Proceedings WARMICE-2012
Sudhir Kumar, et al. (eds)
I.A.H., 2012

Estimation of Potential Evapotranspiration under Changing Environment- A Case Study

[1]Rama Mehta, [2]Kumar Garvit, [2]Hari Om, [2]V. K. Dwivedi and [3]Vipin Kumar

[1]National Institute of Hydrology, Roorkee,
[2]National Institute of Technology, Durgapur
[3]College of Engineering, Roorkee

Abstract : Evapotranspiration constitutes one of the major components of the hydrological cycle and hence its accurate estimation is of vital importance to assess water availability and requirements. This study explores the utility of soft computing techniques to develop the model for evapotranspiration process. An important characteristic of these techniques are that both the model structure and coefficients are simultaneously optimized.

The goal of this study is to analyze the potential evapotranspiration with changing environment using soft computing techniques. A data driven model has been developed to predict the Potential evapotranspiration of Varanasi (a part of Eastern Uttar Pradesh). The Eastern part of Uttar Pradesh includes sixteen districts namely Allahabad, Azamgarh, Ballia, Chandauli, Deoria, Ghazipur, Gorakhpur, Jaunpur, Kushinagar, Maharajganj, Mau, Mirzapur, Sant Kabir Nagar, Sant Ravidas Nagar, Sonbhadra and Varanasi. During this modeling, only data for Varanasi area is considered.

Vapour-pressure, precipitation, cloud-cover, wet day frequency and average temperature of the region are used as the input data while potential evapotranspiration is used as output of the model. A relationship between inputs and output has been developed through the Fuzzy based soft computing modeling. This model has been developed on the basis of Fuzzy Technique which is one of the emerging techniques in the field of modeling. Grid Partitioning and Subtractive Clustering Methods in Adaptive Neuro-Fuzzy Inference System (ANFIS) have been used to develop the models. During the study, data have been taken from the website http://www.indiawaterportal.org/. After developing the models, the effects of temperature increments have been studied over potential evapotranspiration which gives the climate change in this region. Thus the effect of temperature increment (Global Warming) has been studied for the climate change.

Keyword: Evapotranspiration, ANFIS, Fuzzy technique, Subtractive clustering, Grid Partitioning.

INTRODUCTION

The change in temperature is expected to alter precipitation and evapotranspiration, the prime drivers of water availability and agriculture production. Climate change due to evaporation is concern not only to the scientific community but for policy makers as well. The key factor in determining agricultural potential is the adequate supply of water. The water supply may be effected by changes in quantity and timing of precipitation. The possible effect of climate changes in terms of evaporation on surface are investigated using data driven model such as Takagi-Sugeno Grid Partitioning and Subtractive Clustering Techniques for Varanasi city in Eastern Uttar Pradesh.

EFFECTS OF GLOBAL WARMING ON CLIMATE CHANGE

The effects of global warming on climate change are of concern both for the environment and human life. Evidence of observed climate change includes the instrumental temperature record, rising sea levels, and decreased snow cover in the Northern Hemisphere. According to the IPCC fourth assessment report, "most of the observed increase in global average temperature since the mid-20th century is very likely due to the observed increase in greenhouse gas concentrations".

It is predicted that future climate changes will include further global warming (i.e., an upward trend in global mean temperature), sea level rise, and a probable increase in the frequency of some extreme weather events. Ecosystems are seen as being particularly sensitive to climate change. To reduce the risk of large changes in future climate, many countries have implemented policies designed to reduce their emissions of greenhouse gases. Over the last hundred years or so, the instrumental temperature record has shown an increment in global mean temperature i.e. global warming. Moving from global to regional scales, there is increased uncertainty over how climate will change.

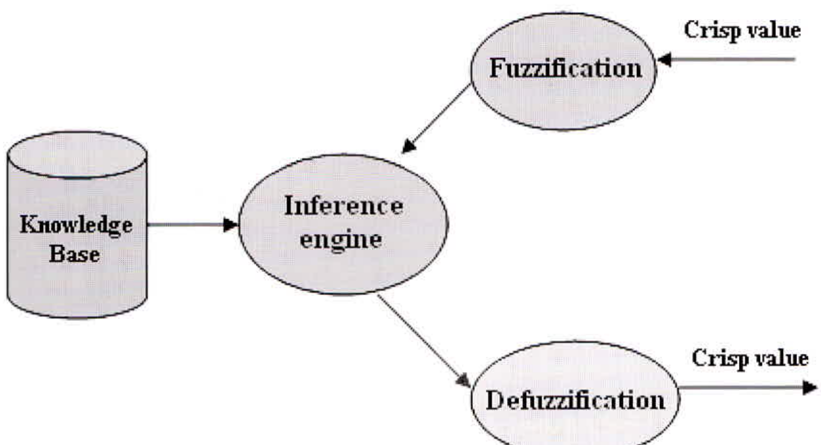
FUZZY LOGIC THEORY (FLT)

Fuzzy logic is the comprehensive form of classical logic. Fuzzy logic is the superset of classical logic with the introduction of "degree of membership". The introduction of degree of membership allows the input to interpolate between the crisp set. The operators in both logics are similar except that their interpretation differs. Fuzzy logic is another area of artificial intelligence that has been applied successfully in different applications. The key idea about fuzzy logic theory is that it allows for something to be partly this and partly that, rather than having to be either all this or all that (Klir 2008). The degree of "belongingness" to a set or category can be described numerically by a membership number between 0 and 1.

FUZZY INTERFERENCE SYSTEMS

Fuzzy inference is the process of formulating the mapping from a given input to an output using fuzzy logic. The mapping then provides a basis from which decisions can be made, or patterns discerned. The process of fuzzy inference involves: Membership Functions, Logical Operations, and If-Then Rules. Fuzzy inference systems have been successfully applied in fields such as automatic control, data classification, decision analysis, expert systems, and computer vision. Because of its multidisciplinary nature, fuzzy inference systems are associated with a number of names, such as fuzzy-rule-based systems, fuzzy expert systems, fuzzy associative memory, fuzzy logic controllers, and simply fuzzy systems.

There are two types of fuzzy inference systems that can be implemented: Mamdani-type and Sugeno-type. These two types of inference systems vary somewhat in the way outputs are determined.



Takagi-Sugeno (TS) Fuzzy Inference System

Takagi-Sugeno FIS is similar to the Mamdani method in many respects. The first two parts of the fuzzy inference process, (fuzzyfying the inputs and applying the fuzzy operator) are exactly the same. The main difference between Mamdani FIS and TS type FIS is, the output membership functions are linear or constant.

Let us consider a function $y = f(x)$ being mapped by the TS model, in which y is the dependent variable and x is the vector (k-dimensional) of independent variables that have a relationship with y. Let us assume that n example pairs [x, y] are available for parameter estimation. Considering m rules, the mathematical functioning of the TS model is

R_i : If x_1 is $A_{i,1}$ AND ... AND x_k is $A_{i,k}$

Then $\quad y_i = a_i^{\,T} x + b_i$

where $x \in R^k$ is the input variable (antecedent) and $y_i \in R$ is the output (consequent) of the ith rule R_i. The number of rules is denoted by m and A is the antecedent fuzzy set (MF) of the ith rule, such that

$$A_i(x) : R^k \rightarrow [0,1]$$

In the case of univariate MFs $\mu_{ij}(x_j)$, the fuzzy antecedent in the TS model is typically defined as an AND-conjugation by means of the product operator

$$A_i(x) = \prod_{j=1}^{k} \mu_{ij}(x)$$

For the lth input x_l, the total output $Y(l)$ of the TS model is computed by aggregating the individual rule's contributions:

$$Y(l) = \sum u_{lj} . y_i(l)$$

where u_{li} is the normalized degree of fulfillment of the antecedent clause of rule R_l

$$u_{li} = \frac{A_i(x_l)}{\sum_{i=1}^{m} A_i(x_l)}$$

Neuro-Fuzzy System

Neuro-fuzzy refers to combination of artificial neural networks and fuzzy logic. Neuro-fuzzy hybridization results in a hybrid intelligent system that synergizes these two techniques by combining the human-like reasoning style of fuzzy systems with the learning and connectionist structure of neural networks. Neuro-fuzzy hybridization is widely termed as Fuzzy Neural Network (FNN) or Neuro-Fuzzy System (NFS). Neuro-fuzzy system incorporates the human-like reasoning style of fuzzy systems through the use of fuzzy sets and model consisting of a set of IF-THEN fuzzy rules (Zimmerman, 1991). The strength of neuro-fuzzy systems involves two contradictory requirements in fuzzy modeling: interpretability versus accuracy. In practice, one of the two properties prevails. The neuro-fuzzy in fuzzy modeling research field is divided into two areas: linguistic fuzzy modeling that is focused on interpretability, mainly the Mamdani model; and precise fuzzy modeling that is focused on accuracy, mainly the Takagi-Sugeno-Kang (TSK) model.

Neuro—fuzzy system display the following properties:

- A neuro-fuzzy system is based on a fuzzy system which is trained by a learning algorithm derived from neural network theory. A neuro-fuzzy system can be viewed as a 3-layer feedforward neural network. The first layer represents input variables, the middle (hidden) layer represents fuzzy rules and the third layer represents output variables. Fuzzy sets are encoded as (fuzzy) connection weights. It is not necessary to represent a fuzzy system like this to apply a learning algorithm to it. However, it can be convenient, because it represents the data flow of input processing and learning within the model. Sometimes 5-layer architecture is used, where the fuzzy sets are represented in the units of the second and fourth layer.

- A neuro-fuzzy system can be always interpreted as a system of fuzzy rules. It is also possible to create the system out of training data from scratch, as it is possible to initialize it by prior knowledge in form of fuzzy rules.
- The learning procedure of a neuro-fuzzy system takes the properties of the underlying fuzzy system into account. This results in constraints on the possible modifications applicable to the system parameters.
- A neuro-fuzzy system approximates unknown function that is partially defined by the training data. The fuzzy rules encoded within the system represent vague samples, and can be viewed as prototypes of the training data.

Jang(1996) introduced a novel architecture and learning procedure for FIS that uses a neural network learning algorithm for constructing a set of fuzzy if-then rules with appropriate MFs from the input-output pairs. This procedure of developing a FIS using the framework of adaptive neural network is called an adaptive neuro fuzzy inference system (ANFIS).

ANFIS operates on Sugeno type systems therefore there are only two types of output MFs: constant and linear. In MATLAB, ANFIS editor GUI invoked using "anfisedit", and training (calibration) and checking (validation) data sets are loaded. Two optimization methods used for FIS training are Hybrid (mixed-least squares and backpropagation) and backpropagation. The error tolerance is used to create a training stopping criterion, which is related to the error size. The training will stop if the designated epoch number is reached or the error goal is achieved, whichever comes first.

The property of the ANFIS model separates the problem into two sub-problem:

(i) grid base parameter identification of the consequent part;

(ii) appropriate clustering partitioning of the space.

Grid Base Partitioning

The parameters for optimization in an ANFIS are the premise parameters which describe the shape of the MFs, and the consequent parameters which describe the overall output of the system. The basic learning rule of an adaptive network, the backpropagation algorithm, which is based on the gradient descent rule, can be successfully applied to estimate these parameters but Hybrid algorithm [combination of the gradient descent method and the least squares estimate (LSE)] is a fast learning algorithm.

The general architecture of the ANFIS is presented in fig. 1. An adaptive network is a multi-layered feed forward structure whose overall output behaviour is determined by the value of a collection of modifiable parameters. Let us consider that the FIS has two inputs x and y and one output. For the first order Sugeno fuzzy model, a typical rule set with two fuzzy if then rules can be expressed as:

Rule1: IF x is A_1 and y is B_1, THEN

$$f_1 = p_1 x + q_1 y + r_1$$

Rule2: IF x is A_2 and y is B_2, THEN $f_2 = p_2 x + q_2 y + r_2$

where A_1, A_2 and B_1, B_2 are the MFs for inputs x and y, respectively; p_1, q_1, r_1 and p_2, q_2, r_2 are the parameters of the input function.

Here the nodes of the layers have similar functions of Mamdani system except the linear or constant output.

Cluster Partitioning

In many situations, Fuzzy rules are difficult to identify by manual inspection and therefore are derived from observed data using techniques collectively known as fuzzy clustering.

The basic purpose of fuzzy clustering is to identify natural grouping of the data from a large data set, producing a concise representation of system behaviour. Various methods have been developed in the literature, such as fuzzy C-means clustering, mountain clustering, Gaustafron-Kessel (GK) fuzzy clustering and subtractive clustering. The subtractive clustering method has been used for present study.

Subtractive Clustering Subtractive clustering method is an extension of the mountain clustering method, where data points are considered as the potential candidates for cluster centres. It is a data clustering technique wherein each data point belongs to a cluster to some degree that is specified by membership grade. Clustering of numerical data forms the basis of many classification and system modeling algorithms. The purpose of clustering is to identify natural groupings of data from a large data set to produce a concise representation of a system's behaviour. Sub clustering tool is used to generate the ANFIS with minimum number of rules. The rules partition themselves according to the fuzzy qualities associated with each of the data clusters.

This approach is used to determine the number of rules and antecedent membership functions by considering each cluster centre (D_i) as a fuzzy rule. In this approach each data point of a set of N data points $\{x_1, x_2, ..., x_n\}$ in a p-dimensional space is considered as the candidate for the cluster centres. After normalizing data points in each direction, a density measure at data point x_i is computed on the basis of its location with respect to other data points and expressed as:

$$D_i = \sum_{j=1}^{N} e^{\left[-\left(\frac{2}{r_a}\right).(x_i - x_j)^2\right]}$$

where r_a is a positive constant called cluster radius.

A data point is considered as a cluster centre when more data points are closer to it. Therefore, the data point (x_1^*) with the highest density measure (D_1^*) is considered as first cluster centre. Now excluding the influence of the first cluster centre, the density measure of all other data points is recalculated as

$$D_i^* = D_i - D_1^*.\mu_i(x_i^*) \tag{1}$$

$$\mu(x_i^*) = e^{\left[-\frac{\|x_i - x_i^*\|^2}{(r_b/2)}\right]} \tag{2}$$

where $r_b (r_b > r_a)$ is a positive constant that results in a measurable reduction in density measures of neighbourhood data points so as to avoid closely spaced cluster centres.

After the density measure for each data point (Eqs.

1 and 2), the data point with the highest remaining density measure is obtained and set as the next cluster D_2^* and all of the density measures for data points are revised again. The process is repeated and the density measures of remaining data points after computation of kth cluster centre are revised by substituting the location $\left(x_k^*\right)$ and density measure (D_k^*) of the kth cluster centre in Eq. (2). This process is stopped when a sufficient number of cluster centers are generated. A sophisticated stopping criterion for automatically determining the number of clusters was suggested by Chau. These cluster centers $\left(D_i^*\right)$ can be reasonably used as the centers for the fuzzy rules' premise and antecedent membership function that describes the system behavior.

STUDY AREA

Eastern part of Uttar Pradesh consist of sixteen districts namely Allahabad, Azamgarh, Ballia, Chandauli, Deoria, Ghazipur, Gorakhpur, Jaunpur, Kushinagar, Maharajganj, Mau, Mirzapur, Sant Kabir Nagar, Sant Ravidas Nagar, Sonbhadra, and Varanasi.

The city of Varanasi is located in the middle Ganges valley of North India, in the Eastern part of the state of Uttar Pradesh, along the left crescent-shaped bank of the Ganges river. The "Varanasi Urban Agglomeration" — an agglomeration of seven urban sub-units — covers an area of 112.26 km ^{2} (approximately 43 mi²). The urban agglomeration is stretched between 82° 56'E - 83° 03'E and 25° 14'N - 25° 23.5'N. Being located in the Indo-Gangetic Plains of North India, the land is very fertile because low level floods in the Ganges continually replenish the soil. Fig. 2 shows the location of Varanasi city in Eastern Uttar Pradesh.

On a *local* level, Varanasi is located on a higher ground between rivers Ganges and Varuna, the mean elevation being 80.71 m. As a result of absence of tributaries and canals, the main land is continuous and relatively dry. In ancient times, this geographic situation must have been highly favourable for forming settlements. But it is difficult to ascertain the original geography of Varanasi because the city's current location is not exactly the same as the one described in some old texts.

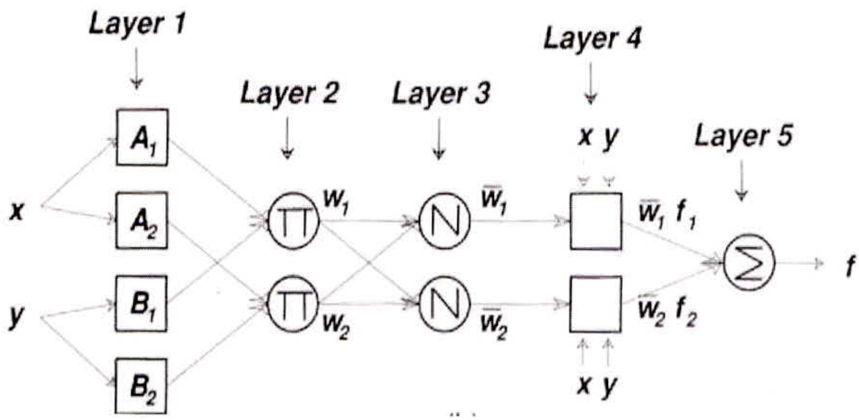


Fig. 1. General Architecture of ANFIS

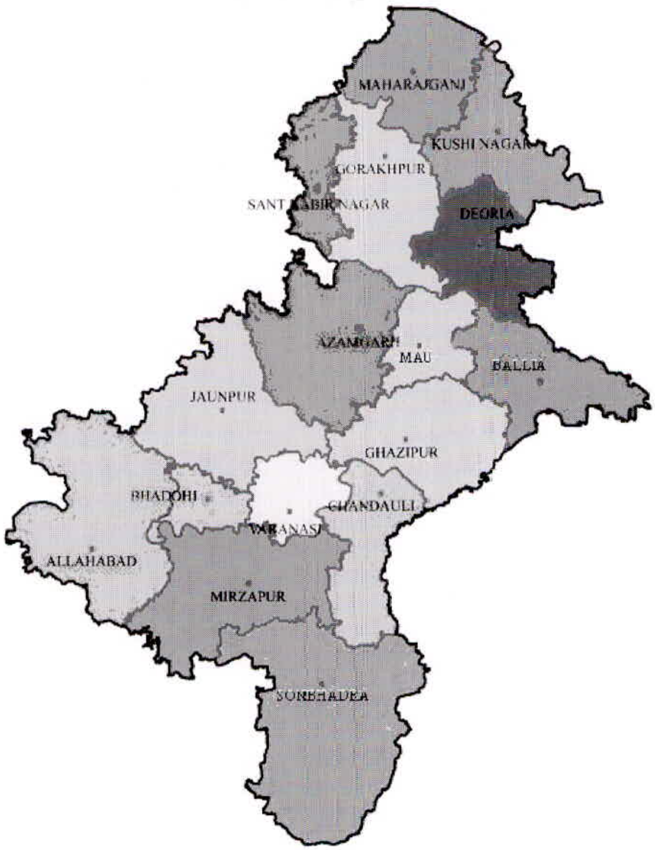


Fig. 2. Location of Varanasi in Eastern U.P.

Varanasi is often said to be located between two confluences: one of the Ganges and Varuna, and other of the Ganges and Assi, (Assi having always been a rivulet rather than a river.) The distance between these two confluences is around 2.5 miles (4.0 km), and religious Hindus regard a round trip between these two places—a *Pancha-kroshi Yatra* (a five mile (8 km) journey) ending with a visit to a *Sakshi Vinayak Temple* as a holy ritual.

Climate

Varanasi experiences a humid subtropical climate with large variations between summer and winter temperatures. Summers are long, from early April to October, with intervening monsoon seasons and are also extremely hot, even by South Asian standards. The temperature ranges between 32°C – 46 °C (90°F – 115 °F) in the summers. Winters in Varanasi see very large diurnal variations, with warm days and downright cold nights. Cold waves from the Himalayan region cause temperatures to dip across the city in the winter from December to February and temperatures below 5 °C are not uncommon. The average annual rainfall is 1,110 mm (44 in). Fog is common in the winters, while hot dry winds, called loo, blow in the summers.

Through a combination of water pollution, new constructions of upstream dams, and increase in the local temperature, the water level of the Ganges has recently gone down significantly, and small islands have become visible in the middle of the river.

DATA AVAILABILITY

We have considered six climate parameters: average temperature (TEMP), cloud cover (CC), potential evapotranspiration (PET), precipitation (PPT), vapor pressure (VP), and wet day frequency (WDF). Out of these parameters, TEMP, CC, PPT, VP and WDF are input parameters and PET is output. We have taken data from the website http://www.indiawaterportal.org/ Monthly data has been taken for the 50 years period (from 1951 to 2000). The available data is partitioned in two parts. 90% values have been used for calibration/training the model and the rest 10% values were used for validation/checking of the model. Twelve month-wise databases have been prepared for five inputs (TEMP, CC, PPT, VP and WDF) and single output (PET).

MODEL DEVELOPMENT

Two models have been developed for Grid partitioning and Subtractive Clustering Method (SCM) in Takagi-Sugeno inference. The models have been developed using ANFIS editor of Fuzzy Logic Tool-Box in MATLAB (2007a).

To generate the FIS file of SCM , following criteria in Table 2 have been adopted to develop the model.

Table 3 give all ANFIS information of the subtractive cluster model. Last two columns give the training/ calibration and checking/ validation errors of the model.

To generate the FIS file of Grid Partitioning Method (GPM), following criteria in Table 4 have been adopted to develop the model.

Table 5 give all ANFIS information of the Grid partitioning method. Last two columns give the training/ calibration and checking/ validation errors of the model.

Table 1. Climate data for Varanasi

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Year
Ave.High Temp °C (°F)	19 (67)	24 (76)	31 (87)	37 (98)		36 (97)	32 (90)	31 (88)	31 (88)	31 (87)	27 (81)	22 (71)	29.9 (85.8)
Ave. Low Temp °C (°F)	8 (47)	12 (54)	17 (62)	22 (72)	25 (77)	27 (80)	26 (78)	26 (78)	24 (76)	21 (70)	15 (59)	11 (51)	19.4 (67.0)

Table 2.

Error tolerance	0.001
Epochs	30
Squash factor	1.25
Reject ratio	0.15
Accept ratio	0.5
Range of influence	0.5

Table 3

Number of nodes	92
Number of linear parameters	42
Total number of parameters	112
Number of training data pairs	540
Number of checking data pairs	60
Number of fuzzy rules	7
Training error	0.190442
Checking error	0.203531

Table 4

Number of inputs	5
Number of output	1
Number of Membership Functions (M.F.) in each input and output	7, 7, 7, 7, 7, 7
Type of M.F.	gbell M.F.
Optimization method	Hybrid
Error Tolerance	0.001

Table 5

Number of nodes	524
Number of linear parameters	243
Total number of parameters	288
Number of training data pairs	540
Number of checking data pairs	60
Number of fuzzy rules	243
Training error	0.198475
Checking error	0.996335

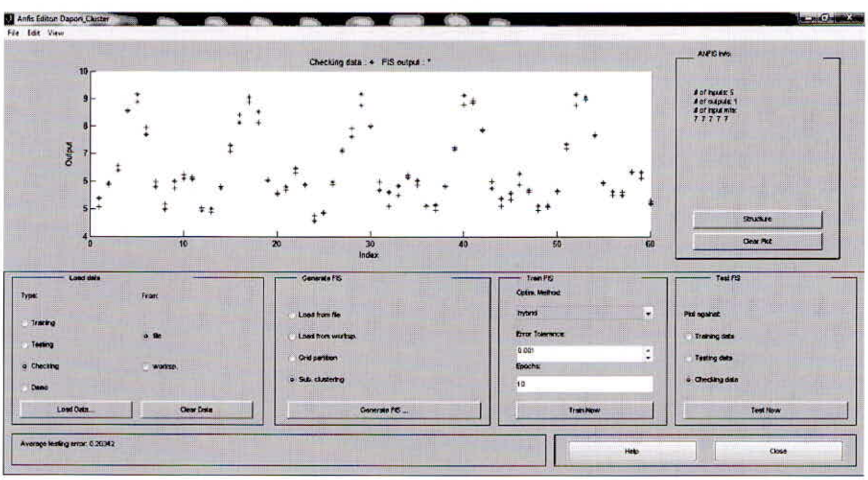


Fig. 3. Comparative output for Valedictory data using GPM .

RESULTS AND DISCUSSION

Models have been developed for calibration data (540 values) using both techniques for their maximum gains. Optimum values of PET have been obtained for valedictory data (60 values). Both models are giving satisfactory results as shown in Fig. 4.

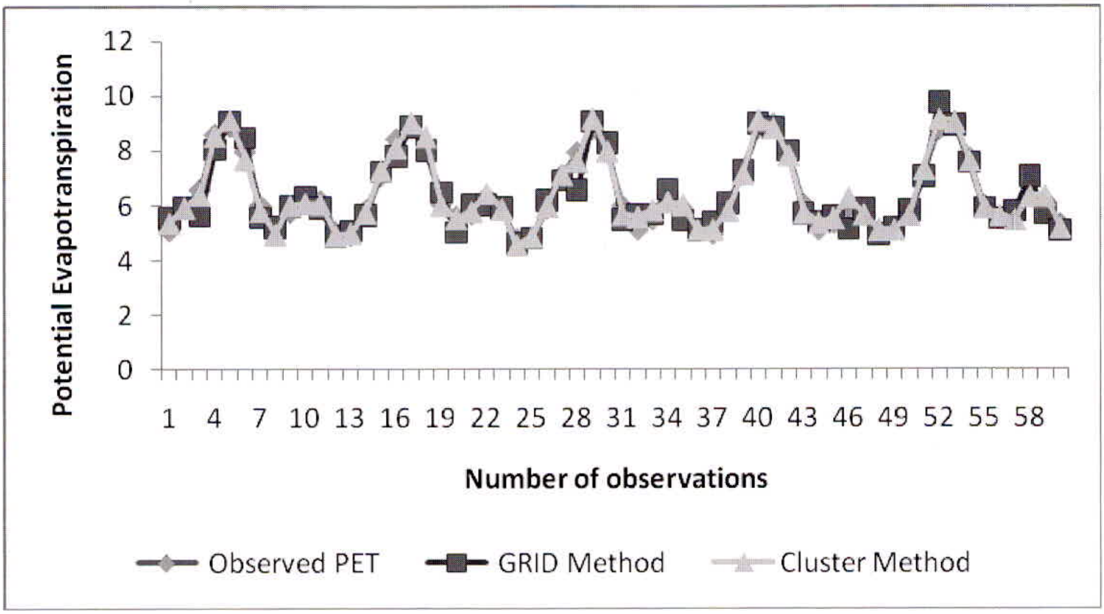


Fig. 4. Comparative plot between observed and modeled values of PET

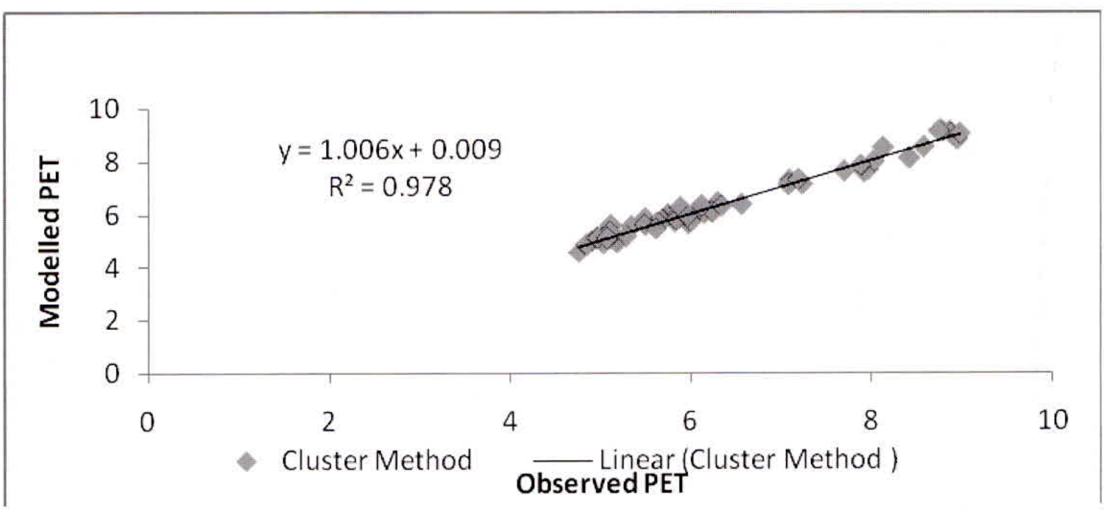


Fig.5. Linear relationship between observed and modeled values of PE for SCM.

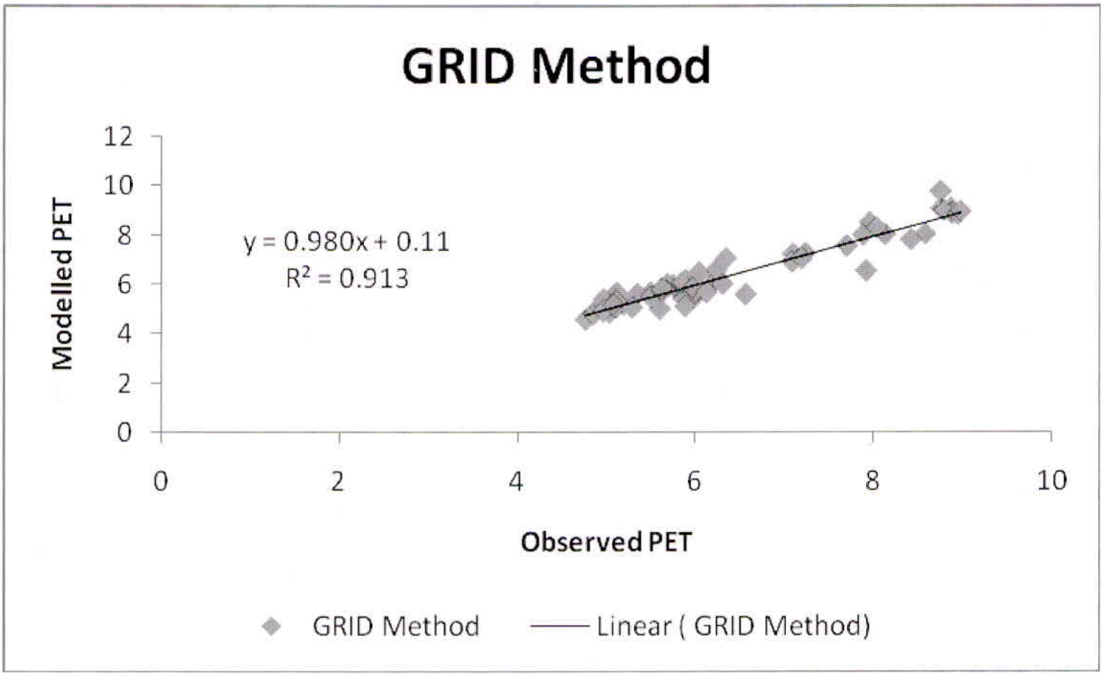


Fig.6. Linear relationship between observed and modeled values of PET for GPM.

CONCLUSIONS

Graphic User Interface (GUI) mode of the model is easy to use for study of climate change in any region for any input data. Modeled values show the consistent results with observed values. The R^2 (co-relation coefficient) show that the SCM model is little bit better than GPM as its R^2 value (0.9781) is higher. Also the number of fuzzy rules in GPM is much higher than SCM which make the model very complex and it takes more computational time. The efficiency of both the models are more than 95%. Still it is concluded that the ANFIS – subtractive cluster method give better results with minimum error.

REFERENCES

Jang Jyh-Shing Roger., *Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence,* Prentice Hall of India, 1996,335-367

Klir George J. – Yuon Bo, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall of India, 2008

MATLAB (R2007b), Fuzzy Logic Toolbox, The Mathworks, USA

Zimmermann HJ (1991) Fuzzy set theory – and its application. Kluwer Academic, Dordrecht, The Netherlands