

Article

Multi-Step-Ahead Forecasting of Wave Conditions Based on a Physics-Based Machine Learning (PBML) Model for Marine Operations

Mengning Wu ^{1,*} , Christos Stefanakos ² and Zhen Gao ^{1,3,4}

¹ Department of Marine Technology, Norwegian University of Science and Technology (NTNU), NO-7491 Trondheim, Norway; zhen.gao@ntnu.no

² SINTEF Ocean, Department of Environment and New Resources, NO-7465 Trondheim, Norway; Christos.Stefanakos@sintef.no

³ Centre for Autonomous Marine Operations and Systems (AMOS), NTNU, NO-7491 Trondheim, Norway

⁴ Centre for Marine Operations in Virtual Environments (MOVE), NTNU, NO-7491 Trondheim, Norway

* Correspondence: mengning.wu@ntnu.no

Received: 29 October 2020; Accepted: 30 November 2020; Published: 5 December 2020



Abstract: Short-term wave forecasts are essential for the execution of marine operations. In this paper, an efficient and reliable physics-based machine learning (PBML) model is proposed to realize the multi-step-ahead forecasting of wave conditions (e.g., significant wave height H_s and peak wave period T_p). In the model, the primary variables in physics-based wave models (i.e., the wind forcing and initial wave boundary condition) are considered as inputs. Meanwhile, a machine learning algorithm (artificial neural network, ANN) is adopted to build an implicit relation between inputs and forecasted outputs of wave conditions. The computational cost of this data-driven model is obviously much lower than that of the differential-equation based physical model. A ten-year (from 2001 to 2010) dataset of every three hours at the North Sea center was used to assess the model performance in a small domain. The result reveals high reliability for one-day-ahead H_s forecasts, while that of T_p is slightly lower due to the weaker implicit relationships between the data. Overall, the PBML model can be conceived as an efficient tool for the multi-step-ahead forecasting of wave conditions, and thus has great potential for furthering assist decision-making during the execution of marine operations.

Keywords: wave forecasting; machine learning modeling; marine operations; uncertainty quantification

1. Introduction

Marine operations such as sea transportation, offshore crane operation and mooring installation are complex operations which are highly weather sensitive. When planning these operations, the response-based operational limits of the systems and equipment are determined and used to estimate the weather window in combination with the long-term statistics of the environmental conditions. However, during the execution phase of the operation, it is important to make short-term forecasts of wave conditions that are characterized by significant wave height (H_s), peak wave period (T_p) and so on. The forecasted wave conditions are compared with the operational sea state limits to decide whether the operation should start or not as long, as the operations last no longer than 72 h [1]. In view of the specific execution position and typical execution duration of marine operations, reliable wave forecasting in a small area from a few hours to several days ahead is necessary.

Currently, wave forecasting methods are generally categorized into three main types: physics-based numerical methods, statistical methods and machine learning methods. Physics-based numerical methods are theory-driven and mimic the behavior of wave evolution based on physical processes such

as wave generation, propagation and dissipation. To date, the most mature physics-based wave model is the third-generation wave model [2]. It applies the wave energy balance equation with various source terms as inputs and uses numerical techniques to solve it. Among source terms, wind forcing is the primary energy source, and therefore, the wind speed and direction of the past, the present and the forecast time horizon are normally required as inputs. Based on the third-generation wave model, different hindcast and forecast products [3–7] have been provided by various institutes, such as the European Centre for Medium-Range Weather Forecasts (ECMWF), the Norwegian Meteorological Institute (MET-Norway) and the National Center for Environmental Prediction (NCEP). However, the application of numerical wave models requires a good understanding of wave theory, insights into the potential implications of coefficient tunings [8], the proper selection of scale and resolution, etc. More importantly, the models usually require massive amounts of computing time and archiving resources, and this computational expense is often a major limitation of physics-based numerical methods for ordinary use.

In contrast, the other two typical methods (statistical method and machine learning method) for forecasting are data-driven. Instead of solving the energy balance equation by numerical techniques, these approaches can provide forecasts by the relation directly found from the historical time series. The former one is the conventional method used in time series wave forecasting by expressing the future wave as an explicit function of its previous data. It does not require the input of wind conditions. There are several commonly used models for this method, such as the autoregressive model [9,10], the autoregressive moving average (ARMA) [10,11] and the autoregressive integrated moving average (ARIMA) [12–14]. These models generally show limitations in modeling nonlinear patterns [15]. The latter (machine learning method) describes a more comprehensive input–output relationship through a large amount of historical data of both input and output. This relationship can be a relation between data at different time steps in a single time series or a relation between different variables with an implicit physical background. The most popular model for wave forecasting is the artificial neural network (ANN) [16–20]. In addition to ANN, other advanced neural networks, such as the recurrent neural network (RNN) [21,22], the convolution neural network (CNN) [23] and the adaptive-network-based fuzzy inference system (ANFIS) [24–26], have also been successfully applied in wave forecasting. However, due to the fluctuation and randomness of wave conditions, most of the above studies have mainly focused on one-step-ahead forecasting of wave conditions, while multi-step-ahead forecasting has rarely been studied. Wu et al. [27] proposed a decomposition-ANFIS model to forecast 24-step-ahead wind and wave conditions with relatively good performance. Similar studies [16,28–30] were also conducted with relatively fewer forecast steps. The results indicate that the forecast performances of the adopted machine learning models decreased significantly with the increasing forecast time horizon. This implies that although highly capable of training on data, the complexity and randomness of the wave series still challenge the usage of machine learning algorithms.

Under these circumstances, a physics-based machine learning (PBML) model is proposed in this study combining the physics-based wave model with machine learning technique for wave forecasting. Specifically, the inputs and outputs of the PBML model are first determined by the primary variables in the physics-based wave model, and then a machine learning algorithm is adopted to train and perform multi-step-ahead forecast. In this paper, we use ten-year three-hourly wind and wave data at the North Sea center (from CERA-20C) [31] to assess the forecast performance of the proposed model. For the requirements of marine operations, for example, the installation of offshore wind turbines, this study takes one-day-ahead wave forecasts in a small spatial domain as an example.

The remainder of this paper is organized as follows: The methodology is introduced in Section 2. Sections 2.1 and 2.2 describe physics-based wave models and machine learning algorithms, respectively. They provide a reference for the PBML model for wave forecasting, which is presented in Section 2.3. In addition, the error metrics utilized to evaluate forecast performance are introduced in Section 2.4. To assess the proposed model, the multi-step-ahead forecasting of wave conditions at the North Sea

center is conducted in Section 3. Section 3.1 presents a brief description of the case study area and dataset. For the dataset, the model architecture is determined in detail in Section 3.2. The forecasting results of the total wave conditions and two wave components (i.e., wind-generated waves and swells) are investigated with a comparison to the actual hindcast data and discussed in Sections 3.3 and 3.4, respectively. Finally, the main conclusions are summarized in Section 4.

2. Methodology

2.1. Physics-Based Wave Models

In recent decades, many physics-based wave models have been adopted in the wave forecasting with the improvement of computational capacity, such as WAM (wave modeling) [2], WaveWatch III [32], SWAN (simulating waves nearshore) [33] and STWAVE (steady state wave model) [34]. Among them, WAM and WaveWatch III are ocean scale models which are predominantly used for predictions in deep water, while SWAN and STWAVE are coastal or shelf-sea models which include the wave modifications in the nearshore area and are more oriented towards predictions in shallow water [8]. Typical forecasted wave parameters are short-term (for example, one-hour or three-hour) wave spectra, from which significant wave height, and the spectral wave period for both wind-generated waves and swells can be obtained. They describe the wave evolution according to the energy balance equation (see Equation (1)) with a range of source terms.

$$\frac{DE}{Dt} = S \tag{1}$$

where $\frac{D}{Dt}$ is a derivative operator and t is the time. E is the wave energy density spectrum. S is the source function describing different physical processes.

Components in the source item S are the key factor for different physics-based wave models. The widely used and reliable SWAN model will be briefly introduced here to show the modeling process and primary variables of numerical wave models. SWAN is a third-generation spectral wave prediction model developed by the Delft University of Technology in 1999. Its source term includes wave generation; dissipation and quadruplet wave-wave interactions; and dissipation due to bottom friction, triad wave-wave interactions and depth-induced breaking. In this model, the action balance equation is utilized instead of the energy balance equation since the action density N (defined as $N = E/\sigma$) is conserved during wave propagation [35]. The action balance equation in the Cartesian coordinate system can be expressed as:

$$\frac{\partial N}{\partial t} + \frac{\partial c_x N}{\partial x} + \frac{\partial c_y N}{\partial y} + \frac{\partial c_\sigma N}{\partial \sigma} + \frac{\partial c_\theta N}{\partial \theta} = \frac{S_{tot}}{\sigma} \tag{2}$$

where N is an abbreviation for $N(x,y,t;\sigma,\theta)$ with respect to frequency σ and propagation directions θ in space x, y and time t . c_x and c_y are the propagation velocities of wave energy in spatial x and y space, respectively. c_σ and c_θ are the propagation velocities in spectral σ and θ space, respectively. S_{tot} is the source term that represents different physical processes, including both deep and shallow water processes.

The first item on the left-hand side represents kinematic change in time, while the second and third items denote spatial change. The last two terms depict the refraction effects and shifting of the relative frequency, respectively. The source item, on the right-hand side, consists of six components describing the generation, dissipation or redistribution of wave energy, which are expressed in Equation (3).

$$S_{tot} = S_{in} + S_{nl3} + S_{nl4} + S_{ds,w} + S_{ds,b} + S_{ds,br} \tag{3}$$

where:

- S_{in} denotes wave growth by the wind;

- S_{nl3} denotes the nonlinear transfer of wave energy through triad wave-wave interactions;
- S_{nl4} denotes the nonlinear transfer of wave energy through quadruplet wave-wave interactions;
- $S_{ds,w}$ denotes wave dissipation due to whitecapping;
- $S_{ds,b}$ denotes wave dissipation due to bottom friction;
- $S_{ds,br}$ denotes wave dissipation due to depth-induced wave breaking.

Regarding wave forecasting based on the SWAN model, the mean wind speed and direction at 10 m above mean sea level are the key inputs that drive the model runs. Initial and boundary conditions of the wave field are also needed in the spectrum form, such as JONSWAP spectra and Pierson–Moskowitz (PM) spectra. The finite difference method is employed together with the Gauss–Seidel technique [36] for spectral wave computation. After calculation, the wave statistics, such as significant wave height H_s , peak wave period T_p and mean wave period T_m , can be derived by integrating the wave spectrum. In addition, from the split spectrum (i.e., the wind-generated and swell ones), wind and swell contributions can be easily revealed. It should be mentioned that the widely used peak period T_p is defined from the energetic concept and is only valid in total waves. For the wind-generated wave or swell, only the mean period T_m is available.

As shown above, a physics-based wave model (such as SWAN) can correctly describe the phenomena of the wave formation and propagation by the key inputs and governing equations. However, a physics-based wave model normally requires large computational and storage resources. On one hand, the space-time evolution of the wave energy needs to be solved based on a complex mathematical system, including differential and integral equations, parameterization schemes, highly nonlinear empirical expressions, etc. On the other hand, modeling usually covers a large region or global domain, which increases the storage burden of input and output information.

2.2. Machine Learning Algorithms

Recently, machine learning algorithms have drawn more attention in forecast problems as alternative methods based only on time series due to their high cost efficiency. Regarding wave forecasting, a popular type of machine learning method is based on the use of historical wave data only, not relying on the input of wind conditions. Such data can be based on measurements or simulations. It can learn the underlying behavior of a system from a large amount of historical data without having to derive the explicit relationships between them. That is, it can provide results without a governing equation by relying solely on the time history of wave data.

In this study, an artificial neural network (ANN), one of the most popular machine learning algorithms, is utilized due to its ability to capture the nonlinear relations between inputs (past and current wave data) and outputs (forecasted wave data) and to produce multiple outputs simultaneously. It mimics the information processing mechanisms and processes of neurons in the human brain [37]. An ANN consists of three types of layers: the input layer, hidden layers and output layer. A simple ANN with one hidden layer is shown in Figure 1.

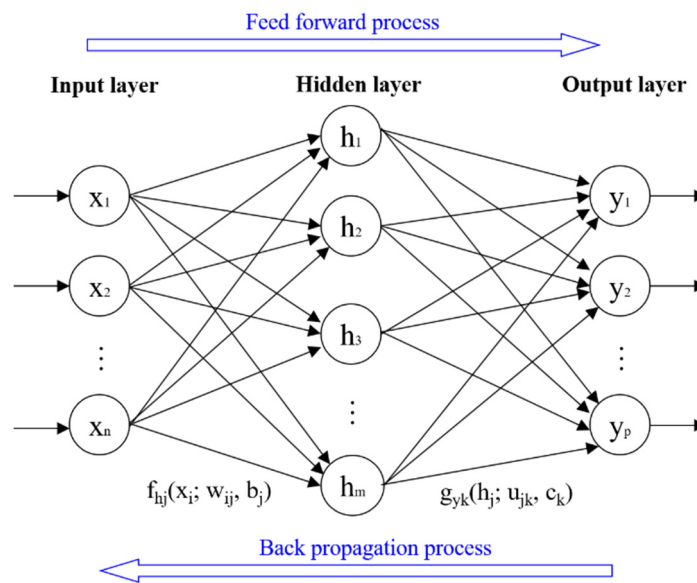


Figure 1. The structure of a typical ANN.

The first and last layers are called the input and output layers, respectively. They contain n and p neurons representing the input and output variables, respectively. The middle layer connecting the input and output layers is the hidden layer, which contains m neurons. The neuron in each layer receives input information (from the neurons in the previous layer based on the corresponding weight, bias parameter and activation function) and then further propagates it forward to the neurons in the following layer. To illustrate this process more clearly, Equations (4) and (5) show the calculations of the hidden neuron and output neuron, respectively.

$$h_j = f_{hj} \left(\sum_{i=1}^n w_{ij} x_i + b_j \right), \quad j = 1, \dots, m \tag{4}$$

$$y_k = g_{yk} \left(\sum_{j=1}^m u_{jk} h_j + c_k \right), \quad k = 1, \dots, p \tag{5}$$

where f and g are activation functions of the hidden and output layers, respectively. w_{ij} and u_{jk} are weight values from input neuron x_i to hidden neuron h_j and from hidden neuron h_j to output neuron y_k , respectively. b_j and c_k are the biases of the corresponding neurons h_j and y_k .

In this study, the back propagation neural network (BPNN) [38] is adopted, which is a model that fine-tunes the weights of a neural network based on the error rate obtained in the previous iteration. In the training phase, the initial weight and bias parameters are selected arbitrarily. Then, they are adjusted iteratively to reduce the error rate by an optimization algorithm (e.g., the gradient descent method). When the overall error is within an acceptable range, the training process stops, and the optimal neural network is determined. In the subsequent testing phase, the forecasts can be obtained by new inputs and this optimal neural network.

Generally, machine learning is an efficient tool that can handle data flexibly. However, the lack of a clear physical background presents a challenge to the performance and reliability of machine learning algorithms for wave forecast using only the historical data of wave conditions.

2.3. The Physics-Based Machine Learning Model

In view of the advantages of physics-based wave models and machine learning algorithms, a hybrid model, termed the physics-based machine learning (PBML) model, is proposed. In the model, physical knowledge from the physics-based wave model is utilized as a guide for designing inputs and

outputs, and the machine learning algorithm is adopted to learn the implicit relationships between them. Afterwards, the trained model can provide wave forecasts directly with new inputs.

The physics-based wave model reveals that the corresponding wind field is the main energy source of the wave field. Thus, wind properties such as the mean speed and direction are introduced as the input variables in the PBML model to forecast wave conditions. In addition, wave boundary conditions at the initial time step are added in the model as well. The wave-wave interaction and wave dissipation effects are implicit in the learned relationship. As a result, the physics-based machine learning (PBML) model is established as follows.

2.3.1. Input and Output Variables

For a typical forecast domain, as shown in the blue part in Figure 2, significant wave height H_s and peak period T_p at the initial time step at the highlighted grid points (by the red lines) are the initial wave boundary conditions. Wind speeds U_w and directions D_u at all concerned grid points are the wind forcing inputs. Similar to the physics-based wave model, the wind forcing should cover the past, the current and the whole forecast time horizon, which can be obtained from the wind forecast models such as numerical weather prediction (NWP) models. The future (forecasted) wave conditions (H_s and T_p) at all grid points are the outputs.

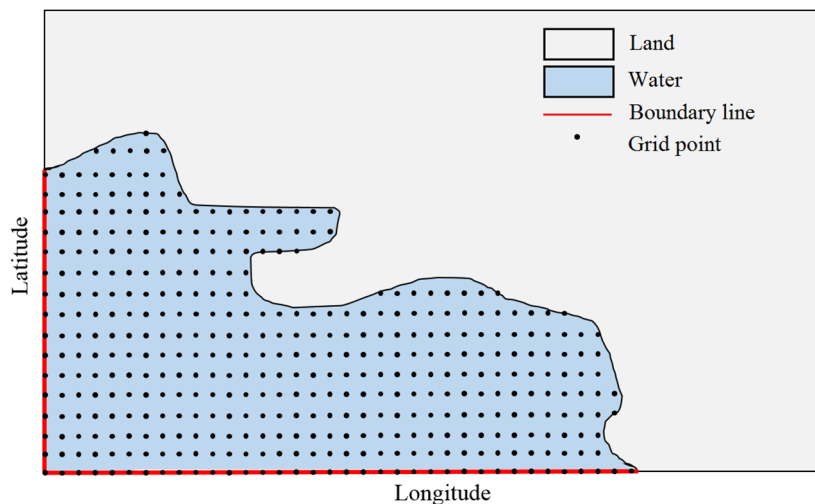


Figure 2. Illustration of the forecast domain.

2.3.2. Calculation Algorithm

Differently from the action balance equation employed in the physics-based wave model, the PBML model provides an alternative method, directly establishing the input-output relationship through machine learning algorithms. In this process, the computational cost can be significantly decreased due to no need to solve partial differential equations by an iteration process. Several machine learning algorithms are available, such as ANN, ANFIS, RNN, and SVM.

2.3.3. Creating the Forecast Model

After deciding the input and output variables and the calculation algorithm, the forecast model can be established. To illustrate this, a forecast model of H_s is shown as an example in Equation (6), and its architecture is shown in Figure 3.

$$\begin{aligned}
 & [H_{s1}(t + N), H_{s2}(t + N), \dots, H_{sn}(t + N)] \\
 & = f_{hN}(H_{s1}(t), T_{p1}(t), \dots, H_{sm}(t), T_{pm}(t), U_{w1}(t + N), D_{u1}(t + N), \dots, \\
 & U_{wn}(t + N), D_{un}(t + N), \dots, U_{w1}(t + N - TU + 1), D_{u1}(t + N - TU + 1), \dots, \\
 & U_{wn}(t + N - TU + 1), D_{un}(t + N - TU + 1))
 \end{aligned} \tag{6}$$

where t is the current time and N is the forecast step. n and m are the number of grid points in the forecast domain and along the boundary lines, respectively. TU is the time range of wind speeds and directions used as wind forcing. The wind input will then include the past and the current values, and the values of the forecast time horizon.

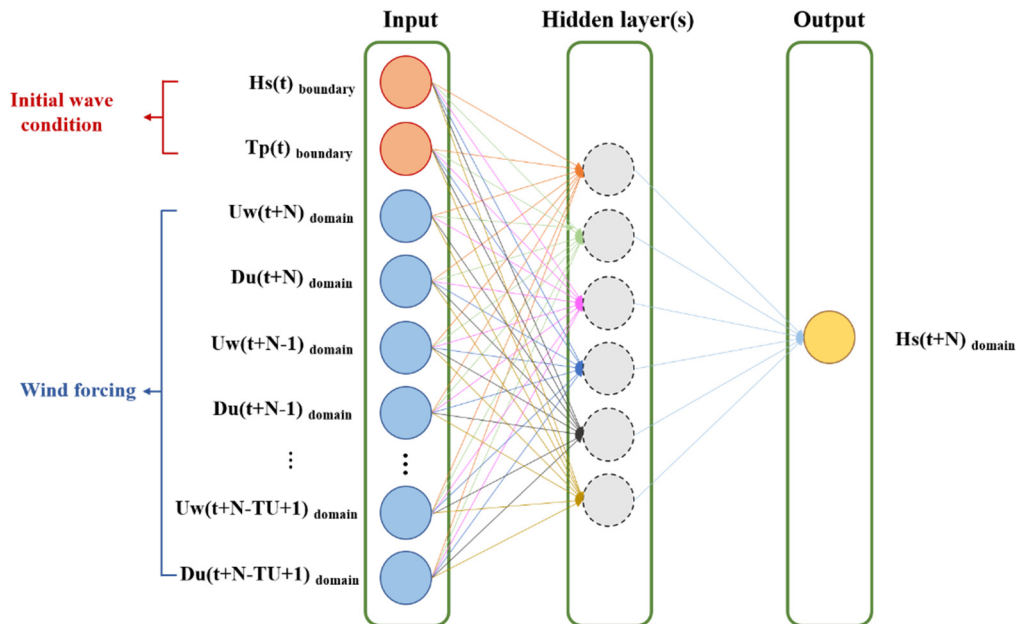


Figure 3. Illustration of the PBML (Physics-Based Machine Learning) model architecture.

As displayed in Equation (6) and Figure 3, the outputs of the model are N -step-ahead H_s in the forecast domain. The input variables include the initial wave boundary conditions (H_s and T_p) at time t and the wind forcing (U_w and D_u) in the forecast domain from $(t + N - TU + 1)$ to $(t + N)$. In addition, f_{hN} indicates the forecast system at N -step-ahead, which depends on the selected machine learning algorithm. In this study, ANN is used.

2.3.4. Training the Established Model

Afterwards, a training process is performed using the training dataset to learn the relationship between inputs and outputs shown in Equation (6). The training process varies according to the selected machine learning algorithm. For instance, for ANN, the weights and bias in the model are updated in each iteration based on the optimization algorithm (e.g., gradient descent algorithm). This process repeats until the stopping criteria are reached.

2.3.5. Evaluating the Performance of the Trained Model

Once training is complete, the model is considered satisfactory and able to make forecasts. By using the inputs in the testing data, the forecasts can be generated from the trained model. Then, the forecast accuracy can be evaluated in terms of several error metrics by comparing forecasted and actual outputs. The metrics used for evaluating the forecast performance are described in detail in Section 2.4.

2.4. Forecast Performance Evaluation

After developing a forecast model, it is important to evaluate the model performance. In this study, conventional error measures, such as MAE (mean absolute error), RMSE (root mean square error), bias, R^2 (correlation coefficient), SI (scatter index), MAPE (mean absolute percentage error) and

a new proposed forecast error factor are used to quantify the forecasting accuracy. The expressions of the error measures are shown in Equations (7)–(12).

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |f_i - a_i| \tag{7}$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (f_i - a_i)^2} \tag{8}$$

$$\text{bias} = \frac{1}{n} \sum_{i=1}^n (f_i - a_i) \tag{9}$$

$$R^2 = \frac{\sum_{i=1}^n (f_i - \bar{f})(a_i - \bar{a})}{\sqrt{\sum_{i=1}^n (f_i - \bar{f})^2} \cdot \sqrt{\sum_{i=1}^n (a_i - \bar{a})^2}} \tag{10}$$

$$SI = \sqrt{\frac{\sum_{i=1}^n [(f_i - \bar{f}) - (a_i - \bar{a})]^2}{\sum_{i=1}^n a_i^2}} \tag{11}$$

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \left| \frac{a_i - f_i}{a_i} \right| \times 100 \tag{12}$$

where n is the number of data points in the testing dataset. a_i and f_i are i th actual and forecasted data, respectively. \bar{a} and \bar{f} are the mean values of the actual and forecasted data, respectively.

A new proposed error factor ε_M shown in Equation (13), which is defined by the difference between the forecasted and the actual value dividing the actual value, can reveal the forecast uncertainty at each time step. More information on evaluating multi-step-ahead forecasts can be found in reference [27].

$$\varepsilon_M(t) = \frac{f(t) - a(t)}{a(t)} \tag{13}$$

where $f(t)$ and $a(t)$ are the forecasted and actual values, respectively, with a forecast time horizon of t .

3. Data, Results and Discussion

3.1. The Case Study Area and Datasets

The North Sea is considered an area with great potential for offshore wind energy development [39]. In the present study, the North Sea center is the main focus, as highlighted in Figure 4. In this area, various marine operations have been studied and analyzed, such as the installation of offshore wind turbine monopiles and transition pieces [40], and single turbine blade [41–44]. Determining suitable weather windows based on the wave forecasts is a major challenge for decision-making during the execution of these operations.



Figure 4. The case study area and research site (Source: Google Maps).

In this study, the detailed site-specific wave and wind hindcast data extracted from the ECMWF CERA-20C datasets [31,45] were used. Several verification studies [31,46–48] have indicated that the generated data from ECMWF are consistent with the measurement data and suitable for climate studies. However, the same method can be applied when measurement data of both wind input and wave output are available. According to the requirements of the PBML model (described in Section 2.3), the wind direction D_u and wind speed U_w at 10m height, significant wave height H_s and peak wave period T_p , are necessary.

A ten-year dataset (from 2001 to 2010) of waves and wind were extracted from CERA-20C with an interval of 3 hours and a spatial resolution of 0.125° (about 13.7 km) in both longitude and latitude. The first nine years of data (from 2001 to 2009) were considered as the training data for establishing and validating the model, while the last year (2010) of data were considered the testing data for evaluating the forecast performance of the model. Regarding the typical duration of marine operations, the aim of this study was to perform one-day-ahead wave forecasts, that is, eight-step-ahead forecasts for the selected data. Given that marine operations are normally carried out at a selected position, in this study, the focus is on a small area. As shown in Figure 5, the forecast domain contains nine grid points numbered from 1 to 9.

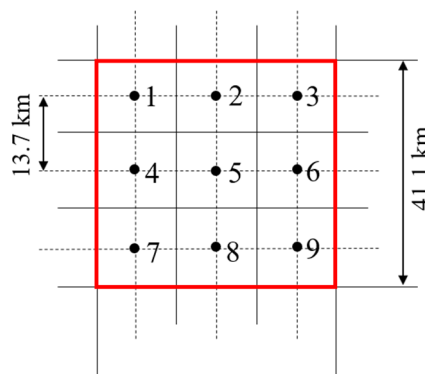


Figure 5. Illustration of a small forecast domain.

3.2. Model Architecture Determination

It is important to determine the model architecture (including the ANN structure and the time range of wind forcing) based on the specific dataset samples before using it. Random selection may result in poor forecast performance. Therefore, the objective of this section is to determine the optimal architecture for the PBML model. For a typical ANN, three key parameters, namely, the number of hidden layers, the number of neurons in each layer and the type of activation function, need to be carefully selected to guarantee a good performance. Regarding the optimal number of hidden layers and neurons, there are no fixed selection formulas, but some rules of thumb are available to determine their scope [49,50]. In the sensitivity study, five different combinations of hidden layers and neurons are considered. In addition, three common activation functions are used for each combination, one of which is the linear function purelin (MATLAB®), and the other two are the nonlinear functions log-sigmoid (logsig, MATLAB®) and hyperbolic tangent sigmoid (tansig, MATLAB®). They are displayed in Figure 6. For point 5 shown in Figure 5, the one-step-ahead forecast results of H_s are summarized in Table 1, where error metrics RMSE and R^2 are used to measure the model's performance. It should be noted that for the sake of simplicity, all results were obtained through the PBML model with $TU = 1$ (see Equation (6)).

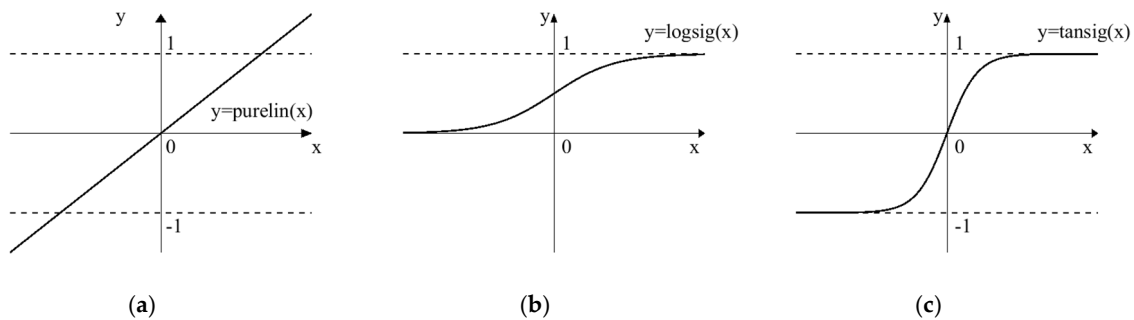


Figure 6. Activation functions used in the sensitivity study: (a) purelin; (b) logsig; (c) tansig.

Table 1. Performance of the PBML model with different ANN architectures (point 5).

Cases	ANN Architecture			Performance			
	Number of Hidden Layers	Number of Neurons	Activation Function	Training Dataset		Testing Dataset	
				RMSE	R^2	RMSE	R^2
1	1	50	purelin	0.0303	0.9966	0.1426	0.9934
2			logsig	0.6496	0	2.7844	0.005
3			tansig	0.0110	0.9995	0.2115	0.9770
4	2	30	purelin	0.0302	0.9966	0.1429	0.9933
5			logsig	0.6512	0.6758	2.7851	0
6			tansig	0.0093	0.9995	0.0972	0.9976
7	3	20	purelin	0.0303	0.9966	0.1426	0.9933
8			logsig	0.6496	0.7080	2.7838	0.0141
9			tansig	0.0092	0.9996	0.0940	0.9978
10	4	15	purelin	0.0303	0.9965	0.1480	0.9925
11			logsig	0.6512	0	2.7851	0
12			tansig	0.0103	0.9995	0.1190	0.9958
13	5	10	purelin	0.0303	0.9966	0.1427	0.9933
14			logsig	0.6512	0.6005	2.7851	0
15			tansig	0.0120	0.9993	0.1134	0.9955

It is clearly seen in Table 1 that the tansig activation function is better than the others, as the RMSE is lower and R^2 is higher for all cases. However, the numbers of hidden layers and neurons have little influence on the forecast performance in all cases with the tansig activation function. By comparison, case 9 obtained the best result. Therefore, a neural network with three hidden layers, twenty neurons in each layer and the tansig type activation function was adopted for the PBML model in this study.

In addition to the ANN structure discussed above, the time range of the input wind field also deserves consideration. As shown in Equation (6), when performing the N-step-ahead forecasting, the wind forcing from $(t + N - TU + 1)$ to $(t + N)$ can be used as input. To investigate the impact of wind forcing at different time steps on H_s forecasts, the ANN was developed in terms of the optimal structure obtained above, and the TU value was selected from one to eight. When TU is equal to 1, only the wind field at the same time as the forecasted H_s is considered. With the increase in TU , an increasing number of past wind fields are taken into account for wave forecasting. The results of the sensitivity study of the time range (TU) of input wind forcing are shown in the Figure 7.

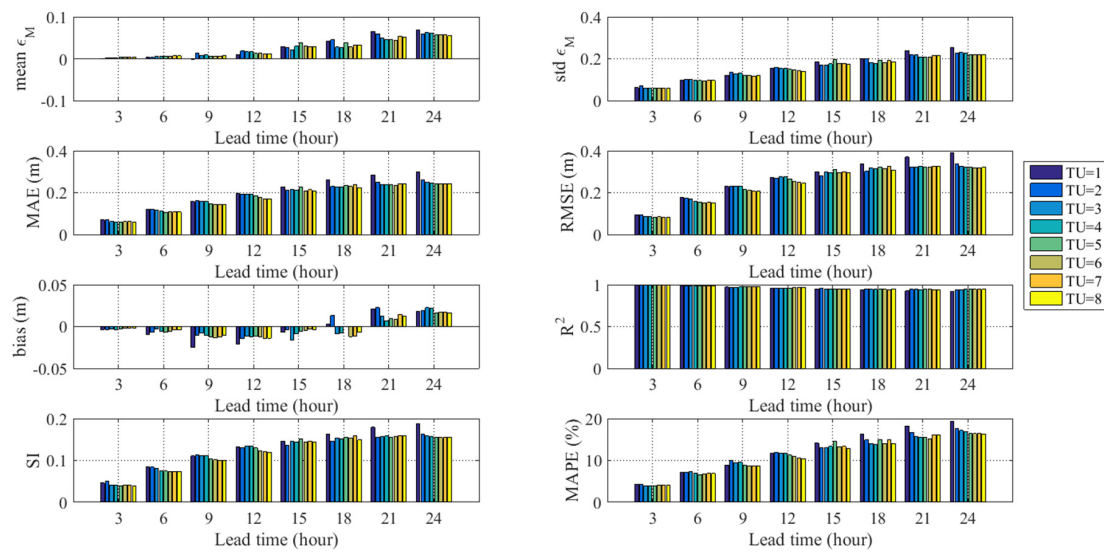


Figure 7. Forecast errors of the PBML model with different durations of wind forcing (point 5).

As seen in Figure 7, in general, the forecast performance will increase with increasing TU values, indicating that the wind variation should be taken into account in wave forecasting by the PBML model. However, the level of improvement in model performance gradually decreases as the TU increases from 2. Meanwhile, increasing the TU value will increase the number of inputs, which makes the forecast model more complicated and requires more computational time. Therefore, $TU = 2$ was chosen in the PBML model in this study. That is, the dynamic characteristics of the winds are provided in the model by applying wind speeds and directions at two consecutive time steps.

In addition, it is interesting to see the importance of the future wind forcing. Figure 8 shows the forecast performance of the model using wind forcing up to $(t + N - TU^*)$ when forecasting wave conditions at $(t + N)$, where TU^* ranges from 0 to 7. From the sensitivity study, it is clear that the forecast performance will gradually decrease as TU^* increases. Therefore, in order to accurately forecast the wave conditions at $(t + N)$, the wind conditions at $(t + N)$ should be included.

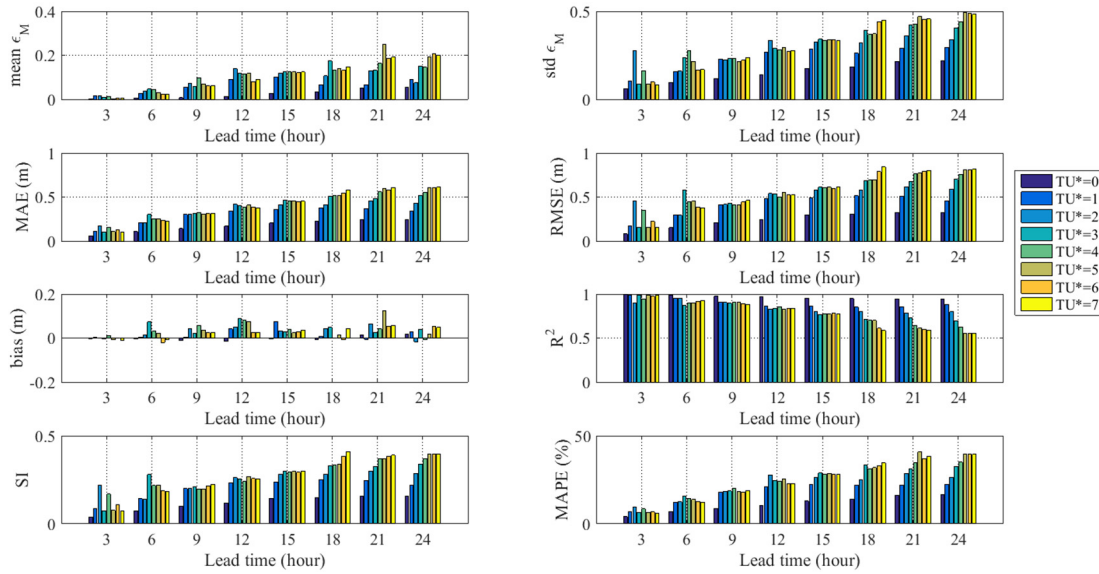


Figure 8. Forecast errors of the PBML model using wind forcing up to $(t + N - TU^*)$ (point 5).

3.3. Forecast of Combined Wind-Generated and Swell Wave Conditions

3.3.1. Significant Wave Height

(1) Small domain

In this subsection, a PBML model is developed for H_s in a small domain given in Figure 5. The model architecture is set up based on the analysis in Section 3.2. Consequently, the forecast model of H_s can be expressed as follows:

$$\begin{aligned}
 & [H_{s1}(t + N), H_{s2}(t + N), \dots, H_{s9}(t + N)] \\
 & = f_{hN}(H_{s1}(t), T_{p1}(t), \dots, H_{s9}(t), T_{p9}(t), U_{w1}(t + N), D_{u1}(t + N), \dots, \\
 & U_{w9}(t + N), D_{u9}(t + N), U_{w1}(t + N - 1), D_{u1}(t + N - 1), \dots, \\
 & U_{w9}(t + N - 1), D_{u9}(t + N - 1)), N = 1, \dots, 8
 \end{aligned} \tag{14}$$

where $H_{s1}(t + N), H_{s2}(t + N), \dots, H_{s9}(t + N)$ are the outputs, which are N-step-ahead forecasts of H_s for points 1 to 9. f_{hN} is the ANN model at the N-step-ahead. The inputs of f_{hN} include the initial wave boundary wave conditions (i.e., $H_s(t)$ and $T_p(t)$ at points 1, 2, 3, 4, 6, 7, 8 and 9) and the wind forcing at two consecutive time steps in the forecast domain (i.e., U_w and D_u for all nine points at N- and N-1-step-ahead). Thus, it can be seen that in this case, the model for each forecast step has a total of 52 inputs and 9 outputs.

The forecast results during the testing period (one year) in the grid, point 5 (in Figure 5), are plotted in Figure 9, in which the actual and forecasted time series are represented by black and red lines, respectively. In addition, the forecast errors at each forecast lead time are summarized in Table 2. To reveal the forecast performance more clearly, two short-time windows are also shown in Figure 9. It is visible from the windows that each forecast contains the forecasting of eight points, from one-step-ahead to eight-step-ahead forecasts. The green and blue points represent one-step-ahead and eight-step-ahead forecast values, respectively, which mean the beginning and end of one forecast. The next forecast starts at the time when the previous forecast horizon ends.

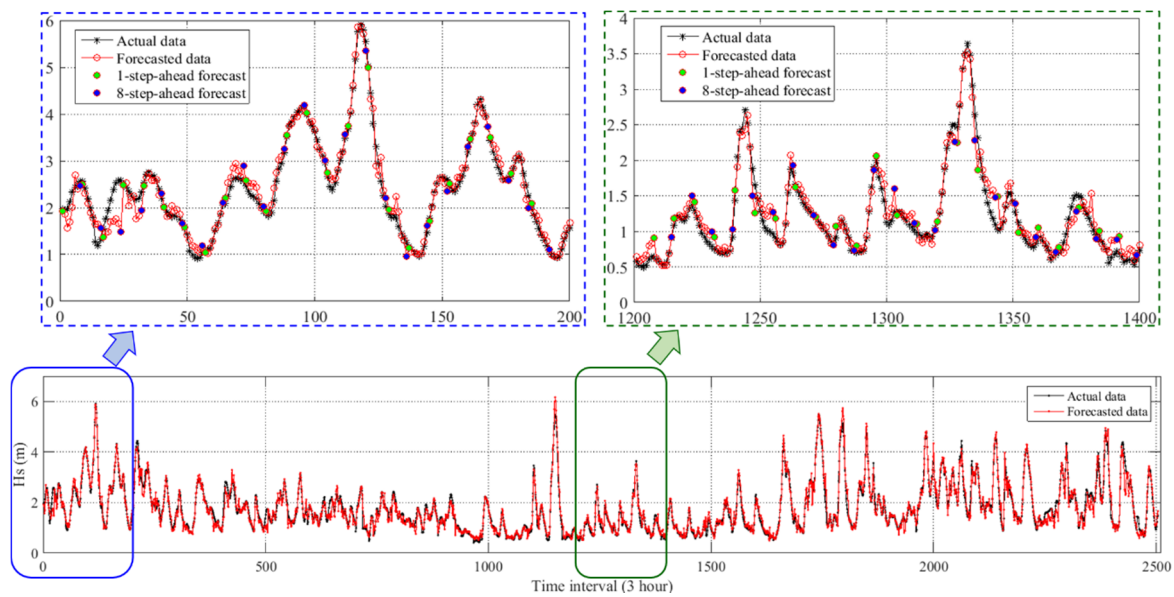


Figure 9. Forecast results of H_s based on the PBML model.

Table 2. Forecast errors of H_s at each lead time based on the PBML model.

Error Measures	Lead Time (h)							
	3	6	9	12	15	18	21	24
m_ϵ	0.0050	-0.0039	0.0054	0.0176	0.0171	0.0158	0.0183	0.0239
std_ϵ	0.0676	0.0862	0.1008	0.1088	0.1290	0.1576	0.1618	0.1741
MAE (m)	0.0617	0.0960	0.1140	0.1303	0.1525	0.1693	0.1827	0.1961
RMSE (m)	0.0927	0.1375	0.1573	0.1828	0.2213	0.2380	0.2587	0.2727
bias (m)	0.0009	-0.0071	-0.0043	0.0130	0.0085	0.0047	-0.0009	-0.0008
R^2	0.9975	0.9903	0.9866	0.9820	0.9742	0.9711	0.9646	0.9611
SI	0.0447	0.0665	0.0767	0.0887	0.1072	0.1155	0.1256	0.1318
MAPE (%)	4.1138	5.9599	7.2648	8.0594	9.4332	10.8368	11.5503	12.6979

Overall, a good agreement between the forecasted and actual data is observed in Figure 9, which reveals the reliability of the PBML model. In addition, there is no significant difference in forecast performance between the two windows. This implies that the proposed PBML model is suitable for any time of year. In Table 2, one can see that the forecast accuracy decreases with the lead time, which is reflected in the decrease in R^2 and the increase in other error metrics. Nevertheless, all 8-step-ahead forecasts can be considered satisfactory because the corresponding errors are quite small. For example, the RMSE and R^2 for all steps are less than 0.3 m and higher than 96%, respectively. Furthermore, regarding the computational time, the PBML model can be quickly executed to train and make forecasts extremely fast. The training process of the PBML model on a common server typically takes few minutes.

The above results only display the forecast at the grid point 5 in the forecast domain. Figure 10 further summarizes the error measures at all nine grid points in the forecast domain done by different colors. Little difference in all error measures can be found regardless of the grid points, which indicates that the PBML model is suitable and reliable in a small forecast domain.

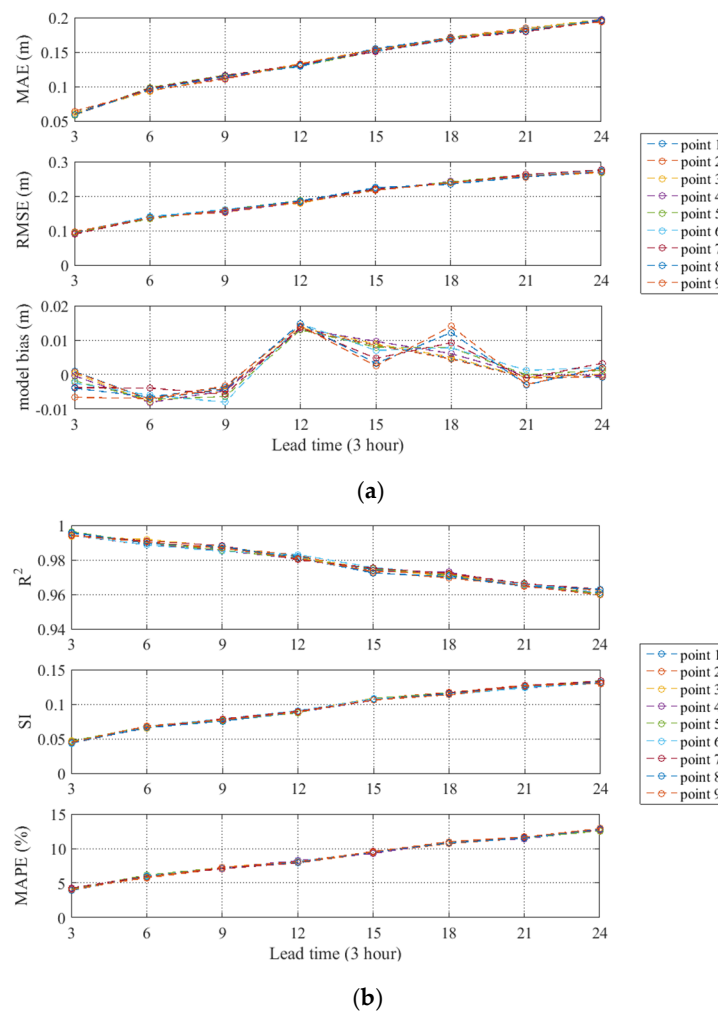


Figure 10. Error measures of H_s for all 9 grid points in the forecast domain. (a) Error measures (MAE, RMSE, bias). (b) Error measures (R^2 , SI, MAPE).

(2) Single position

As shown above, the proposed PBML model exhibits good performance for H_s forecasting in a small domain. In this subsection, the forecast performance of the PBML model for a single point is investigated. This situation is very important to explore whether the PBML model is applicable to the measurement data, since measurement data may only be available at limited positions. In practice, we cannot expect detailed measured data for a large region (which are normally available in reanalysis or hindcasting).

Since the corresponding input information is greatly reduced, the forecast model is simplified to Equation (15). In this case, the inputs are the initial wave conditions and wind forcing at the position of interest (point 5 in Figure 5 is selected as an example here), and the outputs are the future wave conditions at the same position. That is, the model is simplified to have six inputs and one output. The results of forecasted time series and forecast errors are illustrated in Figures 11 and 12, respectively.

$$H_{s5}(t + N) = f_{h5N}(H_{s5}(t), T_{p5}(t), U_{w5}(t + N), D_{u5}(t + N), U_{w5}(t + N - 1), D_{u5}(t + N - 1)), \quad (15)$$

$$N = 1, \dots, 8$$

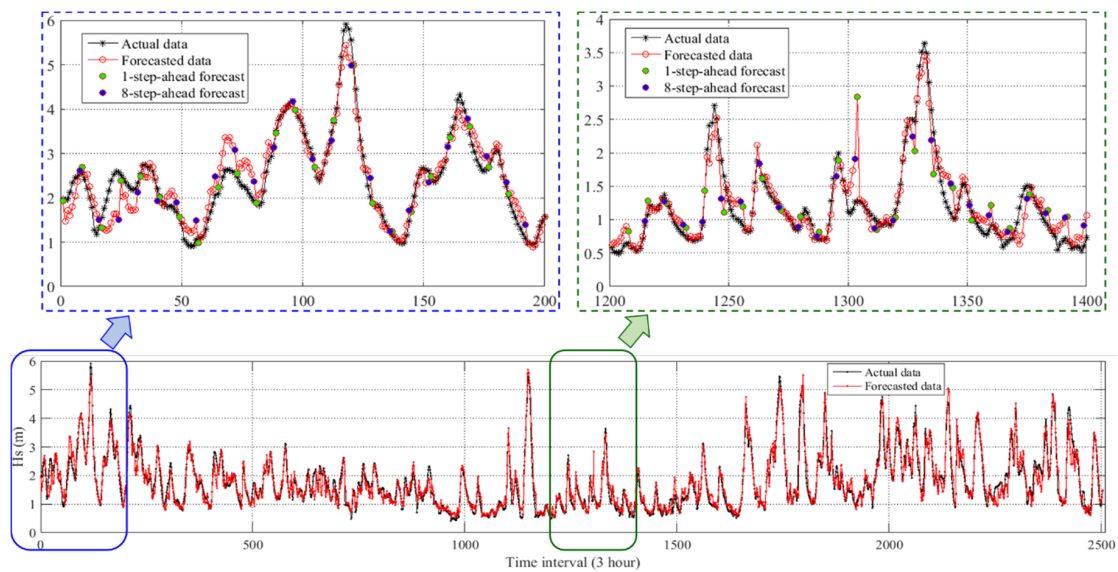


Figure 11. Forecast results of H_s based on the PBML model with one-point domain.

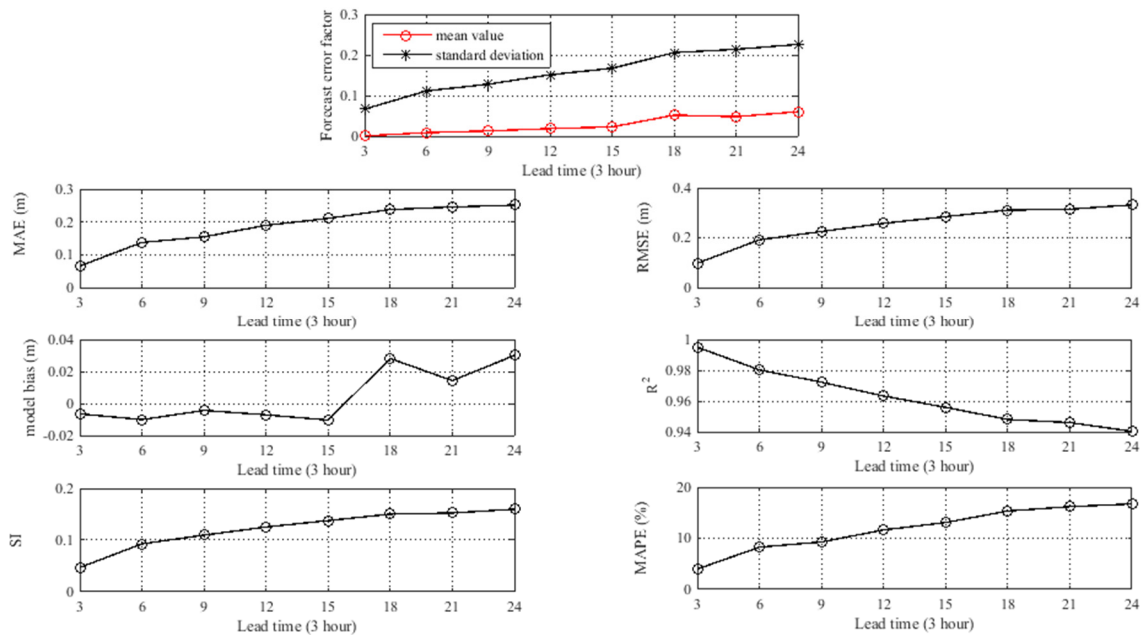


Figure 12. Forecast errors of H_s based on the PBML model with one-point domain.

Compared with the results in Figure 9, the forecast accuracy is reduced to some extent for this simple model due to the significant reduction of the input information. Nevertheless, this model is still acceptable since it exhibits good forecast performance. By observing the forecast errors in Figure 12, one can find that the R^2 and RMSE of all 8-step-ahead forecasts are higher than 94% and lower than 0.33 m, respectively. The results imply that the proposed PBML model can forecast H_s with relatively high accuracy based on the measured data of a single position.

3.3.2. Peak Wave Period

In addition to H_s , the forecast performance of the PBML model for T_p is investigated in this section. Similarly to Equation (14), the corresponding model of T_p is expressed in Equation (16), in which the initial wave boundary conditions and spatial wind forcing at two consecutive time steps are adopted

as the inputs. The setting of the ANN architecture is consistent with that in Section 3.1. The forecast results are shown in Figure 13, and the errors are summarized with respect to lead times in Table 3.

$$\begin{aligned}
 & [T_{p1}(t + N), T_{p2}(t + N), \dots, T_{p9}(t + N)] \\
 & = f_{tN}(H_{s1}(t), T_{p1}(t), \dots, H_{s9}(t), T_{p9}(t), U_{w1}(t + N), D_{u1}(t + N), \dots, \\
 & U_{w9}(t + N), D_{u9}(t + N), U_{w1}(t + N - 1), D_{u1}(t + N - 1), \dots, \\
 & U_{w9}(t + N - 1), D_{u9}(t + N - 1)), N = 1, \dots, 8
 \end{aligned}
 \tag{16}$$

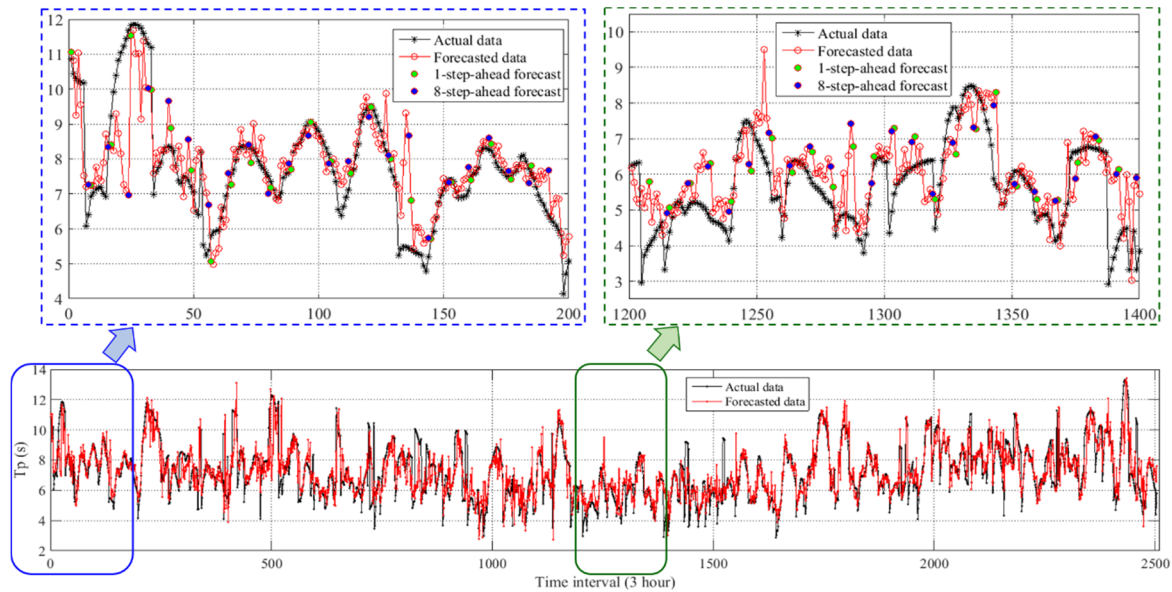


Figure 13. Forecast results of T_p based on the PBML model.

Table 3. Forecast errors of T_p based on the PBML model.

Error Measures	Lead Time (h)							
	3	6	9	12	15	18	21	24
m_ϵ	0.0088	-0.0004	0.0231	0.0081	0.0335	0.0397	0.0321	0.0514
std_ϵ	0.1302	0.1347	0.1412	0.1769	0.1737	0.1868	0.1774	0.1891
MAE (s)	0.3617	0.5276	0.6151	0.7172	0.7771	0.8859	0.9136	0.9605
RMSE (s)	0.7290	0.8633	0.9498	1.1201	1.1186	1.2355	1.2351	1.3026
bias (s)	-0.0038	-0.1092	0.0463	-0.1207	0.0669	0.0575	0.0157	0.1195
R^2	0.9196	0.8855	0.8585	0.8043	0.7941	0.7530	0.7405	0.7125
SI	0.0964	0.1133	0.1259	0.1476	0.1496	0.1655	0.1648	0.1724
MAPE (%)	5.4973	7.6696	8.9811	10.3773	11.6007	13.3601	13.2666	14.1797

In general, PBML model can obtain acceptable T_p forecasts, as the R^2 and RMSE at all lead times are higher than 71% and lower than 1.30 s respectively. However, poor performance can normally be observed at the points on steep gradients (e.g., data near 130 and 1200 in Figure 13), which is less reliable than the forecast of H_s .

Significant fluctuation in T_p might be the key reason for this phenomenon. Differently from H_s , T_p is a wave period associated with the highest energetic waves in the total wave spectrum, and it can be predominated by either wind-generated waves or swells. This uncertainty increases the instability of the series of T_p and further challenges the forecast performance.

To further explore the factors that affect the performance, in Section 3.4, the total waves are separated into wind-generated waves and swells, and the PBML model is utilized to forecast wave conditions of these two components independently.

3.4. Forecast of Separate Wind-Generated Wave Conditions and Swell Conditions

In general, a total wave comprises two components, i.e., the wind-generated wave and the swell. Approximately, the spectral component is considered to be subject to forcing by the wind when the following condition is met [51]:

$$1.2 \times 28(u^*/c) \cos(\theta - \phi) > 1 \tag{17}$$

where u^* is the friction velocity. $c = c(\sigma, h)$ is the wave celerity, as derived from the linear theory of waves, in which σ is wave frequency and h is water depth. θ and ϕ are the wave and wind directions, respectively.

Based on Equation (17), the wind-generated wave and swell can be separated, and the corresponding spectra and wave statistics can be obtained.

In this part, the data of wind-generated waves and swells were also extracted from CERA-20C with the same forecast domain, resolution and duration as those of the total waves. The extracted variables include the significant heights of wind-generated waves H_{sw} and swells H_{ss} ; the corresponding mean periods T_{mw} and T_{ms} ; and the mean wind speed U_w and direction D_u . The forecast performances of mean wave period (T_{mw} and T_{ms}) and significant wave height (H_{sw} and H_{ss}) are discussed in the following subsections, respectively.

3.4.1. Mean Wave Period

The PBML models of T_{mw} and T_{ms} are similar to Equation (14), except for the initial wave condition. In addition, the identical procedure was adopted, which utilized the data from the first nine-year and the last year to train and test the model, respectively. The corresponding results of point 5 (in Figure 5) are shown in Figures 14 and 15 as an example, given the similar forecast performance for all nine points revealed in Section 3.3.1. Meanwhile, Tables 4 and 5 give the forecast errors of T_{mw} and T_{ms} with respect to the lead time, respectively.

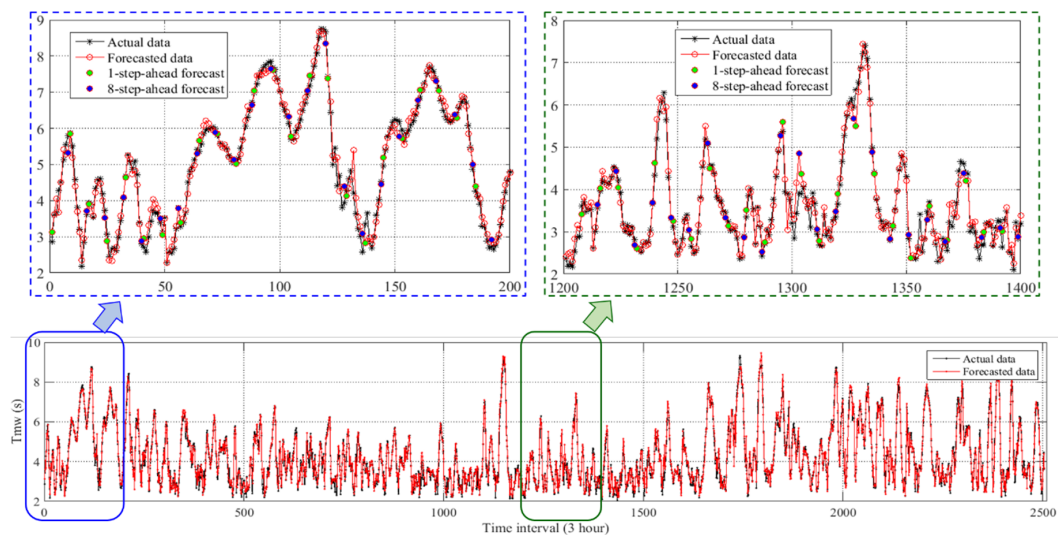


Figure 14. Forecast results of T_{mw} based on the PBML model.

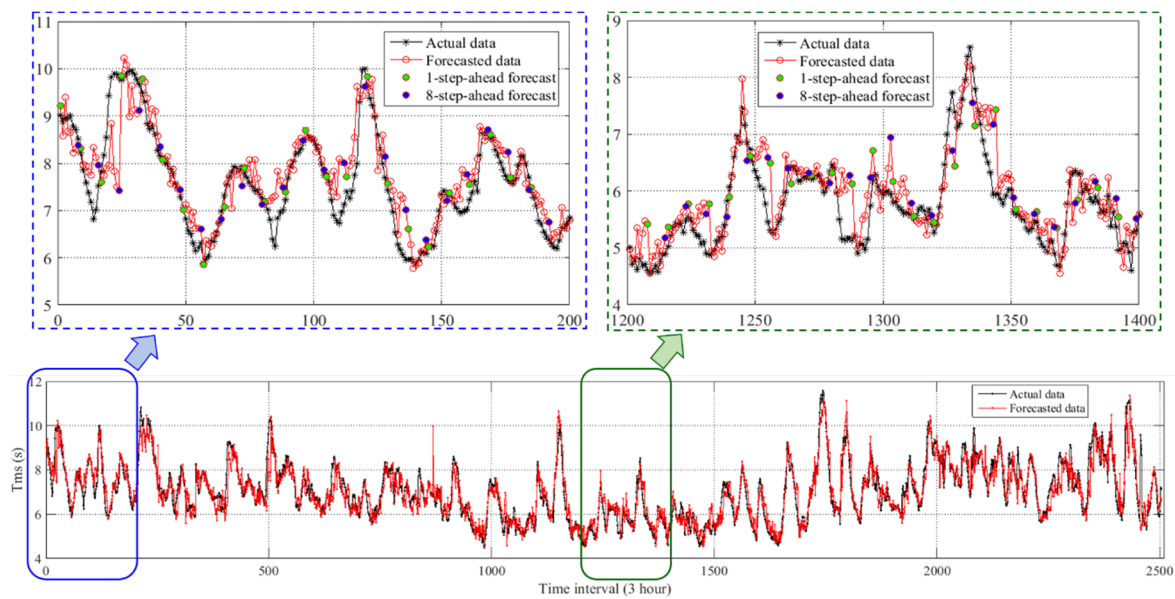


Figure 15. Forecast results of T_{ms} based on the PBML model.

Table 4. Forecast errors of T_{mw} based on the PBML model.

Error Measures	Lead Time (h)							
	3	6	9	12	15	18	21	24
m_ϵ	0.0073	0.0027	0.0087	0.0060	0.0107	0.0051	0.0057	0.0032
std_ϵ	0.0486	0.0601	0.0653	0.0680	0.0685	0.0646	0.0691	0.0716
MAE (s)	0.1061	0.1671	0.1803	0.1938	0.1953	0.1994	0.2076	0.2203
RMSE (s)	0.1572	0.2354	0.2410	0.2551	0.2581	0.2746	0.2791	0.3054
bias (s)	0.0210	-0.0009	0.0168	0.0061	0.0221	-0.0058	0.0037	0.0003
R^2	0.9944	0.9873	0.9864	0.9851	0.9849	0.9825	0.9824	0.9793
SI	0.0342	0.0517	0.0530	0.0562	0.0567	0.0600	0.0611	0.0669
MAPE (%)	2.8555	4.2903	4.6335	4.9302	4.9289	4.7782	5.1044	5.2865

Table 5. Forecast errors of T_{ms} based on the PBML model.

Error Measures	Lead Time (h)							
	3	6	9	12	15	18	21	24
m_ϵ	-0.0013	-0.0046	-0.0017	0.0098	0.0030	0.0171	0.0126	0.0082
std_ϵ	0.0386	0.0515	0.0625	0.0779	0.0766	0.0822	0.0901	0.0960
MAE (s)	0.1471	0.2541	0.3130	0.4186	0.4011	0.4523	0.4969	0.5340
RMSE (s)	0.3164	0.4106	0.4707	0.5595	0.5693	0.6024	0.6648	0.7136
bias (s)	-0.0153	-0.0504	-0.0386	0.0139	-0.0095	0.0695	0.0273	-0.0070
R^2	0.9713	0.9528	0.9373	0.9103	0.9033	0.8870	0.8621	0.8382
SI	0.0438	0.0565	0.0650	0.0777	0.0792	0.0833	0.0926	0.0995
MAPE (%)	2.0241	3.4833	4.3740	5.9968	5.6001	6.4784	7.0630	7.5135

Figure 14 shows good agreement between the forecasted and actual data of wind-generated waves. The small mean (0.003) and standard deviation (0.072) of the forecast error factor shown in Table 4 further verify the high reliability of the PBML model, even at the last forecast step (8-step-ahead/24 h). For the swells, Figure 15 presents a result with less confidence; and a slightly higher mean (0.008) and standard deviation (0.096) of the forecast error factor at eight-steps-ahead also reveal this phenomenon in Table 5. The weak relation between the output (swells) and inputs (wind forcing) is the main reason for this forecast performance. To discuss this phenomenon in detail, Figure 16 shows scatter diagrams of the mean wind speed and different wave periods (i.e., T_p , T_{mw} and T_{ms}) at the same time and position.

The data from year 2001 to 2009 were used to investigate the correlations—about 26,000 data points. The color depicts the absolute frequencies of occurrence. Obviously, the correlation between the mean wind speed and mean wave period for swells (Figure 16c) and total waves (Figure 16a) is much weaker than that for wind-generated waves (Figure 16b) since the swell stems from the propagating waves generated far away from the focused position. The time variant propagating velocity and direction increase the uncertainty of the swells and further increase the forecast complexity of T_p . In contrast, wind-generated waves stemming from the local wind can be exactly described as the wind forcing used as inputs.

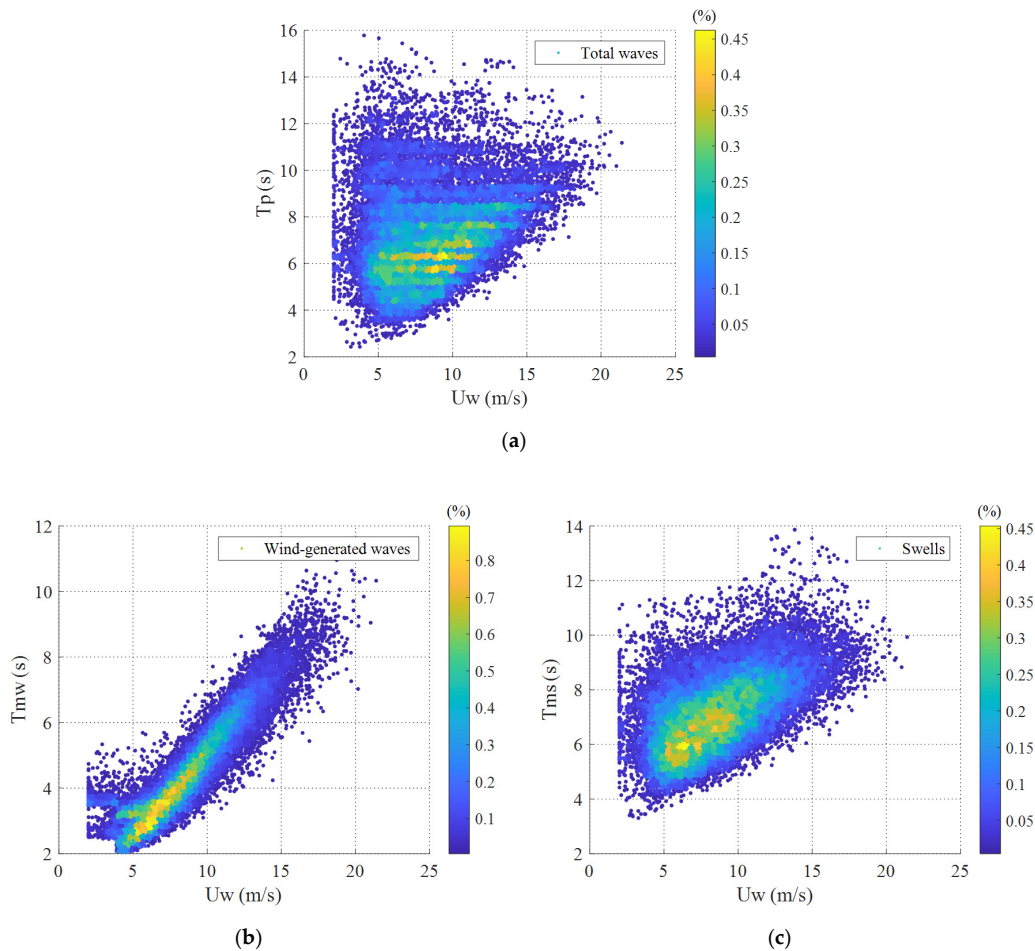


Figure 16. Scatter plot of wave period and mean wind speed: (a) Total waves. (b) Wind-generated waves. (c) Swells.

3.4.2. Significant Wave Height

In line with the study of peak wave period T_p , significant wave height is also treated separately by the sources of wind-generated waves (H_{sw}) and swells (H_{ss}), to further prove the findings observed in the forecast of wave period. The PBML models of H_{sw} and H_{ss} were established respectively, and the corresponding forecast results of point 5 (in Figure 5) are shown in Figures 17 and 18. Meanwhile, Tables 6 and 7 summarize the forecast errors of H_{sw} and H_{ss} , respectively.

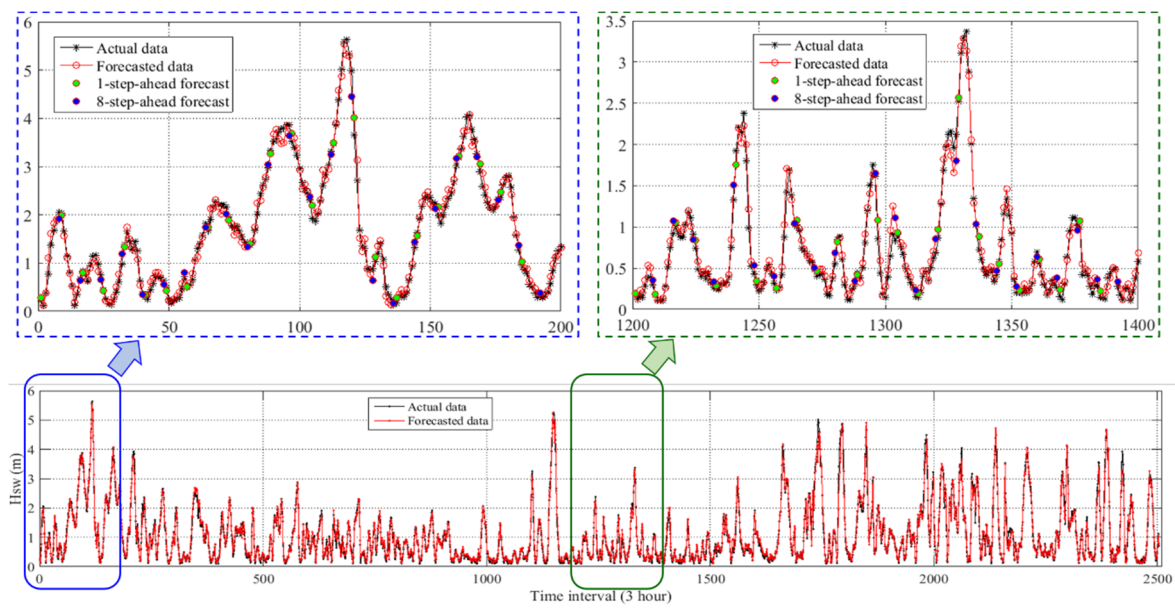


Figure 17. Forecast results of H_{sw} based on the PBML model.

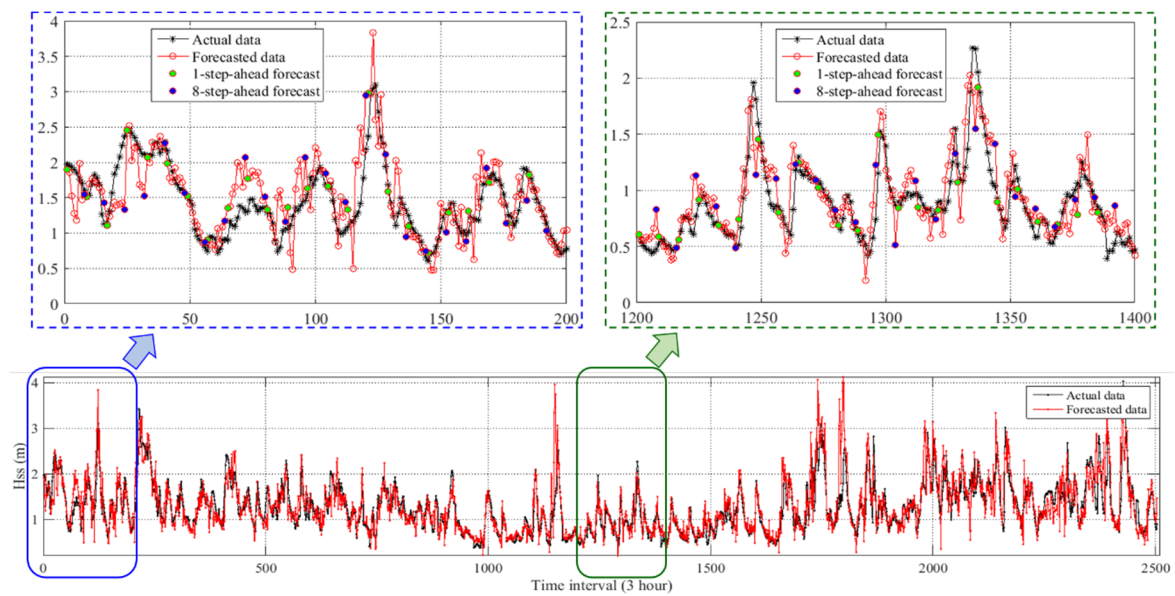


Figure 18. Forecast results of H_{ss} based on the PBML model.

Table 6. Forecast errors of H_{sw} based on the PBML model.

Error Measures	Lead Time (h)							
	3	6	9	12	15	18	21	24
m_ϵ	0.0730	0.0671	0.1060	0.0646	0.0417	0.1160	0.0618	0.0888
std_ϵ	0.1449	0.1799	0.1836	0.1562	0.1647	0.2362	0.1982	0.2320
MAE (m)	0.0747	0.0892	0.1001	0.0888	0.1095	0.1174	0.0910	0.1001
RMSE (m)	0.1027	0.1277	0.1428	0.1283	0.1782	0.1779	0.1356	0.1548
bias (m)	0.0422	0.0405	0.0558	0.0224	-0.0102	0.0028	-0.0098	-0.0036
R^2	0.9951	0.9926	0.9900	0.9917	0.9855	0.9898	0.9919	0.9908
SI	0.0619	0.0810	0.0892	0.0848	0.1178	0.1173	0.0891	0.1011
MAPE (%)	10.9096	12.7095	14.7157	12.0080	12.5973	17.4128	13.5800	15.3229

Table 7. Forecast errors of H_{ss} based on the PBML model.

Error Measures	Lead Time (h)							
	3	6	9	12	15	18	21	24
m_ϵ	0.0238	-0.0057	-0.0234	0.0440	0.0780	0.0741	0.0941	0.0919
std_ϵ	0.1552	0.1636	0.1800	0.2319	0.3136	0.3333	0.2992	0.3328
MAE (m)	0.1129	0.1342	0.1652	0.2040	0.2584	0.2846	0.2730	0.3011
RMSE (m)	0.1758	0.2008	0.2404	0.3031	0.4013	0.4450	0.3848	0.4446
bias (m)	0.0129	-0.0179	-0.0383	0.0372	0.0670	0.0685	0.0661	0.0634
R^2	0.9456	0.9337	0.9087	0.8613	0.7821	0.7438	0.7573	0.7041
SI	0.1283	0.1440	0.1701	0.2156	0.2860	0.3199	0.2775	0.3228
MAPE (%)	9.9648	11.2799	13.1894	16.5625	20.8705	22.8518	22.6173	24.6826

Obviously, Figures 17 and 18 present similar findings to those illustrated in Figures 14 and 15, namely, that the forecast performance of the PBML model for wind-generated waves is better than that for swells. Specifically, the PBML model can generate H_{sw} forecasts accurately and the forecast uncertainty is quite low. For example, Table 6 shows that the RMSE and R^2 for all steps are less than 0.18 m and higher than 98%, respectively. In contrast, the RMSE and R^2 for H_{ss} forecasts are less than 0.45 m and higher than 70% for all steps respectively, as can be seen in Table 7.

In summary, limited knowledge on the source of swells persuaded us to stop at this stage, since merely increasing the control area and inputs would contribute little. In view of the swell being related to the waves propagating beyond the zones of their generation, the spatial domain of inputs is expected to be expanded, and more related input variables such as wave direction, speed and the distance from the position of the input to the output, are expected to be involved to improve the PBML model for T_p forecasting in future studies.

4. Conclusions

In this study, a physics-based machine learning (PBML) model was proposed and developed in detail to do multi-step-ahead forecasting of wave conditions for marine operations. The model attempts to combine the advantages of both physics-based wave models and machine learning algorithms. Specifically, the PBML model takes physical background into account by applying the primary inputs (i.e., initial wave boundary conditions and the future wind forcing) in physics-based models to design the model structure. A machine learning algorithm (the ANN considered in this study is an example) is employed to train the PBML model based solely on historical data.

The PBML model attempts to provide wave forecasts for marine operations, and high computational efficiency is the primary advantage compared with widely used physics-based wave models. On one hand, the PBML model is capable of using the wind forcing parameters in a small area or even from a single position as input to generate forecasts of wave statistics, which do not need the real-time simulations for large domains using directly physics-based numerical wave models. On the other hand, the training process of machine learning is performed mainly through matrix operations rather than solving partial differential control equations. Furthermore, once the model is trained based on the site-specific data, making forecasts on new inputs is straightforward and extremely fast. Based on the forecasted wave conditions and the environmental limits of a specific marine operation, suitable weather windows can be identified, which will further assist decision-making during the execution phase of the operation.

In this study, numerical wave models were used to generate both historical input and historical output data for training the machine learning algorithms. The developed algorithms are very specific for the location in consideration. However, the same methodology can be applied if the data from field measurements or the data from other locations are used.

To assess the performance of the proposed model, three-hourly time series of wave and wind data (such as wind direction D_w , mean wind speed U_w , significant wave height H_s and peak wave

period T_p) from 2001 to 2010 at the North Sea center from CERA-20C were used. As a new hybrid model, the optimal architecture of the PBML model was determined first. Afterwards, eight-step-ahead forecasts of H_s and T_p were generated individually and evaluated by commonly used error measures (MAE, RMSE, SI and R^2) and the newly defined forecast error factor. The results indicate that the forecast performance of H_s is extremely good, while that of T_p is slightly worse due to the weaker implicit relationship between the selected input and output; that was further proved by forecasting wave conditions of wind-generated waves and swells separately. Overall, the PBML model can be conceived as an efficient calculation model for wave forecasting in a small area, which is particularly suitable for marine operations. The forecast uncertainty of the peak period might be further decreased by including some more related factors, such as wave direction, speed and the distance from the position of the input to the output in the PBML model. A feature selection technique such as neighborhood component analysis (NCA) will be adopted to select and weight the related factors. Moreover, how to account for the effect of weather forecast uncertainty on marine operations is a key issue in practice. A suitable method for addressing wave forecast uncertainties in the planning and execution of marine operations can also be expected in the future publications.

Author Contributions: Methodology, M.W., C.S. and Z.G.; software, M.W.; validation, M.W.; writing—original draft preparation, M.W.; writing—review and editing, M.W., C.S. and Z.G.; supervision, Z.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: This work was supported by the Marine Operations in Virtual Environments (MOVE) and Centre for Autonomous Marine Operations and Systems (AMOS), at the Department of Marine Technology, NTNU, Trondheim, Norway. The support is gratefully acknowledged by the authors.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. DNV GL. *DNV-OS-H101: Marine Operations. General, Offshore Standard*; DNV: Oslo, Norway, 2011.
2. Group, T.W. The WAM model-A third generation ocean wave prediction model. *J. Phys. Oceanogr.* **1988**, *18*, 1775–1810. [[CrossRef](#)]
3. Uppala, S.M.; Kållberg, P.; Simmons, A.; Andrae, U.; Bechtold, V.D.C.; Fiorino, M.; Gibson, J.; Haseler, J.; Hernandez, A.; Kelly, G. The ERA-40 re-analysis. *Q. J. R. Meteorol. Soc. A J. Atmos. Sci. Appl. Meteorol. Phys. Oceanogr.* **2005**, *131*, 2961–3012. [[CrossRef](#)]
4. Saha, S.; Moorthi, S.; Pan, H.-L.; Wu, X.; Wang, J.; Nadiga, S.; Tripp, P.; Kistler, R.; Woollen, J.; Behringer, D. The NCEP climate forecast system reanalysis. *Bull. Am. Meteorol. Soc.* **2010**, *91*, 1015–1058. [[CrossRef](#)]
5. Dee, D.P.; Uppala, S.; Simmons, A.; Berrisford, P.; Poli, P.; Kobayashi, S.; Andrae, U.; Balmaseda, M.; Balsamo, G.; Bauer, d.P. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* **2011**, *137*, 553–597. [[CrossRef](#)]
6. Gusdal, Y.; Carrasco, A.; Furevik, B.; Sætra, Ø. *Validation of the Operational Wave Model WAM at met. no-Report 2010*; Report of the Norwegian Meteorological Institute-Oceanography no. 14/2011; Norwegian Meteorological Institute: Oslo, Norway, December 2011; Volume 8.
7. Hersbach, H.; Dee, D. ERA5 reanalysis is in production. *Ecmwf Newsl.* **2016**, *147*, 5–6.
8. Lavidas, G.; Venugopal, V. Application of numerical wave models at European coastlines: A review. *Renew. Sustain. Energy Rev.* **2018**, *92*, 489–500. [[CrossRef](#)]
9. Fusco, F.; Ringwood, J.V. Short-term wave forecasting with AR models in real-time optimal control of wave energy converters. In Proceedings of the 2010 IEEE International Symposium on Industrial Electronics, Bari, Italy, 4–7 July 2010; pp. 2475–2480.
10. Pena-Sanchez, Y.; Ringwood, J.V. A critical comparison of AR and ARMA models for short-term wave forecasting. In Proceedings of the 12th European Wave and Tidal Energy Conference, University College Cork, Cork, Ireland, 27 August–1 September 2017; pp. 961–966.

11. Ge, M.; Kerrigan, E.C. Short-term ocean wave forecasting using an autoregressive moving average model. In Proceedings of the 2016 UKACC 11th International Conference on Control (CONTROL), Belfast, UK, 31 August–2 September 2016; pp. 1–6.
12. Agrawal, J.; Deo, M. On-line wave prediction. *Mar. Struct.* **2002**, *15*, 57–74. [[CrossRef](#)]
13. Kamranzad, B.; Etemad-Shahidi, A.; Kazeminezhad, M. Wave height forecasting in Dayyer, the Persian Gulf. *Ocean Eng.* **2011**, *38*, 248–255. [[CrossRef](#)]
14. Adytia, D.; Yonanta, A.R.; Subasita, N. Wind Wave Prediction by using Autoregressive Integrated Moving Average model: Case Study in Jakarta Bay. *Int. J. Inf. Commun. Technol. (Ijoict)* **2019**, *4*, 33–42. [[CrossRef](#)]
15. Qin, M.; Li, Z.; Du, Z. Red tide time series forecasting by combining ARIMA and deep belief network. *Knowl.-Based Syst.* **2017**, *125*, 39–52. [[CrossRef](#)]
16. Deo, M.; Naidu, C.S. Real time wave forecasting using neural networks. *Ocean Eng.* **1998**, *26*, 191–203. [[CrossRef](#)]
17. Deo, M.C.; Jha, A.; Chaphekar, A.; Ravikant, K. Neural networks for wave forecasting. *Ocean Eng.* **2001**, *28*, 889–898. [[CrossRef](#)]
18. Malekmohamadi, I.; Bazargan-Lari, M.R.; Kerachian, R.; Nikoo, M.R.; Fallahnia, M. Evaluating the efficacy of SVMs, BNs, ANNs and ANFIS in wave height prediction. *Ocean Eng.* **2011**, *38*, 487–497. [[CrossRef](#)]
19. Mandal, S.; Rao, S.; Raju, D. Ocean wave parameters estimation using backpropagation neural networks. *Mar. Struct.* **2005**, *18*, 301–318. [[CrossRef](#)]
20. Peres, D.; Iuppa, C.; Cavallaro, L.; Cancelliere, A.; Foti, E. Significant wave height record extension by neural networks and reanalysis wind data. *Ocean Model.* **2015**, *94*, 128–140. [[CrossRef](#)]
21. Mandal, S.; Prabaharan, N. Ocean wave forecasting using recurrent neural networks. *Ocean Eng.* **2006**, *33*, 1401–1410. [[CrossRef](#)]
22. Sadeghifar, T.; Nouri Motlagh, M.; Torabi Azad, M.; Mohammad Mahdizadeh, M. Coastal wave height prediction using Recurrent Neural Networks (RNNs) in the south Caspian Sea. *Mar. Geod.* **2017**, *40*, 454–465. [[CrossRef](#)]
23. Ni, C.; Ma, X.; Bai, Y. Convolutional Neural Network Based Power Generation Prediction of Wave Energy Converter. In Proceedings of the 24th International Conference on Automation and Computing (ICAC), Newcastle University, Newcastle upon Tyne, UK, 6–7 September 2018.
24. Özger, M.; Şen, Z. Prediction of wave parameters by using fuzzy logic approach. *Ocean Eng.* **2007**, *34*, 460–469. [[CrossRef](#)]
25. Akpınar, A.; Özger, M.; Kömürçü, M.I. Prediction of wave parameters by using fuzzy inference system and the parametric models along the south coasts of the Black Sea. *J. Mar. Sci. Technol.* **2014**, *19*, 1–14. [[CrossRef](#)]
26. Stefanakos, C. Fuzzy time series forecasting of nonstationary wind and wave data. *Ocean Eng.* **2016**, *121*, 1–12. [[CrossRef](#)]
27. Wu, M.; Stefanakos, C.; Gao, Z.; Haver, S. Prediction of short-term wind and wave conditions for marine operations using a multi-step-ahead decomposition-ANFIS model and quantification of its uncertainty. *Ocean Eng.* **2019**, *188*, 106300. [[CrossRef](#)]
28. Basu, S.; Sarkar, A.; Satheesan, K.; Kishtawal, C. Predicting wave heights in the north Indian Ocean using genetic algorithm. *Geophys. Res. Lett.* **2005**, *32*. [[CrossRef](#)]
29. Ahmed, A.; Khalid, M. Multi-step ahead wind forecasting using nonlinear autoregressive neural networks. *Energy Procedia* **2017**, *134*, 192–204. [[CrossRef](#)]
30. Niu, T.; Wang, J.; Zhang, K.; Du, P. Multi-step-ahead wind speed forecasting based on optimal feature selection and a modified bat algorithm with the cognition strategy. *Renew. Energy* **2018**, *118*, 213–229. [[CrossRef](#)]
31. Laloyaux, P.; de Boisseson, E.; Balmaseda, M.; Bidlot, J.R.; Broennimann, S.; Buizza, R.; Dalhgren, P.; Dee, D.; Haimberger, L.; Hersbach, H. CERA-20C: A coupled reanalysis of the Twentieth Century. *J. Adv. Model. Earth Syst.* **2018**, *10*, 1172–1195. [[CrossRef](#)]
32. Tolman, H.L. A third-generation model for wind waves on slowly varying, unsteady, and inhomogeneous depths and currents. *J. Phys. Oceanogr.* **1991**, *21*, 782–797. [[CrossRef](#)]
33. Booij, N.; Ris, R.C.; Holthuijsen, L.H. A third-generation wave model for coastal regions: 1. Model description and validation. *J. Geophys. Res. Ocean.* **1999**, *104*, 7649–7666. [[CrossRef](#)]
34. Smith, J.; Sherlock, A.; Resio, D. *STWAVE: Steady-State Spectral Wave Model, User's Guide for STWAVE Version 3.0*, ERDC; CHL SR-01-01; US Army Engineer Research and Development Center: Vicksburg, MI, USA, 2001.

35. Rusu, E. Strategies in using numerical wave models in ocean/coastal applications. *J. Mar. Sci. Technol.* **2011**, *19*, 58–75.
36. Zijlema, M. Computation of wind-wave spectra in coastal waters with SWAN on unstructured grids. *Coast. Eng.* **2010**, *57*, 267–277. [[CrossRef](#)]
37. Rosenblatt, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **1958**, *65*, 386. [[CrossRef](#)]
38. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Cogn. Model.* **1988**, *5*, 1. [[CrossRef](#)]
39. Ramírez, L.; Fraile, D.; Brindley, G. *Offshore Wind in Europe: Key Trends and Statistics 2019*; Wind Europe: Brussels, Belgium, 2020.
40. Acero, W.G.; Li, L.; Gao, Z.; Moan, T. Methodology for assessment of the operational limits and operability of marine operations. *Ocean Eng.* **2016**, *125*, 308–327. [[CrossRef](#)]
41. Zhao, Y.; Cheng, Z.; Gao, Z.; Sandvik, P.C.; Moan, T. Numerical study on the feasibility of offshore single blade installation by floating crane vessels. *Mar. Struct.* **2019**, *64*, 442–462. [[CrossRef](#)]
42. Jiang, Z.; Gao, Z.; Ren, Z.; Li, Y.; Duan, L. A parametric study on the final blade installation process for monopile wind turbines under rough environmental conditions. *Eng. Struct.* **2018**, *172*, 1042–1056. [[CrossRef](#)]
43. Zhao, Y.; Cheng, Z.; Sandvik, P.C.; Gao, Z.; Moan, T. An integrated dynamic analysis method for simulating installation of single blades for wind turbines. *Ocean Eng.* **2018**, *152*, 72–88. [[CrossRef](#)]
44. Zhao, Y.; Cheng, Z.; Sandvik, P.C.; Gao, Z.; Moan, T.; Van Buren, E. Numerical modeling and analysis of the dynamic motion response of an offshore wind turbine blade during installation by a jack-up crane vessel. *Ocean Eng.* **2018**, *165*, 353–364. [[CrossRef](#)]
45. ECMWF. CERA-20C. Available online: <https://apps.ecmwf.int/datasets/data/cera20c/levtype=sfc/type=an/> (accessed on 10 February 2020).
46. Turgut, A.R. *Derivation of Century-Based Wave Climate and Extreme Wave Analysis along Turkish Coasts*; Izmir Institute of Technology: Urla, Turkey, 2019.
47. Ozbahceci, B.O.; Turgut, A.; Bozoklu, A.; Abdalla, S. Calibration and Verification of Century Based Wave Climate Data Record along the Turkish Coasts Using Satellite Altimeter Data. *Adv. Space Res.* **2020**, *66*, 2319–2337. [[CrossRef](#)]
48. Patra, A.; Bhaskaran, P.K. Temporal variability in wind–wave climate and its validation with ESSO-NIOT wave atlas for the head Bay of Bengal. *Clim. Dyn.* **2017**, *49*, 1271–1288. [[CrossRef](#)]
49. Heaton, J. *Introduction to Neural Networks with Java*; Heaton Research, Inc.: St. Louis, MO, USA, 2008.
50. Masters, T. *Practical Neural Network Recipes in C++*; Morgan Kaufmann: Burlington, MA, USA, 1993.
51. Bidlot, J.-R. *Ocean Wave Model Output Parameters*; ECMWF: Reading, UK, 2016.

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).