

OCEAN WAVE FORECASTING USING RECURRENT NEURAL NETWORKS

Mandal S¹ and Prabaharan N²

National Institute of Oceanography, Dona Paula, Goa 403 004, INDIA

Phone-+91-832-2450265; Fax - +91-832-2450604

Email: 1smandal@nio.org ; 2prabha@nio.org

ABSTRACT

The tremendous increase in offshore operational activities demands improved wave-forecasting techniques. With the knowledge of accurate wave conditions, it is possible to carry out the marine activities such as offshore drilling, naval operations, merchant vessel routing, nearshore construction, etc. more efficiently and safely. This paper describes an artificial neural network, namely recurrent neural network with rprop update algorithm and is applied for wave forecasting. Measured ocean waves off Marmugao, west coast of India are used for this study. Here, the recurrent neural network of 3, 6 and 12 hourly wave forecasting yields the correlation coefficients of 0.95, 0.90 and 0.87 respectively. This shows that the wave forecasting using recurrent neural network yields better results than the previous neural network application.

INTRODUCTION

Ocean waves can be either forecasted using the meteorological conditions or hindcasted from the existing meteorological charts. However, the forecast may not accurately represent the measured values. The parametric or differential equation based on wind wave relationship and a differential equation of wave energy are solved numerically in wave forecasting. This is generally employed to give an estimate over the following 6-72 hours or so. The spatial wave information on numerical wave forecasting schemes are useful and attractive in many applications, but it needs elaborate meteorological and oceanographic data sets and involves an enormous amount of computational effort. Apart from this translation of data from wind to waves results in an element of uncertainty and approximation of forecasts (Herbich, 1990).

The work carried out by Deo and Naidu (1999) describes the application of neural network analysis in forecasting of waves and is carried out for significant wave height (Hs) with 3 hour lead period. They have carried out various combinations of training

patterns to obtain the desired output. Also the work of average 12hr and a day wave forecasting are carried out. In their work the backpropagation, cascade correlation and conjugate gradient methods are used with the input layer of one node. The correlation coefficients obtained for the lead of 3 hour, half day and a day are 0.81, 0.78 and 0.71 respectively. Using updated algorithms in backpropagation neural network, the wave forecasting yielded better results (Subba Rao *et al*, 2001). Improved correlation coefficients for the lead of 3 hour, half-day and a day are obtained as 0.93, 0.80 and 0.73 respectively. Deo *et al* (2001) have demonstrated the use of neural networks for wave forecasting for three different sets of data. It shows that a proper trained network could yield good results in open wider areas, in deep water and when the sampling and prediction interval are large (say, week). More and Deo (2003) have used Elman and Jordan types recurrent neural networks (RNN) for wind forecasting. Ho et al (2002) have carried out a study on predictive compressor failures by autoregressive integrated moving average (ARIMA) and RNN models. They showed that RNN at the optimal weighting factor gives satisfactory performances compared to the ARIMA model. Kermanshahi (1998) has applied RNNs for forecasting next 10 years loads of nine Japanese utilities. Hindcasting of storm waves using back-propagation neural network was carried out by Subbarao and Mandal (2005).

This paper deals with the forecasting of wave heights directly from measured waves. Significant wave heights are forecasted using recurrent neural networks. Here nonlinear autoregressive models with exogenous inputs (NARX) recurrent algorithms are used in back-propagation neural network.

BACKPROPAGATION NEURAL NETWORK

The artificial neural network (ANN) is similar to the biological neurons, works on the input and output passing through a hidden layer. The ANN used here is a data-oriented modeling technique to find relations between input and output patterns by self learning and without any fixed mathematical form assumed beforehand. Here one needs not necessarily have the knowledge of the underlying physical process being involved. In this work, inputs are the consecutive points of significant wave heights and the output is a predicted wave height. The ANN is organised in the form of layers. The first one is the input layer consisting of nodes equal to the number of

input values. The last one is the output layer and the number of nodes is equal to the number of output values. The layer between the input and the output is called the hidden layer having number of nodes. The output of nodes, $f(x)$ is generated using a 'non-linear function', sigmoid shape.

$$f(x) = 1/(1+ e^{-x}) \quad (1)$$

where x is sum of the multiplied inputs to its weights.

Before neural network is used for forecasting, it is required to train using the available data and this is done using various training algorithms like standard backpropagation, Quickprop, Rprop, SuperSAB, Quasi-Newton, Lavenberg-Marquardt, etc. (Londhe and Deo, 2004; Subbarao and Mandal, 2005). In the present work, standard backpropagation with rprop algorithm is used.

In any training process, the objective is to reduce the overall error E , between network output and the actual observation, which is defined as

$$E = 1/p \sum E_p \quad (2)$$

Where,

$$E_p = 1/2 \sum (T_k - O_k)^2 \quad (3)$$

p is the total number of training patterns;

T_k is the actual output and

O_k is the predicted output at k th output node.

In the learning process of backpropagation neural network, the error at the output layer is back-propagated through hidden layer in the network to yield the desired outputs. The gradient descend method is used to calculate the weight of the network and then to adjust the weight of interconnections to reduce the output error. The interconnection weights are adjusted using an error convergence process called *updated algorithm*. There are many such updated algorithms available (Tveter, 2000; Deo and Naidu, 1999). Here we have chosen an updated algorithm namely rprop (Subbarao and Mandal, 2005) which seems to be efficient and consistent in

performance. The use of recurrent neural network will be a better choice in prediction of time series data sets (Principe *et al*, 2000). The recurrent neural networks take values from hidden layer or output layer units or combination of both and copy them down to the input layer for use with the next input. The values that are copied down are a kind of coded record of recent inputs to the network and this gives the network a simple kind of short-term memory, possibly a little like human short-term memory.

Rprop algorithm

Rprop stands for 'resilient propagation' and is an efficient learning scheme which performs a direct adaptation of the weight step based on local gradient information. Riedmiller (1994) introduced for each weight its individual update value, Δ_{ij} , which determines the size of the weight update. This adaptive update value evolves during the learning process as briefly described by Subbarao and Mandal (2005).

At the beginning, all updated values of Δ_{ij} are set to the initial value Δ_0 , then it determines directly the size of the first weight step which is preferably chosen in a reasonable proportion to the size of the initial weights. The good choice is $\Delta_0 = 0.1$. Even for much larger or much smaller values of Δ_0 the fast convergence is reached.

RECURRENT NEURAL NETWORKS

The use of backpropagation recurrent networks improves the forecast results. In this paper the Nonlinear Autoregressive with eXogeneous inputs (NARX) recurrent neural networks are used and results obtained from these are described.

Nonlinear AutoRegressive with exogenous inputs network

One of the classes of discrete time nonlinear system is Nonlinear Autoregressive with eXogeneous inputs (NARX) model (Lin et al, 1998)

$$y(t) = f \{ u(t-D_u), \dots, u(t-1), u(t), y(t-D_y), \dots, y(t-1) \} \quad (4)$$

and $u(t)$ and $y(t)$ represent input and output of the network at time t , D_u and D_y are the input and output orders with the nonlinear function f . The function f is approximated with a Multilayer Perceptron (MLP), and the resulting system is becoming a NARX recurrent neural network.

A NARX network consists of a Multilayer Perceptron which takes input window of past input and output values and computes the current output

$$y(t) = \Psi\{u(t-D_u), \dots, u(t-1), u(t), y(t-D_y), \dots, y(t-1)\} \quad (5)$$

where the Ψ is the mapping performed by the MLP in the network. Figure-1 shows that NARX network corresponds to a set of two tapped delay lines. One consists of D_u tapes in the input values, and other consists of D_y output values. And the states are updated as

$$x_i(t+1) = \begin{cases} u(t) & i = D_u \\ y(t) & i = D_u + D_y \\ x_{i+1}(t) & i < i < D_u \text{ and } D_u < i < D_u + D_y \end{cases} \quad (6)$$

so that at time t the taps correspond to the values

$$x(t) = [u(t-D_u) \dots u(t-1) \quad y(t-D_y) \dots y(t-1)] \quad (7)$$

The MLP consists of a set of two nodes organized into two layers. There are H nodes of the first layer and this performs the function

$$z_i(t) = \sigma \left(\sum_{j=1}^N a_{i,j} x_j(t) + b_i u(t) + c_i \right), \quad i=1,2,\dots, H. \quad (8)$$

Where σ is the nonlinear transfer function, $a_{i,j}$, b_i and c_i are fixed real valued weights.

And the output layer consists of single linear node

$$y(t) = \sum_{j=1}^H w_{i,j} z_j(t) + \theta_i, \quad i=1, 2, \dots, H. \quad (9)$$

where $w_{i,j}$ and θ_i are fixed real valued weights.

The NARX network with $D_u = D_y = 2$ and $H = 3$ is shown in Figure-1a.

Here compared to the normal network combination of 1-7-1, the NARX network with the combination of 7-7-1 is used, with the input coming from the preceding output values from the network itself. For the present study, Figure 1b shows the neural network combination of 7-7-1 where the input value is one and the other values are taken from memorization of earlier sequences.

Study area

A Datawell waverider buoy was deployed at 23m water depth about a distance of 11.2km off Marmugao Port, west coast of India (Latitude 15° 27.9'N and Longitude 73° 41.0' E) as shown in Fig-2. The wave data collected during February 1996 and January 1997 are used for the present studies. Three hourly significant wave heights (Hs) over the period are considered for the neural network.

Data sets

Four data sets are made from the available significant wave heights, namely (a) three hourly significant wave height (Hs), (b) six hourly average Hs, (c) twelve hourly average Hs and (d) a day average Hs. Deo et al (2001) and Agarwal and Deo (2002) have used 80% of the data for training and remaining data was used for prediction. For the present study 80% of the available data is used as training and remaining 20% is kept for comparison with the neural network forecasted data.

RESULTS AND DISCUSSION

The forecasting of next immediate wave height based on the current and the previous values are carried out using the NARX neural networks. The network structures used in wave forecasting are shown in Figs 1a and 1b. The correlation coefficient will provide the performance of the neural network. The correlation coefficient (CC) of forecast data over the measured values is calculated using the formula:

$$cc = \frac{\sum y_1 y}{(\sum y_1^2 \sum y^2)^{1/2}} \quad (10)$$

where

$$y_1 = Y_1 - Y_1'$$

Y_1 = observed wave height

Y_1' = mean of X

$$y = Y - Y'$$

Y = forecasted wave height

Y' = mean of Y

The time series of the forecasted three hourly Hs (NARX network) with the measured three hourly Hs is shown in Figure 3. It can be seen that the forecasted waves are closely matching with the measured values. The three hourly wave forecasting by NARX network has estimated the CC of 0.95. Figure 3 also shows the correlation for NARX network.

Six hourly averaged significant wave height is forecasted and CC is 0.90 for NARX networks. Twelve hourly (half-day) averaged wave forecasting data yields CC of 0.87 for NARX network. The time series profile is shown in Figure 4. In this also the up and down of wave profile is picked up by the forecast profile. The corresponding correlation is also shown in Fig 4. The day averaged forecasting Hs yields the correlation coefficient values of 0.73 for NARX network.

Comparison of results

Results show that the recurrent network predicts better than the earlier neural network methods. It can also be seen that for even averaged wave data the improved results are obtained. The comparative values of correlation coefficients are shown in Table I. The correlation coefficient for 3 hourly Hs has improved from 0.81 (Deo and Naidu, 1999) to 0.95 (present case).

Table-1 Comparison of correlation coefficients

Forecasting lead time	Deo & Naidu (1999)	Subbarao <i>et al</i> (2001)	Present case
3 hour	0.81	0.93	0.95
6 hour	--	0.88	0.90
12 hour	0.78	0.80	0.87
24 hour	0.71	0.73	0.73

The wave forecasting carried out is the next immediate value based on the present values. For some specific operation planning forecasting in advance is useful and the significant wave height will have signature of variation. The averaged forecasted H_s values for six and twelve hours also give better results. Using this method the H_s can be known in advance up to a period of twelve hours which will be useful for planning operation in the sea, such as like pilot deep sea mining in the sea or undertaking any sub sea installation which might need good sea conditions.

CONCLUSIONS

The forecasting of ocean waves is carried out using backpropagation recurrent neural networks with rprop updating algorithm. The forecasted significant wave heights closely agree with the measured wave heights. In the present work, NARX network with rprop algorithm for three hourly wave forecasting yields correlation coefficient of 0.95, which shows an improvement in wave forecasting carried out by others. The forecasting of waves from the averaged data for various periods is carried out and for 12 hours averaged wave data the correlation coefficient obtained is 0.87. This study shows that the wave forecasting can be efficiently carried out using neural networks.

ACKNOWLEDGEMENT

Authors are thankful to the Director, National Institute of Oceanography, Goa for his encouragement. Authors are thankful to the staff of Ocean Engineering division for providing measured wave data. This article is NIO's Contribution No 4028.

REFERENCES

- Agarwal, J. D. and Deo, M. C., 2002. On-line wave prediction. *Marine Structures*, 15(1), 57-74.
- Deo, M. C. and Naidu, C. S., 1999. Real time wave forecasting using neural networks. *Ocean Engineering*, 26, 191-203.
- Deo, M. C; Jha, A; Chaphekar, A. S. and Ravikant, K., 2001. Neural networks for wave forecasting. *Ocean Engineering*, 28, 889-898.
- Londhe, S. N. and Deo, M. C., 2004. Artificial neural networks for wave propagation. *Journal of Coastal Research*, 20(4), 1061-1069.
- Herbich, J. B., 1990. Handbook of Coastal and Ocean Engineering. vol. 1, Gulf Publishing Company, London.
- Ho, S. L.; Xie, M. and Goh, T. N., 2002. A comparative study of neural network and Box-Jenkins ARIMA modeling in the time series prediction, *Computers and Industrial Engineering*, 42(2-4), 371-375.
- Kermanshahi, B., 1998. Recurrent neural network for forecasting next 10 years loads of nine Japanese utilities, *Neurocomputing*, 23(1-3), 125-133.
- Lin, T.; Home, B. G. and Giles, C. L., 1998. How embedded memory in recurrent neural network architectures helps learning long-term temporal dependencies, *Neural Networks*, 11, 861-868.
- More, A. and Deo, M. C., 2003. Forecasting wind with neural networks, *Marine Structures*, 16 (1), 35-39.
- Principe, J. C; Euliano, N. R. and Lefebvre, W. C., 2000. Neural and Adaptive Systems: Fundamentals through simulations, John Wiley and Sons.
- Riedmiller, M., 1994. Advanced supervised learning in multi-layer perceptrons -From backpropagation to adaptive learning algorithms. *Computers Standards and Interfaces*, Special issue on neural networks, (5).
- Subba Rao, Mandal, S. and Prabakaran, N., 2001. Wave forecasting in near realtime basis by neural network, Proc. *International Conference in Ocean Engineering*, IIT Madras Chennai, 105-108.
- Subba Rao and Mandal, S., 2005. Hindcasting of storm waves using neural network. *Ocean Engineering*, 32 (5-6), 667-684.
- Tveter, D. R., 2000. Backpropagator's Review. <http://www.dontveter.com/bpr/bpr.html>

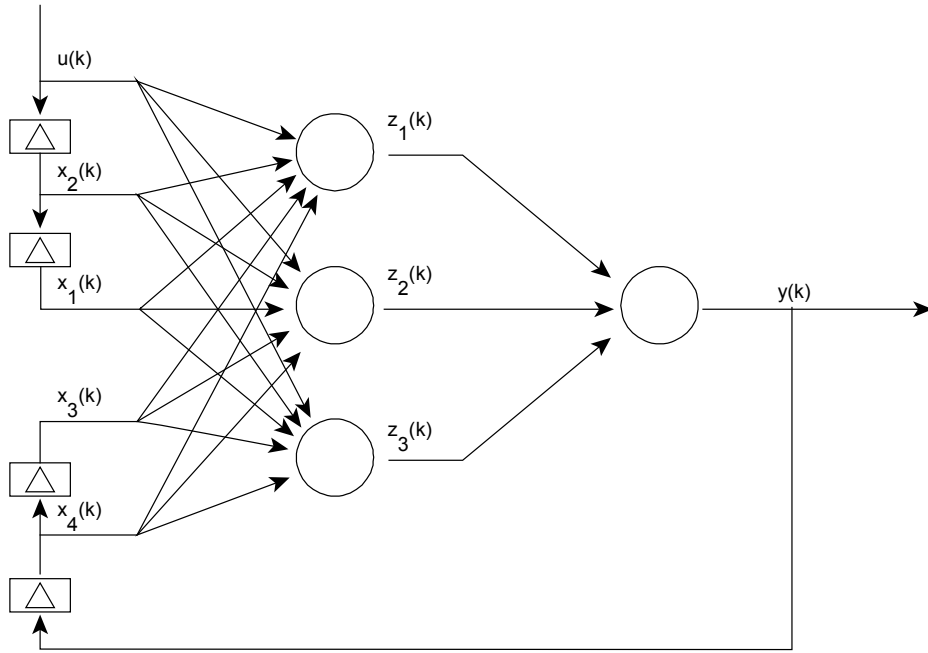
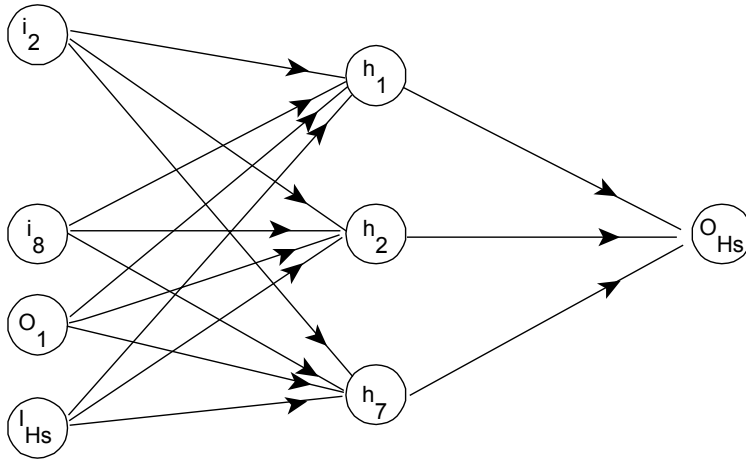


Fig-1a. A NARX network with $D_o=D_y=2$ and $H=3$

INPUT LAYER HIDDEN LAYER OUTPUT LAYER



$i_2 - i_6$ Values taken from earlier output backwards
 O_1 Values taken from immediate output
 I_{Hs} Input H_s value
 O_{Hs} Output H_s value

Fig -1b NARX RECURRENT NETWORK

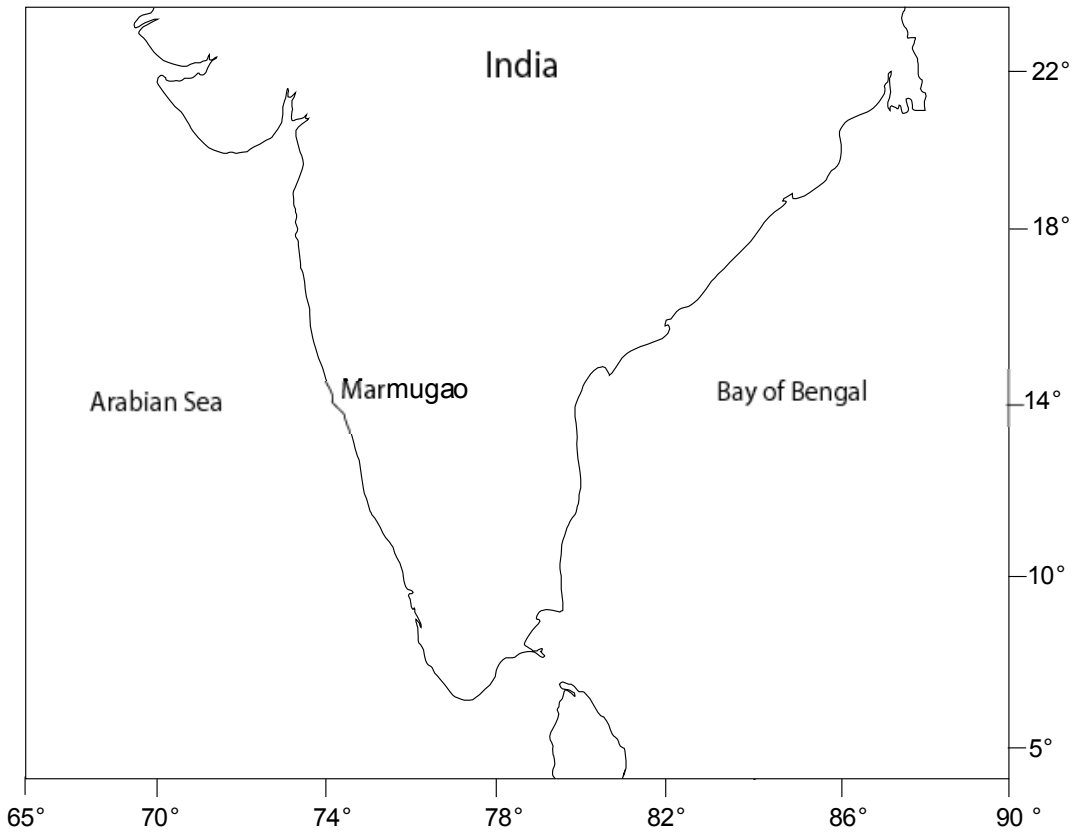


Fig 2 Location of Marmugao, west coast of India.

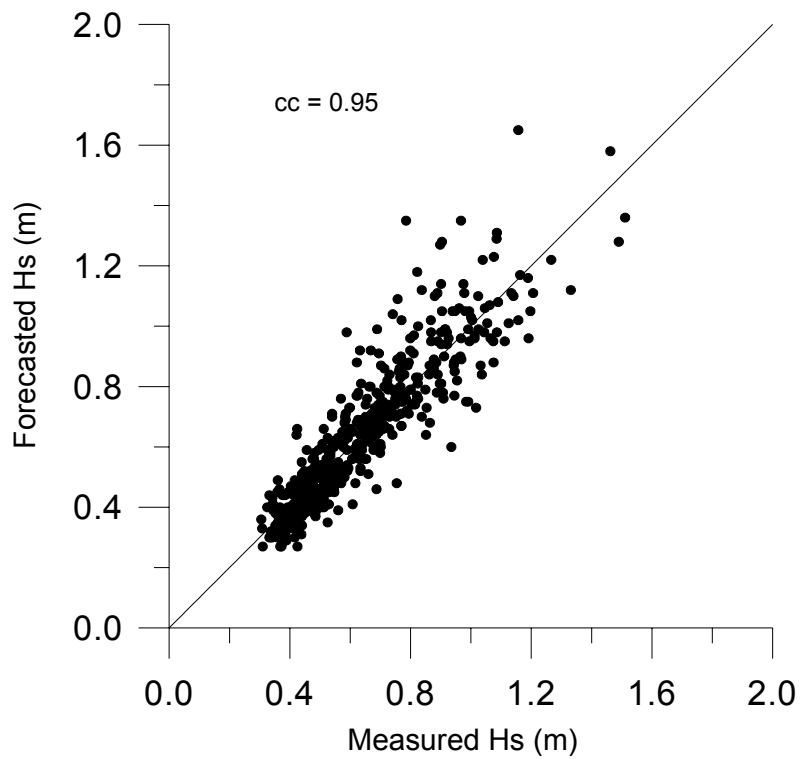
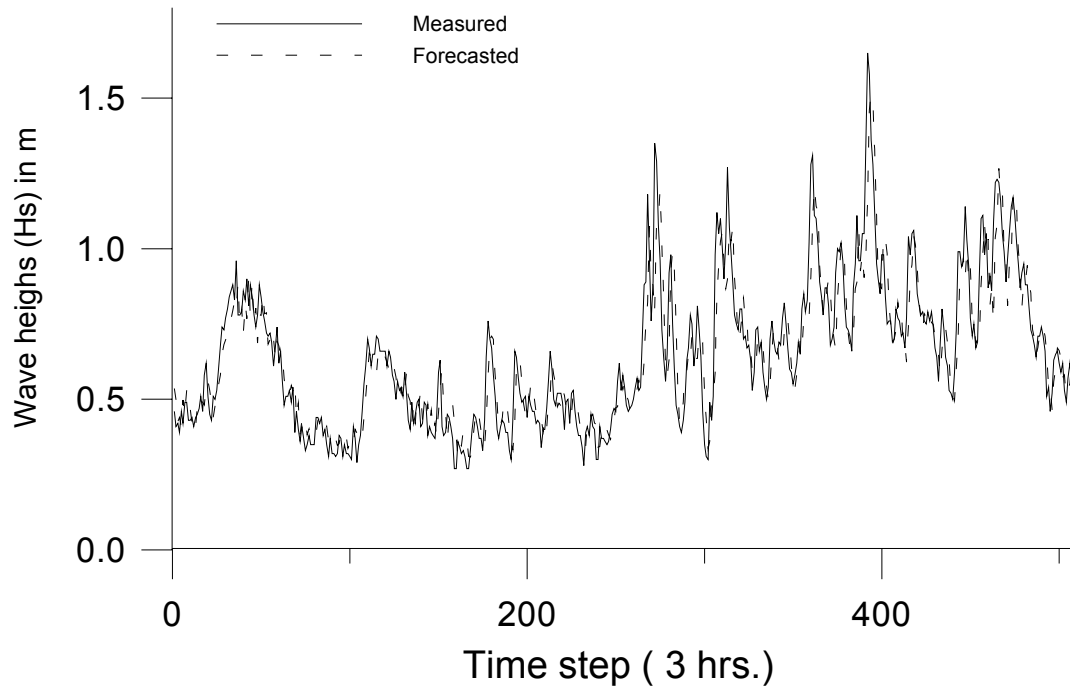


Fig 3. Forecasted (3 hrs) waves off Marmugao by NARX network

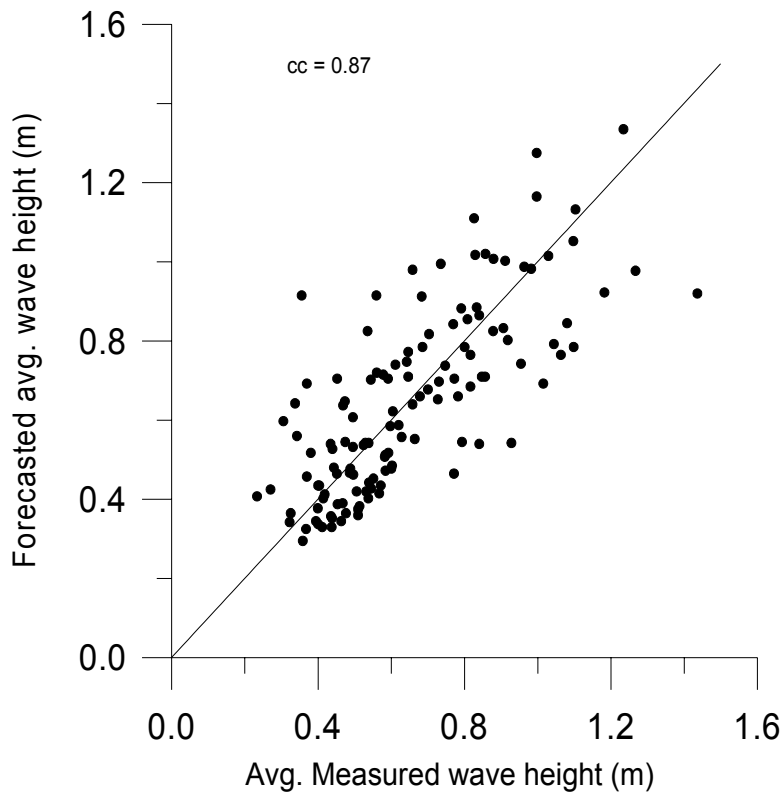
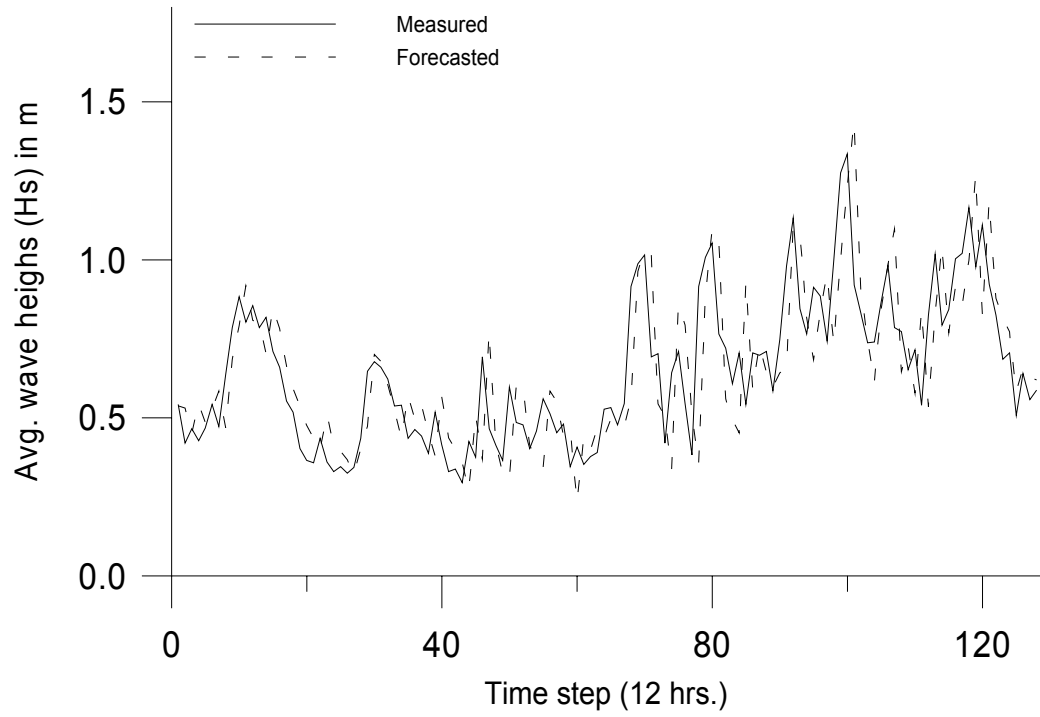


Fig 4 Forecasted (12 hour avg) waves off Marmugao by NARX network