Comparison Of Geostatic Method Based On GIS For Determining The Best Interpolation Method For Zoning The Aquifer Transmissivity In Shabestar Plain

[1]Gholamreza Ahmadzadeh
Department of geology, faculty of basic science,
university of Mohaghegh Ardabili,
Ardabil, Iran
Email:gholamrezaahmadzadeh@yahoo.com

[2]Faeze Jafari
Department of watershed Management, Faculty of Agricultural and Natural Resources
university of Mohaghegh Ardabili,
Ardabil, Iran
Email: Faeze_jafari86@yahoo.com

***Abstract*—Aquifer transmissivity is one of the most important factors in determining suitable areas for floodwater spreading operation and it is significant for creating a suitable map for this factor in order to create the final map of suitable areas. As these data are in the form of points, for creating the intended map, these data should be generalized to the whole area. In this regard, for creating the transmissivity map, the geostatic methods are used. In this article, the inverse distance weighing, local interpolation and Kriging methods were adapted. Results indicated that compared to the inverse distance weighing and local interpolation methods, the Kriging method were of lower errors. Comparing different variograms, it was indicated that the Kriging method had the best data fitness variogram and exponential variogram which with an RMSE equal to 0.3722 had the lowest error level compared to other variograms. Moreover, the ratio of the nugget effect and sill for exponential variogram were equal to 0.325 which shows a moderate spatial structure. However, in this relationship, other variograms were of weak spatial structures. Therefore, the Kriging method with exponential variogram were chosen as the best method of interpolation and map creation.**

Keywords—Transmissivity, Kriging, Variogram, RMSE

I) Introduction

One of the important factors in determining suitable areas for floodwater spread is the aquifer transmissivity. Transmissivity is one of the hydrodynamic coefficients which indicates the fitness of an aquifer in water transition and shows the water movement in porous environment [1]. For providing the transmissivity map, a modeling method is required. Suitable choice of modeling (interpolation) function is significant as the given fitness model encloses the sampling data to reality [2]. In this regard, for creating the transmissivity map, the geostatic methods are utilized. In the recent decades, the science of Geostatic is developed well and the capabilities of this branch of statistics are identified in studying and predicting the spatial variables [3]. In geostatic science, by having the amount of a quantity in defined coordinates, the amount of that quantity can be estimated in another location, if the unidentified amount coordinates is located in the dominant spatial structure [4,5]. In a research, [6] attempted to measure Kriging, Cokriging, radial basis functions and the inverse distance weighing geostatic methods in estimating the spatial spread of snow depth. Results showed that the most suitable variogram model is the exponential type, as this variogram has a nugget effect of 0.05 and a sill of 0.533, and the variogram analysis and the RMSE validation index of the mentioned methods showed that among these methods, the Kriging method is the most suitable technique for zoning. Moreover, results indicated that among the usual Kriging methods, the universal Kriging method is of the lowest error rate (11.49) and has the most accurate estimation of snow depth in the area. In an article, [7] attempted to investigate and compare different interpolation models used for prioritizing the suitable areas for floodwater spreading. The studied and compared interpolation methods are Weighted Moving Average (WMA) (with the power of three), Thin Plate Smooth Spline (TPSS) (with the power of two) and Kriging (Ordinary log). In this regard, the methods above were evaluated and compared for providing the surface permeability and alluvium thickness of Samal watershed areas in the province of Boushehr.

Results gained from this research indicate that among the evaluated methods, the weighted moving average with the power of two with an MAE equal to 1.5 is the best interpolation method for creating the surface permeability map, and the Kriging log method with an MAE equal to 8.8 is the best interpolation method for creating the alluvium thickness map. This study aims to find the best interpolation method for creating the aquifer transmissivity map as one of the most important factors in determining the suitable location for operating floodwater spread in Shabestar plain.

II) Material and method:

The study area is Shabestar plain with a surface of 1290.134 square kilometers with longitudes of “42”, 26’, °45 to 22", 05, 46 ° in east and latitudes of 04', 38° to 13", 23', 38° in north in kilometer 35 of Eastern Azerbaijan province. The study area is limited in north to the Mishodagh[1] heights, in east to the Mouro Mountain, in west to the Tesouj plain and in south to the Coast of Oromia Lake. Since studying a vast surface of an area is usually difficult and costly, usually for such surfaces, several samples are selected in forms of points and then through using particular functions, their data are generalized to the whole studied area. For expanding and generalizing the point information from the sampled points and the location changes of each variable, it is necessary to use models that are capable to simulate the studying variable in unknown points [8].

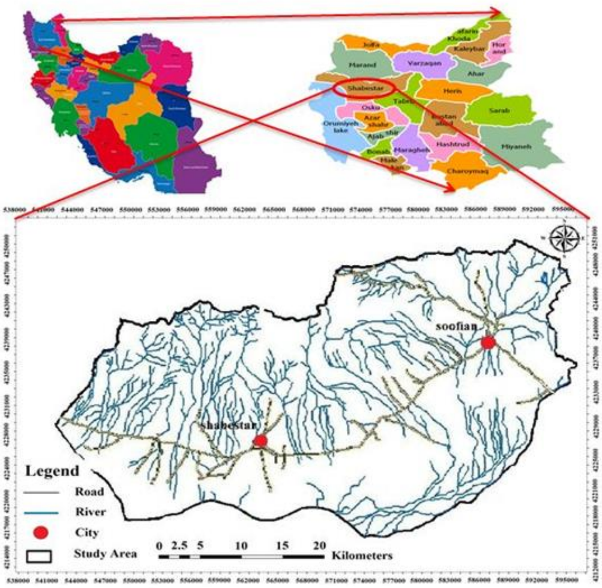

Fig. 1 – location of study area

a) Geostatic methods:

In geostatic methods, estimation is conducted according to the existing spatial structure of the intended environment [9]. As geostatic methods consider the correlation and spatial structure of data, they are highly significant. In geostatic, in addition to a determined amount of a quantity in a sample, the sample location is also considered. Thus, the samples’ location can be analyzed with the amount of the intended quantity [10]. In this article, the transmissivity data of piezo-metric[2] wells and utilization of the area were determined through well pump test. For converting the point data to surface map and creating the plain transmissivity map, the inverse distance weighing, local interpolation and Kriging geostatic methods were adapted. The features of these methods are discussed in details in the next parts.

b) The inverse distance weighing method:

In the inverse distance weighing method, the basic assumption is that the amount of correlation and similarities between neighbors suit their distance to each other, which can be defined as a function with an inverse power from the distance of each point from neighboring ones. The definition of neighboring radius and the power related to the inverse distance are considered as two of the important features of this method [11].

In the inverse distance weighing method, neighbors are located in equations in relation to their distance from the unknown point in a weighted form, and in fact, a kind of weighted average measure is conducted according to the following equation [12]:

$$z(x,y) = \frac{\sum_{i=1}^{n} \frac{z_i}{d_i^p}}{\sum_{i=1}^{n} \frac{1}{d_i^p}}$$

Where p is a power which controls the weight reduction rate with distance and is usually considered as 2. The more p is considered, the higher the effect of closer points will be. d_i is the distance of unknown points to the i^{th} point. In this study, this method is used with the powers of 2, 3 and 4.

c) The local interpolation method:

The multivariate interpolation method is divided into two classes of generic and local multivariate interpolation. The generic multivariate interpolation

Journal of Multidisciplinary Engineering Science Studies (JMESS)
ISSN: 2458-925X
Vol. 2 Issue 10, October - 2016

constantly fits a multivariate equation on a surface. But in the local multivariate interpolation method fits several multivariate equation on a continuous surface. In other words, the changes of the intended phenomenon is not created with a single equation, therefore, the estimation is conducted based on samples on unknown points [13].

d) Kriging method:

The Kriging method is a geostatic method for data interpolation based on spatial variance. Similar to the inverse distance weighing in which closeness to sample points are considered as the estimation weight, in Kriging as well the spatial variance is considered as a function of distance [14]. The Kriging method is an estimation method based on weighed moving average, and this estimator is known as the best linear unbiased estimator [15]. One of the most important and most obvious features of Kriging is that for each estimation, the related error can also be measured (Khosravi and Jabar, 2011). The Kriging general equation is as follows:

$$Z^*(x_i) = \sum_{i=1}^{n} \lambda_i Z(x_i)$$

Where $Z^*(x_i)$ is the estimated amount of variable in the location of x_i, λ_i is the weight of the i^{th} sample, $Z(x_i)$ is the observed amount of the i^{th} variable, and n is the number of observations [16]. Kriging has various types which in this article, the ordinary kriging method were adapted.

e) The ordinary kriging method:

The ordinary kriging method is one of the geostatic estimators which is used for measuring the depth of local scales [17]. In this method, variogram is used for expressing the spatial changes and minimizing the predicted error amounts considering the spatial distribution of the predicted data [18].

1) Variogram

The next step of using geostatic methods is drawing variogram. The main purpose of calculating the variogram is investigation of variable changeability in relation to time or place distance. For this, it is necessary that the total differential square of pair points that are located in the determined h distance from each other is calculated and be drawn in front of h [4]. The variogram function is as follows:

$$\gamma(h) = \frac{1}{2N(h)} + \sum_{i=1}^{N(h)} (Z(x_i) - Z(x_i + h))^2$$

Where $\gamma(h)$ is the variogram amount, N is the pair samples, h is distance, x_i is variable and $Z(x_i)$ is the variable amount [6]. A variogram has characteristics including Sill, Range effect, and Nugget effect [19]. The sill amount is the maximum amount of variogram which is the same as the intended location variance variable. He minimum variogram point is the nugget effect which expresses the calculation error variance, and range effect expresses the distance in which variogram has the highest amount. The ratio of nugget effect to sill can be investigated for assessing the location structure of studied data. When this ratio is lower than 0.25, the intended variable has a robust location structure, when it is between 0.25 and 0.75, the location structure is moderate and when it is higher than 0.75, the location structure is weak [20]. In this research, different variogram models such as circular, spherical, exponential and Gaussian were adapted, and their equations are as bellow [2]:

Circular variogram:

$$\delta(h) = C_0 + C_1 \left[1 - \frac{2}{\pi}\cos^{-1}\left(\frac{h}{a}\right) + \sqrt{1 - \frac{h^2}{a^2}}\right]$$

Spherical variogram:

$$\delta(h) = C_0 + C_1 \left[\frac{3}{2} \times \frac{h}{a} - \frac{1}{2} \times \left(\frac{h}{a}\right)^3\right]$$

Exponential variogram:

$$\delta(h) = C_0 + C_1 \left[1 - \exp\left(-\frac{h}{a}\right)\right]$$

Gaussian variogram:

$$\delta(h) = C_0 + C_1 \left[1 - \exp\left(-\frac{h}{a}\right)^2\right]$$

Where C is the sill threshold, a is the effect domain, and h is the distance between samples.

2) Validation:

In this step, the results of different interpolation methods are compared with each other. For determining the validity of results and comparing and analyzing the methods, the error amount of calculated errors were measured through the root mean square error (RMSE)[3] which is as bellow:

$$RMSE = \frac{\sum (X - X_i)^2}{n}$$

Where X is the observed amount X_i is the measured amount and n is the number of data [2]. The lower the RMSE validity statistic, the more efficient the obtained variogram acts in showing the intended parameter location changes [6]. The 0 amount of RMSE index also shows the lack of error in model measurement [21].

III) Research findings:

Data normality is the condition of using the geostatic method [22]. By investigating the two data histogram and QQplot methods, it was indicated that data distribution is not normal. In a normal distribution, the average number amount should be lower than the mean [23]. When data distribution is not normal, for enclosing the data distribution to normal distribution, the normal log and Box-Cox are used. This is why data are transported into the data logarithm, so that in this way, data distribution becomes normal. Figures 2 to 5 show the transmissivity data histograms and data QQplot diagram before and after normalization. Tables 1 and 2 show the statistical indexes before and after normalization. Studying the tables, it is indicated that mean equals 217.5 and average amount equals 268.84, and after logarithm implementation and normalization, the amount of data average became less than mean (mean equals 2.332 and average equals 2.3023). In the next step, various variogram models such as circular, spherical, exponential and Gaussian models were used.

Table 3 shows the Kriging method variogram features. Comparing the various variograms, the exponential

semivariogram with an RMSE equal to 0.3722 has the lowest error amount compared to other semivariograms. Moreover, the ratio of nugget effect to exponential semivariogram equals 0.325 which indicates a moderate location structure. However, in this relationship, other semivariograms have weak correlations. Furthermore, the Root-Mean-Square Standardized in the exponential smivariogram is closer to one compared to other semivariograms and is also of smaller Average standard error and Mean standardized.

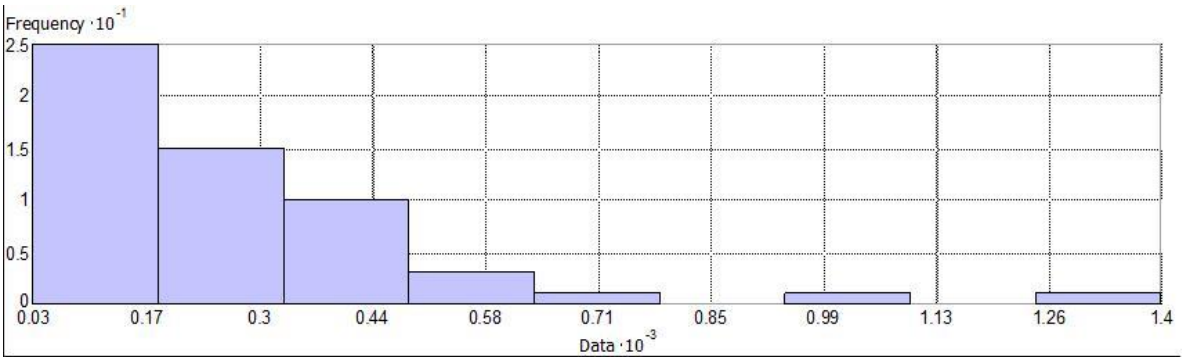

Fig. 2- transmissivity data histogram before normalization.

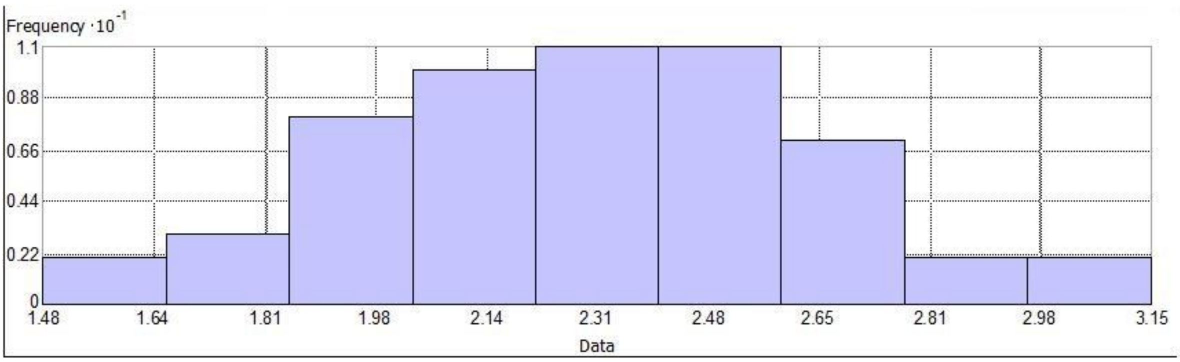

Fig. 3- transmissivity data histogram after normalization and logarithm implementation

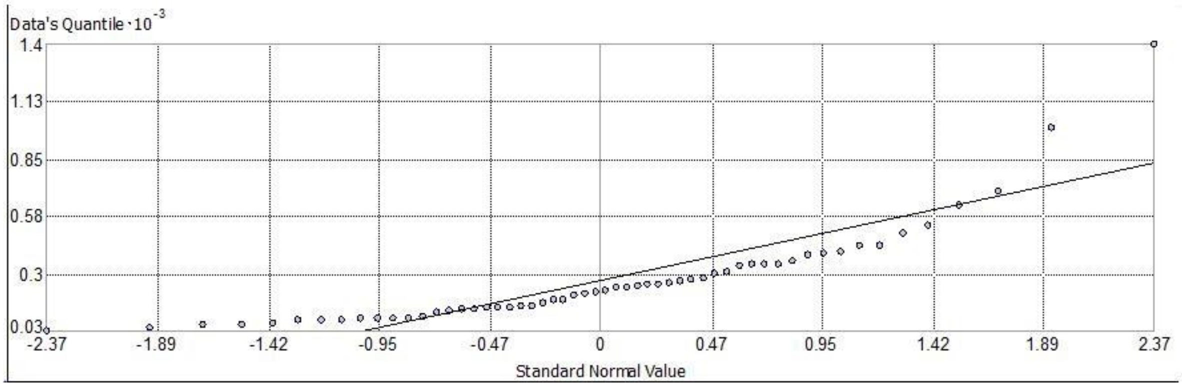

Fig. 4- The QQplot diagram of transmissivity data before normalization.

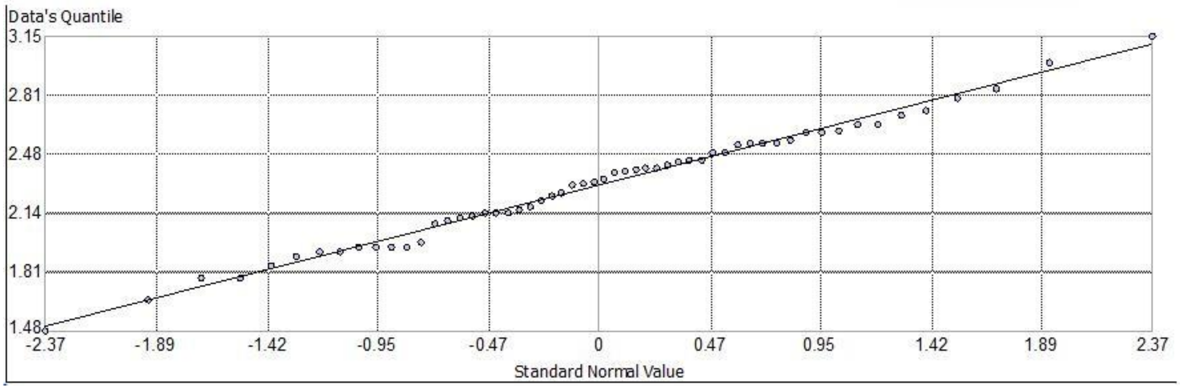

Fig. 5- The QQplot diagram of transmissivity diagram after normalization and logarithm implementation.

Table 1- Data statistical indexes before normalization

min	30
max	1400
mean	268.84
median	217.5
Standard deviation	237.22
skewness	2.5992
kurtosis	11.735

Table 2- Data statistical indexes after normalization

min	1.477
max	3.146
mean	2.3023

median	2.332
Standard deviation	0.33597
skewness	0.013058
kurtosis	2.9031

Table 3- Variogram features of Kriging method for the transmissivity factor

Root-Mean-Square Standardized	Average standard error	Mean standardized	$\frac{nugget}{sill}$	RMSE	Sill	nugget	variogram
0.913	0.417	0.031	1.892	0.3767	0.056	0.106	circular
0.919	0.4168	0.0316	1.745	0.3790	0.059	0.103	spherical
0.927	0.414	0.025	0.325	0.3722	0.123	0.040	exponential
0.911	0.416	0.028	2.48	0.3768	0.045	0.112	Gaussian

Table 4 – The obtained SMSE index amount or various powers of inverse distance weighing method

RMSE	Power
0.3974	2
0.4235	3
0.4424	4

Table 5 - The obtained SMSE index amount for various functions and powers of the local interpolation method

RMSE	power	Kernel function
0.3792	2	Exponential
0.4137	3	
0.3761	2	Gaussian
0.4126	3	

Tables 4 and 5 indicate the calculated RMSE index amount for the inverse distance weighing method in powers of 2, 3, and 4, and the local interpolation method with exponential and Gaussian functions with powers of 2 and 3. It can be explained that in the inverse distance weighing, until the error amounts are moving in a decreasing way, a higher power were used and tested, so that the minimum error can be achieved. And when error amounts becomes increasing, the higher powers are no longer used, which on this basis, the power of 2 here has the lowest error amount (RMSE equals to 0.39740). It is observed that by increasing power, the error also increases. Thus, the higher powers should not be used. In case of the local interpolation method, the Gaussian function with an RMSE equal to 0.3761 has a lower error level.

IV) Conclusion:

In this research, for creating the transmissivity map, the geostatic methods were adapted. The utilized methods include inverse distance weighing, local interpolation and Kriging, and finally, according to the results gained from the calculations and comparison of all various methods used in this research, the Kriging method with the exponential semiviogram[4] had the lowest error (RMSE equal to 0.3722) and the ratio of nugget effect to sill which equals to 0.325 is the most suitable method of creating the transmissivity and zoning map.

References:

[1] B. Ghermez Cheshmeh, J. Ghayoumian, S. Feyznia, J. Porhemmad, Geomorphology usage in floodwater spreading positioning in areas without statistics, Case study: Meymeh plain of Isfahan, National conference of fields management–Soil erosion and continuous development, Arak. 2001.

[2] A. Shabani, M. Shahbakhti, R. Abbaspour, Atmosphere polluters distribution modeling in Tehran Metropolis using statistical and geostatic methods, Geomatics national conference. 2011.

[3] F. Misaghi, K. Mohammadi, Rainfall information zoning using classic statistical and geostatic methods and comparison of artificial neural networks, Agriculture scientific magazine, issue 29, 4th year, 2006, p. 1-14.

[4] A.A. Hassani Pak, Geostatic, Tehran University Publications, 1998, p. 313

[5] H. Asakereh, Usage of the Kriging model in rainfall interpolation, Geography and development magazine, 2008, issue 12, p. 25-42.

[6] S. Ebdam, A. Fathzadeh, assessment of geostatic methods for measuring the location distribution of snow depth in semi-dry areas (Case study: Sakhvid aquifer), water and irrigation research-scientific periodical, fourth year, 2013, issue 13, p. 113-124.

[7] M.J. Soltani, S. Shadfar, P. Baktash, Comparison of different interpolation methods in selecting suitable locations for executing the floodwater spreading operation using geographical information systems, 2010.

[8] S. Maroufi, A. Toranjian, H. Zare Abyaneh, Assessment of Geostatic methods for estimation of Electric guidance and PH current drainage water of Bahar plain – Hamedan, Water and soil preservation researches magazine, 2010, Vol. 16, Issue 2, p. 169-187.

[9] M.R. Haji Hashemi Jazi, M. Atashgahi, A.H. Hamidian, creating local transmittal maps of aquifer waters pollution, a tool for managing these water resources, the seventh national conference of watershed engineering sciences, Isfahan industrial university. 2011.

[10] M. Mahdavi, I. Husseini Chegini, M.H. Mahdian, S. Rahimi Bandarabadi, Comparison of geostatic methods in measuring the annual rainfall local distribution in dry and semi-dry areas of Iran's South-western areas, Iran's natural resources magazine, 2004, Issue 57, issue 2.

[11] Y. Khosravi, M.K. Jabbari, Geographic information systems (GIS) principles and ArcGis 10 self-trainer, Azar Kalak Publications, 2011, P. 366.

[12] N. El-sheimy, Digital Terrain Modeling.Department of Geomatics Engineering. The University of Calgary. 1999.

[13] M. Ghahroudi Tali, O.S. Babai Fini, H. Atai, An introduction to Geographic Information

Systems, Payam-e Nour Publications, 2015, p. 234.

[14] S.H. Anderson, An Evaluation of Spatial Interpolation methods Pn Air Temperature in Phoenix, AZ. 2000.

[15] M.R. Yazdani, Chavoshi, Sattar; Khodagholi, Morteza; Saghafian, Bahram, (2005), Investigation of Meteorological draughts in Isfahan province, Water and watershed magazine, issue 4, p. 41-51.

[16] M. Ayouzi, A. Mosaedi, M. Meftah Halaghi, M. Hesam, Comparison of different interpolation methods in Golestan province draught zoning, the second national conference of draught effects and its management solutions, Isfahan. 2009.

[17] S.S. Carrol, N. Cressie, A comparison of geostatistical methodologies used to estimate snow water equivalent, Water Resources Bull, 1996, 32:267-278.

[18] H.A. Faraji Sabokbar, G. Azizi, assessment of the accuracy of spatial interpolation methods (Case study: rainfall modeling of Mashhad's Kardeh area), Geographical researches magazine, 2006, issue 58, p. 1-15.

[19] M. Vafakhah, M. Mohseni Saravi, M. Mahdavi, S.K. Alavipanah, Usage of geostatic in measuring the snow depth and density in Orazan aquifer, Iran's watershed engineering and scientific magazine, Second period, 2008, issue 4, p. 49-55.

[20] J. Mohammadi, A review on geostatic principles and its usage in geology, Water and soil science magazine, 15th period, 2001, Issue 1, P. 99-121.

[21] Y. Maghami, R. Ghazavi, A.A. Vali, S. Sharafi, Assessment of different interpolation methods for water quality zoning using GIS (Case study: Abadeh county), Geography and environmental planning magazine, 22nd year, 2011, issue 42, p. 171-182.

[22] M. Moradi, H. Vagharfard, A. Khorani, V. Mahmoudinejad, Assessment of different interpolation methods in aquifer salinity zoning through the Cross-validation technique (Case study: Shahre Kurd plain), Remote assessment and GIS magazine of Iran, Third year, 2011, Issue 1, P. 35-44.

[23] S. Hamad, Geostatistical analysis of groundwater Levels in the south Al Jabal Al Akhdar area using GIS. GIS Ostrava, 2009, 25:1-10.