Vol. 48, No. 1, pp 65-69, 2020
Indian Journal of Soil Conservation

Intercomparision of ANN, regression and climate based models for estimation of reference evapotranspiration

D.D. Khedkar[1,*], P.K. Singh[2] and Mahesh Kothari[2]

[1]Interfaculty Department of Irrigation Water Management, MPKV, Rahuri, Maharashtra; [2]Department of Soil Water Engineering, MPUAT, Udaipur, Rajasthan.

**Corresponding author:*
E-mail: devidaskhedkar@gmail.com (D.D. Khedkar)

ARTICLE INFO

Article history:
Received : November, 2019
Revised : April, 2020
Accepted : April, 2020

ABSTRACT

The present study investigates the applicability of linear regression (LR) and artificial neural network (ANN) models for estimating reference evapotranspiration (ET_0) and their intercomparison with climate-based models on the basis of limited data availability in semi-arid environment of Solapur, Maharashtra, India. The eight climate based methods *viz.,* Soil Conservation Service Blaney-Criddle, Thornthwaite, Hargreaves-Samani, Pan evaporation, Jensen-Haise, Priestly-Taylor, Turc, and Radiation were compared with Penman-Monteith (P-M) method for estimation of ET_0. The input combinations for all LR and ANN models were decided on the basis of climatic parameters required for selected climate-based methods. These are *viz.,* Model 1 (evaporation), Model 2 (T_{max} and T_{min}), Model 3 (T_{max} and T_{min} and Sun Shine Hours - SSH), Model 4 (T_{max}, T_{min}, RH_{max}, RH_{min} and SSH). The accuracies of the models were evaluated by using statistical criteria such as: coefficient of determination (R^2), index of agreement d(IA), root mean square error (RMSE), mean absolute error (MAE), mean absolute percentage error (MAPE) and coefficient efficiency (CE), and ranking were assigned of models. All LR and ANN models showed satisfactory performance in development and validation stage and can be accepted to predict ET_0 values. The overall comparison of climate-based, LR and ANN models were carried out using the data from the year 1980 to 2014. The average weekly ET_0 values were estimated using climate-based, LR and ANN models and compared with those of P-M method. It was observed that ANN4 secured first rank and exhibited overall best performance with R^2 = 0.895, d(IA) = 0.972, RMSE = 0.508, MAE = 0.391, MAPE = 7.931 and CE = 0.894 followed by ANN3, LR4, LR3, ANN2, LR2, ANN1, and LR1, while all climate-based methods showed poorer performance than ANN and LR models. It was inferred that all LR models showed satisfactory performance for estimation of ET_0, however the performance has improved marginally with corresponding ANN models. Based on the overall results it was recommended that all ANN models can be used for the prediction of ET_0 followed by all LR models as per data availability and simplicity of users for Solapur region.

Key words:
Reference evapotranspiration
Climate-based models
Regression analysis
ANN
Statistical measures

1. INTRODUCTION

Evapotranspiration (ET) is a key parameter in agro-meteorological studies and water resources management. It includes evaporation of water from land surfaces and transpiration by vegetation, and is essential for estimating irrigation water requirements (Allen *et al.,* 1998; Sahoo *et al.,* 2010; Panigrahi, 2013). Different reference evapotranspiration (ET_0) methods exist for direct measurement of ET_0 *viz.,* Thornthwaite, 1948; Doorenbos and Pruitt, 1977; Hargreaves and Samani, 1985 and Penman–Monteith, FAO 56 (called as P-M model) which was introduced as a standard model to estimate ET_0 (Allen *et al.*, 1998). The major limitation to P-M model is that it requires more meteorological parameters such as maximum and minimum air temperatures and relative humidity values, wind speed and sunshine hours. Hence, its utility is limited in data-sparse areas (Singh *et al.,* 2016). All the weather data needed to solve the P-M model are often incomplete and/or not available in many of the developing countries like India. Sibale *et al.* (2016) evaluated three ET_0 estimation methods at Dapoli, Maharashtra and found that the pan evaporation method performed reasonably well with P-M 56 model. Irmak *et al.* (2003) and Yoder *et al.* (2005) noted that the Turc radiation-based method showed promising results in the south-eastern United States under data-limited conditions. However there is a need to standardize existing climate-base methods on location basis.

Another alternative is the application of mathematical models like artificial neural networks (ANNs). In recent years, ANNs have been applied in the field of ET estimation as its complex non-linear phenomenon. Kumar *et al.* (2002) developed ANN models for the estimation of ET and found that the ANNs could predict ET better than the conventional method. Some ANNs have been compared with empirical equations for ET_0 estimation, with the results showing better performance for the former (Khoob, 2008). Due to some difficulties in working with ANN approaches, some researchers have used fast, simple and straight forward statistical methods such as regression models with measured meteorological parameters as independent variables for ET estimation. Hence in this study it was planned to compare various ANNs, multiple linear regression (MLR) and existing ET equations with P-M method using limited data under the climatic conditions of Solapur region of Maharashtra, India.

2. MATERIALS AND METHODS

Study Area and Climate Dataset

The climatic data for this study of Solapur station were obtained from the IMD, Pune and SAU, Rahuri, India. The data comprised the maximum and minimum air temperatures, and relative humidity values, wind speed, sunshine hours and pan evaporation for the period of 1980-2014. It should be noted that the average weekly values of the weather data were used for the analysis.

Climate Based Methods

The climate-based ET_0 estimation methods were selected as per the ranking given by Jensen *et al.* (1990) and on the basis of minimum data requirement. The selected climate-based methods are listed in Table 1.

Linear Regression (LR)

Regression analysis is commonly used to describe quantitative relationships between a response variable and one or more explanatory variables. It is the function of a linear equation, *i.e.* straight-line, in the form:

$$Y = a + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5$$

Where, Y is the dependent variable and x_1, x_2, x_3, x_4, x_5 are the independent variables, a is interceptor and b_1, b_2, b_3, b_4, b_5 are the partial regression coefficients. In the present study, the ET_0 estimated using P-M method was considered as dependent variable while meteorological parameters were assumed as independent ones for the development of LR models. The basis for combinations of independent variables is presented in Table 2. The SPSS 21.0 software was used to develop statistically optimal models of simple and MLR for estimation of ET_0 values.

Artificial Neural Network (ANN)

A neural network is an artificial intelligence technique that mimics a function of the human brain. It is a capable of identifying complex non-linear relationships between input and output data sets which are difficult to describe using physical equations. Most ANNs have three or more layers: an input layer, which is used to present data to the network; an output layer, which is used to produce an appropriate response to the given input; and one or more intermediate layers, which are used to act as a collection of feature detectors. In the present study, feed forward back propaga-

Table: 1
Selected climate based methods

S.No.	ET_0 methods
1.	Penman-Monteith (P-M) (Allen *et al.,* 1998).
2.	SCS Blaney-Criddle (SCS BC) (Blaney and Criddle, 1962)
3.	Thornthwaite (THOR) (Thornthwait, 1948)
4.	Hargreaves-Samani (H-S) (Hargreaves and Samani, 1985)
5.	Pan evaporation (PAN) (Doorenbos and Pruitt, 1977)
6.	Jensen-Haise (J-H) (Jensen and Haise, 1963)
7.	Priestly-Taylor (P-T) (Priestley and Taylor, 1972)
8.	Turc (TURC) (Turc, 1962)
9.	Radiation (RAD)(Doorenbos and Pruitt, 1977)

Table: 2
Combinations of climatic variables in LR / ANN modelling on the basis of data requirement of climate base methods

S.No.	Variables	Climate based methods							
		PAN	SCS BC	THOR	H-S	J-H	P-T	RAD	TURC
1.	T_{max}	-	Y	Y	Y	Y	Y	Y	Y
2.	T_{min}	-	Y	Y	Y	Y	Y	Y	Y
3.	RH_{max}	-	-	-	-	-	-	-	Y
4.	RH_{min}	-	-	-	-	-	-	-	Y
5.	Wind Speed	-	-	-	-	-	-	-	-
6.	SSH	-	-	-	-	Y	Y	Y	Y
7.	E_{pan}	Y	-	-	-	-	-	-	-

Combinations of independent/input variables for models (LR/ANN)

S.No.	Variables	Model 1	Model 2	Model 3	Model 4
1.	T_{max}	-	Y	Y	Y
2.	T_{min}	-	Y	Y	Y
3.	RH_{max}	-	-	-	Y
4.	RH_{min}	-	-	-	Y
5.	Wind Speed	-	-	-	-
6.	SSH	-	-	Y	Y
7.	E_{pan}	Y	-	-	-

tion type of network was selected for the development of architecture for ET_0 modelling application with *Learngdm* (gradient descent with momentum weight and Bias learning function) adaption learning function for this application. The selections of combinations of inputs of each neuron were based on the meteorological inputs of the ET_0 equations and are presented in Table 2. The number of nodes in the output layer depends on the number of target variables. In this study, the output layer will be single node corresponding to ET_0 estimated using P-M method and neurons in the hidden layer varied alternatively from 3 to 19. The most widely used non-linear activation function *i.e.* log sigmoid for the hidden layer and linear transfer function in output layer were selected. The Neural Network/Data Manager graphical user interface (nntool) of Matlab 7.12.0 (R2011a) software was used to develop ANN architecture for ET_0 modelling.

Combinations of Meteorological Parameters in LR and ANN Modelling

The input combinations for LR and ANN models were decided on the basis of climatic parameters required for the previously selected climate-base methods and are presented in Table 2. Model 1 represents single climatic parameter (*i.e.* evaporation) required for Pan evaporation method. Model 2 consists two parameters (*i.e.* T_{max} and T_{min}) that are required for SCS BC, THOR and H-S methods. Model 3 consist three parameters (*i.e.* T_{max}, T_{min} and SSH) which are required by J-H, P-T and RAD methods. Model 4 consists five parameters (*i.e.* T_{max}, T_{min}, RH_{max}, RH_{min} and SSH) that are required for TURC method. A similar kind of basis for formulating combinations of parameters in modelling of ET_0 was adopted by Tabari *et al.* (2012).

Comparison of Climate-based, LR and ANN Models with Limited Data

The values of ET_0 were estimated from climate-based, LR and ANN models and were compared with the standard P-M method for standardization or calibration. Different statistical measures such as Coefficient of Determination (R^2), Index of Agreement d(IA), Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Coefficient Efficiency (CE) were worked out to test the performance of all models (Singh *et al.,* 2018 and Panigrahi, 2011). The criteria for ranking of models were based on the following conditions (Pandey *et al.,* 2016):

(a) In case of R^2, d(IA) and CE values tending towards one.
(b) In case of RMSE, MAE and MAPE values tending towards zero.

The selection of best method or performance of the methods was decided on the summation of all ranks obtained from all statistical measures. Based on the total ranks obtained by each method, the overall ranking was decided.

3. RESULTS AND DISCUSSION

Development and Validation of LR and ANN Models

In order to compare the performance of ET_0estimation models, it required development of LR and ANN models and comparison of their performance with existing climate based models. Out of total data period (1980-2014) for Solapur station, 80% data (1980-2007) were used for development of model and 20% data (2008-2014) were used for their validation.

LR Models

The statistical criteria for defining and validating LR models with their mathematical expressions are presented in Table 3. During development stage, all LR models showed the performance in the sequence of LR4, LR3, LR2 and LR1. It indicated that all LR models performed satisfactorily and showed marginal difference in their performances. Results also showed that with increase in the number of independent variables, the performance of models increased.

In validation stage, it was found that all models showed numerically at par results for each performance measures. It indicated that all LR models were validated satisfactorily and generalized for estimation of ET_0. Overall, the performance suggests that all LR models can predict ET_0 within acceptable range for Solapur station with reference to the availability of number of meteorological parameters. Most of researchers such as Reddy *et al.,* 2010; Tabari *et al.,* 2012; and Sriram and Rashmi, 2014 also found that LR models with varying independent variable can be adopted for the prediction of ET_0.

Table: 3
Statistical criteria for development and validation of LR models with their mathematical expressions for Solapur station

Model	Statistical Criterias					
	R^2	d(IA)	RMSE	MAE	MAPE	CE
Development Period (1980-2007)						
LR1	0.706	0.908	0.858	0.652	13.541	0.706
LR2	0.774	0.932	0.753	0.597	12.220	0.774
LR3	0.828	0.951	0.657	0.518	10.525	0.828
LR4	0.852	0.958	0.610	0.480	9.706	0.852
Validation Period (2008-2014)						
LR1	0.843	0.925	0.719	0.541	10.842	0.773
LR2	0.843	0.941	0.684	0.557	12.664	0.794
LR3	0.848	0.953	0.618	0.499	11.100	0.832
LR4	0.850	0.953	0.630	0.503	10.946	0.825

Models	Mathematical expressions
LR1	$ET_0 = 1.522+0.471E_{pan}$
LR2	$ET_0 = -7.490+0.346T_{max}+0.037T_{min}$
LR3	$ET_0 = -7.095+0.220T_{max}+0.153T_{min}+0.211SSH$
LR4	$ET_0 = -3.455+0.128T_{max}+0.213T_{min}-0.005RH_{max}-0.024RH_{min}+ 0.165SSH$

ANN Models

Table 4 shows the statistical performance of best fit ANN models with limited data (ANN1 to ANN4) during training and validation period with network of models for Solapur station.

During training mode, it was observed that ANN4 model showed best values of all performance measures, while ANN1 showed lower level performance among them. It was observed that all ANN models satisfied performance criteria well in development mode. Results clearly showed that the performance of model is directly related to the number of input parameters and are in with the findings of Huo *et al.* (2012). A very close difference was observed in the performances of all models during training and validation stages. The similar findings were obtained by other researchers (Kumar *et al.,* 2008; Kale *et al.,* 2013) for predicting ET_0 accurately.

Table: 4
Statistical criteria for development and validation of ANN models with their network for Solapur Station

Model	Network	Statistical Criteria					
		R^2	d(IA)	RMSE	MAE	MAPE	CE
Development Period (1980-2007)							
ANN1	1-3-1	0.719	0.913	0.838	0.637	13.255	0.719
ANN2	2-3-1	0.801	0.942	0.706	0.559	11.449	0.801
ANN3	3-17-1	0.863	0.961	0.587	0.458	9.193	0.862
ANN4	5-19-1	0.893	0.971	0.516	0.395	7.825	0.894
Validation Period (2008-2014)							
ANN1	1-3-1	0.826	0.914	0.748	0.569	11.526	0.754
ANN2	2-3-1	0.865	0.949	0.648	0.523	12.176	0.815
ANN3	3-17-1	0.878	0.963	0.545	0.428	9.567	0.869
ANN4	5-19-1	0.912	0.973	0.487	0.384	8.481	0.895

Comparison of Climate Based, LR and ANN Models

The overall comparison of climate-based, LR and ANN models were carried out on the basis of limited data availability for the entire data period (*i.e.* from 1980 to 2014). The ET_0 values of climate-based, LR and ANN models were determined and compared with P-M method. The results of performance measures for comparison of all models with P-M method for Solapur station are presented in Table 5 and ranking of models are tabulated in Table 6. It was also observed that all climate-based methods showed poorer performance than corresponding ANN and LR models.

Results showed that the proposed LR models can be adopted satisfactorily for the estimation of ET_0. The accuracy in ET_0 estimation may further be improved using corresponding ANN models for Solapur station. Because of the representation of explicit equations that can be easily interpreted and accessible, it was interpreted that the LR models are more advantageous over ANN models. Based on the overall results it was recommended that all ANN models can be used for prediction of ET_0 followed by all LR models as per data availability and simplicity of users for Solapur region.

4. CONCLUSIONS

In this paper, an attempt was made to select the best method for estimating ET_0 in the absence of the full weather data for P-M method application in a semi-arid environment of Solapur region Maharashtra, India. It was found that all LR and ANN models showed satisfactory performance and can be used to predict ET_0 values for Solapur region as per data availability. However, the prediction accuracy may slightly be better in ANN models than corresponding LR

Table: 5
Performance evaluation of ET_0 values of climate based LR and ANN models with those of P-M method for entire period (1980-2014)

S.No.	Models	Statistical criteria					
		R^2	d(IA)	RMSE	MAE	MAPE	CE
1	SCS BC	0.659	0.785	1.488	1.254	29.416	0.096
2	THOR	0.729	0.849	1.613	1.242	24.997	-0.063
3	H-S	0.794	0.899	0.849	0.705	15.821	0.705
4	PAN	0.721	0.908	1.038	0.796	16.836	0.56
5	J-H	0.762	0.826	1.504	1.286	27.723	0.076
6	P-T	0.699	0.775	1.224	0.905	15.901	0.388
7	TURC	0.656	0.437	3.413	3.194	62.667	-3.76
8	RAD	0.686	0.817	1.291	1.124	25.674	0.319
9	LR1	0.721	0.913	0.828	0.627	12.968	0.72
10	LR2	0.780	0.934	0.737	0.586	12.274	0.778
11	LR3	0.830	0.951	0.646	0.512	10.595	0.829
12	LR4	0.848	0.958	0.611	0.483	9.921	0.847
13	ANN1	0.728	0.914	0.818	0.621	12.872	0.727
14	ANN2	0.806	0.943	0.693	0.55	11.565	0.804
15	ANN3	0.865	0.962	0.576	0.45	9.227	0.864
16	ANN4	0.895	0.972	0.508	0.391	7.931	0.894

Table: 6
Ranking of all models as per the performance criteria for Solapur station

S.No.	Models	Ranking as per performance criteria							
		R^2	d(IA)	RMSE	MAE	MAPE	CE	Total	Overall Rank
1	SCS BC	15	14	13	14	15	13	84	15
2	THOR	9	11	15	13	12	15	75	12
3	H-S	6	10	9	9	9	9	52	9
4	PAN	12	9	10	10	11	10	62	10
5	J-H	8	12	14	15	14	14	77	14
6	P-T	13	15	11	11	10	11	71	11
7	TURC	16	16	16	16	16	16	96	16
8	RAD	14	13	12	12	13	12	76	13
9	LR1	11	8	8	8	8	8	51	8
10	LR2	7	6	6	6	6	6	37	6
11	LR3	4	4	4	4	4	4	24	4
12	LR4	3	3	3	3	3	3	18	3
13	ANN1	10	7	7	7	7	7	45	7
14	ANN2	5	5	5	5	5	5	30	5
15	ANN3	2	2	2	2	2	2	12	2
16	ANN4	1	1	1	1	1	1	6	1

models, whereas climate-based methods showed lower performance than corresponding LR and ANN models. Based on the overall results it was recommended that all ANN models can be used for predicting ET_0 followed by all LR models as per data availability and simplicity of users for Solapur region.

REFERENCES

Allen, R.G., Pereira, L.S., Raes, D. and Smith, M. 1998. *Crop Evapotranspiration, Guideline for Computing Crop Water Requirements*. FAO Irrigation and Drainage, Paper 56. FAO Rome, 299p.

Blaney, H.F. and Criddle. M. 1962. Determining consumptive use and irrigation water requirements. *USDA Tech. Bull*., 1275: 1359.

Doorenbos, J. and Pruitt, W.O. 1975. *Guidelines for Prediction of Crop Water Requirement*. FAO Irrigation and Drainage Paper No. 24, FAO Rome, Italy, 179p.

Hargreaves, G.L. and Samani, Z.A. 1985. Reference crop evapotranspiration from temperature. *Appl. Eng. Agric., Trans. ASAE,* 1(2): 96-99.

Huo, Z., S. Feng, S. and Dai, X. 2012. Artificial neural network models for reference evapotranspiration in an arid area of northwest China. *J. Arid Environ.,* 82: 81-90.

Irmak, S., Irmak, A., Allen, R.G. and Jones, J.W. 2003. Solar and net radiation-based equations to estimate reference evapotranspiration in humid climates. *J. Irrig. Drain. Eng. ASCE*, 129(5): 336–347.

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 Engineering, New York.

Jensen, M.E. and Haise, H.R. 1963. Estimating evapotranspiration from solar radiation. *J. Irrig. Drain. Eng. ASCE,* 89:15-41.

Kale M.U., Nagdeve, M.B. and Bagade, S.J. 2013. Estimation of evapotranspiration with ANN technique. *J. Indian Water Resour. Soc.,* 33(1): 23-29.

Khoob, J. 2008. Comparative study of Hargreaves's and artificial neural network's methodologies in estimating reference evapotranspiration in a semi-arid environment. *Irrig. Sci.,* 26: 253-259.

Kumar, M., Raghuwanshi, N.S., Singh, R., Wallender, W.W. and Pruitt, W.O. 2002. Estimating evapotranspiration using artificial neural network. *J. Irrig. Drain. Eng., ASCE,* 128: 224-233.

Kumar, M. and Pandey, V. 2008. Validation of different evapotranspiration models over semi-arid region of India. *Int. J. Math. Model Simul. Appl.,* 1(3): 321-327.

Pandey, P.K., Parmendra P., Dabral and Pandey, V. 2016. Evaluation of reference evapotranspiration methods for the north-eastern region of India. *Int. Soil Water Conserv. Res.,* 4: 52–63.

Panigrahi, B. 2011. *Irrigation Systems Engineering*. New India Publishing Agency, New Delhi, 428p.

Panigrahi, B. 2013. *A Handbook on Irrigation and Drainage*. New India Publishing Agency, New Delhi, 601p.

Priestley, C.H.B. and Taylor, R.J. 1972. On the assessment of the surface heat flux and evaporation using large scale parameters. *Monthly Weather Rev.,* 100: 81-92.

Reddy, K., Chandra Sekhar, J., Aruna and Mallikarjuna, P. 2010. Estimation of weekly reference evapotranspiration using linear regression and ANN models. *Proc. Int. Conf. Adv. Civil Eng.,* pp 83-88.

Sahoo, B.C., Panda, S.N. and Panigrahi, B. 2010. Modelling evaporative water loss from the on-farm reservoir with biological shading. *J. Hydrol. Engi.*, 15(7): 544-553.

Singh, M., Niwas, R., Khichar, M.L. and Godra, A.K. 2016. Trends and variability in potential evapotranspiration over north-west India. *J. Agrometeor.,* 18(2): 335-338.

Singh, M, Singh, J.P. and Singh, K.G. 2018. Development of mathematical models for predicting vapour pressure deficit inside a greenhouse from internal and external climate. *J. Agrometeor.,* 20(3): 238-241.

Sibale, D., Mane, M.S., Patil, S.T. and Ingle, P.M. 2016. Evaluation of three methods for estimating reference evapotranspiration at Dapoli, Maharashtra. *J. Agrometeor.,* 18(1): 157-158.

Sriramm, A.V. and Rashmi C.N. 2014. Estimation of Potential Evapotranspiration by Multiple Linear Regression Method. *IOSR J. Mechanical Civil Engi.,* 11(2IV): 65-70.

Tabari, H., Ozgur, K., Azadeh, E. and Hosseinzadeh, T.P. 2012. SVM, ANFIS, regression and climate based models for reference evapotranspiration modelling using limited climatic data in a semi-arid highland environment. *J. Hydrol.,* 444-445: 78-89.

Thornthwaite, C.W. 1948. An approach towards a rational classification of climate. *Geograph. Rev.,* 30: 55-94.

Turc, L. 1961. Estimation of irrigation water requirements, potential evapotranspiration: A simple climatic formula evolved up to date. *Annuals Agron.,* 12: 13-49.

Yoder, R.E., Odhiambo, L.O. and Wright, W.C. 2005. Evaluation of methods for estimating daily reference crop evapotranspiration at a site in the humid southeast United States. *Appl. Eng. Agric.,* 21:197–202.