Journal of Artificial Intelligence Research & Advances
ISSN: 2395-6720 (Online)
Volume 5, Issue 2
www.stmjournals.com

Evaporation Estimation Using Multilayer Neural Network: A Case Study for Chhattisgarh Plains

Diwakar Naidu**, *Babita Majhi
Department of Computer Science and Information Technology, Guru Ghasidas University, Bilaspur, Chhattisgarh, India

Abstract
Evaporation being one of the most significant elements of the hydrological cycle, it is important to recognize its relationship with other climatic factors for developing schemes for effective water resource management, especially for farming uses. There are few direct and indirect methods in practice for the measurement or computation of evaporation losses. This study is conducted to demonstrate the ability of multilayer neural network (MLNN) with back propagation weight update mechanism to estimate the evaporation losses in Chhattisgarh Plains region using other climatic factors. The meteorological parameters considered for estimation of evaporation are temperature, humidity, vapor pressure, wind speed and sunshine hours. Results indicate that the MLNN model with complete set of meteorological parameters is able to estimate evaporation losses precisely with monthly and weekly time series. However, the results are a little inferior with daily time series. It is also evident that with a smaller number of input features, the performance accuracy of MLNN reduces in estimating evaporation.

Keywords: *Evaporation estimation, multilayer neural network, back propagation*

****Author for Correspondence*** E-mail: dnaidu1971@gmail.com

INTRODUCTION

Under the influence of climate change, the rainfall distribution in different parts of the world has become quite erratic during the recent decades, which leads to drought conditions for agricultural crops. In order to effectively utilize the available water resources, proper understanding of evaporation losses and factors affecting it is extremely essential, especially for water resource management, drought alleviation and studying the crop water requirement for irrigation scheduling in different crops. Evaporation can be measured directly using the instrument called USWB Class-A open pan, recommended by WMO and the International Association of Hydrological Science, worldwide. Despite being the best option for measurement of evaporation using direct method, due to practical difficulties such as maintenance of instrument and availability of trained human resources for accurate measurement of evaporation, estimation of evaporation indirectly using the climatic factors with empirical methods is also in practice subjected to availability of sufficient meteorological data for its computation.The main weather parameters affecting evaporation are temperature, relative humidity, vapor pressure, wind speed and sunshine hours and solar radiation. As the evaporation process is very complex and highly non-linear in nature, it is often very difficult to make precise estimation of evaporation losses for a given location using empirical methods [1]. Hence, attempts are being made by many scientists to find suitable technique which can demonstrate the ability to model non-linear relationship associated between evaporation and other climatic factors[2].

In recent decade, soft computing techniques including fuzzy logic, neural network, evolutionary computing etc., have emerged as an alternate methodology for estimation and prediction of evaporation losses using the climatic factors. Kisi *et al*. had modeled monthly evaporation from climatic data using neural networks and decision tree for different location and concluded that neural network model has outperformed other two decision tree models, viz., chi-squared automatic interaction detector (CHAID) and classification and regression tree (C&RT) investigated in the study [3, 4]. Guven*etal*. had investigated the ability of different soft computing models and empirical methods in estimating total evaporation losses at various locations and found that different variant of linear genetic programming (LGP) models have performed better than fuzzy logic (FG), adaptive neuro-fuzzy inference system (ANFIS), artificial neural network (ANN) andStephens and Stewart empirical model[5]. Keshtegar *et al*. have suggested three nonlinear mathematical models for prediction of daily pan evaporation losses using the regression based conjugate gradient optimization technique [6]. The results obtained using these models are found to be better as compared to ANFIS, M5 Model Tree in predicting daily pan evaporation values. Wang *et al.* have also modeled pan evaporation and examine the ability of different heuristic soft computing techniques (Multilayer perception, generalized regression neural network, Least square support vector machine, fuzzy genetics, multivariate adaptive regression spline and ANFIS with grid partition) and regression methods (Multiple linear regression and Stephens and Stewart Model) and found that soft computing techniques performed better as compared to the regression model [7, 8]. MLNN performed better as compared to other soft computing techniques considered in their study. Kumar *et al.* have examined the ability of the multilayer perceptron model with back propagation technique for different input combinations of diverse locations and performed sensitivity analysis for ranking of input combinations based on their contribution in modeling the weekly evaporation and found the mean temperature as the most influential climatic factor for estimation of evaporation as compared to other variables [9, 10]. Kumar *et al*. has also compared the performance of local linear regression (LLR) with feed forward neural network model with different input combinations and recommended LLR because of its low complexity over ANN in predicting evaporation losses [9, 10].

Being motivated by the reported literatures on the use of soft computing techniques for estimation of evaporation, a case study is carried out in this paper, particularly for Chhattisgarh Plains for estimating evaporation losses using multilayer neural network. The paper is organized in four sections. The first section introduces the need of the present study and a brief review of the available literatures on the topic. The details of the datasets used, the design of the proposed MLNN based predictor, training and testing procedures and performance evaluation criteria are described in the second section. The results and discussions on simulation studies and performance of the proposed model are reported in the third section, followed by concluding remarks in the subsequent section.

METHODOLOGY

Design of Multilayer Neural Network as a Predictor

Artificial Neural Network (ANN) was introduced by McCulloch and Pitts [11]. It is modeled on the neurons present in the human brain, which are interconnected to form a complex network where each neuron acts as a processing unit. Each unit or node produces an output signal, if it receives a sufficiently strong input signal from the other nodes to which it is connected.

The predictor using a multilayer neural network is shown in Figure 1. It consists of an input layer, one or more hidden layers and an output layer. The training of the network is done by back-propagation algorithm which is based on the error-correcting learning rule to update the weights and biases of different layers. Basically, there are two passes through the different layers of the network: forward pass and the backward pass. The forward pass produces an estimated output, which is a linear combiner of weighted sum of the input signal and backward pass adjusts the weights and biases through back propagation algorithm to minimize the error between estimated and desired output.

Data Collection

Chhattisgarh state of east central India has three agro-climatic zones viz., Chhattisgarh plain, Bastar plateau and the Northern Hills. The present investigation is aimed at estimating evaporation losses daily for Raipur

81.38°E) which represents the Chhattisgarh Plains agro-climatic zone region of the state. Daily meteorological data on maximum temperature (°C), minimum temperature (°C), relative humidity (%) for morning and afternoon hours, vapour pressure (mm) for morning and evening hours, wind speed (kmph), sunshine hours and Pan evaporation (mm) during the period 1981–2015 are collected. These weather observations are recorded at Agro-meteorological observatory Raipur which is maintained by Indira Gandhi Agricultural University, Raipur and certified by India Meteorological Department, Pune. Weekly and monthly time series are computed by averaging individual parameters considered for the study for each station from the available daily data.

Summary statistics on mean, minimum, maximum, range, standard deviation (SD), coefficient of variation (CV) and correlation coefficient (R) of different meteorological parameters considered for this study are shown in Table 1. It is evident from the descriptive statistics that there is not much difference between daily mean values of different meteorological parameters as compared to their weekly and monthly average as weekly and monthly times series are obtained by averaging the daily values. However, the range (difference between maximum and minimum values) of each weather parameter decreased, when the daily time series is converted to weekly and monthly time series. Daily evaporation values range between 0.1 and 21.5 mm/day in daily time series, whereas the same for weekly and monthly time series is between 1.5 and 18.2 and 2.3 to 14.8 mm/day respectively. It can be seen from the table that standard deviation values for temperature (°C) and vapor pressure (mm) computed around 5 to 6, whereas wind speed (kmph), sunshine (h) and evaporation (mm) it is around 2.5 to nearly 4. SD values for humidity (%) ranges from 16 to 24. CV of minimum temperature is more as compared to maximum temperature; similar is the case between morning and afternoon humidity and vapor pressure. The higher CV values are observed for both wind speed and evaporation (CV above 0.5) as compared to other parameters, where CV values range from 0.14 to 0.50 for different meteorological parameters. Correlation coefficient (R) between evaporation and other parameters are also worked out. Maximum temperature is highly and positively correlated with evaporation (>0.9). Minimum temperature, wind speed and sunshine hours also showed positive correlation with evaporation. Relative humidity values during morning hours seems to have highly negative (>–0.85) correlation with evaporation data. Relative humidity during afternoon hours, vapour pressure during both morning and afternoon does not exhibit any linear relationship between evaporation and are negatively correlated.

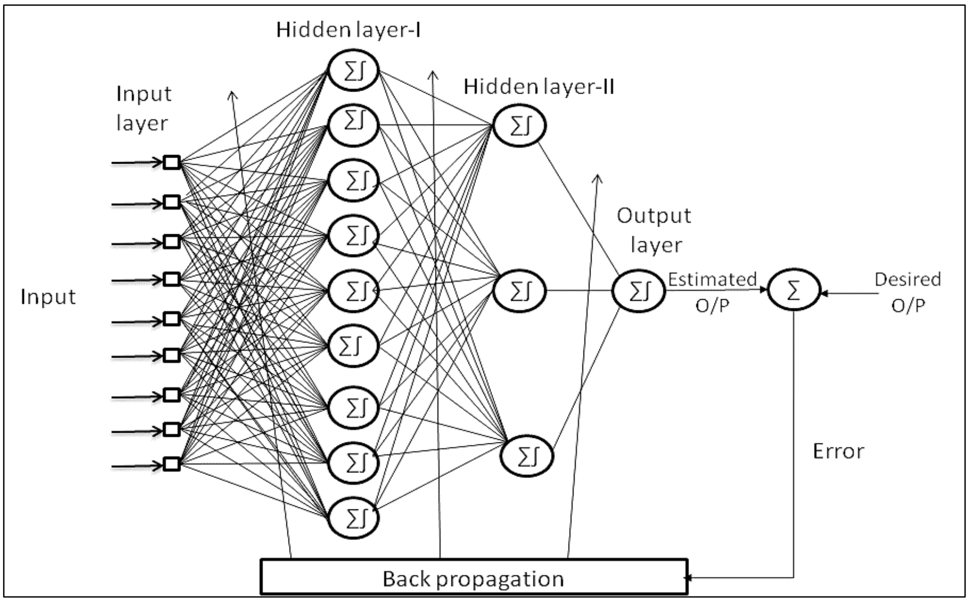

Fig. 1: *Structure of Multilayer Neural Network with Back Propagatation.*

Table 1: *Descriptive Statistics of Meteorological Parameters at Raipur.*

Datasets	Variable	Mean	Minimum	Maximum	Range	SD	CV	R
Daily	Maximum Temperature (°C)	32.67	15.40	47.60	32.2	5.27	0.16	0.90
Weekly		32.69	21.00	46.40	25.4	5.01	0.15	0.94
Monthly		32.67	24.10	44.10	20.0	4.65	0.14	0.96
Daily	Minimum Temperature (°C)	19.80	3.60	34.10	30.5	6.00	0.30	0.48
Weekly		19.83	5.80	31.50	25.7	5.78	0.29	0.52
Monthly		19.78	8.20	28.60	20.4	5.56	0.28	0.54
Daily	Relative Humidity Morning (%)	79.40	23.30	97.00	73.7	17.7	0.20	–0.85
Weekly		79.40	32.70	95.20	62.5	16.4	0.20	–0.93
Monthly		79.40	9.00	100.00	91.0	18.8	0.20	–0.95
Daily	Relative Humidity afternoon (%)	44.20	3.00	100.00	97.0	24.1	0.50	–0.53
Weekly		44.20	5.40	91.00	85.6	22.4	0.50	–0.58
Monthly		44.10	9.70	82.90	73.2	20.6	0.50	–0.56
Daily	Vapour Pressure morning (mm)	15.95	4.50	27.70	23.2	5.46	0.34	–0.12
Weekly		15.97	6.10	24.80	18.7	5.24	0.33	–0.12
Monthly		15.92	7.40	23.80	16.4	5.00	0.31	–0.90
Daily	Vapour Pressure afternoon (mm)	14.50	1.30	29.90	28.6	6.47	0.45	–0.28
Weekly		14.52	3.00	25.60	22.6	6.15	0.42	–0.30
Monthly		14.47	5.20	24.60	19.4	5.79	0.40	–0.27
Daily	Wind Speed (kmph)	5.52	0.10	31.40	31.3	3.90	0.71	0.36
Weekly		5.53	0.50	21.70	21.2	3.44	0.62	0.38
Monthly		5.51	0.90	13.60	12.7	3.14	0.57	0.47
Daily	Bright Sunshine hours	6.91	0.00	12.00	12.0	3.40	0.49	0.33
Weekly		6.90	0.00	11.20	11.2	2.77	0.40	0.32
Monthly		6.92	0.70	10.40	9.7	2.39	0.35	0.27
Daily	Evaporation (mm/day)	5.50	0.10	21.50	21.4	3.55	0.65	1.0
Weekly		5.51	1.50	18.20	16.7	3.33	0.60	1.0
Monthly		5.50	2.30	14.80	12.5	3.09	0.56	1.0

Training and Testing of the Multilayer Neural Network Based Predictor

Out of 35 years of meteorological data of Raipur from 1981 to 2015, 30 years of data (from 1981 to 2010) are used to train the model and the remaining five years data (from 2011 to 2015) are used for testing purposes. The details about the datasets are given in Table 2. A 9-3-1 structure of MLNN (9 neurons in the first hidden layer, three neurons in the second hidden layer and one neuron at the output layers) is considered in this study with different input combinations. Input patterns are normalized to lie between –1.0 and 1.0. Weights and biases are also initialized to random values lying between –1.0 and 1.0 for different layers. In order to train the model, the first pattern is given as input layer of the MLNN and after the forward pass through intermediate hidden and output layer neurons, an estimated output is obtained at the output layer. The estimated output is compared with the desired values for the corresponding input pattern and an error value is computed. During the backward pass, the back-propagation algorithm is applied to update weight and biases of each layer of the MLNN. This process is repeated, and all the remaining input patterns are sequentially fed into the MLNN model and in each case, error values are obtained after each forward pass and weights and biases are updated with each backward pass using back propagation. This completes an epoch or iteration. The training is continued for 10000 epochs. The value of the convergence coefficient (μ) is fixed at 0.01. In order to understand the learning characteristic of the predictor, after each epoch, RMSE values are computed and plotted against iterations.

Table 2: *Details of the Datasets.*

Temporal Scale	Total No. of Patterns Generated	Total No. of Patterns Used for Training	Period of Training Dataset	Total No. of Patterns Used for Testing	Period of Testing Data Set
Daily	12783 days	10958	1981–2010 (30 years)	1825	2011–2015 (05 years)
Weekly	1820 weeks	1560		260	
Monthly	420 months	360		60	

understand the learning characteristic of the predictor, after each epoch, RMSE values are computed and plotted against iterations. The training process is stopped when the RMSE values reached a desired minimum value. This completes the training process and the weights and biases values then obtained represent the model parameters. After the training process, the weights and biases at each layer of the MLNN are fixed to test the model and test patterns are fed sequentially. For each test pattern, estimated output is obtained and compared with the desired values and error is computed. After computing error values for each test pattern, RMSE value is computed to evaluate the model performance.

Performance Evaluation Criteria

Performance of based predictor is evaluated by computing root mean square error (RMSE) coefficient of determination between desired and estimated values of evaporation for the datasets considered. Mathematical expression for different evaluation criteria is as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(Out_d - Out_o)^2} \quad (1)$$

$$R^2 = \frac{\left(\sum_{i=1}^{N}(Out_d - \overline{Out_d})\,(Out_o - \overline{Out_o})\right)^2}{\sum_{i=1}^{n}(Out_d - \overline{Out_d})^2\ \sum_{i=1}^{n}(Out_o - \overline{Out_o})^2} \quad (2)$$

Where, Out_d and Out_o represents the desired and estimated evaporation values respectively. N is the total number of input patterns and I denotes the number of particular instances of input pattern.

SIMULATION RESULTS AND DISCUSSIONS

Simulation studies are conducted with different combinations of input parameters and the performance measure of each model is shown in Table 3 in terms of RMSE values. It can be seen from the table that minimum RMSE values are obtained with monthly time series dataset followed by weekly time series during both training and testing phase, whereas comparatively higher RMSE values are obtained with daily time series datasets with different models. Model-V performed better both during training and testing process with all datasets and includes all the meteorological parameters considered in this study as input combination, i.e. temperature, humidity, vapor pressure, wind speed and sunshine hours. However, the performance of the MLNN decreases as the numbers of meteorological parameters are excluded in subsequent models with same structure. RMSE values with daily dataset are decreased from 2.00 to 1.05 and 1.79 to 1.20 mm/day during training and testing phase respectively. Same with weekly dataset are decreased from 0.91 to 0.47 and 0.67 to 0.57 during training and testing phase respectively. Similarly, the RMSE values with monthly data sets during training and testing phase are reduced from 0.67 to 0.30 and 0.49 to 0.43 respectively. Lowest RMSE values are obtained with monthly dataset. Training RMSE values are decreased from 0.67 to 0.30 mm/day and testing RMSE values are decreased from 0.49 to 0.43 mm/day. Wang *et al.* have reported low training RMSE values in the range of 0.16 to 0.54 mm/day and low testing RMSE values in the range of 0.54 to 0.90 mm/day at different location in China using MLNN model with full meteorological parameters, which is found to be the best model among the six other soft computing models considered for the study [7, 8]. Similarly, Kumar *et al.* have also reported mean absolute percentage error (MAPE) values in the range 2.15 to 0.61 for different locations in India [9, 10]. A comparison of observed and estimated values of evaporation for daily, weekly and monthly time series are shown in Figure 2(a–c) respectively. It can be seen that both observed and estimated values of evaporation during the testing phase are in good agreement with each other. The relationship between observed and estimated values of evaporation is also depicted in Figure 3(a–c) for daily, weekly and monthly time series respectively. Coefficient of determination (R^2) between observed and estimated values of evaporation are calculated as 0.892, 0.964 and 0.979 with daily, weekly and monthly datasets respectively. From the simulation studies it is exhibited that the MLNN is a good candidate for predictor of evaporation.

Table 3: *Training and Testing Performance (RMSE Values in mm/day) of MLNN.*

MLNN Models	Daily		Weekly		Monthly	
	Training	Testing	Training	Testing	Training	Testing
Model-I	2.00	1.79	0.91	0.69	0.67	0.49
Model-II	1.63	1.45	0.78	0.69	0.65	0.50
Model-III	1.66	1.51	0.78	0.67	0.61	0.50
Model-IV	1.03	1.09	0.50	0.58	0.31	0.47
Model-V	1.05	1.20	**0.47**	0.57	**0.30**	**0.43**

Model-I*: Temperature only,* ***Model-II****: Temperature and Humidity,* ***Model-III****: Temperature, Humidity and Vapour Pressure,* ***Model-IV****: Temperature, Humidity, Vapour Pressure and Wind Speed, Model-V: Temperature, Humidity, Vapour Pressure, Wind Speed and Sunshine Hours.*

Journal of Artificial Intelligence Research & Advances
Volume 5, Issue 2
ISSN: 2395-6720 (Online)

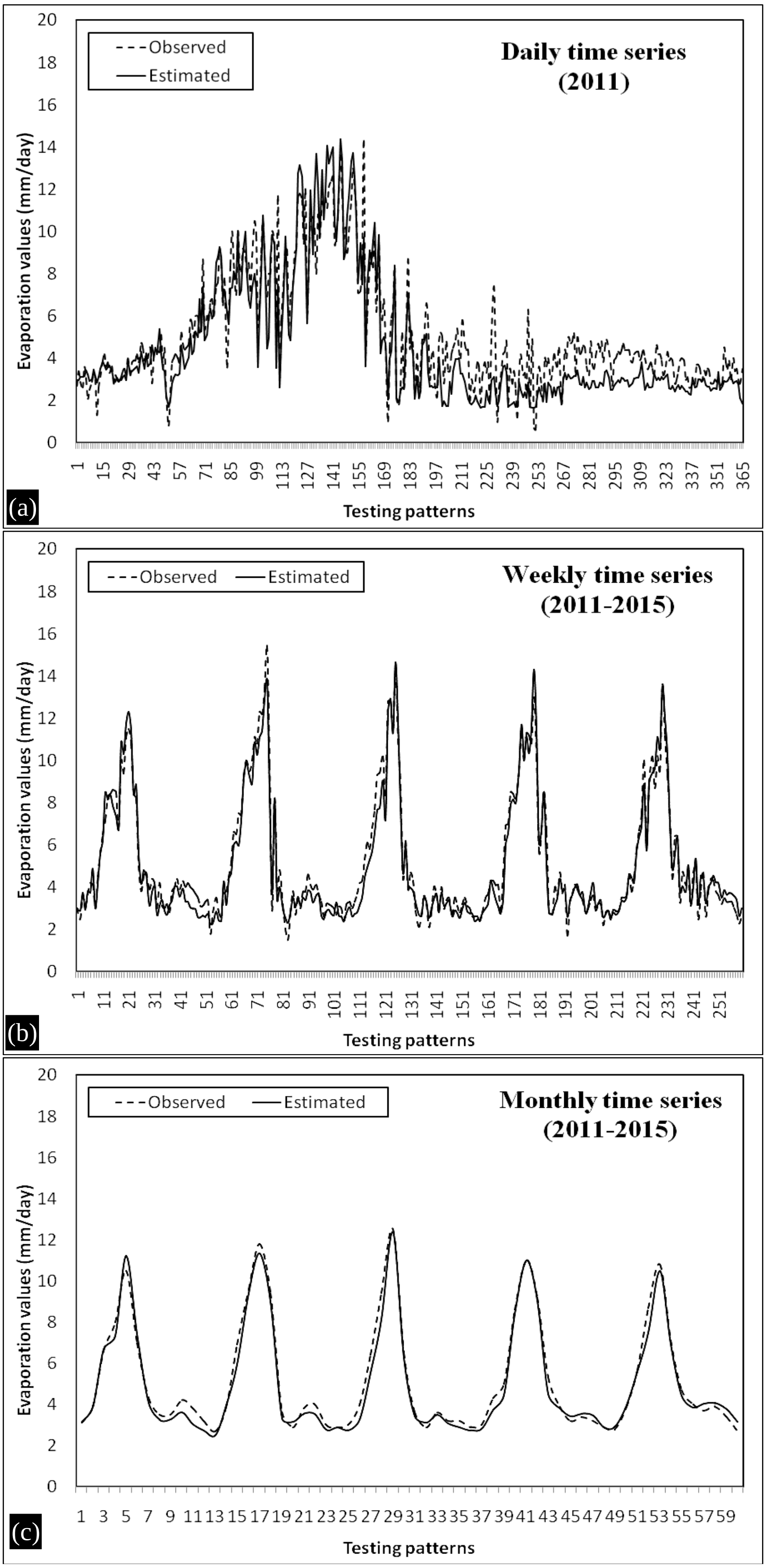

Fig. 2: *Comparison between Observed Predicted Evaporation for MLNN Model in Testing Period.*

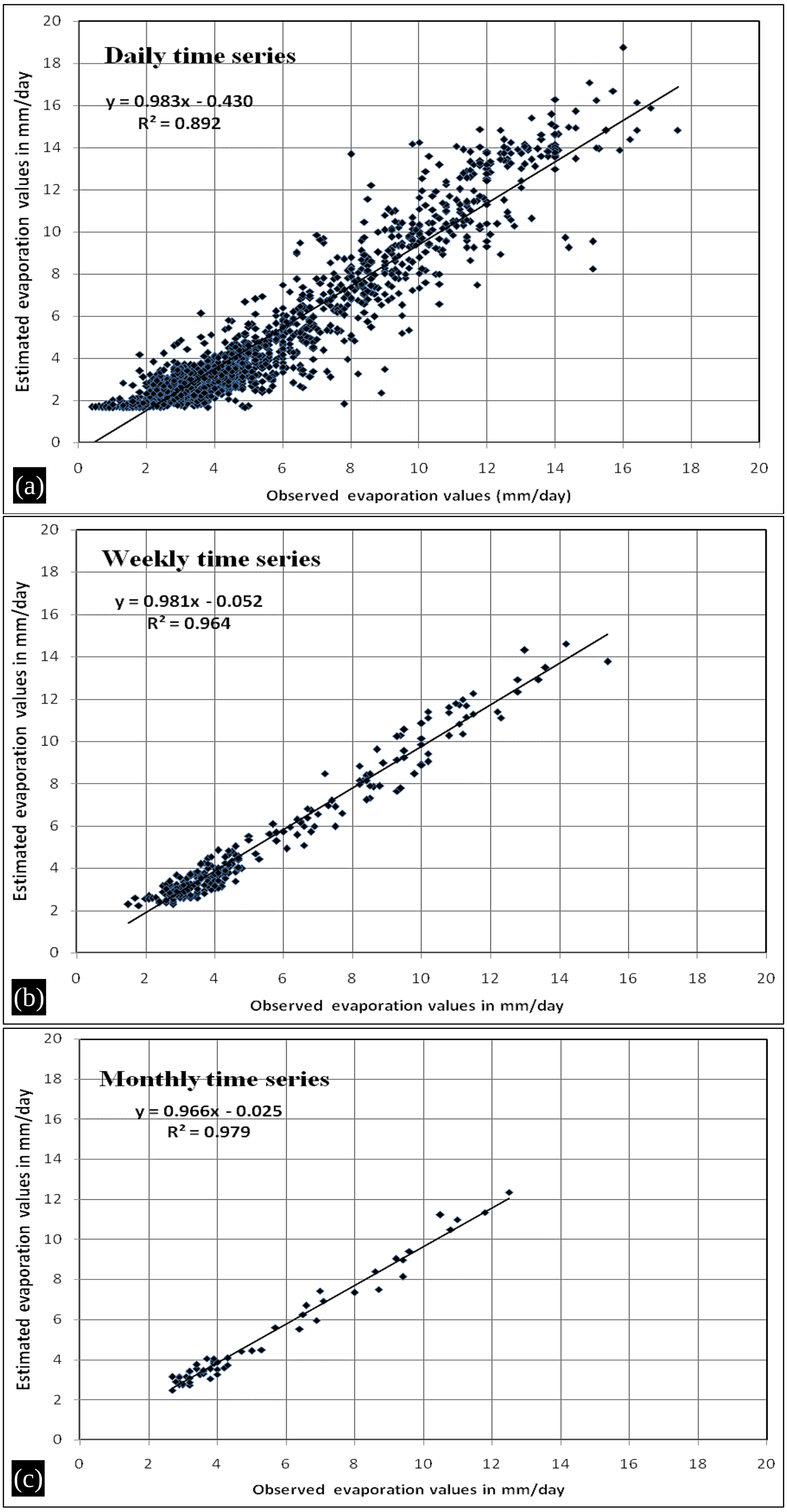

Fig. 3: *Relationship between Observed Predicted Evaporation for MLNN in Testing Period.*

CONCLUSIONS

This investigation is carried out to estimate evaporation using multilayer neural networks with back propagation-based weight update mechanism for Raipur with different input combinations of climatic factors. It has been observed that the MLNN model is able to estimate the evaporation losses more accurately when all the meteorological variables considered in the investigation are included as the input patterns. The results indicate that, MLNN model with monthly and weekly time series performed better when compared with estimated evaporation values with daily time series as input patterns. Hence it is suggested that MLNN with back propagation learning can be used to estimate weekly and monthly evaporation with low RMSE values of 0.47 and 0.30 mm/day respectively, with temperature, humidity, vapor pressure, wind speed and sunshine data as an input combination in Chhattisgarh plains agro-climatic zone of Chhattisgarh state. In future studies, we can use other neural network-based predictors for the same purpose. Also, the nature inspired algorithms like particle swarm optimization (PSO), differential evaluation (DE), cat swarm optimization (CSO) etc. can also be used for updating the weights of the neural network-based predictors.

REFERENCES

1. Kim S, Shiri J, Singh VP, *et al.* Predicting Daily Pan Evaporation by Soft Computing Models with Limited Climatic Data. *Hydrol Sci J*. 2015; 60(6): 1120–1136p.
2. Majid M, Alizadeh A, Farid A, *et al.* Estimating Evaporation from Lakes and Reservoirs under Limited Data Condition in a Semi-Arid Region. *Water Resour Manage.* 2015; 29(10): 3711–3733p.
3. Kisi O, Genc O, Dinc S, *et al.* Daily Pan Evaporation Modeling Using Chi-Squared Automatic Interaction Detector, Neural Network, Classification and Regression Tree. *J Comp Elect Agril*. 2016; 122: 112–117p.
4. Kisi O. Pan Evaporation Modeling Using Least Square Support Vector Machine, Multivariate Adaptive Regression Splines and M5 Model Tree. *J Hydrol*. 2015; 528: 312–320p.
5. Guven A, Kisi O. Monthly Pan Evaporation Modeling Using Linear Genetic Programming. *J Hydrol*. 2013; 503: 178–185p.
6. Keshtegar B, Piri J, Kisi O. A Nonlinear Mathematical Modeling of Daily Pan Evaporation Based on Conjugate Gradient Method. *J Comp Elect Agril.* 2016; 127: 120–130p.
7. Wang L, Niu Z, Kisi O, *et al.* Pan Evaporation Modeling Using Four Different Heuristic Approaches. *J Comp Elect Agril.* 2017; 140: 203–213p.
8. Wang L. Pan Evaporation Modeling Using Six Different Heuristic Computing Methods in Different Climates of China. *J Hydrol.* 2017; 544: 407–427p.
9. Kumar A, Chattopadhyay C, Mishra AK. *et al.* Neural Network Based Prediction Models for Evaporation. *Mausam*. Apr 2016; 67(2): 389–396p.
10. Kumar P, Kumar D, Panwar R. Evaporation Estimation from Climatic Factors. *Mausam.* Oct 2016; 67(4): 897–902p.
11. Mc-Culloch WS, Pitts W. A Logical Calculus of the Ideas Immanent in Nervous Activity. *B Math Biophy.* 1943; 15: 115–133p.

Cite this Article
Diwakar Naidu, Babita Majhi. Evaporation Estimation using the Multilayer Neural Network: A Case Study for Chhattisgarh Plains. *Journal of Artificial Intelligence Research & Advances*. 2018; 5(2):