Artificial Neural Network Modeling for Monthly Evaporation

Aparajita Singh
Department of Farm Engineering
Institute of Agricultural Science, Banaras Hindu University, Varanasi, UP, 221005

A. R. Senthil Kumar
National Institute of Hydrology, Roorkee, Uttarakhand

R. M. Singh
Institute of Agricultural Science, Banaras Hindu University, Varanasi, UP, 221005

Abstract

Most of the available researches established a linear relationship between the parameters that affects the evaporation and the target evaporation. Now it is proved that the parameters which affect the evaporation are nonlinear in nature. This leads the whole researches towards the nonlinear estimation of evaporation. The evaporation affecting parameters are nonlinear and follow a very irregular trend so an Artificial Neural Network (ANN) has evolved a better technique to establish the structural relationship between the various entities. This paper examines the applicability of ANN approach to model the monthly evaporation. The input combinations to model monthly evaporation were selected on the basis of data's statistical properties which were obtained from NIH observatory. The highest correlation coefficient valves during calibration and validation were (0.831, 0.819) with lowest RMSE valves (0.376, 0.261) respectively for best evaporation model [4-5-1]. This was obtained with all 4 input parameters namely, monthly rainfall, monthly maximum temperature, monthly minimum temperature and monthly relative humidity at the same time. The comparison was made between the observed and computed values of evaporation which emphasizes the usefulness of ANN technique for monthly evaporation estimation..

Keywords- ANN, Evaporation, Calibration, Validation

I. INTRODUCTION

As a major component of hydrological cycle, the estimation of evaporation is essential for managing scarce water resources effectively. This fulfills the need of sustainable crop production in the recent time. The various field including hydrology, land resources planning, forestry and many more needs the knowledge of evaporation losses for water balance computation, irrigation management, crop yield forecasting etc (Xu and Singh, 2005). The effective managements of water resources need the basic knowledge of accurate evaporation losses. Hence this arise the need for reliable model for evaporation estimation to deal with wide ranges of problems occurred during water resources management, drainage designing, irrigation management etc. Many conventional models like empirical, regression based models, conceptual models etc. were developed to estimate the evaporation which requires large amount input data with insignificant results as compared to soft computing techniques like ANN, Fuzzy logic etc. (ASCE,2000).

An ANN is a flexible mathematical architecture which is capable to identify the complex nonlinear relationships between input and output data sets (Kumar et. al., 2002). An ANN can model any complex nonlinear problems effectively as compared to conventional model (Zang et.al, 1998). An ANN model is useful because it requires less computational efforts, fewer input data and even less computational time than any conventional mathematical model (McCulloch and Pitt, 1943). A neural network model is even sensitive to tiny input changes in a dynamic environment (sudheer et. al., 2002). The objective of present study was to model the monthly evaporation using Artificial Neural Networks (ANN) with climatic parameters as input variables and monthly evaporation as output obtained from national institute of hydrology (NIH) observatory, Roorkee.

II. MATERIALS AND METHODS

A. Study Area

Roorkee is located at latitude $29^0$51' N and longitude 77^0 53' E in Haridwar district on the south bank of solani River. The upper ganga canal is the most important features which enhances the beauty of the city. As it runs from north to south and divides the city in two distinct parts. The city is located at an altitude of 265 meters above mean sea level and receives the average annual rainfall of 1170 mm with an average monsoon rainfall of 878 mm. It has an average maximum temperature is about 40 ^{0}C with an average minimum temperature of 2^0C. The maximum humidity of city is 100 % with an average minimum humidity about 30 %.

B. Artificial Neural Network

An ANN is a newer soft computing technique, consists of various processing elements interconnected to each other called nodes, similar to neurons of human brain (Fausett, 1994). The selection of significant inputs to the ANN model was the initial step, based correlation analysis between independent and dependent variables (Maier and Dandy, 2000) and as per equation 1.

$$r_k = \frac{\sum_{t=1}^{N-k}(x_t - \overline{x_t})(x_{t+k} - \overline{x_{t+k}})}{\left[\sum_{t=1}^{N-k}\left(x_t - \overline{x_t}\right)^2 \sum_{t=1}^{N-k}\left(x_{t+k} - \overline{x_{t+k}}\right)^2\right]^{1/2}} \quad (1)$$

Where r_k - lag-k correlation coefficient, x_t - time series for t = 1,..,N, x_{t+K} - lagged time series for t = 1,…, N-k, $\overline{x_t}$ - sample mean for t = 1,…,N, $\overline{x_{t+k}}$ - sample mean for t = 1,…,N-k, N - sample size.

The second step was to normalize the input data sets between 0 and 1 because an ANN model was consist sigmoid logistic activation function. The final output of model should be denormalised to provide meaningful results (Dawson and Wilby, 1998). The normalization of input data sets can be done as per equation 2.

$$N_i = \frac{R_i - Min_i}{Max_i - Min_i} \quad (2)$$

Where, R_i - real value applied to neuron i, N_i - subsequent normalized value calculated for neuron i, Min_i - minimum value of all values applied to neuron i, Max_i - maximum value of all values applied to neuron i.

An ANN model consists of feed forward neural network (FFNN) structure, trained with BR algorithm was used in present study to model the monthly evaporation. It consists of input layer, one or more hidden layers and output layer. The data flows from one direction i.e. input layer to output layer through different hidden neurons. The nodes in input layer received inputs from external environment and do not perform any transformations. It sends the weighted values to the immediately adjacent nodes, usually hidden layer. The nodes in hidden layer received the transferred weighted inputs and perform the transformations on it. Hidden layer passes the output to next adjacent layer, may be another hidden layer or the output layer. The node in output layer receives the output result from hidden layer and sends it to the user. The receiving node sums the weighted signals from all nodes to which it was connected in the preceding layer. The net _input x_i to node i was the weighted sum of all the incoming signals and represented as per equation 3.

$$net\_input = \sum x_i w_i \quad (3)$$

Where, x_i - net input coming to node i, w_i - weight stored in node i

The activation function y_i, was a nonlinear function of net_ input and described with sigmoid logistic function as per equation 4

$$y_i = \frac{1}{1 + \exp(-\text{net\_input})} \quad (4)$$

The bayesian regularization (BR) algorithm automatically sets optimum values for the objective function parameters (Anctil *et al.*, 2004). The number of nodes in the input layer was based on the inputs to the model. The hidden neurons were responsible to capture dynamic and complex relationship between input and output variables and were identified by various trials. MATLAB 2010a software was used for the analysis.

C. Performance Evaluation of the Developed ANN Model

The performance evaluation of the developed ANN model was estimated on the basis of visual observations and quantitative evaluation to judge the goodness of fit between observed and estimated values. The visual observation consists of graphical comparison between the observed and the estimated value, simplest method of performance assessment of a model.

Coefficient of efficiency (CE) used to compare the relative performance of the two approaches effectively and expressed as per equation 5

$$CE = \left\{1 - \frac{\sum_{j=1}^{n}(Y_j - X_j)^2}{\sum_{j=1}^{n}(Y_j - \bar{Y}_j)^2}\right\} \quad (5)$$

The coefficient of efficiency into three categories viz. perfectly acceptable simulation (C.E. > 0.90), acceptable simulation (C.E. between 0.60 and 0.90) and unacceptable simulation (C.E. < 0.60) (Chiew et al. 1993)

Root mean square error (RMSE) indicates the discrepancy between the observed and estimated values. The lowest RMSE valve indicates more accurate result and expressed as per equation 6

$$RMSE = \sqrt{\sum_{j=1}^{n}\frac{(Y_j - X_j)^2}{n}} \quad (6)$$

Explained Variance (EV) Explained variation measures the proportion to which a mathematical model accounts for the variation (dispersion) of a given data set and given by equation 7

$$EV = \sqrt{\left(\frac{\sum(X_j - \bar{Y}_j)^2}{\sum(Y_j - \bar{X}_j)^2}\right)} \quad (7)$$

Where, Yj - Observed Evaporation, Xj - Estimated Evaporation, n -Number of observations, Yj -Mean of observed Evaporation, j - Mean of estimated Evaporation.

In the present analysis, ANN model with FFNN structure was developed to model the monthly evaporation with antecedent monthly rainfall; monthly relative humidity, monthly maximum temperature and monthly minimum temperature as input vector obtained at NIH observatory from January 2009 to October 2013.

III. RESULTS AND DISCUSSION

The feed forward ANN model architecture with BR algorithms was considered to estimate the monthly evaporation at Roorkee, with the monthly rainfall, monthly relative humidity, monthly maximum temperature and monthly minimum temperature data as the input to the model. The whole data set divided into two sets, one for calibration and other for validation of ANN model. The data from January 2009 to October 2013 (58 monthly datasets) was considered for training the model. Out of 58 dataset, 35 monthly datasets (60.34%) were used for calibration (training), 23 datasets (39.65%) were used for validation. The number of neurons in the hidden layer was found by a trial and error method, started initially with one hidden neuron and goes up to 10 based on performance criteria of model.

The performance of model EVAP5 (4-5-1) as per table 1with fifth number of hidden neurons was best in respect to other structure with different hidden neuron derived from statistical procedure (Sudheer *et al.*, 2002). The ANN structure 4-10-1 also gives better result but after crossing more number of neurons, the model performance was fluctuated (decreasing and then it is increasing) which leads to over fitting of an ANN structure. The performance of best ANN models for monthly evaporation estimation at NIH, Roorkee during calibration and validation were presented in Figure 1 and Figure 2 respectively. The comparison between observed and estimated evaporation clearly demonstrate the potentiality of an ANN model. The results in terms of various statistical indices during calibration and validation of the best ANN models was presented in the Table 1

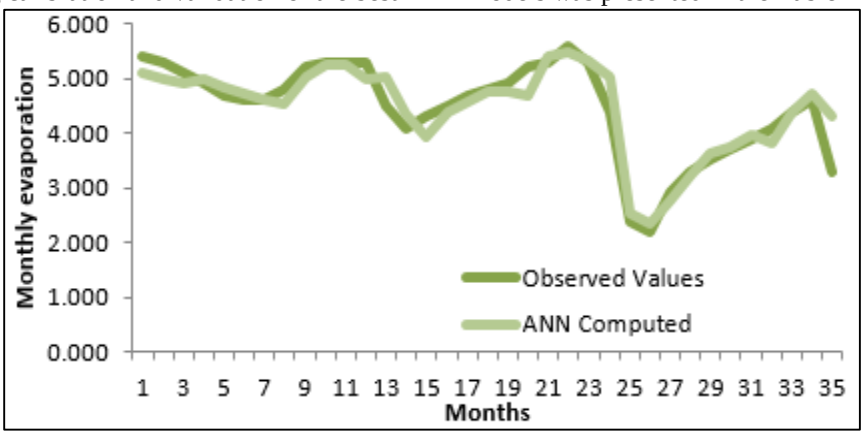

Fig. 1: Observed and computed monthly evaporation during calibration

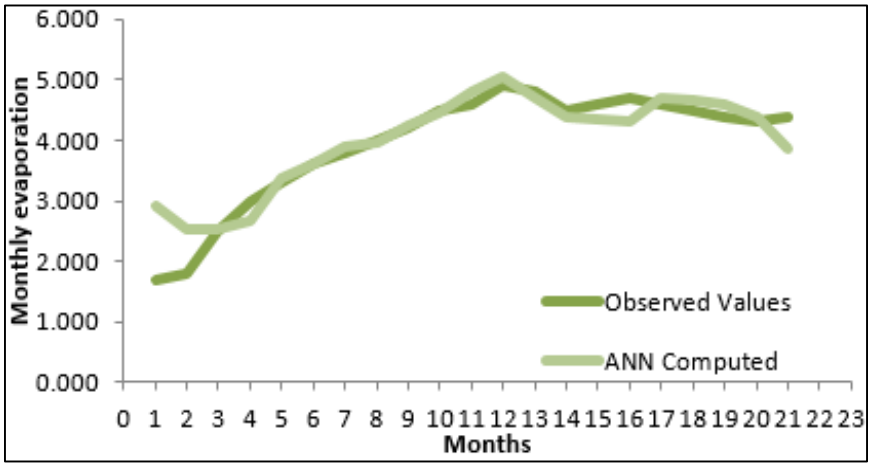

Fig. 2: Observed and computed monthly evaporation during validation

Table 1: Statistical indices during calibration and validation

Model No	*ANN Structure*	*Calibration*			*Validation*		
		CORR	*RMSE*	*EFF%*	*CORR*	*RMSE*	*EFF%*
ANN_EVAP 1	*4-1-1*	*0.765*	*0.284*	*0.664*	*0.804*	*0.301*	*0.704*
ANN_EVAP 2	*4-2-1*	*0.774*	*0.292*	*0.672*	*0.824*	*0.311*	*0.712*

ANN_EVAP 3	*4-3-1*	*0.780*	*0.295*	*0.681*	*0.831*	*0.312*	*0.721*
ANN_EVAP 4	*4-4-1*	*0.785*	*0.295*	*0.683*	*0.804*	*0.321*	*0.734*
ANN_EVAP 5	*4-5-1*	*0.831*	*0.376*	*0.772*	*0.819*	*0.261*	*0.743*
ANN_EVAP 6	*4-6-1*	*0.798*	*0.297*	*0.680*	*0.805*	*0.315*	*0.711*
ANN_EVAP 7	*4-7-1*	*0.794*	*0.292*	*0.685*	*0.801*	*0.306*	*0.721*
ANN_EVAP 8	*4-8-1*	*0.792*	*0.293*	*0.686*	*0.807*	*0.317*	*0.713*
ANN_EVAP 9	*4-9-1*	*0.799*	*0.298*	*0.682*	*0.803*	*0.318*	*0.715*
ANN_EVAP10	*4-10-1*	*0.811*	*0.380*	*0.783*	*0.811*	*0.219*	*0.723*

IV. CONCLUSION

The present study was done to modeled monthly evaporation with an ANN model. The data of monthly rainfall, monthly maximum temperature, monthly minimum temperature and relative humidity obtained at NIH observatory, Roorkee from January 2009 to October 2013 were used for the analysis. The correlation between input vectors and monthly evaporation was analyzed and potential inputs were fed to ANN model. The FFNN structure trained with BR algorithm with 4 input nodes, 1 output node and different hidden nodes. The number of neurons in the hidden layer was optimized to 10 based on the trial and error method. Out of 58 monthly data sets, 35 sets (60.34%) of data for calibration, 23 data sets (39.65%) of data used for validation. The statistical indices such as coefficient of correlation, root mean squared error (RMSE) and model efficiency were used to evaluate the performance of the model. The RMSE of ANN model during calibration and validation were found to be 0.376 and 0.261 respectively, and also the ANN model efficiency during calibration and validation were 0.772 and 0.743 respectively, indicates the substantial improvement in the model performance. In addition, comparison made between observed and estimated evaporation showed that evaporation values estimated by the ANN model was more precise.

REFERENCES

[1] ASCE Task Committee on Application of Artificial Neural Networks in Hydrology, 2000. Artificial neural networks in hydrology. II: Hydrologic applications. Journal of Hydrologic Engineering, 5(2), pp.124-137.
[2] Anctil, F. and Tape, D.G., 2004. An exploration of artificial neural network rainfall-runoff forecasting combined with wavelet decomposition. Journal of Environmental Engineering and Science, 3(S1), pp.S121-S128.
[3] Chiew, F.H.S., Stewardson, M.J. and McMohan, T.A. 1993. Comparison of six rainfall-runoff modelling approaches. , 147: 1-36.
[4] Dawson, C.W. and Wilby, R., 1998. An artificial neural network approach to rainfall-runoff modelling. Hydrological Sciences Journal, 43(1), pp.47-66.
[5] Fausett, L., 1994. Fundamentals of neural networks: architectures, algorithms, and applications. Prentice-Hall, Inc.
[6] Kumar, M., Raghuwanshi, N.S., Singh, R., Wallender, W.W. and Pruitt, W.O., 2002. Estimating evapotranspiration using artificial neural network. Journal of Irrigation and Drainage Engineering, 128(4), pp.224-233.
[7] McCulloch, W.S. and Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics, 5(4), pp.115-133.
[8] Maier, H.R. and Dandy, G.C., 2000. Neural networks for the prediction and forecasting of water resources variables: a review of modelling issues and applications. Environmental modelling & software, 15(1), pp.101-124.
[9] Sudheer, K.P., Gosain, A.K., Mohana Rangan, D. and Saheb, S.M., 2002. Modelling evaporation using an artificial neural network algorithm. Hydrological Processes, 16(16), pp.3189-3202.
[10] Xu CY, Singh VP (2005) Evaluation of three complementary relationship evapotranspiration models by water balance approach to estimate actual regional evapotranspiration in different climatic regions Journal of Hydrology 308(1-4) 105-121
[11] Zhang, Guoqiang, B. Eddy Patuwo, and Michael Y. Hu. "Forecasting with artificial neural networks: The state of the art." International journal of forecasting 14.1 (1998): 35-62.