Modelling of Reference Evapotranspiration for Semi-arid Climates Using Artificial Neural Network

by

Adeeba Ayaz, Sharath Chandra, Pratik Mandlecha, Shaik Rehana

in

Water and Energy Management in India

Report No: IIIT/TR/2021/-1

Centre for Spatial Informatics
International Institute of Information Technology
Hyderabad - 500 032, INDIA
March 2021

Modelling of Reference Evapotranspiration for Semi-arid Climates Using Artificial Neural Network

Adeeba Ayaz, Sharath Chandra, Pratik Mandlecha, and Rehana Shaik

Abstract Reference Evapotranspiration (ET_0) is one of the prominent hydrologic variables affecting water and energy balances and critical factors for crop water requirements and irrigation scheduling. Evapotranspiration is a complex hydrological variable defined by various climatic variables. Various empirical formulations have been developed to estimate ET_0 depending upon the availability of meteorological variables. Such empirical formulations are region-specific and are for particular climatic conditions. In this context, mathematical models have emerged as simple and readily implementable for the estimation of ET_0 with measured meteorological parameters as independent variables. Such data-driven models can be valuable to predict ET_0 when climate data is insufficient. The present study compared various empirical models and data-driven algorithms to predict ET_0 using various climate variables. Artificial neural networks (ANN) were adopted to estimate reference ET_0. Four empirical methods Penman-Monteith, Hargreaves, Turc, and Priestley-Taylor were used to estimate ET_0 at a daily time scale. Dataset consists of daily meteorological data over a period of 51 years (1965–2015) for Hyderabad, the largest city of the Indian state, Telangana, with semi-arid climate. The input variables for the ANN model consist of maximum and minimum air temperatures, relative humidity, solar radiation, and wind speed. The Penman-Monteith method was considered as the standard method to compare the ANN and various empirical models of ET_0. ANN model was trained and tested with climate variables as input variables and various empirical models as reference models. The most influencing climate variables on ET_0 were found in the order of temperature, solar radiation, wind speed, and relative humidity based on correlation coefficients. These variables have formed as the basis to choose different datasets to train over ANN model. Validation has been carried out using the coefficient of determination (R^2) which is obtained for the training (1965–2000) and testing period (2001–2015) period as 0.97 and 0.96 respectively. Temperature and radiation-based models of Turc and Priestley-Taylor methods can be used to estimate ET_0 when all other climate variables are not available as they

A. Ayaz · S. Chandra · P. Mandlecha · R. Shaik (✉)
Lab for Spatial Informatics, International Institute of Information Technology, Hyderabad 500032, Telangana, India
e-mail: rehana.s@iiit.ac.in

M. Majumder and G. D. Kale (eds.), *Water and Energy Management in India*,
https://doi.org/10.1007/978-3-030-66683-5_7

also correlate well with the Penman-Monteith method. Advancement towards artificial intelligence techniques in water resources engineering has motivated to simulate reference ET_0 using limited meteorological variables to produce accurate results. Such data-driven algorithms developed based on standard empirical models can be implemented for prediction with limited climate data.

Keywords Artificial neural network · Penman-Monteith · Evapotranspiration · Turc · Priestley-Taylor · Hargreaves

1 Introduction

Evapotranspiration (ET) is one of the most critical components of the hydrological water cycle affecting terrestrial water-energy balances. Actual evapotranspiration, potential evapotranspiration, and reference evapotranspiration are significant types of evapotranspiration. Actual Evapotranspiration (AET) is a significant component of the water balance and utilized generally in fields such as agronomy, hydrology, climatology, meteorology, ecology, and environmental sciences (Chiew and McMahon 2002; Liu et al. 2018; Peng et al. 2019; Tasumi 2019). Two more closely related types of evapotranspiration are potential evapotranspiration (PET) and reference evapotranspiration (ET_0). Although both PET and ET_0 provide estimates of atmospheric evaporative demand, they are based on different ideas, concepts, application fields and have different equations that can help to differentiate the terms. However, many researchers have treated PET and ET_0 as identical concepts and used similar equations for their estimation (Allen & Food and Agriculture Organization of the United Nations 1998; Irmak and Haman 2003a, b; Yates 1997; Zhang et al. 2017) The first idea of PET was proposed by Thornthwaite (1948) and that core idea with improvements are being used now. The PET equations were classified as mass-transfer, temperature, and radiation-based, while the reference ET_0 equations were classified as temperature, radiation, and pan-evaporation based (Chiew and McMahon 2002). PET has been applied mostly in hydrology, meteorology, and climatology. Whereas, the ET_0 has been applied mostly in agronomy, agriculture, irrigation, and ecology. The accurate estimation of ET_0 is essential in irrigation planning, scheduling, hydrological balance studies, and watershed hydrology (Feng et al. 2018; Yao et al. 2018). It has a broader significance in numerous fields of research including crop yield simulation, optimization of water lost, management and irrigation system design, water usage improvement in agriculture, and hydrologic water balance.

The ET_0 can be estimated based on energy balance and water vapour mass flux transfer methodologies (Rehana et al. 2020). Various methods to estimate reference ET_0 have been developed and are being utilized, depending upon the availability of meteorological variables. Empirical models for ET_0 estimation, i.e., statistical functions of approximation between meteorological variables and values, can overcome the difficulties associated with data availability for ET_0 estimation (Magliulo et al. 2003; Naoum and Tsanis 2003a, b). Among these Priestley-Taylor (Hargreaves and

Samani 1985; Penman 1948; Priestley and Taylor 1972) are well-established models. These empirical models vary in terms of solar radiation, temperature considering the physical processes of radiation and transport characteristics of natural surfaces.

The modified Penman-Monteith 56 equation has been recommended for the calculation of ET_0 and calibration of other equations by various international organizations such as the United Nations Food and Agriculture Organization (FAO) and World Meteorological Organization (Allen & Food and Agriculture Organization of the United Nations 1998; Walter et al. 2001). The Penman-Monteith equation has two critical advantages. First, it can be used in a wide variety of environments and climate scenarios without the need for any local calibrations because of its physical basis. Second, it is a well-documented method that has been validated using lysimeters under a wide range of climate conditions (Landeras et al. 2008). The main drawback of this equation is that it requires data on a large number of climate variables that are unavailable in many regions. An empirical model such as The Priestley-Taylor equation (Priestley and Taylor 1972) can estimate regional monthly ET_0 provided that the adjustment factor is adapted to different site conditions (Castellvi et al. 2001).

Nevertheless, the superiority of the Penman-Monteith method over the Priestley Taylor equation has recently been demonstrated (Alexandris et al. 2006) carried out surface polynomial regression analysis using hourly solar radiation, air temperature, and relative humidity (RH) to estimate ET_0. A much simpler alternative is the Thornthwaite scheme (Thornthwaite 1948) as it requires the only temperature as input data. However, this approach has been found to underestimate ET_0 under arid conditions and overestimate in a humid climate (Pereira and Pruitt 2004). The Hargreaves and Samani equation are an empirical approximation of the ET_0 calculation based on temperature and extraterrestrial radiation data (Gafurov et al. 2018).

Empirical models can be the best choice for estimating the ET_0 given the availability of meteorological variables. However, for ungauged basins where meteorological data is insufficient data-driven algorithms have proven to be valuable tools. Such data-driven models work with various climate factors as input variables and ET_0 estimates as reference models. In the past decades, there has been a widespread interest in the application of data-driven modelling and machine learning techniques in the field of water resources and hydrology (Kumar et al. 2020; Rehana 2019). In this context, the Artificial Neural Network (ANN) has been a widely applied machine learning algorithm in water resources engineering, including evapotranspiration (Kumar et al. 2020). ANNs are mathematical models whose architecture is inspired by biological neural networks and are highly appropriate for the modelling of nonlinear processes and are being used to predict and forecast water variables in the last decades and have been successfully used in hydrological processes, water resources management, water quality prediction and reservoir operation (Antonopoulos and Antonopoulos 2017). In recent years, ANN algorithms have been applied in the field of ET_0 estimation. Kumar et al. (2002) developed ANN models for the estimation of ET_0 and found that the ANNs could predict ET_0 better than the conventional empirical methods. More recently, Kisi (2007) investigated the modelling of ET_0 using ANNs with the Levenberg-Marquardt training algorithm and inferred that ANNs could be employed successfully in modelling ET_0 from available climate data. Jain et al. (2015) interpreted the physical meanings of ANNs for ET_0 estimation. Some of them utilized the comparable climatic data required for the application of the FAO Penman-Monteith method (Kumar et al. 2002; Odhiambo et al. 2001a, b; Trajkovic 2005). These researchers reported that the ANN can anticipate ET_0 ever better than the FAO Penman-Monteith conventional method. Sudheer et al. (2003) and Zanetti et al. (2007) simplified the input variables, and ET_0 was evaluated as a function of air temperature, extraterrestrial solar radiation, and daylight hours. Chauhan and Shrivastava (2009) compared the performance of four climate-based methods and Artificial Neural Networks (ANNs) for estimation of ET_0 when input climatic parameters are insufficient to apply the FAO Penman-Monteith method. They concluded that ANN models performed better than climatic methods. Suryavanshi et al. (2014) examined the trend in temperature and potential evapotranspiration over the Betwa basin, India. Sonali and Nagesh Kumar (2016) analysed the trend of maximum and minimum temperature of annual, monthly, winter, pre-monsoon, monsoon, and post-monsoon. The studies were carried out for three-time slots 1901–2003, 1948–2003, and 1970–2003, for India as a whole and seven homogeneous regions of India. Authors considered the effect of serial correlation, trend detection analysis while applying the Mann-Kendall test, Sen's slope estimator, and other non-parametric methods. Bandyopadhyay et al. (2020) have carried out the trend analysis of ET_0 using the Mann-Kendall trend test for India. The authors indicated that the leading cause of the rising trends of ET is due to an increase in relative humidity and a decrease in wind speed for the study duration. In another study, Rahimikhoob (2010) applied the ANN technique to estimate ET_0 based on air temperature data under humid subtropical conditions on the southern coast of the Caspian Sea situated in the north of Iran. The study showed that ANN successfully estimated the daily ET_0 better than the Hargreaves classical equation. Adamala (2018) made a comparison of developed models with the artificial neural network models and also with the linear and wavelet regression and conventional methods to estimate evapotranspiration using temperature-based generalized wavelet-neural network models. Estimation of the ET_0 of Punjab was done based machine learning models and was compared in predicting daily ET_0 with the performance of the Deep Learning model and was compared to Penman-Monteith model. The Generalized Linear Model (GLM), Random Forest (RF), and Gradient-Boosting Machine (GBM) models were also used in the study as various machine learning algorithms and concluded that the deep learning model performed better than the considered models for training, validation and testing sets. Pal and Deswal (2009) and Saggi and Jain (2019) investigated the different data-driven based regression approaches to model daily ET_0 using four inputs, including solar radiation, average air temperature, average relative humidity, and average wind speed. Results from their study suggested that the different data-driven and machine learning models could successfully be employed in modelling the ET_0.

The present study made efforts to implement the ANN model for the estimation and prediction of ET_0 in a semi-arid climate of India. The objectives of this study are to (1) develop ANN models with available climate factors for ET_0 estimation using long-term meteorological data; (2) to assess the applicability and validity of different ET_0

methods such as Penman-Monteith, Priestley-Taylor, Hargreaves, and Turc methods. Since the maximum and minimum air temperature and relative humidity records are more readily available around the globe, these records with extraterrestrial radiation are being used as input in the above models for the estimation of ET_0. Extraterrestrial radiation reflects the seasonality of ET_0 and can theoretically be calculated as a function of the local latitude and Julian data, according to the equations presented by Allen & Food and Agriculture Organization of the United Nations (1998). Therefore, for the models suggested in this study, only temperature and relative humidity are the parameters that require monitoring. Here, the FAO Penman-Monteith method was used as a substitute for measured ET_0 data, as this is the standard procedure used when no measured lysimeter data is available (Irmak and Haman 2003a, b). The study has been implemented on the semi-arid climate conditions of Hyderabad, Telangana, India.

2 Data and Case Study

The area under study was Hyderabad, the largest city of the Indian state of Telangana which lies between latitude 17.3850° N and longitude 78.4867° E located on the Deccan Plateau in the northern part of South India and covers an area of 650 km^2 at an elevation of 542 m. Based on the Koppen climate classification, the climate is tropical wet and dry bordering on a hot semi-arid, with an average annual precipitation of about 171 mm (Fig 1).

Daily meteorological data were obtained from January 1965 through December 2015 (51 years) (612 months) from weather station situated in Professor Jayashankar Telangana State Agricultural University, Rajendranagar Mandal, Hyderabad, Telangana. The annual average weather data of the meteorological station is presented in Table 1. Five monthly meteorological variables were recorded including: (1) mean maximum air temperature (T_x °C); (2) mean minimum air temperature (T_n °C); (3) mean relative humidity (RH %); (4) mean wind speed (U_2 m s^{-1}); (5) solar radiation (R_s, MJ m^{-2} d^{-1}) and (6) Evapotranspiration (ET_0 mm/day). Measurements were made at the height of 2 m (air temperature and relative humidity) and 10 m (wind speed) above the soil surface. Wind speed data at 2 m (U_2) were obtained from those taken at 10 m using the log-wind profile equation.

3 Materials and Methods

This study mainly implemented the ANN model for estimation and prediction of ET_o. Evapotranspiration is calculated using the following methods with the limited meteorological parameters by considering Penman-Monteith method as standard method as it requires radiation, wind speed, relative humidity and temperature.

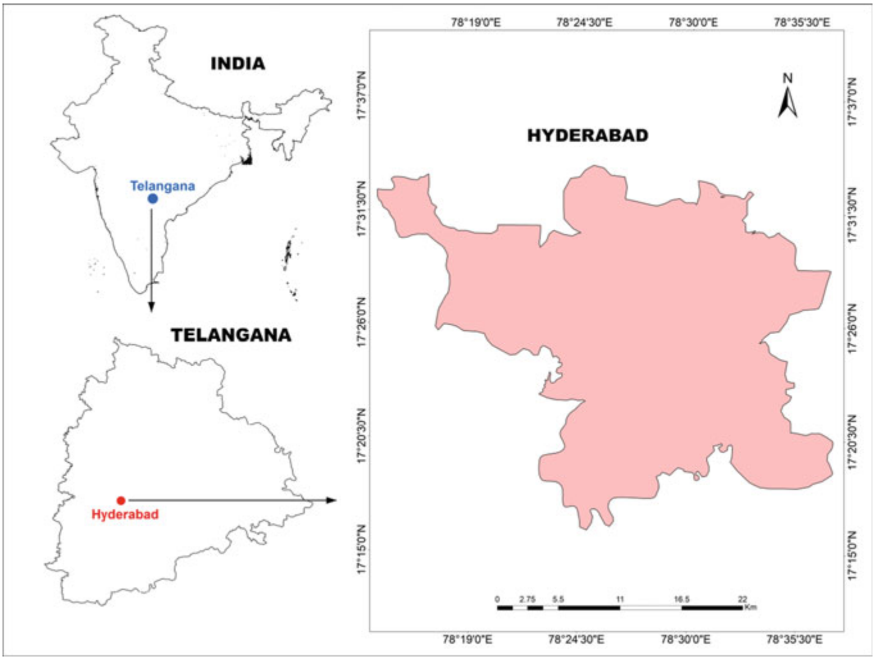


Fig. 1 Case study: Hyderabad, Telangana, India

Table 1 Statistical parameters of available meteorological variables and ET_0 at Hyderabad

Parameters	T_x	T_n	RH_{mean}	U_2	R_s	ET_0
Maximum	45.5	33.0	139	36.0	14.45	13.16
Minimum	17.6	5.0	10	0	4.00	0.48
Mean	32.37	19.88	60.70	4.69	9.32	3.76
Standard deviation	4.1	4.79	14.93	4.62	2.44	1.72

3.1 FAO-56 Penman-Monteith Method

The Penman-Monteith (Penman 1948) method was recommended by the FAO. It is calculated on a daily time scale. The formulation can be expressed as follows:

$$ET_0 = \frac{0.408\Delta(Rn - G) + g\left(\frac{900}{T+273}\right)U_2(e_s - e_a)}{\Delta + g(1 + 0.34U_2)} \tag{1}$$

where, Rn is net radiation (MJ m^{-2} d^{-1}), G is soil heat flux (MJ m^{-2} d^{-1}), T is average temperature at 2 m height (°C), U_2 is wind speed measured at 2 m height (m s^{-1}), ($e_s - e_a$) is pressure deficit for measurement at 2 m height (k Pa), Δ is slope vapor pressure curve (k pa °C^{-1}), g is psychrometric constant (k pa °C^{-1}), 900 is coefficient for the reference crop (l J^{-1} kg K d^{-1}), 0.34 is wind coefficient for the reference crop (s m^{-1}).

3.2 *Turc Method*

Turc (1961) method estimates ET_0 based on mean temperature and solar radiation on daily time scale. The formulation can be expressed as follows

$$ET_0 = 0.013\frac{Tm}{Tm + 15}(23.88Rs + 50) \tag{2}$$

where Tm is mean temperature (°C), solar radiation (Rs) is [0.25 + 0.5 (n/N)] Ra, Ra is extraterrestrial radiation (mm day^{-1}), n is actual hours of bright sunshine (h), N is maximum possible hours of sunshine (h).

3.3 *Priestly and Taylor Method*

Priestley and Taylor (1972) method is calculated using net radiation and latent heat of vaporization on a daily time scale. The formulation can be expressed as follows

$$ET_0 = A\left(\frac{D}{D + g}\right)\left(\frac{Rn - G}{L}\right) \tag{3}$$

$$D = \frac{4098\left[0.6108exp\left(\frac{17.27*Tm}{Tm+237.3}\right)\right]}{(Tm + 237.3)^2} \tag{4}$$

where D is slope vapour pressure curve (k pa °C^{-1}), g is psychrometric constant (k pa °C^{-1}), Rn is the net radiation at crop surface (MJ m^{-2} d^{-1}), A is a calibration constant 1.26, L is the latent heat of vaporization and can be considered as 2.45 (MJ/kg) which is constant.

3.4 *Hargreaves Method*

Hargreaves (1972) method which was modified in 1985 (Hargreaves 1983) estimates ET_0 based on temperature and radiation is calculated on a daily time scale. The formulation can be expressed as follows:

$$ET_0 = 0.0023Ra\left(\frac{Tm}{2} + 17.8\right)(Td^{0.5}) \tag{5}$$

where, Td is difference between maximum temperature and min temperature (°C), Tm is mean temperature (°C), Ra is extra-terrestrial radiation (mm day^{-1})

3.5 Artificial Neural Networks (ANN)

Artificial Neural Networks has gained much attention in hydrology for the prediction of various conceptual processes such as rainfall-runoff, streamflows, water quality and ground water modelling, etc. (Kurian et al. 2020). ANN is a computational model inspired by networks of biological neurons, wherein the neurons compute output values from inputs (Heddam and Kisi 2018). It learns from its past experience and errors in a nonlinear parallel processing manner (Gupta and Singh 2011). ANNs are fully connected neural nets that consist of an input layer, hidden layers (multiple or single), output layer. Each node can be considered as a neuron. The neuron is the basic calculating entity that computes from a number of inputs and delivers one output compared with a threshold value and turned on (fired). The computational processing is done by internal structural arrangement consisting of hidden layers that utilize the backpropagation and feed-forward mechanism to deliver output close to accuracy. Fully connected neural nets are those where each node in a layer is connected to every other node in the next layer (right). Each node takes the weighted sum of its inputs which then passes through a nonlinear activation function (like RELU, sigmoid, tanh, etc.), which then becomes the input of other nodes in the next layer (Rumelhart et al. 1986). In Eq. 6 the function, f, represents the activation function and w is the weight matrix, X is the set of input vectors (Fig. 2).

$$\mathrm{Z} = \mathrm{f}\,(\mathrm{x.\ w}) = \mathrm{f}\left(\sum_{i=1}^{n} x_i w_i\right) \quad x \in d_{1\times n,} w \in d_{n\times 1,} z \in d_{1\times 1,} \tag{6}$$

The present study used a feed-forward backpropagation neural network. The weights are initially randomly assigned. The train: test spilt on the dataset is 7:3. A forward pass is performed for every training data using the current weights, and the output is calculated for each node. At the last node, the final output is acquired, and the error is calculated with a loss function. Now, a backward pass is performed to calculate the contribution of each node in error calculated. The error is propagated to every single node using backpropagation. Once, the contribution of each node has been calculated the weights are adjusted accordingly using gradient descent. The present study used gradient descent with momentum and adaptive linear regression. The procedure is repeated until the loss function gives an error which is less than the threshold value and the weights and bias of the required network are thus obtained. Thus, the model converges, and a definite result can be obtained for any type of testing dataset.

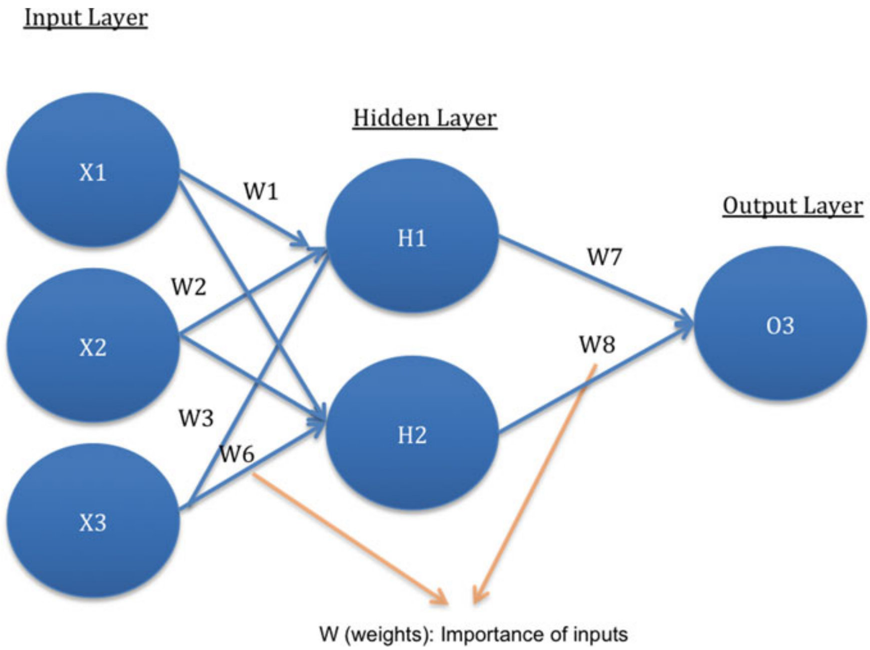


Fig. 2 Structure of ANN used for training a model with hidden layer and weights and the output layer showing a feed-forward pass

3.6 Performance Metrics

The model's performance criteria were validated using different standard statistical methods. In this study coefficient of Determination (R^2) (Krause et al. 2005) Root Mean Square Error (RMSE) (Legates and McCabe 1999) the Mean Absolute Error (MAE) were used as validation criteria. The equations for these methods are as follows:

$$\mathbf{R^2} = \left(\frac{\sum_{i=1}^{n} \left(O_i - O_{av_9}\right)\left(S_i - S_{av_9}\right)}{\sqrt{\sum_{i=1}^{n} \left(O_i - O_{av_9}\right)^2} \sqrt{\sum_{i=1}^{n} \left(S_i - S_{av_9}\right)^2}} \right)^2 \tag{7}$$

$$\mathbf{RMSE} = \sqrt{\sum_{i=1}^{n} (O_i - S_i)^2} \tag{8}$$

$$\mathbf{MAE} = \frac{1}{N} \sum_{i=1}^{n} (O_i - S_i) \tag{9}$$

where O is the observed values (the reference evapotranspiration), S is the simulated values by the other methods, and O_{avg} and S_{avg} are the mean observed and computed values, respectively.

4 Results and Discussions

The climate variables considered for estimating daily ET_0 using ANN and Penman-Monteith methods were the daily maximum temperature, minimum temperature, relative humidity, solar radiation and wind speed. Similarly, for the Turc and corresponding ANN model, the input variable considered is the mean temperature and solar radiation. Whereas, for the Hargreaves and corresponding ANN model, the input variables considered are maximum and minimum temperatures and solar radiation. Furthermore, for the Priestly Taylor method, the input variables used in ANN are temperature, solar radiation, and relative humidity. The ANN model used in the present study is Multi-Layered Perceptron (MLP) imported from the sci-kit-learn library in python. The study used three hidden layers with several neurons same as the number of features or parameters, i.e. 6 (maximum air temperature, minimum temperature, relative humidity, solar radiation), and ran the model for 500 iterations. Convergence was obtained for the datasets of all the four empirical methods. The prediction values have been calculated by fitting the test data on the trained model. As the number of meteorological variables for each empirical method is different, therefore, for each empirical model, an ANN model was trained, and results were tested. Input vector has the features considered in each method (Penman, Hargreaves, Turc, and Priestley-Taylor) 3 hidden layers have been used for each method and the output vector is the expected reference evapotranspiration value calculated from each method. The optimal node number in the hidden layer of the network was determined using a trial and error method by considering the MAE, RMSE, and R^2 values from a test sample. In this study, ANNs were trained for 500 epochs with one to 6 nodes in the hidden layer and mentioned before, statistical parameters were calculated using only the whole test data set after each training run. The training period considered is from 1965 to 2000 and the testing period considered from 2001 to 2015. The validity and efficiency of the model can be seen when the training dataset is fit on the trained model, and high accuracy and minimal values of RMSE were obtained. The performance of each empirical model corresponding with the ANN model in terms of R^2, RMSE, and MAE was listed in Table 2. Figures 3, 4, 5, 6 and 7 shows the comparison between daily ET_0 values form empirical models of Penman-Monteith, Priestley-Taylor, Hargreaves, Turc, and ANN methodologies for training and testing datasets.

Figure 3a shows the comparison of ET_0 daily values predicted by the ANN model versus the ET_0 values of the Penman-Monteith method for both testing and training periods. A good correlation was observed with R^2 values as higher than 0.95, RMSE as 0.03, and MAE as 0.009 between the Penman-Monteith method and ANN for the testing period. The trained and tested ANN model performs very well when compared

Table 2 Statistical summary of testing and training period for ANN

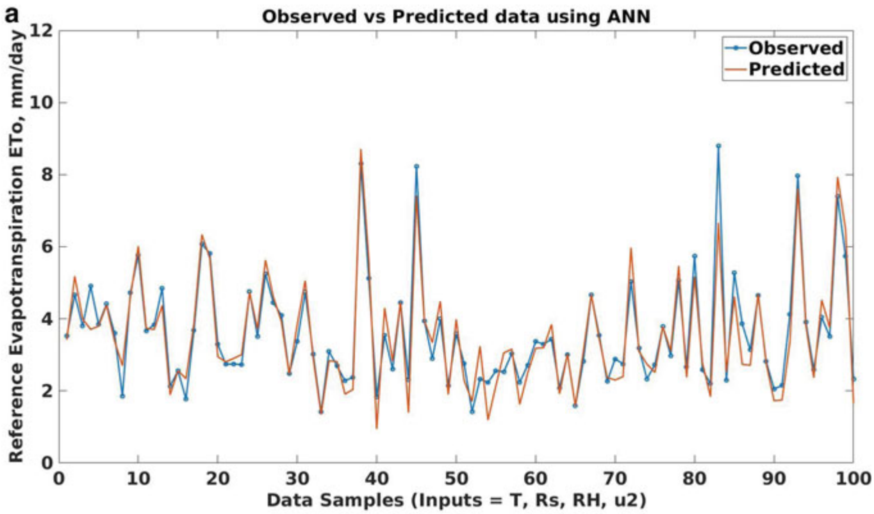
Empirical methods	Artificial neural network (training)			Artificial neural network (testing)		
	R^2	RMSE	MAE	R^2	RMSE	MAE
Penman-Monteith	0.97	0.02	0.008	0.96	0.03	0.009
Turc	0.96	0.03	0.007	0.95	0.04	0.012
Hargreaves	0.94	0.05	0.015	0.94	0.06	0.017
Priestley Taylor	0.91	0.10	0.022	0.92	0.12	0.025

with Penman-Monteith estimates. The comparison shows that neither overestimation nor underestimation was produced in the range of the values studied. This verifies that the ANN models can be used to estimate ET_0 values. Thus, compared to all other empirical models, the Penman-Monteith has been predicted well with the data-driven algorithm of ANN. It can be noted that, as the Penman-Monteith method accounts for all climate variables into modelling, such accuracies were expected to be comparable to other empirical models.

Furthermore, the present study tried to understand the sensitivity and dependency of each meteorological variable on the modelled ET_0 using the Penman-Monteith model. The study plotted the scatter plots between each climate variable and ET_0 modelled based on the Penman-Monteith method, as shown in Fig. 8.

As shown in Fig. 8, the temperature and solar radiation followed by relative humidity have the most substantial influence on ET_0 estimations based on the Penman-Monteith model. Therefore, ANN models that were derived based on temperature, solar radiation, wind speed and relative humidity as input and the ET_0as output variables. The ANN model results, when using (T, RH, R_s and U_2) from the four essential meteorological variables as input, seldom show the same values of coefficient of determination (R^2). These results prove that the relative humidity has a very low contribution to ET_0 when using ANNs models. The overall accuracies of most models were found to be similar to each other.

Furthermore, the results of the ANN can be significantly influenced by the number of input data which can lead to significant error and deviation. On the other hand, lowering the number of neurons in the input layer to three or even two can give us satisfactory results in the estimation of the reference evapotranspiration. The most critical inputs for accurate estimation of ET_0 using an ANN are temperature and radiation data (Jain et al. 2015). The results showed that the proper choice of ANN architecture allows not only error minimization but also maximizes the relationship between the dependent and the independent variables. The results of the study reveal that temperature and solar radiation as the most influencing variables compared to relative humidity and wind speed for semi-arid climate conditions, as demonstrated in the present study. Given the intense data requirements for applying the Penman-Monteith model, the study employed ANN with minimum input variables such as temperature, and solar radiation. The trained and tested algorithms developed based on empirical models can be valuable tools to predict ET_0 for limited data case studies.



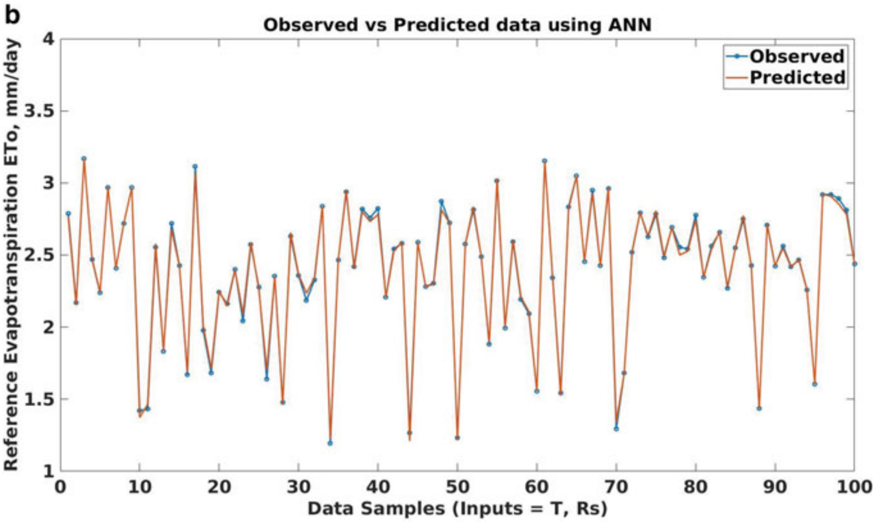
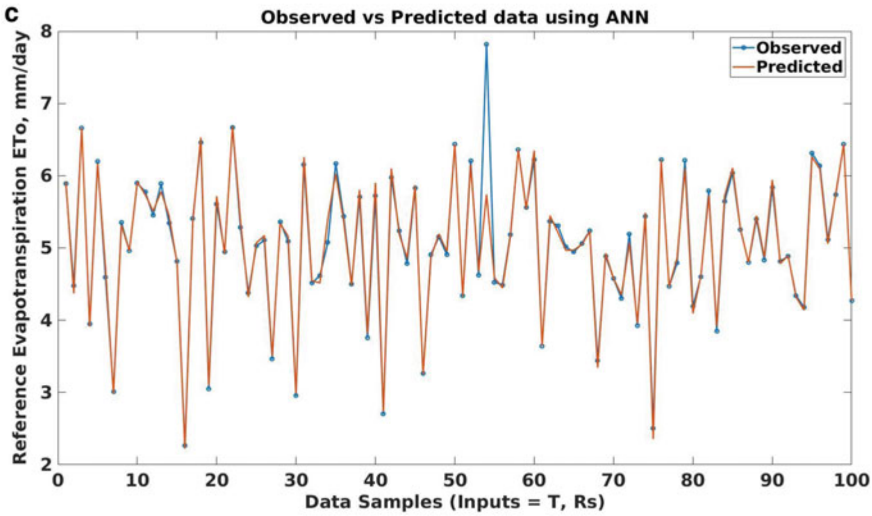


Fig. 3 **a** Variation in reference evapotranspiration (ET_0) from Penman-Monteith for the first 100 data points using ANN. **b** Variation in reference evapotranspiration (ET_0) from Turc for the first 100 data points using ANN. **c** Variation in reference evapotranspiration (ET_0) from Hargreaves for the first 100 data points using ANN. **d** Variation in reference evapotranspiration (ET_0) from Priestley-Taylor for the first 100 data points using ANN



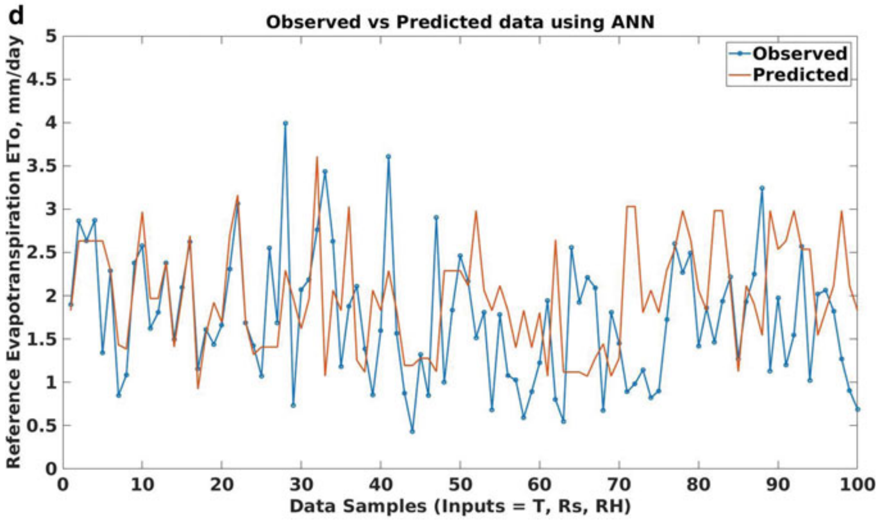


Fig. 3 (continued)

Analysing the sensitivity of each climate variable on ET_0 and testing the statistical dependencies, data pre-processing to acquire relevant information before the development of such data-driven algorithms is of most importance in the implementation. Analysis of compensating accuracies with the inclusion of limited climate input variables in the ET_0 estimates compared to standard empirical models can be a potential area of research.

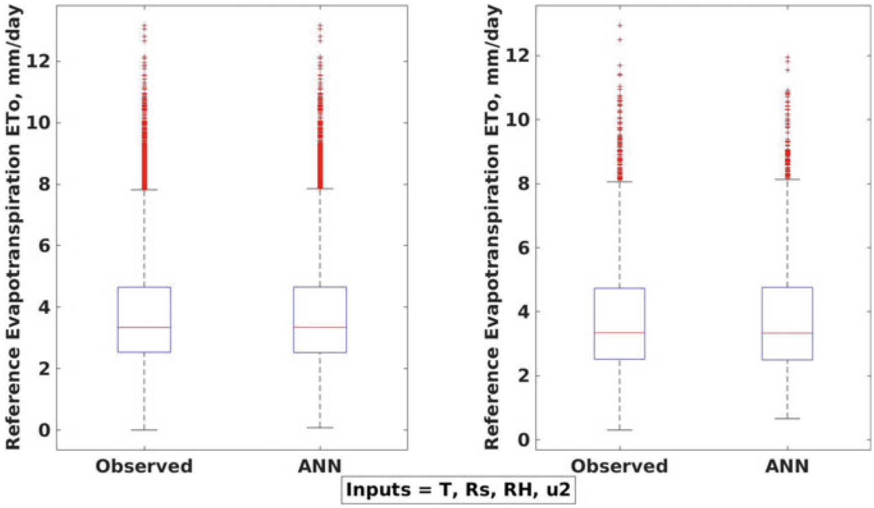


Fig. 4 Comparison of ET_0 predicted by ANN and Penman-Monteith method values for training and testing periods

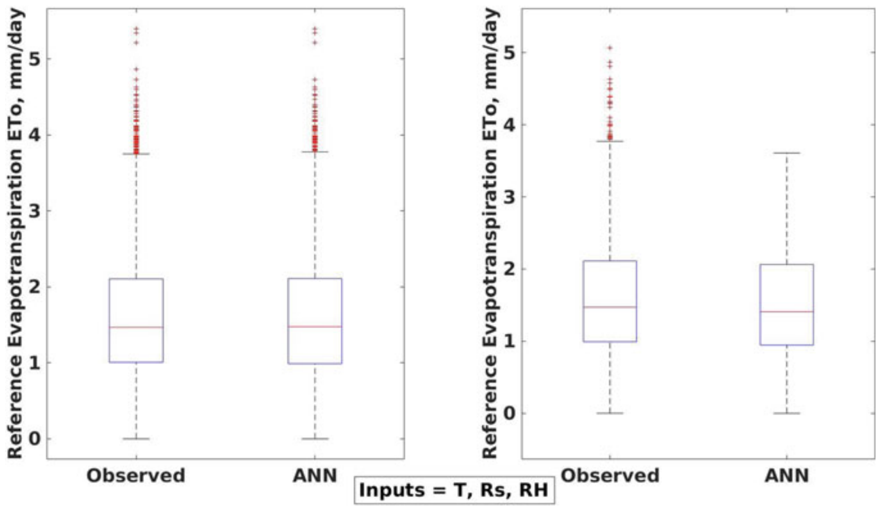


Fig. 5 Comparison of ET_0 predicted by ANN for Priestley-Taylor method values for training and testing periods

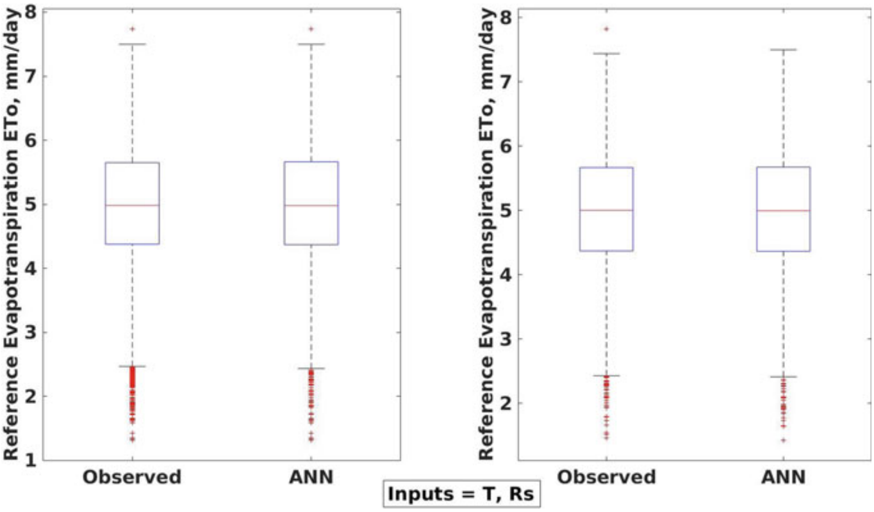


Fig. 6 Comparison of ET_0 predicted by ANN for Hargreaves method values for training and testing periods

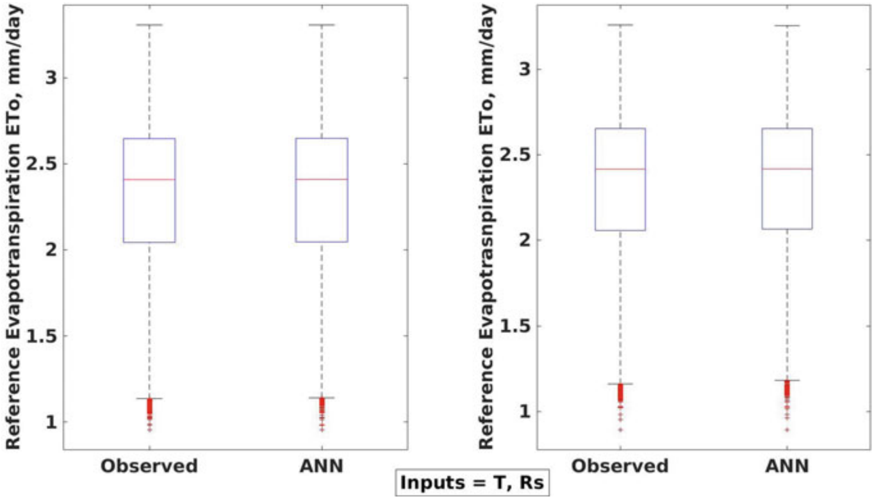


Fig. 7 Comparison of ET_0 predicted by ANN for Turc method values for training and testing periods

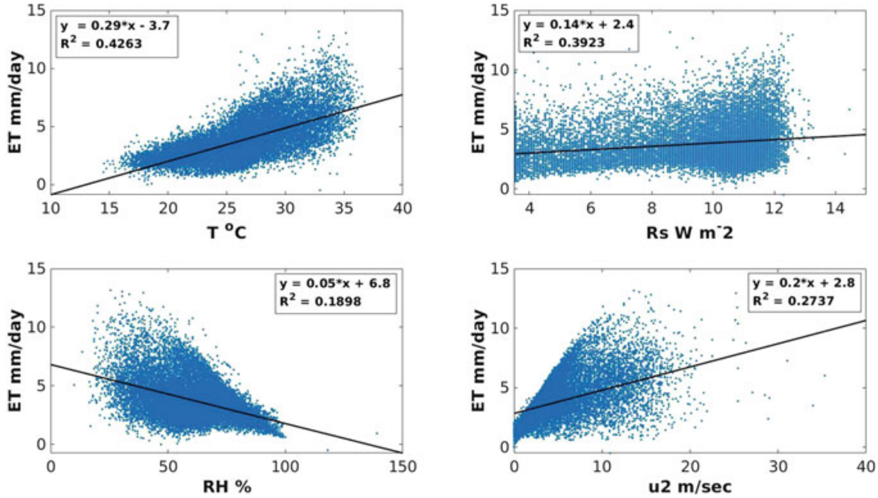


Fig. 8 Correlation of main meteorological parameters such as temperature, relative humidity, solar radiation, wind speed to ET_0

5 Conclusions

The daily reference evapotranspiration over semi-arid climatic conditions over Hyderabad, Telangana, India, were modelled using empirical and data-driven models. The Penman-Monteith model estimates of reference evapotranspiration were considered as standard reference models for various temperature and radiation-based empirical models and also for data-driven models. The daily reference evapotranspiration rates were estimated with ANN modelling technique using four input variables as maximum and minimum air temperatures, relative humidity, solar radiation, and wind speed; three input variables as average air temperature, relative humidity, and solar radiation; two input variables as temperature and solar radiation. The results were discussed with the results of alternative methods of ET_0 calculation, such as the combination-based method of Penman-Monteith, the radiation-based methods of Priestly-Taylor, the temperature-based methods of Hargreaves, and the Turc method. The correlation coefficient values suggest that temperature is the most important factor followed by solar radiation, wind speed, and relative humidity, respectively. ANN with all-climate variables as input was able to simulate ET_0 values estimated using the Penman-Monteith method. Temperature and solar radiation have a maximum correlation with ET_0 estimates of Penman-Monteith models as compared to relative humidity and wind speed. The Turc model uses temperature and solar radiation as input variables and high accuracy with the ANN model. Whereas, the relative humidity has the least correlation with the reference ET_0 estimates. The Priestly-Taylor model considers relative humidity, temperature, and solar radiation as input variables. Due to the lower dependency of relative humidity on the reference ET_0 estimates, the Priestly-Taylor model has lower accuracy with ANN compared to the Turc model. The study concludes that the empirical models work well with data-driven algorithms that consider the climate variables having high dependency with the standard reference ET_0 estimates. Such studies can be implemented for the development of data-driven models statistically dependent with reference model ET estimates. Further, it can be concluded that when a parameter or an input variable with a lower correlation is added to the set of features for training over ANN, the accuracy of prediction will be decreased. The results showed that ANN provides quite good agreement with the ET_0 obtained by the Penman-Monteith method. The study demonstrated that modelling of ET_0 through the use of the ANN technique gave better estimates that proved with their performance criterion, i.e. R^2 as 0.96. The study concludes that the performance of the model varies according to the number of inputs as well as the predicted time step. Overall, results are of significant practical use when limited climate data is available to estimate the reference evapotranspiration.

Acknowledgements The authors sincerely thank Professor Jayashankar Telangana State Agricultural University, Rajendranagar Mandal, Hyderabad, Telangana, for providing meteorological data. The research work presented in the paper is partly funded by Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India through Start-up Grant for Young Scientists (YSS) Project no. YSS/2015/002111 to Dr. S. Rehana.

References

Adamala, S. (2018). Temperature based generalized wavelet-neural network models to estimate evapotranspiration in India. *Information Processing in Agriculture, 5*(1), 149–155. https://doi.org/10.1016/j.inpa.2017.09.004.

Alexandris, S., Kerkides, P., & Liakatas, A. (2006). Daily reference evapotranspiration estimates by the "Copais" approach. *Agricultural Water Management, 82*(3), 371–386. https://doi.org/10.1016/j.agwat.2005.08.001.

Allen, R. G., & Food and Agriculture Organization of the United Nations. (Eds.). (1998). *Crop evapotranspiration: Guidelines for computing crop water requirements*. FAO irrigation and drainage paper. Rome: Food and Agriculture Organization of the United Nations.

Antonopoulos, V. Z., & Antonopoulos, A. V. (2017). Daily reference evapotranspiration estimates by artificial neural networks technique and empirical equations using limited input climate variables. *Computers and Electronics in Agriculture, 132,* 86–96. https://doi.org/10.1016/j.compag.2016.11.011.

Bandyopadhyay, N., Bhuiyan, C., & Saha, A. K. (2020). Drought mitigation: Critical analysis and proposal for a new drought policy with special reference to Gujarat (India). *Progress in Disaster Science, 5,* 100049. https://doi.org/10.1016/j.pdisas.2019.100049.

Castellvi, F., Stockle, C. O., Perez, P. J., & Ibañez, M. (2001). Comparison of methods for applying the Priestley-Taylor equation at a regional scale: Priestley-Taylor equation. *Hydrological Processes, 15*(9), 1609–1620. https://doi.org/10.1002/hyp.227.

Chauhan, S., & Shrivastava, R. K. (2009). Performance evaluation of reference evapotranspiration estimation using climate based methods and artificial neural networks. *Water Resources Management, 23*(5), 825–837. https://doi.org/10.1007/s11269-008-9301-5.

Chiew, F. H. S., & McMahon, T. A. (2002). Modelling the impacts of climate change on Australian streamflow. *Hydrological Processes, 16*(6), 1235–1245. https://doi.org/10.1002/hyp.1059.

Feng, W., Shum, C., Zhong, M., & Pan, Y. (2018). Groundwater storage changes in china from satellite gravity: An overview. *Remote Sensing, 10*(5), 674. https://doi.org/10.3390/rs10050674.

Gafurov, Z., Eltazarov, S., Akramov, B., Yuldashev, T., Djumaboev, K., & Anarbekov, O. (2018). Modifying Hargreaves-Samani equation for estimating reference evapotranspiration in dryland regions of Amudarya River Basin. *Agricultural Science, 09*(10), 1354–1368. https://doi.org/10.4236/as.2018.910094.

Gupta, A. K., & Singh, Y. P. (2011). Analysis of Hamming network and MAXNET of neural network method in the string recognition. *International Conference on Communication Systems and Network Technologies, 2011,* 38–42. https://doi.org/10.1109/CSNT.2011.15.

Hargreaves, D. J., & Bolton, N. (1972). Selecting creativity tests for use in research. *British Journal of Psychology, 63*(3), 451–462. https://doi.org/10.1111/j.2044-8295.1972.tb01295.x.

Hargreaves, G. H. (1983). Closure to "*Estimating potential evapotranspiration*" by Hargreaves, G. H., & Samani, Z. A. (September, 1982). *Journal of Irrigation and Drainage Engineering, 109*(3), 343–344. https://doi.org/10.1061/(asce)0733-9437(1983)109:3(343).

Hargreaves, G. H., & Samani, Z. A. (1985). Reference crop evapotranspiration from temperature. *Applied Engineering in Agriculture, 1*(2), 96–99.

Heddam, S., & Kisi, O. (2018). Modelling daily dissolved oxygen concentration using least square support vector machine, multivariate adaptive regression splines and M5 model tree. *Journal of Hydrology, 559,* 499–509. https://doi.org/10.1016/j.jhydrol.2018.02.061.

Irmak, S., & Haman, D. Z. (2003a). Evapotranspiration: Potential or reference. *IFAS Extension, ABE, 343,* 1–3.

Irmak, S., & Haman, D. Z. (2003b). Evaluation of five methods for estimating class A pan evaporation in a humid climate. *HortTechnology, 13*(3), 500–508. https://doi.org/10.21273/HORTTECH.13.3.0500.

Jain, V. K., Pandey, R. P., Jain, M. K., & Byun, H.-R. (2015). Comparison of drought indices for appraisal of drought characteristics in the Ken River Basin. *Weather and Climate Extremes, 8,* 1–11. https://doi.org/10.1016/j.wace.2015.05.002.

Kisi, O. (2007). Evapotranspiration modelling from climatic data using a neural computing technique. *Hydrological Processes, 21*(14), 1925–1934. https://doi.org/10.1002/hyp.6403.

Krause, P., Boyle, D. P., & Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences, 5,* 89–97. https://doi.org/10.5194/adgeo-5-89-2005.

Kumar, M., Raghuwanshi, N. S., Singh, R., Wallender, W. W., & Pruitt, W. O. (2002). Estimating evapotranspiration using artificial neural network. *Journal of Irrigation and Drainage Engineering, 128*(4), 224–233. https://doi.org/10.1061/(ASCE)0733-9437(2002)128:4(224).

Kumar, N., Shankar, V., & Poddar, A. (2020). Investigating the effect of limited climatic data on evapotranspiration-based numerical modeling of soil moisture dynamics in the unsaturated root zone: A case study for potato crop. *Modeling Earth Systems and Environment, 6*(4), 2433–2449. https://doi.org/10.1007/s40808-020-00824-8.

Kurian, C., Sudheer, K. P., Vema, V. K., & Sahoo, D. (2020). Effective flood forecasting at higher lead times through hybrid modelling framework. *Journal of Hydrology, 587,* 124945. https://doi.org/10.1016/j.jhydrol.2020.124945.

Landeras, G., Ortiz-Barredo, A., & López, J. J. (2008). Comparison of artificial neural network models and empirical and semi-empirical equations for daily reference evapotranspiration estimation in the Basque Country (Northern Spain). *Agricultural Water Management, 95*(5), 553–565. https://doi.org/10.1016/j.agwat.2007.12.011.

Legates, D. R., & McCabe, G. J., Jr. (1999). Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation. *Water Resources Research, 35*(1), 233–241. https://doi.org/10.1029/1998WR900018.

Liu, Y., Hejazi, M., Li, H., Zhang, X., & Leng, G. (2018). A hydrological emulator for global applications—HE v1.0.0. *Geoscientific Model Development, 11*(3), 1077–1092. https://doi.org/10.5194/gmd-11-1077-2018.

Magliulo, V., Andria, R., & Rana, G. (2003). Use of the modified atmometer to estimate reference evapotranspiration in Mediterranean environments. *Agricultural Water Management, 63*(1), 1–14. https://doi.org/10.1016/S0378-3774(03)00098-2.

Naoum, S., & Tsanis, I. K. (2003a). Hydro informatics in evapotranspiration estimation neural networks. *Water Resources Management, 18,* 143–161.

Naoum, S., & Tsanis, I. (2003b). Hydroinformatics in evapotranspiration estimation. *Environmental Modelling and Software, 18*(3), 261–271. https://doi.org/10.1016/S1364-8152(02)00076-2.

Odhiambo, H. O., Ong, C. K., Deans, J. D., Wilson, J., Khan, A. A. H., & Sprent, J. I. (2001a). *Plant and Soil, 235*(2), 221–233. https://doi.org/10.1023/A:1011959805622.

Odhiambo, L. O., Yoder, R. E., Yoder, D. C., & Hines, J. W. (2001b). Optimization of fuzzy evapotranspiration model through neural training with input–output examples. *Transactions of the ASAE, 44*(6), 1625.

Pal, M., & Deswal, S. (2009). M5 model tree based modelling of reference evapotranspiration. *Hydrological Processes: An International Journal, 23*(10), 1437–1443. https://doi.org/10.1002/hyp.7266.

Peng, L., Zeng, Z., Wei, Z., Chen, A., Wood, E. F., & Sheffield, J. (2019). Determinants of the ratio of actual to potential evapotranspiration. *Global Change Biology, 25*(4), 1326–1343. https://doi.org/10.1111/gcb.14577.

Penman, H. L. (1948). Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London*: *Series Mathematical, and Physical Sciences, 193*(1032), 120–146. Retrieved from http://www.jstor.org/stable/98151.

Pereira, A. R., & Pruitt, W. O. (2004). Adaptation of the Thornthwaite scheme for estimating daily reference evapotranspiration. *Agricultural Water Management, 66*(3), 251–257. https://doi.org/10.1016/j.agwat.2003.11.003.

Priestley, C. H. B., & Taylor, R. J. (1972). On the assessment of surface heat flux and evaporation using large-scale parameters. *Monthly Weather Review, 100*(2), 81–92. American Meteorological Society. https://doi.org/10.1175/1520-0493(1972)100%3c0081:otaosh%3e2.3.co;2.

Rahimikhoob, A. (2010). Estimating global solar radiation using artificial neural network and air temperature data in a semi-arid environment. *Renewable Energy, 35*(9), 2131–2135. https://doi.org/10.1016/j.renene.2010.01.029.

Rajashekar, N. R. *Estimation of reference evapotranspiration using empirical methods and CROPWAT model for Rangareddy district*. Professor Jayashankar Telangana State Agricultural University.

Rehana, S. (2019). River water temperature modelling under climate change using support vector regression. In S. K. Singh & C. T. Dhanya (Eds.), *Hydrology in a changing world* (Springer water) (pp. 171–183). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-02197-9_8.

Rehana, S., Sireesha Naidu, G., Monish, N. T., & Sowjanya, U. (2020). Modeling hydro-climatic changes of evapotranspiration over a semi-arid river basin of India. *Journal of Water and Climate Change*.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986) Learning representations by back-propagating errors. *Nature, 323*(6088), 533–536. Nature Publishing Group. https://doi.org/10.1038/323533a0.

Saggi, M. K., & Jain, S. (2019). Reference evapotranspiration estimation and modeling of the Punjab Northern India using deep learning. *Computers and Electronics in Agriculture, 156,* 387–398. https://doi.org/10.1016/j.compag.2018.11.031.

Sonali, P., & Nagesh Kumar, D. (2016). Spatio-temporal variability of temperature and potential evapotranspiration over India. *Journal of Water and Climate Change, 7*(4), 810–822. https://doi.org/10.2166/wcc.2016.230.

Sudheer, K. P., Gosain, A. K., & Ramasastri, K. S. (2003). Estimating actual evapotranspiration from limited climatic data using neural computing technique. *Journal of Irrigation and Drainage Engineering, 129*(3), 214–218. https://doi.org/10.1061/(ASCE)0733-9437(2003)129:3(214).

Suryavanshi, S., Pandey, A., Chaube, U. C., & Joshi, N. (2014). Long-term historic changes in climatic variables of Betwa Basin, India. *Theoretical and Applied Climatology, 117*(3–4), 403–418. https://doi.org/10.1007/s00704-013-1013-y.

Tasumi, M. (2019). Estimating evapotranspiration using METRIC model and Landsat data for better understandings of regional hydrology in the western Urmia Lake Basin. *Agricultural Water Management, 226,* 105805. https://doi.org/10.1016/j.agwat.2019.105805.

Thornthwaite, C. W. (1948). An approach toward a rational classification of climate. *Geographical Review, 38*(1), 55. https://doi.org/10.2307/210739.

Trajkovic, S. (2005). Temperature-based approaches for estimating reference evapotranspiration. *Journal of Irrigation and Drainage Engineering ASCE, 131*(4), 316–323. https://doi.org/10.1061/(ASCE)0733-9437(2005)131:4(316).

Turc, L. (1961). Estimation of irrigation water requirements, potential evapotranspiration: A simple climatic formula evolved up to date. *Annals of Agronomy, 12*, 13–49.

Walter, I. A., Allen, R. G., Elliott, R., Jensen, M. E., Itenfisu, D., Mecham, B., et al. (2001). ASCE's standardized reference evapotranspiration equation. *Watershed Management and Operations Management, 2000*, 1–11. Presented at the Watershed Management and Operations Management Conferences 2000, Fort Collins, Colorado, United States: American Society of Civil Engineers. https://doi.org/10.1061/40499(2000)126.

Yao, J., Zhao, Y., Chen, Y., Yu, X., & Zhang, R. (2018). Multi-scale assessments of droughts: A case study in Xinjiang, China. *Science of the Total Environment, 630,* 444–452. https://doi.org/10.1016/j.scitotenv.2018.02.200.

Yates, D. N. (1997). Approaches to continental scale runoff for integrated assessment models. *Journal of Hydrology, 201*(1–4), 289–310. https://doi.org/10.1016/S0022-1694(97)00044-9.

Zanetti, S. S., Sousa, E. F., Oliveira, V. P., Almeida, F. T., & Bernardo, S. (2007). Estimating evapotranspiration using artificial neural network and minimum climatological data. *Journal of Irrigation and Drainage Engineering, 133*(2), 83–89. https://doi.org/10.1061/(ASCE)0733-9437(2007)133:2(83).

Zhang, D., Chen, P., Zhang, Q., & Li, X. (2017). Copula-based probability of concurrent hydrological drought in the Poyang lake-catchment-river system (China) from 1960 to 2013. *Journal of Hydrology, 553,* 773–784. https://doi.org/10.1016/j.jhydrol.2017.08.046.