

Article

An EMD–PSO–LSSVM Hybrid Model for Significant Wave Height Prediction

Gang Tang ¹, Jingyu Zhang ¹, Jinman Lei ¹, Haohao Du ¹, Hongxia Luo ¹, Yide Wang ² and Yuehua Ding ^{3,*}

¹ Logistics Engineering College, Shanghai Maritime University, Shanghai 201306, China; gangtang@shmtu.edu.cn (G.T.); hxliao@shmtu.edu.cn (H.L.)

² Institut d'Electronique et des Technologies du NumeRique(IETR), CNRS UMR6164, Nantes Universite, F-44000 Nantes, France; yide.wang@univ-nantes.fr

³ School of Electronic and Information Engineering, South China University of Technology, Guangzhou 510640, China

* Correspondence: eeyhding@scut.edu.cn

Abstract: The accurate prediction of significant wave height (SWH) offers major safety improvements for coastal and ocean engineering applications. However, the significant wave height phenomenon is nonlinear and nonstationary, which makes any prediction work a non-straightforward task. The aim of the research presented in this paper is to improve the predicted significant wave height via a hybrid algorithm. Firstly, an empirical mode decomposition (EMD) is used to preprocess nonlinear data, which are decomposed into several elementary signals. Then, a least squares support vector machine (LSSVM) with nonlinear learning ability is adopted to predict the SWH, and a particle swarm optimization (PSO) automatically performs the parameter selection of the LSSVM modeling. The results show that the EMD–PSO–LSSVM model can compensate for the lag in the prediction timing of the prediction models. Furthermore, the prediction performance of the hybrid model has been greatly improved in the deep-sea area; the prediction accuracy of the coefficient of determination (R^2) increases from 0.991, 0.982, and 0.959 to 0.993, 0.987, and 0.965, respectively. The prediction performance results show that the proposed EMD–PSO–LSSVM performs better than the EMD–LSSVM and LSSVM models. Therefore, the EMD–PSO–LSSVM model provides a valuable solution for the prediction of SWH.

Keywords: wave prediction; significant wave height; empirical mode decomposition; particle swarm optimization; least squares support vector machine



Citation: Tang, G.; Zhang, J.; Lei, J.; Du, H.; Luo, H.; Wang, Y.; Ding, Y. An EMD–PSO–LSSVM Hybrid Model for Significant Wave Height Prediction. *J. Mar. Sci. Eng.* **2023**, *11*, 866. <https://doi.org/10.3390/jmse11040866>

Academic Editors: Hua Li and Francisco Haces-Fernandez

Received: 25 March 2023

Revised: 15 April 2023

Accepted: 18 April 2023

Published: 20 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Massive maritime operations increase the requirement for improved wave forecasting techniques. Understanding accurate wave conditions allows for more efficient and safer maritime activities and coastal management, for instance, the installation of marine structures, ports and docks, marine transportation and navigation, and shoreline protection, especially to prevent coastal erosion and more [1,2]. Furthermore, wave research can further develop ocean wave resources, and the use of wave energy converters can convert wave energy into electrical energy [3]. Wave prediction data can help to provide motion compensation, which may prevent the crash of cargo in cargo transfer, improve the firing accuracy of ship-borne weapon systems, and performance of the motion control systems [4]. Significant wave height (SWH) is an effective feature of ocean waves. However, most researchers have only focused on predicting the height of the wave and not on predicting the pair of values of the significant height of the wave and its period, which will make the maritime activities and coastal management more efficient and safer. Therefore, the purpose of this paper is to achieve the efficient prediction of ocean waves by predicting SWH.

Many factors affect wave formation, including air pressure, temperature, wind speed, wind direction, and so on. Over the past few years, several numerical methods have been developed to predict the SWH, such as the Sverdrup, Munk and Bretschneider (SMB), and

Pierson–Neumann–James (PNJ) manual-based methods, and the numerical calculation model is based on differential equations [5]. These methods calculate wave height from the wind information based on wind–wave relationships. However, they need to compute elaborate meteorological and oceanographic data sets and thus involve an enormous amount of computational effort. In addition, due to the uncertainty of the wind–wave relationship, there may be some uncertainty when converting wind energy into wave energy, and the predicted results have not always been very accurate [6]. The complicated coastal geomorphology, coastal erosion, and structure conditions make wave forecasting very difficult.

Later, artificial intelligence methods based on linear and nonlinear models, or hybrid models were applied to predict the SWH [7–10]. The observations based on historical data are often used as the input to predictive models in artificial intelligence. The wind speed is an important parameter for wave prediction. Many scholars had used the wind speed as the input to the prediction model to predict the SWH [11]. Deo pointed out that providing wave information on certain locations should be carried out based on the sea and/or meteorological measurements in that location, or as near to it as possible [12]. Mahjoobi used the wind speed as the model input to predict the SWH, which further illustrated that the hysteresis of wind speed can increase the error of the predicted SWH [13]. Compared with numerical analysis, those methods had achieved good prediction results, and the prediction accuracy had been improved. However, uncertain factors such as wind speed, direction, and wind propagation distance can affect the generation and development of waves; the correlation coefficient of the results between the observed and the predicted data is not always satisfactory. In recent years, many scholars have used the data of past waves as the input of the model to predict the waves at a future moment [14]. Jain only used the effective wave height as the input of the model when studying local wind, which has a good prediction effect [15]. Through the above analysis, it can be found that the historical wind and waves contain different information, so the prediction effect is different when used as the input of the model. Using a wind model as an input works well for predicting waves at specific locations. However, due to the hysteresis of wind speed and other unpredictable factors, when we study a local wave using the observed historical wave data—which contain implicitly all the practical factors, such as the air pressure, temperature, local geomorphology, wind speed, and wind direction—as the input of the prediction model, it can have a better prediction effect.

However, accurate SWH prediction requires a large amount of sensor-based data and high-performance computations, so wave height predictions are often not always very accurate [16–18]. With the development of machine learning, time series analysis provides computationally alternative solutions mainly based on historical wave height data [19,20]. Such modeling approaches have the advantage of being based on previous data and wave patterns, thus avoiding heavy computational resources. Early combinations of wave prediction based on machine learning apply classical time series models, such as the auto-regressive (AR) model, auto-regressive moving average (ARMA) model, and an autoregressive integrated moving average (ARIMA) model [21–23]. Soares applied AR models to describe the SWH time series in two Portuguese coast locations. Later, the AR models were further generalized from the application of univariate models of long-term SWH time series to SWH bivariate series and mean periods [24,25]. However, predictions based on a single AR model in harsh conditions have poor performance. To further improve the prediction performance, Agrawal applied ARMA and ARIMA models to predict the wave heights for 3, 6, 12, and 24 h of offshore location in India, respectively [26,27]. Despite the high efficiency and adaptiveness of classical time series models, the prediction results in severe sea conditions are far from being accurate enough. Since waves are always nonstationary, the linear and stationary classical time series models' assumptions are not rigorously valid. Consequently, these approaches are not suitable for predicting nonlinear and nonstationary waves.

To address the nonlinear component of ocean waves, intelligent-technique-based nonlinear models such as artificial neural networks (ANNs) models have been extensively studied. Such methods can carry out nonlinear simulations without a deep understanding of the relationships between the input and output variables. Deo and Sridhar Naidu were amongst the first to apply an ANN to predict, in real-time, the wave in the next 3–24 h using past wave data [28]. To estimate the large wave height and average wave periods, Deo used wind velocity and fetched data [12]. Tsai applied an ANN-based on data from three wave graph stations in areas with different physical characteristics for short-term estimation of the wave height [29]. Makarynsky used ANN for substantial wave height prediction and for the subsequent 1–24 h forecast times [30]. Mandal and Prabaharan used a recurrent neural network (RNN) for wave height prediction in Marmugao, west coast of India [31]. They showed that the wave prediction using RNN provides better results than the other neural network-based methods. One of the limitations of the neural network approach is that it needs to find network parameters such as the number of hidden layers and neurons by trial and error, which is time-consuming. Mahjoobi and Adeli Mosabbe applied support vector machine (SVM) to predict the wave height; the analysis showed that the SVM model had a reasonable precision compared to ANN-based methods, which took less computation time [32]. Different experiments on the prediction performance in Lake Superior were carried out by Etemad-Shahidi and Mahjoobi, and they compared the model trees and feedforward backpropagation ANNs [13]. Their findings revealed that the model tree system was the most precise. Dixit found the phenomenon of prediction time lag while using ANNs to predict the ocean wave height. They used a discrete wavelet transform to enhance the predicted values and to remove the lag in the prediction timing [33]. Akbarifard and Radmanesh introduced a symbiotic organisms search (SOS) to predict the ocean wave heights. Their findings showed that the SOS algorithm performed better than that of the support vector regression (SVR), ANN [34]. Fan proposed a long short-term network for the quick prediction of the SWH with higher accuracy than the convolutional neural network (CNN) [35].

SWH is impacted by various components in a nonlinear, dynamic way [36]. The time series prediction of non-stationary data by ANN methods can lead to the homogenization of the different characteristics of the original input data, which can affect the prediction accuracy. Accordingly, the non-stationarity of the time series of the SWH and input variables should be reduced. To handle the nonstationary features, the inputs of the corresponding data-driven models should be appropriately preprocessed. Hybrid models combining preprocessing techniques with single prediction models are possible alternatives. The wavelet analysis can be used for nonstationary data [37]. Deka and Prahlada developed a wavelet neural network model by hybridizing ANN with a wavelet transform, and the prediction results suggested that the hybrid models outperformed single models [38]. Kaloop designed the wavelet-PSO-ELM (WPSO-ELM) model for estimating the wave height belonging to coastal and deep-sea stations. The results showed that the WPSO-ELM outperforms other models for wave height prediction in both hourly and daily leading times [39]. Essentially, a linear and nonstationary solution is based on wavelet transform. It represents a signal through a linear combination of functions of the wavelet base. For nonlinear data, it may not be suitable [40]. The other issue with wavelets is that they require a well-suited mother wavelet transform a priori [41]. This is still an unresolved issue and generally requires a lengthy trial and error process [42]. In hybrid prediction models, a more effective decomposition technique is needed to overcome the nonlinearity and non-stationarity instantaneously.

When considering nonlinear and nonstationary data sets, a data-driven methodology known as empirical mode decomposition (EMD) is efficient and adaptive [43]. The EMD multiresolution utility offers self-adaptability by avoiding the need for any basis function and mother wavelets. It functions as a dyadic filter that divides a large frequency band complex signal into relatively essential, time-scale components [44]. Duan proposed EMD-SVR for the short-term prediction of ocean waves. The result showed the EMD-SVR

model shows good model performance and provides an effective method for the short-term prediction of nonlinear and nonstationary waves [44].

Based on the above analysis, we introduce an EMD and particle swarm optimization (PSO) and least-squares SVM (LSSVM) based model whose objective is to improve the SWH prediction performance. The LSSVM with nonlinear learning ability can be used for signal prediction, while EMD provides an empirical analysis tool for processing nonlinear and nonstationary data sets. Preprocessing with EMD can reduce the prediction complexity; PSO is a swarm intelligence optimization algorithm, and by updating the distance between the current and best locations, the important parameters of LSSVM are optimally adjusted by PSO to improve the prediction accuracy of a single LSSVM.

The remainder of this paper is organized as follows. The proposed EMD–PSO–LSSVM-based prediction model is described in Section 2. The wave data and prediction measures are presented in Section 3. The performance of the proposed method is assessed in Section 4. Finally, the conclusion is highlighted in Section 5.

2. Methodology Formulation

2.1. EMD–PSO–LSSVM Prediction Model

Ocean wave time series is a type of complicated nonlinear and nonstationary signal composed of various oscillation scales. When performing wave predictions, the different oscillation scales widely impact the quality of the LSSVM model. Combining an EMD model with the LSSVM model is likely to enhance the wave height prediction. The EMD is adopted to decompose the wave height series into one residual series and several intrinsic mode functions (IMFs). Then, the residual series and IMFs are modeled by the LSSVM model. Finally, wave height prediction can be achieved by summing the prediction outputs of the subseries. Moreover, a PSO algorithm is employed to optimize the important parameters in the LSSVM to increase the prediction accuracy. The specific steps of the EMD–PSO–LSSVM prediction algorithm are displayed in Figure 1. The next part is to present the hybrid technique separately.

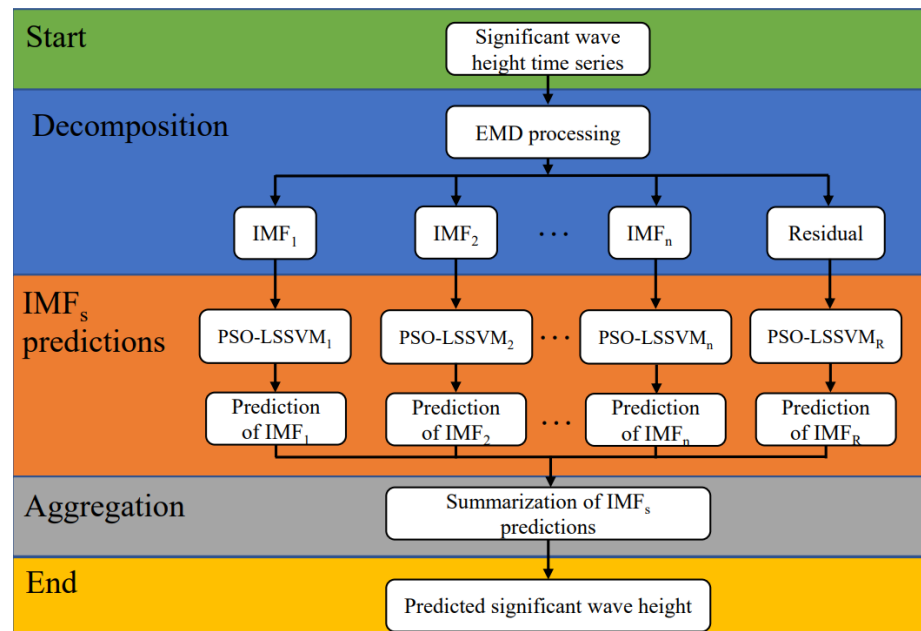


Figure 1. Algorithm flow of the proposed hybrid EMD–PSO–LSSVM prediction model. The flow chart includes three important steps: EMD data preprocessing, PSO–LSSVM prediction, and, finally, SWH predictions.

2.2. Preprocess Data by EMD

Empirical mode decomposition (EMD) is an empirical analysis tool used for processing nonlinear and nonstationary data sets. The main idea of an EMD is to decompose the nonlinear and nonstationary time series into a sum of several simple intrinsic mode function (IMF) components and one residue with individual inherent time scale properties. Each IMF represents a natural oscillatory mode and has to satisfy the following two conditions:

- (a) The number of extremes and the number of zero-crossings should be equal, or differ by one.
- (b) The local average should be null, i.e., the means of the upper envelope defined by the local maxima and the lower envelope defined by the local minima are null.

With a given SWH time sequence $x(t)$, the EMD processing steps are summarized as follows:

- (1) Identify the local extrema.
- (2) Generate the upper envelope $u(t)$ and the lower envelope $l(t)$ via a spline interpolation among all the local maxima and the local minima, respectively. Then, the mean envelope is obtained as follows: $m(t) = [l(t) + u(t)]/2$.
- (3) Subtract $m(t)$ from the signal $x(t)$ to obtain the IMF candidate, i.e., $h(t) = x(t) - m(t)$.
- (4) Verify whether $h(t)$ satisfies the conditions for IMFs, and carry out step (1) to step (4) until $h(t)$ is an IMF.
- (5) Obtain the n th IMF component $imf_n = h(t)$ (after n shifting processes) and the corresponding residue $r(t) = x(t) - h(t)$.
- (6) Repeat the whole algorithm with $r(t)$ obtained in step (5) until the residue is a monotonic function.
- (7) By implementing these algorithms, the decomposition procedure of a signal is expressed as

$$x(t) = \sum_{i=1}^n imf_i(t) + r(t). \tag{1}$$

2.3. Least Squares Support Vector Machine (LSSVM)

SVM is a statistical learning theory-based method with a strong capacity to handle nonlinear problems. Its basic idea is to map nonlinear data into a high dimensional feature space using a nonlinear mapping function, where linear techniques are available. LSSVM is the least-squares formulation of a standard SVM. Unlike the inequality constraints introduced in the standard SVM, LSSVM proposes equality constraints in the formulation. This changes the solution being transformed from one of solving a quadratic program to a set of linear equations known as the linear Karush–Kuhn–Tucker (KKT) systems. LSSVM is a nonlinear prediction model based on SVM theory, widely applied in short-term prediction problems. LSSVM has been retained thanks to its good generalization ability. It has been shown that the performance of an LSSVM model in the prediction problem is better than other nonlinear models. The basic idea of the method can be described as follows.

A training data set of N stations is given $\{(x_i, y_i), i = 1, 2, \dots, N\}$ with input data $x_i \in R^N$ and output data $y_i \in R$. A nonlinear mapping function is defined to map the input data into the high dimensional feature space. In the high dimensional feature space, there theoretically exists a linear function to express the nonlinear relationship between input and output data. Such a linear function, namely the LSSVM function, can be defined as

$$y(x_i) = \omega^T \phi(x_i) + b, \tag{2}$$

where ω and b are adjustable coefficients. The corresponding optimization problem for LSSVM is formulated as

$$\begin{cases} \text{Min } J(\omega, e_i) = \frac{1}{2} \|\omega\|^2 + \frac{1}{2} C \sum_{i=1}^N e_i^2, \\ y(x_i) = \omega^T \phi(x_i) + b + e_i \end{cases}, \tag{3}$$

where C denotes the regularization constant and e_i represents the training data error. The Lagrangian is represented by

$$L(\omega, a_i, b, e_i) = J + \sum_{i=1}^N a_i [y_i - \omega^T \phi(x_i) - b - e_i]. \tag{4}$$

From the Karush–Kuhn–Tucker (KKT) conditions, the following equations must be satisfied:

$$\frac{\partial L}{\partial \omega} = 0; \frac{\partial L}{\partial a_i} = 0; \frac{\partial L}{\partial b} = 0; \frac{\partial L}{\partial e_i} = 0. \tag{5}$$

The solution is found by solving the system of linear equations expressed in the following matrix form:

$$\begin{bmatrix} 0 & 1_v^T \\ 1 & \Psi + C^{-1}I \end{bmatrix} \begin{bmatrix} b \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ y \end{bmatrix}, \tag{6}$$

where $y = [y_1, \dots, y_N]^T$, $1_v = [1, \dots, N]^T$, $a = [a_1, \dots, a_N]^T$; I denotes the identity matrix; $\Psi_{ij} = K(x_i, x_j)$, $i, j = 1, \dots, N$, which satisfies Mercer’s condition.

The LSSVM regression model becomes

$$f(x_i, x_j) = \sum_{i=1}^n a_i K(x_i, x_j) + b, \tag{7}$$

where a_i denotes the Lagrange multipliers that can be obtained by solving the dual problem; $K(x_i, x_j)$ denotes the kernel function which equals the inner product of $\phi(x_i)$ and $\phi(x_j)$.

The most frequently used kernel functions are the polynomial kernel function, sigmoid kernel function, and radial basis function (RBF) kernel. Considering that the RBF kernel is not only easy to implement but also is an efficient tool for dealing with nonlinear problems, the RBF function is adopted in this paper. The RBF function is defined by

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right). \tag{8}$$

The efficiency of the LSSVM generalization (prediction accuracy) depends on the good collection of meta parameters, C , σ , and parameters of the kernel. When the RBF function is selected, the parameters (C and σ) must be optimized using the PSO–LSSVM system. The regularization parameter C and kernel parameter σ of LSSVM have a significant influence on the classification accuracy. The choices of C and σ govern the model’s complexity of the prediction.

2.4. LSSVM Optimization by PSO

To avoid the under-fitting and over-fitting issues, the LSSVM model’s hyper-parameters should be appropriately tuned. The PSO algorithm is used to find the best values of C and σ in LSSVM. The LSSVM fitting process optimized by the PSO is shown in Figure 2.

PSO uses the velocity–position search model. The iteration formula adjusting the position and speed of a particle is as follows:

$$V_i^{t+1} = wV_i^t + c_1r_1(p_{best} - X_i^t) + c_2r_2(g_{best} - X_i^t), \tag{9}$$

$$X_i^{t+1} = X_i^t + V_i^{t+1}, \tag{10}$$

where w denotes the inertial weight; c_1 and c_2 denote the cognition and social learning factors, respectively; r_1 and r_2 are two random numbers; t denotes the t th iteration; X_i^t denotes the position of particle i in d -dimensional space, which denotes the current value of LSSVM parameters C and σ ; V_i^t denotes the velocity of particle i in d -dimensional space, which decides to update the direction and distance of the next generation of C and σ ;

p_{best} denotes the best position that every particle can be obtained during the execution of the PSO method; g_{best} denotes the best situation that particles have obtained during the implementation of the PSO method.

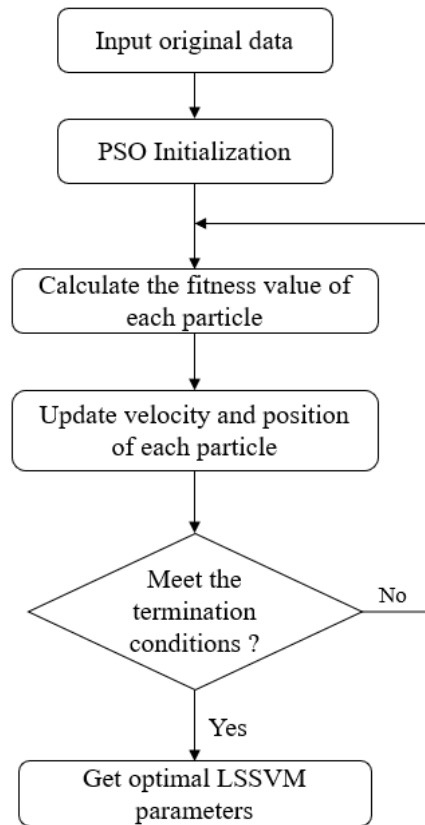


Figure 2. Flowchart of PSO-based parameter selection algorithm.

Some parameter descriptions and parameter settings of the PSO algorithm are listed. The iteration t is set to 50; the default values of c_1 and c_2 are set to 1. These default values can ensure that particles are more affected locally or globally; r_1 and r_2 are two random numbers in the range $[0, 1]$. w is the weight factor. The inertial weight determines the influence of the previous velocity on the current one. A considerable inertia weight facilitates global exploration, while a small one tends to facilitate local exploration. A suitable value of the inertia weight usually provides a balance between the global and regional exploration abilities. We use a linearly decreasing inertia weight, which starts at 0.9 and ends at 0.4. The performance of PSO can be significantly improved. The inertial weight can be expressed as follows:

$$w = w_{max} - \frac{w_{max} - w_{min}}{t_{max}} \times t, \tag{11}$$

where w_{max} and w_{min} denote the initial and terminal weights, respectively; t_{max} denotes the maximum iteration counter. New fitness values of the particles are calculated after the velocity and position updating if required. p_{best} and g_{best} are also updated, and the same procedure is performed continuously until the stop criteria are satisfied. Usually, the velocity of each particle is restricted to a maximum value within the interval $[-0.01, 100]$, which is defined according to the bounds on decision variables.

3. Descriptions of the Wave Data and Prediction Accuracy Measures

3.1. Raw Data

Two North Atlantic Ocean areas have been selected to predict the SWH. The SWH and meteorological series were downloaded from National Data Buoy Center (NDBC) (<https://www.ndbc.noaa.gov> (accessed on 26 October 2020)). To study the prediction performance of SWH in different water areas, the data from the offshore area and deep-sea area, respectively, are studied. Two stations are utilized in this study (see Figure 3), where Station A denotes Station 41025 at 35°1'30" N 75°21'47" W in the offshore, while Station B denotes Station 41048 at 31°49'53" N 69°34'23" W in the deep-sea area. These stations are selected as they have an unimpaired and long series of recorded SWH and meteorological data. The SWH data set used in this study is gathered from the actual marine environment, which contains implicitly all the practical factors, such as the air pressure, temperature, local geomorphology, wind speed, wind direction, etc.

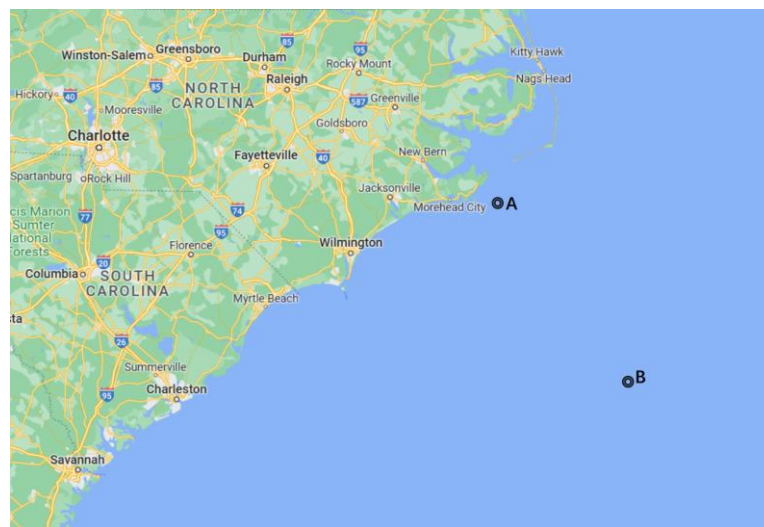


Figure 3. Location map of the stations. Station A denotes the offshore area, while Station B denotes the deep-sea area. The map has been downloaded from © Google Maps, and we have added two marks for our studied locations.

There are three parts of the used SWH data, corresponding to the date of the years 2014, 2015, and 2016, with 1500 h of sample points for each year. Data from the years 2014 and 2015 are used as the training data, and data from the year 2016 are used as the testing data. Figure 4 shows the SWH records of both stations. Table 1 shows the minimum, maximum, and average values of different training parameters and testing data sets.

Table 1. Specific information on the two stations.

Station ID	Water Depth (m)	Year	SWH Average (m)	SWH Range (m)
A	59.4	2014	1.2287	[0.37,3.53]
		2015	1.1938	[0.45,2.81]
		2016	1.3141	[0.50,3.27]
B	5309	2014	2.0715	[0.63,8.01]
		2015	1.9668	[0.66,5.04]
		2016	2.2987	[0.67,6.34]

From Table 1 and Figure 4, it can be concluded that the average SWH at Station A, located near the coast, is around 1.2 m, with the maximum SWH being around 3 m. The sea state is relatively stable. The average SWH at Station B, located in the deep-sea area, is about 2 m, and the maximum SWH is about 6.5 m. The sea conditions are relatively rough. Therefore, it is difficult to predict the SWH in the deep-sea area.

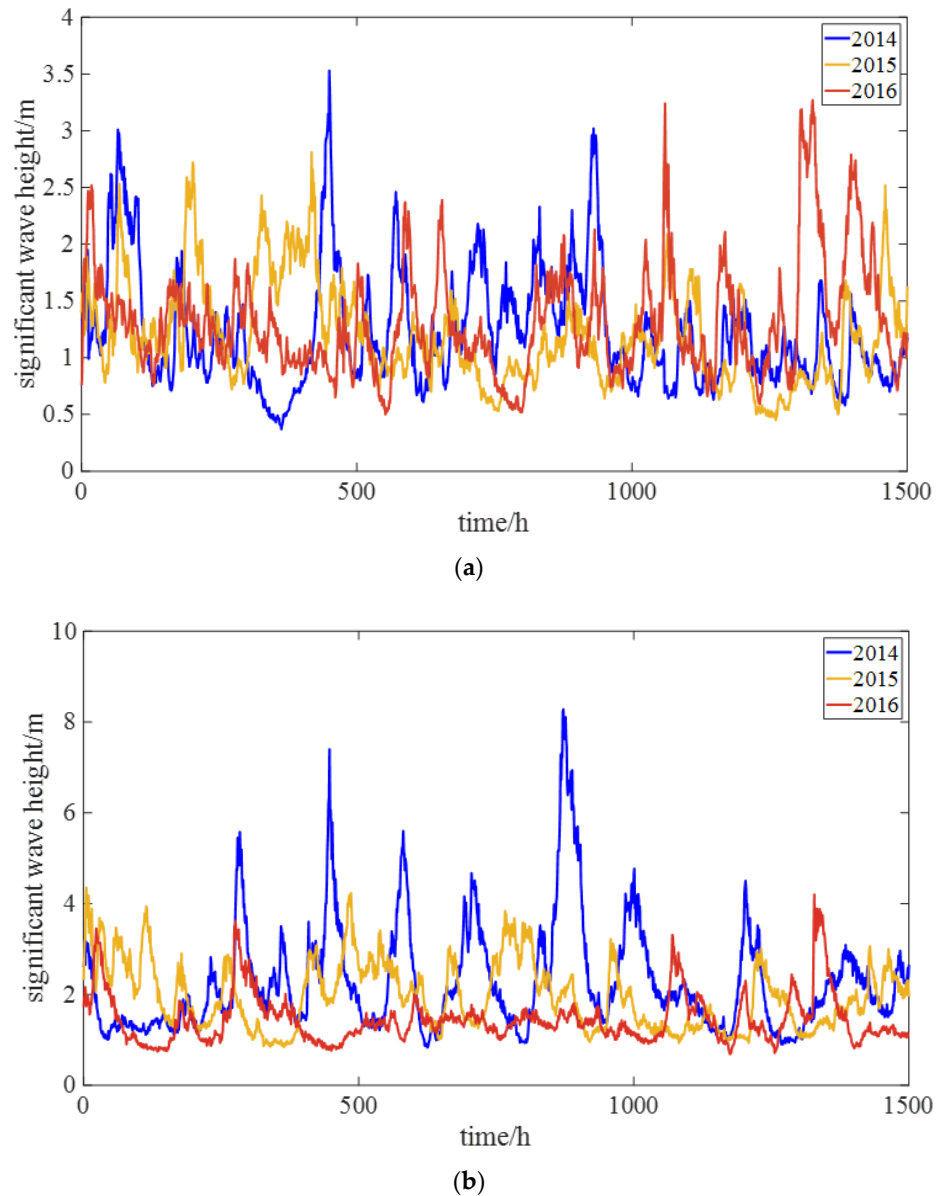


Figure 4. SWH of two stations. (a) Station A, (b) Station B.

The SWH data for the years 2014 and 2015 are used as the input variables for model learning. The relevance of each feature with the SWH needs to be determined before choosing the input features. The autocorrelation coefficient is used to study the dependence between the instantaneous values of the same signal at two time instants. The correlation coefficient $r_{x,y}$ can be calculated as

$$r_{x,y} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}}, \tag{12}$$

where $r_{x,y}$ denotes the correlation coefficient between data sets x and y ; i ($i = 1, 2, 3 \dots n$). $|r| \geq 0.8$ indicates that there is a high correlation between the two features. The correlation coefficient between the input features and the output feature is shown in Table 2. H- i in the table represents the SWH data from i th hours ago; H-2 represents the SWH data from two hours ago as an example. In the H-2 test, $r_{x,y}$ is equal to 0.9435, and $r_{x,y}$ means the correlation coefficient between the data sets from the previous two hours and the data set at the selected time. From the table, it can be seen that the correlation coefficient is lower than

0.8 at H-6, so the data from the previous five hours are used as the input in the proposed prediction model.

Table 2. Correlation coefficient of the input features with the output feature.

	H-1	H-2	H-3	H-4	H-5	H-6
$r_{x,y}$	0.9707	0.9435	0.9093	0.8710	0.8039	0.7680

3.2. Models Evaluations

To evaluate the performance of the models, some statistical and standard metrics are used. The mathematical formulations of the adopted metrics are given as follows:

- (1) Root mean square error (RMSE) is expressed as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2}, \tag{13}$$

- (2) Mean absolute error (MAE) is expressed as follows:

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - y_i|, \tag{14}$$

- (3) Mean square error (MSE) is expressed as follows:

$$MSE = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2, \tag{15}$$

- (4) Coefficient of determination (R^2) is expressed as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}, \tag{16}$$

where x and y denote the observed and the predicted values, respectively; \bar{x} denotes the mean value of the observed values; n denotes the number of observations. The lower the values of RMSE, MAE, and MSE are, the better the accuracy of the models is. The parameter R^2 ranges between 0 and 1, where 1 indicates a perfect prediction performance, and 0 shows the prediction fails totally.

4. Results and Discussion

4.1. Prediction of Single Models

We first consider a single model to predict the SWH. The models LSSVM, ELM, and ANN are used individually to predict the SWH. The prediction time is a set parameter in the system, which can be modified according to different environmental conditions. We consider an input time of 5 h and the prediction times of 1 h and 3 h as examples. In the same station, the data from the years 2014 and 2015 are used as the training set, and the data from the year 2016 are used as the testing data to predict the wave height of the year 2016 for 1 h and 3 h, respectively. The specific parameters of various model networks are shown in Table 3, where IN denotes the number of input layer units; H denotes the number of hidden layer units; O denotes the number of output layer units; σ denotes the confidence; C denotes the penalty coefficient. With the SWH for the five previous hours as the input, the SWH at the next time is the output in the prediction model. For example, the input is the SWH of the previous 5 h, and the output is the SWH at the next moment (means at the sixth hour). Next, the SWH from the second to sixth hour is taken as the model input, and the SWH at the seventh hour is predicted.

Table 3. Parameter of three single models.

Model	Initial Settings
LSSVM	IN = 5, O = 1, $\sigma = 10$, C = 100, kernel function = Radial Basis Function (RBF)
ELM	IN = 5, H = 10, O = 1, activation function = Sigmoid
ANN	IN = 5, H = 10, O = 1, training algorithm = Levenberg–Marquardt

Figures 5 and 6 show the prediction of the SWH of Stations A and B by the three compared single models, respectively. Table 4 shows the numerical analysis of specific evaluation indicators. It can be seen from Table 4 that for the wave height prediction of A near the coast, R^2 can be kept above 0.8 when the 3 h prediction is made. For Station B in the deep-sea area, R^2 can be maintained above 0.9 during the 3 h prediction, and there is a high correlation between the predicted SWH and the observed SWH. In general, the three algorithms have achieved satisfactory results in predicting the SWH, but LSSVM has higher prediction accuracy than the other two models. This clearly shows that among the compared models, the proposed LSSVM model can provide the best prediction performance.

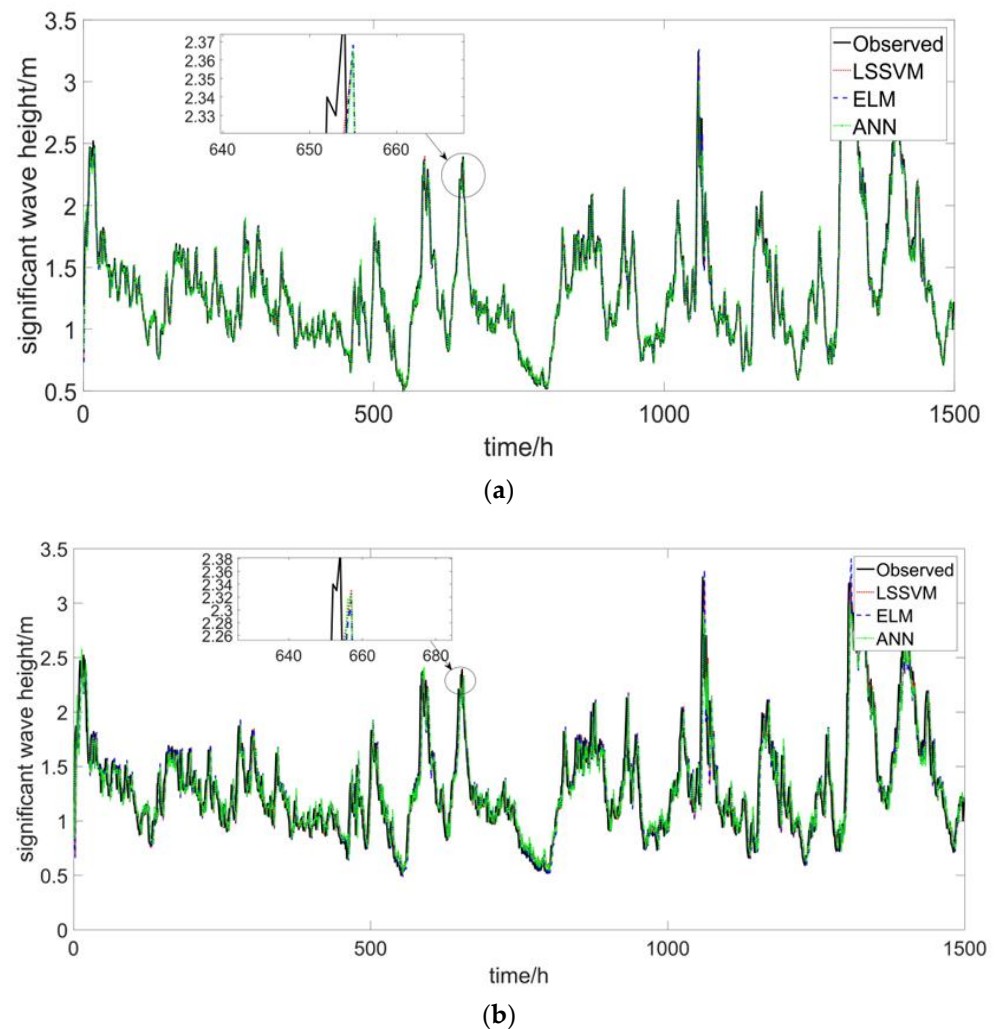


Figure 5. Comparison between the observed and predicted SWH at A. (a) one hour prediction, (b) three hour prediction. The wave prediction is performed on different predicted models during 1500 h. Comparing (a) and (b), it can be seen that as the leading time increases, the prediction timing lag problem becomes greater.

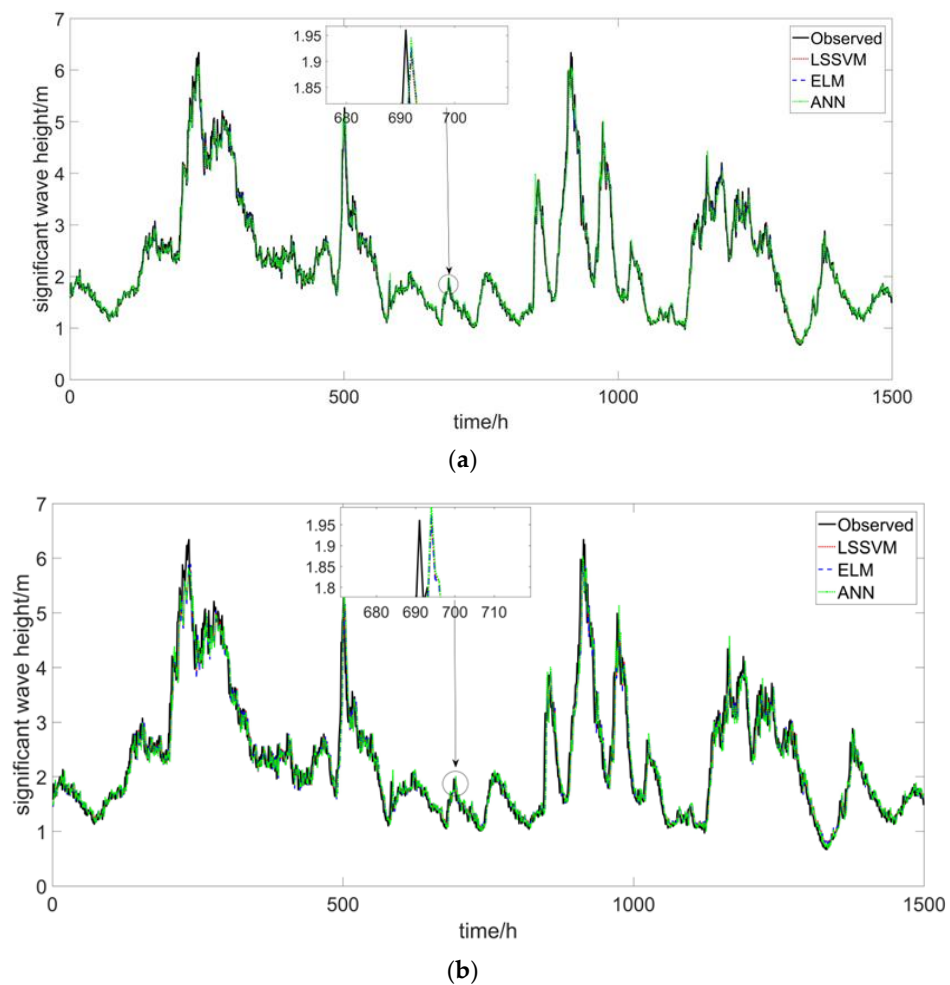


Figure 6. Comparison between the observed and predicted SWH at Station B. (a) one hour prediction, (b) three hour prediction. The wave prediction is performed during 1500 h on different predicted models. Comparing (a) and (b), it can be seen that as the leading time increases, the prediction timing lag problem becomes greater.

Table 4. Analysis of the prediction results of three single models.

Station	Model	Leading Time	RMSE	MAE	MSE	R ²
A	LSSVM	1	0.115	0.082	0.013	0.942
		3	0.200	0.141	0.040	0.826
	ELM	1	0.115	0.082	0.013	0.942
		3	0.201	0.141	0.040	0.826
	ANN	1	0.116	0.082	0.013	0.942
		3	0.201	0.141	0.041	0.825
B	LSSVM	1	0.183	0.126	0.033	0.972
		3	0.276	0.184	0.076	0.936
	ELM	1	0.184	0.127	0.034	0.972
		3	0.279	0.185	0.079	0.936
	ANN	1	0.188	0.128	0.035	0.971
		3	0.278	0.184	0.077	0.935

It can be seen from Figures 5 and 6 that the observed wave height and the predicted wave height are slightly misaligned on the time scale axis. It can be seen from the enlarged view in Figures 5 and 6 that a one-time step significantly shifts the predicted wave heights of the three single models. These wave forecasting models exhibit lag in the prediction

timing, making the univariate time series forecasting a futile attempt. As the leading time increases, these lags become larger. The lag is a type of prediction error that can also be found in other works on wave forecasting using single models. Dixit found the phenomenon of prediction time lag while using ANNs to predict the ocean wave height. They used a discrete wavelet transform to enhance the predicted values and to remove the lag in the prediction timing [33]. The lag mainly results from the nonstationary data hidden in the measured wave time series. Modeling a nonlinear and nonstationary data set by applying a single nonlinear model is very difficult because there are too many possible patterns hidden in the data. A single model may not be general enough to capture all the essential features. Even if the nonlinear ANN is used to forecast the nonlinear and nonstationary wave heights, the lags remain. A single prediction model cannot capture all the components with different scales simultaneously. Therefore, the 'lags' occur in the forecasting results. These apparent lag phenomena affect the accuracy of prediction, so the following work content aims to eliminate this lag phenomenon and improve the prediction accuracy.

4.2. Prediction Based on the Proposed Technique

The non-stationarity and non-linearity of the time series of ocean characteristics (i.e., SWH, wave periods) appear in different oscillation scales. The time series prediction of non-stationary data by using single models can only lead to the homogenization of the original input data's various characteristics, which could affect the prediction accuracy and cause the lag phenomenon. Accordingly, the non-stationarity of the time series of the SWH and input variables should be reduced. The combination of an EMD model with the LSSVM model provides an effective way to improve wave prediction. The EMD is adopted to decompose the SWH series into one residual series and several IMFs. Then, the residual series and IMFs are modeled by the LSSVM model. Finally, the summation of the prediction output of subseries SWH is realized. In addition, the PSO is employed to optimize the LSSVM parameters to increase the prediction accuracy.

In the first step, the wave height time series is decomposed into a couple of meaningful and straightforward IMFs and one residual by EMD (Figure 7).

Significant wave data sets are decomposed into IMFs and residuals when implementing the EMD-based prediction models. Figure 7 displays the decomposition results of the wave height time series measured at A; the EMD decomposition decomposes the nonlinear SWH into 7 IMFs and 1 residual signal, where it is seen that several simple components can represent the complex wave height time series. This would have enabled the single model to extract features during the modeling of the SWH effectively. The decomposed IMF components contain the local characteristic signals of different time scales of the original signal. The positive and negative values of the IMF represent the characteristic information of different scales of the original signal, which is a part of the original data. By superimposing the IMFs, a composite signal equivalent to the original signal is formed. Next, an EMD-based hybrid model can be used to predict the SWH.

From Figure 8, it can be concluded that the single model LSSVM shows a prediction timing lag (red dotted line). The other two models have overcome the lag by using the EMD technique; the prediction results for the nonlinear and nonstationary waves are improved mainly by combining the EMD technique with the single model.

It can be seen from Figure 8 that the preprocessing method of EMD decomposition has solved the lag phenomenon. However, the SWH prediction performance using the LSSVM model is still not very satisfactory. For example, there are errors in predicting the peaks and troughs of the SWH. The next step is to optimize LSSVM parameters to improve the prediction accuracy of the model.

Changing the parameter values of a prediction system can have a significant impact on its performance. Therefore, we should find the optimum parameter values for the prediction system. These optimal parameters are found typically by using a priori knowledge or

through human experiences. However, this approach can be subject to human bias. PSO has emerged as a practical tool for high-quality parameter selection in prediction systems.

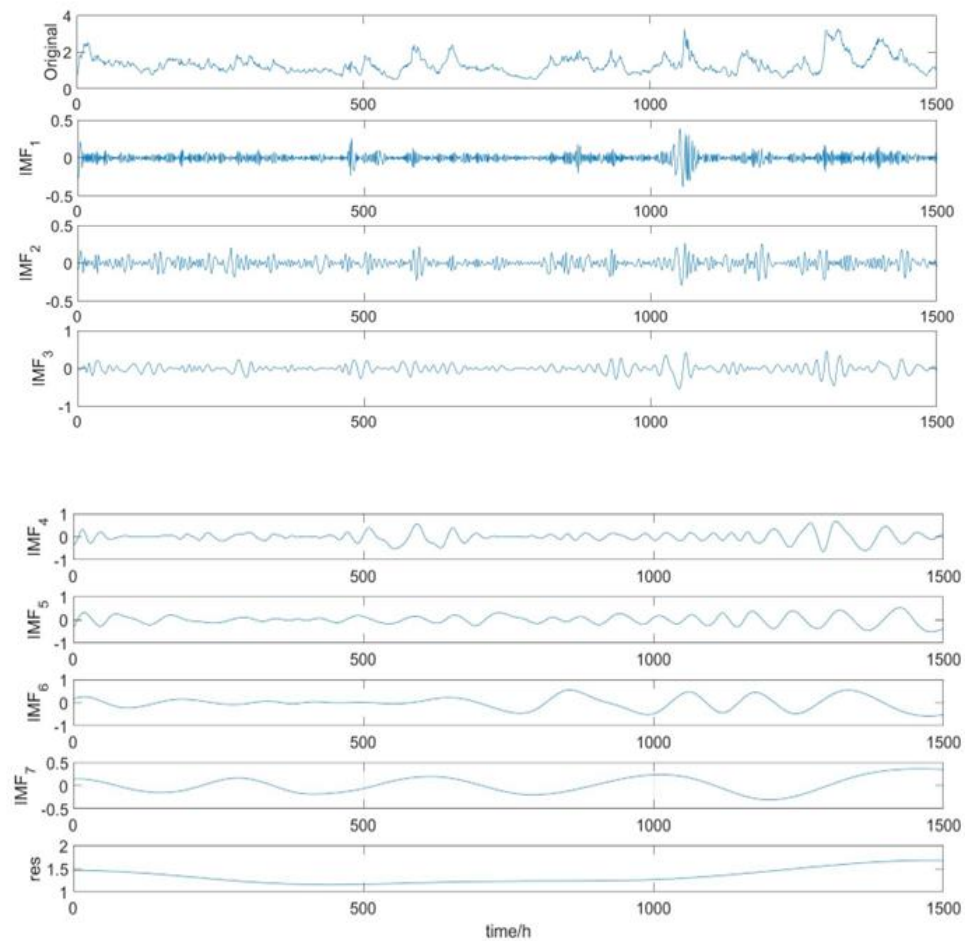


Figure 7. Decomposition results of SWH time series data using the EMD. The EMD decomposition decomposes the nonlinear SWH into 7 IMFs and 1 residual signal, facilitating the following prediction.

PSO is used to optimize the LSSVM parameters. The methodological steps can be found in the description of the method in Section 2. Figure 8 shows that the EMD–PSO–LSSVM model can predict the SWH peaks and troughs very well, significantly improving the prediction accuracy.

Table 5 presents the results obtained by the proposed EMD–PSO–LSSVM method. The prediction of the SWH is accurate with the proposed method, and the performance of using PSO to optimize the LSSVM parameters can be seen with an improvement in the prediction accuracy from $R^2 = 0.972, 0.945$ and 0.888 (EMD–LSSVM) to $R^2 = 0.972, 0.958,$ and 0.902 (EMD–PSO–LSSVM) at A, and an improvement in the prediction accuracy from $R^2 = 0.991, 0.982$ and 0.959 (EMD–LSSVM) to $R^2 = 0.993, 0.987,$ and 0.965 (EMD–PSO–LSSVM) at Station B. Correspondingly, the RMSE, MAE, and MSE predicted by EMD–PSO–LSSVM at the two stations are also the lowest.

There is a good correlation between SWH and the tenth, maximum wave height, and average wave height, so it can represent the characteristics of ocean waves. Near the predicted stations, there are seldom large winds; the predicted SHW hasn’t a big variation, which means that the stations rarely encounter extremely difficult sea conditions. When the SWH is larger, the fitted data stations are more scattered. Station B belongs to the deep-sea area; the SWH range is relatively large, and the training data are also relatively large. It can be seen from Figures 9 and 10 that when the SWH is large, the fitting performance of the data stations is still better.

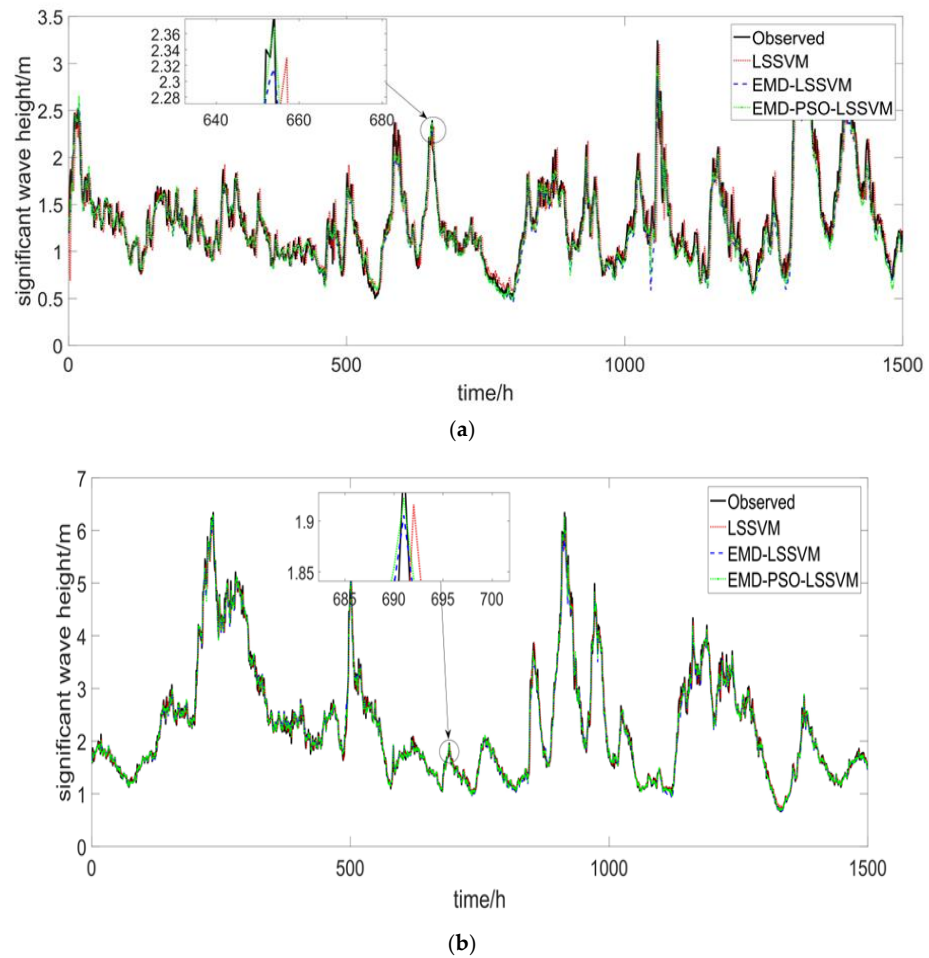


Figure 8. Comparison between LSSVM, EMD-LSSVM, and EMD-PSO-LSSVM. (a) Station A, (b) Station B. By adding the EMD method to preprocess the data, the proposed hybrid model fixes the time lag problem of a single model-based prediction. At the same time, the prediction accuracy of the peak of the SWH has also been improved.

Table 5. Performance results for Station A and Station B.

Station	Algorithm	Leading Time	RMSE	MAE	MSE	R ²
A	EMD-PSO-LSSVM	1	0.089	0.062	0.008	0.972
		3	0.105	0.079	0.011	0.958
		6	0.155	0.112	0.024	0.902
	EMD-LSSVM	1	0.097	0.071	0.009	0.972
		3	0.125	0.092	0.016	0.945
		6	0.169	0.123	0.029	0.888
	LSSVM	1	0.115	0.082	0.013	0.942
		3	0.200	0.141	0.040	0.826
		6	0.287	0.202	0.082	0.645
B	EMD-PSO-LSSVM	1	0.089	0.063	0.008	0.993
		3	0.127	0.091	0.016	0.987
		6	0.205	0.140	0.042	0.965
	EMD-LSSVM	1	0.105	0.074	0.011	0.991
		3	0.150	0.104	0.022	0.982
		6	0.224	0.154	0.050	0.959
	LSSVM	1	0.183	0.126	0.034	0.972
		3	0.278	0.184	0.076	0.936
		6	0.416	0.277	0.173	0.858

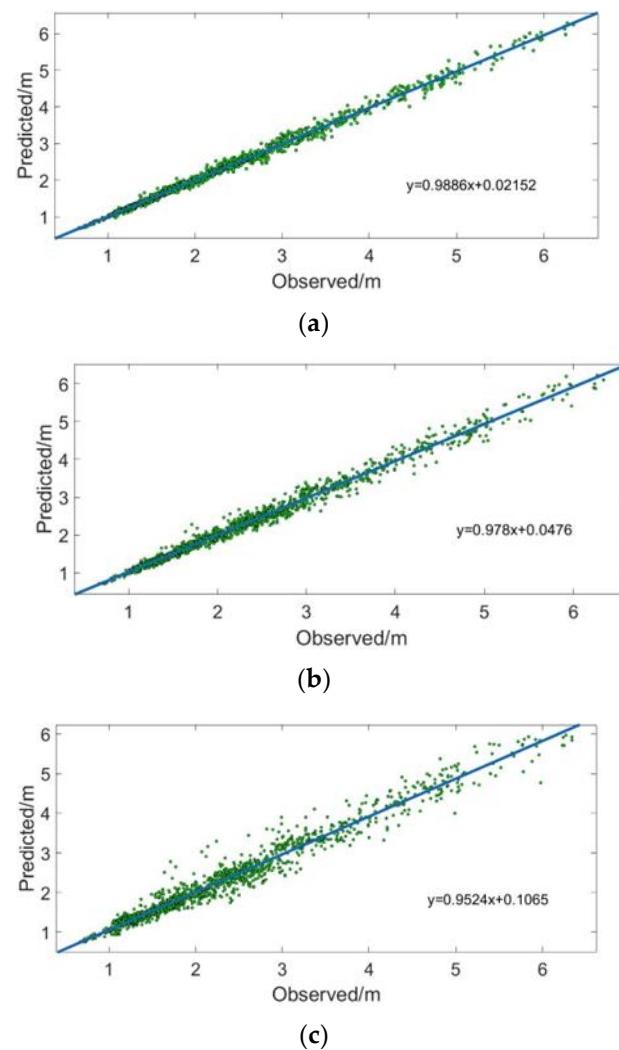


Figure 9. Scatter diagram of the observed and the predicted using the EMD–PSO–LSSVM at Station B: (a) one hour, (b) three hours, (c) six hours. The fitting slopes of EMD–PSO–LSSVM to Station B at 1, 3, and 6 h are 0.9886, 0.978, and 0.9524. Compared with A in the offshore, the combined model is more suitable for the deep-sea area, and the fitted slopes are all above 0.95.

Figures 9 and 11 show a comparison between the observed and predicted values by EMD–PSO–LSSVM at stations A and B, respectively. Figures 10 and 12 show a comparison between the observed and predicted values by EMD–LSSVM at Station A and B, respectively. Referring to the scatter plots, by looking at the distribution of the scatter plot and the slope of the fitted line, the relationship between the predicted and observed values can be seen. The denser the scatter plot distribution and the closer the slope of the fitted line is to 1, the better the prediction result. Through the above comparison, it appears that the estimation of the EMD–LSSVM time series is more scattered and farther than the EMD–PSO–LSSVM model. As the leading time increases, the EMD–LSSVM performance decreases drastically, but EMD–PSO–LSSVM performance decreases gradually, as shown in Figures 9–12. For example, the best-fit line slopes for the scatter of wave predictions of six leading hours at stations A and B are 0.8973 and 0.9524, respectively. Comparatively, the EMD–PSO–LSSVM model performs better than the EMD–LSSVM model. The coefficient of determination for the wave prediction at all stations using the EMD–PSO–LSSVM model is more than 0.902 (see Table 5), while the best-fit line slopes for the scatters are better than 0.8973 (see Figures 9 and 11). Correspondingly, the RMSE, MAE, and MSE predicted by EMD–PSO–LSSVM at the two stations are also the lowest.

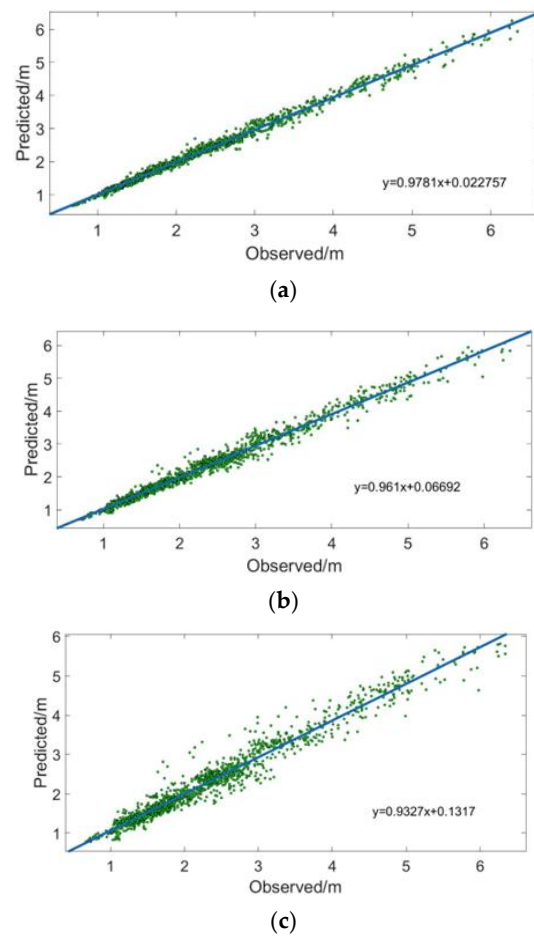


Figure 10. Scatter diagram of the observed and predicted using the EMD–LSSVM at Station B: (a) one hour, (b) three hours, (c) six hours. The fitting slopes of EMD–LSSVM to Station B at 1, 3, and 6 h are 0.9781, 0.961, and 0.9327, respectively. As with the results for Station A, the hybrid model using PSO has higher accuracy.

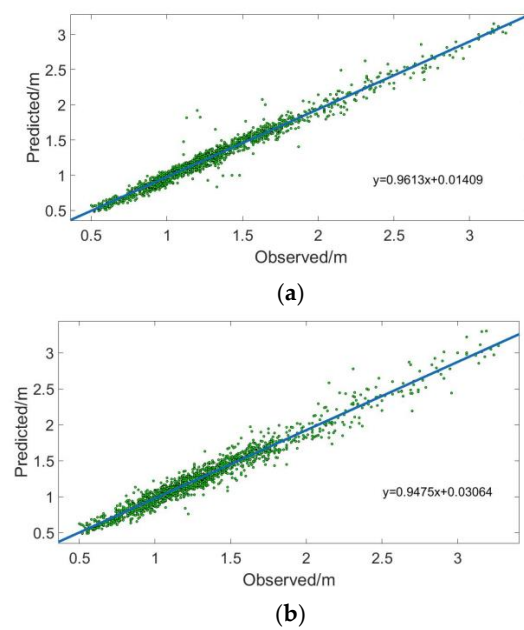
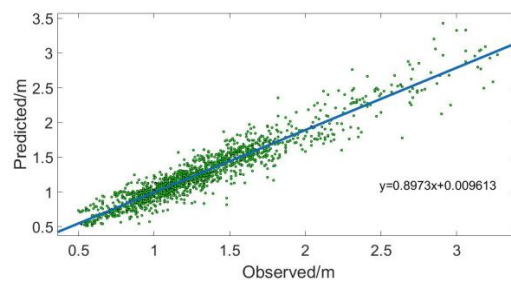
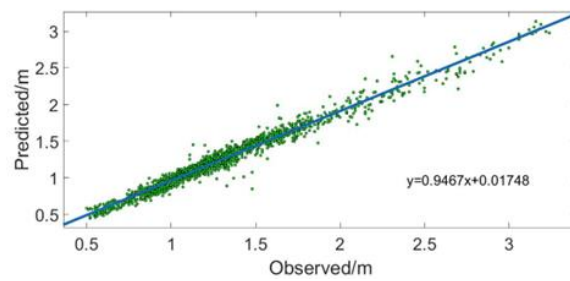


Figure 11. Cont.

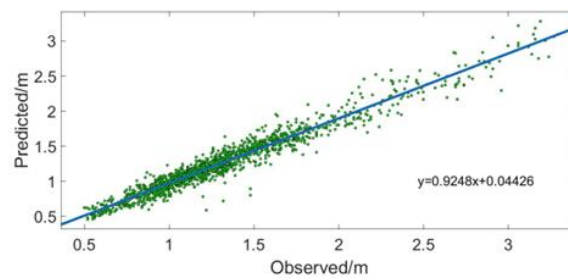


(c)

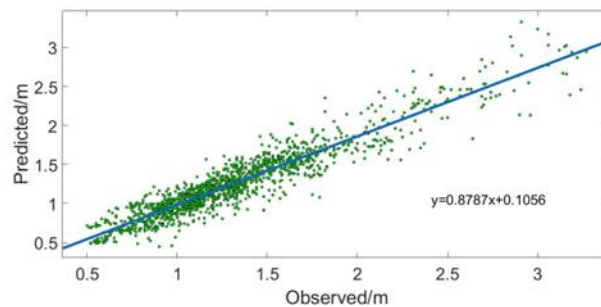
Figure 11. Scatter diagram of the observed and the predicted using the EMD-PSO-LSSVM at Station A: (a) one hour, (b) three hours, (c) six hours. The predicted value and the observed value are linearly fitted. The closer the slope of the fitted line is to 1, the better the prediction performance. The fitting slopes of EMD-PSO-LSSVM to Station A at 1, 3, and 6 h are 0.9613, 0.9475, and 0.8973, respectively.



(a)



(b)



(c)

Figure 12. Scatter diagram of the observed and the predicted using the EMD-LSSVM at Station A: (a) one hour, (b) three hours, (c) six hours. The fitting slopes of EMD-LSSVM to Station A at 1, 3, and 6 h are 0.9467, 0.9248, and 0.8787. Compared with the hybrid model using PSO for parameter optimization, the accuracy of the model is not as good as the improved model, so the effectiveness of the PSO method is proved.

4.3. Prediction Considering Wind Speed as Input

The evolution of waves depends very much on the surface winds. The wind speed is an important factor creating sea-level fluctuation. To further study the performance of wind speed on the prediction of SWH, the wind speed is added to the prediction model as an input parameter and its performance is analyzed. By comparing the prediction performance of prediction models with wind speed as an input parameter, without the wind speed parameter, the effect of wind speed on the forecasting model is discussed.

By taking Station A and Station B as examples, whose locations are shown in Figure 3, WS denotes wind speed. Before using the data, the original data are preprocessed. The WS records of both stations are shown in Figure 13. Table 6 shows the minimum, maximum, and average values of different training parameters and testing data sets.

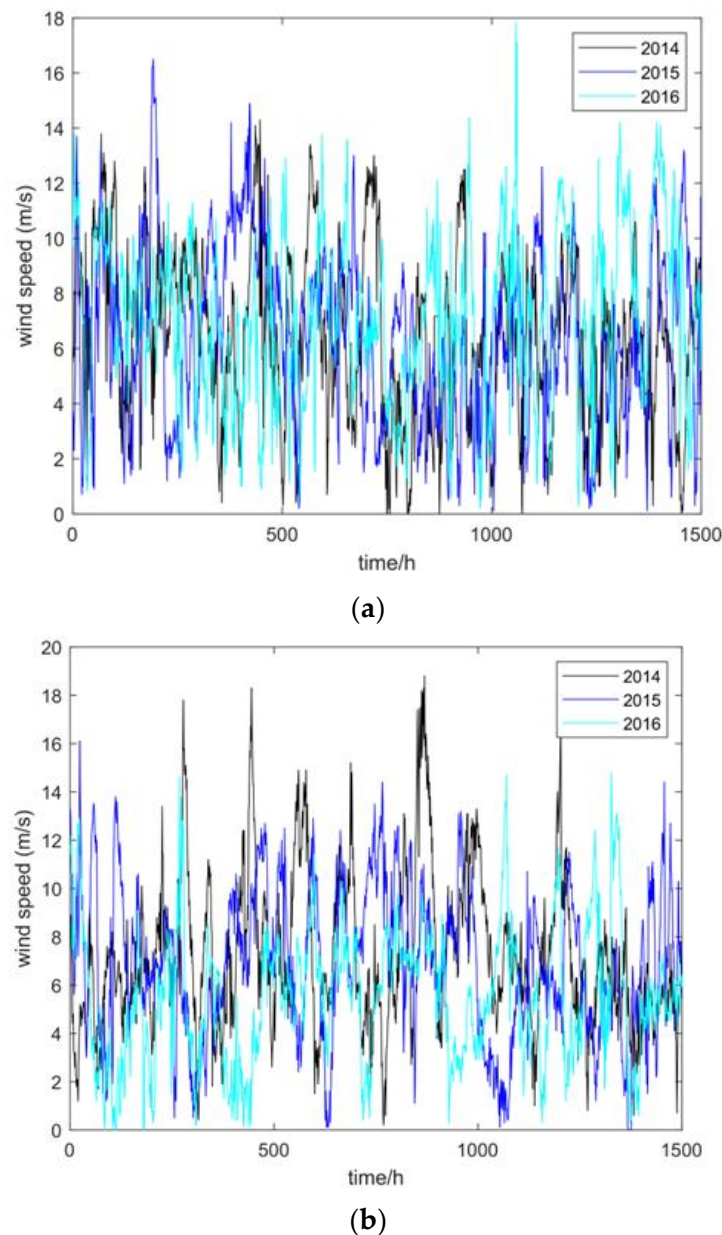


Figure 13. WS of two stations: (a) Station A, (b) Station B. The wind speed and SWH are the input, and the predicted next SWH is the output. Data including WS and SWH from the years 2014 and 2015 are used as the training data, and data including SWH from the year 2016 are used as the testing data.

Table 6. Specific information on Station A and Station B regarding wind speed.

Station ID	Water Depth (m)	Year	WS Average (m)	WS Range (m)
A	59.4	2014	6.4192	[0.0,14.3]
		2015	6.4140	[0.1,16.5]
		2016	7.1163	[0.2,17.8]
B	5309	2014	7.1917	[0.2,18.8]
		2015	6.9334	[0.0,16.1]
		2016	5.5119	[0.0,14.8]

In the prediction model, we input ten parameters, which included five SWHs and five WSs. With SWHs and five WSs at five times being the input, the predicted SWH at the next time is the output. The SWHs are predicted for 1, 3, and 6 h by different models. The predicted results are shown in Table 7.

Table 7. Performance results for Station A and Station B with WS and SWH as inputs.

Station	Algorithm	Leading Time	RMSE	MAE	MSE	R ²
A	EMD-PSO-LSSVM	1	0.093	0.066	0.009	0.970
		3	0.110	0.083	0.012	0.956
		6	0.160	0.117	0.026	0.900
	EMD-LSSVM	1	0.097	0.071	0.009	0.996
		3	0.125	0.092	0.016	0.993
		6	0.169	0.123	0.029	0.888
	LSSVM	1	0.294	0.170	0.087	0.642
		3	0.338	0.228	0.114	0.508
		6	0.397	0.283	0.157	0.341
B	EMD-PSO-LSSVM	1	0.081	0.056	0.007	0.979
		3	0.181	0.118	0.033	0.899
		6	0.332	0.208	0.111	0.717
	EMD-LSSVM	1	0.124	0.084	0.015	0.952
		3	0.230	0.141	0.053	0.843
		6	0.284	0.184	0.081	0.782
	LSSVM	1	0.525	0.332	0.276	0.410
		3	0.559	0.382	0.313	0.356
		6	0.632	0.464	0.399	0.265

By comparing Tables 5 and 7, it can be seen that the prediction performance is not improved by adding the WS parameter as the additional input. The prediction of the SWH is more accurate without the additional wind speed parameter input. The performance with WS and SWH as inputs can be seen by a decrease in the prediction accuracy from $R^2 = 0.972, 0.958, \text{ and } 0.902$ to $R^2 = 0.970, 0.956, \text{ and } 0.900$ at Station A when the leading time is 1, 3, and 6 h. The prediction performance is shown in Figure 14.

When the wind is used as an input to the prediction model, one of the reasons that the prediction accuracy may decrease is because the historic data of SWH implicitly carry information about wind, and that additional input wind might add some redundant data. By adding the wind parameters in the input, the models are prone to overfitting, which reduces the prediction ability of the model. The other reason is that the wind is uncertain or random. It may be inaccurate to measure the position of the waves with the characteristics of the wind, and the influence of wind on waves has the characteristics of time delay, so the prediction accuracy decreases. Therefore, adding wind speed as an input to the prediction model does not necessarily improve the prediction accuracy in predicting local waves.

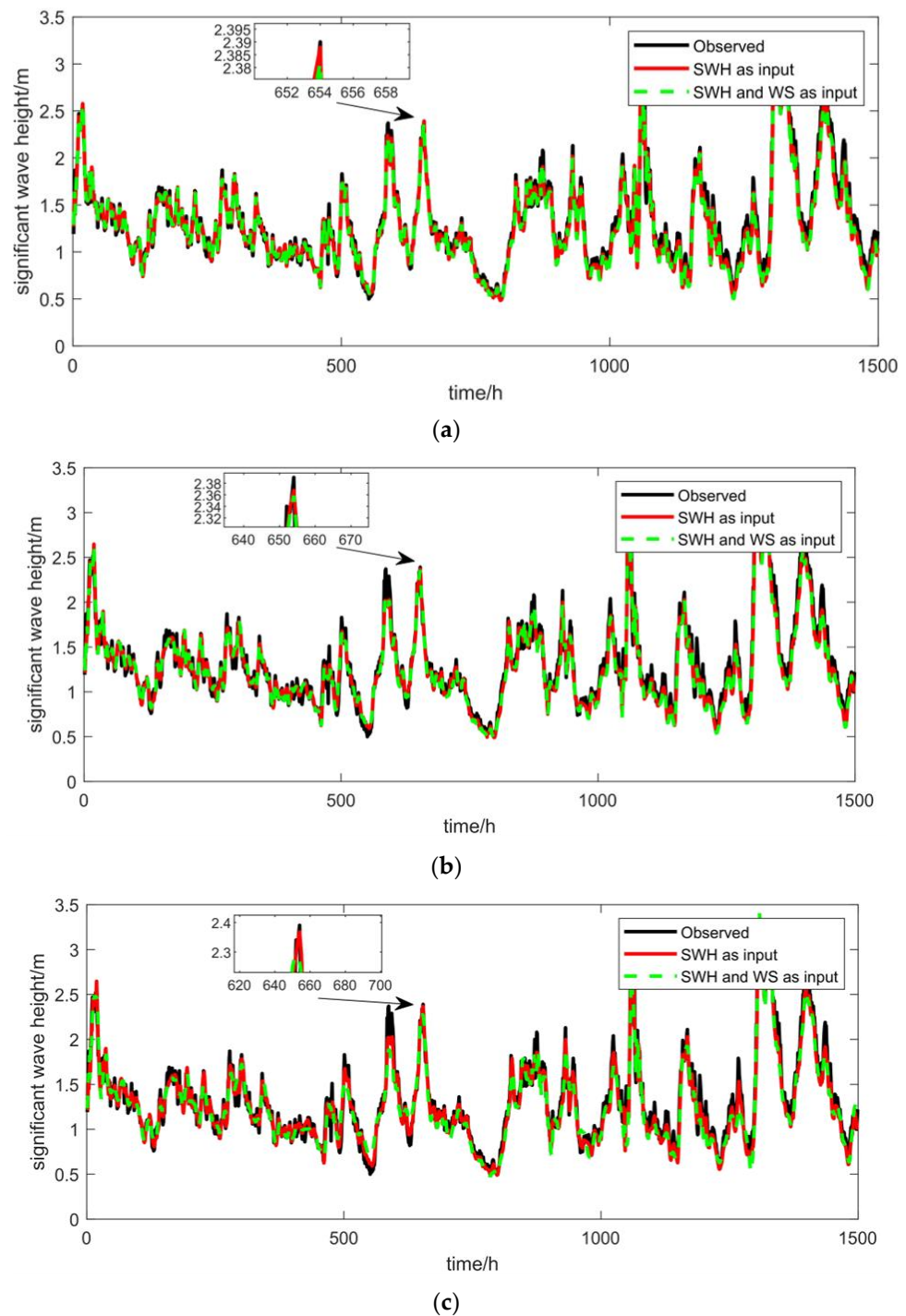


Figure 14. Comparison between the prediction with only SWH as input and with both SWH and WS as inputs by the proposed hybrid EMD–PSO–LSSVM. (a) one hour, (b) three hours, (c) six hours. It can be seen that the model prediction performance is the best when only SWH is used as the input.

4.4. Discussion

This section compares the LSSVM and ANN prediction model performance based on two data sets. The single LSSVM and ANN model will have obvious delay when predicting SWH, and obvious errors when predicting peak and trough data. In order to further improve the prediction performance of LSSVM, the original signal is decomposed by the EMD, and the decomposed value is predicted. Finally, the superimposed composite

signal is equivalent to the significant wave height. PSO is used to optimize the parameters of LSSVM, which further improves the accuracy of the proposed prediction model.

Various data from Table 4 show that the proposed EMD–PSO–LSSVM predictor of the SWH in different locations provides consistent outcomes. The proposed hybrid model can solve the problem of prediction lag and significantly improve the prediction accuracy. In the above analysis, we find that the prediction model is better in deep-sea areas than in shallow water when the wind parameters are not added to all the models. When the wind parameters are added to all the models, the prediction performance is not improved. One of the reasons is that the parameters of SWH include already the influence factors of the wind. By adding wind parameters in the input, the models are prone to overfitting, which reduces the prediction ability of the model. Compared with other models, it can be seen that the proposed hybrid EMD–PSO–LSSVM has a good prediction performance. Not only that, but the hybrid EMD–PSO–LSSVM model is also more stable in the prediction process, which means that it can be applied to more fields in the future.

5. Conclusions

This paper presents a new prediction method based on the hybrid EMD–PSO–LSSVM for the nonlinear and nonstationary SWH prediction. The whole approach is investigated by considering actual forecasting operations on the SWH in the offshore and deep-sea areas of the North Atlantic Ocean, where single and hybrid models are compared using several statistical indices for evaluating the accuracy of the predictions. From the obtained results, due to the nonlinearity and nonstationary data of the SWH, usual single models are prone to a lagging phenomenon that reduces the prediction accuracy. The PSO algorithm is added to the original EMD–LSSVM hybrid method, and the critical parameters of LSSVM are optimized through the PSO algorithm. In this way, a new hybrid model EMD–PSO–LSSVM is developed. The main results are as follows:

1. When local waves are predicted, the prediction effect can be better when only the past observation value of the waves is used in the prediction model, rather than the mixed input of the waves and the wind speed.
2. The proposed hybrid EMD–PSO–LSSVM model has good superiority for the prediction of non-linear and non-stationary waves. EMD is used to decompose the SWH data into a number of IMF components and one residual signal; then, LSSVM is used to forecast these IMFs and residual values individually. PSO is implemented to automatically perform the parameter selection in LSSVM modeling to gain the optimization parameters of LSSVM.
3. Based on the LSSVM model, the performance study results show that EMD–PSO–LSSVM performs better than the EMD–LSSVM and LSSVM models, with higher prediction accuracy in the wave prediction.

Author Contributions: Conceptualization, H.D.; software, H.D.; resources, H.D.; writing—original draft preparation, H.D.; writing—review and editing, J.Z., J.L., H.L., Y.W. and Y.D.; visualization, H.D.; supervision, G.T.; funding acquisition, Y.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by the Guangdong Science and Technology Program under grant 2021A1515011854 and the Guangdong Science and Technology Program under grant 2022A1515011707.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Acknowledgments: The authors would like to thank the funding body for the grant.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

Abbreviation	Definition
SWH	Significant Wave Height
EMD	Empirical Mode Decomposition
LSSVM	Least Squares Support Vector Machine
PSO	Particle Swarm Optimization
SMB	Sverdrup, Munk, and Bretschneider
PNJ	Pierson–Neumann–James
AR	Auto-regressive
ARMA	Auto-regressive Moving Average
ARIMA	Autoregressive Integrated Moving Average
ANNs	Artificial Neural Networks
RNN	Recurrent Neural Network
SVM	Support Vector Machine
SOS	Symbiotic Organisms Search
SVR	Support Vector Regression
CNN	Convolutional Neural Network
WPSO–ELM	Wavelet-PSO–ELM
IMFs	Intrinsic Mode Functions
KKT	Karush–Kuhn–Tucker
RBF	Radial Basis kernel Function
NDBC	National Data Buoy Center

References

- Mubasher, A.; Salah, H.; Elgohary, T. Significant Deep Wave Height Prediction by Using Support Vector Machine Approach (Alexandria as Case of Study). *Int. J. Curr. Eng. Tech.* **2017**, *7*, 135–143.
- Richter, M.; Schaut, S.; Walser, D.; Schneider, K.; Sawodny, O. Experimental validation of an active heave compensation system: Estimation, prediction and control. *Control Eng. Pract.* **2017**, *66*, 1–12. [[CrossRef](#)]
- Cornejo-Bueno, L.; Garrido-Merchán, E.C.; Hernández-Lobato, D.; Salcedo-Sanz, S. Bayesian optimization of a hybrid system for robust ocean wave features prediction. *Neurocomputing* **2018**, *275*, 818–828. [[CrossRef](#)]
- Ra, W.S.; Whang, I.H. Real-time long-term prediction of ship motion for fire control applications. *Electron. Lett.* **2006**, *42*, 1020–1022. [[CrossRef](#)]
- Ippen, A.T. *Estuary and Coastline Hydrodynamics*; IncCatalog Card Number 65–27677; McGraw-Hill Book Company: New York NY, USA, 1966.
- Kim, Y.C. *Handbook of Coastal and Ocean Engineering*; California State University: Los Angeles, CA, USA, 2009.
- Hwang, P.A. Duration- and fetch-limited growth functions of wind-generated waves parameterized with three different scaling wind velocities. *J. Geophys. Res.* **2006**, *111*, C02005. [[CrossRef](#)]
- Casas-Prat, M.; Wang, X.L.; Sierra, J.P. A physical-based statistical method for modeling ocean wave heights. *Ocean. Model.* **2014**, *73*, 59–75. [[CrossRef](#)]
- Janssen, P.A.E.M. Progress in ocean wave forecasting. *J. Comput. Phys.* **2008**, *227*, 3572–3594. [[CrossRef](#)]
- Raj, N.; Brown, J. An EEMD-BiLSTM Algorithm Integrated with Boruta Random Forest Optimiser for Significant Wave Height Forecasting along Coastal Areas of Queensland, Australia. *Remote Sens.* **2021**, *13*, 1456. [[CrossRef](#)]
- Wang, J.; Wang, Y.; Yang, J. Forecasting of Significant Wave Height Based on Gated Recurrent Unit Network in the Taiwan Strait and Its Adjacent Waters. *Water* **2021**, *13*, 86. [[CrossRef](#)]
- Deo, M.C.; Jha, A.; Chaphekar, A.S.; Ravikant, K. Neural networks for wave forecasting. *Ocean. Eng.* **2001**, *28*, 889–898. [[CrossRef](#)]
- Etemad-Shahidi, A.; Mahjoobi, J. Comparison between M5' model tree and neural networks for prediction of significant wave height in Lake Superior. *Ocean. Eng.* **2009**, *36*, 1175–1181. [[CrossRef](#)]
- Jörges, C.; Berkenbrink, C.; Stumpe, B. Prediction and reconstruction of ocean wave heights based on bathymetric data using LSTM neural networks. *Ocean. Eng.* **2021**, *232*, 109046. [[CrossRef](#)]
- Jain, P.; Deo, M.C. Real-time wave forecasts off the western Indian coast. *Appl. Ocean. Res.* **2007**, *29*, 72–79. [[CrossRef](#)]
- Yoon, H.; Jun, S.-C.; Hyun, Y.; Bae, G.-O.; Lee, K.-K. A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *J. Hydrol.* **2011**, *396*, 128–138. [[CrossRef](#)]
- Browne, M.; Castelle, B.; Strauss, D.; Tomlinson, R.; Blumenstein, M.; Lane, C. Near-shore swell estimation from a global wind-wave model: Spectral process, linear, and artificial neural network models. *Coast. Eng.* **2007**, *54*, 445–460. [[CrossRef](#)]
- Smit, P.B.; Houghton, I.A.; Jordanova, K.; Portwood, T.; Shapiro, E.; Clark, D.; Sosa, M.; Janssen, T.T. Assimilation of significant wave height from distributed ocean wave sensors. *Ocean. Model.* **2021**, *159*, 101738. [[CrossRef](#)]
- Demetriou, D.; Michailides, C.; Papanastasiou, G.; Onoufriou, T. Coastal zone significant wave height prediction by supervised machine learning classification algorithms. *Ocean. Eng.* **2021**, *221*, 108592. [[CrossRef](#)]

20. Feng, Z.; Hu, P.; Li, S.; Mo, D. Prediction of Significant Wave Height in Offshore China Based on the Machine Learning Method. *J. Mar. Sci. Eng.* **2022**, *10*, 836. [[CrossRef](#)]
21. Gao, S.; Huang, J.; Li, Y.; Liu, G.; Bi, F.; Bai, Z. A forecasting model for wave heights based on a long short-term memory neural network. *Acta Oceanol. Sin.* **2021**, *40*, 62–69. [[CrossRef](#)]
22. Hao, W.; Sun, X.; Wang, C.; Chen, H.; Huang, L. A hybrid EMD-LSTM model for non-stationary wave prediction in offshore China. *Ocean. Eng.* **2022**, *246*, 110566. [[CrossRef](#)]
23. Gao, R.; Li, R.; Hu, M.; Suganthan, P.N.; Yuen, K.F. Significant wave height forecasting using hybrid ensemble deep randomized networks with neurons pruning. *Eng. Appl. Artif. Intell.* **2023**, *117*, 105535. [[CrossRef](#)]
24. Soares, C.G.; Ferreira, A.M.; Cunha, C. Linear models of the time series of significant wave height on the Southwest Coast of Portugal. *Coast. Eng.* **1996**, *29*, 149–167. [[CrossRef](#)]
25. Guedes Soares, C.; Cunha, C. Bivariate autoregressive models for the time series of significant wave height and mean period. *Coast. Eng.* **2000**, *40*, 297–311. [[CrossRef](#)]
26. Agrawal, J.D.; Deo, M.C. On-line wave prediction. *Mar. Struct.* **2002**, *15*, 57–74. [[CrossRef](#)]
27. Deo, M.C.; Sridhar Naidu, C. Real time wave forecasting using neural networks. *Ocean. Eng.* **1998**, *26*, 191–203. [[CrossRef](#)]
28. Tsai, C.-P.; Lin, C.; Shen, J.-N. Neural network for wave forecasting among multi-stations. *Ocean. Eng.* **2002**, *29*, 1683–1695. [[CrossRef](#)]
29. Makarynsky, O. Improving wave predictions with artificial neural networks. *Ocean. Eng.* **2004**, *31*, 709–724. [[CrossRef](#)]
30. Mandal, S.; Prabakaran, N. Ocean wave forecasting using recurrent neural networks. *Ocean. Eng.* **2006**, *33*, 1401–1410. [[CrossRef](#)]
31. Mahjoobi, J.; Adeli Mosabbe, E. Prediction of significant wave height using regressive support vector machines. *Ocean. Eng.* **2009**, *36*, 339–347. [[CrossRef](#)]
32. Dixit, P.; Londhe, S.; Dandawate, Y. Removing prediction lag in wave height forecasting using Neuro—Wavelet modeling technique. *Ocean. Eng.* **2015**, *93*, 74–83. [[CrossRef](#)]
33. Akbarifard, S.; Radmanesh, F. Predicting sea wave height using Symbiotic Organisms Search (SOS) algorithm. *Ocean. Eng.* **2018**, *167*, 348–356. [[CrossRef](#)]
34. Fan, S.; Xiao, N.; Dong, S. A novel model to predict significant wave height based on long short-term memory network. *Ocean. Eng.* **2020**, *205*, 107298. [[CrossRef](#)]
35. Valamanesh, V.; Myers, A.T.; Arwade, S.R.; Hajjar, J.F.; Hines, E.; Pang, W. Wind-wave prediction equations for probabilistic offshore hurricane hazard analysis. *Nat. Hazards* **2016**, *83*, 541–562. [[CrossRef](#)]
36. Rhif, M.; Ben Abbes, A.; Farah, I.R.; Martínez, B.; Sang, Y. Wavelet Transform Application for/in Non-Stationary Time-Series Analysis: A Review. *Appl. Sci.* **2019**, *9*, 1345. [[CrossRef](#)]
37. Deka, P.C.; Prahlada, R. Discrete wavelet neural network approach in significant wave height forecasting for multistep lead time. *Ocean. Eng.* **2012**, *43*, 32–42. [[CrossRef](#)]
38. Kaloop, M.R.; Kumar, D.; Zarzoura, F.; Roy, B.; Hu, J.W. A wavelet—Particle swarm optimization—Extreme learning machine hybrid modeling for significant wave height prediction. *Ocean. Eng.* **2020**, *213*, 107777. [[CrossRef](#)]
39. Huang, N.E.; Wu, Z. A review on Hilbert-Huang transform: Method and its applications to geophysical studies. *Rev. Geophys.* **2008**, *46*, RG2006. [[CrossRef](#)]
40. Chen, J.; Heincke, B.; Jegen, M.; Moorkamp, M. Using empirical mode decomposition to process marine magnetotelluric data. *Geophys. J. Int.* **2012**, *190*, 293–309. [[CrossRef](#)]
41. Prasad, R.; Deo, R.C.; Li, Y.; Maraseni, T. Input selection and performance optimization of ANN-based streamflow forecasts in the drought-prone Murray Darling Basin region using IIS and MODWT algorithm. *Atmos. Res.* **2017**, *197*, 42–63. [[CrossRef](#)]
42. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. In Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences; The Royal Society Publishing: London, UK; 1998; Volume 454, pp. 903–995.
43. Flandrin, P.; Rilling, G.; Gonçalves, P. Empirical Mode Decomposition as a Filter Bank. *IEEE Signal Process. Lett.* **2004**, *11*, 112–114. [[CrossRef](#)]
44. Duan, W.Y.; Han, Y.; Huang, L.M.; Zhao, B.B.; Wang, M.H. A hybrid EMD-SVR model for the short-term prediction of significant wave height. *Ocean. Eng.* **2016**, *124*, 54–73. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.