Neural networks for wave forecasting

M.C. Deo [*], A. Jha, A.S. Chaphekar, K. Ravikant

Department of Civil Engineering, Indian Institute of Technology, Bombay 400 076, India

Abstract

The physical process of generation of waves by wind is extremely complex, uncertain and not yet fully understood. Despite a variety of deterministic models presented to predict the heights and periods of waves from the characteristics of the generating wind, a large scope still exists to improve on the existing models or to provide alternatives to them. This paper explores the possibility of employing the relatively recent technique of neural networks for this purpose. A simple 3-layered feed forward type of network is developed to obtain the output of significant wave heights and average wave periods from the input of generating wind speeds. The network is trained with different algorithms and using three sets of data. The results show that an appropriately trained network could provide satisfactory results in open wider areas, in deep water and also when the sampling and prediction interval is large, such as a week. A proper choice of training patterns is found to be crucial in achieving adequate training. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Wave prediction; Neural networks; Wave forecasting; Network training

1. Introduction

The knowledge of heights and periods of oscillatory short waves is essential for almost any engineering activity in the ocean. These waves are generated by the action of wind through pressure as well as shear mechanism. Wind–wave relationships have been explored over a period of five decades in the past by establishing empirical equations and also by numerically solving the equations of wave growth (Kinsman, 1965; Shore Protection Manual, 1984; World Meteorological Organization, 1988). However, the complexity and uncertainty of the wave generation phenomenon is

---

* Corresponding author. Tel.: +91-22-576-7330; fax: +91-22-576-7302.
  E-mail address: mcdeo@civil.iitb.ernet.in (M.C. Deo).

such that despite considerable advances in computational techniques, the solutions obtained are neither exact nor uniformly applicable at all sites and at all times. In this connection Fig. 1 reproduced from World Meteorological Organization (1988) may be seen which gives comparison of measured significant wave heights with their values predicted by an advanced numerical model.

Prediction of waves from the knowledge of generating wind is basically an uncertain and random process and hence difficult to model by using deterministic equations. This makes it ideally suited to neural network modeling since it is primarily aimed at recognition of a random pattern in a given set of input values and use the same to make a prediction of the desired property. Neural networks are useful to model a random input with the corresponding random output and their application does not require knowledge of the underlying physical process as a precondition. Being a relatively recent technique of computation, its applicability in different areas in yet to be proved.

An attempt is made in the present paper to predict the values of significant wave height and average wave period from the corresponding values of wind speed and related characteristics. Exploration of wind–wave relationship at a given location is a common exercise having applications in planning, design and operation of any
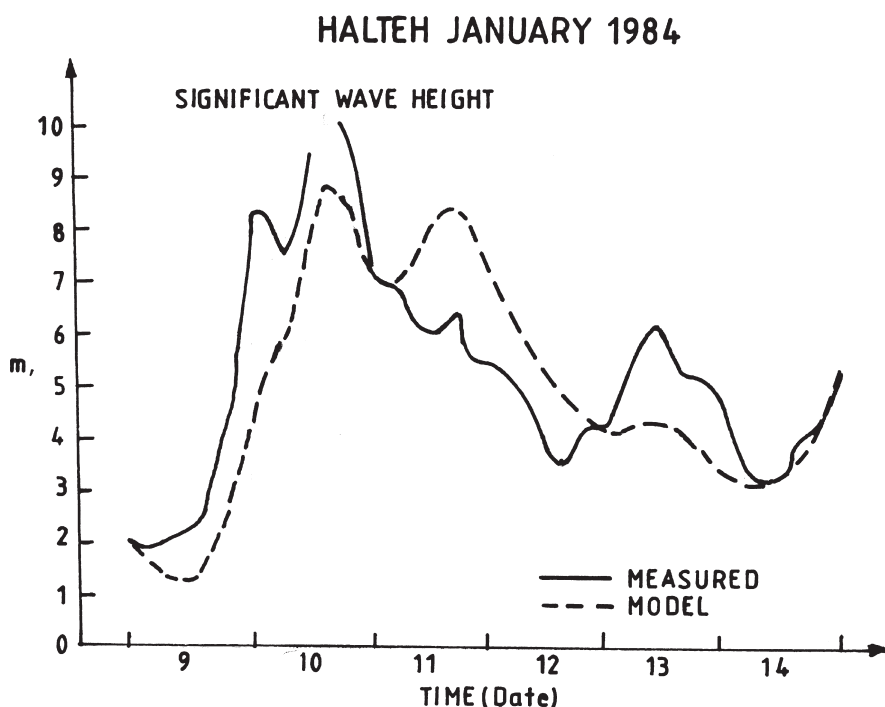

Fig. 1. A sample comparison of measured waves with those predicted by a numerical model. (Site: 65° 02″ N and 07° 33″ E; Period: January, 1984; Model: Norwegian Meteorological Institute. Reference: World Meteorological Organization, 1988).

harbor, coastal or offshore facility. This is done either in preference to or as complementary to the work of actual wave data analysis because the wind measurements are abundantly and cheaply available while wave data are expensive and sparse and also many times more inconvenient to observe. Knowledge of the waves is thus indirectly obtained from that of the generating wind.

Theoretical concepts in neural networks can be seen in many books, such as Kosko (1992) and Wu (1994). Applications of the network to prediction problems in civil engineering include French et al. (1992), Yeh et al. (1993), Kasperkiewicz et al. (1995), Grubert (1995), Thirumalaiah and Deo (1998, 2000) and Deo and Kiran Kumar (1999), and these are related to prediction of rainfall, runoff, concrete strength, estuarine instabilities, river stage and waves, respectively.

2. The neural network

A general sketch of a 3-layered feed forward type of neural network is given in Fig. 2. This is commonly preferred in many engineering applications because of its ability to approximate any nonlinear mathematical dependency structure. Input values are fed to input nodes which transmit them to the hidden layer nodes, each of which sums up the received values from all input nodes, adds a "bias" to this sum and then passes them on through a nonlinear transfer function, like the sigmoid function. The result is fired to the output layer nodes that operate identically to the hidden
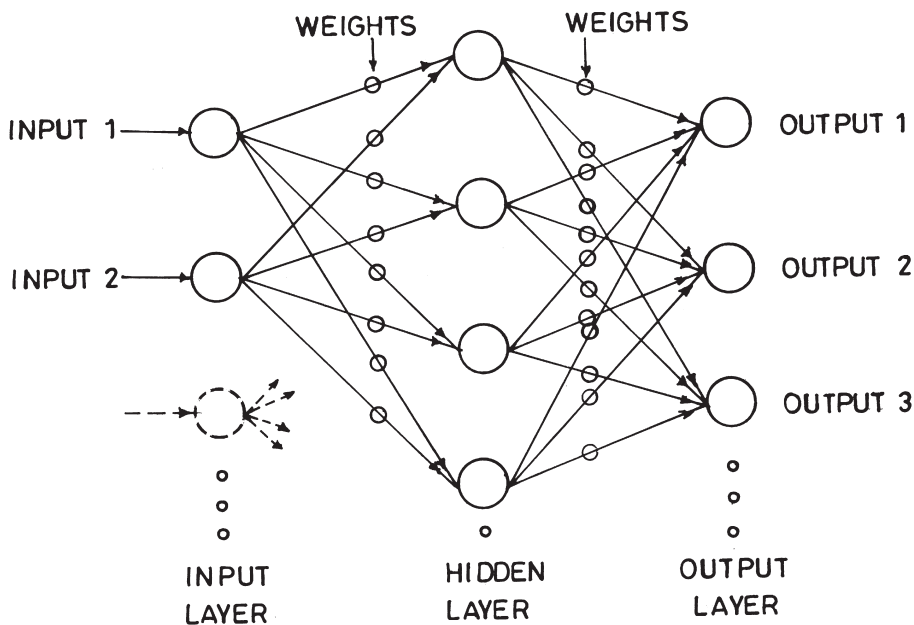

Fig. 2.   A 3-layered feed forward network.

nodes. The resulting transformed output forms the network output. Before its actual application, the network is required to be trained and this is done by using a variety of training algorithms, like Back Propagation, Cascade Correlation and Conjugate Gradient.

All training algorithms are basically aimed at reducing the global error, E, between the network output and the actual observation, as defined below:

$$E = \sum (O_n - O_t)^2 \tag{1}$$

where O_n is the network output at a given output node and O_t is the target output at the same node. The summation is carried out over all output nodes for a given training pattern and then over all training patterns.

In the back propagation method, error minimization is done using the steepest descent or the gradient descent approach wherein the network weights and biases are adjusted by moving a small step in the direction of negative gradient of the error function during each iteration. The iterations are repeated till a specified convergence is reached or until a given number of iterations are over. In order to bring in efficiency in such a back propagation scheme and also to avert undesirable effects arising out of the fixed time step involved in the back propagation iteration, the schemes of cascade correlation and conjugate gradient are used. The cascade correlation algorithm achieves efficiency in training by adding the hidden nodes one-by-one as per the requirement. The weights between the layers of nodes are optimized using the gradient ascent method in which the correlation between output of the hidden unit and the residual error of the network is maximized. In the conjugate gradient approach, the gradient descent is made along a direction which is conjugate or orthogonal to the previous step, unlike the back propagation where it is down the error gradient. The mathematical details of these training schemes can be seen in Fletcher and Reeves, 1964; Rabelo, 1990; Fahlman and Lebiere, 1990; Yeh et al., 1993.

3. Analysis at location 1

The location where wave height and period measurements made by a rider buoy were available was off Karwar in India as shown in Fig. 3 in 16 m deep water. 3-hourly values of significant wave height, H_s, and average zero cross period, T_z, along with those of wind speed and its direction were available from March 1988 to July 1988 and further from December 1988 to May 1989. This enabled the formation of 900 data sets (of simultaneous 3-hourly measurements) or the training patterns where the input side belonged to the wind data while the output side pertained to the wave data. The network training was achieved using 720 patterns while its validation or testing was done with the help of the remaining 180 patterns.

The objective was to obtain 3-hourly values of H_s and T_z based on the input of simultaneous wind speed as well as its fetch and duration. Owing to a tremendous variation in wind as compared to the corresponding wave characteristics, such an
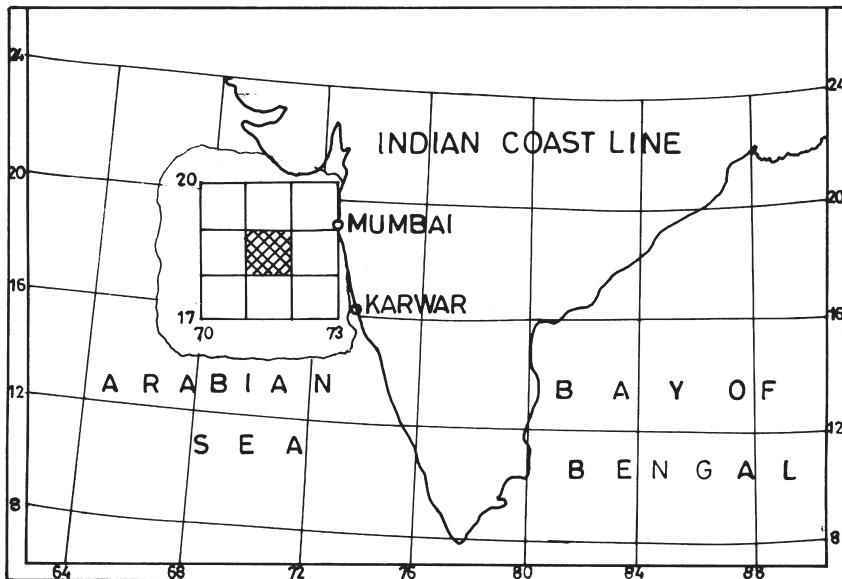
Fig. 3.   Sites of data collection.

input to the network did not yield any satisfactory result, despite following several averaging processes. A sensitivity study showed that consideration of fetch and duration neither helps in achieving training nor in improving accuracy of the output. Hence only wind speed belonging to a few previous time steps was given as input and the fetch and duration were disregarded in a phase-wise manner. Further, the data were separated into monsoon and fair weather season, and the training and testing was done separately. In total, around 50 combinations of feeding data at input level were made. It was found in the end that a network consisting of 2 input nodes (pertaining to wind speeds at the current as well as one previous time step) and 2 output nodes (belonging to forecasted values of H_s and T_z) produced relatively satisfactory values. The number of hidden nodes was fixed as 4 from convergence consideration and by trial in back propagation training while it was internally decided in the cascade correlation based training scheme. Fig. 4 shows time history comparisons between network predicted and the actual observed values of significant wave height and average zero-cross wave period for the non-monsoon season in respect of data earmarked for validation purposes and not considered in the network training. The rising and falling trends in the observed time histories seem to have been picked up by the forecasted series. However, the coefficient of correlation between the measured and the predicted series was as low as 0.52 for wave heights and 0.55 for the wave periods.

With respect to the same data set, the prediction of H_s and T_z values was also made with the help of empirical SMB relationships (Shore Protection Manual, 1984). However this was highly unsatisfactory (Jha, 1997).
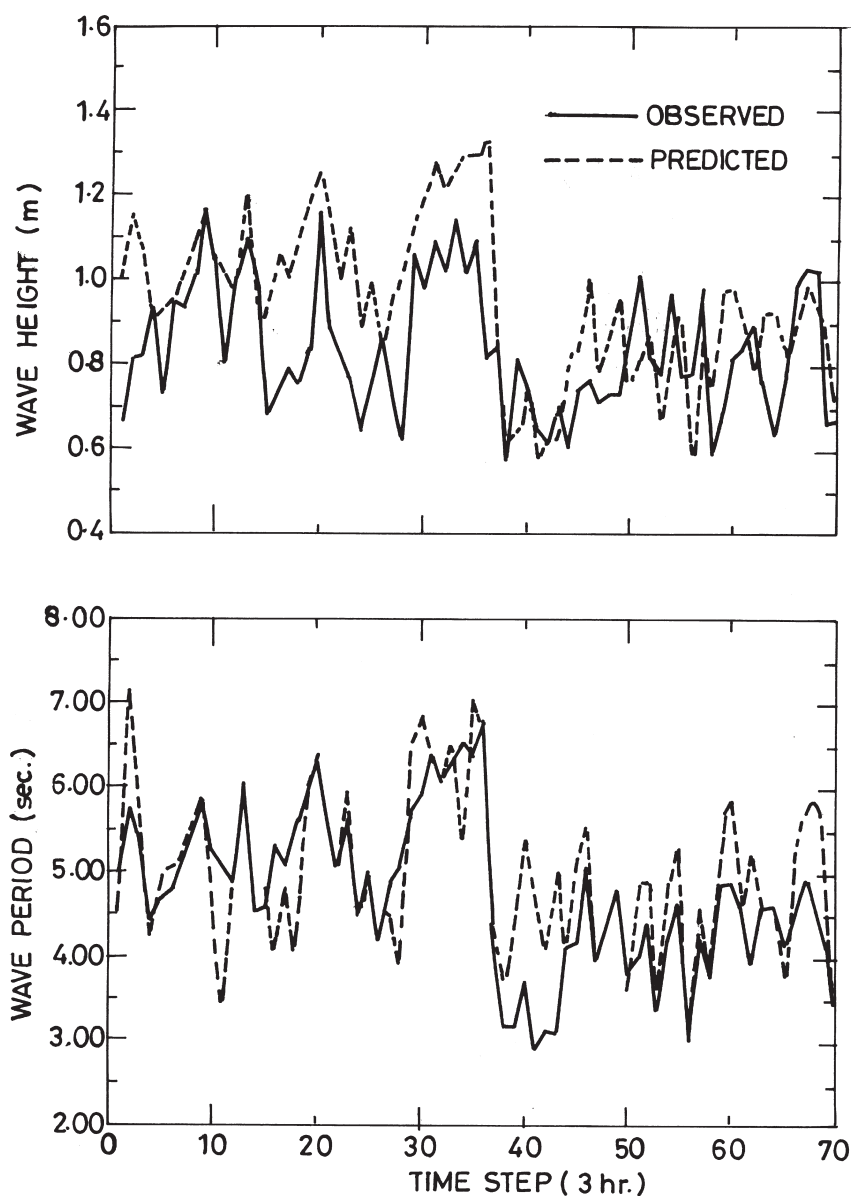
Fig. 4.   Network predicted and actual wave heights and periods (Location 1).

A major reason behind the less satisfactory prediction results could be the fact that the wind was measured at a shore-based station, which was several kilometers away from the location of the wave rider buoy. An attempt was therefore made to consider the case where both the waves as well as the wind are observed at the same station. This is described below.

4. Analysis at location 2

Another set of wave rider buoy based measurements of waves and wind was available near the above mentioned location and in a larger water depth of 75 m. These measurements were in the form of 3-hourly values of significant wave height and average zero-cross wave period, measured simultaneously with the wind speed, over a duration of 18 days from July 1, 1988 onwards. Out of a total of 140 input–output patterns, 80 percent were used for network training while the remainder were utilized for network verification. The previous network, as in the case of location 1, was trained with the help of these training patterns. The validation result is shown in Fig. 5 that gives a time history comparison between the observed as well as predicted values of the significant wave heights. It shows an improved match between the actual observations and the predictions as compared with the previous case. This is reflected in the relatively higher value of the correlation coefficient of 0.68. It is felt that smaller variations in the wind characteristics at this deep water site coupled with the single location measurements of waves and wind had resulted in improved predictions (Chaphekar, 1998).

In view of the lower sample size of the data set under consideration, an attempt was made to make the wave predictions using a larger sample available at a different site. This is described below.
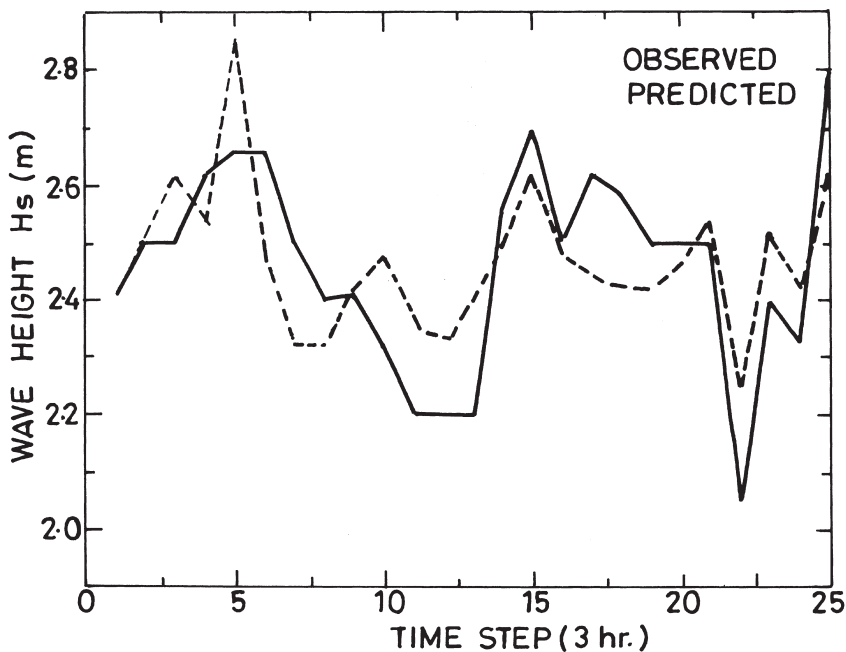

Fig. 5.   Network predicted and actual wave heights (Location 2).

5. Analysis at location 3

Weekly mean wind speed and significant wave height observations made by a radar altimeter onboard satellite TOPEX were available at some locations off the Indian coast. The measurements involved in the current study belonged to the grids of size 1°×1° within the latitudes of 17° to 20° North and the longitudes of 70° to 73° East as shown in Fig. 3. The weekly data covered a period of 42 months from October 1992 onwards. The objective was to develop a network in order to predict weekly mean wave heights and periods when the corresponding wind speed values are given.

After making several combinations of feeding data for training purposes, the network was finally trained using the observations of all surrounding grids and then it was tested with respect to the central grid (shown shaded in Fig. 3) measurements. A combination of 4 input nodes, pertaining to weekly wind speeds in a month, and 4 output nodes, belonging to the corresponding weekly wave heights, was found to be most useful in achieving adequate training in terms of error convergence and testing accuracy. Fig. 6 shows the extent of match between the measured and predicted values of the significant wave height in terms of both time history as well as scatter diagram type of comparisons with respect to the testing data. These results are better in comparison with those obtained for two previous locations as reflected in the higher value of the correlation coefficient, namely, 0.77. This could be due to space as well as time averaging effect in the input series at the current location.

The wave height values for testing part of the data were also calculated using the Pierson–Muskowitz spectrum based relationship (Brebbia and Walker, 1979). However, it yielded extremely low values of the significant wave heights compared to the observed ones (Ravikant, 1999).

Because the underlying phenomenon is highly random and uncertain, a very high accuracy in the predicted values of the wave heights may not be possible. For field and operational applications, the neural networks may provide a simple alternative to the empirical and numerical models, requiring extensive metocean data for their use.

6. Conclusions

Studies on applicability of neural networks to the problem of wave prediction indicated that an appropriately trained network could provide satisfactory results for certain types of predictions. The network is found to be useful in carrying out wind–wave mapping in open wider areas, in deep water and also when the sampling and prediction intervals are large, such as one week. This could be due to an averaging effect over space and time domains.

Short interval wind measurements usually show rapid variations in their values and such an input makes the network training difficult. In the case of short interval data, separate training for both monsoon and fair weather seasons is beneficial.

Unlike the deterministic models, the input of wind fetch and duration does not seem to be necessary in neural networks.
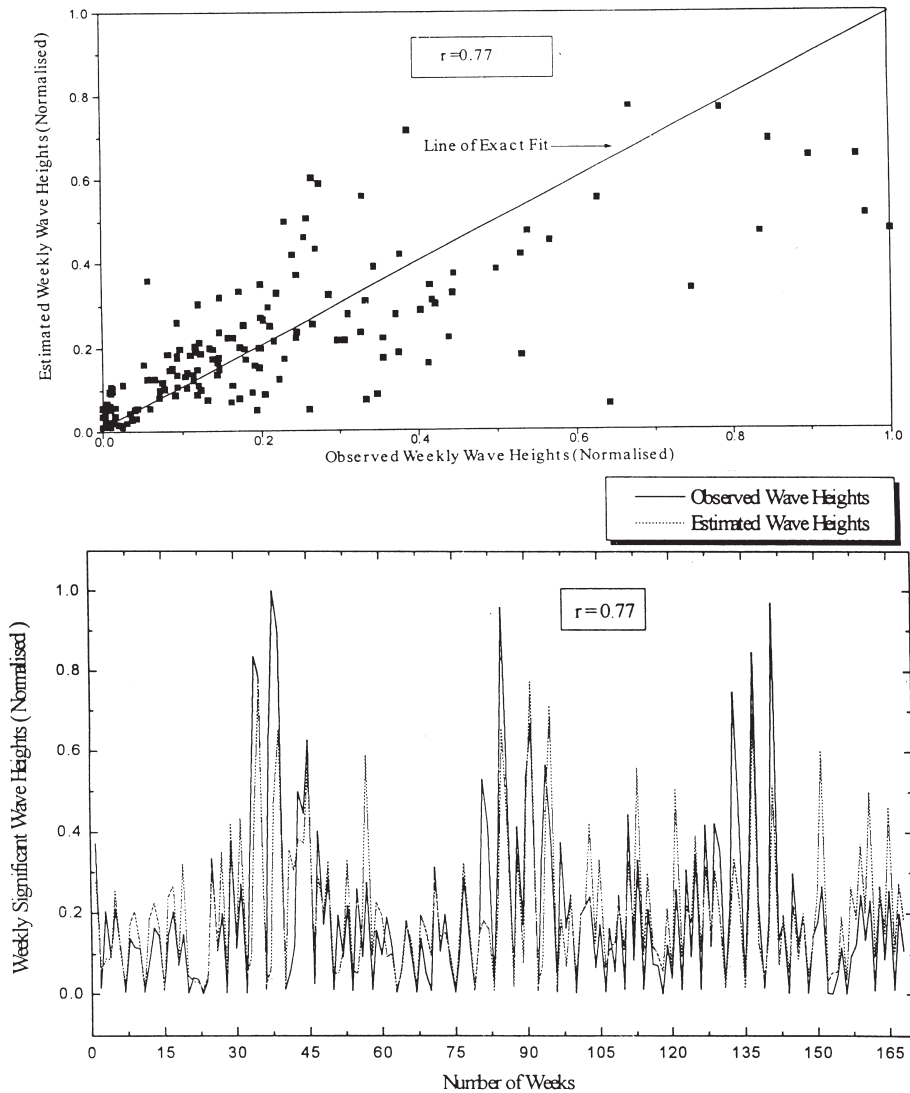
Fig. 6. Network predicted and actual wave heights (Location 3).

An appropriate choice of training patterns is found to be crucial for proper training of the network.

References

Brebbia, C.A., Walker, S., 1979. Dynamic analysis of offshore structures. Newnes–Butterworths, London.
Chaphekar A.S., 1998. Wave forecasting using neural networks, M.Tech. Thesis, Indian Institute of Technology, Bombay.

Deo, M.C., Kiran Kumar, N., 1999. Interpolation of wave heights. Ocean Engineering 27 (9), 907–919.

Fahlman, S.E., Lebiere, C., 1990. The cascade correlation learning architecture. In: Touretsky, D.S. (Ed.), Advances in Neural Information Processing System. Morgan Kaufmann, San Maeto, CA, pp. 524–532.

Fletcher, R., Reeves, C.M., 1964. Function minimization by conjugate gradients. Computer Journal 7 (2), 149–153.

French, M.N., Krajewski, W.F., Cuykendall, R.R., 1992. Rainfall forecasting in space and time using neural networks. Journal of Hydrology 137, 1–31.

Grubert, J.P., 1995. Prediction of estuarine instabilities with artificial neural networks. ASCE Journal of Computing in Civil Engineering 9 (4), 266–274.

Jha A., 1997. Wind–wave mapping using neural networks. M.Tech thesis, Indian Institute of Technology, Bombay.

Kasperkiewiecz, J., Racz, J., Dubrawski, A., 1995. HPC strength prediction using artificial neural networks. ASCE Journal of Computing in Civil Engineering 9 (4), 279–284.

Kinsman, B., 1965. Wind Waves. Prentice Hall, Englewood Cliffs, NJ.

Kosko, B., 1992. Neural Networks and Fuzzy Systems. Prentice Hall, Englewood Cliffs, NJ.

Rabelo, L.C., 1990. A Hybrid Artificial Neural Network and KBES Approach to Flexible Manufacturing System Scheduling. UMI Dissertation Services, Ann Arbor, MI.

Ravikant K., 1999. Estimation of wave heights from wind speeds using neural networks. M.Tech dissertation, Indian Institute of Technology, Bombay.

Shore Protection Manual. Coastal Engineering Research Centre, US Army Corps of Engineers, Washington, DC.

Thirumalaiah, K., Deo, M.C., 1998. River stage forecasting using artificial neural networks. ASCE Journal of Hydrologic Engineering, 3 (1), 26–32.

Thirumalaiah, K., Deo, M.C., 2000. Some studies on hydrological forecasting using neural networks. ASCE Journal of Hydrologic Engineering 5 (2), 180–189.

World Meteorological Organization, Guide to Wave Analysis and Forecasting, WMO no. 702. Secretariat of the World Meteorological Organization, Geneva, Switzerland.

Wu, J.K., 1994. Neural Network and Simulation Methods. Marcel Dekker, New York.

Yeh, Y.C., Kuo, Y.H., Hsu, D.S., 1993. Building KBES for diagnosing PC piles with artificial neural networks. ASCE Journal of Computing in Civil Engineering 7 (1), 71–93.