

## **ESTIMATION OF DAILY REFERENCE EVAPOTRANSPIRATION USING ARTIFICIAL NEURAL NETWORK FOR AMBIKAPUR STATION IN CHHATTISGARH**

**Rekha Bai and \*Jitendra Sinha**

*SV College of Agricultural Engineering & Technology & Research Station,  
IGKV, Raipur-492006*

*\*Author for Correspondence*

### **ABSTRACT**

Determination of reference evapotranspiration ( $ET_0$ ) is a key factor for water balances and irrigation scheduling. The study present the development of artificial neural networks (ANNs) model for estimating reference crop evapotranspiration with climatic data required for Penman-Monteith (P-M) method, to test artificial neural networks (ANNs) for estimating reference evapotranspiration ( $ET_0$ ) with limited climatic data ( $ET_0$ ) and compares the performance of ANNs with P-M method. The ANNs are trained to estimate  $ET_0$  from daily climate data as input and the Penman-Monteith (P-M) estimate as output. The networks were trained with daily two climatic data such as maximum temperature and sunshine hour of past 18 years (1 January 1996 to 31 December 2013) have been considered as input and the Penman-monteith (PM) estimated  $ET_0$  as output. Out of total number of 6575 patterns, 4600 has been used for training the network while remaining 1975 has been used for testing the model. The analysis was carried out in MATLAB software. A feed forward multiple layer networks with sigmoid function are used. Performance evaluation of the models have been carried out by calculating mean absolute deviation (MAD), root mean square error (RMSE), coefficient of correlation (CC), Nash - Sutcliffe coefficient efficiency (CE) and Index of Agreement (IOA). The network are selected based on maximized CC, CE and IOA value and minimized MAD and RMSE values both in training and testing. The optimal ANN (2-6-1) for Ambikapur regions showed a satisfactory performance in the  $ET_0$  estimation as compared to Penman Montieith method. These ANN models may therefore be adopted for estimating  $ET_0$  in the study area with reasonable degree of accuracy.

**Keywords:** *Reference Evapotranspiration; Penman-Monteith Method; Artificial Neural Network; Back-propagation Algorithm; Training; Testing*

### **INTRODUCTION**

Evapotranspiration (ET) is one of the major components of the hydrologic cycle and its accurate estimation is of paramount importance for many studies, such as hydrologic water balance, irrigation system design and management, crop yield simulation, and water resources planning and management. Evapotranspiration (ET) is a term used to describe the sum of evaporation and plant transpiration from the Earth's land surface to atmosphere. Evaporation accounts for the movement of water to the air from sources such as the soil, canopy interception, and water bodies. Transpiration accounts for the movement of water within a plant and the subsequent loss of water as vapor through stomata in its leaves. Evapotranspiration is an important part of the water cycle.

$ET_0$  was introduced to study the evaporative demand of the atmosphere independently of crop type, crop development and management practices. As water is abundantly available at the reference evapotranspiring surface, soil factors do not affect ET. The only factors affecting  $ET_0$  are climatic parameters. Consequently,  $ET_0$  is a climatic parameter and can be computed from weather data.  $ET_0$  expresses the evaporating power of the atmosphere at a specific location and time of the year. Reference evapotranspiration ( $ET_0$ ) estimation methods have been developed and used by researchers and practitioners according to the availability of historical and current weather data.

These methods include empirical equations or Methods based on physical processes of complex. One of the methods that are widely used to estimate evapotranspiration is the Penman-Monteith (Kumar *et al.*,

## **Research Article**

2002). According to research done, the Penman - Monteith is accurate methods of estimating evaporation and it can use for estimate transpiration and watering in different regimes. Empirical models are usually used to estimate ET at local and regional scales. No estimation technique is universal, but a standard method is the Penman-Monteith equation as specified by the UN Food and Agriculture Organization (FAO) in paper number 56. The FAO56 Penman-Monteith method estimates ET rates for a well-watered reference surface based on physical atmospheric observations of solar radiation, temperature, wind speed, and relative humidity.

This estimate is commonly referred to as reference ET. The reference surface is a theoretical grass reference crop with a height of 0.12m, an albedo of 0.23, and a constant surface resistance of 70 s/m. While dependent on time of year and location, the equation is developed for the hypothetical grass reference crop and is thus independent of specific crop characteristics and soil factors.

The Hargreaves and Samani (1985) solar radiation estimate is utilized as an input to the Penman-Monteith equation.

The Hargreaves and Samani (1985) technique is based on the assumption that diurnal temperature range varies indirectly with cloud cover and thus, is related to incoming solar radiation.

Evapotranspiration is a complex and nonlinear phenomenon because it depends on several interacting climatological factors, such as temperature, humidity, wind speed, radiation, type, and growth stage of the crop etc. Artificial neural networks (ANNs) are effective tools to model nonlinear systems. A neural network model is a mathematical construct whose architecture is essentially analogous to the human brain. Basically, the highly interconnected processing elements arranged in layers are similar to the arrangement of neurons in the brain.

According to Sudheer *et al.*, (2003), the main advantage of ANN methods over conventional methods is the ability for solving problems, which are difficult to formalize. In other word, an Artificial Neural Networks (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.

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, computational effort, and less real time control. An ANN can quickly present sensitive responses to tiny input changes in a dynamic environment.

Many researchers have investigated the applicability of ANN in hydrology to rainfall-runoff modeling (Meiaraj, 2008, Sinha, 2011), short-term stream flow (Sinha, 2013).

## **MATERIALS AND METHODS**

### ***Study Area & Collection of Data***

The study site is located in Bhagwanpur khurd Ambikapur, district Sarguja, Chhattisgarh at 23°09' N Latitude, 83°08' E Longitude and altitude of 611 m (above MSL). Data on maximum temperature ( $T_{\max}$ ), minimum temperature ( $T_{\min}$ ), relative humidity (RH), and wind speed (WS), sunshine hour (SSH) of past 18 years (1 January 1996 to 31 December 2013) were collected from Department of Agro-meteorology, Rajmohini Devi College of Agricultural and Research Station Ambikapur IGKV, Raipur (C.G.).

Daily  $ET_0$  values were estimated using the PM method. The PM estimated  $ET_0$  values were considered as standard and used for training and testing of different architectures of ANN. Out of total number of 6576 patterns, 4600 (1 January 1996 to 3 August 2008) has been used for training the network while remaining 1976 (4 August 2008 to 31 December, 2013) has been used for testing the model, in which we have taken 2 input variables i.e.  $T_{\max}$  and SSH, input selection is based on correlation matrix. The training and validation datasets were used to train ANNs against  $ET_0$  estimated by FAO-56 PM method, whereas the test dataset was used to estimate  $ET_0$  by trained ANN models.

## Research Article

### Method

Calculation of Reference Evapotranspiration  $ET_0$

The FAO Penman–Monteith Equation

The FAO-PM equation recommended for daily  $ET_0$  ( $\text{mm day}^{-1}$ ) estimation (Allen *et al.*, 1998) may be written as.

$$ET_0 = \frac{0.408 \Delta (R_n - G) + 900 \gamma u_2 (e_s - e_a) / (T + 273)}{\Delta + \gamma (1 + 0.34 u_2)} \quad \text{----- (1)}$$

Where:

$ET_0$  - Reference evapotranspiration [ $\text{mm day}^{-1}$ ],

$R_n$  - Net radiation [ $\text{MJ m}^{-2} \text{day}^{-1}$ ],

$G$  - Soil heat flux density [ $\text{MJ m}^{-2} \text{day}^{-1}$ ],

$T$  - Mean daily air temperature at 2 m height [ $^{\circ}\text{C}$ ],

$u_2$  - Wind speed at 2 m height [ $\text{m s}^{-1}$ ],

$e_s$  - Saturation vapour pressure [ $\text{kPa}$ ],

$e_a$  - Actual vapour pressure [ $\text{kPa}$ ],

$e_s - e_a$  - Saturation vapour pressure deficit [ $\text{kPa}$ ],

$\Delta$  - Slope of the vapour pressure curve [ $\text{kPa } ^{\circ}\text{C}^{-1}$ ],

$\gamma$  - Psychrometric constant [ $\text{kPa } ^{\circ}\text{C}^{-1}$ ].

Equation (1) determines the  $ET_0$  from an assumed grass reference surface and serves as a standard to which evapotranspiration in different periods of the year. Based on the recommendations of the expert consultation of FAO methodologies for Crop Water Requirements (UNFAO, 1998) and the results of other studies, the United Nations Food and Agriculture Organization (Allen *et al.*; UNFAO, 1998) brought out a comprehensive treatise in FAO Irrigation and Drainage Paper 56, entitled Crop Evapotranspiration, Guidelines for computing Crop Water Requirement.

### Development of Artificial Neural Network Models

In this study back propagation algorithm (Rumelhart and McClelland, 1986) is used to develop ANN models. It includes following steps:

1. Randomize the weights  $\{w_s\}$  to small random values (both positive and negative) to ensure that the network is not saturated by large values of weights.
2. Select an instance  $t$ , that is the vector  $\{x_k^{(i)}\}$ ,  $i = 1, \dots, N_{\text{inp}}$  (a pair of input and output patterns), from the training set.
3. Apply the network input vector to network input.
4. Calculate the network output vector  $\{z_k^{(i)}\}$ ,  $k = 1, \dots, N_{\text{out}}$ .
5. Calculate the errors for each of the outputs  $k$ ,  $k = 1, \dots, N_{\text{out}}$ , the difference between the desired output and the network output.
6. Calculate the necessary updates for weights  $\Delta w_s$  in a way that minimizes this error.
7. Adjust the weights of the network by  $\Delta w_s$ .
8. Repeat steps 2 – 6 for each instance (pair of input–output vectors) in the training set until the error for the entire system is acceptably low, or the pre-defined number of iterations is reached.

It is used in layered feed-forward ANNs. The artificial neurons are arranged in layers, and send their signals “forward”, and then the errors are propagated backwards. The network receives inputs by neurons in the input layer, and the output of the network is given by the neurons on an output layer. There may be one or more intermediate hidden layers. The back propagation algorithm uses supervised learning. The idea of the back propagation algorithm is to reduce the error (difference between actual and expected results), until the ANN learns the training data.

## Research Article

In the present study the BPANN was designed by using MATLAB codes. A programme was written, edited, debugged and run in MATLAB. The programme was suitably modified to accommodate different input patterns and models. The programme is flexible to accommodate different activation functions (tansig, logsig, and purelin), performance functions (mse, sse and msereg), training algorithm (trainbr, trainlm etc.) and a preset number of iterations. The programme takes input data file in the 'delimited text files without header' format with standard extension as 'txt' and gives output in the same format. The output file was then converted into 'ms excel' file with fixed width format.

In this study 'logsig' activation function was used. The training algorithm chosen was Levenberg-Marquardt (trainlm) and the performance function chosen is sum squared error (sse).

### Performance Evaluation of Models

The statistical model evaluation criteria considered in this study are as follows:

#### Mean Absolute Deviation (MAD)

It is a measure of mean absolute deviation of the observed values from the estimated values. It has a unit and is not a normalized criterion. It is expressed as,

$$MAD = \frac{\sum_{j=1}^n |O_j - S_j|}{n} \quad \text{---- (2)}$$

Where,  $O_j$  = Observed  $ET_O$  (mm/day),  $S_j$  = Simulated  $ET_O$  (mm/day),  $n$  = Total number of observations.

#### Root Mean Square Error (RMSE)

It is an alternative to the criterion of residual error (Yu, 1994) and is expressed as the measure of mean of the residual variance summed over the period, that is,

$$RMSE = \sqrt{\frac{\text{residual variance}}{n}} = \left( \frac{\sum_{j=1}^n (O_j - S_j)^2}{n} \right)^{1/2} \quad \text{---- (3)}$$

#### Correlation Coefficient (CC)

The correlation between the observed and simulated values is described by the correlation statistic, called the correlation coefficient. It is estimated by the equation:

$$CC = \frac{\sum_{j=1}^n \left\{ (O_j - \bar{O}) (S_j - \bar{S}) \right\}}{\left\{ \sum_{j=1}^n (O_j - \bar{O})^2 \sum_{j=1}^n (S_j - \bar{S})^2 \right\}^{1/2}} \times 100 \quad \text{----- (4)}$$

Where,  $\bar{O}$  and  $\bar{S}$  are mean of observed and simulated  $ET_O$  values.

#### Coefficient of Efficiency (CE)

Nash and Sutcliffe (1970) proposed the criterion on the basis of standardization of the residual variance with initial variance and named it as the coefficient of efficiency.

$$CE = \left\{ 1 - \frac{\text{residual variance}}{\text{initial variance}} \right\} \times 100 = \left\{ 1 - \frac{\sum_{j=1}^n (O_j - S_j)^2}{\sum_{j=1}^n (O_j - \bar{O})^2} \right\} \times 100 \quad \text{----- (5)}$$

#### Index of Agreement (IOA)

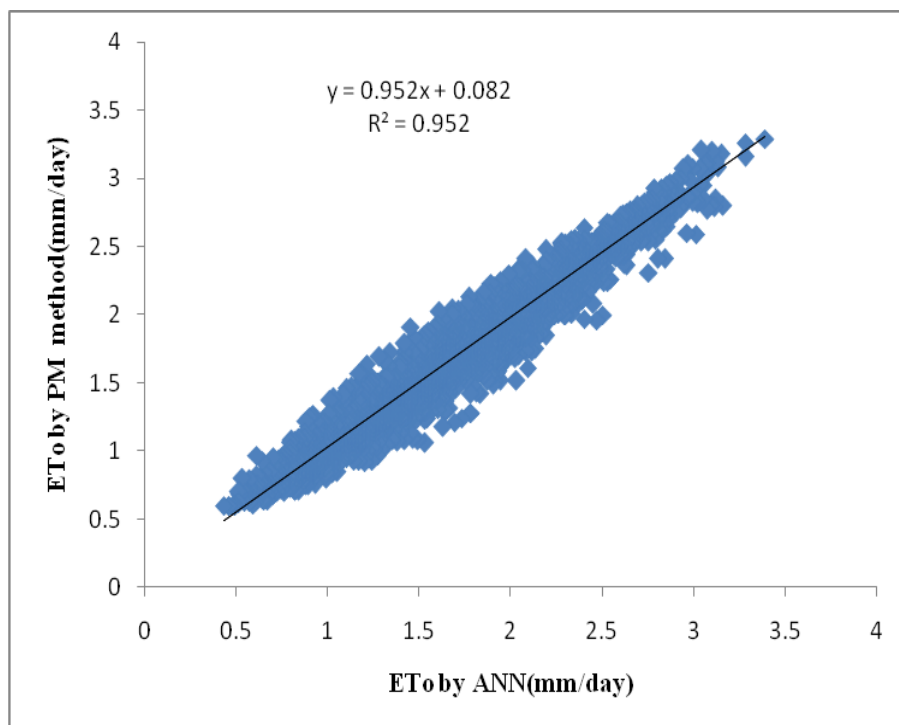
$$IOA = 1 - \left( \frac{\sum (obs - pre)^2}{\sum (abs(pre - mean(obs)) + abs(obs - mean(obs)))^2} \right)$$

## Research Article

$$IOA = 1 - \frac{\sum_{j=1}^n (s_j - o_j)^2}{\sum_{j=1}^n ((|s_j - o|) + (|o - o|))^2} \quad \text{----- (6)}$$

## RESULTS AND DISSCUSSION

ANNs have been successfully implemented to model evapotranspiration in several studies.



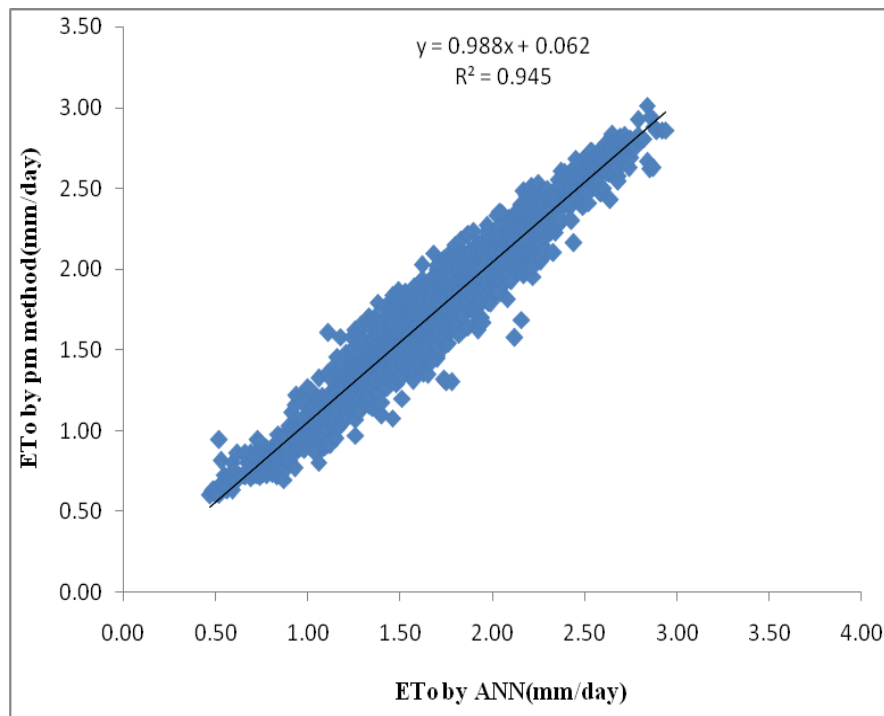
**Figure 1: Relationship between  $ET_0$  estimated by ANN method and  $ET_0$  estimated by PM method during Training period**

These studies indicated that the ANN models can be used as an alternative method to estimate  $ET_0$ . The performance of ANN models reported in these studies was superior to respective conventional methods of  $ET_0$  estimation. Most of the studies considered ANN modeling of FAO-56 PM method  $ET_0$  values. The FAO-56 PM method  $ET_0$  values are the calculated values which were essentially derived from the basic climatic data. Using these calculated values as target, the high model performance of the ANN models could be attained. In this study Artificial Neural Network models have been trained with 70% of the patterns (4600) and tested with 30% (1975) of the patterns 6575. In which we have taken 2 inputs i.e.  $T_{max}$  and SSH. Values of different performance evaluation criteria viz. MAD, RMSE, CC, CE and IOA are analyzed both during training and testing. After detailed numerical experiment the best performance in each model has been selected. The best network was selected using minimum value of SEE and maximum value of model efficiency.

### Comparison of Developed Models

Performance of the developed ANN models has been assessed by comparing the FAO-56 Penman Montieth method. The optimal ANN (2-6-1) for Ambikapur regions showed a satisfactory performance in the  $ET_0$  estimation as compared to Penman Montieth method. It showed the highest CC value as 95.28, highest CE value as 90.79 and highest IOA value as 97.52 during training and highest CC, CE and IOA value as 97.28, 93.67 and 98.43 during testing. Performance of Artificial Neural Network during Training and Testing is given in table-1. Scatter plots of  $ET_0$  estimated by ANN method with that estimated by PM method during training and testing is explained in figure 1 and figure 2, respectively.

## Research Article



**Figure 2: Relationship between  $ET_0$  estimated by ANN method and  $ET_0$  estimated by PM method during Testing period**

**Table 1: Performance of Artificial Neural Network during Training and Testing**

Model Name/ Network Architecture (NOIN- NOHN- NOON)		Training						Testing		
		MAD	RMS E	CC (%)	CE (%)	IOA	MA D	RMS E	CC (%)	CE (%)
Input combination										
2-6-1	Tmax+SS	0.094	0.172	95.2	90.7	97.5	0.100	0.129	97.28	93.67
	H	2	1	8	9	2				98.43

NOIN- Number of input nodes, NOHN- Number of hidden nodes, NOON- Number of output nodes

## Conclusion

In general, ANN's learning process is affected by several factors such as number of input nodes, number of hidden layers, and number of nodes in hidden layer, learning algorithms, and learning cycles. In this study, it was found that single hidden layer networks are sufficient to map the non-linearity between  $ET_0$  and climatic factors. Learning performance did not improve with increased number of nodes in the hidden layer and was identical for both the training schemes tested. Secondly, improvement in ANN's learning performance depends upon the number of input nodes and the climatic variable corresponding to the nodes. Networks corresponding to FAO Penman Montieth method gave lower  $ET_0$  value compared to networks corresponding to Artificial Neural Network. The best ANN architectures corresponding to FAO Penman Montieth method were found to be 2-6-1, for Ambikapur stations. In general, the ANN model

### **Research Article**

developed networks gave better estimates of  $ET_0$  as compared to PM method. Thus, it can be concluded that ANNs are better tools for estimating  $ET_0$  than the conventional methods.

### **REFERENCES**

- Allen RG, Pereira LS, Raes D and Smith M (1998).** Crop evapotranspiration guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56, FAO, Rome.
- Hargreaves GH and Samani ZA (1985).** Reference crop evapotranspiration from temperature. *Applied Engineering In Agriculture* **1**(2) 96–99.
- Kumar M, Raghuwanshi NS, Singh R, Wallender WW and Pruitt WO (2002).** Estimating evapotranspiration using artificial neural network. *Journal of Irrigation and Drainage Engineering ASCE* **128**(4) 224–233.
- McCulloch WS and Pitts WH (1943).** A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics* **5** 115.
- Meiaraj C and Sundararajan R (2008).** Surface Runoff Modelling of Alanthurai Watershed using Artificial Neural Networks. *Journal of Soil and Water Conservation, SCSI New Delhi* **7**(2) 57-62.
- Nash JE and Sutcliffe JV (1970).** River flow forecasting through conceptual models-I: A discussion of principles. *Journal of Hydrology* **10**(3) 238–250.
- Sinha J (2011).** Artificial Neural Network approach to rainfall-runoff modelling for upper Kharun catchment in Chhattisgarh. Ph.D. Thesis, Indira Gandhi Krishi Vishwavidyalaya, Raipur (C.G.).
- Sinha J, Sahu RK, Agarwal A, Senthil AR and Sinha BL (2013).** Levenberg – Marquardt algorithm based ANN approach to rainfall – runoff modeling. *Journal of Soil and Water Conservation, SCSI New Delhi* **12**(1) 48-54.
- Sudheer KP, Gosain AK and Ramasastri KS (2003).** Estimating actual evapotranspiration from limited data using neural computing technique. *Journal of Irrigation and Drainage Engineering ASCE* **129**(3) 214–218.
- Yu PS, Liu CL and Lee TY (1994).** Application of a transfer function model to a storage runoff process. *Proceeding of Stochastic and Statistical Methods in Hydrology and Environmental Engineering, Beijing* 87-97.