BioOne COMPLETE

Estimating Daily Mean Sea Level Heights Using Artificial Neural Networks

Authors: Sertel, E., Cigizoglu, H. K., and Sanli, D. U.

Source: Journal of Coastal Research, 2008(243) : 727-734

Published By: Coastal Education and Research Foundation

URL: https://doi.org/10.2112/06-742.1

BioOne Complete (complete.BioOne.org) is a full-text database of 200 subscribed and open-access titles in the biological, ecological, and environmental sciences published by nonprofit societies, associations, museums, institutions, and presses.

Your use of this PDF, the BioOne Complete website, and all posted and associated content indicates your acceptance of BioOne's Terms of Use, available at www.bioone.org/terms-of-use.

Usage of BioOne Complete content is strictly limited to personal, educational, and non - commercial use. Commercial inquiries or rights and permissions requests should be directed to the individual publisher as copyright holder.

BioOne sees sustainable scholarly publishing as an inherently collaborative enterprise connecting authors, nonprofit publishers, academic institutions, research libraries, and research funders in the common goal of maximizing access to critical research.

Downloaded From: https://bioone.org/journals/Journal-of-Coastal-Research on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Journal of Coastal Research	24	3	727–734	West Palm Beach, Florida	May 2008

Estimating Daily Mean Sea Level Heights Using Artificial Neural Networks

E. Sertel[†], H.K. Cigizoglu[‡], and D.U. Sanli[§]

[†]Istanbul Technical University
Civil Engineering Faculty
Divison of Remote Sensing
Maslak 34469 Istanbul
Turkey
elifsertel@ins.itu.edu.tr

[‡]Istanbul Technical University
Civil Engineering Faculty
Division of Hydraulics
Maslak 34469 Istanbul
Turkey

[§]Bogazici University
Kandilli Observatory and Earthquake Research Institute
Geodesy Division
Cengelkoy 34680 Istanbul
Turkey

ABSTRACT

SERTEL, E.; CIGIZOGLU, H.K., AND SANLI, D.U., 2008. Estimating daily mean sea level heights using artificial neural networks. *Journal of Coastal Research,* 24(3), 727–734. West Palm Beach (Florida), ISSN 0749-0208.

The main purpose of this study is to estimate daily mean sea level heights using five different methods, namely the least squares estimation of sea level model, the multilinear regression (MLR) model, and three artificial neural network (ANN) algorithms. Feed forward back propagation (FFBP), radial basis function (RBF), and generalized regression neural network (GRNN) algorithms were used as ANN algorithms. Each method was applied to a data set to investigate the best method for the estimation of daily mean sea level. The measurements from a single tide gauge at Newlyn, obtained between January 1991 and December 2005, were used in the study. Daily mean sea level estimation was carried out considering the precedent 8-day mean sea level data of the same station, the average and standard deviation of each day for a 15-year period, and 6 monthly and yearly periodicities in tidal variations. Results of the study illustrated that the ANN and MLR models provided comparatively better results than the conventional model used for estimating sea level, least squares estimation. FFBP, RBF, and MLR algorithms produced significantly better results than the GRNN method, and the best performance was obtained using the FFBP algorithm. From the graphs and statistics, it is apparent that neural networks and MLR solution can provide reliable results for estimating daily mean sea level.

ADDITIONAL INDEX WORDS: *Feed forward back propagation, radial basis function, generalized regression neural networks.*

INTRODUCTION

Various geophysical processes within the Earth's system cause changes in global mean sea level: processes such as sea water density variations resulting from temperature and salinity variations, water masses transport among oceans, land and atmosphere, changes in glacial and polar ice sheet mass, terrestrial water storage changes (in soil moisture, snow, and ground water), and atmospheric water vapor variations (CHEN *et al.,* 2000, 2005; DOUGLAS *et al.,* 1990; MINSTER *et al.,* 1999; SCHMITT, 1995). Estimating the mean sea level in coastal zones is important for monitoring and predicting changes in complex marine ecosystems, protecting coastal zone residents, supporting coastal construction plans in these regions, and improving ocean-based technologies (MAKARYNSKYY *et al.,* 2004). It is important to estimate sea level and its variations to study the impact of temporal variation of sea level on coastline and consequently on the engineering works conducted near the coast.

Sea level has been conventionally measured by tide gauges since the mid-nineteenth century. Several types of tide gauges have been designed to date, from a simple tide pole to more sophisticated acoustic and pressure gauges (PUGH, 1987; TOLKATCHEV, 1996). However, recent studies revealed that the records of tide gauges are biased because of land motion that occurs for any reason such as tectonic motion, local subsidence or uplift, and postglacial rebound because they are usually located on the ground (BAKER, 1993; DOUGLAS, 1991; WOODWORTH, SPENCER, and ALCOCK, 1990). The global positioning system (GPS) is frequently used to correct sea level records for vertical land motion, *i.e.,* to obtain true sea levels at tide gauge sites (SANLI and BLEWITT, 2001). This technique provides millimeter-level accuracy; hence yearly sea level rise due to global warming could be monitored at the expected level of precision (NEILAN, VAN SCOY, and WOODWORTH, 1997). A direct measurement of sea level is also possible from space using an altimetry satellite. It is usually used to produce global sea level variations offshore and provides an alternative to that of GPS and tide gauge techniques (NEREM *et al.,* 1997).

Sea level heights are traditionally predicted by least squares estimation using harmonic analysis (DOODSON, 1958; PUGH, 1987). The functional model takes into account the secular variation (long term sea level rise) of the sea level

DOI: 10.2112/06-742.1 received 16 August 2006; accepted in revision 18 January 2007.

This work was funded by the Environment Agency and the Natural Environment Research Council.

Downloaded From: https://bioone.org/journals/Journal-of-Coastal-Research on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Table 1. *Summary statistics of daily mean sea level data set.*

Data	Mean (m)	Median (m)	Standard Deviation (m)	Sample Variance (m^2)	Kurtosis	Skewness	Min (m)	Max (m)	Count
Total	3.193	3.183	0.145	0.021	2.299	0.407	2.043	3.968	5471
Training	3.188	3.177	0.148	0.022	2.324	0.392	2.043	3.968	4383
Testing	3.212	3.203	0.130	0.017	1.948	0.629	2.812	3.944	1088

due to global warming, and a wide spectrum of tidal components as well as atmospheric influences such as pressure, temperature, and wind stress are also included (HANNAH, 1990; VANICEK, 1978). Tidal components are identified by power spectral analysis of the sea level data, and meteorological stations are employed at nearby tide gauges to include the contribution of atmospheric parameters. Refer to LEE (2004) to find reviews of analysis methods employed in sea level prediction.

Artificial neural networks can satisfactorily represent any arbitrary nonlinear function if a sufficient and properly trained neural network is employed. ANNs can find useful relationships between different inputs and outputs. The ANN method has been widely used for multidisciplinary applications such as river flow forecasting, modeling river sediment yield, rainfall-runoff modeling, prediction of the distribution of vegetation online waves (ABRAHART, SEE, and KNEAL, 2001; AGRAWAL and DEO, 2002; CIGIZOGLU, 2003, 2004; CIGIZOGLU and KISI, 2006; HILBERT and OSTENDORF, 2001; MACKAY and ROBINSON, 2000). Recently, ANNs have also been used for the mean sea level studies. ROSKE (1997) used Kohonen networks, one sort of self-organizing neural network, to predict sea levels without using explicit knowledge. MAKARYNSKYY *et al.* (2004) used feed-forward neural networks to predict hourly sea level variations.

In the majority of these studies, the feed-forward backpropagation method (FFBP) was employed to train the neural networks. The performance of the FFBP was found to be superior to conventional statistical and stochastic methods in continuous flow series prediction (CIGIZOGLU, 2003). Though limited, comparison of this method with other ANN algorithms is also available in the literature (CIGIZOGLU, 2005a, 2005b; MASON, PRICE, and TEM'ME, 1996). The FFBP algorithm has some drawbacks such as the local minima problem. In their work, MAIER and DANDY (2000) summarized the methods used in the literature to overcome this problem of training a number of networks, starting with different initial weights, the online training mode used to help the network to escape local minima, the inclusion of the addition of random noise, and the employment of second order (Newton's algorithm, Levenberg-Marquardt algorithm) or global methods (stochastic gradient algorithms, simulated annealing). In the review study of the ASCE TASK COMMITTEE (2000a, 2000b), other ANN methods such as conjugate gradient algorithms, the radial basis function, the cascade correlation algorithm, and recurrent neural networks were briefly explained. The Levenberg-Marquardt algorithm was employed in the FFBP applications included in the presented study.

In this study, three different ANN methods namely (1) least squares estimation method, (2) feed-forward back-propagation radial basis function and generalized regression neural network algorithms, and (3) multilinear regression model were applied to a site-specific mean sea level data set to explore the best method(s) for the estimation of daily mean sea level. Different parameter values for each method were utilized to obtain better performance and accuracy. Results of five methods with different parameter values were presented to figure out the best method and applicability of these methods for daily mean sea level estimation.

DATA ANALYSIS AND FILLING IN THE MISSING DATA

Data used in this study were obtained from the U.K. National Tide Gauge Network. Mean sea level data from the Newlyn tide gauge, collected time interval between January 1, 1991 and December 31, 2005, were used in the research. Newlyn station is located at 50°06′10.8″ N latitude and 5°32′33.9″ W longitude. This station is the oldest station of the U.K. tide gauge network and has been collecting data since 1915. Hourly mean sea level data were collected between 1915 and 1992, and after 1992 the data have been collected for every 15 min. Daily mean sea level heights were derived by averaging the hourly and 15-min data. The data have big gaps in 1984, 1985, and 1986, and many improbable values for several years prior to 1991. To overcome these problems, we selected the 1991–2005 period for the study, and thus we were able to work with continuous and probable data. The data also have missing values for this period, namely between December 15, 1998, and December 27, 1998, and between March 8, 2002, and April 9, 2002. For the interpolation of the missing data in these periods, Equation 1 in the following section was used. After the estimation of unknown parameters in the equation via least square estimation, a common sea level function was formed to interpolate the missing data using the time value of the related data. The total number of data points was 5471; 80% of the data were used for training and 20% of these data were used for testing the artificial neural networks. Summary statistics of training, testing, and total data are given in Table 1.

METHODOLOGY

Least Squares Estimation (LSE) of Sea Level Model

Sea level variation at time t_i is simply described with the following model:

$$SL(t_i) = a_0 + a_1 t_i + \sum_{j=1}^{n} b_j \cos(\omega_j t_i) + \sum_{j=1}^{n} c_j \sin(\omega_j t_i) + R(t_i) \quad (1)$$

Downloaded From: https://bioone.org/journals/Journal-of-Coastal-Research on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

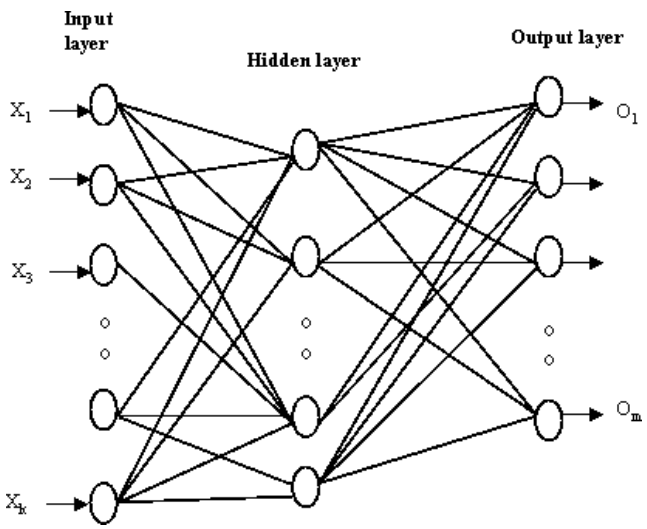


Figure 1. Structure of a feed forward neural network (FFBP).

where a_0 represents mean sea level for a certain time, t_i represents time, a_1 is the sea level rise, b_j and c_j represent the tidal constituents, ω_j is the angular velocity, n is the number of tidal components, and R is the residual term (PUGH, 1987).

Sea level rise (a_1) is the secular component and reveals a possible rise at tide gauges. A wide variety of tidal components need to be removed before the error analysis. Otherwise the remaining component might affect the rate of the estimated sea level trend. Tidal components are determined by the application of Fourier analysis to sea level data and are included in the analysis model. If the frequency of the tide is known, it can be directly incorporated into the least squares model, and the amplitude is solved as an unknown parameter. The linear regression model can be extended by adding as many parameters as possible, considering the wide variety of influences on the sea level. However, adding extra parameters to the model might weaken the least squares solution (DOUGLAS, 1991; SANLI and BLEWITT, 2001).

Feed Forward Back Propagation Method

A FFBP distinguishes itself by the presence of one or more hidden layers, whose computation nodes are correspondingly called hidden neurons of hidden units. A typical feed forward neural network structure is illustrated in Figure 1. The function of hidden neurons is to intervene between the external input and the network output in some useful manner. By adding one or more hidden layers, the network is enabled to extract higher order statistics. In a rather loose sense, the network acquires a global perspective despite its local connectivity because of the extra set of synaptic connections and the extra dimension of NN interconnections (HAYKIN, 1994).

The ability of hidden neurons to extract higher order statistics is particularly valuable when the size of the input layer is large. The source nodes in the input layer of the network supply respective elements of the activation pattern (input vector), which constitute the input signals applied to the neurons (computation nodes) in the second layer (*i.e.,* the first hidden layer). The output signals of the second layer are used as inputs to the third layer, and so on for the rest of the network. Typically, the neurons in each layer of the network have as their inputs the output signals of the preceding layer only. The set of the output signals of the neurons in the output layer of the network constitutes the overall response of the network to the activation patterns applied by the source nodes in the input (first) layer (HAGAN and MENHAJ, 1994). The Levenberg–Marquardt optimization technique was employed for the FFBP method. It is shown that this optimization technique is more robust than the conventional gradient descent technique (CIGIZOGLU and KISI, 2005; HAGAN and MENHAJ, 1994).

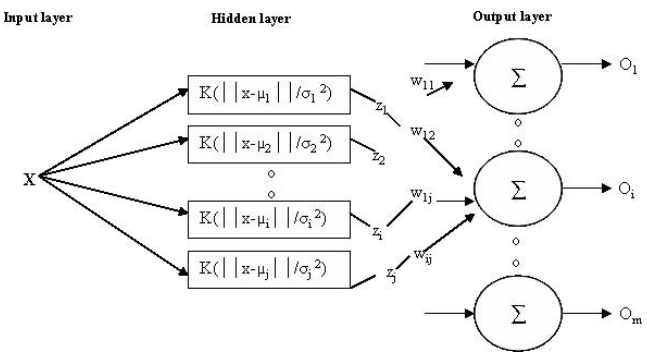


Figure 2. Structure of a radial basis function neural network.

Radial Basis Function (RBF) Networks

RBF networks were introduced into the neural network literature by BROOMHEAD and LOWE (1988). The structure of a radial basis function neural network (RBF) is shown in Figure 2. The RBF network model is motivated by the locally tuned response observed in biological neurons. Neurons with a locally tuned response characteristic can be found in several parts of the nervous system, for example, cells in the visual cortex sensitive to bars oriented in a certain direction or other visual features within a small region of the visual field (POGGIO and GIROSI, 1990). These locally tuned neurons show response characteristics bounded to a small range of the input space. The theoretical basis of the RBF approach lies in the field of interpolation of multivariate functions. The objective of interpolating a set of tuples $(x^s, y^s)_{s=1}^N$ with $x^s \in R^d$ is to find a function $F: R^d \rightarrow R$ with $F(x^s) = y^s$ for all $s = 1, \ldots, N$, where F is a function of a linear space. In the RBF approach the interpolating function F is a linear combination of basis functions

$$F(x) = \sum_{s=1}^{N} w_s \phi(\|x - x^s\|) + p(x) \quad (2)$$

where $\|\cdot\|$ denotes the Euclidean norm, $w_1, \ldots, w_N$ are real numbers, ϕ is a real valued function, and $p \in \Pi_n^d$ is a polynomial of degree at most n (fixed in advance) in d variables. The interpolation problem is to determine the real coefficients $w_1, \ldots, w_N$ and the polynomial term $p =$

Downloaded From: https://bioone.org/journals/Journal-of-Coastal-Research on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

$\Sigma_l^D a_l p_j$, where $p_1, \ldots, p_D$ is the standard basis of Π_n^d and $a_1, \ldots, a_D$ are real coefficients. The interpolation conditions are

$$F(x^s) = y^s, \qquad s = 1, \ldots, N \quad \text{and} \tag{3}$$

$$\sum_{s=1}^{N} w_s p_j(x^s) = 0, \qquad j = 1, \ldots, D \tag{4}$$

The function ϕ is called a radial basis function if the interpolation problem has a unique solution for any choice of data points. In some cases the polynomial term in Equation (2) can be omitted and by combining it with Eq. (3), we obtain

$$\phi w = y \tag{5}$$

where $w = (w_1, \ldots, w_N)$, $y = (y^1, \ldots, y^N)$, and ϕ is an $N \times N$ matrix defined by

$$\phi = [\phi(\|x^k - x^s\|)]_{k,s=1,\ldots,N} \tag{6}$$

Provided the inverse of ϕ exists, the solution w of the interpolation problem can be explicitly calculated and has the form: $w = \phi^{-1}y$. The most popular and widely used radial basis function is the Gaussian basis function

$$\phi(\|x - c\|) = e^{-(\|x-c\|/2\sigma^2)} \tag{7}$$

with a peak at center $c \in R^d$ and deceasing as the distance from the center increases.

The solution of the exact interpolating RBF mapping passes through every data point (x^s, y^s). In the presence of noise, the exact solution of the interpolation problem is typically a function oscillating between the given data points. An additional problem with the exact interpolation procedure is that the number of basis functions is equal to the number of data points, and so calculating the inverse of the $N \times N$ matrix ϕ becomes intractable in practice. The interpretation of the RBF method as an artificial neural network consists of three layers: a layer of input neurons feeding the feature vectors into the network; a hidden layer of RBF neurons, calculating the outcome of the basis functions; and a layer of output neurons, calculating a linear combination of the basis functions (TAURINO *et al.,* 2003). Different numbers of hidden layer neurons and spread constants were tried in the study.

Generalized Regression Neural Networks

A schematic of the GRNN is shown in Figure 4. The basics of the GRNN can be obtained in the literature (SPECHT, 1991). The GRNN consists of four layers: input layer, pattern layer, summation layer, and output layer. The number of input units in the first layer is equal to the total number of parameters, including from one to six previous daily flows. The first layer is fully connected to the second pattern layer, where each unit represents a training pattern and its output is a measure of the distance of the input from the stored patterns. Each pattern layer unit is connected to the two neurons in the summation layer: the S-summation neuron and the D-summation neuron. The S-summation neuron computes the sum of the weighted outputs of the pattern layer

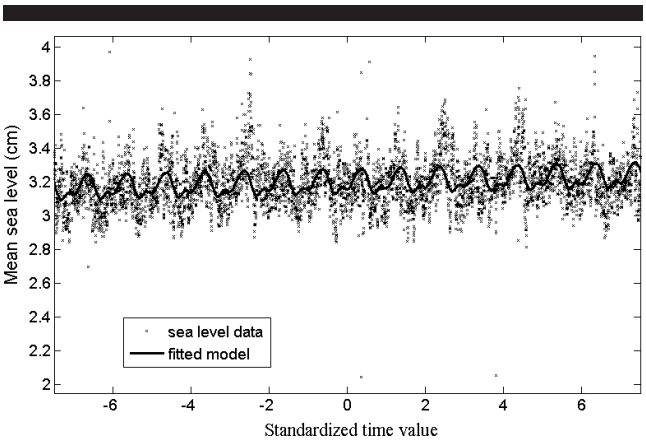


Figure 3. Daily mean sea level estimation by the LSE.

while the D-summation neuron calculates the unweighted outputs of the pattern neurons. The connection weight between the ith neuron in the pattern layer and the S-summation neuron is y_i, the target output value corresponding to the ith input pattern. For the D-summation neuron, the connection weight is unity. The output layer merely divides the output of each S-summation neuron by that of each D-summation neuron, yielding the predicted value to an unknown input vector x as

$$\hat{y}_i(x) = \frac{\sum_{i=1}^{n} y_i \exp[-D(x, x_i)]}{\sum_{i=1}^{n} \exp[-D(x, x_i)]} \tag{8}$$

where n indicates the number of training patterns, and the Gaussian D function in (8) is defined as

$$D(x, x_i) = \sum_{j=1}^{p} \left(\frac{x_j - x_{ij}}{\zeta} \right)^2 \tag{9}$$

where p indicates the number of elements of an input vector. x_j and x_{ij} represent the jth element of x and x_i, respectively. ζ is generally referred to as the spread factor, whose optimal value is often determined experimentally (KIM, KIM, and KIM, 2003). The larger that spread is, the smoother the function approximation will be. Too large a spread means a lot of neurons will be required to fit a fast changing function. Too small a spread means many neurons will be required to fit a smooth function, and the network may not be generalized well. In this study, different spreads were tried to find the best one that gave the minimum MSE for a given problem. Figure 4 illustrates the general structure of GRNNs.

Multiple Linear Regression Model

Multiple linear regression refers to regression applications in which there are two or more independent variables, $x_1, x_2, \ldots, x_k$. A multiple linear regression model with k independent variables has the following equation:

$$y = a_0 + a_1x_1 + a_2x_2 + \ldots + a_kx_k + \varepsilon \tag{10}$$

Downloaded From: https://bioone.org/journals/Journal-of-Coastal-Research on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

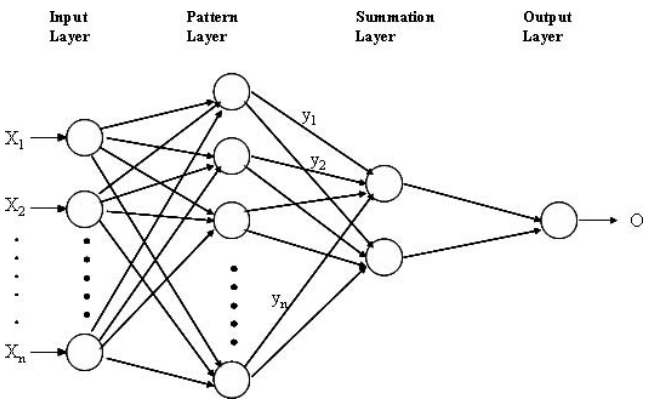


Figure 4. Structure of a GRNN.

The ε is a random variable having mean 0 and variance σ^2. A prediction equation for this model fitted to data can be written as follows:

$$y = a_0 + a_1x_1 + a_2x_2 + \ldots + a_kx_k \quad (11)$$

METHOD APPLICATION

Three MATLAB codes were written: (1) for the feed forward back propagation algorithm; (2) for the radial basis function networks; and (3) for the generalized regression algorithm. LSE and MLR methods were applied using the MATLAB built-in functions. The autocorrelations for the stations were calculated to find the appropriate number of input neurons for ANN analysis. The autocorrelogram shows that the correlation values after 8 days show a significant decrease, and the values until 8 days are considered in the input layer (Table 2). In total, 12 input neurons were used for the ANN applications; 8 neurons belonging to mean sea level data of previous 8 days, 1 neuron belonging to average mean sea level values of each day for 15-yr period, 1 neuron for the standard deviations of each day for 15-yr period, and 2 neurons for periodicities in tidal variations, namely 6-mo and 1-yr periodicities.

After preparing 12 input neurons for further utilization of ANNs to daily mean sea level data, two stages were conducted. The first stage was training of the neural networks, which was performed using 80% of the input data. This stage comprised the presentation of daily mean sea level data describing the input and output to the network and obtaining the interconnection weights. After the complement of the training stage, the formed networks were applied to testing data to analyze the accuracy and performance of suggested ANN approaches. Assessment of the appropriate neural network architecture is an important issue because the network topology directly affects its computational complexity and its generalization capacity (CIGIZOGLU and ALP, 2006). Several network architectures were designed for each method to obtain more accurate and efficient results.

The performances of LSE, FFBP, RBF, GRNN, and MLR methods were compared for estimation of daily mean sea level in Newlyn Tide Gauge Station. The periods selected for training and testing stages were January 1, 1991–December 31, 2002 and January 1, 2003–December 31, 2005, respectively. Root mean square error (RMSE) values and R^2 values between observed and estimated daily mean sea level were used to evaluate the performance of each method.

Table 2. *Autocorrelogram of data.*

	Day $(t-1)$	Day $(t-2)$	Day $(t-3)$	Day $(t-4)$	Day $(t-5)$	Day $(t-6)$	Day $(t-7)$	Day $(t-8)$
Day $(t-1)$	1,000	0,815	0,815	0,815	0,815	0,815	0,815	0,815
Day $(t-2)$	0,815	1,000	0,815	0,815	0,815	0,815	0,815	0,815
Day $(t-3)$	0,676	0,815	1,000	0,814	0,814	0,814	0,814	0,814
Day $(t-4)$	0,572	0,676	0,814	1,000	0,815	0,815	0,815	0,815
Day $(t-5)$	0,462	0,572	0,676	0,815	1,000	0,815	0,815	0,815
Day $(t-6)$	0,364	0,462	0,572	0,676	0,815	1,000	0,814	0,814
Day $(t-7)$	0,291	0,364	0,461	0,572	0,676	0,814	1,000	0,814
Day $(t-8)$	0,241	0,290	0,364	0,461	0,571	0,676	0,814	1,000

LSE Results of Sea Level Model

In this research, a discrete Fourier transform was applied to the data to find out the periodicities and model the tidal variations using Equation 1. Application of a discrete Fourier transform to the data set denoted that there were 6-mo and 1-yr periodicities available in the data. Using calculated periodicities, tide gauge measurements, and time knowns, we formed sea level equations and estimated the unknown parameters using least square adjustment method. The results of the model are shown in Figure 3; the fitted model has an R^2 of 0.19 and an RMSE of 0.130 m. The x axis of the figure shows the standardized time values where -6 and 6 represent July 1, 1992, and July 1, 2004, respectively.

As seen from Figure 5, the daily sea level record is very noisy because of tidal energy and could only be explained to some degree with the harmonic model adopted here. In prac-

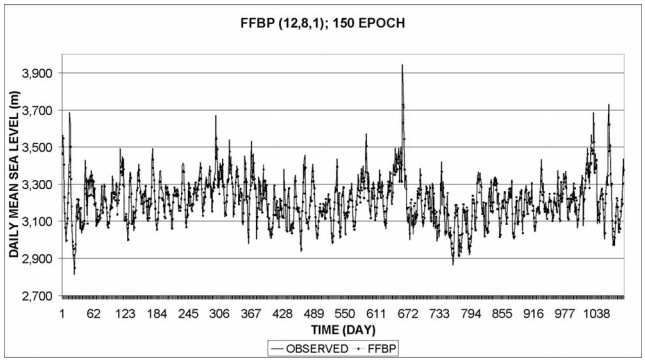


Figure 5. Daily mean sea level estimation by the FFBP method.

Downloaded From: https://bioone.org/journals/Journal-of-Coastal-Research on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Table 3. *Daily mean sea level estimation by FFBP (testing period).*

ANN Configuration	Nodes in Hidden Layer	Epoch Number	RMSE (m/d)	R^2
FFBP (12, 7,1)	7	150	0.072	0.69
FFBP (12, 7,1)	7	150	0.073	0.69
FFBP (12, 7,1)	7	200	0.078	0.65
FFBP (12, 7,1)	7	125	0.071	0.70
FFBP (12, 7,1)	7	100	0.072	0.70
FFBP (12, 8,1)	8	125	0.073	0.69
FFBP (12, 8,1)	8	150	0.071	0.71
FFBP (12, 8,1)	8	200	0.073	0.69

tice, daily mean tide levels are calculated by averaging the heights of low water and high water each day. On the other hand, estimation of secular sea level rise necessitates the use of monthly or yearly sea level data (DOUGLAS 1991; SANLI and BLEWITT, 2001). The monthly and yearly means are simply calculated by averaging hourly heights over a period of a month and a year, or a low-pass filter is applied to hourly levels to remove tidal energy at diurnal and higher frequencies from sea level elevations (PUGH, 1987).

FFBP Results

Various epoch and hidden layer node number combinations were tested for the FFBP algorithm. The first five FFBP combinations included seven hidden layer nodes and the last three included eight hidden layer nodes as shown in Table 3. A FBBP configuration denoted, for example, as FFBP (12, 7, 1) in Table 3 involves 12 input nodes, 7 hidden nodes, and a unique output value, respectively. The output node corresponds to the daily mean sea level value at day t. A network structure with one hidden layer having 8 nodes, and 12 input nodes in the input layer operated with 150 epochs; the FFBP (12, 8, 1)-150 epoch provided the best performance criteria with the lowest RMSE and the highest R^2 for the testing period (RMSE = 0.071 m/d and $R^2 = 0.71$). In all configurations the input layer covered daily mean sea level data within a period of 8 days (days: $t - 1$, $t - 2$, $t - 3$, $t - 4$, $t - 5$, $t - 6$, $t - 7$, and $t - 8$) to estimate the daily mean sea level at day t.

The estimations for the testing period were compared with the observed daily mean sea level values (Figure 5). The observed mean sea level values and estimated values show great consistency with centimeter-level RMSE values and about a 70% R^2 value, which are significantly better than the conventional LSE method. This algorithm can run fast with a large number of data sets. There is a drawback for FFBP simulations encountered during the research; different forecast values were obtained within the same network design after each simulation because of the difference in initial random weight assignment in the beginning of each training. To overcome this problem, simulations were conducted several times, even with the same network structure, until the best performance was obtained. The first and second network structures in Table 3 are an example of this situation.

Table 4. *Daily mean sea level estimation by RBF (testing period).*

ANN Configuration	Speed Parameter (s)	Epoch Number	RMSE (m/day)	R^2
RBF (12, 0.55,1)	0.550	45	0.075	0.68
RBF (12, 0.35,1)	0.350	45	0.084	0.60
RBF (12, 0.65,1)	0.650	45	0.073	0.69
RBF (12, 0.70,1)	0.700	50	0.074	0.69
RBF (12, 0.65,1)	0.650	100	0.073	0.70
RBF (12, 0.65,1)	0.650	200	0.077	0.66

RBF Results

The same study was carried out by the RBF method. Several epoch number and spread parameter combinations were analyzed for the RBF algorithm. A RBF configuration denoted, for example, as RBF (12, 0.55, 1) in Table 4 involves 12 input nodes, 0.55 as spread parameter, and a unique output value. A network architecture with 0.65 spread parameter and 100 epoch number provided the best performance result with the lowest RMSE and the highest R^2 for the testing period (RMSE = 0.073 m/day, $R^2 = 0.70$).

The estimations obtained from the RBF algorithm were compared with the observed daily mean sea level values for the testing period (Figure 6). The observed mean sea level values and estimated values show great consistency with the centimeter level RMSE values and about 70% R^2, as in the FFBP algorithm. This method also has better performance than the LSE method. There are two advantages of this method compared with the previous one. This algorithm can run faster than the FFBP with a large number of data sets. One can obtain same forecast values for the same network architecture, unlike in FFBP. However, there is a disadvantage to RBF simulations encountered during the research; it becomes quite slow, in some cases almost inoperable, when handling a huge amount of data such as 50 years of daily measurements.

GRNN Results

The same data were used to evaluate the GRNN method. Various smoothing parameters were used to obtain the best

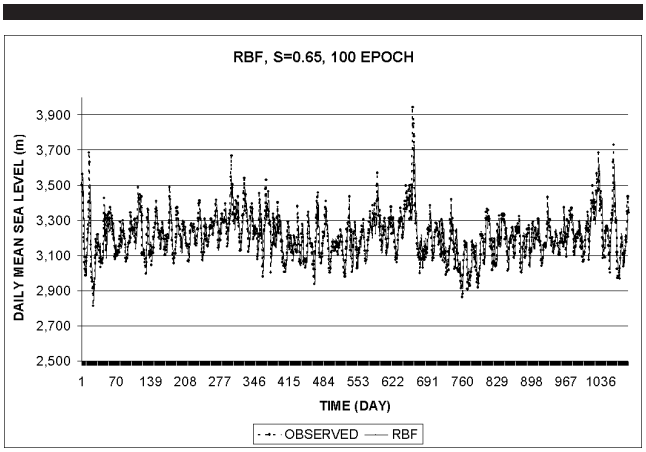


Figure 6. Daily mean sea level estimation by the RBF method.

Downloaded From: https://bioone.org/journals/Journal-of-Coastal-Research on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Table 5. *Daily mean sea level estimation by GRNN (testing period).*

ANN Configuration	Smoothing Parameter (σ)	RMSE (m/day)	R^2
GRNN (12, 0.015,1)	0.015	0.105	0.40
GRNN (12, 0.35,1)	0.350	0.126	0.37
GRNN (12, 0.035,1)	0.035	0.094	0.52
GRNN (12, 0.025,1)	0.025	0.096	0.48
GRNN (12, 0.045,1)	0.045	0.096	0.50

performance results for the GRNN algorithm. A GRNN configuration denoted, for example, as GRNN (12, 0.45, 1) in Table 5 involves 12 input nodes, 0.45 as smoothing parameter, and a unique output value. A network architecture with a 0.035 smoothing parameter presented the best performance result, with the lowest RMSE and the highest R^2, for the testing period (RMSE = 0.094 m/d, R^2 = 0.52). The performance of the GRNN algorithm increased with a decreasing smoothing parameter until 0.035; however, it started to decrease again when the smoothing parameter was smaller than 0.035.

The GRNN algorithm provided better results than the LSE method; however, the estimations obtained from the RBF and FFBP algorithms gave better results than the GRNN algorithm. The correlation between the observed mean sea level values and the estimated values was about 45%. Although the GRNN algorithm could show better performances in some applications like suspended sediment estimations; in our case, while dealing with daily mean sea level data, FFBP and RBF solutions were comparatively better than the GRNN. The disadvantage of the GRNN algorithm is that it worked slower than the other two algorithms.

MLR Results

Sea level measurements of the previous 8 days were used as independent variables to determine mean sea level with the multiple linear regression model. The MLR equation obtained is calculated as follows:

$$y = 0.133 + 0.051X_{(t-8)} + 0.001X_{(t-7)} - 0.023X_{(t-6)} - 0.038X_{(t-5)} - 0.034X_{(t-4)} + 0.083X_{(t-3)} + 0.044X_{(t-2)} + 0.749X_{(t-1)} \quad (12)$$

where $X_{(t-8)}, \ldots, X_{(t-1)}$ represents the measurements of the previous 8 days, respectively, and y is the predicted value of mean sea level and can be represented as $X_{(t)}$. The fitted model has an RMSE of 0.73 m and an R^2 of 0.68. The performance of MLR is similar to FFBP and RBF, and is fairly finer than GRNN and LSE.

DISCUSSION AND CONCLUSION

The knowledge of near-coast sea level variations is an important issue to deal with for several applications, *e.g.,* protecting coastal zone residents, supporting coastal construction plans, and conducting safe navigation. Estimation of the sea level value for a given period could facilitate sea level–dependent studies. In this research, daily mean sea level observations received from Newlyn Tide Gauge Station were used to develop and validate the ANN and MLR methodology for sea level estimations. Three different ANN methods, FFBP, RBF and GRNN, were employed to investigate the applicability of different ANN algorithms for estimating daily mean sea level and examine the performance of each model. Results obtained from ANN and MLR were compared with the conventional LSE of the sea level model. Both ANN and MLR gave comparatively better results than conventional method.

The results of the research illustrated that different ANN algorithms and MLR can be applied to a sea level data set to make successful estimations. While the performance of the conventional method is 0.1312 and 0.1918 for RMSE and R^2, respectively, estimation of daily mean sea level ANN and MLR methods is fulfilled within centimeter accuracy and about 70% R^2 value. The best result among all of the methods was obtained using the FFBP algorithm with a 0.071 m value for RMSE and a 0.71 value for R^2.

The MLR model gave similar results to the FFBP and RBF algorithms; however, it gave significantly better results than the GRNN method.

Each ANN method used in the study has its own pros and cons. FFBP and RBF networks gave better results when compared with the GRNN algorithm. Although the GRNN method can be successfully applied to different datasets from different disciplines, with daily mean sea level data it had the lowest computation speed and the worst estimation results collating with the other two methods. FFBP and RBF had similar satisfactory results for the estimation of mean sea level. As previously mentioned, FFBP method performances are sensitive to the randomly assigned initial weight values; therefore, one can get different estimation and performance values for the same network architecture. This problem, however, was not faced in RBF and GRNN simulations. ANN's main advantage is its ability to model nonlinear processes of the system without any *a priori* assumptions about the nature of the generating processes.

ACKNOWLEDGMENTS

The data were supplied by the British Oceanographic Data Centre as part of the function of the National Tidal & Sea Level facility, hosted by the Proudman Oceanographic Laboratory.

LITERATURE CITED

Abrahart, R.J.; See, L., and Kneal, P.E., 2001. Investigating the role of saliency analysis with neural network rainfall-runoff model. *Computers & Geosciences,* 27, 921–928.

Agrawal, J.D. and Deo, M.C., 2002. On-line wave prediction. *Marine Structures,* 15, 57–74.

ASce Task Committee, 2000a. Artificial neural networks in hydrology I. *Journal of Hydrologic Engineering,* ASCE, 5(2), 115–123.

ASce Task Committee, 2000b. Artificial neural networks in hydrology II. *Journal of Hydrologic Engineering,* ASCE. 5(2), 124–132.

Baker, T.F., 1993. Absolute sea level measurements, climate change and vertical crustal movements. *Global and Planetary Change,* 8, 149–159.

Downloaded From: https://bioone.org/journals/Journal-of-Coastal-Research on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use

Broomhead, D. and Lowe, D., 1988. Multivariable functional interpolation and adaptive networks. *Complex Systems,* 2, 321–355.
Chen, J.L.; Shum, C.K.; Wilson, C.R.; Chambers, D.P., and Tapley, B.D., 2000. Seasonal sea level change from TOPEX/Poseidon observation and thermal contribution. *Journal of Geodsy,* 73, 638–647.
Chen, J.L.; Wilson, C.R.; Tapley, B.D.; Famiglietti, J.S., and Rodell, M., 2005. Seasonal global mean sea level change from satellite altimeter, GRACE, and geophysical models. *Journal of Geodesy,* 79, 532–539.
Cigizoglu, H.K., 2003. Estimation, forecasting and extrapolation of flow data by artificial neural networks. *Hydrological Sciences Journal,* 48(3), 349–361.
Cigizoglu, H.K., 2004. Estimation and forecasting of daily suspended sediment data by multi layer perceptrons. *Advances in Water Resources,* 27, 185–195.
Cigizoglu, H.K., 2005a. Application of the generalized regression neural networks to intermittent flow forecasting and estimation. *ASCE Journal of Hydrologic Engineering,* 10(4), 336–341.
Cigizoglu, H.K., 2005b. Generalized regression neural network in monthly flow forecasting. *Civil Engineering and Environmental Systems,* 22(2), 71–84.
Cigizoglu, H.K. and Alp, M., 2006. Generalized regression neural network in modeling sediment yield. *Advances in Engineering Software,* 37, 63–68.
Cigizoglu, H.K. and Kisi, O., 2005. Flow prediction by three back propagation techniques using k-fold partitioning of neural network training data. *Nordic Hydrology,* 36(1), 1–16.
Cigizoglu, H.K. and Kisi, O., 2006. Methods to improve the neural network performance in suspended sediment estimation. *Journal of Hydrology,* 317(3–4), 221–238.
Doodson, A.T., 1958. The analysis and predictions of tides in shallow water. *International Hydrographic Review,* Monaco, 33, 85–126.
Douglas, B.C., 1991. Global sea level rise, *Journal of Geophysical Research,* 96(C4), 6981–6992.
Douglas, B.C.; Cheney, R.E.; Miller, L.; Carter, W.E., and Robertson, D.S., 1990. Greenland ice sheet; is it growing or shrinking? *Science,* 248–288.
Eberhart, R.C. and Dobbins, R.W., 1990. *Neural Network PC Tools: A Practical Guide,* San Diego, California: Academic Press.
Hagan, M.T. and Menhaj, M.B., 1994. Training feedforward techniques with the Marquardt algorithm. *IEEE Transactions on Neural Networks,* 5(6), 989–993.
Hannah, J., 1999. Analysis of mean sea level data from New Zealand for the period 1899–1988. *Journal of Geophysical Research,* 95(B8), 399–405.
Haykin, S., 1994. *Neural Networks: A Comprehensive Foundation.* Ontario, Canada: IEEE Press.
Hilbert, D.W. and Ostendorf, B., 2001. The utility of artificial neural networks for modelling the distribution of vegetation in past, present and future climates. *Ecological Modelling,* 146, 311–327.
Kim, B., Kim, S., and Kim, K., 2003. Modelling of plasma etching using a generalized regression neural network. *Vacuum,* 71, 497–503.
Lee, T.S., 2004. Back-propogation neural network for long term tidal predictions. *Ocean Engineering,* 31, 225–238.
Mackay, D.S. and Robinson, V.B., 2000. A multiple criteria decision support system for testing integrated environmental models. *Fuzzy Set Systems,* 113(1), 53–67.
Maier, H.R. and Dandy, G.C., 2000. Neural network for the prediction and forecasting of water resources variables: a review of modeling issues and applications. *Environmental Modeling and Software,* 15, 101–124.
Makarynskyy, O.; Makarynska, D.; Kuhn, M., and Featherstone, W.E., 2004. Predicting sea level variations with artificial neural networks at Hillarys Boat Harbour, Western Australia. *Estuarine, Coastal and Shelf Science,* 61, 351–360.
Mason, J.C.; Price, R.K., and Tem'me, A., 1996. A neural network model of rainfall-runoff using radial basis functions. *Journal of Hydraulic Research,* 34(4), 537–548.
Minster, J.F.; Cazenave, A.; Serafini, Y.V.; Mercier, F.; Gennero, M.C., and Rogel, P., 1999. Annual cycle in mean sea level from TOPEX-Poseidon and ERS-1: inference on the global hydrological cycle. *Global Planet Change,* 20, 57–66.
Neilan, R.E.; Van Scoy, P.A., and Woodworth, P.L., 1997. Workshop on Methods for Monitoring Sea Level, Pasadena, California, pp. 202.
Nerem, R.S.; Haines, B.J; Hendricks, J; Minster, J.F; Mitchum, G.T., and White, W.B., 1997. Improved determination of global mean sea level variations using TOPEX/POSEIDON altimeter data. *Geophysical Research Letters,* 24(11), 1331–1334.
Perret, J.S., and Prasher, S.O., 1998. Applications of fuzzy logic in the design of vegetated waterways under uncertainty. *Journal of the American Water Resources Association,* 34(6), 1355–1367.
Poggio, T. and Girosi, F., 1990. Regularization algorithms for learning that are equivalent to multilayer networks. *Science,* 247(4945), 978–982.
Pugh, D.T., 1987. *Tides, Surges and Mean Sea Level: A Handbook for Engineers and Scientists,* Chichester, U.K.: Wiley, 472p.
Roske, F., 1997. Sea level forecasts using neural networks. *German Journal of Hydrography,* 49(1), 71–99.
Sanli, D.U. and Blewitt, G., 2001. Geocentric sea level trend using GPS and >100-year tide gauge record on a postglacial rebound nodal line. *Journal of Geophysical Research,* 106(B1) 713–719.
Schmitt, R.W., 1995. The ocean component of the global water cycle. U.S. National Report to IUGG, 1991–1994. *Review of Geophysics,* 33 (Suppl).
Specht, D.F., 1991. A general regression neural network. *IEEE Transactions on Neural Networks,* 2(6), 568–576.
Taurino, A.M.; Distante, C.; Siciliano, P., and Vasanelli, L., 2003. Quantitative and qualitative analysis of VOCs mixtures by means of a microsensors array and different evaluation methods. *Sensors and Actuators,* 93, 117–125.
Tolkatchev, A., 1996. Global sea level observing system, *Marine Geodesy,* 19, 21–62.
Ultsch, A. and Roske, F., 2002. Self-organizing feature maps predicting sea levels. *Information Sciences,* 144, 91–125.
Vanicek, P., 1978. To the problem of noise reduction in sea level records used in vertical crustal movement detection. *Physics of the Earth and Planetary Interiors, 17,* 265–280.
Woodworth, P.L.; Spencer, W.E., and Alcock, G., 1990. On the availability of European mean sea level data, *International Hydrographic Review, 67*(1), 131–146.

Downloaded From: https://bioone.org/journals/Journal-of-Coastal-Research on 28 Mar 2024
Terms of Use: https://bioone.org/terms-of-use