RESEARCH NOTE

ANN BASED MODELING FOR PREDICTION OF EVAPORATION IN RESERVOIRS

A. Goel

Department of Civil Engineering, National Institute of Technology, Deemed University
Kurukshetra 136119, Haryana, India
drarun_goel@yahoo.co.in

(Received: October 16, 2008 – Accepted in Revised Form: July 2, 2009)

Abstract This paper is an attempt to assess the potential and usefulness of ANN based modeling for evaporation prediction from a reservoir, where in classical and empirical equations failed to predict the evaporation accurately. The meteorological data set of daily pan evaporation, temperature, solar radiation, relative humidity, wind speed is used in this study. The performance of feed forward back propagation neural network model is compared with the linear regression on the basis of performance parameters (correlation coefficient and rmse) having different combinations of input parameters. The comparison of results shows that there is a better agreement when large input parameters are considered for model building and testing as compared to a single parameter. The outcome of study suggests that the feed forward back propagation ANN based modeling can be applied as an alternative approach for estimation of daily evaporation from reservoirs effectively.

Keywords Linear Regression, Neural Network, Correlation Coefficient, Evaporation

چکیده در این مقاله تلاش شده تا امکان و قابلیت‌های مدل سازی بر پایه ANN به منظور پیش‌بینی تبخیر از یک مخزن در حالتی که معادلات کلاسیک و تجربی نتوانسته‌اند تبخیر را به طور دقیق پیش‌بینی کنند، سنجیده شود. یک دسته داده‌های هواشناسی شامل تبخیر سطحی روزانه، دما، تشعشع خورشیدی، رطوبت نسبی و سرعت باد در این مطالعه استفاده شده‌است. ترکیب‌های مختلفی از پارامترها به عنوان ورودی اعمال می‌شود و به‌ازای آن عملکرد پس‌انتشار فید فوروارد در مدل شبکه عصبی با بازگشت خطی بر اساس پارامترهای عملکرد (ضریب همبستگی و rmse) مقایسه می‌گردد. مقایسه نتایج نشان می‌دهد که وقتی برای ساخت و آزمایش مدل، پارامترهای ورودی، بیشتر در نظر گرفته می‌شوند، در مقایسه با حالت تک پارامتری هم‌خوانی بهتری مَشاهده می‌شود. این مطالعه نشان می‌دهد که مدل‌سازی پس‌انتشار فید فوروارد بر پایه ANN می‌تواند به‌عنوان یک رویکرد جایگزین و کاراً برای تخمین تبخیر روزانه از مخازن اعمال گردد.

1. INTRODUCTION

Accurate estimation of evaporation influences water balance studies undertaken for efficient planning, design, operation and management of water resources. In the hydrological practice, the evaporation can be estimated by conventional approaches like direct or indirect methods involving use of the empirical equations. The evaporation is considered to be most difficult to estimate in the hydrologic cycle due to complex interaction between components of land-plant-atmosphere system (Singh, et al [1], Dariane [2]). The sole direct method is the U S weather Bureau Class a pan measurement which gives record with time. The indirect methods use meteorological data to estimate evaporation by empirical based methods or statistical and stochastic approaches. The indirect methods are namely Temperature based formulae, Radiation method; Humidity based relation, Penman formulae, Energy balance approach etc. These methods of evaporation estimation have been applied by (Abtew [3], Choudhury [4], Vallet-Coulomb, et al [5], Terzi, et al [6], Rosenberry [7] etc. Although all these approaches are based on Penman formula, they are sensitive to site-specific evaporation parameters, which can vary from one place to other. Further, it is not possible to consider all the parameters affecting the evaporation estimation by any of the above approaches due to many assumptions and phenomenal constraints.

The literature review showed that these equations vary greatly in their ability to define the magnitude and variability of the evaporation from the reservoirs. It is therefore necessary to develop alternate approaches to estimate the evaporation rates based on metrology variables, which are comparatively easier to measurements and estimation. One of the recent alternate approaches is the use of soft computing modeling techniques, which have better modeling flexibility and capability rather than previous empirical approaches where in each of the metrological parameter takes it share proportionally.

Many researchers have investigated the applicability of ANN in hydrology (ASCE [8,9]) to estimate rainfall-runoff (Zealand, et al [10]), short-term stream flow (Luk, et al [11]), rainfall (Tokar, et al [12]), reservoir inflow (Mohammadi, et al [13]) etc. The ANN models are also applied to estimate the Pan evaporation by Terzi, et al [14], Eskin, et al [15], Kokya, et al [16], Eslamian, et al [17] and recently Moghaddamnia, et al [18,19]. Murthy, et al [20] have applied linear regression approach for prediction of evaporation by using individual parameters from an Indian reservoir. They have used field data of air temperature (T), wind speed (U), sunshine hours (Sh) and relative humidity (Rh) to predict the evaporation loss (E) as individual parameters. This study has a major limitation as they have not considered the combined effect of all the input parameters in determining the evaporation losses from reservoir. This study is aimed to check the potential and application of ANN modeling in predicting the evaporation loss along with linear regression models.

2. STUDY REGION AND DATA USED

The average weekly evaporation and meteorological data of Manasgaon (from 1990 to 2004) are collected from a reservoir in Anand Sagar, Shegaon, India as mentioned in Table 1. The measured meteorological variables include daily observations of evaporation, mean air temperature; sunshine hours, mean relative humidity, and wind speed have been used in the present study.

3. ARTIFICIAL NEURAL NETWORKS

A neural network is an artificial intelligence technique that mimics a function of the human brain. Neural networks are general-purpose computing tools that can solve complex non-linear problems in the field of pattern recognition, classification, speech, vision and control systems. The network comprises a large number of simple processing elements linked to each other by weighted connections according to a specified architecture. A neuron consists of multiple inputs and a single output. The number of neurons in the input and output layers are fixed by the problem being modeled as the number of input variables equals number of input neurons and number of output variables equal number of output neurons. The determination of optimal number of hidden layers and hidden neurons is usually cumbersome, as no general methodology is available for their determination. These networks learn from the training data by adjusting the connection weights. There is a range of artificial neural network architectures designed and used in various fields of hydrology and hydraulics. Most of the studies employing neural networks for water resource problems have used back propagation and radial basis function types of neural networks. In this study, a feed-forward neural network with back propagation learning algorithm is applied. The basic element of a back-propagation neural network is processing node and structure of commonly used back propagation neural network (Figure 1). A three layer feed forward ANN has been shown in Figure 1, which consists of three layers namely input layer neuron x_1, x_2, x_3; hidden layers neurons h_1, h_2, h_3 and output layers neurons O_1, O_2, O_3. A neuron consists of multiple inputs and a single output. The sum of inputs and their weights lead to a summation function. The output of a neuron is decided by an activation function, which can be step, sigmoid, threshold and linear etc.

Each processing node behaves like a biological neuron and performs two functions. First, it sums the values of its inputs. This sum is then passed through an activation function to generate an output. Any differentiable function can be used as activation function. All the processing nodes are arranged into layers, each fully interconnected to the following layer.

TABLE 1. Average Weekly Meterological Data of Reservoir from Year 1999-2004.

Month	Week	Evaporation Rate E (mm/day)	Mean air Temperature T(˚C)	Average Wind Velocity U (m/s. at 2 m height)	Sunshine Hours Sh (hrs/day)	Mean Relative Humidity Rh (%)
Jan.	1	3.4	19.64	3.3	8.8	64.35
	2	3.3	20.29	2.9	8.3	61.45
	3	3.4	21.47	3.2	8.5	60.1
	4	3	20.86	3.8	6.9	60.45
Feb.	1	3.4	21.99	3.5	9	59.35
	2	3.9	23.59	3.8	8.9	55.7
	3	4.3	24.52	4.4	9.1	54.25
	4	4.7	25.62	4.1	8.9	48.4
March	1	6.2	26.31	4.7	10.1	45.45
	2	5.6	26.89	4.9	8.6	45.1
	3	6.5	29.18	5	9.3	40.75
	4	5.9	30.04	5.4	8.2	40.2
April	1	8.1	31.52	5.8	10	42.55
	2	6.8	32.44	6.7	8.3	40.4
	3	8.8	33.36	6.4	9.8	38.05
	4	9.8	34.7	8	9.3	36.4
May	1	12.4	35.46	8.5	11.2	37.6
	2	11.9	34.85	9.5	10.2	44.4
	3	12.2	34.41	11.1	9.3	49.3
	4	11.2	34.68	10.3	8.3	46.6
June	1	15.3	34.05	10.7	9.4	53.3
	2	9.8	31.86	9.1	5	65.25
	3	7.5	30.07	10.4	4.1	74
	4	5.7	29.68	10.8	4.1	73.9
July	1	7.5	29.66	8.8	5.7	73
	2	5.7	29.01	9.3	3.9	77.6
	3	4.9	27.52	8.2	2.6	82.8
	4	3.3	27.37	8.1	3.4	83.95
Aug.	1	3.9	2.1	6.7	3.2	87.65
	2	3.1	26.7	6.8	2.5	86.9
	3	2.9	26.56	6.6	2.4	86.25
	4	3	26.88	6.1	3.7	84.85
Sep.	1	3.1	26.85	6.1	4	84.7
	2	3.3	27.43	5.4	5.6	79.3
	3	4.1	27.8	4.3	5.8	78.7
	4	4.4	28.58	2.9	6.3	78.9
Oct.	1	4.9	27.83	2.9	8	77.5
	2	4.3	27.43	2.8	7.3	75.6
	3	4	25.65	2.2	8.8	69.25
	4	3.5	24.62	2.5	7.6	65.8
Nov.	1	4.5	24.78	2.5	10.1	66
	2	3.5	24.47	2.5	7.4	69.35
	3	3.7	23.03	2.1	8.9	60.5
	4	3.4	22.47	2.2	8.7	58.5
Dec.	1	3.7	20.78	2.4	9.6	61.4
	2	3.3	20.99	2.4	8.3	61.2
	3	3.3	20.21	2.5	8.1	61
	4	2.8	20.38	3.2	6.5	60.2

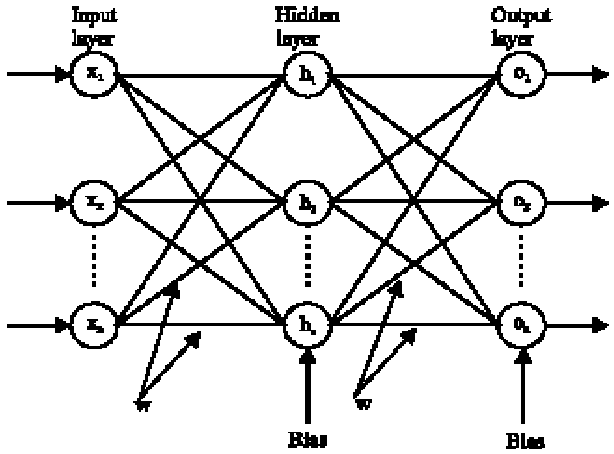

Figure 1. Three layer feed forward neural network.

There is no interconnection between the nodes of the same layer. In a back propagation neural network, generally, there is an input layer that acts as a distribution structure for the data being presented to the network. After this layer, one or more processing layers follow, called the hidden layers. The final processing layer is called the output layer in a network. This process is repeated until the error rate is minimized or reaches to an acceptable level. All the interconnections between each node have an associated weight. The values of the interconnecting weights are not set by the analyst, but are determined by the network during the training process, starting with randomly assigned initial weights. There are a number of algorithms that can be used to adjust the interconnecting weights to achieve minimal overall training error in multi-layer networks. The generalized delta rule, or back-propagation is one of the most commonly used methods as suggested by (Rumelhart, et al [21]). The first derivative of the total error with respect to a weight (Equation 1) determines the extent to which that weight is adjusted.

$\Delta w = -\in \partial E/\partial w$(1)

Where in $\in$ is the learning constant; $\partial E/\partial w$ is the first derivative of the total error with respect to weight and Δw is weight change. A neural network based modeling approach requires setting up several user-defined parameters like learning rate, momentum, optimal number of nodes in the hidden layer and the number of hidden layers, so as to have a less complex network with a better generalization capability.

4. PERFORMANCE COMPARISON CRITERIA

Much success has already been achieved using neural network algorithms in other applications, such as rainfall-runoff modeling, and stage-discharge analysis (ASCE task committee on application of ANNs I and II, 2000; Bhattacharya, et al [22]). Neural networks are now being applied to several other problems related to the hydraulics and hydrologic modeling.

The data sets from Murthy, et al [20] was used in the present study for model building and validation to assess the potential linear regression based modeling and ANN based modeling in predicting the evaporation loss from the reservoir. The neural network is used to calculate correlation coefficient and root mean squared error (RMSE) by using cross-validation to generate the model on different combinations of the input data set in predicting the evaporation loss. Cross validation was used to train/test/validate the models due to the availability of small number of data sets. It is a method of estimating the accuracy of a classification or regression model in which the input data set is divided into several parts (a number defined by the user), with each part in turn used to test a model fitted to the remaining parts. For this study, a ten-fold cross-validation was used. Further, measured evaporation values were plotted against the computed values obtained with linear and ANN algorithms. To study the scatter around the line of perfect agreement, a line at 45 degrees was also plotted for the data set. Due to the availability of small data sets, a cross validation was used to train and test the performance of the ANNs. The input data set is divided into several parts (a number defined by the user), with each part in turn used to test a model fitted to the remaining parts. In this study, the field data were used for both creating and testing the models. The choice of input parameters used in modeling the evaporation may influence the predicting capabilities of ANN. Graphs have been plotted for difference in actual and predicted values of the evaporation loss as shown in Figures 2-6.

5. ANALYSIS OF RESULTS

The first set of analysis was carried out by using ANN and linear regression with input as temperature and output as evaporation loss and results are plotted as shown in Figure 2. The ANN determined a relationship (i.e. create a model) between the input and the output of the available data set of any system. These models are then used to predict the output from the known input values of the same system, thus requiring sufficient number of data to create and test the models. A number of trials were carried out to reach at the various user-defined parameters required for the neural network and linear regression based algorithms using WEKA software (www.cs.waikato.ac.nz/ml/) [23]. The Tables 2 and 3 provide the value different user defined parameters and performance parameters

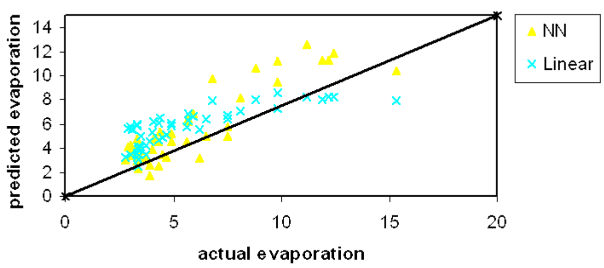

Figure 2. Variation of actual evaporation versus predicted evaporation with temperature.

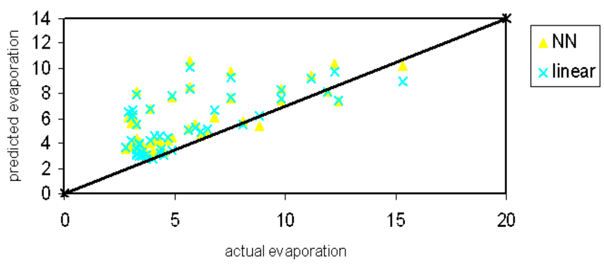

Figure 3. Variation of actual evaporation versus predicted evaporation with U.

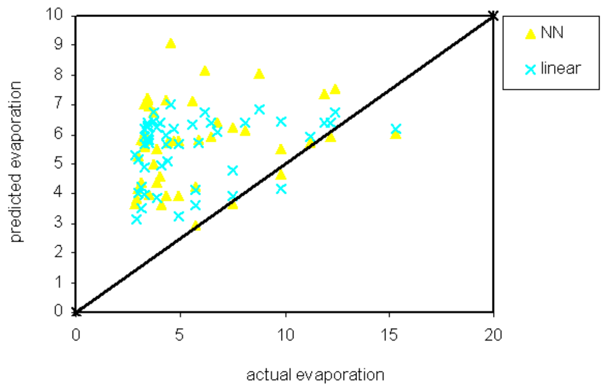

Figure 4. Variation of actual evaporation versus predicted evaporation with Sh.

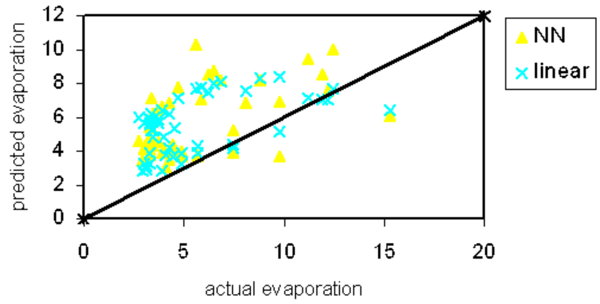

Figure 5. Variation of actual evaporation versus predicted evaporation with Rh.

for the data set.

Similarly wind velocity, sun shine hours, relative humidity were taken as input parameters separately and evaporation loss were predicted and the results obtained are plotted as shown in Figures 3-5 respectively. A graph shown as Figure 6 is plotted when all four parameters were taken for model building and evaporation loss was predicted. Scatter diagrams between actual and predicted daily pan evaporation values are plotted considering different input combinations of meteorological parameters. The results of the different methods obtained by using BP ANN, and linear regression models, are compared with the actual pan evaporation values of the reservoir. The examination of plotted Figures 2-6 shows that there is better agreement of predicted results based on ANN rather than a linear regression.

It is evident from Figure 2-6 that more number

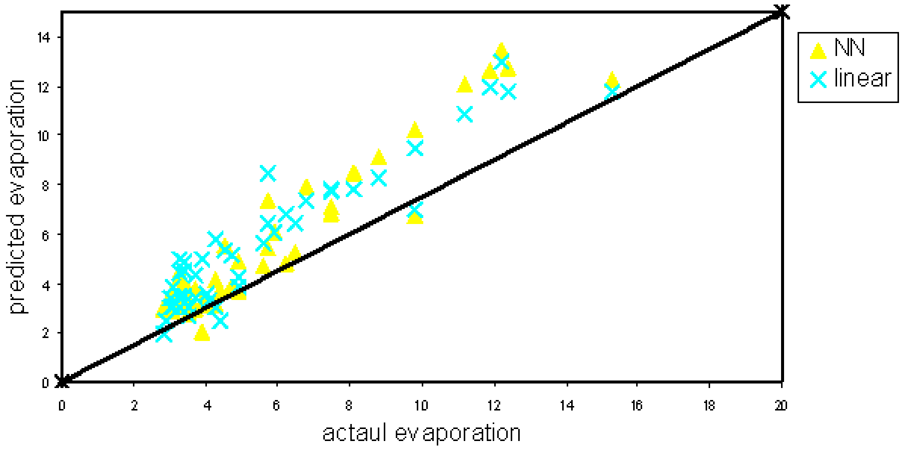

Figure 6. Variation of actual evaporation versus predicted evaporation with full parameters.

TABLE 2. User Defined Parameters of Neural Network Modeling.

S No	Input Parameter	Momentum	Learning Rate	No of Hidden Nodes	No of Iterations
1	T	0.2	0.3	6	500
2	U	0.2	0.05	0	100
3	Sh	0.2	0.3	6	500
4	Rh	0.2	0.3	8	500
5	T + U + Sh + Rh	0.2	0.3	1	500

TABLE 3. Values of Performance Parameters by using ANN and Linear Regression.

Input Parameter	Type of Technique	Correlation Coefficient (r)	Root Mean Squared Error	Total Computed Cost (s.)
T (Figure 1)	ANN (MLP)	0.8973	1.3636	0.25
	Linear	0.5539	2.6868	0.02
U (Figure 2)	ANN (MLP)	0.691	2.186	0.06
	Linear	0.6659	2.2584	0
Sh (Figure 3)	ANN (MLP)	0.2607	2.9337	0.42
	Linear	0.2678	2.9226	0
Rh (Figure 4)	ANN (MLP)	0.5277	2.5926	0.34
	Linear	0.511	2.5984	0
Rh+T+Sh+U (Figure 5)	ANN (MLP)	0.9505	0.9799	0.28
	Linear	0.9022	1.3398	0

of points are lying on the 45˚ line when the ANN was used to predict the evaporation in comparison to linear regression based algorithm for any combination of input parameters. The maximum value of correlation coefficient was obtained for air temperature (r=0.89) followed by wind speed (r=0.69) and relative humidity (r=0.52), while minimum value was obtained for sunshine hours (r=0.26) (Table 3). The performance parameters namely correlation coefficient and RMSE obtained by using ANN and by using linear regression based modeling (Table 3) are plotted in Figures 7 and 8, while considering sunshine hours (minimum correlation coefficient = 0.26 and maximum rmse = 2.9) and all parameters (maximum correlation coefficient = 0.95 and minimum rmse = 0.97). However, all factors influence concurrently (as shown in Figure 8). So it is quite obvious to consider the combined influence of all the parameters in evaporation estimation as indicated by higher values of correlation coefficient (r) and lower rmse values.

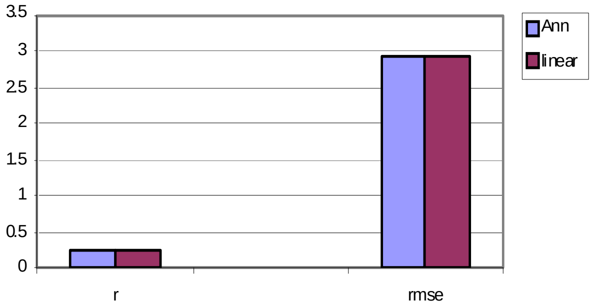

Figure 7. Variation of correlation coefficient(r) and rmse with Sh.

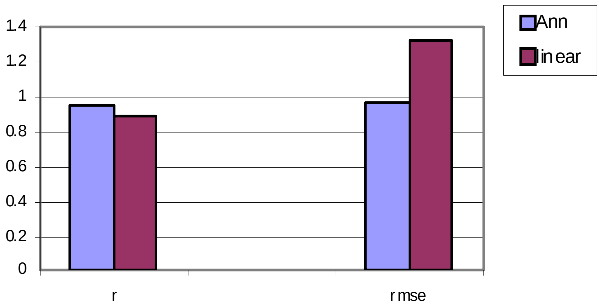

Figure 8. Variation of correlation coefficient (r) and rmse with full parameters.

6. CONCLUSIONS

ANN has been proposed and emerged as an alternative approach of evaporation estimation from a reservoir as compared to linear regression. The back propagation multilayer perception ANN and linear regressions based modeling techniques are performing better when all four parameters are used as input for model building for the prediction of evaporation. Further, the critical examination of plotted figures indicates that the performance of the ANN modeling is data dependent to a great extent. The study also concludes that combination of all input parameters provides better performance of model in estimating the evaporation rather than individual parameters. The study revealed that evaporation loss is best estimated by ANN modeling rather than linear regression techniques. The outcome of the study provided an impetus to the potential use of ANN approach in predicting the evaporation from the reservoirs in water resources projects.

7. ACKNOWLEGEMENTS

The author wishes to gratefully acknowledge authors of Murthy, et al [20] for their data set used in the present study.

8. REFERENCES

1. Singh, V.P. and Xu, C.Y., “Evaluation and Generalization of 13 Mass Transfer Equations for Determining Free Water Evaporation” ***Hydrological Processes***, Vol. 11, (1997), 313-323.
2. Dariane, A.B., “Reservoir Operation During Drought”, ***International Journal of Engineering Transcations B: Applications***, Vol 16, No. 3, (2003), 209-216.
3. Abtew, W., “Evaporation Estimation of Lake Okeechobee in South Florida”, ***J. of Irrigation and Drainage***, Vol. 127, No. 3, (2001), 140-147.
4. Choudhury, B.J.,“Evaluation of an Empirical Equation for Annual Evaporation using Field Observations and Results from Biophysical Model”, ***J. of Hydrology***, Vol. 216, (1999), 99-110.
5. Vallet-Coulomb, Legesse, C., Gasse, D., Tavi, F. and Chernet, T., “Lake Evaporation Estimates in Tropical Africaa (Lake Ziway, Ethopia)”, ***J. of Hydrology***, Vol. 245, (2001), 1-18.
6. Terzi, O. and Keskin, M.E., “Modeling of Daily Pan Evaporation”, ***J. of Applied Sciences***, Vol. 5, (2005),

368-372.
7. Rosenberry, D.O., Winter, T.C., Buso, D.C. and Likens, G.E., "Comparison of 15 Evaporation Methods Applied to a Small Mountain Lake in the Northeastern U.S.A.", ***J. of Hydrology***, Vol. 340, (2007), 149-166.
8. ASCE Task Committee on Application of ANNs in Hydrology "Artificial Neural Networks in Hydrology, I: Preliminary Concepts", ***J. Hydraulic Engineering, ASCE***, Vol. 5, No. 2, (2000a), 115-123.
9. ASCE Task Committee on Application of ANNs in Hydrology, "Artificial Neural Networks in Hydrology, II: Hydrologic Applications", ***J. Hydraulic Engineering***, ***ASCE***, Vol. 5, No. 2, (2000b), 124-137.
10. Zealand, C.M., Burns, D.H. and Simonovic, S.P. "Short Term Stream Forecasting using Artificial Neural Networks", ***J. of Hydrology***, Vol. 214, (1999), 32-48.
11. Luk, K.C., Ball, J.E. and Sharma, A., "A Study of Optimal Model Lag and Spatial Inputs to Artificial Neural Network for Rainfall Forecasting", ***J. of Hydrology***, Vol. 227, (2000), 56-65.
12. Tokar, A.S. and Johnson, P.A., "Rainfall Runoff Modeling using Artificial Neural Network." ***J. of Hydraulic Engineering***, Vol. 4, No. 3, (1999), 232-239.
13. Mohammadi, K, Eslami, H.R. and Dardashti Dayyani, Sh., "Comparison of Regression, ARIMA and ANN Models for Reservoir Inflow Forecasting using Snowmelt Equivalent (A Case Study of Karaj)", ***J. of Agricultural Sciences Technology***, Vol.7, (2005), 17-30.
14. Terzi O., Erol, K.M. and Eliek, T.E., "Estimating Evaporation using ANFIS", ***Journal of Irrigation and Drainage Engineering***, Vol. 132, No. 5, (2006), 503-507.
15. Eskin, M.E. and Erzi, O., "Artificial Neural Network Models of Daily Pan Evaporation", ***Journal of Hydrologic Engineering***, Vol. 11, No. 1, (2006), 65-70.
16. Kokya, B.A. and Kokya, T.A. "Proposing a Formula for Evaporation Measurement from Salt Water Resources", ***International Journal of Hyrological Processes***, Vol 22, (2008), 2005-2012.
17. Eslamian, S.S., Gohari, S.A., Biabanaki, M. and Malekian, R., "Estimation of Monthly Pan Evaporation using Artificial Neural Networks and Support Vector Machines", ***J. of Applied Sciences***, Vol.8, No. 19, (2008) 3497-3502.
18. Moghaddamnia, A., Ghafari, M., Piri, J. and Han, D., "Evaporation Estimation using Support Vector Machines Technique", ***Proceedings of World Academy of Science, Engineering and Technology***, Vol. 33, No. 9, (2008), 14-22.
19. Moghaddamnia, A., Goushehe Ghafari, M., Piri, J., Amin, S. and Han, D., "Evaporation Estimation using Artificial Neural Networks and Adaptive Neuro-Fuzzy Inference System Techniques", ***J. of Advances in Water Resources***, Vol. 32, No. 1, (2009), 88-97.
20. Murthy, S. and Gawande, S., "Effect of Meteorological Parameters on Evaporation in Small Reservoirs Annad Sagar Shegaon-A Case Study", ***Journal of Perdushan Nimulan***, Vol. 3, No. 2, (2006), 52-56.
21. Rumelhart, D.E., Hinton, G.E. and Williams, R.J., "Learning Internal Representation by Error Propagation, In: Parallel Distributed Processing: Explorations, in the Microstructures of Cognition", ***Cambridge***, MIT Press, (1996), 318-362.
22. Bhattacharya, B. and Solomatine, D.P., "Application of Artificial Neural Network in Stage Discharge Relationship", ***Proceedings of 4th International Conference on Hydro Informatics***, Iowa, U.S.A., (2000).
23. Weka 3 Data Mining with Open Source Machine Learning Software, www.cs.waikato.ac.nz/ml/, (2009).