



Journal of Applied Sciences

ISSN 1812-5654

science
alert

ANSI*net*
an open access publisher
<http://ansinet.com>

Modeling of Daily Pan Evaporation

¹Özlem Terzi and ²M. Erol Keskin

¹Faculty of Technical Education, Suleyman Demirel University, Isparta 32260, Turkey

²Faculty of Engineering-Architecture, Suleyman Demirel University, Isparta 32260, Turkey

Abstract: Daily pan evaporation estimations are achieved by a suitable Artificial Neural Network (ANN) model for the meteorological data recorded from Automated GroWeather meteorological station near Lake Eğirdir which lies in the Lake District of western 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 with up to 4 input neurons representing air and water temperatures, air pressure and solar irradiation. Prior to ANN model construction the classical correlation study indicated the insignificance of the wind speed and the relative humidity in the Eğirdir Lake 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 as 3 after various trial and error model running. The ANN model provides good estimations with the least Mean Square Error (MSE).

Key words: Artificial neural networks, daily pan evaporation, Lake Eğirdir

INTRODUCTION

Evaporation is the necessary components in any water balance assessments for different water resources planning, design, operation and management studies including hydrology, agronomy, forestry and land resources, irrigation management, river flow forecasting, investigation of lake ecosystem and modeling, etc. Among the components of the hydrological cycle, evaporation is perhaps the most difficult to estimate owing to complex interactions between the components of the land-plant-atmosphere system^[1]. This is further true for lakes such as in the eastern province of Turkey where naked land-water-mountain composition gives difficulties for detailed evaporation measurement records for long time periods. It is therefore necessary to develop approaches to estimate the evaporation rates from other available meteorology variables, which are comparatively easier for measurements. There are many available direct and indirect methods but it is always preferable to have evaporation measurements simultaneously with other relevant meteorological variables for the development of an effective evaporation model to estimate evaporation. The class A pan and eddy correlation are among the direct methods^[2]. Indirect methods include those that use meteorological data to estimate evaporation from other meteorological variables through empirically developed methodologies or statistical and stochastic approaches in

addition to mass-balance based formulations. Both direct and indirect methods have been used for evaporation estimation studies by many researchers^[3-8].

One of the recent digressions from the classically available approaches is the Artificial Neural Networks (ANN), which provides better modeling flexibility than the previous approaches with its successive adaptive features of error propagation where each meteorological variable takes its share proportionately. Numerous researchers have shown applicability of artificial neural networks in hydrological practices. For instance, solar radiation has been estimated using radial basis function and multilayer perceptron ANN^[9]. They have used latitude, longitude, altitude, sunshine hours and the month of the year as inputs in order to estimate the solar irradiation. The comparison of the model results to observed values indicates that the ANN with radial basis function provides satisfactory estimations. Tasadduq *et al.*^[10] have developed an ANN model for hourly mean ambient temperature prediction 24 h in advance. The comparison of predictions with the actual measurements favored the use ANN architecture in such prediction affairs, which encourages the use of ANN in hourly meteorological predictions. On the other hand, ANN models have also been used by many researchers to estimate short-term streamflow from meteorological variables in order to depict the rainfall-runoff relationship^[11-19].

The main aim of this study was to develop a suitable ANN model by considering the feed-forward back-propagation learning algorithm in the estimation of daily pan evaporation.

ARTIFICIAL NEURAL NETWORKS (ANNs)

Neural networks are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems. As in nature, the network function is determined largely by the connections between elements. A neural network can be trained to perform a particular function by adjusting the values of the connections (weights) between the elements. Commonly neural networks are adjusted, or trained, so that a particular input leads to a specific target output^[20].

Feed forward ANNs comprise a system of neurons, which are arranged in successive layers, namely input and output layers in addition to one or more hidden layers. The neurons in each layer are connected to the neurons in the subsequent layer by a weight *w*, which may be adjusted during training. A data pattern comprising the values *x_i* presented at the input layer *i* is propagated forward through the network towards the first hidden layer *j*. Each hidden neuron receives the weighted outputs *w_{ij}x_i* from the neurons in the previous layer. These are summed to produce a net value (NET_{*j*}), which is then transformed to an output value upon the application of an activation function^[11].

A typical three-layer feed-forward ANN is shown in Fig. 1. A typical ANN consists of three layers, namely input, hidden and output layers. Input layer neurons are *x₀*, *x₁*, *x₂*... *x_n*; hidden layer neurons are *h₁*, *h₂*... *h_n*; and finally output layer neurons are *o₁*, *o₂*... *o_n*.

A neuron consists of multiple inputs and a single output. The sum of the inputs and their weights lead to a summation operation as,

$$NET_j = \sum_{i=1}^n w_{ij}x_i \tag{1}$$

in which *w_{ij}* is established weight, *x_{ij}* is input value and NET_{*j*} is input to a node in layer *j*.

The output of a neuron is decided by an activation function. There are a number of activation functions that can be used in ANNs such as step, sigmoid, threshold, linear etc. The logistic sigmoid function, *f(x)*, commonly used, can be formulated mathematically as:

$$f(x) = 1/[1 + \exp(-x)] \tag{2}$$

$$OUTPUT_j = f(NET_j) = 1/[1 + \exp(-NET_j)] \tag{3}$$

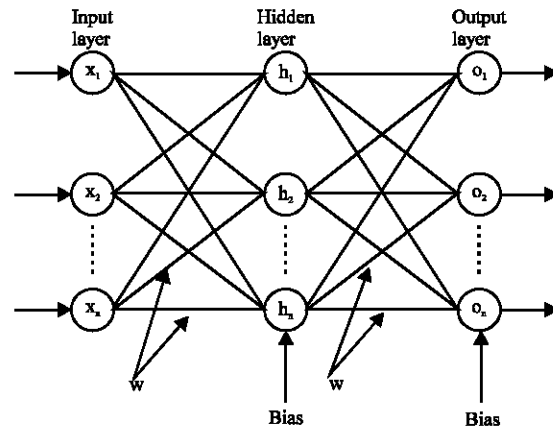


Fig. 1: A typical three-layer feed-forward ANN

The back-propagation learning algorithm is applied to multilayer feed-forward networks consisting of processing elements with continuous and differentiable activation functions. Given a training set of input-output pairs, the algorithm provides a procedure for changing the weights in a back-propagation learning algorithm to classify the given input patterns correctly. The basis for this weight update algorithm is simply the gradient-descent method as used for simple perceptrons with differentiable neurons.

For a given input-output pair, the back propagation algorithm performs two phases of data flow. First, the input pattern is propagated from the input layer to the output layer and, as a result of this forward flow of data, it produces an actual output. Then the error signals resulting from the difference between output pattern and an actual output are back-propagated from the output layer to the previous layers for them to update their weights^[21].

MATERIALS AND METHODS

Study region and data: Lake Egirdir (lat. 37.80° and 38.43°N, lon.30.30° and 31.37°E) is a freshwater lake located in Lakes District of Turkey which is the second largest freshwater lake in the country with a surface area and volume as 470 km² and 4360 hm³, respectively. It is being used as water supply and irrigation purposes. This Lake is of tectonic origin in the northern part of the Egirdir County. The altitude of the lake is about 916 m above mean sea level. Geographically, the lake lies on a 50 km stretch on the north-south direction. The distance between east and west shores is 3 km, at which the depth is around 1.8 m. The mean depth of the lake is 8 to 9 m and the deepest point is 15 m. In the southern part, the width of the lake reaches a maximum of 16 km.

Meteorological data for ANN model were obtained from an Automated GroWeather Meteorological Station near Lake Egirdir. Meteorological parameters included air and water temperature, relative humidity, solar radiation, wind speed and air pressure were logged. Class A pan evaporation values used as output in the ANN models are measured daily by XVIII. District Directorate of State Hydraulic Works. The data used to develop ANN models included 490 daily observations from March 1 to October 31, 2001 and 2002 years.

ANN EVAPORATION MODELS

Scatter diagrams are plotted between daily pan evaporation values and measured meteorological parameters in order to determine the dominating factors affecting evaporation in the lake location. Consideration of correlation coefficients indicates that the dominating factors affecting evaporation are their order of significance, air temperature (T_a), water temperature (T_w), solar radiation (R_C), air pressure (P_a), relative humidity (R_h) and wind speed (U_2). The R_h and U_2 parameters with the least effects are neglected in ANN evaporation model architecture. In order to estimate daily pan evaporation from Lake Egirdir, ANN evaporation models with two (T_a and T_w), three (T_a , T_w and R_C) and four input variables (T_a , T_w , R_C and P_a) are considered.

The adequacy of the ANN evaporation models was evaluated by estimating the coefficient of determination (R^2) defined based on the evaporation estimation errors as:

$$R^2 = \frac{E_o - E}{E_o} \tag{4}$$

where:

$$E_o = \sum_{i=1}^n (E_{i(Pan)} - E_{mean})^2 \tag{5}$$

$$E = \sum_{i=1}^n (E_{i(Pan)} - E_{i(simulated)})^2 \tag{6}$$

where, $E_{i(Pan)}$ and $E_{i(simulated)}$ are daily pan measurement and ANN model evaporation estimation values, respectively with the mean daily pan evaporation, E_{mean} . The Mean Square Error (MSE) is defined as

$$MSE = \frac{1}{n} \sum_{i=1}^n (E_{i(Pan)} - E_{i(simulated)})^2 \tag{7}$$

and used in order to decide about the best model, where n is the number of observed data.

In this study, ANN(i,j,k) indicates a network architecture with i , j and k neurons in input, hidden and output layers, respectively. Herein, i runs from 2, 3, to 4; j assume values of 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 and 12 where as $k = 1$ is adopted in order to decide about the best ANN model alternative. Prior to execution of the model, standardization of the data, X_i , ($i = 1, 2, \dots, n$) is done according to the following expression such that all data values fall between 0 and 1.

$$x_i = (X_i - X_{min}) / (X_{max} - X_{min}) \tag{8}$$

where, x_i is the standardized value but X_{max} and X_{min} are the maximum and minimum measurement values. Such standardization procedures renders the data also into dimensionless form.

An alternative model selection method, which is often employed in the ANN is the cross-validation. It does not remove the uncertainty completely by cross-validation. The motivation for this model selection is similar to the line of arguments leading to information criteria. Model complexity does not result necessarily in a better description of the underlying function due to increasing estimation error. In order to find an appropriate degree of complexity, it is appealing to compare the MSE of different model specifications. Such prediction errors are obtained by sampling the data into M subsets ($M = 5$ in the study), with n observations each. From the M available sets of observed data, $(M-1)$ are used to train the ANN. After the training is finished, the data set left out is predicted and the result is compared to the observed data. This procedure is repeated M times, once for each training data set. The average MSE on the M subsets that have been left out defines the cross-validation error. If, for instance, a large value of this error is obtained, the point excluded during the training process is important and its absence will produce an ANN with poor estimation and generalization capabilities. On the other hand, if the associated error is small, it means that the data set has enough support from its neighbors that its presence is not very important^[22,23].

For ANN models the number of hidden layers considered after trial and cross-validation is only one in all the structures proposed and the numbers of hidden neurons are obtained 5, 6 and 3. These structures are represented by ANN(2,5,1), ANN(3,6,1) and ANN(4,3,1), respectively. The learning rate and momentum parameters affect the speed of the convergence of the back-propagation algorithm. A learning rate of 0.001 and

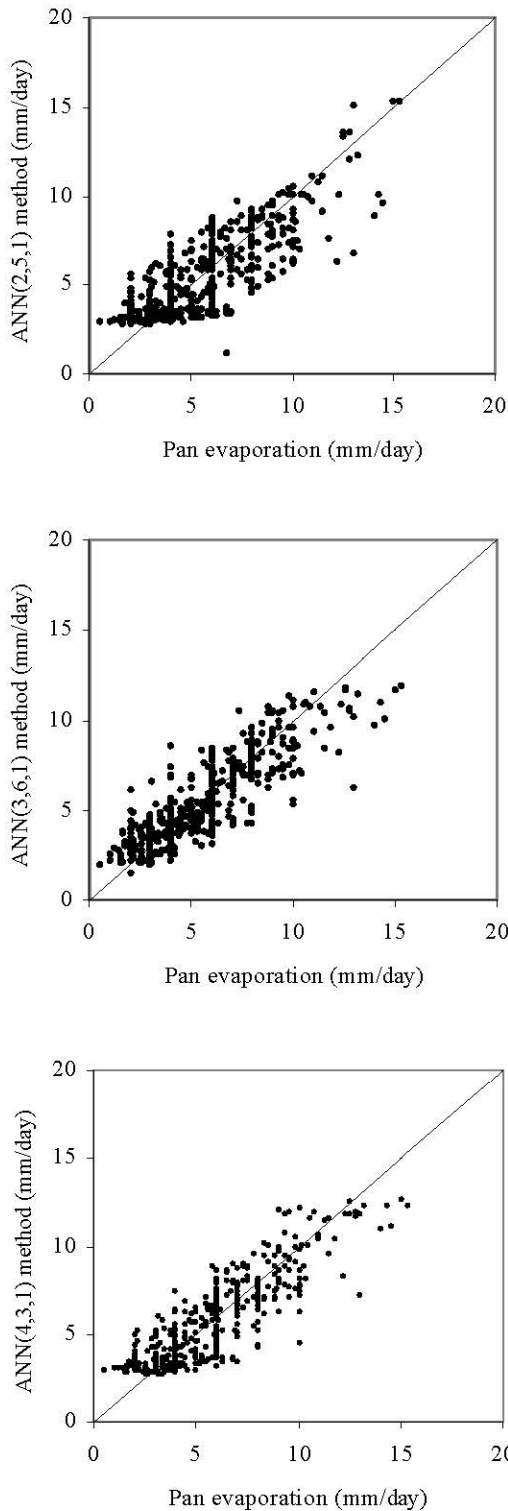


Fig. 2: Comparison daily pan evaporation with ANN(2,5,1), ANN(3,6,1) and ANN(4,3,1) models

momentum 0.1 are fixed for selected network after training and model selection is completed for year 2001. The

Table 1: R² and MSE values between ANN models and daily pan evaporation

Methods	2001 (Training)		2002 (Testing)	
	MSE	R ²	MSE	R ²
ANN (2.5.1)	2.52	0.68	2.60	0.65
ANN (3.6.1)	2.12	0.73	1.90	0.74
ANN (4.3.1)	1.97	0.75	1.61	0.77

trained networks are used to run a set of test data for year 2002. The performance indices reveal that the ANN(4,3,1) model is superior among ANN models. Comparing the performance of the ANN(4,3,1) and ANN(3,6,1) models, it can be observed that these models are performed in a more similar way. The difference between these two models is only in the input variables. The ANN(3,6,1) model depends on daily mean values of air temperature, water temperature and solar radiation, but ANN(4,3,1) is based on air pressure of three parameters. ANN(4,3,1) has better R² and lower MSE than ANN(3,6,1) model for both training and testing as shown in Table 1.

The performance of ANN(4,3,1) model suggests that the evaporation could be estimated easily from available data using ANN approach. The results of ANN models are plotted against measured daily pan evaporation in Fig. 2. ANN(4,3,1) model comparison plot is also around 45° straight lines which imply that there are no bias effects in the model. Therefore, ANN(4,3,1) is selected for daily pan evaporation estimation in this study. Formulations of ANN(4,3,1) model can be written as follows:

$$NET1 = -0,305906 \times Ta - 0,03648 \times Ts - 0,005218 \times Rc - 0,02011 \times Pa + 33,634106 \quad (9)$$

$$F1 = 1 / (1 + e^{-NET1}) \quad (10)$$

$$NET2 = 0,029463 \times Ta + 0,140478 \times Ts + 0,002973 \times Rc + 0,064378 \times Pa - 47,93451 \quad (11)$$

$$F2 = 1 / (1 + e^{-NET2}) \quad (12)$$

$$NET3 = 2,649693 \times Ta - 37,653328 \times Ts + 0,996404 \times Rc + 24,994041 \times Pa - 17240,51962 \quad (13)$$

$$F3 = 1 / (1 + e^{-NET3}) \quad (14)$$

$$NET4 = -3853,4516 \times F1 + 2,17 \times F2 - 0,66719 \times F3 + 3851,4621 \quad (15)$$

$$E = (1 / (1 + e^{-NET4}) + 0.033784) \times 14.8 \quad (16)$$

in which T_a: air temperature (°C), T_w: water temperature (°C), R_c: solar radiation (cal/cm²/day), P_a: air pressure (kPa), E: evaporation (mm/day), NET1-NET4 and

F1-F3 equations are summation and activation function of each neuron at hidden layer, respectively. Although these equations are obtained for Lake Eğirdir, it can be also adapted for various locations.

CONCLUSIONS

Estimating evaporation is important in hydrological practices. There are many available direct and indirect methods used to estimate evaporation. In addition to these methods, alternative model was proposed herein to estimate evaporation using ANN model. Comparing ANN models, ANN(4,3,1) model is superior among ANN models. The comparison shows that there is a better agreement between the results of ANN(4,3,1) model and pan evaporation values than others model. The evaporation could be estimated easily from available data using ANN approach. ANN model can be adapted to estimate, not to measure, daily pan evaporation for Lake Eğirdir where it is either not measured or the measurement system has failed. ANN model formulations can be developed changing input variables to estimate evaporation values in various location.

REFERENCES

1. Singh, V.P. and C.Y. Xu, 1997. Evaluation and generalization of 13 mass-transfer equations for determining free water evaporation. *Hydrological Processes*, 11: 311-323.
2. Ikebuchi, S., M. Seki and A. Ohtoh, 1988. Evaporation from Lake Biwa. *J. Hydrol.*, 102: 427-449.
3. Stewart, R.B. and W.R. Rouse, 1976. A simple method for determining the evaporation from shallow lakes and ponds. *Water Resources Res.*, 12: 623-627.
4. Warnaka, K. and L. Pochop, 1988. Analyses of equation for free water evaporation estimates. *Water Resources Res.*, 24: 979-984.
5. de Bruin, H.A.R., 1978. A simple model for shallow lake evaporation. *J. Applied Meteorol.*, 17: 1132-1134.
6. Andersen, M.E. and H.E. Jobson, 1982. Comparison of techniques for estimating annual lake evaporation using climatological data. *Water Resources Res.*, 18: 630-636.
7. Vallet-Coulomb, C., D. Legesse, F. Gasse, Y. Travi and T. Chernet, 2001. Lake evaporation estimates in tropical Africa (Lake Ziway, Ethiopia). *J. Hydrol.*, 245: 1-18.
8. Abtew, W., 2001. Evaporation estimation for Lake Okeechobee in South Florida. *J. Irrigation and Drainage Eng.*, 127: 140-147.
9. Dorvlo, A.S.S., J.A. Jervase and A. Al-Lawati, 2002. Solar radiation estimation using artificial neural networks. *Applied Energy*, 71: 307-319.
10. Tasadduq, I., S. Rehman and K. Bubshait, 2002. Application of neural networks for the prediction of hourly mean surface temperatures in Saudi Arabia. *Renewable Energy*, 25: 545-554.
11. Imrie, C.E., S. Durucan and A. Korre, 2000. River flow prediction using artificial neural networks: generalisation beyond the calibration range. *J. Hydrol.*, 233: 138-153.
12. Zealand, C.M., D.H. Burn and S.P. Simonovic, 1999. Short term streamflow forecasting using artificial neural networks. *J. Hydrol.*, 214: 32-48.
13. Luk, K.C., J.E. Ball and A. Sharma, 2000. A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting. *J. Hydrol.*, 227: 56-65.
14. Jervase, J.A., A. Al-Lawati and A.S.S. Dorvlo, 2002. Contour maps for sunshine ratio for Oman using radial basis function generated data. *Renewable Energy*, 28: 487-497.
15. Dibike, Y.B. and D.P. Solomatine, 2001. River flow forecasting using artificial neural networks. *Phys. Chem. Earth (B)*, 26: 1-7.
16. Braddock, R.D., M.L. Kremmer and L. Sanzogni, 1998. Feed-forward artificial neural network model for forecasting rainfall run-off. *Environmetrics*, 9: 419-432.
17. Mohandes, M., A. Balghonaim, M. Kassas, S. Rehman and T.O. Halawani, 2000. Use of radial basis functions for estimating monthly mean daily solar radiation. *Solar Energy*, 68: 161-168.
18. Mohandes, M., S. Rehman and T.O. Halawani, 1998. A neural networks approach for wind speed prediction. *Renewable Energy*, 13: 345-354.
19. Keskin, M.E., Ö. Terzi and D. Taylan, 2004. Fuzzy logic model approaches to daily pan evaporation estimation in western Turkey. *Hydrol. Sci., J.*, 49(6) (In press).
20. Demuth, H. and M. Beale, 2001. *Neural Network Toolbox User' Guide-version 4*.
21. Lin, C.T. and C.S.G Lee, 1995. *Neural Fuzzy Systems*. Prentice Hall.
22. Anders, U. and O. Korn, 1999. Model selection in neural networks. *Neural Networks*, 12: 309-323.
23. Sudheer, K.P., A.K. Gosain, R.D. Mohana and S.M. Saheb, 2002. Modelling evaporation using an artificial neural network algorithm. *Hydrol. Processes*, 16: 3189-3202.