Results in Geophysical Sciences 8 (2021) 100032

Contents lists available at ScienceDirect

Results in Geophysical Sciences

journal homepage: www.sciencedirect.com/journal/results-in-geophysical-sciences

Basin scale wind-wave prediction using empirical orthogonal function analysis and neural network models

Mrinmoyee Bhattacharya [a], Mourani Sinha [b,*]

[a] Department of Computer Science and Engineering, Techno India University, Saltlake, Kolkata, West Bengal 700091, India
[b] Department of Mathematics, Techno India University, Saltlake, Kolkata, West Bengal 700091, India

ARTICLE INFO

Keywords:
Empirical orthogonal function analysis
Neural network models
Backpropagation algorithms
Significant wave height
Bay of Bengal

ABSTRACT

A new method is discussed using neural network models in combination with empirical orthogonal function (EOF) analysis for the basin-scale wind-wave forecast. For the Bay of Bengal region EOF analysis has been performed separately on the significant wave height (SWH) data, zonal (U) and meridional (V) components of wind data. For basin scale forecast the dominant principal component (PC) has been subjected to univariate and multivariate neural network models for future predictions. In the univariate approach, only past values of SWH time series are used and in the multivariate approach, U and V time series are used to predict future SWH values. Efficiency in terms of accuracy and speed of four different backpropagation algorithms, namely, Levenberg-Marquardt (LM), Bayesian Regularization (BR), Scaled Conjugate Gradient (SCG) and Fletcher Conjugate Gradient (CGF) have been compared for 1 to 12 multistep ahead time steps and 1 to 13 neurons. After training the models using varied neurons and the PCs, representing the entire basin, the neurons are fixed at which minimum errors are obtained. Further experiments are conducted using the fixed neurons and the PCs for 1 to 12 time steps ahead SWH prediction. Finally independent datasets consisting of normal and cyclonic wind-wave parameters are tested successfully using the above fixed neurons for delays (1 to 12) corresponding to 3 days or 72 h forecast. The novelty of the study lies is the usage of the PCs which represent the entire basin rather than computations at individual locations which are expensive technically and time consuming.

1. Introduction

Accurate extreme and operational wave climate data are essential for various applications of coastal and ocean engineering purposes such as coastal and marine planning, designing, management, and operations. Extreme weather events like the tropical cyclones associated with strong winds, storm surges and heavy rainfall cause destructions and hazards. The Indian coasts are significantly vulnerable to this damaging natural catastrophes causing loss of life and property. Most widely used ocean wave models such as WAM (WAMDIG 1988), WAVEWATCH III (Tolman 1991), and SWAN (Booij et al., 1999; Ris et al., 1999) are forced with surface winds generated by numerical weather prediction models to produce significant wave height forecasts. However, the resolution of the global weather models is often insufficient to capture the steep wind gradients associated with the tropical cyclones. In this study non-linear autoregressive neural network models have been developed for predictions during normal and extreme conditions. The models have been trained with a large data set to achieve the required accuracy and then used for future predictions.

Lecacheux et al. (2012) outline a new approach to perform an analysis of extremes within a complex environment presenting several sources of extreme waves. Several extreme wave scenarios are determined for each regime, based on real historical cases (for cyclonic waves) and extreme value analysis (for non-cyclonic waves). Sampson et al. (2016) described an algorithm to produce significant wave height ensemble. He used forecast ensemble members to produce a probability algorithm related to Monte Carlo wind speed. Chaudhuri et al. (2015) developed multilayer feed-forward neural nets for the North Indian Ocean with different architectures for forecasting the track and intensity of tropical cyclones. For the North Indian Ocean they concluded that the track and intensity of tropical cyclones can be well assessed using neural network models at least 24 h in advance with accuracy levels comparable with existing numerical model output. Richman et al. (2017) evaluated the predictive skill of a machine learning model in order to perform tropical cyclone counts for the North Atlantic region. The machine learning method, which is based on a careful selection of

* Corresponding author.
E-mail address: mou510@gmail.com (M. Sinha).

https://doi.org/10.1016/j.ringps.2021.100032
Received 23 July 2021; Received in revised form 4 November 2021; Accepted 8 November 2021
Available online 14 November 2021

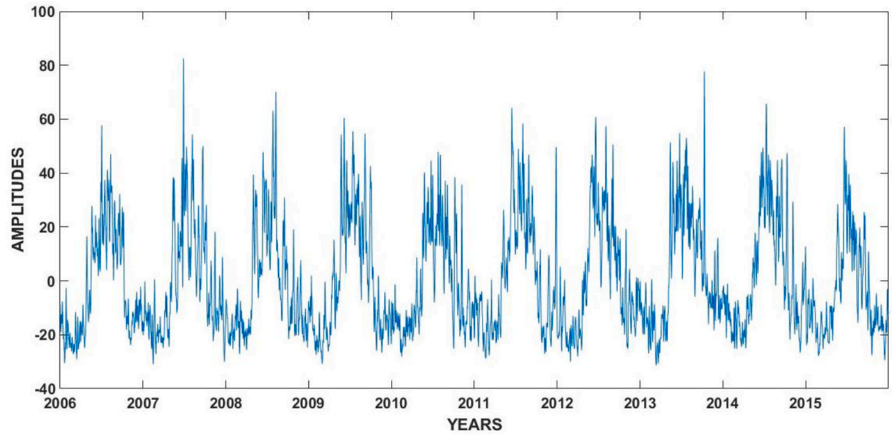

Fig. 1a. The first PC of BOB significant wave height from 2006 to 2015.

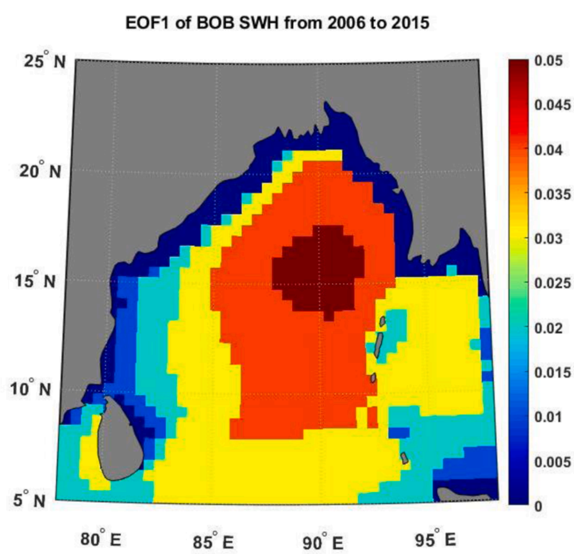

Fig. 1b. The first EOF of BOB significant wave height from 2006 to 2015.

predictors is a support vector regression model, utilizing two different kernels. Pradhan et al. (2018) proposed a deep convolutional neural network model for estimating the hurricane intensity by learning features and using the graphics processing unit. Uncertainties present in tropical cyclone wind fields are demonstrated by Loridan et al. (2017) using machine learning methods. Tropical cyclone wind fields are highly underestimated by the model and thus utility of the same during emergency fails. Several studies show the numerical model forecast parameters lack certain accuracy during extreme conditions. Thus predictive models can be used to estimate ocean parameters more

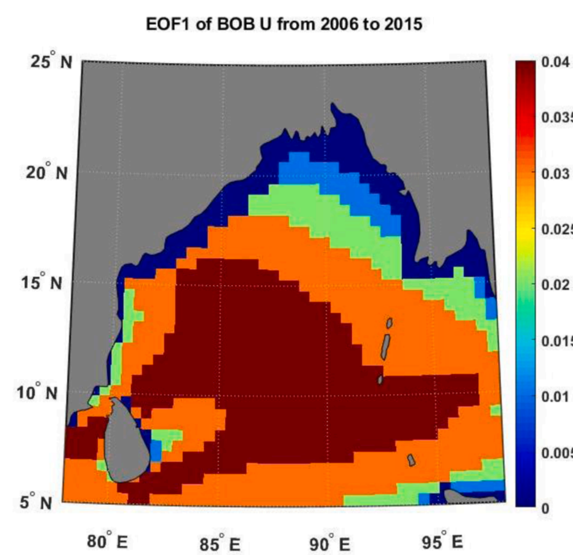

Fig. 2b. The first EOF of BOB zonal (U) wind component from 2006 to 2015.

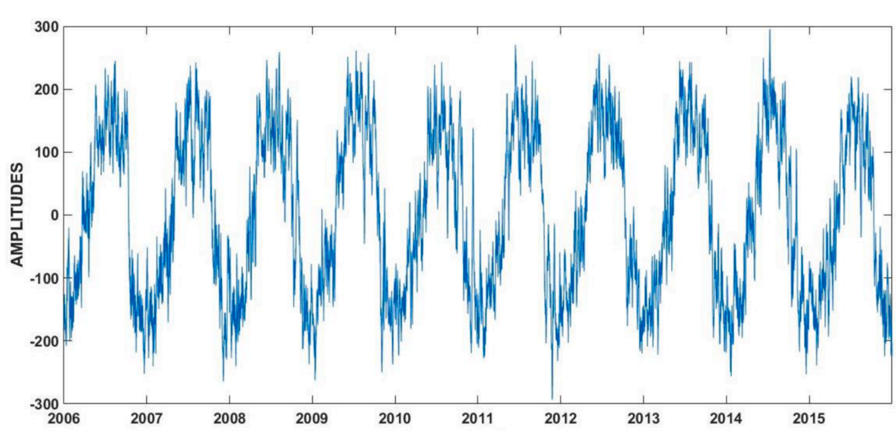

Fig. 2a. The first PC of BOB zonal (U) wind component from 2006 to 2015.

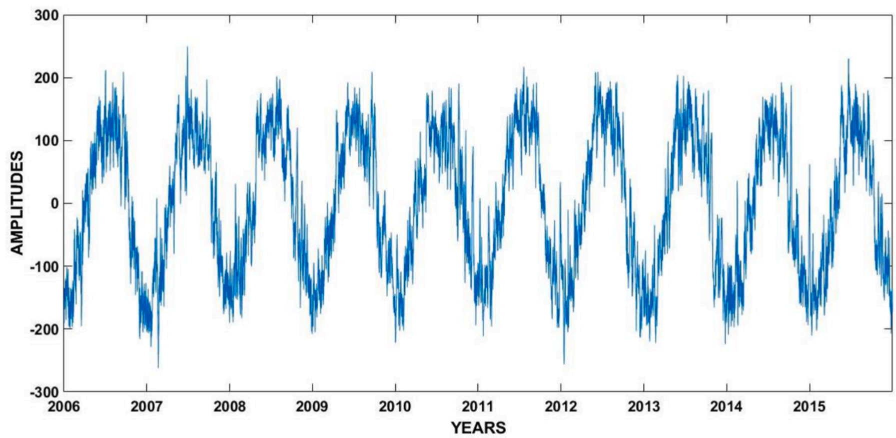

Fig. 3a. The first PC of BOB meridional (V) wind component from 2006 to 2015.

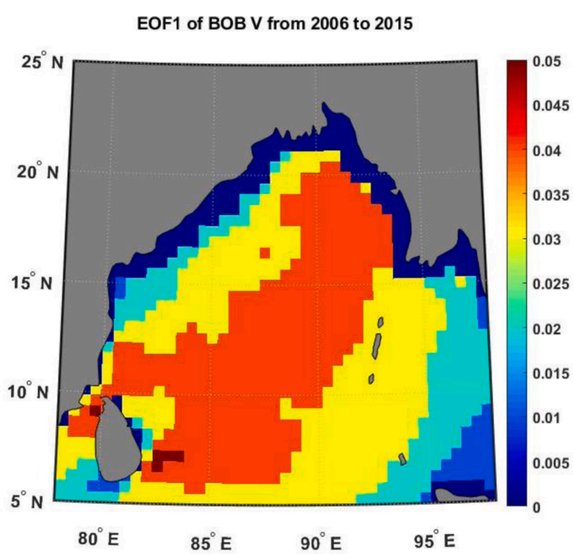

Fig. 3b. The first EOF of BOB meridional (V) wind component from 2006 to 2015.

efficiently. Numerical ocean wave models such as WAM, WAVEWATCH III, and SWAN use bottom topography and surface winds as inputs. Thus the model outputs like the SWH data are largely dependent on the wind vectors. In this paper wind and wave parameters have been obtained from the WAVEWATCH III model which is a globally acknowledged model used operationally worldwide.

Neural networks are built imitating the human brain and the inherent non-linear parallel organization existing are used for accurate future predictions. The artificial neural network (ANN) system can exhibit various tasks like the optimization of data and the function approximation. Wind-wave prediction having wide applications in meteorology and oceanography can be modelled using a certain set of past observed data. The above methodology has the power to change and adjust itself over time under varied conditions, to give better forecasts. An ANN non-linear system is made of three layers, given as, the input, the hidden and the output layers. It is a well-known statistical method that can establish a relationship between different variables. In this method, a backpropagation (BP) training algorithm trains a feedforward network to obtain a minimum error. Recent advances in technology permit the use of the feedforward neural network approach to predict the responses of marine systems with reasonable accuracy (Tiwari et al., 2013). Guo et al. (2011) forecasted wind energy using ANN after removing seasonal effects from the datasets.

Lodge and Yu (2014) developed an ANN model to successfully predict real-time wind speed parameter. Ghanbarzadeh et al. (2009) used ANN models to forecast wind speed from temperature, relative humidity and vapor pressure. Saroha and Agarwal (2014) forecasted wind power using different classes of ANN. They compared linear ANN with time delay, feedforward neural network and Elman recurrent ANN and concluded that all of them exhibited similar results. Models using multi-step ahead forecasting were compared and reviewed with time series data (Taieb et al., 2012). They compared different strategies for

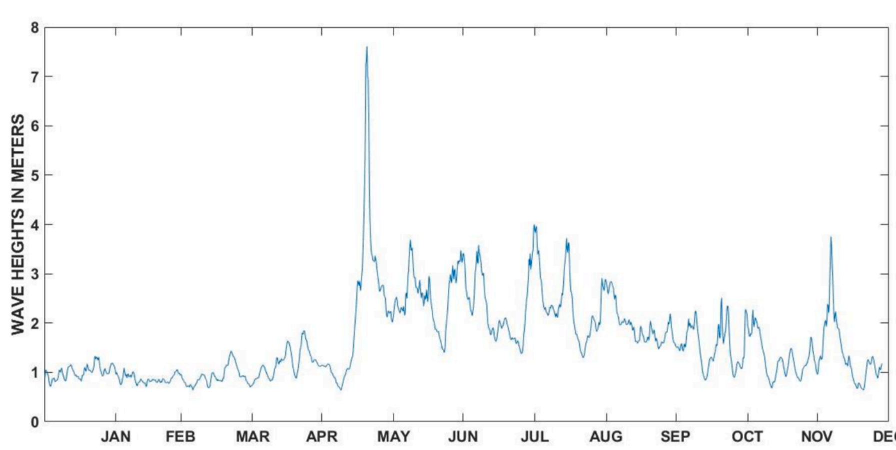

Fig. 4. The time series of 2016 BOB significant wave height (normal grid 15ºN and 90ºE).

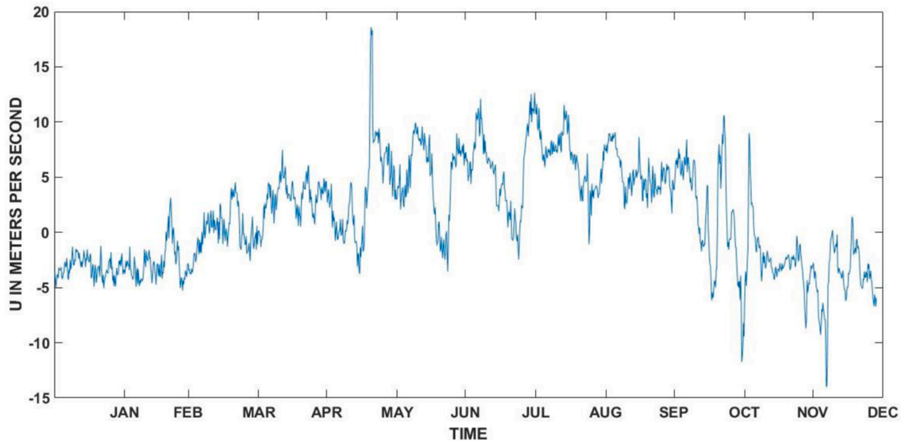

Fig. 5. The time series of 2016 BOB zonal (U) wind component (normal grid 15°N and 90°E).

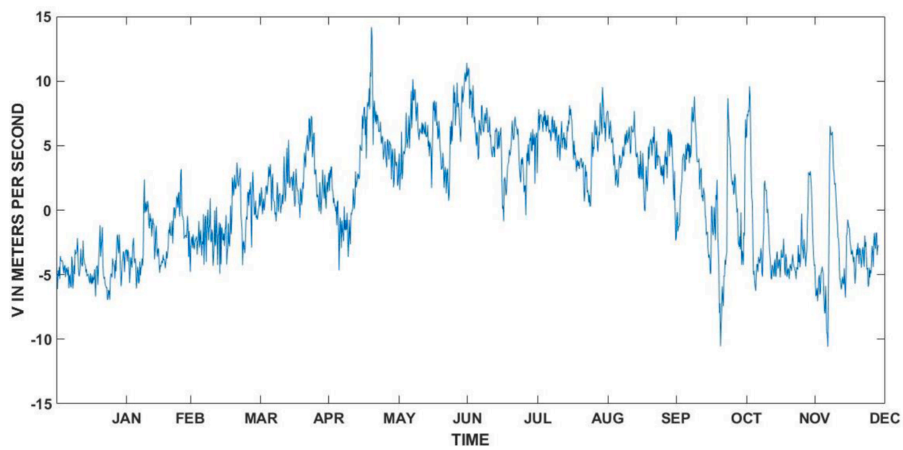

Fig. 6. The time series of 2016 BOB meridional (V) wind component (normal grid 15°N and 90°E).

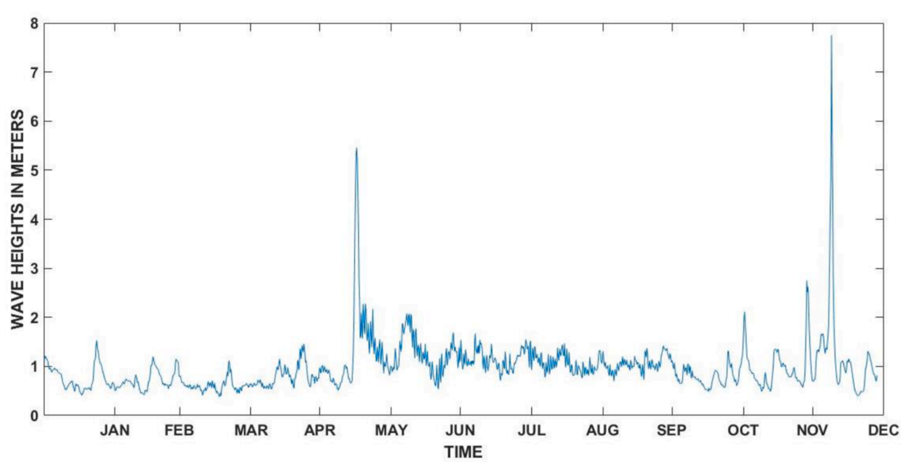

Fig. 7. The time series of 2016 BOB significant wave height (VARDAH grid 13°N and 81°E).

few time-steps ahead forecasting and inferred deseasonalization, selection of input variable and combined strategies lead to better results. Wilamowski (2009) described and compared different neural network architectures and learning algorithms. Pan et al. (2013) compared different BP algorithms for electricity load forecasting. Hagan and Menhaj et al. (1994) found Marquardt algorithm is better than conjugate gradient and learning rate algorithms when few hundred weights are only present. Moller (1993) introduced the scaled conjugate gradient algorithm and showed it to be faster and more effective than the existing standard algorithms like the backpropagation algorithm, conjugate gradient algorithm and Fletcher conjugate gradient algorithm. Li and Shi (2010) compared different ANN models for the purpose of wind speed prediction and discussed the importance of selection of the model for better performance. In another study Parmer et al. (2017) compared

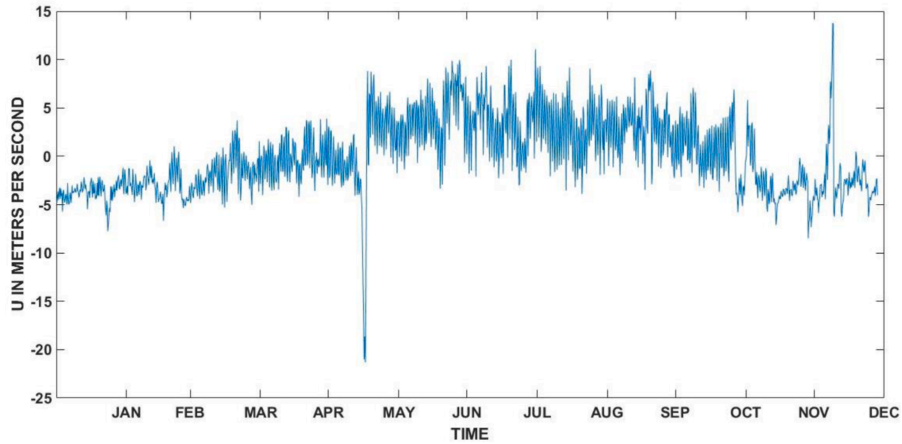

Fig. 8. The time series of 2016 BOB zonal (U) wind component (VARDAH grid 13ºN and 81ºE).

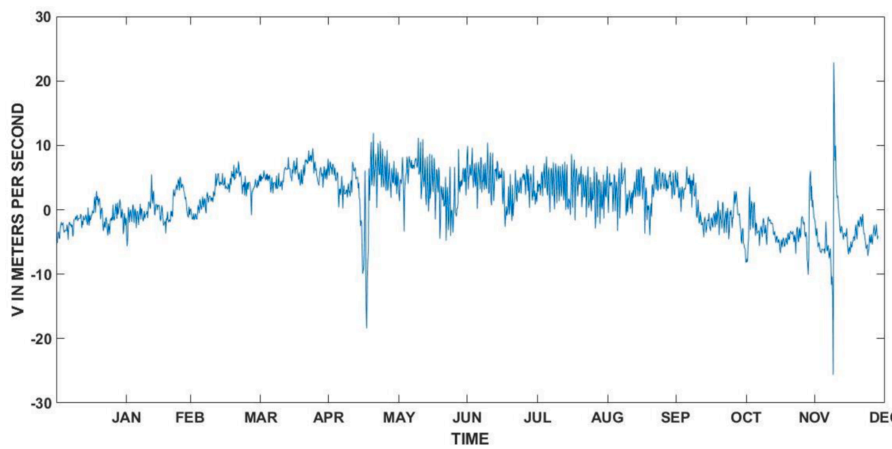

Fig. 9. The time series of 2016 BOB meridional (V) wind component (VARDAH grid 13ºN and 81ºE).

Table 1
Performance results in terms of MSE using first PCs of SWH and U (multivariate case).

NEURONS	LM(MSE)	BR(MSE)	SCG(MSE)	CGF(MSE)
1	0.290	0.356	0.352	0.355
2	0.300	0.286	0.660	0.413
3	0.268	0.214	0.300	0.410
4	0.270	0.246	0.436	0.310
5	0.273	0.243	0.464	0.608
6	0.180	0.127	0.429	0.734
7	0.206	0.156	0.397	0.297
8.	0.156	0.137	0.504	0.332
9.	0.129	0.179	0.521	0.282
10.	0.236	0.191	0.420	0.349
11.	0.0951	0.155	0.540	0.488
12.	0.174	0.225	0.581	0.439
13.	0.112	0.243	0.336	0.499

Table 2
Performance results in terms of MSE using first PCs of SWH and V (multivariate case).

NEURONS	LM(MSE)	BR(MSE)	SCG(MSE)	CGF(MSE)
1	0.391	0.412	0.411	0.344
2	0.431	0.350	0.510	0.413
3	0.329	0.282	0.327	0.434
4	0.322	0.280	0.520	0.455
5	0.253	0.296	0.323	0.403
6	0.287	0.254	0.392	0.705
7	0.518	0.263	0.364	0.340
8.	0.246	0.255	0.356	0.829
9.	0.276	0.223	0.788	0.644
10.	0.301	0.282	0.877	0.511
11.	0.225	0.279	0.627	0.594
12.	0.211	0.296	0.489	0.442
13.	0.287	0.258	0.507	0.563

Levenberg-Marquardt, Scaled Conjugate Gradient and Bayesian Regularization algorithms to forecast wind speed 48 h ahead. In this paper four types of training algorithms, namely, Levenberg-Marquardt (LM), Bayesian Regularization (BR), Scaled Conjugate Gradient (SCG), and Fletcher Conjugate Gradient (CGF) have been used for 72 h ahead forecast of wave height parameter. The evaluation of the algorithms have been done based on mean square error (MSE) in this study. For this purpose, MATLAB is chosen as an experimental environment to perform the required computations and visualizations.

2. EOF analysis, ANN models and BP algorithms

There are various studies relating ANN models and EOF analysis which was introduced by Lorenz (1963) in atmospheric science. Baldi and Hornik (1989) used EOF analysis and ANN models to describe the error function. Tangang et al. (1998) used extended EOFs as inputs to the ANN models which primarily reduced the size of the datasets. For forecasting purposes neural networks using backpropagation algorithms are most widely used. The network is static (feedforward) where the output depends on the input only without any delay or feedback

Table 3
Performance results in terms of MSE using first PC of SWH (univariate case).

NEURONS	LM(MSE)	BR(MSE)	SCG(MSE)	CGF(MSE)
1	0.617	0.622	0.636	0.698
2	0.609	0.630	0.696	0.685
3	0.630	0.611	0.661	0.719
4	0.619	0.624	0.666	0.696
5	0.637	0.626	0.716	0.727
6	0.618	0.620	0.626	0.688
7	0.608	0.611	0.761	0.711
8.	0.611	0.608	0.694	0.680
9.	0.613	0.622	0.657	0.751
10.	0.619	0.611	0.689	0.749
11.	0.626	0.609	0.696	0.918
12.	0.620	0.619	0.688	0.725
13.	0.614	0.617	0.822	0.712

elements. For the dynamic networks which are faster, feedback from other layers are included to the input layer. In this work, forecast techniques using EOF analysis, ANN model and BP algorithms have been developed. Although the method is computationally cheaper, for training the algorithm it has to be dependent on reliable datasets.

2.1. EOF analysis

Given any space-time field, EOF analysis finds a set of orthogonal spatial patterns along with a set of associated uncorrelated time series or principal components (PCs). This technique can compress data by efficiently reducing its dimensionality. The spatial patterns in the data and the temporal variances are identified as the EOF analysis decomposes the given data. EOF techniques were discussed long back by Hotelling (1933) who introduced principal component analysis (PCA), another name for EOFs. A review of PCA or EOFs can be found in Kutzbach

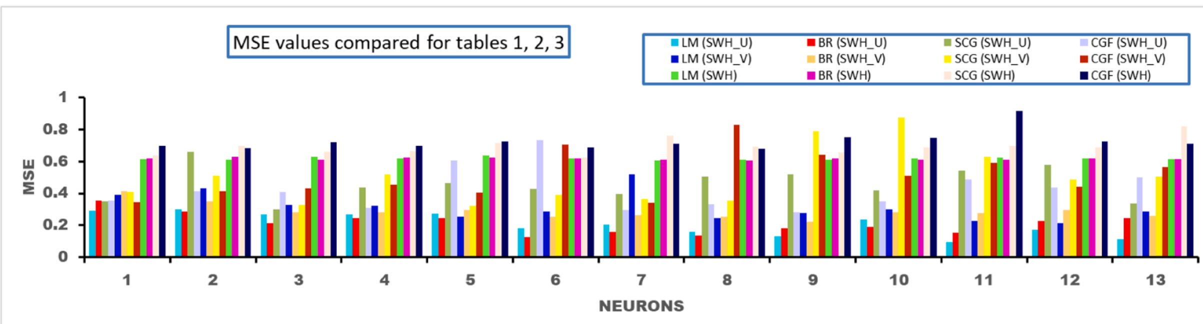

Fig. 10. MSE values compared for tables 1, 2, 3.

Table 4
Prediction results with delay from 1 to 12 using first PCs of SWH and U (multivariate case).

	LM		BR		SCG		CGF	
Neuron→	11		6		13		7	
Delays ↓	MSE	Epoch	MSE	EPOCH	MSE	EPOCH	MSE	EPOCH
1	0.20323	8	0.4573	36	0.42575	12	0.75151	16
2	0.24341	3	0.11753	109	0.80024	27	0.34753	41
3	0.16378	4	0.11461	130	0.98052	16	0.57827	33
4	0.27078	7	0.17724	762	0.2034	52	0.78053	19
5	0.18963	12	0.081923	162	0.58872	25	0.63665	42
6	0.69387	5	0.0071543	212	0.60085	42	0.3608	55
7	0.22644	3	0.042951	225	0.43972	34	0.6393	6
8	0.46298	2	0.090003	798	0.72938	14	0.48233	31
9	0.78458	2	0.016214	289	0.46472	4	0.96951	51
10	0.56268	3	0.048259	106	0.49247	4	0.691	8
11	0.42966	4	0.047566	1000	0.24831	1	0.51731	15
12	0.57757	5	0.0595	249	0.39947	1	0.41265	1

Table 5
Prediction results with delay from 1 to 12 using first PCs of SWH and V (multivariate case).

	LM		BR		SCG		CGF	
Neuron→	12		9		5		7	
Delays ↓	MSE	Epoch	MSE	EPOCH	MSE	EPOCH	MSE	EPOCH
1	0.78148	7	0.69814	24	0.84691	13	0.46054	3
2	0.41058	3	0.2151	682	0.93693	10	0.64787	17
3	0.30865	7	0.19367	110	0.73205	28	0.99631	15
4	0.90717	5	0.11259	177	0.47549	14	0.76591	1
5	0.91275	8	0.10572	464	0.34011	20	0.70867	25
6	0.43437	4	0.099513	186	0.90324	27	0.64738	11
7	0.87179	5	0.00022358	368	0.64344	13	0.79765	1
8	0.25754	4	0.089834	89	0.88046	2	0.49019	4
9	0.93565	4	0.0091432	1000	0.89369	14	0.98942	46
10	0.79209	3	0.044263	37	0.84996	16	0.24235	26
11	0.84173	3	0.00074015	37	0.68787	41	0.38386	33
12	0.45923	1	0.0014599	145	0.68621	30	0.50599	40

Table 6
Prediction results with delay from 1 to 12 using first PC of SWH (univariate).

	LM		BR		SCG		CGF	
Neuron→	7		8		6		8	
Delays ↓	MSE	Epoch	MSE	EPOCH	MSE	EPOCH	MSE	EPOCH
1	2.1422	11	2.2761	14	3.0597	49	2.8154	207
2	0.62166	64	0.61862	1000	0.643	101	0.67835	132
3	0.6212	8	0.5887	1000	0.80762	136	0.84576	67
4	0.62351	13	0.59638	758	0.76086	172	0.71575	155
5	0.60744	214	0.58525	1000	0.79035	157	0.84047	176
6	0.44946	366	0.46369	1000	0.54838	179	0.56795	142
7	0.44248	342	0.44423	1000	0.83106	43	0.61068	194
8	0.43213	67	0.43591	957	0.76272	81	0.59325	81
9	0.4765	7	0.43582	735	0.88371	42	0.66681	97
10	0.4692	236	0.42902	856	0.94798	154	0.50288	250
11	0.52653	29	0.42501	991	0.76216	63	0.55705	276
12	0.44121	28	0.43339	1000	0.64009	191	0.65121	229

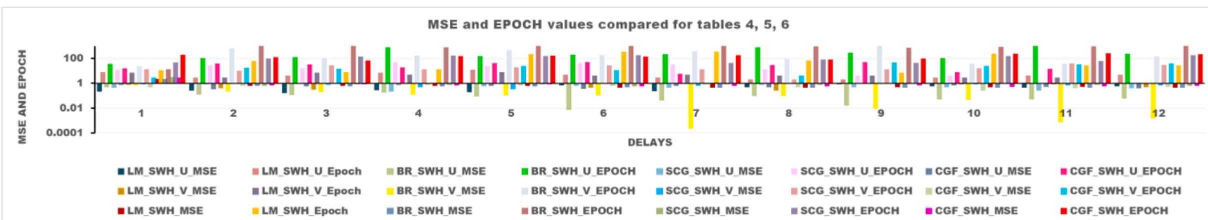

Fig. 11. MSE and EPOCH values compared for tables 4, 5, 6.

Table 7
Prediction results with delay from 1 to 12 using SWH and U for Normal Grid (multivariate case).

	LM		BR		SCG		CGF	
Neuron→	11		6		13		7	
Delays ↓	MSE	Epoch	MSE	EPOCH	MSE	EPOCH	MSE	EPOCH
1	0.0031708	2	0.0020114	47	0.002326	12	0.0026193	23
2	0.0011743	4	0.00099361	45	0.0021729	7	0.0017321	17
3	0.0013594	10	0.00031272	202	0.0017848	21	0.0017589	3
4	0.0020071	2	0.00011179	208	0.0019305	34	0.0065791	17
5	0.0013325	1	0.012125	2	0.001685	14	0.0030352	2
6	0.0015529	2	0.013087	2	0.0014491	16	0.0016862	15
7	0.0029328	2	0.011713	2	0.0034996	12	0.001501	7
8	0.00064895	2	0.013234	2	0.0017421	7	0.0026751	19
9	0.002061	2	0.012991	2	0.0026649	17	0.0014392	22
10	0.0020061	1	0.012128	2	0.0025278	27	0.0037429	38
11	0.0064057	2	0.010868	2	0.0027503	11	0.00085851	11
12	0.0030456	2	0.011444	2	0.0011635	18	0.00082686	14

Table 8
Prediction results with delay from 1 to 12 using SWH and V for Normal Grid (multivariate case).

	LM		BR		SCG		CGF	
Neuron→	12		9		5		7	
Delays ↓	MSE	Epoch	MSE	EPOCH	MSE	EPOCH	MSE	EPOCH
1	0.0035352	3	0.0016553	51	0.0023268	8	0.0046809	3
2	0.0017536	4	0.00094272	253	0.00064388	14	0.0011964	21
3	0.0012013	2	0.00069393	47	0.0038196	14	0.0020339	6
4	0.0033278	15	0.00025019	146	0.0027441	41	0.00085722	6
5	0.002953	2	0.00014845	304	0.0026416	3	0.0026587	3
6	0.0011183	1	0.00055193	67	0.0012325	5	0.0015633	3
7	0.001107	2	0.012994	2	0.00048321	9	0.0015708	13
8	0.002901	3	0.012212	2	0.00064112	12	0.0017363	14
9	0.001536	4	0.01321	2	0.001469	8	0.0016343	12
10	0.002218	2	0.012063	2	0.00075998	17	0.0017195	16
11	0.0016399	2	0.011266	2	0.0017318	7	0.0021606	28
12	0.0031742	2	0.011049	2	0.0010227	12	0.0014721	17

(1967). EOFs, however, are not restricted to multivariate statistics or atmospheric sciences. They extend to the analysis of stochastic fields in the mathematical literature where they are known under the name Karhunen-Lo`eve basis functions (Lo`eve 1978). Detailed analyses of EOFs can be found for example in Wilks (1995), von Storch and Zwiers (1999), and Jolliffe (2002). In this study the PCs are subjected to neural

Table 9
Prediction results with delay from 1 to 12 using SWH for Normal Grid (univariate).

	LM		BR		SCG		CGF	
Neuron→	7		8		6		8	
Delays ↓	MSE	Epoch	MSE	EPOCH	MSE	EPOCH	MSE	EPOCH
1	0.0071111	7	0.0090846	968	0.018238	8	0.13096	41
2	0.012623	6	0.0059592	656	0.030423	54	0.012742	106
3	0.0047111	6	0.0056723	483	0.040017	8	0.016339	70
4	0.02277	3	0.0048328	571	0.0072451	31	0.0069855	48
5	0.0090357	12	0.0051994	512	0.02089	29	0.0292	14
6	0.0086401	15	0.0050633	318	0.039694	21	0.014706	36
7	0.0066282	29	0.0043237	465	0.021771	31	0.006155	73
8	0.008704	6	0.0040026	615	0.039798	12	0.077855	17
9	0.013619	16	0.0042253	285	0.022613	23	0.091517	11
10	0.046085	7	0.0041039	608	0.027201	31	0.011003	71
11	0.0056112	5	0.0035516	473	0.030439	43	0.0091681	83
12	0.010803	12	0.003628	518	0.013342	35	0.0070541	67

Table 10
Prediction results with delay from 1 to 12 using SWH and U for Cyclonic Grid (multivariate case).

	LM		BR		SCG		CGF	
Neuron→	11		6		13		7	
Delays ↓	MSE	Epoch	MSE	EPOCH	MSE	EPOCH	MSE	EPOCH
1	0.0013027	3	0.00095184	19	0.0023012	28	0.022035	25
2	0.0045229	10	0.00077061	184	0.0014327	11	0.00074356	32
3	0.0036226	7	0.00069761	16	0.00057785	28	0.0017648	48
4	0.0016642	2	0.00031996	79	0.0036214	7	0.00089849	23
5	0.0024785	4	0.00022367	1000	0.0028836	6	0.0025898	8
6	0.0026193	4	0.0001731	467	0.002266	21	0.0030139	7
7	0.0028072	2	0.0005849	49	0.0016388	18	0.0010578	22
8	0.001095	3	0.00032362	321	0.0022255	27	0.0013086	12
9	0.0037643	4	0.000325	142	0.00070684	65	0.0013627	30
10	0.0040867	2	0.000121	157	0.0030351	11	0.0053262	4
11	0.0036999	2	0.000207	182	0.0040432	11	0.0019714	12
12	0.0034333	2	0.00969	127	0.0023558	24	0.0077383	17

Table 11
Prediction results with delay from 1 to 12 using SWH and V for Cyclonic Grid (multivariate case).

	LM		BR		SCG		CGF	
Neuron→	12		9		5		7	
Delays ↓	MSE	Epoch	MSE	EPOCH	MSE	EPOCH	MSE	EPOCH
1	0.0016084	20	0.00071733	159	0.001398	27	0.0016826	9
2	0.0010491	3	0.00046379	312	0.0018593	32	0.0027025	4
3	0.0017673	4	0.00050788	52	0.0028108	15	0.00091296	17
4	0.0014231	6	0.00054771	50	0.0022976	24	0.0047081	3
5	0.0014374	2	0.00038716	46	0.0041091	5	0.0010313	25
6	0.0012822	3	0.00023307	91	0.0025661	19	0.0008081	27
7	0.003772	2	0.00028263	182	0.0017938	8	0.00072844	52
8	0.0015069	1	0.00050094	60	0.0013341	56	0.0027934	6
9	0.0022644	6	0.00032867	998	0.00064249	12	0.0027421	3
10	0.0012187	2	0.00028019	998	0.0018504	13	0.0032674	19
11	0.0025088	2	0.000284	263	0.00077328	28	0.0013174	20
12	0.0031934	1	0.00049083	136	0.00088862	15	0.0012438	39

network models to assess the number of hidden neurons for an entire region. The purpose is to use these in future for prediction of particular datasets of the region.

2.2. NARX and NAR models

Modeling of non-linear dynamical systems can be done using the NARX network which has a dynamical neural architecture (Xie et al., 2009). The dynamic models can again have feedforward or recurrent connections. The non-linear neural network with independent inputs (NARX), can be stated as a recurrent dynamic network defined by

$$y(t) = f\left(y(t-1), \ldots y\left(t-n_y\right), z(t-1), \ldots z(t-n_z)\right) \quad (1)$$

where y(i) and z(i) are the outputs and inputs at the i^{th} time. The latest output is regressed for n_y and n_z time steps. A NARX model can be trained using parallel architecture and the series-parallel architecture (Boussaada et al., 2018).

Time series prediction may be performed based on non-linear autoregressive (NAR) models (Chow and Leung, 1996). NAR model applied to time series forecasting is described as a discrete, non-linear, autoregressive model (Chen et al., 1990) and is written as follows:

$$y(t) = f(y(t-1), y(t-2), \ldots y(t-n)) + \lambda(t) \quad (2)$$

The above relationship shows that a NAR model is employed to predict the worth of a knowledge series y at time t, using the n previous values of the series (Potdar and Kinnerkar, 2017). The function f (.) is unknown, and therefore the training of the ANN approximates the above function using network weights and neuron bias. Due to the approximation of the

Table 12
Prediction results with delay from 1 to 12 using SWH for Cyclonic Grid (univariate).

Neuron→	LM		BR		SCG		CGF	
	7		8		6		8	
Delays ↓	MSE	Epoch	MSE	EPOCH	MSE	EPOCH	MSE	EPOCH
1	0.017901	23	0.018814	817	0.18318	9	0.026493	26
2	0.017933	4	0.015162	656	0.012543	18	0.020381	26
3	0.0361	3	0.010822	647	0.016083	34	0.015248	8
4	0.022265	4	0.0093147	381	0.054091	33	0.018969	22
5	0.014858	24	0.0084033	493	0.049102	16	0.01434	34
6	0.011449	4	0.0076164	409	0.075216	32	0.021691	13
7	0.015375	7	0.0073874	471	0.076569	1	0.02583	15
8	0.009165	15	0.0074151	265	0.059441	18	0.021723	15
9	0.020476	6	0.0072042	293	0.096804	19	0.024278	28
10	0.026173	4	0.0066927	282	0.026692	36	0.035568	9
11	0.16793	3	0.006758	270	0.13562	16	0.081053	19
12	0.011782	6	0.0066356	483	0.067072	13	0.033987	45

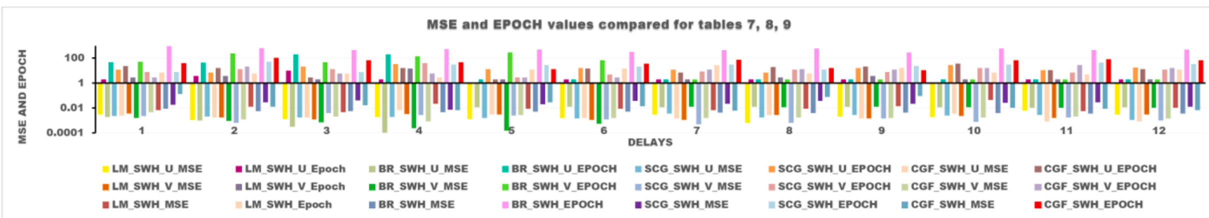

Fig. 12. MSE and EPOCH values compared for tables 7, 8, 9.

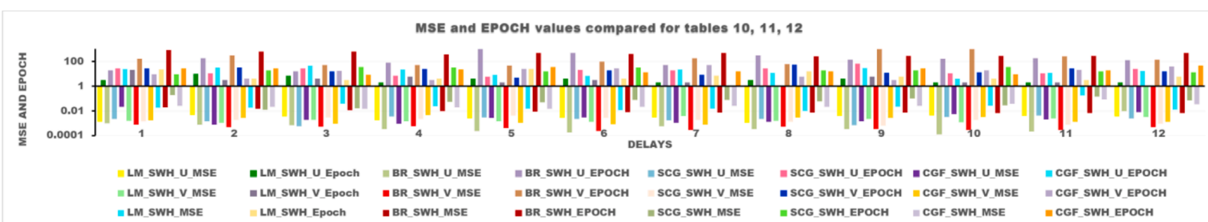

Fig. 13. MSE and EPOCH values compared for tables 10, 11, 12.

series y at time t, the error is given by $\lambda(t)$. The features $y(t - 1)$, $y(t - 2)$, … $y(t - n)$, are called the feedback delays.

2.3. Levenberg-Marquardt and Bayesian regularization backpropagation algorithms

Levenberg-Marquardt backpropagation algorithm (LM) is the most popular optimization BP algorithm designed for a non-linear least square minimization problem. This implies that the algorithm is more efficient and faster than the others. For the LM algorithm the Hessian matrix can be approximated as $H = \mathbf{J}^{\mathrm{T}}\mathbf{J}$. The gradient is given as $g = \mathbf{J}^{\mathrm{T}}\mathbf{e}$, where $\mathbf{J}$ is a Jacobian matrix, and $\mathbf{e}$ is a vector of network errors (Wilamowski and Yu, 2010). The LM algorithm (Reynaldi et al., 2012) uses the following approximation to the Hessian matrix, in which $\mathbf{x}$ represents connection weights.

$$\mathbf{x}_{k+1} = \mathbf{x}_k - - \left[\mathbf{J}^T\mathbf{J} + \mu\mathbf{I}\right]^{-1} \mathbf{J}^T\mathbf{e} \tag{3}$$

The LM algorithm is faster than the conventional gradient descent techniques.

Bayesian regularization (BR) is a training algorithm that provides a rigorous general framework for dynamic state estimation problems. For best results the above algorithm changed the training objective function by introducing network weights (Yue et al., 2011). It is denoted as $F(\theta)$ and given as

$$F(\theta) = \alpha E_\theta + \beta E_\sigma. \tag{4}$$

E_θ is the sum of the squared network weights and E_σ is the sum of squared network errors. The parameters of the objective function are given by α and β. In the BR network, the network weights are regarded as random variables (Foresee and Hagan, 1997), and their distribution as Gaussian. If α is large compared to β the errors are lower although there maybe still over-fitting.

2.4. Scaled conjugate gradient and Fletcher conjugate gradient backpropagation algorithms

The Scaled Conjugate Gradient (SCG) algorithm is given by Moller (1993) and is having super linear convergence rate. SCG is based on optimization techniques known as the Conjugate Gradient (CG) Methods. The network weights are adjusted by the basic back-propagation algorithm along the steepest descent direction. Although the performance function decreases most quickly along the given direction, this may not give the quickest convergence. In the CG algorithms the objective is finding the direction which gives faster convergence keeping the error minimum (Johansson et al., 1991).

CG-based multilayer perceptron neural networks can produce faster speed of convergence but require the computation of the Hessian matrix. The Fletcher–Reeves or CGF is an improved version of the multilayer perceptron neural network-CG algorithm (Chatterjee 2010). CGF is another steepest descent BP algorithm solving unconstrained optimization problems with the finest accuracy. This method is used for solving unconstrained optimization problems. It is well-known that the direction generated by a conjugate gradient method might not be a descent direction of the objective function. A little modification is made to the Fletcher–Reeves method such as the direction generated by the modified method provides a decent direction for the target function. It is also called a line search method where Hessian is not used. It is undifferentiated to the original conjugate direction algorithm in the quadratic case.

In this paper, the dominant mode of the wind and wave parameters are extracted using EOF analysis and then the principal components are trained using BP algorithms. The novelty of the study lies in the multivariate and univariate approaches used for predicting 12 multistep time steps ahead involving the BP algorithms and EOF analysis for the whole Bay of Bengal (BOB) region.

3. Data

For the BOB region, six-hourly SWH, zonal (U) and meridional (V) winds are obtained from NOAA WAVEWATCH III global wave model runs for the period 2006–2015 (https://polar.ncep.noaa.gov/waves/hindcasts/prod-multi_1.php). The multigrid nested model uses ETOPO-1 bathymetry, 10 meter above the surface winds and daily ice fields having half degree resolution as inputs. Winds from Global Forecast System with half degree spatial resolution and three hour temporal intervals have been used as input to the model. The output fields extracted and processed for this work are having 50 km or half degree spatial resolution and six hourly temporal resolution. EOF analysis is performed separately on the three parameters and the first principal components obtained are trained with different BP algorithms. In the univariate approach past SWH data are utilized to predict future SWH values whereas in the multivariate approach zonal and meridional wind values are used to predict future SWH values. NAR model is used for the univariate case and NARX for the multivariate case. After training the models using different backpropagation algorithms optimum values of the hidden neurons are generated which are further used to predict independent datasets. In the BOB region two independent grids are chosen considering normal and cyclonic conditions. U, V and SWH data for the year 2016 are extracted for the grid location (15°N and 90°E) and then the grid location (13°N and 81°E) through which VARDAH (06 −13 December 2016) cyclone passed. These time series data are used as independent sets to test the neural network models using the optimum hidden neurons.

4. Methodology

For the purpose of prediction, training with open loop is far more accurate in comparison to that with closed loop. In the former case while training, the network gets correct feedback inputs in order to produce correct feedback outputs. After the training is over, the network can be converted to closed-loop form, as per requirement. The train network is a network that can calculate good generalization for difficult, small, or noisy datasets using Levenberg-Marquardt (LM) or Bayesian Regularization (BR) or Scaled Conjugate Gradient (SCG) or Fletcher Conjugate Gradient (CGF). Training stops according to adaptive weight minimization (regularization) and the performance is evaluated in terms of mean square error (MSE). The MSE is given by the average squared difference between the outputs and the targets. The lower the value it is more accurate and zero value signifies no error. The methodological framework contains stages like collection of data, training the model, testing the accuracy of various BP algorithms, predictions using varied neurons and delays. MATLAB is used as the platform for the various experiments. The datasets are divided into training, validation and testing sets. For the univariate approach only SWH data is used. For the multivariate approach the inputs are zonal (U) or meridional (V) components of wind data, and the outputs are the SWH data. Different training algorithms are used to train the network for predictions to be held for 1 to 12 ahead time steps.

In this study, a widespread comparison is held for different ANN training algorithms using multistep ahead time step and various neuron values for wave height forecasting. Four types of backpropagation training algorithms (LM, BR, SCG, CGF) for a given feedforward neural network are investigated. The accuracy of the performance is measured in terms of MSE.

5. Results and discussions

EOF analysis is applied to the 10 years (2006–2015) spatially gridded SWH, U and V data separately for the BOB region. The first eigenmode accounts for 80,77 and 72% of the total variability of the respective parameters for the BOB region. The temporal patterns of the first principal components (PC) are shown in Figs. 1a, 2a and 3a respectively. They exhibit annual periodicity with maximum load occurring during the southwest monsoon period. The EOF spatial patterns are given in Figs. 1b, 2b and 3b. Although the second and third modes together account for 10% of the variance, in this study the first eigenmode being the dominant one is only considered. In Fig. 1b the maximum loading of the SWH parameter is in the central Bay region. In Fig. 2b the zonal component of wind data shows a horizontal maxima pattern while in Fig. 3b the meridional component shows a vertical pattern. The zonal component is dominant in the central Bay region whereas the meridional component stretches over the entire basin. Thus the zonal component contributes more towards SWH generation.

Next a normal grid location (15°N and 90°E) and a cyclonic grid location (13°N and 81°E) are chosen in the BOB region for the year 2016 to be used for location specific prediction. Figs. 4, 5 and 6 give the time series plots of the normal grid location for the SWH, U and V parameters respectively. Similar plots are given in Figs. 7, 8 and 9 for the cyclonic grid through which the very severe cyclonic storm VARDAH (06 −13 December 2016) passed. For basin-scale estimation EOF analysis is performed for the BOB region and the dominant PC is fitted with different backpropagation (BP) algorithms. After training the models with the PCs, the neurons are fixed and used for predictions with the time-series data of 2016.

In the univariate approach, the PC of the SWH is trained using different BP algorithms (LM, BR, SCG, CGF). The future values of the principal component time series are predicted from the past values of the same. The input time series is divided into 3 sets, namely, training, validation and testing and using a neural network model predictions are conducted. 70% of the time series (10,136 points) is used for training, 20% for validation (2896 points), and rest 10% (1448 points) for independent testing of network generalization. The number of hidden neurons is set from 1 to 13 and the number of delays from 1to 12. Similarly, for the multivariate approach, the PC of the SWH, U, and V are trained for multistep ahead time steps and varied neurons. The objective of this study is to compare the different BP algorithms for the univariate and multivariate approaches using the PCs for basin scale forecasting. The datasets used here are obtained from NOAA WAVEWATCH III, which contains the zonal (U) wind component, the meridional (V) wind component, and the significant wave height (SWH) data. The performance is evaluated by studying the MSE. Several experiments are conducted with the PCs as input and varied neurons (1 to 13).

Tables 1 and 2 give the performance errors for the multivariate case where the PCs of SWH, U and V have been used and Table 3 gives the errors for the univariate case where only the PC of SWH is used. For the univariate case all the BP algorithms performed uniformly while for the multivariate case LM performed better than the others. From the experiments it can be said predictions of SWH using the zonal wind (U) have less error compared to the meridional wind (V). Thus the zonal component of the wind velocity contributes more than the meridional component for the generation of the wave height on the ocean surface. Fig. 10 gives the comparison of the MSE values given in Tables 1–3.

After fixing the number of neurons from the above experiments, next the MSE are calculated using the PCs for delays from 1 to 12 which corresponds to 3 days forecast for the 06 hourly datasets. Tables 4–6 gives the errors estimated using the various BP algorithms for the given delays and using the PCs which represents the entire basin. BR algorithm performed better than the others for the SWH prediction using U and V components of wind and varied delays. For the univariate case where only PC of SWH is used there is no significant change in the performance errors with the different BP algorithms. Considering all the univariate and multivariate cases (except delay 1 in Table 6) the error is less than a meter. The performance of the neural network models using the PCs representing a basin is discussed here and this can be analysed further in future with different validation datasets. Fig. 11 gives the comparison of the MSE and EPOCH values given in Tables 4–6.

Next the methodology is tested using specific locations in the basin. Time series data are generated for a normal and a cyclonic grid and utilized for predictions. Tables 7–9 gives the performance results for the normal grid for both the multivariate and univariate cases. Similarly for the cyclonic grid the MSE values are given in the Tables 10–12. It is observed the LM algorithm performed with more accuracy than the others. For the normal grid the maximum error was 0.04 meter and for the cyclonic grid it was 0.1 m considering both the multivariate and univariate cases. Thus as per data availability during normal and extreme conditions the neural network models can be used for 3 days forecast for the entire basin or specific locations.

Fig. 12 given below compares the MSE and EPOCH values given in Tables 7–9.

Fig. 13 given below compares the MSE and EPOCH values given in Tables 10–12.

6. Conclusions

This study is conducted to analyze and predict normal and extreme wave events that can damage marine ecosystems, affect coastal infrastructure and ocean industries. Data obtained for the BOB region for the period 2006–2015 has been used for basin scale forecast. Next 2016 normal and cyclonic time series data are chosen for specific location based forecast. From the results, it is observed that the wave height estimation for both the cyclonic and normal data is quite close for all the BP algorithms used. Using numerical models for wave forecasts during extreme conditions is always a challenge. Thus alternative techniques using neural network models are very much advantageous technically and computationally. Using the PCs for basin scale forecast along with the neural network models adds to the novelty of this study. For the univariate case where only SWH data are considered the MSE values are comparatively larger than the multivariate case which involves multiple datasets making the system complex. It can be said using past values of zonal and meridional wind data, multistep ahead prediction of the SWH data can be performed with reasonable accuracy. In this study, a maximum of 12 delay corresponding to 72 h ahead forecast gives reasonable MSE values for all the BP algorithms. However, the accuracy of the BR algorithm is found to be higher in comparison to the other three algorithms for both the univariate and the multivariate approaches. On training the model multiple times different results are generated due to different initial conditions and sampling. In terms of speed of convergence, where the epoch values are considered, the LM algorithm performs the best. Wind-wave forecasting plays a significant role in development and planning of coastal area, and thus accurate wind-wave predictions can substantially reduce the impact of coastal hazards pertaining to coastal area development. As future work, the methodology has to be tested on further datasets involving different cyclones and the predictions can be improved further including the different modes of the PCs.

On behalf of all authors, the corresponding author states that there is no conflict of interest.

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors are thankful to the Environmental Modeling Center for the archived model data from the NOAA WAVEWATCH III runs.

References

Baldi, P., Hornik, K., 1989. Neural networks and principal component analysis: learning from examples without local minima. Neural Netw. 2 (1), 53–58. https://doi.org/10.1016/0893-6080(89)90014-2.

Boussaada, Z., Curea, O., Remaci, A., Camblong, H., Bellaaj, N.M., 2018. A nonlinear autoregressive exogenous (NARX) Neural network model for the prediction of the daily direct solar radiation. Energies 11, 620. https://doi.org/10.3390/en11030620.

Booij, N., Ris, R.C., Holthuijsen, L.H., 1999. A third-generation wave model for coastal regions: 1. Model description and validation. J. Geophys. Res. 104 (C4), 7649–7666. https://doi.org/10.1029/98JC02622.

Chatterjee, A., 2010. A fletcher-reeves conjugate gradient neural-network-based localization algorithm for wireless sensor networks. IEEE Trans. Veh. Technol. 59 (2), 823–830. https://doi.org/10.1109/TVT.2009.2035132.

Chaudhuri, S., Dutta, D., Goswami, S., Middey, A., 2015. Track and intensity forecast of tropical cyclones over the North Indian Ocean with multilayer feed forward neural nets. Meteorol. Appl. 22, 563–575. https://doi.org/10.1002/met.1488.

Chen, S., Billings, S.A., Grant, P.M., 1990. Non-linear system identification using neural networks. Int. J. Control 51 (6), 1191–1214. https://doi.org/10.1080/00207179008934126.

Chow, T.W.S., Leung, C.T., 1996. Neural network based short-term load forecasting using weather compensation. IEEE Trans. Power Syst. 11 (4), 1736–1742. https://doi.org/10.1109/59.544636.

Foresee, F.D., Hagan, M.T., 1997. Gauss-newton approximation to Bayesian learning. In: Proceedings of the International Conference on Neural Networks (ICNN'97). Houston, TX, USA, 3, pp. 1930–1935. https://doi.org/10.1109/ICNN.1997.614194.

Ghanbarzadeh, A., Noghrehabadi, A.R., Behrang, M.A., Assareh, E., 2009. Wind speed prediction based on simple meteorological data using artificial neural network. In: Proceedings of the 7th IEEE International Conference on Industrial Informatics. Cardiff, Wales, pp. 664–667. https://doi.org/10.1109/INDIN.2009.5195882.

Guo, Z., Wu, J., Lu, H., Wang, J., 2011. A case study on a hybrid wind speed forecasting method using BP neural network. Knowl. Based Syst. 24 (7), 1048–1056. https://doi.org/10.1016/j.knosys.2011.04.019.

Hagan, M.T., Menhaj, M., 1994. Training feed-forward networks with the Marquardt algorithm. IEEE Trans. Neural Netw. 5 (6), 989–993. https://doi.org/10.1109/72.329697.

Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components. J. Educ. Psychol. 24 (6), 417–441. https://doi.org/10.1037/h0071325.

Johansson, E.M., Dowla, F.U., Goodman, D.M., 1991. Backpropagation learning for multilayer feed-forward neural networks using the conjugate gradient method. Int. J. Neural Syst. 2 (4), 291–301. https://doi.org/10.1142/S0129065791000261.

Jolliffe, I.T., 2002. Principal Component Analysis, 2nd Ed. Springer-Verlag. https://doi.org/10.1007/b98835.

Kutzbach, J.E., 1967. Empirical eigenvectors of sea-level pressure, surface temperature and precipitation complexes over North America. J. Appl. Meteorol. Climatol. 6, 791–802. https://doi.org/10.1175/1520-0450(1967)006-0791:EEOSLP-2.0.CO;2.

Lecacheux, S., Pedreros, R., Le Cozannet, G., Thiébot, J., De La Torre, Y., Bulteau, T., 2012. A method to characterize the different extreme waves for islands exposed to various wave regimes: a case study devoted to Reunion Island. Nat. Hazards Earth Syst. Sci. 12, 2425–2437. https://doi.org/10.5194/nhess-12-2425-2012.

Li, G., Shi, J., 2010. On comparing three artificial neural networks for wind speed forecasting. Appl. Energy 87 (7), 2313–2320. https://doi.org/10.1016/j.apenergy.2009.12.013.

Loridan, T., Crompton, R.P., Dubossarsky, E., 2017. A machine learning approach to modeling tropical cyclone wind field uncertainty. Mon. Weather Rev. 145, 3203–3221. https://doi.org/10.1175/MWR-D-16-0429.1.

Lo`eve, M., 1978. Probability Theory, 4th ed., II. Springer-Verlag, p. 413.

Lodge, A., Yu, X., 2014. Short term wind speed prediction using artificial neural networks. In: Proceedings of the 4th IEEE International Conference on Information Science and Technology. Shenzhen, pp. 539–542. https://doi.org/10.1109/ICIST.2014.6920535.

Lorenz, E.N., 1963. Deterministic nonperiodic flow. J. Atmos. Sci. 20, 130–141.

Møller, M.F., 1993. A scaled conjugate gradient algorithm for fast supervised learning. Neural Netw. 6 (4), 525–533. https://doi.org/10.1016/S0893-6080(05)80056-5.

Parmar, R., Shah, M., Shah, M.G., 2017. A comparative study on different ANN techniques in wind speed forecasting for generation of electricity. IOSR J. Electr. Electron. Eng. 12 (1), 19–26. https://doi.org/10.9790/1676-1201031926.

Pan, X., Lee, B., Zhang, C., 2013. A comparison of neural network backpropagation algorithms for electricity load forecasting. In: Proceedings of the IEEE International Workshop on Intelligent Energy Systems (IWIES). Vienna, pp. 22–27. https://doi.org/10.1109/IWIES.2013.6698556.

Pradhan, R., Aygun, R.S., Maskey, M., Ramachandran, R., Cecil, D.J., 2018. Tropical cyclone intensity estimation using a deep convolutional neural network. IEEE Trans. Image Process. 27 (2), 692–702.
Potdar, K., Kinnerkar, R., 2017. A non-linear autoregressive neural network model for forecasting Indian index of industrial production. In: Proceedings of the IEEE Region 10 Symposium (TENSYMP), Cochin 1-5. https://doi.org/10.1109/TENCONSpring.2017.8069973.
Reynaldi, A., Lukas, S., Margaretha, H., 2012. Backpropagation and levenberg-marquardt algorithm for training finite element neural network. In: Proceedings of the Proceedings of the Sixth UKSim/AMSS European Symposium on Computer Modeling and Simulation. Valetta, pp. 89–94. https://doi.org/10.1109/EMS.2012.56.
Richman, M.B., Leslie, L.M., Ramsay, H.A., Klotzbach, P.J., 2017. Reducing tropical cyclone prediction errors using machine learning approaches. Procedia Comput. Sci. 114, 314–323. https://doi.org/10.1016/j.procs.2017.09.048.
Ris, R.C., Holthuijsen, L.H., Booij, N., 1999. A third-generation wave model for coastal regions: 2. Verification. J. Geophys. Res. 104 (C4), 7667–7681. https://doi.org/10.1029/1998JC900123.
Sampson, C.R., Hansen, J.A., Wittmann, P.A., Knaff, J.A., Schumacher, A., 2016. Wave probabilities consistent with official tropical cyclone forecasts. Weather Forecast. 31, 2035–2045. https://doi.org/10.1175/WAF-D-15-0093.1.
Saroha, S., Aggarwal, S.K., 2014. Multi-step ahead forecasting of wind power by different class of neural networks. Recent Adv. Eng. Comput. Sci. (RAECS) 1–6. https://doi.org/10.1109/RAECS.2014.6799528.
Taieb, S.B., Bontempi, G., Atiya, A.F., Sorjamaa, A., 2012. A review and comparison of strategies for multi-step ahead time series forecasting based on the NN5 forecasting competition. Expert Syst. Appl. 39 (8), 7067–7083. https://doi.org/10.1016/j.eswa.2012.01.039.
Tangang, F.T., Tang, B., Monahan, A.H., Hsieh, W.W., 1998. Forecasting ENSO events: a neural network-extended EOF approach. J Clim 11 (1), 29–41. https://doi.org/10.1175/1520-0442(1998)011-0029:FEEANN-2.0.CO;2.
Tiwari, S., Naresh, R., Jha, R., 2013. Comparative study of backpropagation algorithms in neural network based identification of power system. Int. J. Comput. Sci. Inform. Technol. 5 (4) https://doi.org/10.5121/ijcsit.2013.5407.
Tolman, H.L., 1991. A third-generation model for wind waves on slowly varying, unsteady, and inhomogeneous depths and currents. J. Phys. Oceanogr. 21 (6), 782–797. https://doi.org/10.1175/1520-0485(1991)021<0782:ATGMFW>2.0.CO;2.
von Storch, H., Zwiers, F.W., 1999. Statistical Analysis in Climate Research. Cambridge University Press, Cambridge.
WAMDIG, 1988. The WAM model-a third generation ocean wave prediction model. J. Phys. Oceanogr. 18 (12), 1775–1810.
Wilamowski, B.M., 2009. Neural network architectures and learning algorithms. Ind. Electron. Mag. IEEE 3 (4), 56–63. https://doi.org/10.1109/MIE.2009.934790.
Wilamowski, B.M., Yu, H., 2010. Improved computation for Levenberg-Marquardt training. IEEE Trans. Neural Netw. 21 (6), 2010.
Wilks, D.S., 1995. Statistical Methods in the Atmospheric Sciences. Academic Press.
Xie, H., Tang, H., Liao, Y., 2009. Time series prediction based on NARX neural networks: an advanced approach. In: Proceedings of the International Conference on Machine Learning and Cybernetics. Hebei, pp. 1275–1279. https://doi.org/10.1109/ICMLC.2009.5212326.