Ocean Engineering 54 (2012) 244–250

Contents lists available at SciVerse ScienceDirect

Ocean Engineering

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

Wave height prediction using the rough set theory

Armaghan Abed-Elmdoust [a], Reza Kerachian [b,*]

[a] School of Civil Engineering, College of Engineering, University of Tehran, Tehran, Iran
[b] School of Civil Engineering and Center of Excellence for Engineering and Management of Civil Infrastructures, College of Engineering, University of Tehran, Tehran, Iran

ARTICLE INFO

Article history:
Received 13 August 2011
Accepted 7 July 2012
Available online 10 August 2012

Keywords:
Wave height prediction
Rough Set Theory
Lake Superior
Uncertainties
Decision rules

ABSTRACT

Integrated interdisciplinary modeling techniques, providing reliable and accurate estimates for wave characteristics, have gained attention in recent years. With the ability to express knowledge in a rule-based form, the Rough Set Theory (RST) has been successfully employed in many fields. However the application of RST has not been investigated in wave height (WH) prediction. In this paper, the RST is applied to Lake Superior in North America to find some simple rules, called decision rules, for WH prediction. Decision rules are derived by expressing WH as functions of wind data gathered by the National Data Buoy Center (NDBC). Comparing results of RST with results of other soft computing techniques such as Support Vector Machines (SVMs), Bayesian Networks (BNs), Artificial Neural Networks (ANNs) and Adaptive Neuro-Fuzzy Inference System (ANFIS) indicates that the RST outperforms other soft computing techniques in WH prediction and provide some simple decision rules which can be accurately used by decision makers and engineers.

1. Introduction

The estimation of wave height (WH) is essential for almost any engineering activity in the ocean. Different empirical, numerical and soft computing approaches have already been proposed for WH height prediction. Recently, artificial neural networks (ANNs) have been widely used to predict wave parameters. Agrawal and Deo (2002), Makarynskyy (2004), and Bazargan et al. (2007) used the ANNs for WH prediction. A review of neural network applications in ocean engineering is given by Jain and Deo (2006). Kazeminezhad et al. (2005), Ozger and Sen (2007), Lin and Chang (2008) and Mahjoobi et al. (2008) used Fuzzy Inference System (FIS) and Adaptive Neuro-Fuzzy Inference Systems (ANFIS) for WH prediction. Egozcue et al. (2005), Scotto and Soares (2007) applied Bayesian Inference technique in WH prediction. Mahjoobi and Mosabbeb (2009) used Support Vector Machines (SVMs) to predict WH. These studies have shown that the wind speed is the most important input data in wave parameters predictions. Malekmohamadi et al. (2011) attempted to use and evaluate several soft computing techniques in order to map wind data to WH. They showed that the ANN, ANSIS and SVM could provide acceptable predictions, while the results of the BNs were unreliable.

The use of the Rough Set Theory (RST) for data analysis has many important advantages. For example, it provides efficient algorithms for finding hidden patterns in data, finds minimal sets of data using a data reduction technique and evaluates the significance of data. It is anticipated that the RST will play a substantial role in data mining and soft computing, especially in large scale and complex problems. Barbagallo et al. (2006) developed reservoir operating rules by using a rough set approach. Ip et al. (2007) studied applications of the RST to river environment quality evaluation. Gong and Sun (2005) Gong et al. 2006 applied the rough set approach to water resources allocation in river basins. Pai and Lee (2010) developed a rough set-based model for water quality analysis. Lashteh-Neshaei and Pirouz (2010) used RST in site selection decision making for water reservoirs.

In this paper, perhaps for the first time, the RST is utilized for WH forecasting in western Lake Superior. Comparing the results of this technique with the results of Malekmohamadi et al. (2011) is the main aim of this paper.

2. Basic concepts of Rough Set Theory (RST)

The RST introduced by Pawlak (1982, 1991) is a mathematical approach to deal with a specific type of uncertainty in data, i.e. situations in which objects having equal description are assigned to the same classes. The main character of the RST is rule induction. Moreover, RST allows an information reduction procedure based on the consequence of particular subsets of attributes.

The RST is based on the supposition that with every object of the universe (U) there is an assured amount of information associated, expressed by means of some attributes (A) used for

* Corresponding author. Tel.: +98 21 61112176; fax: +98 21 66403808.
E-mail addresses: armaghanabed@ut.ac.ir (A. Abed-Elmdoust), kerachian@ut.ac.ir (R. Kerachian).

0029-8018/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.oceaneng.2012.07.020

object explanation. With each attribute $a \in A$ a function is defined as $f_a : U \to \upsilon_a$, where υ_a is the value of a. More exactly, this information can be represented in a data table, named decision-making (DM) table in which rows refer to dissimilar objects and columns refer to the considered attributes. Each cell of this table points to, therefore, a description (quantitative or qualitative) of the object positioned in that row by means of the attribute in the related column.

In a DM, the set of attributes (A) is divided into condition attributes (set C) and decision attributes (set D), such that $C \cup D = A$, $C \cap D = \phi$. Since it shows the dependencies between condition and decision attributes, a DM may also be seen as a set of decision rules. These are logical statements of the type "if…, then…", where the antecedent condition part (if) refers to the value(s) assumed by one or more condition attributes, and the consequence decision part (then) refers to the values assumed by the decision attribute(s). Objects having the same depiction are indiscernible (similar) with respect to the available information. The indiscernibility relation provides a division of the universe into portions of indiscernible objects (elementary sets) that can be used as "blocks" to build knowledge about an actual world.

For example, a DM concerning the records of WH is presented in Table 1, where $U=\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$, $C=\{I_1, I_2, I_3\}$, $D=\{\text{WH}\}$, $\upsilon_a=\{1, 2, 3\}$ for each attribute and $\upsilon_{WH}=\{3, 4\}$. For example, $f_{I_1}(1)=3$ means that the average wind speed during time steps 1 and 0 is in the third class of wind speed and $f_{WH}(1)=4$ denotes that the WH is in the fourth class of WH at time step 1.

An attribute belongs to an equivalence relation in which the members who have the same value according to that attribute are placed in a separate subset. The equivalence relation, related to $a \in A$, is denoted by $R(a)$.

Any subsets S of A verifies a relation $In(S)$ on U, called an *indiscernibility relation*, and is defined as follows: $(x,y) \in In(S)$ if and only if $a(x)=a(y)$, $\forall\, a \in S$, where $a(x)$ denotes the value of attribute a of member x. Clearly, $In(S)$ is an equivalence relation, which is the meeting point of all equivalence relations $R(s)$ related to attribute $s \in S$, i.e., $In(S) = \bigcap_{s \in S} R(s)$. The equivalence class of $In(S)$ will be denoted by $U/In(S)$. For any member x of U, the equivalence of x in relation to $In(S)$ is represented as $[x]_{In(S)}$ (Ip et al., 2007). For example, in Table 1, we may dig out that $R(I_1)=\{\{1, 2, 3, 4, 9, 10\}, \{5, 6, 7, 8\}\}$, $R(I_2)=\{\{1, 2, 3, 4\}, \{7\}, \{5, 6, 8, 9, 10\}\}$, $R(I_3)=\{\{1, 2\}, \{3, 5, 6, 7\}, \{4, 8, 9, 10\}\}$ and $In(WH)=\{\{1, 2, 7, 8, 9, 10\}, \{3, 4, 5, 6\}\}$, $In(I_1, I_2, I_3)=\{\{1, 2\}, \{5, 6\}, \{9, 10\}, \{3\}, \{4\}, \{7\}, \{8\}\}$.

We say that the set of attributes $R \subseteq A$ *depends* on the set of attributes $P \subseteq A$ in the DM table (denotation$P \to R$) if $\tilde{P} \subseteq \tilde{R}$. Finding dependencies between attributes would help the reduction of the set of attributes. Subset $P \subseteq A$ in the DM table is *independent* if for every $P' \subseteq P$, $\tilde{P'} \supset \tilde{P}$; otherwise subset $P \subseteq A$ in the DM table is dependent. For instance, consider the information system in Table 1, we can find the exemplary dependency $\{I_1, I_2, I_3\} \to I_1$. Such dependency may be checked easily by looking at $R(I_1)=\{\{1, 2, 3, 4, 9, 10\}, \{5, 6, 7, 8\}\}$ and $R(I_1, I_2, I_3)=In(I_1, I_2, I_3)=\{\{1, 2\}, \{5, 6\}, \{9, 10\}, \{3\}, \{4\}, \{7\}, \{8\}\}$. Notice that each elementary set of $R(I_1, I_2, I_3)$ is a subset of each elementary set of $R(I_1)$. Other examples of dependencies are $I_1, I_2, I_3 \to I_2$ and $I_1, I_2, I_3 \to I_3$. Such dependencies are not among other subset of attributes. For example, consider $R(I_1)= \{\{1, 2, 3, 4, 9, 10\}, \{5, 6, 7, 8\}\}$ and $R(I_2, I_3)=\{\{1, 2\}, \{3\}, \{4\}, \{5, 6\}, \{7\}, \{8, 9, 10\}\}$. Observe that all elementary set of $R(I_2, I_3)$ is not a subset of any elementary set of $R(I_1)$. More specifically, $\{8, 9, 10\}$ is not a subset of any elementary set of $R(I_1)$.

In practical applications we are attracted in reducing those redundant attributes in the DM table (i.e. we are paying attention in achieving the so called reducts). The least minimal subset which ensures the same quality of classification as the set of all attributes is a *reduct* in the DM table. It is sometimes called a *minimal set of attributes*. Let us notice that an information system may have more than one reducts/minimal set. Intersection of all reducts/minimal sets is called the *core*. The core is a collection of the most significant attributes for the classification in the system. Let us notice that in the information system from Table 1 there are three similar reducts, I_1, I_2, I_3, and a core I_1, I_2, I_3.

Let X be a subset of the universe $U (X \subseteq U)$. The subset X may be categorized by two ordinary sets, called the lower and the upper approximations. The *lower approximation* of X is defined as the union of all the elementary sets completely included in X, more formally:

$$\underline{R}X = x \in U \big| [x]_R \subseteq X \tag{1}$$

The *upper approximation* of X is composed of all the elementary sets which have a non-empty intersection with X (whose elements x, therefore, may belong to X):

$$\overline{R}X = x \in U \big| [x]_R \cap X \neq \phi \tag{2}$$

The distinction between the upper and the lower approximations comprises the *boundary region* of the Rough Set, whose elements cannot be characterized with certainty as belonging or not belonging to X, using the available information. The information about objects from the boundary region is, therefore, conflicting or ambiguous. For this reason, the number of objects from the boundary region may be used as a measure of vagueness of the information about X. Consider the information system in Table 1, let $R = In(I_1, I_2, I_3)=\{\{1, 2\}, \{5, 6\}, \{9, 10\}, \{3\}, \{4\}, \{7\}, \{8\}\}$ and $X=\{2, 3, 5, 7\}$, then the approximations are $\underline{R}X=\{3, 7\}$ and $\overline{R}X=\{1, 2, 3, 5, 6, 7\}$.

Finally, the rough set approach leads to the induction of a set of decision rules representing the knowledge contained in the DM. Each rule is supported by a certain number of objects from U. More precisely, an object $x \in U$ supports a decision rule if its description matches both the condition and decision part of the rule. Let (U,A) be a DM table with $C \cup D=A$, $C \cap D=\phi$, where C is the set of conditional attributes and D is the set of DM attributes. Further:

$$dec(X_i) = (c, attribute\ value\ of\ c) \big| c \in X_i \in U/InC \quad and,$$

$$dec(X_j) = (d, attribute\ value\ of\ d) \big| d \in X_j \in U/InD.$$

A DM rule is defined as follows (Ip et al., 2007):

$$r_{ij} : dec(X_i) = dec(Y_j), \quad X_i \cap Y_j \neq \phi$$

Table 1
A sample of decision-making table including wind and wave height data.

Time	Conditional attributes			DM attribute
	I_1[a]	I_2[b]	I_3[c]	WH
1	2	3	3	4
2	2	3	3	4
3	2	3	2	3
4	2	3	1	3
5	3	2	2	3
6	3	2	2	3
7	3	1	2	4
8	3	2	1	4
9	2	2	1	4
10	2	2	1	4

[a] The average of wind speeds during time steps t and $t-1$.
[b] The average of wind speeds during time steps $t-2$, $t-3$ and $t-4$.
[c] The average of wind speed during time steps $t-5$, $t-6$, $t-7$ and $t-8$.