40

7

# Thesis Work

On

## *Prediction of Evaporation Using ANN and MLR*

*And*

## *Sensitivity analysis for Evapotranspiration*

**Conducted at National Institute of Hydrology, Roorkee Uttarakhand**

**Institute of Agricultural Science ,BHU (Varanasi)**

**Department of Farm Engineering**

**Submitted By**
**Miss. Aparajita Singh**
**(SWCE14295)**

**Under the guidance of**

**Dr. A. R. Senthil Kumar**
**Scientist 'E'**
**SWH Division**
**NIH,ROORKEE**

**Dr. R.M Singh**
**professor**
**Deptt.of Farm**
**Engg,IAS,BHU**

# CERTIFICATE

This is to certify that **Miss. Aparajita Singh** has undergone the thesis work on "***Prediction of Evaporation Using ANN and MLR and Sensitivity Analysis of Evapotranspiration***" from 1st January, 2016 to 8th June 2016 submitted to the Research Management and Outreach Division, National Institute of Hydrology, Roorkee, in partial fulfillment of the requirement for the award of degree of "**Master of Technology"** in **Agricultural Engineering** specialization in **Soil and Water Conservation Engineering** from Department of Farm Engineering, Institute of Agricultural Science at Banaras Hindu University is the original work carried out by her under my supervision and guidance.

A. R. Senthilkumar

**Dr. A. R. Senthil Kumar**
**Scientist 'E'**
**SWH Division**

# ACKNOWLEDGEMENT

I would first like to thank my thesis co- advisor Dr. A.R Senthil Kumar Scientist 'E' of the Surface Water Hydrology Division, at National Institute of Hydrology, Roorkee. The door to Dr. Kumar office was always open whenever I ran into a trouble spot or had a question about my research or writing. He consistently allowed this paper to be my own work, but steered me in the right direction whenever he thought I needed it.

I would also like to thank the experts who were involved in the validation survey for this research project. Shri. Digambar Singh, Scientist 'C' Surface Water Hydrology Division, National Institute of Hydrology, Roorkee. Without their passionate participation and input, the validation survey could not have been successfully conducted.

I would like to acknowledge Prof. Dr. R.M Singh of the Department of Farm Engineering, Institute of Agricultural Science at Banaras Hindu University as the second reader of this thesis, and I am gratefully indebted to him for his very valuable comments on this thesis.

I would also like to thank Prof. Dr.V. K. Chandola , Head, Department of Farm Engineering, Institute of Agricultural Science at Banaras Hindu University for giving me such opportunity to carried out this research work. I am gratefully indebted to his valuable Support for carrying out my thesis work. I would like to thank Prof . Dr. A. K Nema Department of Farm Engineering, Institute of Agricultural Science at Banaras Hindu University to his valuable support. Even I would like to thank all the faculty members of Department and staff of Department of Farm Engineering, Institute of Agricultural Science at Banaras Hindu University.

Finally, I must express my very profound gratitude to my parents, younger brothers and to Mr. B.K Singh for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Aparajita Singh

## CONTENTS


| Sl. No | PARTICULARS | Page No. |
|---|---|---|
| | List of tables | i |
| | List of figures | ii |
| | **ABSTRACT** | 1 |
| I | **INTRODUCTION** | 3 |
| II | **REVIEW OF LITERATURE** | 10 |
| | 2.1 Artificial neural networks (ANNs) | 10 |
| | 2.2 Application of artificial neural networks (ANNs)modeling to evaporation | 11 |
| | 2.3 Multiple Linear Regression(MLR) | 12 |
| | 2.4 Application of Multiple Linear Regression (MLR) modelling to evaporation | 13 |
| III | **STUDY AREA** | 16 |
| | 3.1 Data availability | 16 |
| IV | **MATERIALS AND METHODS** | 21 |
| | 4.1 Artificial Neural Network | 21 |
| | 4.1.1 Model Development | 22 |
| | 4.1.2 Normalization of Input Data | 23 |
| | 4.1.3 Feed forward Neural Network (FNN) | 23 |
| | 4.1.4 Training with Algorithm | 25 |
| | 4.1.5 Bayesian Regularization Algorithm (BR) | 26 |
| | 4.1.6 Network Architecture | 27 |
| | 4.1.7 Performance evaluation of the developed ANN model and MLR model | 27 |
| | 4.1.7.1 Visual observation evaluation | 27 |
| | 4.1.7.2 Quantitative evaluation | 28 |
| | 4.1.7.2(.i) Coefficient of efficiency(CE) | 28 |
| | 4.1.7.2.(ii) Root Mean Squared Error (RMSE) | 28 |
| | 4.1.7.2(iii) Explained Variance (EV) | 28 |
| | 4.2 Multiple Linear Regression (MLR) based evaporation estimation model | 29 |
| | 4.2.1 Performance evaluation of the developed MLR model | 29 |
| | 4.2.1.1 Visual observation evaluation | 29 |
| | 4.2.1.2 Quantitative evaluation | 30 |
| | 4.3 Reference evapotranspiration | 30 |
| | 4.3.1 Temperature based methods | 30 |
| | 4.3.1.1 Hargreaves method | 30 |
| | 4.3.1.2 Thornthwaite method | 31 |
| | 4.3.2 Radiation based method | 31 |

| | | |
|---|---|---|
| | 4.3.2.1 Turc method | 31 |
| | 4.4 Sensitivity analysis | 32 |
| V | **RESULTS AND DISCUSSION** | 34 |
| | 5.1 Development of ANN model | 34 |
| | 5.1.1 Selection of Input Vector | 34 |
| | 5.1.2 Training of ANN model | 37 |
| | 5.2 Analysis of result of ANN and MLR models | 39 |
| | 5.3 Sensitivity analysis by Thornthwaite method | 44 |
| | 5.4 Sensitivity analysis by Turc method | 48 |
| V | **CONCLUSIONS** | 55 |
| VI | **REFERENCES** | 57 |

# LIST OF TABLES


| Table number | Title | Page number |
|---|---|---|
| 3.1 | The average meteorological data for the period from Jan 2001 to Dec 2013 | 20 |
| 5.1 | Result of ANN model during calibration and validation | 39 |
| 5.2 | Comparison of results between ANN and MLR model | 42 |
| 5.3 | Percent increase in ET with increase in temperature (Thornthwaite method) | 44 |
| 5.4 | Average percent increase in ET with increase in temperature (Thornthwaite method) | 45 |
| 5.5 | Percent increase in ET with increase in sunshine (Thornthwaite method) | 46 |
| 5.6 | Average percent increase in ET with increase in sunshine(Thornthwaite method) | 46 |
| 5.7 | Average percent increase in ET with increase in temperature and sunshine(Thornthwaite method) | 47 |
| 5.8 | Percent increase in ET with increase in temperature(Turc method) | 49 |
| 5.9 | Average percent increase in ET with increase in temperature(Turc method) | 50 |
| 5.10 | Percent increase in ET with increase in solar radiation(Turc method) | 51 |
| 5.11 | Average percent increase in ET with increase in solar radiation(Turc method) | 51 |
| 5.12 | Average percent increase in ET with increase in temperature and solar radiation (Turc method) | 52 |
| 5.13 | Average percent increase in ET with increase in independent variables(temperature, sunshine, solar radiation) by both Thornthwaite and Turc method | 53 |

# LIST OF FIGURES


| Figure number | Title | Page number |
|---|---|---|
| 2.1 | Multiple linear regression | 13 |
| 3.1 | Daily maximum temperature at Roorkee from Jan 2001 to Dec 2013 | 17 |
| 3.2 | Daily minimum temperature at Roorkee from Jan 2001 to Dec 2013 | 17 |
| 3.3 | Daily humidity at Roorkee from Jan 2001 to Dec 2013 | 18 |
| 3.4 | Daily rainfall data at Roorkee from Jan 2001 to Dec 2013 | 18 |
| 3.5 | Monthly maximum, minimum, average temperature, relative humidity at Roorkee from Jan 1987 to Dec 2013 for sensitivity analysis of evapotranspiration | 19 |
| 4.1 | Typical artificial neurons | 22 |
| 4.2 | Typical feed forward network | 25 |
| 5.1 | Cross correlation between rainfall and evaporation | 35 |
| 5.2 | Cross correlation between maximum temperature and evaporation | 36 |
| 5.3 | Cross correlation between minimum temperature and evaporation | 36 |
| 5.4 | Cross correlation between relative humidity and evaporation | 37 |
| 5.5 | Graph plot for the result of the best ANN model during calibration (4-4-1) | 40 |
| 5.6 | Graph plot for the result of the best ANN model during validation (4-4-1) | 40 |
| 5.7 | Graph plot for the result of the best MLR model during calibration | 41 |
| 5.8 | Graph plot for the result of the best MLR model during validation | 41 |
| 5.9 | Observed and computed daily evaporation during calibration (comparison between ANN, MLR, observed evaporation) | 42 |
| 5.10 | Observed and computed daily evaporation during validation (comparison between ANN, MLR, observed evaporation) | 43 |
| 5.11 | Average percent increase in ET with increase in temperature(Thornthwaite method) | 45 |
| 5.12 | Average percent increase in ET with increase in sunshine(Thornthwaite method) | 47 |
| 5.13 | Average percent increase in ET with increase in sunshine and temperature | 48 |
| 5.14 | Average percent increase in ET with increase in temperature(Turc method) | 50 |
| 5.15 | Average percent increase in ET with increase in solar radiation (Turc method) | 52 |
| 5.16 | Average percent increase in ET with increase in temperature and solar radiation(Turc method) | 53 |
| 5.17 | Average percent increase in ET with increase in independent variables(temperature,solar radiation,sunshine) by both thornthwaite and turc method | 54 |

# ABSTRACT

Artificial Neural Network based modeling technique has been used to study the influence of different combinations of meteorological parameters on evaporation. The data set used is taken from the records of NIH Observatory. The input combination selected based on the statistical properties of data was used in predicting the evaporation. The prediction accuracy of Artificial Neural Network has also been compared with the accuracy of multiple linear regressions for predicting evaporation. The comparison demonstrated superior performance of Artificial Neural Network over multiple linear regression approach. The findings of the study also revealed the requirement of all input parameters considered together, instead of individual parameters taken one at a time as reported in earlier studies, in predicting the evaporation

The RMSE of ANN model during calibration and validation was found to be 1.0151 and 1.0045 respectively, whereas for the MLR model, RMSE value during calibration and validation was 1.1874 and 1.1523 respectively, and also the ANN model efficiency during calibration and validation was 0.7800 and 0.7830 respectively, whereas the MLR model efficiency during calibration and validation was 0.6988 and 0.8501 respectively, indicates a substantial improvement in the model performance. In addition, comparison of the scatter plots of time series indicates that the values of Evaporation estimated by the ANN model are more precise than those found by the MLR.

The highest correlation coefficient (0.7800), (0.7830) along with lowest root mean square error (1.0151),(1.0045) during calibration and validation of best evaporation model [4-4-1] was obtained with the input combination of rainfall, maximum temperature, minimum temperature, and mean relative humidity. A graph between the actual and predicted values of evaporation suggests that most of the values lie within a scatter of ±15% with all input parameters. The findings of this study suggest the usefulness of ANN technique in predicting the evaporation losses.

The estimation of evapotranspiration is essential in water resources management. Among a group of methods, the Penman–Monteith has been commonly applied to calculate reference evapotranspiration as this method has been also recommended by the Food and Agriculture Organization of the U.N. (FAO). Other methods widely used are: the FAO 24 Penman, the modified Blaney and Criddle, the FAO 24 Makkink, and the Hargreaves, Thornthwaite method, Turc method.

Sensitivity analysis is required to gain a better understanding of the meteorological systems; particularly to indicate the physical meaning of each meteorological parameter used in the estimation of the reference evapotranspiration. The standardized FAO56 Penman-Monteith model, which has been the most reasonable method in both humid and arid climatic conditions, provides reference evapotranspiration (ETo) estimates for planning and efficient use of agricultural water resources. And sensitivity analysis is important in understanding the relative importance of climatic variables to the variation of reference evapotranspiration. In this study, a non-dimensional relative sensitivity coefficient was employed to predict responses of ETo to perturbations of four climatic variables in the NIH Campus, Roorkee. A historical dataset of monthly average temperature, relative humidity, solar radiation, daily sunshine duration from 1987-2013 in the NIH Campus, Roorkee was used in the analysis. Results have shown that monthly sensitivity coefficients exhibited large fluctuations during the growing season, and solar radiation was the most sensitive variable in general for the NIH, followed by average temperature, sunshine hours and relative humidity. According to this study, the response of ETo can be preferably predicted under perturbation of average temperature, sunshine hours, relative humidity and solar radiation by their sensitivity coefficients.

Results show that the influence of the variables to evapotranspiration is not the same for each period, and also the order that the variables influence evapotranspiration is changing. A comparison between the two evapotranspiration methods( THORNWAITE AND TURC) shows that solar radiation and temperature are the main parameters that affect evapotranspiration, while relative humidity and wind speed are not so important for the calculation of evapotranspiration

# 1.INTRODUCTION

Evaporation is one of the major processes of the hydrological cycle in which water is changed from the liquid into vapour through the transfer of heat energy. The management of scarce water resources for sustainable crop production in the face of explosive growth of population is becoming more and more important. In hot climate, the loss of water by evaporation from rivers, canals and open-water bodies is a vital factor as evaporation takes a significant portion of all water supplies. Even in humid areas, evaporation loss is significant, although the cumulative precipitation tends to mask it due to which it is ordinarily not recognized except during rainless period. Design and management of water resources require knowledge of the magnitude and variation of evaporation losses. Therefore, the need for reliable models for quantifying evaporation losses from increasingly scarce water resources is greater than ever before. Water resource development projects and farm irrigation systems are basically designed on the basis of long term mean values of evaporation. Accurate estimation of evaporation is fundamental for effective management of water resources.

The process of evaporation, however, is influenced by number of factors. Meteorological parameters such as solar radiation, temperature, humidity and wind speed are the major parameters affecting evaporation.

Solar radiation affects the temperature and thus the evaporation by heating the air and the water surface. Usually, estimates of evaporation are needed in a wide array of problems in hydrology, agronomy, forestry and land resources planning, such as water balance computation, irrigation management, crop yield forecasting model, river flow forecasting, ecosystem modeling, etc. For example, the widely used Food and Agriculture Organization (FAO) crop monitoring and forecasting model is based on evaporation estimates which are related to crop growth and yield.

Where there is a sufficient water resource, irrigation can substantially increase crop yields, but again the scheduling of the water application is usually based on evaporation estimates. Numerous investigators developed models for estimation of evaporation. The interrelated meteorological factors having a major influence on evaporation have been incorporated into various formulae for estimating evaporation. Unfortunately, reliable estimates of evaporation are extremely difficult to obtain because of complex interactions between the components of the land-plant-atmosphere system.

There is increasingly growing demand for evaporation data for studies of surface water and energy fluxes, especially for the studies, which address the impacts of global warming. Evaporation involves the transformation of water from its liquid state into a gas and the subsequent diffusion of water vapour into the atmosphere. However, the measurement of evaporation in the open environment is difficult and is usually done by proxy.

Potential evaporation is the variable most often used. Potential evaporation is a measure of the ability of the atmosphere to remove water from a surface assuming no limit to water availability, whereas actual evaporation is the quantity of water that is removed from that surface by evaporation (Brutsaert,1982). Therefore, actual evaporation is only equal to potential evaporation when a given surface is saturated. The most widespread measurement method for potential evaporation uses a pan evaporimeter, which quantifies water loss from the instrument itself and not from the surrounding environment. The standard US Class A pan is the most commonly used instrument. It consists of a metal container usually covered by an open wire bird guard that is 1,207 mm across and 254 mm high. Evaporation is the amount of loss (gain) in mm depth with rainfall from an adjacent rain-gauge subtracted. Pan evaporation records may contain many artifacts of measurement (inhomogeneities) caused by equipment changes, exposure changes and location changes (Jones, 1992). More accurate estimates of potential evaporation can be obtained by applying other meteorological data to empirical, water budget, energy budget, and combination approaches.

However, the most accurate approaches tend to be resource-intensive, site specific and do not provide long-term estimates of change. Therefore pan evaporation records are the largest single source of data on historical evaporation trends and models can be helpful for agricultural research.

Empirical methods relate either of pan evaporation, actual lake evaporation or lysimeter measurements to meteorological factors using regression analysis. The most realistic method is to obtain direct evaporation from open water surface, be it from extensive water surface of a lake or from a pan. The evaporation pan is, however, the most widely used instrument for evaporation measurement today. Several types of evaporation pans are available, although the standard US Weather Bureau Class 'A' pan built of unpainted galvanized iron is currently the most popular throughout the world.

For many years, measurements taken on evaporation pans are used to provide estimates of the amount of evaporation from lakes and reservoirs. It has also been used for estimating evapotranspiration from agricultural crops using procedures relating evapotranspiration to pan evaporation (Snyder,1993). As pan evaporation combines the accumulated effects of all the climatic parameters, evaporation from a free water surface of an open pan is widely used as a climatic index for a particular region. Significant problems still exist in the measurement of evaporation.

Many times reliable pan evaporation data are not available because of variations in the shape and size of pans, their exposure, the presence or absence of algae in water, the specific methods of measuring the loss of water from the pans and the protection against use of water by birds and animals. Many studies were therefore undertaken to determine reliable relation between pan evaporation and meteorological factors (Singh et al.,1981). Based on this relationship, pan evaporation has been intensively studied for applications in irrigation scheduling. With the advancement of drip irrigation in horticultural crops like citrus, the irrigation scheduling based on pan evaporation is getting more popular due to improved yield (Shirgure et al., 2001). The management of pan evaporation is proved to be useful in other climatological applications. Chattopadhaya and Hulme (1997) have linked trends in pan evaporation measurements to climate change in India. The effect of various weather parameters on pan evaporation was investigated by Xu and Singh (1998; 2001).

Models developed to date are recognized procedures for estimating evaporation. Since no single model is universally adequate under all climatic conditions, it is difficult to select the most appropriate evaporation model for a given region. This is partly because of the availability of many equations for determining evaporation, the wide range of data types needed and the wide range of expertise needed to use the various equations correctly. More importantly, objective criteria for model selection are lacking. Consequently, the conditions under which one evaporation model would be more suitable are not always spelled out. The models developed from meteorological data involve empirical relationships to some extent. The empirical relationships account for many local conditions. Therefore, most models may give reliable results when applied to climatic conditions similar to those for which they were developed. Without some local or regional calibration, the use of such models for climatic conditions that are greatly different may give results that may differ considerably.

Evaporation is a complex and nonlinear phenomenon because it depends on several interacting climatological factors, such as temperature, humidity, winds speed, bright sunshine hours, etc. Artificial neural networks (ANN) are effective tools to model nonlinear systems (Kumar et. al. 2002,Sudheer et. al. 2003). 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. The ANN have found successful applications in the areas of science, engineering, industry, business, economics and agriculture. Recently, artificial neural networks have been applied in meteorological and agro ecological modeling and applications (Hoogenboom, 2000). Most of the applications reported in literature concern estimation, prediction and classification problems. Neural network applications have diffused rapidly due to their functional characteristics, which provide many advantages over traditional analytical approaches.

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. Instead, ANN, map the implicit relationship between inputs and outputs through training by field observations. 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 are useful, requiring fewer input and computational effort and less real time control. An ANN can quickly present sensitive responses to tiny input changes in a dynamic environment. Forecasting of pan evaporation particularly in water resource projects planning, design and operation is of paramount importance. Pan evaporation varies spatially and temporally. Spatial distributed measurements of pan evaporation are also beneficial for use in various water resources planning and development programs.

Evapotranspiration is an important component of the hydrologic cycle as it can significantly affect the water budget of the natural (i.e. approximately 62% of all precipitation falling on land is evapotranspirated. Consequently, its accurate estimation is essential for, among others, water availability, plant growth, irrigations efficiency, reservoir operation, and water resources management. Several empirical methods have been developed to derive evapotranspiration estimates. Among others, the Penman–Monteith method is recommended by the Food and Agriculture Organization of the U.N. (FAO) as the sole method to calculate reference evapotranspiration, wherever the required input data (i.e. temperature, relative humidity, solar radiation, wind speed) are available (e.g., Allen et al., 1998; Ampas, 2010). Other methods widely used are the FAO 24 Penman method, the FAO 24 Blaney and Criddle method (Doorenbos and Pruitt, 1977), the FAO 24 Makkink method, and the Hargreaves method.

Sensitivity analysis has been an important stage on the evaluation of environmental models; however, current research urges the need to assess the physical meaning of model parameters and their relative influence on the meteorological variables. By definition sensitivity analysis studies the impact of the change of one parameter to another (McCuen, 1973).

Several studies have assessed the parameter sensitivity to estimated evapotranspiration using sensitivity coefficients which were calculated for several independent variables as meteorological parameters, physiological parameters, and climatic conditions. Comparison of sensitivity coefficients has showed the relative importance of each variable. Saxton (1975) conclude that the most important variable for the calculation of ETo, during summer is solar radiation, whereas in autumn and spring the most important variable is the aerodynamic variable. Coleman and DeCoursey (1976) conclude that the most important parameter at the annual scale is relative humidity; during summer both temperature and solar radiation are the most important variables, whereas relative humidity is more important during winter. They also conclude that wind speed has very small importance at the annual scale. Babajimopoulos et al. (1992) conclude that temperature and solar radiation are the most important variables in the summer, whereas the most important parameter in the winter is relative humidity (wind speed has very small importance). Gong et al. (2006) evaluated sensitivity coefficients for the Yangtze River basin and indicated their large spatial variability. Irmak et al. (2006) evaluated sensitivity coefficients for areas under different climatic characteristics. Results showed large spatial variability, and the authors concluded that for areas with strong and dry winds wind speed was the most important variable.

Here, parameter sensitivity to the estimated evapotranspiration based on two methods (Thornwaite and Turc method) has been assessed and evaluated the impact of the change of the measured meteorological variables to the estimated reference evapotranspiration. Finally, relative influence of each meteorological parameter to reference evapotranspiration has been compared. The sensitivity analysis is based on a sensitivity coefficient designed for the comparison of the influence of the independent parameters and uses standard deviation.

In this study, an attempt has been made to develop ANN and MLR based evaporation estimation models using climatic parameters as inputs and evaporation as output for NIH, Roorkee (Uttarakhand), India with the following objectives:

1. To develop and test the pan evaporation prediction models using various weather parameters as input variables with artificial neural networks (ANN) and validate with the independent subset of data for NIH Campus Roorkee (Uttarakhand).
2. To test the suitability of the artificial neural networks for modeling pan evaporation in comparison with multiple linear regression analysis.
3. Sensitivity analysis for evapotranspiration for the data obtained from NIH, Roorkee(uttarakhand)

In this analysis, ANN model with feed forward structure have been developed to predict evaporation with antecedent rainfall, relative humidity, maximum and minimum temperature as input vector. The data of maximum and minimum temperature, relative humidity and antecedent rainfall from Jan 2001 to Dec 2013 at NIH campus have been used to compute the evaporation, which has been fed to ANN as an input. The performance of ANN model with multiple linear regressions (MLR) is compared with the performance of ANN model with real input vector. Data from Jan 1987 to Dec 2013 was used for sensitivity analysis.

# 2. REVIEW OF LITERATURE

An extensive survey of literature on artificial neural network modeling, its applications to estimation of evaporation was made and presented in this chapter. The review of literature indicates the importance of artificial neural networks (ANNs), particularly with application to evaporation for improving the efficiency of the hydrological model.

## 2.1 Artificial neural networks (ANNs)

The development of artificial neural networks (ANNs) began approximately 50 years ago (McCulloch and Pitts, 1943), inspired by a desire to understand the human brain and emulate its functioning. Within the last two decades, it has experienced a huge resurgence due to the development of more sophisticated algorithms and the emergence of powerful computation tools. Extensive research has been devoted to investigate the potential of artificial neural networks (ANNs) as computational tools that acquire, represent, and compute a mapping from one multivariate input space to another (Wasserman, 1989). The ability to identify a relationship from given patterns make it possible for ANNs to solve large scale complex problems such as pattern recognition, nonlinear modeling, classification, association, and control.

Since the early nineties, ANNs have been successfully used in hydrology and water resources engineering such as rainfall-runoff modeling, stream flow forecasting, ground-water modeling, water quality, water management policy, precipitation forecasting, hydrologic time series, and reservoir operations. More concepts and application of ANN models in hydrology has been discussed by Govindraju and RamchandraRao (2000) and by the ASCE task committee on application of artificial neural networks in hydrology (2000 a, b).

There are many different types of ANN models in practice. Multi-layer feedforward neural networks are perhaps the favorite and perform well in most ANN applications. Maier and Dandy (2000) reported that more than 95% of the ANN related papers they reviewed in the water resources area used feed-forward networks. In forecasting time series, the feed-forward network can be viewed as a general nonlinear auto-regressive model. The linear auto-regressive (AR) models are special cases of ANN without hidden nodes (Zhang *et al.* 2001).

## 2.2 Application of artificial neural networks (ANNs) modelling to Evaporation

Prediction of evaporation has significant application in water resource utilization, and management and also for an overall development of the basin. The purpose of observation of evaporation is primarily in studying its temporal and spatial changes.

To date, a wide variety of models have been developed and applied for evaporation forecasting. These models can be categorized into empirical time series model and physical descriptive model. The major drawback of empirical approach is that they are not adequate for forecasting when the dynamical behavior of the hydrological system changes with time. Similarly, physics based model, in practice requires enormous data, in particular data pertaining to atmosphere that is generally difficult or expensive, to simulate evaporation in developing countries like India.

In recent years, artificial neural networks (ANNs) have been used for forecasting in many areas of science and engineering. ANNs have been proven to be effective in modeling virtually any nonlinear function to an arbitrary degree of accuracy. Many researchers used ANN for modelling evaporation. These are presented as follows:

In the last decades, artificial neural networks (ANNs) have been successfully applied in water resources management. Recent experiments have reported that ANN may offer a promising alternative in the hydrological context (Cancelliere *et al.*, 2002; Cigizoglu and Kisi, 2005,2006;Cobaner *et al.*, 2009; Guven and Kisi, 2011; Keskin and Terzi,2006; Kisi, 2008a,b, 2009a,b, 2010; Kisi and Yildirim, 2005a,b; Kumar *et al.*,2004; Piri *et al.*,2009; Sudheer *et al.*, 2002; Supharatid,2003; Tabari *et al.*, 2010; Tan *et al.*, 2007; Tayfur, 2002). Sudheer *et al.* (2002) used a feed forward ANN to estimate evaporation and found that the ANN compared favorably to other conventional approach. Keskin and Terzi (2006) developed feed forward ANN models for modeling daily evaporation and found that the ANN model performed better than the conventional method. Tan *et al.* (2007) used an ANN technique for modeling hourly and daily open-water evaporation rates. Piri *et al.* (2009) estimated daily evaporations in a hot and dry climate by ANN models. Kisi (2009b) modeled daily pan evaporations using three different neural network techniques, multi-layer perceptron's (MLPs), radial basis neural networks (RBNNs) and generalized regression neural networks (GRNNs) and he found that the MLP and RBNN could be employed successfully to model the evaporation process using the available climatic data. Tabari *et al.* (2010) compared ANN and multivariate non-linear regression techniques for modeling daily pan evaporation and found that the ANN performed better than the non-linear regression. Guven and Kisi (2011) modeled daily pan evaporations using linear genetic programming and ANN models. Shiri and Kisi (2011) applied ANN and ANFIS techniques to model daily pan evaporation by using available and estimated climatic data. Nourani *et al.* (2012) applied three different artificial neural networks (ANNs) viz.: multi-layer perceptron (MLP), radial basis neural network (RBNN) and Elman network for estimating daily evaporation and results denoted the superiority of the ANN models on the classic models.

## 2.3 Multiple Linear Regressions

In this study, multiple linear regressions were used to estimate the pan evaporation for the study area. Multiple linear regressions (MLR) is a multivariate statistical technique used to model the linear correlations between a single dependent variable *Y* and two or more independent variables. The regression equation of *Y* can be written as:

$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots\ldots + \beta_k X_n \ldots$ (2.1)

where, $Y$ is the response variable; $X_1, X_2, \ldots\ldots X_n$ are the independent variables; and $\beta_0, \beta_1, \beta_2 \ldots\ldots\ldots \beta_k$ are the regression coefficients.

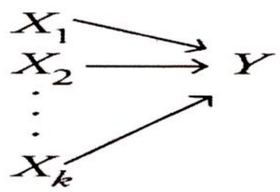

**Fig 2.1 Multiple Linear Regression**

## 2.4: Application of Multiple Linear Regression(MLR) modelling to Evaporation

Senapati *et al.*, (1985) studied the relationship between pan evaporation and various meteorological parameters like air temperature, wind speed, relative humidity and bright sunshine hours. Dalton's type equation was developed on seasonal and annual basis to predict evaporation. A linear multiple regression analysis revealed that the above meteorological parameters were found well correlated with pan evaporation with multiple correlation coefficient of 0.975.

Chandra *et al.*, (1988) studied the effect of meteorological parameters such as air temperature, relative humidity, wind speed and number of sunshine hours on pan evaporation. The prediction equations for the three seasons decided as per the trend of the parameters developed through single and multiple linear regression analysis. The study reveals that the above meteorological parameters can well be correlated with pan evaporation.

Singh *et al.*, (1992) investigated relationship between evaporation from US Class A open pan evaporimeter and meteorological parameters at Hisar. All five meteorological parameters (wind speed, maximum and minimum temperature, per cent relative humidity, bright sunshine hours per day and solar radiation) had significant influence at 1% level on evaporation rate at Hisar.

Wind velocity, sunshine hours, mean air temperature and solar radiation were positively correlated with evaporation and relative humidity was negatively correlated with evaporation. The highest correlation value ($r = 0.78$) were obtained with relative humidity. Multiple regression equation $E = -1.41 + 0.10\,T - 0.03\,RH + 0.20\,S + 0.21\,W + 0.04\,SD$ showed a maximum coefficient of determination ($R2 = 0.96$) when all the major meteorological parameters were considered together.

Sahu *et al.*, (1994) have established quantitative relationship between the pan evaporation and eight different meteorological parameters using stepwise multiple regression analysis. A software package MICROSTAT was used for the analysis. The variables viz. evaporating power of the air (Ea) and solar radiation (Rs) were found to be significant factors influencing pan evaporation for annual and seasonal basis while the weather parameters like air temperature, relative humidity and wind speed showed a significant effect during summer season. The prediction equations fitted for different seasons and annual basis explained more than 91 per cent variation in pan evaporation. With the help of the regression equations developed in this study, it may be possible to estimate pan evaporation in South Saurashtra agro-climatic zone. The multiple regression on annual basis was expressed as $E0 = -3.7629 + 0.5983Ea + 0.7218Rs + 0.1949W$.

Singh *et al.*, (1995) obtained simple correlations between different meteorological parameters and evaporation measured from US Class A open pan evaporimeter. The highest value of correlation coefficient (0.85) was found with maximum temperature followed by wind speed (0.82). The coefficient of determinates for minimum air temperature, relative humidity and bright sunshine hours were 0.70, – 0.56 and 0.15, respectively. After considering all these meteorological parameters together, a multiple regression equation $Y = 26.679 - 0.585\,Tmax + 0.653\,Tmin - 0.277Rh + 0.215\,Ss + 0.336\,Ws$ ($R2 = 0.92$) has been developed for the calculation of evaporation for Hisar region. The estimated values were very close to the observed values of evaporation.

Shrivastava *et al.*, (2000) made correlation between pan evaporation and other climatic parameters for Sundrebans, West Bengal. Weekly climatic data of 25 years (1963-87) have been statistically analyzed. To establish the relationship between pan evaporation and climatic parameters, the method of regression was adopted as suggested by Mendenhall and Sincich. The linear, quadratic and cubic regression equations have been developed for individual parameters and over all utility of the model has been tested considering the significance of leading coefficients for all the parameters. On the basis of overall utility of the model, a linear relationship was obtained between evaporation and parameters named minimum temperature, maximum humidity and wind speed. The relationship was quadratic for maximum temperature and minimum humidity whereas cubic relationship was obtained for sunshine hours. The multiple regression model $E = 17.2646 + 1.6613\,Tmax - 0.6575\,RH1 + 1.8715\,SH + 0.9462\,WS$ ($R2 = 0.921$) was developed. The evaporation values estimated using the existing methods namely, modified Blaney-Criddle Penman, Thornthwaite, Radiation and Christiansen and by the developed regression model were compared with the observed values. It has been observed that the multiple linear regression model gives better estimation of rate of evaporation for Sunderbans.

# 3. STUDY AREA

Roorkee is located in Hardwar district at $29^0 51'$ N and $77^0$ 53' E on the south bank of Solani River. The Upper Ganga Canal is the most important features and adds beauty to the city. Running from north to south, it divides the city in two distinct parts. City is located about 274 Meters above mean sea level and receives the average annual rainfall of 1068 mm, average Monsoon Rainfall of 878 mm and having average Max. Temperature 40 $^0$C and average Min. Temperature 2 $^0$C. Max. Humidity 100 %, Average Min. Humidity 30 %, Average Annual Potential ET 1340 mm, Average Annual Wind Speed 4.9 m/s .

Due to its location away from any major water body and its proximity to the Himalayas, Roorkee has an extreme and erratic continental climate. Temperature begins to rise from March (29.1°C) and reaches to its maximum in June (44°C). The monsoon season starts in July and goes on until October, with torrential rainfall, due to the blocking of the monsoon clouds by the Himalayas. The post monsoon season starts in October and goes on until late November, with average temperatures sliding from 21 °C to 15 °C. Winters start in December, with lows close to freezing and frequent cold waves due to the cold katabatic winds blowing from the Himalayas and the temperature ranges between 10.5°C and 6.1°C. The potential evapotranspiration is maximum in the month of May 198.9 mm and minimum 38.5 mm in the month of December.

## 3.1 Data Availability :

The data obtained at NIH observatory was used for the prediction of Evaporation using ANN and MLR, hence daily rainfall, humidity and maximum, minimum and mean temperature data of Roorkee were collected from Jan 2001 to Dec 2013. The daily maximum temperature, minimum temperature, are presented in figure 3.1, 3.2 respectively. The daily humidity data is represented in figure 3.3 . The daily rainfall data is presented in figure 3.4

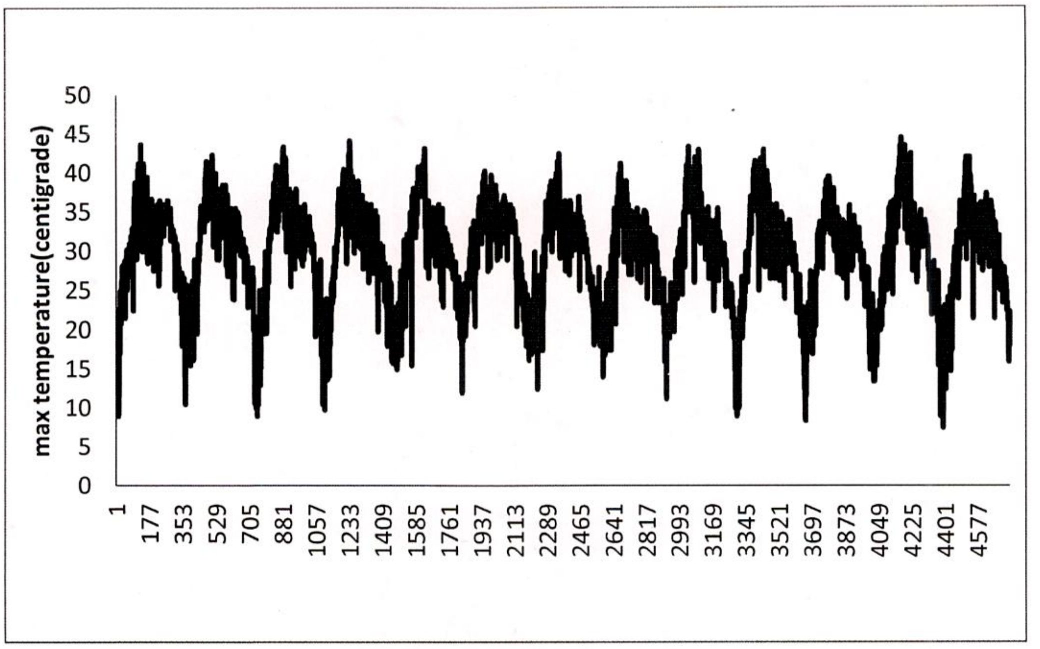


**Figure 3.1 Daily Maximum Temperatures at Roorkee from January 2001 to December 2013**

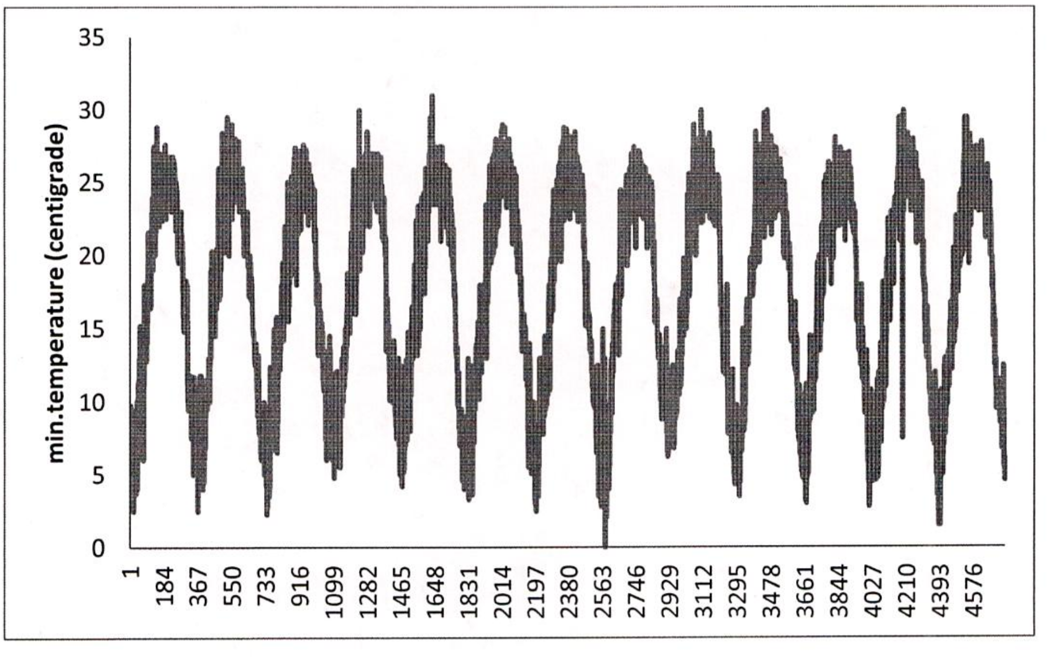


**Figure 3.2 Daily Minimum Temperatures at Roorkee from January 2001 to December 2013**

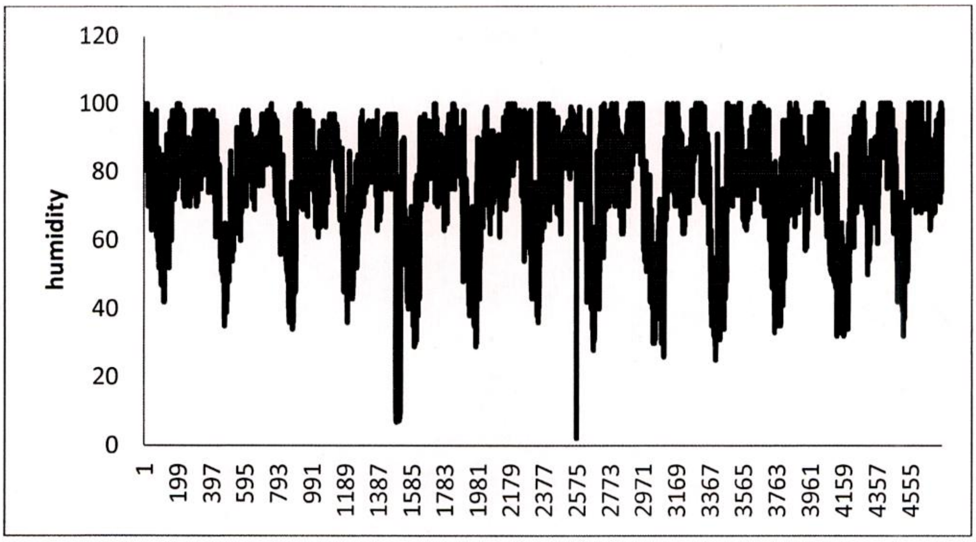


**Figure 3.3 Daily humidity at Roorkee from January 2001 to December 2013**

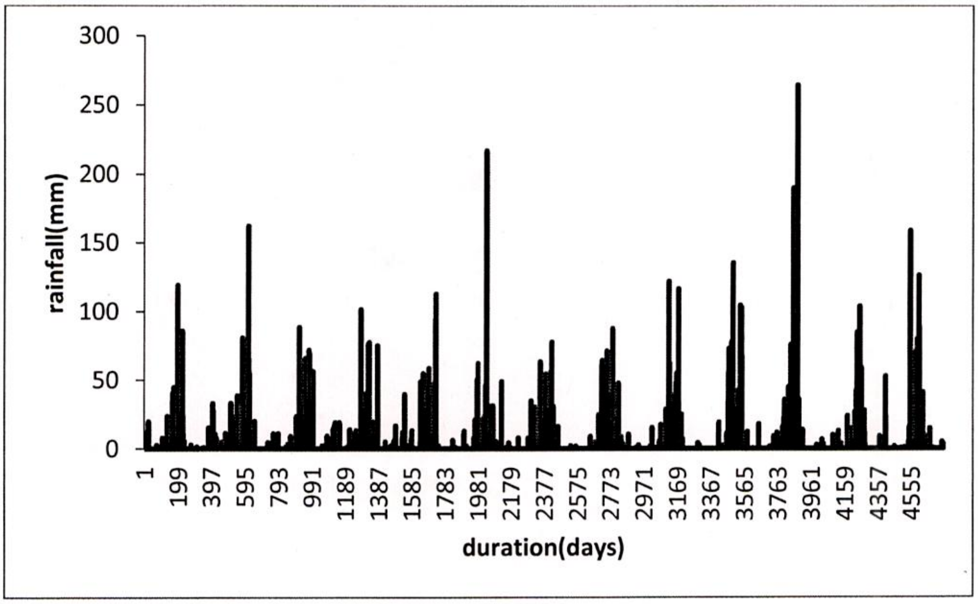


**Figure 3.4 Daily Rainfall data at Roorkee from January 2001 to December 2013**

The data for sensitivity analysis was obtained from NIH observatory ,Roorkee. Monthly humidity, radiation, sunshine hours and maximum, minimum and mean temperature data of Roorkee were collected from Jan 1987 to Dec 2013. The monthly humidity, maximum, minimum, mean temperature data are presented in figure 3.5 and data the is represented in table 3.1

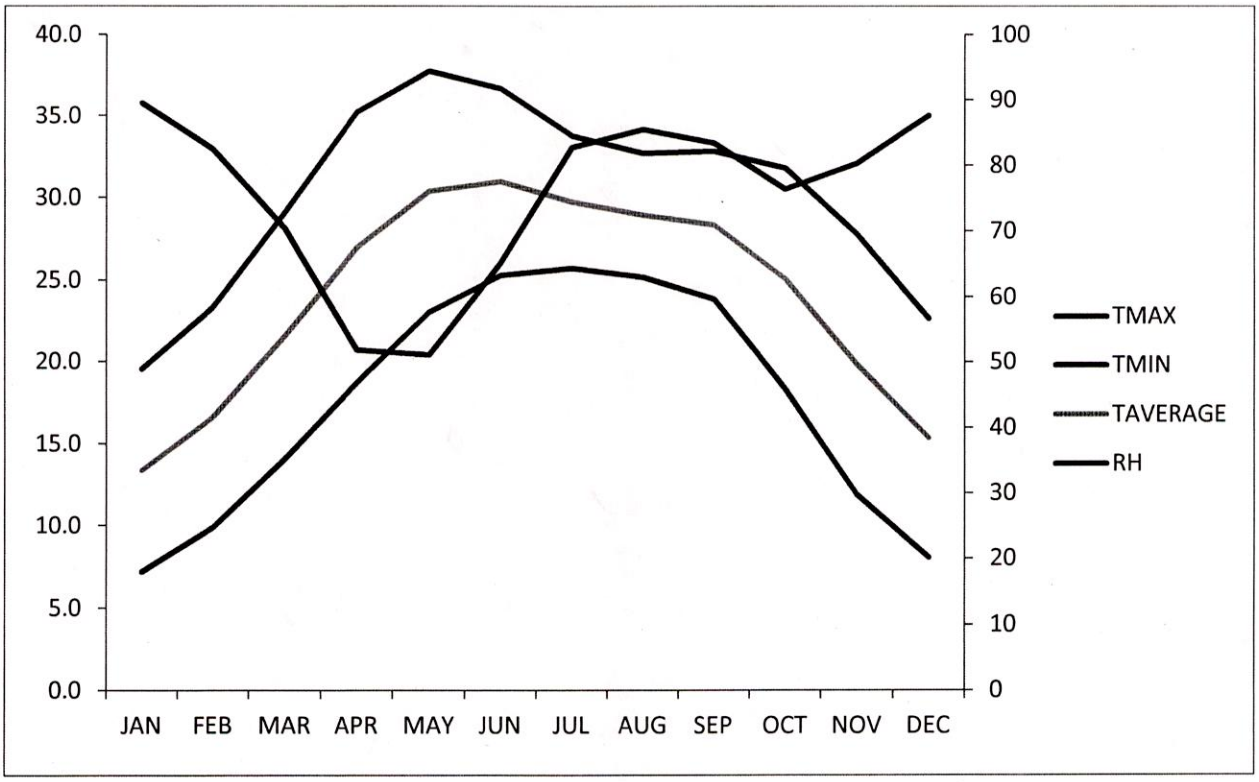


**FIGURE :3.5 Monthly maximum, minimum, average temperature and relative humidity from 1987-2013 for sensitivity analysis of evapotranspiration**

**Table 3.1 The average meteorological data for the period from Jan 1987 to Dec 2013**

| MONTH | Tmax | Tmin | Tavg | RH | sunshine | radiation | Latent heat |
|---|---|---|---|---|---|---|---|
| JAN | 19.5 | 7.2 | 13.4 | 89 | 2.9 | 8.5 | 2.46 |
| FEB | 23.3 | 9.9 | 16.6 | 82 | 6.0 | 10.5 | 2.45 |
| MAR | 29.1 | 14.1 | 21.6 | 70 | 7.0 | 12.7 | 2.44 |
| APR | 35.2 | 18.7 | 27.0 | 52 | 8.4 | 14.8 | 2.43 |
| MAY | 37.8 | 23.0 | 30.4 | 51 | 9.5 | 16.0 | 2.42 |
| JUN | 36.7 | 25.3 | 31.0 | 65 | 7.4 | 16.5 | 2.42 |
| JUL | 33.8 | 25.7 | 29.7 | 83 | 6.3 | 16.2 | 2.42 |
| AUG | 32.7 | 25.2 | 28.9 | 85 | 4.6 | 15.3 | 2.42 |
| SEPT | 32.8 | 23.9 | 28.4 | 83 | 6.7 | 13.5 | 2.43 |
| OCT | 31.8 | 18.3 | 25.1 | 76 | 7.4 | 11.3 | 2.43 |
| NOV | 27.8 | 11.9 | 19.8 | 80 | 6.6 | 9.1 | 2.45 |
| DEC | 22.7 | 8.1 | 15.4 | 88 | 4.3 | 7.9 | 2.46 |

# 4. MATERIALS AND METHODS

Evaporation depends upon many meteorological factors like temperature, relative humidity, rainfall, etc. Many conventional models such as empirical, regression based models and conceptual models have been developed to predict the evaporation with large quantity of data and produce less accurate results compared to recently developed soft computing techniques such as ANN, Fuzzy logic, decision tree algorithms (ASCE, 2000b). The newer techniques such as ANN and Fuzzy logic are applied by many researchers for modelling the evaporation using the available data from the study area. The following sections give an overview about ANN model, the methodology carried out.

## 4.1 Artificial neural network

An ANN is an information-processing construct that consists of a number of interconnected processing elements called nodes, analogous to neurons in the brain. Each node combines a number of inputs and produces an output, which is then transmitted to many different locations, including other nodes (Azhar*et al.,* 2007). A neural network is characterized by its architecture that represents the pattern of connection between nodes, its method of determining the connection weights and the activation function (Fausett, 1994). Figure 4.1 shows a typical single neuron with a sigmoid activation function, three input synapses and one output synapse. Synapses represent the structure where weight values are stored. In this study feed forward neural networks architecture has been used in predicting monthly evaporation.

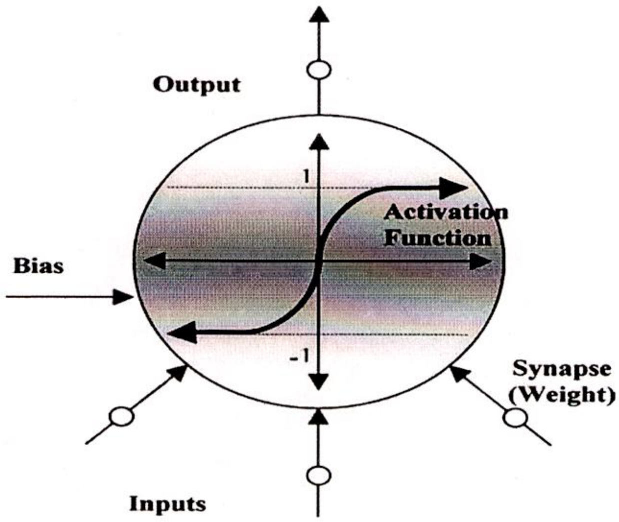


**Figure 4.1 Typical artificial neuron**

### 4.1.1 Model Development

The most important steps in the ANN model development process is the selection of significant input variables. Usually, not all of the potential input variables will be equally informative, because some may be correlated, noisy, or have no significant relationship with the output variable being modelled (Maier and Dandy, 2000). Input variables were selected based on cross-correlation, autocorrelation and partial autocorrelation technique. Many researchers have been successfully using correlation analysis for selection of input variables (Nayak*et al.,* 2006; Sasmita*et al.,* 2013).

The auto-correlation coefficient (Salas *et al.*, 1980) is defined as

$$r_k = \frac{\sum_{t=1}^{N-k}(x_t - \overline{x_t})(x_{t+k} - \overline{x_{t+k}})}{\left[\sum_{t=1}^{N-k}\left(x_t - \overline{x_t}\right)^2 \sum_{t=1}^{N-k}\left(x_{t+k} - \overline{x_{t+k}}\right)^2\right]^{1/2}} \tag{4.1}$$

Where $r_k$ is called the *lag-k* correlation coefficient, the serial correlation coefficient or the auto-correlation function (ACF), $x_t$ is the time series for *t = 1,..,N*, $x_{t+K}$ is the lagged time series for *t = 1,..., N-k*, $\overline{x_t}$ is the sample mean for *t = 1,...,N*, $\overline{x_{t+k}}$ is the sample mean *for t =*

*1,...,N-k, N* is the sample size. The partial auto-correlation coefficient (Salas *et al.*, 1980) may be obtained recursively by the Durbins relations as given below

$$\phi_1(1) = \rho_1, \phi_1(2) = \frac{\rho_1(1-\rho_2)}{(1-\rho_1^2)}, \phi_2(2) = \frac{\rho_2 - \rho_1^2}{(1-\rho_1^2)} \qquad (4.2)$$

$$\phi_k(k) = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_j(k-1)\rho_{k-j}}{1 - \sum_{j=1}^{k-1} \phi_j(k-1)\rho_j} \qquad (4.3)$$

$$\phi_j(k) = \phi_j(k-1) - \phi_k(k)\phi_{k-j}(k-1) \qquad (4.4)$$

### 4.1.2 Normalization of Input Data

The input values should be normalized to the range between 0 and 1 before passing into a neural network since the output of sigmoidal function is bound between 0 and 1. Dawson and Wilby (1998) and many others have emphasised the importance of the normalisation of data and have given the procedure to normalise. The output from the ANN should be denormalised to provide meaningful results. In this study, equation 4.5 is used to normalize the data set:

$$N_i = \frac{R_i - Min_i}{Max_i - Min_i} \qquad (4.5)$$

where$R_i$ is the real value applied to neuron *i*; $N_i$ is the subsequent normalized value calculated for neuron *i*; $Min_i$ is the minimum value of all values applied to neuron *i*; $Max_i$ is the maximum value of all values applied to neuron *i*.

### 4.1.3 Feed forward neural network (FNN)

Feed forward neural networks have been applied successfully in many different problems since the advent of the error back propagation learning algorithm. This network architecture and the corresponding learning algorithm can be viewed as a generalization of the popular least-mean-square (LMS) algorithm (Haykin, 1999).

In feed-forward networks, the data flow through the network in one direction from the input layer to the output layer through the hidden layer(s). Each output value is based solely on the current set of inputs. In most networks, the nodes of one layer are fully connected to the nodes in the next layer; however, this is not a requirement of feed-forward networks. A multilayer perceptron network consists of an input layer, one or more hidden layers of computation nodes, and an output layer.

- *The input layer* consists of nodes that receive an input from the external environment. These nodes do not perform any transformations upon the inputs but just send their weighted values to the nodes in the immediately adjacent, usually 'hidden,' layer.
- *The hidden layer(s)* consists of nodes that typically receive the transferred weighted inputs from the input layer or previous hidden layer, perform their transformations on it, and pass the output to the next adjacent layer, which can be another hidden layer or the output layer.
- *The output layer* consists of nodes that receive the hidden-layer output and send it to the user.

Feed forward means that all the interconnections between the layers propagate forward to the next layer. The type of node being used in the ANN determines the way that total input is calculated as well as the way that the node calculates its output as a function of its net input. In the present study, the activation function used for calculation is a sigmoid logistic function. Each node is a simple processing element that responds to the weighted inputs it receives from other nodes. The receiving node sums the weighted signals from all nodes to which it is connected in the preceding layer. The net input $x_j$ to node j is the weighted sum of all the incoming signals:

$$\text{Net\_input} = x_j = \sum w_{ij} y_i \qquad (4.6)$$

Where, $x_j$ = net input coming to node j

$w_{ij}$ = weight between node i and node j

$y_i$ = activation function at node i.

The activation function, $y_j$, which is a nonlinear function of its net-input, is described by the sigmoid logistic function

$$y_j = \frac{1}{1 + \exp(-x_j)} \tag{4.7}$$

Figure 4.2 shows a typical feed forward network with four input neurons, one hidden layer with three nodes and one output node. The input signal propagates through the network in a forward direction, layer by layer. Their main advantage is that they are easy to handle, and can approximate any input/output map, as established by Horniket *al.* (1989). The key disadvantages are that they train slowly, and require lots of training data.

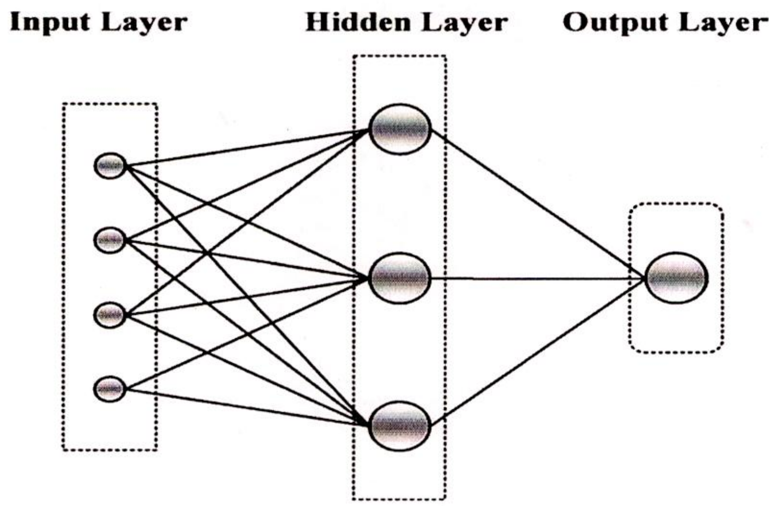


**Figure 4.2 typical feed forward neural network.**

### 4.1.4 Training with Algorithm

Determining the best values of all the weights is called training the ANN. In a so-called supervised learning mode, the actual output of a neural network is compared to the desired output. Weights, which are usually randomly set to begin with, are then adjusted so that the next iteration will produce a closer match between the desired and the actual output. Various learning methods for weight adjustments try to minimize the differences or errors between observed and computed output data.

The main objective of training (calibrating) a neural network is to produce an output vector $Y = (y_1, y_2, ...., y_p)$ that is as close as possible to the target vector (variable of interest or forecast variable) $T = (t_1, t_2, ...., t_p)$ when an input vector $X = (x_1, x_2, ...., x_p)$ is fed to the ANN. In this process, weight matrices *W* and bias vectors *V* are determined by minimizing a predetermined error function as explained as follows:

$$E = \sum_P \sum_p (y_i - t_i)^2 \tag{4.8}$$

Where $t_i$ is a component of the desired output *T*; $y_i$ is the corresponding ANN output; *p* is the number of output nodes; and *P* is the number of training patterns.

The training phase can consume a lot of time. It is considered complete when the artificial neural network reaches a user-defined performance level. At this level the network has achieved the desired statistical accuracy as it produces the required outputs for a given sequence of inputs. When no further learning is judged necessary, the resulting weights are typically fixed for the application.

Bayesian regularization algorithm is used in this study in order to train the given network more efficiently.

### 4.1.5 Bayesian Regularization Algorithm (BR)

The Bayesian regularization is an algorithm that automatically sets optimum values for the parameters of the objective function. In the approach used, the weights and biases of the network are assumed to be random variables with specified distributions. In order to estimate regularization parameters, which are related to the unknown variances, statistical techniques are being used. The advantage of this algorithm is that whatever the size of the network, the function won't be over-fitted. Bayesian regularization has been effectively used in literature (Anctil*et al.*, 2004; Coulibaly*et al.*, 2001a, Porter *et al.*, 2000).

### 4.1.6 Network Architecture

The network geometry is generally highly problem oriented in order to get optimal network geometry trial and error procedure is adopted. The numbers of nodes in the input layer were decided based on the inputs to the model. The number of hidden neurons in the network, which is responsible for capturing the dynamic and complex relationship between various input and output variables, is identified by various trials. For each set of hidden neurons, the network is trained with input datasets in batch mode to minimize the mean square error at the output layer.

Various internal parameters used in the ANN model like learning rate, momentum coefficient, scalar μ, and combination of transfer functions for hidden and output layer were also found out by trial and error. MATLAB 2010a software was used for analysis.

### 4.1.7 Performance Evaluation of the Developed ANN Model and MLR model.

The visual observations and quantitative evaluation of the developed model was performed to judge the goodness of fit between observed and predicted values. The whole data length is divided into two sets based on statistical properties of the time series such as mean and standard deviation, in that one is used for calibration (training) and another for validation of ANN model and the same data set with de-noised values is used as an input for MLR model. The performance during calibration and validation is evaluated by using statistical parameters. They are Coefficient of Efficiency (CE), Root Mean Square Error (RMSE) and Explained Variance (EV) given by following equations:

#### 4.1.7.1 Visual observations evaluation

The visual observation based on the graphical comparison between the observed and the estimated values is one of the simplest methods for the performance assessment of a model. The performance of the model was evaluated by comparing ordinates of observed and estimated evaporation graphs. For assessing the suitability of the model, the estimated evaporation was compared visually with the observed evaporation.

#### 4.1.7.2 Quantitative evaluation:

#### 4.1.7.2(i) Coefficient of efficiency (CE)

Based on the standardization of residual variance with initial variance, the coefficient of efficiency can be used to compare the relative performance of the two approaches effectively. It is expressed as:

$$\text{CE}=\left\{1-\frac{\textit{residual}\ \text{var}\,\textit{iance}}{\textit{initial}\ \text{var}\,\textit{iance}}\right\} =\left\{1-\frac{\sum_{j=1}^{n}(Y_j-X_j)^2}{\sum_{j=1}^{n}(Y_j-\bar{Y})^2}\right\} \tag{4.9}$$

Chiew*et al.* (1993) classified 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).

#### 4.1.7.2(ii) Root mean square error (RMSE)

RMSE indicates the discrepancy between the observed and calculated values. The lowest the RMSE, the more accurate the prediction is. It is expressed as:

$$\text{RMSE} = \sqrt{\frac{\textit{residual}\ \text{var}\,\textit{iance}}{n}} = \sqrt{\sum_{j=1}^{n}\frac{(Y_j-X_j)^2}{n}} \tag{4.10}$$

#### 4.1.7.2(iii) Explained Variance (EV)

Explained variation measures the proportion to which a mathematical model accounts for the variation (dispersion) of a given data set. It is given by

$$EV = \sqrt{\left(\frac{\Sigma(X_j-\bar{Y}_j)^2}{\Sigma(Y_j-\bar{X}_j)^2}\right)} \tag{4.11}$$

Where,

$Y_j$ = Observed water table depth, $X_j$ = Predicted water table depth, n = Number of observations, $\overline{Y}$ j= Mean of observed water table depth, $\overline{X}$ j= Mean of predicted water table depth.

In this study an attempt has been made to predict the evaporation of data obtained from NIH observatory. An ANN model was developed with the historical data of rainfall, maximum, minimum and humidity of Roorkee. The best ANN model was trained with the input data derived from statistical analysis. The MLR model was developed with the same input vector used in the ANN model. The performance of the both the models is evaluated by using statistical parameters.

## 4.2 Multiple Linear Regression (MLR) Based Evaporation Estimation Model

Multiple linear regressions were used to estimate the pan evaporation for the study area. Multiple linear regressions (MLR) is a multivariate statistical technique used to model the linear correlations between a single dependent variable *Y* and two or more independent variables. The regression equation of *Y* can be written as:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots\ldots\ldots + \beta_k X_n \qquad (4.12)$$

where, *Y* is the response variable; $X_1$, $X_2$.....$X_n$ are the independent variables; and $\beta_0, \beta_1, \beta_2 \ldots \beta_k$ are the regression coefficients.

### 4.2.1 Performance Evaluation of the Developed MLR Model

In this study, the effects of meteorological parameters such as maximum and minimum air temperature, relative humidity and rainfall on pan evaporation was analyzed. The prediction equation developed through multiple linear regression analysis computes the evaporation with the decided input parameters. The study reveals the above meteorological parameters can well be correlated with pan evaporation. The rainfall, maximum and minimum air temperature, relative humidity were taken as the independent variables and evaporation was taken as dependent variable. The multiple linear equation was:

**E = -0.01*rain+0.19930*maxt+0.01*mint-0.04*hum+0.31** (4.13 )

#### 4.2.1.1 Visual observation evaluation

The visual observations and quantitative evaluation of the developed model was performed to judge the goodness of fit between observed and predicted values. The visual observation based on the graphical comparison between the observed and the estimated values is one of the simplest methods for the performance assessment of a model.

#### 4.2.1.2 Quantitative evaluation:

Quantitative evaluation is based on the performance during calibration and validation which is evaluated by using statistical parameters. They are Coefficient of Efficiency (CE), Root Mean Squared Error (RMSE) and Explained Variance (EV)

## 4.3 Reference evapotranspiration

The evapotranspiration rate from a reference surface, not short of water, is called reference crop evapotranspiration or reference evapotranspiration (ETo) reference surface is a hypothetical grass reference crop with specific characteristics. ETo expresses the evaporating power of the atmosphere at a specific location and time of the year and does not consider the crop and soil characteristics (Allen *et al.*, 1998, A.S.C.E., 2005).

This index was been introduced to study the evaporative demand of the atmosphere independently of crop type, crop development and management practices, hence it is only affected by meteorological properties (i.e. temperature, Rel. humidity, wind speed, Solar Radiation). Numerous empirical methods have been developed over the last 50 years to estimate evapotranspiration using different climatic variables. However, relationships were often subject to rigorous local calibrations and proved to have limited global validity (Allen *et al.*, 1998).

### 4.3.1 Temperature Based Methods:

#### 4.3.1.1 Hargreaves Method:

The Hargreaves method enables reference crop evapotranspiration (ET0) estimation in areas where meteorological information is scarce. This is an empirical estimation method that uses the average daily air temperature, T (°C), in combination with the extraterrestrial radiation, $R_a$

(MJ/m$^2$/day) as an indicator of the incoming global radiation. The Hargreaves equation is expressed as

$$ET_0 = 0.0023 \; R_a \left[\frac{T_{\max} + T_{\min}}{2} + 17.8\right]\sqrt{T_{\max} - T_{\min}} \quad (4.14)$$

Where, $T_{\max}$ and $T_{\min}$ are average maximum and minimum temperatures

#### 4.3.1.2 Thornthwaite Method:

Thornthwaite correlated mean monthly temperature with ET as determined by east-central United States water balance studies. The Thornthwaite equation is:

$$ET_{0K} = \frac{16 N_K}{360}\left[\left(\frac{10 K_T}{\sum_{K=1}^{12} 0.2 T_K^{1.514}}\right)^{0.016 \sum_{K=1}^{12} 0.2 T_K^{1.514} + 0.5}\right] \quad (4.15)$$

Where, $ET_{0K}$ is potential evapotranspiration in the $K_{th}$ month (mm); $N_K$ is the maximum possible duration of sunshine in the $K_{th}$ month (hours); $T_K$ is the mean air temperature in the $K_{th}$ month (°C) and $k$ = 1, 2, ..... , 12.

### 4.3.2 Radiation Based Methods

#### 4.3.2.1 Turc Method:

Turc developed an equation for potential ET under general climatic conditions of Western Europe. He proposed the following equations for two humidity conditions:

When $RH_{mean}$> 50%,

$$ET_0 = 0.013 \; \frac{T_{mean}}{T_{mean} + 15}(R_S' + 50)\frac{1}{\lambda} \quad (4.16)$$

When $RH_{mean} \leq 50\%$,

$$ET_0 = 0.013 \frac{T_{mean}}{T_{mean} + 15}(R_S' + 50)\frac{1}{\lambda}\left(1 + \left(\frac{50 - RH_{mean}}{70}\right)\right) \quad (4.17)$$

Where, $T_{mean}$ is mean air temperature (ºC), $RH_{mean}$ is mean relative humidity (%), $R_s'$ is solar radiation (cal/cm$^2$/day). If $R_s$ (MJ/m$^2$/day) is known, it can be calculated as

$$R_S' = \frac{R_S}{0.041869} \quad (4.18)$$

$\lambda$ is the latent heat of vaporization (MJ/kg). it can be estimated using mean air temperature as

$$\lambda = 2.501 - 0.002361\, T_{mean} \quad (4.19)$$

### 4.4 Sensitivity analysis:

By definition, sensitivity analysis investigates the effect of change of one factor on another (McCuen,1973). The change of reference evapotranspiration to the change of a meteorological variable, when it tends to zero, is the partial derivative of reference evapotranspiration to this variable. A number of sensitivity coefficients can be defined based on dimensionless values of the reference evapotranspiration change for different purposes of sensitivity analysis (McCuen, 1974, Saxton, 1975, Beven, 1979, Gong *et al.*, 2006). The dimensionless values of sensitivity coefficients for different meteorological parameters allow the comparison between them. Saxton (1975) defined dimensionless sensitivity coefficients for each meteorological variable based on:

$$K_{s_p} = \frac{\delta M_P}{\delta pM} \quad (4.20)$$

where p is the examined independent variable or parameter and M is the modelled value. This coefficient shows the percentage of change in evapotranspiration caused by the percentage change of a meteorological variable. The calculation of the partial derivative of reference evapotranspiration to a variable depends on all the meteorological variables and its value depends on them. However, Equation (4.20) is sensitive to the magnitudes of reference evapotranspiration and p.

In particular, the relative sensitivity coefficient $K_{S_p}$ may not be a good indication of the significance of the variable if either: 1) the value of reference evapotranspiration or the value of the parameter tends to zero independently, or 2) the range of values taken by p is small in relation to its magnitude (Beven, 1979).

Coleman and DeCoursey (1976) provided a more meaningful coefficient when comparing variables some of which may have a range in variability quite different from their mean value; hence the bias caused by the method of measurement is eliminated. The coefficient is given by:

$$K_{S_p} = \frac{\delta Mp - p_{mean}}{\delta pM} \qquad (4.21)$$

where $p_{mean}$ is the minimum observed value of the independent variable. Babajimopoulos et al. (1992) estimated the influence of the meteorological variables to evapotranspiration changing by 10, 20 and 30% the meteorological variables and assessing its impact on the calculated evapotranspiration. However, in this case the variation of a parameter could significantly influence the sensitivity of the parameters to the model. More recently, Ampas (2010) proposed the use of standard deviation and presented a new sensitivity coefficient:

$$K_{S_p} = \frac{\delta M \sigma_p}{\delta pM} \qquad (4.22)$$

where $\sigma_p$ is the standard deviation of the meteorological variable.

# 5. RESULTS AND DISCUSSION

The results of the study with the objective of development of multiple linear regression (MLR) and artificial neural network (ANN) models for forecasting evaporation and sensitivity analysis for evapotranspiration of the study area and the analysis of the results are presented in this chapter.

## 5.1 Development of ANN model

The feed forward ANN model architecture has been considered to simulate the evaporation in Roorkee, using daily rainfall, daily relative humidity and daily maximum and minimum temperature data as the input to the model. ANN models with different hidden neuron structure have been developed and the best ANN model has been selected based on the performance evaluation criteria.

### 5.1.1 Selection of Input Vector

The input vector is selected generally by trial and error. The simple correlation between the dependent and independent variables helps in selecting the significant input vector to the model. To identify the input vector, detailed correlation analysis of the following variables is done.

i. Daily rainfall values with daily evaporation values

ii. Daily maximum and minimum temperature values with daily evaporation values.

iii. Daily relative humidity values with evaporation.

The correlation graph was plotted between the inputs used in analysis and evaporation. From graph, it clearly indicates that the rainfall values are not much correlated with Evaporation values. The correlation analysis helps to find out the possible input variable for the modeling, but it does not give the exact lag values.

Generally, the significant lags of input variable are found out by trial and error. But, Sudheer et al. (2002) have suggested a statistical procedure that avoids the trial and error procedure. They have reported that the statistical parameters such as auto-correlation function (ACF), partial auto-correlation function (PACF) and cross-correlation function (CCF) could be used to find out the significant lag values of input variables.

The CCF between evaporation and rainfall , maximum temperature ,minimum temperature ,relative humidity are presented in **Figure 5.1 , 5.2 , 5.3, 5.4** respectively.

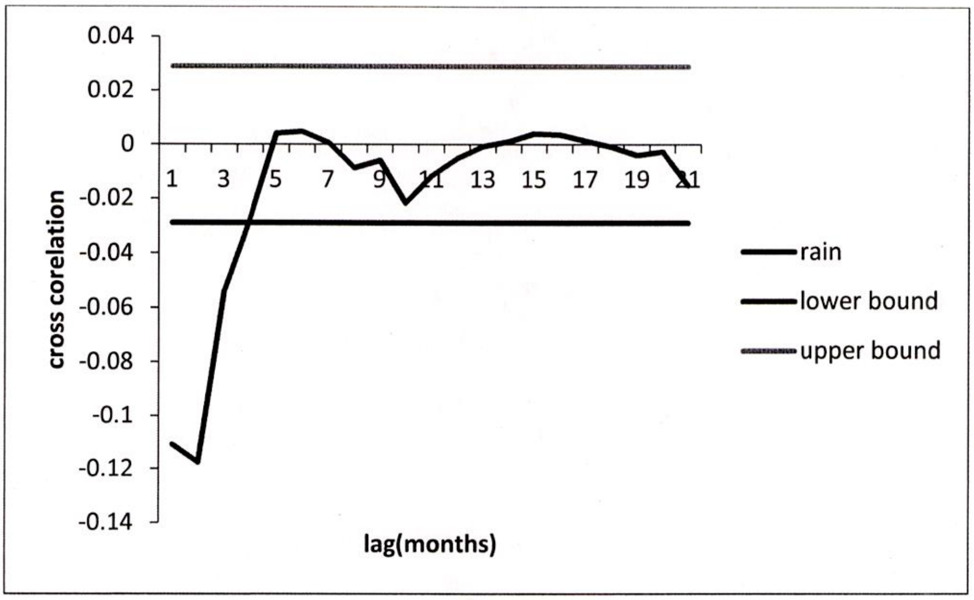


**Figure: 5.1 Cross correlation between rainfall and evaporation**

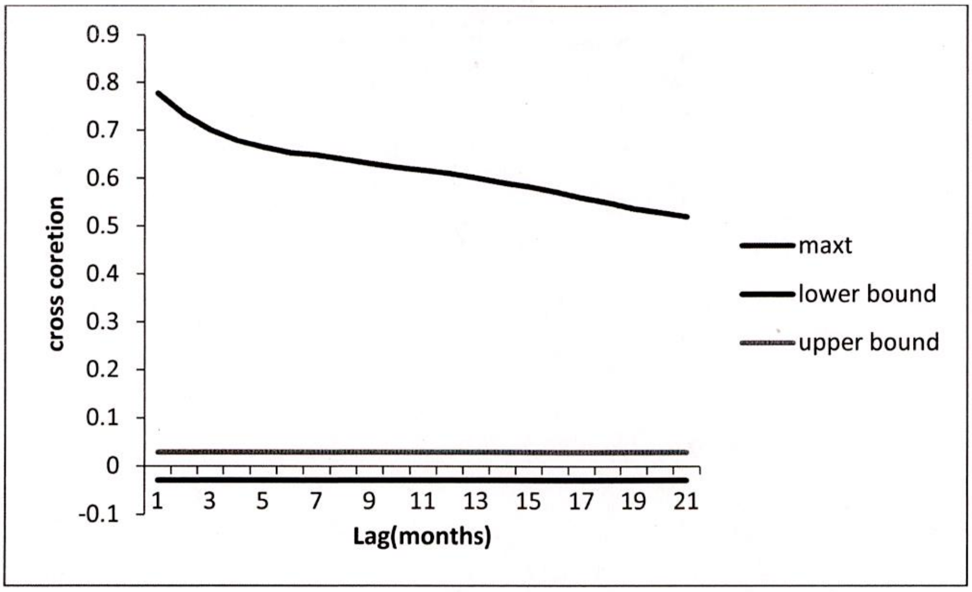


**Figure: 5.2 Cross correlation between maximum temperature and evaporation**

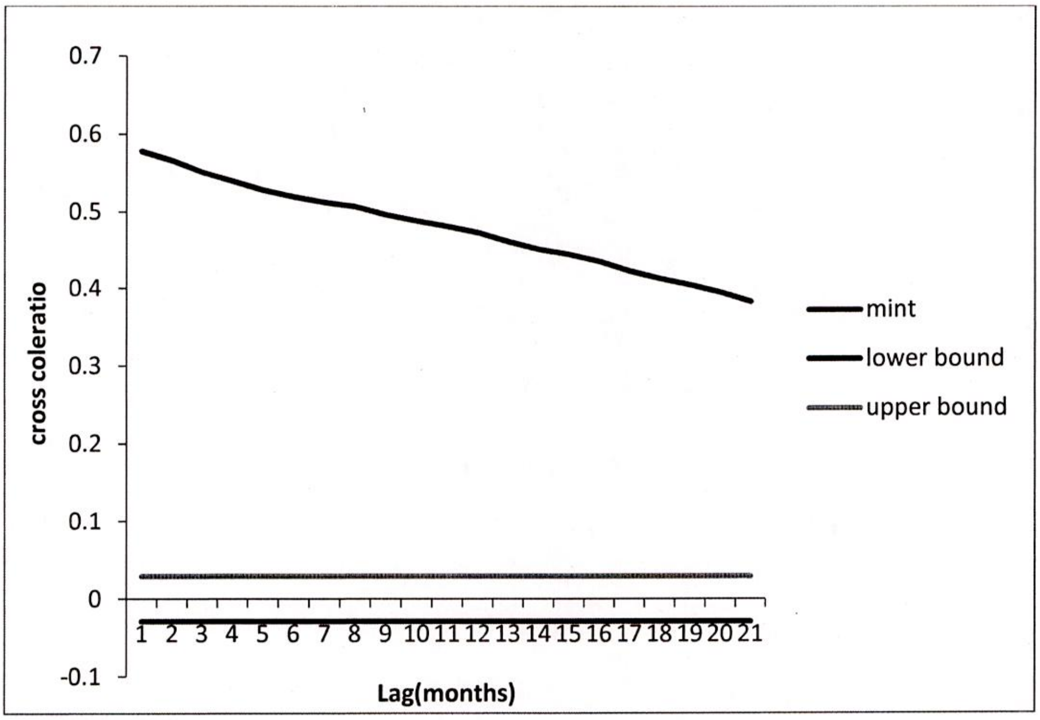


**Figure 5.3 Cross correlation between minimum temperature and evaporation**

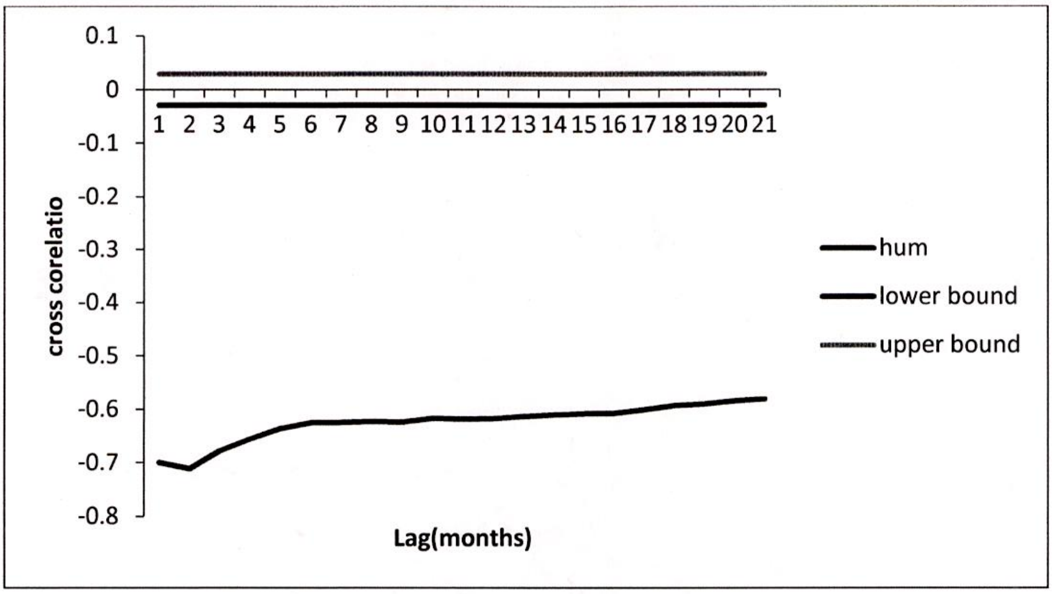


**Figure 5.4 Cross correlation between relative humidity and evaporation**

Cross correlation analysis of daily pan evaporation at (t) with daily rainfall at (t-1) and daily relative humidity at (t-1) has given significant correlation compared to other lags of daily maximum and minimum temperature(t). Hence the inputs for the model are daily rainfall (rain) at (t-1) lag, daily relative humidity (hum) at (t-1) lag and daily maximum temperature (maxt) and daily minimum temperature (mint) based on the partial autocorrelation analysis of daily evaporation and the cross correlation analysis of rainfall and relative humidity .There is one node in the output layer, which predicts evaporation evap (t) for one day at (t).

$$\text{Evap (t)} = f\{\text{Rain (t-1), max(t),min(t),hum(t-1)}\}$$

### 5.1.2 Training of ANN model

The ANN models have been trained using Bayesian regularization algorithm. The whole data set has been divided into two sets for the training and validation of the ANN model. The data from January 2001 to December 2013 (4787 daily datasets) have been considered for the training of the model. Out of 4787 dataset,1 to 3286 datasets (62.50%) of data were used for calibration (training), 3287-4787 datasets (37.50%) of data were used for validation. These data sets were selected by trial and error method.

The number of the neurons in the hidden layer is found by a trial and error based, the trial and error procedure started with one hidden neuron initially, and it has been increased up to 10 based on the performance criteria of the model.

The transfer functions of hidden and output layers have been considered as log sigmoid and pure linear respectively in the training of the ANN model. The performance of the ANN model during calibration and validation with the input combination derived from statistical procedure given by Sudheer et al. (2002) is given in Table 5.1.

The model EVAP4 with ANN structure 4-4-1 is best among all the structure, because the performance of the structure in terms of all the statistical parameters is best among all ANN structure trained as given in the table. Even ANN structure 4-9-1 has given better results, but the difference between the results of these two structures are negligible and also after crossing number of neurons 10 the performance of the model is fluctuating (decreasing and then it is increasing) and it might have led to the over fitting and a large ANN structure.

**Table: 5.1 Results of ANN model during calibration and validation**

| Model No | Input Combination | ANN Structure | Calibration | | | Validation | | | File name |
|---|---|---|---|---|---|---|---|---|---|
| | | | CORR | EFF% | RMSE | CORR | EFF% | RMSE | |
| EVAP1 | [R (t-1), maxt (t), mint (t),hum(t-1)] | 4-1-1 | 0.8753 | 0.7662 | 1.0462 | 0.8852 | 0.7743 | 1.0245 | CAL1 VAL1 |
| EVAP2 | " | 4-2-1 | 0.8790 | 0.7730 | 1.0310 | 0.8883 | 0.7822 | 1.0064 | CAL2 VAL2 |
| EVAP3 | " | 4-3-1 | 0.8809 | 0.7762 | 1.0236 | 0.8890 | 0.7814 | 1.0082 | CAL3 VAL3 |
| EVAP4 | " | 4-4-1 | 0.8828 | 0.7800 | 1.0151 | 0.8897 | 0.7830 | 1.0045 | CAL4 VAL4 |
| EVAP5 | " | 4-5-1 | 0.8826 | 0.7795 | 1.0161 | 0.8899 | 0.7839 | 1.0023 | CAL5 VAL5 |
| EVAP6 | " | 4-6-1 | 0.8833 | 0.7811 | 1.0125 | 0.8897 | 0.7825 | 1.0057 | CAL6 VAL6 |
| EVAP7 | " | 4-7-1 | 0.8846 | 0.7835 | 1.0070 | 0.8896 | 0.7827 | 1.0051 | CAL7 VAL7 |
| EVAP8 | " | 4-8-1 | 0.8838 | 0.7820 | 1.0102 | 0.8903 | 0.7842 | 1.0017 | CAL8 VAL8 |
| EVAP9 | " | 4-9-1 | 0.8858 | 0.7857 | 1.0016 | 0.8894 | 0.7827 | 1.0052 | CAL9 VAL9 |
| EVAP10 | " | 4-10-1 | 0.8829 | 0.7807 | 1.0134 | 0.8829 | 0.7860 | 0.9976 | CAL10 VAL10 |

## 5.2 Analysis of Results of ANN and MLR Models

The performance of best ANN and best MLR models for the prediction of Evaporation at Roorkee during calibration and validation is presented in Fig 5.5 Fig5.6, Fig5.7, Fig5.8 .The graph plots clearly demonstrate the potentiality of the developed ANN and MLR models in the prediction of Evaporation. The performance of best ANN and best MLR model in terms of observed and computed daily evaporation during calibration and validation are presented in Figure 5.9 and 5.10 respectively. The results of the calibration and validation of the best ANN and best MLR models in terms of various statistical indices are presented in the Table 5.2

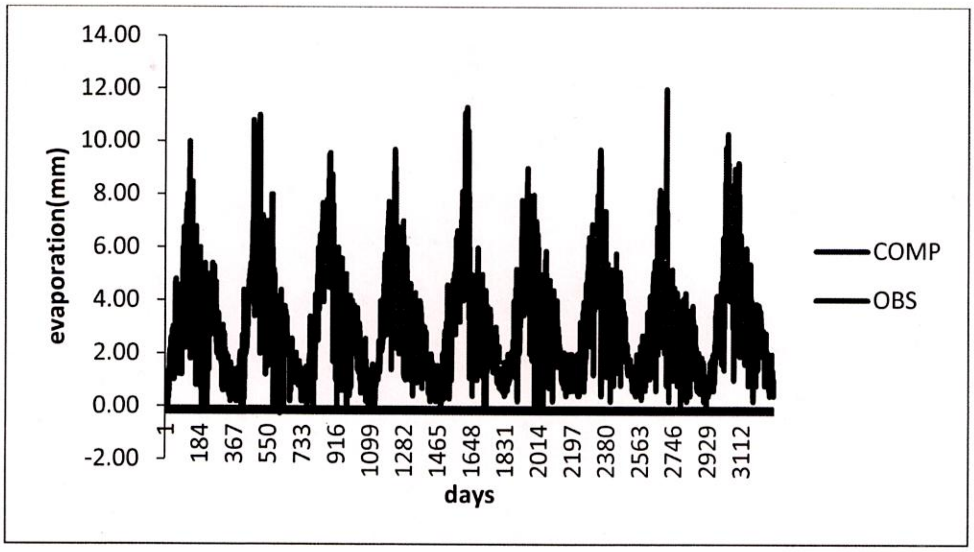


**Figure 5.5 Graph plot for the result of best ANN model during calibration (4-4-1)**

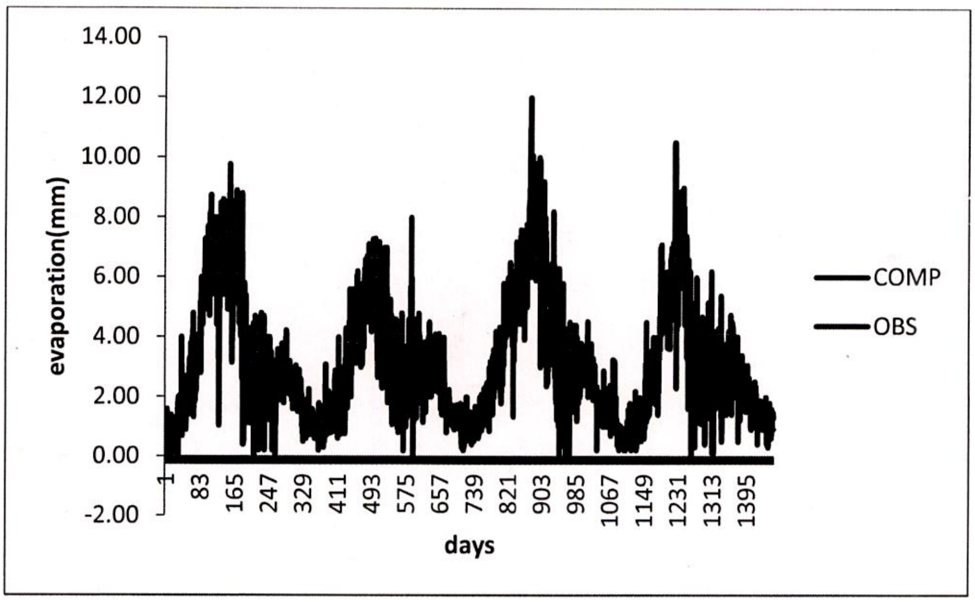


**Figure 5.6 Graph plot for the result of best ANN model during validation (4-4-1)**

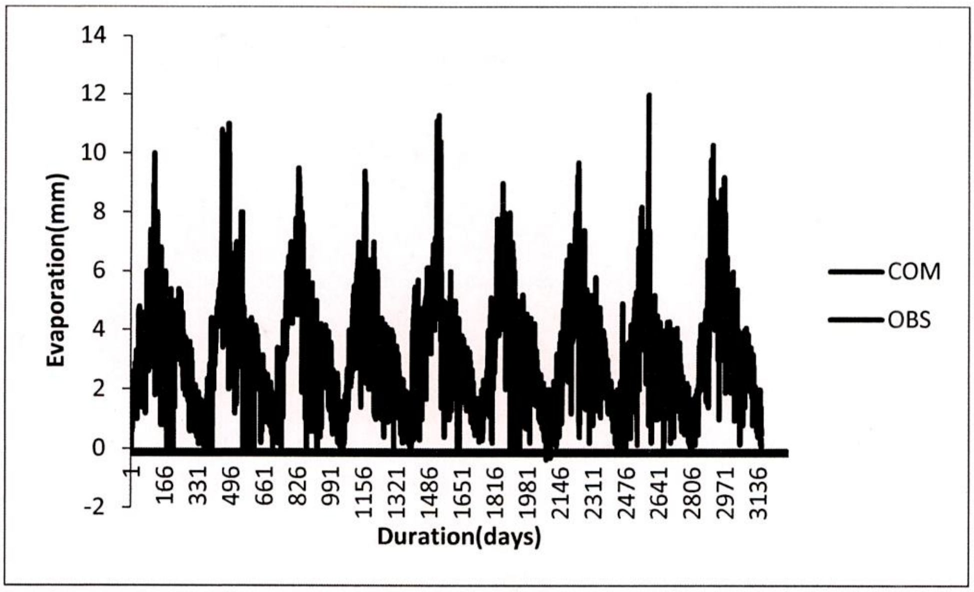


**Figure 5.7 Graph plot for the result of best MLR model during calibration**

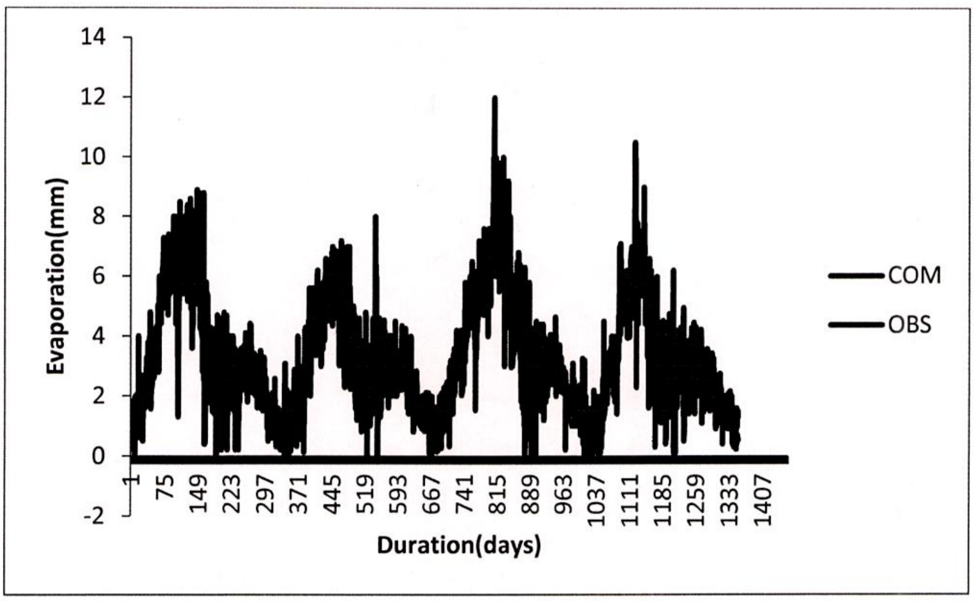


**Figure 5.8 Graph plot for the result of best MLR model during validation**

**Table 5.2 Comparison of results between best ANN and MLR models**

| Model | Calibration | | | Validation | | |
|---|---|---|---|---|---|---|
| | CORR | RMSE | EFF% | CORR | RMSE | EFF% |
| **ANNEVAP4 (4-4-1)** | 0.8828 | 1.0151 | 0.7800 | 0.8897 | 1.0045 | 0.7830 |
| **MLR EVAP** | 0.8359 | 1.1874 | 0.6988 | 0.8563 | 1.1523 | 0.8501 |

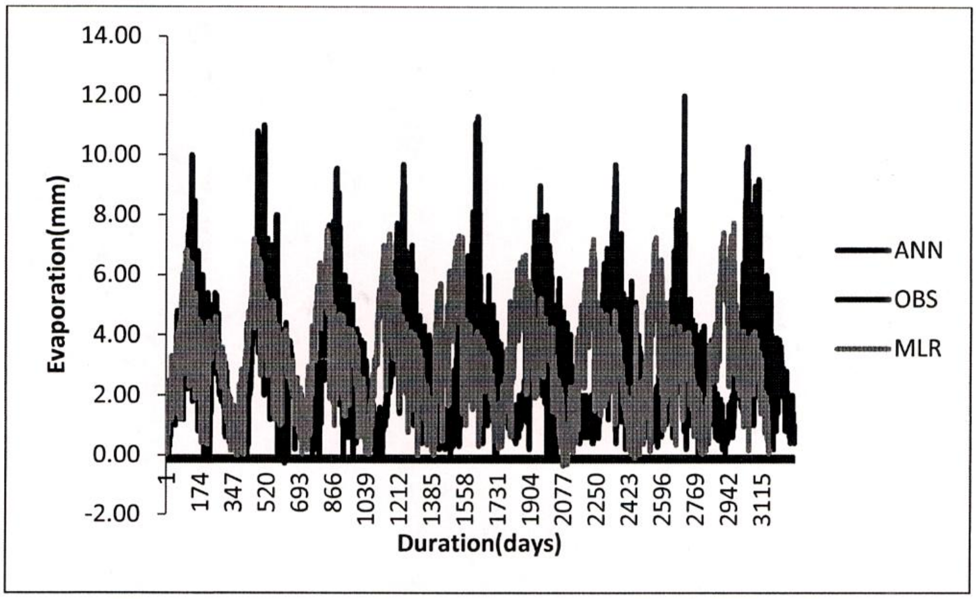


**Figure 5.9 Observed and Computed daily evaporation during calibration(Comparison among ANN, MLR, Observed evaporation)**

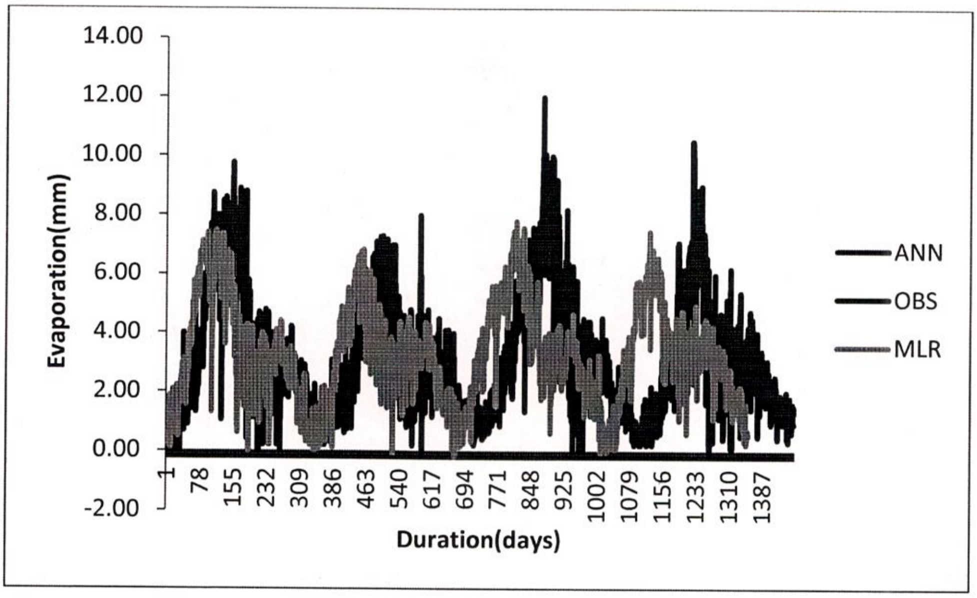


**Figure 5.10 Observed and Computed daily evaporation during validation(Comparison among ANN, MLR, Observed evaporation)**

The coefficients of correlation of the ANN model are higher than that of MLR during calibration and validation. The RMSE values of MLR model, which is residual variance, during calibration and validation, are higher than the values of ANN model. The model efficiency of MLR is deteriorated during the calibration and validation. The model efficiency of ANN is good for both calibration and validation and also there is no much difference between them. The scatter plots of ANN and MLR models during calibration and validation clearly indicates that the higher values of MLR model are under estimated than the ANN models. So, the analysis of the performance of the both the models clearly indicate that the ANN is better than the MLR in predicting the Evaporation

## 5.3 Sensitivity analysis by Thornthwaite method:

Here ET depends upon two independent variables i,e $T_{average}$ and sunshine hours. So by increasing the values of the one of the independent variables to 5%, 10%, 15%, 20% ,25% keeping other constant, the increased values of ET are obtained which are given in below :

**Table: 5.3 Increase in ET with increase in temperature (Thornthwaite method)**

| | Col1 | Col2 | Col3 | Col4 | Col5 | Col6 | Col7 | Col8 | Col9 | Col10 | Col11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | ET (temp 5%) | | ET (temp 10%) | | ET (temp1 5%) | | ET (temp2 0%) | | ET (temp 25%) | | ETO (original ) |
| JAN | 0.125 | -0.065 | 0.116 | -0.130 | 0.108 | -0.194 | 0.099 | -0.258 | 0.091 | -0.319 | 0.134 |
| FEB | 0.472 | -0.032 | 0.456 | -0.066 | 0.438 | -0.102 | 0.419 | -0.141 | 0.399 | -0.181 | 0.488 |
| MAR | 1.140 | 0.010 | 1.149 | 0.018 | 1.154 | 0.023 | 1.155 | 0.024 | 1.154 | 0.022 | 1.128 |
| APR | 2.530 | 0.047 | 2.644 | 0.095 | 2.758 | 0.142 | 2.870 | 0.188 | 2.980 | 0.234 | 2.415 |
| MAY | 3.989 | 0.068 | 4.253 | 0.139 | 4.527 | 0.212 | 4.810 | 0.288 | 5.102 | 0.366 | 3.735 |
| JUN | 3.276 | 0.071 | 3.504 | 0.146 | 3.742 | 0.223 | 3.990 | 0.304 | 4.246 | 0.388 | 3.059 |
| JUL | 2.491 | 0.064 | 2.646 | 0.130 | 2.806 | 0.199 | 2.971 | 0.269 | 3.139 | 0.341 | 2.341 |
| AUG | 1.687 | 0.059 | 1.784 | 0.120 | 1.883 | 0.183 | 1.984 | 0.246 | 2.087 | 0.311 | 1.592 |
| SEPT | 2.320 | 0.056 | 2.445 | 0.113 | 2.572 | 0.171 | 2.701 | 0.229 | 2.830 | 0.288 | 2.197 |
| OCT | 1.825 | 0.035 | 1.885 | 0.069 | 1.943 | 0.102 | 1.997 | 0.132 | 2.047 | 0.161 | 1.763 |
| NOV | 0.850 | -0.003 | 0.845 | -0.010 | 0.837 | -0.020 | 0.826 | -0.032 | 0.812 | -0.048 | 0.853 |
| DEC | 0.273 | -0.044 | 0.260 | -0.089 | 0.246 | -0.137 | 0.232 | -0.185 | 0.218 | -0.234 | 0.285 |
| AVG % INCREASE IN TEMP | | 0.022 | | 0.045 | | 0.067 | | 0.089 | | 0.111 | |

The above table shows the increase in ET with increase in temperature to 5%, 10%, 15%, 20%, 25% and keeping other independent variables i,e sunshine in Thornthwaite equation as constant

**Table: 5.4 Average percent increase in ET with increase in temperature (Thornthwaite method**

| PERCENT INCREASE IN TEMPERATURE | PERCENT INCREASE IN ET WITH INCREASE IN TEMPERATURE |
|---|---|
| 0 | 0 |
| 5% | 0.022 |
| 10% | 0.045 |
| 15% | 0.067 |
| 20% | 0.089 |
| 25% | 0.111 |

The above table shows the average percent increase in ET with in temperature to 5%, 10%, 15%, 20%, 25%, keeping other independent variables i,e sunshine in Thornthwaite equation as constant

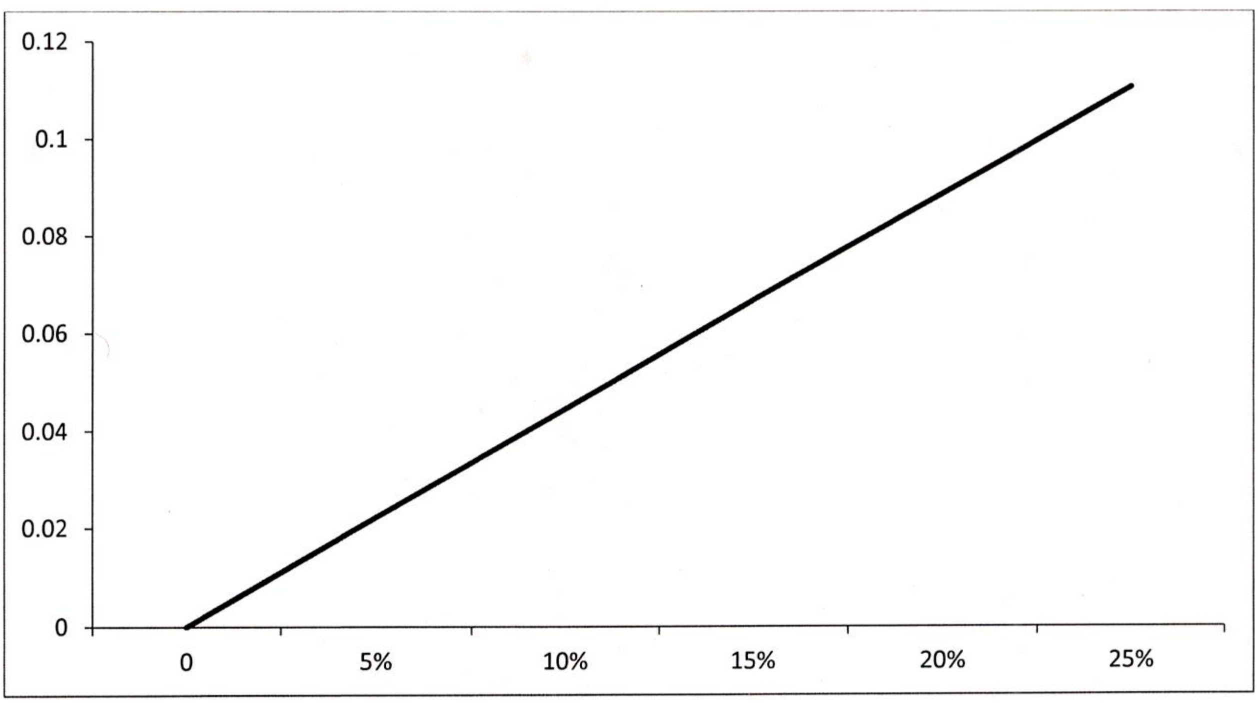


**Figure: 5.11 Average percent increase in ET with increase in temperature (Thornthwaite method)**

**Table: 5.5 percent increase in ET with increase in sunshine (Thornthwaite method)**

| | Col1 | Col2 | Col3 | Col4 | Col5 | Col6 | Col7 | Col8 | Col9 | Col10 | Col11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | ET (sunshine5%) | | ET (sunshine 10%) | | ET (sunshine15%) | | ET (sunshine20%) | | ET (sunshine25%) | | ET (original) |
| JAN | 0.140 | 0.050 | 0.147 | 0.100 | 0.154 | 0.150 | 0.160 | 0.200 | 0.167 | 0.250 | 0.134 |
| FEB | 0.512 | 0.050 | 0.537 | 0.100 | 0.561 | 0.150 | 0.585 | 0.200 | 0.610 | 0.250 | 0.488 |
| MAR | 1.185 | 0.050 | 1.241 | 0.100 | 1.298 | 0.150 | 1.354 | 0.200 | 1.410 | 0.250 | 1.128 |
| APR | 2.536 | 0.050 | 2.657 | 0.100 | 2.778 | 0.150 | 2.898 | 0.200 | 3.019 | 0.250 | 2.415 |
| MAY | 3.922 | 0.050 | 4.109 | 0.100 | 4.296 | 0.150 | 4.482 | 0.200 | 4.669 | 0.250 | 3.735 |
| JUN | 3.212 | 0.050 | 3.364 | 0.100 | 3.517 | 0.150 | 3.670 | 0.200 | 3.823 | 0.250 | 3.059 |
| JUL | 2.458 | 0.050 | 2.575 | 0.100 | 2.692 | 0.150 | 2.809 | 0.200 | 2.926 | 0.250 | 2.341 |
| AUG | 1.672 | 0.050 | 1.752 | 0.100 | 1.831 | 0.150 | 1.911 | 0.200 | 1.990 | 0.250 | 1.592 |
| SEPT | 2.307 | 0.050 | 2.417 | 0.100 | 2.527 | 0.150 | 2.637 | 0.200 | 2.747 | 0.250 | 2.197 |
| OCT | 1.852 | 0.050 | 1.940 | 0.100 | 2.028 | 0.150 | 2.116 | 0.200 | 2.204 | 0.250 | 1.763 |
| NOV | 0.896 | 0.050 | 0.939 | 0.100 | 0.981 | 0.150 | 1.024 | 0.200 | 1.067 | 0.250 | 0.853 |
| DEC | 0.299 | 0.050 | 0.314 | 0.100 | 0.328 | 0.150 | 0.342 | 0.200 | 0.356 | 0.250 | 0.285 |
| AVG % INCREASE IN SUNSHINE | | 0.050 | | 0.100 | | 0.150 | | 0.200 | | 0.250 | |

The above table shows the increase in ET with increase in Sunshine to 5%, 10%, 15%, 20%, 25% and keeping other independent variables i,e Temperature in Thornthwaite equation as constant

**Table: 5.6 Avearage percent increase in ET with increase in sunshine(Thornthwaite method)**

| INCREASE IN SUNSHINE | AVERAGE PERCENT INCREASE IN ET WITH INCREASE IN SUNSHINE |
|---|---|
| 0 | 0 |
| 5% | 0.05 |
| 10% | 0.1 |
| 15% | 0.15 |
| 20% | 0.2 |
| 25% | 0.25 |

The above table shows the average percent increase in ET with in Sunshine to 5%, 10%, 15%, 20%, 25%, keeping other independent variables i,e temperature in Thornthwaite equation as constant

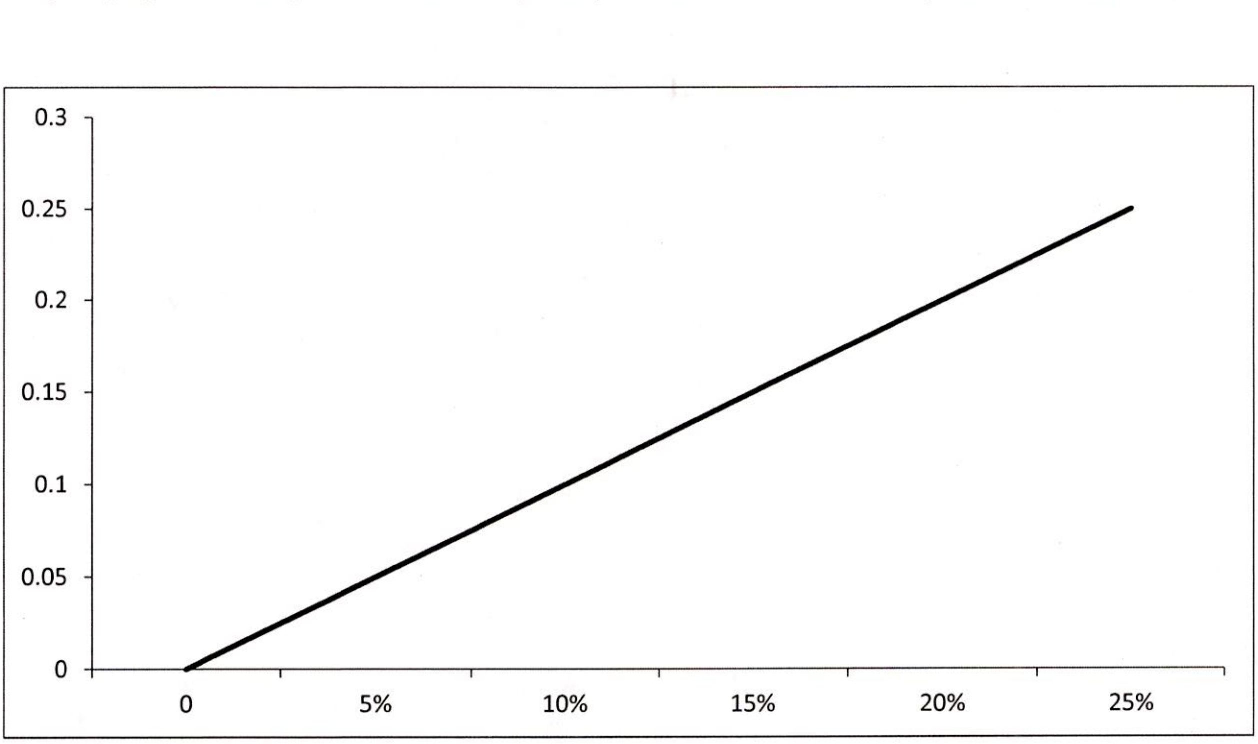


**Figure: 5.12 Average percent increase in ET with increase in sunshine**

**Table: 5.7 Average percent increase in ET with increase in temperature and sunshine (Thornthwaite method)**

| INCREASE IN INDEPENDENT VARIABLES | PERCENT INCREASE IN ET WITH INCREASE IN TEMPERATURE | PERCENT INCREASE IN ET WITH INCREASE IN SUNSHINE |
|---:|---:|---:|
| 0% | 0 | 0 |
| 5% | 0.022 | 0.050 |
| 10% | 0.045 | 0.100 |
| 15% | 0.067 | 0.150 |
| 20% | 0.089 | 0.200 |
| 25% | 0.111 | 0.250 |

The above table shows the average percent increase in ET with increase in both independent variables (i,e temperature and sunshine) to 5%, 10%, 15%, 20%, 25% in Thornthwaite equation

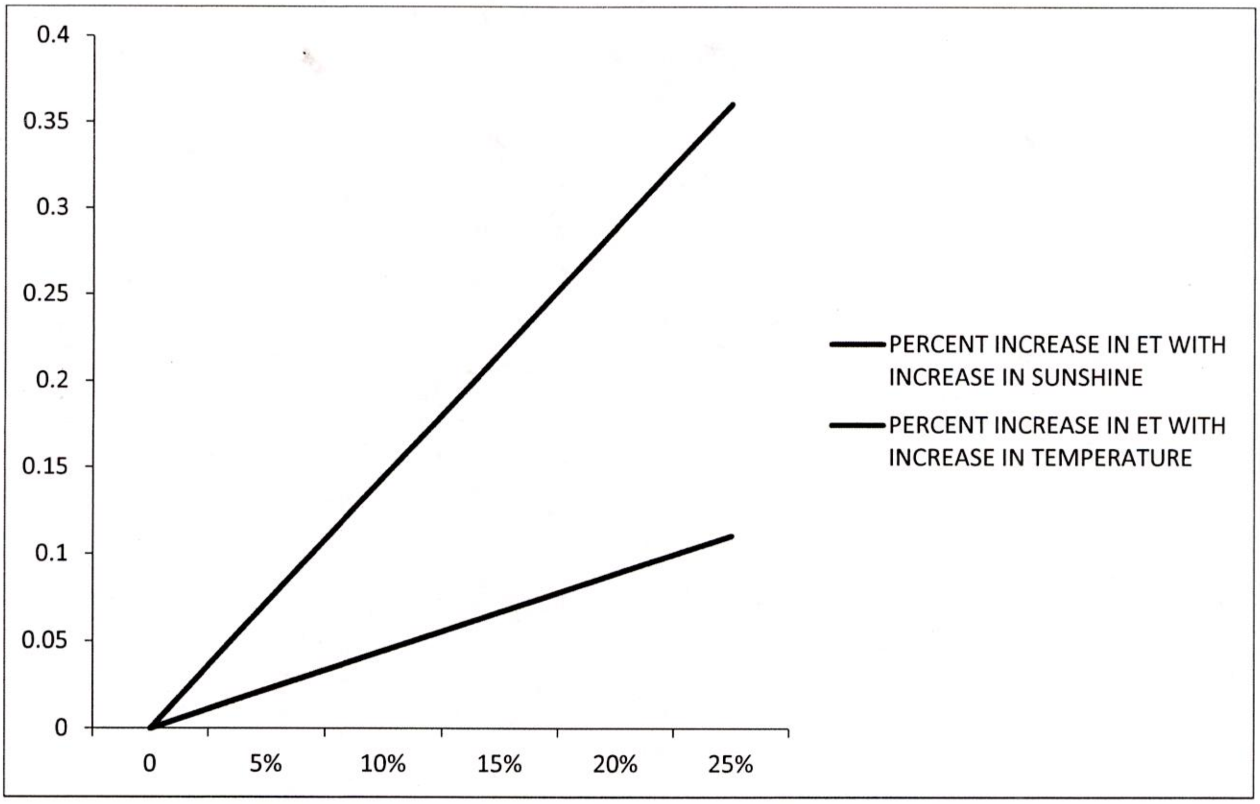


**Figure: 5.13 Average percent increase in ET with increase in sunshine and temperature**

The graph shown above, shows the average percent increase in ET with increase in independent variables i,e temperature and sunshine by Thornthwaite method and the slopes of the graph clearly indicate that impact of sunshine on ET in Thornthwaite method was more than the temperature.

## 5.4 Sensitivity analysis by Turc Method:

Here ET depend upon two independent variables i,e $T_{average}$ and solar radiation. so by increasing the values of the one of the independent variables to 5%, 10%, 15% ,20%, 25% keeping other constant ,we get the values ,which are given in below :

**Table: 5.8 percent increase in ET with increase in temperature(Turc method)**

| | Col1 | Col2 | Col3 | Col4 | Col5 | Col6 | Col7 | Col8 | Col9 | Col10 | Col11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | ET (temp5%) | | ET (temp10%) | | ET1 (temp5%) | | ET (temp20%) | | ET (temp25%) | | ET0 (original) |
| JAN | 1.394 | 0.026 | 1.428 | 0.052 | 1.461 | 0.076 | 1.493 | 0.099 | 1.523 | 0.122 | 1.358 |
| FEB | 1.888 | 0.024 | 1.930 | 0.047 | 1.970 | 0.069 | 2.009 | 0.089 | 2.045 | 0.109 | 1.844 |
| MAR | 2.536 | 0.021 | 2.585 | 0.041 | 2.632 | 0.060 | 2.677 | 0.078 | 2.719 | 0.095 | 2.483 |
| APR | 3.198 | 0.019 | 3.253 | 0.036 | 3.306 | 0.053 | 3.356 | 0.069 | 3.403 | 0.084 | 3.139 |
| MAY | 3.596 | 0.017 | 3.654 | 0.034 | 3.710 | 0.050 | 3.762 | 0.065 | 3.812 | 0.079 | 3.534 |
| JUN | 3.727 | 0.017 | 3.787 | 0.034 | 3.844 | 0.049 | 3.898 | 0.064 | 3.949 | 0.078 | 3.664 |
| JUL | 3.611 | 0.018 | 3.670 | 0.034 | 3.726 | 0.050 | 3.780 | 0.065 | 3.831 | 0.080 | 3.548 |
| AUG | 3.387 | 0.018 | 3.444 | 0.035 | 3.497 | 0.051 | 3.548 | 0.066 | 3.596 | 0.081 | 3.327 |
| SEPT | 2.987 | 0.018 | 3.038 | 0.035 | 3.085 | 0.052 | 3.131 | 0.067 | 3.174 | 0.082 | 2.934 |
| OCT | 2.416 | 0.019 | 2.459 | 0.038 | 2.500 | 0.055 | 2.540 | 0.072 | 2.577 | 0.087 | 2.370 |
| NOV | 1.796 | 0.022 | 1.832 | 0.043 | 1.867 | 0.063 | 1.900 | 0.081 | 1.932 | 0.099 | 1.757 |
| DEC | 1.401 | 0.025 | 1.434 | 0.049 | 1.465 | 0.071 | 1.494 | 0.093 | 1.523 | 0.114 | 1.367 |
| AVG % INCREASE IN TEMP | | 0.020 | | 0.040 | | 0.058 | | 0.076 | | 0.092 | |

The above table shows the increase in ET with increase in Temperature to 5%, 10%, 15%, 20%, 25% and keeping other independent variables i,e Solar Radiation in Turc equation as constant

**Table: 5.9 Average percent increase in ET with increase in temperature (Turc Method)**

| INCREASE IN TEMPERATURE | AVERAGE PERCENT INCREASE IN ET WITH INCREASE IN TEMPERATURE |
|---|---|
| 0 | 0 |
| 5% | 0.020 |
| 10% | 0.040 |
| 15% | 0.058 |
| 20% | 0.076 |
| 25% | 0.092 |

The above table shows the average percent increase in ET with in Temperature to 5%, 10%, 15%, 20%, 25%, keeping other independent variables i,e Solar Radiation in Turc equation as constant

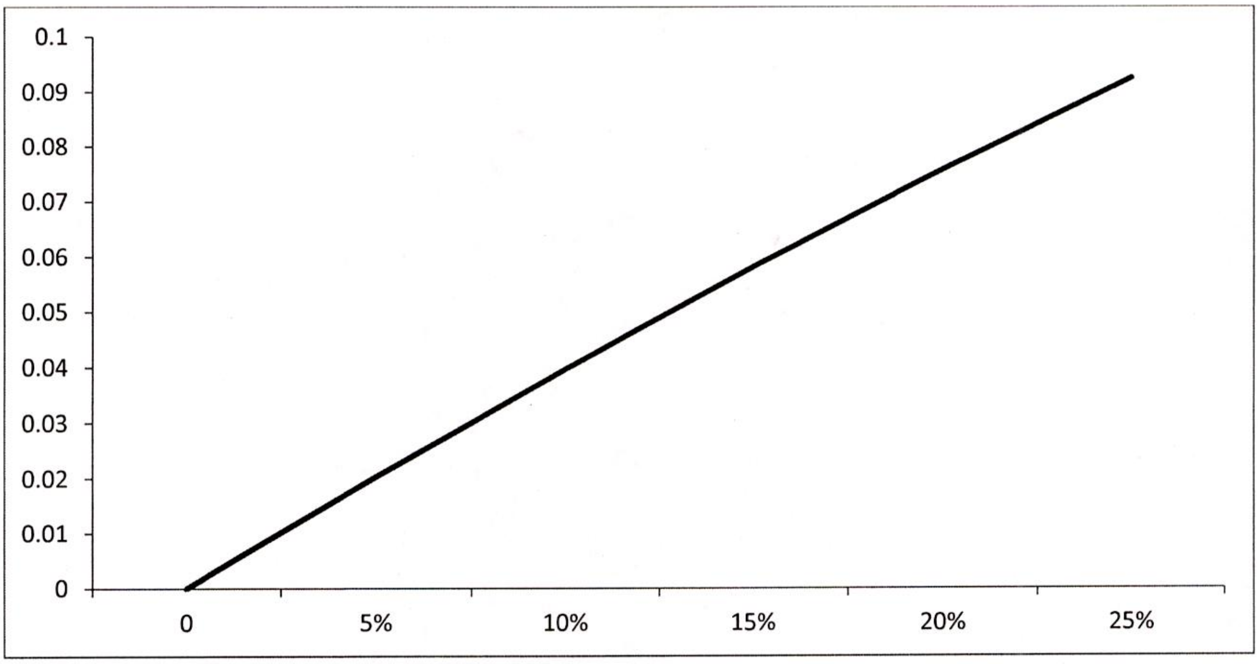


**Figure: 5.14 Average percent increase in ET with increase in temperature ( Turc method)**

**Table: 5.10 Increase in ET with increase in solar radiation (Turc Method)**

| | Col1 | Col2 | Col3 | Col4 | Col5 | Col6 | Col7 | Col8 | Col9 | Col10 | Col11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | ET (Rad 5%) | | ET (rad 10%) | | ET (rad 15%) | | ET (rad 20%) | | ET (rad 25%) | | ETO (original) |
| JAN | 1.419 | 0.045 | 1.481 | 0.091 | 1.543 | 0.136 | 1.604 | 0.182 | 1.666 | 0.227 | 1.358 |
| FEB | 1.929 | 0.046 | 2.014 | 0.092 | 2.100 | 0.139 | 2.185 | 0.185 | 2.270 | 0.231 | 1.844 |
| MAR | 2.600 | 0.047 | 2.716 | 0.094 | 2.833 | 0.141 | 2.949 | 0.187 | 3.065 | 0.234 | 2.483 |
| APR | 3.288 | 0.047 | 3.436 | 0.095 | 3.584 | 0.142 | 3.733 | 0.189 | 3.881 | 0.236 | 3.139 |
| MAY | 3.702 | 0.047 | 3.870 | 0.095 | 4.037 | 0.142 | 4.205 | 0.190 | 4.373 | 0.237 | 3.534 |
| JUN | 3.838 | 0.048 | 4.012 | 0.095 | 4.186 | 0.143 | 4.361 | 0.190 | 4.535 | 0.238 | 3.664 |
| JUL | 3.716 | 0.047 | 3.885 | 0.095 | 4.053 | 0.142 | 4.222 | 0.190 | 4.391 | 0.237 | 3.548 |
| AUG | 3.485 | 0.047 | 3.643 | 0.095 | 3.800 | 0.142 | 3.958 | 0.189 | 4.115 | 0.237 | 3.327 |
| SEPT | 3.072 | 0.047 | 3.210 | 0.094 | 3.348 | 0.141 | 3.486 | 0.188 | 3.624 | 0.235 | 2.934 |
| OCT | 2.480 | 0.046 | 2.590 | 0.093 | 2.700 | 0.139 | 2.810 | 0.186 | 2.921 | 0.232 | 2.370 |
| NOV | 1.838 | 0.046 | 1.918 | 0.091 | 1.998 | 0.137 | 2.079 | 0.183 | 2.159 | 0.229 | 1.757 |
| DEC | 1.429 | 0.045 | 1.490 | 0.090 | 1.552 | 0.135 | 1.614 | 0.180 | 1.676 | 0.226 | 1.367 |
| AVG % INCREASE IN INCREASE | | 0.047 | | 0.093 | | 0.140 | | 0.187 | | 0.233 | |

The above table shows the increase in ET with increase in Solar Radiation to 5%, 10%, 15%, 20%, 25% and keeping other independent variables i,e Temperature in Turc equation as constant

**Table: 5.11 Average percent increase in ET with increase in solar radiation**

| INCREASE IN SOLAR RADIATION | AVGERAGE PERCENT INCREASE IN ET WITH INCREASE IN SOLAR RADIATION |
|---|---|
| 0 | 0 |
| 5% | 0.047 |
| 10% | 0.093 |
| 15% | 0.140 |
| 20% | 0.187 |
| 25% | 0.233 |

The above table shows the average percent increase in ET with in Solar Radiation to 5%, 10%, 15%, 20%, 25%, keeping other independent variables i,e Temperature in Turc equation as constant

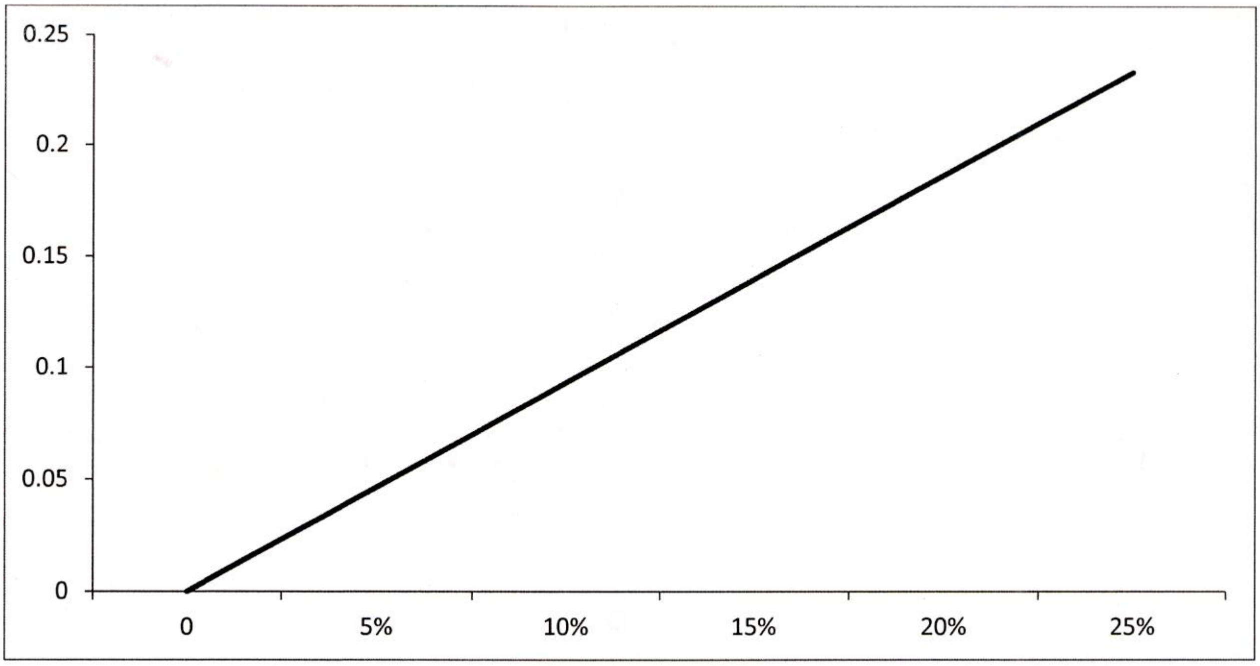


**Figure: 5.15 Average percent increase in ET with increase in solar radiation (Turc Method)**

**Table: 5.12 Average percent increase in ET with increase in temperature and solar radiation (Turc Method)**

| INCRAESE IN TEMPERATURE AND SOLAR RADIATION | AVERAGE PERCENT INCREASE IN ET WITH INCREASING TEMPERATURE | AVEARGE PERCENT INCREASE IN ET WITH INCREASING SOLAR RADIATION |
|---|---|---|
| 0 | 0 | 0 |
| 5% | 0.020 | 0.047 |
| 10% | 0.040 | 0.093 |
| 15% | 0.058 | 0.140 |
| 20% | 0.076 | 0.187 |
| 25% | 0.092 | 0.233 |

The above table shows the average percent increase in ET with increase in both independent variables (i,e temperature and sunshine) to 5%, 10%, 15%, 20%, 25% in Turc equation

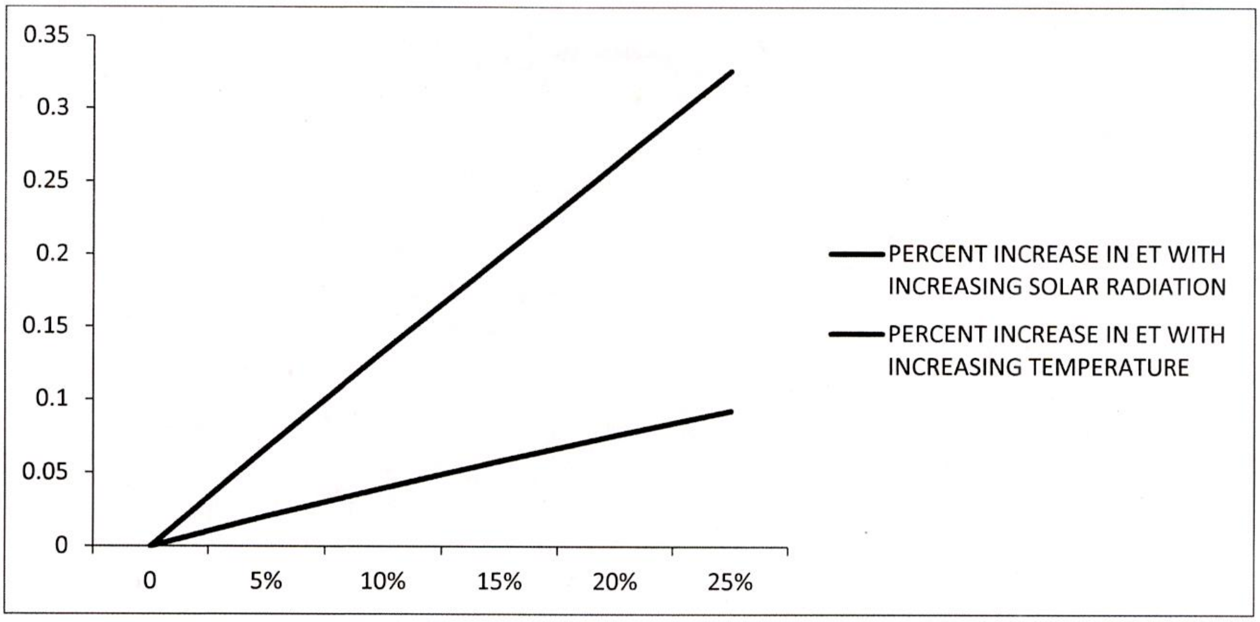


**Figure: 5.16 Average percent increase in ET with increase in temperature and solar radiation (Turc method)**

The graph shown above, shows the average percent increase in ET with increase in independent variables i.e., temperature and solar radiation by Turc method and The slopes of the graph clearly indicate that impact of solar radiation on ET in Turc method was more than the temperature.

**Table: 5.13 Average percent increase in ET with increase in variables (temperature, solar radiation ,sunshine) by both Thornthwaite method and Turc method**

| PERCENT INCREASE IN VARIABLES | AVERAGE PERCENT INCREASE IN ET WITH TEMPERATURE BY THORN | AVERAGE PERCENT INCREASE IN ET WITH INCREASE IN SUNSHINE BY THORN | AVERAGE PERCENT INCREASE IN ET WITH INCREASE IN TEMPERATURE BY TURC | AVERAGE PERCENT INCREASE IN ET WITH SOLAR RADIATION BY TURC |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 |
| 5% | 0.022 | 0.050 | 0.020 | 0.047 |
| 10% | 0.045 | 0.100 | 0.040 | 0.093 |
| 15% | 0.067 | 0.150 | 0.058 | 0.140 |
| 20% | 0.089 | 0.200 | 0.076 | 0.187 |
| 25% | 0.111 | 0.250 | 0.092 | 0.233 |

The above table shows the comparison between average percent increase ET with increase in respective independent variables (i,e temperature, sunshine, solar radiation) in Thornthwaite and Turc Method

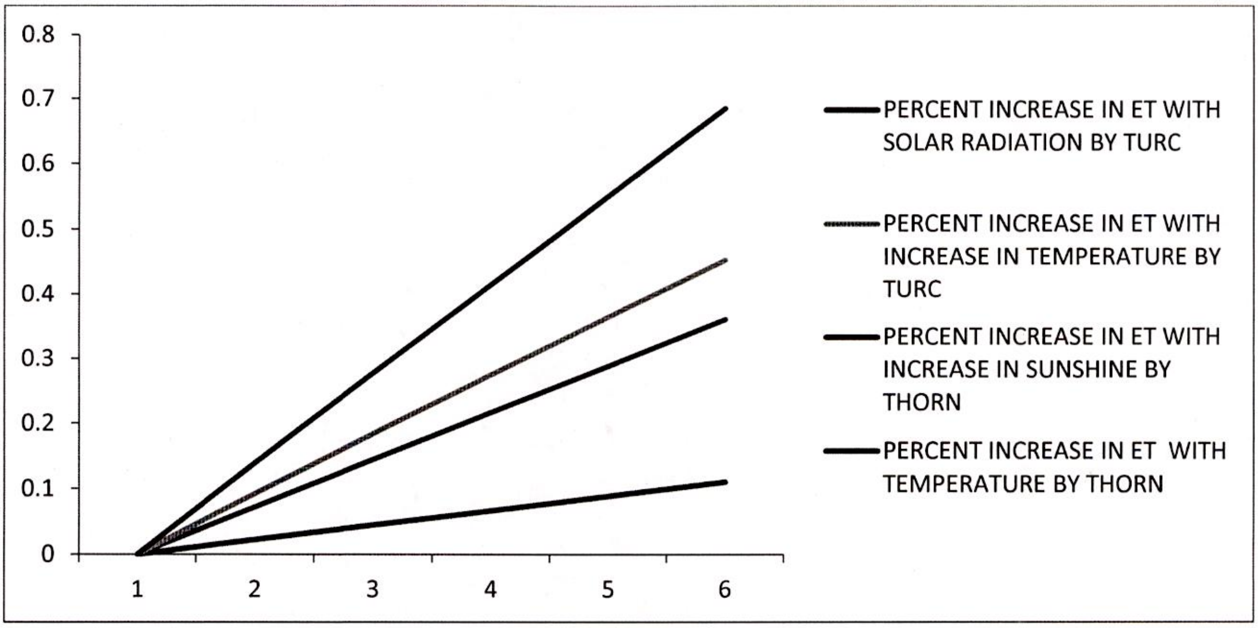


**Figure : 5.17 Average percent increase in ET with increase in variables (temperature, solar radiation ,sunshine ) by Thornthwaite method and Turc method**

The slopes of the graph that are obtained during the sensitivity analysis by Thornthwaite method and Turc method by changing the one independent variables (i,e sunshine hours, average temperature, relative humidity, solar radiation) while keeping others constant indicates the impacts of sensitivity of independent variables on the evapotranspiration. The slopes of the graph clearly indicates that while during the sensitivity analysis by Thornthwaite ,sunshine hours has more impact than that of other independent parameters like average temperature whereas during the sensitivity analysis by Turc method ,solar radiation has more impacts on evapotranspiration .Comparing the slopes of graph made between sensitivity analysis during the Thornthwaite method and Turc method ,we concluded that solar radiation is the most important independent parameter that effects the evapotranspiration followed by average temperature and sunshine hours.

# 6. CONCLUSIONS

ANN model has been developed to predict the evaporation losses for data obtained from NIH campus Roorkee. The antecedent daily rainfall, maximum, minimum temperature, relative humidity and evaporation data is collected from Jan 2001 to Dec 2013..

The statistical parameters ACF, PACF and CCF have been used for selection of the input vector. Based on the results of ACF, PACF and CCF the daily rainfall with one month lag (t-1), daily relative humidity with one month lag (t-1) and daily maximum temperature (t) and daily minimum temperature (t) have been used as the input. The feed forward neural network architecture has been trained with Bayesian regularization algorithm having 4 input nodes and 1 hidden and output node. The number of neurons in the hidden layer is optimized to 4 by trial and error, network parameters is also optimized by trial and error. Out of 4787 daily data sets, 1 to 3286 sets (62.5%) of data is used for training, 3287 to 4787 sets (37.5%) of data is used for calibration.

Similarly MLR model is developed by using data of daily rainfall, daily maximum and minimum temperature, daily relative humidity and a constant as input data and evaporation as output. Here also, out of 4787 daily datasets, 1 to 3286 sets (62.5%) of data is used for training, 3287 to 4787 sets (37.5%) of data is used for validation. The statistical indices such as coefficient of correlation, root mean squared error (RMSE) and model efficiency have been used to evaluate the performance of the both the models.

The analysis of the performance of the both ANN and MLR models clearly indicate that the application of ANN helps in the better prediction of Evaporation. A comparison of results obtained by the ANN model with those of the MLR model shows that ANN is better than that of MLR

The RMSE of ANN model during calibration and validation were found to be 1.0151 and 1.0045 respectively, whereas for the MLR model, RMSE value during calibration and validation were 1.1874 and 1.1523 respectively, and also the ANN model efficiency during calibration and validation were 0..7800 and 0.7830 respectively, whereas for the MLR model efficiency during calibration and validation were 0.6988 and 0.8501 respectively, indicates a substantial improvement in the model performance.

In addition, comparison of time series plots showed that the evaporation values estimated by the ANN model are more precise than those found by the MLR.

The slopes of the graph that are obtained during the sensitivity analysis by Thornthwaite method and Turc method by changing the one independent variables (i,e sunshine hours, average temperature, relative humidity, solar radiation) while keeping others constant indicates the impacts of sensitivity of independent variables on the evapotranspiration. The slopes of the graph clearly indicates that while during the sensitivity analysis by Thornthwaite, sunshine hours has more impact than that of other independent parameters like average temperature whereas during the sensitivity analysis by Turc method ,solar radiation has more impacts on evapotranspiration .Comparing the slopes of graph made between sensitivity analysis during the Thornthwaite method and Turc method ,we concluded that solar radiation is the most important independent parameter that effects the evapotranspiration followed by average temperature and sunshine hours.

# REFERENCES


- ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. 2000a. "Artificial neural networks in hydrology-I: Preliminary concepts". *J. Hydrologic Engineering,* 5(2):115-123.
- ASCE Task Committee on Application of Artificial Neural Networks in Hydrology. 2000b. "Artificial neural networks in hydrology-II: Hydrologic applications". *J. Hydrologic Engineering,* 5(2):124-137.
- Fausett, L. 1994. "Fundamentals of Neural Networks".Prentice Hall, Englewood cliffs, N. J.
- French, M. N., Krajewski,W. F., and Cuykendall,R. R., 1992."Rainfall forecasting in space and time using neural network"*.J. Hydrol.,* 137: 1-31.
- Govindaraju, R. S., RamachandraRao, A., 2000. "Artificial Neural Networks in Hydrology". Kluwer Academic Publishing, The Netherlands.
- Hagan, M.T., H.B. Demuth and Beale, M., 1996."Neural network design".Boston, Mass., PWS.
- Haykin, S., 1999."Neural Networks, A Comprehensive Foundation", second ed. Prentice-Hall, Englewood Cliffs, NJ.
- Hertz, J., A. Krogh and Palmer.R.G., 1991."Introduction to the theory of neural computation.Reading", Mass., Addison Wesley.
- Hopfield, J. J., 1982."Neural networks and physical systems with emergent collective computational abilities"*.Proc., Nat. Academy of Scientists.*79: 2554-2558.
- Hornik, K., Stinchcombe, M., White, M., 1989. "Multilayer feed forward networks are universal approximators". Neural Networks 2, 359–366.
  - http://www.engr.scu.edu/~emaurer/tools/calc_solar_cgi.pl.
- 11. Maier, H. R., and Dandy,G. C., 2000."Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications". *Environ. Modell.Softw.,* 15: 101–124.

- 12. McCulloch, W. S. and Pitts, W., 1943."A logic calculus of the ideas immanent in nervous activity".*Bulletin of Math's Biophysics,* 5: 115-133.
- 13 .Mohanty, S., Jha,M. K., Kumar,A.,Srivastava,R. C. and James,B. K., 2007."Groundwater level forecasting in Kathajodi River basin using artificial neural network". Third International Groundwater Conference (IGC-2007), held at TNAU, Coimbatore, India, TS-5-12.
- 14. Maier, H. R., and Dandy,G. C., 2000."Neural networks for the prediction and forecasting of water resources variables: a review of modeling issues and applications". *Environ. Modell.Softw.,* 15: 101–124.
- 15. McCulloch, W. S. and Pitts, W., 1943."A logic calculus of the ideas immanent in nervous activity".*Bulletin of Math's Biophysics,* 5: 115-133.
- 16. Mohanty, S., Jha,M. K., Kumar,A.,Srivastava,R. C. and James,B. K., 2007."Groundwater level forecasting in Kathajodi River basin using artificial neural network". Third International Groundwater Conference (IGC-2007), held at TNAU, Coimbatore, India, TS-5-12.
- 17 .Singh, B., 1998. "Water resources of India".*J. Indian Water Works Association.*Oct-Dec., pp.229-231.
- 18. VenkataRamana, R., Krishna. B., Kumar. S. R., and Pandey. N. G., 2013. "Monthly Rainfall Prediction Using Wavelet Neural Network Analysis.*Water Resour Manage.* 27:3697–3711.
- 19 .Vijay Shankar, P. S., HimanshuKulkarni, Sunderrajan Krishnan,2011. "India's Groundwater Challenge and the Way Forward". Economic & Political Weekly. 2: 37-45.
- 20. Zhang, G. P., Patuwo,B. E. and Hu,M. Y., 2001."A simulation study of artificial neural networks for nonlinear time-series forecasting".*Comput.Oper. Res.,* 28: 381-396.
- 21. Jhajharia, D., Kithan, S.B., Fancon, A.K., 2006. Correlation between pan evaporation and metrological parameters under the climatic conditions of Jorhat (Assam). *J. of Indian Water Resources,* 26 (1-2), 39-42.

- 22. Khanikar, P.G., Nath, K.K., 1998. Relationship of open pan evaporation rate with some important meteorological parameters. *J. Agril. Sci. Soc. N - E India* 11 (1), 46-50.

- 23. Kisi, O., 2004. Multilayer perceptrons with Levenberg-Marquardt training algorithm for suspended sediment concentration prediction and estimation. *Hydrol. Sci. J.* 49 (6), 1025-40.

- Allen R.G. and Pruitt W.O., (1991). FAO-24 reference evapotranspiration factors, Journal Irrigation and Drainage Engineering, ASCE, 117(IR5), 758-773.

- A.S.C.E. Standardization of Reference Evapotranspiration Task Committee, (2005). The ASCE standardized reference evapotranspiration equation. Environmental and Water Resources Institute of the American Society of Civil Engineers.

- Babajimopoulos Ch., Antonopoulos B., Grigoriadis D. and Ilias A., (1992). Sensitivity analysis of the Penman method. proceedings of 5th conference of H.Y.U., p. 132-140.

- Beven K., (1979). A sensitivity analysis of the penman Monteith actual evapotranspiration estimates,Journal of Hydrology, 44, 169-190.

- Coleman G. and DeCoursey G.D., (1976). Sensitivity analysis applied to some evaporation and evapotranspiration models, Water Resour. Res., 12(5), 873-879

- Doorenbos J. and Pruitt W.O., (1977). Guidelines for predicting crop water requirements. FAO Irrigation and Drainage Paper No 24, 2nd ed., FAO Rome, Italy.

- Gong L., Xu C., Chen D., Haldin S. and Chen Y.D., (2006). Sensitivity of the Penman-Monteith evapotranspiration to key climatic variables in the Changjiang (Yangtze River) basin, Journal of Hydrology, 329, 620-629.

- Irmak S., Payero J.O., Martin D.L., Irmak A. and Howell A.T., (2006). Sensitivity coefficients of standardized daily ASCE Penman –Monteith equation, J. Irrig. Drain. E.-ASCE, 132(6), 564-578.
- McCuen H.R., (1974). A sensitivity and error analysis of procedures used for estimating evapotranspiration. Water Resour. Bull., 10(3), 486-498 Saxton K.E., (1975).
- Sensitivity analysis of the combination evapotranspiration equation, Agricultural Meteorology, 15, 343-353.