Co-active Neuro Fuzzy Inference System for Estimating Reference Evapotranspiration over Indore District of Madhya Pradesh

Madhulika Singh[1*], Ronald Singh[2,] Vikram Singh[3], & Tushar P. Rathod[4]

[1]AKS University, Satna, India
[2]Land-use and Cover Monitoring Division, NRSC-ISRO, Hyderabad, India
[3]Dept. Soil and Water Conservation Engineering, VIAET, SHUATS, Naini, Prayagraj, India
[4]SCAET, Maldad, Maharastra
*Corresponding Author's email: madhulikasingh3u@gmail.com

Abstract
Reference evapotranspiration (ET_o) is defined as the rate at which readily available soil water is vaporized from specified vegetated surfaces (Jensen et al., 1990). The CANFIS technique is used for predicting the desire parameter of a fuzzy system when fewer meteorological variables are available. To conduct the research work, weather data rainfall (Rt), maximum temperature ($Tmax$), minimum temperature($Tmin$), bright sun shine hour (SSH) and wind speed (WS) were collected. Aforementioned data of metrological observatory Indore were taken from Indian Meteorological department Pune Maharashtra. Since Penman Monteith (P-M) equation required maximum numbers metrological variables to estimate reference evapotranspiration, CANFIS model was used for predicting the same using minimum metrological variables. Thus, this research presents the comparison of estimation of reference evapotranspiration using CANFIS model and Penman Monteith (P-M) equation. The performance of the CANFIS models was evaluated with Penman Monteith equation using sum of square error (SSE), root mean square error (RMSE) and R 2 values. It was found that the highest CE value varied from 0.748 to 0.949 during training and 0.971 to 0.949 during testing period, minimum RMSE values varied from 1.640 to 0.692 during training and 0.523 to 0.692 during testing period similarly the 'r' values was varied from 0.864 to 0.981 during training and 0.994 to 0.981 during testing period of the models. This study revealed that the Penman-Monteith (P-M) equation and CANFIS model showed very close similarities in estimation of reference evapotranspiration. On the basis of statistical analysis, the CANFIS-9 model was found best to estimate reference evapotranspiration over others and may therefore be adopted for estimating ET_0 in the regions with reasonable degree of accuracy.

Keywords ANN, CANFIS, Evapotranspiration, Penman Monteith and Metrological observations.

Introduction

Evapotranspiration (ET), the combined process of water vapor transfer from the Earth's surface through evaporation and transpiration, is a critical component of the Earth's hydrological cycle with substantial implications for various sectors, including agriculture, hydrology, and environmental science (Allen et al., 1998). ET's significance in hydrology, climate science, and environmental conservation cannot be overstated. It directly affects local and global climates by influencing temperature, precipitation, and the availability of water resources. In regions with water scarcity, such as arid and semi-arid areas, understanding and accurately estimating ET are pivotal for sustainable water management (Allen et al., 1998). Furthermore, precise ET data contribute to ecosystem preservation by guiding water allocation to wetlands, rivers, and lakes. ET also plays a role in climate change studies, as shifts in ET patterns can impact regional climate variables (Gao et al., 2015). Accurate ET estimation is essential for informed decision-making in water resource management, irrigation planning, climate modeling, and ecological studies.

Historically, several methods have been employed to estimate ET, including the Penman-Monteith equation, energy balance methods, and lysimeter measurements. The Penman-Monteith equation, endorsed by FAO (Allen et al., 1998), is considered the standard for ET estimation, incorporating various meteorological parameters. The ET_0 is commonly estimated by either physically based equations (Penman, Penman-Monteith, etc.) or empirical relationships between meteorological variables (Hargreaves, Hargreaves-Samani, Blaney-Criddle, etc.). Reference evapotranspiration can be obtained by direct and accurate techniques with special equipment, such as lysimeter, or estimated indirectly by mathematical models to provide good results (Alves Sobrinho et al., 2011). Yildirim et al., 2014, prepared daily evaporation prediction models by using empirical Penman equation, Levenberg Marquardt algorithm based on "Feed Forward Back Propagation Artificial Neural Networks (LMANN)", radial basis neural networks (RBNN) and generalized regression neural networks (GRNN). Developed models were compared in this study and it was found that the results of neural network models are statistically more meaningful than the Penman equation. FAO-56 PenmanMonteith method is considered as the best indirect method to estimate ET_0 under various agroclimatic conditions using meteorological data as input variables (Irmak et al. 2003). These traditional methods have demonstrated utility but often require extensive input data and may be impractical in data-scarce regions.

The CANFIS model offers an alternative that leverages the strengths of neural networks and fuzzy logic, enabling a more adaptive and efficient approach to ET estimation (Gao et al., 2015). The CANFIS model is a hybrid computational framework that merges artificial neural networks (ANNs) and fuzzy logic. ANNs excel in capturing complex, non-linear relationships in data, while fuzzy logic can manage uncertainty and linguistic terms (Jang, 1993). The CANFIS model comprises three primary layers: the input layer, the fuzzy layer, and the output layer. The input layer accommodates meteorological and environmental variables, which are then transformed into linguistic variables in the fuzzy layer using membership functions to describe the association of each variable with linguistic terms (Kisi et al., 2010). The architecture of the CANFIS model is pivotal to its effectiveness in ET estimation. The input layer receives meteorological data, including temperature, humidity, wind speed, and solar radiation, which are crucial factors affecting ET. The fuzzy layer transforms these inputs into linguistic variables, such as "low," "moderate," or "high," with the help of membership functions. These linguistic variables are then used to develop fuzzy rules for ET estimation, allowing the model to account for complex interactions between environmental factors (Kisi et al., 2010). One of the distinguishing features of the CANFIS model is its adaptability. It incorporates a learning algorithm that continuously adjusts the membership functions and fuzzy rules within the fuzzy layer as more data becomes available. This adaptability enables the model to evolve and enhance its performance over time, making it suitable for regions with varying climates and land use patterns (Gao et al., 2015).

The CANFIS model offers several advantages for ET estimation, including its ability to capture non-linear relationships, adapt to changing conditions, handle uncertainty, rely on empirical data, and operate with reduced data requirements (Kisi et al., 2010). CANFIS has found applications in various fields, such as agriculture, hydrology, environmental studies, and climate modeling. It assists in optimizing irrigation schedules, water resource management, ecosystem preservation, and improving climate change predictions (Gao et al., 2015). Despite its advantages, CANFIS faces challenges and limitations, including its dependence on data quality, complex model structure, computational demands, and susceptibility to overfitting (Kisi et al., 2010). The CANFIS model presents a promising approach to accurate ET estimation, blending neural networks and fuzzy logic to adapt to changing environmental conditions. This introduction has highlighted the significance of ET, the CANFIS model's structure, its advantages, applications, and limitations. Subsequent sections will delve deeper into the model's components, the process of ET estimation using CANFIS, and case studies illustrating its effectiveness in various contexts.

Materials and Methods

Evaporation can directly be measured by atmometer or can be predicted by empirical equations, (Penman, 1948). This study computed 10 years data from year 2007 to 2016 of Indore region in which 1220 numbers of complete patterns were taken for study after pre-analysis of data. The pre analysis includes removal of outliers using exploratory data analysis. CANFIS model comprised of 4 input variable such as wind speed, sunshine hour, minimum and maximum temperature to estimate the reference evapotranspiration. The flowchart of the methodology has been shown in Figure 2.

Estimation of reference evapotranspiration by Penman-Monteith method:

Many equations have been developed for estimating reference evapotranspiration (ET_0). Penman-Monteith FAO-56 equation has been used for estimating reference evapotranspiration for the present study. The equation represents a basic general description of reference evapotranspiration process. Therefore, the ET_0 computed from meteorological data by the FAO- 56 Penman-Monteith equation has been chosen as true value during training, testing and validating the neural network models. The reference evapotranspiration was calculated by the following Penman-Monteith Equation (Allen et.al., 1998) as shown in equation 1.

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

ET_0 = reference evapotranspiration (mm day^{-1});
G = soil heat flux intensity (MJM^{-2} day^{-1});
R_n = net radiation at crop surface (MJM^{-2} day^{-1});
T = mean daily air temperature at 2 m height (^0C);
γ = psychometric constant (kPa $^0C^{-1}$);
Δ = slope of saturation vapor pressure function (kPa $^0C^{-1}$);
e_s = saturation vapor pressure at temperature T (kPa);

e_a = actual vapor pressure at dew point temperature (kPa);
u_2 = average daily wind speed at 2 m height (m sec^{-1}).

Eq. 1 determines the ET_0 from an assumed grass reference surface in different periods of the year. The CANFIS technique has been used for predicting the desired parameter of a fuzzy system when enough training data are provided. Neuro fuzzy has been successfully applied to solve a number of problems in water recourses planning and management, including: stream flow reconstruction (Chang *et al.,* 2004); estimation of suspended sediment transport (Kisi, 2004).

CANFIS architecture:

CANFIS model consists of 5 inputs viz. rainfall, wind speed, sunshine hour, minimum temperature, maximum temperature and 15 nodes in hidden layer and 1 output node as ET_0. For simplicity, assume that the fuzzy inference system under consideration has two inputs *x* and *y* and one outputs *f.* The first-order Sugeno fuzzy model, a typical rule set with two fuzzy IF-THEN rules for CANFIS architecture, can be expressed in the equation 2 and 3:

Rule 1: IF x is A_1 and y is B_1 THEN $f_1=p_1x+q_1y+r_1$ Eq. (2)
Rule 2: IF x is A_2 and y is B_2 THEN $f_2=p_2x+q_2y+r_2$ Eq. (3)

Where, A_1, A_2 and B_1, B_2 are the membership functions (MF) for inputs *x* and *y* respectively, p_1, q_1, r_1 and p_2, q_2, r_2 are the parameters in the THEN-part (consequent part) of the first-order Sugeno fuzzy model illustrated in Figure 1(a). The architecture of CANFIS of five layers is illustrated in Figure 1(b). The function of each layer is described below:

Layer 1 (Fuzzification layer): Every node i^{th} in this layer is a square node with a node function as on equation 4 and 5;

$O^1_i= \mu A_i(x)$ for i = 1, 2, or Eq. (4)
$O^1_i= \mu B_i(y)$ for i = 1, 2 Eq. (5)

Where, *x* (or *y*) is the input to i^{th} node and A_i (or B_i) is a linguistic label (such as small, large etc.) associated with this node function. In other words, $O^1$is the membership function of A_i (or B_i) and it specifies the degree to which the given input *x* (or *y*) satisfies the quantifier A_i (or B_i). Usually, $\mu A_i(x)$ or $\mu B_i(y)$ are chosen to bell–shaped with maximum equal to 1 and minimum equal to 0 as shown in equation 6, such as;

$$\mu A_i(x; a_i, b_i, c_i) = \frac{1}{1+[(x-c_i)/a_i]^{2b_i}} \quad \text{Eq. (6)}$$

Where, (a_i, b_i, c_i) is the parameter set. As the values of these parameters change, the bell-shaped function varies accordingly, thus exhibiting various forms of membership functions on linguistic label A_i. Parameters in this layer are referred to a premise parameter.

Layer 2 (Rule layer): Every node in this layer is a circle node labelled Π which multiplies the incoming signals and sends the products out. For instance, as shown in equation 7,

$O^2_i = w_i = \mu A_i(x) \times \mu B_i(y)$, i = 1, 2, Eq. (7)

Each node output represents the *firing strength* of a rule.

Layer 3 (Normalization layer): Every node in this layer is a circle node labelled N. The *i*th node calculates the ratio of the i^{th} rule's firing strength to the sum of all rules' firing strengths as shown in equation 8:

$$O_i^3 = \overline{w}_i = \frac{w_i}{w_1 + w_2},\ i = 1, 2, \qquad \text{Eq. (8)}$$

For convenience, the outputs of this layer are called normalized firing strengths.

Layer 4 (Defuzzification layer): Every node i^{th} in this layer is a square node with a node function as shown in equation 9.

$$O_i^4 = \overline{w}_i f_i = \overline{w}_i(p_i x + q_i y + r_i) \qquad \text{Eq. (9)}$$

Where $\overline{w}_i$ I is the output of layer 3, and (p_i, q_i, r_i) is the parameters set. Parameters of this layer are referred to as consequent parameters.

Layer 5: (Summation layer): The single node in this layer is a circle node labelled $\sum$ that computes the overall output as the summation of all incoming signals and the equation used to estimate is shown in equation 10, i.e.,

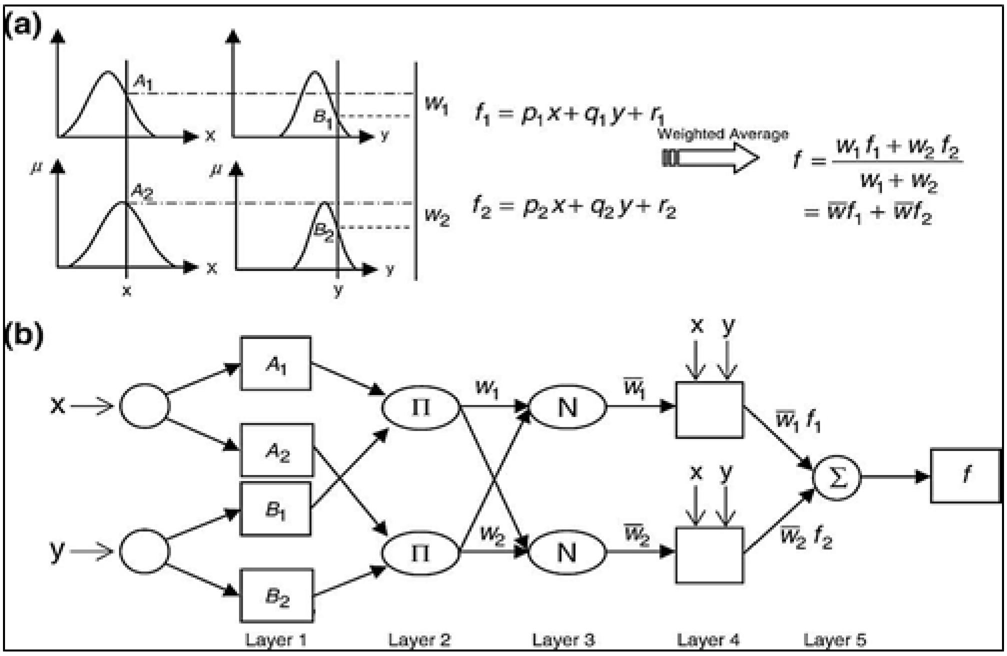

Fig. 1 (a) Sugeno's fuzzy IF-THEN rule. **(b)** the equivalent CANFIS architecture.

$$O_i^5 = \textit{overall output} = \sum_i \overline{w}_i f_i = \frac{\sum_i W_i f_i}{\sum_i W_i} \qquad \text{Eq. (10)}$$

Thus, a CANFIS network can be constructed which is functionally equivalent to a Sugeno first-order fuzzy inference system. The daily data of evapotranspiration was categorized into two major sets i.e. a training data set from 2007 to 2014 and testing data set from 2015 to 2016 of Indore region. The input pairs in the training data set were applied to the network of a selected architecture and training was performed using Gaussian and generalized bell membership functions for CANFIS models. Figure 2 shows the step-by-step procedure followed in research work.

Training and testing of CANFIS models:

The daily data of evapotranspiration was categorized into two major sets i.e. a training data set from 2007 to 2014 and testing data set from 2015 to 2016 of Indore region. The input pairs in the training data set were applied to the network of a selected architecture and training was performed using Gaussian and generalized bell membership functions for CANFIS models. Fig. 2. shows the step-by-step procedure followed in research work.

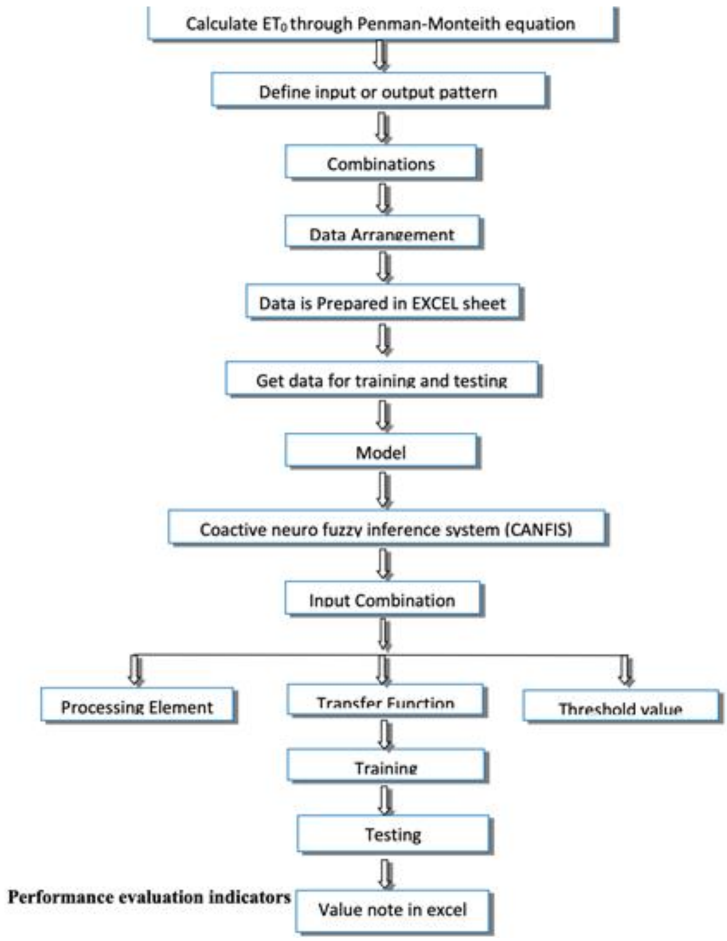

Fig. 2 Flow chart of research Work.

Performance evaluation indicators:

The performance of CANFIS model was evaluated using statistical indices such as root mean square error (RMSE), coefficient of correlation (r) and coefficient of efficiency (CE) for comparison between observed and predicted values. The equations are given by equation 11 to 13:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N} (X_{oi} - X_{pi})^2} \qquad \text{Eq. (11)}$$

$$r = \frac{\sum_{i=1}^{N} (X_{oi} - \overline{X_o})(Y_{pi} - \overline{Y_p})}{\sqrt{\sum_{i=1}^{N} (X_{oi} - \overline{X_o})^2 \sum_{i=1}^{N} (Y_{pi} - \overline{Y_p})^2}} \qquad \text{Eq. (12)}$$

$$\mathrm{CE} = \left[1 - \frac{\sum_{i=1}^{N} (X_{oi} - Y_{pi})^2}{\sum_{i=1}^{N} (X_{oi} - \overline{X_o})^2}\right] \qquad \text{Eq. (13)}$$

Where, Xoi and Xpi are the observed and predicted values for ith dataset and N are the total number of observations. respectively. $\overline{Xo}$ And $\overline{Yp}$ are the mean of observed (ET_0o) and predicted (ET_0p) values,

Results and discussions

The developed CANFIS model was implemented with different combinations of input variables and compared with traditional empirical method. For CANFIS modelling daily climatic data such as maximum temperature, minimum temperature, rainfall, wind speed and sunshine hours were considered as inputs variable while output from FAO PM Method has been taken as target. The simulation potential of CANFIS models for evapotranspiration was determine using various combinations of input variables. The values of statistical indices for the estimation of ET_0 using the selected CANFIS models during training and testing are presented in Table 1. On the basis of lower RMSE, higher CE and 'r' values in the testing phase, the CANFIS-9 model was found best performing model. The comparison of observed (ET_o) and predicted (ET_{0p}) values by CANFIS models during testing period were compared using ET_{0o} graph and scatter plot as shown in Figure 3 to 7.

Table 1. Statistical indices for selected CANFIS evapotranspiration models during training and testing phase for Indore.

S. No	Input Combinations	Model No.	Structure	Training			Testing		
				RMSE	r	CE	RMSE	r	CE
1.	$ET_0 = f(R, W)$	CANFIS-9	(2-3-1)	1.6402	0.8646	0.7482	0.5234	0.9941	0.9712
2.	$ET_0 = f(R)$	CANFIS-2	(1-5-1)	1.6538	0.8624	0.744	0.5316	0.9942	0.9702
3.	$ET_0 = f(T, S, W)$	CANFIS-12	(3-2-1)	1.6372	0.8659	0.7491	0.566	0.9868	0.9663
4.	$ET_0 = f(T, R)$	CANFIS-5	(2-2-1)	1.6308	0.8666	0.7511	0.6292	0.9837	0.9583
5.	$ET_0 = f(R, S)$	CANFIS-8	(2-5-1)	1.6213	0.8681	0.754	0.6275	0.9884	0.9585
6.	$ET_0 = f(R, S, W)$	CANFIS-14	(3-2-1)	1.6555	0.8621	0.7435	0.6312	0.9881	0.9581
7.	$ET_0 = f(T, R, W)$	CANFIS-13	(3-2-1)	1.6554	0.8619	0.7435	0.6675	0.9863	0.9531
8.	$ET_0 = f(T, R, S, W)$	CANFIS-15	(4-2-1)	1.6366	0.8652	0.7493	0.6696	0.9831	0.9528
9.	$ET_0 = f(R, S)$	CANFIS-8	(2-4-1)	1.6389	0.865	0.7486	0.6801	0.9835	0.9513
10.	$ET_0 = f(S, W)$	CANFIS-10	(2-4-1)	0.6921	0.9819	0.9496	0.6921	0.9819	0.9496

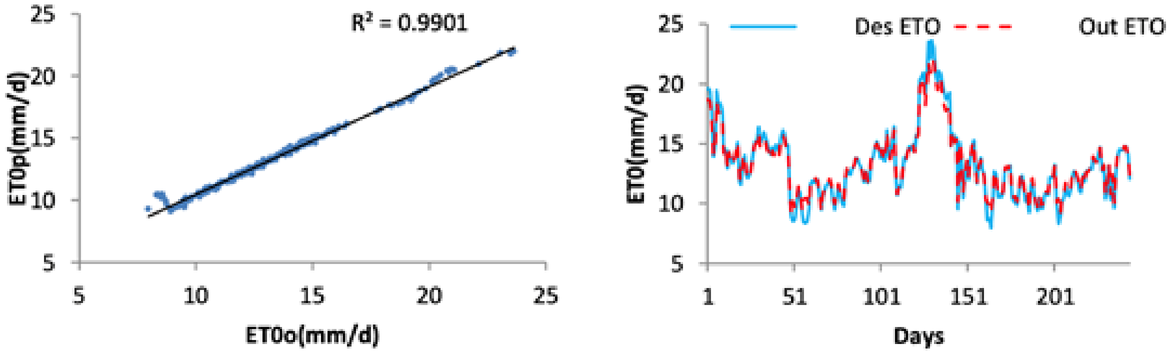

Fig. 3 Comparison of observed (ET0o) and predicted (ET0p) evapotranspiration and corresponding scatter plot in testing period by CANFIS model 9, neuron 3 during the testing period.

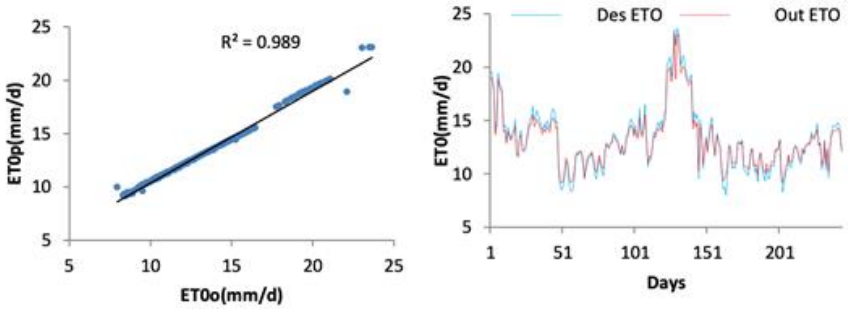

Fig. 4 Comparison of observed (ET0o) and predicted (ET0p) evapotranspiration and corresponding scatter plot in testing period by CANFIS model 2, neuron 5 during the testing period.

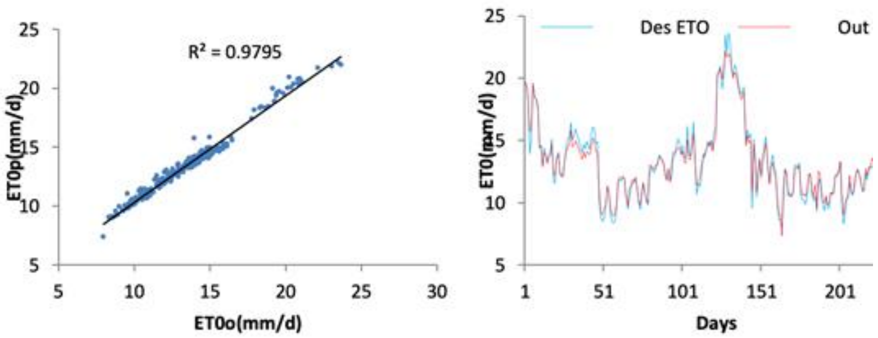

Fig. 5 Comparison of observed (ET0o) and predicted (ET0p) evapotranspiration and corresponding scatter plot in testing period by CANFIS model 12, neuron 2 during the testing period.

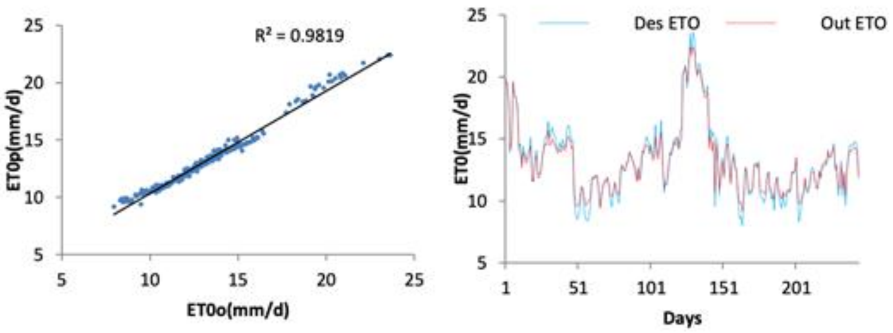

Fig. 6 Comparison of observed (ET0o) and predicted (ET0p) evapotranspiration and corresponding scatter plot in testing period by CANFIS model 5, neuron 2 during the testing period.

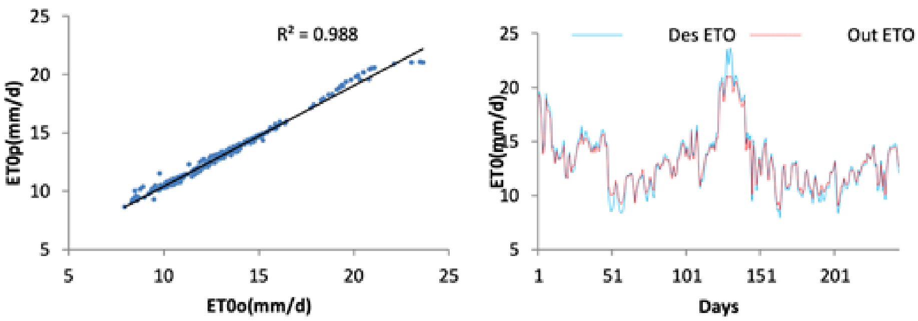

Fig. 7 Comparison of observed (ET_0o) and predicted (ET_0p) evapotranspiration and corresponding scatter plot in testing period by CANFIS model 8, neuron 5 during the testing period.

The scatter plot also indicates that the simulated and observed evapotranspiration values during testing period are well scattered around the best fit regression line. It is evident from the Table 1.1 the RMSE value for the selected CANFIS models varied from 1.640 to 0.692 and 0.523 to 0.692 during training and testing periods respectively whereas the CE values varied from 0.748 to 0.949 and 0.971 to 0.949 for training and testing periods respectively similarly the correlation coefficient (r) varied from 0.864 to 0.981 and 0.994 to 0.981 for training testing periods respectively. The increased values of CE and r by CANFIS model during testing period indicate good generalization capability of the selected CANFIS model. The RMSE, CE and 'r' values of CANFIS-9 model in the testing phase were found 0.5232, 0.9712 and 0.9941 respectively. It was concluded that the CANFIS-9 model was found superior over to other. Similar findings were also reported by Yazid et al., (2019, 2020) and Rekha Bai, (2018).

Conclusion

This research was conducted to compare the reference evapotranspiration using CANFIS model and Penman Monteith (P-M) equation. The performance of these methods was evaluated with sum of squared errors (SSE), root mean squared error (RMSE) and R2 values. The following conclusions are derived from this study:

- The RMSE value for the selected CANFIS models was varied from 1.640 to 0.692 and 0.523 to 0.692 during training and testing periods respectively whereas the CE values was varied from 0.748 to 0.949 and 0.971 to 0.949 for training and testing periods respectively similarly the correlation coefficient (r) was varied from 0.864 to 0.981 and 0.994 to 0.981 for training testing periods respectively.
- The RMSE, CE and 'r' values of CANFIS-9 model in the testing phase were found 0.5232, 0.9712 and 0.9941 respectively.
- Based on the statistical and visual comparisons, it can be concluded that the CANFIS-9 model was found superior over other model to estimate the reference evapotranspiration.
- The proposed CANFIS models may provide benefit to irrigation engineers and agriculturist for better estimation of reference evapotranspiration at the study stations with limited data availability.

References

12. Alves Sobrinho, T., D.B.B. Rodrigues, P.T.S. Oliveira, L.C.S. Rebucci, e C.A. Pertussatti. (2011). Estimativa da evapotranspiração de referência através de redes neurais artificiais. Revista Brasileira de Meteorologia 26:197-203.
13. Chang, F.J. and Chen, Y.C. (2004). A counter propagation fuzzy-neural network modelling approach to real time stream flow prediction. Journal of Hydrology, 245: 153-164.
14. Irmak, S., Allen, R. G., and Whitty, E. B. (2003). "Daily grass and alfalfa-reference evapotranspiration estimates and alfalfa-to-grass evapotranspiration ratios in Florida." J. Irrig. Drain. Eng., 129(5), 360–370.
15. Jensen, M. E., Burman, R. D. and Allen, R. G. (1990). Evapotranspiration and Irrigation Water Requirements. ASCE Manuals and Reports on Engineering Practice No. 70. American Society of Civil Engineers, New York.
16. Kisi, O. (2004). River flow modelling using artificial neural networks. ASCE Journal of Hydrologic Engineering, 9: 60-63.
17. Pandey, P. K., Dabral, P. P., & Pandey, V. (2016). Evaluation of reference evapotranspiration methods for the northeastern region of India. International Soil and Water Conservation Research, 4(1), 52-63.
18. Penman, H.L. (1948) Natural Evaporation from Open Water, Bare Soil and Grass. Proceedings of the Royal Society London A, 194, 120-145. http://dx.doi.org/10.1098/rspa.1948.0037.
19. Raziei, T., and L.S. Pereira. (2013). Estimation of ET_0 with Hargreaves Samani and FAOPM temperature methods for a wide range of climates in Iran. Agricultural and Water Management 121:1-18.
20. Rekha Bai. Estimation of daily reference evapotranspiration using Artificial Neural Network for different agroclimatic zone of Chhattisgarh. Trends in Biosciences (2018) Vol.11 No.8 pp.1905-1910 ref.10. http://trendsinbiosciencesjournal.com.
21. Yazid Tikhamarine, Anurag Malik, Anil Kumar, Doudja Souag-Gamane & Ozgur Kisi. Estimation of monthly reference evapotranspiration using novel hybrid machine learning approaches. Hydrological Sciences Journal. ISSN: 0262-6667 (Print) 2150-3435 (Online) Journal homepage, 2019: https://www.tandfonline.com/loi/thsj20.
22. Yazid Tikhamarine1 & Anurag Malik2 & Doudja Souag-Gamane1 & Ozgur Kisi3. Artificial intelligence models versus empirical equations for modelling monthly reference evapotranspiration. Environmental Science and Pollution Research, 2020.
23. Yildirim Dalkiliç, Umut Okkan, Nesrin Baykan. Comparison of Different Ann Approaches in Daily Pan Evaporation Prediction. Journal of Water Resource and Protection Vol.6 No.4, March 2014.

Citation

Singh, M., Singh, R., Singh, V., Rathod, T.P. (2024). Co-active neuro fuzzy inference system for estimating reference evapotranspiration over Indore district of Madhya Pradesh. In: Dandabathula, G., Bera, A.K., Rao, S.S., Srivastav, S.K. (Eds.), Proceedings of the 43rd INCA International Conference, Jodhpur, 06–08 November 2023, pp. 213–222, ISBN 978-93-341-2277-0.