Using Empirical Relationships and Neural Network in GIS for Developing Rainfall-Runoff Model

Shahrbanu FIRUZI, Mohamadbagher SHARIFI, Kambiz BORNA and MohamadReza RAJABI, Iran

Key words: Run-off, Artificial neural networks, GIS, Un-gagged catchment

SUMMARY

The estimation of run-off volume from a catchment is required for planning and design of water resources projects such as the design of storage facilities, assessment of water available for municipal, agricultural or industrial purposes, etc. but proper planning is possible when the accuracy of hydrological process output would be sure. Simulation and prediction of runoff is very important because of two reasons: the first one is the nonlinear interaction between rainfall and runoff on a catchment and the second one is recorded runoff data are not available for many catchments. So, in the present study were used artificial neural networks as a model which has ability of extracting nonlinear relationships from the analysis of available information in GIS. The Method that is presented in this research identifies that using of empirical equations in the form of neural network models causes not only to add spatiotemporal effective parameters in model but also to increase accuracy of prediction and generalization of model for homogeneous catchment effectively.

Using Empirical Relationships and Neural Network in GIS for Developed Rainfall-Runoff Modeling

Shahrbanu FIRUZI, Mohamadbagher SHARIFI, Kambiz BORNA and MohamadReza RAJABI, Iran

1. INTRODUCTION

Rainfall-runoff models play an important role in water resource management planning and therefore, different types of models with various degrees of complexity have been developed for this purpose. These models, regardless to their structural diversity generally fall into three broad categories; namely, black box or system theoretical models, conceptual models and physically-based models [2].Conceptual models try to express known physical processes which are occurred in formation of rainfall- runoff using simplify the problem with a linear or nonlinear mathematical equations while black box models normally contain no physically-based input and output transfer functions and therefore are considered to be purely empirical models[12]. However Structures of simple general models have the problem in description of dynamics and nonlinear process of rainfall – runoff that issue can be expressed via using artificial neural network. While the capability of ANNs to capture nonlinearity in the rainfall-runoff process remains attractive features comparing with other modeling approaches which has high ability in description of complexity of converting rainfall to runoff process in the catchment. Accordingly, ANN models as illustrated in numerous previous studies, essentially belong to system theoretical model category and bear the weaknesses of this category [4]. Extensive researches have been done about investigation of performance of these models for prediction of runoff in a river catchment. Shamseddin examined the performance of rainfall-runoff modeling with the ANN and the results proved ANN forecasts runoff more accurate than many conventional models.[19] Kisi used neural network and regression methods to predict monthly stream runoff in Gosudor catchment located in the Isaco state of Japan. The results indicated higher accuracy of neural network compared to regression method[9].Nilsson et al,were simulated output values of monthly runoff for two catchments in Norway, using conceptual models and neural network and combining both of them and concluded that the neural network method and its combination with conceptual models offers better estimate of the monthly runoff[13].Firat used three methods, Anfis, ANN, and generalized regression neural network to predict river flow in Seyhan catchment in Turkey. The results indicate the superiority Anfis method than the other methods[3]. According to the cases mentioned, in most researches that is used ANN for rainfall-runoff modeling, input vectors of ANN model, consist of meteorological variable and some geomorphological properties of the catchment. In present study, spatial and temporal parameters are combined in order to optimize the network input vectors for model with using empirical equations of catchment runoff and Automated Geospatial Watershed Assessment (AGWA) tools in GIS.

2. MATERIALS AND METHODS

2.1. Principles of Artificial Neural Network

Artificial neural networks (ANNs) are intelligent dynamical systems. These systems were based on experimental data and learn general lows with calculations on the data. In fact, An Artificial Neural Network (ANNs) is mathematical tools, capable of representing the arbitrarily complex non-linear relationship between dependent and independent variables in any system. ANNs possess the desirable attributes of universal approximation, ability to learn from examples without the need of explicitly representing physical processes and the capability of processing large volumes of data at high speed. It is a highly interconnected network of many of simple processing units called neurons, which are analogous to the biological neurons in the human brain. Neurons having similar characteristics in an ANN are arranged in groups called layers. The strength of connection between two neurons in adjacent layers is represented by what is known as a connection strength or weight. An ANN normally consists of three layers, an input layer, a hidden layer and an output layer. Each layer is made up of several nodes, and layers are interconnected by sets of correlation weights. The pattern of connectivity and the number of processing units in each layer may vary within some constrain. No communication is permitted between processing units within a layer, but the processing units in each layer may send their output to the processing units in the succeeding layers. The nodes receive input either from the initial input or either from the interconnections (ASCE, 2000). In a feed-forward network, the weighted connections feed activations only in the forward direction from an input layer to the output layer. The structure of a feed-forward ANN is shown in fig1.

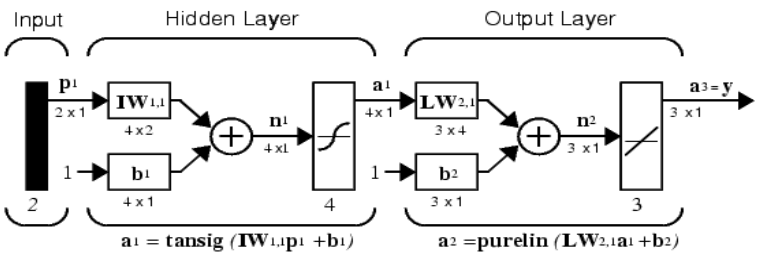


Fig1. Structure of a feed-forward ANN

2.2. Study site description

The purpose of this section is to briefly describe the data preparation and the study area. The proposed methodologies were applied to the Darkesh catchment system. It is a sub catchment with drainage area of 118 square km which is located in the North Khorasan province in Iran. Darkesh river is one of the branches of Atrak river. so, in order to determine the exact boundaries its catchment used AGWA tools in GIS software.

AGWA is a hydrological models used to evaluate the effects of land use and land cover on catchment response. It is an extension for Environmental Systems Research Institute's ArcView versions 3.x, which is directly available for downloading as an individual program suite. In this study, AGWA is applied to divide the catchment and determine sub catchment boundaries based on introduction of specific outlet point and to extract the spatial parameters of the catchment.

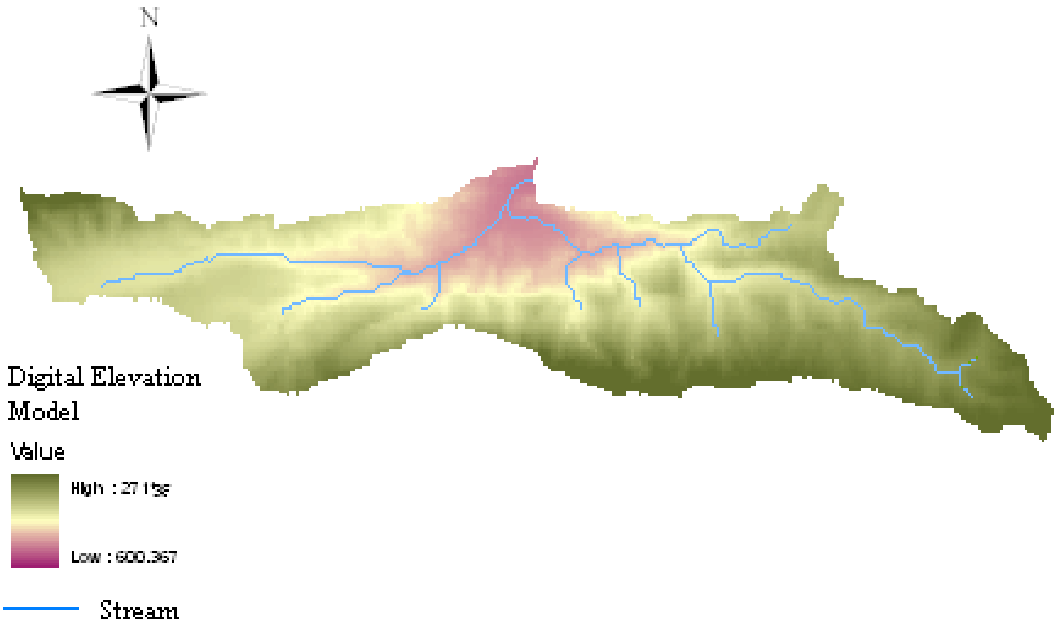


Fig2: DEM and catchment network map of the study area delineated by AGWA tools

The sources of river water include spring discharge and individual rain events. The average of annual precipitation is approximately 450 mm with rainfall durations of 1 day or less than one day.The used data for study included monthly recorded rainfall from the rain-gage of Darkesh and the average monthly temperature recorded from the climate station which is located near the catchment and monthly stream flow data. The Duration of these recorded data was 35 years from 1975-2010 and using vegetation cover parameter and the physical characteristics are extracted by AGWA tools for the catchment.

3.2. Proposed neural network design

In this research, the design of the network is multi-layer perceptron network (MLP) that performs via back propagation error law. Network structure includes input and output vectors, the number of neurons in the middle layer and calculation of model performance evaluation. The relationships between the runoff discharge of a catchment and the corresponding hydro-meteorological and catchment variable is highly non-linear and complex. So, in hydrological forecasts, regardless of type and application of models, must have a correct and clear vision from data and available information. Hence empirical equations of runoff calculation have been used to combine the network input vectors (see Fig3). Indeed, the effective parameters divide by two categories: temporal parameters such as Precipitation and air temperature which are varied during the time period and spatial parameters of area, slope, stream length, etc. The parameters inter to the network as synthetic vectors. Then impact of these vectors in training and testing stages is evaluated.

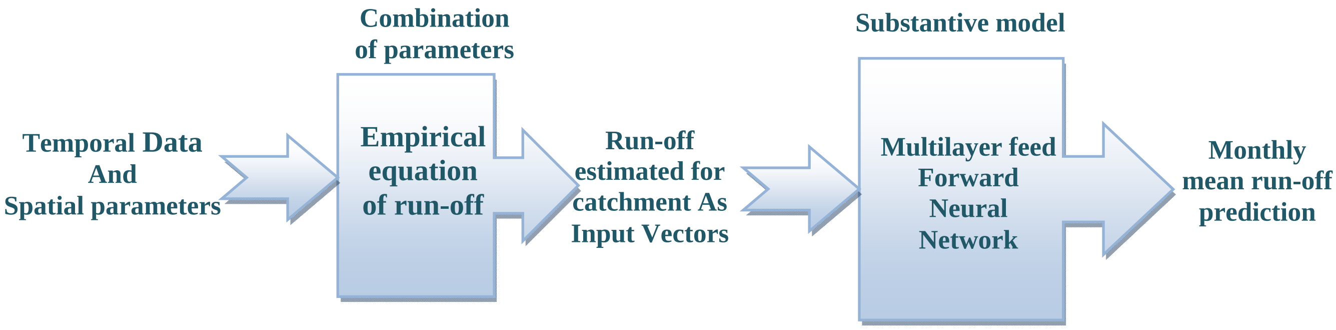


Fig3: Diagram showing linkage between the empirical equations of runoff estimation and ANN model.

4.2. Review of empirical equations of runoff estimation

The connection between rainfall events and physiographic characteristics such as slope, stream length, soil type and vegetation cover can show by using simple empirical equations. These equations consider the parameters that are the most important controlling factors of surface runoff production potential.Hence, some more conventional equations were evaluated in regional conditions and were compared the results with observations evidence so, four methods that provide acceptable results in terms of catchment were selected as the major equations. They were used in computing of network input vectors. These equations are presented in the following:

Khuzla [8] proposed following equation to calculate runoff using temperature and rainfall parameters

$$R = p - \frac{T}{3.74}$$

Laisy [16] presented the following equation for estimating annual runoff based on reviews in several catchments.

$$R = 0.284.SH^{0.155}.\frac{P^2}{(1.8T + 32)}$$

$$SH = \frac{\Delta H}{A^{0.5}}$$

Justin [5] investigated extensively on the relationships between annual rainfall and runoff in many catchments with different climatic conditions and presented their results as a following general equation:

$$R = \frac{P}{1 + \frac{304.8}{p}.\frac{F}{Z}}$$

Also, India irrigation department presented following equation between annual rainfall and river runoff [21].

$$R = p - 1.17P^{0.86}$$

All variables in above equations as the following:
P; average rainfall (cm), R; the corresponding runoff, T; average air temperature, A; catchment area (km), ΔH; maximum elevation difference of catchment, F; Rainfall duration parameter, Z; Coefficient for vegetation

5.2. Network training

The learning process or training forms the interconnection between neurons and is accomplished through known inputs and outputs, and by presenting these to the ANN in some ordered manner. The weights of the interconnections are adjusted using an error convergence technique so that a desired output will be produced for a known input pattern. Training phase as an important step in developing the ANN model is the determination of its weight matrix. There are primarily two types of training mechanisms, supervised and unsupervised. A supervised training algorithm requires an external teacher to guide the training process. The primary goal in supervised training is to minimize the error at the output layer by searching for a set of connection strengths that cause the ANN to produce outputs that are equal to or closer to the targets. A supervised training mechanism called back-propagation training algorithm [17, 23]. It is one of the most commonly used procedures that presented in several forms and algorithms. An efficient algorithm is Levenberg-Marquart, which was greatly increases convergence speed and to accelerate the conclusions. It provides a feed forward neural network, giving the capacity to capture and represent relationships between patterns in a given data set. This technique was used as training procedure throughout the present study.
In order to provide adequate training, network efficiency was evaluated during the training and validation of stages, as suggested by Rajurkar *et al.* [15]. In this case, if the calculated errors of both stages keep decreasing, the training period is increased. This is continued to the

point of the training stage error starting to decrease, but the validation stage error starting to increase. At this point, training is stopped to avoid overtraining and optimal weights and biases are determined [15, 1]. For each one of the developed models available data must be divided into two parts to use in training and validation stages [20].here, total data were separated as 80% for training and 20% for testing.

3. RESULTS AND DISCUSSION

3.1. Criteria for model Evaluation

The model performance is measured with some efficiency terms. Each term is estimated from the predicted values of the ANN and the measured discharges (targets) as follows:

1-The correlation coefficient (R-value) has been widely used to evaluate the goodness-of-fit of hydrologic and hydrodynamic models [10]. This is obtained by performing a linear regression between the ANN-predicted values and the targets and is computed by

$$R = \frac{\sum_{i=1}^{N} t_i p_i}{\sqrt{\sum_{i=1}^{n} t_i^2} \sqrt{\sum_{i=1}^{n} p_i^2}}$$

Where R is correlation coefficient; N is the number of samples; $t_{i=}T_i - \overline{T}$; $P_i = P_i - \overline{P}$ and T_i, P_i are the target and predicted values for i=1,….., N and $\overline{T}$, $\overline{P}$ are the mean values of the target and predicted data set, respectively.

2-The ability of the ANN-predicted values to match measured data is evaluated by the Mean Square Error (MSE). It is defined [18].

$$MSE = \frac{\sum_{i=1}^{N} (T_i - P_i)^2}{N}$$

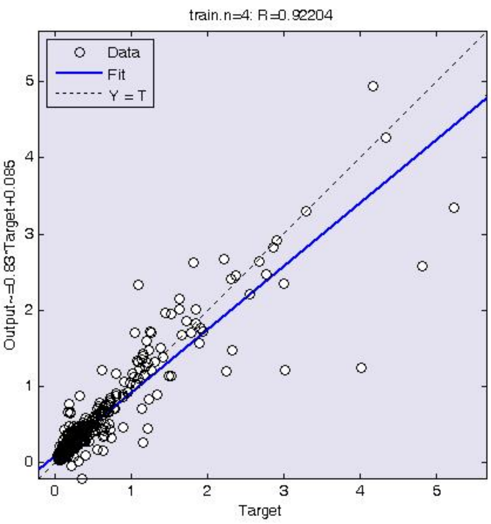
The ANN responses are more precise if R, MSE are found to be close to 1, 0, respectively. In the present study, R and MSE are used for network training and testing stages. Performance efficiencies of each trial have been recorded and compared.

3.2. Model performance evaluation

In neural network methodology, learning , which extracts information from the input data, is a crucial step that is unfavorable affected through the selection of initial weights and the stopping criteria of learning [14,16]. If a well – designed neural network is poorly trained, the weight values will not be close to their optimum and the performance of the neural network will suffer[20]. Little research has been conducted to find good initial weights. In general, initial weight is implemented with a random number generator that provide a random value[14,7]. So, to stop the training process, we could either limit the number of iterations or set an acceptable error level for the training phase. There is no guarantee that coefficients which are close to optimal values will be found during the learning phase even though the number of iterations is capped at a predefined value. Therefore, to ensure that overtraining does not occur, MSE is predefined and the training is conducted until the MSE decreases to the threshold value. In this study, it was observed that the neural network error decreases as low as threshold MSE within 100 epochs for good initial weights without overtraining; however, the threshold values may never be achieved poor initial weights, even after a large number of epochs. The result of model performance level during training and testing stage based on the evaluation of criteria is presented in table 1.Best state for network prediction usually is that structure with the superior correlation coefficient (R) and minimum sum of squares error (MSE) in training stage. So, as presented in table 2, the best structure for model is 4-4-1 with the correlation coefficient R and sum of squares error MSE: 0.93, 0.085. Increasing error coefficients in testing stage are normal and here, the superior of correlation coefficient in testing stage indicates the well- training network.

MSE		ضریب همبستگی R		آرایش شبکه
تست	آموزش	تست	آموزش	
0.634	0.152	0.928	0.902	4-3-1
0.208	0.084	0.931	0.922	4-4-1
0.648	0.992	0.895	0.914	4-5-1
0.222	0.091	0.921	0.916	4-6-1
0.158	0.089	0.921	0.923	4-7-1
0.197	0.086	0.860	0.920	4-8-1
0.929	0.089	0.895	0.917	4-9-1
0.415	0.090	0.864	0.916	4-10-1

Table 1: Result of model performance evaluation during training and testing stages



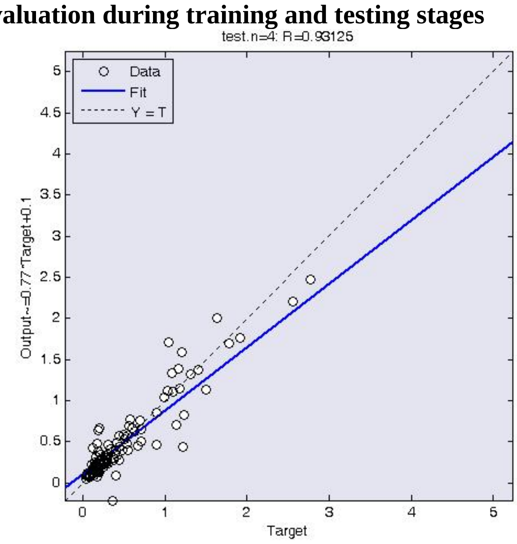


Fig4. Comparison of observed runoff with predicted values by using ANN model in training and testing

4. CONCLUSIONS

In this research were studied the performance of MLP neural network for rainfall – runoff modeling and were investigated the impact of combining spatiotemporal parameters in input vectors using empirical equations to simulate monthly runoff in the Darkesh river. Based on network analysis and comparison of the mean square error MSE in the best network structure, the model predict monthly runoff flow according to training data with a correlation coefficient 0.93. Comparison of observed with computed values in testing stage, is presented in Figure 5.

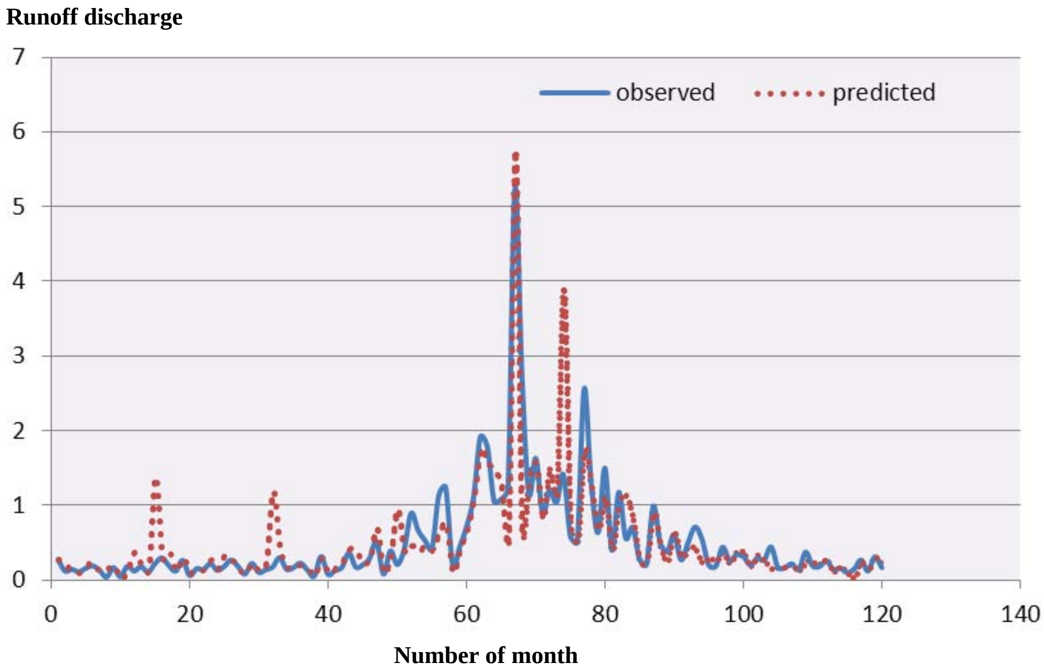


Fig5. Comparison of observed runoff with predicted values by using ANN model

Although the ability of present model to simulated monthly runoff is important, Combination of effective parameters in the network have increased generalization capability of model for estimation of runoff amount in ungagged catchment, that have similar hydro-meteorological conditions. So, the model can be generalized to homogeneous catchments in the study area to cover the lack of gauging station in this catchments.Utilization of AGWA to determine watershed boundaries and automated extraction of catchment structure and spatial parameters was an important step for increasing the accuracy of model inputs that has certainly had considerable impact on results.

REFERENCES

[1] Chiang, Y.-M., L.C., Chang and F.J. Chang, 2004. Comparison of static-feedforward and dynamicfedback neural networks for rainfall-runoff modeling. Journal of Hydrology, 290:97-11.

[2] Dooge, J.C.I., 1977. Problems and methods of rainfall-runoff modeling. In: Ciriani, T.A., U. Maione and J.R. Wallis (Eds.). Mathematical Models for Surface Water Hydrology: The Workshop Held at the IBM Scientific Center, Pisa.Wiley, London, pp: 71-108.

[3] Firat,m.,(2007) "Artificial intelligence techniques for river flow forecasting in the Seyhan river catchment"Turkey.J Hydro1 Earth Syst Sci 4:1369-1406.

[4] Gautam, M.R., K. Watanabe and H. Saegusa, 2000.Runoff analysis in humid forest catchment with artificial neural network. J. Hydrol., 235: 117-136.

[5] Gupta, B. L.(1992). "Engineering Hydrology", Second Edition, Estimation of Run off, Vol
4 , pp. 63-64

[6] Hagan, M.T., H.B. Demuth and M. Beale, 1996.Neural Network Design. 2nd Edn. PWS Publishing Company, Boston, USA runoff model using an artifcial neural network. Journal of Hydrology, 216: 32-35.

[7] Haykin, S., 1999. Neural Networks: A Comprehensive Foundation. Macmillan, NewYork.

[8] Khosla A, N. 1949. Appraisal of water resources (Analysis and utilization of data). Proc. of UNESCO. Conference on the Conservation and Utilization of Resources - Lake Success New York.

[9] Kisi,O., (2004) " River flow modeling using artificial neural networks" J hydrol Eng 9(1) 60-63.

[10] Legates, D.R. and G.J. McCabe, 1999. Evaluating the use of goodness-of-fit measures in hydrologic and hydro-climatic model validation. Water Resources Research, 35 (1): 233-241.

[11] Miller,S., Semments,D., Goordich,D., Hernandez,M.(2007) "The Automated Geospatial Watershed Assessment Tool" Environmental Modelling & Software 22 (2007) 365-377

[12] Najafi, M.R.2001.Hydrologic systems rainfall-runoff modeling. Translated from V.P.singh.Tehran:university of Tehran)

[13] Nilsson,W.E., (2006) "Comparing and combining conceptual and neural network

models" journal of Hydrology, 317:63-80.

[14] Principe, J.C., N.R. Euliano and W.C. Lefebvre,1999. Neural and Adaptive Systems: Fundamentals through Simulations. Wiley, New York.

[15] Rajurkar, M.P., C. Kothyari and U. Chaube, 2004.Modeling of daily rainfall-runoff relationship with artificial neural network. Journal of Hydrology,285: 96-113.

[16] Ray, C. and K.K. Klindworth, 2000. Neural networks for agrichemical vulnerability assessment of rural private wells. Journal of Hydrologic Engineering, 5 (2): 162-171.

[16] Ray,C., Klindworth,K.k., (2000) " Neural networks for agrichemical vulnerability assessment of rural private wells" journal of Hydrologic Engineering, 5(2) 162-171.

[17] Rumelhart, E., G. Hinton and R. Williams, 1986. Learning internal representations by error propagation. Parallel Distributive Process, 1:218-362.

[18] Schaap, M.G. and F.J. Leij, 1998. Database related accuracy and uncertainty of pedotransfer functions. Soil Science, 163: 765-779.

[19] Shamseldin, A.y. (1997) " Application of a neural networks Technique to Rainfall-Runoff modeling" Journal of Hydrology, 199(3-4) 272-294.

[20] Soleimani, k. (2009) "Rainfall-Runoff prediction based on Artificial neural networks" American-Eurasianj.Agric&Environ.Sci, 5(6) 856-865.

[21] U P. I.R.I. 1960. Rainfall Runoff studies for a few Himalayan and Bundelkhand Catchments of U.P. TM 30-PR (Hy-31).

[22] Varshney, R. S. 1979. Engineering Hydrology. Nem Chand & Bros. 917.

[23] Werbos, P., 1974. Beyond regression: new tools for prediction and analysis in the behavioral sciences. Ph.D. Thesis, Harvard University, Harvard.

CONTACTS

Shahrbanu FIRUZI
Islamic Azad University Branch of Tehran Markaz
Punak cross, Ashrafi st., Tehran, IRAN
Tel. +98915-309-6983
Email: shahrbanufiruzi@yahoo.com

Dr. mohamadBagher SHARIFI
Faculty of Civil Eng. Ferdowsi University of Technology
Azadi cross, Kosar st., Mashhad, IRAN
Email: mbsharif@gmail.com

Kambiz BORNA
Islamic Azad University Branch of Bojnourd
Tel. +98915-584-2179
Email: kambizborna@ymail.com

MohamadReza RAJABI
Faculty of Geodesy and Geomatics Eng. K._.Toosi University of Technology
Mirdamad cross, Valiasr st., Tehran, IRA_ 19967-15433
Tel: +98911-343-6025
Email: mreza.rajabi@gmail.com