Journal of Agricultural Science; Vol. 6, No. 3; 2014
ISSN 1916-9752 E-ISSN 1916-9760
Published by Canadian Center of Science and Education

Hybrid of Artificial Neural Network-Genetic Algorithm for Prediction of Reference Evapotranspiration (ET_0) in Arid and Semiarid Regions

Shafika Sultan Abdullah[1,5], M. A. Malek[2], A. Mustapha[3] & Alihosein Aryanfar[4]

[1] Department of Civil Engineering, Universiti Tenaga Nasional, Malaysia

[2] The Institute of Energy, Policy and Research (IEPRe), Universiti Tenaga Nasional, Malaysia

[3] Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Malaysia

[4] Islamic Azad University, Zahed Shahr Branch, Iran

[5] Akre Technical Institute, Dohuk polytechnic, Dohuk, Iraq

Correspondence: Shafika Sultan Abdullah, Department of Civil Engineering, Universiti Tenaga Nasional, Malaysia; Akre Technical Institute, Dohuk polytechnic, Dohuk, Iraq. Tel: 60-14-934-1736. E-mail: sha_akre@yahoo.com

Received: November 22, 2013 Accepted: December 20, 2013 Online Published: February 15, 2014

doi:10.5539/jas.v6n3p191 URL: http://dx.doi.org/10.5539/jas.v6n3p191

Abstract

Evapotranspiration is a principal requirement in designing any irrigation project, especially in arid and semiarid regions. Precise prediction of Evapotranspiration would reduce the squandering of huge quantities of water. Feedforward Backpropagation Neural Network (FFBPNN) model is employed in this study to evaluate the performance of Artificial Neural Networks (ANNs) in comparison with Empirical FAO Penman-Monteith (P-M) Equation in predicting reference evapotranspiration (ET_0); later, a hybrid model of ANN-Genetic Algorithm (GA) is proposed for the same evaluation function. Daily averages of maximum air temperature (T_{max}), minimum air temperature (T_{min}), relative humidity (R_h), radiation hours (R), and wind speed (U_2) from Mosul station (Nineveh, Iraq) are used as inputs to the ANN simulation model to predict ET_0 values obtained using P-M Equation. The main performance evaluation functions for both models are the Mean Square Errors (MSE) and the Correlation Coefficient (R^2). Both models yield promising results, but the hybrid model shows a higher efficiency in prediction of Evapotranspiration and could be recommended for modeling ET_0 in arid and semiarid regions.

Keywords: evapotranspiration, FAO Penman-Monteith Equation, artificial neural network, genetic algorithm

1. Introduction

Increasing demand for water and scarcity of water supply are growing concerns in both arid and semiarid regions of the world. Iraq, which consists of both arid and semiarid climates, suffers from a low rainfall rate of about 160 mm per year and tremendous amount of water loss from evaporation and transpiration (Evapotranspiration). The condition is further aggravated by other climate events such as drought and salinity (Kerr, 1998). Native date palm, cotton, barley, and wheat are common products in Iraq that depend on irrigation because of scarcity of rain water. An understanding of the precise prediction of Evapotranspiration would allow for optimization of water use in irrigation projects.

The specific concept of Evapotranspiration (ET) is influenced by alteration of weather parameters, crop types, stage of growth, and other environmental conditions. The precise determination of ET produced the need for another comprehensive concept, namely, Reference Evapotranspiration (ET_0), which can be defined as “the rate of Evapotranspiration from an extensive surface of 8 to 15 cm tall, green grass cover of uniform height, actively growing, completely shading the ground and not short of water” (Doorenbos & Pruitt, 1977). The extensive surface resembles the reference surface indicated by FAO experts as the “A hypothetical reference crop with an assumed crop height of 0.12 m, a fixed surface resistance of 70 s m^{-1} and an albedo of 0.23” (Allen et al., 1998). Methods for predicting evapotranspiration are either Direct or Indirect methods. Direct Methods, such as Lysimeters (Wright, 1988), or field experiments, are mostly for scientific researches. Indirect Methods are usually empirical and depend on weather parameters such as solar radiation; air temperature; air humidity; and wind speed, measured or estimated at meteorological stations. Many empirical methods have been studied, issued, and applied in the prediction of ET, but the Food and Agriculture Organization (FAO) of the United Nations assumes that the combination of energy balance/aerodynamic equations provides the most accurate results for prediction of ET_0, and it adopted the FAO Penman-Monteith (P-M) Equation as the only standard equation for estimation of ET_0 (Allen et al., 1998).

Artificial Intelligence (AI) applications distinguished itself in many scientific fields during the last decades; however, researchers in hydrology, among many other fields, are still attracted to apply Artificial Neural Networks (ANNs) in modeling stream flow, rainfall, suspended sediment and ET. Kumar et al. (2002) found that application of Multiple Layers Perception (MLP) of ANN, with backpropagation algorithm; gives an accurate estimation of ET_0. In 2005, Trajkovic approved the possibility of achieving reliable results of ET_0 on the basis of temperature data only; FAO P-M Equation is used in comparison with another three temperature-based empirical equations and Radial-Based Function (RBF) Network. The RBF model better predicted FAO P-M ET_0 than the other calibrated empirical methods. Kisi (2006) investigated the possibility of modeling ET_0 using the technique of Generalized Regression Neural Network. The results of the intelligent model were successful. In 2007, Kisi and Öztürk investigated the accuracy of Adaptive Neurofuzzy Inference System Models (ANFIS) in modeling ET_0. Values of ET_0 were obtained using FAO P-M Equation with four years records of daily climate parameters of Pomona Station and Santa Monika Station, operated by the California Irrigation Management Information System (CIMIS), the results were compared with ANFIS and ANN, and also in comparison with Hargreaves and Ritchie empirical methods. The comparative results proved the superiority of ANFIS with inputs of T, U_2, RH and R in modeling daily ET_0 over the ANN and empirical methods. Kisi (2008) examined the accuracy of three ANN techniques, namely, the Generalized Regression (GR), MLP and Radial Basis Neural Networks (RBNNs) in a model of P-M Evapotranspiration. Results proved that both MLP and RBNN techniques could be successfully used in modeling ET_0. Landeras et al. (2008) implemented Seven ANN techniques with different inputs and compared the results to ten empirical and semi-empirical ET_0 equations calibrated to FAO P-M equation using meteorological data as inputs. The comparisons criteria are the statistical error techniques; using PM56 daily ET_0 values as a reference. ANN techniques have obtained better results than the calibrated equations. El-Baroudy et al. (2010) compared the Evolutionary Polynomial Regression (EPR) to ANN and Genetic Programming (GP). The EPR model provided a performance comparable to that of GP and ANN models. Tabari and Talaee (2012) indicated that the main obstacle in application of FAO P-M Equation is the wide range of meteorological data essential as an input for calculation of ET_0, and because of the nonlinearity of ET phenomenon. They used a multi-layer neural network (MLNN) with variable inputs of data sets for modeling of ET_0 in the semiarid region in Hamedan, the model with all required climate parameters used as inputs performed best among other MLP models. Khoshhal and Mokarram (2012) evaluated different structures of MLP for estimation of ET_0, using meteorological data of Eghlid station in Iran for the period 2000-2010 as inputs and P-M ET_0, obtained using the same meteorological data, as output. The performance of 10 ANN models with different inputs were evaluated, the functions used for evaluation is root mean square error (RMSE), mean absolute error (MAE), and determination coefficient (R^2). The model with T_{min}, T_{max}, R_h, R, U_2 is more accurate in predicting ET_0 than other models.

This study investigates the performance of ANN and a hybrid of Artificial Neural Network-Genetic Algorithm (ANN-GA) techniques in predicting ET_0 in comparison with values obtained using P-M Equation with historical data from Mosul meteorological stations in Iraq. Evaluation of the proposed Artificial Intelligence techniques will be carried out against estimations done using the empirical method.

2. Study Area

The study area, named Mosul, is located at the center of Nineveh Governorate, northern Iraq, between latitude 36° 22' 00" N, longitude 43° 07' 00" E, and altitude 222.6 m above sea level, with an overall area of 37,323 km^2 (Figure 1) (Statistical report, 2009). The weather data used in this study are obtained from the main meteorological station in Mosul (Global Station Code 608), which includes the daily averages of: maximum air temperature (T_{max}), minimum air temperature (T_{min}), relative humidity (R_h), radiation hours (R), and wind speed (U_2) from 1980–2005.

Figure 1. Study area in Iraq

3. Techniques in Calculating Reference Evapotranspiration (ET_0)

3.1 Penman-Monteith Empirical Equation (P-M Equation)

The precise determination of ET_0 is a fundamental requirement in planning and scheduling any irrigation project. The empirical method in estimating ET_0 is a practical application if used with the proper crop coefficient (Kc) to find the crop Evapotranspiration (Etc) (i.e., Etc = $ET_0 \times$ Kc) (Yoder et al., 2005).

ET_0 is measured in millimeters per unit time where the time could be extended from seconds to minutes, hours, or even the overall growing season. Allen et al. in 1998 issued FAO paper 56 with the latest form of the FAO P-M Equation adopted by the organization as the only present method for calculation of ET_0. The climatic parameters are the only factors affecting ET_0, thus, ET_0 could be computed using records of weather data for a specific location, at any time intervals such as daily, ten days, or monthly. The equation formula shows all weather data that control energy exchange in the evaporation and transpiration process.

$$ETo = \frac{0.408\Delta(R_n - G) + \gamma \frac{900}{T+273} u_2 (e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)} \tag{1}$$

ET_0 = Reference Evapotranspiration (mm/day), Δ = slope of vapor pressure curve (kPa °C^{-1}), Rn = net radiation (MJ m^{-2}day^{-1}), G = density of soil heat flux (MJ m^{-2}day^{-1}), ϒ = psychometric constant (kPa °C^{-1}), T = average daily air temperature (°C), U_2 = wind speed at height 2 m (m s^{-1}), (e_s-e_a) = deficit in saturation vapor pressure.

The equation does not take into consideration the crop characteristics; Kc combines all the physical and physiological differences between crops for calculation of ETc.

The altitude and latitude data of study location should be specified. These data are essential for computation of radiation, daylight hours (N), or to adjust some weather parameters for the local average value of atmospheric pressure

Standard weather data, including (T_{max}), (T_{min}), (R_h), (R), and wind speed (U_2), over a period of 26 years were collected and organized in daily averages of monthly intervals form, then utilized in P-M equation for estimation of ET_0.

3.2 Feedforward Backpropagation Neural Networks

ANN architecture is a simulation of information processing that occurs in the biological brain. It starts with receiving, learning, adapting, recognizing the pattern, and then performing a desired function (Target) by trial of different weights of the information elements in a computation model. A typical ANN model contains an input layer that receives the input data. The hidden layers, with number of nodes that would satisfy the problem requirements, would recognize patterns and organize these data through multiple trial processes to predict the output (El-Baroudy et al., 2010). In this study, the Supervised Feedforward Backpropagation Neural Network (FFBPNN) is used for prediction of ET_0. In the FFBPNN, neurons are connected forward where each layer of the neural network connects to the next layer. In our work, the network model has an input layer, one hidden layer, and an output layer, using one hidden layer to represent the nonlinear relationship of ET_0 is sufficient (Kumar et al., 2002). The number of actual input variables is seven (7), representing weather parameters, altitude, and latitude values. The weather parameters used as inputs are the historical records obtained from the main meteorological station in Mosul for the period 1980–2005. The total available data is divided into two main categories; 80% of the data, records of 240 monthly averages of each of the weather parameters, is for training the model; the other 20% of data records, representing 72 monthly averages of each of the weather parameters, is used in testing the performance of the model. In the training period, the input data are designed to be classified again into three subclasses, training; testing; and valuating. The backpropagation algorithm is then employed to activate the ten (10) neurons in the hidden layer, determined using trial-and-error methods with application of 5, 7, 9, 10 and 15 neurons), for weighing and training of information elements to achieve the desired target as shown in Figure 2 (Badde et al., 2013). Backpropagation is a form of supervised training in ANN, whereby the network has to be provided with both sample inputs and anticipated outputs during training. In this work, all input and output variables were normalized before training. The function Min and Max (premnmx, postmnmx, tramnmx) were used to standardize the data within a range of [-1, 1] (Malek, 2008). The model was implemented using MATLAB.

$$XN = (X - MinX) / (MaxX - MinX) * 2 - 1 \tag{2}$$

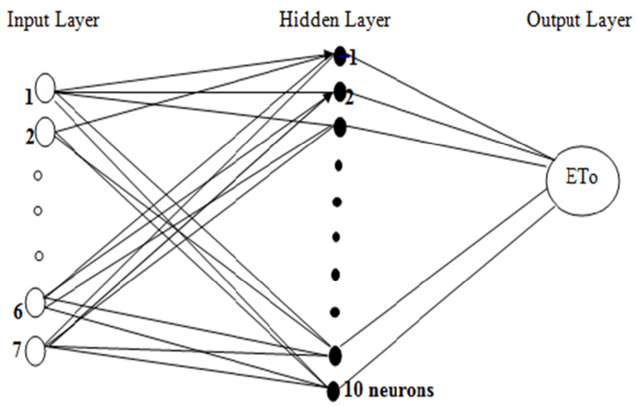

Figure 2. Feedforward backpropagation neural network

After training and testing performance of the simulation model, the FFBPNN model is implemented also to prepare for making prediction, and then to predict ET_0 for any time interval depending on new input data that have not been used in the training or testing periods.

The same ANN model was tested also for inputs to include weather data only, excluding altitude and latitude data; which are constant parameters for the identified study location.

3.3 Genetic Algorithm

Genetic Algorithm (GA) is a method of probabilistic optimization based on evolution theory. Evolution from one generation to another is composed of three main stages: Selection of strings according to their fitness, crossover of strings, and random mutation for selected strings to compute new generation (Azzini, 2007). The records of weather data used in FFBPNN model are used in this study, in the same pattern, in a hybrid model of Binary GA with ANN (GA-ANN). The GA technique is used in this model to improve the performance of ANN by changing the number of hidden layers to three (3) instead of only one (1) layer used in the ANN prediction method. The standard performance function is the Mean Square Errors (MSE), which is used to minimize the connection weights depending on cost of elements and accuracy.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(ETi_{calculated} - ETi_{predicted}\right)^2 \tag{3}$$

The second function is the Correlation Coefficient, which proportionally clarifies the statistical cost of each element. The range of Correlation Coefficient used is from -1 to 1 (Adeloye et al., 2012).

$$R = \frac{N\sum xx' - \sum x \sum x'}{\sqrt{[N\sum x^2 - (\sum x)^2][N\sum x'^2 - (\sum x')^2]}} \tag{4}$$

The mean Absolute Errors is also taken into consideration in the model performance evaluation.

The binary FFBP-GA model is capable; even better then FFBPNN, of preparing itself for prediction depending on the stored efficiencies of the training and testing periods, and making actual prediction for any location or time interval.

4. Results and Discussion

The nonlinear relationship between weather parameters and the effect of the study area location according to latitude and altitude, in addition to other environmental factors, inspires this study to produce precise and dependable ET_0 values with time, effort, and cost consuming. The average values of weather parameters for the period 1980–2005 were employed in the FAO P-M Equation for the calculation of average monthly ET_0 as shown in Figure 3. It shows the inequality of the ET rates at different periods along the year. The design of the water-used system should accommodate peak values, which are in June, July, and August, to satisfy the minimum supply of water according to its monthly consumptive use.

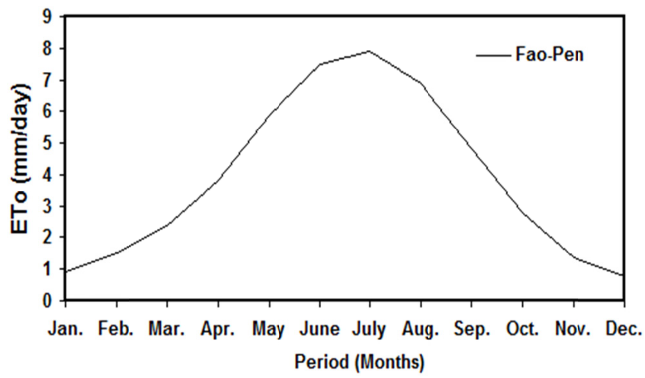

Figure 3. Average monthly P-M ET_0 at Mosul Station for the period 1980–2005

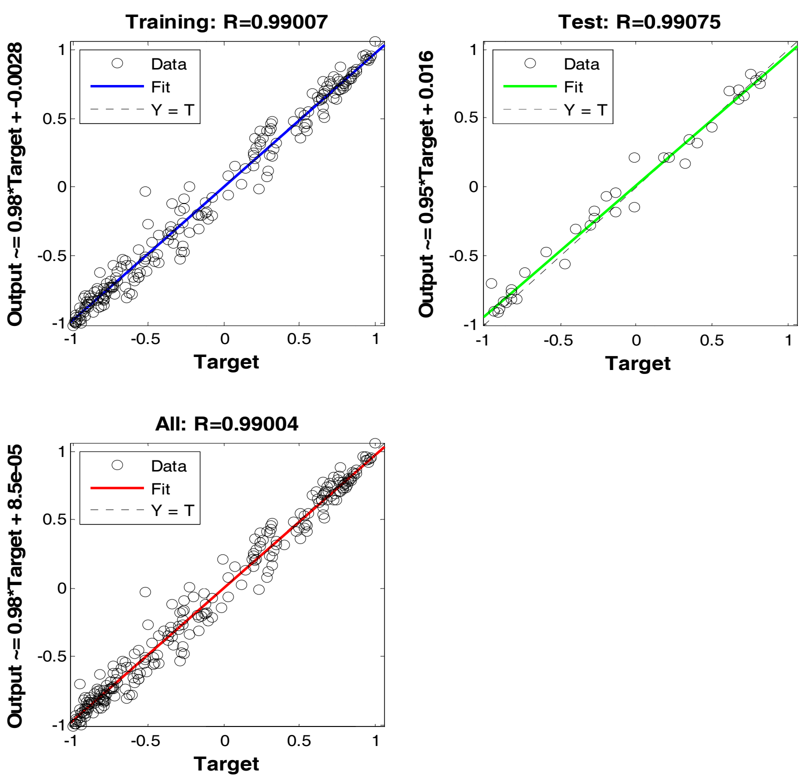

Figure 4. Relationship between targeted FAO ET_0 and ANN predicted ET_0 values for training period at Mosul station

The results of ET obtained from application of FAO P-M Equation were employed in the FFBPNN model to test the capability of the network system in predicting ET_0. The best results, obtained at 233 0f designed 1000 epochs, are shown in Figures 4 and 5 were encouraging. Eighty percent of total observed weather variables, together with their related output data, were employed in the system training, whereas the other 20% were considered in the testing period. In the training period, the data will again be classified into training, testing, and evaluation, as shown in Figure 4.

The ANN model performance at both training and testing periods was evaluated by functions of MSE, MAE and R^2 as presented in Table 1, which indicate the efficiency of the proposed ANN model in predicting ET_0.

Table 1. The performance functions for training and testing periods in ANN Model for years 1980–2005

Data Set	MSE (mm2/day^2)	MAE (mm/day)	R^2
(1) Training 80%	0.0084	0.0679	0.9900
(2) Testing 20%	0.0139	0.0971	0.9900

The correlation between calculated ET_0 using FAO P-M Equation and ANN predicted ET_0 is illustrated in Figure 5, which reveals a significant consistency between the values.

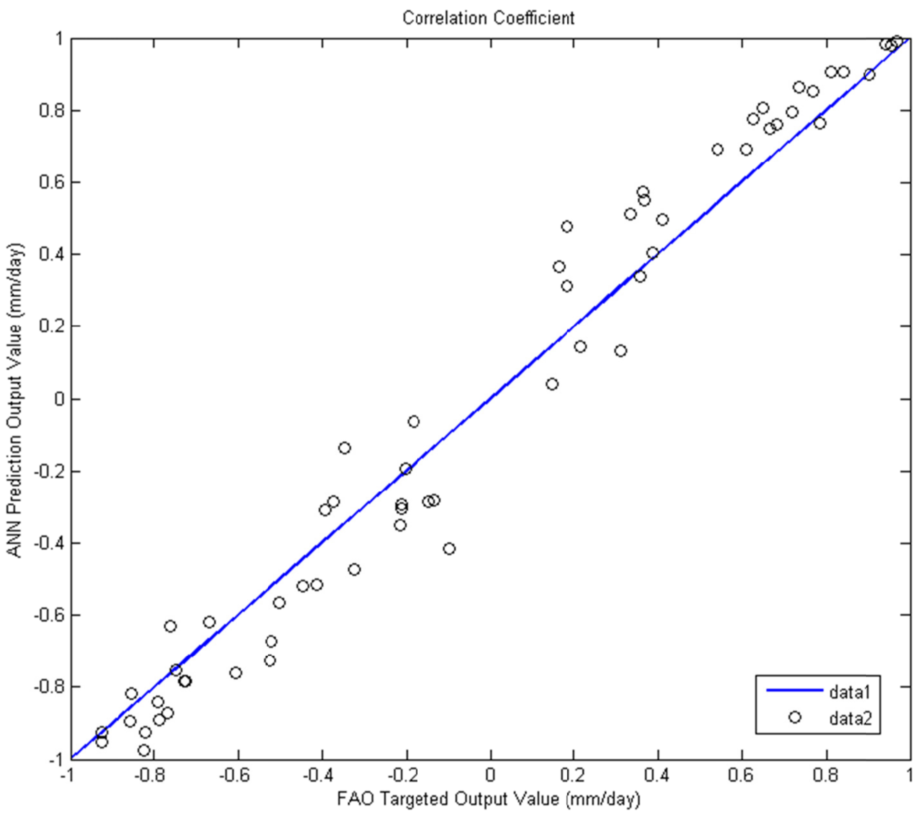

Figure 5. The correlation coefficient between calculated ET_0 using FAO P-M Equation and ANN Predicted ET_0 at Mosul Station

The ANN model was then implemented to investigate the efficiency of predicting P-M ET_0; when the input data are only the records of weather parameters employed previously, and to exclude the data of common latitude and altitude parameters.

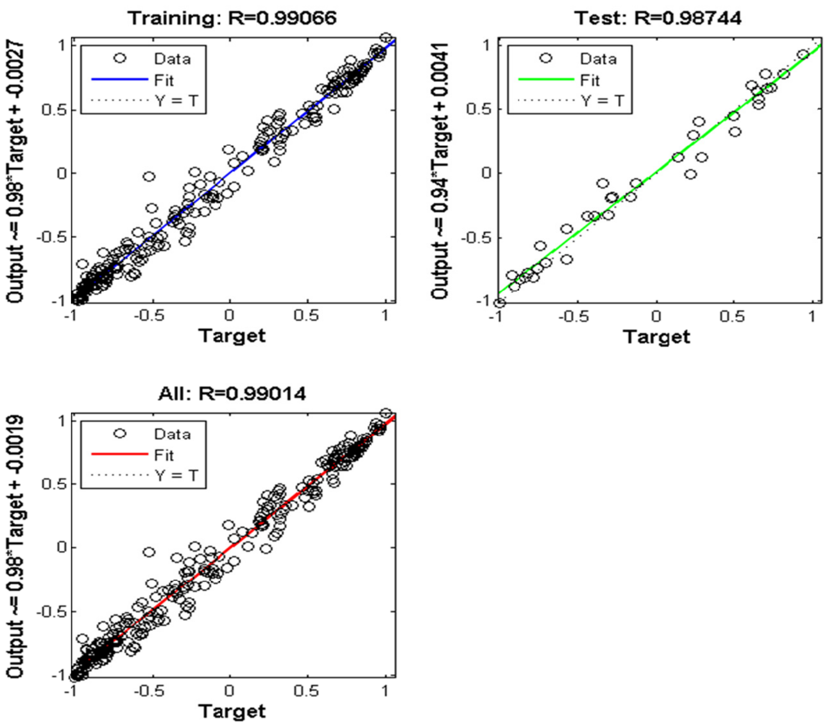

Figure 6. Relationship between targeted FAO ET_0 and ANN predicted ET_0 values without location data for training period at Mosul station

The ANN model performance at both training and testing periods was evaluated also by functions of MSE, MAE and R^2 as presented in Table 2, which indicate the efficiency of the proposed ANN model in predicting ET_0 without location data.

Table 2. The performance functions for training and testing periods in ANN Model without altitude and latitude data for years 1980–2005

Data Set	MSE (mm2/day^2)	MAE (mm/day)	R^2
(1) Training 80%	0.0083	0.0674	0.9900
(2) Testing 20%	0.0139	0.0971	0.9897

The correlation between calculated ET_0 using FAO P-M Equation and ANN predicted ET_0, with weather data only, is illustrated in Figure 7.

The results of implementing ANN model with only 5 inputs (weather data) and P-M ET_0 as output; show slight effect on both training and testing evaluation functions.

The same values of ET results obtained by application of FAO P-M Equation were also used against a hybrid model of FFBPNN with a Binary GA (ANN-GA). In this hybrid ANN-GA model, the maximum number of iteration, after which no significant changes were noticed, is found to be one hundred (100); population size is found to be twenty (20), and mutation rate is 0.15. The optimization function used is the cost function; by minimizing the MSE, The best cost function is 0.0067828 at generation number of 100. The hybrid of ANN-GA enhanced the model performance and minimized the fitness function represented by the Cost, which is the Mean Squared Error.

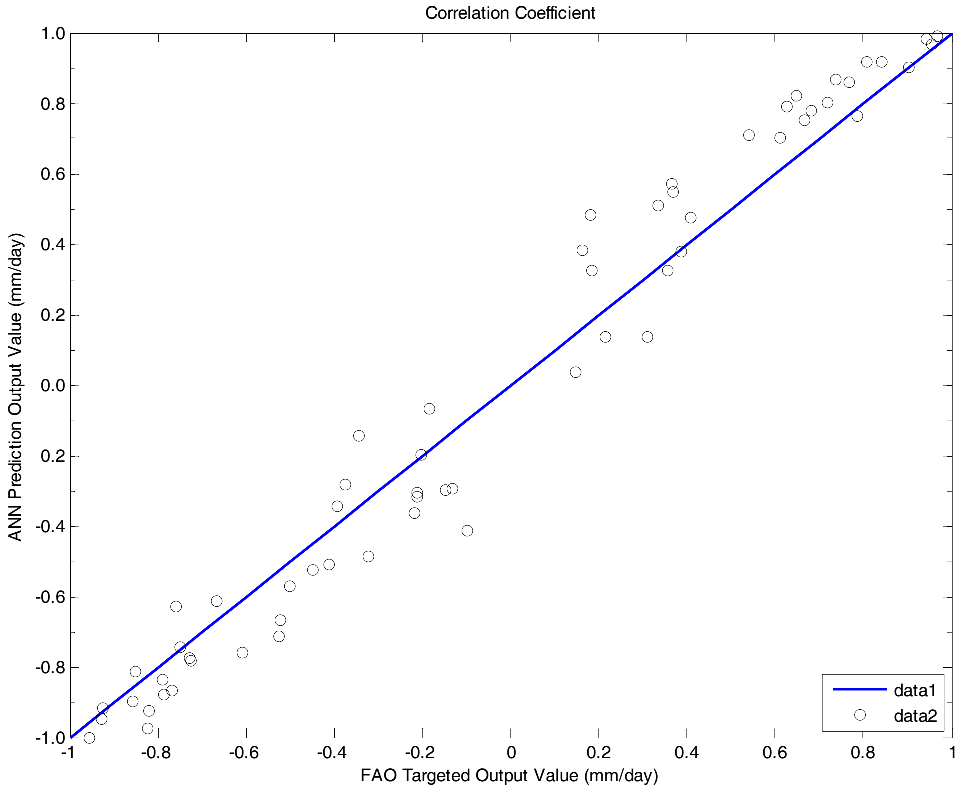

Figure 7. The correlation coefficient between calculated ET_0 using FAO P-M Equation and ANN Predicted ET_0 with weather data only at Mosul Station

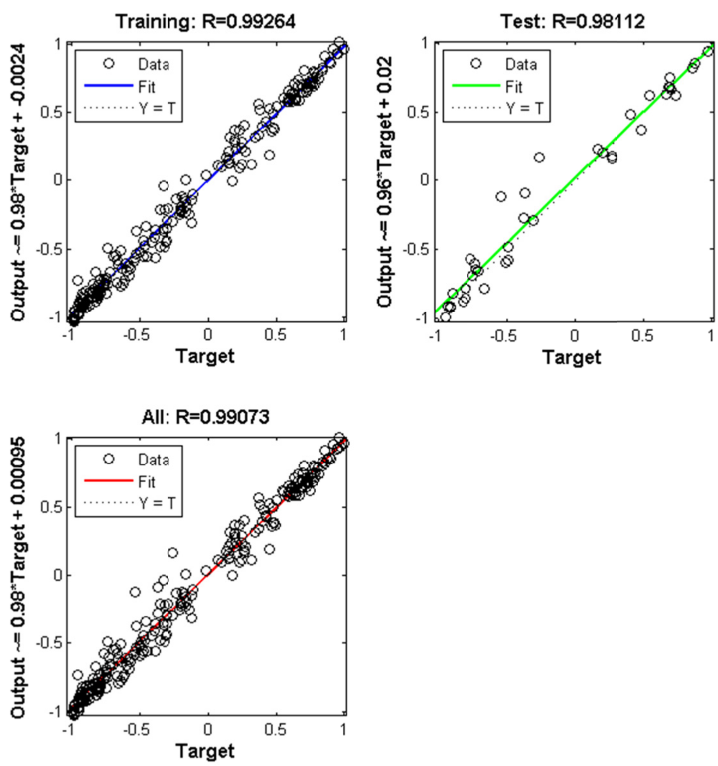

Figure 8. The relationship between targeted FAO ET_0 and Hybrid ANN-GA predicted ET_0 values for the training period at Mosul station

Figure 8 shows that the Hybrid ANN-GA simulation has modify the result of the Correlation Coefficient (R^2). It exhibited a higher consistency rather than using ANN by itself.

The main obstacle in the application of P-M Equation is the comprehensive and rich weather data needed for specific location or project in the design and operation stages. Latitude and altitude are to be specified as well; however, as the location specifications have only slight effect on implementation of intelligent models; these models would allow using historical records of weather data available in any alternative location for training the system and using the limited records available of the exact location for testing the system and even predicting actual values of ET_0 for any period.

5. Conclusion

ET_0 is the key parameter in designing any irrigation project. Building an intelligent model of ANN and Hybrid ANN-GA would provide an effective alternative to the empirical method used in the estimation of ET. The proposed Artificial Intelligence models have significant superiority as compared to the limitations embedded in the traditional empirical FAO P-M Equation method. This study shows that both ANN and Hybrid ANN-GA models can be used for prediction of ET_0 in Iraq, taking into consideration the unstable and sudden changes of weather conditions.

The normalization of input and output parameters would give the researchers a chance to overcome the limitations of application of relevant data to a specific location. In P-M Equation, both latitude and altitude values are considered properties of the study location and have great effect on ET_0 values. Normalization of input and output parameters in the FFBPNN and FFBPNN-GA models will eliminate the effect of those two parameters as they are constants; shared in all time intervals; the AI model could be recommended to be used without restrictions of location specifications.

References

Adebayo, J., Adeloye, A. J., Rustum, R., & Kariyama, I. D. (2012). Neural computing modeling of the reference crop Evapotranspiration. *Journal of Environmental Modelling & Software, 29*, 61-73. http://dx.doi.org/10.1016/j.envsoft.2011.10.012

Allen, R. G., Pereira, L., Raes, S. D., & Smith, M. (1998). *Crop Evapotranspiration, guidelines for computing crop water requirements*. FAO Irrigation and Drainage Paper No. 56, Rome, Italy.

Azzini, A. (2007). *A New Genetic Approach for Neural Network Design and Optimization*. Ph.D. thesis, Universita Degli Studi Di Milano, Italy. Retrieved from http://www.dti.unimi.it/azzini/wwwmat/TesiAzziniAntonia.pdf

Badde, D. S., Gupta, A. K., & Patki, V. K. (2013). Cascade and Feedforward Backpropagation Artificial Neural Network Models for Prediction of Compressive Strength of Ready Mix Concrete. *IOSR Journal of Mechanical and Civil Engineering (IOSR-JMCE),* 01-06.

Doorenbos, J., & Pruitt, W. O. (1977). Guidelines for predicting crop water requirements. *FAO Irrigation and Drainage Paper 24 (revised), part I, 1*. Rome, Italy.

El-Baroudy, A. Elshorbagy, Carey, S. K., Giustolisi, O., & Savic, D. (2010). Comparison of Three Data—Driven Techniques in Modelling the Evapotranspiration Process. *Journal of Hydro Informatics, 12*(4), 365-379. http://dx.doi.org/10.2166/hydro.2010.029

Kerr, R. A. (1998). Sea-Floor Dust Shows Drought Felled Akkadian Empire. *Science Magazine, 279*(5349), 325-326. Retrieved from http://www.sciencemag.org/content/279/5349/325

Khoshhal, J., & Mokarram, M. (2012). Model for Prediction of Evapotranspiration Using MLP Neural Nettwork. *Internetional Journal of Environmental Sciences, 3*(3), 1000-1009.

Kisi, Ö. (2006). Generalized regression neural networks for Evapotranspiration modeling. *Journal of Hydrological Sciences, 51*(6), 1092-1105. http://dx.doi.org/10.1623/hysj.51.6.1092

Kisi, Ö. (2008). The Potential of Different ANN Techniques in Evapotranspiration Modelling. *Hydrology Process, 22*, 2449-2460. http://dx.doi.org/10.1002/hyp.6837

Kisi, Ö., & Öztürk, Ö. (2007). Adaptive Neurofuzzy Computing Technique for Evapotranspiration Estimation. *Journal of Irrigation and Drainage Engineering,* 368-379. http://dx.doi.org/10.1061/(ASCE)0733-9437(2007)133:4(368)

Kumar, M., Raghuwanshi, N. S., Singh, R., Wallender, W. W., & Pruitt, W. O. (2002). Estimating Evapotranspiration Using Artificial Neural Network. *Journal of Irrigation and Drainage Engineering, ASCE, 128*(4), 224-233. http://dx.doi.org/10.1061/(ASCE)0733-9437(2002)128:4(224)

Landeras, G., Ortiz-Barredo, A., & Lo´pez, J. J. (2008). Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain). *Agricultural water management, 95*, 553-565. http://dx.doi.org/10.1016/j.agwat.2007.12.011

Malek, M. A. (2008). *DeveloP-Ment of Rainfall Data in-Filling Model with Expectation Maximization and Artificial Neural Network*. Ph.D. thesis, Universiti Teknologi Malaysia, Malaysia.

The Ministry of Planning Central Bureau of Statistics Directorate of Environment Statistics. (2009). Environmental statistics report for Iraq for the year 2009. Retrieved from http://www.cosit.gov.iq/en/climate

Tabari, H., & Talaee, P. H. (2012). *Multilayer Perceptron for Reference Evapotranspiration Estimation in Semiarid Region*. Neural Compute & Appic. Springer-Verlag London Limited.

Trajkovic, S. (2005). Temperature-Based Approaches for Estimating Reference Evapotranspiration. *Journal of Irrigation and Drainage Engineering, ASCE, 131*(4), 316-323. http://dx.doi.org/10.1061/(ASCE)0733-9437(2005)131:4(316)

Wright, J. L. (1988). Daily and Seasonal Evapotranspiration and yield of irrigated alfalfa in southern Idaho. *Agronomy Journal, 80*, 662-669. http://dx.doi.org/10.2134/agronj1988.00021962008000040022x

Yoder, R. E., Odhiambo, L. O., & Wright, W. C. (2005). Evaluation of methods for estimating daily reference crop Evapotranspiration at a site in the humid southeast United States. *American Society of Agricultural Engineering, Soil and Water Division, 21*(2), 197-220.

Copyrights