Indian Journal of Fundamental and Applied Life Sciences ISSN: 2231– 6345 (Online)
An Open Access, Online International Journal Available at www.cibtech.org/sp.ed/jls/2015/04/jls.htm
2015 Vol. 5 (S4), pp. 1090-1103/Fereydooni and Mooseloo
Research Article

INVESTIGATION OF THE SPATIAL AND TEMPORAL CHANGES IN GROUNDWATER LEVEL USING GEOSTATISTICAL METHODS(KRIGINGS) AND VALIDATION OF MODEL INPUTS BY THE USE OF ARTIFICIAL NEURAL NETWORK TECHNIQUE(MLP-RNN) (CASE STUDY: JAHROM'S PLAIN)

***Mehrdad Fereydooni and Habibollah Zarei Mooseloo**
Department of Civil Engineering, Larestan Branch, Islamic Azad University, Larestan, Iran
**Author for Correspondence*

ABSTRACT
Given the limited water resources and the risk of water crisis in Iran and that a lot of water resources reserves of the country are annually decreased as a result of irregular removal that leads to crisis of plains and thus creation of a negative balance, finding ways to predict the water level before drilling is a necessary action. The groundwater level is a variable which changes over time and space. Therefore the ability to predict the process has a considerable importance. Investigation of the statistics and data of the groundwater level in Jahrom's plain over the past years indicates a declining trend and emergence of continuous drop and reduction of the amount of groundwater reserves. In this study, in order to present a statistical model, the feed-forward artificial neural network (MLP-RNN) was used. In order to prepare the statistical model, the information and data related to the amount of precipitation, evaporation, temperature, and groundwater level of meteorological data of 19 years (from 1995 to 2014) was used. At first, the statistical data were sorted and normalized and then, 80% of the data was used to train the network while 20% was used to evaluate it. Finally, modeling of water level oscillations was conducted by neural networks with three delays in the groundwater level and R_2=87.80% and RMSE=2.71% to predict the groundwater level. In the second stage, having entered the data into the geographical information system (GIS) and using Kriging geostatistical model, contour lines maps and zoning of groundwater level drop were prepared. The results obtained from the study indicate that artificial neural networks are of the most practical models for prediction and modeling of complex hydrogeological and hydrological problems. The more attention is paid to the selection and the way of fitting of models, the more accurate will predictions be. In addition, given the prepared zoning maps and created procedures, tangible excellence of the Kriging model to inverse distance weighting method is evident.

Keywords: *Neural Network, Geostatistics, Zoning, Prediction, Groundwater Level, Jahrom, Feed-Forward, Network*

INTRODUCTION
1- Artificial neural network is a very effective tool for solving non-linear models of different hydrological, hydraulic etc applications. The advantage of artificial neural network, which is no need to complex forms of mathematical calculations, is used as an effective tool for modeling the groundwater level (Jamshidi *et al.*, 2012). Out of artificial neural networks architectures, the feed-forward propagation algorithm, as an efficient algorithm in the field of prediction of groundwater level, has been used to simulate the groundwater level of different plains (Mohammadi *et al.*, 2009). Statistical modeling of phenomena that are changing and evolving over the space and time is necessary, especially in the areas such as geology, hydrology and the environment. Einon, BP and Switzer, 1983, were among those who used spatial and temporal analyses to study atmospheric pollution. In the last decade, the use of artificial neural networks in various fields of engineering science such as water engineering has become increasingly popular, the reason of which is capability of the methods in simulation and accurate estimation of non-linear functions. In fact, processing empirical event, artificial neural networks are able to learn the rule hidden at the heart of data (Toth *et al.*, 2000).

Different types of neural networks have been invented in accordance with the learning algorithm and also network architecture. One of the most useful types of the networks in water engineering is Multi-Layer Perceptron (MLP) with the error propagation algorithm (Coulibaly *et al.*, 1999).

Statement of the Problem

Groundwater constitutes the biggest available reserve of fresh water on the Earth. In the regions where surface water resources are limited or are not easily accessible to humans, human's need to water can be fixed by the use of groundwater which is widespread and spread everywhere (Office of Water Resources Studies, 2004).

In many hydrological problems as well as groundwater resources studies, availability of groundwater statistics and data is very important. In recent years, irregular removal of groundwater resources on the on hand and activities for the supply of food on the other hand has caused irreparable damages to the resources both quantitatively and qualitatively. In order to evaluate the effects of development in current situation and present management methods of groundwater resources both quantitatively and qualitatively, mathematical and computer simulation of the resources is regarded as a powerful tool for optimum utilization of the resources.

The city of Jahrom, as one of the largest cities of Fars province, is located East of Shiraz city. Given the water shortage crisis in the country and successive droughts in Fars province during the past few years, the necessity of dealing with low cost efficient solutions to investigate the groundwater conditions is necessary more than ever. In this study, the data of groundwater level of Jahrom's plain wells over 19 years were used. Accordingly, out data are related to the period October 1995 to March 2014 (234 months), around 80% of which (187 months) were used for training in the neural network and the remaining 20% (46 months) were used for evaluation.

MATERIALS AND METHODS

Artificial Neural Networks

Artificial neural networks are elements of dynamic systems that transfer the knowledge or rule behind data to the network structure by processing the data. For this reason, the systems are called intelligent. However, an artificial neural network is a data processing system which has definite performance characteristics similar to a biological network such as receiving data and signals in parallel as well as sensing, processing and generalization.

- Elements of an artificial neural network

1- Input layer that receives input data.

2- Hidden layer

3- Output layer that generates output data.

Feed-forward networks (MLP)

In these networks, neurons receive their input from previous layer only and transfer their output to the next layer only. Studies on feed-forward multi-layer neural networks date back to early works of Frank Rosen Blot on single layer Perceptron neural networks and early works of Bernard Vidoro and Marianne Hoff.

- Feedback or return networks (RNN)

In feedback networks, at least the output of one neuron is used as the input of the same neuron or neurons of the same layer or previous layer.

- Types of algorithms in a network

- Levenberg-Marquardt algorithm

The above algorithm has a high speed and its results is quite close the minimum error. The algorithm has been used successfully in main studies in accordance with conditions and has a high efficiency and stability.

- Gradient Descent with Momentum (GDX)

In this method, the algorithm of backward propagation of error is used to calculate the network error and determine the weight vector and the network bias so that the network has the least possible error. The network results are obtained considering the amount of changes in the weight and bias by multiplying in momentum and backward propagation of error.

- Evaluation criterion of the model performance and error

In this study, two types of numerical criteria are used to evaluate the performance and error of each network as well as its ability to predict accurately.

1- Root Mean Square Error (RMSE)

Root mean square error is calculated by the following formula:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_t - y_o)^2}$$

In the equation, y_0, y_t and N is respectively observation results, calculation results and the total number of observations. RMSE shows the difference between the observed and calculated values. The lowest amount of RMSE shows the highest accuracy of prediction.

2- R^2 that indicates the network efficiency and is presented as follows:

$$R^2 = 1 - \frac{\sum(y_t - y_o)}{\sum y_t^2 - \frac{\sum y_o^2}{n}}$$

The most desirable result is obtained when RMSE and R^2 respectively approaches zero and one.

- Performing sensitivity analysis to determine the parameters affection groundwater level oscillations of the study region

The evaluation was carried out using five available structures of artificial neural networks consisting of a combination of three algorithms and two feed-forward and return networks (MLP, RNN). In this study, the algorithms and neural networks used were respectively (LM, GDX, BR) and (MLP, RNN). Performing sensitivity analysis on parameters, in addition to determination of the effective factors on groundwater level oscillations, the best structure of the artificial neural network can be selected for the studied plain and used in the next stages.

According to the conducted studies and sensitivity analysis, precipitation, average temperature, total evaporation and groundwater level of the wells can be known as the factors affecting the groundwater level in the modeling of groundwater level oscillations of Jahrom's plain.

The only station with relatively accurate and complete statistics in the studied region is the study region of Jahrom synoptic station.

The amount of precipitation, temperature, evaporation and water level of the observation wells in a period of 19 years (water year of 1995-2014) were used as the input data of the network.

Table (4-1) shows the correlation coefficient of groundwater level of Jahrom's plain with its previous months' data.

Table 4-1: Correlation coefficient of groundwater level of Jahrom's plain with its previous months, evaporation, precipitation and temperature

Correlation coefficient of groundwater level of Jahrom's plain with its previous months, evaporation, precipitation and temperature							
W-1	0.9685	E-1	0.0388	P-1	0.0078	T-1	0.0082
W-2	0.9217	E-2	0.076	P-2	0.0178	T-2	0.0285
W-3	0.8651	E-3	0.1107	P-3	0.0312	T-3	0.046
W-4	0.8164	E-4	0.1233	P-4	0.0327	T-4	0.0448
W-5	0.7811	E-5	0.1033	P-5	0.0214	T-5	0.0267

Patterns

Pattern 1: precipitation (Pt), temperature (Tt), evaporation (Et) and water level in the studied observation well (Wt)

Pattern 2: precipitation (Pt), temperature (Tt), evaporation (Et) and water level in the studied observation well (Wt-1, Wt)

Pattern 3: precipitation (Pt), temperature (Tt), evaporation (Et) and water level in the studied observation well (Wt-2, Wt-1, Wt)

Pattern 4: precipitation (Pt, Pt-1), temperature (Tt), evaporation (Et) and water level in the studied observation well (Wt)

Pattern 5: precipitation (Pt-1, Pt), temperature (Tt), evaporation (Et) and water level in the studied observation well (Wt-1, Wt)

Pattern 6: precipitation (Pt, t-1), temperature (Tt), evaporation (Et) and water level in the studied observation well (Wt,t-1,t-2)

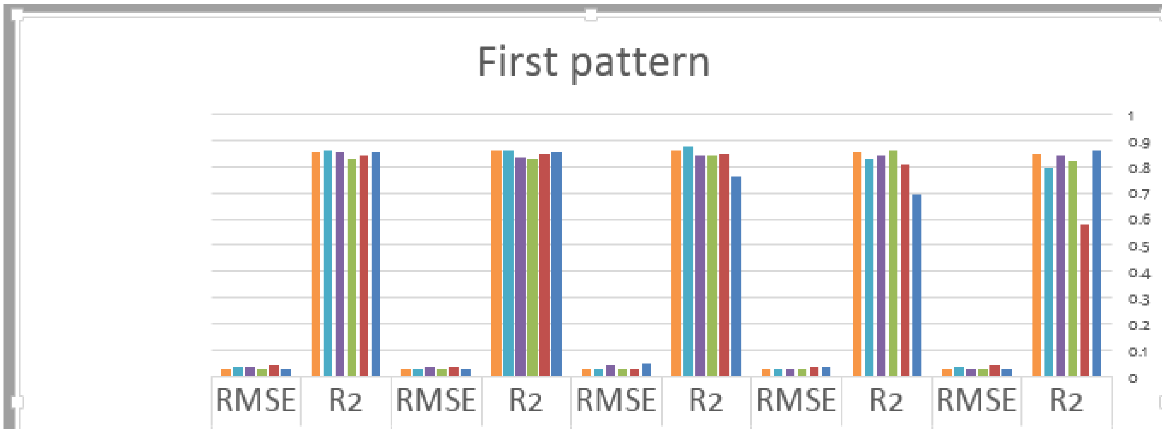
These different patterns were investigated by five existing artificial neural networks (MLP-RNN) and algorithms (LM-GDX-BR) and with various structures with 7, 4, 5 and 6 input nodes, 6, 4, 3, 2, and 1 middle nodes and one output. The results related to RMSE and R^2 have been given for each pattern with definite nodes in percentage.

The results obtained from different structures of the six stated patterns using the error evaluation criteria of R^2 and RMSE, which are given in table (4-2), indicate that the first pattern of data has resulted in the best outcomes.

Table (4-2): Best results of the structures of the six stated patterns

Patterns Ranking						
Patterns	Ranking	Type of pattern with delay	Number of Layers	Network type And Algorithm	R2	RMSE
First	First	P(t)T(t)E(t)W(t)	4-3-1	RNN-GDX	0.878	0.0271
Second		P(t)T(t)E(t)W(t,t-1)	5-6-1	RNN-BR	0.868	0.0288
Third		P(t)T(t)E(t)W(t,t-1,t-2)	6-2-1	MLP-BR	0.848	0.0346
Foruth		P(t,t-1)T(t)E(t)W(t)	5-6-1	MLP-BR	0.8650	0.0283
Fifth	Second	P(t,t-1)T(t)E(t)W(t,t-1)	6-6-1	RNN-BR	0.862	0.0263
Sixth		P(t,t-1)T(t)E(t)W(t,t-1,t-2)	7-4-1	RNN-GDX	0.875	0.0358

Research Article



	4-1-1 RMSE	4-1-1 R2	4-2-1 RMSE	4-2-1 R2	4-3-1 RMSE	4-3-1 R2	4-4-1 RMSE	4-4-1 R2	4-6-1 RMSE	4-6-1 R2
mlp-lm	0.032	0.853	0.028	0.854	0.048	0.765	0.034	0.698	0.027	0.863
mlp-gdx	0.045	0.84	0.034	0.846	0.029	0.849	0.035	0.806	0.039	0.581
mlp-br	0.028	0.832	0.031	0.83	0.026	0.844	0.030	0.863	0.027	0.82
rnn-lm	0.035	0.858	0.036	0.836	0.043	0.845	0.027	0.845	0.026	0.845
rnn-gdx	0.037	0.861	0.027	0.861	0.027	0.878	0.031	0.831	0.039	0.798
rnn-br	0.028	0.859	0.026	0.864	0.028	0.861	0.028	0.855	0.026	0.848

Diagram (4-1): Results of the first pattern structure

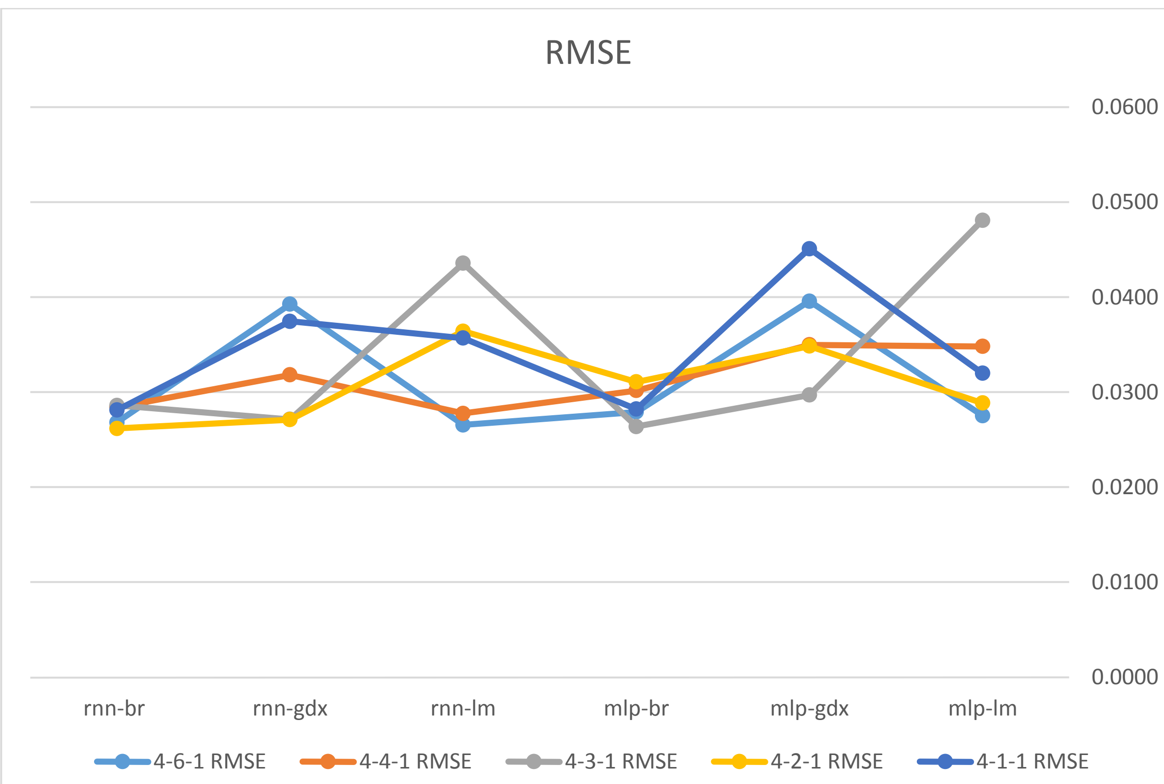


Diagram (4-2): RMSE of the first pattern

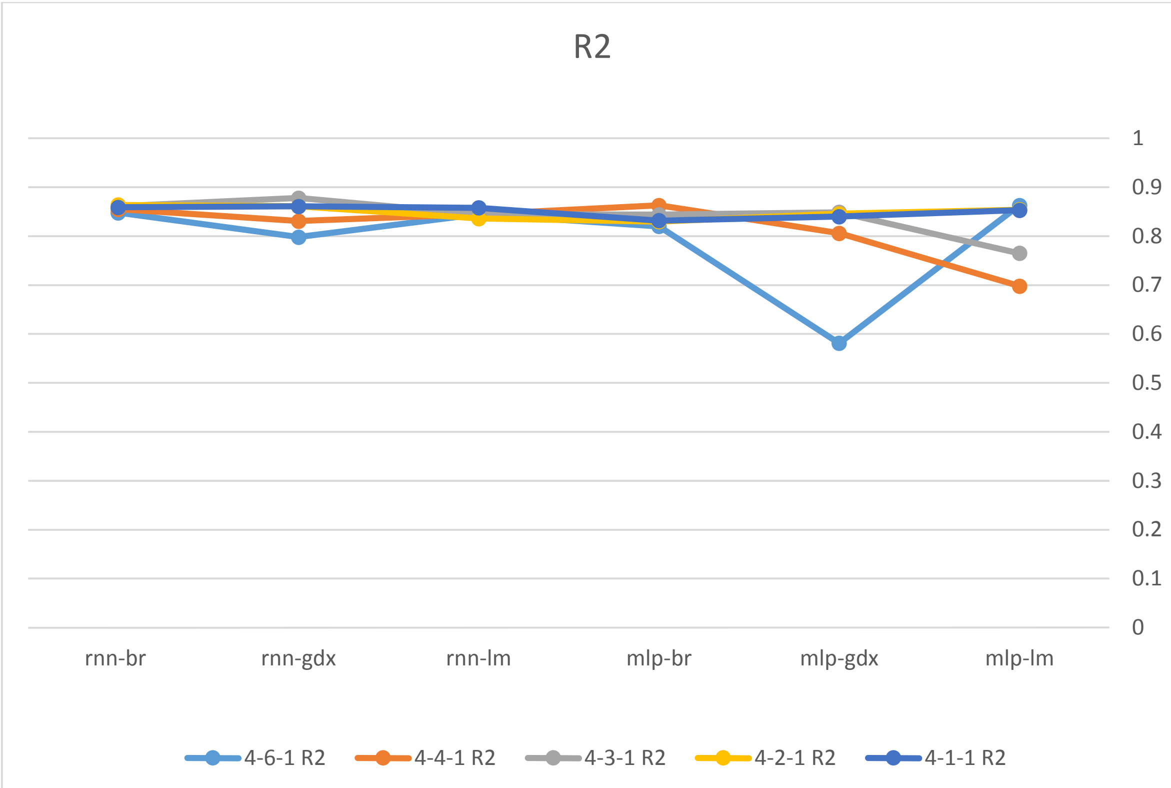


Diagram (4-3): Correlation coefficient of the selected pattern

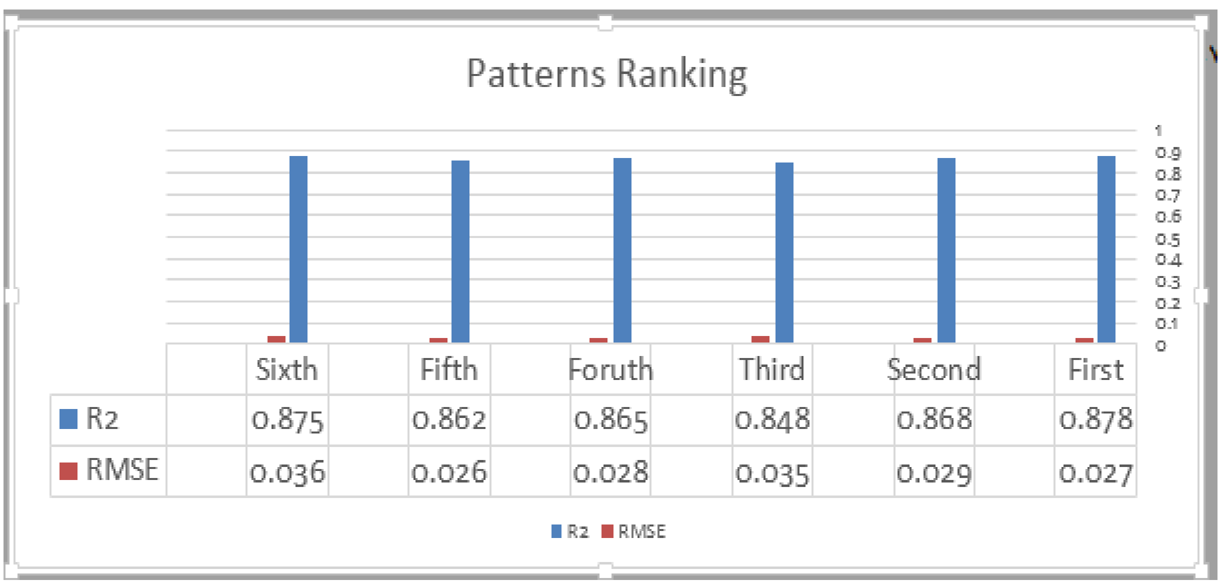


Diagram (4-4): Diagram of R^2 and RMSE of the best structure of the six stated patterns

Generally, the efficiency of different structures decreases with the increase of prediction period, and the best acceptable structure was the structure whose efficiency dropped with the lowest rate with the increase of prediction period.

The best efficiency and prediction of groundwater level for the studied plain is related to the first pattern of RNN-GDX with the number of neurons (4-3-1) and the second prediction is related to the fifth pattern of RNN-BR with the number of neurons (6-6-1) (Diagram 4-5).

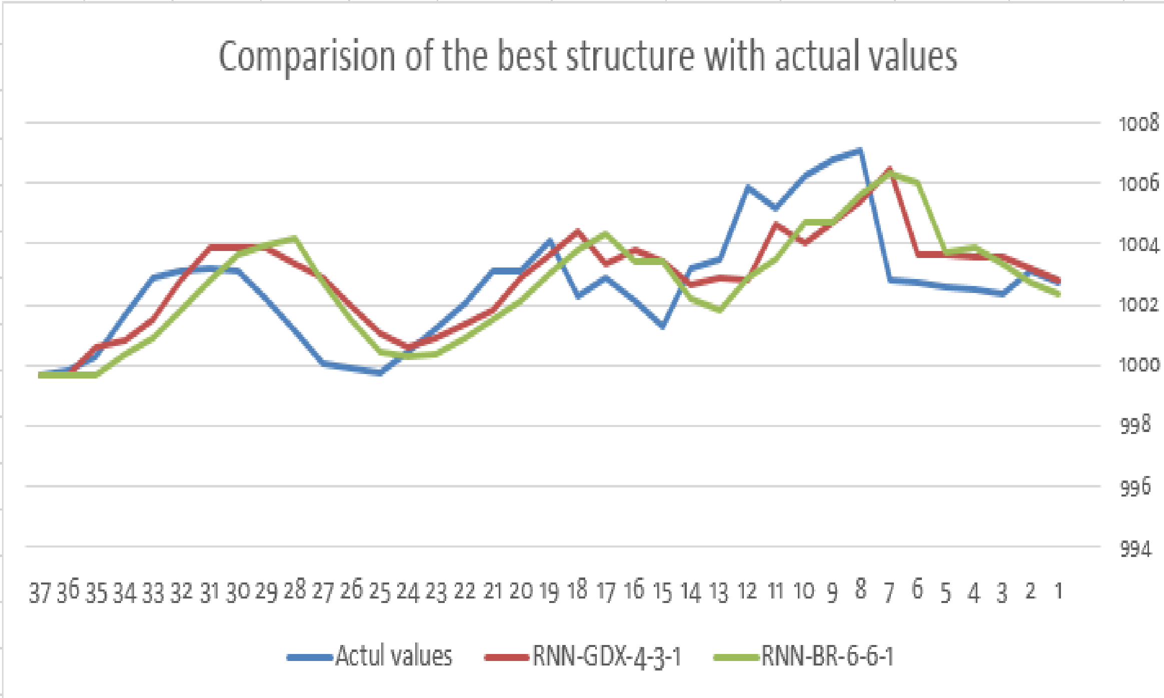


Diagram (4-5): Comparison of the best structures with actual values

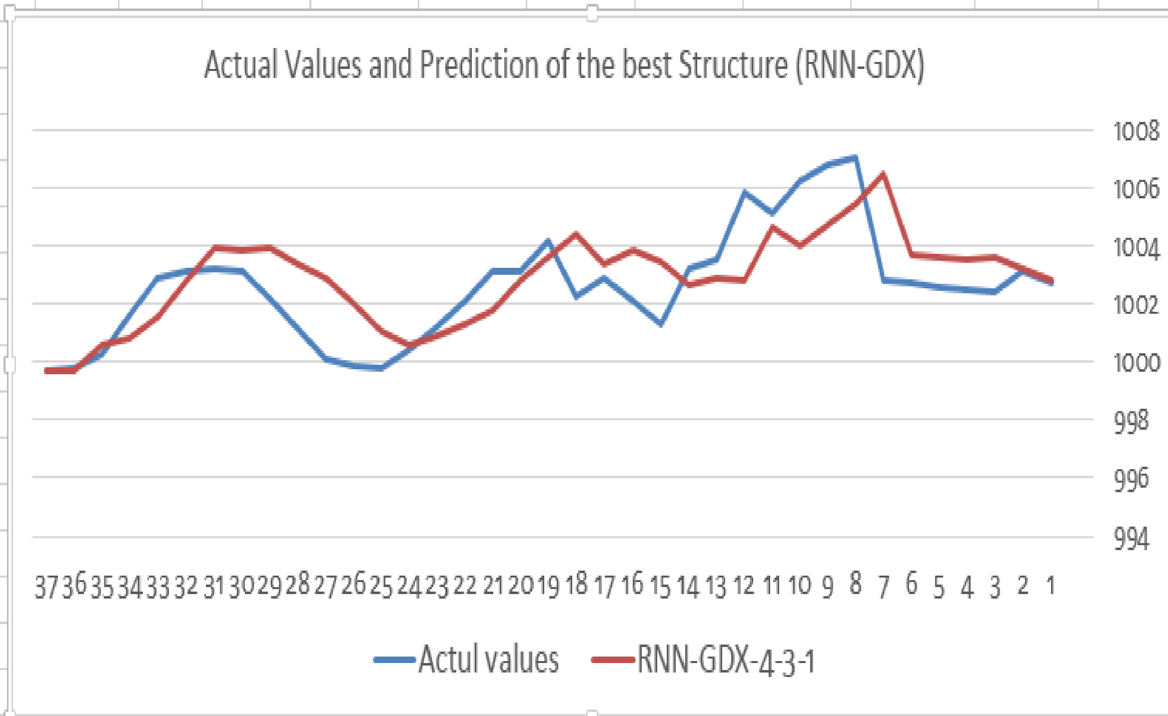


Diagram (4-6): Actual values and prediction of the best structure (RNN-GDX)

Now, given the existence of meteorological and groundwater level statistics and data up to March 2014, prediction of groundwater level of the next month (April 2014) was accomplished by the prediction software as the following table.

Station	Xt	Yt	pridict
Taghi Abad	52.961	29.119	1001.55
Heydar Abad	52.888	28.974	1001.32
Nirugah	53.211	29.018	1000.02
Hossein Abad	53.118	28.907	999.93
Ghal'e Hana	52.971	28.73	1001.3
Forudgah	53.06	28.587	999.25
Emamzade Hassan	53.208	28.706	1002.1
Emamzade Abolfazl	53.36	28.522	999.35
Turan Gaz	53.521	28.441	998.37
Mohammad Abad	53.648	28.522	999.96
Ticheng	53.695	28.582	1062.25
Ticheng sahraeeian	53.653	28.491	1061.30
Arkian Ticheng	53.74	28.524	1059.65
Baba Arab	53.84	28.509	1059.77
Dehriz	53.748	28.428	1062.35
Hava Shenasi	53.985	28.365	1063.02
Bukhen	53.643	28.332	1062.62

Figure (4-3): One month prediction results (April 2014) of the groundwater level of Jahrom's plain

Geostatistics

In order to prepare an integrated and continuous map of the site, there is a need for interpolation methods to predict unknown values. In this respect, there are a variety of interpolation and geostatistical estimation methods. Important factors such as number, spatial distribution of sampling points as well as the ability of interpolation model play an effective role in the accuracy of zoning map preparation. Numerous methods have been proposed for a variable interpolation and Thiessen polygon method and the inverse distance weighting method are among the common ones. However, the methods don't have enough accuracy as a result of not considering the correlation between data and the size and form of the used neighborhood not being optimized.

The theory of geostatistics is a branch of applied statistics that plays and effective role in description and analysis of the Earth observations. Development of geostatistics theories and techniques in recent years has created many developments in the analysis of spatial data in the GIS environment. The use of geostatistics based analyses greatly reduces the deficiencies and limitations of the classical statistics which have been developed mostly based on the random distribution of variables and processes. Data interpolation is produced by interpolation in the spatial analysis by GIS. In fact, generation of smooth and continuous models of the spatial and temporal distribution of the data is possible by interpolation. In the present study, the Kriging model was investigated for zoning of the plain and their advantages and disadvantages were dealt with.

The first stage of the analysis of the geostatistical model is to sort input data based on coordinates and level of wells groundwater, which was obtained using neural networks. Two files are needed for the inverse distance weighting and Kriging geostatistical models. The first one that is an Excel file consists of

Jahrom's plain wells as well as its groundwater coordinates and level while the second file is the map of Jahrom's plain.

The Inverse Distance Weighting Geostatistical Model

For the zoning using the inverse distance weighting method based on the GIS software, the map of Jahrom's plain together with the groundwater coordinates and level file is called (Figure 4-4).

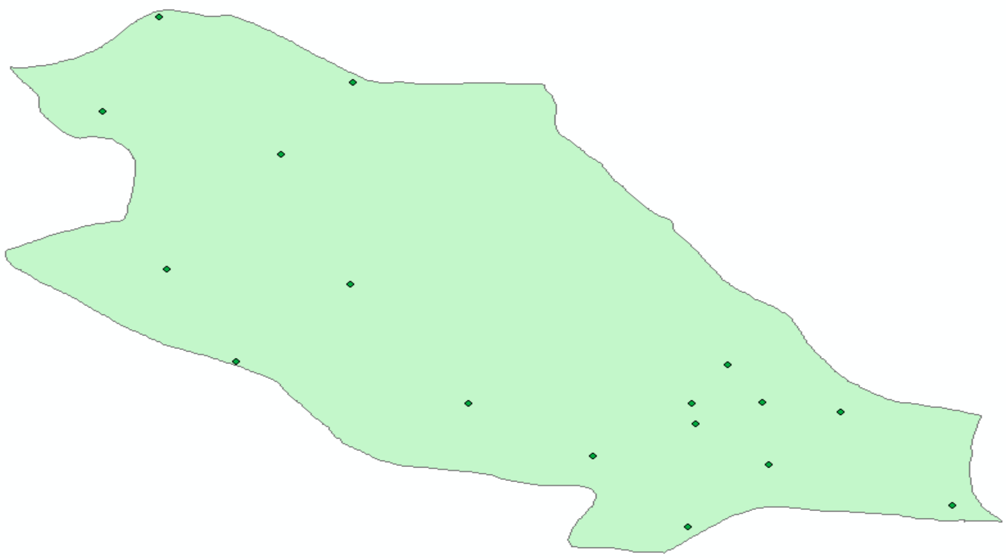


Figure (4-4): Map of Jahrom's plain together with the studied wells

Then, the power governing the network should be normalized and optimized for the model prediction so that the best case is selected based on the primary data of the network and the relationship between them. This function is carried out by the network. Figure (4-5) shows the prediction diagram of the inverse distance weighting model.

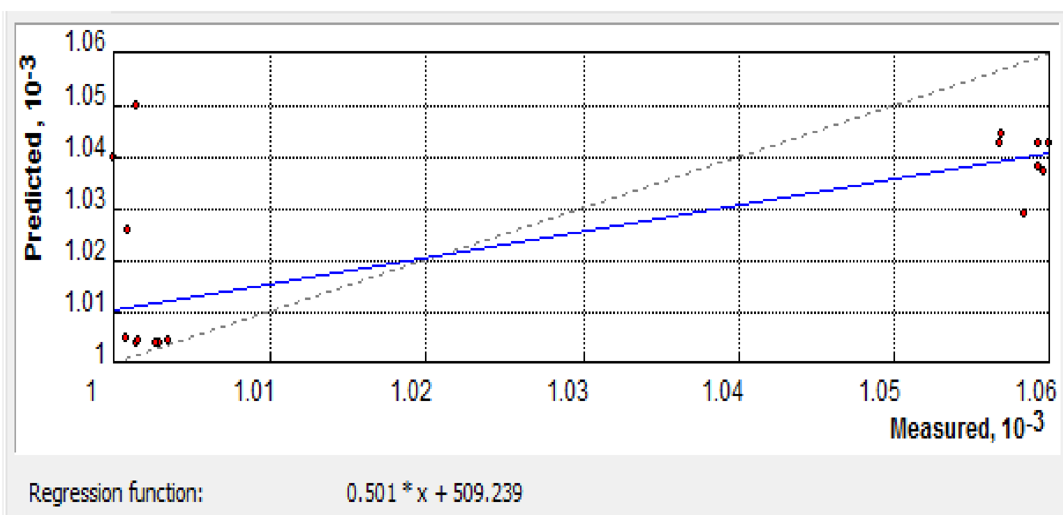


Figure (4-5): Prediction diagram of the inverse distance weighting model

In the next step, the prepared model should be converted to raster to implement the procedure and perform zoning. Then, in order to the draw the procedure, the polygon modeling and clips operation is implemented on it and the final procedure will be the software output according to figure (4-6).

Research Article

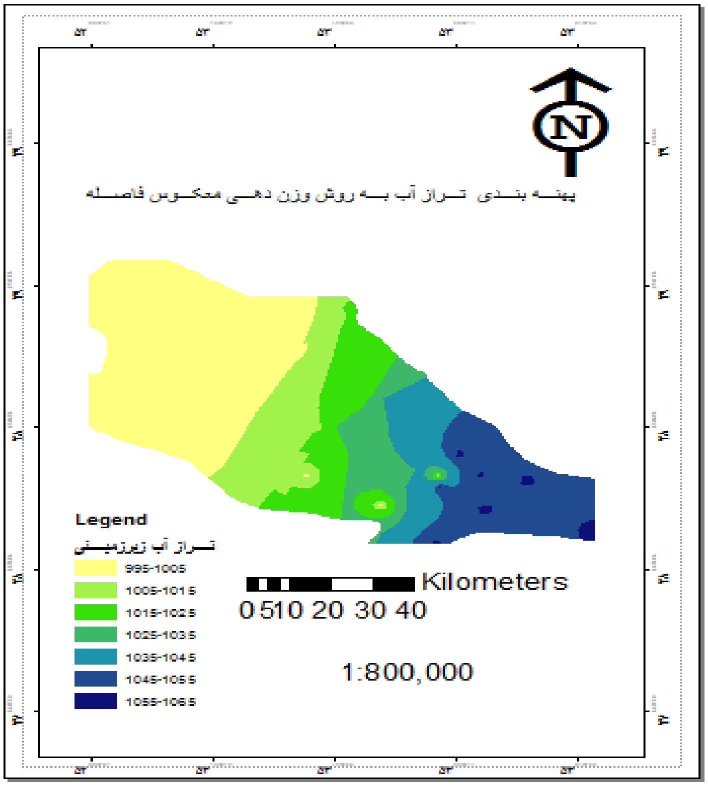


Figure (4-6): Groundwater level zoning using the inverse distance weighting model

The Kriging Geostatistical Model

Kriging is an unbiased estimator with the least estimation variance. The condition of being unbiased is implemented on other estimation methods such as polygon, inverse distance and the inverse square of the distance methods. However, the feature of Kriging is such that it determines the coefficients in such a way that while being unbiased, the estimation error will be minimum. Thus, Kriging generates estimates as well as their errors.

Using the feature, it is possible to determine the parts with high error levels and increase sampling there. On the other hand, before sampling at each point, we can estimate the variance reduction for the sample and determine the best sampling points. Kriging error is a function of variogram characteristics (spatial structure) of the estimated blocks geometry and the blocks that are used for estimation, but is not a function of the actual values of data.

The other feature of Kriging is that it causes changes to soften (reduces oscillations); that is variable functions in estimated blocks have less variations than the actual values of blocks. Therefore Kriging had better not be used when changes should not get softened.

As the distance increases, the spatial structure weakens and eventually disappears. The points located at a further distance of the "impact radius" of the estimation point practically don't have any effect on the point and it is not necessary to be entered into the point estimation. The maximum distance where the points are participated in the estimation is called the "search radius". The radius is normally assumed equal to the impact radius or equivalent to two thirds of it.

When running the Kriging program, the best model can be chosen out of the existing models based on the Nugget/Partial Sill ratio. The less the ratio than 0.25, the more powerful is the model and the more accurate is the prediction.

Thus, the above ratio was first derived for the following models and the Gaussian model was selected based on the obtained solutions.

discripction	model	percent
Nugget/Sill	Circular	0.58
	spherical	0.75
	Exponential	0.88
	Gaussian	0.469

Figure (4-7): Nugget/Sill ratio

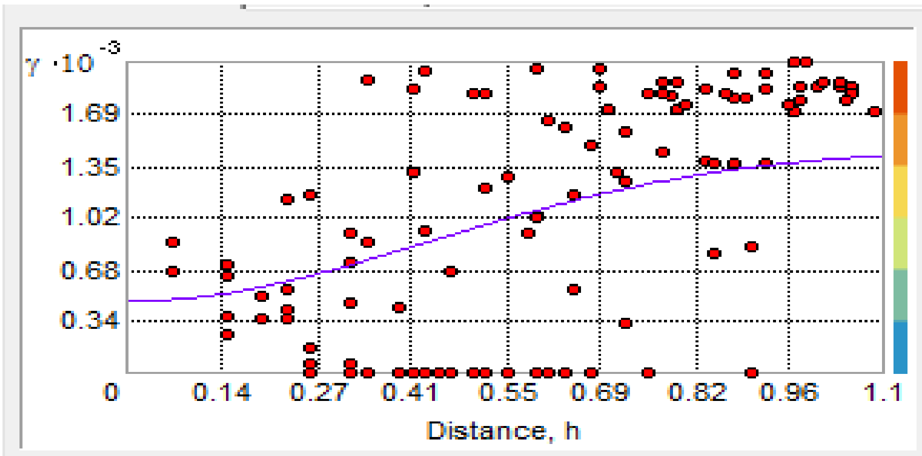


Diagram (4-7): Diagram of gamma ratio to the distance between points by the Gaussian model

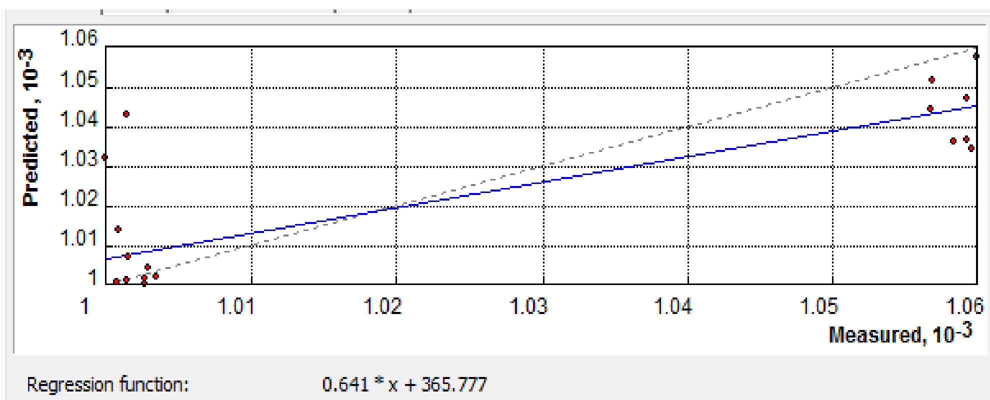


Diagram (4-8): Prediction diagram by the Gaussian model

Selection of the most suitable model for Kriging				
	Gaussian	Exponential	spherical	Circular
Root Mean Square	18.4	19.27	18.86	18.83
Avrage Standard Error	25.1	28.86	27.45	26.91
Mean Standardized	-0.0077	0.013	0.008	0.006
Root Mean Square standardized	0.77	0.7	0.727	0.74

Figure (4-8): Selection of the most suitable model for Kriging

Research Article

Investigating the above table and comparing the error level of different models, the Gaussian method was selected for Kriging as the chosen model, based on which zoning maps were prepared.

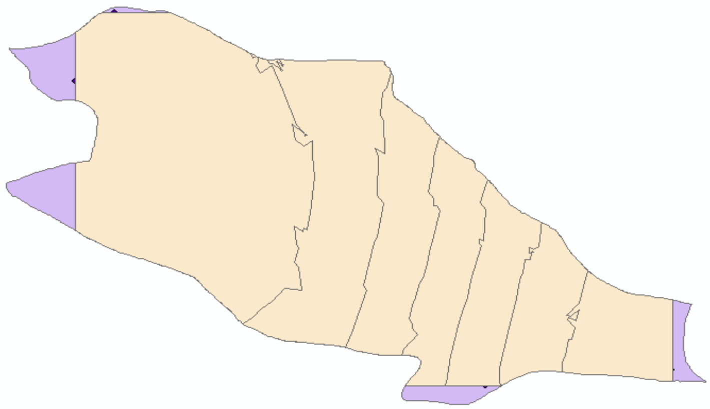


Figure (4-9): Zoning map of Jahrom's plain in the stage of conversion to raster

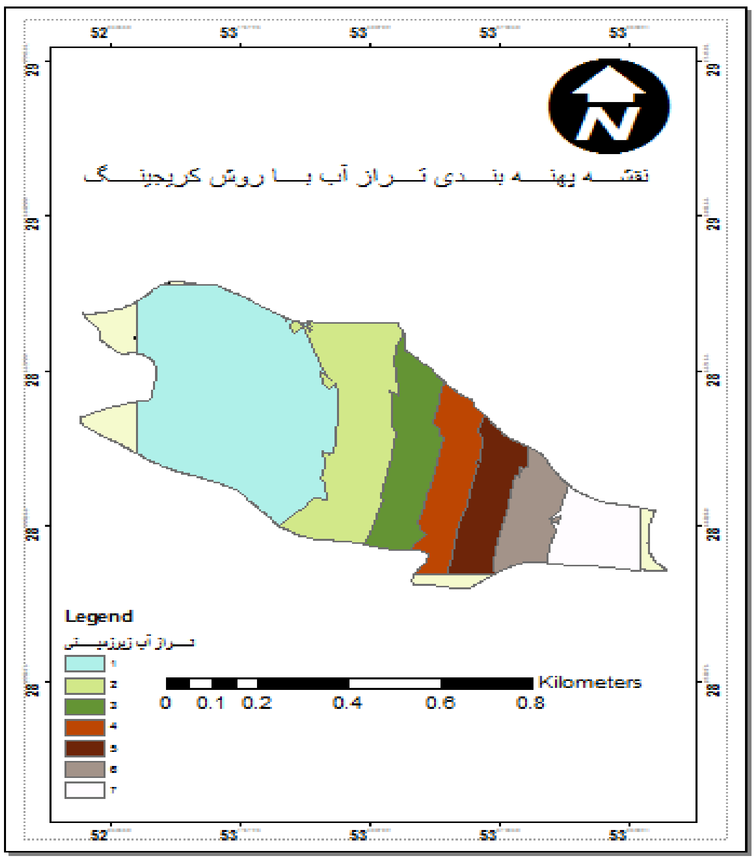


Figure (4-10): Zoning map of groundwater level of Jahrom's plain using the geostatistical model of Kriging

RESULTS AND DISCUSSION

Conclusion

According to the results obtained from the research, the following points can be mentioned:

1- Modeling and simulation of groundwater level by artificial neural networks has a special position because none of mentioned constraints exists by this method and it also has learning capability through presenting examples without the need to equations governing a phenomenon. In fact, artificial neural networks are the most applicable models for prediction and modeling of complex hydrogeological and hydrological problems.

2- The most important issue in the GIS software and the spatial-temporal analysis of data is determination of the correlation structure of data. The more accurate the selection of models type and fitness of models, the more accurate will be predictions. Thus, in this study, having empirically estimated the spatial change logs only and temporal change logs alone, they were fitted with different valid models so that the most accurate models for spatial and temporal correlations of data are obtained as much as possible.

3- From prediction point of view, it should be noted that as the time distance between the last observation and prediction time frame increases, the accuracy of predictions falls.

4- The main objective of this study was to precisely express some of the required statistical methods and present some of their results for use in various fields of climate and an objective example was used to make the results more tangible. It should be noted that the data used in this study was incomplete and up to September 2013 and updating the data more accurate predictions for upcoming years can also be conducted. In addition, if we are supposed to dig a well in a special point of the studied region in the present time or even after several months, it is possible to predict the well depth to water before wasting a lot of expenditure on digging the well and accordingly, make a decision on whether or not to dig the well and then estimate the cost of digging in accordance with its depth.

5- Based on the obtained results it can generally be stated that in spite of having more complexity of equations for artificial intelligence models, the use of linear regression model has been accompanied with better results for the problem of this study.

6- According to the prepared zoning maps and created procedures, the tangible excellence of Kriging model to the inverse distance weighting method is evident.

Suggestions

The main action that should be taken for the management method of water consumption is to prevent irregular consumption of water, especially in the agriculture sector.

1- Control of human factors affecting groundwater resources.

2- Optimization of irrigation methods and increase of efficiency

Water consumption in farmland according to modification of cultivation pattern and increase of irrigation efficiency and integration of land

3- Modification of conventional irrigation structure, promotion of use of appropriate irrigation methods and expediting implementation of the integration plan and also prevention of illegal wells excavation and excess removal of allowable wells and installation of intelligent counters on wells. Public training and promotion of culture for reduction of excess consumption and saving can be effective on this matter. Development of transparent and clear rules to restrict irregular removal of resources, serious and decisive confrontation with violators, implementation of artificial nourishing projects; and execution of the vegetation project

4- Given that each one of the interpolation models has different accuracy depending on factors such as number and distribution of sampling points and capability of models, a definite model cannot be suggested for all conditions. Thus, it is suggested that for each case, the accuracy of different models is evaluated first and the optimum method is implemented.

5- Most areas of the country, especially interior deserts, lack sensing and measurements stations. Furthermore, neighboring areas with higher climatic conditions should acquire an appropriate distribution and number so that interpolation estimates have better precision and accuracy.

REFERENCES

Copola E, Szidarovszky F, Poulton M and Charles E (2003). Artificial neural network approach for predicting transient water levels in multilayered groundwater system under variable state, pumping, and climate conditions. *Journal of Hydrologic Engineering* **8**(6) 348-380.

Coulibaly P, Anctil F and Bobe´e B (1999). Hydrological forecasting usingartificial neural networks: The state of the art. *Canadian Journal of Civil Engineering* **26**(3) 293– 304.

Coulibaly P, Anctil F, Aravena R and Bobée B (2001). Artificial neural network modeling of water table depth fluctuation. *Water Resources Research* **37** 885.

Dehghani M and Asgari Saedi (2008). Comparison of three methods of artificial neural networks, adaptive fuzzy-neural inferential system and geostatistics for interpolation of groundwater level-case study: Ghazvin's plain aquifer. *The Second Conference on Planning and Management of the Environment* **26**(4) 125–201.

Einon BP and Switzer P (1983). The variability of rainfall acidity. *Canadian Journal of Statistics* **11** 11-24.

Jahanshahi A and Moghaddamnia A (2012). Prediction of groundwater level using artificial neural networks, case study: Shahrebabak's plain- Kerman province. *The Second Conference on Planning and Management of the Environment* **22**(2) 117–151.

JamshidiAvanaki M and Ebrahimi K (2012). Comparison of artificial neural networks and geostatistical approach in prediction of groundwater level- case study: Mashhad's plain aquifer. *The Second Conference on Planning and Management of the Environment* **13**(3) 101–211.

Maier HR and Dandy GC (1996). The use of artificial neural networks for the prediction of water quality parameters. *Water Resources Research* **32**(4) 1013– 1022.

Mohammadi, Salajegheh, Mahdavi and Bagheri (2009). Investigation of spatial and temporal changes of groundwater level of Kerman's plain using an appropriate geostatistical method- case study: Kerman. *Scientific Research Quarterly of Range and Desert Research of Iran* **21**(3) 221–304.

Mohtasham M Dehghani, Akbarpoor A, MeftahHalghi M and Etebari B (2010). Prediction of water level using artificial neural networks (case study: Birjand's plain). *Journal of Irrigation and Drainage of Iran* **19**(2) 88–111.

Toth E, Brath A and Montanari A (2000). Comparison of short-term rainfall prediction models for real-time flood forcasting. *Journal of Hydrology* **239** 132-147.

Yang CC, Prasher SO, Lacroix R, Sreekanth S and Patni NKL (1997). Artificial neural network model for subsurface drained farmland. *Journal of Irrigation and Drainage Engineering - ASCE* **123**(4) 285-292.