Simultaneous optimization of clustering and fuzzy IF-THEN rules parameters by the genetic algorithm in fuzzy inference system-based wave predictor models

Morteza Zanganeh

ABSTRACT

Prediction of wave parameters is of great importance in the design of marine structures. In this paper, two shortcomings with the adaptive network-based fuzzy inference system (ANFIS) model for prediction of wave parameters are remedied by employing a genetic algorithm (GA). The first shortcoming in the ANFIS model goes back to its problem for automatic extraction of fuzzy IF-THEN rules and the second one is related to its gradient-based nature for tuning the antecedent and consequent parameters of fuzzy IF-THEN rules. To deal with these shortcomings, in this study a combined FIS and GA model is developed in which the capability of the GA as an evolutionary algorithm is used for simultaneous optimization of the subtractive clustering parameters and the antecedent and consequent parameters of fuzzy IF-THEN rules. Following the development of the combined model, this model is used to predict wave parameters, i.e., significant wave height and peak spectral period at Lake Michigan. The obtained results show that the developed model outperforms the ANFIS model and the Coastal Engineering Manual (CEM) method to estimate the function representing the generation process of the wind-driven waves.

Key words | ANFIS, fuzzy membership functions parameters, genetic algorithm, subtractive clustering parameters, wave prediction

Morteza Zanganeh
Department of Civil Engineering, Faculty of Engineering,
Golestan University,
Golestan,
Iran
E-mail: *m.zanganeh@gu.ac.ir*

NOMENCLATURE

A_i, B_i, C_i, D_i, E_i	fuzzy set variables
C_D	drag coefficient
a_i	standard deviations of the Gaussian membership functions
c_i	mean values of the Gaussian membership functions
D	total number of input and output variables in data set
D_{wi}	wind direction at the ith hour
$\bar{D}_w$	average wind direction for consecutive preceding i hours
E	mean square error
g	gravitational acceleration
H_s	significant wave height
N_{trn}	number of training data
N_{test}	number of testing data
$N_{validation}$	number of validation data
$Numrule$	number of fuzzy IF-THEN rules
$MaxNumrule$	maximum number of fuzzy IF-THEN rules
$o_i, p_i, q_i, r_i, s_i, t_i$	the linear consequent parameters
w_i	firing strength of rule i
$\mu(x)$	degree of membership
O	observed value
P	predicted value
PV_k	the potential of kth data point
PV'_k	the modified potential value of kth data point
PV_1^*	the first cluster center potential value

doi: 10.2166/hydro.2017.045

ra_i	clustering radius for the ith input variable
$ra_{\min_i}$	the minimum clustering radius for the ith variable
η	learning rate
n_e	the number of training epochs
S_1	the set of nonlinear antecedent parameters
S_2	the set of linear consequent parameters
T_p	peak spectral wave period
$t_{\min}$	required time for fetch-limited condition
U	wind speed
U_i	wind speed at the ith hour
$\bar{U}$	the average of wind speed for consecutive preceding i hours
$u*$	shear velocity
F	fetch length
t	wind duration
α	fuzzy IF-THEN rule parameters
d	distance between the candidate clusters
r_{ai}	radii of clusters
γ	quash factor
$\bar{\varepsilon}$	acceptance ratio
$\underline{\varepsilon}$	rejection ratio
X	corrected fetch length in meter
x	value of data point
$\boldsymbol{x}_{C1}$	cluster center with highest potential value

INTRODUCTION

Prediction of wave parameters, including significant wave height and peak spectral period, plays an important role in ocean activities. In that regard, several methodshave been introduced to predict wave parameters. Simplified methods like SPM (Shore Protection Manual 1984) and CEM (Coastal Engineering Manual) (US Army 2003) were the first methods. These techniques are based on constant wind definition and their evaluation versus observed data set in various places proves their deficiency to predict wave parameters (Bishop 1983; Etemad-Shahidi *et al.* 2009).

Numerical models like the SWAN (simulating wave nearshore) model, which work based on solving energy equilibrium equations are the second type of model to predict wave parameters. In the application of these methods, not only the bathymetry of a lake is needed but also several additional parameters such as bottom roughness are needed (Ris *et al.* 1996).

The emergence of soft computing-based methods motivated scientists to employ these methods for prediction of wave parameters. These tools have been broadly used as an alternative method for modeling various complex civil engineering problems. Adaptive neural network (ANN) (Agrawal & Deo 2002; Tsai *et al.* 2002; Altunkaynak 2013) and fuzzy inference systems (FISs) are examples of these applications. The FIS models have been widely employed by many researchers in the field of water engineering including, modeling of rainfall–runoff events (Sen & Altunkaynak 2004), estimating scour uncertainty around bridge piers (Johnson & Ayyub 1996), predicting scour depth at abutments of armored beds (Muzzammil 2010; Muzzammil & Alam 2011), optimizing water allocation systems (Kindler 1992), controlling reservoir operation (Pesti *et al.* 1996), analyzing regional drought (Pongracz *et al.* 1999), estimating pile group scour (Chen & Dao 2007), predicting stream flow (Shi *et al.* 1999; Özger 2009), finding scour location at the downstream of a spillway (Russo 2000), and forecasting flood flow (Rezaeianzadeh *et al.* 2014). Most of the reviewed studies prove that the FIS models outperform the nonlinear regression approaches.

In the field of coastal engineering, Kazeminezhad *et al.* (2005) used FIS to estimate wave parameters while its structure is being optimized by a hybrid model. Özger & Sen (2007) applied fuzzy logic to investigate the relationship between the wind speed and previous and current wave characteristics in the Pacific Ocean. Mahjoobi *et al.* (2008) applied ANN and ANFIS models to hindcast wave parameters while their models' input variables were wind speed, wind direction, fetch length, and wind duration. Their findings showed the ANFIS models' superiority to the FIS and ANN models. Zanaganeh *et al.* (2009) used genetic algorithm–adaptive network-based fuzzy inference systems (GA-ANFIS) for prediction of wave parameters. In their model, a GA was used as the optimizer to tune subtractive clustering parameters, including radii of clustering and the quash factor, while a hybrid gradient-based method known as the ANFIS was used to tune the consequent and antecedent parameters of the fuzzy IF-THEN rules, simultaneously. Their obtained results indicated that the

GA-ANFIS model is more accurate than the ANFIS model in which the clustering parameters were generated randomly. In another study, Zanganeh *et al.* (2011) employed a PSO-FIS-PSO model to estimate the equilibrium depth of scouring beneath pipelines. In their applied model, two particle swarm optimization (PSO) algorithms were employed to tune the subtractive clustering parameters and the antecedent and consequent parameters of fuzzy IF-THEN rules, simultaneously. Their model outperformed the empirical methods to estimate the equilibrium depth of scour. An important deficiency in the developed PSO-FIS-PSO model is tuning the parameters of the two employed PSO algorithms such as initial population, cognitive and social parameters, and so on. Therefore, modifying the model by using one evolutionary algorithm can be beneficial. Akpinar *et al.* (2014) employed FIS and empirical models along the south coast of the Black Sea to predict wave parameters and the obtained results proved the superiority of the ANFIS model compared to empirical models including the CEM and SPM methods. More recently, Zanganeh *et al.* (2016) applied ANN and ANFIS models to estimate coastal current velocities on the Ogata coast. The obtained results showed the models' efficiency to capture the physical complexity of coastal currents' generation process in both longshore and cross-shore directions.

As mentioned above, FISs are the models in which a phenomenon is estimated by mapping a nonlinear relationship among effective input and output variables. These models represent a phenomenon by some known fuzzy IF-THEN rules that their optimization is important to capture the nature of the phenomenon. Fuzzy IF-THEN rules optimization process can be either accomplished by a gradient-based method (ANFIS model) or by an evolutionary algorithm like the GA (combined FIS and GA model). In the combined FIS and GA model, due to the domain-irrelevant behavior of the GA a global optimization can be achieved, whereas hill-climbing methods like the ANFIS require a specific domain in order to guide their search area. To date, various efforts have been devoted to develop combined FIS and evolutionary algorithms either to estimate functions (Homaifar & McCormick 1995; Shi *et al.* 1999; Russo 2000; Hidalgo *et al.* 2012; Zacharia & Nearchou 2012) or to optimize fuzzy logic controllers (Seng *et al.* 1999; Chen & Dao 2007; Poursamad & Montazeri 2008).

The main objective of this paper is to employ the features of the GA as an evolutionary algorithm to optimize structures of fuzzy IF-THEN rules, that the resulting model is termed as a combined GA and FIS model. In the model, the gradient-based learning algorithm in the ANFIS is replaced by a GA in order to tune fuzzy nonlinear antecedent and linear consequent parameters. Also, subtractive clustering parameters are being optimized within the training process. Finally, the model developed to estimate any function is used to predict wind-driven wave parameters. Note that the learning process in the ANFIS model could be either a hybrid learning algorithm introduced by Jang (1993) or the steepest descent (SD) method.

In this study, the introduction above provides a brief review of the previous works in the field of water engineering and wave predictions, with the section below giving an overview of the features of the FIS and the ANFIS methods. Next, are sections describing the GA as an optimizer model and an outline of the developed combined FIS and GA model to estimate every function. Then the performance of the model is evaluated to predict the wind-driven wave parameters followed by a section in which the CEM method is outlined to predict wave parameters. The final section contains the evaluation of developed models to predict wave parameters.

FUZZY INFERENCE SYSTEMS

FIS structures

FISs are mathematical theories allowing one to model a natural process through some linguistic expressions. These methods are suitable to find relationships among effective input variables and the desired output of a system. These models can assign qualitative aspects of human knowledge by some linguistic expressions, so-called fuzzy IF-THEN rules. The rules are usually extracted from data sets representing the phenomenon.

From a mathematical point of view, the following expressions introduce a simple Takagi and Sugeno type (Takagi & Sugeno 1985) of FIS with two fuzzy IF-THEN rules, in which x and y are inputs variables and f is the output variable described as follows:

Rule 1: If x is A_1 and y is B_1 then $f_1 = p_1x + q_1y + r_1$,
Rule 2: If x is A_2 and y is B_2 then $f_2 = p_2x + q_2y + r_2$,

where A_1, A_2 and B_1, B_2 are the fuzzy linguistic variables associated with x and y as input variables, respectively; and p_1, q_1, r_1 and p_2, q_2, r_2 are consequent parameters of fuzzy IF-THEN rules. Figure 1 illustrates the considered FIS architecture associated with the following five layers.

Layer 1: In this layer as the first stage of the learning process, the membership function associated with each input variable is calculated as follows:

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

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

where O_i^1 is the Gaussian membership function in the first layer for ith fuzzy variable (A_i), a_i and c_i are adjustable parameters associated with the membership function. The Gaussian membership function has a better performance in comparison with the other ones (Zanaganeh *et al.* 2009).

Layer 2: This layer is known as the product layer by which previously calculated degrees of memberships are multiplied as:

$$w_i = \mu_{A_i}(x) \times \mu_{B_i}(y), \;\; i = 1, 2 \tag{3}$$

where w_i is called the firing strength of rule i, $\mu_{A_i}(x)$ and $\mu_{B_i}(y)$ are the degrees of membership of x and y, respectively. Also, each circle node output labeled by Π represents the firing strength.

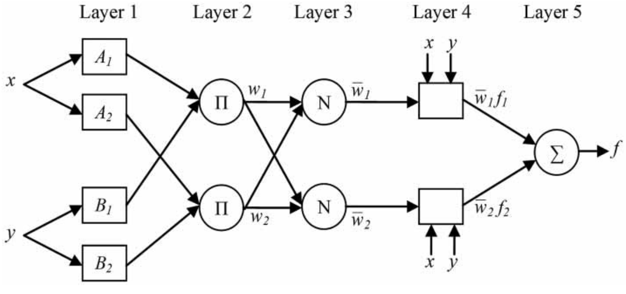

Figure 1 | Schematic diagram of the FIS model for two input variables (Jang 1993).

Layer 3: In this layer, the normalized layer, the ratio of each weight to the total weights is calculated as follows:

$$\overline{w_i} = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2 \tag{4}$$

and in this layer every node is circle like and labeled as N.

Layer 4: This layer is called the defuzzification layer which is an adaptive square like node and is expressed as follows:

$$\overline{w}_i f_i = \overline{w}_i(p_ix + q_iy + r_i), \quad i = 1, 2, \ldots, n \tag{5}$$

where p_i, q_i, and r_i are the adjustable linear consequent parameters.

Layer 5: In this layer, the final output of the network, f, is estimated by summation of all incoming signals as the following:

$$\begin{aligned} f &= \frac{w_1}{w_1 + w_2} f_1 + \frac{w_2}{w_1 + w_2} f_2 = \bar{w}_1 f_1 + \bar{w}_2 f_2 \\ &= \bar{w}_1(p_1x + q_1y + r_1) + \bar{w}_2(p_2x + q_2 + r_2) \\ &= (\bar{w}_1x)p_1 + (\bar{w}_1y)q_1 + (\bar{w}_1)r_1 + (\bar{w}_2x)p_2 + (\bar{w}_2y)q_2 + (\bar{w}_2)r \end{aligned} \tag{6}$$

The known ANFIS model is functionally a Takagi and Sugeno's (TS) type FIS whose parameters are tuned by a learning algorithm. In this model, parameters are divided into two separate categories as:

$$S = S_1 \oplus S_2 \tag{7}$$

where S is total parameters of the ANFIS, S_1 is the set of nonlinear antecedent parameters; S_2 is the set of linear consequent parameters and $\oplus$ is the summation operator.

Tuning the parameters in the ANFIS model is done based on having minimum error of predicting in the training process which is introduced as follows:

$$E = \frac{1}{N_{trn}} \sum_{k=1}^{N_{trn}} (O^k - P^k)^2 \tag{8}$$

$$\Delta\alpha = -\eta \frac{\partial E}{\partial \alpha} \tag{9}$$

where E is the mean square error (MSE), O^k is the kth network output at a given output node, P^k is the kth target output, N_{trn} is the number of training data, α is a given parameter (linear and nonlinear parameters) and η is the learning rate expressed as follows:

$$\eta = \frac{n_e}{\sqrt{\sum_{\alpha} (\partial E/\partial \alpha)^2}} \tag{10}$$

where n_e is the number of training epochs. The above process is called the SD method in the ANFIS model in order to update both antecedent and consequent parameters of fuzzy IF-THEN rules. More details about the ANFIS models can be found in Jang (1993), and in his work a hybrid algorithm was developed to tune the fuzzy antecedent and consequent parameters. The hybrid method combines the least square error (LSE) with a SD back propagation algorithm. The LSE method is used to tune fuzzy linear consequent parameters in the forward path and SD is used to tune fuzzy nonlinear antecedent parameters in the backward path.

Subtractive clustering method

The subtractive clustering method is one of the known methods for fuzzification of any input variable (Chiu 1994). In this method, clustering parameters, categorized as radii and quash factor, control the number and structure of fuzzy IF-THEN rules. In this clustering method, at first the cluster centers are extracted based on their potential values to be a cluster center. Then, the fuzzy IF-THEN rules associated with the clusters are determined. For a given collection of K data points x_k, k = 1, 2, …, K specified by a D-dimensional vector, where D is the total number of input and output variables, the potential value for each data point is estimated as follows:

$$PV_k = \sum_{j=1}^{K} \exp\left(-4\sqrt{\sum_{i=1}^{D}\left(\frac{x_k^i - x_j^i}{r_{ai}}\right)^2}\right) \tag{11}$$

where PV_k is the potential of kth data point and r_{ai} is the radius of cluster associated with ith dimension of the point. Based on the above equation the data point having the highest potential value is selected as the first cluster center while x_{C1} is the point and PV_1^* is its appropriate potential value.

In the second step, the potential value of each data point x_k is modified by using the following equation:

$$PV'_k = PV_k - PV_1^* \exp\left(-4\sqrt{\sum_{i=1}^{D}\left(\frac{x_k^i - x_{C1}^i}{\gamma r_{ai}}\right)^2}\right) \tag{12}$$

where PV'_k is the modified potential value of the kth data point and γ is quash factor, which is multiplied by the radii to determine the other neighboring clusters. Modifying the potential value of all data points, new cluster centers are determined based on some threshold parameters including acceptance ratio $\bar{\varepsilon}$, rejection ratio $\underline{\varepsilon}$, and the relative distance criterion. In this process, a data point having a potential value greater than the acceptance threshold is directly qualified as a cluster center. The acceptance level of data points with potential values between the upper and lower threshold is dependent on the relative distance equation, expressed as:

$$d_{\min} + \frac{PV_k^*}{PV_1^*} \geq 1 \tag{13}$$

where $d_{\min}$ is the minimum distance between the candidate cluster center and all previously found cluster centers, that is described as follows:

$$d_{k,c} = \left(\sqrt{\sum_{i=1}^{D}\left(\frac{x_k^i - x_{C1}^i}{r_{a_i}}\right)^2}\right) \tag{14}$$

where $d_{k,c}$ is the distance of kth data point from c_ith previously found cluster center. This process would continue until the potential values of all data points reach zero.

GENETIC ALGORITHM

The GA is a global stochastic optimization method which works based on the idea of 'survival best fitness' and 'natural selection'. GAs became popular through the work of John

Holland in the early 1970s, and in particular, his book *Adaptation in Natural and Artificial Systems* (Holland 1975). The GA is a kind of robust optimization algorithm well-suited for discontinuous behavior, domain-irrelevant and multimodal functions, even in noisy environments (Beyer 2000). In addition, it is a parallel search algorithm with a certain learning ability, which repeats selection, crossover and mutation operators in each generation after an initialization until the given stopping criteria are met (Figure 2). The following paragraphs outline the operators to produce the next generation children.

Selection

This operator selects the individuals to contribute to the next generation. In this study, *stochastic uniform selections* have been applied by lying individuals on a line according to their scaled value. Then, the algorithm moves along the line in equal steps. At each step, the algorithm selects parents from the section that they are lying in.

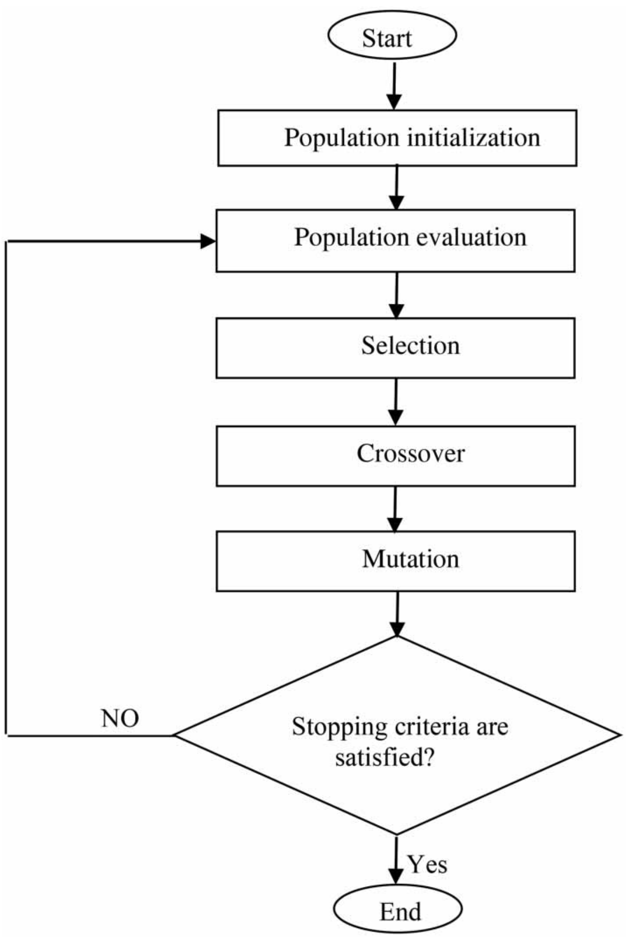

Figure 2 | Flow diagram of the GA optimization model.

Crossover

This operator combines two parents to form the next generation children. In this paper, the *scattered* method is used as the crossover operator, in which a binary string is created randomly while its length is equal to the length of the solution. After that, string 1 values are replaced by the values of the first parent whereas 0 values are substituted by values of the second parent.

Mutation

This operator assigns sudden change in the parents to form new random children. This operator generally increases the robustness of the algorithm for sticking in local optima. In this paper, the Gaussian mutation is used by adding a random number to each vector entry of an individual. This random number is taken from the Gaussian distribution centered on zero. The variance of this distribution can be controlled with two parameters. The **Scale** parameter determines the variance at the first generation. The **Shrink** parameter controls how variance shrinks during the generations. If the Shrink parameter is 0, the variance is constant and if the Shrink parameter is 1, the variance shrinks to 0 linearly as reaching to the last generation (Chipperfield *et al.* 1994).

The GA operators would be controlled by population size, crossover fraction and mutation fraction to find reasonable settings for the problem class being worked on. A very small mutation rate may lead to genetic drift. A recombination rate that is too high may lead to premature convergence of the GA. A mutation rate that is too high may lead to loss of good solutions, unless elitism selection is employed. Note that the elitism selection goes back to the copy of some of the best children into the next generation unchanged.

THE COMBINED MODEL OF FIS AND GA

As mentioned before, the ANFIS is a FIS-based model in which extracting of fuzzy IF-THEN rules structure is not an automatic process. In addition, in the ANFIS model, the antecedent and consequent parameters associated with the fuzzy IF-THEN rules are tuned by a gradient-based method. Thus, entrapping the answer in a local optimum may be possible. Accordingly, in order to overcome the above-mentioned deficiencies, the ANFIS model employing evolutionary algorithms like the GA could be a remedy. On the basis of the discussion in this paper, a GA is combined with a FIS to optimize the structure of the fuzzy IF-THEN rules and their appropriate antecedent and consequent parameters. Figure 3 schematically indicates the working process of the developed combined FIS and GA model.

In order to optimize the structure of fuzzy IF-THEN rules and their antecedent and consequent parameters, two groups of data sets are used. The first one is the training data set directly used in the tuning process to achieve the minimum value of the objective function (root-mean-square error, *RMSE*). The second data set is validation data typically used to prevent the overtraining of the estimator models. Note that in this model, the generation that has the minimum error of validation and training data sets and the fuzzy IF-THEN rules associated with them, is selected as the final answer. However, after the generation, selected only training data error is on the decrease and the error of validation data is on the increase (over-training or over-fitting problems). The objective function of all predictor models is the *RMSE* error as indicated in Figure 4.

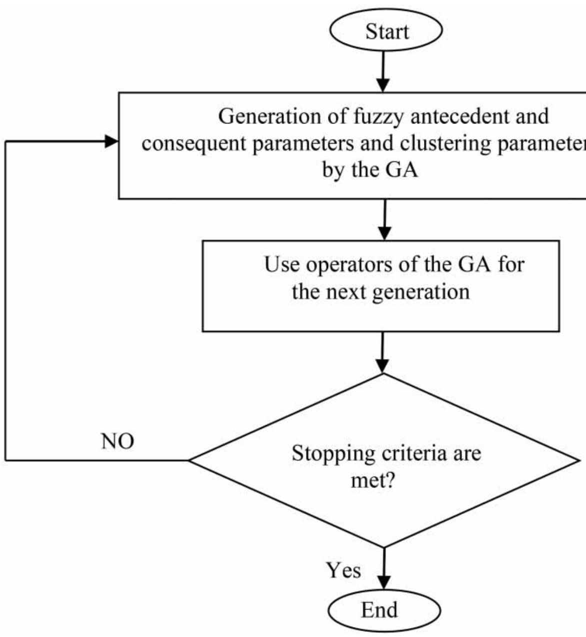

Figure 3 | Schematic diagram of the combined FIS and GA model.

In the figure, ra_i is the clustering radius for the ith input variable and the output variable ($n = 1, \ldots, D$), $ra_{\min i}$ is the minimum clustering radius for the ith variable. *MaxNum-rule*, is the maximum number of fuzzy IF-THEN rules determined based on the prediction errors for the training and validation data expressed as $RMSE_{trn}$ and $RMSE_{validation}$, respectively.

Note that in this paper as mentioned before, fuzzy IF-THEN rules associated with clustering parameters are extracted based on having the lowest similarities among them. The number of rules and the linguistic variables for each input variable are equal to the number of clusters related to clustering parameters. In order to meet minimum similarities in the fuzzy IF-THEN rules, only linguistic variables in the same levels are chosen (Chiu 1994). For example, 'A_1' as the first linguistic parameter for input variable A makes a rule with the first linguistic of variable B as 'B_1'. It is the same for other rules and final obtained rules can be expressed as follows:

Rule 1: If x is A_1 and y is B_1
Rule 2: If x is A_2 and y is B_2
....
Rule N: If x is A_{n} and y is B_{n}.

The construction of initial FIS is inspired by MATLAB GENFIS 2 commands that cause lower similarity in the rule to decrease the time of execution.

In addition to the development of combined FIS and GA models, subtractive clustering parameters and fuzzy IF-THEN rules, antecedent and consequent parameters can be optimized separately for clarifying how important is their optimization in the prediction of wave parameters. To achieve this, in this paper the GA also is used to optimize the subtractive clustering parameters, including radii of clustering and quash factor, without optimization of fuzzy antecedent and consequent parameters. In this optimizing process, two groups of data sets are used as the training and

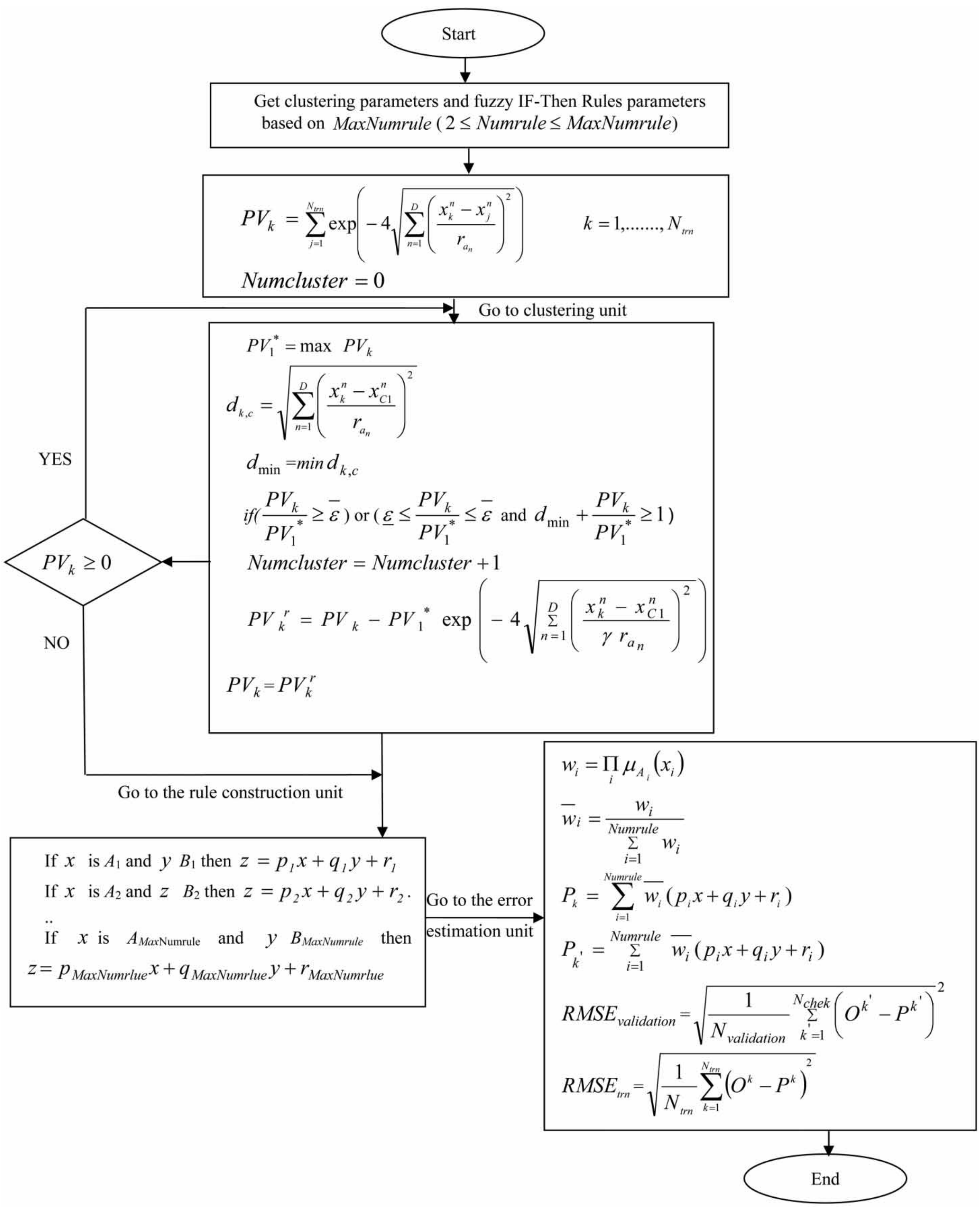

Figure 4 | Objective function of the GA model.

validation data sets similar to other data-driven algorithms. Training data sets are used directly to optimize subtractive clustering parameters whereas validation data sets are used to evaluate the model generalization capability and escape from the 'curse of dimensionality' deficiency in the FIS. In the tuning process, the generation that has the minimum error of training and validation data sets simultaneously is chosen as the final solution. The same scenario can be employed to optimize fuzzy antecedent and consequent parameters for a determined fuzzy IF-THEN rule. In this viewpoint, also two groups of data sets including the training and validation data sets should be used to have a model with suitable generalization capability. In all these scenarios, the *RMSE* error of the training data is the objective function, whereas the *RMSE* error of the validation data is calculated to control the known over-fitting problem.

THE COMBINED FIS AND GA MODEL AT LAKE MICHIGAN FOR WAVE PREDICTION

Study area and data selection

Following on from the introduction of the developed combined FIS and GA model, in this section the model is employed for prediction of wave parameters at Lake Michigan, one of the five Great Lakes of North America. The wave data used to develop the wave predictor models have been collected by the National Data Buoy Center (NDBC) at station 45007 of the lake (Figure 5). This station is located at $42^\circ 40'30''N$, $87^\circ 01'30''W$ at a depth of 159.1 m with a maximum measured peak spectral period 7.3 s; therefore, a deep water condition prevails. Wind and wave data were gathered by a 3-m discus buoy. The wind speed was measured at a height of 5 m above the mean sea level with ± 1 m/s accuracy while its direction was measured with $\pm 10^\circ$ accuracy. The buoy measured and transmitted barometric pressure; wind direction, speed, and guest; air and sea temperature; and wave energy spectra. Significant wave height, dominant wave period and average wave period were derived from wave energy spectra. To measure the wave characteristics, the accelerometers and inclinometers of the buoy measure the heave acceleration and the vertical displacement of the buoy hull during the wave acquisition time. The accuracy of

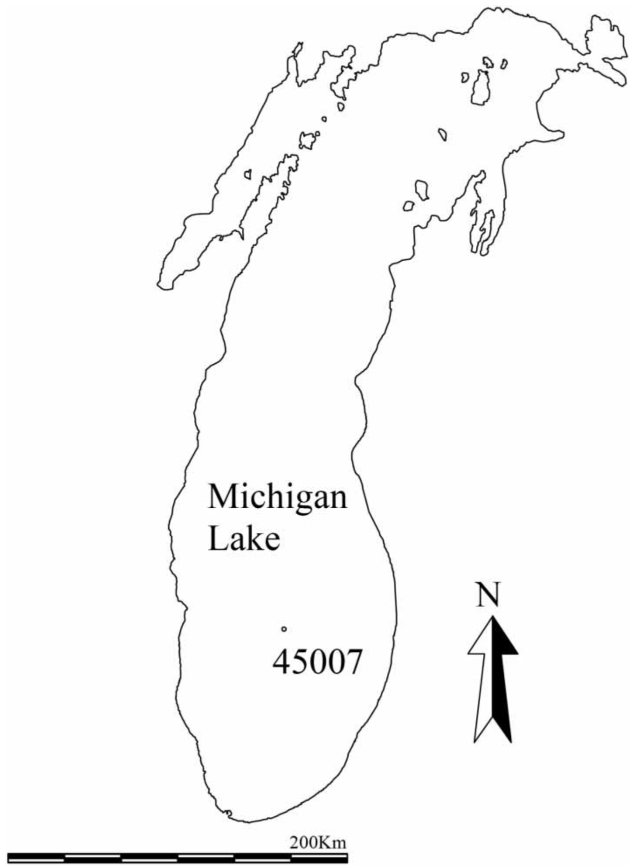

Figure 5 | Map of Lake Michigan and location of buoy 45007.

measurement for significant wave height and peak spectral period is ± 0.2 m and ± 1 s, respectively. The data set belongs to the periods of March to January 2001 and from January to December 2013, with 1-hr intervals. Based on simplified methods, the wind-driven wave parameters mostly are related to the wind speed, the fetch length, and the wind duration. In order to calculate the wind speed and the wind duration, so-called 'constant wind' definition is used (Bishop 1983). Based on the definition, a wind is called a constant wind when meeting the following criteria:

$$\left|U_i - \bar{U}\right| \leq 2.5\, m/s \tag{15}$$

$$\left|D_{wi} - \overline{D_w}\right| \leq 15^\circ \tag{16}$$

where U_i and D_{wi} are wind speed and wind direction at the ith hour, respectively, while $\bar{U}$ and $\bar{D}_w$, respectively, are the average of wind speed and wind direction for consecutive preceding i hours.

To estimate the fetch length at the study area, the CEM (US Army 2003) method criterion was used, by which a fetch length for a certain direction was estimated by considering 30 radials from the point of interest (at 1-degree intervals) and extending them until they intersected the coastline. The fetch length is the arithmetic average of the obtained lengths.

Following the extraction of constant winds, 1,200 hourly data were selected, of which, 1,080 data points (the data set of year 2001) were used as the training data set and the remaining data (data set of year 2013) were selected as the testing data set. Selection of another year for the testing is to provide the fair evaluation of the developed model in a different climate. Out of 1,080 data points, 800 data points were chosen as the training data points and the remaining 220 data points used as the validation data to avoid over-training of the model. Statistical characteristics such as the minimum, maximum, average and range of all data points are reported in Table 1. In the selected intervals of the data set, the maximum and minimum of the recorded wind speed are 16.52 m/s and 5.75 m/s, respectively. This proves that the data set covers a wide range of wind climate at the lake. These data have been selected among a total of 4,554 hourly data, in which data points with wave height less than 0.5 m (the common calm condition in maritime design) have been eliminated from the data set. In addition, existing gaps in the data set either have been excluded from the data set or interpolated. Also, wind speeds have been converted to 10 m above sea level wind speeds. In order to evaluate the performance of the combined FIS and GA model in this section, the GA model is applied in three states. The first state of the GA model application goes back to its application for optimizing the subtractive clustering parameters, i.e., radii of inputs and output variables and the quash factor, leading to extraction of fuzzy IF-THEN rules. In the second state, the GA is employed only to tune the antecedent and consequent parameters of the resultant fuzzy IF-THEN rules from the first step. In the third state, the GA model is used for simultaneous optimization of the subtractive clustering parameters and the antecedent and consequent parameters associated with the selected fuzzy IF-THEN rules from clustering parameters. Note that in this form of the GA application for optimizing fuzzy IF-THEN rules, the number of the antecedent and consequent parameters are related to subtractive clustering parameters. Therefore, the number of the decision variables changes during the execution of the GA model.

Table 1 | Statistical properties of data sets

Parameter	Min.	Max.	Average	Range
Wind speed (m/s)	5.75	16.52	8.37	14.77
Fetch length (CEM) (km)	76	329	129	253
Wind duration (hr)	3	37	6.37	34
Significant wave height (m)	0.51	4.75	1.22	4.24
Peak spectral period (s)	2.98	7.3	4.21	4.32

GA application to extract the fuzzy IF-THEN rules

According to the above discussion in this subsection, the GA is only employed to extract fuzzy IF-THEN rules to predict wave parameters whereas input variables for the FISs are wind duration, fetch length, and wind speed. In this model, the GA is combined with a FIS, such that their structures are being extracted by optimizing the subtractive clustering parameters including radii of input and output variables and the quash factor. In the training process previously selected, training data points and validation data points are used. In the FIS model, each cluster center represents a fuzzy IF-THEN rule and the cluster center i is the mean value parameter (c_i) of the Gaussian membership functions as shown in the following relation:

$$\mu_{A_i}(x) = \exp\left[-\left(\frac{x - c_i}{a_i}\right)^2\right] \tag{17}$$

while the deviation parameter of the Gaussian membership function (a_i) is calculated as follows:

$$a_i = r_{ai}\left(\frac{\max(x_i) - \min(x_i)}{\sqrt{8}}\right) \tag{18}$$

This equation has been selected according to GENFIS 2 commands and is proven by previous work (Zanganeh *et al.* 2016).

Fuzzy IF-THEN rules, as reported in the section 'The combined model of FISs and GA', are selected based on having the lowest similarities among themselves. An example of the fuzzy IF-THEN rules for the prediction of significant wave height is expressed as:

$$\text{If } U_{10} \text{ is } A_1 \,\&\, F \text{ is } B_1 \,\&\, t \text{ is } C_1 \text{ Then } H_S = p_1 t + q_1 U_{10} + r_1 F + s_1 \tag{19}$$

$$\text{If } U_{10} \text{ is } A_2 \,\&\, F \text{ is } B_2 \,\&\, t \text{ is } C_2 \text{ Then } H_s = p_2 t + q_2 U_{10} + r_2 F + s_2 \tag{20}$$

……

$$\text{If } U_{10} \text{ is } A_{\mathrm{n}} \,\&\, F \text{ is } B_{\mathrm{n}} \,\&\, t \text{ is } C_{\mathrm{n}} \text{ Then } H_s = p_n t + q_n U_{10} + r_n F + s_n \tag{21}$$

where U_{10} is the wind speed, F is the fetch length, and t is the wind duration, $A_1, \ldots, A_{\mathrm{n}} B_1, \ldots, B_{\mathrm{n}}$, and $C_1, \ldots, C_{\mathrm{n}}$ are the Gaussian fuzzy values defined for the wind speed, fetch length, and wind duration, respectively, H_s is the significant wave height and p_1, q_1, r_1, s_1, p_2, q_2, r_n, p_n, q_n, r_n, and s_n are the linear consequent parameters in a given fuzzy IF-THEN rule.

Following the above descriptions, in this part the fuzzy IF-THEN rules are extracted to predict wave parameters. These rules are extracted based on having minimum errors of training and validation data with respect to optimum subtractive clustering parameters.

Figures 6 and 7 indicate the optimization process of clustering parameters for the wave height and peak spectral predictor models, respectively. In these figures, the minimum, maximum, and average values of the objective functions are demonstrated. Note that in the GA employed to optimize parameters, the number of populations is 20, the crossover fraction is 0.85, and the number of elitism chromosomes is 5 while the remaining chromosomes are selected for the mutation operator. Also, the number of the fuzzy IF-THEN rules associated with the clustering parameters is equal to 4 for both predictor models.

The optimum values of the clustering parameters are as follows:

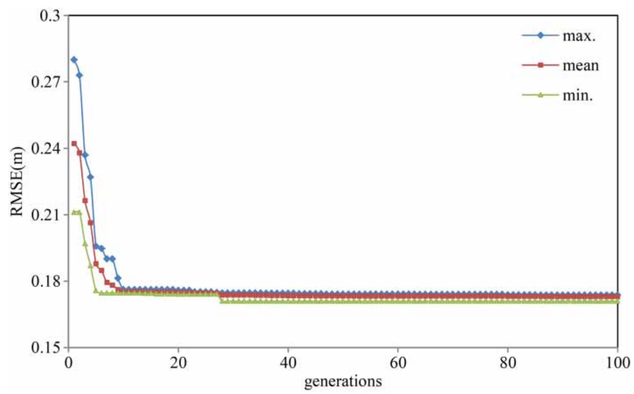

Figure 6 | Optimization process of the clustering parameters in the wave height predictor model.

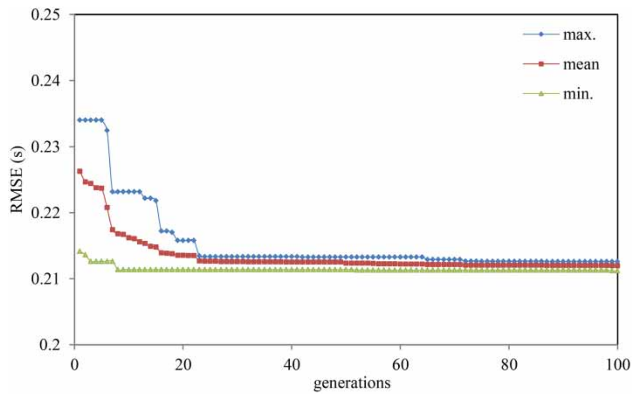

Figure 7 | Optimization process of the clustering parameters in the peak spectral predictor model.

- For the significant wave height predictor model

$$\text{clustering parameters} = [r_{tr}, r_F, r_{U_{10}}, r_{H_s}, \gamma_{Hs}] = [1.288, 0.502, 0.171, 0.290, 0.639]$$

- For the peak spectral predictor model

$$\text{clustering parameters} = [r_{tr}, r_F, r_{U_{10}}, r_{T_p}, \gamma_{Tp}] = [0.805, 0.503, 0.157, 0.796, 0.551]$$

These obtained clustering parameters are associated with the best execution out of ten runs of the predictor models. As is evident from Figure 6, the training error for the significant wave height predictor model has varies from 0.273 m to 0.1705 m, while in Figure 7, this error changes from 0.2331 s to 0.2114 s for the peak spectral period predictor model. In addition, the error of validation data for the appropriate clustering parameters in the wave height predictor model is 0.3101 m while its appropriate generation is 87. In the peak spectral period predictor model, the error of the validation data is equal to 0.3815 s and its appropriate generation is 54. These two steps are the steps after which there is no decrease in the training and validation errors. Thus, improving the obtained FISs should be investigated in order to know whether more optimization is possible or not in the predictor models. However, in this step, there is no optimization on fuzzy antecedent and consequent parameters in both predictor models and so generally it can be concluded that the obtained FISs are local answers. In that regard, optimization of the fuzzy antecedent and consequent parameters considering some fixed fuzzy IF-THEN rules for both predictor models should be undertaken.

Referring to the optimized clustering parameters reveals that the wind speed has the lowest value of the clustering radius in both the wave height predictor model ($r_{U_{10}} = 0.171$) and peak spectral period predictor model ($r_{U_{10}} = 0.157$), whereas the wind duration has the highest value of the clustering radius ($r_{tr} = 1.288$ for wave height and $r_{tr} = 0.805$ for peak spectral period).

The GA application to tune the extracted fuzzy IF-THEN rules parameters

As mentioned above, optimizing the obtained fuzzy antecedent and consequent parameters in both predictor models is of great importance to improve the training and validation errors. Therefore, in this subsection following the extraction of fuzzy IF-THEN rules in the previous section, the appropriate antecedent and consequent parameters extracted by the fuzzy IF-THEN rules are optimized. The number of fuzzy IF-THEN rules associated with the obtained clustering parameters for both wave height and peak spectral period predictor models is 4. As a result, the number of appropriate fuzzy antecedent and consequent parameters is 40 which are considered as the decision variables of the GA. Of them, 16 parameters are linear consequent parameters and 24 parameters are nonlinear antecedent parameters. In the generation process, the population size of the GA for both predictor models is 400, the crossover fraction is 0.7, the number of elitism chromosomes is 20, and the remaining children are considered for the mutation process.

In order to highlight the GA performance to optimize fuzzy antecedent and consequent parameters, first the SD method is used to tune the parameters for prediction of wave parameters. The tuning processes by the SD method are indicated in Figures 8 and 9 to predict significant wave height and peak spectral period, respectively. As shown in the figures, the training process by the SD method is not suitable and entrapping in the local optimum for both wave parameters predictor models is apparent due to no change in answers. Note the figures clarify that the significant wave height predictor has stuck to the point where its associated training error is equal to 0.1705 m,

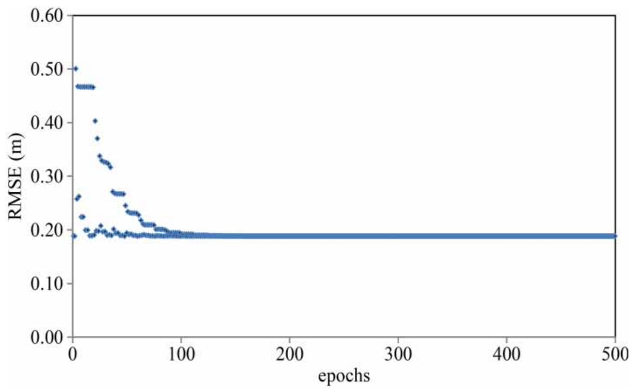

Figure 8 | Training error in the ANFIS model versus epoch number for the prediction of significant wave height prediction.

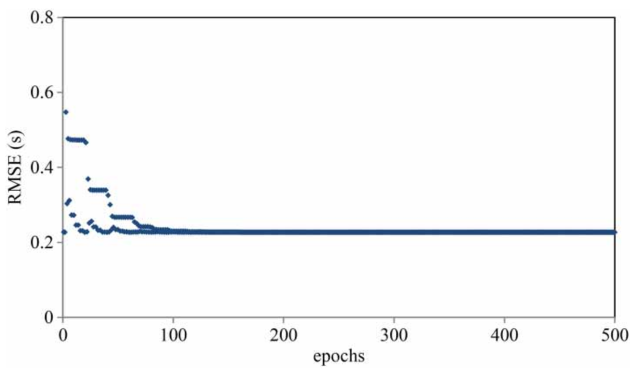

Figure 9 | Training error in the ANFIS model versus epoch number for the prediction of peak spectral period.

while the training error of peak spectral period predictor model is 0.2114 m.

Due to the SD method deficiency to optimize fuzzy antecedent and consequent parameters in wave predictor models, in this stage the GA is used for optimization of fuzzy IF-THEN rules antecedent and consequent parameters. Figures 10 and 11 show the optimization process by the GA in the models. In these figures, results of minimum, average, and maximum of the *RMSE* for ten executions are reported while Figures 12 and 13 show initial and optimized membership functions for both wave height and peak spectral period predictor models' input variables. As apparently shown in these figures, membership functions have been changed significantly by the GA. In addition, the *RMSE* errors of validation and training data are presented in Table 2. As reported in the table, the GA model employed here has decreased the *RMSE* error successfully for both the significant wave height and the peak spectral predictor models. The GA not only has decreased the *RMSE* error from 0.1705 m to 0.1604 m for the wave height predictor model, but it has also improved the *RMSE* error for the peak spectral period predictor model from 0.2114 s to 0.2018 s. Although the GA has improved the results, it can be concluded the second process of optimization to extract fuzzy antecedent and consequent parameters is less effective in comparison with the optimization of subtractive clustering parameters. In other words, tuning clustering parameters is more important than fuzzy antecedent and consequent parameters in the developed predictor models. The validation error has also reached 0.2920 m and 0.3421 s for significant wave height and peak spectral period, respectively.

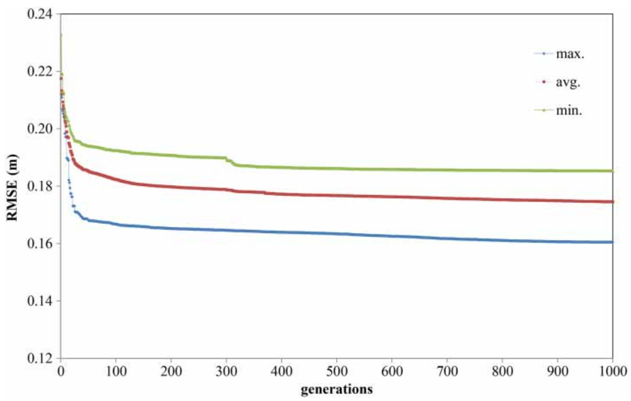

Figure 10 | Variation of *RMSE* error in the combined FIS and GA model versus number of generations for wave height prediction.

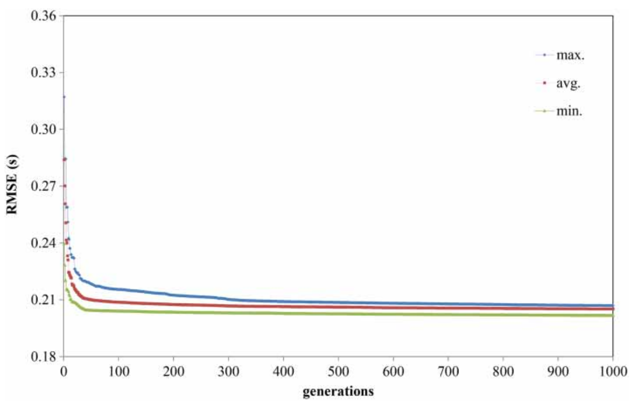

Figure 11 | Variation of *RMSE* error in combined FIS and GA model versus number of generations for peak spectral period prediction.

Application of the combined FIS and GA model

Evaluating the GA model efficiency for separate optimization of the subtractive clustering parameters and the antecedent and consequent parameters of fuzzy IF-THEN rules reveals that simultaneous optimization of the parameters by the GA is a beneficial issue to unify the optimization of subtractive clustering parameters and fuzzy antecedent and consequent parameters. In that regard, herein, the combined FIS and GA model is implemented with the same training and validation data sets used in the previous predictor models. In this model, the number of the GA decision variables changes due to the variation of subtractive clustering parameters. In this step, two models are used to predict wave parameters as well. The first one is significant wave height predictor model and the second is peak spectral period predictor model. Note that in these models the maximum number of fuzzy IF-THEN rules is restricted to the *Max-Numrule* related to the subtractive clustering parameters. The following expression as an example outlines a set of fuzzy IF-THEN rules considered to predict significant wave height:

$$\text{If } U_{10} \text{ is } A_1 \ \& \ F \text{ is } B_1 \ \& \ t \text{ is } C_1 \text{ Then } H_s = p_1 t + q_1 U_{10} + r_1 F + s_1 \tag{22}$$

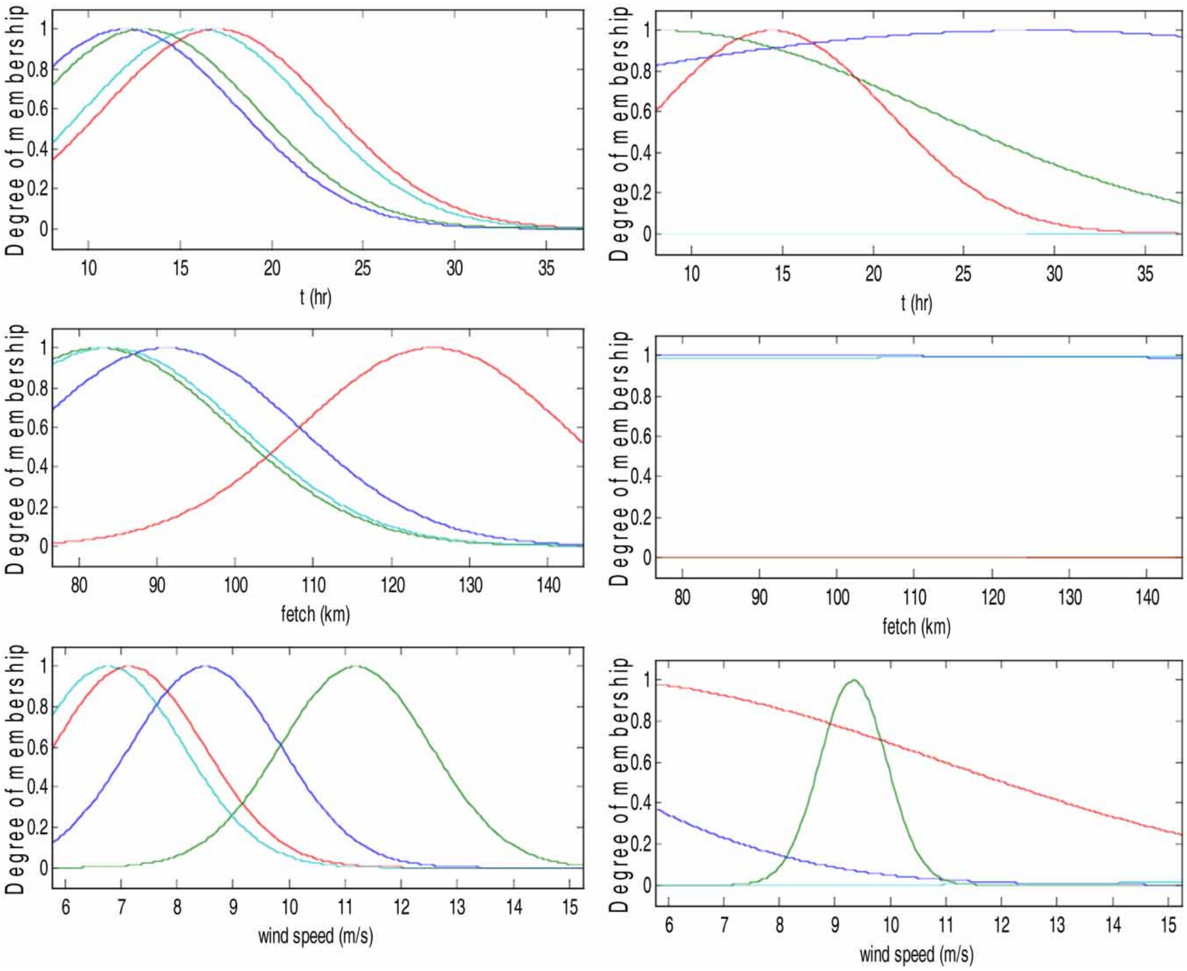

Figure 12 | Initial and optimized fuzzy membership functions by the GA appropriate by each input variable for the wave height predictor model.

$$\text{If } U_{10} \text{ is } A_2 \ \& \ F \text{ is } B_2 \ \& \ t \text{ is } C_2 \text{ Then } H_s = p_2 t + q_2 U_{10} + r_2 F + s_2 \quad (23)$$

……

$$\text{If } U_{10} \text{ is } A_{\text{MaxNumrule}} \ \& \ F \text{ is } B_{\text{MaxNumrule}} \ \& \ t \text{ is } C_{\text{MaxNumrule}} \text{ Then } H_s = p_{\text{MaxNumrule}} t + q_{\text{MaxNumrule}} U_{10} + r_{\text{MaxNumrule}} F + s_{\text{MaxNumrule}} \quad (24)$$

As shown above, the rules associated with the combined FIS and GA model are extracted in the same way as previous models. However, the GA decision variables are the subtractive clustering parameters and antecedent and consequent parameters of fuzzy IF-THEN rules. That makes the optimization process more difficult because of the domain-irrelevant behavior of the GA for the production of the initial population.

Following the development of the models to predict wave parameters, the minimization process of the training error for both significant wave height and peak spectral predictor models are demonstrated in Figures 14 and 15. In these figures, the results of minimum, average, and maximum of the *RMSE* for ten executions are shown. In addition, Figures 16 and 17 show the final obtained membership functions of input variables for both significant wave height and peak spectral period predictor models. As shown in the figures, final membership functions have a value less than 1. This at first may show a contradiction with fuzzy logic arithmetic, as out of the range that we work, fuzzy membership functions have taken the value of 1 as maximum.

The *RMSE* errors of validation and training data sets are also presented in Table 3. As reported in the table, the GA

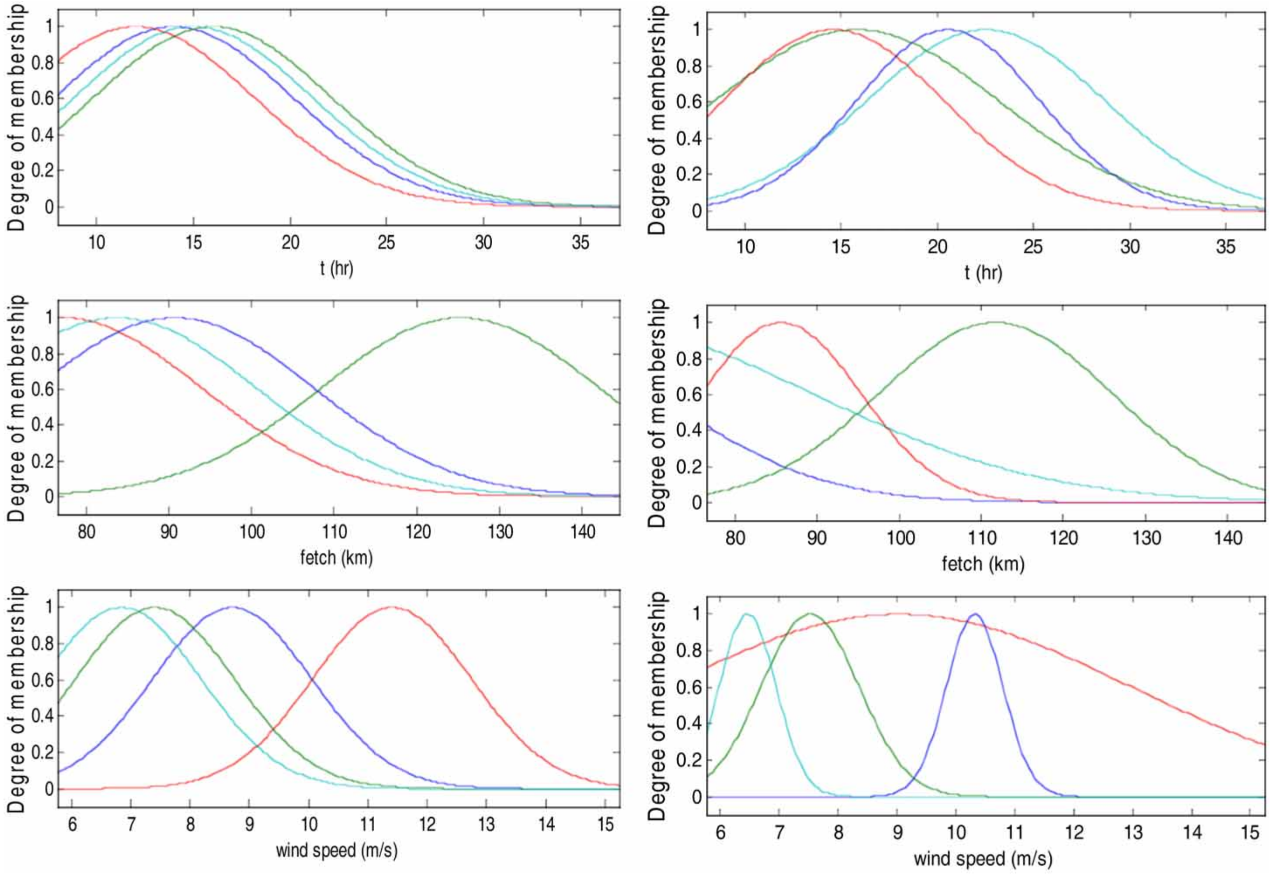

Figure 13 | Initial and optimized fuzzy membership functions by the GA appropriate by each input variable for the wave peak spectral predictor model.

Table 2 | *RMSE* errors for the wave predictor models in the FIS model in which fuzzy antecedent and consequent parameters are optimized

	Method	Validation error (m)	Training error (m)
Significant wave height predictor model	ANFIS	0.3101	0.1705
	FIS and GA	0.2920	0.1604
Peak spectral period predictor model	ANFIS	0.3815	0.2114
	FIS and GA	0.3421	0.2018

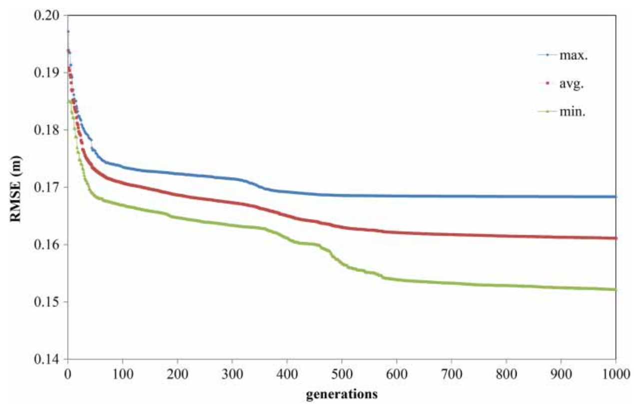

Figure 14 | Variation of *RMSE* error in the combined FIS and GA model versus number of generations for significant wave height prediction.

model employed here has decreased the *RMSE* error successfully for both significant wave height and peak spectral predictor models. The GA not only has decreased the *RMSE* error to 0.1533 m for the wave height predictor model but it has also improved the *RMSE* error for the peak spectral period predictor model to 0.2045 s in the best run. In the generation process, the population size of the GA for both predictor models is 400, the crossover fraction is 0.7, the number of elitism chromosomes is 20, and the remaining children are taken for the mutation process. The obtained validation errors for both predictor models are, respectively, 0.2911 m and 0.3461 s. The obtained results in this part show the combined GA and FIS models' efficiency

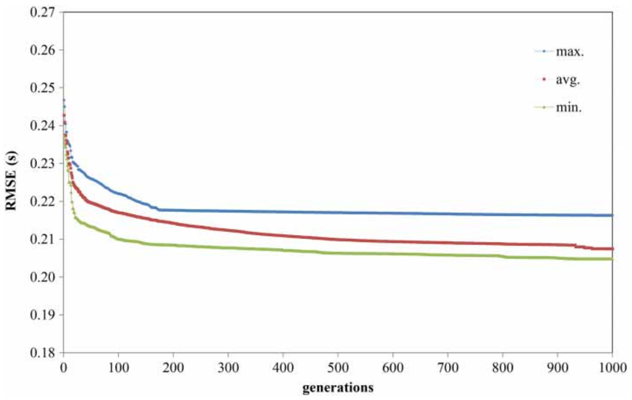

Figure 15 | Variation of *RMSE* error in the combined FIS and GA model versus number of generations for peak spectral period prediction.

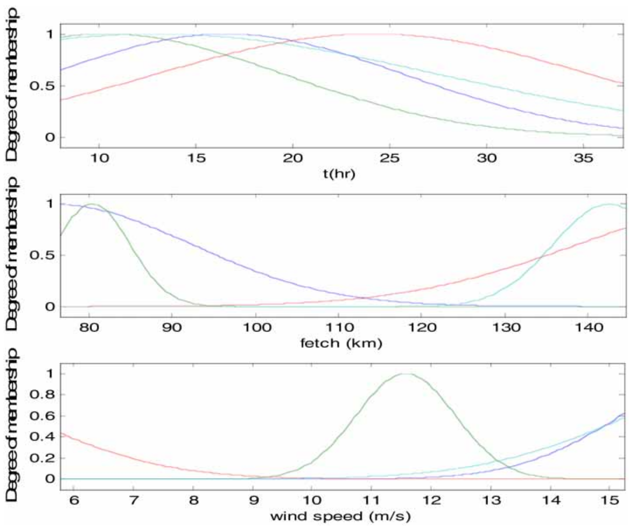

Figure 16 | Final fuzzy membership functions appropriate by any input variables for wave peak spectral predictor model.

to predict wave parameters, although final evaluation of the developed models is related to their evaluation versus the testing data never used during the training process.

THE CEM METHOD

Evaluation of presented approaches against known methods with identical input variables is needed to verify the models and also to create a sound conclusion. To achieve this, in this paper the CEM formulas are employed for prediction of wave parameters and the following paragraphs outline this empirical method.

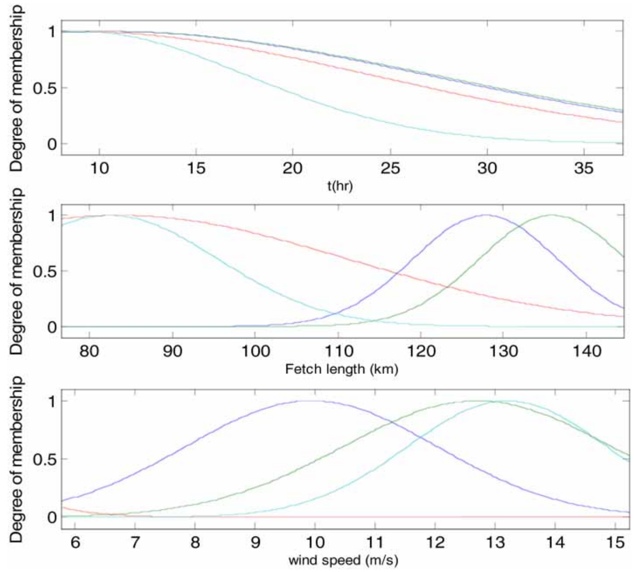

Figure 17 | Final fuzzy membership functions appropriate by any input variables for wave peak spectral predictor model.

Table 3 | *RMSE* errors for the wave predictor models in which both clustering and fuzzy antecedent and consequent parameters are optimized simultaneously

	Method	Validation error (m)	Training error (m)
Significant wave height predictor model	Combined FIS and GA	0.2911	0.1533
Peak spectral period predictor model	Combined FIS and GA	0.3461	0.2045

On the basis of the CEM manual, the generation of wind waves occurs in three states categorized as fetch-limited, duration-limited, and fully developed conditions. In order to determine the generation condition, the estimation of the real wind duration by the following equation is essential:

$$t_{\min} = 77.23 \frac{X^{0.67}}{U_{10}^{0.34} g^{0.33}} \tag{25}$$

where $t_{\min}$ is the required time for the fetch-limited condition in the wave generation process which is in seconds. X is the fetch length in meters, g is the gravitational acceleration which is equal to 9.81 (m/s^2) and U_{10} is the wind speed 10 m above the sea water level (m/s).

If the real wind duration exceeds the wind duration estimated by Equation (25) the fetch-limited condition is prevailing, in which wave parameters are estimated as follows:

$$\frac{gH_s}{u_*^2} = 4.13 \times 10^{-2} \left(\frac{gX}{u_*}\right)^{1/2} \quad (26)$$

$$\frac{gT_p}{u_*} = 0.651 \times \left(\frac{gX}{u_*^2}\right)^{1/3} \quad (27)$$

where H_s is significant wave height, T_p is peak spectral period, u_* is shear velocity calculated as follows:

$$u_* = U_{10} C_D^{0.5} \quad (28)$$

where C_D is drag coefficient estimated as follows:

$$C_D = 0.001(1.1 + 0.035U_{10}) \quad (29)$$

If the wind duration is less than the estimated one by Equation (25), the duration-limited condition is dominant. In this condition, wave parameters are predicted by modifying the fetch length in accordance with the real wind duration by Equation (25). In the empirical methods, another condition, the so-called fully developed method, is considered to estimate wave parameters. At this condition, wave parameters are estimated by the following relationships:

$$\frac{gH_s}{u_*^2} = 2.115 \times 10^2 \quad (30)$$

$$\frac{gT_p}{u_*} = 2.398 \times 10^2 \quad (31)$$

EVALUATION OF THE COMBINED FIS AND GA MODELS

Now that the combined FIS and GA models for prediction of wave parameters have been developed, the models should be verified against some testing data that have not been used in the training processes. This evaluation is to ensure the generalization capability of the developed models. For this reason, the evaluation of the models is accomplished by two statistical indexes. The first index is *bias* showing the mean error (*ME*) caused by overestimating and underestimating the observed value. The second index is the scatter index (*SI*) showing the scattering of the observed and predicted values around the line $y = x$. Indeed, this index is a normalized form of the *RMSE* by dividing mean values of the observed data points. Since the *RMSE* is the objective function of the predictor models, this index is more important than the other one. These two indexes are calculated by the following equations:

$$bias = \frac{1}{N_{test}} \sum_{k=1}^{N_{test}} (O^k - P^k) \quad (32)$$

$$SI = \frac{\text{RMSE}}{\text{average observed value}} \times 100 \quad (33)$$

where O^k is the observed value, P^k is the predicted value, and N_{test} is the number of testing data.

Following the estimation of parameters, the results of wave prediction models are reported in Table 4. From this table, it can be inferred that in the studied case, the combined FIS and GA models are even more accurate than the ANFIS models where their structures have been optimized by the GA. All of the models have reasonable *bias*, indicating the accuracy of models for prediction of the phenomenon.

Figures 18 and 19 also show the estimated values of wave parameters against the observed ones. As is apparent from the figures, the correlation ratio for the combined FIS and GA significant wave height predictor

Table 4 | Final results of the developed combined FIS and GA model versus the ANFIS model

	Combined FIS and GA		CEM		ANFIS	
Wave parameter	**SI (%)**	**Bias**	**SI (%)**	**Bias**	**SI (%)**	**Bias**
$H_s(m)$	20.03	0.0034	42.85	0.241	22.1	0.045
$T_p(s)$	8.12	0.0131	58.2	0.347	9.97	0.0199

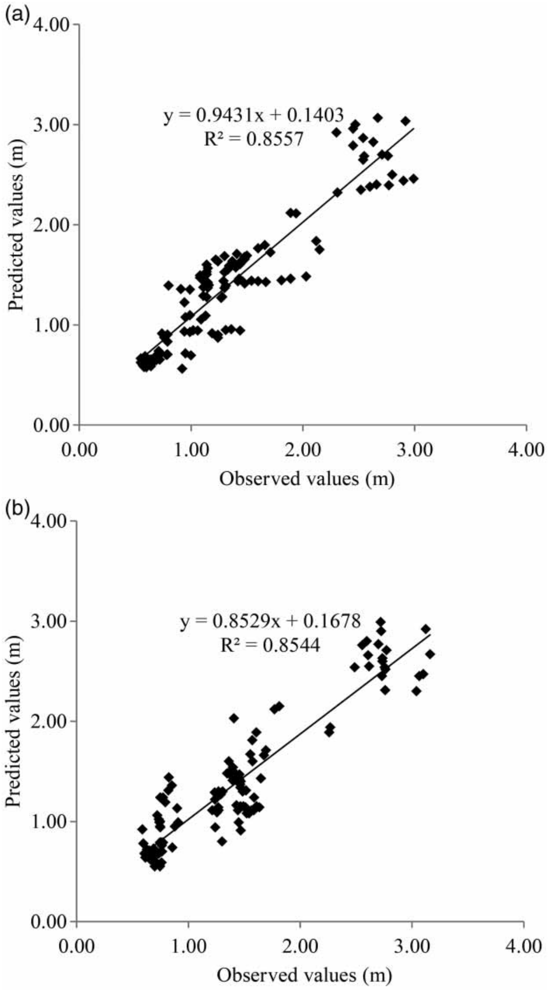

Figure 18 | Observed values of significant wave height versus the predicted ones: (a) by the combined FIS and GA model and (b) by the ANFIS.

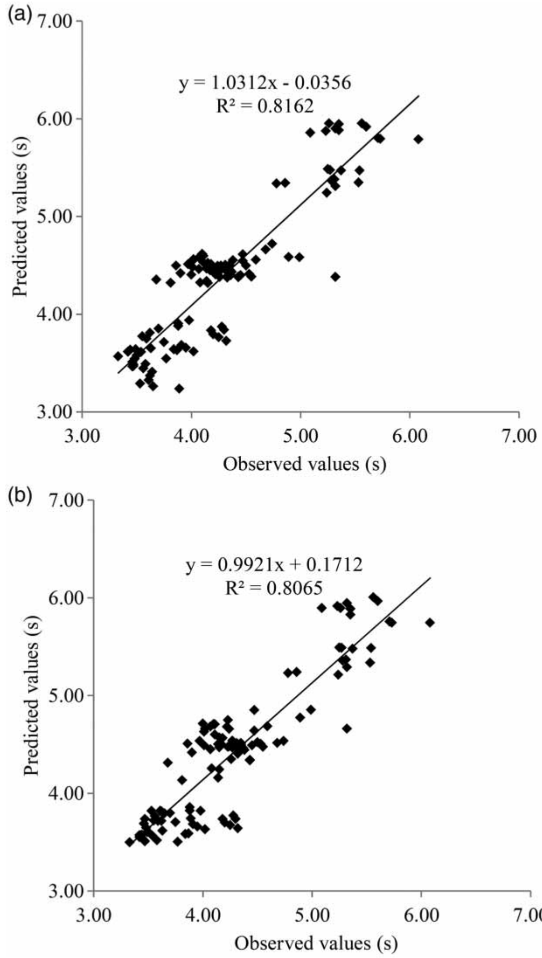

Figure 19 | Observed values of peak spectral period versus the predicted ones: (a) by the combined FIS and GA model and (b) by the ANFIS.

model ($R^2 = 0.8557$) is closer to 1 than that of the ANFIS model ($R^2 = 0.8544$). Furthermore, the correlation ratio in the combined GA and FIS model to predict peak spectral period ($R^2 = 0.8162$) is closer to 1 than the ANFIS model ($R^2 = 0.8065$). Therefore, based on the above results, it can be concluded that the developed combined FIS and GA model can predict the wave parameters with an acceptable accuracy. Also, Figure 20 shows predicted wave parameters by the CEM versus observed wave parameters. The estimated *SI* of the CEM method is 42.85 for the prediction of significant wave height while it is 58.2 for the prediction of peak spectral period. Note the *bias* parameters prove the superiority of the combined GA and FIS model to the ANFIS and CEM methods.

SUMMARY AND CONCLUSIONS

Recently, soft computing tools such as ANFISs and ANNs have been used in the prediction of wave parameters, i.e., significant wave height and peak spectral period. The ANFIS is a gradient-based method in which optimizing of the antecedent and consequent parameters of fuzzy IF-THEN rules can be accomplished by a gradient, thereby entrapping in a local optimum is possible. Therefore, in this study the GA was used to extract fuzzy IF-THEN rules and to optimize the antecedent and consequent parameters of fuzzy IF-THEN rules simultaneously in a model called the combined FIS and GA model. Finally, the combined FIS and GA models were employed

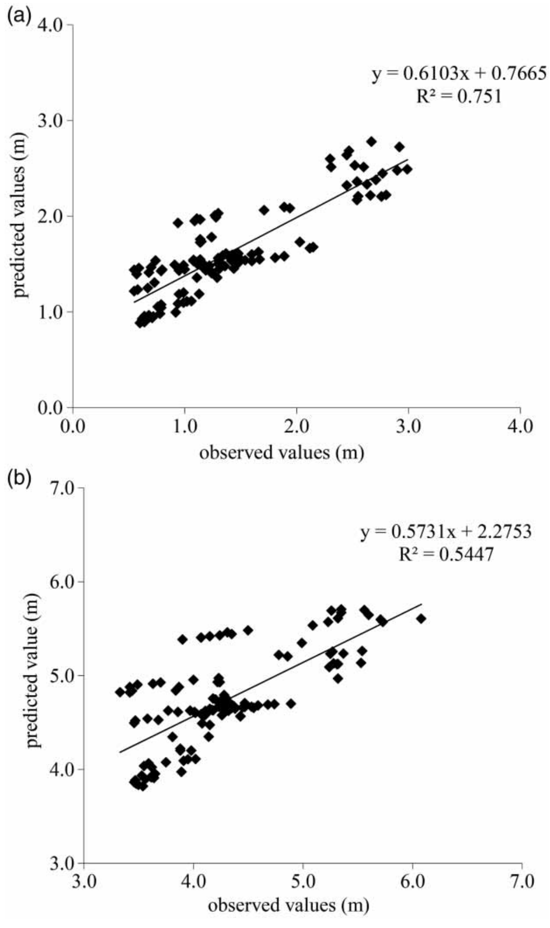

Figure 20 | Predicted values by the CEM method versus observed values: (a) for significant wave heights and (b) for the peak spectral period.

for prediction of wave parameters. Results not only indicated the combined FIS and GA models' accuracy for prediction of wave parameters but also showed that the GA could optimize fuzzy IF-THEN rules and fuzzy antecedent and consequent parameters. In addition, it was inferred that in the wave predictor models optimizing of clustering parameters had a more important effect than fuzzy antecedent and consequent parameters' optimization. For future works, we can consider two viewpoints. The first one can focus on obtaining a FIS-based model capturing three known conditions, fetch-limited, duration-limited, and fully developed sea conditions in the study area. The second viewpoint goes back to define a new membership function to minimize the prediction errors more in order to produce a more robust algorithm.

ACKNOWLEDGEMENTS

This study was partially supported by the Deputy of Research at Golestan University (GU) and the author would like to sincerely thank them for their support during the study. Also, the author wishes to thank Dr Mahmood Hajiani, faculty member of Birjand University, for his constructive comments on the manuscript.

REFERENCES

Agrawal, J. & Deo, M. 2002 On-line wave prediction. *Marine Structures* **15** (1), 57–74.

Akpınar, A., Özger, M. & Kömürcü, M. I. 2014 Prediction of wave parameters by using fuzzy inference system and the parametric models along the south coasts of the Black Sea. *Journal of Marine Science and Technology* **19** (1), 1–14.

Altunkaynak, A. 2013 Prediction of significant wave height using geno-multilayer perceptron. *Ocean Engineering* **58**, 144–153.

Beyer, H.-G. 2000 Evolutionary algorithms in noisy environments: theoretical issues and guidelines for practice. *Computer Methods in Applied Mechanics and Engineering* **186** (2), 239–267.

Bishop, C. T. 1983 Comparison of manual wave prediction models. *Journal of Waterway, Port, Coastal, and Ocean Engineering* **109** (1), 1–17.

Chen, C.-K. & Dao, T.-S. 2007 Genetic fuzzy control for path-tracking of an autonomous robotic bicycle. *Journal of System Design and Dynamics* **1** (3), 536–547.

Chipperfield, A. J., Fleming, P. & Pohlheim, H. 1994 *Genetic Algorithm Toolbox: For Use with MATLAB; User's Guide (Version 1.2)*. Department of Automatic Control and Systems Engineering, University of Sheffield, UK.

Chiu, S. L. 1994 Fuzzy model identification based on cluster estimation. *Journal of Intelligent and Fuzzy Systems* **2** (3), 267–278.

Engineers, U. A. C. O. 1984 *Shore Protection Manual*. Army Engineer Waterways Experiment Station, Vicksburg, MS, USA. 2v, pp. 37–53.

Etemad-Shahidi, A., Kazeminezhad, M. & Mousavi, S. 2009 On the prediction of wave parameters using simplified methods. *Journal of Coastal Research* **SI56**, 505–509.

Hidalgo, D., Melin, P. & Castillo, O. 2012 An optimization method for designing type-2 fuzzy inference systems based on the footprint of uncertainty using genetic algorithms. *Expert Systems with Applications* **39** (4), 4590–4598.

Holland, J. H. 1975 *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence.* University of Michigan Press, Ann Arbor, MI, USA.

Homaifar, A. & McCormick, E. 1995 Simultaneous design of membership functions and rule sets for fuzzy controllers using genetic algorithms. *Fuzzy Systems, IEEE Transactions on* **3** (2), 129–139.

Jang, J.-S. 1993 ANFIS: adaptive-network-based fuzzy inference system. *Systems, Man and Cybernetics, IEEE Transactions on* **23** (3), 665–685.

Johnson, P. A. & Ayyub, B. M. 1996 Modeling uncertainty in prediction of pier scour. *Journal of Hydraulic Engineering* **122** (2), 66–72.

Kazeminezhad, M., Etemad-Shahidi, A. & Mousavi, S. 2005 Application of fuzzy inference system in the prediction of wave parameters. *Ocean Engineering* **32** (14), 1709–1725.

Kindler, J. 1992 Rationalizing water requirements with aid of fuzzy allocation model. *Journal of Water Resources Planning and Management* **118** (3), 308–323.

Mahjoobi, J., Etemad-Shahidi, A. & Kazeminezhad, M. H. 2008 Hindcasting of wave parameters using different soft computing methods. *Applied Ocean Research* **30** (1), 28–36.

Muzzammil, M. 2010 ANFIS approach to the scour depth prediction at a bridge abutment. *Journal of Hydroinformatics* **12** (4), 474–485.

Muzzammil, M. & Alam, J. 2011 ANFIS-based approach to scour depth prediction at abutments in armored beds. *Journal of Hydroinformatics* **13** (4), 699–713.

Özger, M. 2009 Comparison of fuzzy inference systems for streamflow prediction. *Hydrological Sciences* **54** (2), 261–273. doi:10.1623/hysj.54.2.261.

Özger, M. & Şen, Z. 2007 Prediction of wave parameters by using fuzzy logic approach. *Ocean Engineering* **34** (3), 460–469.

Pesti, G., Shrestha, B. P., Duckstein, L. & Bogárdi, I. 1996 A fuzzy rule-based approach to drought assessment. *Water Resources Research* **32** (6), 1741–1747.

Pongracz, R., Bogardi, I. & Duckstein, L. 1999 Application of fuzzy rule-based modeling technique to regional drought. *Journal of Hydrology* **224** (3), 100–114.

Poursamad, A. & Montazeri, M. 2008 Design of genetic-fuzzy control strategy for parallel hybrid electric vehicles. *Control Engineering Practice* **16** (7), 861–873.

Rezaeianzadeh, M., Tabari, H., Yazdi, A. A., Isik, S. & Kalin, L. 2014 Flood flow forecasting using ANN, ANFIS and regression models. *Neural Computing and Applications* **25** (1), 25–37.

Ris, R. C., Booij, N., Holthuijsen, L. H. & Padilla-Hernandez, R. 1996 *SWAN Cycle 2 User Manual*. Department of Civil Engineering, Delft University of Technology, Delft, The Netherlands.

Russo, M. 2000 Genetic fuzzy learning. *Evolutionary Computation, IEEE Transactions on* **4** (3), 259–273.

Sen, Z. & Altunkaynak, A. 2004 Fuzzy awakening in rainfall-runoff modeling. *Nordic Hydrology* **35**, 31–43.

Seng, T. L., Bin Khalid, M. & Yusof, R. 1999 Tuning of a neuro-fuzzy controller by genetic algorithm. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* **29** (2), 226–236.

Shi, Y., Eberhart, R. & Chen, Y. 1999 Implementation of evolutionary fuzzy systems. *Fuzzy Systems, IEEE Transactions on* **7** (2), 109–119.

Takagi, T. & Sugeno, M. 1985 Fuzzy identification of systems and its applications to modeling and control. *Systems, Man and Cybernetics, IEEE Transactions on* **1**, 116–132.

Tsai, C.-P., Lin, C. & Shen, J.-N. 2002 Neural network for wave forecasting among multi-stations. *Ocean Engineering* **29** (13), 1683–1695.

US Army 2003 *Coastal Engineering Manual*. Chapter II-2, Meteorology and Wave Climate. Engineer Manual 1110-2-1100. US Army Corps of Engineers, Washington, DC, USA.

Zacharia, P. T. & Nearchou, A. C. 2012 Multi-objective fuzzy assembly line balancing using genetic algorithms. *Journal of Intelligent Manufacturing* **23** (3), 615–627.

Zanaganeh, M., Mousavi, S. J. & Shahidi, A. F. E. 2009 A hybrid genetic algorithm–adaptive network-based fuzzy inference system in prediction of wave parameters. *Engineering Applications of Artificial Intelligence* **22** (8), 1194–1202.

Zanganeh, M., Yeganeh-Bakhtiary, A. & Bakhtyar, R. 2011 Combined particle swarm optimization and fuzzy inference system model for estimation of current-induced scour beneath marine pipelines. *Journal of Hydroinformatics* **13** (3), 558–573.

Zanganeh, M., Yeganeh-Bakhtiary, A. & Yamashita, T. 2016 ANFIS and ANN models for the estimation of wind and wave-induced current velocities at Joeutsu-Ogata coast. *Journal of Hydroinformatics* **18** (2), 371–391.

First received 2 May 2016; accepted in revised form 8 November 2016. Available online 3 February 2017