Journal of Applied Sciences Research, 8(5): 2785-2799, 2012
ISSN 1819-544X
This is a refereed journal and all articles are professionally screened and reviewed

ORIGINAL ARTICLES

Estimation of Evaporation Losses from Lake Nasser: Neural Network based Modeling versus Multivariate Linear Regression

[1]Essam A. Kalifa, [2]Reda M. Abd El-Hady Rady, [3]Sayed A. Alhayawei

[1]Associate Professor, National Water Research Center, Egypt,
[2]Researcher, National Water Research Center, Egypt,
[3]Junior Civil Engineer, Ministry of Water Resources and Irrigation

ABSTRACT

This paper presents an attempt to apply Artificial Neural Network ANN based modeling to accurately estimate the evaporation from Lake Nasser, Egypt making use of data collected from instrumented hydrometeorological stations at seven locations on the lake. Several input combinations were tried to find out the importance of different input parameters in predicting the evaporation. Water and air temperatures, wind speed, relative humidity, and solar radiation data were used as independent variables. Moreover, the prediction accuracy of ANN was also compared with the accuracy of linear and Multivariate Linear Regression (MLR) as well as with three conventional methods for predicting evaporation. The comparison demonstrated superior performance of ANN over other approaches. A graph between the calculated evaporation values by high Aswan Dam authority and predicted values of evaporation suggests that most of the values lie within a scatter of ±7% and of ±10% for ANN and MLR respectively with all input parameters. The calculations revealed a rate of 7.64 mm/ day as the annual mean evaporation rate.

Key words: Evaporation; Multivariate Linear Regression; Artificial Neural Network

Introduction

Accurate estimation of evaporation influences water balance studies undertaken for efficient planning, design, operation and management of water resources. In the hydrological practice, the evaporation can be estimated by conventional approaches like direct or indirect methods involving use of the empirical equations.

The sole direct method is the U.S weather Bureau Class a pan measurement that gives record with time. The indirect methods use meteorological data to estimate evaporation by empirical based methods or statistical and stochastic approaches. The indirect methods are namely Temperature based formulae, Radiation method; Humidity based relation, Penman formulae, Energy balance approach, and etc. These methods of evaporation estimation have been applied by (Abtew, 2001); (Choudhury, 1999); (Vallet- Coulomb *et al*., 2001); (Terzi, *et al*., 2005); and (Rosenberry, 2007). Although all these approaches are based on Penman formula, they are sensitive to site-specific evaporation parameters, which can vary from one place to other. Further, it is not possible to consider all the parameters affecting the evaporation estimation by any of the above approaches due to many assumptions and phenomenal constraints.

The literature review showed that these equations vary greatly in their ability to define the magnitude and variability of the evaporation from reservoirs. It is therefore necessary to develop alternate approaches to estimate the evaporation rates based on metrology variables, which are comparatively easier to measurements and estimation. One of the recent alternate approaches is the use of soft computing modeling techniques, which have better modeling flexibility and capability rather than previous empirical approaches

Many researchers proposed simple linear relationships between evaporation and individual meteorological parameters for predicting evaporation from reservoirs by using linear regression approach. They used field study data comprised of a single dependent variable (i.e. evaporation, E) and independent variables, describing meteorological parameters that affect evaporation, including water temperature (WST) air temperature (AT), wind speed (WS), solar radiation (SR) and relative humidity (RH). However, they had not considered the combined effect of all the meteorological parameters (WST+AT WS+SR+RH) on evaporation loss by using Multiple Linear Regression (MLR), which seems to be the major limitation of their studies.

Most of the current models for predicting evaporation use the principles of the deterministically based combined energy balance – vapor transfer approach or empirical relationships based on climatologic variables. This resulted in relationships that were often subjected to rigorous local calibrations and therefore proved to have limited global validity. Evaporation is a complex and nonlinear phenomenon because it depends on several

Corresponding Author: Essam A. Kalifa, Associate Professor, National Water Research Center, Egypt,

interacting climatologic factors, such as temperature, humidity, winds speed, bright sunshine hours, etc. An Artificial Neural Network (ANN) is a flexible mathematical structure, which is capable of identifying complex nonlinear relationships between input and output data sets. The ANN models have been found useful and efficient, particularly in problems for which the characteristics of the processes are difficult to describe using physical equations. An ANN model can compute complex nonlinear problems, which may be too difficult to represent by conventional mathematical equations. These models are well suited to situations where the relationship between the input variable and the output is not explicit. Instead, ANN, map the implicit relationship between inputs and outputs through training by field observations. The model may require significantly less input data than a similar conventional mathematical model, since variables that remain fixed from one simulation to another do not need to be considered as inputs. The ANN is useful, requiring fewer input and computational effort and less real time control.

In this study, both ANN and MLR techniques were used to estimate evaporation losses from Lake Nasser. Lake Nasser is located in the lower Nile River Basin at the border between Egypt and Sudan at 182 m above mean sea level. The total capacity of the reservoir is 162 BCM at its highest water level (Sadek *et al*., 1997); and (Omar and El-Bakry, 1981). For many years, the Egyptian Ministry of Water Resources and Irrigation adopted the figure of 7.54 mm/day as the annual mean evaporation rate with a maximum value in June of 10.8 mm/day and a minimum in December of 3.95 mm/day (Whittington and Guariso, 1983).

Water loss from Lake Nasser is a national problem. The evaporated water loss ranges between 10 and 16 BCM every year, which is equivalent to 20–30% of the Egyptian income from Nile water. The evaporation loss from Lake Nasser is an influential factor in the Egyptian water budget. Therefore, the aim of this paper is to assess the potential and usefulness of ANN based modeling for evaporation prediction from Lake Nasser, where in classical and empirical equations failed to predict the evaporation accurately. Hence, in this study a suitable ANN model was developed by considering the feed-forward back propagation learning algorithm in the estimation of daily evaporation from meteorological parameters and its performance comparison with simple and multiple linear regression approaches. The meteorological data set of daily evaporation, temperature, solar radiation, relative humidity, wind speed were collected with the help and under the supervision of High Aswan Dam Authority (HADA) from the 7 hydrometeorological stations scattered along Lake Nasser in order to enhance the estimation of evaporation and reach more accurate results. The performance of feed forward back propagation neural network model was then compared with both linear and multiple linear regression on the basis of performance parameters; correlation coefficient and Root Mean Square Error (RMSE) having different combinations of input parameters.

Study Region and Data Used:

High Aswan Dam Reservoir extends for 500 km along the Nile River and covers an area of 6,000 km^2, of which two-thirds about 350 Km (known as Lake Nasser) is in Egypt and one-third about 150 Km (called Lake Nubia) in Sudan (Atkins, 2002). Lake Nasser lies in the extreme southern part of Egypt occupying a considerable area behind the High Aswan Dam (HAD). Due to the large area of extent and the huge mass of water in this lake, it is the second largest artificial lake in the world. The Lake is founded mainly on Precambrian granitic terrain and extends southward towards the Egyptian-Sudanese border. High Dam Lake is bounded by latitudes 24°N in Egypt and 21°S in Sudan. Lake Nasser water is a major source for drinking, irrigation, and domestic purposes in Egypt. The shoreline of Lake Nasser at 160 m (AMSL) is 5416 Km and at 180 m (AMSL) level is 7875 Km length. The length of eastern shoreline is almost double that of the western shoreline.

The surface area of the entire reservoir is 3084 Km^2 at level of 160 m, when the reservoir is nearly full at the level of 180 m it has a surface area of 6276 Km^2 (Jeongkon and Mohamed, 2002). The long reservoir has 100 side arms called khors, more on the eastern shore than on the western shore. The total capacity of the reservoir (162 BCM) consists of the dead storage of 31.6 BCM (85-147 m), the active storage of 90.7 BCM (147-174 m) and the emergency storage for flood protection of 41 BCM (175-182 m). The reservoir is surrounded by rocky desert terrain. To the west is the great Sahara Desert, and the Eastern Desert on the east side extends to the Red Sea (Fig. 1).

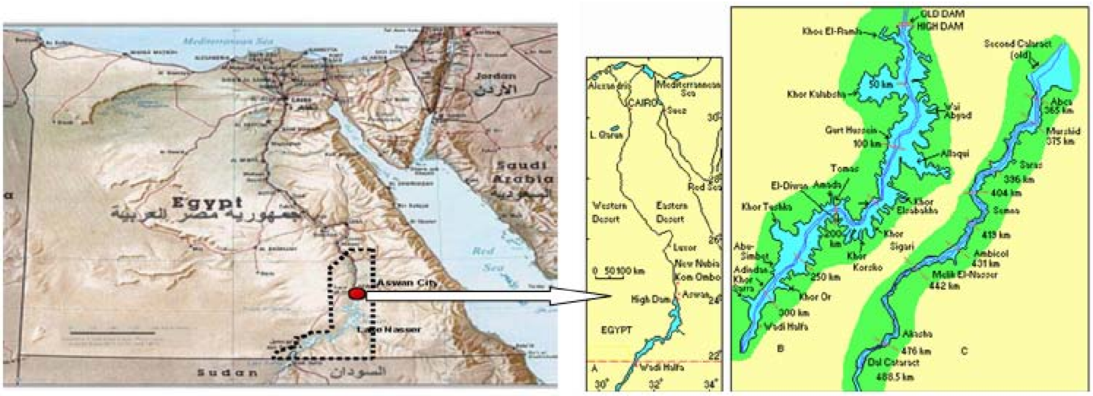

Fig. 1: Lake Nasser location

All measurements were made at 1-h intervals and averaged to daily values. The hourly measurements were collected from seven floating stations (Fig. 2), operated by HADA. The seven stations have worked with full capacity since 1995 for the stations at Raft (2 km upstream of the HAD) and Allaqi (75 km upstream of the AHD), from 1999 for the station at Abusembel (280 km upstream of the HAD), from 2005 for the stations of Amada (185 km upstream of the HAD) and Toshka (240 km upstream of the HAD), , and finally from 2010 for Khor Kallabsha (40 km upstream of the HAD) and Al-Malky (140 km upstream of the HAD).Measured variables included net radiation, water temperature, air temperature, relative humidity, wind speed, and air pressure. The evaporation data was collected for two years; 2010 and 2011. Dealing with data quality, with respect to missing data, it was also important to evaluate the length of consecutive periods when data is not available. To recover the missing data for any station, even in an extreme case, the data coming from the nearest station with almost complete set of data and show a strong seasonal patterns was used. Thus, one specific year for that has a good coverage was chosen to reconstruct the time series on the basis of the variance of the entire data set.

Fig. 2: Location map of the hydrometeorological floating stations along Lake Nasser

Evaporation in Reservoirs:

Evaporation from open-water bodies is quite different from these above-land surfaces (Eichinger *et al*., 2003). The sun's energy penetrates the water to depths of as much as 30 m in clear water, somewhat less in turbid water, and is stored throughout the water column (Eichinger *et al*., 2003). The water column is mixed by surface motion and becomes the source of energy that drives evaporation. Because of the large heat storage capacity of water (1.006×10^{6}J m^{-3}) and the fact that water is approximately 1,000 times more dense than air, the temperature of deep, clear, water bodies does not change significantly throughout the day when compared to the atmosphere (Eichinger *et al*. 2003). The amount of available energy at the surface is nearly constant throughout the day and night, leading to a nearly constant evaporation rate. The water in Nasser Lake lies between the two extremes of a clear water body and a land surface (desert). The Lake is deep over much of its area and is relatively clear, resulting in massive storage of solar energy in a considerably greater volume.

Accounting for the above considerations, observational studies of Lake evaporation have used a variety of different methods to estimate evaporation rates. These include the Bowen Ratio Energy Budget (BREB) method and eddy correlation techniques, the water-budget method, methods of the so-called Dalton group such as the bulk aerodynamic method, the methods in the so-called combination group such as the Penman, Priestley–Taylor, and deBruin–Keijman methods, and the methods in the temperature group such as the Papadakis method among others (Lenters *et al*., 2005); (Winter, 1981); and (Winter *et al*., 1995). Another method that is used in numerous studies is the mass transfer method (Yu and Knapp, 1985); (Ikebuchi *et al*., 1988); and Laird and Kristovich, 2002) because of its ease of application and suitability for modeling (Hostetler and Bartlein, 1990); (Blodgett *et al*., 1997); and Lenters *et al*., 2005). The water-budget method (Myrup *et al*., 1979) can potentially provide a most reliable estimate of evaporation, as long as each water budget component is accurately measured or modeled, which is often a difficult task, especially for the groundwater losses (Lenters *et al*., 2005). As reported by Sene *et al*. (1991) and Stannard and Rosenberry (1991), except for a recent study by Blanken *et al*. (2000), most applications of the eddy-correlation technique have been limited to applications with short-term data series.

Artificial Neural Network:

A neural network is an artificial intelligence technique that mimics a function of the human brain. Neural networks are general-purpose computing tools that can solve complex non-linear problems in the field of pattern recognition, classification, speech, vision and control systems. The network comprises a large number of simple processing elements linked to each other by weighted connections according to a specified architecture. A neuron consists of multiple inputs and a single output. The number of neurons in the input and output layers are fixed by the problem being modeled as the number of input variables equals number of input neurons and number of output variables equal number of output neurons. The determination of optimal number of hidden layers and hidden neurons is usually cumbersome, as no general methodology is available for their determination. Most of the studies employing neural networks for water resource problems have used back propagation and radial basis function types of neural networks.

There is no interconnection between the nodes of the same layer. In a back propagation neural network, generally, there is an input layer that acts as a distribution structure for the data being presented to the network. After this layer, one or more processing layers follow, called the hidden layers. The final processing layer is called the output layer in a network. This process is repeated until the error rate is minimized or reaches to an acceptable level. All the interconnections between each node have an associated weight. The values of the interconnecting weights are not set by the analyst, but are determined by the network during the training process, starting with randomly assigned initial weights. There are a number of algorithms that can be used to adjust the interconnecting weights to achieve minimal overall training error in multi-layer networks. The generalized delta rule, or back-propagation is one of the most commonly used methods as suggested by (Rumelhart, *et al*., 1986). The first derivative of the total error with respect to a weight (Equation 1) determines the extent to which that weight is adjusted.

$$\Delta w = -\in \partial E/\partial w \qquad (1)$$

Where $\in$ is the learning constant; $\partial E/\partial w$ is the first derivative of the total error with respect to weight and Δw is weight change. A neural network based modeling approach requires setting up several user-defined parameters like learning rate, momentum, optimal number of nodes in the hidden layer and the number of hidden layers, so as to have a less complex network with a better generalization capability.

Neural Network and Evaporation:

Evaporation reflects the influence of several meteorological parameters like air temperature, sunshine hours, wind speed, relative humidity, solar radiation, evaporating power of the air and vapor pressure deficit of a locality. But measurement of evaporation with accuracy is difficult task. In such cases, it becomes assertive to use formulae or neural network model that can estimate evaporation from available climatic data, may give more accurate results than the measured pan evaporation.

Artificial Neural Network evaporation models:

The review of literature on ANN and evaporation models is presented here. Bruton (2000) developed ANN models to estimate daily pan evaporation using measured weather variables as inputs. Weather data from Rome, Plains and Watkinville, Georgia, consisting of 2044 daily records from 1992 to 1996 were used to develop the models of daily pan evaporation. Additional weather from these locations, which included 720 daily records from 1997 and 1998, served as an independent evaluation data set for the models. The measured variables included daily observations of rainfall, temperature, relative humidity, solar radiation, and wind speed. The ANN models of daily pan evaporation with all available variables as an input was the most accurate model delivering an r^2 of 0.717 and a RMSE of 1.11 mm for the independent evaluation data set. ANN models were developed with some of the observed variables eliminated to correspond to different levels of data collection as well as for minimal data sets. The accuracy of the models was reduced considerably when variables were eliminated to correspond to weather stations. Pan evaporation estimated with ANN models was slightly more accurate than the pan evaporation estimated with other methods.

Sudheer (2002) investigated the prediction of Class A pan evaporation using ANN technique. The ANN back propagation algorithm has been evaluated for its applicability for predicting evaporation from minimum climatic data. Four combinations of input data were considered and the resulting values of evaporation were analyzed and compared with those of existing models. The results from this study suggest that the neural computing technique could be employed successfully in modeling the evaporation process from the available climatic data set. However, an analysis of the residuals from the ANN models developed revealed that the models showed significant error in predictions during the validation, implying loss of generalization properties of ANN models unless trained carefully. The study indicated that evaporation values could be reasonably estimated using temperature data only through the ANN technique. This would be of much use in instances where data availability is limited.

Ozlem (2005) estimated daily pan evaporation by a suitable ANN model for the meteorological data recorded from the Automated GroWheather meteorological station near Lake Egirdir, Turkey. In this station six meteorological variables are measured simultaneously, namely, air temperature, water temperature, solar radiation, air pressure, wind speed and relative humidity. Since the purpose is the estimation of evaporation the ANN architecture has only one output neuron withup to 4 input neurons representing air and water temperature, air pressure and solar radiation. Prior to ANN model construction the classical correlation study indicated that the insignificance of the wind speed and the relative humidity in the Lake Egirdir area. Hence the final ANN model has 4 input neurons in the input layer with one at the output layer. The hidden layer neuron number is found 3 after various trial and error models running. The ANN model provides good estimate with the least AMSE.

Keskin and Terzi (2006) studied the ANN models and proposed an alternative approach of evaporation estimation for Lake Eirdir. This study has three objectives: (1) to develop ANN models to estimate daily pan evaporation from measured meteorological data; (2) to compare the ANN models to the Penman model; and (3) to evaluate the potential of ANN models. The results of the Penman method and ANN models are compared to pan evaporation values. The comparison shows that there is better agreement between the ANN estimations and measurements of daily pan evaporation than for other model.

Forward Neural Network:

Forward neural networks, trained with a back-propagation learning algorithm, are the most popular neural networks. They are applied to a wide variety of chemistry related problems (J. Zupan and J. Gasteiger, 1993). A forward neural network consists of neurons, that are ordered into layers. The first layer is called the input layer, the last layer is called the output layer, and the layers between are hidden layers. For the formal description of the neurons one can use the so-called mapping function, that assigns for each neuron *i* a subset, which consists of all ancestors of the given neuron. A subset consists of all predecessors of the given neuron *i*. Each neuron in a particular layer is connected with all neurons in the next layer. The connection between the i^{th} and j^{th} neuron is characterized by the weight coefficient w_{ij} and the i^{th} neuron by the threshold coefficientv_i. The weight coefficient reflects the degree of importance of the given connection in the neural network. The output value (activity) of the i^{th} neuron x_i is determined by Eqs. (2) and (3). It holds that:

$$X_i = f(\xi_i) \tag{2}$$

$$\xi_i = v_i + \sum_{j\epsilon T_i^{-1}} w_{ij} x_j \tag{3}$$

where ξ_i is the potential of the i^{th} neuron and function $f(\xi_i)$ is the so-called transfer function (the summation in Eq. (2) is carried out over all neurons *j* transferring the signal to the i^{th} neuron). The threshold coefficient can be understood as a weight coefficient of the connection with formally added neuron *j*, where x_j = 1 (so-called bias). For the transfer function it holds that:

$$f(\xi) = \frac{1}{1+\exp(-\xi)} \tag{4}$$

Back-propagation training algorithm:

In back-propagation algorithm the steepest-descent minimization method is used. For adjustment of the weight and threshold coefficients it holds that:

$$w_{ij}^{(k+1)} = w_{ij}^{(k)} - \lambda \left\{\frac{\partial E}{\partial w_{ij}}\right\}^{(k)} \tag{5}$$

$$v_i^{(k+1)} = v_i^{(k)} - \lambda \left\{\frac{\partial E}{\partial v_i}\right\}^{(k)} \tag{6}$$

where λ is the rate of learning ($\lambda > 0$). The key problem is calculation of the derivatives

$$\frac{\partial E}{\partial w_{ij}}, \frac{\partial E}{\partial v_i}$$

Methods:

Evaporation from water surface is a continuous process affected by many meteorological parameters usually used in the computation of the amount of water lost by evaporation. Three different techniques were used to estimate the evaporation losses from Lake Nasser. They are respectively: atmospheric empirical formulas; linear regression, and ANN based modeling.

Empirical Formulas:

The most common methods and procedures used to estimate the evaporation losses depend on theoretical analyses, whereas others depend on formulae based on atmospheric elements. Some of these methods are applied below to compute the amount of water lost by evaporation from Lake Nasser. In addition, both linear and nonlinear analysis were used to predict the mount of evaporation losses for the 7 stations. Three methods from the literature are used to estimate the evaporation rate as follows:

Method (I) Bulk Aerodynamic:

This method is applicable on Lake Nasser taking into consideration the values of average air temperature at the lake. It is found that the appropriate coefficient (N) in equation (1) equals to 0.1296 (Omar and El-Bakry 1981).

$$E = N\, U_2\, (es – ed) \tag{7}$$

in which:

E : Evaporation (mm/day),
N : A constant equal to 0.1296,
U_2 : The wind speed at height 2.0 m above water surface (m/sec),
es : A saturated vapor pressure at water temperature (Hectopascal) ,and
ed : A vapor pressure of air 2.0 m above water surface (Hectopascal).

Method (II) Modified Bulk Aerodynamic:

This method represents the bulk Aerodynamic method with some modifications added by the Russian. These modifications can be expressed as follows:

$$E = 0.1296\ (1 + 0.7\ U2)\ (ea – ed) \tag{8}$$

in which:
E : evaporation rate (mm/day),
U_2 : wind speed 2.0 m above water (m/sec),
ea : saturated vapor pressure at air temperature (Hectopascal), and

ed : vapor pressure in air 2.0 m above water (Hectopascal).

The main difference between this method and the previous one is that this method always gives a value for evaporation even if the air velocity is equal to zero.

Method (III) Penman Method:

Penman (1948) presented a theory and a formula for the estimation of evaporation from weather data, which allow the computation of evaporation from a free water surface using readily available standard meteorological data only as follows:

$$E = \frac{\frac{\Delta}{\gamma}(R_n + E_n)}{\frac{\Delta}{\gamma} + 1} \tag{9}$$

$$R_n = \left[(1-\alpha)RG - e\,\frac{(352.8 - 0.195\sigma T_a{}^4)(0.1 + 0.9\frac{n}{N})}{59}\right] \tag{10}$$

$$E_a = 0.26\,(\,0.5 + 0.7\,U2\,)\,(\,ea - ed\,) \tag{11}$$

in which:
E : evaporation rate (mm/day),
R_n : solar balance for water surface,
R_n : $R_{ns} - R_{nlg}$,
R_{ns} : short wave radiation. = RG (1 – α),
RG : short wave radiation actually received at the earth from sun and sky,
α : reflection coefficient = 0.04 to 0.09 ,
Rnlg : long wave radiation,
σ: theoretical black body radiation, and
$\frac{\Delta}{\gamma}$: empirical parameter depending on temperature.

This method is considered to be the best method in estimating the evaporation losses because it takes into consideration the energy budget method and the bulk aerodynamic method. The results revealed from this method were in good accordance with the calculated evaporation by HADA. An example of the results of the 3 equations is illustrated for both Aswan and Abusembel stations in Tables 1-2 and Figs. 3-4.

Table 1: Calculated Evaporation Rate at Raft Station in 2011

	Evaporation (mm/day)		
Method	I	II	III
Jan	5.25	5.27	5.42
Feb	6.76	6.83	6.95
Mar	10.34	10.50	9.18
Apr	13.78	14.12	10.97
May	16.45	16.80	11.88
Jun	17.00	17.91	14.05
Jul	16.50	16.75	13.21
Aug	15.85	16.15	11.90
Sep	15.31	15.65	10.95
Oct	12.12	12.65	9.25
Nov	7.90	7.55	6.56
Dec	5.59	5.60	5.41

Table 2: Calculated Evaporation Rate at Abusembel Station in 2010

	Evaporation (mm/day)		
Method	I	II	III
Jan	5.22	5.08	5.42
Feb	5.45	5.58	6.25
Mar	9.20	9.15	8.77
Apr	12.52	12.68	10.57
May	16.25	16.25	11.89
Jun	17.25	17.12	13.55
Jul	15.21	16.04	12.00
Aug	16.75	17.26	11.88
Sep	16.71	16.55	11.20
Oct	15.42	14.71	10.15
Nov	9.27	8.90	7.35
Dec	6.49	6.18	5.85

J. Appl. Sci. Res., 8(5): 2785-2799, 2012

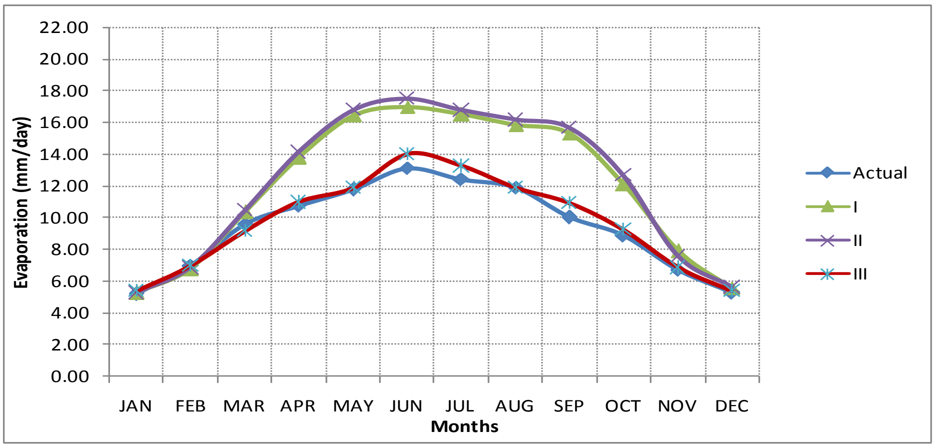

Fig. 3: Calculated evaporation Rate at Raft Station in 2011 using Empirical Formulas

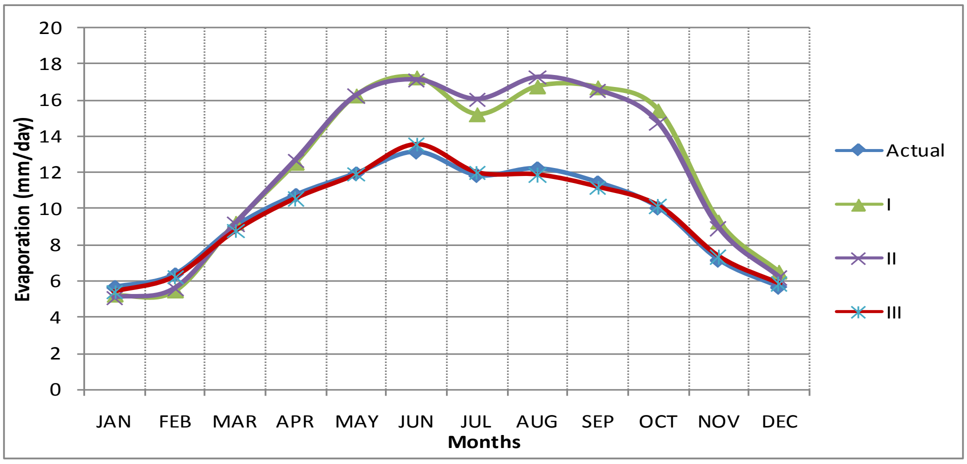

Fig. 4: Calculated Evaporation rate at Abusembel Station in 2010 using empirical formulas

Multivariate Linear Regression:

In this study, Multivariate Linear Regression (MLR) analysis was used in order to reveal the variables which affect evaporation measured at Lake Nasser. XLStat program was used for the regression analysis. Regression represents a mathematical equation expressing one random variable as being correlatively related to another random variable, or to several random variables. The association of more than two variables can be investigated by multiple regression and correlation analysis. The general multiple regression relation may be expressed as the fallowing:

$x1=f(x2,x3,\ldots,xm)+e$

in which $x1,x2,\ldots,xm$ are variables and e are the residuals. The estimation of $x1$ for given values of all other values can be calculated by this equation. If there are m variables to correlate, one dependent and m-1 independent, the equation is $x1=b1+b2x2+\ldots+bixi+\ldots+bmxm$ in which bi coefficients are considered both as the population parameters and their sample estimates. The multiple correlation coefficients show the strength of the relationship between a dependent variable and two or more independent variables. The partial correlation coefficient signifies the strength of the relationship between a dependent variable and one or more independent variables with the effects of independent variables held constant.

In this study, the relationship between the evaporation and both water and air temperatures, net radiation, wind speed, and relative humidity were analyzed. MLR analysis was used to see the effects of variables on evaporation in Lake Nasser. Model summary and coefficients of the MLR analysis for Raft and Abusembel stations are illustrated in Tables 3-4 as selected examples from the seven stations.

J. Appl. Sci. Res., 8(5): 2785-2799, 2012

Table 3: Linear regression results for Abusembel Station (2010)

Variables	Coefficients		Correlation Coefficient
	Value	Standard Error	
Intercept	29.050	3120	
Air Temperature	0.717	23.165	0.576
Wind speed	0.176	3.841	0.735
Water Temperature	0.792	2.151	0.905
Net Radiation	0.106	0.038	0.692
Relative Humidity	-1.453	2.012	-0.265

Table 4: Linear regression results for Raft Station (2011)

Variables	Coefficients		Correlation Coefficient
	Value	Standard Error	
Intercept	97.123	73.115	
Air Temperature	0.821	7.456	0.542
Wind speed	0.381	3.541	0.681
Water Temperature	1.411	1.883	0.907
Net Radiation	0.017	0.013	0.652
Relative Humidity	-0.511	0.361	-0.321

When correlations between dependent and independent variables were examined; a high level relationship (r=0.905) and positive correlation between Lake Nasser evaporation values and temperature values was observed. When other variables were checked, it could be seen that there was a positive correlation between two variables, higher than others, but moderate (r=0.692) for net radiation and (r=0.735) for wind speed.. A low level relationship (r=-0.265) and negative correlation between Lake Nasser evaporation values and relative humidity values was observed. Temperature, wind speed, relative humidity, and net radiation variables together with Lake Nasser evaporation values has highly significant relationship (r=0.945). These variables explained the 94% of the total variance of measured evaporation in Lake Nasser. According to standardized regression coefficient, the order of the relative importance of the independent variables for changes in evaporation of Lake Nasser was temperature, wind speed, solar radiation, and relative humidity. Following the same methodology, regression analysis as in Tables 3-4 was performed for the other 5 stations in both 2010 and 2011. Regression equations were then developed and used to calculate evaporation values of 2010-2011 years. The average error between the calculated values and the HADA evaporation rates values was found to be 0.12 mm/day. Figs. 5-6 show the relation between HADA calculated evaporation and predicted evaporation for both Raft and Abusembel stations.

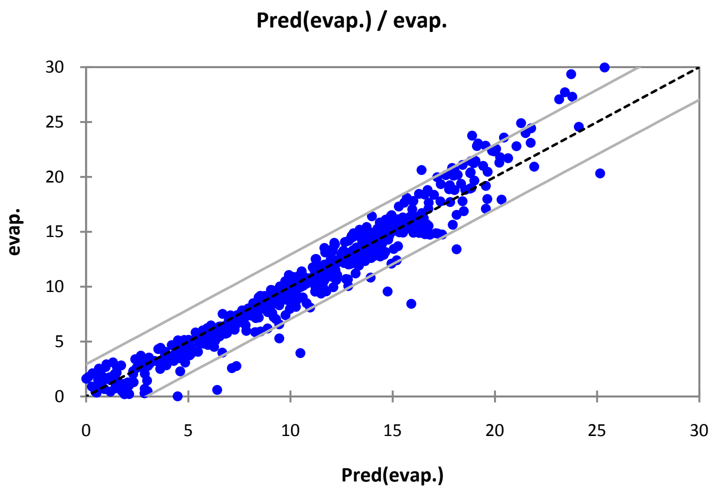

Fig. 5. HADA calculated evaporation vs. predicted evaporation at Abusembel Station in 2010

J. Appl. Sci. Res., 8(5): 2785-2799, 2012

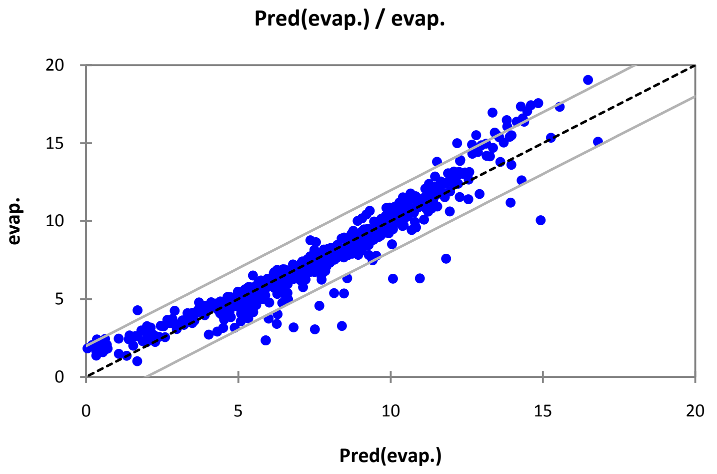

Fig. 6: HADA calculated evaporation vs. predicted evaporation at Raft Station in 2011

Artificial Neural Network:

In this study, a feed-forward neural network with back propagation learning algorithm is used. The basic element of a back-propagation neural network is the processing node. Each processing node behaves like a biological neuron and performs two functions. First, it sums the values of its inputs. This sum is then passed through an activation function to generate an output. Any differentiable function can be used as activation function, *f*. All the processing nodes are arranged into layers, each fully interconnected to the following layer. There is no interconnection between the nodes of the same layer. In a back propagation neural network, generally, there is an input layer that acts as a distribution structure for the data being presented to the network. This layer is not used for any type of processing. After this layer, one or more processing layers follow, called the hidden layers. The final processing layer is called the output layer. Fig. 7 shows the structure of a the ANN model used in this study.

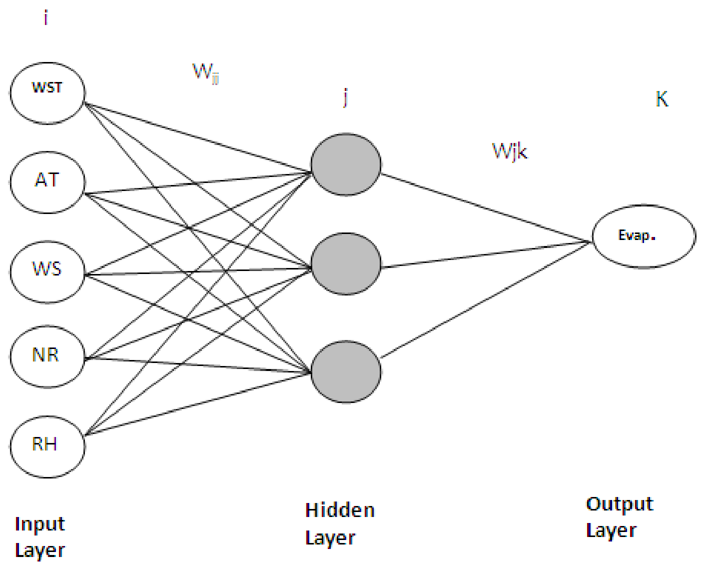

Fig. 7: Structure of the ANN model

To assess the usefulness of neural network in predicting the evaporation losses, Meteorological data from Lake Nasser consisting of 615 daily records from 2010 to 2011 were used to develop the model for daily evaporation estimation. The measured meteorological variables include daily observations of air and water temperature, sunshine hours, solar radiation, air pressure, relative humidity, and wind speed. The neural network is used to calculate correlation coefficient and RMSE by using cross-validation to generate the model on different combinations of the input data set in predicting the evaporation. Cross validation was used to train/test/validate the model. The cross-validation is a method of estimating the accuracy of a classification or regression model in which the input data set is divided into several parts with each part in turn used to test a model fitted to the remaining parts.

For this study, a ten-fold cross-validation was used. One of the important factors in using a neural network for prediction of evaporation requires setting up of the appropriate user defined parameters as the accuracy of a neural network model is largely dependent on the selection of the model parameters. In present study three hidden layers were used as it work well for this data set. Other user-defined parameters used were – momentum = 0.0, learning rate =0.3, hidden layer nodes = 6, training time = 600 and iterations = 1000. These values were obtained after a large number of trials by using different combination of these parameters carried out on used data set (Fig. 8). Both Abusembel and Aswan stations were selected as examples for the whole analysis of results. As shown in Tables 5-6, the first set of analysis was carried out by using ANN and linear regression with input as water temperature and output as evaporation loss. The ANN determined a relationship (i.e. create a model) between the input and the output of the available data set of any system. These models are then used to predict the output from the known input values of the same system, thus requiring sufficient number of data to create and test the models. A number of trials were carried out to reach at the various user defined parameters required for the neural network based algorithms using WEKA software (www.cs.waikato.ac.nz/ml/).

Results and Discussion

Two parameters namely correlation coefficient and RMSE values were used for the performance evaluation of the model and comparison of the results for prediction of evaporation. The higher value of correlation coefficient and the smaller value of RMSE mean a better performance of the model. The results of the neural network based modeling of evaporation using different combination of input parameters with the used data set are provided in Tables 5-6 in terms of correlation coefficient and RMSE. MLR was used and results in the form correlation coefficient and RMSE were obtained with different combinations of input parameters as well (Tables 5-6). As far as the significance of individual meteorological parameters is concerned, the study revealed that the highest value of correlation coefficient and least value of RMSE were obtained for evaporation with water temperature, air temperature, followed by wind speed, solar radiation and relative humidity. While the lowest correlation coefficient was obtained with relative humidity. The effect of water and air temperatures, wind speed, and solar radiation was found to be positive; whereas a negative correlation exists between evaporation and relative humidity (that is evaporation decreases with increase in relative humidity). These results from the application of ANN are in concurrence with the linear regression.

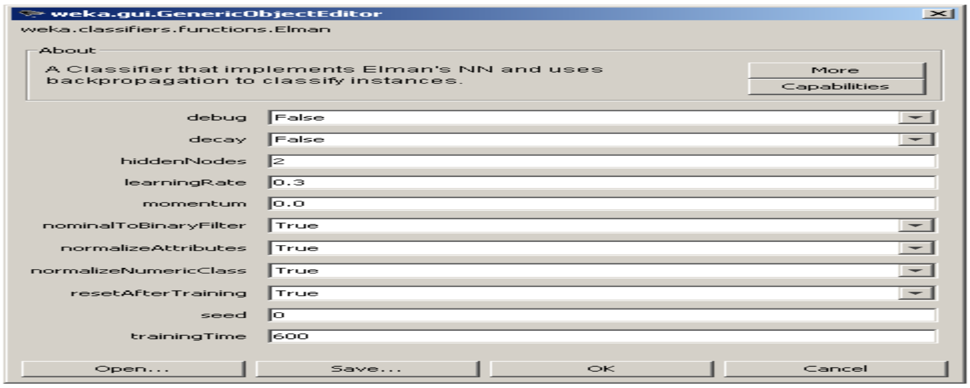

Fig. 8: ANN model parameters

Figs. 9 and 10 provide plot between HADA calculated values of evaporation loss and predicted values of evaporation from the combination of all the meteorological parameters taken together (WST +AT + WS + NR + RH) as inputs by ANN and MLR respectively. Figs. 9-10 indicates that most of the predicted values are lying within ±7% error from the line of perfect agreement with this combination of input parameters. Thus, suggesting the usefulness of all input parameters, instead of single parameter, in modeling the evaporation from a reservoir using ANN approach. The results suggest better performances by ANN as well as MLR approaches in comparison to the simple linear regression approach. Further, ANN is relatively more accurate than MLR in predicting evaporation losses in reservoirs from meteorological parameters (Tables 5-6).

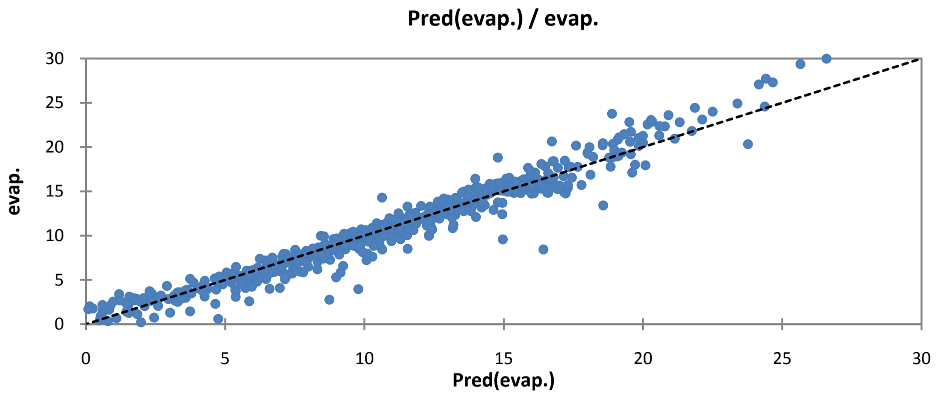

Fig. 9: Actual vs. ANN predicted evaporation at Abusembel Station in 2010.

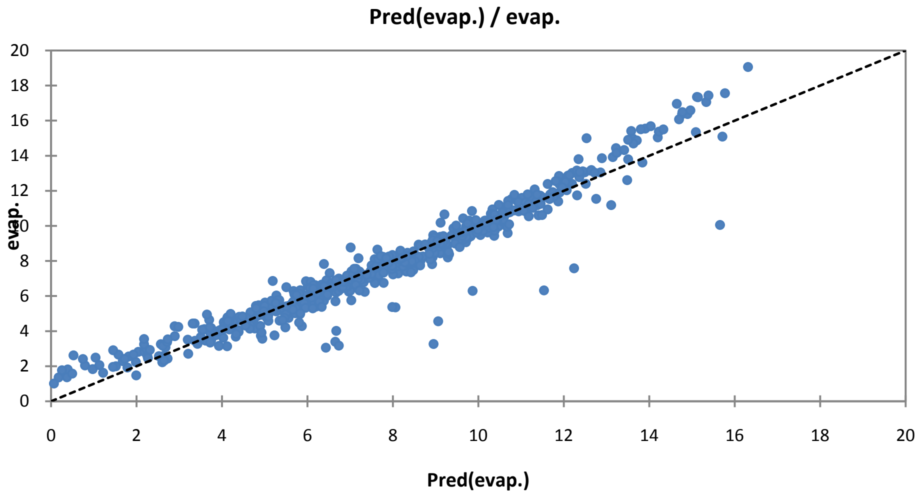

Fig. 10: Actual vs. ANN predicted evaporation at Raft Station in 2011

Table 5: ANN vs. linear and MLR at Abusembel in 2010

Input combinations	ANN		Linear Regression		MLR	
	Correlation Coefficient	RMSE	Correlation Coefficient	RMSE	Correlation Coefficient	RMSE
WST	0.911	1.412	0.905	1.662		
AT	0.652	1.271	0.576	2.231		
WS	0.634	2.423	0.735	2.845		
NR	0.182	3.313	0.692	2.749		
RH	-0.523	2.641	-0.265	2.531		
WST+TA	0.89	1.250			0.831	1.843
WST+WS	0.912	1.163			0.882	1.423
WST+NR	0.926	1.121			0.895	0.956
WST+TA+WS	0.953	0.984			0.914	0.961
WST+TA+WS+NR	0.958	0.956			0.942	0.984
WST+TA+WS+NR+RH	0.964	0.864			0.957	0.993

Figs. 11-12 were plotted when all five parameters were taken for model building and evaporation loss was predicted. The examination of Figs. 11-12 shows that there is better agreement of predicted results based on ANN rather than a linear regression. It is evident from Figs. 11-12 that more number of points are lying on the 45° line when the ANN was used to predict the evaporation in comparison to linear regression based algorithm for any combination of input parameters. The maximum value of correlation coefficient was obtained for water temperature followed by wind speed. Thus, it is quite obvious to consider the combined influence of all the parameters in evaporation estimation as indicated by higher values of correlation coefficient (r) and lower RMSE values.

Table 6: ANN vs. linear and MLR at Aswan in 2011

Input combinations	ANN		Linear Regression		MLR	
	Correlation Coefficient	RMSE	Correlation Coefficient	RMSE	Correlation Coefficient	RMSE
WST	0.923	1.452	0.907	1.692		
AT	0.671	1.283	0.542	2.253		
WS	0.639	2.354	0.681	2.946		
NR	0.193	3.146	0.658	2.716		
RH	-0.567	2.655	-0.321	2.614		
WST+TA	0.88	1.271			0.842	1.833
WST+WS	0.911	1.182			0.891	1.489
WST+NR	0.931	1.136			0.905	0.962
WST+TA+WS	0.948	0.983			0.913	0.958
WST+TA+WS+NR	0.956	0.958			0.943	0.987
WST+TA+WS+NR+RH	0.966	0.884			0.952	0.986

In view of the calculations carried out in sections 5.1, 5.2, and 5.3, the yearly average of the evaporation rate at each of the seven meteorological stations was calculated using the monthly average. Then, the total yearly average for the lake was computed by taking the average of the yearly average of the seven stations. The calculations revealed a rate of 7.64 mm day–1 as the annual mean evaporation rate with a maximum value in June of 10.63 mm/ day and a minimum in December of 4.25 mm/ day.

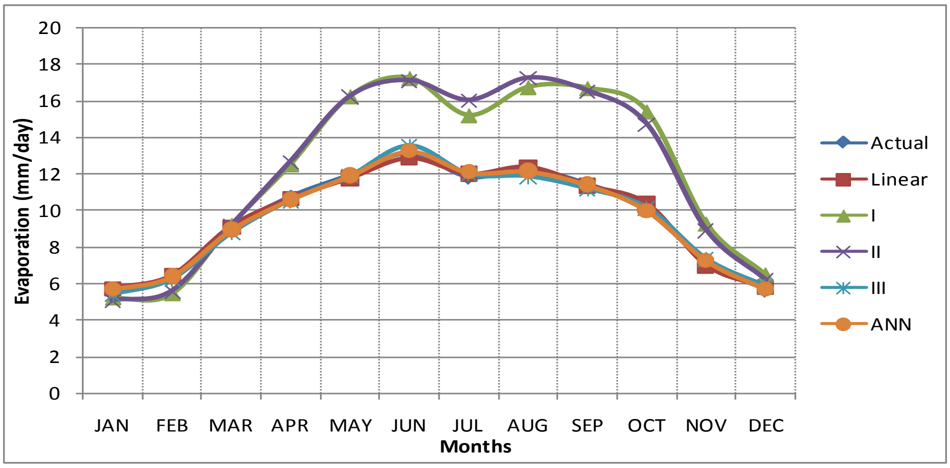

Fig. 11: Calculated evaporation rate at Abusembel Station in 2010 with all methods

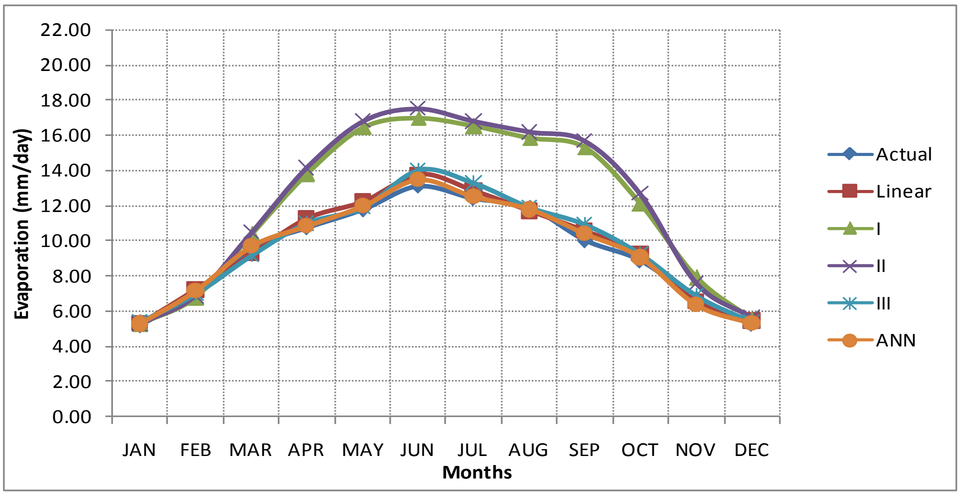

Fig. 12: Calculated evaporation rate at Raft Station in 2011 with all methods

Conclusions:

The comparison of results shows that there is a better agreement when large input parameters are considered for model building and testing as compared to a single parameter. The outcome of study suggests that the feed forward back propagation ANN based modeling can be applied as an alternative approach for estimation of daily evaporation from reservoirs effectively. A comparison among the conventional empirical methods delineated that Penman method is the most accurate among others and revealed evaporation values close to the value predicted by ANN based modeling.

ANN has been proposed and emerged as an alternative approach of evaporation estimation from a reservoir as compared to linear regression. The back propagation multilayer perception ANN and linear regressions based modeling techniques are performing better when all parameters are used as input for model building for the prediction of evaporation. Further, the critical examination of plotted figures indicates that the performance of the ANN modeling is data dependent to a great extent. The study also concludes that combination of all input parameters provides better performance of model in estimating the evaporation rather than individual parameters. The study revealed that evaporation loss is best estimated by ANN modeling rather than linear regression techniques. The outcome of the study provided an impetus to the potential use of ANN approach in predicting the evaporation from the reservoirs in water resources projects. Further, this study also concludes that most of the predicted values with ANN are lying near the 45 line and the scatter range is within ±7% line.

Acknowledgment

The authors are very grateful to High and Aswan Dams Authority staff for their cooperation and support throughout the research.

References

Abtew. W., 2001.Evaporation Estimation of Lake Okeechobee in South Florida. J. of Irrigation and Drainage, 127: 140-147.

Atkins, S., 2002. Comprehensive Development Plan for Aswan and Lake Nasser. UNDP, Ministry for Planning, Aswan Governorate.

Blanken, P.D., W.R. Rouse, A.D. Culf, C. Spence, L.D. Boudreau, J.N. Jasper, B. Kochtubajda, W.M. Scertzer, P. Marsh and D. Verseghy, 2000. Eddy covariance measurements of evaporation from Great Slave Lake, Northwest Territories, Canada. Water Resour Res., 36: 1069-1077.

Blodgett, T.A., J.D. Lenters and B.L. Isacks, 1997. Constraints on the origin of Paleolake expansions in the Central Andes. Earth Interact., 1: 1-28.

Bruton, J.M., R.W. McClendon and G. Hoogenboom, 2000. Estimating daily pan evaporation with artificial neural networks. Trans. ASAE., 43: 491-496.

Choudhury, B.J., 1999. Evaluation of an empirical equation for annual evaporation using field observations and results from biophysical model. J. of Hydrology, 216: 99-110.

Eichinger, W., J. Nichols, J. Prueger, L. Hipps, C. Neale, D. Cooper and A. Bawazir, 2003. Lake evaporation estimation in arid environments. IIHR-Hydroscience & Engineering, The University of Iowa. IIHR Report No. 430.

Hostetler, S.W. and P.J. Bartlein, 1990. Simulation of lake evaporation with application to modeling lake level variations of Harney–Malheur Lake, Oregon. Water Resour Res., 26: 2603-2612.

Ikebuchi, S., M. Seki and A. Ohtoh, 1988. Evaporation from Lake Biwa. J. of Hydrol., 102: 427-449.

Jeongkon, K. and S. Mohamed, 2002. Assessment of long-term hydrologic impacts of Lake Nasser and related irrigation projects in Southwestern Egypt. Elsevier Science, Journal of Hydrology, 262: 68-83.

Keskin, M.E. and O. Terzi, 2006. Artificial neural network models of daily pan evaporation. J. Hydrologic Engrg., 11: 65-70.

Laird, N.F. and D.A.R. Kristovich, 2002. Variations of sensible and latent heat fluxes from a great lakes buoy and associated synoptic weather patterns. J. of Hydrometeorol., 3: 3-12.

Morton, F.I., 1983. Climatological estimates of lake evaporation. Water Resour Res., 15: 64-76.

Myrup, L.O., T.M. Powell, D.A. Godden and C.R. Goldman, 1979. Climatological estimate of the average monthly energy and water budgets of Lake Tahoe, California–Nevada. Water Resour Res., 15: 1499-1508.

Omar, M.H. and M.M. El-Bakry, 1981. Estimation of evaporation from the lake of the Aswan High Dam (Lake Nasser) based on measurements over the lake. Agric Meteorol., 23: 293-308.

Ozlem, T. and M. Evolkesk, 2005. Modeling of daily pan evaporation. J.of Applied Sciences, 5: 368-372.

Rosenberry, D.O., T.C. Winter, D.C. Buso and G.E. Likens, 2007. Comparison of 15 evaporation methods applied to a small mountain lake in the Northeastern U.S.A. J. of Hydrology, 340: 149-166.

Sadek, M.F., M.M. Shahin and C.J. Stigter, 1997. Evaporation from the reservoir of the high Aswan dam, Egypt: a new comparison of relevant methods with limited data. Theor Appl Climatol., 56: 57-66.

Stannard, D.I. and D.O. Rosenberry, 1991. A comparison of short-term measurements of lake evaporation using eddy correlation and energy budget methods. J. of Hydrol., 122: 15-22.

Sudheer, K.P., A.K. Gosain, R.D. Mohana and S.M. Saheb, 2002. Modeling evaporation using an artificial neural network algorithm. Hyd. Processes, 16: 3189-3202.

Terzi, O. and M.E. Keski, 2005. Modeling of daily pan evaporation. J. of Applied Sciences, 5: 368-372.

Vallet, C., C. Legesse, D. Gasse, F. Tavi and T. Chernet, 2001. Lake evaporation estimates in tropical Africa (Lake Ziway, Ethopia). J. of Hydrology, 245: 1-18.

Whittington, D. and G. Guariso, 1983. Water management models in practice: a case study of the Aswan High Dam. Development in environmental modeling, 2. Elsevier, Amsterdam, p: 246.

Winter, T.C., 1981.Uncertainties in estimating the water balance of lakes. Water Resour, 17: 82-115.

Winter., T.C., D.O. Rosenberry and A.M. Sturrock, 1995. Evaluation of 11 equations for determining evaporation for a small lake in the north central United States. Water Resour Res., 31: 983-993.

Yu, Y.S. and H.V. Knapp, 1985. Weekly, monthly, and annual evaporations for Elk City Lake. J. of Hydrol., 80: 93-110.

Zupan, J., J. Gasteiger, 1993. Neural Networks for Chemists, VCH, New York.