Evaluation of Interpolation Methods for Mapping pH of Groundwater

M.H.J.P. Gunarathna, M.K.N. Kumari and K.G.S. Nirmanee

Department of Agricultural Engineering and Soil Science, Faculty of Agriculture, Rajarata University of Sri Lanka, Puliyankulama, Anuradhapura, Sri Lanka

Abstract - **Groundwater is one of the most important natural resources in the world, which is threatenedon its quantity and quality. pH of water is a basic quality parameter and it determines the solubility and biological availability of chemical constituents such as nutrients and heavy metals. Excessively high and low pH can be detrimental in the use of water. Mapping the current situation of groundwater quality providesfor better management of resources. Interpolation methods facilitate to estimate the values for unsampled points and create a continuous dataset to study the spatial distributions. In this study the geostatistical analyst and spatial analyst tools were used to compare the accuracy of different interpolation methods by interpolatingbased on the spatial pattern of groundwater pH in *Malwathu Oya* cascade-I in Anuradhapura District, using Arc GIS 10.2. The total area of the cascade was divided into 1 km^2 grids and forty wells were purposely selected to include two wells per grid,in assessing the groundwater pH during the dry period. Inverse Distance Weighted (IDW), Radial Basis Function (RBF), which are deterministic interpolation methods and kriging which is a geostatistical interpolation method were used with different parameters in both spatial and geostatistical analyst. Empirical Bayesian Kriging was used additionally for geostatistical analyst. The method which shows least root mean square error (RMSE) was selected as the best method to interpolate the spatial variation of groundwater pH. As a spatial analyst tool, Universal Kriging method was given the least RMSE value.As a geostatistical analyst tool, Empirical Bayesian Kriging with linear semivariogram model recordedthe least RMSE value.It can be concluded that geostatistical interpolation method performs better than deterministic interpolation methods for mapping groundwater pH in the study area.**

Keywords **-Geostatistical, GIS, Groundwater, Interpolation, pH, Spatial**

I. INTRODUCTION

Water is essential for sustenance of life. In this, groundwater is one of the most valuable natural resources, which supports human health, economic development and ecological diversity. Due to population explosion, industrial improvement and agricultural development, extraction of groundwater has increased.When sourcing, groundwater across the world often need to address sustainability as indicated by falling water tables, drying of wetlands, increasing sea water intrusion and general deterioration of water quality.Thus practical actions to protect the natural quality of groundwater is essential. As many professionals point out, groundwater quality mapping over extensive areas is the first step in water resources planning [1] and groundwater can be optimally used and sustained only when the quantity and quality is properly assessed [2].

The spatial distribution of quality groundwater shows some heterogeneity and the measurement of quality parameters at every location is not always feasible on account of time as well as cost of the data collection. Therefore, prediction ofvalues based on selectively measured values is one alternativewhile minimizing errors and enhanced rate of calculation accuracy. Geographical Information System (GIS) is a leading tool and has great potential for use in environmental problem solving in several areas, including engineering and environmental fields [3]. Due to the emergence ofgeostatistical analyst as an innovative tool to fill up the gap between geostatistics and GIS, many researchers widelyused it for the analysis of spatial variation of groundwater characteristics.

Reference [4] stated that spatial interpolation is a procedure of predicting the value of attributes at unsampled sites from measurements made at point locations within the same area. Several researcheshave been undertaken to compare different interpolation methods in a variety of situations, using GIS in areassuch asgroundwater depth, groundwater contamination, groundwater quality, etc.,[5],[6]. Kriging, Inverse Distance Weighting (IDW), and Radial Basis Functions (RBF) are three well-known spatial interpolation techniques commonly used for characterizing the spatial variability and interpolation between sampled points and generating prediction maps [7].As revealed by [8] local polynomial method and IDW were the best methods to estimate EC and pH, respectively in a study carried out in Hamedan-Bahar plain, west of Iran. Reference [9] showed that kriging and co-kriging methods are superior to IDW. Therefore it is important to identify a suitable interpolation method for groundwater quality mapping as different studies in different places showed different results. The present study investigated the best interpolation method to illustrate the spatial distribution of the water quality parameters in shallow regolith aquifer present in north central region of Sri Lanka.

II. MATERIALS AND METHODS

A. Study Area

Malwathu Oya cascade-I is one of the meso catchments of *Nuwarawewa* catchment, which is located in the DL_{1b} agro-ecological region of Sri Lanka. It belongs to *Nuwaragam Palatha* - East and *Mihinthale* divisional secretariat areas in Anuradhapura District. The total area of the cascade was divided into 1 km^2 grids and two wells werepurposively selected from each grid to assess the pH level of groundwater.

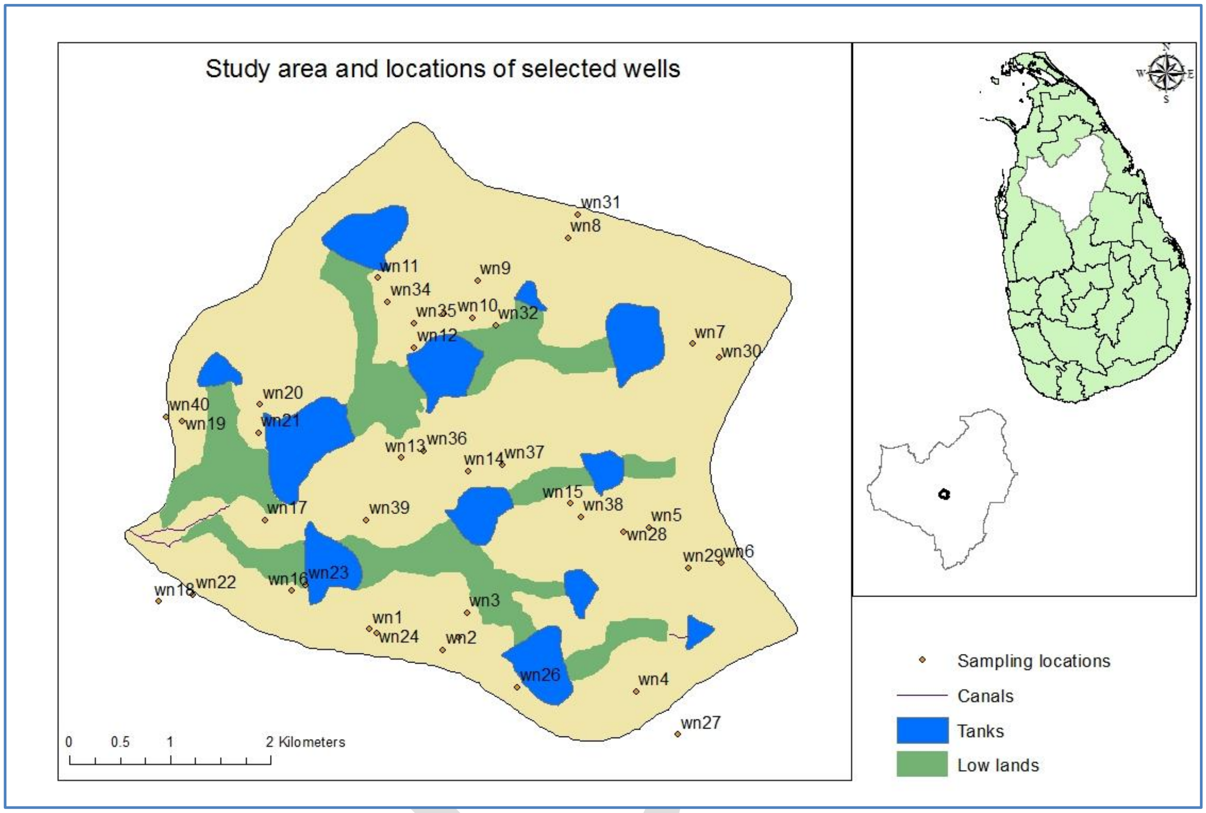

Fig. 1 Study area and selected well locations

B. Assessment of pH of Groundwater

Groundwater pH of the selected wells wasmeasured *insitu*, using multi parameter analyser (Hach -HQ40D)and the locations of the sample sites were taken by using handheld global positioning system unit (Magellan - eXplorist 510).

C. Interpolation Methods

The goodness of interpolation is characterized by the difference of the interpolated value from the true value. Because the true value is not known, some measured points had to be select as reference points for testing the accuracy of the interpolation method.In cross validation process in ArcGIS gives an idea of how well the model predicts the values at the unknown locations, using reference points.

Therefore, using geostatistical and spatial analyst tools, surface variation of groundwater pHwas created by using different interpolation methods such as Inverse Distance Weighting (IDW), Radial Basis Function (RBF), Kriging and Empirical Bayesian Kriging.

D. Inverse Distance Weighted (IDW) Method

Inverse Distance Weighted(IDW) is a deterministic estimation interpolator by which unknown values are computed by a linear combination of values at known points [10]. IDW assumes that each input point has a local influence that diminishes with distance[11], [12] and IDW produce surfaces by establishing a neighborhood search of points and weighting these points by a power function. Often, with the increase of power the effect of the points that are farther diminishes. Lesser power distributes the weights more uniformly between neighboring points [13]. The advantage of IDW is that it is intuitive and efficient and it works best with evenly distributed points and it is sensitive to outliers.Unevenly distributed data clusters result in introduced errors [14]. Equation 01 presents how the IDW interpolation technique calculates the value for an unknown location.

$$Z_j = \frac{\sum_i \frac{Z_i}{d_{ij}^n}}{\sum_i \frac{1}{d_{ij}^n}} \qquad \text{Equation 01}$$

Z_j- estimated value for the unknown point at location j
d_{ij}- distance between known point i and unknown point j
Z_i- value at known point i
n - user-defined exponent for weighting

E. Kriging Method

Kriging is a linear interpolation procedure that provides linear unbiased estimation for quantities, which vary in space and it is an advanced geostatistical procedure that generates an estimated surface from a scattered set of points with z-values (Equation 02). The semivariogram plays a central role in the analysis of geostatistical data in this kriging method. It takes into account the spatial autocorrelation in data to create mathematical models of spatial correlation structures commonly expressed by variograms [15], [16].

$\boldsymbol{Z}\,(\boldsymbol{s_0}) = \sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \lambda_{\boldsymbol{i}} \boldsymbol{Z}(\boldsymbol{s_i})$Equation 02

$Z(s_i)$ - measured value at the i^{th} location
λ_i - unknown weight for the measured value at the i^{th} location
s_0 - prediction location
N - number of measured values

F. Empirical Bayesian Kriging

Empirical Bayesian Kriging (EBK) is a geostatistical interpolation method that automates the most difficult aspects of building a valid kriging model. Other kriging methods in Geostatistical analyst, it is required to manually adjust the parameters in order to receive accurate results, but EBK automatically calculates these parameters through a process of sub-setting and simulations. This method is more accurate than other kriging methods for small datasets [17]. EBK use different semivariogram types.

Power $\gamma(h)$= Nugget + $b|h|^{\alpha}$ Equation 03

Linear $\gamma(h)$= Nugget + $b|h|$ Equation 04

Thinplate spline $\gamma(h)$= Nugget + $b|h^2|*ln(|h|)$Equation 05

h = distance
b= slope

G. Radial Basis Function (RBF)

Radial Basis Function (RBF) methods have a series of exact interpolation techniques and the surfaces created from this method go through each measured sample value. Each basis function has a different shape and results in a slightly different interpolation surface[13].Splines consist of polynomials, which describe pieces of a line or surface, and they are fitted together so that they join smoothly [18].

H. Comparison of Different Interpolation Methods

The validation and the competence of the interpolation method can be tested via a technique called cross validation. Cross validation estimation is obtained by leaving one sample out and using the remaining data to estimate the value. This test allows evaluating the goodness of fit of the method and the appropriateness of neighborhood, whilethe interpolation values are compared to the real [16], [19].To compare observed and predicted values, there are a number of criteria and among them Root Mean Square error (RMSE) was selected for this study.

III. Results and Discussion

Summary statistics and histogram of groundwater pH are shown in.figure 01 and figure 02 shows the normal QQ plot of the data set. The pH of groundwater samples ranged from 6.5 to 8.8 with mean and standard deviation 7.6 and 0.5 respectively. All groundwater samples, except well number 27, had pH values within the desirable level for irrigation (6.5 - 8.5) [20].

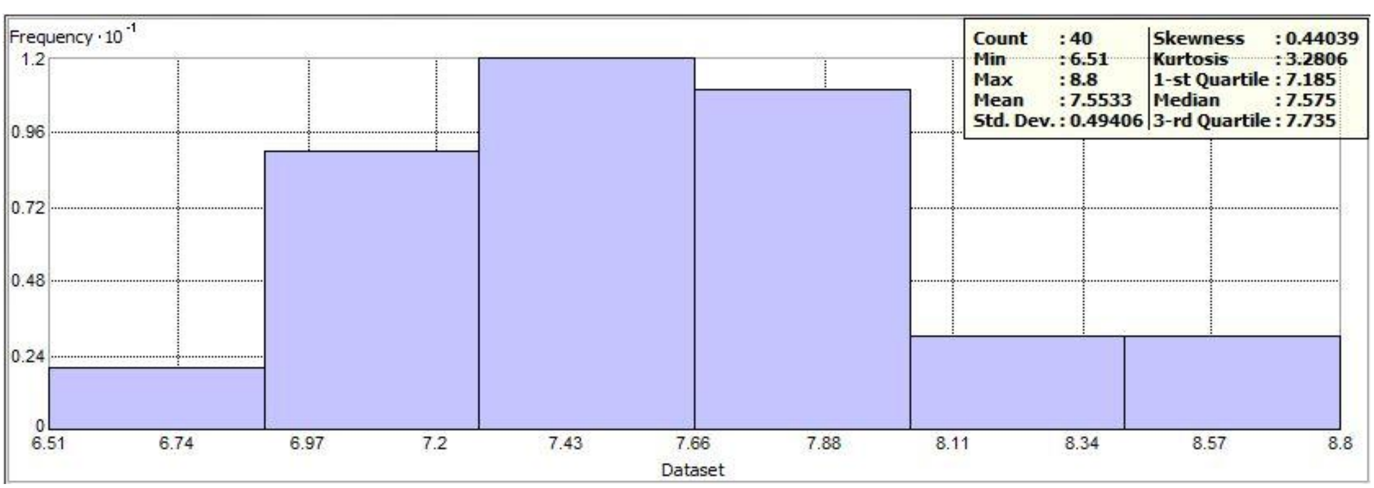

Fig. 2 Summary statistics of pH in groundwater samples

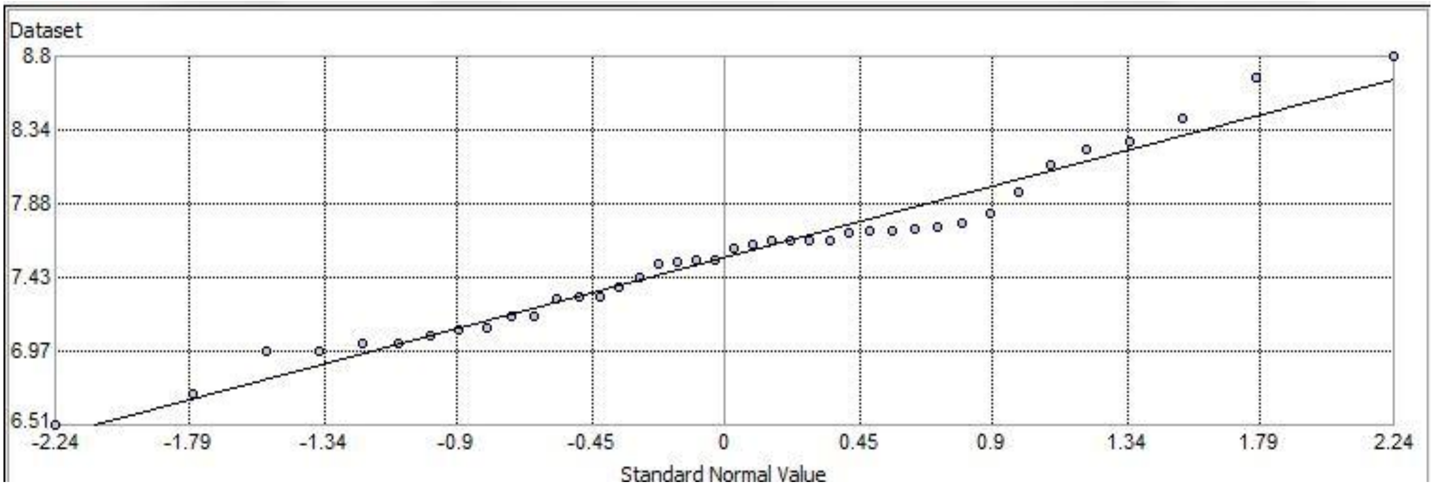

Fig. 3 Normal QQ plot

A. Geo-statistical Analysis

Groundwater pH was interpolated using IDW, Kriging, EBK and RBF interpolation methods using spatial and geostatistical analyst tools, further changing the parameters of each method to estimate the spatial variation of groundwater pH. Table 01 shows the cross validation performance of prediction maps generated by IDW, Kriging, EBK and RBF methods.

According to the data, for mapping the pH of groundwater in *Malwathu Oya* cascade-I, using geostatistical analyst tool, EBK has given the lowest RMSE value with the linear semivariogram type among all the methods tested in this study. When using spatial analyst tool, Universal kriging has given the least RMSE value.

Figures 4 and 5 show the spatial variation of groundwater pH in *Malwathu Oya* cascade-I, using Empirical Bayesiayan Kriging and Universal Kriging.

TABLE I

RMSE VALUES FOR INTERPOLATION METHODS WITH DIFFERENT PARAMETERS

			RMSE value
IDW	Standard	Power 2	0.63248
		Power 3	0.683749
	Smooth	Power 2	0.640349
		Power 3	0.684828
Empirical Bayesian Kriging	Standard circular	Power	0.545651
		Linear	0.54595
		Thin plate spline	0.573473
	Smooth circular	Power	0.535092
		Linear	**0.534907**
		Thin plate spline	0.563479
Kriging/ Cokriging	Standard	Simple	0.5383
		Universal	**0.54103**
		Ordinary	0.541778
	Smooth	Simple	0.555508
		Universal	0.561916
		Ordinary	0.563822
Radial Basis Function	Completely regularized Spline		0.558865
	Spline with tension		0.553069
	Thin Plate Spline		0.689166

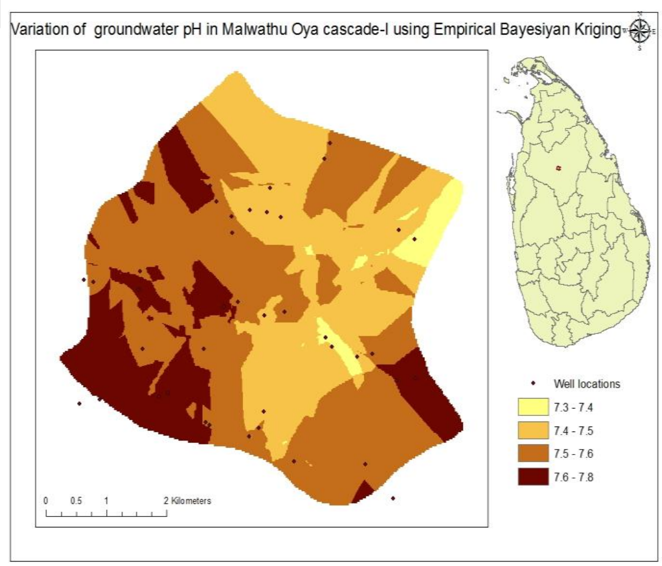

Fig. 4 Variation of groundwater pH in *Malwathu Oya* cascade-I, using Empirical Bayesiyan Kriging

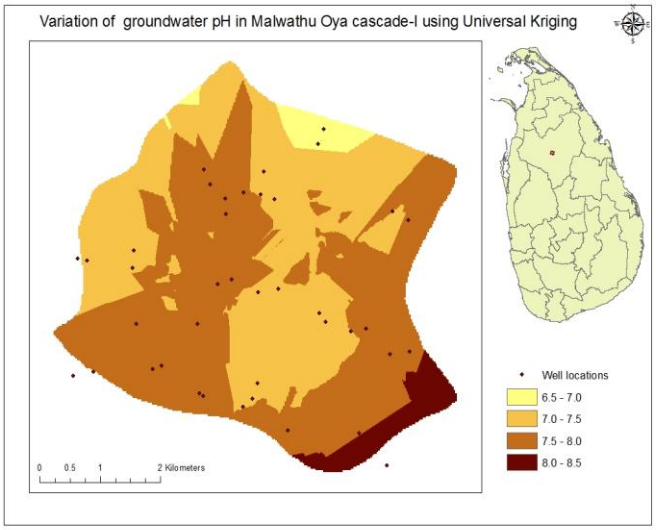

Fig. 5 Variation of groundwater pH in *Malwathu Oya* cascade-I, using Universal Kriging

IV. CONCLUSIONS

The pH of groundwater in *Malwathu Oya* cascade-I was studied by using the geostatistical and spatial analyst tools in ArcGIS. As a spatial analyst tool, Universal Kriging method was given the least RMSE value. As a geostatistical analyst tool, Empirical Bayesian Kriging with linear semivariogram model recordedthe least RMSE value.It can be concluded that geostatistical interpolation method performs better than deterministic interpolation methods for mapping of groundwater pH in the study area.

REFERENCES

[1]. Todd, D.K., (1980). Groundwater Hydrology. 2nd Edition, John Wiley and Sons, New York, USA.

[2]. Kharad, S.M., Rao, K.S., Rao, G.S., (1999). GIS based groundwater assessment model, GIS@development, Nov–Dec 1999. http://www.gisdevelopment.net/application/nrm/water/ground/watg0001.htm. Accessed 21 July 2015.

[3]. Goodchild, M.F., (1993). The state of GIS for environmental problem-solving. In Goodchild, M.F., Parks, B.O. and Steyart, L.T. (Eds.) Environ. Modeling GIS, New York: Oxford University Press, pp. 8–15.

[4]. Burrough, P.A. and McDonnell, R.A., (1998). Principles of Geographical Information Systems. Oxford University Press, Oxford.

[5]. Chao,T, Chou, M., Yang, C., Chung, H., and Wu, M., (2009). Effects of interpolation methods in spatial normalization of diffusion tensor imaging data on group comparison of fractional anisotropy. Magnetic resonance imaging, 27(5), pp. 681-690.

[6]. Chiang, P., Musa, G., Hsieh, D., Liou, D., Wen, C., Chan, T., Chen, H., (2010). Spatial interpolation of cadmium contamination of agricultural soils in Changhua County, Taiwan. International Journal of Environment and Pollution, 40(4), pp. 322-336.

[7]. Zandi, S., Ghobakhlou, A. and Sallis, P., (2011). Evaluation of Spatial Interpolation Techniques for Mapping Soil pH, 19th International Congress on Modelling and Simulation, Perth, Australia, 12–16 December 2011.

[8]. Maroofi, S., Toranjeyan, A., and Zare Abyaneh, H., (2009). Evaluation of geostatistical methods for estimating electrical conductivity and pH of stream drained water in Hamedan-Bahar Plain. Journal of Water and Soil Conservation 16 pp. 169-187.

[9]. Mehrjardi, R.T., A. Akbarzadeh, S. Mahmoodi, A. Heidari and Sarmadian, F.,(2008). Application of geostastical methods for mapping groundwater quality in Azarbayjan Province, Iran. Am. Eurasian J. Agric. Environ. Sci., 3: 726-735.

[10]. Collins, J.R.., Fred, C., (1995). A Comparison of Spatial Interpolation Techniques in Temperature Estimation. Doctoral Dissertation. Virginia Polytechnic Institute and State University, Blacksburg, VA.

[11]. Johnston, Kevin,. Ver Hoef, Jay M., Krivoruchko, Konstantin and Lucas Neil (2001). Using ArcGIS Geostatistical Analyst. Environmental Systems Research Institute, Redlands, CA.

[12]. Anderson, S., (2015). An evaluation of Spatial Interpolation Methods on Air Temperature in Phoenix, AZ. Dept. of Geography, Arizona State University.

[13]. Poshtmasari, H.K., Sarvestani, Z.T., Kamkar, B., Shataei, S. and Sadeghi, S., (2012). Comparison of interpolation methods for estimating pH and EC in agricultural fields of Golestan province (north of Iran). International Journal of Agriculture and Crop Sciences. pp.157-167.

[14]. Balakrishnan, P., Abdul Saleem and Mallikarjun, N. D., (2011).Groundwater quality mapping using geographic information system (GIS): A case study of Gulbarga City, Karnataka, India.

[15]. Gundogdu, K.S. and Guney, I., (2007). Spatial analyses of groundwater levels using universal kriging. J Earth Sys Sci 116 (1): pp. 49-55.

[16]. Uyan, M. and Cay, T. (2010). Geostatistical methods for mapping groundwater nitrate concentrations. Paper presented at the 3rd international conference on cartography and GIS. Nessebar, Bulgaria.

[17]. Environmental Systems Research Institute (ESRI), (2013). ArcGIS Desktop Help 10.2 Geostatistical Analyst.

[18]. Webster, R. and Oliver, M., (2001). Geostatistics for Environmental Scientists. John Wiley and Sons, Ltd, Chichester.

[19]. Leuangthong O., McLennan J.A., Deutsch C.V.,(2004). Minimum acceptance criteria for geostatistical realizations. Nat Resour Res 13:131–141.

[20]. FAO, (1985). Water quality for agriculture; Irrigation and Drainage Paper 29 Rev. 1. FAO, Rome, p.174.