Applied Ocean Research 30 (2008) 172–177

Contents lists available at ScienceDirect

Applied Ocean Research

journal homepage: www.elsevier.com/locate/apor

An alternative approach for the prediction of significant wave heights based on classification and regression trees

J. Mahjoobi [a,b], A. Etemad-Shahidi [b,*]

[a] *Ministry of Energy, Water Research Institute, Hakimieh, Tehran, P.O.Box: 16765-313, Iran*

[b] *School of Civil Engineering, Iran University of Science and Technology, Narmak, Tehran, P.O. Box 16765-163, Iran*

ARTICLE INFO

Article history:
Received 23 March 2008
Received in revised form
29 October 2008
Accepted 1 November 2008
Available online 11 December 2008

Keywords:
Wave prediction
Data mining
Machine learning
Classification
Decision trees
Regression trees
C5 algorithm
CART

ABSTRACT

In this study, the performances of classification and regression trees for the prediction of significant wave heights were investigated. The data set used in this study is comprised of 5 years of wave and wind data gathered from a deep water location in Lake Michigan. Training and testing data include wind speed and wind direction as the input variables and significant wave heights (H_s) as the output variable. To build the classification trees, a C5 algorithm was invoked. Then, significant wave heights for the whole data set were grouped into wave height bins of 0.25 m and a class was assigned to each bin. For evaluation of the developed model, the index of each predicted class was compared with that of the observed data. The CART algorithm was employed for building and evaluating regression trees. Results of decision trees were then compared with those of artificial neural networks (ANNs). The error statistics of decision trees and ANNs were nearly similar. Results indicate that the decision tree, as an efficient novel approach with an acceptable range of error, can be used successfully for prediction of H_s. It is argued that the advantage of decision trees is that, in contrast to neural networks, they represent rules.

1. Introduction

Waves, the most significant maritime phenomenon, due to their complicated and stochastic behavior are known as one of the most difficult subject in coastal and maritime engineering practice. The effect of waves on coastal and marine activities urges us to identify the wave characteristics. Different approaches such as field measurements, theoretical studies and numerical simulations have been used for this purpose. Coastal and offshore engineers generally use these approaches to identify wave climate and extreme wave characteristics as well as annual attributes of waves. Different methods such as empirical, numerical and soft computing approaches have been proposed for significant wave height prediction.

Soft computing techniques such as artificial neural networks have been widely used to predict wave parameters [e.g. [1–5]]. A review of neural network applications in ocean engineering is given in [6]. Recently, other soft computing techniques such as the Fuzzy Inference System (FIS) and the Adaptive-Network-based Fuzzy Inference System (ANFIS) have been used to develop wave prediction models (e.g. [7,8]). These studies have shown that the wind speed is the most important parameter in wave parameter's prediction. Recently, Mahjoobi et al. [9] compared different soft computing methods such as Artificial Neural Networks (ANNs), the Fuzzy Inference System (FIS) and the Adaptive-Network-based Fuzzy Inference System (ANFIS) to hindcast wave parameters. Their results showed that the performances of these methods are nearly the same. Furthermore, using sensitivity analysis, they showed that the wind speed and direction are the most important parameters for wave hindcasting.

As mentioned before, the prediction of significant wave height that is essentially an uncertain and random process is not easy to accomplish by using deterministic equations. Therefore, it is ideally suited to decision trees since they are primarily aimed at the recognition of a random pattern in a given set of input values. Decision trees are helpful in predicting the value of the output of a system from its corresponding random inputs as the application of decision trees does not require knowledge of the underlying physical process as a precondition. Examples of decision tree applications include potential profit analysis of new drugs in pharmaceutical companies [10], medical diagnosis [11] and risk management analysis in petroleum pipeline construction [12]. However, to the authors' knowledge this method has not been applied in wave prediction. In this work, the performances of classification and regression trees as

* Corresponding author. Tel.: +98 21 7391 3170; fax: +98 21 77454053.
E-mail addresses: jmahjoobi@gmail.com (J. Mahjoobi), etemad@iust.ac.ir (A. Etemad-Shahidi).

0141-1187/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.apor.2008.11.001

a new soft computing method for prediction of significant wave height were investigated. For building a classification tree, the C5.0 algorithm [13] is selected due to its speed, small memory requirement and boosting and cross-validation features. Since C5.0 can generate rules that have a straightforward interpretation, it is also quite robust in problems such as missing data and large numbers of fields. The CART algorithm [14] was employed for building and evaluating regression trees. CART builds classification and regression trees for predicting continuous (regression) and categorical predictor variables (classification). CART analysis has a number of advantages over other classification methods, including multivariate logistic regression [14]. First, it is inherently non-parametric. In other words, no assumption is made regarding the underlying distribution of values of the predictor variables. Thus, CART can handle numerical data that are highly skewed or multi-modal, as well as categorical predictors with either ordinal or non-ordinal structure. In this work, results of decision trees were also compared with those of artificial neural networks (ANNs) using statistical error measures. This paper is prepared as follows: Section 2 describes the decision trees, C5 and CART algorithms. Section 3 describes the study area and data set. Section 4 gives the description of results and statistical error analysis and Section 5 covers the summary and conclusions.

2. Decision trees

The patterns and relationships in data can be found using machine learning, statistical analysis, and other data mining techniques. Activities to discover hidden knowledge contained in data sets have been attempted by researchers in different disciplines for a long time. Through a variety of techniques, data mining identifies nuggets of information in bodies of data. Data mining extracts information in such a way that it can be used in areas such as decision support, prediction, forecast, and estimation. Data mining takes advantage of advances in the fields of artificial intelligence (AI) and statistics. Both disciplines have been applied in pattern recognition and classification. Decision trees are data mining methodologies applied in many real-world applications as a powerful solution to classification and prediction problems [15]. A decision tree is an arrangement of tests that prescribes an appropriate test at every step in an analysis. A decision tree is a tree in which each branch node represents a choice between a number of alternatives and each leaf node represents a classification or decision. In general, decision trees represent a disjunction of conjunctions of constraints on the attribute-values of instances. Each path from the tree root to a leaf corresponds to a conjunction of attribute tests and the tree itself to a disjunction of these conjunctions. More specifically, decision trees classify instances by sorting them down the tree from the root node to some leaf node, which provides the classification of the instance. Each node in the tree specifies a test of some attribute of the instance, and each branch descending from that node corresponds to one of the possible values for this attribute [15,16]. The application of decision trees to classification was popularized in machine learning by Quinlan [17,18]. Quinlan's ID3 [17], is a well-known tree-growing algorithm for generating decision trees based on univariate splits. An extended version of this algorithm, called C4.5 [18] and its successor C5.0 [13] use Greedy search methods. They involve growing and pruning decision-tree structures and are typically employed in these algorithms to explore the exponential space of possible models. The algorithm basically chooses the attribute that provides the maximum degree of discrimination between classes locally. Theoretical concepts related to decision trees can be found in many text books (e.g. [15,16,19–21]).

2.1. C5 algorithm

C5.0 [13] is a commercial machine learning program developed by RuleQuest Research and is the successor of the widely used ID3 [17] and C4.5 [18] algorithms. A C5.0 decision tree is constructed using GainRatio. GainRatio is a measure incorporating entropy which measures how unordered the data set is. Entropy is denoted by the following equation:

$$Entropy(S) = \sum_{i=1}^{N} -P(S_i)\log_2(P(S_i)) \tag{1}$$

where, $P(S_i)$ is the probability of class i occurring in the data set S and N is the number of classes. Using the *Entropy*, it is possible to calculate the Information Gain (*Gain*).

Gain is a measure of the improvement in the amount of order given by:

$$Gain(S, A) = Entropy(S) - \sum_{V \in Value(A)} \frac{|S_V|}{|S|} \times Entropy(S_V)$$
$$S_V = \{s \in S | A(S) = V\} \tag{2}$$

where $|S_V|$ and $|S|$ are the number of data points in data sets S_V and S, respectively and A is an attribute.

Gain has a bias towards variables with many values that partition the data set into smaller ordered sets. In order to reduce this bias, the entropy of each variable over its m variable values is calculated as *SplitInformation*:

$$SplitInformation(S, A) = \sum_{i=1}^{m} -\frac{|S_i|}{|S|} \times \log_2\left(\frac{|S_i|}{S}\right). \tag{3}$$

Finally, *GainRatio* is calculated by dividing Gain by *SplitInformation* so that the bias towards variables with large value sets is dampened:

$$GainRatio(S, A) = \frac{Gain(S, A)}{SplitInformation(S, A)}. \tag{4}$$

C5.0 builds a decision tree greedily by splitting the data on the variable that maximizes *GainRatio*.

2.2. CART algorithm

The Classification and Regression Trees (CART) method of Breiman et al. [14] generates binary decision trees. CART is a non-parametric statistical methodology developed for analyzing classification issues either from categorical or continuous dependent variables. If the dependent variable is categorical, CART produces a classification tree. When the dependent variable is continuous, it produces a regression tree. The CART tree is constructed by splitting subsets of the data set using all predictor variables to create two child nodes repeatedly, beginning with the entire data set. The best predictor is chosen using a variety of impurity or diversity measures. The goal is to produce subsets of the data which are as homogeneous as possible with respect to the target variable. In the CART algorithm for each split, each predictor is evaluated to find the best cut point (continuous predictors) or groupings of categories (nominal and ordinal predictors) based on improvement score or reduction in impurity [14]. Then the predictors are compared and the predictor with the best improvement is selected for the split. The process repeats recursively until one of the stopping rules is triggered.

2.2.1. Building a regression tree

The process of constructing a regression tree is similar to that for building a classification tree [14]. Regression tree building centers on three major components: (1) a set of questions of the form: is $X \leq d$? where X is a variable and d is a constant. As with the classification trees, the response to such questions is yes or no; (2) goodness of split criteria for choosing the best split on a variable and (3) the generation of summary statistics for terminal nodes (unique to a regression tree).

In regression tree, the least squared deviation (LSD) impurity measure is used for splitting rules and goodness of fit criteria. The LSD measure $R(t)$ is simply the weighted within node variance for node t, and it is equal to the resubstitution estimate of risk for the node [14]. It is defined as:

$$R(t) = \frac{1}{N_W(t)} \sum_{i \in t} \omega_i f_i \left(y_i - \bar{y}(t)\right)^2 \tag{5}$$

$$\bar{y}(t) = \frac{1}{N_W(t)} \sum_{i \in t} \omega_i f_i y_i \tag{6}$$

$$N_W(t) = \sum_{i \in t} \omega_i f_i \tag{7}$$

where $N_W(t)$ is the weighted number of records in node t, ω_i is the value of the weighting field for record i (if any), f_i is the value of the frequency field (if any), y_i is the value of the target field, and $\bar{y}(t)$ is the mean of the dependent variable (target field) at node t. The LSD criterion function for split s at node t is defined as:

$$Q(s,t) = R(t) - R(t_L) - R(t_R) \tag{8}$$

where, $R(t_R)$ is the sum of squares of the right child node and $R(t_L)$ is the sum of squares of the left child node. The split s is chosen to maximize the value of $Q(s,t)$.

Stopping rules control how the algorithm decides when to stop splitting nodes in the tree. Tree growth proceeds until every leaf node in the tree triggers at least one stopping rule. Any of the following conditions will prevent a node from being split:

a. All records in the node have the same value for all predictor fields used by the model.
b. The number of records in the node is less than the minimum parent node size (user defined).
c. If the number of records in any of the child nodes resulting from the node's best split is less than the minimum child node size (user defined).
d. The best split for the node yields a decrease in impurity that is less than the minimum change in impurity (user defined).

In regression trees, each terminal node's predicted category is the weighted mean of the target values for records in the node ($\bar{y}(t)$).

3. Study area and data set

The data set used in this study comprises of wind and wave data gathered in Lake Michigan from 14 March 2000 to 6 December 2004. The data set was collected by National Data Buoy Center (NDBC) in station 45007 at 42°40′30″ N and 87°01′30″ W (Fig. 1), where water depth is 164.6 m. Wind and wave data were collected using 3-m discus buoy at 1-h intervals. The wind speed at buoy was measured at a height of 5 m above the mean sea level.

In order to select more important waves, all measured significant wave height greater than 0.5 m, were selected (13 243 data points). Then, the data set was divided into two groups. The first one that is comprised of 10 114 wind and wave records (from14 March 2000 to 9 December 2003) was used as training data to develop the models. The second one that is comprised of

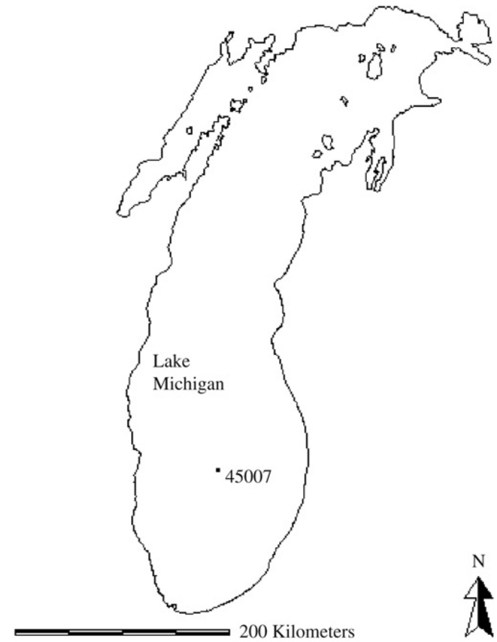

Fig. 1. Lake Michigan bathymetry and location of NDBC buoy 45007 located at 43°37′09″ N and 77°24′18″ W.

Table 1
Ranges and average values of different parameters in training data.

Parameter	Range	Average
Wind speed (m/s)	0.1–18.9	7.155
Wind direction (degree)	2–360	194.65
Significant wave height (m)	0.5–4.25	1.043

Table 2
Ranges and average values of different parameters in testing data.

Parameter	Range	Average
Wind speed (m/s)	0.1–16.8	7.45
Wind direction (degree)	2–360	184.85
Significant wave height (m)	0.5–3.25	1.1

3129 wind and wave records (from 24 March 2004 to 6 December 2004) was used to verify the models. Training and testing data include wind speed and wind direction as input variables [9] and significant wave heights (H_s) as the output variable. Tables 1 and 2 show ranges and average values of different parameters of training and testing data sets, respectively.

4. Building and evaluating classification and regression trees

For building classification trees, the C5 algorithm was invoked. Here, significant wave heights for whole data set were grouped into wave height bins of 0.25 m. Then, a class was assigned to each bin. Tables 3 and 4 show characteristics of this classification for training and testing data, respectively. For evaluation of the C5 algorithm's performance, the index of each predicted class was compared with that of the observed data.

The CART algorithm was employed for building and evaluating regression trees. In addition, a three-layer feed-forward artificial neural network (the back-propagation network) [22] with the sigmoid transfer functions in the hidden layer and a linear transfer function in the output layer was used. From 2 to 10 nodes for the hidden layer were examined. These networks were then trained in a pseudo-parallel fashion. At the end of training process, the model with the lowest error was selected as the final model. The best topology was found to be 2 × 5 × 1 (neurons in the

Table 3
Characteristics of significant wave height classification for training data.

Ranges of wave height (m)	Class	Index of each class (m)	Number of data points in each class (total = 10 114)
0.50–0.75	C1	0.625	3739
0.75–1	C2	0.875	2412
1–1.25	C3	1.125	1499
1.25–1.5	C4	1.375	934
1.5–1.75	C5	1.625	563
1.75–2	C6	1.875	357
2–2.25	C7	2.125	223
2.25–2.5	C8	2.375	119
2.5–2.75	C9	2.625	71
2.75–3	C10	2.875	73
3–3.25	C11	3.125	37
3.25–3.5	C12	3.375	26
3.5–3.75	C13	3.625	23
3.75–4	C14	3.875	22
4–4.25	C15	4.125	16

Table 4
Characteristics of significant wave height classification for testing data.

Ranges of wave height (m)	Class	Index of each class (m)	Number of data points in each class (total = 3129)
0.50–0.75	C1	0.625	1047
0.75–1	C2	0.875	684
1–1.25	C3	1.125	472
1.25–1.5	C4	1.375	315
1.5–1.75	C5	1.625	200
1.75–2	C6	1.875	170
2–2.25	C7	2.125	85
2.25–2.5	C8	2.375	62
2.5–2.75	C9	2.625	54
2.75–3	C10	2.875	31
3–3.25	C11	3.125	9

input × hidden × output layers). For statistical comparison of predicted and observed values bias, root mean square error (RMSE) and scatter index were used. The bias shows the mean error and the scatter index (SI), defined as the root mean square error (RMSE) normalized by the mean of observed values of the reference quantity, is a good non-dimensional error measure:

$$\mathrm{bias} = \bar{y} - \bar{x} \tag{9}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum (x_i - y_i)^2} \tag{10}$$

$$\mathrm{SI} = \frac{\mathrm{RMSE}}{\bar{x}} \tag{11}$$

where, x_i is an observed value, y_i is a predicted value and n is the number of observations, Finally, $\bar{x}$ is the mean of x and $\bar{y}$ is the mean of y.

Comparisons between the observed and predicted significant wave height are shown in Figs. 2–4 for C5, CART and ANNs algorithms, respectively. As seen, all models predicted the trend of the observed data reasonably well. Table 5 shows the error statistics of calculated significant wave heights by C5, CART and ANNs algorithms. As can be seen, the classifications model marginally underpredicted the mean significant wave height (bias = −0.1 m) while the regression tree and ANNs predicted the mean significant wave height accurately. In addition, the scatter index of H_s predicted by C5 algorithm was higher than those of CART and ANNs. Moreover, RMSE of H_s predicted by the C5 algorithm was 0.35 m and RMSE of H_s predicted by the CART and ANNs algorithms were 0.33 m and 0.305 m, respectively. The obtained error statistics are within the range of previous works [e.g. [7,9,23]]. Therefore, it could be concluded that regression trees were more accurate than classification trees. Also,

Table 5
Error statistics of predicted significant wave heights by classification, regression trees and ANNs models for testing data.

Algorithm	bias	RMSE (m)	SI (%)
C5	−0.1	0.35	31.86
CART	−0.003	0.33	29
ANN (2 × 5 × 1)	0.009	0.305	27.74

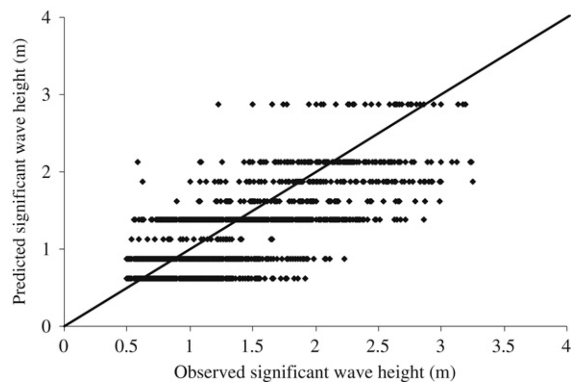

Fig. 2. Comparison between observed and predicted significant wave height by classification tree (C5 algorithm) for testing data.

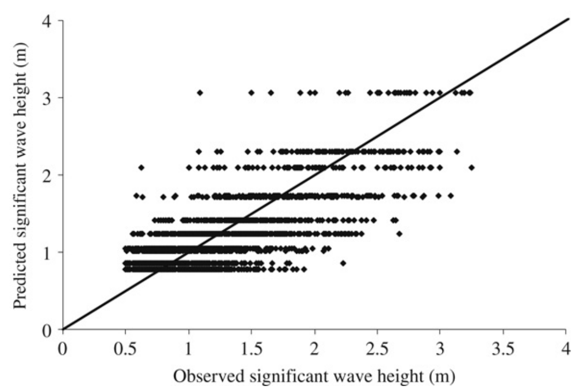

Fig. 3. Comparison between observed and predicted significant wave height by regression tree (CART algorithm) for testing data.

results show that ANNs models were marginally more accurate than classification and regression trees. It is also interesting to investigate the accuracies of the models in predicting the H_s in different height ranges. As seen in Table 6, the prediction of significant wave height by C5, CART and ANNs algorithms for $H_s \leq 1$ m were slightly overestimated. While, for $H_s > 1$ m the predictions were underestimated. In these models, bias and RMSE parameters were increased and scatter index was decreased by increment of significant wave height. Similar trend was obtained by Kazeminezhad et al. [7] and Mahjoobi et al. [9] in Lake Ontario, giving support to this finding.

As a result of Table 6, it can be said that the C5 algorithm leads to better results with significant wave heights of less than 1 m while CART and ANNs algorithms are more applicable for significant wave heights greater than 1 m. For example, in predicting significant wave heights over 2.5 m the scatter indices were 18.92, 23.47 and 28.47 percent and biases were −0.21 m, −0.43 m and −0.64 m for ANNs, CART and C5 algorithms, respectively. Thus, it can be concluded that regression trees are more appropriate and precise than classification trees for prediction of significant wave heights over 1 m. Also regression trees are more applicable than classification trees since it is not necessary to classify significant wave heights in regression trees.

Table 6
Statistics of predicted significant wave heights by C5, CART and ANNs algorithms in different height ranges for testing data.

Wave height ranges (m)	C5 algorithm			CART algorithm			ANN ($2 \times 5 \times 1$)		
	bias	RMSE	SI (%)	bias	RMSE	SI (%)	bias	RMSE	SI (%)
$0.5 \le H_s \le 1$	0.01	0.21	28.52	0.13	0.22	30.45	0.13	0.22	30.12
$1 < H_s \le 1.5$	−0.14	0.37	29.84	−0.09	0.32	26	−0.07	0.30	24.77
$1.5 < H_s \le 2.5$	−0.33	0.53	28.42	−0.26	0.51	27.03	−0.21	0.46	24.72
$H_s > 2.5$	−0.64	0.78	28.47	−0.43	0.65	23.47	−0.33	0.51	18.92

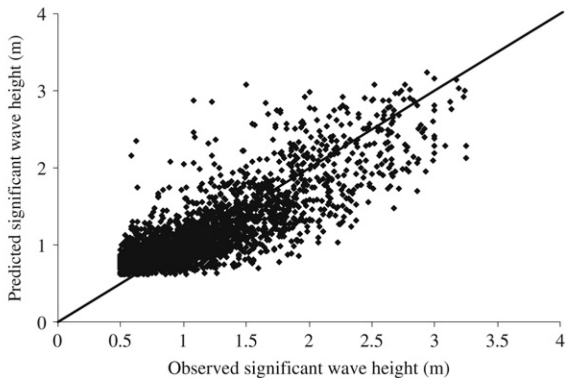

Fig. 4. Comparison between observed and predicted significant wave height by ANNs ($2 \times 5 \times 1$) for testing data.

Table 7
Error statistics of predicted significant wave height by classification, regression trees and ANNs models for testing data after eliminating wave direction.

Algorithm	bias	RMSE (m)	SI (%)
C5	−0.14	0.367	33.86
CART	−0.005	0.335	30.5
ANN ($1 \times 3 \times 1$)	−0.004	0.32	29.43

The advantage of decision trees is due to the fact that, in contrast to ANNs that work more like a black box, decision trees represent rules. Rules can readily be expressed so that humans can understand them. In addition, to find the relative importance of input parameters in ANNs we need to perform sensitivity analysis after the ANNs is completely built [9]. However, in decision trees the algorithm itself determines the important parameters through branching of inputs. Also in ANNs we need to find the best topology, both the number of the hidden layers and the number of neurons in each hidden layer. The process of finding these parameters could be performed via trial and error, which is a time-consuming sequence of actions. On the contrary, decision trees are non-parametric and therefore are more convenient to use. Besides, they need lower run-time and are automatic. Figs. 5a and 5b show the regression tree diagrams. As can be seen, the greatest number of branchings were performed using wind speed. Therefore, wind speed is the most important parameter for significant wave height prediction. This result is consistent with the existing understanding of importance of wind speed [e.g. [9, 23,24]].

In order to investigate the importance of wind direction, it was eliminated from the input variables and C5, CART and ANNs models were built and evaluated again. Table 7 shows the error statistics of calculated significant wave heights by these algorithms after eliminating the direction. As seen, error statistics slightly increased after elimination of the wind direction.

Table 6 also indicates that the higher values of H_s(>2.5 m) are under-predicted. Therefore, following [25] it was decided to perform separate training for higher values using CART and ANNs models. Table 8 shows the error statistics of calculated significant wave heights by CART and ANNs algorithms, for $H_s > 2.5$ m. As can be seen, error statistics decreased more. For example, the

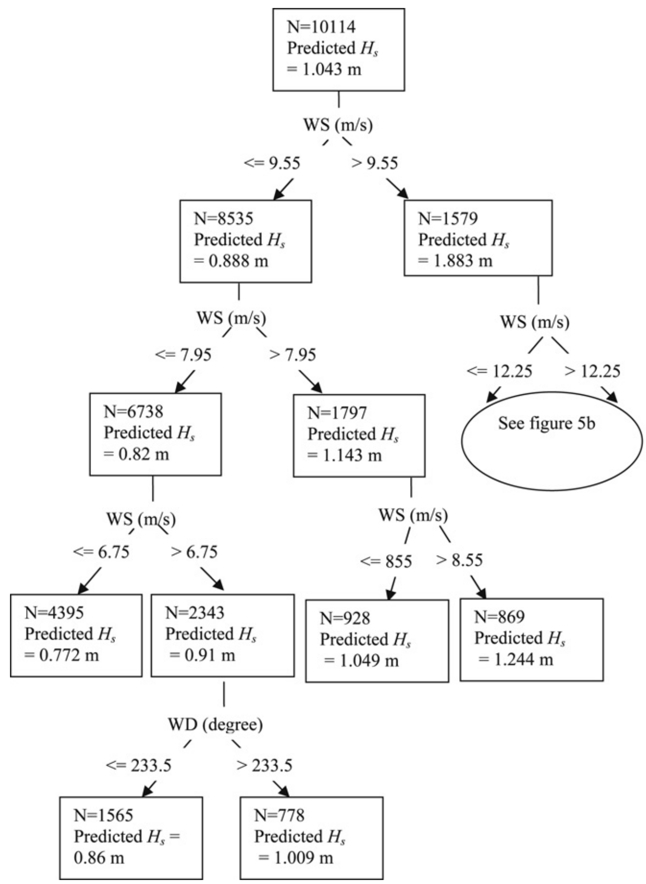

Fig. 5a. Regression tree model generated by CART algorithm, N is the number of data in each node, WS is the wind speed and WD is the wind direction.

Table 8
Error statistics of calculated significant wave height by CART and ANNs algorithms, for higher values of $H_s > 2.5$ (for testing data).

Algorithm	bias	RMSE (m)	SI (%)
CART	0.21	0.38	13.9
ANN ($2 \times 5 \times 1$)	0.21	0.32	11.6

scatter indices are 11.6 and 13.9 and RMSE's are 0.32 m and 0.38 m for ANNs and CART algorithms, respectively. Therefore, the development of separate regression trees and ANNs models for higher values of significant wave heights improved the accuracy of their prediction.

5. Summary and conclusions

Significant wave height is the most important parameter in the design of coastal and offshore structures. In this study, classification and regression trees were used successfully for prediction of the significant wave height. The C5 algorithm was invoked for building and evaluating classification trees and the CART algorithm was employed for building and evaluating

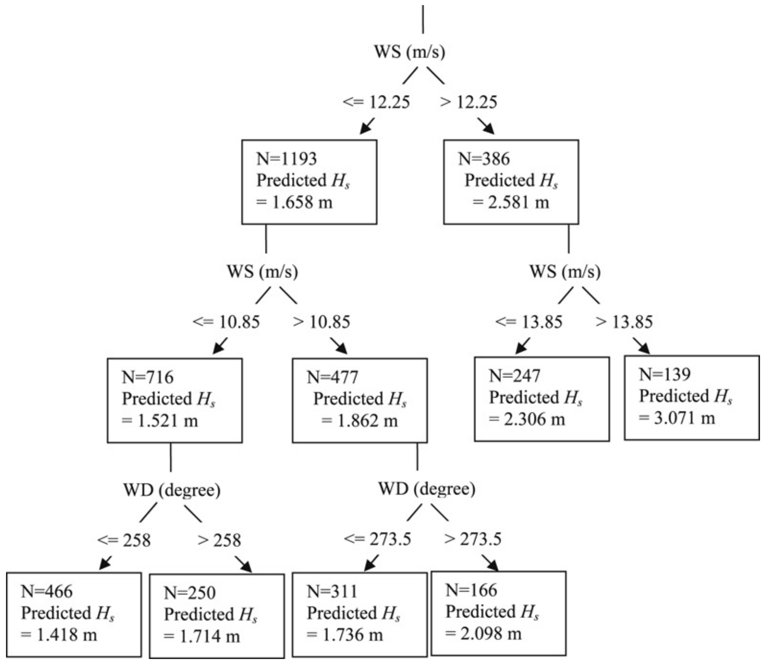

Fig. 5b. Regression tree model generated by CART algorithm, N is the number of data in each node, WS is the wind speed and WD is the wind direction.

regression trees. The data set used in this paper is comprised of wave and wind data gathered from a deep water location in Lake Michigan. Results show that the decision trees algorithm is skillful in the prediction of H_s in the studied case. Furthermore, it was found that regression trees are more appropriate and precise than classification trees for the prediction of significant wave heights over 1 m. In addition, it was argued that regression trees are more applicable than classification trees as they do not require height classification. Results of decision trees were also compared with those of artificial neural networks (ANNs). Results show that error statistics of decision trees and ANNs were similar, while ANNs models were marginally more accurate than classification and regression trees. The advantage of decision trees is due to the fact that, in contrast to neural networks, decision trees represent rules; while ANNs work more like a black box. Also, in ANNs to find the relative importance of input parameters we need to perform sensitivity analysis after the ANNs is completely built. However, in decision trees the algorithm itself determines the relative importance of parameters through branching of inputs. Also, in ANNs the number of the hidden layers and the number of neurons in each hidden layer should be determined. This process can be performed via trial and error, which is a time-consuming sequence of actions. While decision trees are non-parametric and therefore are more convenient for wave prediction. They also need less run-time and are more automatic. Hence, regression trees can be used as a cost effective and easy to use tool for engineers and scientists with much less effort.

Acknowledgements

The authors are grateful to Mohammad Hossein Kazeminezhad for his helpful comments. In addition, the first author expresses special thanks to Ehsan Adeli Mosabbeb for his great efforts.

References

[1] Makarynskyy O, Pires-Silva AA, Makarynska D, Ventura-Soares C. Artificial neural networks in wave predictions at the west coast of Portugal. Comput Geosci 2005;31(4):415–24.
[2] Makarynskyy O. Improving wave predictions with artificial neural networks. Ocean Eng 2004;31(5–6):709–24.
[3] Jain P, Deo MC. Real-time wave forecasts off the western Indian coast. Appl Ocean Res 2007;29:72–9.
[4] Deo MC, Naidu CS. Real time wave forecasting using neural networks. Ocean Eng 1999;26:191–203.
[5] Deo MC, Jha A, Chaphekar AS, Ravicant K. Neural networks for wave forecasting. Ocean Eng 2001;28:889–98.
[6] Jain P, Deo MC. Neural networks in ocean engineering. Int J Ships Offshore Struct 2006;1(1):25–35.
[7] Kazeminezhad MH, Etemad-Shahidi A, Mousavi SJ. Application of fuzzy inference system in the prediction wave parameters. Ocean Eng 2005;32:1709–25.
[8] Ozger M, Sen Z. Prediction of wave parameters by using fuzzy logic approach. Ocean Eng 2007;34:460–9.
[9] Mahjoobi J, Etemad-Shahidi A, Kazeminezhad MH. Hindcasting of wave parameters using different soft computing methods. Appl Ocean Res 2008;30:28–36.
[10] Boer FP. Financial management of R & D. Res Technol Management 2002;45(4):23–35.
[11] Baker JJ. Medicare payment system for hospital inpatients: Diagnosis-related groups. J Health Care Financ 2002;28(3):1–13.
[12] Dey PK. Project risk management: A combined analytic hierarchy process and decision tree approach. Cost Eng 2002;44(3):13–26.
[13] Quinlan JR. C5.0: An Informal Tuorial. RuleQuest. www.rulequest.com/see5-unix.html1998.
[14] Breiman L, Friedman JH, Olshen RA, Stone CJ. Classification and regression trees. Belmont: Wadsworth Statistical Press; 1984.
[15] Kantardzic M. Data Mining: Concepts, models, methods, and algorithms. John Wiley & Sons; 2003.
[16] Hand D, Heikki M, Padhraic S. Principles of data mining. MIT press; 2001.
[17] Quinlan JR. Induction of decision trees. Mach Learn 1986;1:81–106.
[18] Quinlan JR. C4.5: Programs for machine learning. San Mateo, (CA): Morgan Kaufman; 1993.
[19] Nemati N, Barko C. Organizational data mining: Leveraging enterprise data resources for optimal performance. Idea Group Publishing; 2004.
[20] Witten I, Eibe F. Practical machine learning tools and techniques. 2th ed. Elsevier Inc; 2005.
[21] Wang J. Data mining: Opportunities and challenges. Idea Group Publishing; 2003.
[22] Haykin S. Neural networks: A comprehensive foundation. 2th ed. NJ: Prentice-Hall; 1999. p. 842.
[23] Deo MC, Jha A, Chaphekar AS, Ravicant K. Neural networks for wave forecasting. Ocean Eng 2001;28:889–98.
[24] Moeini MH, Etemad-Shahidi A. Application of two numerical models for wave hindcasting in Lake Erie. Appl Ocean Res 2007;29:137–45.
[25] Kalra R, Deo MC. Genetic programming for retrieving missing information in wave records along the west coast of India. Appl Ocean Res 2007;29:99–111.