International Journal of Agricultural Science, Research and Technology Available online on: www.ijasrt.com

Daily Pan Evaporation Estimation Using Artificial Neural Network-based Models

Karimi-Googhari, Sh

Assistant Professor, Department of Water Engineering, Faculty of Agriculture, Shahid-Bahonar University of Kerman, Kerman, I.R. of Iran
(E-mail: shkarimi1352@gmail.com; Tel: 0098-341-3202660)

Received: 25 November 2011, Reviewed: 5 December 2011, Revised: 10 December 2011, Accepted: 20 December 2011

Abstract

Accurate estimation of evaporation is important for design, planning and operation of water systems. In arid zones where water resources are scarce, the estimation of this loss becomes more interesting in the planning and management of irrigation practices. This paper investigates the ability of artificial neural networks (ANNs) technique to improve the accuracy of daily evaporation estimation. Four different ANNs model comprising various combinations of daily climatic variables, that is, air temperature, daily sunshine hours, wind speed, and relative humidity are developed to evaluate degree of effect of each mentioned variables on evaporation for two stations located in central part if I.R. of Iran. A comparison is made between the estimates provided by the ANNs model and the multiple linear regression models. Various statistic measures are used to evaluate the performance of the models. Based on the comparisons, it was revealed that the ANNs computing technique could be employed successfully in modeling of evaporation process from the available climatic data. The ANN also increased dramatically the accuracy of evaporation estimation compare to the multiple linear regression models. **[SH, Karimi-Googhari. Daily Pan Evaporation Estimation Using Artificial Neural Network-based Models. International Journal of Agricultural Science, Research and Technology, 2011; 1(4):159-163].**

Key words: Evaporation, irrigation, estimation, neural networks, arid region

1. Introduction

Evaporation (E) is a complex and non-linear process since it depends on several interacting factors, such as sunshine hour, temperature, humidity and wind speed. Accurate estimation of E is of paramount importance for many studies, such as hydrologic water balance, irrigation system design and management, and water resources planning and management. For many years engineers and researchers have used loss from evaporation pans, multiplied by a coefficient applicable to the particular pan, as an estimate of the evaporation loss from reservoirs. The most widely used pan is the US Weather Bureau Class A pan which is 4 ft in diameter and 10 in. deep and is mounted on a timber grill about 6 inches above the soil surface. Pan evaporation has been widely used as an index of evapotranspiration and for estimating lake and reservoir evaporation (Kisi, 2006).

It is impractical to place evaporation pans at every point where there is a planned or existing reservoir and irrigation project. It is also highly unlikely to have in inaccessible areas where accurate instruments cannot be established or maintained. A practical means of estimating the amount of pan evaporation where no pans are available is of considerable significance to the hydrologists, agriculturists, and meteorologists. A number of methods have developed to estimate the evaporation values from climatic variables and most of these methods require data that are not easily available (Stephens and Stewart, 1963; Reis and Dias, 1998; Irmak et al, 2002; Gavin and Agnew, 2004). Simple methods that are reported (Stephens and Stewart, 1963) try to fit a linear relationship between the variables. However, the process of evaporation is highly non-linear in nature, as it is evidenced by many of the estimation procedures. Many researchers have emphasized the need for accurate estimates of evaporation in hydrologic modeling studies (Sudheer et al, 2002). This requirement could be addressed through better models that will address the inherent non-linearities in the process.

Artificial neural networks (ANNs) are empirical models, which are quite perfect in modeling complex nonlinear phenomenon. ANN architecture is a massively parallel-distributed information processing system that has certain performance characteristics resembling the biological arrangement of neurons in human brain. The neural network typically consists of an input layer, an output layer and at least one layer of nonlinear processing elements, known as the hidden layer. In many respects, ANNs are similar to regression-based models in hydrology, except that ANNs do not require specification of a mathematical form (Hsu et al., 1995). Neural networks approaches have been successfully applied in a number of diverse fields, including water resources. In the hydrological context, recent experiments have reported that artificial neural network (ANN) may offer a promising alternative (Kisi, 2005; Kumar et al, 2004; Kisi, 2005; Supharatid, 2003). Even when there are missing data values, ANN methods can be applied for infilling missing hydrological records (Khalil et al., 2001). Some authors have compared Box–Jenkins and ANN methods (Hsu et al., 1995) confirming, in most cases, the superiority of ANNs. Using ANNs models for short term river flows forecasting have been investigated in the Xallas river (Castellano, 2004), Winipeg river system (Zealand, 1999) and Geer catchment (Vos and Rientjes, 2005) and all of them concluded the feed forward ANNs could forecast the streamflows accurately. However, the application of ANN to evaporation modeling is limited in the literature. To the knowledge of the author, no study has been carried out to utilize the input–output mapping capability of artificial neural network technique in evaporation modeling in arid zones of Iran. This provided an impetus for the present investigation. The potential of the ANNs_based model for estimation of the evaporation using climatic variables is investigated and discussed in the study. The performance of the ANN is compared with multiple regression method.

2. Materials and methods

Artificial neural networks (ANNs) are essentially semi-parametric regression estimators and are suitable for simulate the behavior of complicate physical phenomena. A significant advantage of the ANN approach in system modeling is that there is no need to have a well-defined physical relationship for systematically converting an input to an output. It is needed for most networks provide a collection of representative examples (input–output pairs) of the desired mapping. The ANN then adapts itself to reproduce the desired output when presented with training sample inputs. One of the most popular ANN architectures is multi layer perceptron (MLP). A typical MLP has neurons arranged in a distinct layered topology, as shown in Figure 1. The input layer simply sends the values of the input variables into the hidden layer. The hidden and output-layer neurons are fully connected to all of the units in the preceding layer. Each hidden neuron in an ANN receives a number of inputs from original data or other layer nodes. Each input comes via a connection that has a strength (or weight) attached. The weighted sum of the inputs is formed, to compose the activation of the neuron. The activation signal is passed through an activation (transfer) function produce the output of the neuron. A feed-forward MLP network, where nodes in one layer are only connected to nodes in the next layer, was used for modeling.

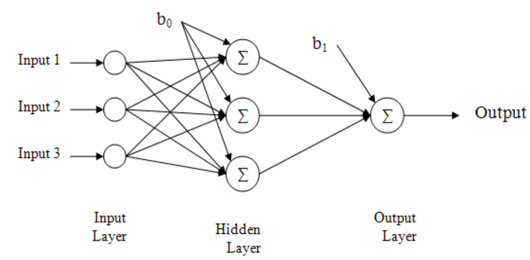

Figure 1: Structure of a typical MLP

Network geometry determines the number of connection weights and how these are arranged. This is generally done by fixing the number of hidden layers and choosing the number of nodes in each of these. It has been shown that ANNs with one hidden layer can approximate any function. The number of nodes in the input layer is fixed by the number of model inputs, whereas the number of nodes in the output layer is equal to the number of model outputs. In this study, there was only one output (reservoir inflow). For selecting the final structure of the ANN model, it being a trial and error procedure, started with a minimum number of nodes in the hidden layer and the network was trained until a minimum mean square error will be attained. The number of nodes in a hidden layer has been increased gradually until such an increase did not significantly improve the performance of the neural network. The process of optimizing the connection weights is known as 'training' or 'learning'. Here, the Levenberg–Marquardt backpropagation training (LMBP) has been used for train a Feed-forward Neural Network (FNN) where nodes in one layer are only connected to nodes in the next layer. The sigmoidal transfer functions that are most common, was used.

3. Results and discussion

The daily climatic data of two weather stations, Esfehan Station (latitude 32° 37′ N, longitude 51° 40′ E) and Kashan station (latitude 33° 59′ N, longitude 51° 27′ E) operated by Iranian meteorological organization are used in the study. The elevations are 1550 and 982 m for the Esfehan and Kashan Stations, respectively. The data sample consisted of ten years (1995–2005) of daily records of air temperature (T), sunshine hour (SH), wind speed (W), relative humidity (RH) and pan evaporation (E). For each station, the 70 percent if data were used to train the models and the remaining data were used for validation and testing. The statistics properties of collected data are presented in Table 1.

A- Kashan station

The weather parameters considered in this study were the air temperature (T), sunshine hour (SH), wind speed (W), and humidity (H). The study examined various combinations of these parameters as inputs to the ANN models so as to evaluate the degree of effect of each of these variables on evaporation. Building the ANN model several times with one different variable added into the input combination per time. Thus, the input combinations evaluated in the present study are: (i) T; (ii) T and SH; (iii) T, SH and W; (iv) T, SH, W and RH.

The ANN model implementation was carried out using the MATLAB routines. The connection weights, threshold and the number of neurons in the hidden layer, which can be interpreted as the model parameters, were adjusted during the calibration or training process through minimization of the mean square error (MSE) using the *Traingdm* function in MATLAB which is based on the gradient descent method using the error back-propagation algorithm whit considering the momentum. Each network has been trained many times to obtain minimum of MSE for a fixed number of neurons in hidden layer. Each network was trained at least 15 times and the number of neurons was varied between 2 to 10. The best model was selected for each subset using raw and transformed data and results were compared. The performance of selected models which were developed in this study was evaluated by using a variety of standard statistical performance evaluation measures. Specifically, 4 different statistical performance indices have been employed: average absolute relative error (AARE), Pearson's correlation coefficient (R), Nash–Sutcliff efficiency (CE), and normalized root mean square error (NRMSE). For more information about these indices you can refer to Karimi and Lee (2011). The final architectures of the ANN models and R, AARE, CE and NRMSE statistics of each ANN model in train and test period for the Kashan station are given in Table 2. Result indicates that the number of neurons in hidden layer is more than input variables for all models. The table 2 indicates that the ANN model whose inputs are the T, SH, W and RH (EANN4) has the smallest AARE (0.287), NRMSE (19.96), the highest R (0.0.918) and CE (0.843) during training process. This emphasizes the factors influencing evaporation, since the model considered all the parameters. The statistics increase non-significant during test process with superiority of KANN4 model like training process. In order to assess the ability of ANN model compare to the multiple regression model a linear model is developed using the KANN4 model inputs. The linear model for Kashan station is (KMLR):

$$E=-1.17+0.402(T)+0.055(SH)+0.353(W)+0.022(RH) \quad (1)$$

(R=0.9)

The R in training period for linear model using all four input variables is equal to KANN1 which uses only temperature variable. The results of estimating evaporation for testing period are illustrated in figure 2 and 3 for neural network and linear models. Results show that the KANN1 model performs better than KMLR model.

Table 1. The statistics properties of the data

Parameter	Unit	Kashan station					Esfehan station				
		Ave.	Min.	Max	Std.	Skewness	Ave.	Min	Max	Std.	Skewness
T	˚C	23.36	0.1	38.5	8.22	-0.441	20.1	0.1	38.4	7.15	-0.4
H	%	33.98	7	93	15.8	1.217	29.83	8	88	13.5	1.33
SH	hour	9.3	0.1	13.5	2.98	-1.39	10.16	0.1	13.9	2.64	-1.63
W	m/sec	0.348	0	4.59	0.6	2.45	1.16	0.1	9.95	1.2	1.56
E	mm	8.02	0.1	17.3	4.14	-0.092	8.007	0.1	18.2	3.84	-0.04

T=Air Temperature, SH=Sunshine Hour, W=Wind Speed, RH=Relative Humidity, E=Pan Evaporation

Table 2. Final architectures of the ANN models and their statistics values for Kashan station

Model Inputs	Best Architecture	Name	R	CE	AARE	NRMSE
During Training						
T	1-6-1	KANN1	0.903	0.816	0.35	21.57
T, SH	2-6-1	KANN2	0.915	0.836	0.333	20.39
T, SH, W	3-4-1	KANN3	0.914	0.833	0.334	20.6
T, SH, W, RH	4-7-1	KANN4	0.918	0.843	0.287	19.96
During Testing						
T	1-6-1	KANN1	0.914	0.83	0.43	21.3
T, SH	2-6-1	KANN2	0.92	0.845	0.418	20.59
T, SH, W	3-4-1	KANN3	0.923	0.847	0.416	20.43
T, SH, W, RH	4-7-1	KANN4	0.923	0.848	0.36	20.37

T=Air Temperature, SH=Sunshine Hour, W=Wind Speed, RH=Relative Humidity, E=Pan Evaporation R= Pearson's Correlation Coefficient, CE= Nash–Sutcliff Efficiency, AARE= Average Absolute Relative Error, NRMSE= Normalized Root Mean Square Error

Table 3. The final architectures of the ANN models and their statistics values for Esfehan station.

Model Inputs	Best Architecture	Name	R	CE	AARE	NRMSE
During Training						
T	1-7-1	EANN1	0.823	0.68	0.413	26.91
T, SH	2-5-1	EANN2	0.845	0.712	0.382	25.41
T, SH, W	3-5-1	EANN3	0.856	0.733	0.37	24.4
T, SH, W, RH	4-7-1	EANN4	0.868	0.754	0.345	23.5
During Testing						
Model Inputs	Best Architecture	Name	R	CE	AARE	NRMSE
T	1-7-1	EANN1	0.817	0.642	0.453	27.43
T, SH	2-5-1	EANN2	0.842	0.69	0.405	25.52
T, SH, W	3-5-1	EANN3	0.838	0.613	0.46	28.5
T, SH, W, RH	4-7-1	EANN4	0.85	0.695	0.40	24.53

T=Air Temperature, SH=Sunshine Hour, W=Wind Speed, RH=Relative Humidity, E=Pan Evaporation
R= Pearson's Correlation Coefficient, CE= Nash–Sutcliff Efficiency, AARE= Average Absolute Relative Error, NRMSE= Normalized Root Mean Square Error

The linear model for Esfehan station (EMLR) is:

$$E= 0.541+0.35 (T)+0.102(SH)+0.508(W)+0.046(RH) \quad (2) \quad (R=0.827)$$

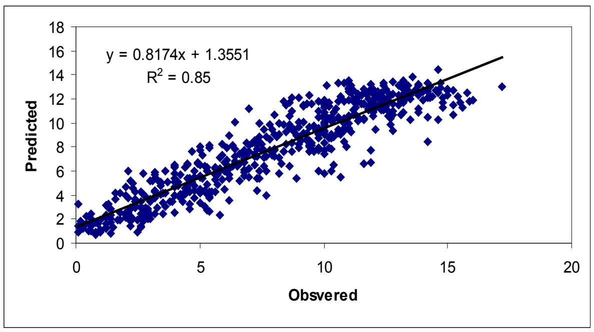

Figure 2: The observed and estimated evaporation of Kashan station in the test period using KANN4 model

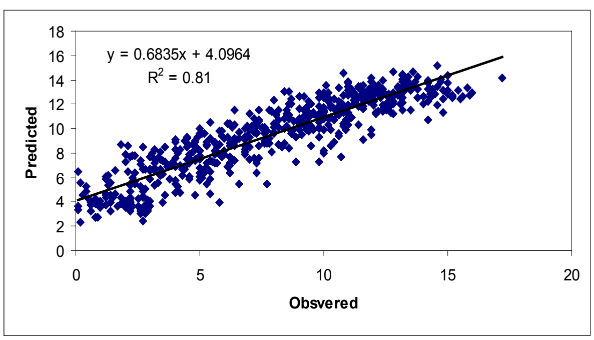

Figure 3: The observed and estimated evaporation of Kashan station in the test period using linear model

B- Esfehan station

The final architectures of the ANN models and R, AARE, CE and NRMSE statistics of each ANN model in train and test period for the Esfehan Station are given in Table 3. This table indicates the number of neurons in hidden layer of each input variable. The number of neurons in hidden layer is more than input variables for all models. The table indicates that the ANN model whose inputs are the T, SH, W and RH (EANN4) has the smallest AARE (0.345), NRMSE (23.5), the highest R (0.868) and CE (0.754). This emphasizes the factors influencing evaporation, since the model considered all the parameters.

The R in training period for linear model using all four input variables is equal to EANN4 which uses same variables. The results of estimating evaporation for testing period are illustrated in figure 3 and 4 for neural network and linear models. Results show that the both models, EANN4 and linear model, could not estimate evaporation as well as Kashan station. The generalization of EANN4 is superior and in term of using less input data the ANN models could perform better.

Although this study was locally applied and its results could not be implemented to other locations, the ANNs model with temperature-based data as only climatology variable which can be found in any area, could be used in irrigation management problem when sufficient or reliable data were not available to estimate reference evapotranspiration Models or evaporation from water bodies. The ANN technique could also be of use in water budgeting of basins, design of reservoirs and various other hydrological analyses where other models may be inappropriate. The study only used data from two areas and further work using more data from various areas may be required to strengthen the results of this study.

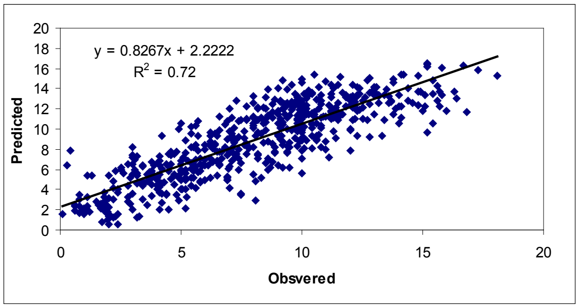

Figure 4: The observed and estimated evaporation of Esfehan station in the test period using EANN4 model.

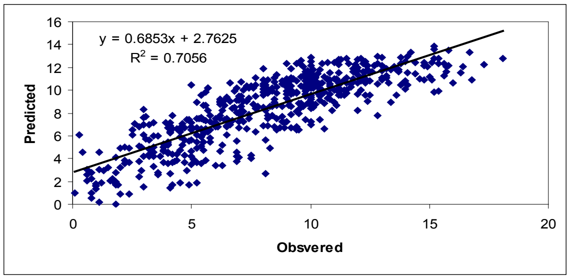

Figure 5: The observed and estimated evaporation of Esfehan station in the test period using linear model.

4. Conclusion and Recommendations

The present study demonstrates the capabilities of artificial neural networks technique (ANNs) for evaporation modeling, however the choice of ANNs architecture and input parameters are crucial for obtaining good estimate accuracy. The ANN model whose inputs are the air temperature, sunshine hour, wind speed, and relative humidity performed the best among the input combinations tried in the study. Results indicate that all these variables are needed for better evaporation modeling. It was found that using only the air temperature input gives poor estimates. In order to assess the ability of ANN model relative to that of the linear regression technique, the comparison was made. The ANN model in both stations that using air temperature, sunshine hour, wind speed, and humidity were found to perform better than the multiple linear regression model that using same meteorological variables.

References

1. Castellano, M. M., González, M . W., Febrero, B. B., Prada, S. J. M and Lozano, C. R. (2004). Modeling of the monthly and daily behavior of the runoff of the Xallas River using Box–Jenkins and neural networks methods. Journal of Hydrology, 296: 38-55.
2. Gavin, H and Agnew, C. A. (2004). Modeling actual, reference and equilibrium evaporation from a temperate wet grassland. Hydrological Processes, 18: 229–246.
3. Hsu, K., Gupta, H. V and Sorooshian, S. (1995). Artificial neural network modeling of the rainfall-runoff process. Water Resources Research, 31(10):2517–2530.
4. Irmak, S., Haman, D and Jones, J. W. (2002). Evaluation of class ‘A’ pan coefficients for estimating reference evapotranspiration in a humid location. Journal of Irrigation and Drainage Engineering, ASCE 128(3): 153–159.
5. Karimi-Googhari, Sh and Lee T. S. (2011). Applicability of Adaptive Neuro-Fuzzy Inference Systems in Daily Reservoir Inflow Forecasting. International Journal of Soft Computing, 6: 75-84.
6. Khalil, M., Panu, U. S and Lennox, W. C. (2001). Groups and neural networks based stream flow data infilling procedures. Journal of Hydrology, 241:153–176.
7. Kisi, O. (2005). Daily river flow forecasting using artificial neural networks and auto-regressive models. Turkish Journal of Engineering and Environmental Sciences, 29: 9–20.
8. Kisi, O. (2006). Daily pan evaporation modelling using a neuro-fuzzy computing technique. Journal of Hydrology, 329: 636– 646.
9. Kumar, D. N., Raju, K. S and Sathish, T. (2004). River flow forecasting using recurrent neural networks. Water Resources Management, 18: 143–161.
10. Reis, R. J and Dias, N. L. (1998). Multi-season lake evaporation: energy budget estimates and CRLE model assessment with limited meteorological observations. Journal of Hydrology, 208: 135–147.
11. Stephens, J. C and Stewart, E. H. (1963). A comparison of procedures for computing evaporation and evapotranspiration. Publication 62, International Association of Scientific Hydrology. International Union of Geodynamics and Geophysics, Berkeley, CA, 123–133.
12. Sudheer, K. P., Gosain, A. K., Rangan, D. M and Saheb, S. M. (2002). Modeling evaporation using an artificial neural network algorithm. Hydrological Processes, 16: 3189–3202.
13. Supharatid, S. (2003). Application of a neural network model in establishing a stage–discharge relationship for a tidal river. Hydrological Processes, 17: 3085–3099.
14. Vos, N. J and Rientjes, T. H. (2005). Constraints of artificial neural networks for rainfall-runoff modeling: trade-offs in hydrological state representation and model evaluation. Hydrology and Earth System Sciences, 2: 365–415.
15. Zealand, C. M., Burn, D. H and Simonovic SP (1999). Short term streamflow forecasting using artificial neural networks. Journal of Hydrology, 214: 32–48.