Prediction of littoral drift with artificial neural networks

A. K. Singh¹, M. C. Deo¹, and V. Sanil Kumar²

¹Department of Civil Engineering, Indian Institute of Technology, Bombay, Mumbai 400 076, India
²Ocean Engineering, National Institute of Oceanography, Dona Paula, Goa 403 004, India

Received: 11 June 2007 – Published in Hydrol. Earth Syst. Sci. Discuss.: 31 July 2007
Revised: 10 December 2007 – Accepted: 18 January 2008 – Published: 21 February 2008

Abstract. The amount of sand moving parallel to a coastline forms a prerequisite for many harbor design projects. Such information is currently obtained through various empirical formulae. Despite so many works in the past an accurate and reliable estimation of the rate of sand drift has still remained as a problem. The current study addresses this issue through the use of artificial neural networks (ANN). Feed forward networks were developed to predict the sand drift from a variety of causative variables. The best network was selected after trying out many alternatives. In order to improve the accuracy further its outcome was used to develop another network. Such simple two-stage training yielded most satisfactory results. An equation combining the network and a non-linear regression is presented for quick field usage. An attempt was made to see how both ANN and statistical regression differ in processing the input information. The network was validated by confirming its consistency with underlying physical process.

1 Introduction

Littoral drift indicates movement of sediments parallel to a coastline caused by the breaking action of waves. Ocean waves attacking the shoreline at an angle produce a current parallel to the coast. Such longshore current is responsible for the longshore movement of the sediment (Komar, 1976). Littoral drift poses severe problems in coastal and harbor operations since it results in siltation of deeper navigation channels due to which ships cannot enter or leave the harbor area. An accurate estimation of the drift is needed in order to know the amount of excavation required so that corresponding budgetary provisions could be made in advance. Unfortunately this is easier said than done because the underlying physical process is too complex to model in the form of mathematical equations – either parametric or differential. Despite this, workable empirical formulae that relate the drift to a set of causative variables are currently in use. They are based on collection of measurements made in the field or on a hydraulic model followed by a curve fitting exercise. The technique of fitting normally employed is non-linear statistical regression. It is well known by now that the soft tools like artificial neural networks (ANN) many times provide better alternative to the statistical methods (see e.g. ASCE Task Committee, 2000; Kambekar and Deo, 2003) and hence a variety of investigators have applied the technique of ANN to solve problems in coastal engineering. These works typically include (a) wave height predictions (Deo and Naidu, 1999; Tsai et al., 2002; Makarynskyy, 2004; Altunkaynak and Özger, 2004; Tolman et al., 2004), (b) evaluating tidal levels and timings of high and low water (Deo and Chaudhari, 1998; Lee, 2004), (c) predicting sea levels (Vaziri, 1997; Cox et al., 2002), (d) forecasting wind speeds (Lee and Jeng, 2002; More and Deo, 2003) (e) establishing estuarine characteristics (Grubert, 1995) and (f) predicting coastal currents (Babovic et al., 2001), (g) other met-ocean parameters (Krasnopolsky et al., 2002; Refaat, 2001). A comprehensive review of ANN applications in related areas can be seen in Jain and Deo (2006). The application of ANN however generally suffers from problems like lack of guarantee of success, arbitrary accuracy, and difficult choices related to training schemes, architectures, learning algorithms, and control parameters. Any new application of the ANN that addresses these issues therefore deserves attention of the potential user community. The current study is directed along this line and discusses an application of the ANN to determine the littoral drift. Novel methods of network training are employed. An equation combining the ANN and the non-linear regression is presented for those desirous of making hand calculations. An attempt is made to see how both ANN and the statistical
regression differ in processing the input information. The consistency of the network with underlying physical process is further checked.

2 Network development

2.1 Training schemes

Most of the previous applications of ANN to water flows involved use of the feed forward network (ASCE Task Committee, 2000). The current study was also based on the same. Both multi-layered perceptron network (MLP) as well as its variant radial basis function (RBF) was used. Training of the MLP was achieved with the help of alternative learning schemes like Conjugate Gradient Polak-Rebiere Update (CGP), Powell Beale Restarts (CGB), Scaled Conjugate Gradient (SCG), One Step Secant algorithm (OSS) and Resilient Backpropagation (RP). Demuth et al. (1998) may be referred to understand details of these training algorithms.

2.2 The database used

The network was trained with the help of field observations. The location belonged to a four-km stretch of the coast off Karwar along the western coast of India. These field measurements were done daily from 5 February 1990 to 31 May 1990 by the National Institute of Oceanography at Goa, India. The sediment load was measured along a cross section of the surf zone at six stations at the same time and at a number of points vertically at each station. In each day, the traps were deployed for 6 h during 07:00 to 13:00 h and the average sediment load per hour was calculated. Two different traps were used to measure the littoral drift rates. Mesh traps having circular openings were used for measuring the suspended load transport and streamer traps were used for measuring the bed-load transport. The opening of the trap was 0.2 m wide, 0.15 m high, and rectangular in shape. The filter cloth mesh opening size was 90 µm and the opening of the mesh trap was 0.034 m. The procedure of Kraus (1987) was used to determine the total sediment transport and this was based on the trapezoidal rule. The measurements of the significant wave height and average zero cross wave period along with the wave direction corresponding to the spectral peak were made with the help of a WAVEC buoy. The breaking wave height and corresponding angle were derived as per the procedure in Skovgaard et al. (1975) and Weishar and Byrne (1978) and also visually confirmed. The width of the surf zone was measured daily using a graduated rope. The average longshore currents were measured daily (in terms of the distance covered in two minutes) using the Rhodanine-B type dye injected at the trap locations. A standard sieve analysis gave the median size distribution. When all the parameters such as wave height, wave period, wave direction, longshore current speed and direction and sediment load at different trap locations along the surf zone, were not collected in a day due to malfunctioning of instruments or due to overtopping of traps, then the data of that day were not used in the analysis. The details on the data collection and the estimation of measured sediment load are presented in Kumar et al. (2003). The tides were predominantly semi-diurnal with an average spring tide of 2 m and neap tide of 0.25 m during the measurement period. The longshore current velocities measured at the trap locations varied from 0.1 m/s to 0.6 m/s with an average value of 0.3 m/s. Table 1 shows ranges of the significant wave height ($H_s$), average zero cross wave period ($T_c$), breaking wave height ($H_b$), breaking angle ($\alpha_b$), surf zone width ($W$) as well as rate of the drift ($Q_s$) along with their mean values and standard deviations involved during the training and testing exercises. The rate of littoral drift was found to be randomly varying with the independent causative variables. This is illustrated in the scatter plot of Fig. 1.

### Table 1. Statistics of the training and testing data set.

<table>
<thead>
<tr>
<th>Data</th>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Range</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>$H_b$ (m)</td>
<td>0.767</td>
<td>0.255</td>
<td>0.33</td>
<td>1.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$T_c$ (s)</td>
<td>5.975</td>
<td>1.520</td>
<td>0.58</td>
<td>9.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$H_b$ (m)</td>
<td>0.888</td>
<td>0.192</td>
<td>0.51</td>
<td>1.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha_b$ (deg)</td>
<td>5.198</td>
<td>1.725</td>
<td>1.14</td>
<td>10.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$W$ (m)</td>
<td>41.067</td>
<td>4.635</td>
<td>30.00</td>
<td>50.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$Q$ (kg/s)</td>
<td>14.969</td>
<td>7.585</td>
<td>1.86</td>
<td>31.96</td>
<td></td>
</tr>
<tr>
<td>Testing</td>
<td>$H_b$ (m)</td>
<td>0.728</td>
<td>0.278</td>
<td>0.42</td>
<td>1.40</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$T_c$ (s)</td>
<td>4.714</td>
<td>0.849</td>
<td>3.50</td>
<td>7.30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$H_b$ (m)</td>
<td>0.834</td>
<td>0.201</td>
<td>0.52</td>
<td>1.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\alpha_b$ (deg)</td>
<td>4.124</td>
<td>6.377</td>
<td>-17.93</td>
<td>10.06</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$W$ (m)</td>
<td>41.048</td>
<td>3.074</td>
<td>35.00</td>
<td>45.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$Q$ (kg/s)</td>
<td>15.650</td>
<td>7.015</td>
<td>5.78</td>
<td>26.43</td>
<td></td>
</tr>
</tbody>
</table>

![Fig. 1. Variation of rate of the drift with breaking wave height.](image-url)
Table 2. Effect of changing input on the testing data set.

<table>
<thead>
<tr>
<th>Input</th>
<th>Training algorithm</th>
<th>r</th>
<th>RMSE (kg/s)</th>
<th>MAE (kg/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( H_s, T_z, H_b, \alpha_b, W )</td>
<td>CGB</td>
<td>0.867</td>
<td>3.664</td>
<td>2.929</td>
</tr>
<tr>
<td>( H_s, T_z, H_b, \alpha_b, V )</td>
<td>RP</td>
<td>0.793</td>
<td>4.443</td>
<td>3.510</td>
</tr>
<tr>
<td>( H_s, T_z, H_b, \alpha_b, d_{50} )</td>
<td>RP</td>
<td>0.831</td>
<td>5.137</td>
<td>4.390</td>
</tr>
</tbody>
</table>

2.3 Network formulation

The phenomenon of littoral drift is influenced by a variety of causative factors—some of which could be of importance while some others may not be so influential in determining the rate of drift. The Shore Protection Manual (Department of the Army, 1984) as well as the Coastal Engineering Manual (Department of the Army, 2002) list these variables as incident significant wave height (\( H_s \)), breaking wave height (\( H_b \)), significant or average zero cross period (\( T_z \)), angle of the wave at the time of breaking (\( \alpha_b \)), width of the wave breaking (surf) zone (\( W \)), sediment size (\( d_{50} \)), and, long-shore current (\( V \)). A network with these parameters as input and the rate of drift, \( Q \) as output was considered. In total 81 input-output patterns were available through the measured data; out of which 75 percent selected randomly were used for training. Such a trained network was tested with the help of remaining 25 percent patterns. A typical plot showing how the training error reduced with increasing number of training epochs is shown in Fig. 2. It is to be noted here that collection of all these parameters simultaneously in the fierce oceanic conditions is a difficult task due to variety of instruments and equipments involved and hence most of the times investigators have to work with a limited sample size. An alternative to this is to resort to laboratory measurements. But this is always associated with problems like scale effects and ignorance of complex real sea conditions.

All of the causative variables listed above may not be equally influential in producing the drift at a given location. A sensitivity analysis of the input was done using the pruning method in which all causative variables were considered and given as input. The network was trained and the testing performance in terms of the various error measures described subsequently was noted. Thereafter each input was omitted one by one and the training and testing was repeated. This exercise revealed that exclusion of any of the parameters of \( H_s, T_z, H_b, \) and \( \alpha_b \) resulted in low performance. However it was also noted that in addition to these if we include \( W \) in preference to V and \( d_{50} \) then the best performance is seen. Table 2 shows resulting performance over the testing pairs (when the best learning algorithm was employed) in terms of the multiple error criteria of the coefficient of correlation (\( r \)), root mean square error (RMSE), and mean absolute error (MAE). It is mentioned that random selections of training and testing pairs were done innumerable number of times till the one that produced the best outcome in terms of the error statistics was arrived at.

From Table 2 it is clear that a network that includes the width of the surf zone, \( W \), in addition to that of \( H_s, T_z, H_b, \) and \( \alpha_b \) gives the best accuracy for testing. However it needs to be mentioned here that this accuracy resulted after resorting to training by alternative schemes like SCG, RP, OSS, CGP, CGB and not by adoption of any one of these randomly. The algorithm of CGB resulted in the best performance. This scheme of training achieves its efficiency using minimum orthogonality between the current and the preceding error gradient. The CGB algorithm performed consistently well for almost all trials.

The number of hidden nodes in case of the above network (input: \( H_s, T_z, H_b, \alpha_b, W \)) was 6. This was decided by trials conducted by increasing the number of hidden nodes one by one and every time noticing performance of the trained network by the error statistics and stopping when such performance did not change with further addition of the hidden nodes. A scatter plot checked the testing performance of this network (Fig. 3), which further qualitatively indicates a satisfactory result.
3 Regression models

In order to check how the neural network performs vis-à-vis the statistical regression three new regression equations (linear multiple (LM) as well as non-linear (NL1 and NL2)) were fitted to the training set of data. The resulting equations respectively are:

\[ Q = -18.7152 - 13.7319H_S - 0.3759T_Z + 39.4895H_b + 0.3455\alpha_b + 0.2340W \]  \hspace{1cm} (1)

\[ Q = 0.28 \times H_S^{-0.7693} \times T_Z^{-0.0704} \times H_b^{2.7935} \times \alpha_b^{0.0005} \times W^{1.1005} \]  \hspace{1cm} (2)

The last 3 rows of Table 3 show the testing performance of these regression-fits vis-à-vis the ANN, which confirms the necessity of employing ANN for this problem in place of the traditional regression (higher \( r \) and lower RMSE and MAE).

The adequately selected network thus yielded a higher level of accuracy compared with the traditional regression models; the major underlying reasons being, model-free estimation procedure and flexibility in the mapping process involved.

3.1 Traditional formulae

The above study discussed how the network performed vis-à-vis those statistical regression models that were newly derived and based on the especially collected data by authors. Traditionally however most of the harbor and coast development works in India are carried out by using an empirical equation known as the Coastal engineering Research Centre (CERC) formula and also by the Walton and Bruno equation. The CERC formula (Shore Protection Manual, Department of the Army, 1984) assumes that the drift (\( Q \)) is proportional to the longshore energy flux (\( P_l \)), i.e.

\[ Q = K \cdot P_l \]  \hspace{1cm} (4)

where \( K \) is a dimensionless constant. The flux, \( P_l \), in turn depends on the sediment characteristics, (like its mass density, \( \rho_s \) and porosity, \( p \)), the breaking wave height \( H_b \) and its angle \( \alpha_b \) with the shore and the wave period \( T \). Specifically

\[ P_l = \frac{1}{64\pi} \left[ (\rho_s - \rho_w) g (1 - p) \right]^{-1} \rho_w g^2 H_b^2 T \sin 2\alpha_b \]  \hspace{1cm} (5)

In the above equation \( \rho_w \) is the mass density of seawater and \( g \) is the acceleration due to gravity.

The Walton and Bruno formula on the other hand relies more on the derived parameters rather than the actual measured ones. The introduction of the surf zone width is also a specialty of this formula. Accordingly the longshore flux (\( P_{ls} \)) is given by:

\[ P_{ls} = 0.008 \left( \frac{v_0}{v_0} \right) \left[ (\rho_s - \rho_w) g (1 - p) \right]^{-1} \rho_w H_b W V \]  \hspace{1cm} (6)
The above equation is based on the assumption that the friction factor is 0.005 and that the theoretical non-dimensional longshore current velocity \((v/v_0)\) is calculated with a mixing parameter of 40%. Equation (6) also uses the actual longshore current speed \(V\).

The drift predicted by the above formulae was compared with its corresponding value actually measured in the field for the testing data conditions. Figure 4 shows the outcome. It clearly indicates that the field observations of the rate of sediment transport are different than the corresponding values suggested by the two traditional formulae. The empirical constants in these equations were earlier derived on the basis of measurements made at those locations where the coastal environment, geomorphology and topographic characteristics were different than the same at the Indian sites. Based on a comparison of the measured values with the commonly used and existing formula, Kumar et al. (2003) had found that the CERC formula and Walton and Bruno formula over predict the longshore sediment transport rate. The difference between the measured and calculated values is attributed due to the use of the empirical formulas developed for the high-energy coast during relatively low wave conditions as the average wave height during the measurement period was 0.8 m. Currently research is on to develop new empirical formulae based on different data sets collected at different parts of the world including the data used in the present study (Bayram et al., 2007).

The unacceptable predictions obtained in the above exercise further confirm the necessity of the ANN or ANN-regression hybrid models (described later) developed in this study.

4 Extended two-stage network

In order to increase accuracy of the network prediction further the outcome of network (architecture: 5-6-1) was given as input to another network with one input node and one output node (and 2 hidden nodes selected after trials as mentioned earlier in the section: Network Formulation) as shown in Fig. 5. Such a two-stage network, where a cause-effect network carries out the basic function approximation in the beginning and the recycler network later does the fine-tuning, was trained with the help of the training pairs and tested with the help of the testing pairs, as earlier. The testing results (Fig. 6) indicated that such a two-stage network performs much better than the earlier single-stage one, with \(r\) as high as 0.913 and RMSE and MAE as low as 3.006 (kg/s) and 2.222 (kg/s) respectively.

The use of two networks made in this way seems to work better than that of an equivalent single network with three or so hidden layers since in the case of the two-stage network a pool of hidden neurons is first allowed to learn independently and later on by capturing the finer details that may have been left out during the basic learning process of the main network. The first network can be expected to have done major regression, but leaving relatively large difference between the actual output and the evaluated one. The second network may build upon such an outcome (or difference) and may learn the target output more efficiently since now the learning process became more simplified.
5 ANN-regression hybrid model

In the light of the fact that the NL1 regression was next in line in terms of the testing performance (Table 3) and that for quick field applications or for making hand calculations an equation would be preferred by the practitioners rather than the complex matrix of trained weight and bias, a new and simple network with one-input node belonging to the littoral drift rate, $Q$, given by Eq. (2), or NL1 model, and one output node belonging to the output value of $Q$ was trained and further tested on the basis of the testing data set. The result was encouraging ($r=0.832$, RMSE=4.349 (kg/s), MAE=3.411), though not as satisfactory as the two-stage ANN, and this is given in an equation form below:

$$Q = f (-0.0555 f (-16.8 Q_{NL1} + 16.8)$$
$$+ 0.5738 f (16.8Q_{NL1} - 8.4)$$
$$+ 0.312 f (16.8Q_{NL1}) - 0.9999)$$

(7)

where, $Q_{NL1}$ = output from Eq. (2), and in general for any $x$, $f(x) = [1 + \exp(-x)]^{-1}$

(8)

The authors also carried out ANN calibration using the n-fold validation in which the total training set was divided into subsets and the training and validation was carried out for varying subset sizes. But this did not provide any further improvement in the testing results. Alternative and probably simpler machine learning methods like the M5 model tree of Quinlan, or instance-based learning methods (for example the k-nearest neighbor) can also be explored, but it is feared that all these methods including the n-fold validation of the ANN may get “saturated” for the given sample size and produce similar results.

6 Consistency in following the physical process

The breaking waves mobilize the sediments that are subsequently moved by the wave induced longshore current. The parameters which influence the sediment transport rate at a location are breaking wave height, wave period, breaker angle, sediment size and the nearshore profile or the surf zone width. The longshore sediment transport in the study area is induced mainly by wave breaking, than due to tide or wind-driven currents, hence the input parameters arrived at after pruning are related to wave breaking. During the study period, the variations in sediment size were relatively small with median grain size varying from 0.15 to 0.2 mm. Hence inclusion of the sediment size did not yield good results.

The ANN developed by training cannot be put into practice unless its performance after training is found to be consistent with the underlying physical process. This may be
viewed as especially necessary when one works with rather limited sample size. A parametric study was therefore performed in which one input variable was varied over its full range keeping all other input quantities as constant. The idea was that the variation drawn in this way must match with the one that can be expected from the known physics of the underlying process. Thus increase in magnitude of the wave height should yield larger drift owing to the increase in the resulting longshore current. This can be clearly seen in Figs. 7a and 8a which indicate what happens in the trained network when significant wave heights and breaking wave heights become higher. Many studies in the past (e.g., Narasimhan and Deo, 1979) have shown that there is only a weak correlation between the wave height $H_s$ and the wave period $T_z$. A given wave height can occur in association with any value of the wave period and thus can be associated with a range of values of the wave period. However as $H_s$ starts increasing from a low value $T_z$ also increases, but this trend continues only up to a certain higher value of $H_s$ after which a reverse trend is noticed. Thus very high $H_s$ values usually correspond to some middle range of $T_z$ values. Higher $H_s$ would mean larger drifts and thus it can be guessed that the maximum drift would correspond to some middle range of $T_z$ values. Figure 9 confirms this. Similarly higher values of the breaking angle, $\alpha_b$, should mean higher longshore current component and hence a larger drift. A clear tendency towards this is not seen in Fig. 10 (although a weak trend may be speculated). This may probably be due to a limited range of $\alpha_b$ values involved during the period of data collection. The developed
network thus can be seen to be generally consistent with the physical process of coastal sediment movement.

In order to understand why the ANN performed better than the regression a parametric variation of Q against all causative variables was studied. Figure 7a and b as well as Fig. 8a and b show examples of how the trained network as well as the regression Eq. (2) processed the input of increasing $H_t$ and $H_b$ values respectively. The relatively low spread of points around the fitted line in case of the regression (Figs. 7b and 8b) indicate that the regression performs rigid approximations with changing input compared with the ANN, due to which its accuracy drops down.

7 Conclusions

Feed forward networks were developed to predict the rate of littoral drift from a variety of causative variables. The use of a two-stage network system in which a regular network trained in the best possible manner first carries out a cause-effect modeling and another one later on fine tunes its outcome resulted in an improved accuracy of predictions.

New regression Eqs. (1–3) derived in this study can also be used to forecast the value of the littoral drift although with less accuracy than the ANN.

An Eq. (7) combining the ANN and the non-linear regression is presented for quick field usage, although it may not predict the drift with as equal accuracy as that of the ANN.

An analysis showing how both ANN and statistical regression process the input is also presented. It is found that the regression performs rigid approximations with changing input compared with the ANN and due to this its accuracy drops.

The developed network was found to be consistent with the underlying physical process and generally followed expected trends in the variables of the drift with an increase in the values of causative parameters.

It is recognized that the findings reported in this paper are based on a limited set of field observations and hence it would be desirable to confirm the same further by collecting additional samples. However the latter is too difficult since it calls for collection of a large number of field parameters simultaneously in the fierce monsoon weather.

References


