Journal of Agrometeorology

ISSN : 0972-1665 (print), 2583-2980 (online)
Vol. No. 25 (1) : 145 - 150 (March- 2023)
DOI: https://doi.org/10.54386/jam.v25i1.1914
https://journal.agrimetassociation.org/index.php/jam

Research Paper

Estimation of reference evapotranspiration using artificial neural network models for semi-arid region of Haryana

RAM NARESH*, MUKESH KUMAR, SANDEEP KUMAR, KULDEEP SINGH and PARMOD SHARMA

Department of Soil & Water Engineering, CCS Haryana Agricultural University, Hisar-125004, India
**Corresponding author's email: ramnaresh.hau@gmail.com*

ABSTRACT

The study was conducted to evaluate performance of artificial neural network (ANN) models for estimating reference evapotranspiration (ET_0) for semi-arid region of Haryana state. Ten years (2011-2020) daily weather data of maximum and minimum temperature, relative humidity, wind speed and sun shine hours was collected from the meteorological observatory at CCS HAU, Hisar. Multilayer perceptron feed forward back propagation ANN models were evaluated for different training algorithms (10), number of hidden layers (1-3) and number of neurons in hidden layers (1-30). Training algorithms compared in the study were heuristic techniques (GDA, GDX, RP), conjugate gradient (CGF, CGP, CGB, SCG), quasi-Newton (BFG, OSS) and Levenberg-Marquardt (LM). Results were compared against standard FAO Penman-Monteith method. The study revealed that best performance for ANN was found with LM algorithm in single layer of 13 neurons exhibiting RMSE, R, ME and RPD values 0.306, 0.986, 0.976 and 6.63, respectively. ANN models showed good performance in prediction of reference evapotranspiration.

***Key words*:** ANN, Penman-Monteith, reference evapotranspiration, semi-arid, training method

The management of scarce water resources for sustainable crop production in the face of explosive growth of population is becoming more and more important especially in hot climate where the loss of water by evaporation and transpiration takes a significant portion of the irrigated water. Design and management of irrigation water resources require knowledge of the magnitude and variation of evapotranspiration losses. Besides irrigation management, estimates of evapotranspiration is needed in hydrology, agronomy, forestry and land resources planning, such as water balance computation, crop yield forecasting model, river flow forecasting , ecosystem modeling, studies on the impacts of global warming etc. (Huang et al., 2019; Wu et al., 2019; Yama and Todorovic, 2020). Therefore, the need for easy and accurate models for quantifying evapotranspiration losses for better management of scarce water resources is greater than ever before.

Evapotranspiration (ET) can be directly measured using lysimeter or by water balance approach or can be estimated indirectly using the climatic data. Many times field measurement of ET using lysimeter is not possible since it is time consuming and requires precise and carefully planned experiments. Moreover, installation and maintenance of lysimeter requires skilled manpower, correct instrumentation and finance. Owing to the difficulty of obtaining accurate field measurements, ET is commonly computed from weather data. A large number of empirical or semi-empirical equations have been developed for assessing evapotranspiration from meteorological data. An expert consultation held in May 1990, recommended the FAO Penman-Monteith (PM) method as the sole standard method for computation of the reference evapotranspiration (Allen et al., 1998). The PM method requires radiation, air temperature, air humidity and wind speed data and incorporates both the aerodynamics and thermodynamics aspects, therefore proved more accurate than other empirical methods (Fan et al., 2019). The PM model has been evaluated against various other methods under diverse areas, climates and time steps and performance better than other empirical equations (Pereira et al., 2015). However, the PM requires numerous features for evaporation estimation, including the geological variables such as elevation and latitude besides meteorological variables which bring a major drawback to the application of the PM model.

Artificial Intelligence (AI) techniques such as artificial neural network (ANN), fuzzy logic (FL) etc. have been successfully used extensively for the prediction of ET_0 and major advantage

***Article info - DOI*:** *https://doi.org/10.54386/jam.v25i1.1914*

Received: 28 October 2022; Accepted: 06 January 2023; Published online : 17 February 2023

is that these models are fully nonparametric and do not require a priori concept of the relations between the input variables and the output data (Kisi et al., 2016; Fahimi et al., 2017; Yama et al., 2020). The ANNs in particular have received extensive attention from researchers in estimation of ET_0 since 2000 (Bruton et al., 2000) and in subsequent studies (Traore et al., 2010; Sibale et al., 2016; Qasem et al., 2019; Ingle and Purohit, 2020). The ANN models needs to be fine tuned in terms of various hyper-parameters to get the minimum estimation error for a particular site. Considering the limited irrigation facility, uneven and ill-distributed rainfall and possibility of drought in the semi-arid region of Haryana, accurate estimation of evapotranspiration is highly desirable. Thus there is a needs of accurate method for evapotranspiration estimation so that a handy tool could be made available to the developmental agencies involved in planning and utilization of water resources in the region. Therefore in present study ANN model was evaluated for reference evapotranspiration prediction.

MATERIALS AND METHODS

The aim of this study was to develop and evaluate the model for reference evapotranspiration prediction for Hisar. Ten years (2011-2020) daily weather data of maximum temperature (T_{max}), minimum temperature (T_{min}), relative humidity (RH), wind speed (WS) at height of 2.0 m and sunshine hours (SH) were collected from the Department of Agricultural Meteorology, CCS Haryana Agricultural University, Hisar (Latitude 21° 15' N, Longitude 75°68' E and elevation as 215 m msl. 70 % data was utilized for model training (1-2556 days), next 15 % data was used for model validation (2557- 3104 days) and remaining 15 % for model testing (3105-3652 days). Neural network training can be done more efficiently if certain preprocessing steps like data normalization are performed on the network inputs and targets. Data normalization scales the inputs and target in such a way that they fall within a specific range. The normalization process removes the effect of variation in scale of magnitude of different input parameters and the cyclicity of the data.

Reference evapotranspiration (ET_0) was estimated using the reference FAO Penman-Monteith equation (Allen et al. 1998).

Artificial neural network (ANN)

Multi-layer perceptron ANN has been most widely used for regression type of problems. In this study MLP with 5 input parameters viz. maximum temperature (T_{max}), minimum temperature (T_{min}), relative humidity (RH), wind speed (WS) and sunshine hours (SH) were used to estimate reference evapotranspiration as single output. ANN structure has three parts, an input layer, an output layer and in between three hidden layers. Neurons in input layer only act as buffers for distributing the input signals x_i (i=1, 2 …5) to neurons in the next hidden layers. Each neuron j in the first hidden layer sums up its input signals x_i after weighting them with the strengths of the respective connections W_{ij} from the input layer and computes its output y_j as a function 'f' of the sum Σ . 'f ' can be a simple threshold function or a sigmoidal, hyperbolic tangent or radial basis function.

Training method

Multilayer perceptron feed forward back propagation ANN models were evaluated for different training algorithms, number of neurons in hidden layers and number of hidden layers. There are several high performance algorithms that can converge to the global minima. These algorithms fall into two main categories. The first category uses heuristic techniques, which were developed from an analysis of the performance of the standard steepest descent algorithm. The heuristic technique included the gradient descent with variable learning rate (GDA), the variable learning rate with momentum (GDX) and resilient backpropagation (RP).

The second category of fast algorithms uses standard numerical optimization techniques. Most common three types of numerical optimization techniques for neural network training are conjugate gradient, quasi-Newton and Levenberg-Marquardt (LM). The conjugate gradient algorithms includes Fletcher-Reeves update (CGF), Polak-Ribiére update (CGP), Powell-Beale restarts (CGB) and Scaled conjugate gradient (SCG) while quasi-Newton includes BFGS algorithm (BFG) and One-step secant algorithm (OSS).

Each training algorithm went 30 layer deep (number of neurons in hidden layer) and three layers wide (number of hidden layers).

Model performance criteria

Different model performance criteria were used for testing the performance of ANN models for their capabilities in estimation of reference evapotranspiration based on climatic data.

Root mean square error (RMSE)

The root mean square error (RMSE) is a good measure of accuracy showing the differences between values predicted by a model and the observed values. These individual differences are also called residuals, and RMSE serves to aggregate them into a single measure of predictive power, which was calculated by using following equation:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2}$$

Where, P_i and O_i are the predicted and observed values, respectively, and N is the total number of observations.

Correlation coefficient (R)

The value of coefficient of correlation was determined using following equation:

$$R = \frac{\sum_i^N (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_i^N (O_i - \bar{O})^2 \sum_i^N (P_i - \bar{P})^2}}$$

Where, and are mean of the predicted and observed values.

The value of the coefficient of correlation always lie between –1 to +1. When R is +1, there is a perfect positive correlation between the variables, whereas, R equals to -1 indicates perfect negative correlation between the variables. R equal to zero

Table 1: Comparison of best of ANN training architectures

Training algorithm	ANN architecture	Performance criteria			
		RMSE	R	ME	RPD
GDA	5-1-1	0.522	0.967	0.931	3.89
GDX	5-1-1-1	0.439	0.977	0.951	4.62
RP	5-14-14-1	0.350	0.979	0.970	5.80
CGF	5-18-18-1	0.325	0.982	0.973	6.25
CGP	5-18-18-1	0.330	0.987	0.970	6.15
CGB	5-18-18-1	0.320	0.984	0.970	6.34
SCG	5-18-18-1	0.326	0.984	0.973	6.23
BFG	5-20-1	0.316	0.977	0.974	6.42
OSS	5-18-1	0.350	0.984	0.971	5.80
LM	5-13-1	0.306	0.986	0.976	6.63

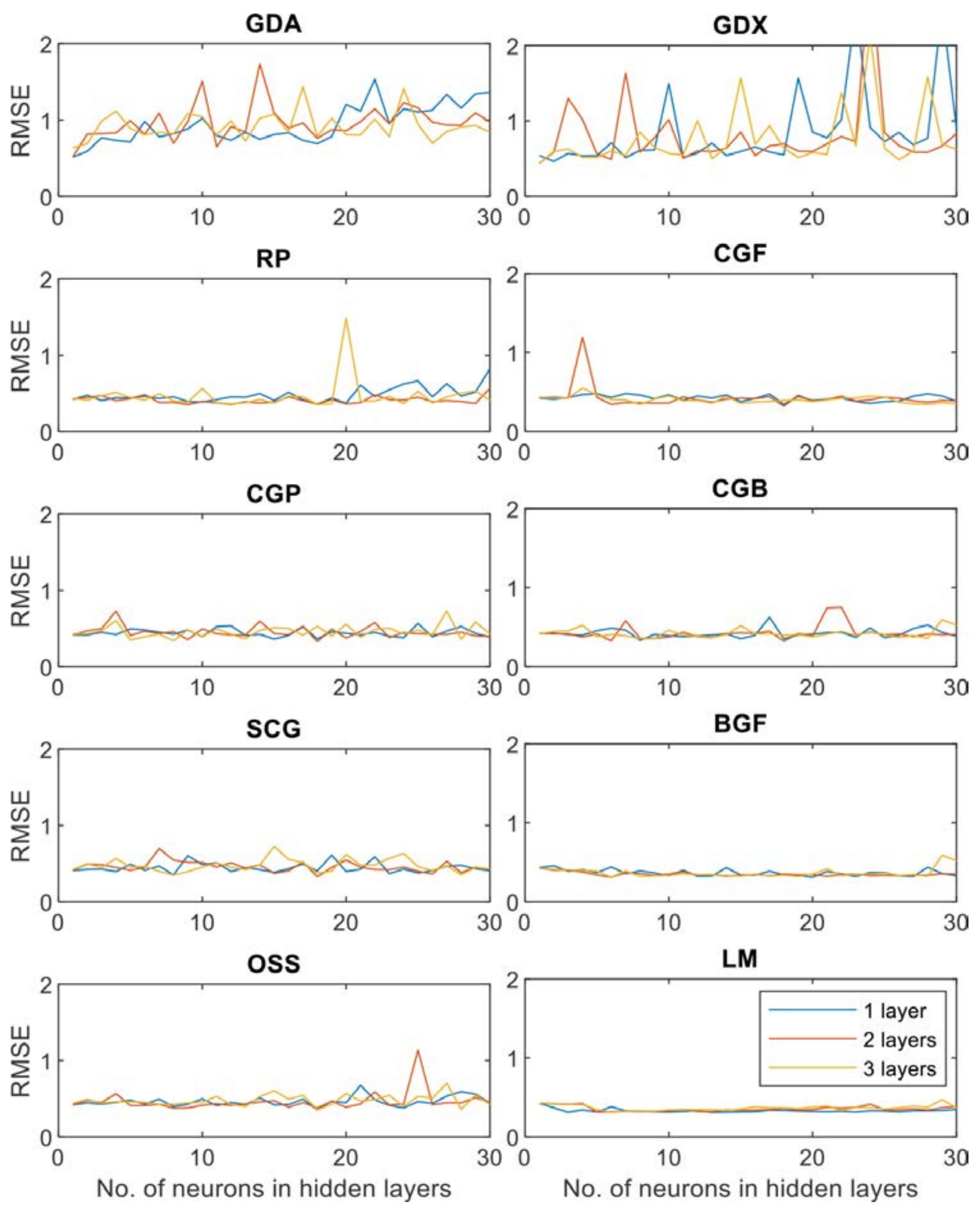

Fig. 1: Comparison of various training algorithms with respect to number of hidden layers and number of neurons

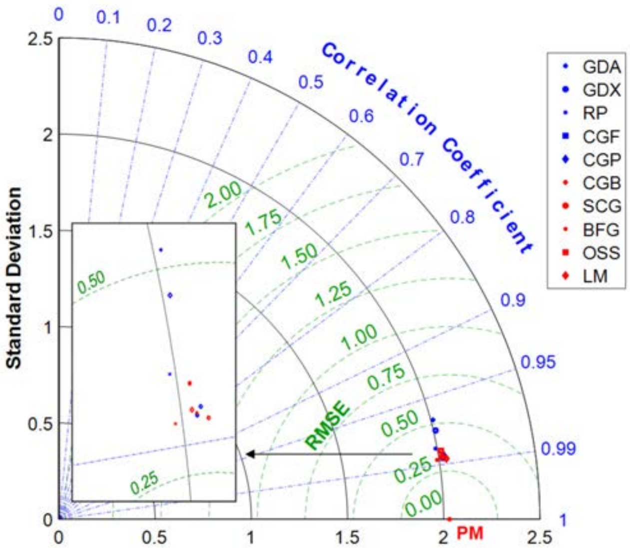

Fig. 2: Taylor diagram of performance of ANN with different training algorithms

shows that there is no relationship between the two variables.

Model efficiency (ME)

Model efficiency is calculated using the following equation:

$$ME = 1 - \frac{\sum_{i=1}^{N}(P_i - O_i)^2}{\sum_{i=1}^{N}(\bar{O} - O_i)^2}$$

In the situation of a perfect model with an estimation error variance equal to zero, the resulting model efficiency equals 1.

Ratio of performance to deviation (RPD)

Ratio of performance to deviation is calculated using the following equation:

$$RPD = \frac{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(O_i - \bar{O}_i)^2}}{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_i - O_i)^2}}$$

The greater the RPD, the better the model's predictive capacity.

RESULTS AND DISCUSSION

Reference evapotranspiration calculated from FAO PM method

The monthly average value of reference evapotranspiration was calculated using FAO PM model. The monthly average value of reference evapotranspiration varied from 1.63 mm day^{-1} in January to 7.91 mm day^{-1} in June for the period 2011-2020. The standard deviation in monthly average of reference evapotranspiration varied from 0.22 mm day^{-1} in December to 0.61 mm day^{-1} in March.

Performance of ANN models

Prediction errors (in terms of RMSE) of ANN models with different training algorithms are depicted in Fig. 1. Heuristic type training algorithms (GDA, GDX and RP) exhibited overall higher RMSE compared to the numerical optimization techniques. Lowest value of RMSE for heuristic algorithms was found at lower number of neurons (1 neuron for GDA and GDX, 14 neurons for RP). Heustric techniques are self evolving by nature and adding the complexity of number of neurons as well as number of hidden layers, reduced the performance as the complexity overburdened these algorithms. All of the conjugate gradient algorithms (CGF, CGP, CGB and SCG) performed their individual best at exactly same number of neurons (18) and with two hidden layer architecture. This behaviour exhibits that selection of the conjugate direction methods does not affect the performances but the ANN architecture affect the performances. Conjugate gradient algorithms overall exhibited better performances having lower RMSE values than the Heuristic type training algorithms. The Quasi-Newton algorithms (BFG and OSS) performed their best with single hidden layer and almost identical number of neurons (20 for BFG and 18 for OSS). The LM algorithm, with single hidden layer and 13 neurons in the hidden layer, performed best amongst all ten training algorithms for estimation of reference evapotranspiration for the study area. The second best performing algorithm (BFG) also showed the lowest RMSE with single hidden layer. For more complex type of problems having large number of input and output variables, adding depth (the number of neurons) or adding the width (number of hidden layers) might have improved the results but for this particular type of problem, having small number of input variable (5) and single output, the best results were obtained with single hidden layer and keeping fewer number of neurons in hidden layer. Performances parameters of the ANN model with different training algorithms are compiled in Table 1 and comparative Taylor diagram is presented in Fig. 2. Amongst all the training algorithms, the best results were obtained with LM having RMSE, R, ME and RPD values of 0.306, 0.986, 0.976 and 6.63, respectively, while the least performance was found with GDA having RMSE, R, ME and RPD values of 0.522, 0.967, 0.931 and 3.89, respectively.

Kayri (2016) in comparison of the predictive ability of different training algorithms, found good performance of LM training algorithm. In another study for estimation of monthly evaporation prediction for Mediterranean region of Turkey (Kermani et al., 2021), R^2 values for BFG (0.925) was found slightly better than LM (0.914). Evaporation simulation at Qassim, Saudi Arabia (Ghumman et al., 2021) with LM backpropagation as the training function showed highest values of the performance indicators, although, there was very small difference in the statistical parameters over Quasi-Newton BFG and SCG. Similarly, highest performance of LM training algorithm was observed in regression study by Nguyen et al., (2021) while for classification, Quasi-Newton training algorithm showed better accuracy (Karim et al., 2018). LM algorithm was designed to approach second-order training speed without having to compute the Hessian matrix which makes it best training method for training moderate-sized feed-forward neural networks for regression.

In this study, the absolute error of best ANN having LM with 13 neurons in single hidden layer varied from 0.006 mm day-1 to 1.304 mm day-1 (Fig. 3), with mean value of 0.2176 mm day-1 and standard deviation of 0.213 mm day-1. ANN slightly

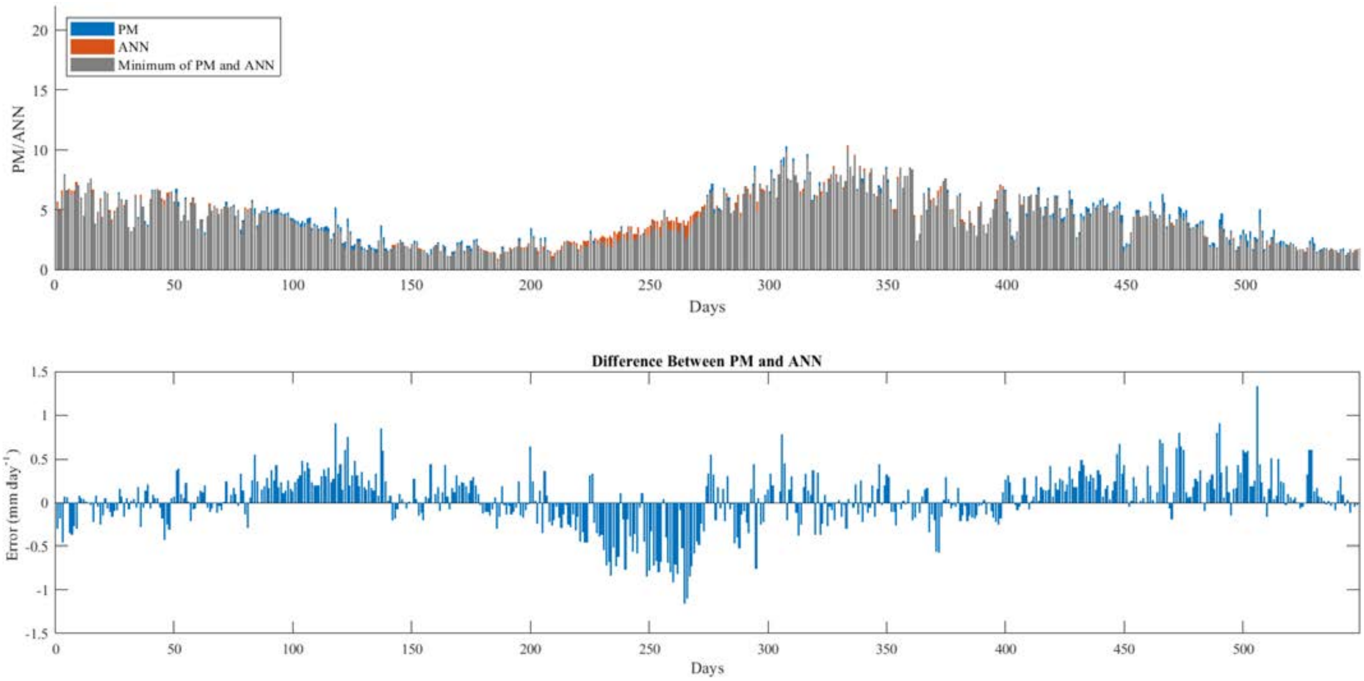

Fig. 3 : Error plot of ANN

overestimated the reference evapotranspiration than the PM. A maximum deviation of 1.33 mm day^{-1} was observed on Nov 19, 2018 but the day did not exhibit any extreme climate (T_{max} 28.4 ^{0}C, T_{min} 17.5 ^{0}C, RH 62.0 %, WS 7.8 km/h and SH as 7.2 h).

CONCLUSION

The present study concluded that amongst all evaluated ANN models, best performance on the basis of model accuracy parameters was exhibited by LM algorithm with 13 neurons in single hidden layer. Performance of numerical optimization techniques was found better than heuristic techniques in estimation of reference evapotranspiration. Increasing the number of hidden layers beyond two layers decreased the performance of all the training algorithms. All ANN models showed good performance in prediction of reference evapotranspiration but efficiency could be further improved by fine tuning it with training methods, number of hidden layers and number of neurons in the hidden layers.

ACKNOWLEDGMENTS

The authors wish to acknowledge support from the Department of Agricultural Meteorology, CCS HAU, Hisar for providing meteorological data for the study.

Conflict of Interest Statement: The author(s) declare(s) that there is no conflict of interest.

Disclaimer: The contents, opinions, and views expressed in the research article published in the Journal of Agrometeorology are the views of the authors and do not necessarily reflect the views of the organizations they belong to.

Publisher's Note: The periodical remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

Allen, R.G., Pereira, L.S., Raes, D. and Smith, M. (1998). Crop evapotranspiration-Guidelines for computing crop water requirements. Irrigation and Drainage Paper 56, FAO, *Rome,* 9, 105-109.

Bruton, J, Hoogenboom, G and McClendon, R. (2000). A comparison of automatically and manually collected pan evaporation data. Trans. *ASAE,* 43(5): 1097–1101.

Fahimi, F., Yaseen, Z.M. and El-shafie, A., (2017). Application of soft computing based hybrid models in hydrological variables modeling: a comprehensive review. Theor. *Appl. Climatol.,* 128, 875–903.

Fan, J., Ma, X., Wu, L., Zhang, F., Yu, X. and Zeng, W. (2019). Light Gradient boosting Machine: an efficient soft computing model for estimating daily reference evapotranspiration with local and external meteorological data. *Agric. Water Manag.,* 225, 105758.

Ghumman, A. R., Jamaan, M., Ahmad, A., Shafiquzzaman, M., Haider, H., Al Salamah, I. S. and Ghazaw, Y. M. (2021). Simulation of pan-evaporation using Penman and Hamon equations and artificial intelligence techniques. *Water,* 13, 793.

Huang, G., Wu, L., Ma, X., Zhang, W., Fan, J., Yu, X., Zeng, W. and Zhou, H. (2019). Evaluation of catboost method for prediction of reference evapotranspiration in humid regions. *J. Hydrol.*, 574, 1029 – 1041.

Ingle, P. M. and Purohit, R. C. (2020). Assessment of reference evapotranspiration using ANN at Mulde, Maharashtra.

J. Agrometeorol., 22(2): 220-225 DOI: https://doi.org/10.54386/jam.v22i2.175

Kayri, M. (2016). Predictive abilities of bayesian regularization and Levenberg–Marquardt algorithms in artificial neural networks: A comparative empirical study on social data. *Math. Comput. Appl.*, 21(20), 1–11.

Karim, H., Niakan, S. and Safdari, R. (2018). Comparison of neural network training algorithms for classification of heart diseases. *Int. J. Artif. Intell.*, 7, 185

Kermani, M., Keshtegar, B., Kisi, O. and Scholz, M (2021). Towards a comprehensive assessment of statistical versus soft computing models in hydrology: Application to monthly pan evaporation prediction. *Water,* 13, 2451.

Kisi, O., Genc, O., Dinc, S. and Zounemat-Kermani, M. (2016). Daily pan evaporation modeling using chi-squared automatic interaction detector, neural networks, classifcation and regression tree. Comput. *Electron Agric.*, 122, 112–117.

Nguyen, T.A., Mai, H., and Tran, V. Q. (2021). On the training algorithms for artificial neural network in predicting the shear strength of deep beams. *Complexity,* 3, 1-18.

Pereira, L.S., Allen, R.G., Smith, M. and Raes, D. (2015). Crop evapotranspiration estimation with FAO56: past and future. *Agric. Water Manag.*, 147, 4–20.

Qasem, S.N., Samadianfard, S., Kheshtgar, S., Jarhan, S., Kisi, O., Shamshirband, S. and Chau, K.W. (2019). Modeling monthly pan evaporation using wavelet support vector regression and wavelet artificial neural networks in arid and humid climates. *Eng. Appl. Comput. Fluid Mech.*, 13, 177–187

Sibale D., Mane, M.S., Patil, S.T. and Ingle, P.M. (2016). Evaluation of three methods for estimating reference evapotranspiration ETo at Dapoli, Maharashtra. *J. Agrometeorol.*, 18(1):157-158 DOI: https://doi.org/10.54386/jam.v18i1.925

Traore, S, Wang, Y. M. and Kerh, T. (2010). Artificial neural network for modeling reference evapotranspiration complex process in Sudano-Sahelian zone. *Agric. Water Manag.*, 97, 707–714.

Wu, L., Zhou, H., Ma, X., Fan, J. and Zhang, F. (2019). Daily reference evapotranspiration prediction based on hybridized extreme learning machine model with bio-inspired optimization algorithms: Application in contrasting climates of china. *J. Hydrol.*,577, 123960.

Yama, S. S. and Todorovic, M. (2020). Estimation of daily potato crop evapotranspiration using three different machine learning algorithms and four scenarios of a vailable meteorological data. *Agric. Water Manag.*, 228,105875.