Prediction of Salt Load Flowing to Lake El Manzala Using Artificial Neural Networks

Eman Abdel Ghaffar Hasan
National Water Research Center, Delta Barrage, Egypt

Abstract

Lake El Manzala is the largest of the northern Nile Delta coastal lagoons. It is an important and valuable natural resource area for fish catch, wildlife, and hydrological and biological regime. It contains Ashtume El Gamile protected area. Lake El Manzala is a highly dynamic aquatic system that has undergone considerable physical, chemical and biological changes during the past century. This was as a result of different aspects of human impacts such as continuous drying processes for cultivation purpose, human settlement and pollution with different kinds of water discharge into the lake. Water body of Lake El Manzal is affected qualitatively and quantitatively by drainage water that flow into the lake through the drainage system of the Eastern delta. The long – term changes in water quality condition are untimely linked to salinity and water level in the lake. One of the Drainage Research Institute products is the yearbook, which describes the drainage water status in the Nile delta as a request of the Ministry of Water Resources and Irrigation. Intensive monitoring networks, field trips for collecting in-situ data and extensive cost are needed to gain the knowledge of drainage water quantity and quality. New tools are urgently needed to improve the process of acquiring knowledge over time to control surface water pollution. In this aspect non-linear statistical data-modeling tool namely Artificial Neural Networks (ANNs) can be used. This paper presents the framework for developing ANN model for the prediction of salt load flowing to Lake El Manzala as it affects the regional biodiversity. Representing drainage water quantity, salinity and total dissolved salt that flows to the lake; model can predict the salt load and assess the effects of various drainage water disposal options. The model was calibrated and validated against the measured drainage water data for the period from year 1984 to 2005. Salt load prediction by model generally agreed with the calculated one from the measured data. The model performance was satisfactory with yearly efficiency equal 0.98.

Keywords: Artificial Neural Networks, Salt load, Drainage water, Feed-forward, Back-propagation

Introduction

Economically, Lake El Manzala in Egypt is the most important fishing ground amongst the northern Delta lakes. Its annual fish production represents about half of the total fish yield of Delta lakes, and about one fifth the (non-marine) fish yield of Egypt. It contains Ashtume El Gamile protected area by law 102/1983. Water body of Lake El Manzal is affected quantitatively and qualitatively by drainage water that flow into it. The hydrological properties of coastal lagoon are determined by its configuration and link to the sea, balance between fresh water sources and tidal inflow range, and finally on the prevailing meteorological conditions. The long – term changes in water quality condition are untimely linked to salinity and water level in the lake (Fares, 2006). Although Lake El Manzala is severely degraded, there still remain some high value aquatic resource areas within the lake that provide important services for local human populations and contribute substantially to regional biodiversity.

ANNs have been developed as a generalization of mathematical models of human cognition or neural biology. Neural networks are models of biological neural structures; the starting point for most neural networks is a model neuron. This neuron consists of multiple inputs and a single output. Each input is modified by a weight, which multiplies with the input value. The neural will combine these weighted inputs and will refer to a threshold value and activation function, use these to determine its output (ASCE, 2000). Since the early nineties, ANNs have been successfully used in hydrology – related areas such as stream flow forecasting, water quality and water management policy.

(Abdeen, 2001) built a neural network model for predicting flow characteristics in irregular open channels, the developed technique was capable with small computational effort and high accuracy of predicting flow depths and average flow velocities along the channel reach when the geometrical properties of the channel cross sections were measured and visa versa. (Ali, et al, 2006) developed the ANN – based model for the prediction of Lake Qaroun water levels using the Feed-Forward Multilayer Perceptions network. The developed model was a useful tool for analyzing water levels of the lake. (El – kholy, 2006) proved the efficiency of ANN model in prediction of the unsteady behavior of water quality physical, chemical and biological relations. The neural networks were able to produce highly accurate values of target output variables for Chemical Oxygen Demand and 6% for nitrate (NO_3) absolute deviation from the actual measures.

The objective of this paper is to examine the ANNs capability in handling the problem of estimating the annual and monthly salt load flowing to Lake El Manzala based on the measured drainage water quantity and salinity. The model was trained (calibrated) and simulated by comparing model results with field measurements of the study area (1984-2005).

Drained Water to Lake El Manzala

Lake El Manzala is situated at the eastern margin of the Nile Delta between 31° 00` - 31° 30` N latitude and 31° 45` - 32° 22`E longitude. The lake occupies an area of 1,410 km2 (47 km long by 30 km wide). The lake lies within the borders of five Egyptians governorates (Dakahliya, Damietta, Port Said, Ismailia and Sharkiya). It is bordered by the Suez Canal to the East, the Damietta Branch of Nile to the West, the Mediterranean Sea to the North and agricultural land in the south. The lake is connected to the Mediterranean via an artificial opening, El- Gamil outlet (100 meters wide) about 10 km west of Port Said. The northwestern side of the lake is connected with the Damietta Estuary by two canals - El-Suffara and El-Ratama, which lie north of the city of Damietta. Through these outlets, the exchange of water and biota between the lake and the adjoining Mediterranean Sea is possible. The northern part of the lake is affected by marine water invasion through the El-Gamil outlet (Boughaz El-Gamil) as shown in figure (1). The drainage systems of the Eastern Delta, except few catchments, drain their water to Lake El Manzala, which in turn discharge freely to the Mediterranean Sea. Two main drainage systems, Baher El Baqer and Baher Hadus drains, flow about 75 % of East delta drainage water into lake El Manzala. Each system consists of several sub-catchments. However the drainage systems of Mataraya, Lower Serw and Farsqur are contained in one catchment. Their pumping stations deliver the drainage directly to the Lake. Bahr El Baqar drain carries water for a distance of 170 kilometers from eastern Cairo to the lake. The drain carries largely untreated sewage from Cairo and other delta cities and industries that contain a large amount of matter, nutrients, bacteria, heavy metals, and toxic organics. The lake faces some environmental problems. Fish production is chemically and microbiology contaminated. Moreover, fish and bird species have been declined substantially in the area. Such problems mean that the hydrological condition of the lake is clearly deteriorating. (DRI, 2006) reported that the annual drainage water flowing to the Lake equals about 4.1 billion cubic meters, which lead to salt load of about 13.237 million tons per year. Increased salinity concentration is likely a major contributing factor to this decline in fishers as salinity at 40 mg/L provides unsuitable habitat for most species (Kilada, 1997).

The pattern of monthly variation in drainage water quantity and salinity is shown in figure (2). It could be observed from the figure that the maximum drainage water discharged to the lake take a place in August. while the minimum discharge occurs in February. The maximum drainage water salinity is 3.75 dS/m in March when the drainage water discharge is 382.87 $m.m^3$.

Artificial Neural Networks

A neural network is characterized by its architecture that represents the pattern of connection between nodes (neurons), its method of determining the connection weights, and the activation function (Fausett, 1994). A typical ANN consists of a number of nodes that are organized according to a particular arrangement. In feed forward networks, which are commonly used in the hydrology, the nodes are generally arranged in layers, starting from a first input layer and ending at the final output layer. There can be several hidden layers, with each layer having one or more nodes.

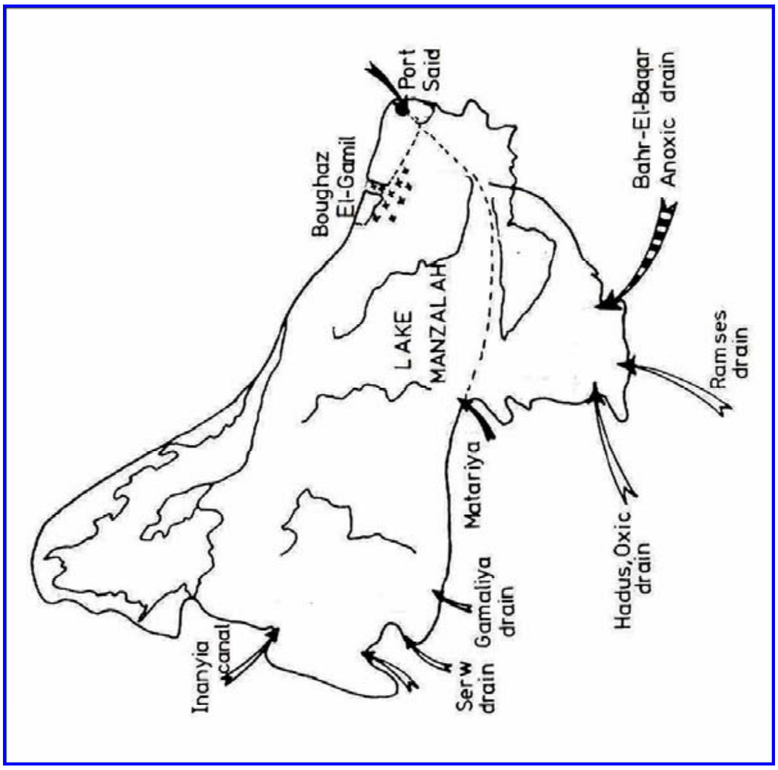

Figure (1) Catchment Area of El Manzala Lake

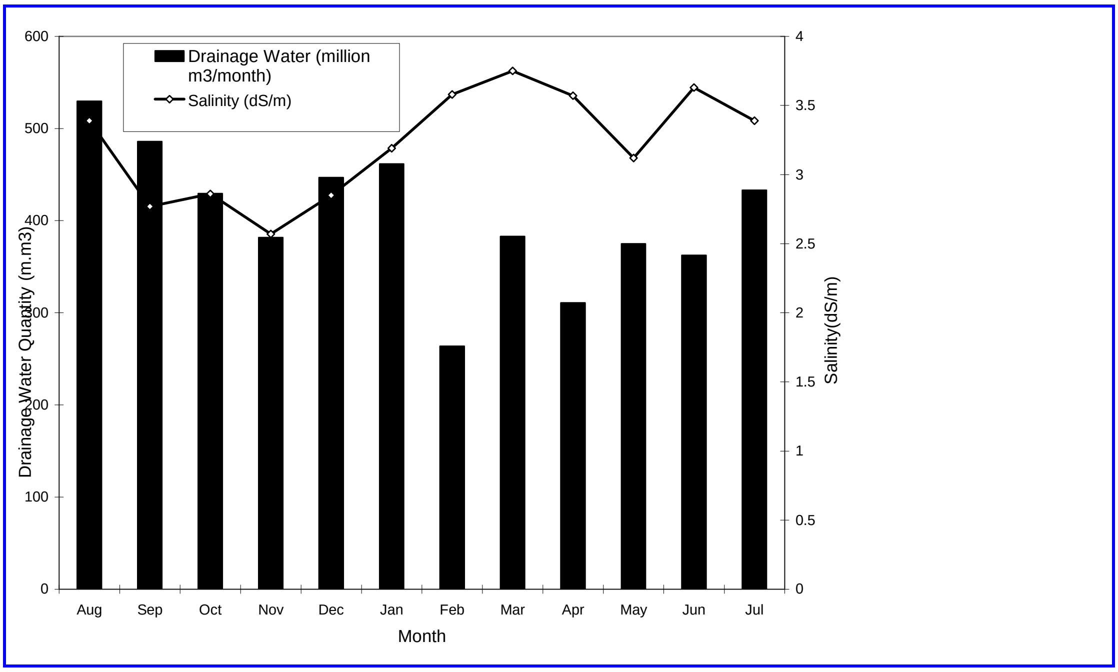

Figure (2) Monthly Drainage Water Quantity and Salinity Flowing To Lake El Manzala of Year 2005/2006

Configuration of Feed Forward Three-Layers ANN

In most networks**,** the input (first) layer receives the input variables for the problem at hand. This consists of all quantities that can influence the output. The input layer is thus transparent and is a mean of providing information to the network. The last or output layer consists of values predicted by the network and thus represents model output. The number of hidden layers and the number of nodes in each hidden layer are usually determined by a trial-and-error procedure. The nodes within neighboring layers of the network are fully connected by links. A synaptic weight (wij) is assigned to each link to represent the relative connection strength of two nodes at both ends in predicting the input-output relationship. Figure (3) shows the configuration of a feed forward three-layer ANN, where X is a system input vector composed of a number of causal variables that influence system behavior, and Y is the system output vector composed of a number of resulting variables that represent the system behavior. The output of node j, yj, is obtained by computing the value of function f with respect to the inner product of vector X and Wj minus bj, where bj is the threshold value, also called the bias, associated with this node. In ANN parlance, the bias bj of the node must be exceeded before it can be activated. The following equation defines the operation:

$$Y_j = f\left(X.W_j - b_j\right) \qquad (1)$$

The function f is called an activation function. Its functional form determines the response of a node to the total input signal it receives. The most commonly used form of f in equation (1) is the sigmoid function. The sigmoid function is a non decreasing function that provides a graded, nonlinear response. This function enables a network to map any nonlinear process. The popularity of the sigmoid function is partially attributed to the simplicity of its derivative that will be used during the training process.

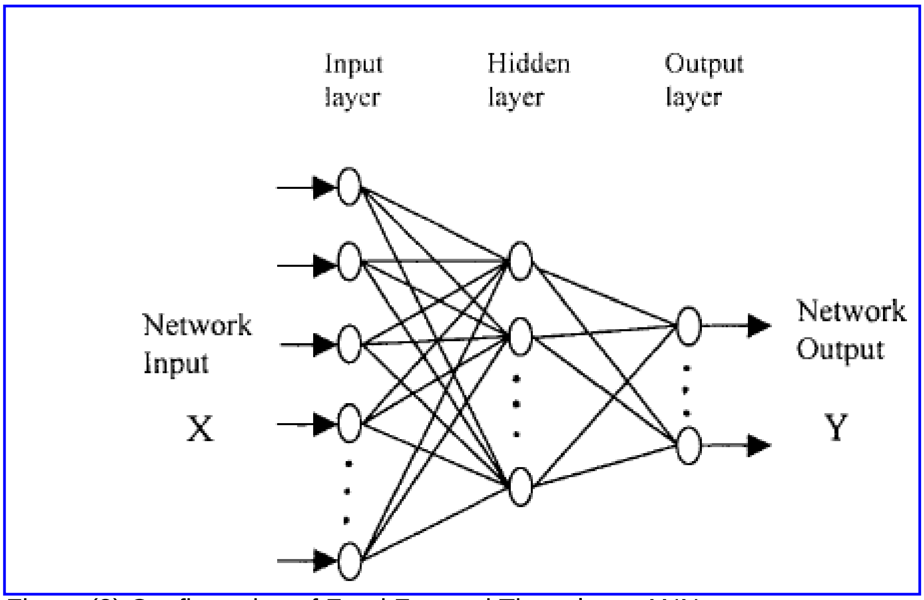

Figure (3) Configuration of Feed Forward Three-layer ANN

Network Training

ANN is characterized by its ability to learn (train) from examples. When a set of inputs is presented, a network adjusts weights in order to approximate the target output based on predefined procedures. Learning in ANN consists of three parts: the relative importance of the inputs, the control of the output generation from a neuron, and adjustments description of the weight during training.

In order for an ANN to generate an output vector Y = (y_1, y_2, . . . , y_p) that is as close as possible to the target vector T = (t_1, t_2, . . . , t_p), a training process, also called learning, is employed to find optimal weight matrices W and bias vectors V, that minimize a predetermined error function that usually has the form:

$$E = \sum_{p} \sum_{P} \left(y_i - t_j \right)^2 \qquad (2)$$

where, t_i is a component of the desired output T; y_i = corresponding ANN output; p = number of output nodes; and P = number of training patterns.

Back-Propagation

Back-propagation paradigm trains a neural network using a gradient descent algorithm in which the mean square error between the network's output and the desired output is minimized. This creates a global cost function which minimizes iteratively by "back propagating", the error from the output nodes to the input nodes. Once the network's error has decreased to less than or equal to the specified threshold, the network has converged and is considered to be trained (Albet Tebo, 1994).

Developed Model Description

The model network is created with multiple layers of neurons by using Neuralyst version 1.41. The particular version used to run the simulation is an Excel environment module. The network here consists of three layers with one hidden layer that has two neurons. The input layer contains three neurons and the output layer contains one neuron. Each neuron contains sigmoid function. This because the ability of Sigmoid function to approximate any function between any training data, but in limited area. ANN model is constructed to estimate the monthly and annual salt load discharged to lake El Manzala. The required input data to run the model are drainage water flowing to Lake El Manzala from the Eastern delta, salinity and total dissolved salt and the model output is the predicted salt load. This model assists the planner and manager of lakes in controlling water salinity that affects fisheries and biodiversity.

The network is trained to obtain the nearest output to the target (measured output). During training the weights (the input weight and layer weight) and biases of the network are iteratively adjusted to minimize the network performance function. The default performance function for feed-forward is the mean square error - the average squared error between the network outputs and the target output. The network is probably learning with the aid of the training data (Input/Target) pairs. The training process is done with a certain number of iterations (epochs). Finally, the network can be simulated to deduce any output with certain inputs for all the ranges in the training data.

Results and Discussion

Calculated and predicted salt load flowing to Lake El Manzala was compared using data from 1984 to 2005. Figure (4) shows the calculated (from measured data) and predicted annul salt load flowing to lake El Manzala. From the figure, it is clear that the trend between the measured and predicted salt load is similar. The Root Mean Square Error (RMSE) value is 0.026. This value shows that the simulation is the best fit with the measured data. The model efficiency is calculated according to Hack-ten Broke, et al, 1996 equation as follows:

$$EF = .\frac{\left[\sum_{i=1}^{n}(O_i - O_m)^2 - \sum_{i=1}^{n}((P_i - O_i)^2\right]}{\left[\sum_{i=1}^{n}(O_i - O_m)^2\right]} \quad \text{...(3)}$$

where:

EF: model efficiency
Oi : the calculated salt load form measurements
Pi : the prediction salt load
Om: the mean value of Oi
n : the number of records

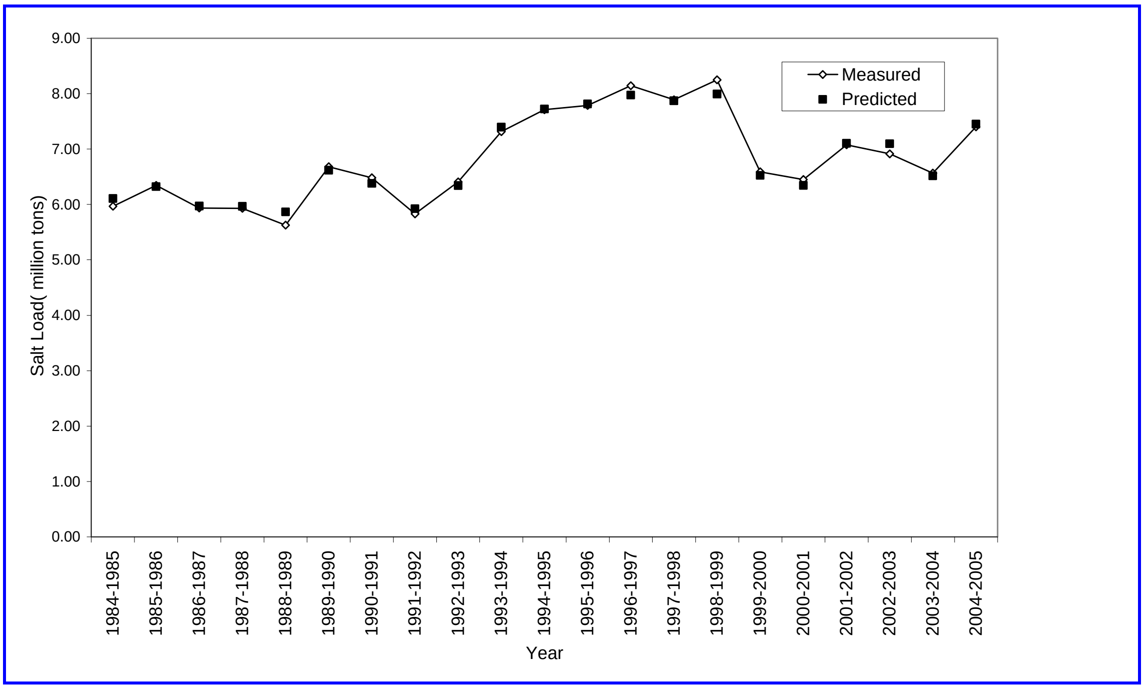

Figure (4) Annual Evaluation of ANN Model- Salt Load Prediction (1984-2005)

The model efficiency on an annual base was 0.98 which is close to 1. This shows that the model is in good agreement with the measured data. Moreover, over the study period, the annual percentage of relative error was calculated as shown in figure (5) which indicates the high degree of accuracy of the developed ANN model in predicting the salt load flowing to lake El Manzala as the maximum percentage of relative error between the target required value and the ANN prediction one is 3.15.

Figure (6) shows the calculated and predicted monthly salt load flowing to the lake for the year 2005-2006. The RMSE value is 0.037. The model efficiency on monthly bases ranges from 90 to 100%. Meanwhile, the monthly percentage of relative error during the year 2005-2006 ranges from -8.63 to 5 which confirmed the accuracy of the developed ANN model on monthly bases as shown in figure (7).

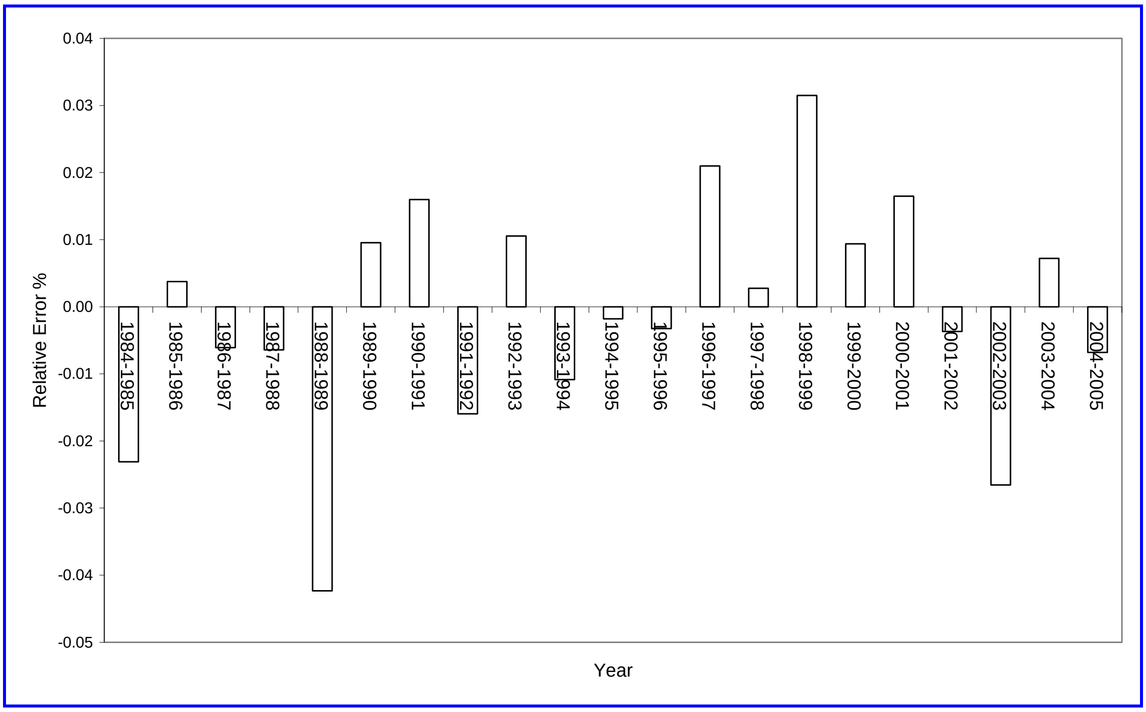

Figure (5) Annual Percentage of Relative Error for ANN Model- Salt Load Prediction (1984-2005)

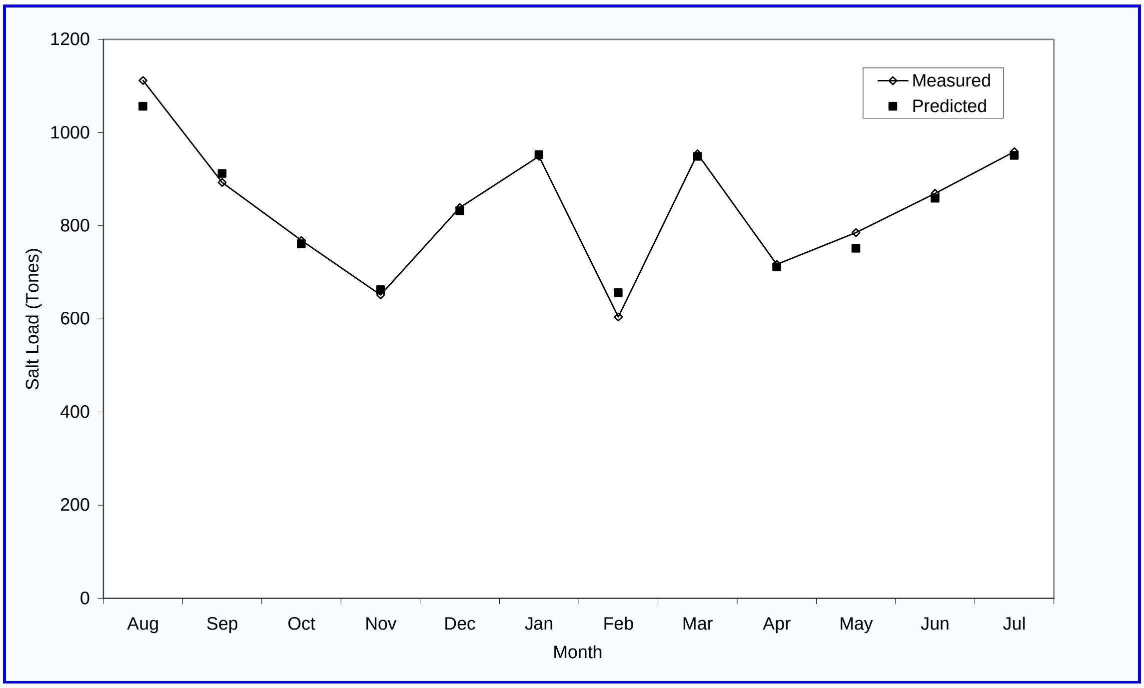

Figure (6) Monthly Evaluation of ANN Model- Salt Load Prediction for year 2005/2006

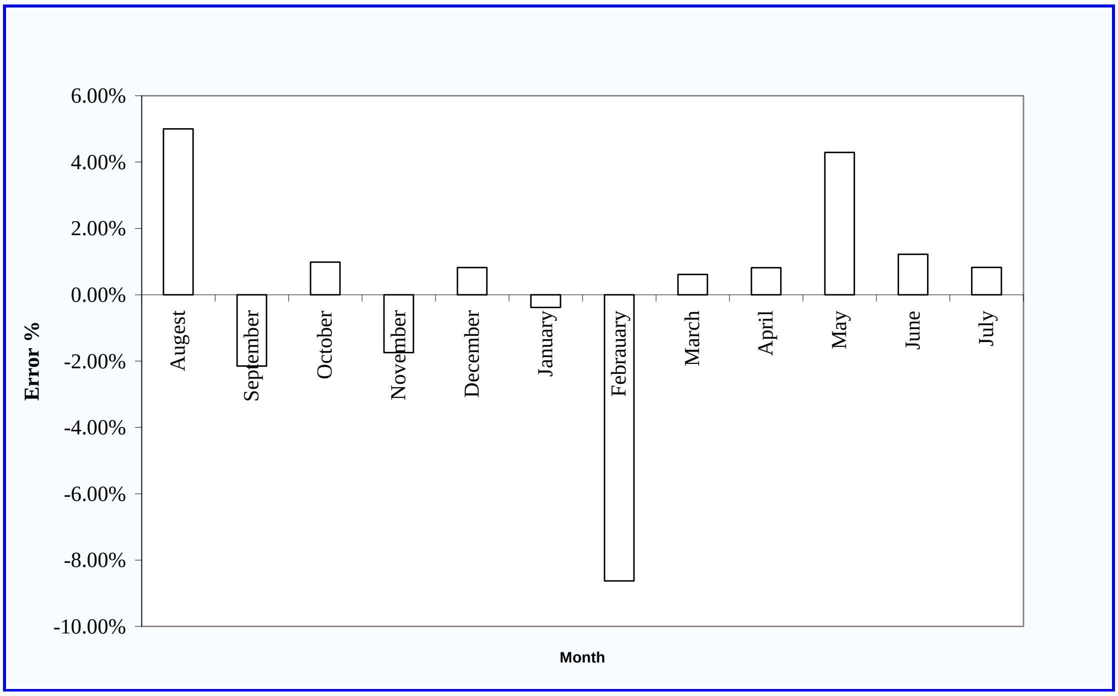

Figure (7) Monthly Percentage of Relative Error for ANN Model- Salt Load Prediction for year 2005/2006

Conclusion and Recommendation

The ANN model for salt load prediction has been developed using the Feed - Forward Multilayer network based on the monitoring network for drainage water of the Eastern delta. The model performed satisfactory with efficiency equaled to 0.98 on annual bases. While on monthly bases the modeling efficiency varied from 0.90 to 1. Comparison between simulation and measurements proved that the model is a useful tool for prediction of salt load flowing to lake El Manzala for various drainage disposal options. The percentage of relative error during the study period confirmed the accuracy of the developed ANN model since the maximum percentage of relative error between the target required value and the ANN prediction one is 3.15.

This model can assist planner and manger of lakes in controlling water salinity that affects fisheries and biodiversity of lake El Manzala as it contains one of the protected areas namely Ashtume El Gamile.

References

Abdeen M.A, 2001. Neural Network Model for Predicting Flow Characteristics in Irregular Open Channels, Scientific Journal, Faculty of Engineering-Alexandria University, Vol. 40, No. 4, pp. 539-546.

Albet Tebo, 1994. Artificial Neural Networks: a Developing Science. OE Reports. www.spie.org/ neural_net.html.

Ali, H, E Khalifa and M El- Beshry, 2006. Prediction of Lake Qaroun Water Levels Using Artificial Neural Networks. the 3rd Arab Water Regional Conference Research Advancement Managing, Limited Water Resources, National Water Research Center (NWRC) and Arab Water Council (AWC), 9-11 December 2006. Intercontinental Cairo Hotel, Cairo, Egypt.

ASCE Task Committee. 2000. Artificial Neural Networks in Hydrology: Preliminary Concepts. Journal of Hydrologic Engineering, Vol.5, No. 2.

DRI, 2006. Drainage Water Status in the Nile Delta Yearbook 2004/2005, Technical Report No.75.

El-kholy R. M.S., 2006 Artificial Neural Network Model for Predicting Drainage Water Quality in Egypt. the 3rd Arab Water Regional Conference Research Advancement Managing, Limited Water Resources, National Water Research Center (NWRC) and Arab Water Council (AWC), 9-11 December. Intercontinental Cairo Hotel, Cairo, Egypt.

Fausett, L, 1994. Fundamentals of Neural Networks. Prentice Hall, Englewood Cliffs, N.J.

Hack-ten Broke, M, J. D and Hegmans, J.H.B, 1996. Use of Saline Physical Characteristics from Laboratory Measurements or Saturated Series for Modeling Unsaturated Water Flow. Agricultural Water Management. VOL.29, P201-213.

Kilada, Rauf. 1997. Fisheries Development in Fayoum Water Management Project II. Workshop Proceedings on Fayoum Water Policy and Salinity Management of Qaroun and Wadi Rayan Lakes. ARCDIS Eurocosult and Darwish Consulting Engineers, Egypt.

Fares. Y.R, 2006. Conceptualization of Hydrodynamics and Water Quality Changes in Coastal Brackish Lagoons. The First International Conference Environmental Change of Lakes, Lagoons & Wetlands in the Southern Mediterranean Region 4-7 January Cairo, Egypt.