Ocean Engineering 146 (2017) 352–362

Contents lists available at ScienceDirect

Ocean Engineering

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

Forecasting hurricane wave height in Gulf of Mexico using soft computing methods

Somayeh Mafi, Gholamreza Amirinia *

Department of Civil and Environmental Engineering, Florida A&M University – Florida State University College of Engineering, Tallahassee, FL 32310, USA

ARTICLE INFO

Keywords:
Hurricane
Wave height
Artificial neural network
Support vector machine
Random forest

ABSTRACT

The main objective of this paper was to use different soft computing methods to forecast hurricane wave height over Gulf of Mexico. For this purpose, three well-known soft computing methods; Support Vector Machine (SVM), Artificial Neural Network (ANN), and Random Forest (RF) were used. The friction velocity was also used instead of wind speed at 10 m height. For this case, three different assumptions were considered for computing friction velocity. For the analysis, one month (from August 20th, 2008 to September 20th, 2008) data including passage of hurricanes Gustav and Ike over Gulf of Mexico were collected from six different buoys. Next, seven different parameters were used to train the models based on the data observed from five buoys. Later the trained models were used to forecast observed wave height in the sixth buoy in different lead times. The results showed that using more number of input parameters leads to better performance of the methods, especially in longer lead times. Moreover, in longer lead times, the effect of other parameters such as air pressure were increased and assisted the models to outperform previous models. In addition; different assumptions for friction velocity eventuated comparable results.

1. Introduction

Accurate analysis and forecasting wave field is crucial in offshore engineering issues such as renewable energy, wave loading on offshore structures, and marine operations (Amirinia et al., 2017a, 2017b; Hassanzadeh and Naughton, 2016; Yeganeh-Bakhtiary and Rezaee, 2017). In this case, forecasting hurricane wave characteristics in offshore regions is more important in different aspects. The operation of offshore oil and gas platforms as well as riser emergency shutdown valves (RESDVs) in platforms are fully depending on sea state, especially wave height (Goff, 2016). Fast and accurate analysis of the hurricane sea state is essential to manage the evacuation of manned platforms, stopping the platform operation, and preventing disasters such as oil spills.

The effect of hurricane winds on sea state is one of the differences between hurricane winds and regular boundary layer winds. This is because hurricane wind fields are characteristically intense, spatially inhomogeneous and directionally varying (Young and Burchell, 1996). Several studies focused on simulating the hurricane wave field by implementing different methods (Chen et al., 2013; Dietrich et al., 2011; Holthuijsen et al., 2012; Hu and Chen, 2011; Huang et al., 2013; Kennedy et al., 2010; Lynett et al., 2010). Young and Burchell (1996) used satellite wind data to present hurricane wave. They confirmed that both the maximum wind speed in the hurricane and the velocity of forward movement, play significant roles in determining the significant wave height. Young (1998) examined an extensive database of one-dimensional spectra observed during hurricanes and showed that for one-dimensional spectra, the previous well-known spectra for wind-induced waves, approximate the data well. Later, Young (2003) by reviewing several studies showed that hurricane spectra are typically unimodal (single peak) and consistent with forms reported for fetch limited growth; however, he pointed that his model provides no information on the directional spreading of the waves within a hurricane. Dietrich et al. (2012) implemented a coupled SWAN-ADCRIC model for address hurricane waves. They validated their analysis with data from hurricanes Katrina and Rita and showed that the wave-circulation interactions are important in hurricane wave modeling.

For analyzing the wind-induced wave fields, different methods have been proposed. The SMB method (Bretschneider, 1951; Sverdrup and Munk, 1947), Coastal Engineering Manual (CEM) (US Army, 2003), Shore Protection Manual (SPM) (US Army, 1984), and Donelan (1980) are examples of fast and simple empirical methods. Although the empirical methods were fast and accurate, they are applicable for limited cases (Kamranzad et al., 2011). In order to cover a wider range of the problems, numerical models such as Wave Analysis Model (WAM) (Komen et al., 1996) and Simulating WAves Nearshore (SWAN) (Booij

* Corresponding author.
E-mail address: gamirinia@fsu.edu (G. Amirinia).

https://doi.org/10.1016/j.oceaneng.2017.10.003
Received 6 June 2017; Received in revised form 18 August 2017; Accepted 1 October 2017

et al., 1999) were introduced for deep and shallow water conditions respectively. Numerical models are generally more accurate than the empirical methods; however, they are costly and time-consuming (Goda, 2003; Kamranzad et al., 2011). By increasing the number of real observations in the last decade, soft computing methods such as Artificial Neural Network (ANN) (Mahjoobi et al., 2008; Makarynskyy, 2004; Tsai et al., 2002; Vimala et al., 2014), Fuzzy Inference Systems (FIMs) (Mahjoobi et al., 2008; Malekmohamadi et al., 2011; Özger, 2010), and regression trees (Etemad-Shahidi and Mahjoobi, 2009; Samadi et al., 2014) became popular for predicting wave characteristics. The most important advantages of soft computing method in competition with empirical and numerical methods are simple application as well as less required computational costs (Kamranzad et al., 2011).

Soft computing methods have been used widely in different engineering problems (Fatehnia et al., 2016; Mafi et al., 2013; Mahjoobi and Etemad-Shahidi, 2008; Nagy et al., 2002) including simulating or forecasting wave height. Makarynskyy et al. (2005), Kamranzad et al. (2011), and Krishna kumar et al. (2017) used ANN to predict significant wave height. These studies mostly used wind speed and direction for training and testing the models. In addition, Lee (2006) used ANN to predict the storm surge by adding pressure, wind speed, wind direction and harmonic analysis tidal level. Kazeminezhad et al. (2005) and Mahjoobi et al. (2008) applied fuzzy inference system (FIS) and adaptive-network-based fuzzy inference system (ANFIS) for predicting wave parameters. These studies used wind speed, direction, and duration as well as fetch length for the analysis. Mahjoobi and Etemad-Shahidi (2008) compared a classification method (C5), ANN, and classification and regression tree (CART) models for predicting significant wave height. They used wind speed and direction for prediction and found that regression trees are more appropriate methods than classification tree. Also, they found that the accuracy of regression tree and ANN are almost similar. More recently, Mahjoobi and Mosabbeb (2009) and Salcedo-Sanz et al. (2015) applied SVM method for estimating the significant wave height. In this case, Mahjoobi and Mosabbeb (2009) used wind speed to predict wave height by using support vector machine (SVM) and found that SVM results in slightly better accuracy than ANN. The mentioned studies mostly used limited input parameters for hindcasting non-hurricane wave characteristics; however, these method has been rarely used for predicting hurricane wave characteristics during incoming hours. In addition, Fernández et al. (2015) and Cornejo-Bueno et al. (2016) implemented different learning classifiers and genetic algorithm-extreme learning machine methods to predict significant wave height and energy flux predictions.

The main objective of this paper is to forecast hurricane wave height by implementing different soft computing methods instead of costly and time-consuming numerical models. Also, in addition to well-known methods such as SVM and ANN, another soft computing method, Random Forrest (RF), was implemented for the analysis. Moreover, the effects of different friction velocities based on the regular and hurricane conditions were targeted. For these purposes, in section 2, three different machine learning methods (SVM, ANN, and RF) were briefly presented. In section 3, one month data from August 20th, 2008 to September 20th, 2008 were collected from six different offshore buoys in Gulf of Mexico (GoM) from National Buoy Data Center (NBDC). The time period was selected in such a way that contains the occurrence of the hurricane Ike in offshore GoM (September 1st, 2008 to September 15th, 2008). In addition, all buoys locations were close to hurricanes Ike's and Gustav's path within offshore GoM. In this section, summary of the observed data were presented. Later in section 3, the mentioned machine learning methods were used to train models based on the observed data from buoys. The parameters considered in the analysis were friction velocity, wave height and period, air pressure, air temperature, water temperature, and dew point temperature. At this step, three different assumptions for friction velocity based on the regular and hurricane conditions were used. In section 4, the models were trained to forecast the wave height in 3 h, 6 h, 12 h, and 24 h based on the observed data from five buoys. Later, the trained model was used to forecast the observed wave height in the sixth buoy. At the end, the results from three models by considering different friction velocity assumptions were presented and compared to previous studies.

2. Methods

Three different machine learning method have been used in this paper for predicting hurricane waves in GoM. In this section, a brief description of each method is presented.

2.1. Support vector machine (SVM)

SVM is a supervised learning method which is commonly used for classification, regression, and density estimation problems. In this method, in order to classify the data, a hyperplane is designed in such a way that minimize the misclassification error and maximize the margin between hyperplane and the nearest data point (support vectors). Moreover, SVM can use a nonlinear transformation to map input space into higher feature dimensional space. For this purpose, the training data sets can be defined as $[x_i, y_i]$ where $x_i \in R^n$ is the input vector, n is the dimension of input vector, and $y_i \in [-1, 1]$ is the output vector. The SVM uses the quadratic programming techniques to find optimal hyperplanes that separates classes. The quadratic programming can be expressed as (Malekmohamadi et al., 2011):

$$\min \frac{1}{2} w^t w + C \sum_{i=1}^{N} \xi_i \tag{1}$$

$$y_i(w\phi(x_i) + b) + \xi_i - 1 \geq 0 \tag{2}$$

where $\phi(x_i)$ maps the training data in high dimensional feature space, w is the weight vector, b is the bias term, C is the penalty for the error term, and $\xi_i \geq 0$ is the slack variable (Vapnik, 2013). The slack parameter, ξ_i, and parameter C are used for misclassification of noisy data and controlling over-fitting respectively. Fig. 1 illustrates the theory of SVM for choosing the optimal hyperplane that maximizing the margin.

The presented formulation in eq. (1) and eq. (2) can be solved by means of Lagrange techniques. After constructing the optimal hyperplane which separates the classes, the classification decision can be carried out by eq. (3) as:

$$f(y) = sign\left(\sum_{i=1}^{N} y_i c_i k(x_i, x_j) + b\right) \tag{3}$$

where $sign()$ represents the sign function, c_i is the Lagrange multiplayer

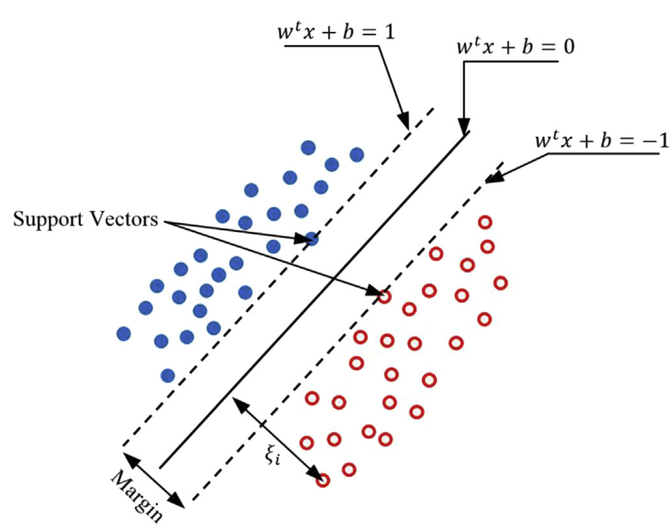

Fig. 1. SVM graphical hypothesis.