Ocean Wave Height Prediction using Ensemble of Extreme Learning Machine

N. Krishna kumar[a,*], R. Savitha[b], Abdullah Al Mamun[a]

[a] *School of Electrical and Computer Engineering, National University of Singapore, Singapore*
[b] *Institute of Infocomm Research, Agency for Science, Technology and Research, Singapore*

Abstract

The intense increase in offshore operational activities warrants periodical and accurate prediction of the wave characteristics. Usually, complex numerical models that require high computational power are used in this prediction. To overcome these challenges of these numerical models, in this paper, we propose the use of an ensemble of Extreme Learning Machine (Ens-ELM) to predict the daily wave height. We exploit the randomness of initialization in ELM to obtain better generalization performance. This is done by constructing an Ensemble of ELM, with the parameters of each ELM initialized in distinct regions of the input space. For each sample in the data set, the output of the ELM with the least mean square for each sample in the data set is reported as its output. We study the performance of the Ens-ELM to predict the daily wave height in 10 stations of varying terrains from Gulf of Mexico, Brazil and Korean region. The Ens-ELM network is trained using the past wave data and the measured atmospheric conditions obtained in these stations between Jan 1, 2011 and Dec 31, 2014 and is tested with data in these stations between Jan 1, 2015 and Aug 30, 2015. In this study, the performance of Ens-ELM is evaluated in comparison with ELM, Online Sequential ELM (OS-ELM), and Support Vector Regression (SVR). From this study, we infer that the Ens-ELM out performs ELM, OS-ELM and SVR in the daily wave height prediction.

Keywords: Wave characteristics, SLFN, Extreme Learning Machine, SVR, OS-ELM.

1. Introduction

The intense increase in various offshore operations has spurred an interest in accurate prediction of wave characteristics [1]. There has been widespread

*Corresponding author

Email addresses: `a0107324@u.nus.edu` (N. Krishna kumar), `ramasamysa@i2r.a-star.edu.sg` (R. Savitha), `eleaam@nus.edu.sg` (Abdullah Al Mamun)

research in developing wave models and the guide to wave analysis and forecasting [2] from World Meteorological Organization classifies these models into three generation of wave models. It also claims that the wave energy at a given location is influenced by advection, external environment and dissipation and that wave modeling is based on the spectral energy-balance equation defined as:

$$\frac{\partial E(f,\theta,t)}{\partial t} = S = S_{in} + S_{nl} + S_{ds} \tag{1}$$

where $E(f,\theta,t)$ is the wave spectrum that depends on frequency f and propagation direction θ. S is the net source function that depends on S_{in}, the external wave making factors such as local wind and local current, S_{nl} the non-linear energy conduction by wave-wave interactions and S_{ds} the dissipation related to wave-disperse processes and its reaction with turbulence of the water layer on the surface.

In numerical models, energy input term S_{in} is generally expressed as the resonant interaction [3] and feedback [4] between waves and turbulent pressure patterns in the air. The dissipation term S_{ds} is usually formulated in the wave model as a function of wave height [5]. Representation of the nonlinear term S_{nl} poses a tough challenge to the task of wave modeling [2], [6], [7]. While the most primitive methods of wave models did not have an accurate representation S_{nl}, parametric, analytic and numerical approximations were adapted to represent this term in the latter methods. Wave models using numerical methods to represent S_{nl} include WAM [8], JONSWAP [5], SWAN[9], WIS, WAVEWATCH III [10] etc. Despite the computational load and complexity ([11], [12], [13], [14], [15]) involved in all these methods, the wave prediction by these models are not accurate, and do not generalize over various oceans and seas. Therefore, there is a need to develop computationally efficient nonlinear methods for wave modeling and forecasting, which can be applied uniformly across different oceans and seas. In recent years, there has been a tremendous increase in application of soft-computing methods for prediction of wave height. Of the soft-computing method, neural networks have the ability to approximate any complex nonlinear process without priori system knowledge. They have been extensively used for solving wave parameters prediction problems [16], [17], [18], [19], [20], [21], [22], [23], [24], [25]. [7],[26], [27]. Numerous principles like resilient back propagation [16], nonlinear autoregressive model recurrent algorithm [19] etc., are used to deal this problem. Wave forecasting using neural network models are widely reviewed in the literature [7], [21], [20], [23], [24], [25]. In [28], [29], [30], Adaptive Network-bases Fuzzy Inference System (ANFIS) is used to predict wave parameters from data set consisting of wind and wave observations collected from buoy 45012 in deep water location of Lake Ontario. However, the performance of ANFIS deteriorates for wave heights above 1m. The performance of Support Vector Machines (SVM), artificial neural networks (ANN), ANFIS and Bayesian Networks (BN) in predicting wave heights at western Lake superior in the United States of America (USA) is presented in [31]. The long-term predictions of significant wave height at Northern North sea using BN and extreme value techniques has also been studied [32]. However, the soft attributes of BN leads to ambiguity in the three parameters namely, location, scale and shape of Generalized Extreme Value (GEV) distribution. In [33], rough set theory (RST) has been used to predict wave height in Lake superior, USA. Performance of RST exceeds SVM, BN, ANN and ANFIS, but it requires tedious mathematical computations. A hybrid work using numerical wave models and typical ANN has proven to improve the efficiency of the former [11], [34], [35], [36]. Notably among these works, [37] recorded the ability of ANN in simplifying the tedious and intricate calculations of the numerical models.

Soft-computing methods have also been efficiently used to reconstruct the missing time series of ocean buoy data [38], [39]. X- band radar images are used for prediction of near by wave field, but it is influenced by shadowing effects, radar location on offshore structures. The efficiency of significant wave height prediction from X-band radar images is improved by combining these images with Support Vector Regression (SVR) and Multi-Layer Perceptron (MLP) neural network [40]. The improvement in prediction accuracy is achieved at the cost of need for reliable radar images that requires detailed simulation works on wave field. In [41], a combined approach using genetic algorithm and ELM has been used for significant wave height prediction, but the correlation coefficient of prediction is fairly less. Eight different ordinal and nominal classifiers together with SVR is used for prediction of significant wave height at 6h time-horizon in [42]. Grouping Genetic Algorithm- Extreme Learning Machine (GGA-ELM) approach is used in [43]. Although GGA-ELM has been shown to perform accurate predictions, it requires careful selection of operators for GGA crossover and mutation process. Moreover, [43] does not provide a generalized approach that is uniform across all oceans and seas.

Earlier, it has been shown a maximum of N hidden neurons [44], [45], [46] with randomly chosen radial basis function centers [47], [48], [49], [50], [51] is sufficient to learn any function defined by N samples. These research findings paved the way for the development of a computationally less intensive and a fast learning Extreme Learning Machines (ELM). ELM uses analytical methods to determine output parameters of a feedforward network with randomly chosen input parameters. The universal approximation capability of ELM has been proved by illustrating the equivalent relationship between ELM and the orthonormal method [52], [44], [53]. As the input parameters are chosen randomly, the performance of ELM is heavily influenced by the choice of input parameters and the number of hidden neurons. Researchers have attempted to devise methods for optimum initialization of input parameters [54], [55], [56] and to fix optimum number of hidden neurons [57], [58], [59], [60]. Further, previous studies have shown the advantages of ensembling several ELM to obtain stable, generalized performance [61], [62], [63], [64], [65]. In [65], a cognitive Ens-ELM classifier based on hinge loss error function has been developed for steganalysis. The ensemble classifier combines the outputs of individual classifier by calculating the weighted sum of outputs of the classifiers for all the samples. Ens-ELM methods have been used to solve regression tasks, too. For example, an ensemble of ELM constructed by obtaining the mean of individual

ELM is used to forecast the price of crude oil in [63]. Studies in [66], [67] show that ensemble of extreme learning machines is more stable than a single extreme learning machine. Ensemble learning in ELM has been widely reviewed in [68], and it has been reinforced that the stability of performance in ELM is enhanced through ensemble learning techniques.

In this paper, we use an Ensemble of ELM (Ens-ELM) to forecast wave heights. We collect data across three distinct geographical locations of increased offshore activity. Our regions of interest are: (a) Gulf of Mexico, (b) Brazil and (c) Korean region. Initially, we locate stations of varying terrains in the three selected regions. We collect six hourly data from these stations. The data set used in this work covers different wind speeds such as gentle, breeze, storm, hurricane, cyclones and typhoons. The Ens-ELM is constructed from several individual ELM, each initialized in different regions of the input space. The response of each ELM to each sample varies based on its initialization. We identify the ELM with minimum mean square error for each sample, and construct the ensemble by combining these ELM. At the end, the ELM that does not contribute to the ensemble is pruned. Thus trained ELM is used predict wave height at these stations for every 12AM based on the atmospheric condition and wave height at 6AM, 12PM and 6PM. This study is localized to the region considered. Thus, we predict the 24-hourly wave height at individual stations, based on the earlier data.

This paper is organized as follows : In Section 2 we present a detailed description of the datasets used in this paper. Section 3 describes the Ens-ELM predictor in detail. Performance of Ens-ELM in wave height prediction is studied and results of this study is presented in Section 4. Finally, Section 5 summarizes the study.

2. Materials

Marine data is usually acquired periodically at 6 hour intervals through ships, ocean buoys and land weather stations. Remote measurements of surface wind speed is enabled through microwave sensors such as scatterometers, radiometers and radar altimeters from satellites. Of all these data, the data obtained from moored ocean buoys are a reliable source of data owing to the following factors: (a) Measurement errors in ocean buoys are carefully evaded through thoughtful consideration of sensors location so as to avoid atmospheric exposure problems (b) Calculation of the sampling and averaging periods considering the buoy motion, (c) A priori calibration of multiple sensors that account for redundancy.

Considering the advantages of using data acquired through ocean buoys, we collect ocean buoys data from the following locations to enable this study:

(a) Gulf of Mexico; Source: National Oceanic and Atmospheric Administration

(http://www.ndbc.noaa.gov/)

(b) Brazil Waters; Source: European Marine Observation and Data Network (http://www.emodnet-physics.eu/map/) and Pacific Marine Environmental Laboratory (http://www.pmel.noaa.gov/)

(c) Korean region; Source: European Marine Observation and Data Network (http://www.emodnet-physics.eu/map/)

The measurements are in acceptable accuracy limits as specified by the National Data Buoy Center which also meets World Meteorological Organization (WMO) regulations.

Eight offshore stations at various terrains from the three regions are considered in this study. The details of the eight stations, including the latitude, longitude, and water depth obtained from the various organizations listed above, are presented in Table 1. From the table, it can be seen that the chosen locations are geographically widely distributed. Further, the water depths at these stations range from as small as a few meters to a few thousand meters below sea level.

Table 1: Details of the Selected Buoy/ Ship Stations

Region	Station Location	Latitude	Longitude	Water depth(m)	Type of Buoy / Ship	Station Id
Gulf of Mexico	Mid Gulf	25.888	-89.658	3365	3-m Discus	42001
	East of Brownsville Texas	26.091	-93.758	3125.1	3-m Discus	42002
	South of Freeport Texas	27.907	-95.353	125	3-m Discus	42019
	East of Galveston	29.232	-94.413	83.2	3-m Discus	42035
Brazil	Rio Grande Do Sul	-33.434	-52.077	80	3-m Discus	31053
	Rio Grande Rise	-28.492	-47.537	90	3-m Discus	31262
Korean Region	Korea Strait	34	127.5	130	3-m Discus	22103
	Tsushima Basin	37.53	130	1729	3-m Discus	22105
	Tsushima Basin	37.46	131.11	2277	3-m Discus	21229
	Yellow Sea	36.25	125.75	60	3-m Discus	22108

The data's are selected , such that it covers wide range wind speeds such as calm weather, tropical storms, cyclones and hurricanes. The various types of wind relating wind speed to observed conditions at sea are represented by the Beaufort number [2] defined by Francis Beaufort.

The study is conducted for the period between Jan 1, 2011 and Aug 30, 2015. The information of storms, hurricanes and cyclones that occurred during the year 2010 to 2015 in the each of these regions are collected from National Hurricane Center and WMO as listed in Table 2. Although a number of parameters are available in the meteorological data for each of the selected stations, only the parameters that are vital for wave generation, namely, the latitude and longitude of the stations, wind speed [3],[4], [69], month, air to sea temperature difference [2], water depth [70], [71], atmospheric pressure [72], [73], [2] wave heights at previous 6th hour, 12th hour and 18th hour are used in this study. We develop an ensemble of Extreme Learning Machine that is capable to predict the daily wave heights based on these parameters. In the next section 3, we present the Ens-ELM predictor in detail.

Table 2: List of Cyclones,Storms and Hurricanes (Period 2010-2014)

Region	Year	Period	Name	Type	Max.Wind Speed (Knots)
Gulf of Mexico	2010	Jun 25-Jul 2	Alex	Hurricane	87
	2010	Sep 5-9	Hermine	Tropical Storm	60
	2011	Jul 27-30	Don	Cyclone	45
	2011	Sep 2-5	Lee	Tropical Storm	50
	2012	Aug 21-Sep 2	Issac	Cyclone	55
	2013	Aug 25-26	Fernand	Tropical Storm	50
	2013	Sep 12-17.	Ingrid	Hurricane	75
	2013	Oct 2-6	Karen	Tropical Storm	55
	2014	Sep 1-3	Dolly	Tropical Storm	45
Brazil Waters	2011	Mar 14	Arani	Near Gale	29
	2015	Feb 5	Bapo	Hurricane	91
Korean Region	2011	Jul 21-24	Nock-ten	Storm	65
	2011	Jul 27-Aug 15	Muifa	Hurricane	95
	2011	Sep 23-30	Nesat	Hurricane	78
	2012	Jul 14-20	Khanun	Storm	52
	2012	Jul 27-Aug 4	Damrey	Hurricane	70
	2012	Aug 17-Sep 1	Tembin	Hurricane	78
	2012	Sep 10-12	Sanba	Hurricane	108
	2014	Jul 2-13	Neoguri	Hurricane	100
	2014	Jul 16-26	Matmo	Hurricane	70
	2014	Sep 17-25	Fung-Wong	Strong Gale	44
	2015	Jun 25-Jul 12	Chan-Hom	Hurricane	91
	2015	Jul 29-Aug 12	Sudelor	Hurricane	113

3. Ensemble of Extreme Learning Machine Predictor

In this Section, we describe the learning algorithm of the Ens-ELM in detail. First, we present the ELM algorithm in Section 3.1 and then describe the Ens-ELM in Section 3.2.

3.1. Extreme Learning Machine

ELM is a fast learning algorithm that uses analytical methods to determine output parameters of a feedforward network with randomly chosen input parameters. For any data set with L training samples, $\{(\mathbf{u}^1, \mathbf{v}^1), \cdots, (\mathbf{u}^t, \mathbf{v}^t), \cdots, (\mathbf{u}^L, \mathbf{v}^L)\}$; $\mathbf{u}^t \in \Re^m$, $\mathbf{v}^t \in \Re^n$, the regression problem is defined as approximating the functional relationship $f : \mathbf{u}^t \rightarrow \mathbf{v}^t$, as accurately as possible, enabling output prediction for new samples with better accuracy. An ELM with m input neurons and n output neurons is required to solve this regression problem.

Let us consider a feed-forward network with K hidden neurons. In the study conducted in this paper, each neuron in the hidden layer employs the Gaussian activation function. Therefore, the response of the k-th hidden neuron for the t-th sample is:

$$h_k^t(\mathbf{u}^t, \boldsymbol{\mu}_k, \sigma_k) = exp\Bigg(- \Bigg(\frac{(\mathbf{u}^t - \boldsymbol{\mu}_k)^T(\mathbf{u}^t - \boldsymbol{\mu}_k)}{2\sigma_k^2} \Bigg) \Bigg); \quad k = 1, \cdots, K \quad (2)$$

where, $\boldsymbol{\mu}_k \in \Re^m$ is the Gaussian center of the k-th hidden neuron, and $\sigma_k \in \Re$ is its width.

The neurons in the output layer of the ELM are linear. Hence, the output of the l-th neuron for the t-th sample ($\widehat{v}_l^t$) is given by:

$$\widehat{v}_l^t = w_{lk} h_k^t; \ l = 1, \cdots, n; \ k = 1, \cdots, K \quad (3)$$

where $w_{lk} \in \Re$ is the output weight connecting the k-th hidden neuron and the l-th output neuron.

Eq. (3) can be expressed in matrix form as

$$\widehat{V} = WH \quad (4)$$

where, $W \in \Re^{n \times K}$ is the output weight matrix and the hidden layer output matrix (H) is:

$$H(\mathbf{U}, \boldsymbol{\mu}, \boldsymbol{\sigma}) = \begin{pmatrix} h_1^1(\mathbf{u}^1, \boldsymbol{\mu}_1, \sigma_1) & \cdots & h_1^L(\mathbf{u}^L, \boldsymbol{\mu}_1, \sigma_1) \\ \vdots & \cdots & \vdots \\ h_K^1(\mathbf{u}^1, \boldsymbol{\mu}_K, \sigma_K) & \cdots & h_K^L(\mathbf{u}^L, \boldsymbol{\mu}_K, \sigma_K) \end{pmatrix} \quad (5)$$

The objective of the ELM is to minimize the error between the actual and predicted outputs. In other words, it has to estimate the optimal output weights corresponding to the minimal error, i.e., Estimate W such that:

$$V - \widehat{V} = 0 \quad (6)$$

From Eqs. (4) and (6), the optimal output weight can be estimated as:

$$W = VH^{\dagger} \quad (7)$$

The learning algorithm of ELM can be summarized as:

1 Given a training data set $\{(\mathbf{u}^1, \mathbf{v}^1), \cdots, (\mathbf{u}^t, \mathbf{v}^t), \cdots, (\mathbf{u}^L, \mathbf{v}^L)\}$, choose the number of hidden neurons (K)according to the incremental-decremental procedure discussed in [74].

2 For each neuron, choose the Gaussian centers $(\boldsymbol{\mu}_k;\ k = 1, \cdots, K)$ and their width $(\sigma_k;\ k = 1, \cdots, K)$ randomly.

3 Calculate the hidden layer output matrix (H) according to Eq. (5).

4 Estimate the optimal output weights (W) using Eq. (7).

It must be noted here that the Gaussian centers of the hidden neurons $(\boldsymbol{\mu}_k)$ and their width (σ_k) are chosen randomly. As the performance of the ELM depends on this random initialization, we present an Ens-ELM to improve the prediction accuracy in Section 3.2.

3.2. Ensemble of Extreme Learning Machine

The generalized performance of ELM depends on the initialization of the centers $(\boldsymbol{\mu})$ and width (σ). Therefore, by training multiple ELM initialized in different regions of the input space, one can effectively improve the overall prediction performance.

The block diagram of the proposed ensemble of ELM is presented in Fig. 1. We first train M ELM, namely, $\{ELM[1], ELM[2], \cdots, ELM[M]\}$ with all the samples in the training data set. The Gaussian center and width parameters of the neurons in the hidden layer for each ELM is initialized in different regions of the input space. The ELM that presents the minimum mean square error for each sample is identified. The networks those present a high error for all the samples are discarded from the ensemble. The output of each ELM in the ensemble are then combined using the $min(MSE)$ function, which computes the minimum mean square error for each sample.

In the following section, we present the performance study of the wave height prediction of the Ens-ELM.

4. Performance Study

In this section, we present the performance study results for wave height prediction in the 10 stations listed in Table 1 for the period between Jan 1, 2014 and Aug 31, 2015, located at the three independent regions, viz., Gulf of Mexico, Brazil and Korean region. The total number of samples used in training and testing in each region is tabulated in Table 3.

We perform two performance studies: First, we compare the performances of Ens-ELM with state-of-the-art regression models in Section 4.2. The regression models used in comparison are Support Vector Regression (SVR), Extreme Learning Machines (ELM) and Online Sequential Extreme Learning Machines (OS-ELM). The number of hidden neurons and supprt vectors for all these networks are selected using a constructive-destructive procedure similar to that presented in [74]. The upper bound on the number of hidden neurons is the

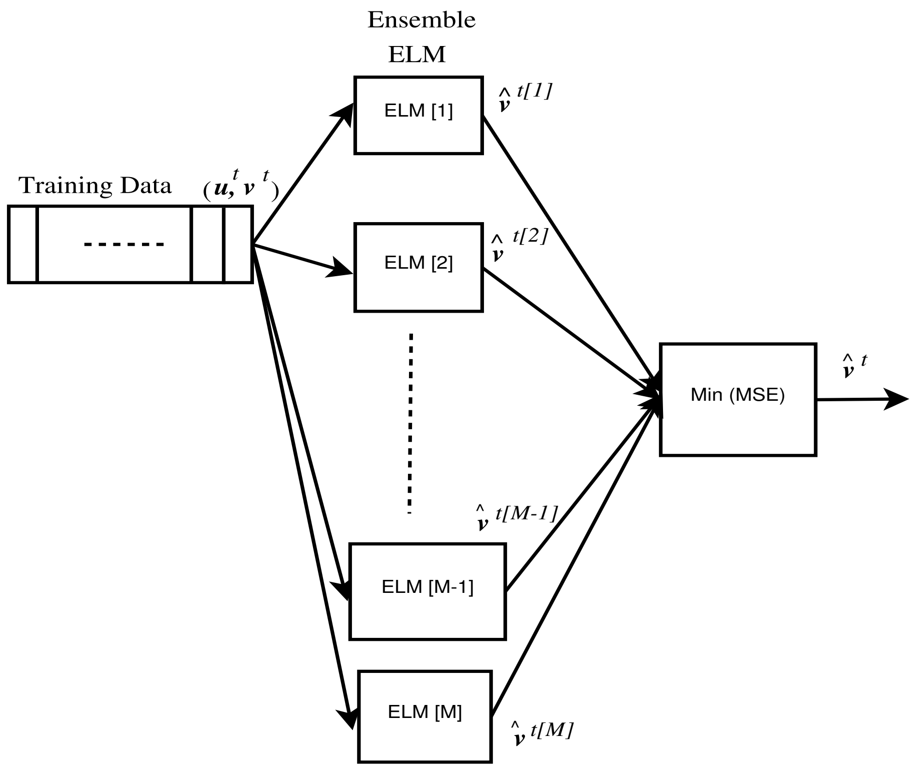

Figure 1: Block Diagram of the Ens-ELM

number of distinct training samples, that is $N \leq K$, where N is the total number of training samples. The number of ELM in the Ens-ELM is chosen based on the distribution of samples. We first train a pool of 50 ELM ($M = 50$) with data collected from each region. Of these, the ELM that consistently performed poorly for all the samples are deleted from the pool. Several combinations of type of SVR, and the kernel for SVR were studied and we chose the best performing epsilon-SVR with an RBF kernel. The optimal values of epsilon and gamma of SVR are chosen as 0.2 and 0.125 through a grid-search technique. Secondly, we compare the wave height prediction performance of the Ens-ELM with the numerical models in the literature listed in Section 4.1. In all these studies, the Root Mean Square Error (RMSE) and the Correlation Coefficient (CC) are used as the performance measures.

The RMSE represents the accuracy of prediction and is defined by

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\mathbf{v}_i - \widehat{\mathbf{v}}_i)^2} \tag{8}$$

$$\tag{9}$$

Table 3: Training and Testing Sample Size

Region	No. of Samples	
	Training	Testing
Gulf of Mexico	4747	465
Brazil	638	66
Korean Region	3244	866

The CC measures the strength and the linear dependency of two random variables. A CC greater than 0.8 is generally described as *strong*, whereas a correlation less than 0.5 is described as *weak*. It is defined as:

$$CC = \frac{Cov(\mathbf{v}, \widehat{\mathbf{v}})}{\sqrt{Var(\mathbf{v})\, Var(\widehat{\mathbf{v}})}} \tag{10}$$

where $Cov(,)$ and $Var()$ refers to the covariance and variance operator, respectively.

4.1. Performance study with state-of-the-art regression models

The number of hidden neurons (K), the Root Mean Square Error (RMSE) and the correlation coefficients (CC) of OS-ELM, Ens-ELM, ELM and SVR are presented in Table 4, where ^{a} indicates support vector in SVR.

From the table 4, it can be seen that Ens-ELM and OS-ELM predict the daily wave height accurately compared with ELM and SVR. Of these two algorithms, the Ens-ELM predicts the daily wave heights and peak wave height more accurately than the OS-ELM. Thus, it can be inferred that the Ens-ELM performs better than the OS-ELM, individual ELM and SVR in predicting the daily wave heights in individual regions. Further, using the Ens-ELM for daily wave height prediction avoids the tedious numerical integration and approximation approach of the existing numerical wave models and is also uniformly applicable across different oceans and seas.

4.2. Comparison with existing works in literature

In this section, we compare the results of wave height prediction using Ens-ELM and OS-ELM, with existing results in the literature for the stations considered in our study. Earlier, the wave height prediction performances of the second generation numerical model WISWAVE, and the third generation models, namely WAM and WAVEWATCH III, in stations 42001 and 42035 for October 1995 has been reported in [10]. In addition to these, the conventional MLP with BP algorithm (MLP-BP) was used to predict wave heights at station 42035 for Feb 2004 in [23]. As Ens-ELM and OS-ELM outperform the SVR and ELM in the wave height prediction as shown in Section 4.1, we compare these results only with Ens-ELM and OS-ELM. The results for WISWAVE, WAM and WAVEWATCH III are reproduced from [10] and that of MLP-BP is reproduced from [23].

Table 4: Performance Study Results

Region	Station	Algorithm	K	RMSE	CC
Gulf of Mexico	42001	SVR	283[a]	0.5721	0.8283
		ELM	700	0.4742	0.8018
		OS-ELM	**25**	**0.2307**	**0.9172**
		Ensemble ELM	**55**	**0.1730**	**0.9591**
	42002	SVR	283[a]	0.4936	0.8646
		ELM	700	0.2243	0.8878
		OS-ELM	**25**	**0.3092**	**0.9057**
		Ensemble ELM	**55**	**0.2203**	**0.9501**
	42019	SVR	283[a]	0.6858	0.8063
		ELM	700	0.2977	0.8481
		OS-ELM	**25**	**0.3961**	**0.8209**
		Ensemble ELM	**55**	**0.3293**	**0.8660**
	42035	SVR	283[a]	0.7247	0.7914
		ELM	700	0.1878	0.9363
		OS-ELM	**25**	**0.3886**	**0.8296**
		Ensemble ELM	**55**	**0.2668**	**0.8921**
Brazil Region	31053	SVR	571[a]	0.5512	0.9184
		ELM	150	0.3768	0.8794
		OS-ELM	**15**	**0.3941**	**0.8795**
		Ensemble ELM	**15**	**0.3383**	**0.9207**
	31262	SVR	571[a]	0.6775	0.8454
		ELM	150	0.7820	0.7450
		OS-ELM	**12**	**0.6810**	**0.8352**
		Ensemble ELM	**15**	**0.0093**	**0.9956**
Korean Region	22103	SVR	195[a]	0.5207	0.5865
		ELM	150	0.3843	0.7867
		OS-ELM	**25**	**0.3793**	**0.8615**
		Ensemble ELM	**30**	**0.2277**	**0.9462**
	22105	SVR	195[a]	0.7037	0.6624
		ELM	150	0.6586	0.8157
		OS-ELM	**25**	**0.6218**	**0.8353**
		Ensemble ELM	**30**	**0.3808**	**0.9193**
	21229	SVR	195[a]	0.6194	0.6511
		ELM	150	0.5199	0.8426
		OS-ELM	**25**	**0.3774**	**0.8600**
		Ensemble ELM	**30**	**0.3466**	**0.9342**
	22108	SVR	195[a]	0.4082	0.7298
		ELM	150	0.3616	0.8629
		OS-ELM	**25**	**0.3418**	**0.8872**
		Ensemble ELM	**30**	**0.2117**	**0.9588**

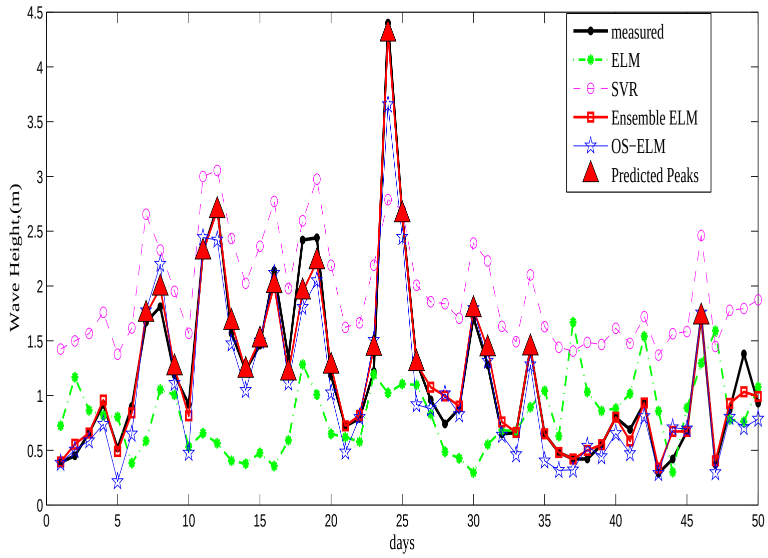

Figure 2: Predicted Wave Height for the period between Jan 1, 2015 and Aug 30, 2015 at station 42002, Gulf of Mexico

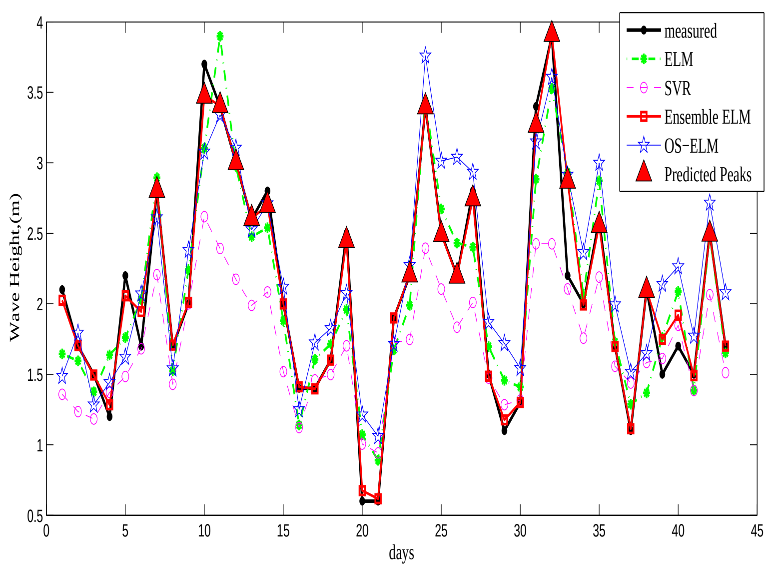

Figure 3: Predicted Wave Height for the period between Nov 1 and Dec 13, 2011 at station 31053, Brazil Waters

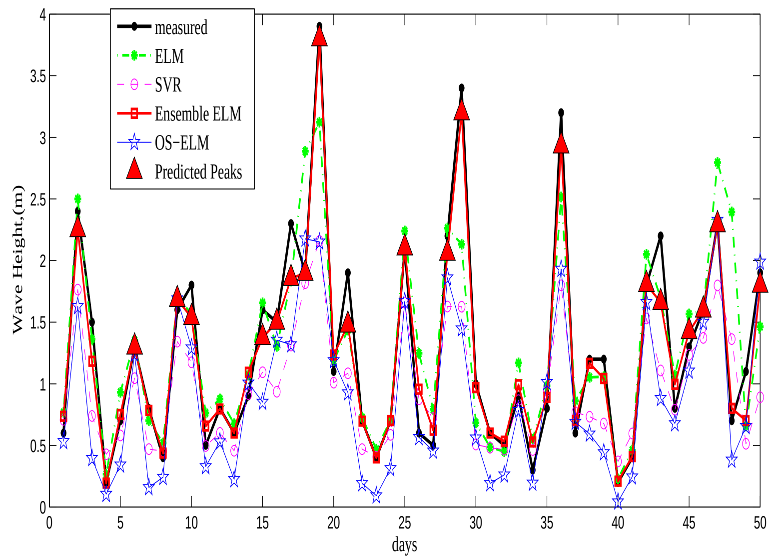

Figure 4: Predicted Wave Height for the period between Jan 1, 2015 and Aug 30, 2015 at station 22108, Korean Region

The wave height prediction performances of the various models are compared using the performance measures namely RMSE and CC. The RMSE and CC of OS-ELM, Ens-ELM in comparison with those of the WISWAVE, WAM and WAVEWATCH III for the stations 42001 and 42035 in October 1995, is shown in Table 5. From the table, we can observe that the performances of OS-ELM and Ens-ELM are almost similar to those of WISWAVE, WAM and WAVEWATCH III in predicting wave heights at station 42001 in October 1995. Moreover, it can also be observed that at station 42035, the OS-ELM and Ens-ELM outperform the numerical models. It must also be noted that the numerical models involve huge computational effort and requires high processing time [11]. This also limits their ability to be applicable across different oceans and seas.

The wave height prediction performances of OS-ELM and Ens-ELM are compared against that of the MLP-BP in station 42035 for February 2004, and the results are presented in Fig. 4.2. The RMSE of the MLP-BP, OS-ELM and Ens-ELM are **0.4305**, **0.3047** and **0.1881**, respectively. From the figure and RMSE of the various algorithms, it can be seen that OS-ELM and Ens-ELM outperform MLP-BP in the prediction of wave heights.

From the results in this section, we can observe that OS-ELM and Ens-ELM outperform the existing method in literature for the problem of ocean wave height prediction.

Table 5: Comparison of OS-ELM and Ens-ELM with Numerical wave models

Station Id	Model/Algorithm	RMSE	CC
42001	WISWAVE	0.28	0.91
	WAM	0.26	0.91
	WAVEWATCH III	0.28	0.90
	OS-ELM	**0.2541**	**0.9202**
	Ensemble ELM	**0.2309**	**0.9495**
42035	WISWAVE	0.21	0.87
	WAM	0.21	0.86
	WAVEWATCH III	0.24	0.83
	OS-ELM	**0.2344**	**0.8903**
	Ensemble ELM	**0.1728**	**0.9307**

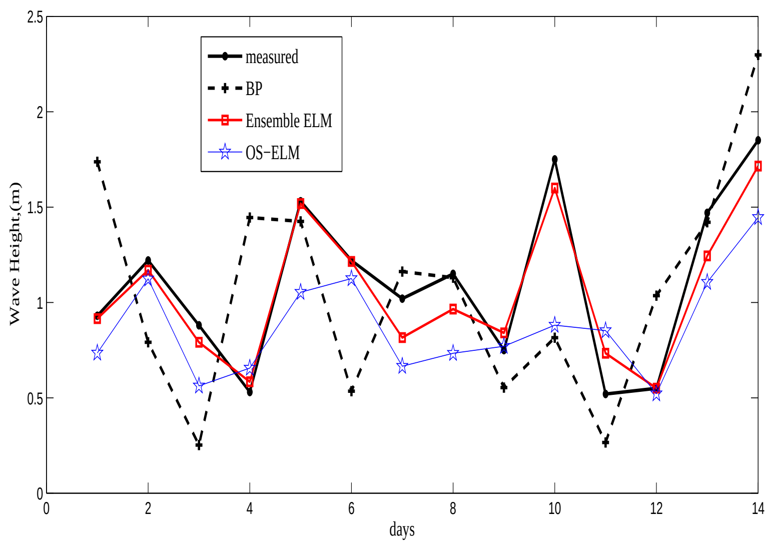

Figure 5: Predicted Wave Height for Feb 2004 at station 42035, Gulf of Mexico

5. Conclusion

In this paper, we develop an ensemble of Extreme Learning Machine (Ens-ELM) to predict the daily wave conditions at 10 different stations from the Gulf of Mexico, Brazil and Korean region. We construct an Ens-ELM, each with input parameters initialized at different regions of the input space. For each sample in the training data set, the ELM that produces the least mean square error is identified, and the output of that ELM is considered as the output for that sample. Thus, the randomness of the initialization in ELM is exploited to achieve superior generalization performance. The network is trained using the past wave data and the measured atmospheric conditions obtained from the selected stations between Jan 1, 2011 and Dec 31, 2014 and is tested with data in these stations between Jan 1, 2015 and Aug 30, 2015. The performance of the Ens-ELM in predicting the daily wave height in these stations is studied in comparison with those of SVR, ELM and OS-ELM. From this study, we infer that the Ens-ELM outperforms OS-ELM, ELM and SVR in the daily wave height prediction. The results of Ens-ELM also outperforms results in [23] and numerical second, third generation wave models. The future work scope involves in prediction of wave height at any time of the day and also in developing a network that is location independent and universal for any sea states. This work shall be extended to other major offshore activity regions across the world to enable support in off-shore operations and marine control applications.

6. Acknowledgements

The authors would like to thank the National Data Buoy Center, European Marine Observation and Data Network, Pacific Marine Environmental Laboratory and National Hurricane Center for providing the measured atmospheric and wave data used in this work.

References

[1] R. Hashim, C. Roy, S. Motamedi, S. Shamshirband, D. Petkovic, 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.

[2] WMO, Guide to Wave Analysis and Forecasting, World Meterological organization,WMO-No.702, 1998.

[3] O.M.Philips, On the generation of waves by a turbulent wind, Journal of Fluid Mechanics 2 (5) (1957) 417–445.

[4] J.W.Miles, On the generation of surfaces waves by shear flows, Journal of Fluid Mechanics 3 (2) (1957) 185–204.

[5] K.Hasselmann, T.P.Barnett, E.Bouws, H.Carlson, D.E.Cartwright, K.Enke, J.A.Ewing, H.Gienapp, D.E.Hasselmann, P.Kruseman, A.Meerburg, P.Mller, D.J.Olbers, K.Ritcher, W.Sell, H.Walden, Measurements of wind-wave growth and swell decay during the joint north sea wave project (jonswap), Ergnzungsheft zur Deutschen and Hydrographischen Zeitschrigt Reihe 8 (12) (1973) 95.

[6] J. Thomas, G. Dwarakish, Numerical wave modelling - a review, Proceedings of Internation Conference on Water Resources, Coastal and Ocean Engineering (ICWRCOE'15), Aquatic Procedia 4 (2015) 443–448.

[7] S.Mandal, N.Prabaharan, Ocean wave prediction using numerical and neural network models, The Open Ocean Engineering Journal 3 (7) (2010) 12–17.

[8] W. Group, The wam model- a third generation ocean wave prediction model, Journal of Physical Oceanography 18 (12) (1988) 1775–1810.

[9] N. Booij, R. Ris, L. Holthuijsen, A third generation wave model for coastal regions. 1. model description and validation, Journal of Geophysical Research: Oceans 104 (C4) (1999) 7649–7666.

[10] B. Tracy, A. Cialone, Comparison of gulf of mexico wave information studies (WIS) 2-G hindcast with 3-G hindcast, 8th International Workshop on Wave Hindcasting and Forecasting/Coastal Hazards Symposium, North Shore, Oahu, Hawaii, USA (2004) 1–15.

[11] M. Browne, B. Castelle, D. Strauss, R. Tomlinson, M. Blumenstein, Nearshore swell estimation from a global wind-wave model: Spectral process, linear, and artificial neural network models, Coastal Engineering 54 (5) (2007) 445–460.

[12] P. Galiatsatou, P. Prinos, Modeling non-stationary extreme waves using a point process approach and wavelets, Stochastic Environmental Research and Risk Assessment 25 (2) (2010) 165–183.

[13] M. Hoda, R. Krishnamurti, Arabic calligraphy: A computational exploration, Mathematics and Design'01, Deakin University, Geelong, Australia.

[14] A. Osborne, Nonlinear Ocean Waes and the Inverse Scattering Transform, Academic Press, 2010.

[15] M. Zijlema, Computation of wind-wave spectra in coastal waters with swan on unstructured grids, Coastal Engineering 57 (3) (2010) 267–277.

[16] S.Rao, S.Mandal, Hindcasting of storm waves using neural networks, Ocean Engineering 32 (5-6) (2005) 667–684.

[17] P.Jain, M.C.Deo, Artificial intelligence tools to forecast ocean waves in real time, The Open Ocean Engineering Journal 1 (1) (2008) 13–20.

[18] H.L.Tolman, V. Krasnopolsky, D. Chalikov, Neural network approximations for nonlinear interactions in wind wave spectra:direct mapping for wind seas in deep water, Ocean Modelling 8 (3) (2005) 253–278.

[19] S. Rao, S.Mandal, N.Prabaharan, Wave forecasting in near real time using neural networks, International Conference on Ocean Engineering,IIT Madras 1 (1) (2001) 103–108.

[20] J.D.Agarwal, M.C.Deo, On-line wave prediction, Marine Structures 15 (1) (2002) 57–54.

[21] M.C.Deo, A.Jha, A. Chaphekar, K.Ravikant, Neural networks for wave forecasting, Ocean Engineering 28 (7) (2001) 889–898.

[22] R. Kalra, M.C.Deo, R. Kumar, V. K. Agarwal, Rbf network for spatial mapping of wave heights, Marine Structures 18 (3) (2005) 289–300.

[23] S.N.Londhe, V. Panchang, One-day wave forecasts based on artificial neural networks, Journal of Atmospheric and Oceanic Technology 23 (11) (2006) 1593–1603.

[24] I.Malekmohamadi, R.Ghiassi, M.J.Yazdanpanah, Wave hindcasting by coupling numerical model and artificial neural networks, Ocean Engineering 35 (3-4) (2008) 417–425.

[25] O.Makarynskyy, Improving wave predictions with artificial neural networks, Ocean Engineering 31 (5-6) (2004) 709–724.

[26] J. Pooja, M. Deo, G. Latha, V. Rajendran, Real time wave forecasting using wind time history and numerical model, Ocean Modelling 36 (1-2) (2011) 26–39.

[27] C. Tsai, C. Lin, J. Shen, Neural network for wave forecasting among multistations, Ocean Engineering 29 (13) (2002) 1683–1695.

[28] M. Kazeminezhad, A. Shahidi, S. Mousavi, A neuro-fuzzy based approach to wave modeling, In: Proceedings of the 7th international conference on hydro informatics, HIC 2006.

[29] M. Kazeminezhad, A. Shahidi, S. Mousavi, Application of fuzzy inference system in the prediciton of wave parameters, Ocean Engineering 32 (14-15) (2005) 1709–1725.

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

[31] I. Malekmohamadi, M. Bazargan-Lari, R. Kerachian, M. R. Nikoo, M. Fallahnia, Evaluating the efficacy of svms, bns, anns and anfis in wave height prediction, Ocean Engineering 38 (2-3) (2011) 487–497.

[32] M. Scotto, C. G. Soares, Bayesian inference for long-term prediction of significant wave height, Coastal Engineering 54 (5) (2007) 393–400.

[33] A. Abed-Elmdoust, R. Kerachian, Wave height prediction using the rough set theory, Ocean Engineering 54 (2012) 244–250.

[34] A. Zamani, D. Solomatine, A. Azimian, A. Heemink, Learning from data for windwave forecasting, Ocean Engineering 35 (10) (2008) 953–962.

[35] W. Kanbua, S. Supharatid, I. M. Tang, Ocean wave forecasting in the gulf of thailand during typhoon linda 1997 : Wam and neural network approaches, Science Asia 31 (2005) 243–250.

[36] K. Gunaydin, The estimation of monthly mean significant wave heights by using artificial neural network and regression methods, Ocean Engineering 35 (14-15) (2008) 1406–1415.

[37] V. Krasnopolsky, D. Chalikov, H. Tolman, A neural network technique to improve computational efficiency of numerical oceanic models, Ocean Modelling 4 (3-4) (2002) 363–383.

[38] E. Alexandre, L. Cuadra, J. Nieto-Borge, G. Candil-Garcia, M. D. Pino, S. Salcedo-Sanz, A hybrid genetic algorithm- extreme learning machine approach for accurate significant wave height reconstruction, Ocean Modelling 92 (2015) 115–123.

[39] A. Duran-Rosal, C. Hervas-Martinex, A. Tallon-Ballesteros, A. Martinex-Estudillo, S. Salcedo-San, Massive missing data reconstruction in ocean buoys with evolutionary product unit neural networks, Ocean Engineering 117 (2016) 292–301.

[40] S. Salcedo-Sanz, J. Nieto-Borge, E. Alexandre, L. Cuadra, L. Carro-Calvo, K. Hessner, Significant wave height estimation using svr algorithms and shadowing information from simulated and real measured x-band radar images of the sea surface, Ocean Engineering 101 (2015) 244–253.

[41] L. Cornejo-Bueno, J. Nieto-Borge, P. Garcia-Diaz, G. Rodriguez, S. Salcedo-Sanz, Significant wave height and energy flux prediction for marine energy applications: A grouping genetic algorithm- extreme learning machine approach, Renewable Energy 97 (2016) 380–389.

[42] J. Fernandez, S. Salcedo-Sanz, P. Gutierrez, E. Alexandre, C. Hervas-Martinez, Significant wave height and energy flux range forecast with machine learning classifiers, Engineering Applications of Aritifical Intelligence 43 (2015) 44–53.

[43] L. Cornejo-Bueno, A. Aybar-Ruiz, S. Jumenez-Fernandez, E. Alexandre, , J. Neito-Borge, S. Salcedo-Sanz, A grouping genetic algorithm extreme learning machine approach for optimal wave energy prediction, IEEE Congress on Evolutionary Computation (CEC) (2016) 3817–3823.

[44] G. B. Huang, L. Chen, C. K. Siew, Universal approximation using incremental constructive feedforward networks with random hidden nodes, IEEE Transactions on Neural Networks 17 (4) (2006) 879–892.

[45] G. B. Huang, Q. Y. Zhu, C. K. Siew, Real-time learning capability of neural networks, IEEE Transactions on Neural Networks 17 (4) (2003) 863–878.

[46] G. B. Huang, Learning capability and storage capacity of two-hidden layer feedforward networks, IEEE Transactions on Neural Networks 14 (2) (2003) 274–281.

[47] E. B. Baum, On the capabilities of multilayer preceptrons, Journal of Complexity 4 (3) (1988) 193–215.

[48] D. S. Broomhead, D. Lowe, Multivariable functional interpolation and adaptive networks, Complex Systems 2 (1988) 321–355.

[49] B. Igelnik, Y. H. Pao, Stochastic choice of basis functions in adaptive function approximation and the function-link net, IEEE Transactions on Neural Networks 6 (6) (1995) 1320–1329.

[50] D. Lowe, Adaptive radial basis function nonlinearities and the problem of generalisation, Proceedings of first IEE international conference on artificial neural network (1989) 171–175.

[51] S. Ferrari, R. F. Stengel, Smooth function approximation using neural networks, IEEE Transactions on Neural Networks 16 (1) (2005) 24–38.

[52] L. Chen, L. Zhou, H. Pung, Universal approximation and qos violation application of extreme learning machine, Neural Processing Letter 28 (2) (2008) 81–95.

[53] G. B. Huang, L.Chen, Convex incremental extreme learning machine, Neurocomputing 70 (16–18) (2007) 3056–3062.

[54] S. Saraswathi, S. Sundaram, N. Sundararajan, M. Zimmermann, M. Nilsen-Hamilton, ICGA-PSO-ELM approach for accurate multiclass cancer classification resulting in reduced gene sets in which genes encoding secreted proteins are highly represented, IEEE/ACM Transactions on Computational Biology and Bioinformatics 8 (2) (2011) 452–463.

[55] S. Suresh, S. Saraswathi, N. Sundararajan, Performance enhancement of extreme learning machine for multi-category sparse data classification problems, Engineering Applications of Artificial Intelligence 23 (7) (2010) 149–1157.

[56] S. Suresh, R. V. Babu, H. J. Kim, No-reference image quality assessment using modified extreme learning machine classifier, Applied Soft Computing 9 (2) (2009) 541–552.

[57] G. B. Huang, C. K. Siew, Extreme learning machine with randomly assigned RBF kernels, Int. J. Inf. Technol 11 (1).

[58] N.-Y. Liang, G. B. Huang, P. Saratchandran, N. Sundararajan, A fast and accurate on-line sequential learning algorithm for feedforward networks, IEEE Transactions on Neural Networks 17 (6) (2006) 1411–1423.

[59] H. J. Rong, G. B. Huang, N. Sundararajan, P. Saratchandran, Online sequential fuzzy extreme learning machine for function approximation and classification problems, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics 39 (4) (2009) 1067–1072.

[60] J. Zhao, Z. Wang, D. S. Park, Online sequential extreme learning machine with forgetting mechanism, Neurocomputing 87 (2012) 79–89.

[61] S. C. Stevica, B. S. Milos, S. V. Nikolic, Multichannel descriptors and ensemble of extreme learning machines for classification of remote sensing images, Signal Processing: Image Communication (In Press).

[62] S. Huang, B. Wang, J. Qiu, J. Yao, G. Wang, G. Yu, Parallel ensemble of online extreme learning machine based on mapreduce, Neurocomputing (In Press).

[63] L. Yu, W. Dai, L. Tang, A novel decomposition ensemble model with extended extreme learning machine for crude oil price forecasting, Engineering Applications of Artificial Intelligence (In Press).

[64] S. Shukla, R. N. Yadav, Voting based extreme learning machine with accuracy based ensemble pruning, International Journal of Computer Applications 115 (22) (2015) 14–18.

[65] S. Vasily, R. Savitha, S. Suresh, H. J. Kim, H. J. Hwang, A cognitive ensemble of extreme learning machines for steganalysis based on risk-sensitive hinge loss function, Cognitive Computation 7 (1) (2015) 103–110.

[66] Yuan, Ensemble of online sequential extreme learning machine, Neurocomputing 72 (13–15) (2009) 3391–3395.

[67] J. Cao, Z. Lin, G. B. Huang, N. Liu, Voting based extreme learning machine, Information Sciences 185 (1) (2012) 66–77.

[68] G. Huang, G. B. Huang, S. Song, K. You, Trends in extreme learning machines: A review, Neural Networks 61 (2015) 32–48.

[69] H.Charnock, Wind stress on a water surface, Quarterly Journal of the Royal Meteorological Society 81 (350) (1955) 639–640.

[70] L.Cavaleri, P. M. Rizzol, Wind wave prediction in shallow water: theory and applications, Journal of Geophysical Research 86 (C11) (1981) 1096110973.

[71] M. B.Abbott, H. H. Petersen, O. Skovgaard, Computations of short waves in shallow water, Proceedings of 16th Coastal Engineering Conference, Hamburg, Germany 1 (1) (1978) 414–433.

[72] R.L.Snyder, F. W. Dobson, J. A. Elliott, R. B. Long, Array measurements of atmospheric pressure fluctuations above surface gravity waves, Journal of Fluid Mechanics 102 (1) (1981) 1–59.

[73] R. M.Brooks, W. D. Corson, Summary of archived Atlantic coast wave information study pressure, wind,wave and water level data, WIS Report 13,Waterways Experiment Station,Vicksburg, MS, USA, 1984.

[74] Suresh, S. N. Omkar, V. Mani, T. N. G. Prakash, Lift coefficient prediction at high angle of attack using recurrent neural network, Aerospace Science and Technology 7 (8) (2003) 595–602.