International Journal of Geology, Earth & Environmental Sciences ISSN: 2277-2081 (Online)
An Open Access, Online International Journal Available at http://www.cibtech.org/jgee.htm
2017 Vol. 7 (2) May-August, pp.51-62/Fereydooni and Karimi

Research Article

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

Mehrdad Fereydooni and *Omid Karimi
Department of Civil Engineering, Larestan Branch, Islamic Azad University, Larestan, Iran
**Author for Correspondence*

ABSTRACT

In recent decades due to increased water demand and reduced water resources capitation has been a significant concern and how to use these resources as appropriate, effective and efficient to ensure sustainable development, is one of the most important topics in international circles. Groundwater level is a variable that changes over time and space. So, the ability to predict this process is very important.

The method in this article is based on computer modeling using artificial neural networks and the model is kriging. In this study, meteorological data such as temperature, precipitation, evaporation and Larestan groundwater level has been collected for 180 months. On the basis of data and raw data from October 1995 to March 2014, 80% in order to train the neural network and 20% of information is used to evaluate the model. First of all numerical data will be normalized based on maximum and minimum of data. A network will be created by using a neural network (MLP-RNN) and will be trained by 80% of data. Then, the program will be applied by setting and determining of repeated training parameters, error percentage and display information processing in many times in training. And evaluation of the results taken from underground water level and information used to model kriging geostatistical method is given. Due to changes in groundwater levels geostatistical map will be prepared in raster which continues become the polygons, and the polygon maps, showing the spatial variation of water level, will be the output.

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

INTRODUCTION

In recent decades due to increased water demand and reduced water resources capitation has been a significant concern and how to use these resources as appropriate, effective and efficient to ensure sustainable development, is one of the most important topics in international circles.

Iran is a semi-arid region and average of its annual rainfall is about a third of world's rainfall. In recent decades, population growth has been high in the country. This is due to the limited amount of the water, threatens potential of water precipitation seriously. In this research by use of statistical models Fluctuations in groundwater level of Lar's plain has been reviewed and possible future situation has been predicted and provided the right solution with the model. According to the importance and sensitivity of the control of surface water, especially, in our country which the most are seasonal rivers of different regions and lack of water in vast areas of the country, the need to identify and to model the behavior of rivers and long-term plans for water arteries and more and better use of their potentials deeply felt. Newly being of the most hydrometric stations, existing defects in statistics of most of the stations, location of most rivers in arid regions, critical situation from Groundwater Withdrawal and paying more attention to surface waters, all more and more cause a subtle reason that Effects and gives a more complete expression in the issue of prediction and artificial production in the catchment areas of the country. Artificial Neural Networks in Engineering Sciences is a new name that introduced in the world by Frank Rosen blot in 1962 and with new influence shape by innovating and improving the model of Perceptron considered in the world by Rommel Hart and Mcclelland in 1986.

This method of structural neurons and intelligent with proper modeling of neurons in the human brain tries to defined through mathematical functions to simulate the behavior of intracellular neurons in the brain and Synaptic function in normal neurons can be modeled through calculated weights existed in the communication lines of artificial neurons. Experimental and flexible nature of this method makes on some issues such as prediction, such an attitude can be seen in their structure and are non-linear behavior and unbridled, is well used.

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. In this study, the data of groundwater level of Lar'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 were used for training in the neural network and the remaining 20% 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 Lar's plain.

The only station with relatively accurate and complete statistics in the studied region is the study region of Lar 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 1 shows the correlation coefficient of groundwater level of Lar’s plain with its previous months' data.

Table 1: Correlation Coefficient of Groundwater Level of Lar's Plain with its Previous Months, Evaporation, Precipitation and Temperature

Correlation Coefficient of Groundwater Level of Lar’s Plain with its Previous Months, Evaporation, Precipitation and Temperature							
W-1	0.9044	E-1	0.0002	P-1	0.0017	T-1	0.001
W-2	0.8089	E-2	0.0002	P-2	0.006	T-2	0.0005
W-3	0.7248	E-3	0.0013	P-3	0.0005	T-3	0.0004

Patterns

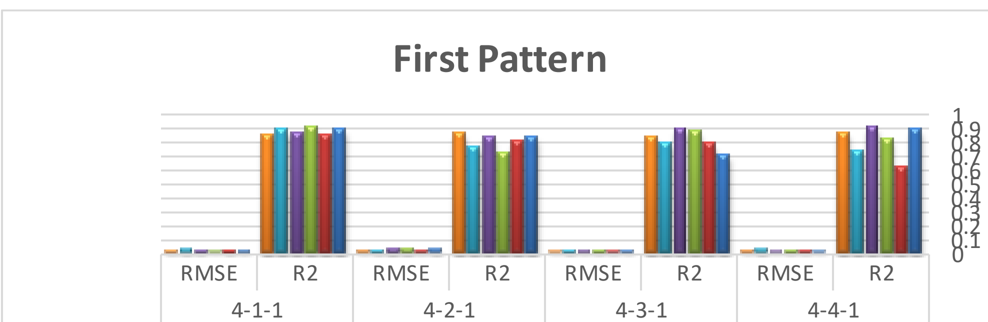
Pattern 1: precipitation (Pt), temperature (Tt), evaporation (Et) and water level in the studied observation well (Wt), input, surface of water with one delay (Wt+1) as the output (target)

International Journal of Geology, Earth & Environmental Sciences ISSN: 2277-2081 (Online)
An Open Access, Online International Journal Available at http://www.cibtech.org/jgee.htm
2017 Vol. 7 (2) May-August, pp.51-62/Fereydooni and Karimi
Research Article

Pattern 2: precipitation (Pt), temperature (Tt), evaporation (Et) and water level in the studied observation well (Wt), surface of with one delay (Wt+1) as input, surface of water with two delays (Wt+2) as the output (target)
Pattern 3: precipitation (Pt), precipitation with one delay (Pt+1), temperature (Tt), evaporation (Et), water level in the studied observation well (Wt) as the input and surface of water with one delay (Wt+1) as the output(target)
Pattern 4: precipitation (Pt+1, Pt+2), temperature (Tt), evaporation (Et) and water level in the studied observation well (Wt) as the input and surface of water with one delay (Wt+1) as the output (target)
These different patterns were investigated by five existing artificial neural networks (MLP-RNN) and algorithms (LM-GDX-BR) and with various structures with 4, 5, 6 input nodes, 1, 2, 3, 4, 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 four stated patterns using the error evaluation criteria of R^2 and RMSE, which are given in table 2, indicate that the third pattern of data has resulted in the best outcomes.

Table 2: Best Results of the Structures of the Four Stated Patterns

Patterns Ranking						
Pattern	**Ranking**	**Type of Pattern with Delay**	**No. of Layers**	**Network Type and Algorithm**	**R2**	**RMSE**
1		P(t)T(t)E(t)W(t)	4-1-1	MLP-BR	0.917	0.0149
2	2	P(t)T(t)E(t)W(t,t-1)	5-4-1	RNN-GDX	0.9364	0.0187
3	1	P(t,t-1)T(t)E(t)W(t)	5-2-1	MLP-GDX	0.9423	0.0139
4		P(t,t-1,t-2)T(t)E(t)W(t)	6-5-1	MLP-LM	0.88	0.0201



	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
mlp-lm	0.0187	0.897	0.0315	0.843	0.0192	0.713	0.0156	0.897
mlp-gdx	0.0265	0.857	0.0215	0.819	0.0200	0.795	0.0215	0.625
mlp-br	0.0149	0.917	0.0352	0.734	0.0218	0.882	0.0218	0.833
rnn-lm	0.0245	0.877	0.0382	0.846	0.0188	0.907	0.0154	0.911
rnn-gdx	0.0301	0.895	0.0241	0.772	0.0255	0.799	0.0311	0.745
rnn-br	0.0204	0.854	0.0179	0.872	0.0146	0.848	0.0189	0.875

Diagram 1: Results of the First Pattern Structure

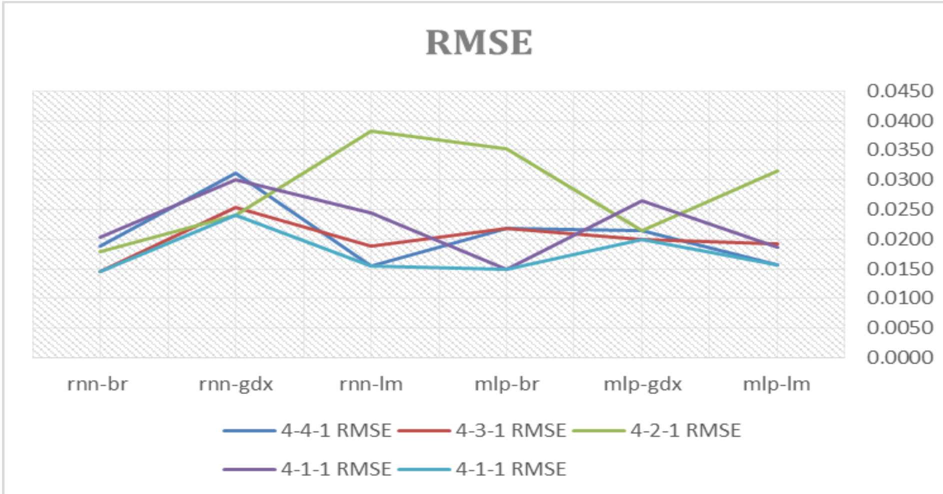


Diagram 2: RMSE of the First Pattern

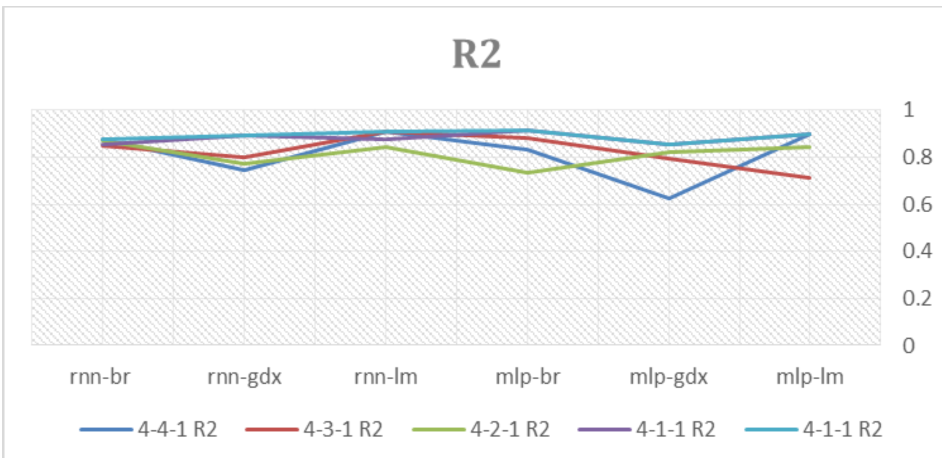
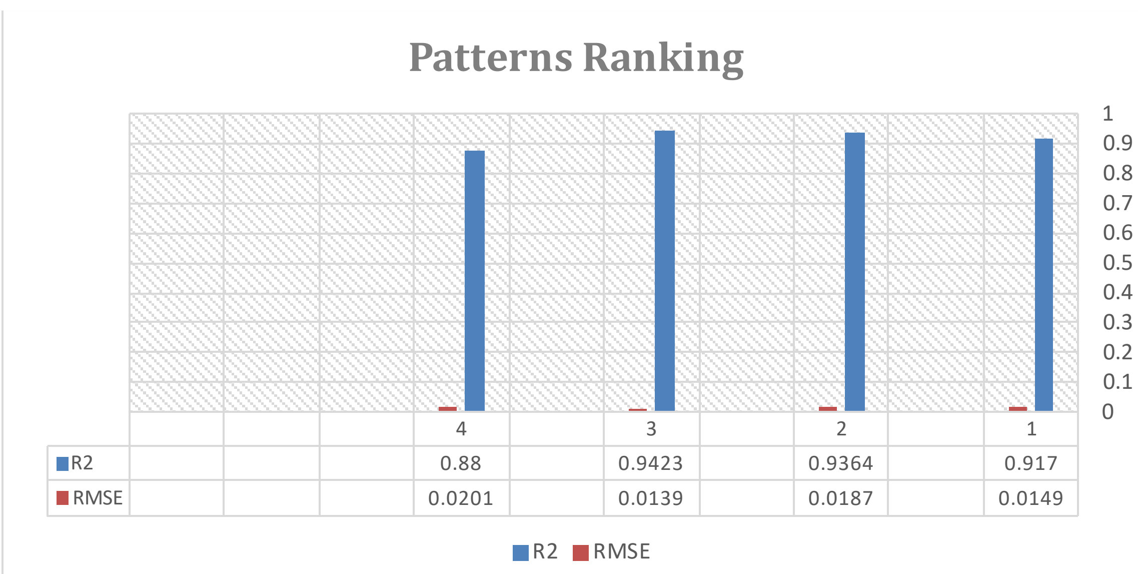


Diagram 3: Diagram of R^2 of the First Pattern



				4	3	2	1
R2				0.88	0.9423	0.9364	0.917
RMSE				0.0201	0.0139	0.0187	0.0149

Diagram 4: Diagram of R^2 and RMSE of the Best Structure of the Four Stated Patterns

International Journal of Geology, Earth & Environmental Sciences ISSN: 2277-2081 (Online)
An Open Access, Online International Journal Available at http://www.cibtech.org/jgee.htm
2017 Vol. 7 (2) May-August, pp.51-62/Fereydooni and Karimi
Research Article

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 third pattern of MLP-GDX with the number of neurons (5-2-1) and the second prediction is related to the second pattern of RNN-BR with the number of neurons (5-4-1) (Diagram 5).

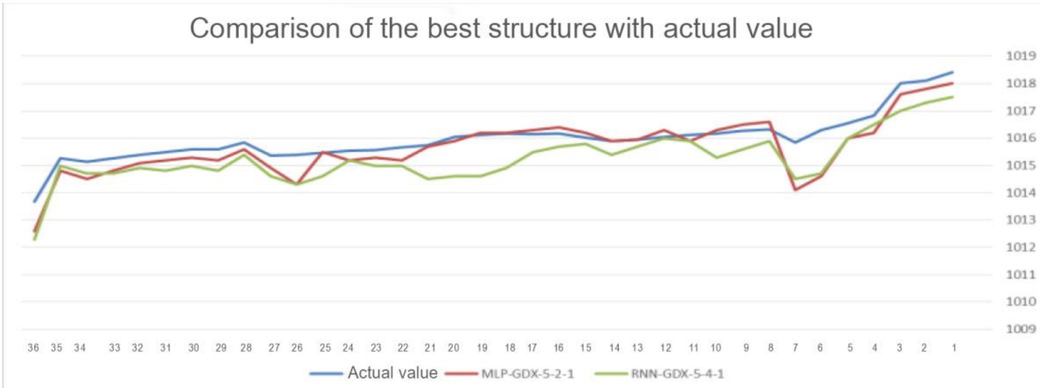


Diagram 5: Comparison of the Best Structures with Actual Values

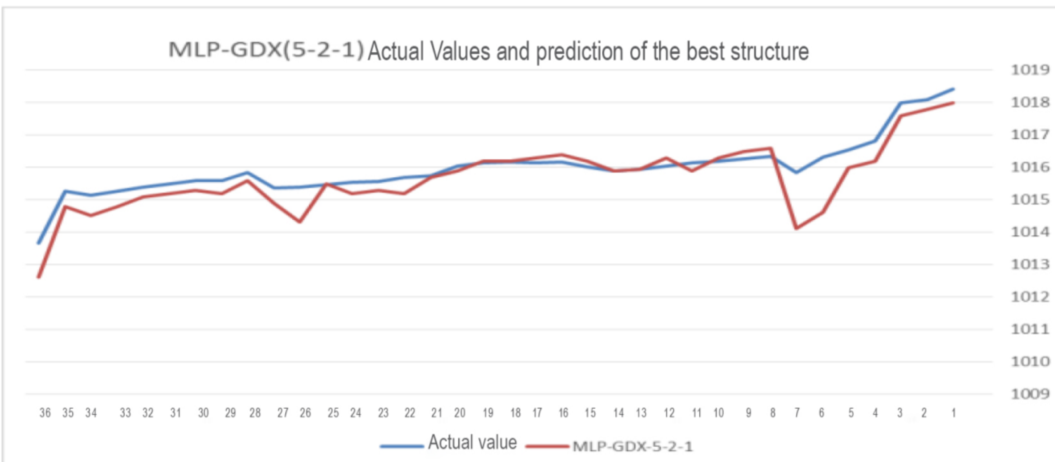


Diagram 6: Actual Values and Prediction of the Best Structure (RNN-GDX)

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 table 3.

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

Figure 3: One Month Prediction Results (April 2014) of the Groundwater Level of Lar's Plain

Station	XT	YT	Predict
A	53.55	28.21	1003.26
B	53.89	28.29	1006.84
C	53.59	28.06	1003.65
D	53.75	27.89	1010.44
E	53.96	27.95	996.85
F	53.73	27.72	997.3
G	53.68	27.55	1002.65
H	53.82	27.47	1000.3
I	54	27.55	1002.41
J	54.16	27.6	1001.35
K	54.29	27.61	995.11
L	54.33	27.45	994.32
M	54.29	27.69	992.65
N	54.6	27.54	990.31
O	54.37	27.75	998.14
P	54.96	27.47	996.02
Q	54.91	28.06	995.88
R	55.39	27.93	997.41
S	55.26	27.59	991.2
T	54.89	27.68	995.36

The Inverse Distance Weighting Geostatistical Model

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

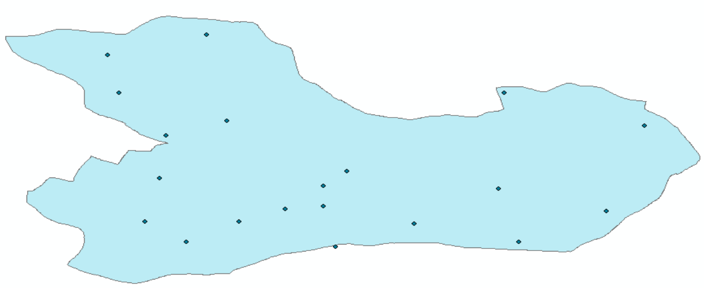


Figure 4: Map of Lar's Plain together with the Studied Wells

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 5 shows the prediction diagram of the inverse distance weighting model.

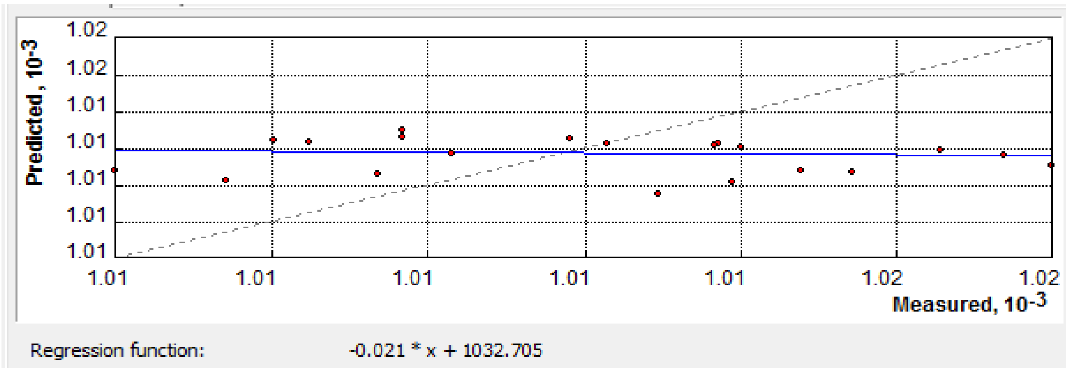


Figure 5: Prediction Diagram of the Inverse Distance Weighting Model

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 6.

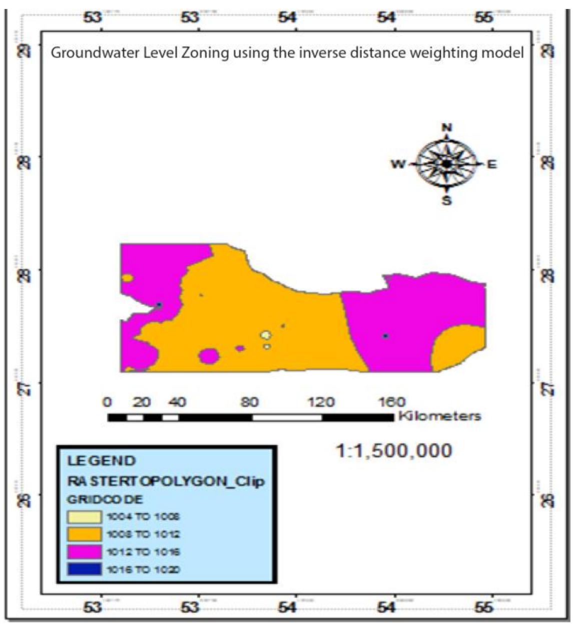


Figure 6: Groundwater Level Zoning using the Inverse Distance Weighting Model

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 Circular model was selected based on the obtained solutions.

Discription	model	percent
Nugget/sill	Circlar	0
	spherical	0.015
	Exponential	0.0135
	Gaussian	0.201

Figure 7: Nugget/Sill Ratio

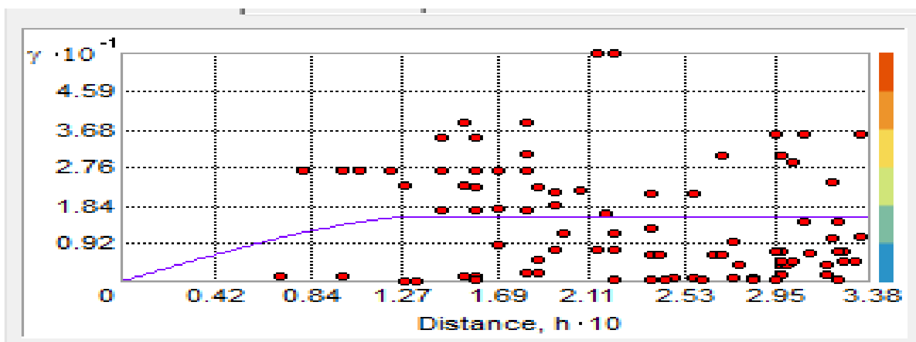


Diagram 7: Diagram of Gamma Ratio to the Distance between Points by the Circular Model

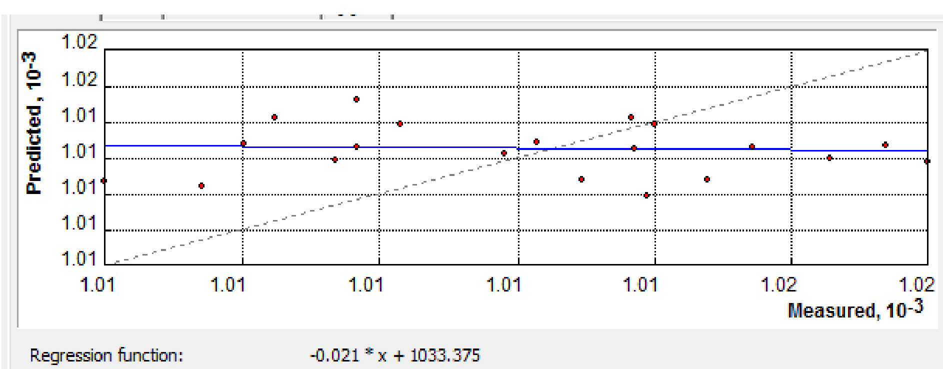


Diagram 8: Prediction Diagram by the Circular Model

Selection of the most suitable model for Kriging				
	Gaussian	Exponential	Spherical	Circular
Root Mean Square	3.981	4.02	3.98	0.22
Avrage Standard Error	4.27	4.28	4.96	3.01
Mean Standardized	0.09	0.1	0.098	0.001
Root Mean Square Standardized	0.92	0.945	0.998	0.75

Figure 8: Selection of the most Suitable Model for Kriging

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

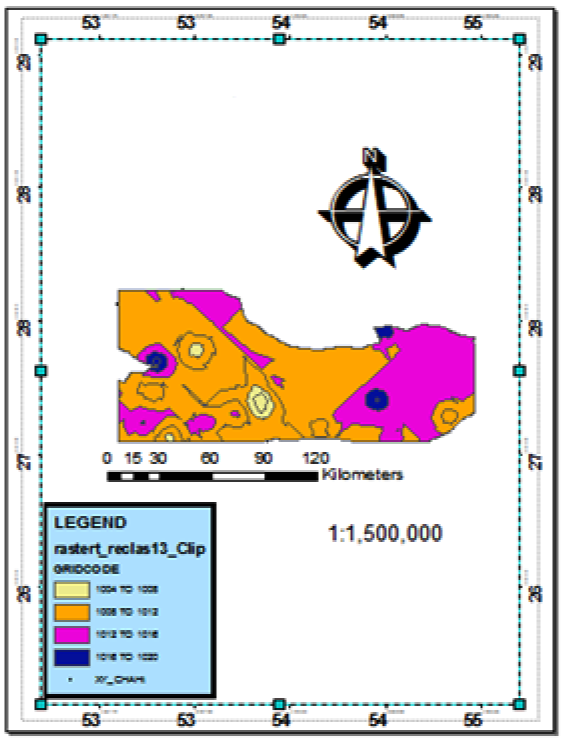


Figure 9: Zoning Map of Groundwater Level of Lar's Plain Using the Geostatistical Model of Kriging

Conclusion

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

1- In working with monthly data in the study area of Lar's plain, the monthly neural network displayed a successful performance. It could be said that the period of a month to simulate groundwater level by the neural network is a time range that is relatively large. The neural network can be well learned information in such intervals and with such good information can be provided relatively good results.

2- 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.

3- 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.

4- 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.

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

Suggestion

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- Monthly statistics from surface water and groundwater in order to prepare water balance and review the changes and make more accurate models to analyze instability situation more efficient management of water resources also, studies and seismic investigations to get information from the layered aquifer, the depth of difficult terrain and how to expand aquifers, drill several new piezometers wells with sufficient depth in different parts of Lar’s plain specially around the rivers in order to determine the hydraulic behavior of rivers and aquifers which can be can be used in future studies to increase the quality of similar studies.

2- 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

3- 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.

REFERENCES

Akbari M, Jorgeh M and Madani Sadat H (2009). Study of groundwater drawdown using geographic information system (GIS), Case Study: Mashhad plain. *Research Journal of Soil and Water Conservation* **16**(4).

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 using artificial 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 S (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*.

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

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.

Salajegheh S, Khorasani N, Mirseyed Hosseini H and Zahedi Amiri GH (2011). The spatial distribution of groundwater quality using geostatistical methods, case study in Karaj. *Journal of Watershed Management* **100**

Shamsnia S, Honar M, Shahidi N and Shirfath N (2010). Spatial analysis and mapping groundwater quality using kriging models in GIS. *The First International Conference on Modeling, Plant, Water, Soil, Air*.

Toth EA, 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.