Available online at www.sciencedirect.com

ScienceDirect

Applied Ocean Research 29 (2007) 72–79

Applied Ocean Research

www.elsevier.com/locate/apor

Real-time wave forecasts off the western Indian coast

P. Jain, M.C. Deo*

Indian Institute of Technology, Bombay, Mumbai 400 076, India

Received 16 February 2007; received in revised form 4 March 2007; accepted 17 May 2007
Available online 27 June 2007

Abstract

The wave observations at three locations off the west coast of India have been analyzed using artificial neural network (ANN) to obtain forecasts of significant wave heights at intervals of 3, 6, 12 and 24 h. The most appropriate training method requiring an input of four observations spread over previous 24 h has been selected after considerable trials. Further, the networks are trained after filling in the missing information. Larger gaps in data are filled in using spatial mapping involving observations at nearby locations, while relatively smaller gaps are accounted for by the statistical technique of multiple regressions in temporal mode. It is found that by doing so the long-interval forecasting is tremendously improved, with corresponding accuracy levels becoming close to those of the short-interval forecasts. If the amount of gaps is restricted to around 2% per year or so it is possible to obtain 12 h ahead forecasts with 0.08 m accuracy on an average and 24 h ahead forecast with a mean accuracy of 0.13 m. However, in harsher environments the prediction accuracy can change.

Keywords: Neural networks; Data gaps; Gap in-filling; Data buoys; Wave forecasting

1. Introduction

The phenomenon of generation of ocean waves from wind depends on a number of atmospheric, meteorological and oceanographic factors and hence becomes very complex to describe mathematically. Operational forecasting of waves was until recently based on the conversion of wind information to waves. In order to effect such conversion empirical models like SMB, and Dartyshire [20] were routinely used in the 1960's and 1970's. Subsequently numerical models like WAve Modeling (WAM) and Simulation of WAves in Nearshore (SWAN) based on the energy balance equation [23,10] became popular in the 1980's and 1990's due to their mathematical rigor and also owing to a large temporal and spatial coverage yielded by them. Since the early 1990's, however, wave rider buoys, deployed at several locations where ocean projects were desired, have become routinely operational e.g. the U.S. National Data Buoy Program, the Italian National Data Buoy program, and the Data Buoy Program of National Institute of Ocean Technology (NIOT) of India. Station-specific forecasts in real time or in an online basis thus started coming into vogue [5,13,14].

Real-time forecasting of waves over a time step of a few hours or days at a specific site is required in operational planning of any engineering activity in the ocean. Such works include facility operation (like the work of a pipe-laying barge or ship routing), recreational activity planning (like wind surfing) and the issuing of warnings to coastal populations and fishermen. While this can be done using numerical models like WAM and SWAN, such models are more useful for forecasting over a large spatial and temporal domain. When point forecasts at a specified location are needed the time series based models like Auto-Regressive (AR), Auto-Regressive Moving Average (ARMA), Auto-Regressive Integrated Moving Average (ARIMA), Kalman Filter and recently the artificial neural network (ANN) should be preferred. The advantages of these time series based schemes over the numerical models include lack of any exogenous data (apart from the wave time series), relatively simple modeling, smaller requirement of computer memory and time, and absence of errors arising out of wind-to-wave conversion. The application of the time series based methods to wave forecasting being relatively new, compared to numerical models, it requires further exploration

* Corresponding address: Indian Institute of Technology, Department of Civil Engineering, Mumbai 400 076, India. Tel.: +91 22 2576 7330; fax: +91 22 2576 7302.
E-mail address: mcdeo@civil.iitb.ac.in (M.C. Deo).

0141-1187/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.apor.2007.05.003

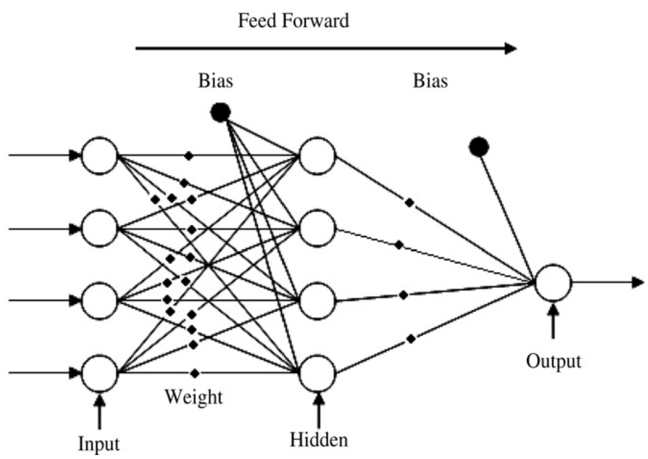

Fig. 1. A typical feed-forward network.

with the aid of the latest techniques made available to computing engineers.

The time series of significant wave heights (H_s) is essentially random in nature. In order to analyze random data a new approach called soft computing has been gaining popularity in the last decade. Soft computing essentially utilizes the tolerance of the real world to uncertainties, imprecision, inaccuracies and partial truth associated with the input information in order to come up with robust solutions. The ideal before soft computing is the human mind. An understanding of the cognition process of the human brain is imitated in a soft approach like ANN. ANN's have been successfully applied to make hydrological predictions in the last 15 years, while their use in oceanic predictions has started only in the last decade. There is scope to exploit the full potential of ANN's in wave analysis and forecasting.

A typical artificial neural network (exemplified in Fig. 1) consists of the interconnection of computational elements called neurons or nodes, each of which basically carries out the task of combining the input, determining its strength by comparing the combination with a bias (or alternatively passing it through a nonlinear transfer function) and firing out the result in proportion to such a strength. Mathematically,

$$O = \frac{1}{[1 + \mathrm{e}^{-S}]} \tag{1}$$

$$S = (x_1 w_1 + x_2 w_2 + x_3 w_3 + \cdots) + \theta \tag{2}$$

where O = output from a neuron; $x_1, x_2, \ldots$ = input values; $w_1, w_2, \ldots$ = weights along the linkages connecting two neurons that indicate strength of the connections; θ = bias value. Eq. (1) indicates a transfer function of sigmoid nature, commonly used, although there are other forms available, like sinusoidal, Gaussian, hyperbolic tangent. (Readers can refer to the text books by Kosko [11], Wassermann [22] and Wu [24] to understand the theoretical details of working of a neural network.) The majority of the applications made in ocean engineering so far have involved a feed-forward type of network as against the feedback or recurrent one. A feed-forward multi-layer network would consist of an input layer, one or more hidden layers and an output layer of neurons, as shown in Fig. 1 referred to earlier.

Applications of neural networks in ocean engineering are mainly oriented towards estimating or predicting values of some particular random variable. Details can be seen in [8]. The ANN applications dealing with real time wave forecasting include works of Deo and Naidu [5], Makarynskyy [13], Makarynskyy et al. [15], Londhe and Panchang [12] and Makarynskyy and Makarynska [14]. These investigators include analysis of wave rider buoy data collected at different locations in the world including the west coast of India, the Atlantic Ocean and the Irish Sea, and the coasts off Tasmania and Portugal as well as the Gulf of Mexico. While all of these works endorse the usefulness of ANN's for online wave forecasting, they also uniformly report almost unacceptably low prediction accuracies at larger intervals like 12 and 24 h. In terms of the error statistics, the correlation coefficient (R) hardly reached a value of 0.70 or so when the prediction of the actual significant wave height (rather than its average) was predicted as in [5,1].

The current study is therefore directed towards improving the real-time forecasts at larger lead times. It involves analysis of time series of significant wave heights (H_s) at three locations along the west coast of India, not reported so far in any of the previous studies. The study would show probably for the first time that the filling in of missing information by well-investigated methods is crucial in realizing accurate forecasts at longer intervals, and further, that this should be accompanied by many trials on imparting training to the selected ANN.

2. The database used

The measurements of the wave height time series considered in this work were made at locations code-named DS1, SW2 and SW4 off Indian west coast, shown in Fig. 2. The observations in the form of 3-hourly H_s values were made by the National Institute of Ocean Technology of India. More information on the underlying data collection can be seen on the website: http://www.niot.res.in/ndbp/bm.pdf [6]. The site DS1 is in deep water (3800 m) whereas the locations SW2 and SW4 are in shallower waters, having 80 m and 24 m water depths, and being off the major coastal cities of Ratnagiri and Mangalore, respectively. The period of data collection varied from 2.5 years (03.02.1998 to 23.08.2000) at SW2 to about 4 years (01.02.98 to 22.02.2002) at DS1 and to around 5.5 years (24.10.98 to 06.05.2004) at SW4.

3. The network development

In this study the usual feed-forward (FF) type of network (Fig. 1) was first considered. The network was trained using variants of the back-propagation (BP) learning scheme like resilient back propagation and quick propagation. In a general back propagation the error between the observed and the predicted values is propagated back using the delta rule, and the unknown connection weights and bias are optimized in an iterative manner. Search-based methods like the conjugate