Application of Machine Learning Techniques in Wave Height Forecasting In Marine Environment – A Review

Akhil C P¹ & Paresh Chandra Deka²
P.G. Student, Department of Applied Mechanics and Hydraulics, NITK Surathkal, Karnataka, India¹
Associate Professor, Department of Applied Mechanics and Hydraulics, NITK Surathkal, Karnataka, India²

ABSTRACT: Prediction of wave height is one of the most important issues in coastal engineering and for coastal structures. Over the past few years, advancements in the prediction of significant wave height using soft computing techniques have been developed drastically. Apart from traditional prediction of significant wave height, soft computing technique has explored a new way for prediction of significant wave height. This paper explains studies carried out in prediction of significant wave height using different soft computing techniques such as Support Vector Regression (SVR), Extreme Learning Machine (ELM), Adaptive Neuro-Fuzzy Inference System (ANFIS), Neuro-Wavelet Technique, Artificial Neural Network (ANN), Genetic Algorithm (GA) and other models. Every method has its own advantages and disadvantages and is explained in this paper.

KEYWORDS: Wave forecasting, Machine learning, Soft Computing, prediction, ANN, SVR, ANFIS, GA.

I. INTRODUCTION

Wave height is one of the most important factors in coastal processes and coastal engineering studies. In case of energy extraction from waves, sediment movements, harbour design and soil erosion, wave height plays a vital role in these. Long-term observed data’s are needed for all the practical applications. There are different methods for finding wave heights such as field measurements, theoretical studies and numerical stimulation. But in most of these cases there won’t be long-term measurements, so prediction of wave height is essential.

Now day’s soft computing based models have been used for wave prediction. Soft computing based methods give results with high accuracy and time taken for training and prediction is very less compared to other traditional methods. This paper explains different studies carried out in prediction of significant wave height using different soft computing techniques such as Support Vector Regression (SVR), Extreme Learning Machine (ELM), Adaptive Neuro-Fuzzy Inference System (ANFIS), Neuro-Wavelet technique, Artificial Neural Network (ANN), Genetic Algorithm (GA) and other models. Every method has its own advantages and disadvantages and is explained in this paper.

II. SOFT COMPUTING STUDIES IN WAVE HEIGHT FORECASTING

Artificial neural network has been used by Deo and Chaudhari (1998) [1] for forecasting tides on either coast of India. From monthly mean wave height data Deo and Kumar (2000) [3] obtained weekly wave height data from it. Back propagation has been implemented to tide prediction by Mandal (2001) [2]. Feed forward type artificial neural network with three layered has been developed by Deo et al. (2001) [4] for predicting wave height and average wave period of three hourly values. Current wind speed and previous one time step wind speed has been used as the inputs. For short interval data, he has trained separately for monsoon and post monsoon weather condition. They also conclude that wind fetch and duration doesn’t have an important role as input in neural networks. Artificial neural network has been used by Makarynskyy (2004) [5] for forecasting significant wave heights and zero-up-crossing wave periods for a time interval of 1 hour up to 24 hour. He also increased the accuracy of forecasting through data merging neural nets. Feed forward back propagation method and radial bias function has been used by Kalra et al. (2005) [6] for predicting
significant wave height. Wind speed, significant wave height and average wave period has been given as input. Wave period and wave height of cyclone generated wave has been hindcasted by Rao and Mandal (2005) [7] by the help of two input configuration in ANN. Speed of forward motion of cyclone, radius of maximum wind speed, and central pressure was used as input in first configuration, while wind speed and fetch was used as input in second configuration. Storm tidal level in Taiwan during typhoon has been predicted Lee (2006) [9]. Wind direction, wind speed, harmonic analysis tidal level, pressure of that station was used as the input parameter for ANN. Wave height prediction of 3, 6, 12, and 24 hours has been done by Mandal and Prabaharan (2006) [10] using recurrent neural network. A review of ANN applications in different ocean engineering scenarios has been done by Jain and Deo (2006) [11].

Gradient descent back propagation algorithm and ANN was used by Makarynskyy et al. (2005) [8] for forecasting wave height and period 3, 6, 12, and 24 in advance. Better simulation of wave parameter was for shorter lead time than for longer interval prediction. MIKE21 and SWAN has been used for wave parameter prediction by Moeini and Etemad (2007) [15]. Mean wave direction, significant wave height and peak spectral period of Lake Erie has been predicted. The result shows that MIKE21 is more accurate than SWAN model. Moeini et. al. (2012) used a hybrid SWAN and ANN model. Mean wind speed, wind duration, means wind directions are the best input parameter for error prediction network. The results were compared with plain ANN and hybrid SWAN-ANN model, and hybrid model shows better error prediction than plain ANN.

Regression trees and ANN was used by Mahjoobi and Etemad-Shahidi (2008) [16] for predicting significant wave height. They used wind speed and wind direction as the input parameters and significant wave height as the output parameter. Regression tree was built and evaluated by CART algorithm. Result shows that error of both ANN and Regression tree was almost same, regression tree is a novel approach with tolerable range of error. They argue that the main advantage of regression tree than ANN is that they represent rules. Kemal (2008) [19] used ANN and regression method for wave forecasting. He used monthly mean wind speeds, air temperature and sea level pressures as input parameters and used seven different combinations of these to get the best result. The result shows that the ANN with all input parameter shows the best result. In these input parameters wind speed has the maximum effect on output wave height. Air temperature showed better performance when used as input parameter in single and in combination, but pressure has effect only when it is added into input combinations. Etemad and Mahjoobi (2009) [20] have compared neural network and M5' model for predicting significant wave height. ANN needs to find network parameters such as neurons and number of hidden layers and which need to be finding through trial and error method which will take time. Model tree has an advantage, that it will represent set of rules which can be easy understand by humans. Wind speed is used as input parameter and significant wave height as the output parameter for comparing M5' and ANN. The result shows that model tree was more accurate than feed-forward network. SenayAsma et. al. (2012) [28] has used Multiple linear regression (MLR) and ANN for significant wave height forecasting. They used different variables such as air temperature, wind direction, wind speed, wind gust, air pressure, water temperature, wave height and average wave direction at west coast of India, Goa. They concluded that for better result of significant wave height, non-linear model with wind gust and wind speed of previous time step and water temperature, air pressure, and air temperature at the same time step should be used. The main advantage of ANN is that it has smaller error rates and the advantage of MLR is that with less parameter they show smaller standard error. Both methods proved that wind guest and wind speed of previous time step is the main variable in predicting significant wave height and they suggest for using ANN since it has less mean squared error compare to MLR.

Coastal Engineering Manual (CEM) and Adaptive-Network-Based Fuzzy Inference System (ANFIS) method was used for wave forecasting by M.H. Kazeminezhad et al. (2005). ANFIS is more accurate than CEM in terms of predicting significant wave height. Fuzzy logic approach has been used by Özger and Sen (2007) [14] for prediction of wave height and period. They investigated the relation between wind speed, current and previous wave characteristics. They compared there model with classical Auto Regressive Moving Average with exogenous input (ARMAX) model and the result shows that fuzzy logic worked better with a high difference level. Various soft computing techniques has been used by Mahjoobi et al. (2008) [17] such as artificial neural networks (ANN), fuzzy inference system (FIS) and adaptive network-based fuzzy inference system (ANFIS) for wave parameter prediction. They obtained result such that almost all are performing more or less same. They also showed that wind speed and wind direction is the main parameters in wave prediction through sensitivity analysis. Özger (2009) [21] used hybrid neuro-fuzzy approach for estimating wave characteristics of ocean. To overcome the problem of missing data, he shows that the significant wave
height of one buoy location can be predicted using the data from adjacent buoys using a neuro-fuzzy approach. This approach shows a possibility of retrieving missing and also to determine the optimal position of buoys. Georgiose et al. (2009) [22] has done wind wave modelling through fuzzy interface system (FIS) and developed FIS act as a valuable tool for forecasting wave parameters. The proposed method showed a quick convergence of observed and predicted data. Ozger et al. (2014) [38] used fuzzy inference system for predicting wave parameters. They used fetch length, duration, wind speed and wave heights as variables for wave height prediction, and used different combinations of these. They compared these results with Jonswap, CEM, Wilson, SPM and found that ANFIS is the best model with less error. They also found that combination of wind speed and duration of wind blowing as input gives better results. Hashim et al. (2016) [44] done selection of climatic parameters using enhanced Takagi-Sugeno-based fuzzy which affects wave height forecasting. They used wind speed, wind direction, sea surface temperature and air temperature as input variables and significant wave height as output. The main aim of that study was to identify the most predominant input parameters influencing wave height prediction. They found that wind speed, air temperature, sea surface temperature and wind direction has most to least influence in wave height prediction respectively. They also found that the combination of wind speed, air temperature and wind direction are the most influencing set of input parameters. To improve the wave height prediction they suggest considering air temperature and sea surface temperature. The advantage of ANFIS model is that its adaptability to optimization, computationally efficient, adaptive methods and is faster compared to other control strategies. Stefanakos (2016) [45] forecasted non stationary wind and wave data using combination of Fuzzy Inference Systems (FIS) and Adaptive Network- based Fuzzy Inference Systems (ANFIS), it will remove non-stationary character of wind wave data before we predict the data. The data is first decomposed by means of the above-mentioned non stationary modelling into a residual and seasonal mean value time series multiplied by a seasonal standard deviation and FIS or ANFIS model is applied. By that method they obtained point wise prediction for specific data points and field-wise prediction for whole field of wave parameters. The results were compared with only FIS and ANFIS and it shows that the combination outperforms these methods.

Genetic programming, a novel approach has been used for wave forecasting by Surabhi and Deo (2008) [18]. Genetic algorithm has expanded to Genetic Programming (GP) is suitable to explore the relation between input and output data. For a lead time of 3, 6, 12 and 24 h, significant wave height has been forecasted. They have compared the results of GP with ANN and GP showed a better result, it was because of relatively efficient handling of large amount of data. Genetic programming showed that it is a promising tool for predicting wave parameters. Morteza (2009) [23] developed a hybrid genetic algorithm and ANFIS (GA–ANFIS). In this clustering and rule base parameters has been simultaneously optimized by using GAs and artificial neural nets (ANNs). Clustering of parameter in which structure and number of fuzzy rules are optimized and controlled by GA and ANFIS is used tuning the parameter of rule base generated by GA. Results were compared with ANFIS, GA-ANFIS model and also with Shore Protection Manual (SPM), GA-ANFIS showed high accuracy in terms of prediction. Canellas et al. (2010) [25] has predicted wave height using genetic algorithm. They have used 44 years meteorological and oceanographic records to train the algorithm. They obtained a correlation between numerical model and genetic algorithm output as 0.80<=r<=0.97. They have predicted significant wave height at each buoy with 1 hour ahead forecast time by only using 12 hour previous wind data with a correlation of 0.80<=r<=0.97. They shows that both GA and numerical models can be combined together to increase the prediction accuracy. Raoet al. (2013) [32] used genetic algorithm for wave forecasting. They used a combination of genetic algorithm (GA) and empirical orthogonal function (EOF) analysis. Two different variants of GA are used, univariate GA and multivariate GA to the time series. The main drawback of GA is that it requires large computation and data support. To overcome this various data adaptive approaches has been proposed, these approaches only require few time series wave height data at a particular location for predicting data. The result shows that univariate GA performs better than multivariate GA. Abdusselam (2013) [37] used geno-multilayer perceptron for wave forecasting. With using genetic algorithm identifying the weighting coefficient of multilayer perceptron is called as geno-multilayer perceptron. They predicted significant wave height using current wind speed and previous significant wave height as input variables. The results were compared with genetic algorithm and back propagation algorithm, and genetic algorithm with optimized weighting coefficient rivals better result than back propagation with observed data. They found a good consistency between the observed significant wave height and predicted wave height using geno-multilayer perceptron. The advantages of GA over other optimizing techniques are that, it simultaneously search many point in solution space, it is mainly based on basic arithmetical operations, it can even work with discontinuous data, many solution points can be produced by it, it is a stochastic and probabilistic processes but not a
Deterministic process. (Buckles and Petry, 1994; Altunkaynak, 2008, 2009; S-en, 2004; Abdusselam, 2013). The main drawback of GA is that it takes longer running time compared to other optimization techniques. Abdüsselam and Wang (2012) used hybrid Genetic Algorithm (GA) and Kalman Filtering (KF) for prediction significant wave height. GA is used for optimization and KF is used for prediction. KF used to predict based on dynamic and stochastic structures. The result of GKF was compared with ANN and GKF shows better result than ANN. Cornejo et al. (2016) [43] predicted significant wave height and energy flux with hybrid Grouping Genetic Algorithm Extreme Learning Machine approach (GGA-ELM). GGA solve the problem by looking for several subsets with important features, and is used to predict reduced number of features. Significant wave height was predicted by extreme learning machine (ELM) and the main advantage of this method is that it is extremely fast- training approach with excellent result based on prediction quality. The results were compared between ELM and SVM, both obtained best results using GGA.

Mahjoobi and Ehsan (2009) [20] has investigated the use Support Vector Machine (SVM) in significant wave height forecasting. They used wind wave data from deep water of Lake Michigan. Current wind speed and data up to previous hours has been used as input. Error statistics for significant wave height prediction decreases as the lag in wind speed increases. The results of SVM was compared with ANN, Radial basis function (RBF), Multi-layer perceptron (MLP) models and SVM outperforms ANN with very less computational time. Maziar and Mostafa (2011) [27] have used Support Vector Regression (SVR) for gap-filling and predicting wave parameters. They have investigated the method of SVR in wave forecasting. The developed model SVR was done in order to estimate the wave surface spectral density for extensive range of wave frequencies out of wave parameters of significant wave height and period. Randomly sea states were selected for training and testing of model, and this different approach in training showed better result. They also proposed a new method of gap-filling. They show that SVR model can be used as the powerful wave forecasting tool and also a best model for gap-filling. Salcedo-Sanz et al. (2015) [39] showed how to predict significant wave height using Support vector regression (SVR) from x-band radar images. They show that SVR algorithm removes part, Arabian Sea, and the predicted results was found satisfactory for two time step ahead. Deka and Prahlada (2012) [31] used hybrid wavelet and ANN network naming Wavelet Neural Network (WLNN) for predicting significant wave height up to 48 hour lead time. Time series significant wave height has been decomposed to multi resolution time series data by using discrete wavelet transformation (DWT). Time series multi resolution data were used as the input data for ANN model for predicting significant wave height. The result shows that WLNN model worked better than ANN model using multi resolution time series input data. It shows that for lower lead time lesser decomposition level of DWT is needed for better result but for higher lead time due to uncertainty more decomposition levels of DWT are needed. Prediction lag in wave height forecasting has been removed by using hybrid neuro wavelet model, Pradnya et al. (2015) [40]. Neuro wavelet model is a combination of Discrete wavelet transformation and Artificial neural network. They have predicted significant wave height up to 36hr in advance. Discrete wavelet transformation (DWT) has used and it will decompose the signal to approximate (low) and detail (high) frequency components without any correlation between wave height data. This decomposed data is used for training and the output of testing data is reconstructed using inverse DWT. They showed that using hybrid neuro-wavelet technique prediction lag in forecasting wave height has removed completely. Pradnya and Londhe (2016) [46] used neuro wavelet technique to predicted extreme wave height such as Katrina 2005, Dean 2007, Gustav 2008, Ike 2008, Irene 2011. Wavelet with third, fifth and seventh level of decompositions makes it possible for predicting extreme events with a high accuracy. Wavelet coefficients after 7th decomposition performs better than the models trained with 3rd and 5th
level of decomposition. They accurately predicted extreme events such as Katrina of 16.91m in 36hr advance by using 5d and 7d models and concluded that neuro wavelet can accurately predict extreme wave height.

III. DISCUSSION AND CONCLUDING REMARKS

- The review focused on the application of various soft computing techniques in predicting significant wave height.
- The predictive efficiency of a machine learning approach depends upon quality and size of the data set available.
- ANN takes more computational time and finds difficulty in determining the effective structure of the network.
- Hybrid models give better result than plain model.
- Wind speed, air temperature, sea surface temperature and wind direction has most to least influence in wave height prediction respectively.
- Combination of wind speed, air temperature and wind direction are the most influencing set of input parameters
- The advantage of ANFIS model is that its adaptability to optimization, computationally efficient, adaptive methods and is faster compared to other control strategies.
- Hybrid FIS and ANFIS can remove non-stationary character of wind wave data
- The advantages of GA is that it can even work with discontinuous data
- The main drawback of GA is that it requires large computation data support and longer running time
- Advantage of Extreme learning machine is that it is an extremely fast-training approach with excellent result based on prediction quality.
- Advantage of SVR algorithm is that it removes calibration necessity and can predict wave height accurately after training.

REFERENCES


Copyright to IJIRSET DOI:10.15680/IJIRSET.2017.0605300 9653


RoslanHashima,b.n, ChandrabhusanRoy a, ShervinMotamemia,b, ShahaboddinShamshirbande,c,n, DaliborPeković, 2016. Selection of climatic parameters affecting wave height prediction using an enhanced Takagi-Sugeno-based fuzzy methodology. Renewable and Sustainable Energy Reviews 60(2016)246–257


Copyright to IJIRSET DO:10.15680/IJIRSET.2017.0605300 9654

ISSN(Online): 2319-8753
ISSN (Print): 2347-6710

International Journal of Innovative Research in Science, Engineering and Technology
(An ISO 3297: 2007 Certified Organization)
Website: www.ijirset.com
Vol. 6, Issue 5, May 2017


