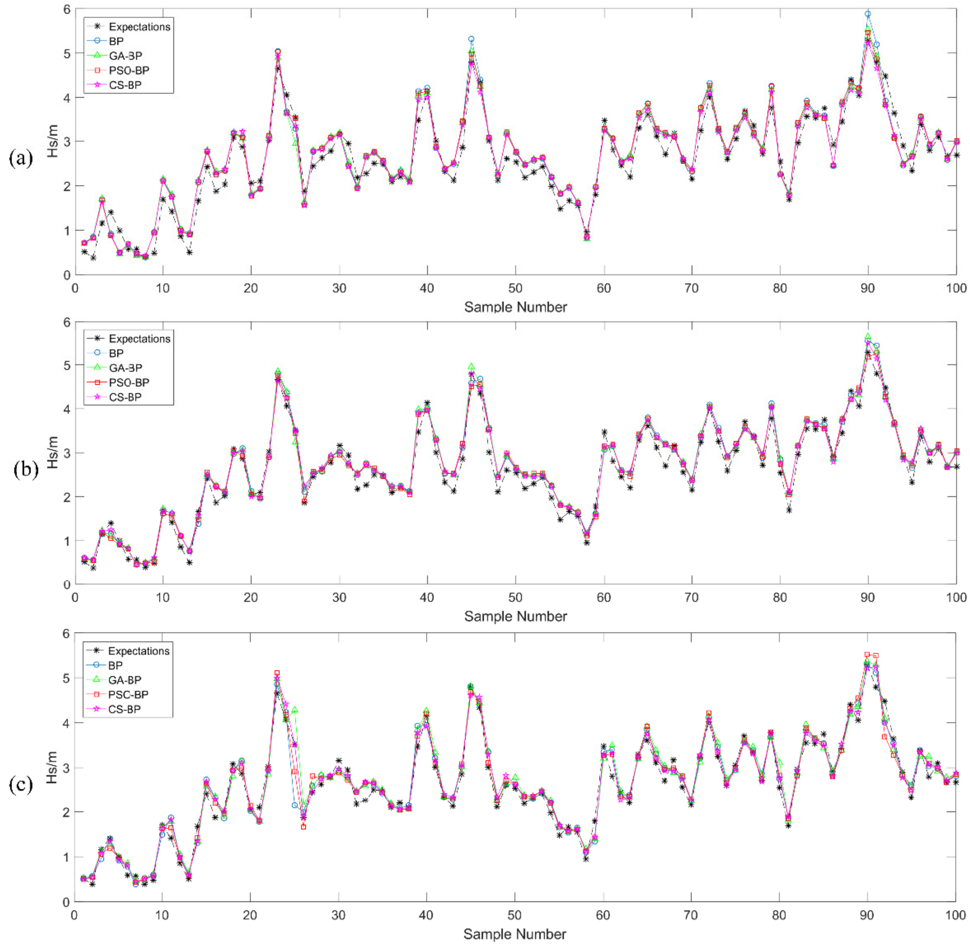


FIGURE 5. The expected and predicted significant wave heights based on BP, GA-BP, PSO-BP and CS-BP obtained by (a) (3-7-1), (b) (6-13-1), and (c) (9-19-1) network structure.

data set to predict Hs. In the modeling process, the default samples are randomly divided into three categories: training samples, validation samples and test samples. In Figure 4. four graphs show the correlation coefficients between the prediction output data and the simulated output data at three stages of training, verification and testing based on BP, GA-BP, PSO-BP and CS-BP model. The BP model is effective in training stage, the effect of the validation stage becomes worse, and the effect of the test stage is poor. These indicate that the prediction results of BP model for the unknown data are worst in four models and the generalization ability of BP needs to be improved. The prediction performance of the GA-BP model is relative equilibrium in three stages of training, verification, testing. In test stage, the effect of GA-BP model is better than the BP, and the prediction performance of PSO-BP and CS-BP model are better than that of other models, as well as their respective training stage. All these indicate that GA, PSO and CS can not only improve global searching ability but also improve the generalization ability of BP model, and the generalization ability of CS-BP is the best with the correlation coefficient of 0.97298 in the test stage.

B. PERFORMANCE COMPARISON OF THE BP, GA-BP, PSO-BP AND CS-BP MODEL

The Hs in the South China Sea are predicted by CS-BP model. In order to better compare the prediction effect of CS-BP model, not only CS-BP model, but also ARIMA, ELM, BP and GA-BP model are proposed. The accuracy in predicting the Hs was then judged by four performance criteria [37], [38]. They are mean absolute error (MAE), mean absolute percentage error (MAPE), root mean square error (RMSE) and root mean squared percentage error (RMSPE), and mathematical expressions of the performance criteria are given by:

The MAE can be calculated as

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|X_i - P_i| \tag{16}$$

The MAPE can be calculated as

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{X_i - P_i}{X_i}\right| \times 100\% \tag{17}$$

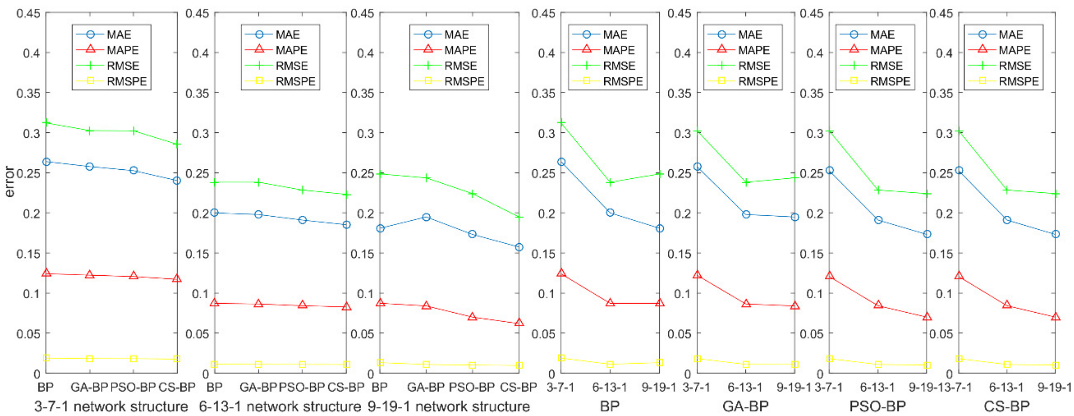


FIGURE 6. The error changes in different cases.

TABLE 2. The MAE, MAPE, RMSE and RMSPE of the ARIMA, ELM, BP, GA-BP, PSO-BP and CS-BP model.

Model	MAE	MAPE	RMSE	RMSPE
ARIMA	0.1787	0.0868	0.2152	0.0122
ELM	0.1796	0.0809	0.2290	0.0112
BP(3-7-1)	0.2639	0.1241	0.3121	0.0189
GA-BP(3-7-1)	0.2577	0.1223	0.3026	0.0183
PSO-BP(3-7-1)	0.2526	0.1206	0.3023	0.0182
CS-BP(3-7-1)	0.2402	0.1171	0.2854	0.0177
BP(6-13-1)	0.1999	0.0871	0.2382	0.0113
GA-BP(6-13-1)	0.1979	0.0864	0.2382	0.01112
PSO-BP(6-13-1)	0.1910	0.0842	0.2284	0.0110
CS-BP(6-13-1)	0.1852	0.0827	0.2231	0.0109
BP(9-19-1)	0.1808	0.0873	0.2485	0.0133
GA-BP(9-19-1)	0.1949	0.0840	0.2436	0.0110
PSO-BP(9-19-1)	0.1731	0.070	0.2240	0.0102
CS-BP(9-19-1)	0.1575	0.0620	0.1950	0.0097

The RMSE can be calculated as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i - P_i)^2} \tag{18}$$

The RMSPE can be calculated as

$$RMSPE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i - P_i}{X_i}\right)^2} \times 100\% \tag{19}$$

where X_i is the simulated output of the test data, and P_i is the predicted output of the test data.

Four performance criteria are proposed to eliminate the influence of weather and other uncertainties. The expected and predicted Hs based on BP, GA-BP, PSO-BP and CS-BP model are shown in Figure 5. From Figure 5, the predicted results are consistent with the expectation. The table 2 shows four performance criteria of BP, GA-BP, PSO-BP and CS-BP model. The error changes in different cases are shown in Figure 6. In the PSO-BP and CS-BP model, the MAE, MAPE, RMSE and RMSPE all decrease with the increase of input parameters and number of hidden layers. However, the MAPE, RMSE and RMSPE of 9-19-1 neural network are greater than 6-16-1 neural network in the BP model and, the RMSE of 9-19-1neural network is greater than 6-16-1 neural network in the GA-BP model. It can be concluded that the proper increase of the number of input parameters and hidden layers can reduce the prediction error and improve the prediction accuracy. In the same network structure, MAPE, RMSE and RMSPE of BP network are the largest, that of CS-BP is the smallest, and that of GA-BP and PSO-BP are between them. It can be seen that GA, PSO and CS can optimize BP neural network to improve prediction accuracy. In addition, in 9-19-1 network structure, the MAE of the GA-BP model is the largest. This is because of cross and mutation operations in genetic algorithms, which can destroy original genes. The destroyed genes can lead to bigger errors, namely worse predictions. However, both PSO and CS can avoid this problem. The former has the function of memory, the particles move towards the direction of local optimal and global optimal in each iteration and movement, and can approach gradually according to the current speed. The latter can approach gradually the optimal value based on the location of the updated nest. Therefore, PSO and CS have better optimization ability than genetic algorithm. From table 2, we can clearly see that the optimal prediction results of ARIMA and ELM are slight better that the 9-19-1 BP neural network and the prediction results of the 9-19-1 CS-BP model are the best compared with any other model. This is because that Lévy flights is used in the global search for CS, unlike the standard random walks used in other algorithms. In the algorithm flow, we found that CS is simple, which GA needs selection, crossover and mutation to make it complicated. In addition, the CS-BP model has another advantage that once the model is trained, it can be exploited as a quick, accurate tool for indirect estimation of Hs. In summary, all the results show that the CS-BP model is feasible and reliable in predicting Hs.

VII. CONCLUSION

In this study, the CS-BP model for Hs prediction was proposed. The motivation behind this research is the significance of rapid and accurate Hs prediction in many applications including marine engineering construction and emergency prevention. The study covers a wide range of geographical locations and different weather, making the data universal and available. This study based on CS-BP model for Hs prediction in South China Sea, has led to several conclusions.

1) GA, PSO and CS can improve the generalization ability of BP model. However, compared with BP, GA-BP and PSO-BP model, the CS-BP model has the best prediction performance in the test stage, with the correlation coefficient of 0.97298, which is better than its own training stage. The CS-BP model has the best generalization ability.

2) In the 9-19-1 network structure, the MAE of GA-BP model is the largest. The reason causing this phenomenon is that cross and mutation operation in genetic algorithm can sometimes destroy good genes, which lead to large error. Except for this phenomenon, under the same input parameters and network structure, the prediction accuracy of CS-BP model is the best, followed by PSO-BP, GA-BP and BP. These mean that CS-BP model is best compared with BP, GA-BP and CS-BP model.

3) With the appropriate increase of the number of input parameters and hidden layers, the prediction error of the model is reduced and the prediction accuracy is improved. The study revealed that compared with other models (including ARIMA and ELM), predicted results in the South China Sea using the 9-19-1 CS-BP model are best according to the performance criteria obtained: MAE of 0.1575, MAPE of 0.0620, RMSE of 0.1950 and RMSPE of 0.0097. The reasons for good results are that the nest approach gradually the optimal value based on the location of the updated nest, and unlike the standard random walks used by other algorithms, Lévy flights is used in the global search for CS.

Through comparison, the CS-BP model can predict the ocean Hs quickly and accurately in some degree.

ACKNOWLEDGMENT

The authors would like to thank the NASA Earth scientific enterprise for providing CCMP wind data and the NECP, The WAW for the support on the WW3 model. The author would also like to thank the editor and anonymous reviewers for their constructive comments on the manuscript.

REFERENCES

[1] P. Jain and M. C. Deo, "Real-time wave forecasts off the western Indian coast," *Appl. Ocean Res.*, vol. 29, nos. 1–2, pp. 72–79, 2007.
[2] P. Jain, M. C. Deo, G. Latha, and V. Rajendran, "Real time wave forecasting using wind time history and numerical model," *Ocean Model.*, vol. 36, nos. 1–2, pp. 26–39, 2011.
[3] S. Gaur and M. C. Deo, "Real-time wave forecasting using genetic programming," *Ocean Eng.*, vol. 35, nos. 11–12, pp. 1166–1172, 2008.
[4] A. Akpınar, B. Bingölbali, and G. P. Van Vledder, "Long-term analysis of wave power potential in the Black Sea, based on 31-year SWAN simulations," *Ocean Eng.*, vol. 130, pp. 482–497, Jan. 2017.
[5] M. L. Khandekar, *Operational Analysis and Prediction of Ocean Wind Waves*. Springer, 1989, p. 214. [Online]. Available: https://link.springer.com/book/10.1007%2F978-1-4613-8952-1
[6] H. Yoon, S.-C. Jun, Y. Hyun, G.-O. Bae, and K.-K. Lee, "A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer," *J. Hydrol.*, vol. 396, nos. 1–2, pp. 128–138, 2011.
[7] W. Wang, R. Tang, C. Li, P. Liu, and L. Luo, "A BP neural network model optimized by mind evolutionary algorithm for predicting the ocean wave heights," *Ocean Eng.*, vol. 162, pp. 98–107, Aug. 2018.
[8] J. D. Agrawal and M. C. Deo, "On-line wave prediction," *Mar. Struct.*, vol. 15, no. 1, pp. 57–74, Jan./Feb. 2002.
[9] J. Mahjoobi and E. A. Mosabbeb, "Prediction of significant wave height using regressive support vector machines," *Ocean Eng.*, vol. 36, no. 5, pp. 339–347, Apr. 2009.
[10] M. Özger, "Significant wave height forecasting using wavelet fuzzy logic approach," *Ocean Eng.*, vol. 37, no. 16, pp. 1443–1451, 2010.
[11] S. P. Nitsure, S. N. Londhe, and K. C. Khare, "Wave forecasts using wind information and genetic programming," *Ocean Eng.*, vol. 54, pp. 61–69, Nov. 2012.
[12] A. R. Kambekar and M. C. Deo, "Wave prediction using genetic programming and model trees," *J. Coastal Res.*, vol. 279, pp. 43–50, Jul. 2012.
[13] D. Q. Truong and K. K. Ahn, "Wave prediction based on a modified grey model MGM(1,1) for real-time control of wave energy converters in irregular waves," *Renew. Energy*, vol. 43, pp. 242–255, Jul. 2012.
[14] J. C. Fernández, S. Salcedo-Sanz, P. A. Gutiérrez, E. Alexandre, and C. Hervás-Martínez, "Significant wave height and energy flux range forecast with machine learning classifiers," *Eng. Appl. Artif. Intell.*, vol. 43, pp. 44–53, Aug. 2015.
[15] S. Shahabi, M.-J. Khanjani, and M.-R. H. Kermani, "Significant wave height modelling using a hybrid wavelet-genetic programming approach," *KSCE J. Civil Eng.*, vol. 21, no. 1, pp. 1–10, 2016.
[16] L. Cornejo-Bueno, J. C. Nieto-Borge, P. García-Díaz, G. Rodríguez, and S. Salcedo-Sanz, "Significant wave height and energy flux prediction for marine energy applications: A grouping genetic algorithm—Extreme learning machine approach," *Renew. Energy*, vol. 97, pp. 380–389, Nov. 2016.
[17] S. Yang, Z. Zhang, L. Fan, T. Xia, S. Duan, C. Zheng, X. Li, and H. Li, "Long-term prediction of significant wave height based on SARIMA model in the South China Sea and adjacent waters," *IEEE Access*, vol. 7, pp. 88082–88092, 2019.
[18] J. Wang, Z. Sheng, B. Zhou, and S. Zhou, "Lightning potential forecast over Nanjing with denoised sounding-derived indices based on SSA and CS-BP neural network," *Atmos. Res.*, vol. 137, pp. 245–256, Feb. 2014.
[19] L.-M. Zhou, L.-Y. Wu, P.-F. Guo, and A.-F. Wang, "Simulation and study of wave in South China Sea using wavewatch-III," *J. Tropical Oceanogr.*, vol. 26, no. 5, pp. 3–10, 2007.

[20] C. Zheng, H. Zhuang, X. Li, and X. Li, "Wind energy and wave energy resources assessment in the East China Sea and South China Sea," *Sci. China Technol. Sci.*, vol. 55, no. 1, pp. 163–173, 2011.

[21] R. M. Atlas, J. V. Ardizzone, R. N. Hoffman, and J. C. Jusem, "The cross-calibrated, multi-platform (CCMP) ocean surface wind product: Current status and plans," in *Proc. AMS 13th Conf. Integr. Observing Assimilation Syst. Atmos., Oceans, Land Surf.*, Phoenix, AZ, USA, Jan. 2009.

[22] E. Ozgan, "Artificial neural network based modelling of the marshall stability of asphalt concrete," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 6025–6030, 2011.

[23] J.-R. Zhang, J. Zhang, T.-M. Lok, and M. R. Lyu, "A hybrid particle swarm optimization–back-propagation algorithm for feedforward neural network training," *Appl. Math. Comput.*, vol. 185, no. 2, pp. 1026–1037, 2007.

[24] R. Adhikari, R. K. Agrawal, and L. Kant, "PSO based neural networks vs. traditional statistical models for seasonal time series forecasting," in *Proc. 3rd IEEE Int. Adv. Comput. Conf.*, Feb. 2013, pp. 719–725.

[25] J. L. McClelland and D. E. Rumelhart, *Explorations in Parallel Distributed Processing: A Handbook of Models, Programs, and Exercises*. 1st ed. Cambridge, MA, USA: MIT Press, 1988.

[26] H.-M. Lee, C.-M. Chen, and T.-C. Huang, "Learning efficiency improvement of back-propagation algorithm by error saturation prevention method," *Neurocomputing*, vol. 41, nos. 1–4, pp. 125–143, Oct. 2001.

[27] I. A. Basheer and M. Hajmeer, "Artificial neural networks: Fundamentals, computing, design, and application," *J. Microbiol. Methods*, vol. 43, no. 1, pp. 3–31, Dec. 2000.

[28] W. Bi, X. Wang, Z. Tang, and H. Tamura, "Avoiding the local minima problem in backpropagation algorithm with modified error function," *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, vol. E88-A, no. 12, pp. 3645–3653, 2005.

[29] H. Li, C.-X. Hu, and Y. Li, "Application of the purification of materials based on GA-BP," *Energy Procedia*, vol. 17, pp. 762–769, Jan. 2012.

[30] S. Wang, N. Zhang, L. Wu, and Y. Wang, "Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method," *Renew. Energy*, vol. 94, pp. 629–636, Aug. 2016.

[31] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, Perth, WA, Australia, Nov./Dec. 1995, pp. 1942–1948.

[32] A. Ismail, D. S. Jeng, and L. L. Zhang, "An optimised product-unit neural network with a novel PSO–BP hybrid training algorithm: Applications to load–deformation analysis of axially loaded piles," *Eng. Appl. Artif. Intell.*, vol. 26, no. 10, pp. 2305–2314, 2013.

[33] D. J. Armaghani, M. Hajihassani, E. T. Mohamad, A. Marto, and S. A. Noorani, "Blasting-induced flyrock and ground vibration prediction through an expert artificial neural network based on particle swarm optimization," *Arabian J. Geosci.*, vol. 7, no. 12, pp. 5383–5396, 2013.

[34] E. T. Mohamad, D. J. Armaghani, E. Momeni, A. H. Yazdavar, and M. Ebrahimi, "Rock strength estimation: A PSO-based BP approach," *Neural Comput. Appl.*, vol. 30, no. 5, pp. 1635–1646, 2016.

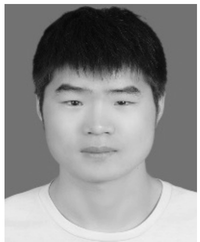
[35] X.-S. Yang and S. Deb, "Cuckoo Search via Lévy flights," in *Proc. World Congr. Nature Biol. Inspired Comput.*, Dec. 2009, pp. 210–214.

[36] A. H. Gandomi, X.-S. Yang, and A. H. Alavi, "Cuckoo search algorithm: A metaheuristic approach to solve structural optimization problems," *Eng. Comput.*, vol. 29, no. 1, pp. 17–35, 2011.

[37] G. W. Chang, H. J. Lu, Y. R. Chang, and Y. D. Lee, "An improved neural network-based approach for short-term wind speed and power forecast," *Renew. Energy*, vol. 105, pp. 301–311, May 2017.

[38] J. A. Rodger, "A fuzzy nearest neighbor neural network statistical model for predicting demand for natural gas and energy cost savings in public buildings," *Expert Syst. Appl.*, vol. 41, no. 4, pp. 1813–1829, 2014.

SHAOBO YANG received the Ph.D. degree in instrument and meter engineering from the School of Precision Instrument and Opto-Electronics Engineering, Tianjin University, where he is currently a Teacher. His current research interest includes ocean observation and exploration technology, such as atmosphere duct measuring technology, deep Argo float, and wave assessment.



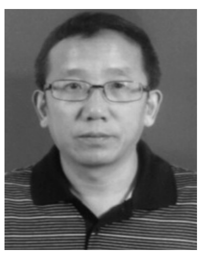
TIANLIANG XIA received the bachelor's degree in electrical engineering and automation from the Henan Institute of Science and Technology, in 2017. He is currently pursuing the M.S. degree in instrument and meter engineering from the School of Precision Instrument and Opto-Electronics Engineering, Tianjin University. His current research interest includes the assessment and prediction of the height of ocean waves.

ZHENQUAN ZHANG received the degree from the Hefei University of Technology, in 2017. He is currently pursuing the master's degree in instrument and meter engineering with the School of Precision Instrument and Opto-Electronics Engineering, Tianjin University. His current research interests include machine learning algorithm and marine data processing.

CHONGWEI ZHENG received the Ph.D. degree in atmospheric from the National University of Defense Technology. He is with the Dalian Naval Academy. He is also a Visiting Scholar with the Shandong Provincial Key Laboratory of Ocean Engineering, Ocean University of China and the State Key Laboratory of Estuarine and Coastal Research. His current research interests include marine new energy evaluation and physical oceanography.

XINGFEI LI is currently a Professor with the School of Precision Instrument and Opto-Electronics Engineering, Tianjin University. His current research interests include multisensor fusion technology and measurement technology.

HONGYU LI received the Ph.D. degree in instrument and meter engineering from the School of Precision Instrument and Opto-Electronics Engineering, Tianjin University. He is currently a Teacher with the Shandong University of Science and Technology. His current research interest includes ocean observation and exploration technology, such as atmosphere duct measuring technology, deep Argo float, and wave assessment.



JIANJUN XU is currently a Professor with the South China Sea Institute of Marine Meteorology, Guangdong Ocean University, and a Researcher with the Southern Marine Science and Engineering Guangdong Laboratory, Zhanjiang. His current research interests include marine meteorology, air-sea interaction, satellite measurement technology, and data assimilation.