Geostatic and its Applications in the Creating Iso-Maps in Underground Water Studies in a GIS Framework
A Case Study of Razan-Ghahavand Plain in Hamedan Province, Iran

[1]Amir Mostafa Ahmadi ,[2]Mohammad Ahmadi,

1. Modares Environmental Research Center Co-Worker, Tehran, Iran
2. Dam Department of Bandab Consulting Engineers, Tehran, Iran

Abstract

With taking discrete sampling from the water resource parameters in quantity and quality and the use of procedures of turning discrete spots to connected area one can study the process of the surface changes of selected area. There are different procedures for turning broken discrete spots to continuity of areas such as procedure of statistic ground which conclude Kriging procedure, Inverse Distance Weighting (IDW), Radial Basis Functions (RBF), and Local Polynomial Interpolation, Global Polynomial Interpolation, Co-kriging. In this research with the use of Kriging, IDW and RBF procedures the curves of equal level of underground water is drawn by GIS. And with the application of statistic signs such as Mean Absolute Error (MAE), Mean Absolute Relative Error (MARE), Root Mean Square Error (RMSE), MBE, Coefficient of Correlation and also the graph of the average of the percent dispersion of absolute value of relative error, the researcher has compared these procedures and drawn the direction of the flow of underground water which is the major relevant issue to Hydro geologic studies in Dasht Razan-Ghahavand.

Key words**:** *Statistic ground, the level of underground water, statistic signs, GIS.*

1-Introduction

Examining the existent situation and observing the quantities of target value in area limit of study can be the first step in the studies of surface resource water and under ground flow. In this regard extracting and recording data related to evident resource periodically and regularly and the formation of the regular data base are great help with analyzing and evaluating existent situation and comparing between statistical periods. So with having access to recorded data base from the set of bench marks and the use of Geostatistics procedures one can draw variation curves of target value in the limit of study area. It is remarkable to note that whatever contours curves become attentive the study of further stages also grow more attentive.

Geostatistics experts regarding the needs, provide different procedures for drawing the area changes and contours curves which each procedure has its own strengths and weaknesses. In this study under ground flow direction which is an important factor in relation to hydraulic connection of aquifers will be drawn with the use of Geostatistics procedure in Razan-Ghahvand plain. The information of the surface level of the underground flow in stated plain in water year of 2001-2002 is used as raw input. The stages of this study will be explained completely further.

2-The situation of study area

Study area of Razan-Ghahavand is located in Hamedan province with the area of 3083 square kilometer. Razan-Ghahavand plain has covered 72 percent of the study area which is about 2205 squire kilometer. This study area is located between longitudes of 48 45 till 49 26 east and latitudes of 34 42 till 35 37 north. Map no.1 shows the situation of the study limit along with the limit of Razan-Ghahavand plain.

3-Methodology

Stages of this research concludes of collecting data, providing geo database in GIS environment, using of Geostatistics procedures, comparing of the results of each noted procedure and selecting better model. For model selection any Statistical indices have been presented, too.

4-Introduction related to the procedures of Geostatistics

In general Geostatistics procedures are based on Regionalized Variable theory. Regionalized Variable refers to every environmental feature distributed in two or three dimensional space. The changes of this set of variables from one point to another are clear and their continuity is obvious. The features such as the level of the surface of underground water, texture of soil and/or the amount of different elements in soil are examples of the regionalized variables. The major difference between classic statistics and Geostatistics is that it is assumed that the samples collected form society are not depend on each other in classic statistics ,therefore the existence of one sample does not give any information about the other samples located in certain distance. So regionalized variable theory is based on the different environmental features which has regionalized dependency in the way that the amount of one environmental variable in close distance are similar and with distance increase this similarity decreases. Regionalized dependency between samples can be shown through a mathematical model entitled variogram.

4-1-Kriging procedure

This procedure based on models and statistical procedures is auto-correlation. It is an estimator based on the logic of weighted moving average and it is an unbiased estimator and it is determined by the use of Krige's formula. The equation no.1 shows how it is estimated in Kriging procedure.

$$\hat{z}(x) = \sum_{i=1}^{n} \lambda_i . z(x_i) \qquad (1)$$

In which z(x) is the estimated parameter and ..is the weight or the significance of the quantity that depended on i^{th} sample and $z(x_i)$ is known parameter.

4-2-The procedure of Inverse Distance Weighing or IDW

In this procedure the amount of one quantity in spots with known coordinate can be attained by the use of quantity of the same amount in other spots with known coordinate. In other words in this procedure the value of one variable is counted based on the mean of neighbors in specific zones. In this way the inverse of distance from one known point is the base of weighing and whatever the distance of unknown point decreases from known spot the weight of that point increases and the points with unknown values with the use of around points in a specific radius can be measured based on the equation no. (2).

$$\hat{z}(x_i) = \frac{1}{n} \sum_{i=1}^{n} z(x_i) \qquad (2)$$

In that $\hat{z}(x)$ is the estimated parameter and $z(x_i)$ is known parameter.

4-3-The procedure of Radial Basis Functions or RBF

Radial Basis functions are a procedure contains 5 kinds of radial functions as explained through following words. There is not a big difference among the results of different functions in RBF procedure and the selection of the radial basis function happens by validating the estimate results. It is worth to note that there is a parameter which put direct influence on the estimate results in each function of the summation of radial basis functions and the best quality of that parameter is the correspond amount of the minimum of Root Mean Squared Prediction Error (RMSPE). In this procedure, with taking distance from known points, the estimated amount, with considering the selected radial function will change. The equations no.3 to no.7 introduces these functions in order: TPS[1], MQ, IMQ, CRS, and SWT. The CRS function is used in this research. The procedures and the functions of RBF are the especial form of Artificial Neural Networks (ANN).

$$\hat{z}(r) = (\sigma . r)^2 . \ln(\sigma . r) \qquad (3)$$

$$\hat{z}(r) = (r^2 + \sigma^2)^{0.5} \qquad (4)$$

[1]Thin-Plate Spline

$$\hat{z}(r) = (r^2 + \sigma^2)^{-0.5} \quad (5)$$

$$\hat{z}(r) = -\sum_{n=1}^{\infty} \frac{(-1)^n . r^{2n}}{n! n} = \ln(\frac{\sigma . r}{2})^2 + E_1(\frac{\sigma . r}{2})^2 + C_E \quad (6)$$

$$\hat{z}(r) = \ln(\frac{\sigma . r}{2}) + K_0 (\sigma . r)^2 + C_E \quad (7)$$

In these equations σ is Tension Parameter, E_1 is Exponential Integral Function, C_E is Euler Constant, and K_0 is Modified Bessel Function.

5-Establishing and Validating Geostatistics Model

In this stage with the use of GIS and random procedure, establishing samples, and validating samples of models are 80 and 20 percent of the whole selected data bank, and are recalled in cyber space in the form of two separate layers. After this stage, Kriging models, IDW, and RBF are established and validated. The results of models are shown in the figure no.2. One of the criteria in examining of the validity of the attained results from models is the criteria of Coefficient of correlation (r); whatever its absolute value gets closer to 1 the better adaptation between the observed amounts is shown. But with the use of just Coefficient of correlation one can not declare anything about the efficiency and the accuracy of the model; therefore other parameters and statistics will be used for examining the designed models. In this regard the criteria of Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Mean Absolute Relative Error (MARE) and Mean Bias Error (MBE) have been used. The equations number (8) and (12) shows these parameters in order. The amount of stated parameters in both stages of establishing (education) and validating (test) of Geostatistics models are shown in table no. 1.

$$r = \frac{\sum_{i=1}^{n} (p_i - \overline{p})(O_i - \overline{O})}{\sqrt{\sum_{i=1}^{n} (p_i - \overline{p})^2 (O_i - \overline{O})^2}} \quad (8)$$

$$MARE = \frac{1}{n} \sum_{i=1}^{n} \frac{\left| O_i - P_i \right|}{O_i} \quad (11)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (O_i - P_i)^2} \quad (9)$$

$$MBE = \frac{1}{n} \sum_{i=1}^{n} (O_i - P_i) \quad (12)$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left| O_i - P_i \right| \quad (10)$$

In these equations, O_i is the observed amount, P_i is the predicted amount and n is the number of observations. The above introduced features are the whole statistical indices which do not provide any information about the procedure of error distribution; therefore for evaluating the capacity of the models, statistical features are needed which specifies how the error distribution in the established models. For this reason the relative absolute error distribution diagram for final evaluation of used models (besides statistical parameter) is used.

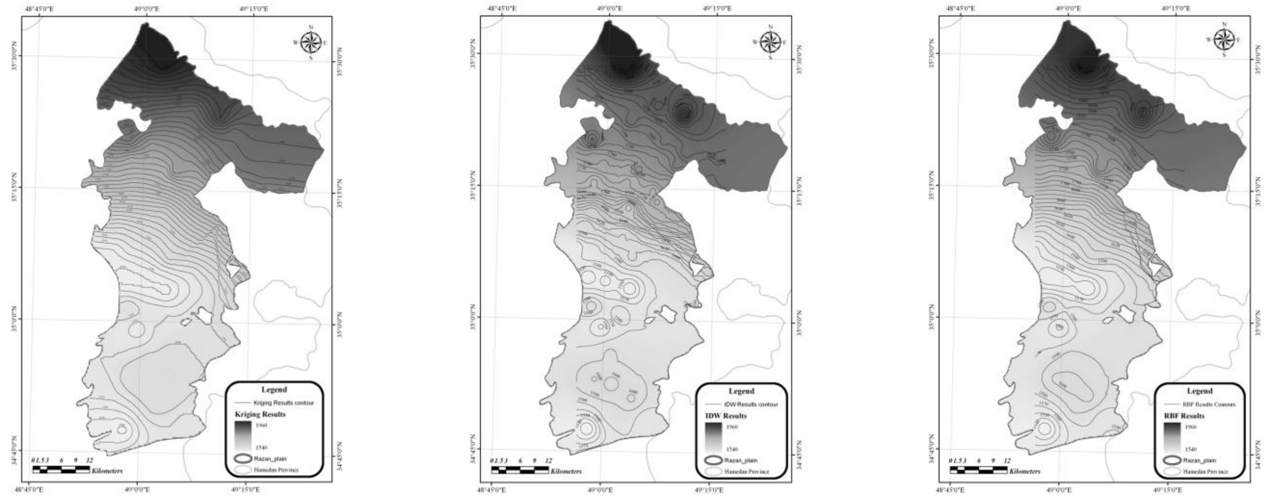


Figure no. 2: The comparison curves of the same level of the underground flow surface , the result of Kriging Geostatistics procedures, IDW, and RBF (from left to right).

Table no. 1: the comparison of different Geostatistics procedures with the use of statistical features

Model Testing	Model Training	Statistical Index	
0.9828	0.9462	R2	Kriging
12.65	17.3	MAE	
0.0072	0.01	MARE	
17.92	26.51	RMSE	
4.79	2.49	MBE	
0.9617	0.9321	R2	IDW
19.91	20.37	MAE	
0.011	0.011	MARE	
28.91	31.63	RMSE	
14.77	3.39	MBE	
0.9797	0.9437	R2	RBF
13.31	17.28	MAE	
0.0076	0.01	MARE	
20	27.51	RMSE	
8.71	2.11	MBE	

Statistical indices educating and establishing stage testing and validating stage Figures number 3 to 5 shows a comparison between the observed amounts and predicted amount in validating stage of models. As it is obvious in these figures, Kriging procedure always has provided better results. The earlier proof to this issue is table no. 1.

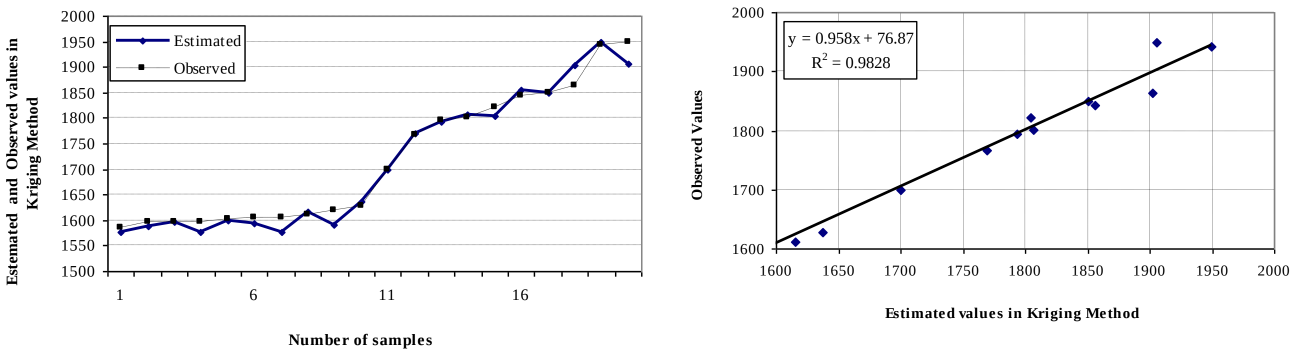


Figure no. 3: Comparison between the observed and predicted amounts in validating stage of Kriging model

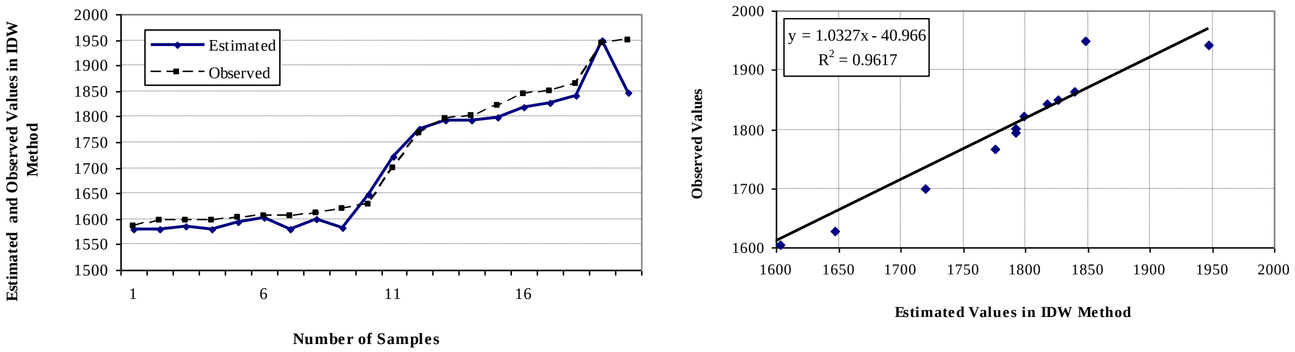


Figure no.4: comparison between the observed and predicted amounts in the evaluating stage of model IDW

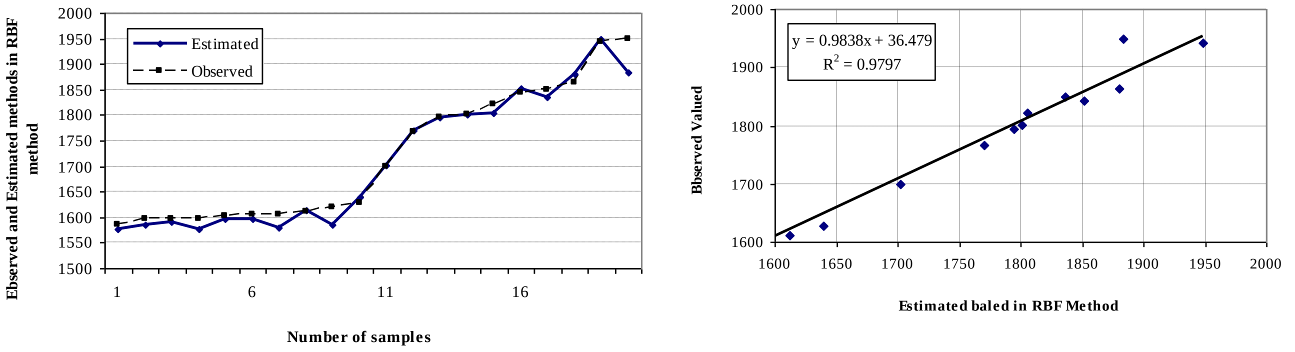
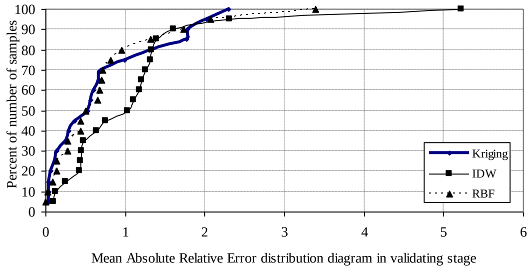


Figure no.5: Comparison between the observed and predicted amounts in the stage of validating RBF model

5-1-The selection of better model

In this research not only dispersal diagram of absolute values of relative error is used but also statistical parameters (table no.1), and the diagrams of the comparison between the observed and predicted amounts (figure number 3 to 5). The result of this analysis in figure no. 6 for stage of the validating models. It is easily will be defined that how much is of the error of each percent of the number of predicted samples of the surface level of underground water in the stage of establishing and validating in each model. For instance it is possible to declare that about 80 percent of the predicts of the establishing stage of Kriging model and RBf, have relative error less than 1.5. 60 percent of the predicts of the establishing stage of Kriging model and RBF have relative error less than 1 percent, and this error in IDW model is about 1/2 percent. 100 percent of Kriging predicts in the evaluating stage, has less relative error than 2.5 but the relative error in IDW models and RBF is about 5.2 and 3.4 percent. Considering of what is said before and the statistical analysis, Kriging model is the selected one for estimating the level of under ground water.



Shape No. 6: the scattering of amounts of relative absolute error in the stage of testing (validating) models

6-Conclusion

With the use of Geostatistics procedures and statistical analysis of results it is tried to find out the best estimator of underground water table in the points apart from the measured points in Razan-Ghahavand plain .The best estimator is kriging model respecting part no.6.

The use of MAE shows the amount of absolute error, With regard to table No.1 kriging model has the best estimate which has the least absolute estimation error.

The amount of diagonal of the index of MBE and respecting equation no.12 one can see this point that if it is negative it is overestimated and if it is positive it is underestimated of the observed amounts. This feature in all the Geostatistics models is positive which shows that all models estimate the surface level of underground flow less than whatever it is observed; this issue is for the case of safety and it increases the factor of safety in estimating the volume of extractable water from aquifer this issue leads to the optimized underground flow resource management, because always the volume of extractable from underground flow is estimated less than the real amount of that and the problem of improper extraction of underground flow will resolved to some extend.

Respecting to this point that determining place changes of some quantity and quality parameters are counted as the input data of other stages of water resource study, selecting the best estimator model and its careful estimation has a direct influence on the carefulness of further stages.

References

Ghohroudi Tali, M. (2005). Geographic Information System in Three Dimensional Environment, Three Dimensional GIS in Arc Gis Environment, Jahad Daneshgahi publication of Tarbiat Moalem University, first edition.

Sanjeri. S. (2008). GIS Training in use, Abed publications, 3rd edition, 2008.

Mohammadi J. (2007). Geostatistical Evaluation of salt changes of soil in Ramhormoz zone, 1- Kriging Method, Scientific and technical studies of Agriculture and Natural Resources Journal, Vol 2, Number 4.

Lo C.P., Yeung A.K.W. (2005). Concepts and Techniques of Geographic Information Systems, Prentice Hall of India Private Limited, New Dehli.

Ahmadi M., Partani, S., Zabihi M. (2009). GIS Application in Climate-Microzonation in hot and arid areas, Case study: Raze2 Dam, The 1st International conference on Water crisis, Zabol University, Iran.

Mohammadi J. (1999). Geostatistical Evaluation of salt changes of soil in Ramhormoz zone, 1- Kriging Method, Scientific and technical studies of Agriculture and Natural Resources Journal, Vol 3, Number 1.

Ghohroudi Tali, M. (2006). Establishing and Modification methods of creating elevated models, Case study: Golestan 2 Dam, Geographic Researches Journal, Number 57.

Noori, R. (2008). Comparison of Multi-Variable Regression and Artificial Neural Networks to Estimate Air pollution of Tehran.