

# A Comparison of Soft Computing Methods for the Prediction of Wave Height Parameters

Rifat Tur<sup>1,\*</sup> and Serbay Yontem<sup>2</sup>

<sup>1</sup> Dept. of Civil Engineering, Faculty of Engineering, Akdeniz University, Antalya, Turkey

<sup>2</sup> 13<sup>th</sup> Regional Directorate, General Directorate of Highways, Antalya, Turkey; srbyyntm@gmail.com

\* Corresponding: rifattur@gmail.com

Manuscript received: 03-03-2021, revised: 27-04-2021, accepted: 28-04-2021.

## Abstract

In the previous studies on the prediction of wave height parameters, only the significant wave height has been considered as the unknown parameter to be predicted. However, the other wave height parameters, which may be required for the design of coastal structures depending on their importance level, have been neglected. Therefore, in this study, novel soft computing methods were used to predict all wave height parameters required for the design of coastal structures. To this end, wave data were derived from a buoy located in Southwest Black Sea Coast. Then, Multi-layer Perceptron Neural Network (MLPNN) and Adaptive-Neuro Fuzzy Inference System (ANFIS) models were developed to predict wave height parameters. Various input combinations were selected to create seven different sub-models. These sub-models were applied using developed MLPNN and ANFIS models. Accuracy of sub-models were evaluated for each wave height parameters in terms of performance evaluation criteria. The results showed that the wave height parameters predicted by the MLPNN and ANFIS methods are similar and both methods yield results acceptable for design purposes. However, for maximum wave height,  $H_{max}$ , ANFIS sub-model yields slightly better results.

**Keywords:** Wave height parameters; wave prediction; MLPNN; ANFIS; Filyos region; Black Sea.

## 1. Introduction

Coastal activities such as design of coastal and offshore structures, sediment transport, coastal erosion, and environment related coastal studies require an accurate prediction of wave parameters. These parameters can be obtained by using long term measurement or observed wave data. However, in practice, the wave data required for design are either recorded for short periods or missing and hence the use of complicated wave prediction models is inevitable. This situation was mentioned by Mahjoobi et al. (2008) as field observations of wave height parameters can be extremely difficult and little wave data are available for engineering purpose [1]. Several empirical-based and numerical-based methods have been widely used in literature to predict wave characteristics. Coastal Engineering Manual [2], SMB [3], Shore Protection Manual [4], JONSWAP [5] and Goda [6] can be classified the empirical-based methods. On the other hand, numerical-based methods such as SWAN [7] and WAM [8] can be used for deep water conditions, simulating waves and wave prediction studies. Numerical methods are generally based on an energy transfer equation or action balance equation [9]. However, the application of both empirical and numerical-based methods requires high amount of time for calculation and accurate local bathymetry surveys for implementation. In addition, these methods do not entirely account for the complexity and uncertainty of wave phenomena, as stated by Makarynskyy et al. (2005) [10]. Therefore, in recent decades, soft computing methods such as Artificial Neural Networks (ANN), Adaptive Neuro-Fuzzy Inference System (ANFIS), Genetic Programming (GP) and others have been used to predict wave height parameters. They make it possible to obtain a great deal of wave data quickly with great success. Soft computing methods are considered as a robust technique for coastal studies such as prediction of missing wave data [11]–[13], forecasting of tidal level [14], [15], reliability assessment [16]–[18], prediction of wave energy [18]–[20], and coastal structural stability analysis [21]–[27].

Focusing on the ANN and ANFIS-based wave parameter prediction models, our review showed that several applications exist in the related literature. For example, Tsai et al. (2002) forecasted significant wave height at three stations in Taichung Harbor, Taiwan using ANN [28]. Makarynskyy et al. (2005) collected wave data from an offshore buoy at the west coast of Portugal. An ANN model was developed to predict significant wave height at the location [10]. The performance of the model was evaluated using three different parameter such as root mean square (RMSE), correlation coefficient (R) and scatter index (SI). The results showed that the developed ANN model can be used to predict short term significant wave height. Mandal and Prabaharan (2006) presented a recurrent neural network approach to predict ocean wave [29]. To this aim, wave data were

collected from a buoy at 23-meter water depth located in the west coast of India. To evaluate the model performance, R parameter was used. Authors indicated that recurrent neural network approach outperform the classical ANN models. Zamani et al. (2008) compared the ANN model with instance-based learning (IBL) for wind-wave forecasting [30]. Therefore, two sets of meteorological and wave data were collected from two different buoy in the southern part of Caspian Sea. Input parameters were selected for the models using average mutual information. A comparison was performed between the models based on the R, SI, and RMSE. The results showed that the ANN model is better than the IBL model for significant wave height prediction. Günaydın (2008) compared the significant wave height prediction performance of the ANN and regression methods [31]. The required data were collected from three different buoys located in Atlantic region. The results of the study showed that ANN is superior to regression methods for prediction purpose. Kamranzad et al. (2011) proposed an ANN model that use the wind speed, wind direction and significant wave height as input [32]. The proper input combinations were selected using trial and error method. Wave and wind data were measured from a buoy and a meteorological station located in Persian Gulf respectively. R, mean square error (MSE), and index of agreement ( $I_a$ ) parameters were used to evaluate model performance. The results showed that the previous wind speed is more effective on the significant wave height prediction in long lead times, while previous wave height is more effective in short lead times. Mafi and Amirinia (2017) presented three soft computing models such as support vector machine (SVM), ANN, random forest (RF) to predict hurricane wave height [33]. To this aim, a huge set of meteorological and wave data were collected from six different buoys located in Gulf of Mexico. The SVM, ANN, and RF models were developed using the huge dataset. R, RMSE, SI parameters were used to evaluate the model performance. Authors indicated that developed models to predict hurricane wave heights accurately. A recent comparison between the ANN and other soft computing methods such as extreme learning machine (ELM), and SVM were performed by Shamshirband et al. (2020) for significant wave height prediction [34]. The wave data were collected from the two-station located in Persian Gulf. Wind data for the study area were collected from the source of European Center for Medium-range Weather Forecast (ECMWF). These data were used as input to feed three models and wave data were used as the target values of model. Furthermore, the bathymetry maps were obtained, and a numerical-based SWAN model was developed for significant wave height prediction. Although the results showed that the performance of soft computing models can be considered as similar, it was observed that the ELM model slightly better than others. On the other hand, hybrid neuro-wavelet models that is the combination of wavelet decomposition and ANN were also presented by a few researchers in literature [35].

Another soft computing method that used in significant wave height prediction studies is ANFIS. Architecture and learning procedure underlying ANFIS were first presented by Jang (1993) [36]. It combines the optimization and learning capabilities of neural networks with fuzzy logic linguistic IF-THEN rules which consist of membership functions. ANFIS has been frequently used in literature to predict wave parameters. For example, Kazeminezhad et al. (2005) predicted significant wave height developing an ANFIS and CEM model [37]. The required wave and wind data was collected from a station located in Lake Ontario, USA. The results showed that the ANFIS is superior to CEM model based on the performance evaluation. TÜR and BALAS (2010) developed several ANFIS model to predict significant wave height at Filyos region, Turkey [38]. The different input combinations that include significant wave height and wave periods were used. Performance evaluation of the model was done using RMSE, mean absolute error (MAE), and R. In addition, developed model were also compared with the other models presented by Dixit et al. (2015), Kazeminezhad et al. (2005), Londhe (2008), Makarynsky et al. (2005), Mandal and Prabakaran (2006) and Zamani et al. (2008) [10], [12], [29], [30], [35], [37]. The comparison showed that the developed ANFIS model superior to other models. Authors also showed that the daily significant wave height can be predicted using daily average wave height and wave period data recorded at different time intervals. Akpınar et al. (2014) developed an ANFIS model and compared it with the four different empirical-based methods (Wilson, SPM, Jonswap, and CEM) [9]. To this end, this study focused on the Southern Black Sea. Wind speed, wind duration, significant wave height and mean wave period data were obtained for the study. Fetch length, duration, and wind speed were selected as input parameters for ANFIS model, whereas significant wave height values were selected as the target. The performance evaluation of the predicted results gathered from the models were done using MAE, and R. Authors indicated that ANFIS model is superior to empirical-based models. Hashim et al. (2016) developed ANFIS model to predict offshore significant wave height [39]. To this end, meteorological and wave dataset were gathered from three different buoy located in Atlantic Ocean. Meteorological data such as wind speed, wind direction, air temperature, and sea surface water temperature were selected as the input parameters of the model. The target of the model was selected as significant wave height. Authors indicated that the including air temperature, and sea surface temperature parameters as input, cause an increase on the model performance in terms of RMSE values. More recently, a hybrid FIS-based ANFIS significant wave height model was introduced by Stefanakos (2016) [40]. Authors indicated that for the first time in literature, fuzzy logic and ANFIS were

combined. The results showed that the proposed hybrid model has more accuracy than the standalone FIS and ANFIS models.

The reviewed articles revealed that ANN model is a highly preferred methods to prediction modeling of significant wave height. However, ANN-based MLPNN algorithm can be still investigated for prediction of wave height. Aforementioned studies were presented to predict significant wave height by researchers. However, there is a scarcity on the literature to predict other wave height parameters required for coastal activities. In the related literature, studies are found, where revision necessity for the coastal structure design formulae due to the rise in the wave height parameters as a consequence of climate change, is reported [41]–[44]. Therefore, it may be plausible to use wave parameters other than significant wave height during the determination of design parameters such as the armor stone weight and wave overtopping. To this purpose, two soft computing methods; MLPNN and ANFIS were developed to predict four wave height parameters including the significant wave height,  $H_s$ ; the mean wave height,  $H_{mean}$ ; the mean value of the 1/10 highest wave heights,  $H_{1/10}$ ; and  $H_{max}$ , the maximum value of wave heights (within the wave train), in this study. These parameters are given below in detail:

$$H_s = \frac{\sum_{i=1}^n H_i}{\frac{n}{3}} \tag{1}$$

$$H_{mean} = \frac{\sum_{i=1}^n H_i}{n} \tag{2}$$

$$H_{\frac{1}{10}} = \frac{\sum_{i=1}^{\frac{n}{10}} H_i}{\frac{n}{10}} \tag{3}$$

where  $H_i$  denotes the individual wave heights in a certain time series and  $n$  denotes the amount of wave height data.

Another novelty of the present study is that several sub-models were defined and used to obtain the numerical results presented. The best sub-models for each wave parameter studied, and for both numerical methods were determined. Also, the error parameters for MLPNN and ANFIS sub-models were presented.

## 2. Study Area and Data

The data set used in this study comprises wave height parameters;  $H_s$ ,  $H_{mean}$ ,  $H_{1/10}$ ,  $H_{max}$  and the wave periods;  $T_s$ ,  $T_{mean}$ ,  $T_{1/10}$ ,  $T_{max}$ . These data were gathered from a buoy station located in Southwest Black Sea Coast. This station is located on the latitude of  $41^{\circ}32'50''$  and longitude of  $32^{\circ}01'00''$  in Filyos Region, Turkey (see Figure 1). The measurements on the station are performed by Acoustic Wave and Current Profiler device (AWAC). These devices are located at 12.5-m depth.

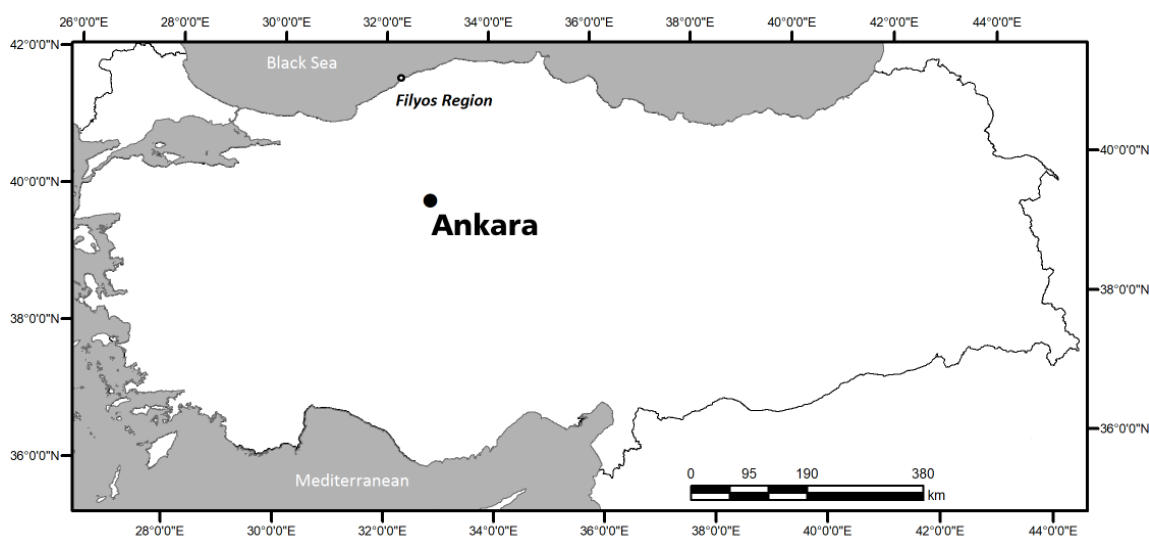


Figure 1. The Study Area and the Location of the AWAC.

The wave data covers 2-year measurement of the station in mention. 7949 different readings for each wave height parameters obtained. Bilyay et al. (2011) used the same dataset on their study. Authors indicated that the dataset has some gaps due to battery and memory cassette changes [45]. The minimum, maximum, mean value, skewness, and standard deviation values for the data used were presented in Table 1.

**Table 1.** Statistical Properties of Representative Wave Height Parameters.

Wave Height Parameters	Min (m)	Max (m)	Mean (m)	Skewness	Std. Deviation (m)
$H_s$	0.03	5.00	0.62	1.98	0.59
$H_{\text{mean}}$	0.03	3.32	0.40	2.07	0.38
$H_{1/10}$	0.03	6.06	0.78	1.94	0.73
$H_{\text{max}}$	0.03	7.55	1.01	1.88	0.96

Filyos is a region that have some extreme wind and wave characteristics due to nature of Black Sea. Study of Özhan and Abdalla (2002) showed that the Black Sea coast including Filyos has higher wave height and wind speed values than the other coast of Turkey [46]. Their studies showed that the buoy measurement in the region is done in short durations. Therefore, general attempt is estimating wave height using wind information of the region. Also, TÜR and BALAS (2010b) studied reliability-based risk assessment of revetment in the Filyos region [47]. Design wave height of 5.8-m and 7.50-m was selected for study.

### 3. Materials and Methods

The proposed soft computing forecasting models for the wave prediction consist of two models, namely (1) Adaptive Neuro-Fuzzy Inference System (ANFIS), (2) Multi-layer perceptron neural network (MLPNN). In this part, selection criteria of input combination, MLPNN theory, ANFIS theory, and model performance evaluation criteria were described in an order.

#### 3.1 Input Combination Selection

In this study, wave height parameter was selected as the target parameter, while previous values of both wave height and period were used as the input of the model. To this end, seven different sub-model were created to obtain best prediction of target parameter. In literature several researchers used trial/error manner to select input combinations [32], [38], [48]. This manner allows to select different input combinations and then, each combination is tested. Similarly, input combinations were selected with trial/error manner in this study. The alternative sub-model architectures built by different predictor configurations. These procedure Sub-models were given in Table 2, where  $H_{(t)}$  is wave height (2-hourly values) and  $T_{(t)}$  is corresponding wave period as inputs.

**Table 2.** Sub-model Architectures for Output  $H_{(t)}$ .

Sub-Model	Model Architecture	
	Output	Inputs
Mod - 1	$H_{(t)}$	$H_{(t-1)}, T_{(t-1)}$
Mod - 2	$H_{(t)}$	$H_{(t-1)}, T_{(t-1)}, T_{(t-2)}$
Mod - 3	$H_{(t)}$	$H_{(t-1)}, H_{(t-2)}$
Mod - 4	$H_{(t)}$	$H_{(t-1)}, H_{(t-2)}, T_{(t-1)}$
Mod - 5	$H_{(t)}$	$H_{(t-1)}, H_{(t-2)}, H_{(t-3)}$
Mod - 6	$H_{(t)}$	$H_{(t-1)}, H_{(t-3)}, T_{(t-1)}$
Mod - 7	$H_{(t)}$	$H_{(t-1)}, H_{(t-3)}, H_{(t-5)}$

For example, the 2<sup>nd</sup> sub-model (Mod-2) consists of output variable  $H_{(t)}$  and input variables  $H_{(t-1)}$ ,  $T_{(t-1)}$ , and  $T_{(t-2)}$ , which denote one time step wave height, one time step past period and two time step past period respectively.

#### 3.2 Multi-layer Perceptron Neural Networks (MLPNN) Model

An ANN model imitates human brain activity, and it consists of neurons, which are processing units, as stated Balas et al. (2006) [49]. A multi-layer perceptron neural networks (MLPNN) is a type of ANN algorithm that consists of two or more layers with neurons in each layer. It is a frequently used supervised learning method. The network response is:

$$y_i = f\left(\sum_{i=1}^n w_i x_i - t_j\right) \tag{4}$$

where  $n$  denotes the total number of input values and their weights,  $x_i$  parameter are the input values of a perceptron, parameter  $w_i$  are weights for each input,  $t_j$  value is the bias of the perceptron and  $y_j$  refers to output of a perceptron that states activation function values of summation. It is shown in Figure 2.

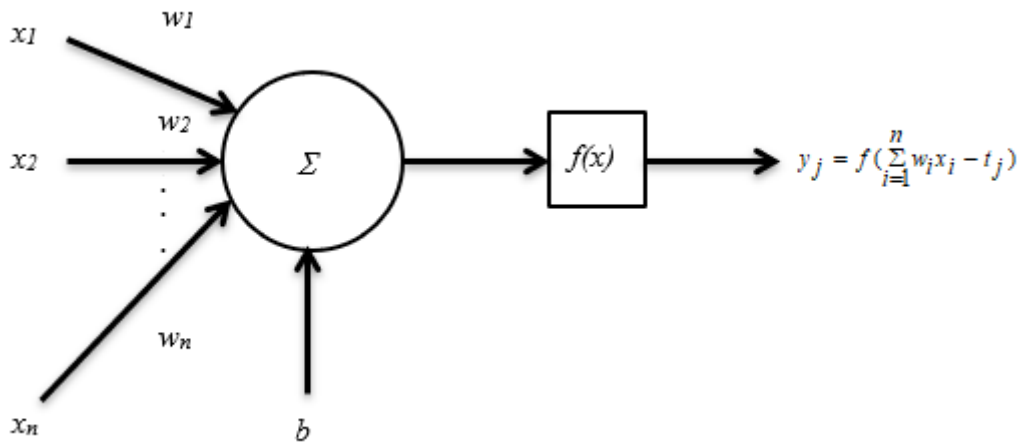


Figure 2. The Detail of Each Artificial Neuron.

Two kind of classical activation function is used in MLPNN. The first is hyperbolic tangent function and the second is logistic function. They range between -1 to 1, and 0 to 1, respectively. There are hidden layers between the input and the output layers. Each neuron layer receives input from the previous layer and gets in contact with the next layer by sending an output. Figure 3 shows the activation process of a hidden layer.

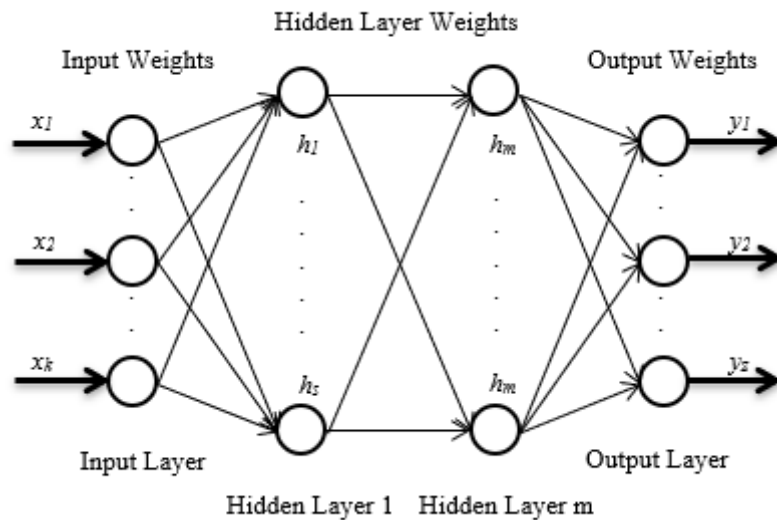


Figure 3. The Structure of a General MLPNN Model.

The output  $y$  of a perceptron for hyperbolic tangent sigmoid, or tanh, function in the hidden layers is denoted as;

$$y = f(z) = \tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (5)$$

Parameter  $e$  denotes natural algorithm, and parameter  $z$  denotes the output of a perceptron summation.

A training data set, which consists of randomly selected input and output values is used to train the MLPNN and a testing data set is used to check the success of the model. The performance of the model is determined by calculating the performance index,  $\nabla R$ , as:

$$\nabla R(w) = \frac{1}{zN} \sum_{n=1}^N \nabla S(w, n) \quad (6)$$

where,  $N$ ,  $n$ ,  $z$ , and  $\nabla S(w, n)$  denote the total sample number, the total epoch number, the output layer neuron total number, the gradient of total error values, respectively. The updates for weight and bias updates are made by;

$$\nabla S(w, n) = \left[ \frac{\partial S}{\partial w_{11}^h} \cdots \frac{\partial S}{\partial w_{ji}^h} \cdots \frac{\partial S}{\partial w_{11}^y} \cdots \frac{\partial S}{\partial w_{lm}^y} \right] \quad (7)$$

The hidden and output layer weights are denoted by  $W^h$  and  $W^y$ ,

$$W^h = [w_{11}^h \cdots w_{1i}^h \cdots w_{ji}^h] j = 1, \dots, s; i = 1, \dots, k \quad (8)$$

$$W^y = [w_{11}^y \cdots w_{1i}^y \cdots w_{ji}^y] j = 1, \dots, z; m = 1, \dots, s \quad (9)$$

The weight factors are calculated according to the GDM (Gradient Descent Method),

$$w(k + 1) = w(k) - \eta \nabla R(w) + \alpha \nabla w(k - 1) \quad (10)$$

where,  $\eta$  denotes the step size re-evaluated at  $k^{\text{th}}$  iteration,  $w(k + 1)$  denotes the weight vector at the  $(k + 1)^{\text{th}}$  iteration, and  $\alpha$  denotes the momentum constant selected by the user.

### 3.3 Adaptive Neural Network Based Fuzzy Inference – ANFIS

ANFIS model was introduced by Jang and it makes use of ANN and fuzzy logic [37]. ANFIS can be described as a network structure that has the capacity of neural learning of Sugeno type fuzzy systems. This network structure is formed of layers of nodes associated with separate nodal functions [50]. For example, a Sugeno type FIS including 1 output value ( $f$ ), 2 input values ( $x, y$ ) and a rule base with 2 fuzzy rules is given below:

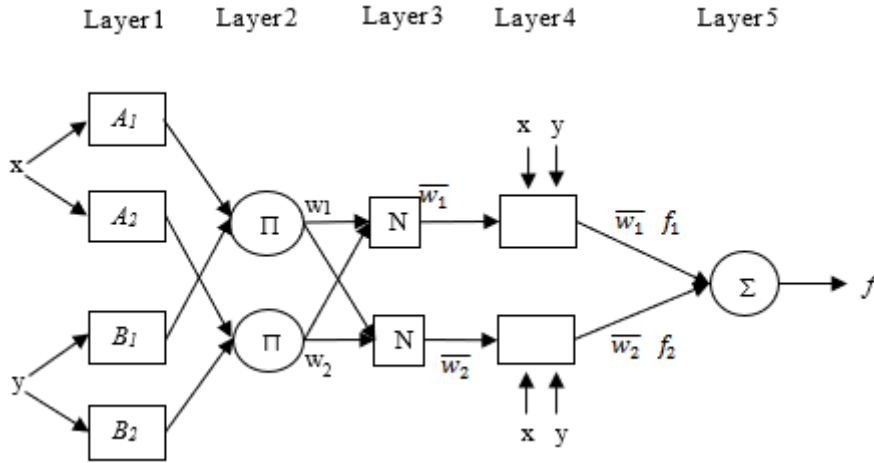
**1<sup>st</sup> rule:**

$$\text{If } x \text{ is } A_1 \text{ and } y \text{ is } B_1 \text{ then } f_1 = p_1x + q_1y + r_1$$

**2<sup>nd</sup> rule:**

$$\text{If } x \text{ is } A_2 \text{ and } y \text{ is } B_2 \text{ then } f_2 = p_2x + q_2y + r_2$$

where  $A_i$  and  $B_i$  ( $i=1, 2$ ) are the membership function for input  $x$  and  $y$ , and the  $q_i, p_i$  and  $r_i$  ( $i = 1, 2$ ) are the parameters used for the Sugeno type first degree fuzzy model. The ANFIS structure is shown in Figure 4, where adaptive nodes are indicated by squares and fixed nodes are indicated by circles [37].



**Figure 4.** ANFIS Architecture for Two-input Sugeno Fuzzy Model.

ANFIS comprises five layers (see Figure 4) and description of each layer can be seen below:

**1<sup>st</sup> Layer:** For a given node, membership degrees, which describe the belonging degree to the fuzzy set, are calculated using the membership functions,  $O_i^1$ , is;

$$O_i^1 = \mu_{A_i}(x) \tag{11}$$

where,  $x$  denotes the crisp input value for node  $i$ ,  $A_i$  and  $B_i$  are linguistic terms, and  $\mu_{A_i}$  and  $\mu_{B_i}$  denote the membership functions.

For example, bell-shaped curve membership function can be written as follows:

$$\mu_{A_i} = \frac{1}{1 + \left[ \frac{(x - c_i)^2}{a_i} \right]^{b_i}} \tag{12}$$

where  $a_i$ ,  $b_i$ ,  $c_i$  denote the membership functions curve parameters for the fuzzy if-then rule structure. This function yields values between 0 and 1.

There are several choices for membership functions in the related literature. In the present study, the Gaussian curve membership function is used. As it is the most extensively used function in similar studies.

**2<sup>nd</sup> Layer:** The nodes in the 2<sup>nd</sup> layer, labeled by  $\mu$ , is multiplied by the incoming values, the output is transferred to the next layer; see Figure 4. These outputs show the rule strength. The multiplication operation for this layer is;

$$w_i = \mu_{A_i}(x) \times \mu_{B_i}(y), \quad i=1, 2. \tag{13}$$

**3<sup>rd</sup> Layer:** The equation below shows that the  $i^{\text{th}}$  node of layer 3 yields the firing strength ratio for the  $i^{\text{th}}$  rule to the firing strength ratios of all of the rules.

$$\bar{w}_i = \frac{w_i}{w_1 + w_2}, \quad i=1, 2 \tag{14}$$

**4<sup>th</sup> Layer:** Weighted mean value of the  $i^{\text{th}}$  rule is;

$$Q_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i x + r_i) \tag{15}$$

where,  $p_i$ ,  $q_i$ ,  $r_i$  are the parameters and  $\bar{w}_i$  is the output.

**5<sup>th</sup> Layer:** The final output is obtained by summing up all of the incoming values.

$$Q_i^5 = \text{overall output} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \tag{16}$$

The ANFIS can be trained by a hybrid learning algorithm presented by Jang (1993) [36]. The aim of the learning process is to determine the premise parameters (see Layer 1) and consequent parameters (see Layer 4) based on error measures which should be minimize. Error measurement is performed by using the difference between the observed and predicted outputs. Root mean square error (RMSE) is the source of the ANFIS error measurement process. The deeper information about the ANFIS can be found in the studies of Jang (1993) [36].

### 3.4 Performance Evaluation Criteria

The performance of a model can be evaluated by using standard statistical performance evaluation criteria. In this study, RMSE, MAE, R, and Bias were selected as the performance criteria. According to Legates and McCabe (1999), the R only evaluates linear relationship between the variables [51]. It sometimes makes poor model performance due to it measure relative error between the measured and predicted values. Therefore, RMSE and MAE techniques which measure the absolute errors were also adopted in this study. The performance of the RMSE, MAE, and Bias were investigated by Willmott (1982) [52]. The results showed that these criteria can be used for the hydrological studies. On the other hand, the literature review we did at the beginning of the paper showed that these performance criteria are widely used in the coastal engineering studies. Based on the relevant literature, Root Mean Square Error (RMSE), Mean Absolute Error (MAE), the correlation coefficient (R) and Bias can be defined as follows:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (y_t^o - y_t^p)^2}{n}} \times 100 \tag{17}$$

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t^o - y_t^p| \tag{18}$$

$$R = \frac{\sum_{t=1}^n (y_t^o - \bar{y}^o)(y_t^p - \bar{y}^p)}{\sqrt{\sum_{t=1}^n (y_t^o - \bar{y}^o)^2 \sum_{t=1}^n (y_t^p - \bar{y}^p)^2}} \tag{19}$$

$$Bias = \frac{1}{n} \sum_{t=1}^n (y_t^p - y_t^o) \tag{20}$$

where,  $y_t^o$ ,  $y_t^p$ , and  $n$  denote observed and the predicted values and the total number of observations,  $\bar{y}$  denotes the mean value.

## 4. Results

In this study, wave height and wave period data were derived from a buoy station located in Filyos region (see Figure 1). The wave height parameters ( $H_s$ ,  $H_{mean}$ ,  $H_{1/10}$ ,  $H_{max}$ ) and their corresponding periods ( $T_s$ ,  $T_{mean}$ ,  $T_{1/10}$ ,  $T_{max}$ ) were obtained. After arranging the obtained data in an excel sheet, seven sub-models were built that use the previous values of the selected parameters as predictors of wave height parameters (see Table 2). These proposed sub-models were used for both ANFIS, and MLPNN models, which are described in Tables 3, and 4, respectively. The wave data used was divided into two parts; first part being the training set consisting of 5565 readings (70% of the wave data available) and the second part being the testing set consisting of 2384 readings (30% of the wave data available).

**Table 3.** Characteristics of the MLPNN Models.

Wave Parameters	Model No	Network Configuration	Algorithm	Transfer Func.
$H_s$	Mod-6	3-3-1	Back propagation	Log. sig.
$H_{mean}$	Mod-7	3-3-1	Back propagation	Hyp.tan.sig.
$H_{1/10}$	Mod-3	2-2-1	Cascade correlation	Hyp.tan.sig.
$H_{max}$	Mod-1	2-2-1	Cascade correlation	Hyp.tan.sig.

**Table 4.** Characteristics of the ANFIS Models.

Wave Parameters	Model No	Membership Func.		Consequent Part	Epoch
		Type	Number		
$H_s$	Mod-6	Gumbell	5	Constant	300
$H_{mean}$	Mod-7	Gauss	8	Constant	100



$H_{1/10}$	Mod-7	Gumbell	5	Constant	500
$H_{max}$	Mod-7	Gauss	8	Constant	150

The sub-models developed in this study for the prediction of different representative wave height parameters were used to evaluate various statistical performance indicators to assess the accuracy and performance of these sub-models.

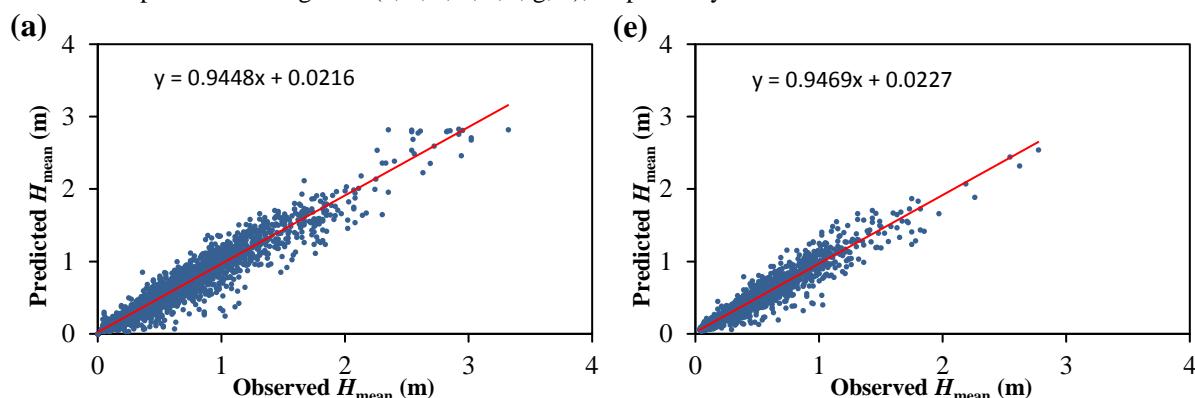
4.1 Results obtained using MLPNN sub-models

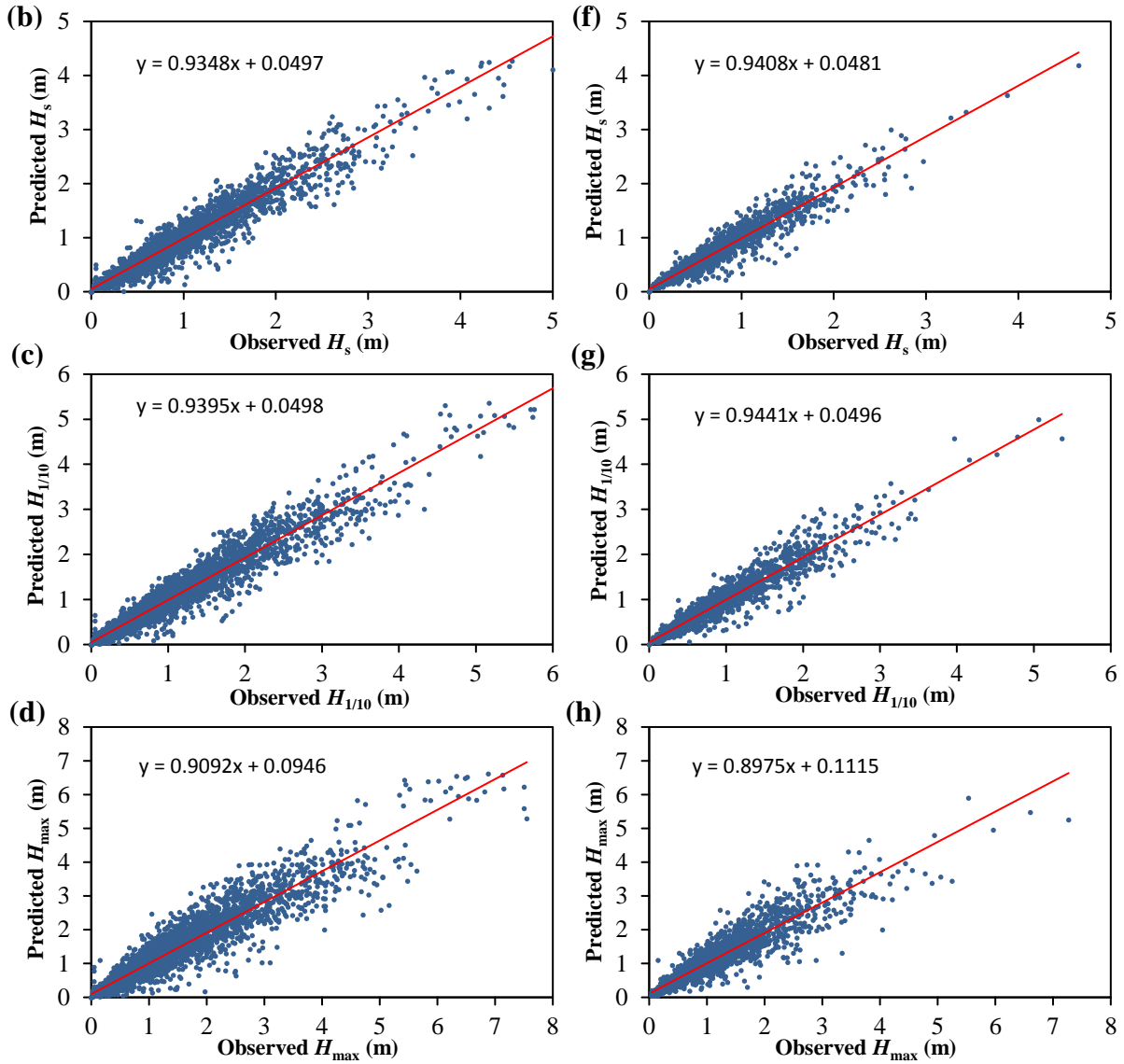
The MLPNN sub-models were trained and tested. All inputs were normalized between 0 and 1. The seven sub-models explained in Table 2 were used to predict the wave height parameters and acceptable results were obtained from all sub-models. However, Mod-7 for  $H_{mean}$ , Mod-6 for  $H_s$ , Mod-3 for  $H_{1/10}$ , and Mod-1 for  $H_{max}$  performed better than the others. In these predicted sub-models back propagation and cascade correlation algorithm were used. In addition, hyperbolic tangent sigmoid transfer function performed favorably, (ie. the error is lower) than logarithmic sigmoid transfer function as can be seen from Table 5.

Table 5. The Summary of Model Performance Measures.

Method	Predicted Parameter	Sub-Model	Training				Testing			
			RMSE (%)	MAE (%)	R	Bias (m)	RMSE (%)	MAE (%)	R	Bias (m)
MLPNN	$H_s$	Mod-6	3.33	1.82	0.95	0.09	3.16	1.90	0.93	0.09
	$H_{mean}$	Mod-7	3.19	1.76	0.95	0.08	3.07	1.71	0.93	0.08
	$H_{1/10}$	Mod-3	3.52	1.98	0.95	0.09	3.34	1.92	0.93	0.09
	$H_{max}$	Mod-1	4.48	2.53	0.91	0.19	4.30	2.52	0.87	0.18
ANFIS	$H_s$	Mod-6	3.15	1.79	0.95	0.09	3.15	1.80	0.94	0.09
	$H_{mean}$	Mod-7	2.95	1.65	0.96	0.05	3.45	1.73	0.94	0.06
	$H_{1/10}$	Mod-7	3.35	1.92	0.95	0.12	3.42	1.94	0.93	0.12
	$H_{max}$	Mod-7	3.94	2.50	0.93	0.17	4.28	2.47	0.89	0.19

The observed and predicted wave parameters for training and testing for the developed MLPNN sub-models are presented in Figure 5 (a, b, c, d, e, f, g, h), respectively.



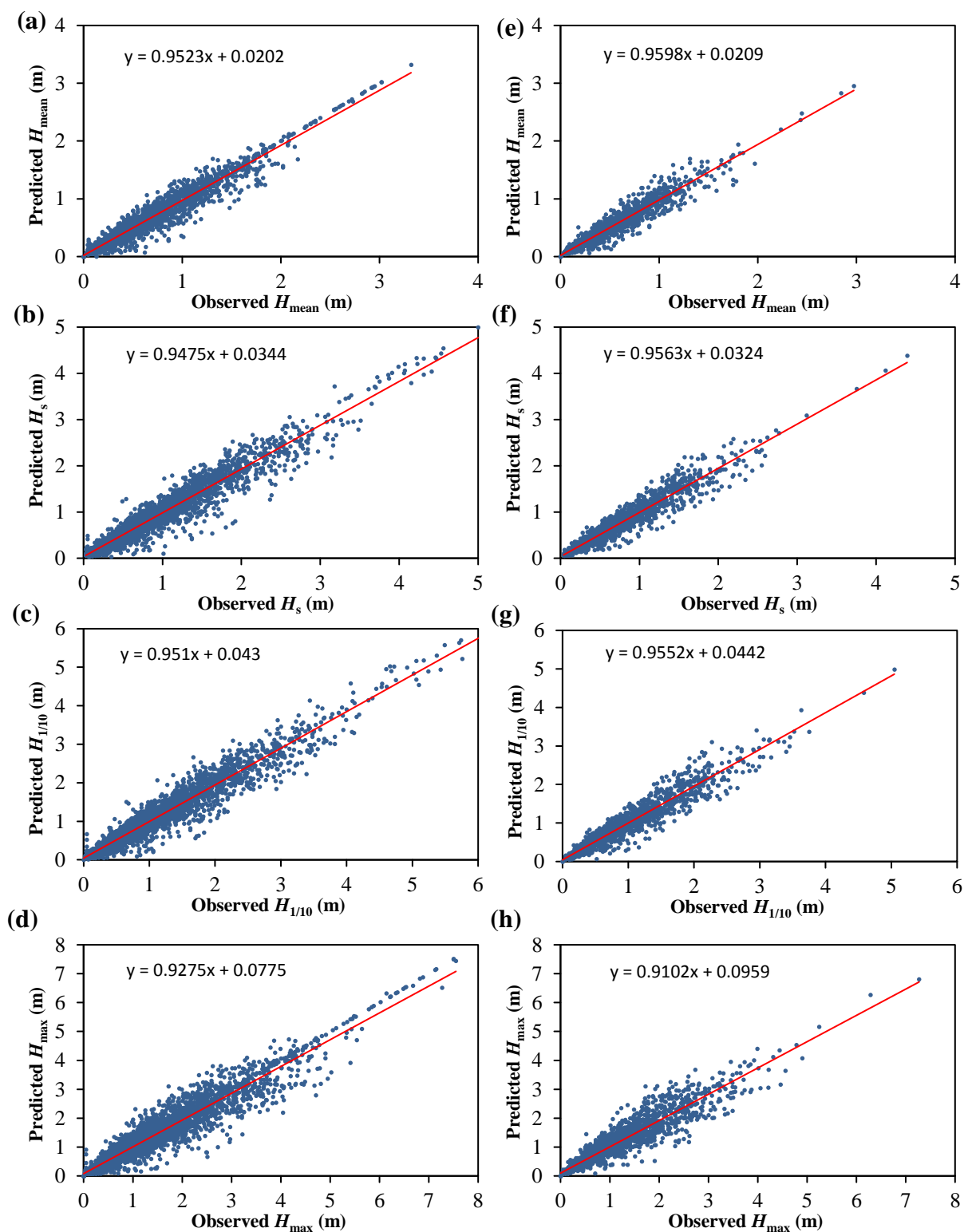


**Figure 5.** Comparison Between the Observed and Predicted Wave Parameters for; (a) Training Using MLPNN  $H_{mean}$ ; (b) Training Using MLPNN  $H_s$ ; (c) Training Using MLPNN  $H_{1/10}$ ; (d) Training Using MLPNN  $H_{max}$ ; (e) Testing Using MLPNN  $H_{mean}$ ; (f) Testing Using MLPNN  $H_s$ ; (g) Testing Using MLPNN  $H_{1/10}$ ; (h) Testing Using MLPNN  $H_{max}$ .

These results were also summarized in Table 5 by selecting the best performing sub-models for each wave height parameter. In Table 5, the  $R$  values for  $H_{max}$  were obtained as 0.91 for training and 0.87 for testing, which means that the results may be used for design purposes. For the rest of the predicted parameters ( $H_{mean}$ ,  $H_s$ ,  $H_{1/10}$ ),  $R$  values were obtained as 0.95 for training and 0.93 for testing. The lowest  $R$  was obtained for  $H_{max}$ , which is an expected result as  $H_{max}$  varies more than the other wave parameters. In addition to  $R$ , the performance indicators, such as RMSE, MAE, and Bias were presented in Table 5. According to these indicators for both testing and training the error, it was observed that a decrease in the following order: 1<sup>st</sup>:  $H_{max}$ , 2<sup>nd</sup>:  $H_{1/10}$ , 3<sup>rd</sup>:  $H_s$ , and 4<sup>th</sup>:  $H_{mean}$ .

#### 4.2 Results obtained using the ANFIS sub-models

Similarly, the MLPNN model seven sub-models of ANFIS were developed using same input combinations. It was observed that Mod-6 and Mod-7 provides better results by taking all seven sub-models into account. Moreover, Gumbell and Gauss membership functions were preferred for these sub-models thanks to better performances (Table 4). For the ANFIS sub-models, Figure 6 (a, b, c, d, e, f, g, h), illustrated the accuracies of the trained and the tested data sets, respectively.



**Figure 6.** Comparison Between the Observed and Predicted Wave Parameters for; (a) Training Using ANFIS  $H_{\text{mean}}$ ; (b) Training Using ANFIS  $H_s$ ; (c) Training Using ANFIS  $H_{1/10}$ ; (d) Training Using ANFIS  $H_{\text{max}}$ ; (e) Testing Using ANFIS  $H_{\text{mean}}$ ; (f) Testing Using ANFIS  $H_s$ ; (g) Testing Using ANFIS  $H_{1/10}$ ; (h) Testing Using ANFIS  $H_{\text{max}}$ .

The modeling results were presented separately for RMSE, MAE and R in Figure 7(a, b), 8(a, b), and 9(a, b), respectively. Figure 7, 8 and 9 showed that the observations and sub-model outputs are sufficiently close to each other. The best performing sub-models for each wave height parameter were presented in Table 5 where, the R values for  $H_{\text{max}}$  are obtained as 0.93 for training and 0.89 for testing, which means that the results are acceptable for design purposes. For  $H_{\text{mean}}$ ,  $H_s$ , and  $H_{1/10}$ , R values are obtained as 0.96 for training and 0.94 for

testing. Again, as was the case for the MLPNN sub-models, the lowest  $R$  for ANFIS sub-models is obtained for  $H_{max}$ . The best results in terms of performance indices were obtained for three inputs and results are presented in Figure 7(a, b), 8(a, b), and 9(a, b).

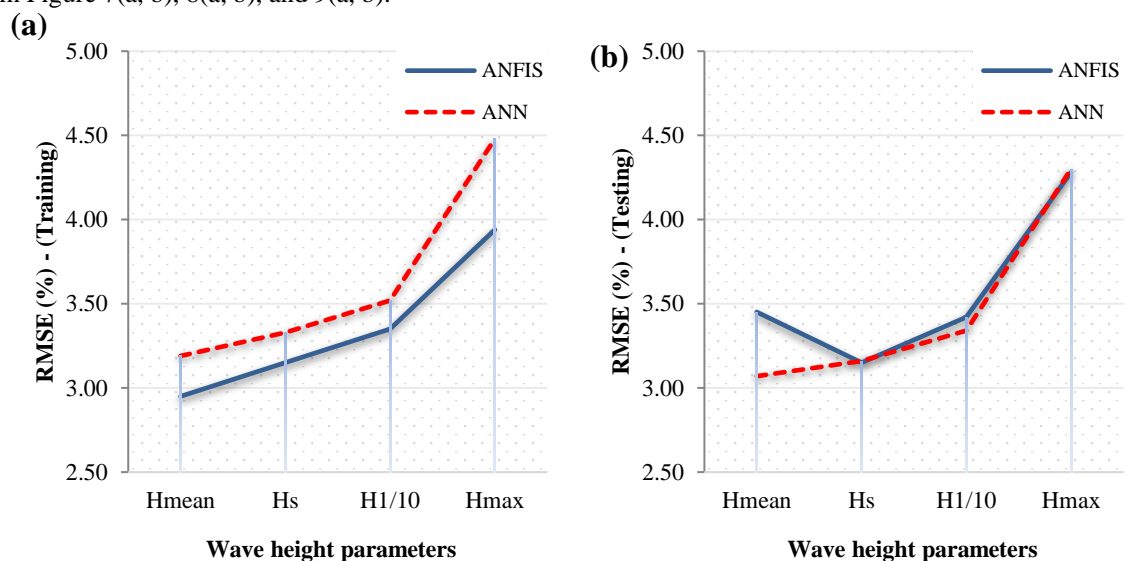


Figure 7. Root Mean Square Error Values for the MLPNN and ANFIS Models; (a) Training Data; (b) Testing Data.

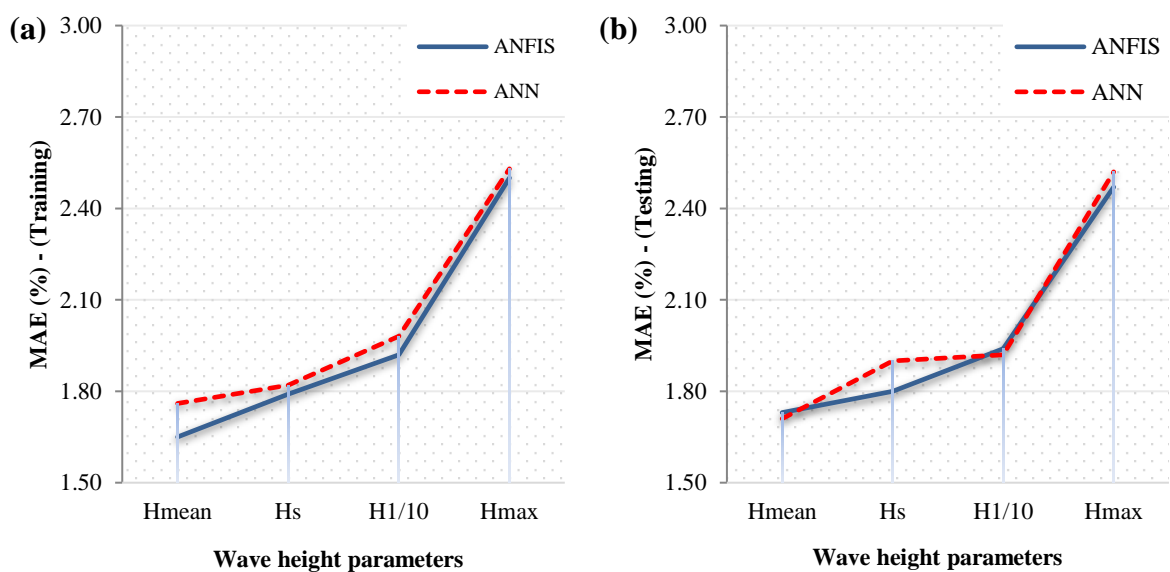
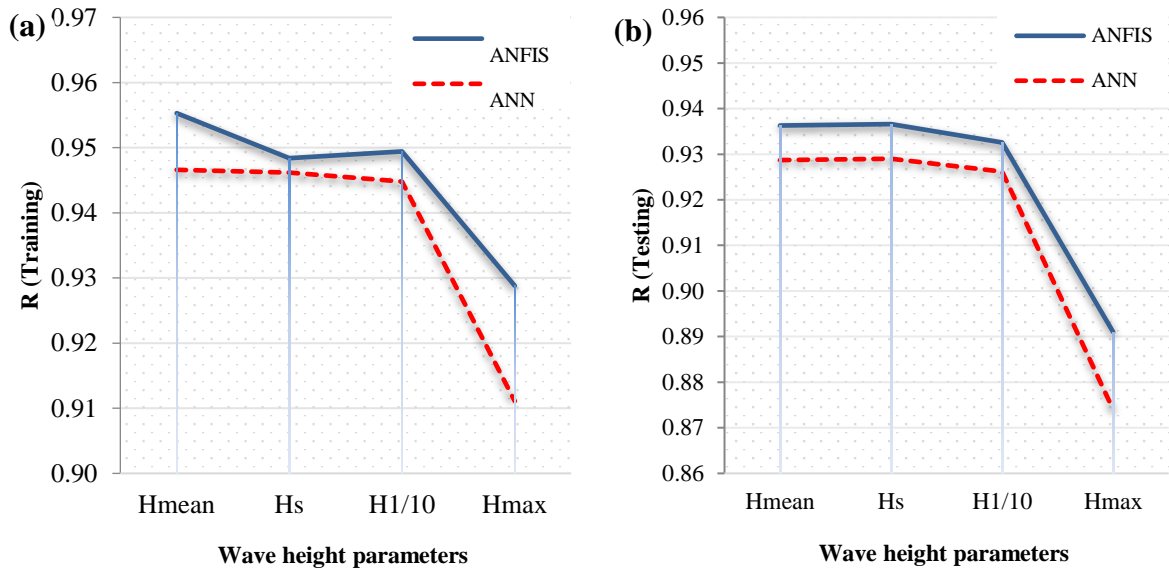


Figure 8. Mean Absolute Error Values for the MLPNN and ANFIS Models; (a) Training Data; (b) Testing Data.



**Figure 9.** Correlation Coefficients for the MLPNN and ANFIS Models; (a) Training Data; (b) Testing Data.

It can be observed from these figures that, RMSE is less than 4% for the training data set and about 4.3% for the testing data set. Also, the MAE values were found out to be approximately the same for both data sets, (i.e., training and testing.) As a result, it can be concluded that these sub-models give reasonable prediction performance for the represented wave heights.

## 5. Discussion

There are two different issues to discuss in this study. First one is the prediction of significant wave height,  $H_s$ , which has been commonly studied in the literature. Several soft computing models for the prediction of significant wave height were published by researchers. Almost all sub-models used in this study for MLPNN and ANFIS have an R value of greater than 0.90. Similarly, low error can be seen in terms of RMSE, MAE, and Bias parameter for the developed sub-models. Therefore, our findings are well in agreement with relevant literature. In addition, the performance of the MLPNN (which is an ANN-based method) was tested in this study. It can be accepted as an accurate modeling method for the significant wave height prediction. Although both the MLPNN and ANFIS models performed reasonably well in training and testing as evident by the presented representative wave height parameters, the accuracy of the ANFIS sub-models is found to be better than the MLPNN sub-models.

The other issue is the prediction of wave height parameters other than significant wave height such as  $H_{mean}$ ,  $H_{1/10}$ , and  $H_{max}$ . In the related literature, studies are found where revision necessity for the coastal structure design formulae due to the rise in the wave height parameters as a consequence of climate change is reported [41]–[44]. Also, we realize that there is a scarcity in the literature to predict other wave height parameters required for coastal activities. To this end, the prediction of  $H_{mean}$ ,  $H_{1/10}$ , and  $H_{max}$  has a paramount importance. Both MLPNN and ANFIS models used in this study showed good performance for prediction in terms of performance evaluation criteria. The results presented in Table 5 show that the error parameters for both MLPNN and ANFIS sub-models in predicting the representative wave height parameters decrease as the wave height parameters, such as  $H_{max}$  and  $H_{1/10}$  increase. It can also be observed from Table 5 that, for both training and testing data sets, relatively weaker correlation, comparing with the other wave height parameters, is obtained for  $H_{max}$ . The most accurate sub-model out of the proposed sub-models is  $\{H_{(t)}, H_{(t-1)}, H_{(t-3)}, T_{(t-1)}\}$  for  $H_s$  and  $\{H_{(t)}, H_{(t-1)}, H_{(t-3)}, H_{(t-5)}\}$  for  $H_{mean}$  for both ANFIS and MLPNN.

## 6. Conclusions

In this study, representative wave height of the Filyos region along the Southwest Coast of the Black Sea were predicted by using a 2-hourly wave dataset. Trial and error algorithm was used to select input combinations. Seven sub-models were created that include the previous wave height and wave period data as predictor. The target or output of these sub-models were selected as the one-time ahead wave height parameters. Entire data set divided into two as train and test as a data pre-process. Train data set includes 70% of the entire data, whereas test data set includes 30%. The predictor and predictand parameters were selected randomly to apply models.

Then, MLPNN and ANFIS models were developed, and each sub-model were tested individually. Before the MLPNN model was created, all train and test data set were normalized. The performance evaluation of the models was performed in terms of R, RMSE, MAE, Bias. The results showed that the wave height parameters predicted by the MLPNN and ANFIS models are similar and both methods yield results, which may be acceptable for design purposes. However, for maximum wave height,  $H_{\max}$ , ANFIS sub-models yield results having slightly lower error. It is shown in this study that, the error parameters decrease for both MLPNN and ANFIS sub-models as the wave height parameters increase and relatively weaker correlation, comparing with the other wave height parameters, is obtained for  $H_{\max}$ . The most accurate sub-model out of the proposed sub-models is also determined and presented.

As the results obtained in this study are very promising due to the randomness and uncertainty involved in predicting the wave height parameters. Both MLPNN and ANFIS may be better alternatives than the related conventional deterministic or statistical techniques used in coastal engineering design. Among these models, the ANFIS model may be chosen as it yields slightly better results and also is less time consuming than the MLPNN model. This study only considers the previous wave heights and wave periods as a predictor of target. Meteorological parameters such as wind speed, wind direction, air temperature etc. are not considered in this study. In the future works, it will be a good practice to investigate the contribution of meteorological factors on wave height parameters including  $H_{\text{mean}}$ ,  $H_{1/10}$ , and  $H_{\max}$ .

**Acknowledgements:** The authors are grateful to the Republic of Turkey Ministry of Transport and Infrastructure -General Directorate of Understructure Investments Department of Research for the data sets they have generously given.

**Conflict of interest:** The authors declare no conflict of interest.

## References

- [1] J. Mahjoobi, A. Etemad-Shahidi, and M. H. Kazeminezhad, "Hindcasting of wave parameters using different soft computing methods," *Appl. Ocean Res.*, vol. 30, no. 1, pp. 28–36, 2008.
- [2] Usac. CEM, "Coastal engineering manual." US Army Engineer Research and Development Center, Coastal and Hydraulics Laboratory, Vicksburg, MS, 2003.
- [3] H. U. Sverdrup and W. H. Munk, *Wind, sea and swell: Theory of relations for forecasting*, no. 303. Hydrographic Office, 1947.
- [4] C. E. R. C. (US), *Shore Protection Manual*. Department of the Army, Waterways Experiment Station, Corps of Engineers ..., 1984.
- [5] K. F. Hasselmann *et al.*, "Measurements of wind-wave growth and swell decay during the Joint North Sea Wave Project (JONSWAP).," *Ergaenzungsh. zur Dtsch. Hydrogr. Zeitschrift, R. A*, 1973.
- [6] Y. Goda, "Revisiting Wilson's formulas for simplified wind-wave prediction," *J. Waterw. port, coastal, Ocean Eng.*, vol. 129, no. 2, pp. 93–95, 2003.
- [7] N. Booij, R. C. Ris, and L. H. Holthuijsen, "A third-generation wave model for coastal regions: 1. Model description and validation," *J. Geophys. Res. Ocean.*, vol. 104, no. C4, pp. 7649–7666, 1999.
- [8] T. W. Group, "The WAM model—A third generation ocean wave prediction model," *J. Phys. Oceanogr.*, vol. 18, no. 12, pp. 1775–1810, 1988.
- [9] A. Akpınar, M. Özger, and M. I. Kömürçü, "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.*, vol. 19, no. 1, pp. 1–14, 2014.
- [10] O. Makarynsky, A. A. Pires-Silva, D. Makarynska, and C. Ventura-Soares, "Artificial neural networks in wave predictions at the west coast of Portugal," *Comput. Geosci.*, vol. 31, no. 4, pp. 415–424, 2005.
- [11] C. E. Balas, L. Koç, and L. Balas, "Predictions of missing wave data by recurrent neuronets," *J. Waterw. port, coastal, Ocean Eng.*, vol. 130, no. 5, pp. 256–265, 2004.
- [12] S. N. Londhe, "Soft computing approach for real-time estimation of missing wave heights," *Ocean Eng.*, vol. 35, no. 11–12, pp. 1080–1089, 2008.
- [13] S. N. Londhe and V. Panchang, "Correlation of wave data from buoy networks," *Estuar. Coast. Shelf Sci.*, vol. 74, no. 3, pp. 481–492, 2007.
- [14] M. C. Deo and G. Chaudhari, "Tide prediction using neural networks," *Comput. Civ. Infrastruct. Eng.*, vol. 13, no. 2, pp. 113–120, 1998.
- [15] C.-P. Tsai and T.-L. Lee, "Back-propagation neural network in tidal-level forecasting," *J. Waterw. Port, Coastal, Ocean Eng.*, vol. 125, no. 4, pp. 195–202, 1999.
- [16] D. H. Kim and W. S. Park, "Neural network for design and reliability analysis of rubble mound breakwaters," *Ocean Eng.*, vol. 32, no. 11–12, pp. 1332–1349, 2005.
- [17] M. L. Koç and C. E. Balas, "Reliability analysis of a rubble mound breakwater using the theory of fuzzy random variables," *Appl. Ocean Res.*, vol. 39, pp. 83–88, 2013.
- [18] C. E. Balas and M. L. Koc, "Reliability analysis of rubble mound breakwaters by neural network model," *J. Coast. Res.*, pp. 1506–1509, 2006.
- [19] S. Hadadpour, A. Etemad-Shahidi, and B. Kamranzad, "Wave energy forecasting using artificial neural networks

- in the Caspian Sea,” *Proc. Inst. Civ. Eng. - Marit. Eng.*, vol. 167, no. 1, pp. 42–52, 2014.
- [20] B. Aydođan, B. Ayat, and Y. Yüksel, “Black Sea wave energy atlas from 13 years hindcasted wave data,” *Renew. energy*, vol. 57, pp. 436–447, 2013.
- [21] H. Mase, M. Sakamoto, and T. Sakai, “Neural network for stability analysis of rubble-mound breakwaters,” *J. Waterw. port, coastal, Ocean Eng.*, vol. 121, no. 6, pp. 294–299, 1995.
- [22] S. Mandal, S. Rao, Y. R. Manjunath, and D. H. Kim, “Stability analysis of rubblemound breakwater using ANN,” 2007.
- [23] D. Kim, D. H. Kim, and S. Chang, “Application of probabilistic neural network to design breakwater armor blocks,” *Ocean Eng.*, vol. 35, no. 3–4, pp. 294–300, 2008.
- [24] T. Erdik, “Fuzzy logic approach to conventional rubble mound structures design,” *Expert Syst. Appl.*, vol. 36, no. 3, pp. 4162–4170, 2009.
- [25] A. Etemad-Shahidi and M. Bali, “Stability of rubble-mound breakwater using H50 wave height parameter,” *Coast. Eng.*, vol. 59, no. 1, pp. 38–45, 2012.
- [26] C. E. Balas, M. L. Koç, and R. Tür, “Artificial neural networks based on principal component analysis, fuzzy systems and fuzzy neural networks for preliminary design of rubble mound breakwaters,” *Appl. Ocean Res.*, vol. 32, no. 4, pp. 425–433, 2010.
- [27] M. L. Koç and C. E. Balas, “Genetic algorithms based logic-driven fuzzy neural networks for stability assessment of rubble-mound breakwaters,” *Appl. Ocean Res.*, vol. 37, pp. 211–219, 2012.
- [28] C.-P. Tsai, C. Lin, and J.-N. Shen, “Neural network for wave forecasting among multi-stations,” *Ocean Eng.*, vol. 29, no. 13, pp. 1683–1695, 2002.
- [29] S. Mandal and N. Prabakaran, “Ocean wave forecasting using recurrent neural networks,” *Ocean Eng.*, vol. 33, no. 10, pp. 1401–1410, 2006.
- [30] A. Zamani, D. Solomatine, A. Azimian, and A. Heemink, “Learning from data for wind–wave forecasting,” *Ocean Eng.*, vol. 35, no. 10, pp. 953–962, 2008.
- [31] K. Günaydın, “The estimation of monthly mean significant wave heights by using artificial neural network and regression methods,” *Ocean Eng.*, vol. 35, no. 14–15, pp. 1406–1415, 2008.
- [32] B. Kamranzad, A. Etemad-Shahidi, and M. H. Kazeminezhad, “Wave height forecasting in Dayyer, the Persian Gulf,” *Ocean Eng.*, vol. 38, no. 1, pp. 248–255, 2011.
- [33] S. Mafi and G. Amirinia, “Forecasting hurricane wave height in Gulf of Mexico using soft computing methods,” *Ocean Eng.*, vol. 146, pp. 352–362, 2017.
- [34] S. Shamshirband, A. Mosavi, T. Rabczuk, N. Nabipour, and K. Chau, “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.*, vol. 14, no. 1, pp. 805–817, 2020.
- [35] P. Dixit, S. Londhe, and Y. Dandawate, “Removing prediction lag in wave height forecasting using Neuro - Wavelet modeling technique,” *Ocean Eng.*, vol. 93, pp. 74–83, 2015.
- [36] J.-S. Jang, “ANFIS: adaptive-network-based fuzzy inference system,” *IEEE Trans. Syst. Man. Cybern.*, vol. 23, no. 3, pp. 665–685, 1993.
- [37] M. H. Kazeminezhad, A. Etemad-Shahidi, and S. J. Mousavi, “Application of fuzzy inference system in the prediction of wave parameters,” *Ocean Eng.*, vol. 32, no. 14–15, pp. 1709–1725, 2005.
- [38] R. TÜR and C. A. N. BALAS, “Neuro-fuzzy approximation for prediction of significant wave heights: the case of Filyos region,” vol. 25, no. 3, pp. 505–510, 2010.
- [39] R. Hashim, C. Roy, S. Motamedi, S. Shamshirband, and D. Petković, “Selection of climatic parameters affecting wave height prediction using an enhanced Takagi-Sugeno-based fuzzy methodology,” *Renew. Sustain. Energy Rev.*, vol. 60, pp. 246–257, 2016.
- [40] C. Stefanakos, “Fuzzy time series forecasting of nonstationary wind and wave data,” *Ocean Eng.*, vol. 121, pp. 1–12, 2016.
- [41] P. Lionello, S. Cogo, M. B. Galati, and A. Sanna, “The Mediterranean surface wave climate inferred from future scenario simulations,” *Glob. Planet. Change*, vol. 63, no. 2–3, pp. 152–162, 2008.
- [42] P. Galiatsatou, C. Anagnostopoulou, and P. Prinos, “Modeling nonstationary extreme wave heights in present and future climates of Greek Seas,” *Water Sci. Eng.*, vol. 9, no. 1, pp. 21–32, 2016.
- [43] J. B. Debernard and L. P. Røed, “Future wind, wave and storm surge climate in the Northern Seas: a revisit,” *Tellus A Dyn. Meteorol. Oceanogr.*, vol. 60, no. 3, pp. 427–438, 2008.
- [44] M. Casas-Prat and J. P. Sierra, “Projected future wave climate in the NW Mediterranean Sea,” *J. Geophys. Res. Ocean.*, vol. 118, no. 7, pp. 3548–3568, 2013.
- [45] E. Bilyay, B. O. Ozbahceci, and A. C. Yalciner, “Extreme waves at filyos, southern black sea,” *Nat. Hazards Earth Syst. Sci.*, vol. 11, no. 3, pp. 659–666, 2011.
- [46] E. Özhan and S. Abdalla, *Türkiye kıyıları rüzgar ve derin deniz dalga atlası*. MEDCOAST, ODTÜ, 2002.
- [47] R. TÜR and C. A. N. BALAS, “RELIABILITY-BASED RISK ASSESSMENT OF REVETMENTS,” *J. Fac. Eng. Archit. Gazi Univ.*, vol. 25, no. 3, pp. 511–516, 2010.
- [48] J. Shiri, O. Makarynsky, O. Kisi, W. Dierickx, and A. F. Fard, “Prediction of Short-Term Operational Water Levels Using an Adaptive Neuro-Fuzzy Inference System,” *J. Waterw. Port, Coastal, Ocean Eng.*, vol. 137, no. 6, pp. 344–354, 2011.
- [49] C. E. Balas, A. T. Williams, A. Ergin, and M. L. Koc, “Litter categorization of beaches in Wales, UK by multi-layer neural networks,” *J. Coast. Res.*, vol. 38, pp. 1515–1519, 2006.
- [50] L. H. Tsoukalas and R. E. Uhrig, “Fuzzy and neural approaches in engineering,” 1997.
- [51] D. R. Legates and G. J. McCabe, “Evaluating the use of ‘goodness-of-fit’ measures in hydrologic and hydroclimatic model validation,” *Water Resour. Res.*, vol. 35, no. 1, pp. 233–241, 1999.

- [52] C. J. Willmott, "Some Comments on the Evaluation of Model Performance," *Bull. Am. Meteorol. Soc.*, vol. 63, no. 11, pp. 1309–1313, 1982.