





## Article

# Modeling Significant Wave Heights for Multiple Time Horizons Using Metaheuristic Regression Methods

Rana Muhammad Adnan Ikram <sup>1</sup>, Xinyi Cao <sup>2,\*</sup>, Kulwinder Singh Parmar <sup>3</sup>, Ozgur Kisi <sup>4,5,6,\*</sup>, Shamsuddin Shahid <sup>6</sup> and Mohammad Zounemat-Kermani <sup>7</sup>

<sup>1</sup> School of Economics and Statistics, Guangzhou University, Guangzhou 510006, China; rana@gzhu.edu.cn

<sup>2</sup> College of Environmental Sciences, Sichuan Agricultural University, Chengdu 611130, China

<sup>3</sup> Department of Mathematical Sciences, IKG Punjab Technical University, Jalandhar 144603, India; dr.kulmaths@ptu.ac.in

<sup>4</sup> Department of Civil Engineering, Lübeck University of Applied Science, 23562 Lübeck, Germany

<sup>5</sup> Department of Civil Engineering, School of Technology, Ilia State University, 0162 Tbilisi, Georgia

<sup>6</sup> School of Civil Engineering, Faculty of Engineering, Universiti Teknologi Malaysia (UTM), Johor Bahru 81310, Malaysia; sshahid@utm.my

<sup>7</sup> Department of Water Engineering, Shahid Bahonar University of Kerman, Kerman 76169-14111, Iran; zounemat@uk.ac.ir

\* Correspondence: caoxinyi@hhu.edu.cn (X.C.); ozgur.kisi@th-luebeck.de or ozgur.kisi@iliauni.edu.ge (O.K.)

**Abstract:** The study examines the applicability of six metaheuristic regression techniques—M5 model tree (M5RT), multivariate adaptive regression spline (MARS), principal component regression (PCR), random forest (RF), partial least square regression (PLSR) and Gaussian process regression (GPR)—for predicting short-term significant wave heights from one hour to one day ahead. Hourly data from two stations, Townsville and Brisbane Buoys, Queensland, Australia, and historical values were used as model inputs for the predictions. The methods were assessed based on root mean square error, mean absolute error, determination coefficient and new graphical inspection methods (e.g., Taylor and violin charts). On the basis of root mean square error (RMSE), mean absolute error (MAE), and coefficient of determination ( $R^2$ ) statistics, it was observed that GPR provided the best accuracy in predicting short-term single-time-step and multi-time-step significant wave heights. On the basis of mean RMSE, GPR improved the accuracy of M5RT, MARS, PCR, RF and PLSR by 16.63, 8.03, 10.34, 3.25 and 7.78% (first station) and by 14.04, 8.35, 13.34, 3.87 and 8.30% (second station) for the test stage.

**Keywords:** significant wave height; short-term prediction; Gaussian process regression; partial least square regression

**MSC:** 90C59; 90C90



**Citation:** Ikram, R.M.A.; Cao, X.; Parmar, K.S.; Kisi, O.; Shahid, S.; Zounemat-Kermani, M. Modeling Significant Wave Heights for Multiple Time Horizons Using Metaheuristic Regression Methods. *Mathematics* **2023**, *11*, 3141. <https://doi.org/10.3390/math11143141>

Academic Editors: Siddhartha Bhattacharya and Mohamed Abd Elaziz

Received: 5 June 2023

Revised: 5 July 2023

Accepted: 12 July 2023

Published: 16 July 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

Accurate measurement of wave properties is crucial for designing coastal and offshore structures, undertaking maritime projects, estimating sediment transport, and performing other coastal-engineering-related tasks. Winds acting on the ocean's surface are the single most important factor in determining wave heights; however, other influential variables such as ocean currents, environmental changes, and earth systems also affect them [1,2]. Hence, proper estimation of wave properties is a challenge in coastal and offshore engineering. Various factors represent the wave characteristics, of which significant wave height (HSW) is the most important property [2–4].

Several methods exist for short- and mid-term prediction of HSW, including wave energy balance-based models [3], numerical models [4], chaos-theory-based models [5], empirical models [6], time-series and stochastic models [7,8], and soft computing and machine

learning methodologies [9,10]. The spectral energy or action balance equation is the foundation for some widely used models, including empirical and numerical models. However, their implementation poses several challenges due to their complexity, extensive computational cost, and precise local bathymetric data requirements, which are the main drawbacks of these methodologies. In the field of chaos, solitons and fractals, several researchers have conducted mathematical modeling to understand complex dynamical systems [11,12].

It is worth noting that artificial intelligence systems are powerful tools for modeling intricate relationships and identifying patterns in complex datasets. In a relevant study, Pourzangbar et al. [13] demonstrated that the prediction of scour depth due to non-breaking waves with the aid of Machine Learning (ML) models such as regression trees and support vector regression achieves the highest accuracy. Therefore, soft computing and ML methods, which fall under artificial intelligence methodologies, have been successfully utilized in modeling wave characteristics and wave heights [14–16]. Besides the beneficial features of the ML methodologies, it should be mentioned that the majority of ML models do not offer a simple, explicit description of the mathematical structure of the constructed models. This trait can sometimes be considered a shortcoming of the ML models.

Applications of ML models in ocean and coastal engineering have received a great deal of attention recently [17]. This area includes several objectives, such as prediction of wave height [18], prediction of water level and tides [19], breakwater simulation [13], and ocean current simulation [20]. In the realm of HSW modeling, which is the main topic of this study, Mahjoobi and Mosabbeh [21] utilized machine learning models, such as a regressive support vector machine (SVM) model, a multilayer perceptron (MLP) neural network, and a radial basis function (RBF) neural network, to predict short-term HSW based on waves and wind in Lake Michigan. The study revealed that all ML models successfully predicted HSW ( $R > 0.93$ ), but the SVM with an RBF kernel provided the most accurate predictions ( $R = 0.96$ ). In a separate study, Krishna Kumar et al. [22] developed two types of hybrid sequential learning machine models, namely the minimal resource allocation network (MRAN) and the growing and pruning radial basis function (GAP-RBF) network, to forecast wave heights at 13 sites. Results revealed more accurate predictions using MRAN and GAP-RBF than traditional ML models. Additionally, Kaloop et al. [23] employed a hybrid model, wavelet-particle swarm optimization (PSO)—extreme learning machine (ELM), to predict HSW. They utilized wavelet analysis for the frequency content analysis of wave signals and then applied particle swarm optimization (PSO) to train the ELM model. The relative performance of several regular ML models indicates the better ability of the developed hybrid model to simulate HSW for short-term (hourly) and mid-term (daily) lead times. Table 1 summarizes the applications of ML and soft computing techniques in predicting HSW. An overview of the literature in Table 1 demonstrates that most researchers applied network-based ML models for predicting HSW, such as artificial neural networks (ANNs), ELMS, recurrent neural networks (RNNs), and long short-term memory (LSTM). In contrast, the SVM/SVR and the tree-based/regression-based models have been used less frequently.

The complex and nonlinear nature of HSW time series has a tremendous impact on forecasting and predicting accuracy [24]. Although ML models are powerful and capable tools for modeling wave height, the structure of most network-based ML models, like ANNs and adaptive network-based fuzzy inference systems (ANFIS), is not as transparent as regression-based ML models. Furthermore, network-based ML models necessitate trial and error to determine network hyperparameters, like hidden-layer and neuron numbers, which is time-consuming [25,26]. It is worth noting that other ML models, such as genetic programming and gene expression programming methods, along with regression-based models (e.g., model trees), could be a beneficial alternative for modeling nonlinear phenomena such as HSW [27].

The prediction/forecasting accuracy of HSW worsens as the lead time expands. Hence, having different prediction strategies based on various forecasting lead times can greatly enhance the models' capability and reliability. This study focuses on the implications

of regression-based models for achieving this purpose. In this regard, several models are employed to predict HSW, including tree-based methodologies (e.g., multivariate adaptive regression spline (MARS), random forest (RF), and M5 model tree (M5RT) and the statistics-based partial least square regression (PLSR) and Gaussian process regression (GPR)). The main reasons for opting for these ML models lie in their high capabilities for modeling complicated phenomena and their diverse architecture, which makes them good candidates for reaching a general perspective in choosing more appropriate ML structures for simulating wave height. The modeling strategies are based on lagged HSW data and forecasting HSW multiple lead times. This study appraises the prospects and competency of six regression-based ML models in predicting (viz., forecasting) HSW for different hourly lead times up to 24 h.

**Table 1.** Literature review—applications of ML models for predicting significant wave height.

Researcher(s)	Models Applied	Target Parameter/ Prediction Interval	Remarks
Londhe and Panchang [28]	Feed-forward artificial neural network (ANN)	HSW/mid-term (daily)	The model's ability to identify interannual variability and provide more reliable forecasts was shown to be critical; this was achieved by considering several years of data and carefully selecting the training set.
Mahjoobi and Etemad-Shahidi [29]	Classification and regression trees (CART) and ANNs	HSW/short-term (hourly)	Results indicated that the CART model could be used successfully to predict HSW.
Mahjoobi and Mosabbeq [21]	Regressive support vector machines, MLP and RBF neural networks.	HSW/short-term (hourly)	The cross-validation and non-cross-validation results showed that the SVMs (RBF kernel and polynomial kernel) performed slightly better than ANNs.
Etemad-Shahidi and Mahjoobi [30]	M5 model tree and ANN	HSW/short-term (hourly)	The results implied that the M5 tree is marginally better than the ANN's.
Özger [7]	Wavelet fuzzy logic (WFL), ARIMA, and ANN	HSW/short-term to mid-term (hourly)	The WFL outperformed other models. Also, its performance improved with longer lead times.
Altunkaynak [31]	Geno-multilayer perceptron ANN	HSW/Short-term (15 min)	Good consistency was reported between the observed and predicted results from the geno-multilayer perceptron model.
Salcedo-Sanz et al. [32]	Support vector regression	HSW/Short-term (hourly)	The SVR model provided good results in HSW estimation from sea-surface X-band radar images.
Duan et al. [33]	Support vector regression	HSW/Short-term (hourly)	The statistical indices showed that the empirical model decomposition-SVR provided proper short-term prediction performance.
Cornejo-Bueno et al. [34]	Grouping genetic algorithm and extreme learning machine (GGA-ELM), GGA-MLP, and SVR	HSW/Short-term (hourly)	The outcomes of the models showed that the proposed GGA-ELM improved the predicted results.

Table 1. Cont.

Researcher(s)	Models Applied	Target Parameter/ Prediction Interval	Remarks
Berbić et al. [35]	ANNs and SVMs	HSW/Short-term (half-hourly)	HSW was predicted for 0.5 to 5.5 h lead time. The SVM performed slightly better than the ANN.
Nikoo et al. [36]	Fuzzy KNN-based, SVR, regression tree induction (M5P), Bayesian network model	HSW/Short-term (hourly)	The Fuzzy KNN-based model performed better than other applied ones, followed by the M5P model.
Ali and Prasad [37]	Hybridized extreme learning machine	HSW/Short-term (hourly)	The hybrid ELM model exhibited high accuracy in predicting HSW and was identified as a promising tool.
Kaloop et al. [23]	Wavelet—particle swarm optimization—extreme learning machine	HSW/short-term (hourly) and mid-term (daily)	The study suggested that WPSO-ELM performed better than ANN, fuzzy logic, SVM, MLR, and ELM in predicting HSW for different lead times.
Demetriou et al. [38]	Ensemble tree modeling and ANN	HSW/short-term (minutely)	The decision tree ensemble provided better outcomes.
Feng et al. [39]	Recurrent neural network (RNN), long short-term memory network (LSTM), and gated recurrent unit network (GRU)	HSW/(hourly)	The findings indicated that GRU and LSTM networks outperformed traditional RNNs.
Gao et al. [40]	Hybrid-ensemble deep randomized networks	HSW/(hourly, four-hourly)	The neuron pruning strategy for removing noisy information from the random features enhanced the prediction accuracy of the ML model.
Minuzzi and Farina [10]	Long short-term memory algorithm (LSTM)	HSW/(hourly)	The findings indicated that a data-driven approach can serve as a substitute for computationally intensive physical models, with an accuracy level of around 95% compared to reanalysis data and 87% in comparison to buoy data.

## 2. Case Study

Two wave monitoring sites in Queensland, Australia were selected for predicting HSW. The Townsville wave monitoring buoy is located in South East Queensland at coordinates of 19°09.550' S latitude and 147°03.560' E longitude, and the Brisbane wave monitoring buoy is situated in North Queensland, with a geographical location of 27°29.230' S latitude and 153°37.900' E longitude (Figure 1). The Queensland environmental department monitored HSW data at both wave monitoring sites for 30 min intervals from 1 January 2022 to 31 December 2022. The data were collected for both wave monitoring sites from the online portal of the Queensland government (<https://www.data.qld.gov.au/dataset>, accessed on 20 February 2023). For single and multistep wave height prediction, only hourly data for significant wave height (HSW) were used for both stations. For the applications of models, all of the HSW variable data were divided into training and testing parts. However, the selection of a more critical training and testing dataset from the data is a key step. Therefore, a 4-fold cross-validation technique was used as a preliminary step for the selection of the training and testing dataset using the MARS model. For this purpose, the full data set was

divided into four equal test datasets parts, and each part was tested against the remaining training dataset. After the preliminary results of the 4-fold cross-validation technique, hourly HSW data for 1 January 2022 to 5 October 2022, were utilized as training data, whereas the hourly HSW data from 6 October 2022 to 31 December 2022, were utilized as the testing dataset. Table 2 shows the basic HSW statistics for the stations.



Figure 1. Station locations.

Table 2. Statistical parameters of the applied data.

	Mean	Min.	Max	Skewness	Std. Dev.
Station 1					
Full Dataset	1.8795	0.4380	5.9680	1.0417	0.7914
Training Dataset	1.9179	0.4380	5.9680	0.8876	0.7788
Testing Dataset	1.7638	0.5290	5.3090	1.5229	0.8174
Station 2					
Full Dataset	0.6726	0.1060	2.2850	0.7850	0.3552
Training Dataset	0.6786	0.1060	2.2850	0.7329	0.3494
Testing Dataset	0.6526	0.1060	2.0640	0.9499	0.3731



### 3. Methods

#### 3.1. Multivariate Adaptive Regression Splines (MARS)

The multivariate adaptive regression spline (MARS) model is a common prediction algorithm. Friedman [41] developed and introduced this model in 1991. Subsequently, several amendments were made to boost the model’s quality and reduce the prediction error. The remarkable enhancement of the MARS method lies in its capability of input-output mapping without any specific assumptions. Unlike other prediction models, which require assumptions for mapping, MARS considers segment endpoints as nodes. As a nonparametric regression model, it enables time-series predictions. Because of the model’s adaptability, it can handle additive interactions or include connections to other model inputs. MARS can handle the compound mapping of the predictor and response easily. It can also use both forward and backward stepwise measures. Andres et al. [42] proposed a backward stepwise procedure to eliminate preventable variables and enhance the model’s capability. In addition, the stepwise forward procedure can be used to determine which variables serve as inputs to the MARS model. Unlike previous prediction models that consider segment terminals as nodes, the MARS model does not require any assumptions about the mapping between input and output variables. This nonparametric regression model can predict continuous numerical outcomes and has adaptable procedures for handling relationships that are almost additive or include relations with other input variables.

Using two basis functions with a variety of inputs,  $Y$  is mapped from  $X$  using  $c$  as the threshold value:

$$Y = \max(0, X - c) \tag{1}$$

$$Y = \max(0, c - X) \tag{2}$$

MARS uses two adjacent splines to ensure the continuity of the basis function at the lump. The MARS algorithm has been extensively employed in various research fields due to its ability to predict accurately. This study utilized MARS to predict HSW at the sample sites. Figure 2 shows the flow chart of the MARS model used in this study.

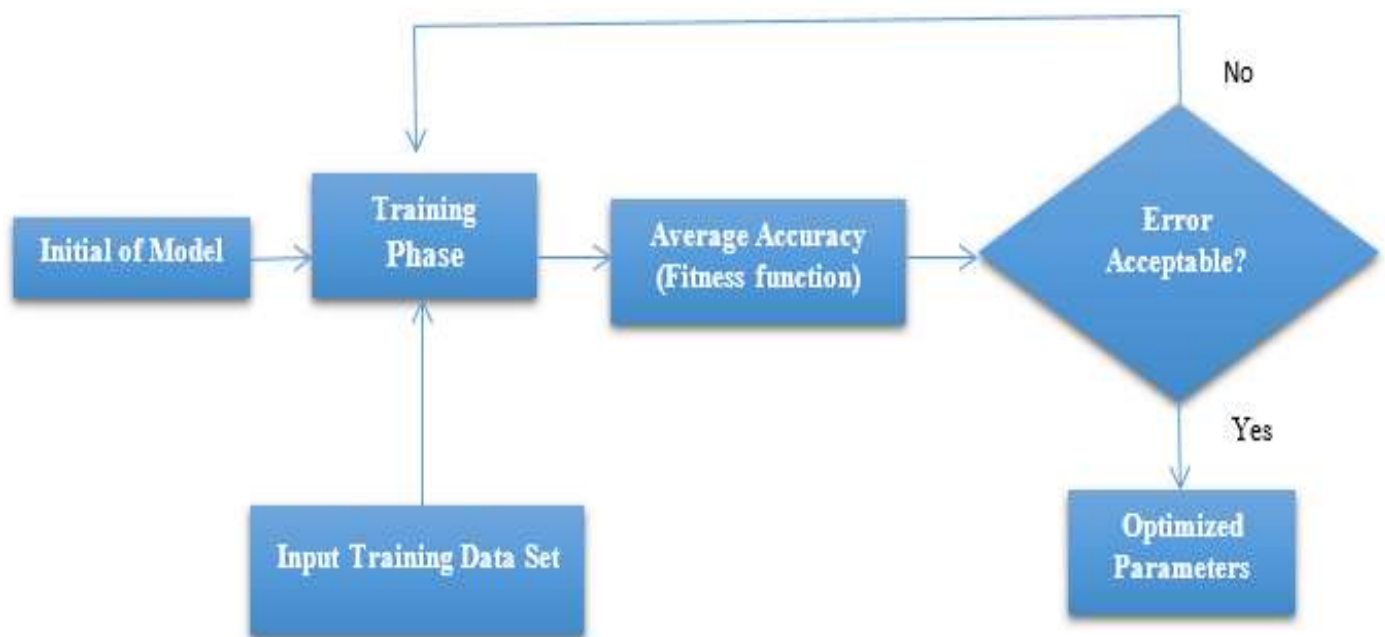


Figure 2. Flow chart for the MARS model.

### 3.2. Gaussian Process Regression (GPR)

GPR is crucial in defining the initial distribution for a versatile regression model and classification. It is a nonparametric method used to identify the presence of hesitation in the model to further improve its capability [43]. The GPR model is a crucial supervised learning tool to use after Bayesian machine learning models, because it performs well with small datasets without underfitting or overfitting. Additionally, it excels at analyzing nonlinear datasets by utilizing probabilistic approaches and approximating posterior degradation by requiring prior data distribution. The diversity of the covariance function is an essential component of the Gaussian process, which helps to develop tasks with different structures [44–46]. Also, the standard property of the GPR model makes this model vital for statistical modeling. The covariance is defined in the Gaussian process with the mean function  $m(x)$ .

$$m(x) = E(f(x)) \tag{3}$$

$$k(x, x') = E((f(x) - m(x))(f(x') - m(x')))) \tag{4}$$

Here,  $k(x, x')$  represents the kernel or covariance function evaluated at the points  $x$  and  $x'$ . The Gaussian process function is expressed as:

$$f(x) \sim GP(m(x), k(x, x')) \tag{5}$$

The average function value is 0, and the following form can recognize the input-target relationship.

$$y = f(x) + \delta \tag{6}$$

Here,  $\delta$  is Gaussian white noise uncorrelated with  $f(x)$ , with a mean value of 0 and a variance of  $\sigma^2$ . Also,  $y$  and  $f(x)$  follow Gaussian distribution, and the set of the joint distribution of finite observations is a Gaussian process:

$$y \sim GP(m(x), k(x, x') + \sigma^2 \gamma_{ij}) \tag{7}$$

Here,  $\gamma_{ij}$  is the Kronecker delta function. Also, it is assumed that  $y = [y_1, y_2, y_3, \dots, y_n]^T$  and  $f = [f(x_1), f(x_2), f(x_3), \dots, f(x_n)]$  depend upon the Gaussian process

$$p(f) = N(0, K) \tag{8}$$

where  $k$  is the covariance matrix with the elements:

$$k = \begin{bmatrix} k(x_1, x_1) & k(x_1, x_2) & \dots & k(x_1, x_n) \\ k(x_2, x_1) & k(x_2, x_2) & \dots & k(x_2, x_n) \\ & & \dots & \\ & & & \dots \\ k(x_n, x_1) & k(x_n, x_2) & \dots & k(x_n, x_n) \end{bmatrix} \tag{9}$$

Here,  $k_i$  is the covariance between the values of the eigenfunctions  $f(x_i)$  and  $f(x_j)$ . Gaussian process regression (GPR) is used to calculate the predicted distribution for the function values of the  $f^*$  at the test point  $X^* = [x^*_1, x^*_2, \dots, x^*_m]$ . GPR gives the marginal distribution:

$$p(f) = N(f|0, K) \tag{10}$$

$$K_y = K + \sigma^2 I \tag{11}$$

$$\begin{bmatrix} y \\ y^* \end{bmatrix} = \left( \begin{bmatrix} f \\ f^* \end{bmatrix} + \begin{bmatrix} \delta \\ \delta^* \end{bmatrix} \right) N\left(0, \begin{bmatrix} K_y & K^* \\ K_n^T & K^{**} + \sigma^2 \end{bmatrix}\right) \tag{12}$$

for the eigenfunction for the variable, and

$$K^* = [k(x^*, x_1) \ k(x^*, x_2) \ \dots \ k(x^*, x_n)]^T \tag{13}$$

$$K^{**} = k(x^*, x^*) \tag{14}$$

Using the Gaussian condition laws, the predictive distribution  $p(y^*|y)$  is a Gaussian distribution with the variance and mean  $m(x^*) = k^{*T}k_y^{-1}y$  [47]. The determination of the covariance is an essential component in the Gaussian process regression.

3.3. Random Forest (RF)

Multiple decision trees are the basis of an RF, which can be used for both classification and regression. Researchers have turned to the bagging technique to create forecasting training sets that are independent of one another [48–52]. The output of the RF is based on an average of the predictions made with each tree during the regression phase [41,42]. The fundamental procedures of this paradigm are as follows:

- (a) First, k decision tree models are constructed using the bootstrap sampling approach, where k random samples are drawn from the original training set.
- (b) A split feature set of n (nm, where m is the total number of features in the sample) features is chosen at random from each sample. Nodes are generated after evaluating the optimum characteristics, with the minimal Gini coefficient serving as the dividing line.
- (c) The trees reach their full potential without trimming. The procedures above yield random forests when repeated.
- (d) A democratic vote on each record determines the final results based on the k categorization outcomes. It is helpful to look at the average decline in the Gini coefficient at the node split to determine the classification.

Figure 3 shows the working procedure of the RF model.

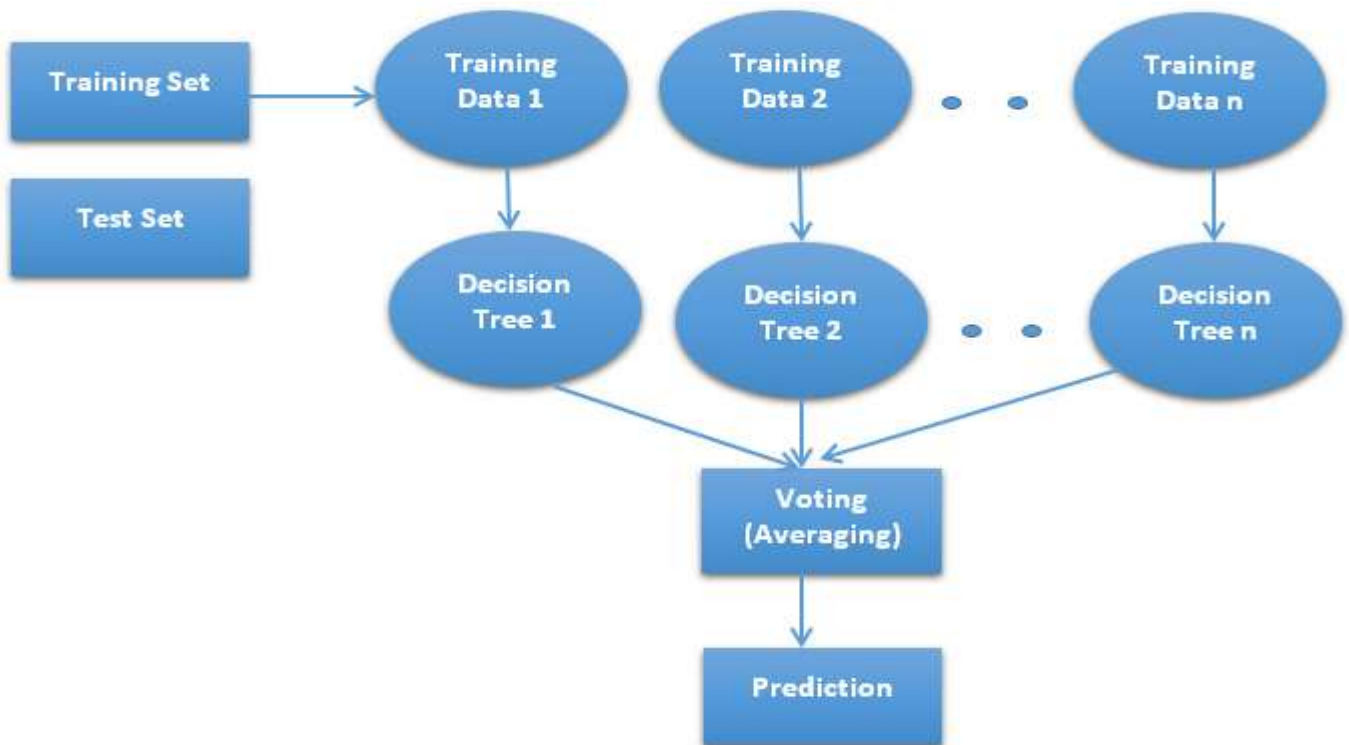


Figure 3. Flow chart for the RF model.



### 3.4. Principal Component Regression (PCR)

The PCR provides a linear regression model relating inputs with the output. Suppose the input dataset  $X \in R^{n \times m}$  has a sample size  $n$  and a number of variables  $m$ . Also,  $Y \in R^{n \times l}$  is the output data set with  $n$  observations, and  $l$  is the quality variable [53–55].

Singular value decomposition (SVD) is applied to the covariance matrix  $\frac{X^T X}{n-1}$  to find the loading and eigenvalue matrices  $\Lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_m)$ .

$$\frac{X^T X}{n-1} = \begin{bmatrix} \widehat{P} & \widetilde{P} \end{bmatrix} \begin{bmatrix} \Lambda_{pc} & 0 \\ 0 & \Lambda_{res} \end{bmatrix} \begin{bmatrix} \widehat{P}^T \\ \widetilde{P}^T \end{bmatrix} \approx \widehat{P} \Lambda_{pc} \widehat{P}^T \tag{15}$$

Here, the process variables are divided into residual subspace and subspace.

$$X = \widehat{X} + \widetilde{X} = \widehat{T} \widehat{P}^T + \widetilde{X} \tag{16}$$

Equation  $Y^T \approx \phi X^T$  multiplies  $\frac{X}{n-1}$  by both sides, resulting in  $\frac{Y^T X}{n-1} \approx \phi \frac{X^T X}{n-1}$ . The regression coefficient is expressed as:

$$\phi = \frac{Y^T X \widehat{P} \Lambda_{pc}^{-1} \widehat{P}^T}{n-1} = \frac{Y^T \left( \Lambda_{pc}^{-\frac{1}{2}} \widehat{P}^T X^T \right)^T \Lambda_{pc}^{-\frac{1}{2}} \widehat{P}^T}{n-1} \tag{17}$$

$$\bar{\phi} = \Lambda_{pc}^{-\frac{1}{2}} \widehat{P}^T \tag{18}$$

The PCR regression model is obtained as:

$$\widehat{y} = \phi x = \frac{Y^T \left( \Lambda_{pc}^{-\frac{1}{2}} \widehat{P}^T X^T \right)^T \Lambda_{pc}^{-\frac{1}{2}} \widehat{P}^T}{n-1} x = \bar{\phi} \bar{x} \tag{19}$$

where  $\bar{x} = \Lambda_{pc}^{-\frac{1}{2}} \widehat{P}^T x$  is the normalized input variable and  $\bar{\phi}$  is the corresponding regression coefficient vector. The singular value decomposition (SVD) is again applied but now on the  $\bar{\phi}$ . The decomposed findings are given below.

$$\bar{\phi} = \frac{Y^T \left( \Lambda_{pc}^{-\frac{1}{2}} \widehat{P}^T X^T \right)^T}{n-1} = \widehat{U} \begin{bmatrix} D^{1/2} & 0 \end{bmatrix} \begin{bmatrix} \widehat{V}^T \\ \widetilde{V}^T \end{bmatrix} = \widehat{U} D^{1/2} \widehat{V}^T \tag{20}$$

The projection  $x$  onto  $y$  is shown as:

$$\bar{y} = \widehat{V}^T \bar{x} = \widehat{V}^T \Lambda_{pc}^{-\frac{1}{2}} \widehat{P}^T x \sim N(0, I_{l \times l}) \tag{21}$$

where  $\bar{y}$  is a key variable vector for quality. The prediction  $y$  is as follows:

$$\widehat{y} = \bar{\phi} \bar{x} = \widehat{U} D^{1/2} \widehat{V}^T \Lambda_{pc}^{-\frac{1}{2}} \widehat{P}^T x = \widehat{U} D^{1/2} \bar{y} \tag{22}$$

The PCR model’s prediction results are applied to predict short-term significant wave height.

### 3.5. Partial Least Square Regression (PLSR)

PLSR is developed by merging two models, the principal component analysis and multiple linear regression models. It finds the linear transformation of the input features, which are highly covariant with the response variable and are also uncorrelated with themselves [56]. Such variables are known as latent components. This model employs the regression method to predict the response value and rebuild the original data matrix using latent features. The objective function is designed to maximize the covariance of this latent component with the response variable. The cross-validation approach minimizes prediction error [56–58]. These latent variables are called X-scores ( $t_a$ ) and are Y predictors:

$$X = TP' + E \tag{23}$$

where  $T$  is the score matrix, columns are  $t_a$ ,  $P'$  is the loading matrix, and  $E$  is the matrix of the X-residuals.

$$T = XW^* \tag{24}$$

Here,  $W^*$  is a matrix of transformed PLSR weights.

$$Y = TC^* + F \tag{25}$$

$C^*$  is the Y-weight matrix, and  $F$  is the matrix of the Y-residuals.

### 3.6. M5 Regression Tree (M5RT)

The M5 regression tree model uses a decision tree to establish a connection between inputs and output. Quinlan first proposed this method [59] and it was later refined by Wang and Witten [60]. There are two classes in the M5 tree model. The first class sorts input variables into categories according to linear regression in an attempt to minimize the approximation error between observed and forecasted values. This initial process constructs the decision tree based on the input evidence and uses the standard deviation drop to establish the division rule for the M5 tree model [61,62]. The second set consists of the compiled trees from each leaf. Each age involves more granular categorization, down to the level of individual branches and leaves. The M5 tree model was developed from the popular classification and regression tree (CART) technique. Following is a detailed explanation of the standard deviation (SDR) formula.

$$SDR = sd(T) - \sum \frac{|T_i|}{|T|} sd(T_i) \tag{26}$$

where  $T_i$  is the  $i$ -th result of the possible group, SDR is the standard deviation reduction,  $sd$  is the standard deviation,  $T$  is a collection of samples reaching the node, and  $i$  is an integer. Overfitting occurs when the data's  $sd$  is less than that of the parent nodes, leading to a massive, tree-like structure that doesn't generalize effectively.

## 4. Accuracy Assessment

The main aim of this study was to compare several metaheuristic regression methods, M5RT, MARS, PCR, RF, PLSR and GPR, in predicting HSW. Root mean square error (RMSE), mean absolute error (MAE), and determination coefficient ( $R^2$ ) criteria were used to assess the methods employed.

$$RMSE : \text{Root Mean Square Error} = \sqrt{\frac{1}{N} \sum_{i=1}^N [(HSW_o)_i - (HSW_C)_i]^2} \tag{27}$$

$$MAE : \text{Mean Absolute Error} = \frac{1}{N} \sum_{i=1}^N |(HSW_o)_i - (HSW_C)_i| \tag{28}$$

$$R^2 : \text{Determination Coefficient} = \left[ \frac{\sum_{t=1}^N (HSW_o - \overline{HSW_o})(HSW_c - \overline{HSW_c})}{\sqrt{\sum_{t=1}^N (HSW_o - \overline{HSW_o})^2 (HSW_c - \overline{HSW_c})^2}} \right]^2 \quad (29)$$

where  $HSW_c$ ,  $HSW_o$ ,  $\overline{HSW_o}$ ,  $N$  are calculated, observed, and mean significant wave height and the number of data, respectively.

### 5. Results and Discussion

This section compares the six metaheuristic regression approaches in predicting HSW for multiple horizons from  $t + 1$  (one hour ahead) to  $t + 24$  (one day ahead). The following section explains the results in detail.

#### 5.1. Results

Before application of the models, the data were divided into two parts, training (80%) and testing (20%), and then simulations were performed in the MATLAB environment. This study used MATLAB statistics and ML toolbox for the simulations. The training and testing results for the M5RT, MARS, PCR, RF, PLSR and GPR models are provided in Tables 3–8 for the first station. In the tables,  $t + 1$  indicates 1 h ahead, or a 1 h time horizon, and  $HSW_t$  refers to significant wave height at time  $t$  (current hour). This study used three lags as inputs to the regression models because, beyond this, lag did not produce considerably better accuracy. It is visible from Tables 3–8 that all regression models had the best accuracy in the prediction of  $t + 1$  HSWs with three lagged inputs ( $HSW_t$ ,  $HSW_{t-1}$ ,  $HSW_{t-2}$ ). Therefore, this study used this input case to predict HSW for the higher horizons from  $t + 2$  (2 h ahead) to  $t + 24$  (1 day ahead). As expected, the models' accuracies decreased with increasing time horizons, because the relationship between inputs and output is more complex in the case of farther horizons. For example, the RMSE and MAE of M5RT increased from 0.1254 m and 0.0860 m, respectively, to 0.5929 m and 0.4224 m, and  $R^2$  decreased from 0.9769 to 0.5022 in the test stage. A comparison of the metaheuristic regression models revealed that GPR provided the best accuracy in predicting HSW 1 h ahead with the lowest RMSE (0.1115 m) and MAE (0.0771 m) and the highest  $R^2$  (0.9817) in the test stage. The relative RMSE differences between GPR and M5RT, MARS, PCR, RF and PLSR were 11.08, 6.38, 2.02, 1.24 and 0.27%, respectively. There was a slight difference between GPR and PLSR models in predicting HSW for  $t + 1$  horizon.

Table 3. Model performance in training and testing for multiple-step HSW predictions—M5RT for Station 1.

Time Horizon	Input Combination	Training Period			Test Period		
		RMSE (m)	MAE (m)	$R^2$	RMSE (m)	MAE (m)	$R^2$
$t + 1$	$HSW_t$	0.1270	0.0915	0.9755	0.1296	0.0895	0.9733
	$HSW_t, HSW_{t-1}$	0.1190	0.0850	0.9765	0.1284	0.0896	0.9746
	$HSW_t, HSW_{t-1}, HSW_{t-2}$	0.1176	0.0839	0.9771	0.1254	0.0860	0.9769
$t + 2$	$HSW_t, HSW_{t-1}, HSW_{t-2}$	0.1356	0.0930	0.9726	0.1390	0.0982	0.9680
$t + 4$	$HSW_t, HSW_{t-1}, HSW_{t-2}$	0.1786	0.1241	0.9504	0.2010	0.1385	0.9402
$t + 8$	$HSW_t, HSW_{t-1}, HSW_{t-2}$	0.2579	0.1787	0.8900	0.3058	0.2098	0.8656
$t + 12$	$HSW_t, HSW_{t-1}, HSW_{t-2}$	0.3156	0.2211	0.8353	0.3941	0.2701	0.7764
$t + 24$	$HSW_t, HSW_{t-1}, HSW_{t-2}$	0.4451	0.3177	0.6723	0.5939	0.4224	0.5022

**Table 4.** Model performance in training and testing for multiple-step HSW predictions—MARS for Station 1.

Time Horizon	Input Combination	Training Period			Test Period		
		RMSE (m)	MAE (m)	R <sup>2</sup>	RMSE (m)	MAE (m)	R <sup>2</sup>
t + 1	HSW <sub>t</sub>	0.1245	0.0901	0.9773	0.1233	0.0853	0.9744
	HSW <sub>t</sub> , HSW <sub>t-1</sub> ,	0.1167	0.0837	0.9784	0.1204	0.0833	0.9775
	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1161	0.0831	0.9788	0.1191	0.0826	0.9777
t + 2	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1352	0.0927	0.9727	0.1384	0.0980	0.9683
t + 4	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1770	0.1226	0.9517	0.1934	0.1321	0.9442
t + 8	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.2538	0.1759	0.8934	0.2798	0.1927	0.8833
t + 12	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.3098	0.2167	0.8413	0.3535	0.2456	0.8149
t + 24	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.4409	0.3143	0.6785	0.5007	0.3619	0.6280

**Table 5.** Model performance in training and testing for multiple-step HSW predictions—PCR for Station 1.

Time Horizon	Input Combination	Training Period			Test Period		
		RMSE (m)	MAE (m)	R <sup>2</sup>	RMSE (m)	MAE (m)	R <sup>2</sup>
t + 1	HSW <sub>t</sub>						
	HSW <sub>t</sub> , HSW <sub>t-1</sub> ,	0.1140	0.0803	0.9778	0.1182	0.0818	0.9786
	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1093	0.0794	0.9808	0.1138	0.0795	0.9781
t + 2	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1349	0.0924	0.9728	0.1382	0.0978	0.9684
t + 4	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1745	0.1205	0.9537	0.1823	0.1258	0.9472
t + 8	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.2432	0.1704	0.9021	0.2735	0.1857	0.8884
t + 12	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.3017	0.2107	0.8494	0.3412	0.2356	0.8264
t + 24	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.4349	0.3086	0.6872	0.4741	0.3388	0.6564

**Table 6.** Model performance in training and testing for multiple-step HSW predictions—RF for Station 1.

Time Horizon	Input Combination	Training Period			Test Period		
		RMSE (m)	MAE (m)	R <sup>2</sup>	RMSE (m)	MAE (m)	R <sup>2</sup>
t + 1	HSW <sub>t</sub>	0.1213	0.0876	0.9783	0.1210	0.0841	0.9756
	HSW <sub>t</sub> , HSW <sub>t-1</sub> ,	0.1069	0.0757	0.9802	0.1146	0.0789	0.9791
	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1037	0.0742	0.9811	0.1129	0.0781	0.9793
t + 2	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1330	0.0915	0.9736	0.1368	0.0969	0.9691
t + 4	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1726	0.1207	0.9539	0.1799	0.1247	0.9482
t + 8	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.2413	0.1689	0.9037	0.2689	0.1824	0.8921
t + 12	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.3011	0.2101	0.8501	0.3332	0.2291	0.8345
t + 24	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.4316	0.3040	0.6918	0.4709	0.3335	0.6598

**Table 7.** Model performance in training and testing for multiple-step HSW predictions—PLSR for Station 1.

Time Horizon	Input Combination	Training Period			Test Period		
		RMSE (m)	MAE (m)	R <sup>2</sup>	RMSE (m)	MAE (m)	R <sup>2</sup>
t + 1	HSW <sub>t</sub>						
	HSW <sub>t</sub> , HSW <sub>t-1</sub> ,	0.1022	0.0709	0.9827	0.1132	0.0784	0.9808
	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0928	0.0653	0.9858	0.1118	0.0773	0.9813
t + 2	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1302	0.0874	0.9799	0.1551	0.1072	0.9642
t + 4	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1676	0.1122	0.9687	0.1763	0.1232	0.9497
t + 8	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.2369	0.1584	0.9124	0.2562	0.1749	0.9021
t + 12	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.2979	0.2064	0.8648	0.3208	0.2207	0.8465
t + 24	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.4246	0.2954	0.7084	0.4623	0.3365	0.6681

**Table 8.** Model performance in training and testing for multiple-step HSW predictions—GPR for Station 1.

Time Horizon	Input Combination	Training Period			Test Period		
		RMSE (m)	MAE (m)	R <sup>2</sup>	RMSE (m)	MAE (m)	R <sup>2</sup>
t + 1	HSW <sub>t</sub>	0.1125	0.0819	0.9791	0.1197	0.0826	0.9786
	HSW <sub>t</sub> , HSW <sub>t-1</sub> ,	0.0944	0.0671	0.9853	0.1120	0.0775	0.9815
	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0911	0.0627	0.9863	0.1115	0.0771	0.9817
t + 2	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1297	0.0863	0.9801	0.1425	0.0988	0.9697
t + 4	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1648	0.1094	0.9688	0.1758	0.1221	0.9508
t + 8	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.2346	0.1535	0.9128	0.2553	0.1741	0.9028
t + 12	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.2965	0.2047	0.8651	0.3182	0.2193	0.8490
t + 24	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.4191	0.2889	0.7149	0.4467	0.3312	0.6749

The RMSE and MAE for GPR predictions at multiple horizons, from 1 h ahead (t + 1) to 1 day ahead (t + 24), ranged from 0.1115 m and 0.0771 m, respectively, to 0.4467 m and 0.3312 m, while the corresponding values for PLSR ranged from 0.1118 m and 0.0773 m to 0.4623 m and 0.3365 m, respectively. When comparing GPR to other methods for forecasting HSW at the t + 24 horizon, the RMSE differences were as follows: 24.79, 10.78, 5.78, 5.14, and 3.37%. When the prediction horizon was expanded, the gaps between GPR and other regression approaches widened dramatically.

The results of training and testing of the M5RT, MARS, PCR, RF, PLSR, and GPR models at the second station are summarized in Tables 9–14. All regression models at this station achieved their highest accuracy when using three lagged inputs to predict t + 1 HSWs. Consequently, this input set was employed to forecast HSW at various time scales. However, when the time horizon was extended from t + 1 to t + 24, the performance of metaheuristic regression approaches deteriorated significantly. During the testing phase, the RMSE and MAE values for the M5RT model increased from 0.0418 and 0.0295 m to 0.28 and 0.2122 m, respectively, while the R2 value decreased from 0.9873 to 0.4780. GPR, on the other hand, proved to be the best model for predicting HSWs one hour in advance at this station, exhibiting the lowest RMSE (0.0391 m), the best MAE (0.0274 m), and the highest R2 (0.9890) during the test phase. Comparing GPR to other methods, the RMSE values were 6.46%, 3.46%, 1.76%, 1.26%, and 0.76% for M5RT, MARS, PCR, RF, and PLSR,



respectively. It should be noted that for the  $t + 1$  horizon, HSW predictions by GPR and the PLSR/RF models differed to some extent.

**Table 9.** Model performance in training and testing for multiple-step HSW predictions—M5RT for Station 2.

Time Horizon	Input Combination	Training Period			Test Period		
		RMSE (m)	MAE (m)	R <sup>2</sup>	RMSE (m)	MAE (m)	R <sup>2</sup>
t + 1	HSW <sub>t</sub>	0.0416	0.0293	0.9853	0.0467	0.0322	0.9843
	HSW <sub>t</sub> , HSW <sub>t-1</sub>	0.0412	0.0290	0.9861	0.0434	0.0304	0.9865
	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0397	0.0285	0.9869	0.0418	0.0295	0.9873
t + 2	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0590	0.0411	0.9714	0.0635	0.0442	0.9710
t + 4	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0947	0.0666	0.9263	0.0988	0.0702	0.9298
t + 8	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1494	0.1096	0.8167	0.1561	0.1147	0.8258
t + 12	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1753	0.1313	0.7475	0.2029	0.1526	0.7099
t + 24	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.2069	0.1568	0.6481	0.2810	0.2122	0.4780

**Table 10.** Model performance in training and testing for multiple-step HSW predictions—MARS for Station 2.

Time Horizon	Input Combination	Training Period			Test Period		
		RMSE (m)	MAE (m)	R <sup>2</sup>	RMSE (m)	MAE (m)	R <sup>2</sup>
t + 1	HSW <sub>t</sub>	0.0415	0.0292	0.9855	0.0444	0.0310	0.9859
	HSW <sub>t</sub> , HSW <sub>t-1</sub>	0.0411	0.0289	0.9861	0.0425	0.0296	0.9870
	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0386	0.0274	0.9878	0.0405	0.0287	0.9882
t + 2	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0581	0.0405	0.9722	0.0590	0.0416	0.9751
t + 4	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0921	0.0647	0.9302	0.0918	0.0657	0.9395
t + 8	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1449	0.1058	0.8275	0.1547	0.1151	0.8281
t + 12	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1687	0.1254	0.7661	0.1906	0.1442	0.7389
t + 24	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.2055	0.1548	0.6529	0.2527	0.1910	0.5490

**Table 11.** Model performance in training and testing for multiple-step HSW predictions—PCR for Station 2.

Time Horizon	Input Combination	Training Period			Test Period		
		RMSE (m)	MAE (m)	R <sup>2</sup>	RMSE (m)	MAE (m)	R <sup>2</sup>
t + 1	HSW <sub>t</sub>						
	HSW <sub>t</sub> , HSW <sub>t-1</sub>	0.0409	0.0288	0.9862	0.0418	0.0292	0.9872
	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0366	0.0254	0.9890	0.0398	0.0281	0.9886
t + 2	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0553	0.0391	0.9780	0.0566	0.0396	0.9740
t + 4	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0871	0.0611	0.9376	0.0901	0.0652	0.9418
t + 8	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1332	0.0959	0.8543	0.1484	0.1101	0.8417
t + 12	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1635	0.1203	0.7804	0.1868	0.1405	0.7498
t + 24	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.2012	0.1510	0.6674	0.2472	0.1861	0.5651

**Table 12.** Model performance in training and testing for multiple-step HSW predictions—RF for Station 2.

Time Horizon	Input Combination	Training Period			Test Period		
		RMSE (m)	MAE (m)	R <sup>2</sup>	RMSE (m)	MAE (m)	R <sup>2</sup>
t + 1	HSW <sub>t</sub>	0.0414	0.0291	0.9857	0.0437	0.0310	0.9863
	HSW <sub>t</sub> , HSW <sub>t-1</sub> ,	0.0407	0.0286	0.9864	0.0412	0.0289	0.9878
	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0342	0.0239	0.9904	0.0396	0.0279	0.9887
t + 2	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0545	0.0386	0.9787	0.0560	0.0391	0.9743
t + 4	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0865	0.0608	0.9385	0.0875	0.0631	0.9450
t + 8	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1328	0.0958	0.8551	0.1440	0.1063	0.8511
t + 12	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1625	0.1198	0.7829	0.1805	0.1360	0.7659
t + 24	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.2009	0.1504	0.6683	0.2429	0.1828	0.5810

**Table 13.** Model performance in training and testing for multiple-step HSW predictions—PLSR for Station 2.

Time Horizon	Input Combination	Training Period			Test Period		
		RMSE (m)	MAE (m)	R <sup>2</sup>	RMSE (m)	MAE (m)	R <sup>2</sup>
t + 1	HSW <sub>t</sub>						
	HSW <sub>t</sub> , HSW <sub>t-1</sub> ,	0.0404	0.0285	0.9866	0.0408	0.0288	0.9881
	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0339	0.0232	0.9906	0.0394	0.0277	0.9888
t + 2	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0472	0.0321	0.9817	0.0543	0.0387	0.9777
t + 4	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0806	0.0581	0.9491	0.0838	0.0598	0.9470
t + 8	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1284	0.0888	0.9119	0.1362	0.0997	0.8671
t + 12	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1592	0.1087	0.8782	0.1755	0.1310	0.7790
t + 24	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1932	0.1428	0.7038	0.2363	0.1797	0.6024

**Table 14.** Model performance in training and testing for multiple-step HSW predictions—GPR for Station 2.

Time Horizon	Input Combination	Training Period			Test Period		
		RMSE (m)	MAE (m)	R <sup>2</sup>	RMSE (m)	MAE (m)	R <sup>2</sup>
t + 1	HSW <sub>t</sub>	0.0413	0.0290	0.9859	0.0435	0.0305	0.9864
	HSW <sub>t</sub> , HSW <sub>t-1</sub> ,	0.0402	0.0284	0.9867	0.0405	0.0286	0.9882
	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0330	0.0227	0.9911	0.0391	0.0274	0.9890
t + 2	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0450	0.0305	0.9833	0.0530	0.0374	0.9799
t + 4	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.0801	0.0573	0.9596	0.0835	0.0592	0.9499
t + 8	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1263	0.0824	0.9124	0.1358	0.0996	0.8678
t + 12	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1573	0.1079	0.8808	0.1754	0.1308	0.7792
t + 24	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	0.1919	0.1414	0.7085	0.2322	0.1776	0.6134

The GPR predictions demonstrated varying RMSE and MAE values across different time horizons, ranging from 0.0391 m and 0.0274 m to 0.2322 m and 0.1776 m. In comparison, the corresponding values for PLSR ranged from 0.0394 m and 0.0277 m to 0.2363 m and 0.1797 m, respectively, while for RF, they ranged from 0.0396 m and 0.0279 m to 0.2429 m and 0.1828 m. When comparing GPR to other methods in predicting HSW at the  $t + 24$  horizon, the relative RMSE differences were 17.37%, 8.11%, 6.07%, 4.41%, and 1.74% for M5RT, MARS, PCR, RF, and PLSR, respectively. These findings are consistent with those of the first station, indicating a larger disparity between GPR and other regression methods when forecasting HSW one day in advance.

Figures 4 and 5 compare the results visually through scatterplots for the first and second stations. While GPR’s accuracy in predicting HSW for the next hour is somewhat better than that of other models (excluding M5RT), the difference was not statistically significant. Figures 6 and 7 are Taylor diagrams illustrating GPR’s superior RMSE and correlation values. Moreover, compared to other approaches, the sd of the GPR predictions was more in line with the actual values. Figures 8 and 9 show a comparison of the distribution of the predictions produced by different regression methods. The GPR distribution more closely resembles the observed one in both stations. The computational times of the methods are compared in Table 15. The simulations were done in the MATLAB environment (MATLAB R2017b) using a computer with a Windows 10 (64 bit) operating system with an Intel(R) Core (TM) i5-10500 CPU @ 3.10 GHz processor with 16 GB RAM. All three input cases are included in Table 15. It is clear that the methods are fast enough; however, GPR seems to be slightly faster than the other methods in simulating HSW, with an average computational time of 0.1058 min.

**Table 15.** Computational time of the applied models on the basis of RMSE fitness function (in minutes).

Models	HSW <sub>t</sub>	HSW <sub>t</sub> , HSW <sub>t-1</sub>	HSW <sub>t</sub> , HSW <sub>t-1</sub> , HSW <sub>t-2</sub>	Mean Time
M5RT	0.1638	0.1724	0.1782	0.1715
MARS	0.1437	0.1519	0.1576	0.1511
PCR	0.1346	0.1482	0.1543	0.1457
RF	0.1083	0.1176	0.1217	0.1159
PLSR	0.1158	0.1264	0.1359	0.1260
GPR	0.0979	0.1068	0.1126	0.1058

### 5.2. Discussion

In the presented study, six different metaheuristic regression methods were compared in predicting multiple-step HSWs using historical values as inputs for the models. Comparison results revealed that the Gaussian process regression provided the best predictions compared to the other alternatives for predicting HSWs 1 h ahead. The improvement in RMSE and MAE was 11.08% and 0.27%, respectively, in the first station and 6.46% and 0.76% in the second station when comparing the GPR to the M5RT and PLRS methods. The M5RT model offered the worst predictions at multiple horizons from  $t + 1$  to  $t + 24$ . This result can be explained by the previous literature [63,64]. Srivastava et al. [63] compared the M5RT, MARS and RF methods for predicting solar radiation and found that the RF provided the best accuracy while the M5RT performed worse than the MARS method. Wang et al. [64] compared the MARS, M5RT and RF for predicting soil salinity and found RF to provide the best accuracy.

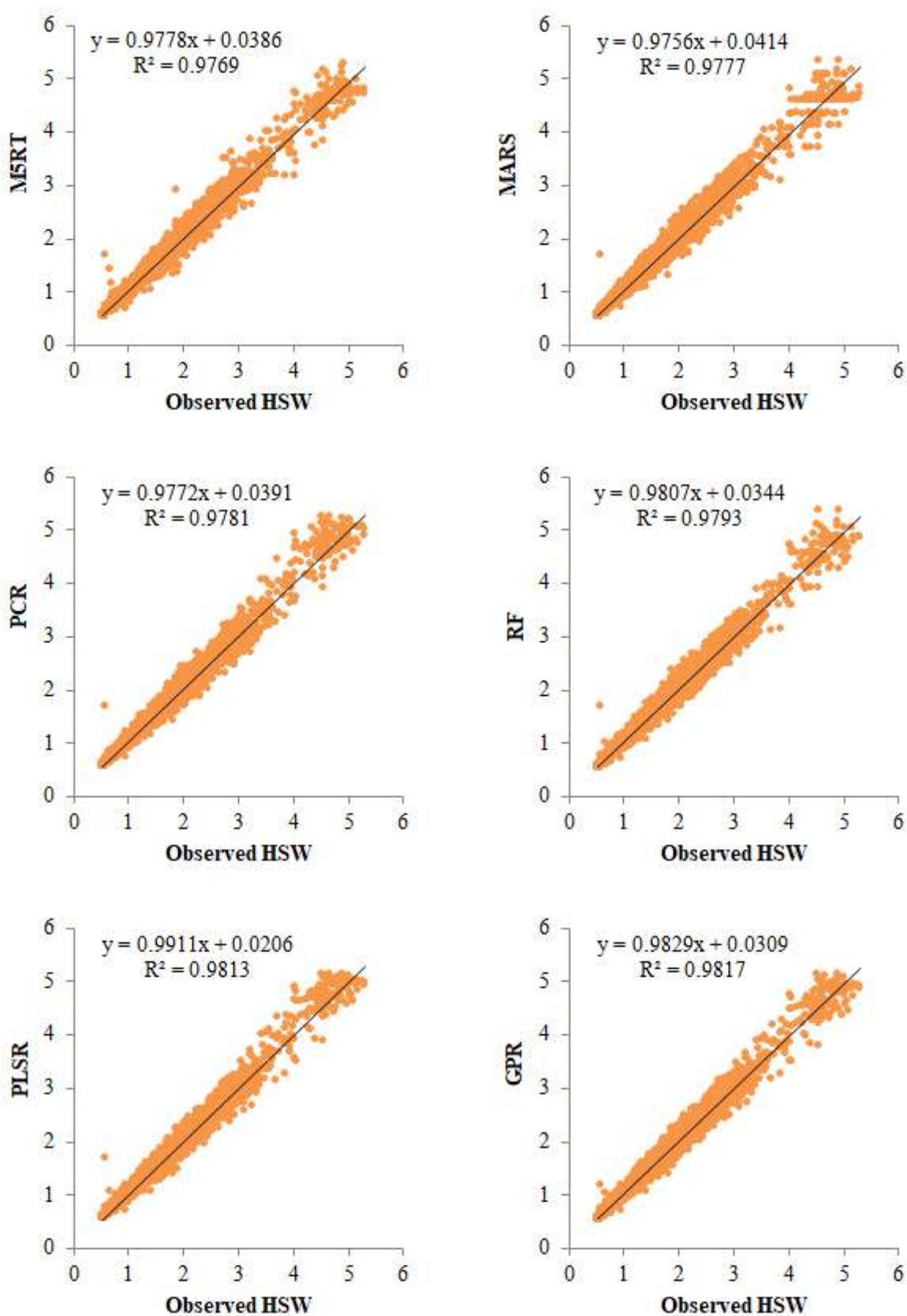


Figure 4. Scatterplots of observed and model-predicted HSWs during testing using the best input combination—Station 1.

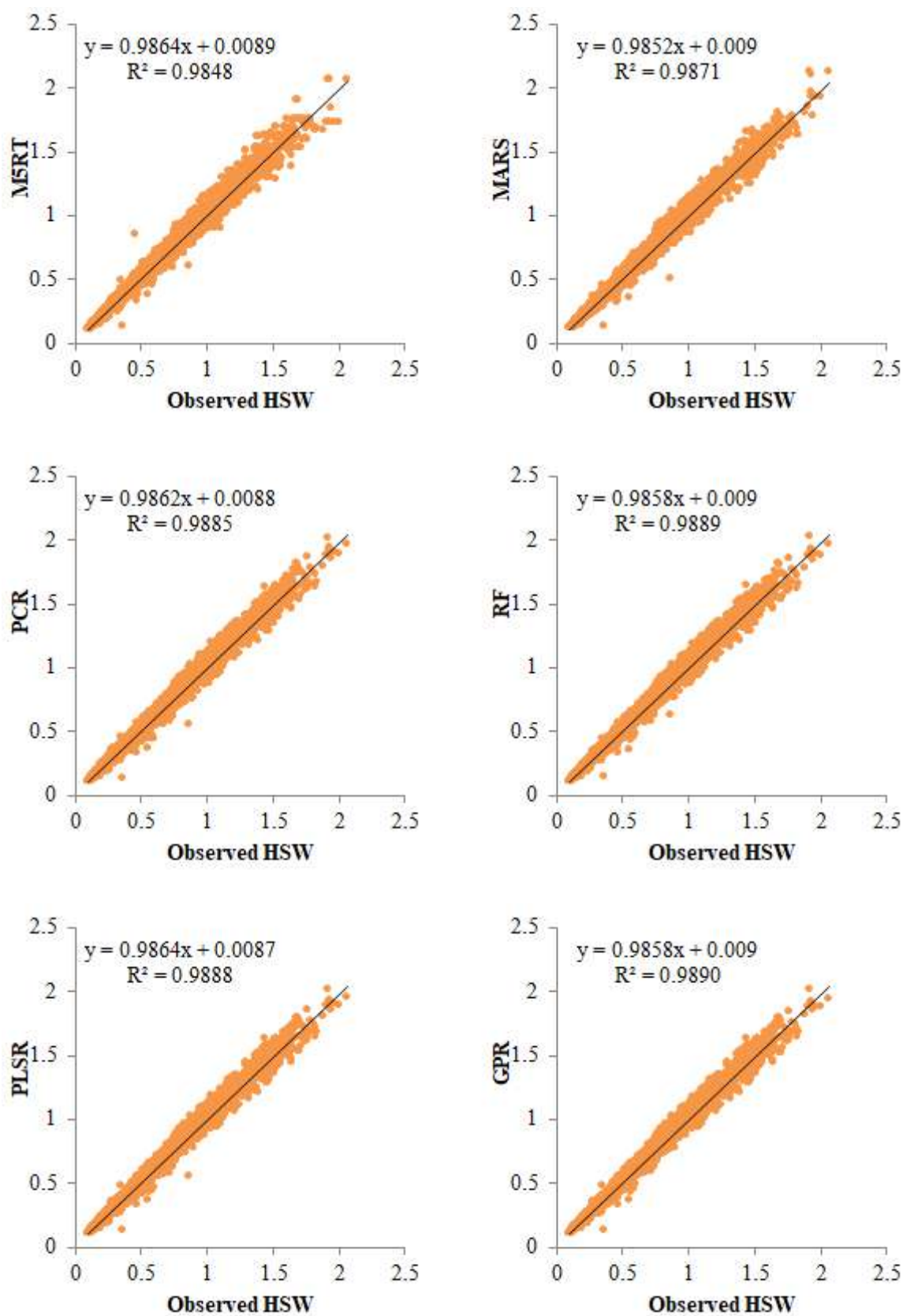


Figure 5. Scatterplots of observed and model-predicted HSWs during testing using the best input combination—Station 2.



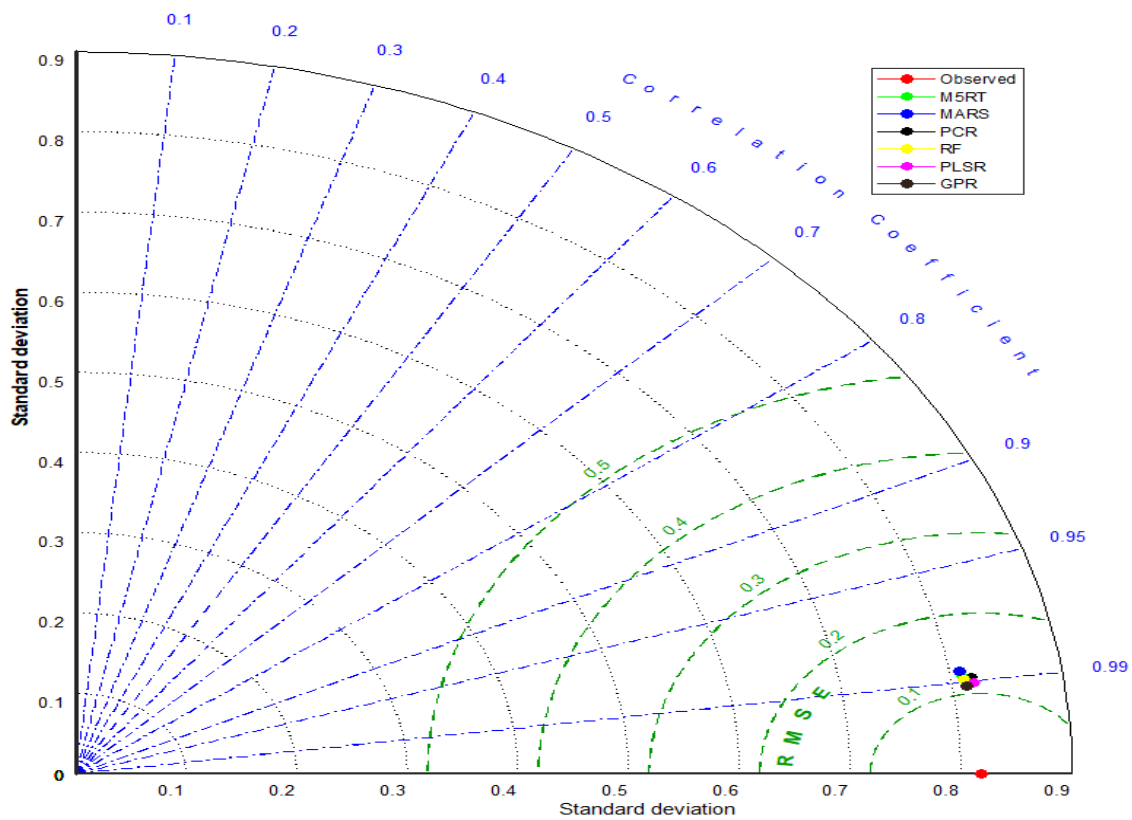


Figure 6. Taylor diagrams of HSWs predicted by different models using the best input combination—Station 1.

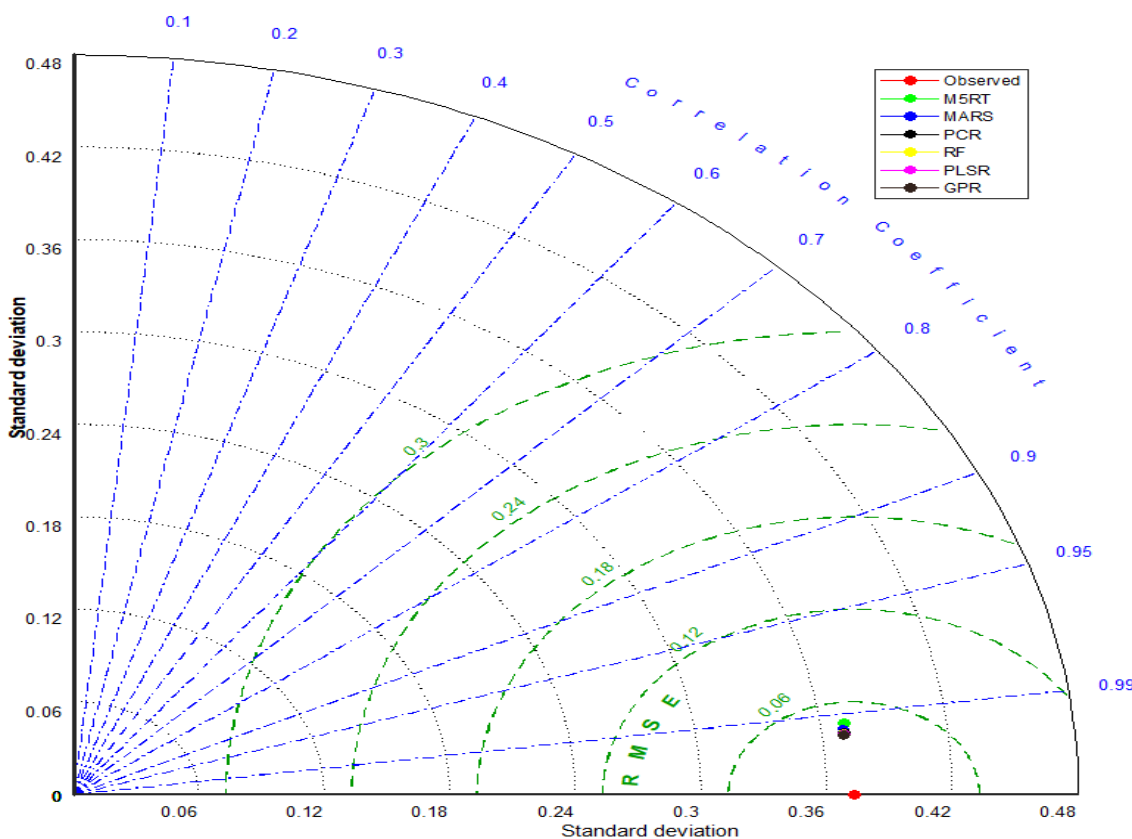
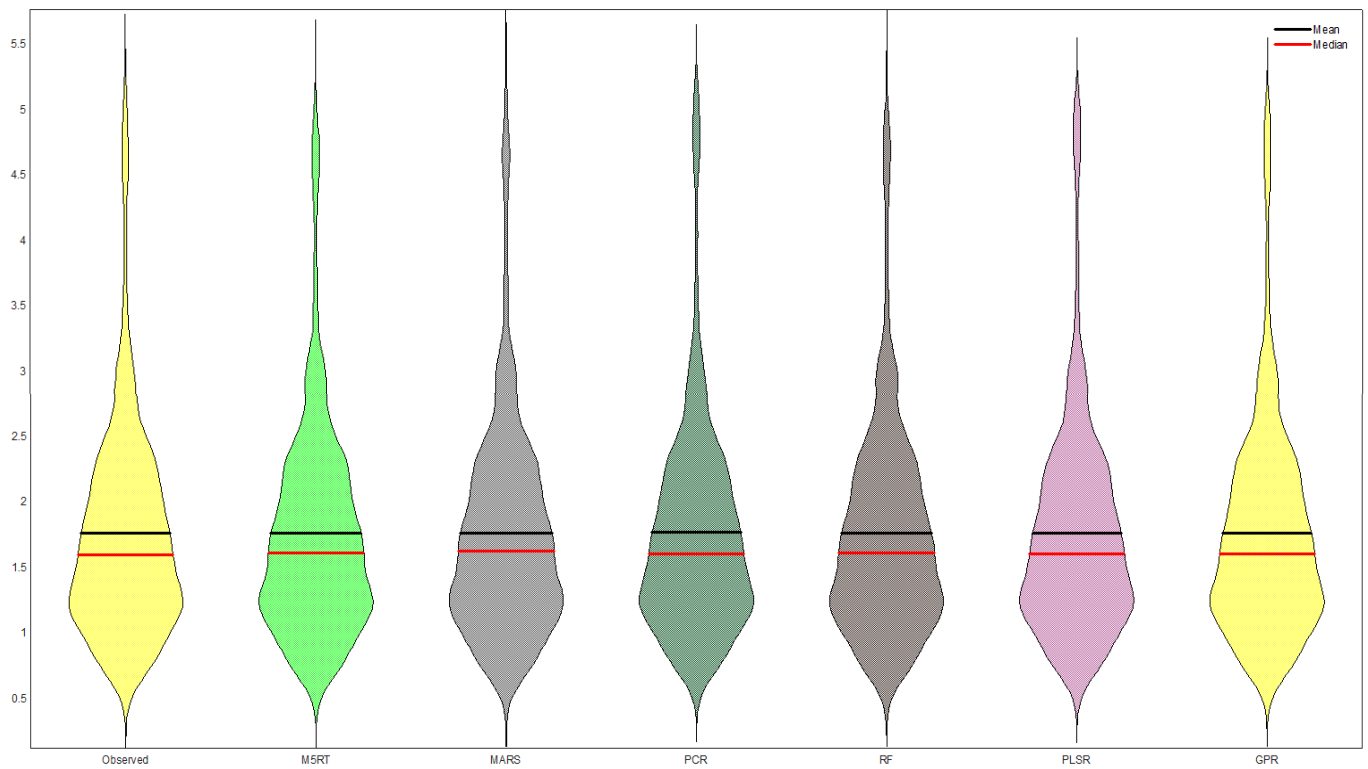
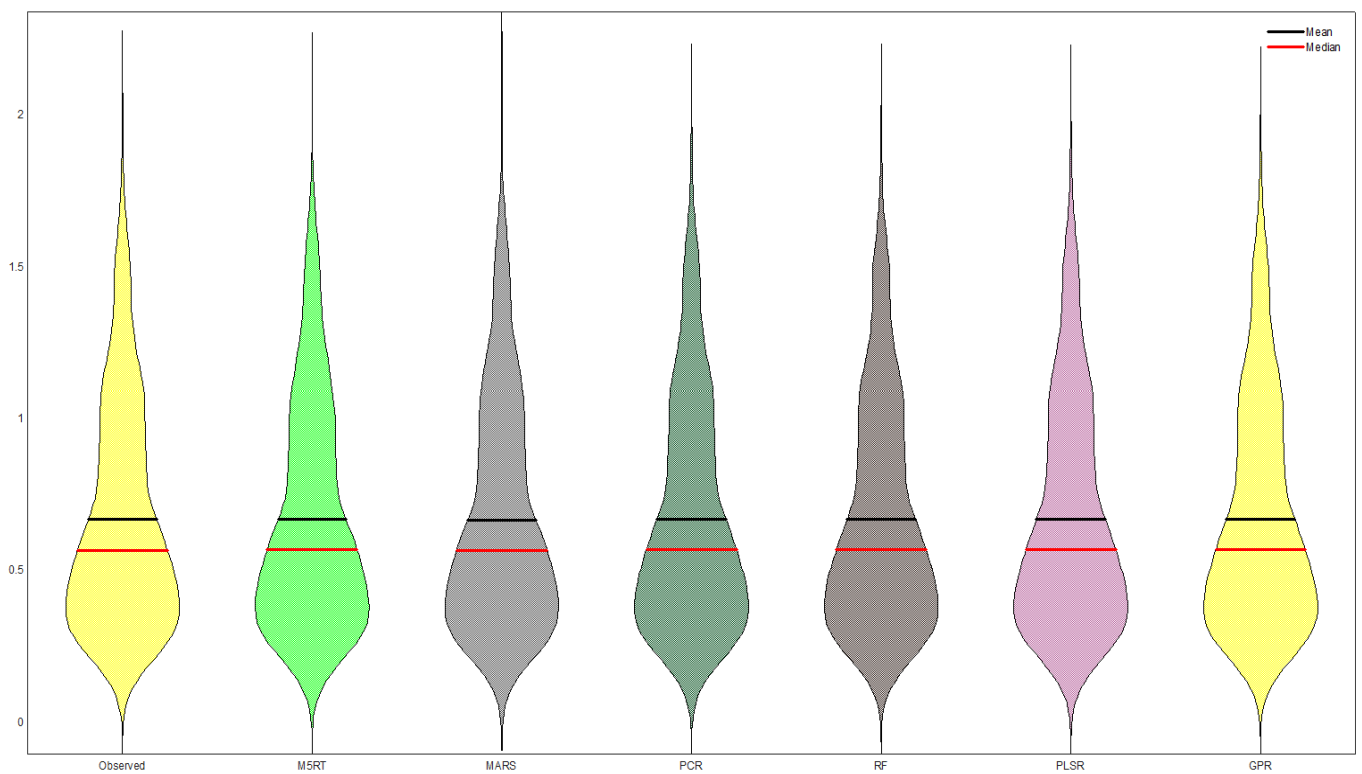


Figure 7. Taylor diagrams of model-predicted HSWs using the best input combination—Station 2.



**Figure 8.** Violin charts of model-predicted HSWs using the best input combination—Station 1.



**Figure 9.** Violin charts of model-predicted HSWs using the best input combination—Station 2.

With substantial wave height predictions across numerous horizons, GPR showed the greatest accuracy for both stations. This approach’s primary benefit is that it provides precise prediction results and expresses the uncertainty of such results. It uses the Bayesian method, providing a principled approach to dealing with uncertainty [65]. Previous studies

also found superior accuracy for GPR in different research areas. Asante-Okyere et al. [66] applied GPR for predicting reservoir porosity and permeability and obtained superior accuracy over the back-propagation ANN, generalized-regression ANN (GRANN), and radial-basis-function ANN. Shabani et al. [67] used four ML methods—GPR, K-nearest neighbors (KNN), RF, and support vector regression—to predict pan evaporation and found that GPR performed best. Singh et al. [68] compared the accuracy of GPR and GRANN for modelling hybrid micro-electric discharge machining and found that GPR provided better accuracy.

Accurate real-time predictions of HSW characteristics are crucial for various short-term management tasks, including energy power generation. Marine engineering applications such as ship-movement forecasting, construction of maritime structures, dredging operations, and disaster warnings all rely on precise hourly predictions of HSW [33]. As the forecasting horizon extends from 1 h to 24 h, there is a noticeable decline in the accuracy of the models. However, GPR (gaussian process regression) generally exhibits superiority in such scenarios, offering potential benefits for monitoring HSW. Machine learning enables us to uncover connections between physical parameters that may be hidden from or unknown to us. Wave formation involves a nonlinear and intricate physical mechanism, and HSW is influenced by various factors, including wind speed, sea surface temperature, water depth, air humidity, and other weather parameters. In this particular study, only HSW data were utilized as inputs due to the unavailability of other relevant influencing parameters.

## 6. Conclusions

The presented work studied the applicability of six metaheuristic regression methods—M5RT, MARS, PCR, RF, PLSR and GPR—in predicting short-term significant wave height at multiple horizons. Data from two stations, the Townsville and Brisbane buoys, were used, and models used historical data as inputs. Among the regression methods, GPR provided the best accuracy in predicting 1 h prediction at both stations. It was closely followed by the PLSR method, while the M5RT produced the worst outcomes. The differences among the methods increased with the increase of the prediction horizon. However, GPR was found to be the best model for all prediction horizons. GPR improved the RMSE of M5RT, MARS, PCR, RF and PLSR by 17.37, 8.11, 6.07, 4.41 and 1.74% (first station) and by 24.79, 10.78, 5.78, 5.14 and 3.37% (second station), respectively, in predicting HSW at the  $t + 24$  horizon (one day ahead) in the test stage. The study's outcomes showed that GPR is a useful tool for predicting significant wave height at multiple time horizons from one hour to one day ahead, and can be used to monitor significant wave height for the study area. Generalization of the implemented methods can be explored in future studies by employing data from other sites.

**Author Contributions:** Conceptualization, R.M.A.I. and O.K.; formal analysis, R.M.A.I.; validation, O.K., X.C., K.S.P., M.Z.-K., S.S. and R.M.A.I.; supervision, O.K. and X.C.; writing—original draft, O.K., K.S.P., M.Z.-K., S.S. and R.M.A.I.; visualization, R.M.A.I. and K.S.P.; investigation, O.K., M.Z.-K. and K.S.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data presented in this study will be available on request from the corresponding author.

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

## References

1. Zhou, S.; Bethel, B.J.; Sun, W.; Zhao, Y.; Xie, W.; Dong, C. Improving significant wave height forecasts using a joint empirical mode decomposition–long short-term memory network. *J. Mar. Sci. Eng.* **2021**, *9*, 744. [[CrossRef](#)]
2. Yang, S.; Deng, Z.; Li, X.; Zheng, C.; Xi, L.; Zhuang, J.; Zhang, Z. A novel hybrid model based on STL decomposition and one-dimensional convolutional neural networks with positional encoding for significant wave height forecast. *Renew. Energy* **2021**, *173*, 531–543. [[CrossRef](#)]

3. Rijnsdorp, D.P.; Smit, P.B.; Guza, R.T. A nonlinear, non-dispersive energy balance for surfzone waves: Infragravity wave dynamics on a sloping beach. *J. Fluid Mech.* **2022**, *944*, A45. [[CrossRef](#)]
4. Shamshirband, S.; Mosavi, A.; Rabczuk, T.; Nabipour, N.; Chau, K.W. Prediction of significant wave height; comparison between nested grid numerical model, and machine learning models of artificial neural networks, extreme learning and support vector machines. *Eng. Appl. Comput. Fluid Mech.* **2020**, *14*, 805–817. [[CrossRef](#)]
5. Zounemat-Kermani, M.; Kisi, O. Time series analysis on marine wind-wave characteristics using chaos theory. *Ocean. Eng.* **2015**, *100*, 46–53. [[CrossRef](#)]
6. Wang, Z.; Xu, H.; Xia, L.; Zou, Z.; Soares, C.G. Kernel-based support vector regression for nonparametric modeling of ship maneuvering motion. *Ocean Eng.* **2020**, *216*, 107994. [[CrossRef](#)]
7. Özger, M. Significant wave height forecasting using wavelet fuzzy logic approach. *Ocean. Eng.* **2010**, *37*, 1443–1451. [[CrossRef](#)]
8. Huang, W.; Dong, S. Improved short-term prediction of significant wave height by decomposing deterministic and stochastic components. *Renew. Energy* **2021**, *177*, 743–758. [[CrossRef](#)]
9. Fernández, J.C.; Salcedo-Sanz, S.; Gutiérrez, P.A.; Alexandre, E.; Hervás-Martínez, C. Significant wave height and energy flux range forecast with machine learning classifiers. *Eng. Appl. Artif. Intell.* **2015**, *43*, 44–53. [[CrossRef](#)]
10. Minuzzi, F.C.; Farina, L. A deep learning approach to predict significant wave height using long short-term memory. *Ocean. Model.* **2023**, *181*, 102151. [[CrossRef](#)]
11. Wang, Z.; Ahmadi, A.; Tian, H.; Jafari, S.; Chen, G. Lower-dimensional simple chaotic systems with spectacular features. *Chaos Solitons Fractals* **2023**, *169*, 113299. [[CrossRef](#)]
12. Adnan, R.M.; Meshram, S.G.; Mostafa, R.R.; Islam, A.R.M.T.; Abba, S.I.; Andorful, F.; Chen, Z. Application of Advanced Optimized Soft Computing Models for Atmospheric Variable Forecasting. *Mathematics* **2023**, *11*, 1213. [[CrossRef](#)]
13. Pourzangbar, A.; Brocchini, M.; Saber, A.; Mahjoobi, J.; Mirzaaghasi, M.; Barzegar, M. Prediction of scour depth at breakwaters due to non-breaking waves using machine learning approaches. *Appl. Ocean. Res.* **2017**, *63*, 120–128. [[CrossRef](#)]
14. Ikram, R.M.A.; Cao, X.; Sadeghifar, T.; Kuriqi, A.; Kisi, O.; Shahid, S. Improving Significant Wave Height Prediction Using a Neuro-Fuzzy Approach and Marine Predators Algorithm. *J. Mar. Sci. Eng.* **2023**, *11*, 1163. [[CrossRef](#)]
15. Zhu, D.; Zhang, J.; Wu, Q.; Dong, Y.; Bastidas-Arteaga, E. Predictive capabilities of data-driven machine learning techniques on wave-bridge interactions. *Appl. Ocean. Res.* **2023**, *137*, 103597. [[CrossRef](#)]
16. Adnan, R.M.; Sadeghifar, T.; Alizamir, M.; Azad, M.T.; Makarynsky, O.; Kisi, O.; Barati, R.; Ahmed, K.O. Short-term probabilistic prediction of significant wave height using bayesian model averaging: Case study of chabahar port, Iran. *Ocean Eng.* **2023**, *272*, 113887. [[CrossRef](#)]
17. Juan, N.P.; Valdecantos, V.N. Review of the application of Artificial Neural Networks in ocean engineering. *Ocean. Eng.* **2022**, *259*, 111947. [[CrossRef](#)]
18. Sadeghifar, T.; Lama, G.F.C.; Sihag, P.; Bayram, A.; Kisi, O. Wave height predictions in complex sea flows through soft-computing models: Case study of Persian Gulf. *Ocean. Eng.* **2022**, *245*, 110467. [[CrossRef](#)]
19. Yang, C.H.; Wu, C.H.; Hsieh, C.M.; Wang, Y.C.; Tsen, I.F.; Tseng, S.H. Deep learning for imputation and forecasting tidal level. *IEEE J. Ocean. Eng.* **2021**, *46*, 1261–1271. [[CrossRef](#)]
20. Park, S.; Byun, J.; Shin, K.S.; Jo, O. Ocean current prediction based on machine learning for deciding handover priority in underwater wireless sensor networks. In Proceedings of the 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIC), Fukuoka, Japan, 19–21 February 2020; pp. 505–509.
21. Mahjoobi, J.; Mosabbeh, E.A. Prediction of significant wave height using regressive support vector machines. *Ocean. Eng.* **2009**, *36*, 339–347. [[CrossRef](#)]
22. Krishna kumar, N.; Savitha, R.; Al Mamun, A. Regional ocean wave height prediction using sequential learning neural networks. *Ocean. Eng.* **2017**, *129*, 605–612. [[CrossRef](#)]
23. 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](#)]
24. Miky, Y.; Kaloop, M.R.; Elnabwy, M.T.; Baik, A.; Alshouny, A. A Recurrent-Cascade-Neural network-nonlinear autoregressive networks with exogenous inputs (NARX) approach for long-term time-series prediction of wave height based on wave characteristics measurements. *Ocean. Eng.* **2021**, *240*, 109958. [[CrossRef](#)]
25. Ikram, R.M.A.; Mostafa, R.R.; Chen, Z.; Islam, A.R.M.T.; Kisi, O.; Kuriqi, A.; Zounemat-Kermani, M. Advanced Hybrid Metaheuristic Machine Learning Models Application for Reference Crop Evapotranspiration Prediction. *Agronomy* **2023**, *13*, 98. [[CrossRef](#)]
26. Han, L.; Ji, Q.; Jia, X.; Liu, Y.; Han, G.; Lin, X. Significant Wave Height Prediction in the South China Sea Based on the ConvLSTM Algorithm. *J. Mar. Sci. Eng.* **2022**, *10*, 1683. [[CrossRef](#)]
27. Pourzangbar, A.; Saber, A.; Yeganeh-Bakhtiary, A.; Ahari, L.R. Predicting scour depth at seawalls using GP and ANNs. *J. Hydroinformatics* **2017**, *19*, 349–363. [[CrossRef](#)]
28. Londhe, S.N.; Panchang, V. One-day wave forecasts based on artificial neural networks. *J. Atmos. Ocean. Technol.* **2006**, *23*, 1593–1603. [[CrossRef](#)]
29. Mahjoobi, J.; Etemad-Shahidi, A. An alternative approach for the prediction of significant wave heights based on classification and regression trees. *Appl. Ocean. Res.* **2018**, *30*, 172–177. [[CrossRef](#)]



30. 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](#)]
31. Altunkaynak, A. Prediction of significant wave height using geno-multilayer perceptron. *Ocean. Eng.* **2013**, *58*, 144–153. [[CrossRef](#)]
32. Salcedo-Sanz, S.; Borge, J.N.; Carro-Calvo, L.; Cuadra, L.; Hessner, K.; Alexandre, E. Significant wave height estimation using SVR algorithms and shadowing information from simulated and real measured X-band radar images of the sea surface. *Ocean. Eng.* **2015**, *101*, 244–253. [[CrossRef](#)]
33. 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](#)]
34. Cornejo-Bueno, L.; Nieto-Borge, J.C.; García-Díaz, P.; Rodríguez, G.; Salcedo-Sanz, S. Significant wave height and energy flux prediction for marine energy applications: A grouping genetic algorithm–Extreme Learning Machine approach. *Renew. Energy* **2016**, *97*, 380–389. [[CrossRef](#)]
35. Berbić, J.; Ocvirk, E.; Carević, D.; Lončar, G. Application of neural networks and support vector machine for significant wave height prediction. *Oceanologia* **2017**, *59*, 331–349. [[CrossRef](#)]
36. Nikoo, M.R.; Kerachian, R.; Alizadeh, M.R. A fuzzy KNN-based model for significant wave height prediction in large lakes. *Oceanologia* **2018**, *60*, 153–168. [[CrossRef](#)]
37. Ali, M.; Prasad, R. Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition. *Renew. Sustain. Energy Rev.* **2019**, *104*, 281–295. [[CrossRef](#)]
38. 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](#)]
39. 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](#)]
40. 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](#)]
41. Friedman, J.H. Multivariate adaptive regression splines. *Ann. Statist.* **1991**, *19*, 1–67. [[CrossRef](#)]
42. De Andrés, J.; Lorca, P.; de Cos Juez, F.J.; Sánchez-Lasheras, F. Bankruptcy forecasting: A hybrid approach using Fuzzy c-means clustering and Multivariate Adaptive Regression Splines (MARS). *Expert Syst. Appl.* **2011**, *38*, 1866–1875. [[CrossRef](#)]
43. Rasmussen, C.E.; Bousquet, O.; Luxburg, U.V.; Rätsch, G. *Gaussian Processes in Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2004.
44. Bai, W.; Ren, J.; Li, T. Modified genetic optimization-based locally weighted learning identification modeling of ship maneuvering with full scale trial. *Future Generat. Comput. Syst.* **2019**, *93*, 1036–1045. [[CrossRef](#)]
45. Moreno, R.; Moreno-Salinas, D.; Aranda, J. Black-box marine vehicle identification with regression techniques for random manoeuvres. *Electronics* **2019**, *8*, 492. [[CrossRef](#)]
46. Ikram, R.M.A.; Goliatt, L.; Kisi, O.; Trajkovic, S.; Shahid, S. Covariance Matrix Adaptation Evolution Strategy for Improving Machine Learning Approaches in Streamflow Prediction. *Mathematics* **2022**, *10*, 2971. [[CrossRef](#)]
47. Hartigan, J.A.; Wong, M.A. Algorithm AS 136: A K-Means Clustering Algorithm. *J. R. Stat. Soc. Ser. C (Appl. Stat.)* **1979**, *28*, 100–108. [[CrossRef](#)]
48. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
49. Xia, Z.; Stewart, K.; Fan, J. Incorporating space and time into random forest models for analyzing geospatial patterns of drug-related crime incidents in a major U.S. metropolitan area. *Comput. Environ. Urban Syst.* **2021**, *87*, 101599. [[CrossRef](#)]
50. Biau, G.; Scornet, E. A random forest guided tour. *Test* **2016**, *25*, 197–227. [[CrossRef](#)]
51. Fan, G.-F.; Yu, M.; Dong, S.-Q.; Yeh, Y.-H.; Hong, W.-C. Forecasting short-term electricity load using hybrid support vector regression with grey catastrophe and random forest modeling. *Util Policy* **2021**, *73*, 101294. [[CrossRef](#)]
52. Cutler, A.; Cutler, D.R.; Stevens, J.R. Random forests. In *Ensemble Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 157–175.
53. Wang, G.; Jiao, J. Quality-Related Fault Detection and Diagnosis Based on Total Principal Component Regression Model. *IEEE Access* **2018**, *6*, 10341–10347. [[CrossRef](#)]
54. Jiang, Q.; Yan, X. Chemical processes monitoring based on weighted principal component analysis and its application. *Chemom. Intell. Lab. Syst.* **2012**, *119*, 11–20. [[CrossRef](#)]
55. Fei, Z.; Liu, K. Online process monitoring for complex systems with dynamic weighted principal component analysis. *Chin. J. Chem. Eng.* **2016**, *24*, 775–786. [[CrossRef](#)]
56. Abdi, H. Partial least square regression (PLS regression). In *Encyclopedia of Measurement and Statistics*; Salkind, N., Ed.; Sage: Thousand Oaks, CA, USA, 2007; pp. 792–795.
57. Wold, S.; Sjostrom, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [[CrossRef](#)]
58. de Jong, S.D. SIMPLS: An alternative approach to partial least squares regression. *Chemom. Intell. Lab. Syst.* **1993**, *18*, 251–263. [[CrossRef](#)]
59. Quinlan, J.R. Learning with continuous classes. In Proceedings of the 5th Australian Joint Conference on Artificial Intelligence, Hobart, Australia, 16–18 November 1992; World Scientific: Singapore, 1992; Volume 92, pp. 343–348.



60. Wang, Y.; Witten, I.H. *Induction of Model Trees for Predicting Continuous Classes*. 1996. Available online: <https://researchcommons.waikato.ac.nz/handle/10289/1183> (accessed on 3 March 2023).
61. Adnan, R.M.; Parmar, K.S.; Heddham, S.; Shahid, S.; Kisi, O. Suspended Sediment Modeling Using a Heuristic Regression Method Hybridized with Kmeans Clustering. *Sustainability* **2021**, *13*, 4648. [[CrossRef](#)]
62. Kisi, O.; Parmar, K.S.; Soni, K.; Demir, V. Modeling of air pollutants using least square support vector regression, multivariate adaptive regression spline, and M5 model tree models. *Air Qual. Atmos. Health* **2017**, *10*, 873–883. [[CrossRef](#)]
63. Srivastava, R.; Tiwari, A.N.; Giri, V.K. Solar radiation forecasting using MARS, CART, M5, and random forest model: A case study for India. *Heliyon* **2019**, *5*, e02692. [[CrossRef](#)]
64. Wang, F.; Shi, Z.; Biswas, A.; Yang, S.T.; Ding, J.L. Multi-algorithm comparison for predicting soil salinity. *Geoderma* **2020**, *365*, 114211. [[CrossRef](#)]
65. Richardson, R.R.; Osborne, M.A.; Howey, D.A. Gaussian process regression for forecasting battery state of health. *J. Power Sources* **2017**, *357*, 209–219. [[CrossRef](#)]
66. Asante-Okyere, S.; Shen, C.; Yevenyo Ziggah, Y.; Moses Rulegeya, M.; Zhu, X. Investigating the Predictive Performance of Gaussian Process Regression in Evaluating Reservoir Porosity and Permeability. *Energies* **2018**, *11*, 3261. [[CrossRef](#)]
67. Shabani, S.; Samadianfard, S.; Sattari, M.T.; Mosavi, A.; Shamshirband, S.; Kmet, T.; Várkonyi-Kóczy, A.R. Modeling Pan Evaporation Using Gaussian Process Regression K-Nearest Neighbors Random Forest and Support Vector Machines; Comparative Analysis. *Atmosphere* **2020**, *11*, 66. [[CrossRef](#)]
68. Singh, S.K.; Mali, H.S.; Unune, D.R.; Wojciechowski, S.; Wilczyński, D. Application of Generalized Regression Neural Network and Gaussian Process Regression for Modelling Hybrid Micro-Electric Discharge Machining: A Comparative Study. *Processes* **2022**, *10*, 755. [[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.