Modeling the Rainfall-Runoff Data in Snow-Affected Watershed

M. Vafakhah, F. Sedighi, and M. R. Javadi, *Member, IACSIT*

Abstract—**In this study, rainfall-runoff modeling was carried out in Latyan dam watershed using artificial neural networks (ANNs) and adaptive neuro-fuzzy interface system (ANFIS). For this reason, 92 MODIS instrument images have obtained from NASA website for 2003, 2004 and 2005 years. Snow cover area (SCA) was extracted from all images. Then, snow water equivalent (SWE) was computed using SCA and SWE for mentioned years. Rainfall, temperature and SWE were used as inputs for ANN and ANFIS. Root mean square error (RMSE), Nash–Sutcliffe efficiency coefficient (NS) and determination coefficient (R^2) statistics are employed to evaluate the performance of the ANN and ANFIS models for forecasting runoff. Comparison of the obtained results reveals that the performance of ANN and ANFIS was very good for snowmelt runoff prediction. Based on the results of test stage, ANN with RMSE=0.04 m^3 s^{-1}, NS=0.85 and R^2=0.68 and is superior to rainfall-runoff modeling than the ANFIS with RMSE=0.05 m^3 s^{-1}, NS=0.65 and R^2=0.62. The combination of ANN and ANFIS by using daily SWE as input proved to be an excellent alternative to perform high quality daily snowmelt runoff prediction.**

Index Terms—**ANN, ANFIS, rainfall- runoff modeling, SWE, latyan watershed.**

I. Introduction

Snow is important in cold regions and in these areas snowmelt is of importance to many aspects of hydrology including water supply, erosion, and flood control. Also, the hydrological importance of snow is not restricted to areas where it lies for months; many dryland rivers in areas with little or no snow are fed largely by meltwater from high mountains 100-1000 km away [1]. Modelling snowmelt is important for water resources management and the assessment of spring snowmelt flood risk [2]. Modelling snowmelt is especially problematic because an incorrectly simulated melt event not only incorrectly predicts flow on that day, but also on the day when the real melt occurs.

Rainfall-runoff is a nonlinear, spatially and temporarily, completely stochastic process and could not be described easily by simple models [3]. Therefore, exact prediction of the amount of runoff produced by rainfall or snowmelt is very difficult. Using these methods, though simplicity has a great amount of errors in calculations and in many locations, the results could not be relied and verified [4]. ANN and ANFIS in modeling complex nonlinear systems, successful applications of these methods in rainfall-runoff modeling have been extensively reported [3]-[13]. A comprehensive review of the application of ANNs to hydrology can be found in the ASCE Task Committee [14], [15]. However, soft computing applications in streamflow forecasting still has been advancing to provide novel and robust. Satellites now provide valuable data with higher spatial and temporal resolution. The moderate-resolution imaging spectrometer MODIS provides a good opportunity to study snow distribution on daily basis [16].

Latyan dam is one of the main sources of water for the Tehran metropolitan area [17]. Most of the precipitation occurs from January to March, and around 48% of the annual precipitation is snow which plays a significant role in providing the water resource for drinking and agricultural uses. This study is a comparative evaluation of ANN and ANFIS models and using daily SWE as input for rainfall-runoff modeling in a watershed snow affected.

II. Description of the Models

The most commonly used neural network structure is feed-forward back propagation network (FFBP) [18]. The feed-forward back propagation network (FFBP) details can be found in Hagan and Menhaj [19]. The methodology used for adjusting the weights of the ANN model was Levenberg–Marquardt because this technique is more powerful than conventional gradient descent techniques (19). Sigmoid, hyperbolic tangant and linear activation functions were used for the hidden and output node (s), respectively. The ANFIS used in the study is a fuzzy inference model of Sugeno type, and is a composition of ANNs and fuzzy logic approaches [20]. The root mean square error (RMSE), Nash–Sutcliffe efficiency coefficient (NS) and determination coefficient (R^2) were considered as statistical performance evaluation.

III. Description of the Study Area and the Data

The Jajrood River basin is located in the southern part of the central Alborz mountain range which almost entirely covers the northern part of Iran (Fig. 1). The drainage area is 435.3 km^2 up to Roodak hydrometric station at the entrance of the Latyan dam reservoir. The basin is mountainous with elevations ranging from 1700 to 4212 m. The mean elevation is 2830 m and the mean basin slope is about 45.6%. The maximum length of flow in the basin is 32.5 km.

A. Moderate Resolution Imaging Spectroradiometer (MODIS)

The MODIS products are being produced to obtain daily

Manuscript received May16, 2013; revised July 25, 2013.

M. Vafakhah is with the Department of Watershed Management Engineering, Faculty of Natural Resources, Tarbiat Modares University, P. O. Box 46417-76489, Noor, Mazandaran Province, Iran; (e-mail: vafakhah@ modares.ac.ir).

F. Sedighi and M. R. Javadi are with Islamic Azad University, Noor Branch, (e-mail: sadighi.fatemeh@yahoo.com, javadi.desert@gmail.com).

DOI: 10.7763/IJCEE.2014.V6.790

snow cover data grids with a resolution of 500 × 500 m and they are distributed The MODIS images used in this study were obtained from the National Snow and Ice Data Center Distributed Data Archive (NSIDC, http://www.nsidc.org) in hierarchical data format (HDF). The images have been saved and displayed into the ENVI. The georeferencing has been carried out automatically. Atmospheric modifications have been applied to the images by the amount of wave reflexed from the lake of Latyan dam. It has been tried to use the images with no cloud coverage on the study area. Totally, 92 MODIS images were suitable for 2003, 2004 and 2005 years. The detailed description of the MODIS snow algorithm is presented in the Algorithm Theoretical Basis Document [22].

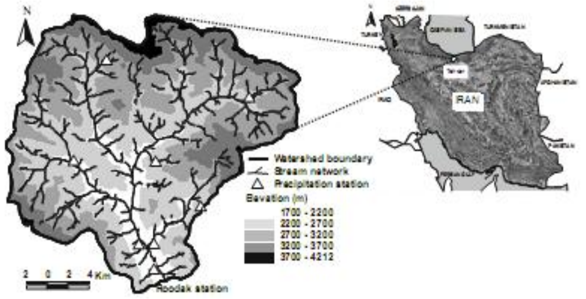


Fig. 1. Location of the Latyan dam watershed in Iran [21].

IV. SNOWMAP

The basic principle of the MODIS snow detection algorithm uses the difference between the infrared reflectance of snow in visible and short-wave wavelengths. The main distinctive feature of snow properties is a strong reflectance in the visible and strong absorption capacity in the short-wave part of the spectrum. The measure of snow reflectance difference in the MODIS snow mapping procedure is the normalised difference snow index (NDSI). The NDSI allows to distinguish snow from many other surface features and is adaptable for a number of illumination conditions. The discrimination between snow and clouds is based on differences between cloud and snow/ice reflectance and emittance properties. Clouds, typically, have high reflectance in visible and near-infrared wavelengths, while the reflectance of snow decreases towards the short-wave infrared wavelengths [23].

The mapping of snow cover is limited in areas where snow cover is obscured by dense forest canopies [22]. In the MODIS product, mapping snow in forested locations is based upon a combination of the normalised difference vegetation index (NDVI) and the NDSI [23]. Application of the NDVI index allows for the use of different NDSI thresholds for forested and non-forested pixels without compromising the algorithm performance for other land cover types. The snow cover algorithm would consider a pixel as snow only if the below conditions are satisfied:

1) Band 2 has a reflection more than 11 percent.
2) Band 4 has a reflection greater or equal to 10 percent.
3) NDSI amount should be estimated greater than 0.4.

It should be noted that final snow map is in binary format and follows from Boolean logic. In this model, the whole image is divided into two parts, with and without snow.

A. Snow Covered Area (SCA) in Days without Satellite Images

SCA in days without any MODIS images was obtained employing the cumulative depth of melting snow (ΔM). ΔM is a function of degree-day factor (α) and the number of degree-days above the critical degree-days (T^+) and is obtained from Eq. (1) in range [t_1, t_2]:

$$\Delta M(t_1, t_2) = \sum_{t_1}^{t_2}(aT^4) \qquad t_1 < t_x \tag{1}$$

$$\boldsymbol{\alpha} = 1.1\,\frac{\rho_s}{\rho_w} \tag{2}$$

In Eq. (2), ρ_s and ρ_w are snow and water densities, respectively and in recent snowfalls, the degree-day factor would be modified and introduced to the model. It has been assumed that two satellite images in times t_1 and t_2 and the SCAs extracted by these two images are $SCA(t_1)$ and $SCA(t_2)$, respectively. If the temperature falls below the critical temperature between times t_A and t_E, melting snow stops, the case in which the SCA in time t_K will be obtained

$$SCA(t_x) = SCA(t_{x-1}) - \frac{SCA(t_1) - SCA(t_2)}{\Delta M(t_1, t_A) + \Delta M(t_E, t_2)}\Delta M(t_{x-1}, t_x) \tag{3}$$

where: $SCA(t_x) =$ SCA in time t_x, $SCA(t_{x-1}) =$ SCA in time t_{x-1}, $SCA(t_1) = SCA$ in time t_1, $SCA(t_2) = SCA$ in time t_2, $\Delta M(t_1, t_A) =$ cumulative depth of melting snow between t_1 and t_A, $\Delta M(t_E, t_2) =$ cumulative depth of melting snow between t_E and t_2 and $\Delta M(t_{x-1}, t_x) =$ cumulative depth of melting snow between t_{x-1} and t_x.

Spatially distributed SWE was estimated daily at a 0.25 km^2 resolution by point SWE measurements from snow survey station namely Imameh station (station no: 41-007, 51 °36′E, 35 °54′ N, elevation: 2350 m a.sl.) operated by Iranian Water Research Institute (WRI). A linear regression was computed between SCA and SWE for determination of SWE within the days without SWE measurements.

B. Streamflow and Meteorological Data

The time series of daily streamflow data collected from Roodak station by Iranian WRI The time series of daily streamflow data collected from Roodak station by Iranian Water Research Institute were used in this study. The rainfall and temperature data were comprised the observations belonging to eight meteorological stations (Roodak, Imameh, Galookan (Kamarkhani), Rahat Abad, Ahar, Garmabdar, Shemshak and Roodbar Ghasran) and three meteorological stations (Imameh, Rahat Abad and Galookan), respectively. The average rainfall and temperature of Latyan watershed was computed using Thiessen polygon. The used data spans a period of 3 years from 24.09.2003 to 23.090.2004 (1096 days) for the mentioned station.

In the modeling process, the data sets of streamflow, rainfall, temperature and SWE were scaled to the range between 0.1 and 0.9. The 70%, 15% and 15% of the whole data set was used for training, testing and validation, respectively. The daily streamflow, rainfall, temperature and SWE statistics of training, test, validation and entire data set are presented in Table I.

TABLE I: THE STATISTICAL CHARACTERISTICS OF THE DAILY RAINFALL, TEMPERATURE AND SWE AND STREAMFLOW DATA

Variable	Data set	Numbers of data	Average	Standard deviation	Maximum	Minimum
Rainfall (mm)	Training	768	2.08	5.47	68.90	0
	Test	164	3.09	6.28	31.67	0
	Validation	164	0.95	2.55	36.68	0
	Entire	1096	2.06	5.39	68.90	0
Temperature (℃)	Training	768	10.94	8.46	29.26	-5.45
	Test	164	2.64	5.20	11.50	-9.32
	Validation	164	20	4.16	26.92	4.16
	Entire	1096	11.06	9.03	29.26	-9.32
SWE (mm)	Training	768	80.87	113	292.68	0
	Test	164	154.77	116.83	292.23	0
	Validation	164	4.07	9.74	81.71	0
	Entire	1096	80.43	112.66	292.68	0
Streamflow ($m^3 s^{-1}$)	Training	768	8.99	10.21	119	2.32
	Test	164	12.14	10.82	38.7	2.86
	Validation	164	7.82	7	33.7	2.17
	Entire	1096	9.28	9.97	119	2.17

V. RESULTS AND DISCUSSIONS

In this study, feedforward back-propagation (FFBP) and ANFIS models was accomplished using algorithms written in MATLAB, respectively. Hidden layer unit number was found separately for each of the input layer scenarios. The partial correlation analyses of the data of rainfall, temperature and SWE are employed for selecting appropriate input vectors in ANN and ANFIS runoff estimation models. The input combinations evaluated in the study were; (i) R_t, R_{t-1},T_t,SWE_t (ii) R_t,T_t,T_{t-1}, SWE_t, (iii) R_t, R_{t-1}, T_t, SWE_t, SWE_{t-1}, (iv) R_t,T_t,T_{t-1},SWE_t,SWE_{t-1}, (v)R_t, (vi)R_t,R_{t-1}, (vii) R_t,R_{t-1},R_{t-2},T_t,T_{t-1},T_{t-2}, SWE_t,SWE_{t-1}, (viii) R_t ,T_t , SWE_t,SWE_{t-1}, (ix)T_t, (x) R_t,R_{t-1},T_t,T_{t-1}, (xi) R_t,T_t,T_{t-1}, (xii) R_t,R_{t-1},R_{t-2}, T_t,T_{t-1},T_{t-2}, (xiii) R_t, R_{t-1}, R_{t-2}, R_{t-3}, T_t, T_{t-1}, T_{t-2}, T_{t-3}, (xiv) Rt, Tt, SWE_t, (xv) R_t, R_{t-1}, T_t, T_{t-1}, SWE_t, SWE_{t-1}, (xvi) R_t, SWE_t, (xvii) R_t, R_{t-1}, R_{t-2}, T_t, T_{t-1}, T_{t-2}, SWE_t, (xvii) R_t, R_{t-1}, SWE_t, SWE_{t-1}, (xviii) R_t, R_{t-1}, R_{t-2}. In all cases, the output was the streamflow Qt for the current. The RMSE, R^2 and NS statistics of ANN models in training, test and validation periods as well as the optimum ANN structures are given in Table II.

TABLE II: THE RMSE, R2 AND NS STATISTICS OF THE BEST ANN MODELS

Model inputs	ANN structure	Activation function		Training period		Test period		Validation period		
		Hidden layer	Output layer	RMSE (m^3/s)	R^2	RMSE (m^3/s)	R^2	RMSE (m^3/s)	R^2	NS
Rt,Rt-1,Rt-2	3-4-1	Sigmoid	Linear	0.05	0.23	0.06	0.51	0.06	0.09	0.06
Rt,Rt-1,Rt-2,Rt-3, Tt,Tt-1,Tt-2,Tt-3	8-4-1	Sigmoid	Linear	0.04	0.49	0.05	0.44	0.05	0.53	0.30
Rt, Tt, SWEt, SWEt-1	4-10-1	Hyperbolic tangent	Linear	0.03	0.69	0.04	0.79	**0.04**	**0.68**	**0.85**
Rt, SWEt	2-6-1	Hyperbolic tangent	Linear	0.04	0.52	0.04	0.26	0.04	0.45	0.55

By studying Table II, it is obvious that the ANN (4-10-1) model whose inputs were the rainfall of current day, temperature of current day and SWE of current and previous day (input combination viii), has the smallest RMSE and the highest R^2 and NS. Here, the ANN (4-10-1) denotes an ANN model comprising 4 inputs, 10 hidden and 1 output nodes. The training parameters of ANFIS model are given in Table III.

TABLE III: THE TRAINING PARAMETERS OF THE ANFIS

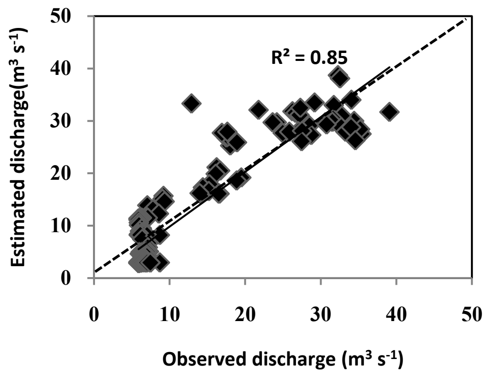
Parameters	Roodak station
Membership function	gbellmf
AND method	Prod
Or method	Maximum
Imp. method	Prod
Aggr. method	Maximum
Defuzzification method	wtaver

The RMSE, NS and R^2 statistics of ANFIS models in training and validation periods are given in Table VI. By studying Table IV, we can see that the ANFIS model used whose inputs were the rainfalls, temperature and SWE of two previous days (input combination xv) has the smallest RMSE, and the highest NS and R^2.

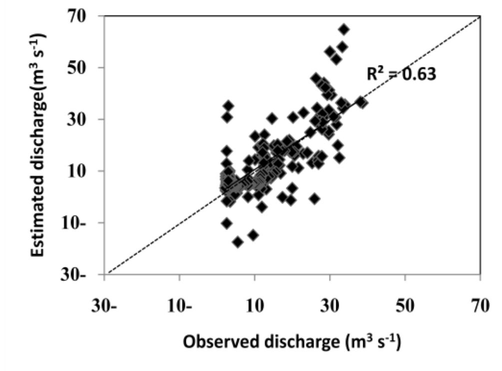
TABLE IV: THE RMSE, R2 AND NS STATISTICS OF ANFIS MODELS

Model inputs	Training period		Validation period		
	RMSE (m3/s)	R^2	RMSE (m3/s)	NS	R^2
Rt, Rt-1	0.06	0.56	0.08	0.46	0.20
Rt, Rt-1, Tt, Tt-1	0.05	0.72	0.07	0.56	0.21
Rt, SWEt	0.06	0.64	0.07	0.51	0.16
Rt, Rt-1, Tt, Tt-1, SWEt, SWEt-1	0.23	0.88	**0.05**	**0.65**	**0.62**

Fig. 2 demonstrates the streamflow forecasts of the ANN and ANFIS models in the validation period. There, the ANN and ANFIS models predict the maximum peak as 39.08 m^3s^{-1} and 36.37 m^3s^{-1} instead of measured 38.7 m^3s^{-1} with overestimation of 0.98% and underestimation of 6.02%. However, the ANN and ANFIS prediction of the second maximum peak 38.10m^3s^{-1} are 35.39m^3s^{-1} and 36.76m^3s^{-1}, respectively with underestimations of 7.11% and 3.51%, respectively. The ANN and ANFIS seem to be the best at forecasting peak flows.



(a)



(b)

Fig. 2. The scatterplot of streamflow forecasts of the ANN (4,10,1) (a) and ANFIS (b) models in validation period

REFERENCES

[1] R. I. Ferguson, "Snowmelt runoff models," *Progress in Physical Geography*, vol. 23, no. 2, pp. 205-227, 1999.

[2] J. You, D. G. Tarboton, and C. H. Luce, *Modeling the Snow Surface Temperature in an Energy Balance Snowmelt Model*, Lincoln, USA, 2004.

[3] M. Vafakhah, M. Mohseni Saravi, M. Mahdavi, and S. K. Alavipanah, "Snowmelt runoff prediction by using artificial neural network and adaptive neuro – fuzzy inference system in taleghan watershed," *Journal Iran-Watershed Management Science & Engineering*, vol. 5, no. 14, pp. 23-36, Spring 2011.

[4] M. T. Dastorani, H. Sharifi, Darani, A. Talebi, and A. M. Nia, "Evaluation of the application of artificial neural networks and adaptive neuro-fuzzy inference systems for rainfall-runoff modelling in zayandeh-rood dam basin," *Journal of Water & Wastewater*, vol. 80, no. 4, pp. 114-125, 2010.

[5] A. Dorum, A. Yarar, M. Faik Sevimli, and M. Onüçyildiz, "Modelling the rainfall–runoff data of Susurluk basin," *Journal of Expert Systems with Applications*, vol. 37, no. 9, pp. 6587-6593, 2010.

[6] A. S. Farahmand, F. V. Golkar, and M. Farahmand, "Modeling of rainfall - runoff in a river basin using artificial neural network Choir," *The First Conference of Applied Research of Water Resources. Kermanshah University of Technology*, May 11, 2010, pp. 141-147.

[7] B. Kurtulus and M. Razack, "Modeling daily discharge responses of a large karstic aquifer using soft computing methods: Artificial neural network and neuro-fuzzy," *Journal of Hydrology*, vol. 381, no. 1, pp. 101–111, Elseveir, Feb. 5, 2010.

[8] H. Z. Abyaneh and M. B. Varkeshi, "Evaluation of Artificial Intelligent and Empirical Models in Estimation of Annual Runoff," *Journal of Water and soil*, vol. 25, no. 2, pp. 365-379, 2011.

[9] M. S. Khan and P. Coulibaly, "Bayesian neural network for rainfall-runoff modeling," *Journal Water Resources Research*, vol. 42(W07409), pp. 18, 2006.

[10] N. I. Ahmat. Nor, "Rainfall-runoff modelling using artificial neural network method," Ph.D. Dissertation, Faculty of Civil Engineering, Malaysia Technology University, 2005.

[11] N. Pustizadeh and N. Najafi, "Discharge Prediction by Comparing Artificial Neural Network with Fuzzy Inference System (Case study:Zayandeh rud River)," *Journal Iran-Water Resources Research*, vol. 7, no. 2, pp. 92-97, 2011.

[12] P. R. Deshmukh and A. GhatolShort, "Term Flood Forecasting Using Recurrent Neural Networks a Comparative Study," *IACSIT International Journal of Computer Applications*, vol. 5, no. 12, pp. 0975 – 8887, August 2010.

[13] V. Nourani, Ö. Kisi, and M. Komasi, "Two hybrid Artificial Intelligence approaches for modeling rainfall–runoff process," *Journal of Hydrology*, vol. 402, no. 1-2, pp. 41-59, 2011.

[14] ASCE Task Committee, "Artificial neural networks in hydrology—I: preliminary concepts," *J. Hydrol. Eng*., vol. 5, pp. 115–123, 2000.

[15] ASCE Task Committee, "Artificial neural networks in hydrology—II: hydrologic applications," *J. Hydrol. Eng*., vol. 5, pp. 124–137, 2000.

[16] S. K. Jain, A. Goswami, and A. K. Saraf, "Role of Elevation and Aspect in Snow Distribution in Western Himalaya," *Water Resour Manage*, DOI 10.1007/s11269-008-9265-5, 2008.

[17] A. Mahmodian, H. Ghasemi, Gh. Hoshmand Fini, and M. Sarmadi, *Shemiran Township*, 3rd Edition, Iran Encyclopedia Compiling Foundation, Iranian Encyclopedia Foundation Press, 2009, ch.1, pp. 330.

[18] P. Coulibaly, F. Anctil, and E. B. Bobé, "Prévision hydrologique parréseaux de neurones artificiels: étatdelárt," *Can. J. Civ. Eng.*, vol. 26, no. 3, pp. 293-304, 1999.

[19] M. T. Hagan and M. B. Menhaj, "Training feed forward networks with the Marquardt algorithm," *Journal IEEE Transactions on Neural Networks*, vol. 5, pp. 989–993, 1994.

[20] J. S. R. Jang, "ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans. *System Management in Cybernetics*," vol. 23, no. 3, pp. 665–685, 1993.

[21] H. Zeinivand and F. De Smedt, "Hydrological modeling of snow accumulation and melting on river basin scale," *Journal of Water Resour Manage*, vol. 23, pp. 2271-2287, 2009.

[22] D. K .Hall, G, A. Riggs, and V. V. Salomonson. (2001). *Algorithm theoretical Basis Document (ATBD) for the MODIS Snow and Sea Ice Mapping Algorithms*. [Online]. Available: http://modiz.gsfc.nasa.gov/data/atbd/atbd_mod10.pdf.

[23] D. K. Hall, J. L Foster, D. L. Verbyla, A. G.Klein, and C. S. Benson, "Assessment of snow-cover mapping accuracy in a variety of vegetation-cover densities in central Alaska, Remote Sens," *Environ.*, vol .66, no. 2, pp. 129–137, 1998.

Mehdi Vafakhah received the B.Sc. degree from the Department of Rangeland and Watershed Management at Agricultural Sciences & Natural Resources University of Gorgan, Gorgan, Iran, in 1996, and the MSc and Ph.D. degrees from the Departement of Watershed Management Engineering at Tarbiat Modares University and from the Department of Rehabilitation of Arid and Mountainous Regions at Tehran University in 1999 and 2008, respectively. He is currently an assistant professor of Hydrology in Tarbiat Modares University. His research interests include surface hydrology, snow hydrology, geostatistic and parameter estimation with artificial neural networks, adaptive neuro-fuzzy inference system and several data-driven techniques. He has published many national and international scientific papers and is also involved many national watershed management projects.

Fatemeh Sedighi received the B.Sc. degree from the Islamic Azad University, Noor Branch, Iran, in 2010. She is currently an MSc student of watershed management in Islamic Azad University Noor Branch.

Mohamadreza Javadi received the B.Sc. degree from the Department of Rangeland and Watershed Management at Agricultural Sciences & Natural Resources University of Gorgan, Gorgan, Iran, in 2000, and the MSc and Ph.D. degrees from the Department of Rehabilitation of Arid and Mountainous Regions at Tehran University and from Islamic Azad University, Science and Technology Branch in 2004 and 2010, respectively. He is currently an assistant professor in Islamic Azad University, Noor Branch.