

Geo-Statistics and its Application for Creating Iso Maps in Hydrogeological Studies, Technical Study: Electric Conductivity contours of Atrak Plain, Golestan Province, Iran

Mohammad Ahmadi*, Peyman Shirin Zade¹, Behrouz Yaghoubi², Mostafa Safari Komeil³

*Corresponding author: MSc. Student of civil Engineering, Bu Ali Sina University, Hamedan, Iran.

E-mail address: mohammad_ahmadi8m@yahoo .

Tel: +98-912-2974529

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 Geostatistics which conclude Kriging procedure, Inverse Distance Weighting (IDW), Radial Basis Functions (RBF), Local Polynomial Interpolation, Global Polynomial Interpolation and Co-kriging. This research is going to explain any applications of Geostatistics for creating iso-maps such as underground water table contours, iso-quality (for example Ec and pH) contours, and place changes of too many other hydro-geological parameters. In this way, statistic signs such as Mean Absolute Error (MAE), Mean Absolute Relative Error (MARE), Root Mean Square Error (RMSE), MBE and Coefficient of Correlation, could be employed to select the best Geo-Statistic model in a GIS framework. Eventually, the power in IDW method has been optimized and also best Geo-Statistic method has been introduced for predicting the Electric Conductivity (EC) of underground water in Atrak plain. Finally any conclusions have been extracted.

Keywords: *Geo-statistics; underground water quality; statistic signs; GIS.*

1-Introduction

Change procedure assessment of objective value in study area could be obtained by discrete sampling of variable. In this regard, increase or decrease evaluation of objective value and extracting the critical points in study area could be classified as one the most important technical problems for geo-science researchers. The accuracy of future studies which have been based on these sampling and also change procedure studies of target value in study area, have directly depended on the accuracy and efficiency of data gathering and the procedure of turning discrete spots to continuous surface. It means that, the exactitude of future studies has widely depended on the certitude of spot point to contour map conversion method. There are too many ways to create iso-maps from spot sampling, Geo-Statistics (G.S) could be explained as one of these methods. The Geo-Statistics method divided into some sub-methods such as Inverse Distance Weighting (IDW), Radial Basis Functions (RBF), Global Polynomial Interpolation (GPI), Local Polynomial Interpolation (LPI), Kriging and Co-Kriging methods. This prediction approach, could estimate target value on non-measured points in the study area more accurate than the classic procedures such as Triangular Irregular Network (TIN). It should be mentioned that in Geo-Statistics, there are some constant

1 River Engineering Department of Yekom Consulting Engineers, Tehran, Iran.

2 Director of Integrating and Balance of Water Resources, Hamedan Regional Water Co., Iran.

3 M.S.c Student of Hydrogeology, Shahid Beheshti University, Tehran, Iran.

coefficient that they are effective on the accuracy of estimation. In this respect, a statistical index, which has called Root Mean Squared Prediction Error (RMSE), could be defined for selecting the best Geo-Statistic estimator and finding the optimized constant coefficient. Mentioned index, which has show differences between observed target value and the predicted one in non-measured points, could help the researcher to find the best estimator and optimized constant coefficients. Best estimator should have minimum value of RMSE. In presented study, by employing the results of qualitative analysis of underground water samples of Atrak plain in Golestan province, the best G.S estimator has been selected for predicting the Electrical Conductivity of underground water in mentioned study area.

2-Methodology

Selecting the best Geo-Statistics estimator to predict Electric Conductivity¹, and also power optimization of IDW method could be considered as the main purpose of presented study. Figure No. 1 describes main steps of this research.

2-The Location of study area

Atrak plain is located in Golestan province with the area of 3257 square kilometers. This study area has a plain in its central zone. Figure No.1 shows the location of the study limit along with the limit of Atrak plain.

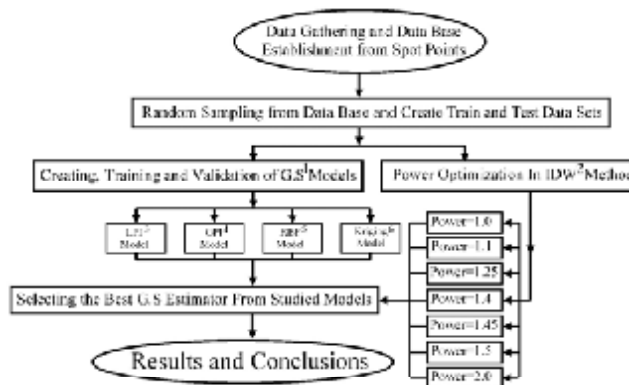


Figure 1. Research Methodology

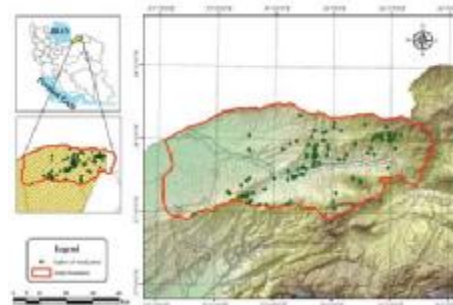


Figure 2. Location of Study Area

1-Geo-Statistics

2-Inverse Distance Weighting

3-Local Polynomial Interpolation

4-Global Polynomial Interpolation

5-Kriging model (Based on D. G. Krige theorem)

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 Electrical Conductivity, 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

1. EC

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. For example, Kriging 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. In Inverse Distance Weighting or IDW method, 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. Equation No. (2) describes the IDW prediction procedure. 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. The equations no.3 and 4 introduces some of these functions such as CRS and SWT respectively. The CRS function is used in this research. The procedures and the functions of RBF are the especial form of Artificial Neural Networks (ANN).

$$\hat{z}(x) = \sum_{i=1}^n \lambda_i . z(x_i) \quad (1) \quad \hat{z}(r) = -\sum_{n=1}^{\infty} \frac{(-1)^n . r^{2n}}{n!n} = \ln\left(\frac{\sigma . r}{2}\right)^2 + E_1\left(\frac{\sigma . r}{2}\right)^2 + C_E \quad (3)$$

$$\hat{z}(x_i) = \frac{1}{n} \sum_{i=1}^n z(x_i) \quad (2) \quad \hat{z}(r) = \ln\left(\frac{\sigma . r}{2}\right) + K_0(\sigma . r)^2 + C_E \quad (4)$$

In which $z(x)$ is the estimated parameter and λ_i is the weight or the significance of the quantity that depended on i^{th} sample and $z(x_i)$ is known parameter and $\hat{z}(x)$ is the estimated parameter and $z(x_i)$ is known parameter [1,2,4,5,6, and 7]. In these equations No. 3 and 4, σ is Tension Parameter, E_1 is Exponential Integral Function, C_E is Euler Constant, and K_0 is Modified Bessel Function [8].

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.4. 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), and Mean Bias Error (MBE) have been employed. The equations number (5) to (8) 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 [7 and 8].

$$r = \frac{\sum_{i=1}^n (p_i - \bar{p})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^n (p_i - \bar{p})^2 (O_i - \bar{O})^2}} \quad (5) \quad RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2} \quad (6) \quad MAE = \frac{1}{n} \sum_{i=1}^n |O_i - P_i| \quad (7) \quad MBE = \frac{1}{n} \sum_{i=1}^n (O_i - P_i) \quad (8)$$

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 Mean Absolute Relative Error distribution diagram for final evaluation of used models (besides statistical parameter) is used.

5-1-Optimization of Power in IDW Model and selection of best G.S¹ Model

In this section, by taking the results of Kriging, RBF, GPI, LPI and optimized IDW method, the best estimator which belongs to the minimum of RMSE² has been extracted. For this purpose, first of all optimization of power in IDW method has been carried out.

5-1-1- Power Optimization

The best and accurate estimation of objective value could be mentioned as the main purpose of G.S. In this regard, constant coefficient optimization to reach the minimum RMSE plays the main rule. This part of research wants to optimize the power in IDW method as an example for constant coefficient optimization. Comparison between results of optimized models with non-optimized models expresses the necessity of linear or nonlinear optimization in such these studies. Figure No. 3 shows the relationship between RMSE and power variation in IDW results. By having in hand the results which has presented in mentioned Figure, The optimized power in IDW method is 1.38 and the results that related to this power have the minimum RMSE.

5-1-2- Best Estimator Selection

As it noted in paragraph No. (5-1-1), this research aims to provide an optimal method for calculating the objective value in non-measured locations, change procedure studies and create Iso-Maps for presenting the change surface of variable in Raster maps in a GIS frame work. For this purpose, having in hand the estimated results with different G.S methods and incorporate statistical index Correlation Coefficient (r), RMSE, MAE and MBE, the best estimator could be extracted. In order to table No. 1, Kriging Method has the best correlation between estimations and measurement and the minimum RMSE. Also, by employing the results which has been presented in fig no. 4, kriging method has the best Mean Absolute Relative Error (MARE) (the MARE of more than %89 of datasets in test stage is less than %38), so the Kriging Geo-Statistic method could be introduced as the best EC estimator in Atrak plain.

Table No. 1: Comparison of Statistical Indices for Best Estimator Selection

| G.S Method | r | RMSE | MBE | MAE |
|---------------------|-------|------|------|------|
| Kriging | 0.81 | 2200 | 485 | 1740 |
| RBF | 0.79 | 2263 | 427 | 1662 |
| GPI | 0.63 | 2981 | 347 | 2031 |
| LPI | 0.64 | 2942 | 349 | 2002 |
| Power Optimized IDW | 0.459 | 4816 | -183 | 2845 |

1 Geo-Statistics

2 Root Mean Squared Prediction Error

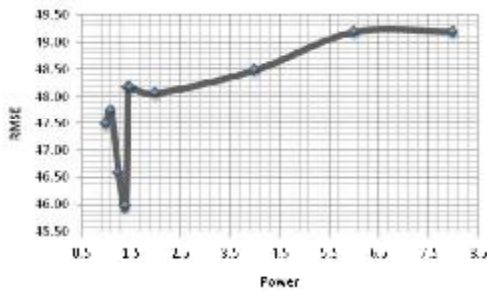


Figure No. 3: Power optimization in IDW method and the minimum of RMSE

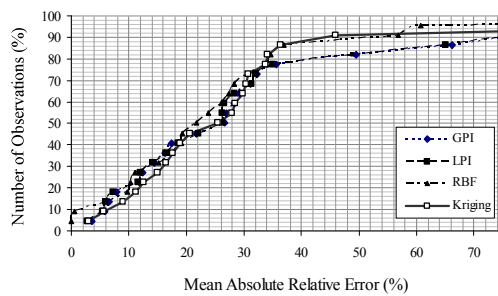


Figure No.4: Evaluation of Mean Absolute Relative Error change in Number of Data sets for evaluated Methods

6-Results and Conclusions

- With the use of Geostatistics procedures and statistical analysis of results it is tried to find out the best estimator of Electric Conductivity in the points apart from the measured points in Atrak plain .The best estimator is kriging model respecting part No.5-1-2.
- 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.

7-Suggestions

- Optimization of constant coefficients such as tension parameter in RBF procedure, in Geostatistics models and the comparison of different procedures happen in the presence of optimized factor.
- The comparison of the results of Geostatistics procedures and the procedure of Triangular Irregular Network (TIN) in providing the map of the iso-level of EC.

References

- [1] Ghohroudi Tali, Manije, "*Geographic Information System in Three Dimensional Environment, Three Dimensional GIS in Arc Gis Environment*", Jahad Daneshgahi publication of Tarbiat Moalem University, first edition, spring 2005.
- [2] Sanjeri. Sara, "*GIS Training in use*", Abed publications, 3rd edition, 2008.
- [3] Mohammadi Jahangard, "*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, Winter 2007.
- [4] C.P. Lo, Albert K. W. Yeung, "*Concepts and Techniques of Geographic Information Systems*", Prentice Hall of India Private Limited, New Dehli, 2005.
- [5] Ahmadi M., Partani, S., Zabihi M., "*GIS Application in Climate-Microzonation in hot and arid areas, Case study: Raze2 Dam*", Proceeding of the 1st International conference on Water crisis, Zabol University, winter 2009.

- [6] Mohammadi Jahangard, "*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, spring 1999.
- [7] Ghohroudi Tali, Manije, "*Establishing and Modification methods of creating elevated models, Case study: Golestan 2 Dam*", Geographic Researches Journal, Number 57, fall 2006.
- [8] Asadian F., Ahmadi M., Arzjani Z., Partani S., Pirmoradi R., "*A Comparison of Different Procedures of Geostatic in the Study of Place Changes of the Level of Underground Water by GIS (A Case Study of Razan-Ghahavand Plain in Hamedan Province, Iran)*", Proceeding of International Conference on Water Resources (ICWR2009), Shahroud University, Summer 2009.