Sjournals
Scientific Journals
Scientific Journal of Agricultural (2012) 1(5) 126-137
ISSN 2322-2425

Contents lists available at Sjournals

Scientific Journal of Agricultural

Journal homepage: www.Sjournals.com

Original article

Prediction of daily pan evaporation using neural networks models

P.S. Shirgure*, G.S. Rajput

National Research Centre for Citrus, Nagpur, Maharashtra, India – 440 010.

*Corresponding author: National Research Centre for Citrus, Nagpur, Maharashtra, India – 440 010.

A R T I C L E I N F O

Article history:
Received 01 October 2012
Accepted 25 October 2012
Available online 30 October 2012

Keywords:
Artificial neural networks
Back propagation algorithm
Coefficient of determination
Evaporation
Index of agreement
Modeling
Multiple linear regression
Pan evaporation

A B S T R A C T

The investigation was carried out to develop and test the daily pan evaporation prediction models using various weather parameters as input variables with artificial neural network (ANN) and validated with the independent subset of data for five different locations in India. The measured variables included daily observations of maximum and minimum temperature, maximum and minimum relative humidity, wind speed, sunshine hours, rainfall and pan evaporation. In this general model (GM) model development and evaluation has been done for the five locations viz. NRCC, Nagpur (M.S.); JNKVV, Jabalpur (M.P.); PDKV, Akola (M.S.); ICRISAT, Hyderabad (A.P.) and MPUAT, Udaipur (Raj.). The daily data of pan evaporation and other inputs for two years was considered for model development and subsequent 1-2 years data for validation. Weather data consisting of 3305 daily records from 2002 to 2006 were used to develop the GM models of daily pan evaporation. Additional weather of Nagpur station, which included 2139 daily records from 1996 - 2004, served as an independent evaluation data set for the performance of the models. The model plan strategy with all inputs has shown better performance than the reduced number of inputs. The General ANN models of daily pan evaporation with all available variables as a inputs was the most accurate model delivering an R^2 of 0.84 and a root mean square error 1.44 mm for the model development data set. The GM evaluation with Nagpur model development (NMD) data shown lowest RMSE (1.961 mm), MAE (0.038 mm) and MARE (2.30 %) and highest r (0.848), R^2 (0.719) and d (0.919) with ANN GM M-1with all input variables.

1. Introduction

The evaporation is a complex and nonlinear phenomenon involving several weather factors in the hydrological process. It is the necessary components of any water balance assessments for different water resources planning, design, operation and management studies including hydrology, agronomy, horticulture, forestry, land resources, irrigation management, flood forecasting, investigation of agro-ecosystem and modeling etc. and perhaps the most difficult to estimate owing to complex interactions between the components of the soil-water-plant-atmosphere system. Evaporation is an important variable in making the crop management decision and modeling crop response to weather conditions and climate change. It has been extensively used for estimating the potential and reference evapo-transpiration of various crops (Gavin and Agnew, 2004). Based on the limitations in measuring the pan evaporation research has been performed to model pan evaporation using various meteorological variables. It is therefore necessary to develop approaches to estimate the evaporation rates from other available meteorology variables, which are comparatively easier for measurements. The class 'A' pan evaporation is direct method. Indirect method includes those that use meteorological data to estimate evaporation from other weather variables through empirically developed methodologies or artificial neural network and statistical approaches.

The basic aim of this research investigation is to improve the effectiveness and operational ability of classic and sophisticated methods and discover perspectives of modern and sophisticated approaches such as artificial neural networks (ANN) for the prediction and estimation of pan evaporation. The regression models developed from meteorological data involve empirical relationships to some extent accounts for many local conditions. Therefore, most regression models gave reliable results when applied to climatic conditions similar to those for which they were developed. The correlations and regression studies between different meteorological parameters and evaporation measured from US Class 'A' open pan evaporimeter under Indian conditions. Linear regression methods for prediction of evaporation using weather variables has been applied for decades and are well and understood (Jhajariya *et al.,*2006; Shirgure and Rajput, 2011). The highest correlation was found with maximum temperature followed by wind speed. The coefficient of determination for maximum temperature, minimum temperature, wind speed and soil temperature at 5 cm depth was positively correlated. The relative humidity was negatively correlated (Khanikar and Nath 1998). The pan evaporation and the meteorological parameters recorded at Jabalpur revealed that the morning relative humidity and the maximum temperature have a significant influence on the rate of evaporation (Shrivastava *et al*., 2001). There was a need for developing the model which can generalize for the diversified Indian conditions. Artificial neural network (ANN) is effective tool to model non-linear process like evaporation from open pan. Pan evaporation has been widely used as an estimate for irrigation scheduling of vegetables and horticultural crops (Shirgure *et al.,* 2001).

The artificial neural networks provide better modeling flexibility than the other statistical approaches. Numerous researchers have shown applicability of multiple linear regression technique for estimating the evaporation, but very few have been seen on artificial neural networks in agricultural and hydrological processes in India. For instance multi-layered feed forward ANN with error back propagation techniques has been used for estimating evaporation (Arca *et al.,* 1998; Sudheer *et al.;* 2002; Ozlem and Evolkesk, 2005; Keskin and Terzi, 2006), evapo-transpiration (Kumar *et al.,* 2002; Sudheer *et al.;* 2003), soil water evaporation (Han and Felkar, 1997) and various neuro-computing techniques for predicting the various atmospheric processes and parameters (Gardner and Dorling, 1998). Bruton *et al.* (2000) developed ANN models to estimate daily pan evaporation using measured weather variables as inputs. The measured variables included daily observations of rainfall, temperature, relative humidity, solar radiation, and wind speed. Daily pan evaporation was also estimated using multiple linear regression and compared to the results of the ANN models. The ANN models of daily pan evaporation with all available variables as a inputs was the most accurate model delivering an r^2 of 0.717 and a root mean square error 1.11 mm for the independent evaluation data set. The accuracy of the models was reduced considerably when variables were eliminated to correspond to weather stations. Pan evaporation estimated with ANN models was slightly more accurate than the pan evaporation estimated with a multiple linear regression models. Sudheer *et al.* (2002) investigated the prediction of Class A pan evaporation using the ANN technique. The ANN back propagation algorithm has been evaluated for its applicability for predicting evaporation from minimum climatic data. The study indicated that evaporation values could be reasonably estimated using temperature data only through the ANN technique. This would be of much use in instances where data availability is limited. These features provide neural networks the potential to model complex non-linear phenomenon like prediction of daily pan evaporation using meteorological measured variables. Observations of relative humidity, solar radiation, temperature, wind speed and evaporation for the past 22 years have been used to train and test the developed networks. Results revealed that the networks were able to well learn the events they were trained to recognize. These encouraging results were supported by high values of coefficient of correlation and low mean square errors reaching 0.98 and 0.00015 respectively (Taher, 2003). Daily pan evaporation was estimated by a suitable ANN model for the meteorological data recorded from the Automated *GroWheather* meteorological station near Lake Egirdir, Turkey (Ozlem *et al.,*2005). The ANN models were developed and validated a simulation model of the evaporation rate of a Class A evaporimeter pan located near Cartagena (Southeast Spain) (Molina Martinez *et al.* 2006). Keskin and Terzi (2006) studied the ANN models and proposed as an alternative approach of evaporation estimation for prediction of daily pan evaporation estimation. The comparison shows that there is better agreement between the ANN estimations and measurements of daily pan evaporation than for other model. Deswal and Mahesh Pal (2008) studied an ANN based modeling and the influence of meteorological parameters on evaporation from a reservoir. The findings of the study also revealed the requirement of all input parameters considered together, instead of individual parameters taken one at a time. The highest correlation coefficient (0.960) along with lowest root mean square error (0.865) was obtained with the input combination of air temperature, wind speed, sunshine hours and mean relative humidity.

The basic objectives of this research work is to develop three layered feed forward with error back propagation neural network models to estimate daily pan evaporation values based on different meteorological data. The other objectives are to test the suitability of the artificial neural networks for modeling daily pan evaporation and to test and validate the developed ANN evaporation models for the input variables using the independent subset of data for other locations.

2. Materials and methods

2.1. Study area and dataset

Data were taken from agro-meteorological stations at Nagpur (National Research Centre for Citrus, Experiment Station, lat. 21.09°N, long. 79.22°W. 311.3 m amsl), Jabalpur(Jawaharlal Nehru Agricultural University, lat. 23.10°N, long. 79.58°W, 410 m amsl), Akola (Dr. Panjabrao Deshmukh Agricultural University, lat. 20.42°N. long. 77.04°W. 309 m amsl), Hyderabad (International Crop Research Institute for Semi Arid Tropics, lat. 17.53°N. long. 78.27°W. 545 m amsl) and Udaipur (Maharana Pratap University of Agriculture and Technology, lat.24.35 °N. long.73.42 °W. 582 m amsl), These locations represent the dry sub-humid, sub-humid, semi-arid and arid climatic regions within the state of Maharashtra, Madhya Pradesh, Andhra Pradesh and Rajasthan. The intent was to develop a single model that could be used for any of the agro-climatic zones of India. The data were partitioned into a model development data set and an independent evaluation data set. The model development data set consisted of 607, 731, 507, 730 and 730 days of data (3305 total observations) for Nagpur, Jabalpur, Akola, Hyderabad and Udaipur respectively, from 2002 to 2006. The model development data set was further divided by randomly placing 80% of the observations in a training data set and the remainder 20 % in a testing data set. The training data set was used to develop the neural network models. The testing set was used to evaluate the accuracy of the ANN models during training in order to determine when to stop the training, Training was continued as long as the minimum goal of error of the ANN estimate of pan evaporation on was kept as 0.001. The model development data set was used to choose preferred ANN training parameters and to develop ANN models of pan evaporation based on various numbers of inputs. With this procedure, all model development and parameter selection were done with the model development data set. The Nagpur model development (NMD) data size was 2139. The model GM evaluation data set consisted of 2066 observations for Nagpur, Jabalpur, Akola, Hyderabad and Udaipur respectively, from 2004 to 2007 for the five locations. Weather variables included in the evaluation data set were the same as those in the model development data set. The evaluation data set was not used in model development in any way and was only used to assess the accuracy of the ANN models.

A three-layer back propagation ANN architecture was employed in all models. The MATLAB (Ver.7.4, 2007) software package was used to develop the ANN pan evaporation models, The data set included daily records of eight measured variables viz. maximum temperature (TMAX), minimum temperature (TMIN), maximum relative humidity (RMAX), minimum relative humidity (RMIN), wind speed (WISP), sunshine hours (SNHR), rainfall (RAIN) and pan evaporation (EPAN). The number of measurable weather variables used as inputs in the development of the ANN pan evaporation models (M-1 to M-7) was varied with emphasis placed on investigating models with reduced numbers of inputs, based on positive and negative effects of the parameters on evaporation process. Seven model strategies (combination of variable inputs) to be evaluated in this study is shown in Table 1.

Table 1
The modeling strategies in development of ANN pan evaporation models

Variables	**Modeling strategies**						
	m-1	**m-2**	**m-3**	**m-4**	**m-5**	**m-6**	**m-7**
tmax	√	√	√	√	√	√	√
tmin	√	√	√	√	√	-	-
rmax	√	√	√	-	√	-	-
rmin	√	√	√	-	√	-	-
wisp	√	√	√	√	-	√	√
snhr	√	√	-	√	-	√	-
rain	√	-	√	√	-	-	-

- indicates the inclusion of the variable in neural network modeling

Modeling strategy in M-1 included all seven available daily weather input variables. Model strategy M-2 considered the effect of removing rainfall from the variables in modeling strategy M-1creating the situation of non-rainy period evaporation process. Model strategy M-3 considered removing sunshine hours and inclusion of rainfall crating the weather with rainy season and without sunshine hours. The effect of absence of maximum and minimum relative humidity in the evaporation process is tested in model strategy M-4. This is like without humidity days. In modeling plan M-5 the maximum and minimum temperature as well as relative humidity was as variables, modeling the situation with temperature and humidity. Modeling strategy M-6 included maximum temperature, wind speed and sunshine hours as input variables. The model M-7 included maximum temperature and wind speed as the only observed weather values.

2.2. Model architecture

The best neural network architecture according to de Villiers and Barnard (1993), ANN models with one hidden layer can approximate any continuous function. Thus, in this study, model architecture of one hidden layer was used. The number of nodes in the hidden layer(s) should be large enough to ensure a sufficient number of degrees of freedom for the network function and small enough to minimize the risk of loss of the network's generalization ability. Furthermore, it's important to note that a useless increase of the neural network size will lead to a significant increase in training and running time. As recommended by Maier and Dandy (2001), the node number upper limit in the hidden layer(s) was fixed as follows:

$$N^H = \min (2N^I+1; N^{TR} /N^I+1)$$

Where, N^H is the number of nodes in the hidden layer(s), N^I number of input nodes and N^{TR} the number of training sample. The number of hidden nodes is therefore taken from 5 to 20 as the inputs vary from seven in model M-1 to two in model M-7. The number of training samples is large. So, the ratio of the training samples to training parameters is large and $2N^I+1$ is minimum number of hidden nodes of the networks. The accuracy increase with increasing the nodes, but the generalization ability of the model goes down. So, the optimization of nodes is done with minimum squared and absolute error. The final GM models were developed with 15 hidden nodes, which gave minimum errors. The developed model evaluation was done with the help of the Nagpur model development dataset and GM evaluation dataset. The number of records used for evaluation is 484 for NRCC, Nagpur; 487 for JNKVV, Jabalpur; 365 each for PDKV, Akola, ICRISAT, Hyderabad and MPUAT, Udaipur.

2.3. ANN modeling software source code program using MATLAB

The neural network utility file was edited in MATLAB (ver 7.4, 2007) source code program. The main GUI consisting of all the input variables selection, input data source file, network options, training and testing functions, setting the data for training, evaluation, plotting the predicted values and saving the network is created and run in MATLAB software. The mainframe is executed by running the MainGUI.mat file in MATLAB program. The Graphical User Interface (GUI) appears by changing the directory option, which gives the details of the ANN modeling selections. The program displays the title, the input variables to be selected by making the tick marks on/off. The source data input file (.csv) is browsed for the training, testing and evaluation. The variables were selected according to the model plan M-1 to M-7 for developing and evaluating the ANN models. The ANN model architecture is three layered feed forward with error back propagation. Which is most commonly used neural network for the prediction of the non-linear process like prediction of pan evaporation. The number of hidden nodes selection is done from 5 to 20. The transfer function from input to hidden layer is logistic sigmoid and from hidden layer to output layer is linear (as the output is one i.e. EPAN). The training function is *Levenberg-Marquardt*, which is most common and accurate is selected.

The performance functions for training and testing the networks used are MAE (mean absolute error) and MSE (mean squared error). The learning rate, which decides the accuracy of the training is used for optimization are 0.1, 0.3, 0.5 and 0.7. The momentum of the training the network, which gives the speed of the training varied from 0.1, 0.3, 0.5, 0.7 and 0.9. The various combinations of hidden nodes, learning rate and momentum is done to arrive at optimum combinations to give less error. The network iterations (epochs) were kept at 1000. The ANN model development and evaluation with saved model plans is done as follows. The mainframe of the ANN modeling with MATLAB is shown in Figure 1.

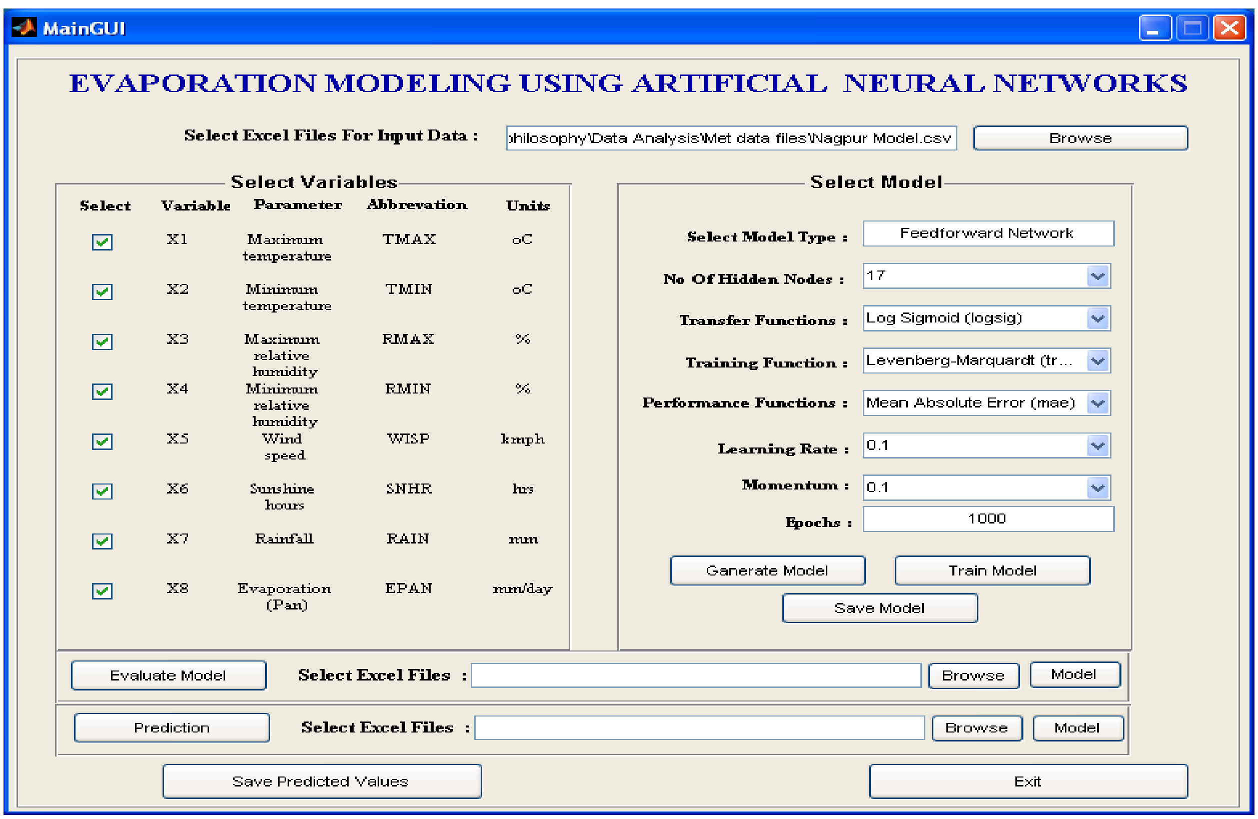

Fig. 1. The Graphical User Interface (GUI) of ANN modeling for EPAN prediction.

The GUI.mat file is opened in MATLAB software. The model development input data file (saved in .csv format) is browsed to the main GUI frame. The hidden nodes, training and transfer functions are selected. The model is generated on clicking the button Generate model. This generated model is displayed in command window of the

MATLAB program. This gives the network architecture, input, hidden and output layer weights and bias weights. The generated ANN model is trained by clicking on the Train model option. Training is undertaken by setting the percentage training, testing and validation values as 80 %, 20 % and 100 % respectively. The network training process is displayed on the screen with minimum error in comparison with the set 0.001goal. The program gives the graphical output with set goal value. The model when gives the minimum value of MSE/MAE is saved as a model with particular identification name. The GM models (GM) were saved in the name as GM 15 M-1 to GM 15 M-7. This also indicates the General model with 15 hidden nodes, 0.1 learning rate and momentum each and for the model plans.

These saved models were used for validation of the same data on screen or the other source data of the independent locations. This was done by selecting the source file and clicking on the button Evaluate model on mainframe of the program. The saved neural network model was used for evaluation of new input file, which is browsed in evaluation module of the mainframe. The output of the pan evaporation is predicted in command window of the MATLAB program. The same is copied and used for estimating the model performance like RMSE, MAE, MARE, r, R^2 and d. The graphical output of the training, testing and validation of the network as well as the plotting of the observed and ANN predicted EPAN is displayed on the desktop, which was saved for the comparison of the different ANN model plans. In this way model plans M-1 to M-7 of the General model was developed and evaluated with five different locations in India.

2.4. Model evaluation performance parameters

The ANN model development and evaluation using the independent datasets of various locations was done with help of the MATLAB source code program and the datasets concerned. The developed and evaluated models were verified with the squared and absolute statistics of different performance functions. Mean squared error (MSE), Root mean squared error (RMSE), Mean absolute relative error (MARE), Correlation coefficient (r) and coefficient of determination (R^2) are used as model development as well as evaluation criteria's. Based on these criteria's the superiority of the model is judged. The R^2 measures the degree to which two variables are linearly related. MSE and MARE provide different types of information about the predictive capabilities of the model. The MSE and RMSE measures the goodness of fit relevant to high evaporation values, whereas the MARE yields a more balanced perspective of the goodness of fit at moderate evaporation. In addition an index of agreement (d) between the observed and predicted values was calculated.

2.5. General model development

The General model (GM) is developed using daily records of the weather inputs from Nagpur, Jabalpur; Akola; Hyderabad and Udaipur. The program in MATLAB codes is opened and run for the main frame of the ANN modeling. The main frame which helped in selecting the input data file, selecting the input variables, number of the hidden nodes, training, learning and transfer functions, learning rate and momentum. The model input file which is saved in spreadsheet (with csv format) is browsed and selected. The number of hidden nodes, learning rate and momentum is kept as 15, 0.1 and 0.1 respectively, as it has shown optimum parameters. The training function is back propagation with *Levenberg-Marquardt* algorithm and transfer function is logistic sigmoid. The number of iterations (epochs) were kept 1000. The network performance functions are kept as MAE and MSE. The goal set for testing the neural networks was 0.001 (0.1 %). The data was set to train, test and validate in the ratios of 80 %, 20 % and 100 % respectively. The training network figure for lowest MSE / MAE is also saved. The process of training the network was repeated many times and when the MSE/MAE was shown minimum the ultimate network is saved in .mat file with some name. Like this GM models with 15 hidden nodes and 0.1 learning rate and 0.1 momentum were saved optimally in the names of GM 15 M-1 to GM 15 M-7. The ANN General models M-1 to M-7 has input layer and bias weights. The input weights, layer weights and bias weights of these ANN models are explored in window of the MATLAB program and saved.

3. Results and discussion

The model development data set was used to determine preferred values for number of hidden nodes, learning rate, and momentum. Data for all five sites were included to obtain the best model parameters, as well as all 7 input parameters. The ANN parameters for learning rate and momentum have values which are typically less than one. Learning rate and momentum combinations using 0.1, 0.3, 0.5, and 0.9 were tested with hidden node numbers ranging from 5 to 20. The training data were used for training and the testing data were used to determine when to stop training. The accuracy of the model on the model development data set (training data and testing data) was then determined. In the initial experimentation, it was found that the accuracy of the models was reduced for number of hidden nodes above 15, and Learning rate and momentum below 0.1. Otherwise, there was little variation in accuracy for the various combinations of parameters. The accuracy was higher for 15 hidden nodes, a learning rate of 0.1 and a momentum of 0.1. Experimentation was performed due to the limited effect of the ANN parameters on model accuracy. The accuracies of the models in feed forward mode for the development and evaluation data sets are shown in Table 2. The results for the evaluation data set provide an independent evaluation of the prediction accuracy of the models.

3.1. Error statistics of General model (GM) development

The ANN General model development data shows that the RMSE, MAE and MARE is increased and r, R^2 and d is decreased when the input variables were sequentially removed from seven variables in M-1 to two variables in M-7. The lowest RMSE (1.439 mm), MAE (0.0207 mm) and MARE (0.915 %) and highest r (0.916), R^2 (0.839) and d (0.956) was observed with GM M-1 model, in which all the input variable are taken into consideration. In model M-2, where all the six input variables excluding RAIN was considered in ANN model has also shown lower RMSE, MAE and MARE and higher r, R^2 and d. The RMSE, MAE and MARE were 1.478 mm, 0.0218 mm and 1.076 %, which is slightly more than the model M-1, but not a significant difference exists. The r, R^2 and d in model M-2 is 0.911, 0.831 and 0.953 respectively. These values are less than the values of model M-1. The RMSE, MAE and MARE of GM M-3 model are higher and r, R^2 and d are lower as compared to ANN model M-1 and M-2. This is mainly due to the replacement of sunshine hours input of model M-2 with RAIN variable. This indicates that the performance of all variables without SNHR (but with RAIN variable) shows higher RMSE, MAE and MARE and lower r, R^2 and d as compared to model M-2. The RMSE, MAE and MARE of model M-3 are 1.559 mm, 0.024 mm and 1.362 % respectively. The r, R^2 and d of model M-3 are 0.901, 0.812 and 0.947 respectively. The error statistics of model GM M-4 in which the relative humidity (RMAX and RMIN) are not the input variables. The RMSE, MAE and MARE of model M-4 is 1.625 mm, 0.026 mm and 1.506 %, which is higher than the models M-1, M-2 and M-3. The r, R^2 and d of M-4 is 0.892, 0.796 and 0.942 respectively and these values are higher than model M-1 M-2 and M-3. It is mainly due to the absence of RMAX and RMIN input variables and have higher negative correlation with the pan evaporation. Although five variables are there in model M-4, but the important variables i.e. RMAX and RMIN are not considered for modeling has shown more error.

The model M-5 performance is better than the model M-4. The RMSE, MAE and MARE in model M-5 is increased to 1.685 mm, 0.0284 mm and 1.549 % from 1.625 mm, 0.026 mm and 1.50 % in model M-4. The r, R^2 and d is also decreased to 0.883, 0.78 and 0.935 in model M-5 as compared to model M-4.The input variables are TMAX, WISP and SNHR in model M-6. The RMSE, MAE and MARE of model M-6 is 1.702 mm, 0.0289 mm and 1.996 % respectively. The r, R^2 and d of model M-6 were 0.88, 0.776 and 0.935 which are lower as compared to M-5 and higher than model M-7. The RMSE, MAE and MARE is increased to 1.759 mm, 0.0309 mm and 2.52 % in GM model M-7, in which only TMAX and WISP are the input variables of the model. The r, R^2 and d were also lowest in model M-7 as compared to M-1 to M-6. The r, R^2 and d in model M-7 was 0.872, 0.76 and 0.929 respectively. The evaluation of general model development was graphically presented in figure 2.

3.2. General model evaluation with Nagpur model development (NMD) dataset

The Nagpur model development (NMD) data from April, 1996 to March, 2004 (2139 observations) was used for evaluating the ANN General model with M-1 to M-7 model strategies. The optimum number of hidden nodes, learning rate and momentum for General model (GM) is 15, 0.1 and 0.1 respectively. The trained ANN models M-1 to M-7 with minimum error reaching towards the goal. Each saved model architecture indicates the input variables, hidden layers and number of nodes in it, biases, input, hidden and output layer weights, bias weights. The corresponding model output values of the predicted EPAN as used for evaluation of these developed models. The output resulted from the ANN models at minimum MSE / MAE was compared with observed EPAN and the various model evaluation parameters were analysed to decide the best predictive model for pan evaporation. The absolute and squared statistics of the ANN models is tabulated and presented in Table 2. ANN General model evaluation using Nagpur model development data shows that the RMSE, MAE and MARE is increased and r, R^2 and d is decreased when the input variables were sequentially removed from seven inputs in model M-1 to six inputs in model M-2 and M-3. The lowest RMSE (1.961 mm), MAE (0.038 mm) and MARE (2.30 %) and highest r (0.848), R^2

(0.719) and d (0.919) was observed with GM M-1 model, in which all the input variables are taken into consideration. However, in model M-2 and M-3, where six input variables are considered in GM modeling has also shown lower RMSE, MAE and MARE and higher r, R^2 and d. This shows that the GM model generalization capacity is better when tested with Nagpur model development data. The RMSE, MAE and MARE in model plan M-2 was 1.973 mm, 0.0389 mm and 2.49 %, which is slightly more than the model M-1. The r, R^2 and d in model M-2 is 0.845, 0.715 and 0.917 respectively. These values are also less to the values of model M-1. The RMSE, MAE and MARE of GM M-3 model are higher and r, R^2 and d are lower as compared to model M-1 and M-2. This is mainly due to the replacement of sunshine hours input of model M-2 with RAIN variable. This indicates that the performance of all variables without SNHR (but with RAIN variable) shows higher RMSE, MAE and MARE and lower r, R^2 and d as compared to model M-2. The RMSE, MAE and MARE of model M-3 are 1.984 mm, 0.0394 mm and 2.86 % respectively. The r, R^2 and d of model M-3 are 0.844, 0.712 and 0.915 respectively. The error statistics of model GM M-4 in which the relative humidity (RMAX and RMIN) are not the input variables. The RMSE, MAE and MARE of model M-4 is 2.081 mm, 0.0433 mm and 3.27 %, which is higher than the models M-1, M-2 and M-3. The r, R^2 and d of M-4 is 0.826, 0.683 and 0.906 respectively and these values are higher than model M-3. It is mainly due to the absence of RMAX and RMIN input variables and have higher negative correlation with the pan evaporation.

The model M-5 performance is better than the model M-4. The RMSE, MAE and MARE in model M-5 is reduced to 2.061 mm, 0.0425 mm and 3.24 % from 2.081 mm, 0.0433 mm and 3.27 % in model M-4. The r, R^2 and d is also increased to 0.830, 0.689 and 0.910 in model M-5 as compared to model M-4. The model M-5, in which four input variables is assumed to be one of the better models with minimum number of input variables. The input variables are TMAX, WISP and SNHR in model M-6. The RMSE, MAE and MARE of model M-6 is 2.127 mm, 0.0453 mm and 4.24 % respectively. The r, R^2 and d of model M-6 were 0.818, 0.669 and 0.904, which are lower as compared to M-5. The RMSE, MAE and MARE is decreased to 2.111 mm, 0.0445 mm and 5.01 % in GM model M-7, in which only TMAX and WISP are the input variables of the model. The r, R^2 and d was also slightly higher in model M-7 as compared to M-6. The r, R^2 and d in model M-7 is 0.821, 0.674 and 0.902 respectively. The evaluation of General models using Nagpur model development data with ANN model M-1 to M-7 are graphically presented in figure 3. These results indicated that maximum temperature and relative humidity observations were very beneficial in accurate estimation of pan evaporation, when compared to any other single variable.

Table 2

RMSE, coefficient of determination (R^2) and index of agreement (d) of ANN General model development and evaluation dataset.

ANN GM models	GM model development			GM model evaluation with NMD data		
Model	RMSE, mm/d	R^2	d	RMSE, mm/d	R^2	d
GM M-1	1.44	0.84	0.95	1.96	0.72	0.92
GM M-2	1.48	0.83	0.95	1.97	0.71	0.91
GM M-3	1.56	0.81	0.94	1.98	0.71	0.91
GM M-4	1.62	0.79	0.94	2.08	0.68	0.90
GM M-5	1.68	0.78	0.93	2.06	0.69	0.91
GM M-6	1.70	0.77	0.93	2.12	0.67	0.90
GM M-7	1.76	0.76	0.92	2.11	0.67	0.90

4. Conclusions

The three layered feed forward neural network with back propagation prediction method that uses Artificial Intelligence (AI) to model the daily pan evaporation was developed and evaluated using independent datasets. From these research results, it is concluded that the constructed ANN models successfully identifies the evaporation process and accurately predict the pan evaporation in a next time step. Three layered feed forward neural networks with error back propagation was found suitable for modeling the evaporation. Once the network is trained, with a satisfactory amount of input data, it can accurately predict the daily pan evaporation. The network architectures performed better with one hidden layer and 15 hidden nodes (2^{n+1}) for GM modeling problems.

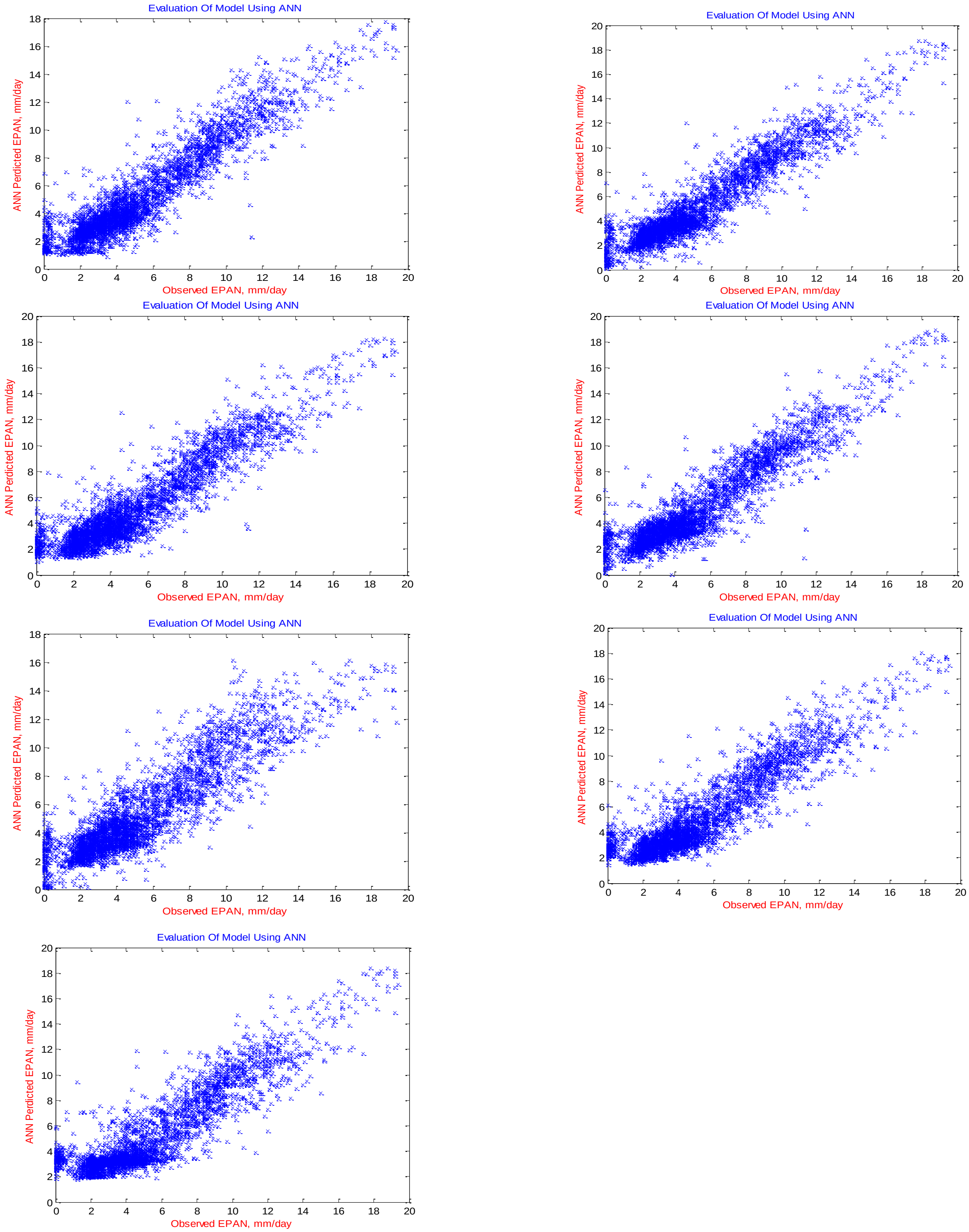

Fig. 2. Observed and predicted EPAN with General models (GM M1- M7) development data.

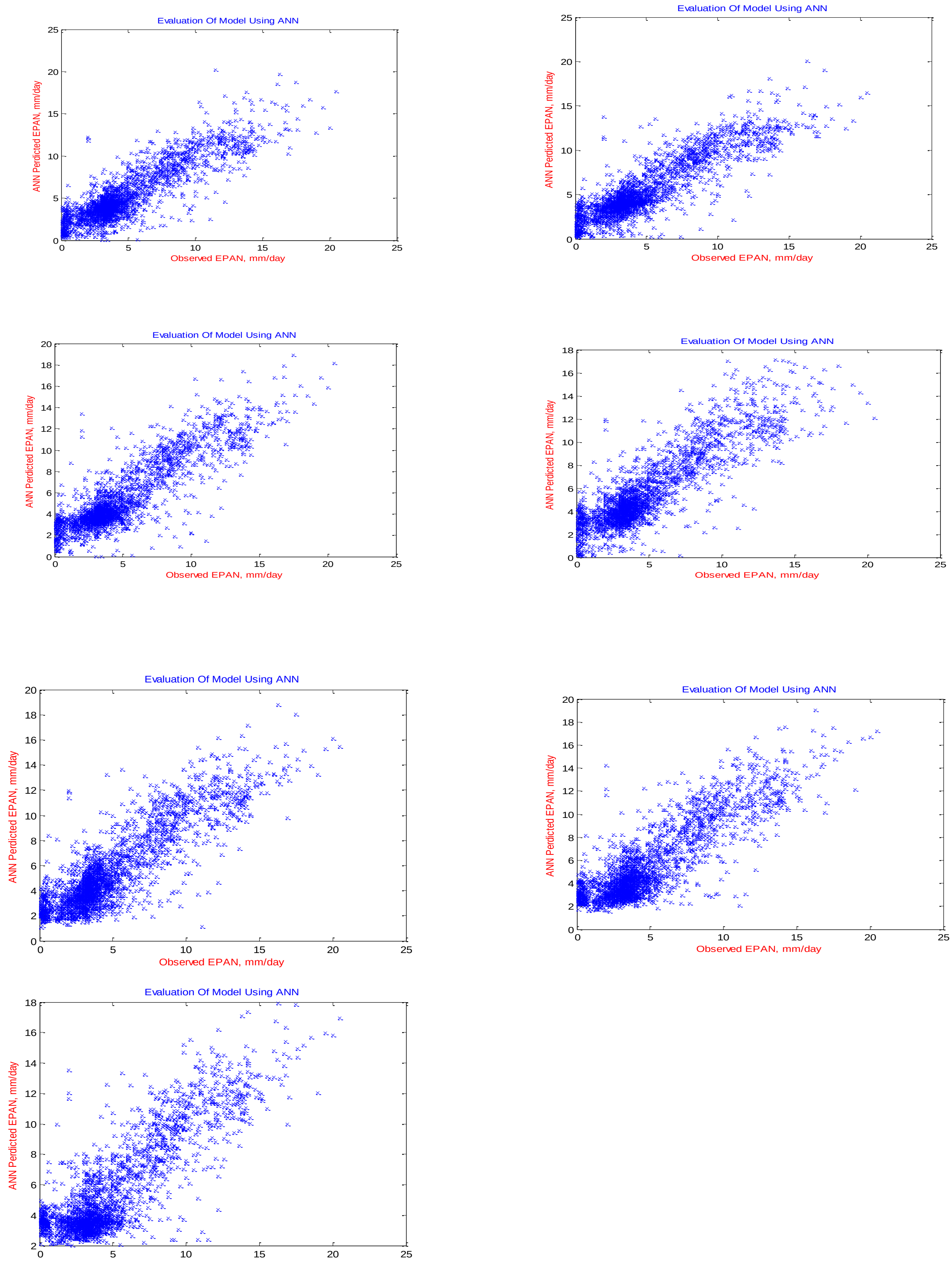

Fig. 3. ANN Models evaluation with Nagpur model development data (NMD).

Single hidden layer performed good for both the models. The *Levenberg Marquardt* minimization training function gave the best network training results when used in the back propagation algorithms. While evaluating the data the model plan strategy with all inputs has shown better performance than the reduced number of inputs. The daily pan evaporation is estimated for the five locations using three layer feed forward neural network with error back propagation. Seven model strategies with combinations of input variables to develop the ANN models of pan evaporation were developed. Strategy M-1, which includes all 7 input variables, had the highest accuracy. This ANN model had an R^2 of 0.78 and an RMSE of 1.61 mm on the independent GM evaluation data set. The models evaluation with Nagpur model development (NMD) dataset had an R^2 of 0.72 and an RMSE of 1.96 mm. The ANN model has a strong correlation with observed pan evaporation. Differences between the ANN model prediction and observed pan evaporation do not appear to be biased as the observation of pan evaporation changes.

Acknowledgement

The author express sincerely thanks to the Indian Council of Agricultural Research (ICAR), New Delhi for providing the Ph. D. study leave and the opportunity to learn the new area like Artificial Neural networks and non-linear process, which is essentially required in the agricultural research system. Thanks are also extended to the Director, National Research Centre for Citrus, Nagpur for providing the dataset and other office facilities. Author also express sincere thanks to the Dean, Professors and staff of the College of Agricultural Engineering, JNKVV, Jabalpur (MP) for their kind co-operation, help and guidance during the study period.

References

Arca, B., Benincasa F., De Vincenzi M., Ventura A., 1998. Neural network to simulate evaporation from Class A pan. In: Proc. 23rd Conf. on Agricultural and Forest Meteorology, 2-6 November 1998, Albuquerque, New Mexico, Am. Meteorol. Soc., Boston, MA.

Bruton, J.M., McClendon R.W., Hoogenboom, G., 2000. Estimating daily pan evaporation with artificial neural networks. Trans. ASAE, 43 (2), 491-496.

de Villiers, J., Barnard, E., 1993. Back propagation neural nets with one and two hidden layers. IEEE Transactions on Neural Networks, 4(1), pp. 136-141.

Deswal S., Mahesh, Pal., 2008. Artificial Neural Network based Modeling of Evaporation Losses in Reservoirs. Proc. of World Academy of Science, Engineering and Technology, Vol. 29, 279-283.

Gavin, H., Agnew, C.A., 2004. Modeling actual, reference and equilibrium evaporation from a temperate wet grassland. Hydrol. Processes, 18, 229-246.

Gardner, M.W., Dorling, S.R., 1998. Artificial neural network (Multilayer Perceptron) a review of applications in atmospheric sciences. Atmospheric Environment. 32**,** 2627-2636.

Han, H., Felkar P., 1997. Estimation of daily soil water evaporation using an artificial neural network. J. Arid Environ. 37 (2), 251-260.

Jhajharia, D., Kithan, S.B., Fancon, A.K., 2006. Correlation between pan evaporation and metrological parameters under the climatic conditions of Jorhat (Assam). J. of Indian Water Resources, 26 (1-2) , 39-42.

Keskin, M.E., Terzi, O., 2006. Artificial neural network models of daily pan evaporation. ASCE J. of Hydrologic Engineer. 11 (1) , 65-70.

Khanikar, P.G., Nath K.K., 1998. Relationship of open pan evaporation rate with some important meteorological parameters. J. Agril. Sci. Soc. N - E India. 11(1), 46-50.

Kumar, M., Raghuwanshi, N.S., Singh, R., Wallender W.W., Pruitt, W.O., 2002. Estimating evapo-transpiration using artificial neural network. J. Irri. Drain. Engg. ASCE, 128 (4), 224-233.

Maier, H.R., Dandy, G.C., 2001. Neural network based modelling of environmental variables: a systematic approach. Mathematical and Computer Modelling, 33, pp. 669-682.

Molina Martinez J.M., Alvarezv, M., Gonzalez Real M.M., Baille, A., 2006. A simulation model for predicting hourly pan evaporation from meteorological data. J. hydrol. 318 (1-4), 250-261.

Ozlem, T., Evolkesk, M., 2005. Modeling of daily pan evaporation. J. Appl. Sci. 5 (2), 368-372.

Shirgure, P.S., Srivastava A.K., Singh S., 2001. Effect of pan evaporation based irrigation scheduling on yield and quality of drip irrigated Nagpur mandarin. Indian J. Agri. Sci. 71 (4), 264 – 266.

Shirgure, P.S., 2011. Evaporation modeling with neural networks – A Research review. International Journal of Research and Reviews in Soft Intelligent Computing. 1 (2),37-47.

Shrivastava, S.K., Sahu, A.K., Dewangan, K. N., Mishra, S. K., Upadhyay A. P., Dubey, A. K., 2001. Estimating pan evaporation from meteorological data for Jabalpur. Indian J. Soil. Cons. 29 (3), 224-228.

Sudheer, K.P., Gosain, A.K., Mohana, R.D., Saheb, S.M., 2002. Modeling evaporation using an artificial neural network algorithm. Hyd. Processes, 16, 3189-3202.

Sudheer, K.P., Gosain A.K., Ramasastri, K.S., 2003. Estimating actual evapotranspiration from limited climatic data using neural computing technique. J. Irri. Drain. Engg. ASCE. 129 (3), 214-218.

Taher, S.A., 2003. Estimation of potential evaporation: Artificial Neural Networks versus Conventional methods. J. King. Saud. Univ. Eng. Sci. 17(1), 1-14.