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

Shaobo Yang1,2, Tianliang Xia1, Zhenquan Zhang1, Chongwei Zheng4,5,6, Xingfei Li1,3, Hongyu Li7 and Jianjun Xu2,8

ABSTRACT Forecasting the significant wave heights (Hs) is indispensable in Hs-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 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].

Deo [8] et al. 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 Mosabbeb [9] et al. proposed a support vector machine (SVM) method for predicting wave heights. Compared with the traditional neural network, the prediction error of SVM is small. Ozger [10] et al. applied the hybrid method of wavelet and fuzzy to predict the significant wave heights. The results showed that the model has higher prediction reliability. Cornejo-Bueno [16] et al. 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 [7] et al. 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 [17] et al. 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 $H_s$, 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 $H_s$ in the South China Sea, transforming the uncertainty of $H_s$ into predictability.

II. RESEARCH METHODOLOGY

According to the theory that wave is created by wind interacting with the ocean surface, $H_s$ and wind data are combined with neural network to predict $H_s$. The Figure 1. shows the research method for predicting $H_s$, 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 $H_s$ 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 $H_s$ using neural network model and analysis of prediction results. The $H_s$ are predicted by the neural network model. It is compared with the $H_s$ from WAVESWATCH-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 $H_s$ 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 $H_s$ at the nine prediction points, which cover a wide range of geographical locations and different weather conditions. This could make
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 16,434 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. 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).

$$q = 2p + 1$$

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 I.

V. PREDICTION MODEL

Recently, the neural network has developed rapidly and achieved good results in economy, biology, ocean and other aspects. Zhang [23] et al. studied a hybrid algorithm combining PSO and BP together to train the weights and thresholds of the network. 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 16,434 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. 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).

$$q = 2p + 1$$

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 I.
learning problem of connection weight of hidden layer in multi-layer neural network

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Number of input layers</th>
<th>Number of hidden layers</th>
<th>Network structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>array of all ones, V10(t) and U10(t)</td>
<td>3</td>
<td>7</td>
<td>3-7-1</td>
</tr>
<tr>
<td>array of all ones, V10(t), U10(t), V10(t-1), U10(t-1) and Hs(t-1)</td>
<td>6</td>
<td>13</td>
<td>6-13-1</td>
</tr>
<tr>
<td>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)</td>
<td>9</td>
<td>19</td>
<td>9-19-1</td>
</tr>
</tbody>
</table>

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:

$$X = \sum_{i=1}^{n} x_i o_i + \theta$$  \hspace{1cm} (2)

Where $n$ is the number of inputs, and $x_i$ and $o_i$ demonstrate respectively the value of $i$th input and weight. The thresholds applied to the neurons is denoted by $\theta$.

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. OPTIMIZED 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.

$$f_i = \frac{k}{F_i}$$  \hspace{1cm} (3)

$$p_i = \frac{f_i}{\sum_{j=1}^{n} f_j}$$  \hspace{1cm} (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 $k$th chromosome $a_k$ and the $l$th chromosome $a_l$ at position $j$ is performed as follows.

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

$$a_{lj} = a_{lj} (1-b) + a_{kj} b$$  \hspace{1cm} (5)

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

The $j$th gene $a_{kj}$ of the $i$th individual was selected to be mutated by the following methods.

$$a_{kj} = \begin{cases} a_{kj} + (a_{max} - a_{min}) * f(g) & r > 0.5 \\ a_{kj} + (a_{min} - a_{kj}) * f(g) & r \leq 0.5 \end{cases}$$  \hspace{1cm} (6)

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

Where $a_{max}$ and $a_{min}$ are upper and lower bounds of gene $a_{kj}$ 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:

$$X_i^{k+1} = X_i^k + V_i^{k+1}$$ (8)

$$V_i^{k+1} = \omega V_i^k + c_1 r_1 (P_{i,j}^k - X_i^k) + c_2 r_2 (G_{i}^k - X_i^k)$$ (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 $P_{i,j}$ and $G_i$, in Eq. (9). Besides, $\omega$ 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.

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

Where $\omega_{\text{start}}$ is the initial inertia weight. $\omega_{\text{end}}$ is the inertia weight at the maximum number of iterations. $k$ and $T_{\text{max}}$ are respectively the number of current iteration and the maximum number of iteration. In this paper, $\omega_{\text{start}}$ is 0.9 and $\omega_{\text{end}}$ is 0.4.

The cuckoo search algorithm (CS) was proposed in 2009 by Yang and Deb, Cambridge scholars [35]. It is a meta-heuristic algorithm that simulates cuckoo parasitic reproduction and combines the cuckoo reproduction process with the Lévy flight search method [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:

$$x_i = r \times (Ub - Lb)$$ (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:

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

Where $x_i^{(t+1)}$ and $x_i^{(t)}$ denote the new solution and current location unit, respectively. $\alpha$ is the step size associated with the optimized problem, and the step size $\alpha = 1$ is selected. The notation $\otimes$ 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):

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

where $x_b^{(t)}$ is the best location, $\beta=1.5$, $\mu$ and $\nu$ are deduced 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

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

Where $r$ and $\gamma$ are the uniform random numbers within the interval [0, 1]. $x_i^{(t)}$ and $x_b^{(t)}$ 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):

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

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

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 to obtain a new location of the 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 \(\gamma\) generated by Eq. (14) with \(P_a\), 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 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 (a) BP, (b) GA-BP, (c) PSO-BP and (d) CS-BP model.

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.

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 [38], [39]. 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

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |X_i - P_i|$$  \hspace{1cm} (16)$$

The MAPE can be calculated as
\[ \text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{X_i - P_i}{X_i} \right| \times 100\% \quad (17) \]

The RMSE can be calculated as
\[ \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - P_i)^2} \quad (18) \]

The RMSPE can be calculated as
\[ \text{RMSPE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left( \frac{X_i - P_i}{X_i} \right)^2} \times 100\% \quad (19) \]

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 \( H_s \) 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 II 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 9-19-1 neural 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 the 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 II, 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 \( H_s \). In summary, all the results show that the CS-BP model is feasible and reliable in predicting \( H_s \).

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.

FIGURE 6. The error changes in different cases.

TABLE II

<table>
<thead>
<tr>
<th>Model</th>
<th>MAE</th>
<th>MAPE</th>
<th>RMSE</th>
<th>RMSPE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARIMA</td>
<td>0.1787</td>
<td>0.0868</td>
<td>0.2152</td>
<td>0.0122</td>
</tr>
<tr>
<td>ELM</td>
<td>0.1796</td>
<td>0.0809</td>
<td>0.2290</td>
<td>0.0112</td>
</tr>
<tr>
<td>BP(3-7-1)</td>
<td>0.2639</td>
<td>0.1241</td>
<td>0.3121</td>
<td>0.0189</td>
</tr>
<tr>
<td>GA-BP(3-7-1)</td>
<td>0.2577</td>
<td>0.1223</td>
<td>0.3026</td>
<td>0.0183</td>
</tr>
<tr>
<td>PSO-BP(3-7-1)</td>
<td>0.2526</td>
<td>0.1206</td>
<td>0.3023</td>
<td>0.0182</td>
</tr>
<tr>
<td>CS-BP(3-7-1)</td>
<td>0.2402</td>
<td>0.1171</td>
<td>0.2854</td>
<td>0.0177</td>
</tr>
<tr>
<td>BP(6-13-1)</td>
<td>0.1999</td>
<td>0.0871</td>
<td>0.2382</td>
<td>0.0113</td>
</tr>
<tr>
<td>GA-BP(6-13-1)</td>
<td>0.1979</td>
<td>0.0864</td>
<td>0.2382</td>
<td>0.0112</td>
</tr>
<tr>
<td>PSO-BP(6-13-1)</td>
<td>0.1910</td>
<td>0.0842</td>
<td>0.2284</td>
<td>0.0110</td>
</tr>
<tr>
<td>CS-BP(6-13-1)</td>
<td>0.1852</td>
<td>0.0827</td>
<td>0.2231</td>
<td>0.0109</td>
</tr>
<tr>
<td>BP(9-19-1)</td>
<td>0.1808</td>
<td>0.0873</td>
<td>0.2485</td>
<td>0.0133</td>
</tr>
<tr>
<td>GA-BP(9-19-1)</td>
<td>0.1949</td>
<td>0.0840</td>
<td>0.2436</td>
<td>0.0110</td>
</tr>
<tr>
<td>PSO-BP(9-19-1)</td>
<td>0.1731</td>
<td>0.0700</td>
<td>0.2240</td>
<td>0.0102</td>
</tr>
<tr>
<td>CS-BP(9-19-1)</td>
<td>0.1575</td>
<td>0.0620</td>
<td>0.1950</td>
<td>0.0097</td>
</tr>
</tbody>
</table>

VII. CONCLUSIONS

In this study, the CS-BP model for \( H_s \) prediction was proposed. The motivation behind this research is the significance of rapid and accurate \( H_s \) 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 \( H_s \) 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.

VIII. ACKNOWLEDGEMENTS

This work was supported by the National Key R&D Program of China (No. 2018YFC1506902), the Fund of Southern Marine Science and Engineering Guangdong Laboratory (Zhanjiang) (No. ZJW-2019-XX), State Key Laboratory of Estuarine and Coastal Research (Grant number SKLEC-KF201707), the State Key Laboratory of Precision Measuring Technology and Instruments of Tianjin University, the National Natural Science Foundation of China and the Innovation and Entrepreneurship Foundation of Tianjin University, the National Natural Science Key Laboratory of Precision Measuring Technology and Instruments of Tianjin University, the National Natural Science Foundation of China and the Innovation and Entrepreneurship Project of Ocean Engineering Institute of Tianjin University (No. 20190201-5), the Opening Fund of Shandong Provincial Key Laboratory of Ocean Engineering. The authors acknowledged 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


