IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 25, 2019, accepted June 20, 2019, date of publication June 26, 2019, date of current version July 18, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2925107

Long-Term Prediction of Significant Wave Height Based on SARIMA Model in the South China Sea and Adjacent Waters

SHAOBO YANG[1,2,3], ZHENQUAN ZHANG[1], LINLIN FAN[1], TIANLIANG XIA[1], SHANHUA DUAN[1], CHONGWEI ZHENG[4,5], XINGFEI LI[1,2,3], AND HONGYU LI[2,6]

[1]State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China
[2]Pilot National Laboratory for Marine Science and Technology, Qingdao 266237, China
[3]Qingdao Institute for Ocean Technology, Tianjin University, Qingdao 266200, China
[4]State Key Laboratory of Estuarine and Coastal Research, Shanghai 200241, China
[5]State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing 100029, China
[6]School of Mechanical and Electronic Engineering, Shandong University of Science and Technology, Qingdao 266590, China

Corresponding author: Chongwei Zheng (chinaoceanzcw@sina.cn)

This work was supported in part by the Pilot National Laboratory for Marine Science and Technology, Qingdao, through the Wenhai Project, under Grant ZR2016WH01, in part by the National Natural Science Foundation of China under Grant 61733012, in part by the Innovation and Entrepreneurship Project of Ocean Technology Institute of Tianjin University under Grant 20190201-5, in part by the State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, and in part by the Scientific Research Fund for Young Teachers under Grant Pilq1702.

ABSTRACT Nowadays, the world is facing the dual crisis of the energy and environment, and renewable energy, such as wave energy, can contribute to the improvement of the energy structure of the world, enhance energy supply and improve the environment in the framework of sustainable development. Long-term prediction of the significant wave height (SWH) is indispensable in SWH-related engineering studies and is exceedingly important in the assessment of wave energy in the future. In this paper, the spatial and temporal characteristics of wave energy in the South China Sea (SCS), and adjacent waters are analyzed. The results show that there are abundant wave energy resources in the waters around the Taiwan Strait, the Luzon Strait, and the north part of the SCS with annual average SWH (SWH) of over 1.4 m and obvious increasing trend. Then, the SARIMA approach considers the relationship between the current time and the values, residuals at some previous time and the periodicity of the SWH series are proposed to forecast the SWH in the SCS and adjacent waters. The results obtained are promising, showing good performance of the prediction of monthly average SWH in the SCS and adjacent waters.

INDEX TERMS SARIMA, long-term prediction, significant wave height (SWH).

I. INTRODUCTION

As the second decade of the 21st century passing away, the two problems, crisis of resource and environmental pollution, are approaching gradually. The development and utilization of renewable energy sources will play an important role in the future sustainability of human society [1]–[3]. Wave energy has received much attention recently due to its large storage capacity, wide distribution and pollution-free, as well as the conservation of terrestrial resources and needing no immigration [4], [5]. Since the temporal and spatial variability of wave height, highly irregular frequency and direction, it is often described as an unreliable source of energy. In addition to this, if the risk and uncertainty level increases in expected generation, it will function as an inhibiting factor toward energy security. Therefore, the availability of accurate SWH forecasts in the management routines involved in Wave Energy Converters (WECs) provides considerable support for optimizing the operational costs and improving their reliability [6], [7]. In the other words, by applying accurate SWH forecasting, the wave energy can be scheduled and wave power penetration will be increased. This has significant economic impact on the system operation and can substantially reduce costs. Therefore, applying SWH prediction methods offering the best possible accuracy over a number of time scales is required [8].

The associate editor coordinating the review of this manuscript and approving it for publication was Arup Kumar Goswami.

As mentioned, wave prediction is a difficult and vital subject for development and utilization of wave energy resource due to its stochastic nature, so the research work on this topic has been intense in the last few years [9]–[11]. For instance, Mahjoobi and Etemand-Shahidi investigated the performances of classification and regression trees for prediction of wave height in 2008. Wind speed and wind direction are given as input variables, and the SWH is the output parameter. Results indicate that the decision tree, as an efficient novel approach with an acceptable range of error, can be used successfully for forecasting SWH [12]. And then, Mahjoobi and Ehsan Adeli Mosabbeb employed support vector machine (SVM) to predict SWH and compared SVM results with those of artificial neural network (ANN), multilayer perceptron (MLP) and radial basis function (RBF) models. Comparisons indicate that the error statistics of SVM model marginally outperforms ANN even with much less computational time required [13]. Kamranzad et al. employed ANN as a robust data learning method to forecast the wave height for the next few hours in 2011. Meanwhile, they evaluated the effects of different parameters using different models with various combinations of wind and wave parameters [14]. In the same year, Ozger investigated the relationship between ocean wave energy and meteorological variables such as wind speed, air temperature, and sea temperature and proposed a new approach based on an expert system of fuzzy logic modeling to predict wave power [15]. Nitsure et al. obtained satisfactory results by using genetic programming (GP) based models for 12h and 24h ahead to forecast wave heights [16]. The performance of different ordinal and nominal multiclass classifiers is evaluated by J.C. Fernandez et al. The results obtained show an acceptable reconstruction by ordinal methods with respect to nominal ones in terms of wave height and energy flux [17]. L. Cornejo-Bueno et al. presented a hybrid Grouping Genetic Algorithm-Extreme Learning Machine approach (GGA-ELM) for SWH and wave energy flux prediction and obtained good results. This approach can solve feature selection problems and may be applied to alternative regression approaches [18]. Wenxu Wang et al. compared the predicted performances and generalization capabilities of Mind Evolutionary Algorithm-BP neural network strategy (MEA-BP) with the Genetic Algorithm-BP neural network model (GA-BP) and Standard BP neural network model (St-BP). The performance study results demonstrate that MEA-BP performs better than others with faster running time and higher prediction accuracy [19].

The researches mentioned above have made great progress in short-term prediction of wave energy, and they usually focus their attention on the prediction of a single point. These researches can help improve the efficiency of wave energy conversion. So far, the short-term forecasting protocols are relatively well developed for wave energy resources [4], [20]. However, medium- to long-term predictions of wave energy resources, which are central to the development of future operating and trading strategies, for example, when considering whether it is possible to manage the wave power requirements of a remote reef population up to a year in advance, the long-term prediction of wave energy must play a key role in the decision-making process [4], remain scare. In this paper, the model for the long-term and regional prediction of wave energy resources is proposed, which will contribute to the construction of future WECs network and provide a significant tool for optimizing operating costs and improving reliability.

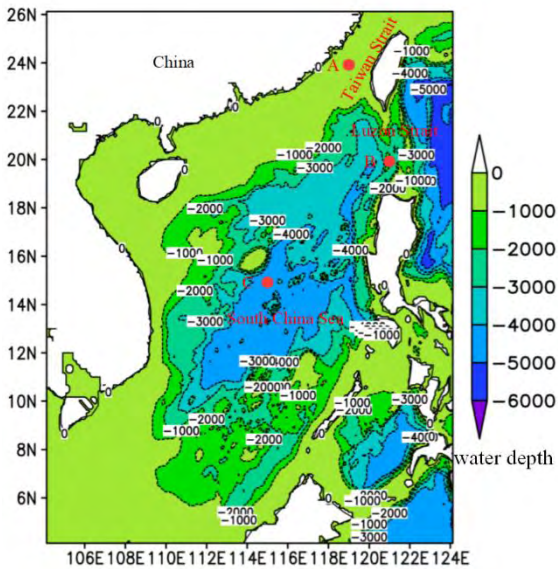

FIGURE 1. Topography of the South China Sea and adjacent waters.

In this paper, we construct SARIMA model based on the periodicity of monthly average SWH and predict it in the South China Sea and adjacent waters in 2017. The paper is structured as follows. Section II presents a brief introduction of the source of the data materials. Then the spatial and temporal characteristics of SWH in the SCS and adjacent waters are analyzed in Section III. In Section IV, we introduce the principle of time series analysis method in detail and the framework of SARIMA model for predicting SWH in the SCS and adjacent waters. Finally, the performance of SARIMA approach is evaluated.

II. STUDY AREA AND METJODLOGY

In this study, the wave field in the South China Sea and adjacent waters were simulated over the period from 0000 UTC on January 1st, 1990 to 1800 UTC on December 31st, 2017 using the third-generation wave model WAVEWATCH-III (WW3), with Cross-Calibrated, Multi-Platform (CCMP) wind filed as driving field. The CCMP data set comes from ESE (NASA Earth Science Enterprise), it combines cross-calibrated satellite microwave winds and instrument observations using a Variational Analysis Method (VAM) to produce high-resolution (0.25 degree) gridded analyses. The range selected for model calculation is 4.125°N-26.125°N, 104.125°E-124.125°E (Figure 1). The spatial resolution takes

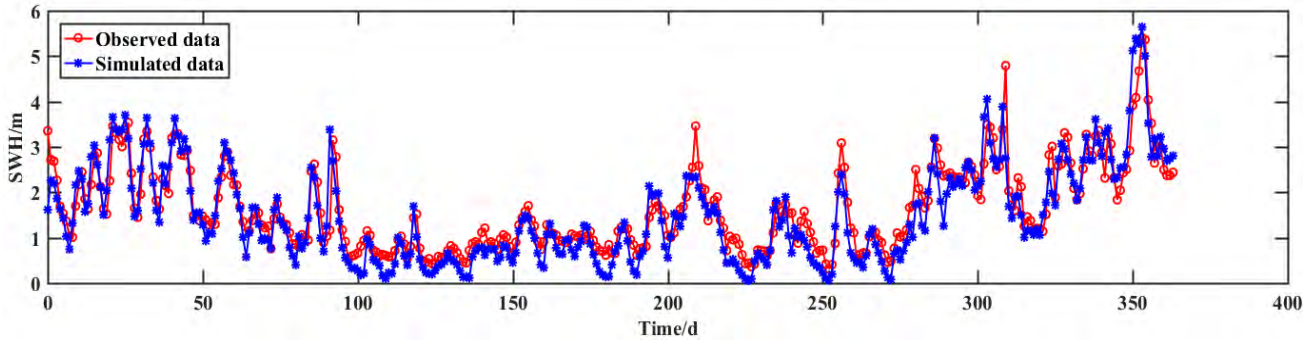

FIGURE 2. Observed and simulated daily-averaged SWH in point C in 2017.

0.25° × 0.25° and the time step takes 900s, with outputs logged every 3h. In this paper, the SWH data from Jason-2 altimetry is regarded as the in situ wave data (Observed data), because of sparse voluntary ship data and no wave buoy data in the SCS and adjacent waters. And then, the WW3 simulated data to be continuous in time and high resolution in space is more suitable for building prediction models compared with the data mentioned above. To validate the precision of the simulated SWH data, three points including A (119°E, 24°N), B (121°E, 20°N) and C (115°E, 15°N) are taken respectively in the Taiwan Strait, the Luzon Strait and the South China Sea. The temporal and spatial resolution of observed and simulated data are inconsistent. Do average of simulated SWH in a day, the daily-averaged SWH of every point is obtained. As shown in Figure 2, the simulated data and observed data have a good consistency in the trend of the curve in point C. And the scatter plot of satellite observed and simulated SWH in the whole domain also is shown in Figure 3.

The Bias, the correlation coefficient (CC) and the root mean square error (RMSE) are calculated individually in point A, B, C and the whole domain in order to better evaluate the accuracy of the simulated SWH. The specific content is shown in Table 1.

$$Bias = \bar{y} - \bar{x} \tag{1}$$

$$CC = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}} \tag{2}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{n} (y_i - x_i)^2} \tag{3}$$

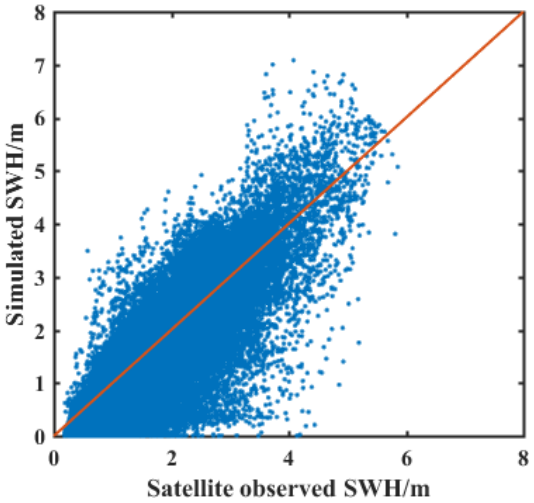

FIGURE 3. Scatter plot of satellite observed and simulated SWH in the whole domain.

where, $\bar{y}$ represents the average value of simulated data, $\bar{x}$ represents the average value of Jason-2 altimetry data, y_i and x_i respectively represent observed SWH and simulated SWH, and N is the total number of samples. As can be seen from Table 1, there is a negative bias of less than 0.15 between the simulated and the observed data in the three points and the bias is −0.254 in the whole domain. It indicates that the simulated data are slightly smaller than the observed data. The CC values of the three points and the whole domain are all above 0.85, and their RMSE values are less than 0.6. From the statistical analysis, we can find that the simulated data is reliable. Previous researches also

TABLE 1. Statistical error of the simulated swh in point A, B, C, and the whole domain.

Location	Longitude/°E	Latitude/°N	Bias	CC	RMSE
A	119	24	-0.140	0.850	0.587
B	121	20	-0.065	0.921	0.463
C	115	15	-0.138	0.932	0.414
The whole domain	-	-	-0.254	0.883	0.521

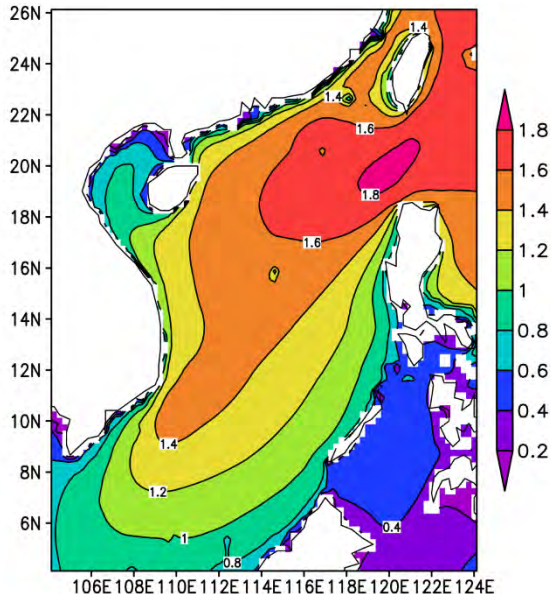

FIGURE 4. Annual average SWH from 1990 to 2017.

indicate the ability of WW3 on wave field simulation by using TOPEX/Poseidon (T/P) satellite SWH data and wave buoy data [21]–[23]. Among the three points, the sea area at point C is the broadest, the bias, the correlation coefficient and the root mean square error all reach the ideal results. Compared with point B and C, point A has the narrowest sea area, which is slightly worse than point B and C in Bias, CC and RMSE. This result indicates that WW3 wave model is more suitable for the study of wide sea areas, correspond with previous research [24].

III. ANALYSIS OF WAVE CHARACTERISTICS IN THE SOUTH CHINA SEA AND ADJACENT WATERS

Before discussing the prediction of SWH in the SCS and adjacent waters, the characteristics of SWH in this region need to be fully understood, including spatial differences and temporal differences. These analyses can improve our understanding of the variation of SWH and provide a foundation for improving the long-term prediction.

The SCS and adjacent waters locate in the edge of the Pacific Ocean, so the annual average SWH in this region is clearly lower than that in the ocean, basically within 2.0 m [22]. As shown in Figure 4, the SWH in the SCS and adjacent waters has obvious spatial differences. The large value center of SWH locates in the Luzon Strait, basically between 1.6m and 2.0 m, and the SWH in the north part of the SCS and the eastern part of Taiwan Island is 1.4-1.6 m. It is observed that the large value regions of the SCS and adjacent waters are distributed in the northeast and southwest zones. Finally, the small value regions of SWH are found in the edge of the SCS, of about 1.0 m, due to the complex geographical environment.

The annual average SWH in the SCS and adjacent waters has been analyzed in above section. We can learn that the wave energy reserves in different regions. Also, the potential of wave energy resources in different regions should be evaluated. The long-term variation of the SWH is analyzed using linear regression method shown in Figure 5. Similarly, we analyze the long-term trend of SWH in the SCS and adjacent waters from 1990 to 2017 in Figure 6. It is obvious that the SWH in most areas of the SCS and adjacent waters exhibits a significant increasing trend, of about 0.2-1.2 cm/yr. The larger growth area is mainly distributed in the waters around the Taiwan Island and the northern part of the SCS, of about 0.5-1.2 cm/yr, especially in the waters of the Taiwan Strait, the annual average growth of the SWH reaches 1.1-1.4 cm/yr.

IV. PREDICTION OF SWH

A. TIME SERIES ANALYSIS

The Autoregressive Moving Average (ARMA) model is one of the most popular approaches for understanding and predicting the future value of a specified time series [25]. ARMA model based on two parts: autoregressive (AR) part and moving average (MA) part.

Autoregressive (AR) model specifies that the output variable x_t depends linearly on its own previous values: $x_{t-1}, x_{t-2}, \ldots x_{t-p}$ and on a stochastic term ε_t, where p determines the number of steps into the past needed to predict the current value. The notation AR(p) indicates an autoregressive model of order p. The AR(p) model is defined as:

$$x_t = a_0 + a_1 x_{t-1} + a_2 x_{t-2} + \ldots + a_p x_p + \varepsilon_t \qquad (4)$$

where $a_0, a_1, \ldots, a_p$ are the autoregressive coefficients of the model, and they are constants. ε_t is a Gaussian white noise.

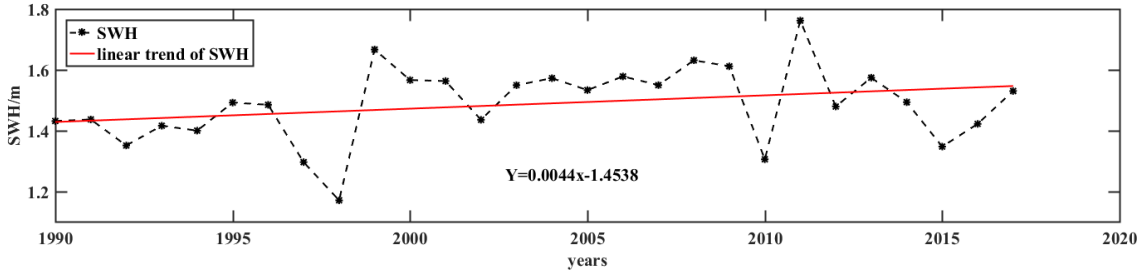

FIGURE 5. The long-term trend of SWH.

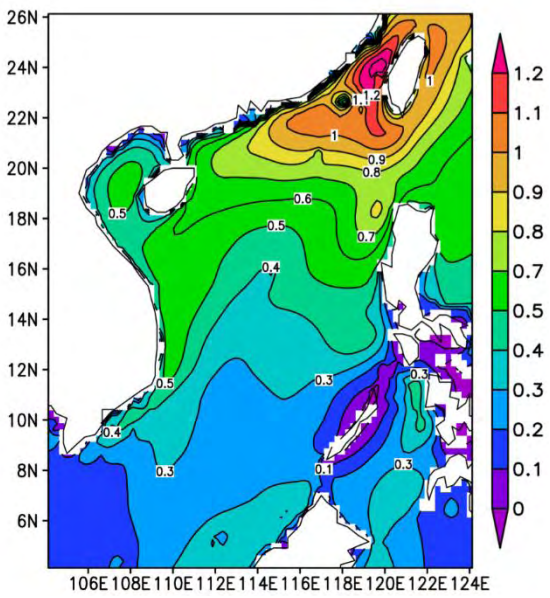

FIGURE 6. Long term trend of SWH from 1990 to 2017.

Different from the AR model, which believes that there is an autoregressive relationship between the series, and the residual series is not correlated at different moments, and it requires that the residuals are not correlated with the values of previous moments, moving average (MA) model based on the idea that the current value of the series, x_t can be explained as a function of q past external disturbances: $\varepsilon_t, \varepsilon_{t-1}, \ldots \varepsilon_q$, An moving average model of order q, abbreviated MA(q), is of the form:

$$x_t = \mu + \varepsilon_t + b_1\varepsilon_{t-1} + b_2\varepsilon_{t-2} + \ldots + b_q\varepsilon_q \tag{5}$$

where μ is the mean of the series, ε_t is the Gaussian white noise, also is the disturbance of current time. $b_1, b_2, \ldots, b_q$ are the parameters of the model and $\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots \varepsilon_{t-p}$ are white noise error terms.

For a time series $\{x_t : t \in Z\}$, if a p-order autoregressive model is modeled, and there is still some correlation between residual sequences, then we need to mix the autoregressive model and moving average model to get the ARMA model. In an ARIMA model, the future value of a variable is supposed to be a linear combination of past values and past errors [26], expressed as follows:

$$x_t = c + a_1x_{t-1} + a_2x_{t-2} + \ldots + a_px_p + \varepsilon_t + b_1\varepsilon_{t-1} + b_2\varepsilon_{t-2} + \ldots + b_q\varepsilon_q \tag{6}$$

where c is a constant, a_i (i $= 1, 2 \ldots$ p) and b_j (j $= 1, 2 \ldots$ q) are constants representing the auto-regressive AR, and the moving average MA parameters of order p, q, respectively. x_t is the actual value and ε_t expresses the Gaussian white noise with mean zero in time t.

AR model, MA model and ARMA model are generally used in stationary time series. Non-seasonal ARIMA model and seasonal ARIMA (SARIMA) are applied in some cases where data show evidence of non-sationarity, where an initial differencing step (as shown in equation (7-8)) can be applied one or more times to eliminate the non-stationarity. The former is an important method to study non-stationary time series with trend only, and the latter is an important method to study non-stationary time series with trend and periodicity. Non-seasonal ARIMA model and seasonal ARIMA are generally denoted ARIMA(p, d,q) and SARIMA$(p, d, q) \times (P, D, Q)_m$, respectively. Where parameters p, d, q are non-negative integers, p is the order of the AR model, d is the degree of differencing, and q is the order of the MA model. P, D, Q refer to the autoregressive, differencing, and moving average terms for the seasonal part of the ARIMA model, and m is the period of the time series.

$$x_t^{'} = x_t - x_{t-1} \tag{7}$$

$$y_t = x_t - x_{t-m} \tag{8}$$

Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) [27] are the criterions to measure the goodness of statistical model fitting and the effective methods to determine the orders of p and q. These values can be given qualitatively as follows:

$$AIC\,(k,l) = ln\left(\hat{\sigma}^2\,(k,l)\right) + \frac{2\,(k+l+1)}{N} \tag{9}$$

$$BIC\,(k,l) = ln\left(\hat{\sigma}^2\,(k,l)\right) + \frac{ln\,(N)\,(k+l+1)}{N} \tag{10}$$

where, the parameter N is the length of the observed series, $\hat{\sigma}^2\,(k,l)$ is estimated value of the error variance when $p = k, q = l$. In this paper, the upper bound p_0 and q_0 of (p, q) are determined to be 3 and $H = \{(k,l) \mid 0 \le k \le p_0, 0 \le l \le q_0\}$. For any $(k,l) \in H$, calculate AIC function or BIC function according to the equation (5-6) to minimize $AIC\,(k,l)$ or $BIC\,(k,l)$, in this case, (k,l) is the order of the model.

B. PREDICTION METHOD: SARIMA MODEL

A brief flowchart of the SARIMA algorithm is shown in Figure 7. Take the data in point C as an example, the process of SARIMA modeling is described in detail in the following steps.

Step 1: Acquisition of SWH data. The time series of monthly average SWH from January 1990 to December 1999 is shown in Figure 8(a). It can be seen that the time series has a period with a length of 12, that is, the monthly average SWH is related to the season. In one period, the SWH always rises from high to low, that is, the SWH reaches the maximum in winter, followed by spring and autumn, and the smallest in summer. Therefore, it may be non-stationary and seasonal differencing method was used to make it stationary.

Step 2: Making 12-step differencing to the SWH series. The specific formula as equation (8), where m $= 12$. The result is shown in Figure 8(b).

Step 3: Testing the stationarity of the series after 12-step differencing. The seasonal differencing of SWH time series jump around zero and it has no obvious trend and periodicity,

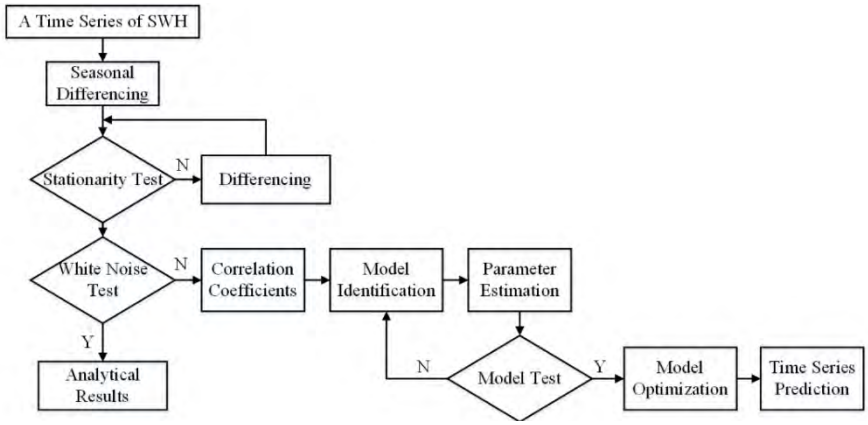

FIGURE 7. Program flowchart of the SARIMA algorithm.

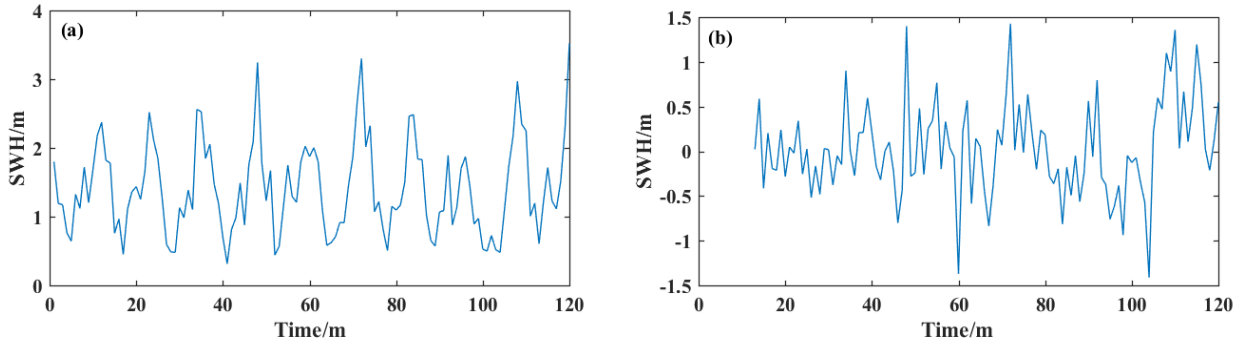

FIGURE 8. Time series diagram of SWH (a) and seasonal differencing (b).

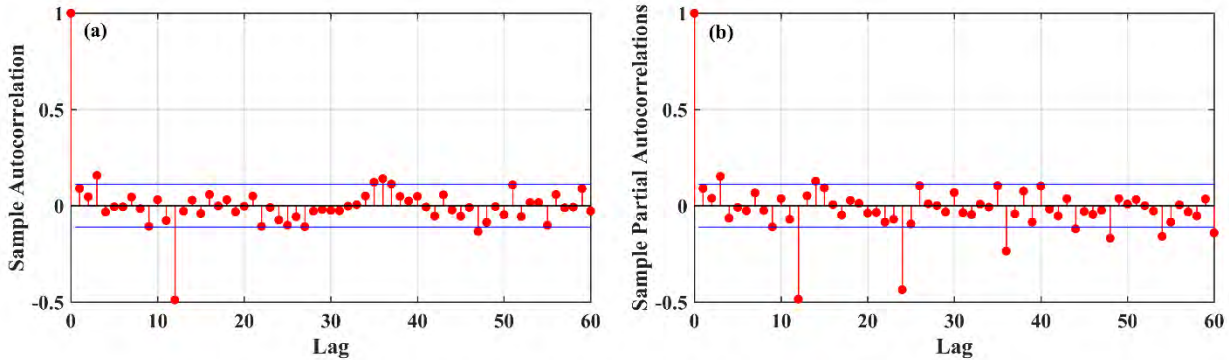

FIGURE 9. Sample autocorrelation (a) and Sample partial autocorrelation (b) of seasonal differencing time series.

which can be seen from Figure 8(b). And the differencing series passed the Augmented Dickey-Fuller (ADF) test, which means that there is not unit root in the series. Therefore the sequence can be considered as stable.

Step 4: White noise test. It can be seen from the sample autocorrelation graph Figure 9(a) and sample partial autocorrelation graph Figure 9(b), the autocorrelation coefficient does not fall within the range of 2 times the standard deviation. So the data is no-white noise.

Step 5: Model selection. Both the autocorrelation coefficient and partial autocorrelation coefficient are trailing, SARIMA $(p, 0, q) \times (0, 1, 0)_{12}$ model is established.

Step 6: Estimating the orders p, q of the model SARIMA $(p, 0, q) \times (0, 1, 0)_{12}$. As described in Section IV-A, the value of AIC can be able to compute for different models by varying p and q each SWH series from January 1990 to December 2016. Here, the upper bound p_0 and q_0 are both set 3.

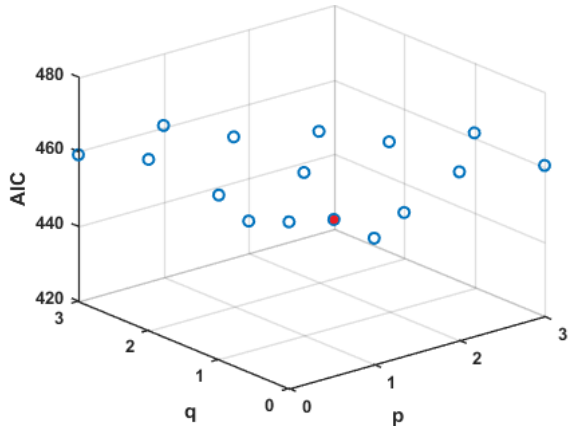

FIGURE 10. AIC values of different orders.

The AIC values of different orders are shown in Figure 10. We can see that the minimum value of AIC has been found by red circle, with the orders (3, 3). Therefore, the SARIMA $(3, 0, 3) \times (0, 1, 0)_{12}$ model is established.

Step 7: The selection of prediction step. Forecasting performances of the SARIMA $(3, 0, 3) \times (0, 1, 0)_{12}$ model have been evaluated on the basis of one-step ahead 1, 2, 3, 4, 6, 12 out of sample forecasts. Actual and predicted values of 12 months with different steps as shown in Figure 11. The initial values of prediction of different steps are the same, slightly than the actual value. As the predicted time increases, their values gradually separate. However, their overall trends are consistent, all similar to the trend of actual values.

Still, Bias, CC and RMSE are used to measure the goodness of prediction with different steps. The formulas have been shown in section II, equation (1-3). The Bias, CC, and RMSE of these models are as shown from Table 2. From Bias, we can see that the predicted values is smaller than actual values. Except for 6-step prediction, the Bias is all around -0.1. There is a close relationship between predicted values and actual values, among these models, 12-step and 1-step prediction are best, 6-step, 4-step and 3-step prediction are inferior, 2-step prediction is worst. It can been seen from RMSE that the 6-step, 4-step predicted results are worse than the other 4 models. Through the above analysis, we can find the 12-step prediction has the best performance in prediction of SWH.

Taken together, we choose SARIMA $(3, 0, 3) \times (0, 1, 0)_{12}$ model and 12-step prediction as the final prediction model to predict SWH in the South China Sea and adjacent waters.

V. PREDICTED RESULTS AND DISCUSSION

The monthly average SWH of simulated data and predicted data in the SCS and adjacent waters in 2017 is presented in the form of contour maps in Figure 12. From the figure, we can see that the SWH in the SCS and adjacent waters reaches its peak in winter (January, February and December), and most of the areas are above 1.2 m. The lowest SWH appears in spring (March, April and May) and summer (June, July and August), basically between 0.3 m and 1.2 m.

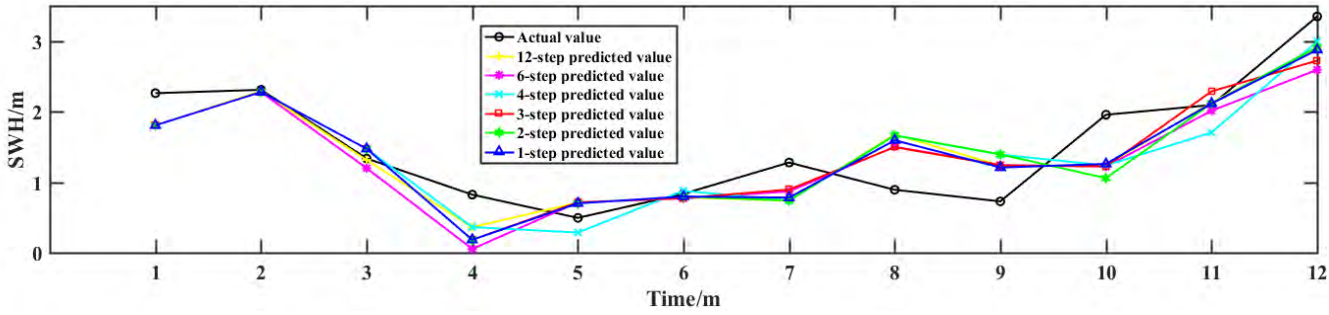

FIGURE 11. Actual and predicted values with different steps.

TABLE 2. Statistical analysis of prediction models of different steps.

Calculation method	12-step	6-step	4-step	3-step	2-step	1-step
Bias	-0.0979	-0.1745	-0.1289	-0.1053	-0.1028	-0.1072
CC	0.8521	0.8378	0.8424	0.8481	0.8099	0.8538
RMSE	0.3392	0.6046	0.4466	0.3648	0.3562	0.3714

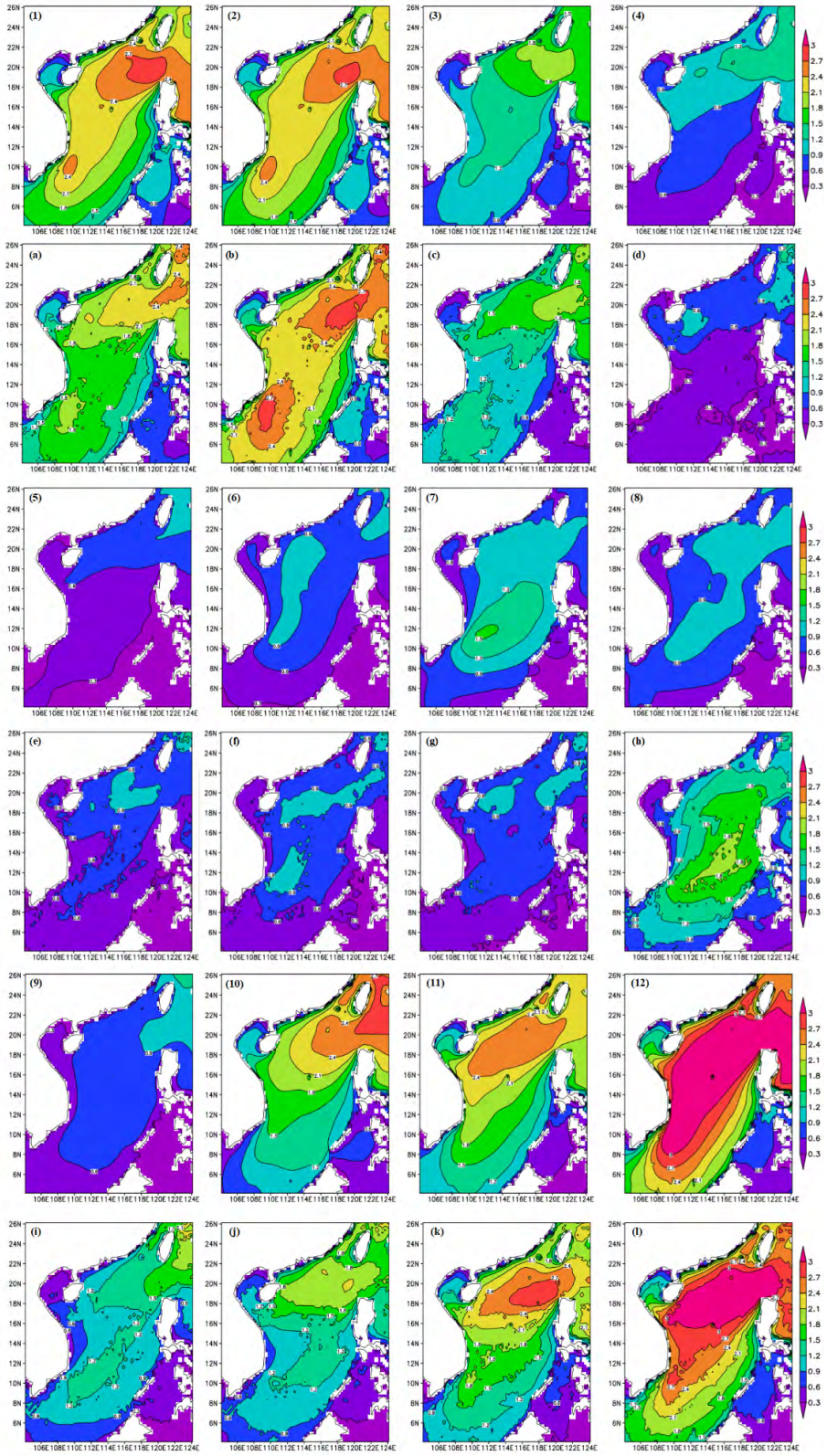

FIGURE 12. Monthly average SWH of simulated data (1-12) and predicted data (a-l) in 2017.

In autumn (September, October and November), the SWH of the first month, September, is low, only a small part of the area receives 0.9 m. The SWH in October and November is slightly lower than in winter, but it is much higher than other times. While in the view of time line, the trend of simulated values is consistent with that of the simulated values.

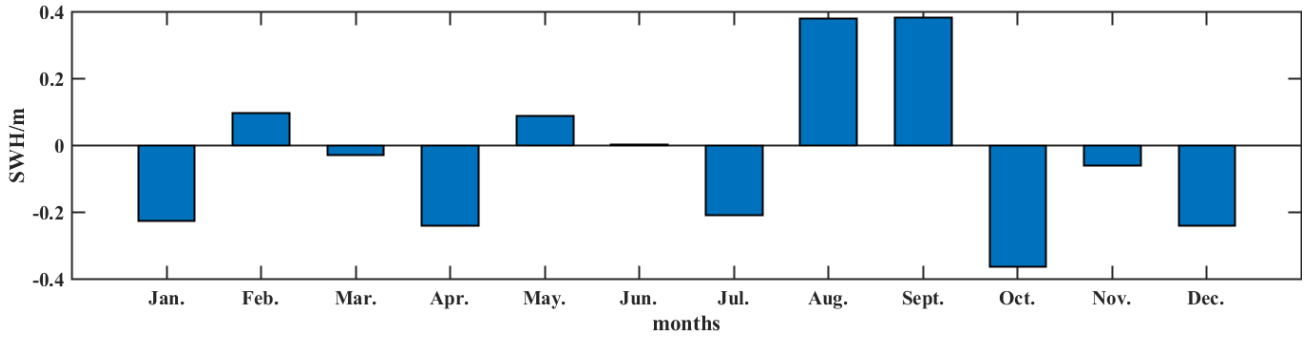

FIGURE 13. The overall Bias between predicted and simulated data in 2017.

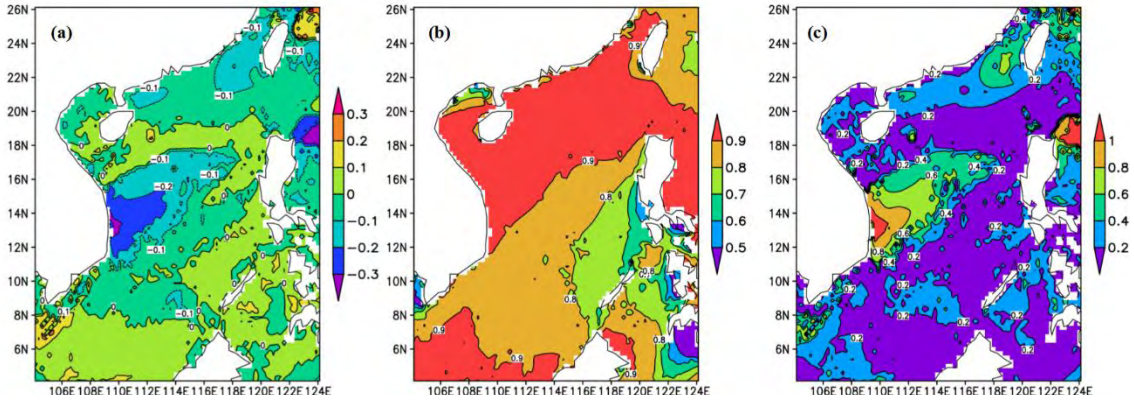

FIGURE 14. The Bias (a), CC (b), and RMSE (c) between predicted data and simulated data in 2017.

From the perspective of physical space, the spatial distribution of predicted and simulated SWH is similar, and is consistent with the analysis in the section III. The large values are mainly concentrated in the waters near the Taiwan Island, the Luzon Strait and the northern part of the SCS, showing a zonal distribution of northeast to southwest. In the waters around the Taiwan Island, predicted values are slightly lower than simulated values, expect for July and September. In the Luzon Strait, there is fluctuation between predicted SWH and simulated SWH. In January, April and October, the predicted values are slightly lower, in February, June, August, September and November they are higher, and more accurate results are obtained in other months. In the south-central part of the SCS, the predicted results are complex, but the bias between predicted values and simulated values are within 0.3 m.

In order to further verify the accuracy of the prediction, we analyzed the Bias, CC and RMSE between predicted data and observed data using equation (1-3). Where, $\bar{y}$ represents the average value of predicted data, $\bar{x}$ represents the average value of simulated data, y_i and x_i respectively represent predicted SWH and simulated SWH, and $N = 12$.

Do average of Bias of all grid points in every month, the overall Bias is shown in Figure 13. The Bias is bigger than other months in August, September and October. Similarly, the Bias, CC and RMSE at every $0.25° \times 0.25°$ grid point in the SCS and adjacent waters are obtained, as shown in Figure 14. From Figure 14(a), it is clearly that the predicted SWH are slightly bigger than simulated values in the northern and southern part of the SCS, of about 0 to 0.1 m. However, in other regions, the differences between predicted and simulated values show the opposite phenomenon, the Bias is about -0.2 to 0 m. The Bias of our prediction results within -0.1 to 0.1 m accounts for 77.32% of study region. The overall Bias is relatively small. As can be seen from Figure 14(b), the correlation between simulated values and predicted values is relatively high, and the correlation coefficient basically above 0.8, of about 79.24 % of study area. Only in the southeast of the verification area, the results of CC are unsatisfactory, which may be due to complex geography. Figure 14(c) presents the RMSE between predicted and simulated values. The RMSE of the SCS and adjacent waters is basically below 0.4, of about 82.18% of study area. Therefore, on the whole, SARIMA method has achieved good results in predicting SWH of the SCS and adjacent waters.

VI. CONCLUSIONS

In this paper, the wave field in the SCS and adjacent waters was simulated using the third-generation wave model WW3, with high resolution and continuous time. And the comparison between WW3 simulated data and satellite altimeter data proves that WW3 numerical model is capable of simulating

the wave field in the SCS and adjacent waters. Based on WW3 simulated data, we analyzed the wave characteristics of the SCS and adjacent waters from space and time. The northern part of the SCS, Luzon strait and Taiwan Strait are advantageous areas for the exploitation and utilization of wave energy resources. The SWH is 1.4-2 m, with an average annual growth of 0.5-1.4 cm/yr.

Due to the complicated and stochastic behavior of ocean waves, long-term prediction of SWH is full of challenge. By analyzing the time series of SWH in the SCS and adjacent waters, it is found that the SWH in this area is non-stationary with a period of 12. A novel approach based on SARIMA model is proposed for the SWH prediction. This approach takes into account the relationship between current time and the values, residuals at some previous time and the periodicity of the SWH series. The statistical analysis shows that SARIMA model has a good performance for the long-term prediction of SWH in the SCS and adjacent waters.

ACKNOWLEDGEMENTS

The authors would like to thank the NASA Earth Science Enterprise for providing CCMP wind data and the NOAA/NCEP, the WAM for the support on the WW3 model. They would also like to thank anonymous reviewers for their careful work and thoughtful suggestions that will make this paper more improvement.

REFERENCES

[1] P. A. J. Bonar, I. G. Bryden, and A. G. L. Borthwick, "Social and ecological impacts of marine energy development," *Renew. Sustain. Energy Rev.*, vol. 47, pp. 486–495, Jul. 2015.

[2] C. P. Collazo, D. Greaves, and G. Iglesias, "A review of combined wave and offshore wind energy," *Renew. Sustain. Energy Rev.*, vol. 42, pp. 141–153, Feb. 2015.

[3] D. Silva, P. Martinho, and C. G. Soares, "Wave energy distribution along the Portuguese continental coast based on a thirty three years hindcast," *Renew. Energy*, vol. 127, pp. 1064–1075, Nov. 2018.

[4] C. W. Zheng, Q. Wang, and C. Y. Li, "An overview of medium-to long-term predictions of global wave energy resources," *Renew. Sustain. Energy Rev.*, vol. 79, pp. 1492–1502, Nov. 2017.

[5] C. W. Zheng and C. Y. Li, "Analysis of temporal and spatial characteristics of waves in the Indian Ocean based on ERA-40 wave reanalysis," *Appl. Ocean Res.*, vol. 63, pp. 217–228, Feb. 2017.

[6] T. G. Barbounis, J. B. Theocharis, M. C. Alexiadis, and P. S. Dokopoulos, "Long-term wind speed and power forecasting using local recurrent neural network models," *IEEE Trans. Energy Convers.*, vol. 21, no. 1, pp. 273–284, Mar. 2006.

[7] T. G. Barbounis and J. B. Theocharis, "Locally recurrent neural networks for long-term wind speed and power prediction," *Neurocomputing*, vol. 69, nos. 4–6, pp. 466–496, 2006.

[8] H. B. Azad, S. Mekhilef, and V. G. Ganapathy, "Long-term wind speed forecasting and general pattern recognition using neural networks," *IEEE Trans. Sustain. Energy*, vol. 5, no. 2, pp. 546–553, Apr. 2014.

[9] J. D. Agrawal and M. C. Deo, "On-line wave prediction," *Mar. Struct.*, vol. 15, no. 1, pp. 57–74, 2002.

[10] A. R. Kambekar and M. C. Deo, "Wave prediction using genetic programming and model trees," *J. Coastal Res.*, vol. 28, no. 1, pp. 43–50, Jan. 2012.

[11] 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.

[12] J. Mahjoobi and A. Etemad-Shahidi, "An alternative approach for the prediction of significant wave heights based on classification and regression trees," *Appl. Ocean Res.*, vol. 30, no. 3, pp. 172–177, Jul. 2008.

[13] 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.

[14] B. Kamranzad, A. Etemad-Shahidi, and M. H. Kazeminezhad, "Wave height forecasting in Dayyer, the persian gulf," *Ocean Eng.*, vol. 38, no. 1, pp. 248–255, Jan. 2011.

[15] M. Özger, "Prediction of ocean wave energy from meteorological variables by fuzzy logic modeling," *Expert Syst. Appl.*, vol. 38, no. 5, pp. 6269–6274, May 2011.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] G. Reikard, "Integrating wave energy into the power grid: Simulation and forecasting," *Ocean Eng.*, vol. 73, no. 4, pp. 168–178, 2013.

[21] P. C. Chu, Y. Q. Qi, Y. C. Chen, P. Shi, and Q. W. Mao, "South China sea wind-wave characteristics. Part I: Validation of WAVEWATCH-III using TOPEX/Poseidon data," *J. Atmos. Ocean. Technol.*, vol. 21, no. 11, pp. 1718–1733, Nov. 2004.

[22] C. W. Zheng and C. Y. Li, "Variation of the wave energy and significant wave height in the China Sea and adjacent waters," *Renew. Sustain. Energy Rev.*, vol. 43, pp. 381–387, Mar. 2015.

[23] C.-W. Zheng, J. Pan, and J. X. Li, "Assessing the China Sea wind energy and wave energy resources from 1988 to 2009," *Ocean Eng.*, vol. 65, no. 2, pp. 39–48, Jun. 2013.

[24] M. M. Amrutha, V. S. Kumar, K. G. Sandhya, T. M. B. Nair, and J. L. Rathod, "Wave hindcast studies using SWAN nested in WAVEWATCH III—Comparison with measured nearshore buoy data off Karwar, eastern Arabian Sea," *Ocean Eng.*, vol. 119, pp. 114–124, Jun. 2016.

[25] G. E. P. Box and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 3rd ed. Englewood Cliffs, NJ, USA: Prentice-Hall, 1994.

[26] P.-F. Pai and C.-S. Lin, "A hybrid ARIMA and support vector machines model in stock price forecasting," *Omega*, vol. 33, pp. 497–505, Dec. 2005.

[27] K. P. Burnham and D. R. Anderson, "Multimodel inference: Understanding AIC and BIC in model selection," *Sociol. Methods Res.*, vol. 33, no. 2, pp. 261–304, 2016.

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 interests include ocean observation and exploration technology such as atmosphere duct measuring technology, deep Argo float, and wave assessment.

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

LINLIN FAN received the bachelor's degree in automation from the North China University of Science and Technology, in 2016. She is currently pursuing the M.S. degree in instrument and meter engineering from the School of Precision Instrument and Opto-Electronics Engineering, Tianjin University. Her current research interests include marine data processing and the assessment of wave energy.

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 interests include the assessment and the prediction of the height of ocean waves.

SHANHUA DUAN received the B.S. degree in geographic information sciences from Jilin Jianzhu University, in 2017. She is currently pursuing the master's degree in instrument and meter engineering from the School of Precision Instrument and Opto-Electronics Engineering, Tianjin University. Her current research interests include marine data processing and remote sensing technology.

CHONGWEI ZHENG received the Ph.D. degree in atmospheric from the National University of Defense Technology. He is a Visiting Research Fellow of the College of Marine Science, Sun Yat-Sen University, where he is also a Senior Visiting Scholar with 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 multi-sensor 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 interests include ocean observation and exploration technology such as atmosphere duct measuring technology, deep Argo float, and wave assessment.