Evaluation of soft computing and regression-based techniques for the estimation of evaporation

Aparajita Singh, R. M. Singh, A. R. Senthil Kumar, Ashish Kumar, Subodh Hanwat and V. K. Tripathi

ABSTRACT

The estimation of evaporation in the field as well as the regional level is required for the efficient planning and management of water resources. In the present study, artificial neural network (ANN) and multiple linear regression (MLR)-based models were developed to estimate the pan evaporation on the basis of one day-lagged rainfall (P_{t-1}), one day-lagged relative humidity (RH_{t-1}), current day maximum temperature (T_{max}) and minimum temperature (T_{min}). These were selected as the most effective parameters on the basis of cross-correlation. The performance of models was evaluated using correlation coefficient (r), root-mean-square error (RMSE) and Nash–Sutcliffe efficiency (coefficient of efficiency, CE) during calibration and validation periods. Based on the comparison, the ANN model (4-9-1), with sigmoid as activation function and Levenberg–Marquardt as a learning algorithm, was selected as the best performing model among all ANN models. The values of r, CE and RMSE for training and validation periods were found as 0.885, 0.785 and 1.00 mm/day and 0.889, 0.782 and 1.01 mm/day, respectively, through the ANN model (4-9-1). The values of r, CE and RMSE for training and validation periods were found as 0.835, 0.698 and 1.19 mm/day and 0.866, 0.750 and 1.15 mm/day, respectively, through the selected MLR model. Based on the sensitivity analysis, RH_{t-1} is selected as the most effective parameter followed by P_{t-1}, T_{max} and T_{min}. The developed model can be utilized as an alternative for the estimation of the evaporation at the regional level with limited input data.

Key words | artificial neural network, feedforward neural network, log-sigmoid activation function, multilinear regression

Aparajita Singh
R. M. Singh
Ashish Kumar
Subodh Hanwat
V. K. Tripathi (corresponding author)
Department of Farm Engineering,
  Institute of Agricultural Science,
Banaras Hindu University,
Varanasi, Uttar Pradesh 221005,
India
E-mail: tripathiwtcer@gmail.com

A. R. Senthil Kumar
National Institute of Hydrology,
Roorkee 247667,
Uttarakhand

INTRODUCTION

Evaporation is the major, diffusive process of the hydrological cycle in which water is changed from the liquid state into the vapor state due to the transfer of heat energy. It plays a vital role in water resource planning and development in arid and semi-arid climatic regions (Shirsath & Singh 2010). The management of scarce water resources is important for sustainable crop production (Shirgure & Rajput 2012). In a hot climate, water loss plays an important role through evaporation from rivers, canals and open-water bodies. Evaporation losses are even significant in humid areas, but the cumulative mean precipitation over these areas can mask it, meaning it is not ordinarily recognized except during the rainless period. The estimation of evaporation is required in water balance computations, irrigation management, crop yield forecasting, river flow forecasting and ecosystem modeling in hydrology, agronomy, forestry and land resource planning (Terzi & Keskin 2005).

Irrigation scheduling is based on evapotranspiration estimated through evaporation. It is a systematic approach for efficient irrigation scheduling. Farm irrigation systems,

as well as water resource development projects, are also designed with a long-term mean value of evaporation and their magnitude along with variation in evaporation losses (Shiri *et al.* 2011). There is a growing demand for evaporation data for the studies of surface water and energy fluxes, especially addressing the impacts of global warming (Xu & Singh 2005). The accurate estimation of evaporation is fundamental for the effective management of water resources. Therefore, the need for reliable models for quantifying evaporation losses from increasingly scarce water resources is greater than ever before (Tabari *et al.* 2010).

An appropriate and consistent measurement of evaporation estimated by direct and indirect methods for a longer duration is always challenging for the researchers. In the direct evaporation measurement method, the United States Weather Bureau (USWB) Class A pan evaporimeter and eddy correlation techniques were used, whereas, in indirect methods, the meteorological variables were used to estimate the evaporation. The most widely used crop monitoring and forecasting models by the Food and Agriculture Organization (FAO) are based on evaporation estimates (Gommes 1998). Yihdego & Webb (2017) evaluated the differences in the measured pan evaporation and estimated evaporation seasonally of Lake Burrumbeet, Australia and showed that evaporation is fully radiation driven and the effect of wind is minimal. Although various empirical formulas are available, their performances are not satisfactory because of the complex nature of the evaporation process and non-availability of the input data. Therefore, the estimation of evaporation from easily available meteorological parameters is another good alternative.

Evaporation is a complex and nonlinear phenomenon as it depends on several interacting meteorological factors such as temperature, humidity, wind speed and bright sunshine hours (Xu & Singh 1998, 2001). Unfortunately, reliable estimates of evaporation are extremely difficult to obtain because of complex interactions between the components of the soil–plant–atmosphere system. This is partly due to reasons such as a wide range of data type and expertise are required (Goyal *et al.* 2014). Many conventional models such as empirical, regression-based models and conceptual models have been developed to estimate the evaporation with a large quantity of data but produced less accurate results. The models developed from meteorological data like multiple regression models involve empirical relationships to some extent which accounts for local conditions. Therefore, most of the models may give reliable results when applied to climatic conditions similar to those for which they were developed. But without local or regional calibration, the use of such models under greatly different climatic conditions may give results that differ considerably. No single model is universally adequate under all climatic conditions, so it is difficult to select the most appropriate evaporation model for a given region.

Climate-based models (Priestley & Taylor 1972) like regression models (Kisi 2009), artificial neural network (ANN) models (Sudheer *et al.* 2002) and adaptive neural-based fuzzy inference systems (Guven & Kisi 2011) were used for estimating daily evaporation. Of all the models, the recently developed soft computing techniques such as ANN, fuzzy logic, decision-tree algorithms (ASCE 2000) with a different network structure and meteorological variable combinations were effectively applied by many researchers for modeling the evaporation. The ANN models have been found to be useful and efficient, particularly in problems for which the characteristics of the processes are difficult to describe using physical equations (Zhang *et al.* 1998). Neural network applications have diffused rapidly due to their functional characteristics which provide many advantages over traditional analytical approaches (ASCE 2000). The model requires significantly minimum input data than a similar conventional mathematical model since variables remain fixed from one simulation to another. An ANN can quickly present sensitive responses to tiny input changes in a dynamic environment (Moghaddamnia *et al.* 2009). The estimation of the reference evapotranspiration using the least square support vector machine gamma test (LSSVM-GT) model with different kernels of radial basis function, linear and polynomial, ANN-GT, adaptive neuro-fuzzy inference system (ANFIS-GT), and empirical equations was carried out and it was found that the LSSVM-GT approach is highly capable of providing favorable predictions with high precision in arid regions (Seifi & Riahi 2018). In the present study, an attempt has been made to perform the following objectives: (a) development of the relationship between independent variables and dependent variables using correlation coefficient, cross-correlation techniques for selecting the best performing

input variables, (b) development of the ANN and multiple linear regression (MLR)-based evaporation estimation models, (c) qualitative and quantitative performance evaluation of the developed model and (d) sensitivity analysis to determine the most effective parameter.

MATERIALS AND METHODS

Study area

The present study was carried out in the Roorkee, having an area of 811 ha, which comes under the Haridwar District of Uttarakhand, India as shown in Figure 1. The Upper Ganga Canal divides the city into two distinct parts that enhance the beauty of the city. The city is located at a latitude of 29°51′N and longitude of 77°53′E with an elevation of 274 m above the mean sea level. It receives an average annual rainfall of 1,032 mm. The average temperature in January and June is 13.8 and 32.2 °C, respectively, with annual temperature difference of 18.4 °C. The highest temperature of 45.5 °C (on 9 May 1956) was recorded and the lowest temperature of 3.3 °C was recorded (on 26 January 1964). The temperature ranges from 5 to 20 °C in Winter (December to March), 25 to 40 °C in Summer (April to June) and 20 to 40 °C in the rainy season (July to September). The humidity ranges from 30% to 100%. The soil type is alluvial of the Ganga plain derived from the soft dolomitic rocks of the Himalayas. Due to its location away from any major water body and its proximity to the Himalayas, the study area of Roorkee has an extreme and erratic continental climate, with average annual wind speed of 4.9 m/s (http://nihroorkee.gov.in/location.html; accessed 27 September 2018). Daily rainfall data (P), minimum temperature (T_{min}), maximum temperature (T_{max}) and relative humidity (RH) from January 2001 to December 2013 were collected from NIH (2018), Roorkee Observatory.

Artificial neural network

An ANN is an information processing structure that consists of several interconnected processing elements called nodes which are analogous to human brain neurons. Each node combines some inputs and produces an output

which is then transmitted to many different locations including other nodes (Affandi *et al.* 2008). Synapses are the structures where weight values are stored. An ANN model can compute complex nonlinear problems that may be too difficult to represent by conventional mathematical equations (McCulloh & Pitts 1943). An ANN model with the feedforward neural network (FFNN) architecture was used in the present study to estimate the daily pan evaporation.

Feedforward neural network

In the FFNN, the data flow through the network in one direction from the input layer to the output layer through the hidden layer(s). Each output value was based on the current set of inputs. The input layer consists of nodes that receive inputs from the external environment and do not perform any transformations on the inputs. It just sends the weighted values to the immediately adjacent nodes, usually the hidden layer. The hidden layer(s) consists of nodes that receive the transferred weighted inputs from the input layer or the previous hidden layer. It performs the transformations on the weighted inputs and passes the output to the next adjacent layer which can be either another hidden layer or the output layer. The output layer consists of nodes that receive the hidden layer output and send it to the user. Feedforward means that all the interconnections between the layers propagate forward to the next layer. The main advantage of FFNN structures is that they are easy to handle and can approximate any input/output map (Hornik 1991). The log-sigmoid activation function was used in this study to introduce nonlinearity in the model. Each node is a simple processing element that responds to the weighted inputs that it receives from other nodes. The receiving node sums the weighted signals from all nodes to which it is connected in the preceding layer. The net input to the node was the weighted sum of all the incoming signals as represented in the following equation:

$$\text{net} = \sum x_i w_i \tag{1}$$

where x_i ($i = 1, 2, 3\ldots n$) = net input, w_i ($i = 1, 2, 3\ldots n$) = respective weights.
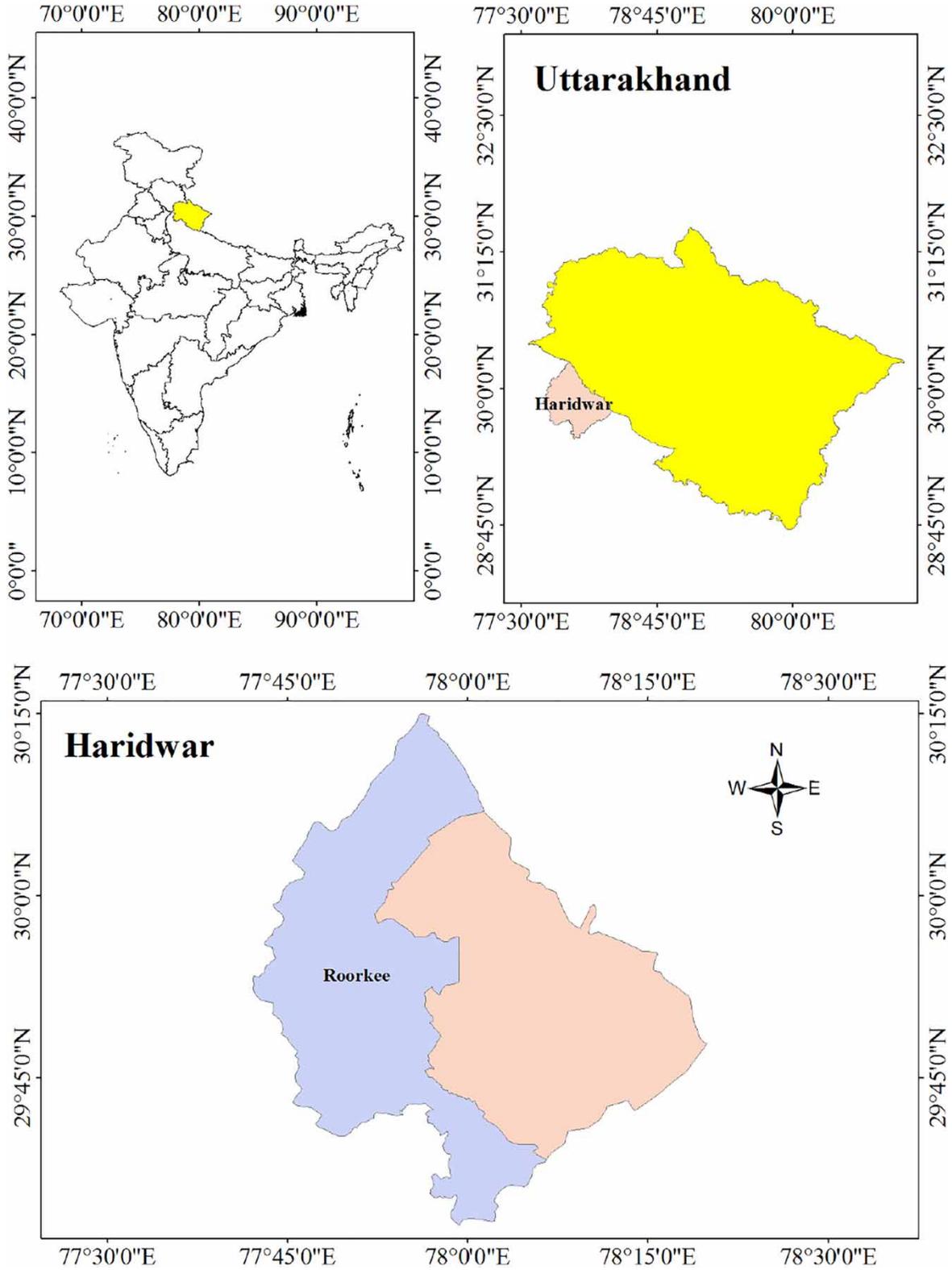
Figure 1 | Location of the study area.

The activation function, *y*, is a nonlinear function of its net input and described by the sigmoid logistic function as per the following equation:

$$y = f(\text{net}) = \frac{1}{1 + e^{-\text{net}}} \tag{2}$$

Network architecture

The network geometry is generally highly problem oriented to obtain optimal network geometry through the trial-and-error procedure. The number of hidden neurons in the network, which are responsible for capturing the dynamic and complex relationship between various input and output variables, is identified by various trials. For each set of hidden neurons, the network is trained with input datasets in the batch mode to minimize the mean squared error at the output layer. Various internal parameters used in the ANN model like learning rate, momentum coefficient, and the combination of transfer functions for hidden and output layers were also found out by trial-and-error. The MATLAB 2010 software was used for ANN model analysis (Higham & Higham 2016). In this study, the ANN model with four input layers, one output layer and one hidden layer with 1–10 neurons in the hidden layer was used. The basic architecture is shown in Figure 2. The equation for
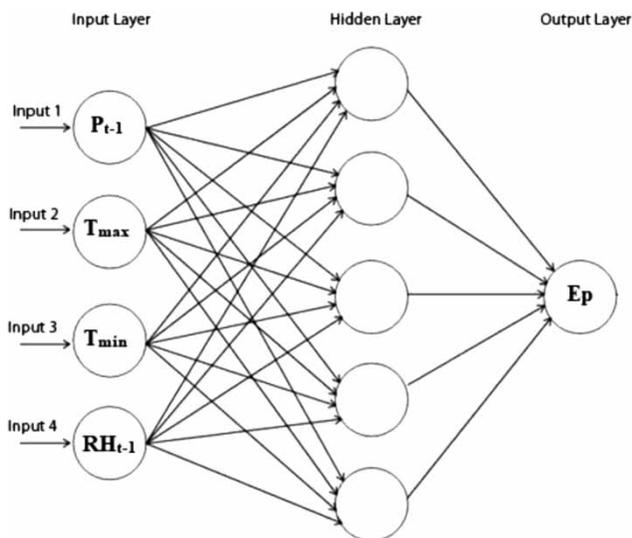

Figure 2 │ Artificial neural network structure.

an ANN model used in the present study is expressed in the following equation:

$$\text{Net input} = x_1 w_{11} + x_2 w_{12} + x_3 w_{13} + x_4 w_{14} \tag{3}$$

Multiple linear regressions

MLRs are a multivariate statistical technique used to model the linear correlations between the dependent variable *y* and two or more independent variables by fitting a linear equation to observed data. The value of the dependent variable *y* is associated with each response of independent variables. The regression equation of *y* can be written as follows:

$$y = m_1 x_1 + m_2 x_2 + m_3 x_3 \ldots m_n x_n + C \tag{4}$$

where *y* is the response variable or the dependent variable, x_1, $x_2 \ldots x_n$ are the independent variables and m_1, $m_2 \ldots m_n$ are the regression coefficients and *C* is constant.

When there was a linear relationship between the dependent variable and independent variables, then the MLR models were the standard method for the estimation of responses between a dependent variable and various independent variables.

Data formulation

The daily observed data of relative humidity (RH), rainfall (*P*), minimum temperature ($T_{\min}$) and maximum temperature ($T_{\max}$) of the current day and previous days from January 2001 to December 2013 (4,787 daily datasets) were explored to estimate the current day pan evaporation. Out of 4,787 datasets, 1–3,286 datasets (62.50%) were used for calibration, while datasets from 3,287 to 4,787 (37.50%) were used for validation. These datasets were selected based on a trial-and-error method.

Normalization of input data

The input values were normalized (using Equation (5)) between 0 and 1 before passing into a neural network since the output of the sigmoid function is bound between

0 and 1. Dawson & Wilby (1998) and many others have emphasized the importance of the normalization of data and have given the procedure to normalize. The output from the ANN should be denormalized to provide meaningful results.

$$N_i = \frac{R_i - \text{Min}_i}{\text{Max}_i - \text{Min}_i} \tag{5}$$

where R_i is the real value applied to neuron i, N_i is the subsequent normalized value calculated for neuron i, Min_i is the minimum value of all values applied to neuron i, Max_i is the maximum value of all values applied to neuron i.

Performance evaluation of the developed ANN model and the MLR model

The qualitative and quantitative evaluation of the developed models was performed to judge the goodness of fit between observed and estimated values. The qualitative observation is based on the visual comparison of graphs between the observed and estimated evaporation values. The performances of the model based on quantitative evaluation during calibration and validation were evaluated by using statistical parameters, namely correlation coefficient (r), root-mean-square error (RMSE) and coefficient of efficiency (CE).

Coefficient of efficiency

Based on the standardization of residual variance with initial variance, the CE was used to compare the relative performance of the two approaches effectively and is given as per the following equation:

$$CE = \left[1 - \frac{\sum_{j=1}^{n} (y_{ej} - y_j)^2}{\sum_{j=1}^{n} (y_{ej} - \overline{y}_{ej})^2} \right] \tag{6}$$

Chiew *et al.* (1993) classified the CE into three categories: perfectly acceptable simulation (CE >0.90), acceptable simulation (CE between 0.60 and 0.90) and unacceptable simulation (CE <0.60).

Root-mean-square error

RMSE is used to evaluate the comparison between the model's simulated responses with that of recorded watershed responses. The RMSE is zero for perfect fit, whereas its increased values indicate higher discrepancies between predicted and observed values. The lowest the RMSE, the more accurate the estimation is and given by the following equation:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{j=1}^{n} (y_{ej} - y_j)^2} \tag{7}$$

Correlation coefficient

The correlation coefficient (r) is an indicator of the degree of closeness between observed and predicted values. If observed and predicted values are completely independent, the correlation coefficient will be zero and is given as per the following equation:

$$r = \frac{\sum_{j=1}^{n} \{(y_j - \overline{y}_j)(y_{ej} - \overline{y}_{ej})\}}{\sum_{j=1}^{n} (y_j - \overline{y}_j)^2 \sum_{j=1}^{n} (y_{ej} - \overline{y}_{ej})^2} * 100 \tag{8}$$

where y_j is a predicted variable, y_{ej} is observed variable, $\overline{y}_j$ is the mean of the predicted valve, $\overline{y}_{ej}$ is the mean of the observed valve and n is the number of observation.

RESULTS AND DISCUSSION

Selection of input vector

The cross-correlation analysis between the dependent and independent variables was carried out to select the significant input vector. To identify the input vector, a detailed cross-correlation analysis of the following variables was done.

(i) Daily rainfall (P) valves with daily evaporation valves.
(ii) Daily maximum ($T_{\max}$) and minimum temperature ($T_{\min}$) valves with daily evaporation valves.

(iii) Daily relative humidity (RH) valves with daily evaporation valves.

The cross-correlation graph was plotted between inputs used in analysis and pan evaporation. It was observed that the daily rainfall (P) valves and RH are negatively correlated with evaporation valves (Figures 3(a), 3(b), 4(a) and 4(b)).

The correlation analysis was also done to find the exact lag between the dependent and independent variables. The correlation coefficients were estimated between daily pan evaporation and different input parameters such as daily rainfall (P), daily relative humidity (RH), daily maximum temperature (T_{max}) and daily minimum temperature (T_{min}). The maximum positive correlation coefficient value estimated between T_{max} and evaporation was 0.72, whereas that estimated value between T_{min} and evaporation was 0.53. The negative correlation coefficient with values −0.030 and −0.610 was found between evaporation and rainfall (P) and relative humidity (RH), respectively, as shown in Table 1. Negative correlation means that evaporation decreases with an increase in rainfall and RH.

The inputs for the ANN model based on correlation analysis were one day-lagged rainfall (P_{t-1}), one day-lagged relative humidity (RH_{t-1}) current day maximum temperature (T_{max}) and current day minimum temperature (T_{min}) as shown in the following equation. There was one node in the output layer which estimates evaporation $E(t)$.

$$E(t) = f\{P_{t-1}, RH_{t-1}, T_{min}, T_{max}\} \tag{9}$$

Development of ANN- and MLR-based models

The estimation equation developed through MLR analysis computes the daily pan evaporation with the same input parameters used in the ANN model. The one day-lagged rainfall (P_{t-1}), one day-lagged RH (RH_{t-1}, daily T_{max} and daily T_{min}) were taken as the independent variables and pan evaporation was taken as a dependent variable for the MLR model. The regression equation developed for the present study was given in the
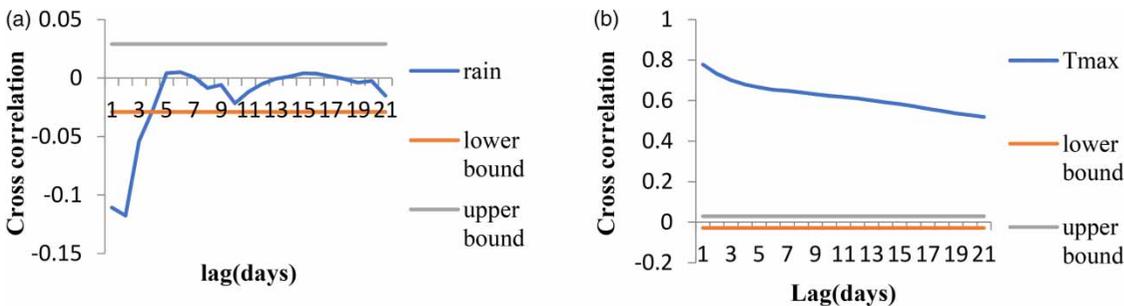

Figure 3 | (a) Cross-correlation graph plotted between daily rainfall and evaporation. (b) Cross-correlation graph plotted between daily maximum temperature and evaporation.
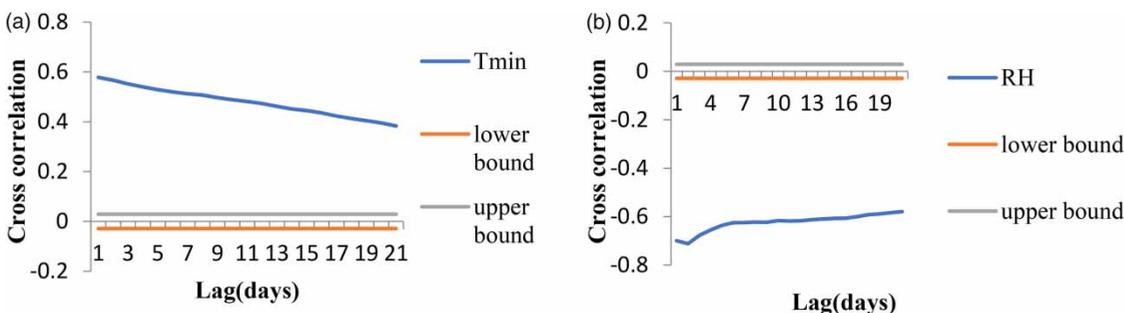

Figure 4 | (a) Cross-correlation graph plotted between daily minimum temperature and evaporation. (b) Cross-correlation graph plotted between daily relative humidity and evaporation.

Table 1 | Relation of input variables with the output variable

S. no.	Data	Correlation coefficient with evaporation
1	Rainfall (mm)	−0.03
2	Relative humidity (%)	−0.61
3	Maximum temperature (°C)	0.72
4	Minimum temperature (°C)	0.53

following equation:

$$EVAP(t) = m_1 * P_{t-1} + m_2 * RH_{t-1} + m_3 * T_{min} + m_4 * T_{max} + C \tag{10}$$

The coefficients m_1, m_2, m_3 and m_4 were found according to the regression analysis between the dependent variable and independent variables. The regression coefficient values $m_1 = -0.010$ observed between evaporation and daily rainfall, $m_2 = 0.199$ and $m_3 = 0.008$ observed between daily evaporation and daily maximum temperature, daily minimum temperature, respectively, and $m_4 = -0.044$ observed between evaporation and daily RH. From regression analysis, it was observed that rainfall and RH were negatively related to evaporation. The multiple linear equation was developed as per the following equation:

$$E = -0.010 * P_{t-1} + 0.199 * RH_{t-1} + 0.008 * T_{min} \\ - 0.044 * T_{max} + 0.313 \tag{11}$$

The number of neurons in the hidden layer is explored from 1 to 10 based on the trial-and-error procedure. The transfer functions of hidden and output layers have been considered as log sigmoid and pure linear, respectively, in the training of the ANN model.

Performance evaluation of ANN and MLR models

On the basis of performance of the ANN models during calibration and validation periods, the EVAP9 model with structure ANN (4-9-1), with sigmoid as activation function and Levenberg–Marquardt as a learning algorithm, was selected as the best performing model among the all ANN models, and the values of r, CE and RMSE for training and validation periods were found to be 0.885, 0.785 and 1.001 mm/day and 0.889, 0.782 and 1.005 mm/day, respectively. The MLR model with values of r, CE and RMSE for training and validation periods, 0.835, 0.698 and 1.187 mm/day, and 0.866, 0.750 and 1.152 mm/day, respectively, as given in Table 2 was selected from all the structures. Even the ANN structure 4-7-1 with seven neurons in the hidden layer had given a better result, but the differences between the results of these two structures with different numbers of neurons in the hidden layer were negligible, and also further iterating the model with more than 10 neurons in the hidden layer, the performance of the model was fluctuating (decreasing and then is increasing) which might have led to the overfitting. The performance of the

Table 2 | Comparison of ANN models during calibration and validation periods

Model no.	ANN structure	Calibration			Validation		
		r	CE	RMSE	r	CE	RMSE
EVAP1	4-1-1	0.875	0.766	1.05	0.885	0.774	1.02
EVAP2	4-2-1	0.879	0.773	1.03	0.888	0.782	1.01
EVAP3	4-3-1	0.880	0.776	1.02	0.889	0.781	1.01
EVAP4	4-4-1	0.882	0.780	1.02	0.889	0.783	1.00
EVAP5	4-5-1	0.882	0.779	1.02	0.889	0.783	1.00
EVAP6	4-6-1	0.883	0.781	1.01	0.889	0.782	1.01
EVAP7	4-7-1	0.884	0.783	1.01	0.889	0.782	1.01
EVAP8	4-8-1	0.883	0.782	1.01	0.890	0.784	1.00
EVAP9	4-9-1	0.885	0.785	1.00	0.889	0.782	1.01
EVAP10	4-10-1	0.882	0.780	1.01	0.882	0.786	0.99

best ANN structure (4-9-1) with nine neurons in the hidden layer and the MLR model for the estimation of daily pan evaporation at Roorkee during calibration and validation were plotted against observed daily pan evaporation as depicted in Figures 5–10. The scatter plots and graphs between the ordinates of the best ANN model, the MLR model and observed evaporation during calibration and validation demonstrate the potentiality of the developed ANN model over the MLR model in the estimation of the daily

pan evaporation. The results of the calibration and validation of the best ANN and MLR models in terms of various statistical indices are presented in Table 3.

Sensitivity analysis

To check the effect of each input parameter on the pan evaporation, sensitivity analysis has been performed by utilizing the selected ANN (4-9-1) model. The effectiveness of each
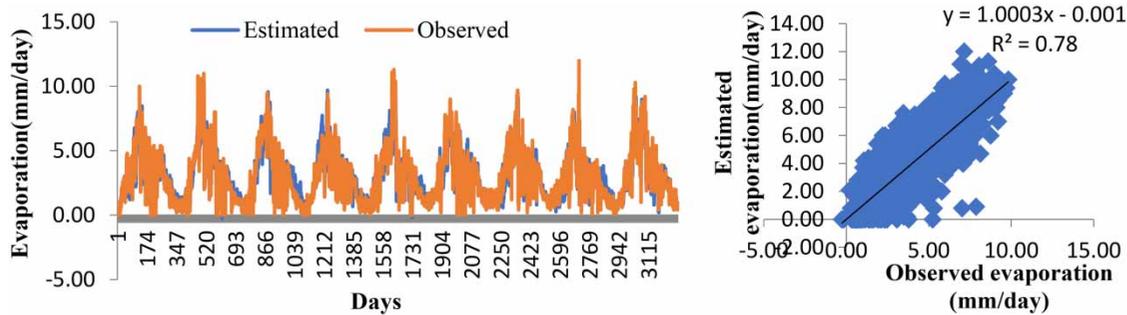

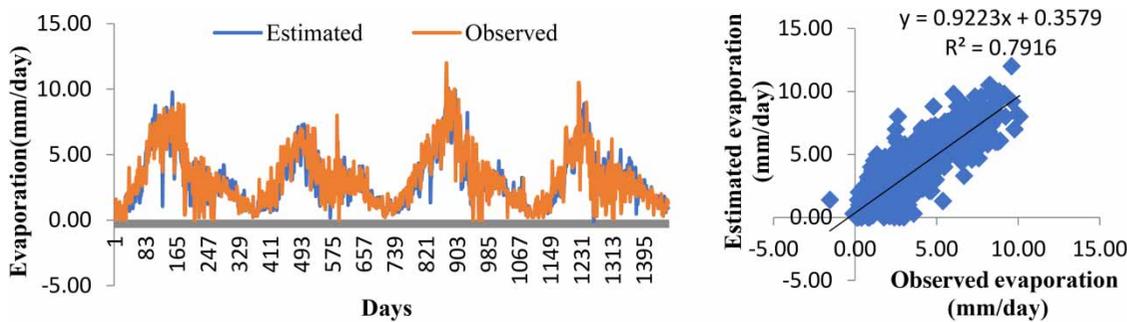**Figure 5** | Line diagram and scatter plot of the selected ANN (4-9-1) model during the training period.

Figure 6 | Line diagram and scatter plot of the selected ANN (4-9-1) model during the validation period.
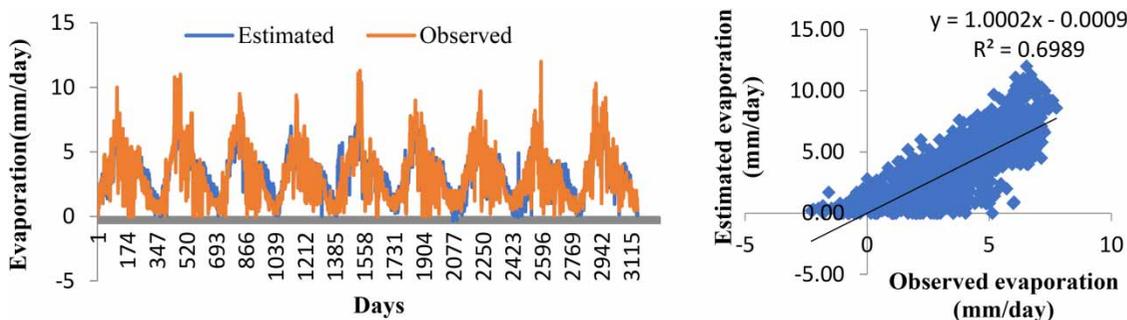

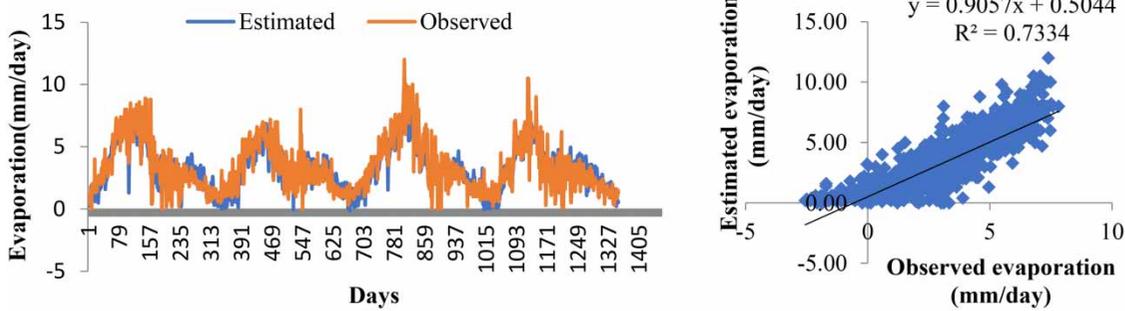**Figure 7** | Line diagram and scatter plot of the MLR model during the training period.

Figure 8 | Line diagram and scatter plot of the MLR model during the validation period.
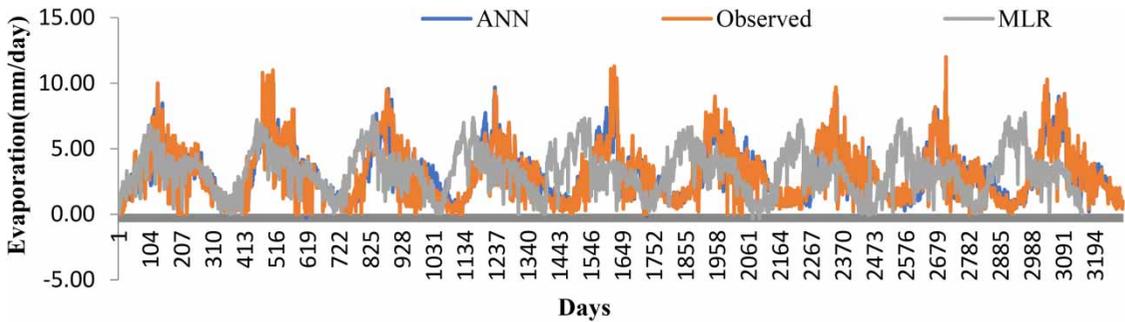

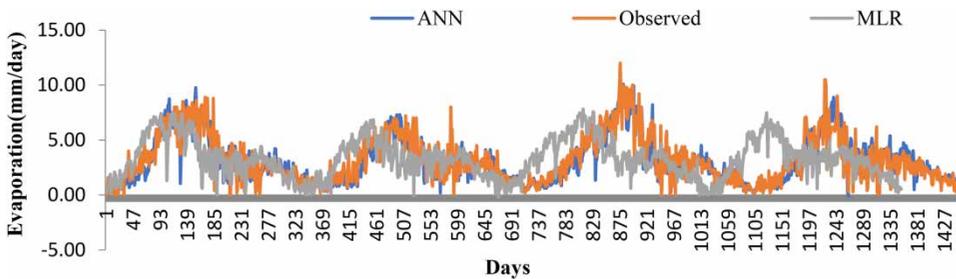**Figure 9** | Comparison of selected ANN (4-9-1) and MLR models during the training period.

Figure 10 | Comparison of selected ANN (4-9-1) and MLR models during the validation period.

Table 3 | Comparisons between selected ANN and MLR models

| Model | r | Calibration | | r | Validation | |
		RMSE (mm/day)	CE		RMSE (mm/day)	CE
ANNEVAP4 (4-9-1)	0.885	1.00	0.785	0.889	1.01	0.782
MLR EVAP	0.835	1.19	0.698	0.856	1.15	0.750

parameter was judged by skipping parameters one by one from the selected model with four inputs as depicted in Table 4. These models were explored in the ANN with the sigmoid activation function and Levenberg–Marquardt as a learning algorithm with nine neurons in the hidden layer. The model without RH_{t-1} has a very low value of r (0.78) and CE (0.723) during training. Similar values of r and CE were observed during testing periods. The higher value of

Table 4 | Comparison of results for sensitivity analysis of selected model ANN (4-9-1)

Input variables	Calibration			Validation		
	r	RMSE (mm/day)	CE	r	RMSE (mm/day)	CE
P_{t-1}, RH_{t-1}, T_{max}, T_{min}	0.885	1.00	0.785	0.889	1.01	0.782
RH_{t-1}, T_{max}, T_{min}	0.810	1.26	0.763	0.742	1.18	0.615
P_{t-1}, T_{max}, T_{min}	0.780	1.30	0.723	0.729	1.26	0.737
P_{t-1}, RH_{t-1}, T_{min}	0.854	1.12	0.743	0.786	1.28	0.715
P_{t-1}, RH_{t-1}, T_{max}	0.860	1.16	0.743	0.792	1.18	0.745

RMSE was observed with the variables one day-lagged rainfall (P_{t-1}), current day maximum temperature (T_{max}) and current day minimum temperature (T_{min}) by skipping RH_{t-1} as compared with a model with all variables.

CONCLUSIONS

The estimation of evaporation at the regional level is cumbersome with limited meteorological data due to the complex nature of the evaporation process. It can easily be performed using an ANN with available measured meteorological parameters. Evaporation is highly correlated with current day minimum and maximum temperatures as compared with one day-lagged rainfall and RH. The ANN model with a single hidden layer having nine neurons was the best architecture among several architectures tested using log-sigmoid activation function and Levenberg–Marquardt learning algorithm. The selected ANN model showed greater capability for the estimation of evaporation, when compared with the MLR-based model, due to higher optimizing capacity to reach the desired result with fewer requirements of input parameters under validation. The model efficiency and correlation coefficient of the ANN model was higher than the MLR model during calibration and validation, but the RMSE value was higher in the MLR model. One day-lagged RH is the most sensitive variable followed by day-lagged rainfall, current day maximum temperature and current day minimum temperature. Results

of the study for the estimation of evaporation are of great importance for planning and management of irrigation management projects.

REFERENCES

Affandi, A. K., Watanabe, K. & Tirtomihardjo, H. 2008 Application of an artificial neural network to estimate groundwater level fluctuation. *Journal of Spatial Hydrology* **7** (2), 23–46.

ASCE Task Committee on Application of Artificial Neural Networks in Hydrology 2000 Artificial neural networks in hydrology II: hydrologic applications. *Journal of Hydrologic Engineering* **5** (2), 124–137.

Chiew, F. H. S., Stewardson, M. J. & McMahonn, T. A. 1993 Comparison of six rainfall-runoff modeling approaches. *Journal of Hydrology* **147** (1–4), 1–36.

Dawson, C. W. & Wilby, R. 1998 An artificial neural network approach to rainfall-runoff modeling. *Hydrological Sciences Journal* **43** (1), 47–66.

Gommes, R. 1998 *Agrometeorological Crop Yield Forecasting Methods*. International Statistical Institute, Voorburg, The Netherlands.

Goyal, M. K., Bharti, B., Quilty, J., Adamowski, J. & Pandey, A. 2014 Modeling of daily pan evaporation in subtropical climates using ANN LS-SVR, Fuzzy Logic, ANFIS expert systems with applications. *Expert Systems with Applications* **41** (11), 5267–5276.

Guven, A. & Kisi, O. 2011 Daily pan evaporation modeling using linear genetic programming technique. *Irrigation Science* **29** (2), 35–145.

Higham, D. J. & Higham, N. J. 2016 *MATLAB Guide*, Vol. 150. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, pp. xxvi+476.

Hornik, K. 1991 Approximation capabilities of multilayer feed forward networks. *Neural Networks* **4** (2), 251–257.

Kisi, O. 2009 Daily pan evaporation modelling using multi-layer perceptrons and radial basis neural networks. *Hydrological Processes* **23** (2), 213–223.

McCulloh, W. S. & Pitts, W. 1943 A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics* **5** (4), 115–133.

Moghaddamnia, A., Gousheh, M. G., Piri, J., Amin, S. & Han, D. 2009 Evaporation estimation using artificial neural networks and adaptive neuro-fuzzy inference system techniques. *Advances in Water Resources* **32** (1), 88–97.

NIH 2018 Roorkee, India. Available from: http://nihroorkee.gov.in/location.html (accessed 27 September 2018).

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.

Seifi, A. & Riahi, H. 2018 Estimating daily reference evapotranspiration using hybrid gamma test-least square support vector machine, gamma test-ANN, and gamma

test-ANFIS models in an arid area of Iran. *Journal of Water and Climate Change*. https://doi.org/10.2166/wcc.2018.003.

Shirgure, P. S. & Rajput, G. S. 2012 Prediction of daily pan evaporation using neural networks models. *Scientific Journal of Agricultural* **1** (5), 126–137.

Shiri, J., Dierickx, W., Baba, A. P. A., Neamati, S. & Ghorbani, M. A. 2011 Estimating daily pan evaporation from climatic data of the State of Illinois USA using adaptive neuro-fuzzy inference system (ANFIS) and artificial neural network (ANN). *Hydrology Research* **42** (6), 491–502.

Shirsath, P. B. & Singh, A. K. 2010 A comparative study of daily pan evaporation estimation using ANN, regression and climate based models. *Water Resources Management* **24** (8), 1571–1581.

Sudheer, K. P., Gosain, A. K., Mohana, R. D. & Saheb, S. M. 2002 Modelling evaporation using an artificial neural network algorithm. *Hydrological Processes* **16** (16), 3189–3202.

Tabari, H., Marofi, S. & Sabziparvar, A. A. 2010 Estimation of daily pan evaporation using artificial neural network and multivariate non-linear regression. *Irrigation Science* **28** (5), 399–406.

Terzi, O. & Keskin, M. E. 2005 Modelling of daily pan evaporation. *Journal of Applied Science* **5** (2), 368–372.

Xu, C. Y. & Singh, V. P. 1998 Dependence of evaporation on meteorological variables at different time scales and inter comparison of estimation methods. *Hydrological Processes* **12** (3), 429–442.

Xu, C. Y. & Singh, V. P. 2001 Evaluation and generalization of temperature-based methods for calculating evaporation. *Hydrological Processes* **15** (2), 305–319.

Xu, C. Y. & Singh, V. P. 2005 Evaluation of three complementary relationship evapotranspiration models by water balance approach to estimate actual regional evapotranspiration in different climatic regions. *Journal of Hydrology* **308** (1–4), 105–121.

Yihdego, Y. & Webb, J. A. 2017 Comparison of evaporation rate on open water bodies: energy balance estimate versus measured pan. *Journal of Water and Climate Change* **9** (1), 101–111. https://doi.org/10.2166/wcc.2017.139.

Zhang, G., Patuwo, B. E. & Hu, M. Y. 1998 Forecasting with artificial neural networks: the state of the art. *International Journal of Forecasting* **14** (1), 35–62.