African Journal of Agricultural Research Vol. 6(14), pp. 3402-3408, 18 July, 2011
Available online at http://www.academicjournals.org/AJAR
DOI: 10.5897/AJAR11.027
ISSN 1991-637X ©2011 Academic Journals

Full Length Research Paper

Application of kriging and cokriging in spatial estimation of groundwater quality parameters

Abdolrahim Hooshmand[1], Mahdi Delghandi[2]*, Azizollah Izadi[3] and Khaled Ahmad Aali[4]

[1]Water Sciences Engineering Faculty, Shahid Chamran University, Ahwaz, Iran.
[2]Department of Irrigation and Drainage, Shahid Chamran University, Ahwaz, Iran.
[3]Department of Water Engineering, Ferdowsi University of Mashhad, Iran.
[4]Department of Irrigation and Reclamation, University of Tehran, Iran.

Accepted 21 February, 2011

Excessive use of groundwater aquifers may result in low quality groundwater. The chloride content and the sodium adsorption ratio (SAR) are among the most important water quality parameters, that is, the estimation of these two parameters and, consequently, Na^{+}, Ca^{2+} and Mg^{2+} concentration are much more time-consuming and expensive than water salinity measurement. Thus, it seems necessary to find a guideline to estimate SAR and chloride using salinity extent. For this purpose, geostatistic (cokriging and kriging) methods can be very helpful. In this study, the chloride content and SAR of groundwater in Boukan area were determined by irrigation water salinity (ECw) and the two kriging and cokriging methods were applied. The comparison of the obtained results indicated that for the estimation of both parameters of SAR and Cl, the cokriging method was more accurate than kriging method. However, they differ slightly and, in general, the two methods have suitable accuracy to estimate SAR and Cl-based on water salinity parameter.

Key words: Sodium adsorption ratio, chloride, geostatistic methods, groundwater.

INTRODUCTION

In the recent decade, population growth, rising demand for food and the deficiency of surface water for agricultural crops has increased the surface area of irrigation lands in Iran. This has led to change in the exploitation policy of water resources and soil. Also, this has resulted in the excavation of many wells in most regions in Iran. Therefore, excessive use of groundwater aquifers is growing. The increased exploitation of groundwater resources can decrease regional water quality as a whole (Kardovani, 2007). Hu et al. (2005) pointed out that drought and groundwater level drop are the main causes of soil salinity and alkalinity. Sodium content in the irrigation water affects particles dispersion, soil structure demolition and crop production. If irrigation water with high sodium is applied to a soil for years, the sodium in the water can displace the calcium and magnesium in the soil. This will cause a decrease in the ability of the soil to form stable aggregates and a loss of soil structure and tilth. This will also lead to a decrease in infiltration and permeability of the soil to water leading to problems with crop production. The most common method to assess the effects of sodium is the applying sodium adsorption ratio (SAR).

SAR is a measure of the suitability of water for use in agricultural irrigation, as determined by the concentrations of solids dissolved in the water. The formula for calculating sodium adsorption ratio is (Suarez et al., 2006):

$$SAR = \frac{Na^{+}}{\sqrt{Ca^{2+} + Mg^{2+} / 2}} \quad (1)$$

Where sodium, calcium, and magnesium are in mq/l.

Although SAR is only one factor in determining the suitability of water for irrigation, in general, the higher

*Corresponding author. E-mail: delghandi@gmail.com. Tel: +989151735854. Fax: +986113365670.

Figure 1. The study area.

the sodium adsorption ratio, the less suitable the water is for irrigation. Irrigation using water with high sodium adsorption ratio may require soil amendments to prevent long-term damage to the soil (Michael et al., 2008). As some plants show specific sensitivity to the chloride ion and some others decrease their relative yield in chloride threshold, it is necessary to assess them on the basis of Cl^- variable. Since the determination of Cl^- and the sodium adsorption ratio, and also, the Na^+, Ca^{2+} and Mg^{2+} concentration are much more time- consuming and expensive, it is necessary to find a guideline to estimate the SAR and the Cl^- by the salinity extent. The science involved in the spatial and temporal variations, as well as the variables modeling is referred to as "geostatistics". Geostatistics is a branch of statistics focusing on spatial or spatiotemporal datasets (Einax and Soldt, 1999). Istock and Cooper (1998) applied kriging method to estimate heavy metals. They found that the mentioned method is the best estimator for spatial prediction of lead. Dagostino et al. (1998) studied spatial and temporal variability of nitrate, using kriging and cokriging methods in groundwater. They used microbial data as auxiliary variable in cokriging method. The results showed that cokriging method has suitable accuracy to estimate groundwater quality.

Ahmad (2002) used kriging method to estimate TDS in groundwater and demonstrated accuracy of this method to prediction of TDS. Gaus (2003) studied the pollution of Bangladesh groundwater in view of heavy metal. They used disjunctive kriging method to estimate arsenic concentration and to prepare risk map. Their results showed that 35 milion people are exposed to high concentration of “arsenic” (50 ppm) and 50 million people are exposed to 10 ppm. Finke et al. (2004) used simple kriging to estimate water surface changes in Netherlands and introduced it as a suitable method for mapping of water surface. Barca and Passarella (2008) used disjunctive kriging and simulation methods to make nitrate risk map in 10, 50 (mgr/lit) thresholds, in Modena plain of Italy. Their results showed that disjunctive kriging method is the suitable method to study deterioration level of groundwater. The mentioned method has been widely used for the spatial variability of groundwater quality (Dieleman, 1962; Dick and Heuvelink, 2007; Marengo et al., 2008; Ahmadi and Sedghamiz, 2007). The study area (Boukan region) was not protected against the invasion to groundwater aquifer so that in 1996 there were 9 wells in the whole area to exploit groundwater resources.

At present, the number of wells excavated for agricultural field's irrigation has reached about 2500 wells. This increasingly growing invasion to groundwater aquifer can cause very serious problems, either in natural resources and desertification or in human resources, that is, immigration and unemployment. In this study, considering the significance of the chloride content and the SAR and, also, their impact on soil and agricultural crops, their spatial distribution and the estimation of their values via water salinity have been dealt. For this reason, two cokriging and kriging estimators were used and, finally, their efficiency was assessed and the superior choice was determined.

MATERIALS AND METHODS

Description of the study area

The Boukan region is located in southern part of west Azerbaijan Province, Iran. The study area is located between 36° 32´ N latitude and 46° 13´ E longitude. The total geographical area is 47300 ha (Figure 1). Average elevation in this region is 1330 m above sea level. The average annual precipitation of study area with regard to its semi-arid climate is 517 mm. The major limitation of irrigation lands in this area is included in factors like soil, topography and salinity. In this study, several samples were taken from 80 wells throughout the region. Latitude and longitude of the samples location was recorded via GPS. Then, the sample was collected from each well in a black-colored plastic container and transferred to the laboratory to be analyzed and to measure sodium, calcium, magnesium (for sodium adsorption ratio measurement) and chloride contents, as well as the salinity content.

Cokriging and kriging methods were applied to estimate the parameters including the chloride content and the sodium adsorption ratio. For this reason, ArcGIS 9.2 and GS+ 5.1 software were used. After applying kriging and cokriging methods, in order to make an assessment: firstly, an empirical variogram was drawn from data taken and a theory variogram was fitted there on. Each time a measured value is omitted in a point and another amount is estimated for it from the neighboring points. Then, the real value is returned to the previous position and this is repeated for all measurement points. The assessment was carried out using determination coefficient R^2 and the root mean square error (RMSE). Data variogram was analyzed to examine the spatial correlation and the spatial structure of variables. To make analysis on the data variogram of the SAR and the chloride content after normalization, initially, the variogram of each variable was drawn by using GS+ and, then, an appropriate model selected to be fitted on the empirical variogram.

Geostatistical methods

Geostatistics is a branch of statistics focusing on spatial or spatiotemporal datasets. Structures in a data set can be disclosed. Geostatistics was originally developed for estimation of ore reserves in mining (Einax and Soldt, 1999). The mathematical fundamentals are described in detail by Cressie (1991) and Journel and Huijbregts (1978). The main tool in geostatistics is the semi-variogram, which expresses the spatial dependence between neighboring observations. The semivariogram quantifies the relationship between the semivariance and the distance between sampling pairs by the following equation (Isaaks and Srivastava, 1989, Kitanidis, 1997):

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

Where N(h) is the number of all pair-wise Euclidean distances, and z (x_i) and z (x_i+h) are observations of the variable Z at spatial locations x_i and x_i+h, respectively.

A plot of semi-variance versus h is called the semi-variogram. The semi-variogram is called isotropic if semi- variograms for various directions are the same. When h approaches zero, the experimental variogram extrapolated may approach zero. But in practice it will approach a value which is called the nugget, C_o, caused by measurement error or variability at distances less than the sampling distance. The semi- variograms is expected to increase with increase in h. At a certain value of h the semi-variance may remain constant which is called sill and is equal to the total variance. The range of a semi-variogram is the distance over which the observations are assumed correlated (Alemi et al., 1988). Prior to the geostatistical estimation, we require a model that enables us to compute a variogram value for any possible sampling interval (Ahmadi and Sedghamiz, 2008). The most commonly used models are spherical, exponential, gaussian, and pure nugget effect (Isaaks and Srivastava, 1989).

The mathematical terms of the models are described in several papers (Myers, 1991; Delay and Marsily, 1994).

Kriging method

Kriging method is a precise interpolation estimator used to find the best linear unbiased estimate. Detailed discussions of kriging methods can be found in Goovaerts (1997). The general form of kriging equation is:

$$Z^*(x_p) = \sum_{i=1}^{n} \lambda_i Z(x_i) \tag{3}$$

In order to achieve unbiased estimations in kriging the following set of equations should be solved simultaneously:

$$\begin{cases} \sum_{i=1}^{n} \lambda_i \gamma(x_i, x_j) - \mu = \gamma(x_i, x) \\ \sum_{i=1}^{n} \lambda_i = 1 \end{cases} \tag{4}$$

where $Z^*(x_p)$ is the estimated value at location x_p, $Z(x_i)$ is the known value at location x_i, λ_i is the weight associated with the data, μ is the Lagrange coefficient, and $\gamma(x_i, x_j)$ is the value of variogram corresponding to a vector with origin in x_i and extremity in x_j.

Cokriging method

Cokriging is the multivariate equivalent to kriging. By using multiple data sets it is a very flexible interpolation method, allowing the user to investigate graphs of cross-correlation and autocorrelation. The general equations of cokriging estimator are:

$$\begin{cases} \sum_{l=1}^{v} \sum_{i=1}^{n} \lambda_{il} \gamma_{lv}(x_i, x_j) - \mu_v = \gamma_{uv}(x_j, x) \\ \sum_{i=1}^{nl} \lambda_{il} = \begin{cases} 1, & 1 = u \\ 0, & 1 \neq u \end{cases} \end{cases} \tag{5}$$

Where *u* and *v* are the primary and covariate (secondary) variables, respectively.

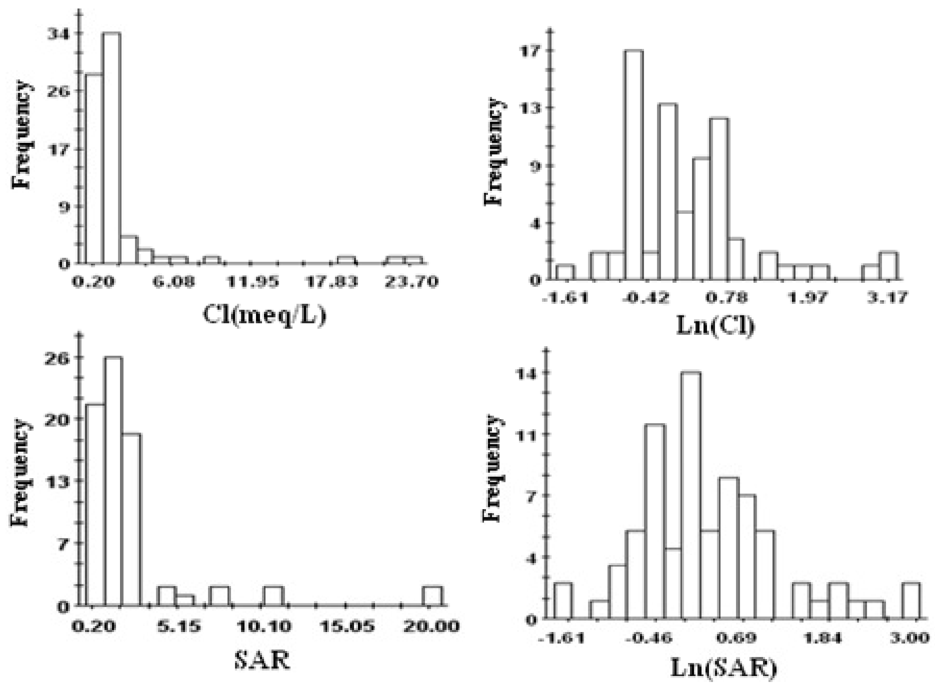
In the cokriging method, the *u* and *v* variates are cross-correlated and the covariate contributes to the estimation of the primary variate. Generally, measuring the covariate is simpler than measuring the primary variate. For cokriging analysis, the cross semi-variogram (or cross-variogram) should be determined in prior. Provided that there are points where both *u* and *v* have been measured, the semi cross-variogram is estimated by (Ahmadi and Sedghamiz, 2008):

$$\gamma_{uv}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \{Z_u(x_i) - Z_u(x_i + h)\}\{Z_v(x_i) - Z_v(x_i + h)\} \tag{6}$$

To evaluate the performance of interpolation methods is used the cross validation method. In this procedure, an observed value is temporarily discard from the sample data set, and one estimated value at that location is determined using the other sample points. This results in a series of observed and estimated values that can be used to assess the validity of the interpolation method. In this study, estimated and observed values were compared using

Table 1. Statistical parameters of studied variables.

Variables	Mean	Min	Max	Coefficient of variation	Standard deviation	Skewness	Kurtosis
SAR	2.32	0.2	20	1.54	3.58	3.64	14.01
Cl(meq/l)	2.27	0.2	23.7	1.89	4.3	4.06	16.07



(a) (b)

Figure 2. Data histogram of SAR and the chloride before and after normalization of data distribution. a) Before normalization of data distribution. b) After normalization of data distribution.

determination coefficient (R^2) and root mean square error (RMSE). The smallest RMSE indicate the most accurate predictions. The RMSE was derived according to Equation 7 (Taghizadeh et al., 2008):

$$RMSE = \sqrt{[\sum_{i=1}^{n} \left(Z(x_{i\,p}) - Z(x_i)\right)^2] / n} \quad (7)$$

Where Z (x_i) is observed value at point x_i, Z (x_{ip}) is predicted value at point x_i, n is number of samples and *i* is the index for the number of data. Data variogram was analyzed to examine the spatial correlation and the spatial structure of variables. To make analysis on the data variogram of the SAR and the chloride content after normalization, initially, the variogram of each variable was drawn by using GS+ and, then, an appropriate model selected to be fitted on the empirical variogram.

RESULTS AND DISCUSSION

Some statistical characteristics such as mean, standard deviation, minimum, maximum, coefficient of variation, skewness, and kurtosis are presented in Table 1 for the variables of the chloride and the sodium adsorption ratio. For data analysis, a historgram was drawn for each study variable as shown in Figure 2a. The histograms demonstrate that the two variables of the SAR and the chloride have skewness and, therefore, logarithmic method was used for data normalization. The variables histogram, after normalization, can be seen in Figure 2b.

Geostatistical analysis

Table 2 shows variogram parameters fitted to the data of the SAR and the chloride content. After evaluating different models, it was demonstrated that the Gaussian Model best suited for the both variables and therefore, it was selected as a best fitted model on the data. Theory and empirical semi- variogram were prepared for the SAR and the chloride in GS+ media as shown in Figures 3 and 4 respectively. Since some parameters are affected by environmental factors, they can be involved in

Table 2. Variogram parameters fitted to the data of the SAR and the chloride content.

Variable	Model	Nugget effect	Sill	Radius impact(m)	R^2
SAR	Gaussian	0.485	1.629	42200	0.9
Cl(meq/l)	Gaussian	0.418	1.2	26200	0.83

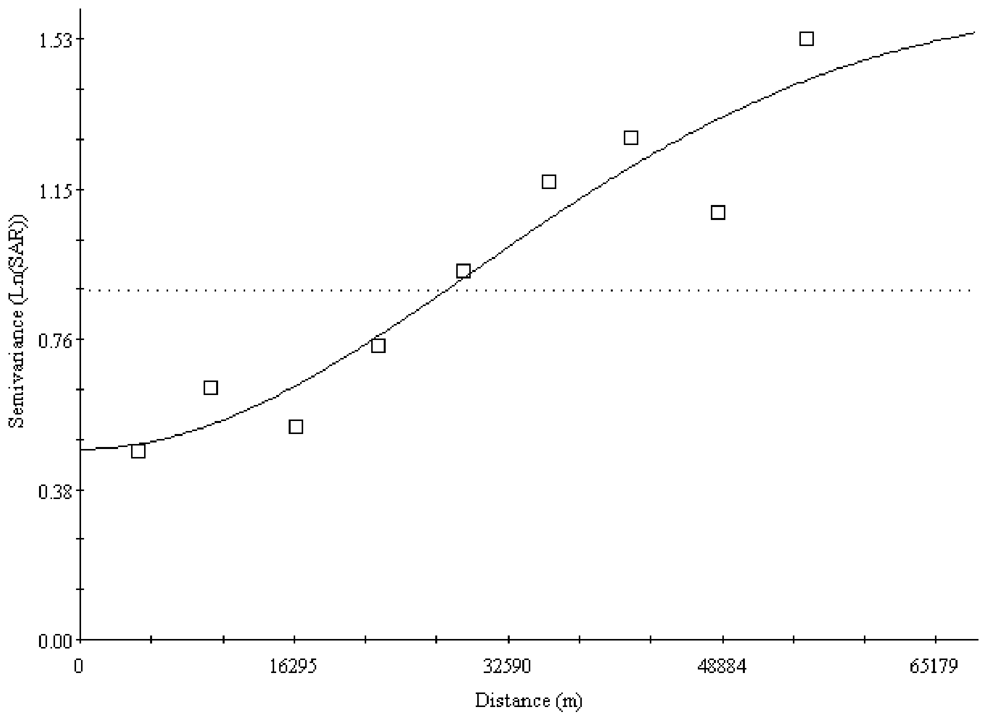


Figure 3. Empirical semi-variance and theory model fitted to SAR.

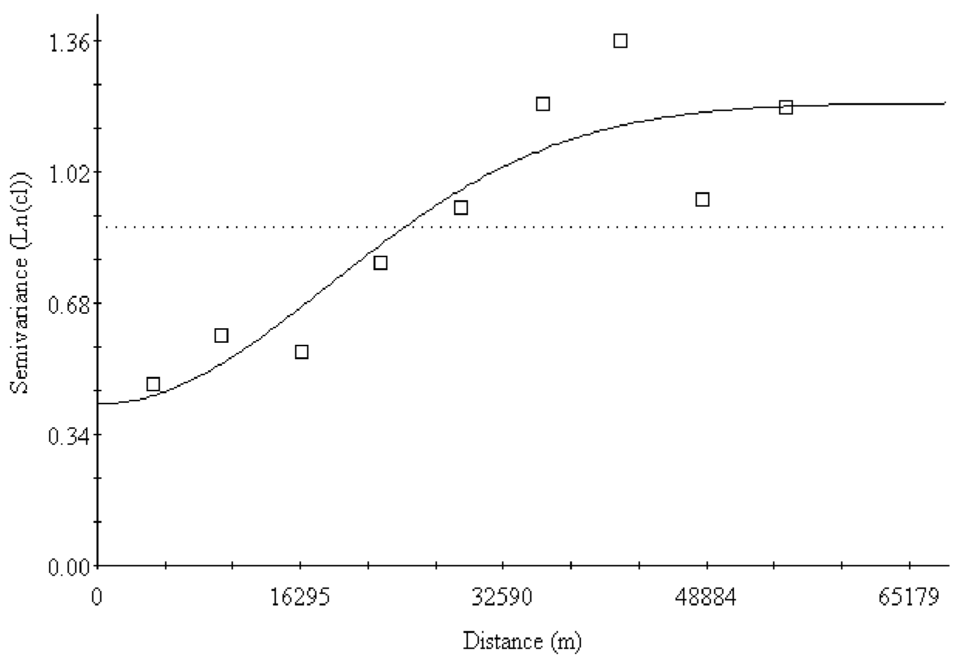


Figure 4. Empirical semi-variance and theory model fitted to Cl.

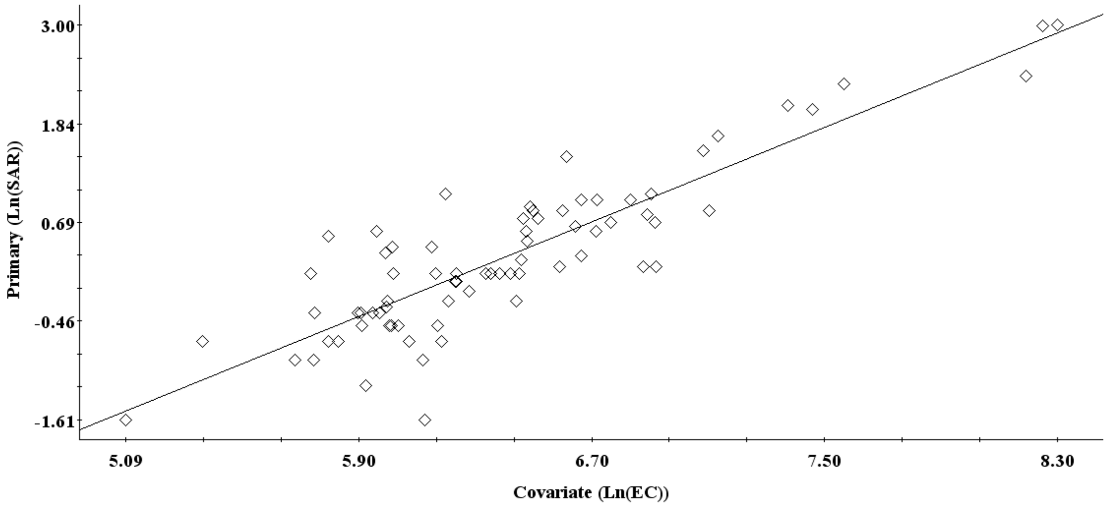


Figure 5. Correlation diagram of Ln (EC) with auxiliary variable Ln (SAR).

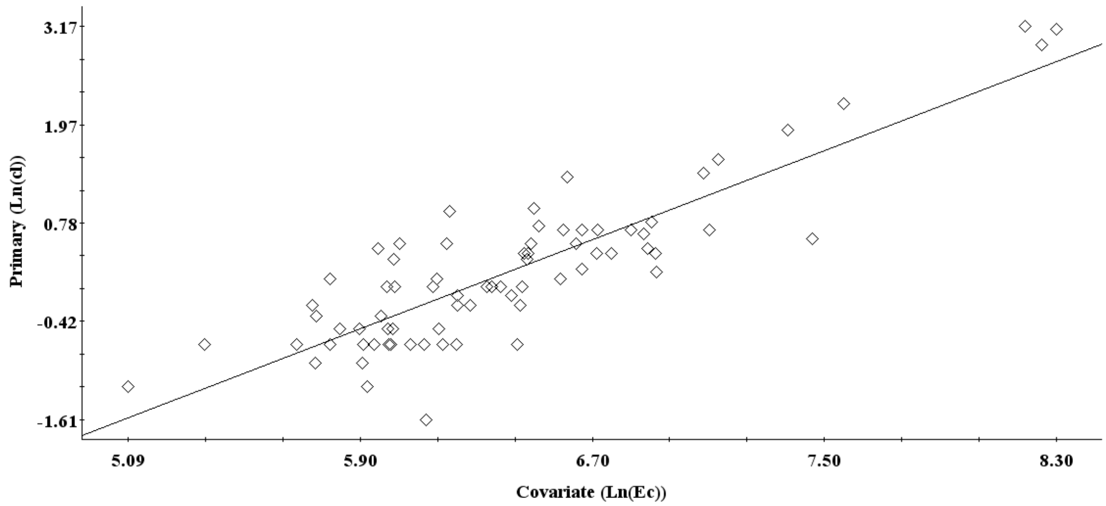


Figure 6. Correlation diagram of Ln (EC) with auxiliary variable Ln (Cl).

the estimation of the main variable by using the cokriging estimator, if a correlation exists. By the examination of the correlation among the studied variables, it was observed that there was a high correlation between the variables of Cl and the SAR with the EC. R^2 = 0.77 between SAR and EC and R^2 = 0.75 was obtained between Cl and EC. Therefore, the estimation of the chloride content and the sodium adsorption ratio through EC by using the cokriging estimator will be entirely reasonable. Figures 5 and 6 illustrate the correlation between SAR and EC as well as Cl and EC, respectively.

Evaluation of geostatistical methods

With regard to RMSE criterion as depicted in Table 3, the estimation of sodium adsorption ratio by cokriging with a RMSE = 3.07 was obviously more precise than kriging method, though the two methods were acceptably accurate. For the estimation of the chloride content, the accuracy of cokriging was higher than that of kriging (RMSE = 3.82).

Conclusion

In this research kriging and cokriging methods were used to estimate the sodium adsorption ratio (SAR) and the chloride content of groundwater on the basis of groundwater salinity data. Results obtained from cokriging method were compared with those from kriging method. For the data of the SAR and the Cl, the Gaussian

Table 3. Assessment of cokriging and kriging methods in the estimation SAR and Cl according to RMSE.

Parameter	Cokriging	Kriging
SAR	3.07	3.15
Cl	3.82	3.89

variogram was selected as the best fitted model. Radius impacts of Gaussian model for the data of the SAR and the Cl determined 42200 and 26200 m, respectively. Considering the high correlation between the study variables ($R^2 = 0.77$ between SAR and EC, and $R^2 = 0.75$ between SAR and EC were obtained), the applying of cokriging estimator to estimate the sodium adsorption ratio and the chloride content is entirely reasonable. According to RMSE, the accuracy of geostatistical methods in the estimation of groundwater quality parameters is assessed as being very good.

The results showed that this method can be applied as a tool to estimate the sodium adsorption ratio and the chloride content of groundwater in areas with data restriction.

REFERENCES

Ahmed S (2002). Groundwater monitoring network design: Application of Geostatistics with a few Case studies from a granitic aquifer in a semiarid region. Groundwater Hydrology. Sherif, M.M., V.P. Singh and M. Al-Rashed (Eds.). Balkema, Tokyo, Japan, 2: 37-57.

Ahmadi SH, Sedghamiz A (2007). Geostatistical analysis of spatial and temporal variations of groundwater level. Environ. Monit. Assess., 129: 277-294.

Ahmadi SH, Sedghamiz A (2008). Application and evaluation of kriging and co-kriging methods on groundwater depth mapping. Environ. Monit. Assess., 138: 357-368.

Alemi MH, Shahriari MR, Nielsen DRD (1988). Krigong and cokriging of soil water properties. Soil Technol., 1: 117-132

Barca E, Passarella G (2008). Spatial evaluation of the risk of groundwater quality degradation. A comparison between disjunctive kriging and geostatistical simulation, Environ. Monit. Assess., 137: 261-273.

Cressie NAC (1991). Statistics for Spatial Data. Wiley. New York.

Dagostino V, Greene EA, Passarella G, Vurro M (1998). Spatial and temporal study of nitrate concentration in groundwater by means of coregionalization. Environ. Geol., 36: 285-295.

Delay F, Marsily G (1994). The integral of the semivariogram: A powerful method for adjusting the semivariogram in geo- statistics. Math. Geol., 26: 301-321.

Dick JB, Heuvelink BM (2007). Optimization of sample patterns for universal kriging of environmental variables. Geoderma, 138: 86-95.

Dieleman PJ (1962). Reclamination of salt-affected soils in Iraq. International Institue of Land Reclamination and Improvement, Pub. No. 5, Wagningen The Netherland.

Einax JW, Soldt U (1999). Geostatistical and multivariate statistical methods for the assessment of polluted soils-merits and limitations. Chemom. Intell. Lab. Syst., 46: 79-91

Finke PA, Brus DJ, Bierkens MFP, Hoogland T, Knotters M, Vries FD (2004). Mapping groundwater dynamics using multiple sources of exhaustive high resolution data. Geoderma, 123: 23-39.

Gaus I, Kinniburgh DG, Talbot JC, Webster R (2003). Geostatistical analysis of arsenic concentration in groundwater in Bangladesh using disjunctive kriging. Environ. Geol., 44: 939-948.

Goovaerts P (1997). Geostatistics for natural resources evaluation. New York: Oxford University Press.

Hu K, Huang Y, Li H, Li B, Chen D, White RE (2005). Spatial variation of shallow groundwater level, electrical conductivity and nitrate concentration, and risk assessment of nitrate contamination in North China Plain. Environ. Int., 31: 896-903.

Isaaks E, Srivastava RM (1989). An introduction to applied geostatistics. Oxford University Press, 561 P. N.Y.

Istok JD, Cooper RM (1998). Geostatistics Applied to Groundwater Pollution. III: Global Estimates. J. Environ. Eng., 114(4): 915-928.

Journel AG, Huijbregts Ch (1978) Mining Geostatistics, Academic Press, London.

Kardovani p (2007). The problems of waters in iran: Salin water, Tehran University Press, 2: 237(in Farsi)

Marengo E, Gennaro CM, Robotti E, Maiocchi A, Pavese G, Indaco A, Rainero A (2008). Statistical analysis of ground water distribution in Alessandria Province (Piedmont Italy). Microchem. J., 88: 167-177.

Michael AM, Kherpar SD, Sondhi SD (2008). Water wells and pumps. McGraw-Hill, New Delhi.

Myers DE (1991). Interpolation and estimation with spatially located data, Chemom. Intell. Lab. Syst., 11: 209-228.

Suarez DL, Wood JD, Lesch SM (2006). Effect of SAR on water infiltration under a sequential rain–irrigation management system. Agric. Water Manage., 86: 150-1 64

Taghizadeh Mehrjardi R, Zareian Jahromi M, Mahmodi Sh, Heidari A (2008). Spatial Distribution of Groundwater Quality with Geostatistics (Case Study: Yazd-Ardakan Plain). World Appl. Sci. J., 4(1): 09-17.