Desalination and Water Treatment
www.deswater.com
doi: 10.1080/19443994.2013.803937

52 (2014) 1980–1989
February

Taylor & Francis
Taylor & Francis Group

Multivariate statistical and geostatistical techniques for assessing groundwater salinization in Sfax, a coastal region of eastern Tunisia

Ibtissem Triki*, Nadia Trabelsi, Moncef Zairi, Hamed Ben Dhia

Laboratoire" eau, energie et environnement", Ecole nationale d'Ingénieurs de Sfax, route de soukra Route Soukra BP 1173, Sfax 3038, Tunisia
Tel. +216 22668399; email: Triki_ibtissem@yahoo.fr

Received 2 January 2013; Accepted 28 April 2013

ABSTRACT

In this study, we investigate the ability to combine a multivariate statistical analysis with the cokriging method to point out the groundwater salinization in the coastal Sfax aquifer (eastern Tunisia). First, multivariate statistical analysis such as principal component analysis (PCA) and cluster analysis were performed on 75 water samples. PCA identifies three main processes influencing groundwater chemistry which are seawater intrusion, water–rock interaction, and contamination by nitrates, these three factors accounted for 76% of total variance of the groundwater. Furthermore, cokriging is applied to take into account spatial dependence between the studied variables. Five variables were processed: concentration of sulfates, chlorides, sodium and the sodium adsorption ratio, as primary variables, and the more numerous data for total dissolved solid, as auxiliary variables. The generated spatial variability maps highlighted the high-risk zone of groundwater contamination of the superficial aquifer of Sfax. The effectiveness of the high estimation capability of the cokriging is demonstrated by cross-validation. Compared with ordinary kriging for a single variable, cokriging can provide an improvement of the uncertainty in terms of reducing the mean-squared error and mean error.

Keywords: Groundwater quality; Principal component analysis; Cluster analysis; Geostatistics; Cokriging; Cross-validation.

1. Introduction

Groundwater is the principal *source* of fresh water in arid areas to meet the demands of domestic, agricultural, and industrial needs [1]. However, excessive use of groundwater aquifers may result in low-quality groundwater [2,3]. In recent years, groundwater quality analysis gained importance regarding the understanding of the processes contributing to pollution. The factors behind this contamination may be natural or anthropogenic. Important natural processes contributing to pollution in groundwater are rock–water interactions, dissolution, precipitation, sorption, and geochemical reactions. Anthropogenic activities such

*Corresponding author.

Presented at the 6th International Conference on Water Resources in Mediterranean Basin (WATMED6), 10–12 October 2012, Sousse, Tunisia

as waste disposals, leaching of salts, fertilizers, pesticide from the agricultural fields, and salt intrusion due to over exploitation contribute to groundwater pollution [4].

The early characterization of groundwater facies utilized graphical representations of the major compositions of groundwater. These classical classification techniques such as Stiff and Piper diagrams only consider selected major water constituents in determining the groundwater type [5]. In recent years, with the increasing number of chemical and physical variables of groundwater, a wide range of statistical methods such as principal component analysis (PCA), factor analysis (FA), cluster analysis (CA), and discriminate analysis (DA) are applied for the interpretation of complex monitoring data matrices. They offer a better understanding of water quality of the studied systems, allow identification of the possible factors/sources that influence the water systems and present a valuable tool for reliable management of water resources as well as rapid solutions to improve water quality [6–9]. These methods of cluster analysis and principle component were used with remarkable success to evaluate the temporal and spatial characteristics of groundwater salinization in the coastal areas [10–14].

Geostatistics offer a variety of techniques to make optimal use of measurement information for interpolating groundwater chemistry and pollutant concentration levels in space [15]. Kriging has been applied to quantify variability of groundwater quality variables [16–19]. For example, Nas [18] applied ordinary kriging to determine spatial distribution of groundwater quality parameters in Konya (Turkey). Dash et al. [19] assessed the risk of pollutant concentrations in shallow groundwater exceeding anthropogenic limits and found that the indicator kriging was better than ordinary kriging estimation to study deterioration level of groundwater. Whereas kriging is used to evaluate the spatial distribution of a variable based on sampled data of the variable itself, cokriging is applied to improve estimation of undersampled variables by taking into account their spatial cross-correlation with better sampled ones [20]. As cokriging has been shown to be successful in constructing more precise maps of transmissivities and other aquifer parameters [21,22], it is consequently expected to be a more convenient tool for mapping groundwater quality variables [23–26]. Istok et al. [25] applied cokriging to estimate pesticide concentrations with the auxiliary data of nitrates. Mehrjardi et al. [26] studied spatial variability of some groundwater quality indices, using inverse distance weighting (IDW), kriging, and cokriging. Their results showed that cokriging is the best method to estimate groundwater quality indices. Castrignanò et al. [23] used multivariate geostatistics and GIS to delineate the zones at high risk of groundwater contamination in Apulia region. Jang [27] compared two nonparametric multivariate kriging methods (multiplying indicator variables and averaging indicator variables) to probabilistically determine extents of pollutants in groundwater.

The main objective of the present study is to examine and estimate the spatial variability of groundwater salinization in the coastal Sfax aquifer. For this purpose, multivariate statistical approaches in conjunction with a geostatistical method of cokriging are performed. The CA and PCA techniques are used to understand the interrelation of the groundwater quality parameters and to identify the underlying major factors that are responsible for groundwater pollution.

Afterward, ordinary kriging and cokriging are used to estimate the spatial variability and distribution of the groundwater quality, and finally, their efficiency was assessed and the superior choice was determined.

2. Materials and methods

2.1. Study area

The study area, located in eastern Tunisia around Sfax city (Fig. 1(a)), shows a semi-arid climate with an annual precipitation mean of 230 mm, an average annual temperature of 20°C and a potential evapotranspiration of 328 mm/yr [28]. This region has shown in recent decades a significant development of agricultural and industrial activities, associated with high population growth [29].

In general, the aquifers of the study area are divided into two main systems: the deep confined aquifer and the shallow aquifer [28]. The superficial aquifer, object of this study, covers an area of 8,500 Km^2 (Fig. 1(b)). It is limited east by the Mediterranean Sea, the N-S Axis mountain chain west, the Korj, Bouthadi, Chorbane, Zeramdine and Djemmel Hills north, and Mezzouna Mountain south. A number of sebkhas (salt plain) are spread west. The geologic formations of the studied aquifer belong to the Mio-Pliocene and Quaternary layer system, and they are composed of alternations of sandy clay and sandy layers. These deposits present several productive layers separated by semi-permeable layers. The thickness of the reservoir is ranging from 20 m to 140 m; the transmissivity varies from $2\ 10^{-4}\ m^2/s$ to $5\ 10^{-3}\ m^2/s$, and the storativity from $7\ 10^{-5}$ to $0.3\ 10^{-3}$. The annual pumping flux increased from 36.5 Mm^3 in 2000 to 56.6 Mm^3 in 2005 through 9,547 pumping wells [12]. The piezometric maps built up by groundwater

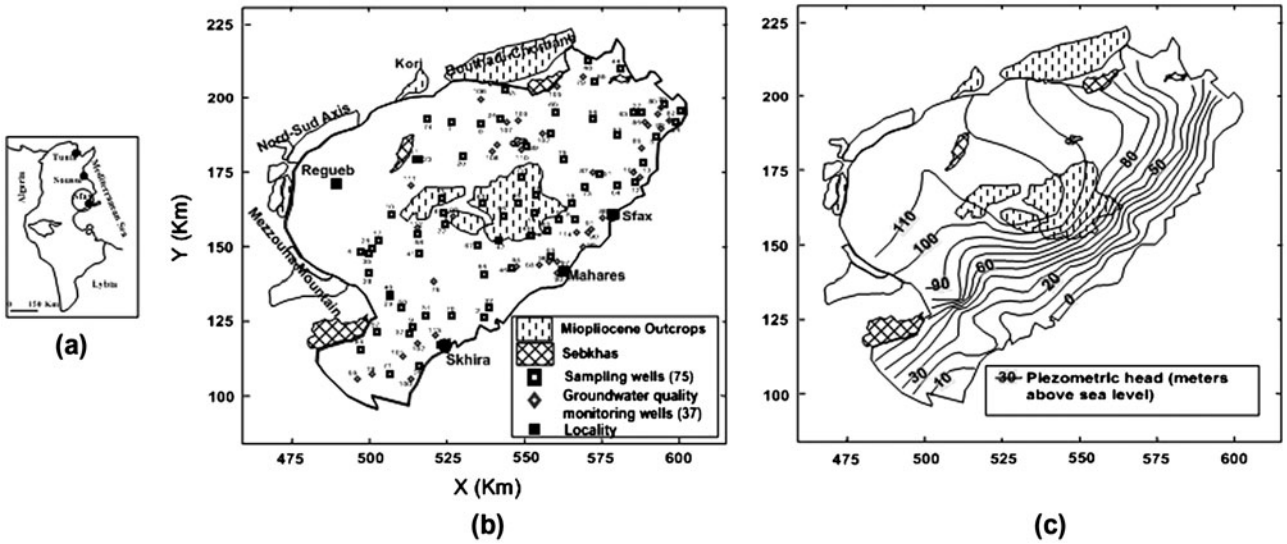

Fig. 1. Locations of study area in the eastern Tunisia (a), sampling sites (b) and piezometric maps of the superficial aquifer of Sfax measured in December 2004 (c).

levelling survey during December 2004 (Fig. 1 (c)) show a recharge area located in the middle of the study area and a high groundwater discharge in the Mediterranean Sea. In the northern part of the aquifer, the groundwater flows towards the zone of topographic depressions occupied by sebkhas El Ghorra and Sebkhat El Jem. The mean hydraulic gradient of the aquifer is ranging from 0.4% in the eastern part to 0.15% in the septentrional part of the basin [28].

2.2. Groundwater sampling and measurement

In this study, 75 representative groundwater samples were collected during December 2004 from farm, monitoring and public supply wells in the Sfax basin and were analyzed for major chemical constituents (Fig. 1 (b)).

Unstable parameters such as pH and electrical conductivity (EC) were measured *in situ* and the hydrochemical parameters were measured in laboratory. The analyses of K and SO_4 were undertaken by gravimetry, those of HCO_3, Cl, NO3, and Ca by the titrimetry, whereas Mg and Na were analyzed by atomic adsorption spectrometry.

An important chemical parameter for judging the degree of suitability of water for irrigation is sodium adsorption ratio (SAR), which can be calculated by Eq. (1) given by the author [30], where all the ions are expressed in me/L.

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

Available information regarding the salinity of groundwater in the study area consisted of data from 37 monitoring wells measured in October 2004 (Fig. 1 (b)). The reported data by the local authorities were used as auxiliary variables in geostatistical estimation.

Basic statistics of major ions concentrations, EC, pH, TDS, and SAR are presented in Table 1.

The coefficient of variance (CV) shows the degree of variability of each ion concentration in the groundwater [31]. It is noteworthy that the CV value reflects a moderate to high variability among samples of the variables. The highest variability was for NO_3, followed by Cl, and K, with a coefficient of variation values above 80%, reflecting spatial variation of groundwater quality in the Sfax superficial aquifer. Furthermore, most of the coefficients of kurtosis and skewness are respectively close to 3 and 0 indicating a normal distribution of data. In contrast, the distribution of pH values and K and NO_3 concentrations are most skewed and showed some shifting from normality.

In the present study, the R-mode hierarchical CA was adopted to the standardized data using Ward's method, with squared Euclidean distance as a measure of similarity.

2.3. Theory of cokriging

In this study, a brief description of cokriging is presented below and further details are given in textbooks [32–35] or papers [36–39]. All geostatistical analyses were conducted using geostatistical software package ISATIS.

Table 1
Statistics of hydrochemical parameters in groundwater and the irrigation water quality standard

	Units	Number of samples	Min	Max	Mean	CV (%)	Skewness	Kurtosis
pH		73	6.5	11	8	75	2.4	11
EC	mS/cm	72	0.5	13.8	5.86	58	0.54	2.5
K^+	mg/L	72	1.2	47	12.5	81	1.41	4.7
HCO_3^-	mg/L	71	25	450	174.3	47	0.6	3.7
Mg^{2+}	mg/L	75	0.56	455	140	65	0.79	3.4
Ca^{2+}	mg/L	75	40	1,042	517	46	−0.13	2.4
NO_3^-	mg/L	66	0.79	360	48.9	110	3.2	17.76
SO_4^{2-}	mg/L	75	183	3,700	1933	43	−0.12	2.66
Na^+	mg/L	75	21	2,110	709	68	0.8	3.1
SAR		75	0.14	13.95	5.6	58	0.5	2.5
Cl^-	mg/L	75	19.5	3,000	883.1	87	1.1	3.4
TDS	g/L	113	0.46	11.4	4.22	55	0.69	3.13

Note: Min: minimum; Max: maximum; CV: coefficient of variance.

The first step in multivariate geostatistics is to establish a suitable model for cross-continuity and dependency between two or more variables. This positive correlation between variables is called cross-regionalization or co-regionalization and it can be estimated by cross-covariance and cross-variogram [39]. These models are used to describe and interpret the cross-continuity and dependency between two or more variables. As an example, let $Z_i(x)$ and $Z_j(x)$ be two random variables. Hence, under the second-order stationarity, the cross-variogram as:

$$\gamma_{ij}(h) = \frac{1}{2}E\{[Z_i(x+h) - Z_i(x)][Z_j(x+h) - Z_j(x)]\} \qquad (2)$$

For modeling the co-regionalization a possible model is the so-called linear model of co-regionalization (LMC) [32]. In this model, all direct ($\gamma_{ii}(h)$) and cross-variograms ($\gamma_{ij}(h)$) are expressed as a linear combination of the same basic structures. In the situation of three basic structures,

$\gamma_1(h)$, $\gamma_2(h)$, and $\gamma_3(h)$, the MLC is written as:

$$\begin{aligned}\gamma_{ii}(h) &= b_{ii}^1\gamma^1(h) + b_{ii}^2\gamma^2(h) + b_{ii}^3\gamma^3(h)\\ \gamma_{ij}(h) &= b_{ij}^1\gamma^1(h) + b_{ij}^2\gamma^2(h) + b_{ij}^3\gamma^3(h)\end{aligned} \qquad (3)$$

where the coefficients b are the contributions to the sill from each structure.

All these structures must be modeled simultaneously and should be adequate to all experimental, direct, and cross-variograms, to guarantee positive definiteness and proper regionalized correlation coefficients [22].

Cokriging extends the principle of optimal estimation using regionalized variable theory from that of a single property to situations where there are two or more co-regionalized properties. In other words, co-kriging takes advantages of intervariable correlation. Cokriging is more efficiently used where one variable may not have been sampled sufficiently (owing to experimental difficulties, high costs, etc.) to provide estimates of acceptable precision. Estimation precision can be improved by utilizing the spatial correlation between the under-sampled (primary) variable and other, more frequently sampled co-variables. The concepts of co-kriging discussed here assume only one co-variable, but the equations are readily expanded to include additional co-variables [24].

The co-kriged value of the undersampled, or primary, variable Z_2, is computed as a weighted average of the observed values of the co-variable, Z_1, and Z_2 occurring in the estimation neighborhood of each kriged point. The co-kriged value $\hat{z_2}$ at point zero is [21]:

$$\hat{Z}(0) = \sum_{i=1}^{n_1} W_{1i}Z_1(i) + \sum_{j=1}^{n_2} W_{2j}Z_2(j) \qquad (4)$$

where W_{1i} and W_{2j} are the weights associated with Z_1 and Z_2, respectively, and n_1 and n_2 are the number of neighbors of Z_1 and Z_2 involved in estimating $\hat{z_2}$at point of interest zero, respectively.

The weights on observed values of Z_1 and Z_2 are chosen so that the estimate is unbiased with minimum variance.

Like kriging, solution of the co-kriging system also yields the co-kriging estimation variance for interpolated location zero. The equations of co-kriging system have been presented in full by authors [21,38,39].

3. Results and discussion

3.1. Principal component analysis

Variables for PCA analysis in this study were pH, EC, HCO_3^-, Ca^{2+}, Mg^{2+}, Na^+, K^+, SO_4^{2-}, NO_3^-, Cl^-, TDS, and SAR. By applying the Pearson's correlation matrix for the 12 variables (Table 2), there are significant positive correlations between TDS with EC, Mg^{2+}, Cl^-, Na^{2+}, and SO_4^{2-} which indicate that such ions in the groundwater were likely from the same sources.

Based on the eigen values (>1), the first three factors are selected to represent the hydrogeochemical processes of the groundwater, without the loss of significant information. The analysis generated three factors which together account for 76% of variance. The rotated loading, eigen values, percentage of variance, and cumulative percentage of variance of all the three factors are given in Table 3.

The first eigen value is 5.9 which accounts for 49.7% of the total variance and this constitutes the first and main factor. The second and third eigen values are 1.67 and 1.47 and these accounts 14 and 12.25%, respectively, of the total variance. Factor 1 is mainly associated with very high loading of Na, TDS, EC, SAR, Cl, SO_4, and Mg. This factor can be associated with the seawater intrusion affecting the studied aquifer, which is considered to be an important contamination process affecting the quality of groundwater. The second factor, accounting for 14% of total variance, is mainly associated with negative loading (−0.7) for HCO_3 and positive loading for pH and K. PC_2 is assumed to be indicative of the natural processes and water-rock interaction. The third factor explains 12.25% of the variance and is mainly related to NO_3. The contribution of NO_3^- to the groundwater quality likely originated from the excessive application of agricultural fertilizers associated with infiltration of irrigation return waters.

3.2. Cluster analysis

The *R*-mode hierarchical cluster analysis was performed for the set of 75 samples and 12 variables. It yielded a dendrogram (Fig. 2), grouping all of the 12 descriptors into two statistically significant clusters.

From this dendrogram, one can find the relationship between different variables, the dendrogram shows a high correlation between major ions (Na^{2+}, SAR, EC, TDS, and Cl^-), which indicated the processes of marine intrusion. The second cluster shows the similarity between, nitrate, bicarbonate, and pH as one group, which probably indicate surface water recharge, water–rock interaction in addition of agriculture fertilizers.

Overall, the CA result confirms the PCA classification, but the second cluster corresponded to relatively non-marine origin variables.

3.3. Geostatistical estimation of groundwater quality

In order to highlight the importance of the groundwater salinization in Sfax coastal aquifer, the geostatistical method of cokriging is applied to take into account spatial dependence of ground water quality parameters.

In this method, the parameters of groundwater quality with the highest correlation coefficients in the correlation matrix were selected [23]. Hence, a significant correlation between TDS, Cl, Na, SO_4, and SAR is evidenced. However, Mg^{2+} shows weak correlations with SAR ($r = 0.43$). In such a way, only five groundwater quality variables, including, Cl, Na, TDS, SO4, are considered as primary variables and TDS as auxiliary variable for conducting cokriging.

Table 2
Correlation matrix of groundwater quality parameters of Sfax superficial aquifer

	EC	TDS	pH	Cl^-	NO_3^-	SO_4^{2-}	K^+	HCO_3^-	Na^+	Ca^{2+}	Mg^{2+}	SAR
EC	1											
TDS	0.87	1										
pH	−2.21	−1.15	1									
Cl^-	0.85	0.83	−0.02	1								
NO_3^-	0.11	0.12	−0.04	0.14	1							
SO_4^{2-}	0.65	0.79	−0.18	0.5		1						
K^+	0.4	0.52	0.04	0.44	−0.12	0.56	1					
HCO_3^-	−0.13	−0.18	−0.14	−0.17	0.03	−0.29	−0.35	1				
Na^+	0.85	0.9	−0.18	0.79	−0.03	0.73	0.44	−0.08	1			
Ca^{2+}	0.52	0.58	−0.03	0.59	0.35	0.55	0.44	−0.44	0.33	1		
Mg^{2+}	0.7	0.8	0.03	0.69	0.1	0.76	0.53	−0.13	0.65	0.51	1	
SAR	0.71	0.73	−0.23	0.6	−0.13	0.57	0.3	0.07	0.93	0.04	0.43	1

Table 3
Rotated factor pattern of three factors after varimax rotation

	Factor 1	Factor 2	Factor 3
Na^+	**0.96**	−0.1	-9.4×10^{-2}
TDS	**0.95**	-7.2×10^{-2}	0.16
EC	**0.91**	-5.2×10^{-2}	0.17
SAR	**0.84**	−0.31	−0.29
Cl^-	**0.84**	8.4×10^{-2}	0.23
SO_4^{2-}	**0.81**	0.24	5.72×10^{-2}
Mg^{2+}	**0.78**	0.24	0.17
HCO_3^-	−0.12	−0.77	-9.38×10^{-2}
K^+	0.55	0.56	−0.11
pH	−0.22	0.52	−0.11
NO_3^-	-10^{-2}	−0.16	0.89
Ca^{2+}	0.47	0.47	0.62
Eigen value	5.97	1.67	1.47
% total variance	49.75	13.96	12.25
Cumulative% variance	49.75	63.72	76

Bold values indicate strong loading (⩾0.75).

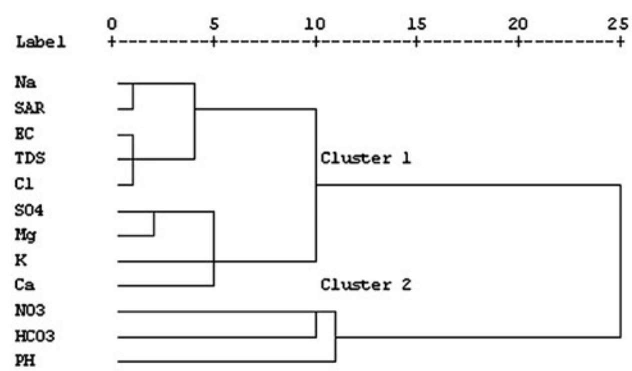

Fig. 2. Dendrogram generated from hydrochemical data showing relations between variables.

To study the cross-correlation effect between the five variables, the experimental variograms for each of the variables as well as the cross-variogram were computed. In total, variogram functions were estimated. In the co-regionalization analysis, the structure of spatial correlation was assumed to be independent from direction. Therefore, the omnidirectional variograms were used for analyses.

The next step consists of fitting to the experimental variograms and cross-variograms analytical models. The best-fit variogram models of these five groundwater quality variables were displayed in Fig. 3.

Three basic structures were used: (1) a nugget-effect; (2) an exponential model with a practical range of 20 km; (3) spherical models with ranges of 14, 14, and 16 km, respectively.

Fig. 3 shows the simple and cross-experimental semivariograms with the fitted models. The hulls, limiting values of model that would hold if correlation was perfect are also reported [34,40].

All direct semivariograms generally show some spatial structure and are bounded, but most of them have a large component of the nugget effect. Indeed, the ratio of nugget to total variation of the variograms models of SO_4, Cl, TDS, Na, and SAR was about 2.3–62%, indicating that the spatial correlation of the studied variables was moderately dependent, except in the case of SO_4. This situation indicates that the groundwater quality properties vary at a scale smaller than the minimum lag distance.

The experimental cross-variograms indicated generally a positive spatial cross-correlation, whereas the spatial correlation for the small distance was negative for the couples SO_4–Na, SO_4–TDS, and SO_4–SAR. The traditional simple product–moment correlation coefficient does not reveal the real relationship among the variables because it (1) averages out distinct changes in the correlation structures occurring at different spatial scales, and (2) includes the measurement errors inherent in the nugget effect as reported by Sollitto et al. [41].

In addition, most of the cross-variograms models are generally quite close to the upper dotted lines of maximum correlation, except in the cases of Cl–SAR and Cl–SO_4. The reduction of spatial cross-correlation is related to their low correlations Cl–SO_4 (0.5) and Cl–SAR (0.60).

The pattern of spatial variation in groundwater quality indices was assessed by interpolating the data

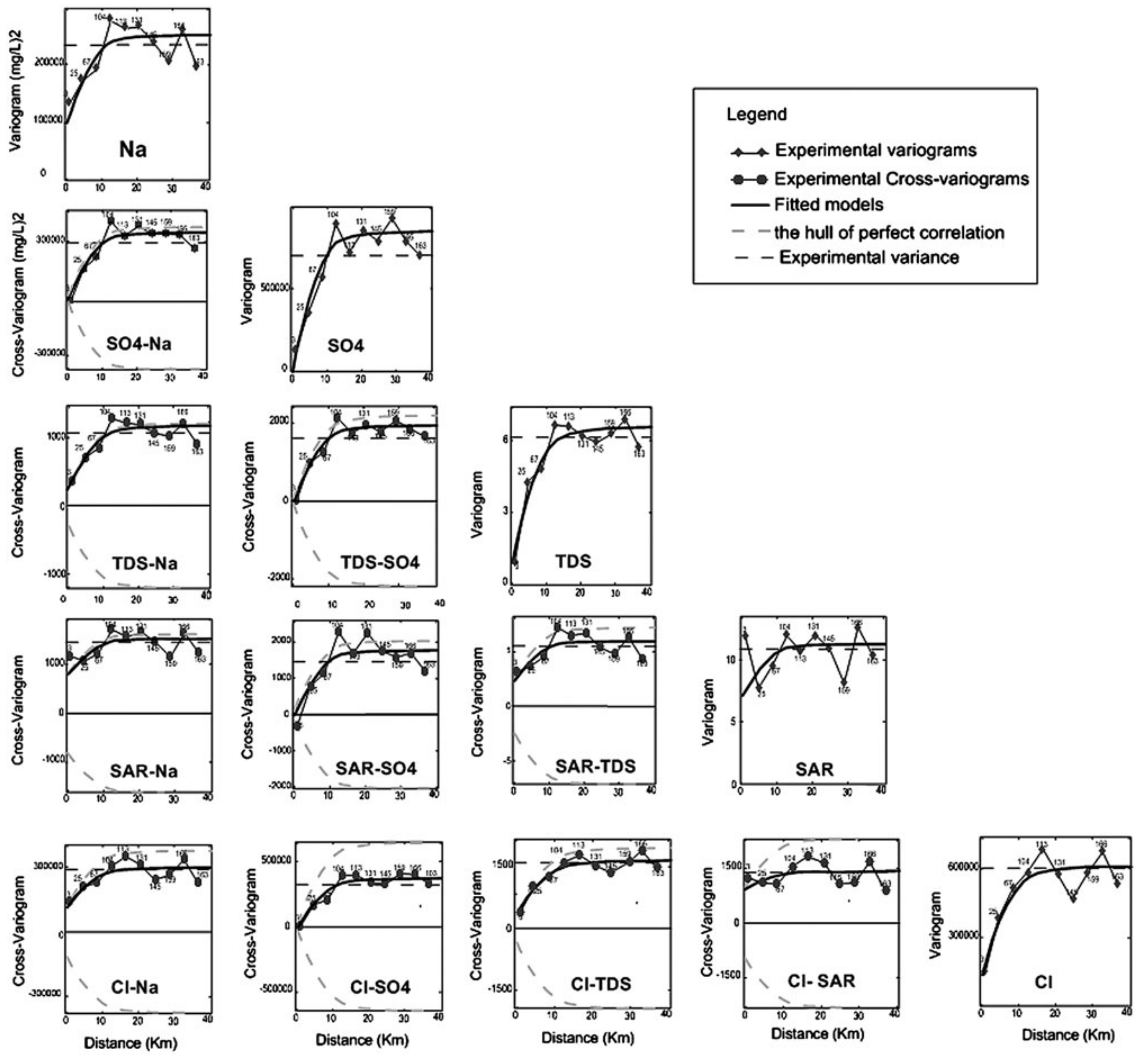

Fig. 3. The Simple and cross-variogram and their model fit for 5 groundwater quality parameters in the Sfax superficial aquifer.

by cokriging with a 2 km × 2 km cell grid, for a total of 3,600 grid nodes.

Cokriging estimates for Na (Fig. 4(a)), SO_4 (Fig. 4 (b)), TDS (Fig. 4(c)), SAR (Fig. 4(d)), and Cl (Fig. 4(e)), exhibit similar spatial patterns and reveal the zones at high risk of groundwater contamination. The higher values are indicated by darker colors, representing high salinization regions, and lower values are indicated by lighter colors, denoting low salinization regions. It is noteworthy that only the recharge area located essentially in the region of Bir Ali shows low salinity values. However, the concentrations of TDS, Cl, Na, SAR, and SO_4 exceeded the agricultural water quality recommendation set by Tunisian Water Code [42] in the coastal band of the study area, where the salinity locally exceeds 5 g/L and may reach 9 g/L particularly in the zone of Djebeniana (Fig. 4 (c)) due to seawater intrusion in groundwater [12].

3.4. Performance of predicted errors in cross-validation procedure

This study adopted a cross-validation procedure to evaluate the accuracy of the variogram and the

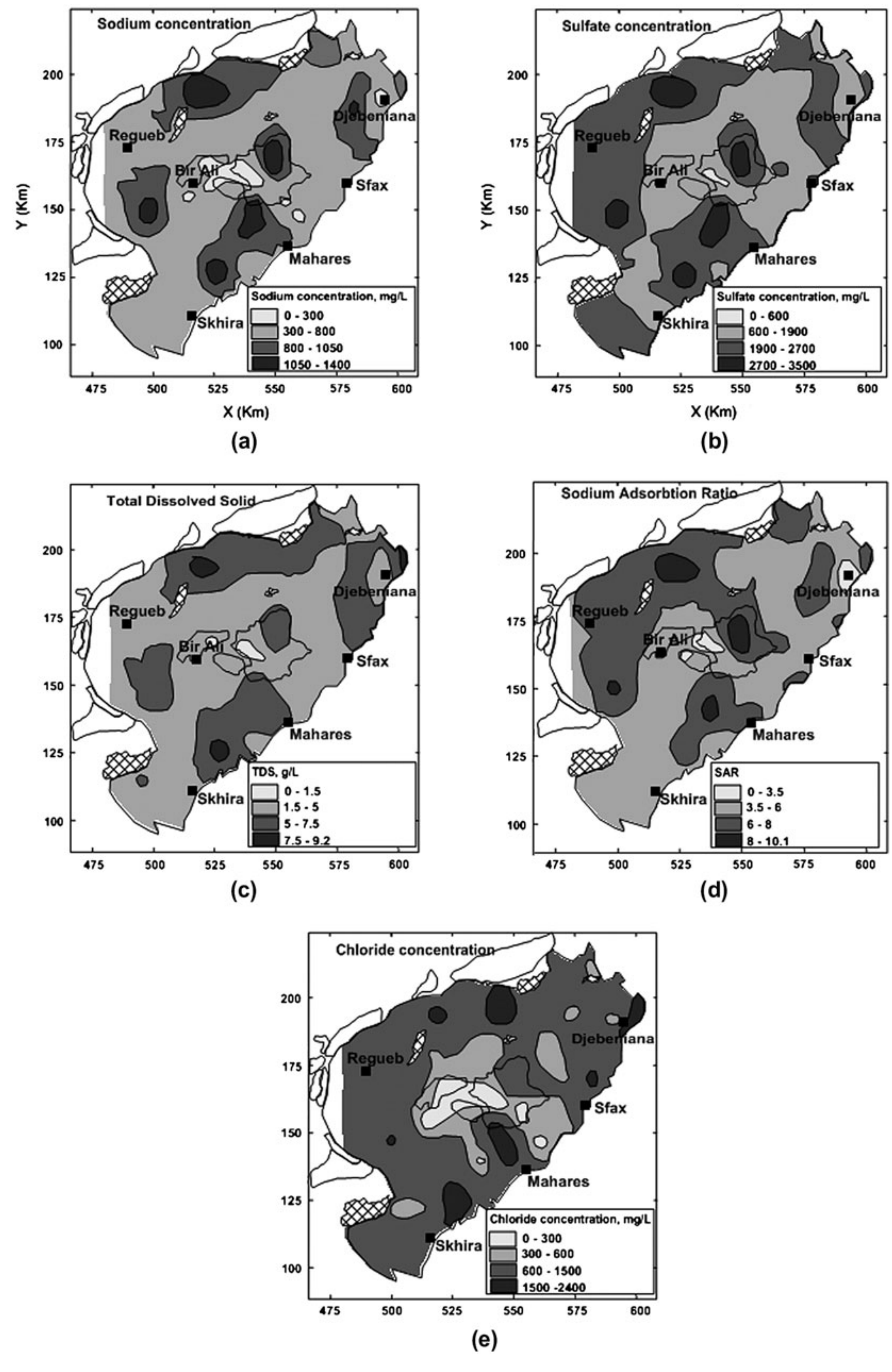

Fig. 4. Spatial variability maps of groundwater quality parameters generated by cokriging in the superficial aquifer of Sfax.

Table 4
Results of the cross-validation method applied for kriging and cokriging

Variable	Summary statistics of kriging				Summary statistics of cokriging			
	ME (mg/L)	MSE $(mg/L)^2$	MSSE	R^2	ME (mg/L)	MSE $(mg/L)^2$	MSSE	R^2
Na	−5.78	235,011.5	1.05	0.14	−2.3	5,504	1.01	0.98
SO_4	−10.96	611,378.8	0.91	0.36	2.3	138,463	1.1	0.89
Cl	−17.59	559,205.8	1.2	0.3	0.66	94,838.8	1.14	0.91
SAR	0.04	10.6	0.96	0.09	0.03	0.65	0.97	0.97
TDS	−0.09	4.49	1.2	0.42	0.01	0.51	0.97	0.95

ME: mean error; MSE: mean-squared error; MSSE: mean-square standard error; R^2: correlation coefficient between estimated and measured data.

cross-variogram models for kriging and cokriging. In this procedure, every known point is estimated using the values at the neighborhood around it, but not itself [20]. The mean error (ME), mean-squared error (MSE), and mean-square standard error (MSSE) were adopted to evaluate the performance of predicted errors in a cross-validation procedure for kriging and cokriging, and defined as follows [21].

$$\mathrm{ME} = \frac{1}{n}\sum_{i=1}^{n}(Z_i^* - Z_i) \approx 0 \tag{5}$$

$$\mathrm{MSSE} = \frac{1}{n}\sum_{i=1}^{n}(Z_i^* - Z_i)^2 \quad \text{minimum} \tag{6}$$

$$\mathrm{MSSE} = \frac{1}{n}\sum_{i=1}^{n}\left\{\frac{(Z_i^* - Z_i)^2}{\sigma_i^2}\right\} \approx 1 \tag{7}$$

where Z_i^* and Z_i are the estimated and observed values of the variable at the location i, and σ_i is the associated standard deviation of the estimation error.

Additionally, crossing plot of the estimate vs. the true value shows the correlation coefficient (R^2). The most appropriate variogram was chosen based on the highest correlation coefficient by trial-and-error procedure [20].

The results obtained through cross-validation (Table 4) show that cokrigng provides much better estimation results than kriging, because (1) the ME and MSE predicted by cokriging are smaller than ME and MSE predicted by kriging, (2) yhe MSSE predicted by cokriging is closer to one than MSSE predicted by kriging, and (3) the correlation coefficient (R^2) of estimated value associated with kriging and cokriging methods against the real groundwater quality for Na, SO_4, Cl, SAR, and TDS is closer to one for cokriging method than kriging one.

3.5. Conclusions

The purpose of this study has been to demonstrate the use of multivariate statistical analysis and cokriging technique for mapping groundwater quality contamination zones.

Principal components analysis and cluster analysis assisted the extraction and recognition of the factors responsible for groundwater quality variations. The ionic composition of the groundwater in the superficial aquifer of the Sfax was explained by three major factors in which Cl, Na, SO_4, SAR, and TDS represent the higher contribution to the first factor.

Considering the high correlation between the above mentioned parameters, the application of multivariate geostatistical technique appears entirely reasonable. Thus, cokriging is used to take into account spatial dependence between these parameters and to delineate the zones at high risk of groundwater contamination in coastal Sfax aquifer. Results showed that for the studied variables, estimation can be significantly improved using cokriging.

This work illustrates the feasibility of this approach to obtain a representative assessment of groundwater salinization risk, which forms a basis for sustainable management or improvement of the quality of groundwater in agricultural fields.

References

[1] B. Agoubi, A. Kharroubi, H. Abida, Saltwater intrusion modelling in Jorf coastal aquifer, southeastern Tunisia: Geochemical, geoelectrical and geostatistical application, Hydrol. Process. 27 (2012) 1191–1199.
[2] R. Reghunath, T.R. Sreedhara Murthy, B.R. Raghavan, The utility of multivariate statistical techniques in hydrogeochemical studies: An example from Karnataka. India, Water Res. 36 (2002) 2437–2442.
[3] A. Hooshmand, M. Delghandi, A. Izadi, K.A. Aali, Application of kriging and cokriging in spatial estimation of groundwater quality parameters, Afr. J. Agric. Res. 6 (2011) 3402–3408.

[4] P.J. Sajil Kumar, P. Jegathambal, E.J. James, Multivariate and geostatistical analysis of groundwater quality in palar river basin, Int. J. Geol. 4 (2011) 108–119.
[5] A.M. Subyani, M.E. Al Ahmadi, Multivariate statistical analysis of groundwater quality in wadi ranyah, Saudi Arabia, Earth Sci. 21 (2010) 29–46.
[6] A. Astel, M. Biziuk, A. Przyjazny and J. Namies´nik, Chemometrics in monitoring spatial and temporal variations in drinking water quality, Water Res. 40 (2006) 1706–1716.
[7] M.L. Wu, Y.S. Wang, C.C. Sun, H. Wang, J.D. Dong, J.P. Yin, S.H. Han, Identification of coastal water quality by statistical analysis methods in Daya Bay, South China Sea. Mar. Pollut. Bull. 60 (2010) 852–860.
[8] F. Sanchez-Martos, R. Jimenez-Espinosa, A. Pulido-Bosch, Mapping groundwater quality variables using PCA and geostatistics: A case study of Bajo Andarax, southeasten Spain, Hydro. Sci. J. 46 (2001) 227–242.
[9] F. Zhou, Y. Liu, H. Guo, Application of multivariate statistical methods to water quality assessment of the watercourses in northwestern new territories, Hong Kong, Environ. Monit. Assess. 132 (2007) 1–13.
[10] H. Arslan, Application of multivariate statistical techniques in the assessment of groundwater quality in seawater intrusion area in Bafra Plain, Turkey, Environ. Monit. Assess. 185 (2013) 2439–2452.
[11] A. Kharroubi, F. Tlahigue, B. Agoubi, C. Azri, S. Bouri, Hydrochemical and statistical studies of the groundwater salinization in mediterranean arid zones: Case of the Jerba coastal aquifer in southeast Tunisia, Environ. Earth Sci. 67 (2012) 2089–2100.
[12] R. Trabelsi, M. Zairi, H. Ben Dhia, Groundwater salinization of the Sfax superficiel aquifer, Tunisia, Hydrogeol. J. 15 (2007) 1341–1355.
[13] Y.C. Huang, C.P. Yang, Y.C. Lee, P.K. Tang, W.M. Hsu, Variation of groundwater quality in seawater intrusion area using cluster and multivariate factor analysis, ICNC 6 (2010) 3021–3025.
[14] M. Bahar, S. Reza, Hydrochemical characteristics and quality assessment of shallow groundwater in a coastal area of southwest Bangladesh, Environ. Earth Sci. 61 (2010) 1065–1073.
[15] M.S. Yeh, Y. Lin, L.C. Chang, Designing an optimal multivariate geostatistical groundwater quality monitoring network using factorial kriging and genetic algorithms, Environ. Geol. 50 (2006) 101–121.
[16] B. Agoubi, A. Kharroubi, H. Abida, Hydrochemistry of groundwater and its assessment for irrigation purpose in coastal Jeffara Aquifer, southeastern Tunisia, Arab. J. Geosci. 6 (2013) 1163–1172.
[17] P.P. Adhikary, C. Jyotiprava Dash, H. Chandrasekharan, T.B. S. Rajput, S.K. Dubey, Evaluation of groundwater quality for irrigation and drinking using GIS and geostatistics in a periurban area of Delhi, India, Arab. J. Geosci. 5 (2012) 1423–1434.
[18] B. Nas, Geostatistical approach to assessment of spatial distribution of groundwater quality, Polish. J. Environ. Stud. 18 (2009) 1073–1082.
[19] J.P. Dash, A. Sarangi, D.K. Singh, Spatial variability of groundwater depth and quality parameters in the national capital territory of Delhi, Environ. Manage. 45 (2010) 640–650.
[20] L. Pozdnyakova, R. Zhang, Geostatistical analyses of soil salinity in a large field, Precis. Agric. 1 (1999) 153–165.
[21] S. Ahmed, G. de Marsily, A. Talbot, Combined use of hydraulic and electrical properties of an aquifer in a geostatistical estimation of transmissivity, Groundwater 26 (1988) 78–86.
[22] M.N.M. Boezio, J.F.C.L. Costa, J.C. Koppe, Accounting for extensive secondary information to improve watertable mapping, Nat. Resour. Res. 15 (2006) 33–47.
[23] A. Castrignanò, G. Buttafuoco, C. Giasi, Assessment of groundwater salinisation risk using multivariate geostatistics, Geo. Env. VI 15 (2008) 191–202.
[24] F. Ben Jemaa, M.A. Marino, Effect of parameter cross-correlation on groundwater sampling design, IAHS-AISH P 202 (1991) 239–246.
[25] J.D. Istok, J.D. Smyth, A.L. Flint, Multivariate geostatistical analysis of groundwater contamination: A case history, Groundwater 31 (1993) 63–74.
[26] R.T. Mehrjardi, M.Z. Jahromi, S. Mahmodi, A. Heidari, Spatial distribution of groundwater quality with geostatistics (case study: Yazd-Ardakan plain), World Appl. Sci. J. 4 (2008) 09–17.
[27] C.H. Jang, Use of multivariate indicator kriging methods for assessing groundwater contamination extents for irrigation, Environ. Monit. Assess. 185 (2012) 4049–4061.
[28] I. Triki, M. Zairi, H. Ben Dhia, A geostatistical approach for groundwater head monitoring network optimisation: Case of the Sfax superficial aquifer (Tunisia), Water Environ. J. (2012). Available from: http://onlinelibrary.wiley.com/journal/10.1111/%28ISSN%291747-6593/earlyview
[29] B. Chulli, A. Davraz, J. Makni, M. Bedir, H. Ben Dhia, Hydrogeological investigations of thermal waters in the Sfax basin (Tunisia), Environ. Earth Sci. 66 (2012) 1–16.
[30] L.A. Richards, Diagnosis and Improvement of Saline and Alkali Soils; USDA Agriculture Handbook, Number 60, U.S. Department of Agriculture, Washington, DC, 1954.
[31] K.H. Kim, S.T. Yun, B.Y. Choi, G.T. Chae, Y. Joo, K. Kim, H.S. Kim, Hydrochemical and multivariate statistical interpretations of spatial controls of nitrate concentrations in a shallow alluvial aquifer around oxbow lakes (Osong area, central Korea), J. Contam. Hydrol. 107 (2009) 114–127.
[32] J.P. Chiles, P. Delfiner, Geostatistics: Modeling Spatial Uncertainty, Wiley, New York, 1999.
[33] X. Emery, Géostatistique Linéaire, Ecole des Mines de Paris, Paris, 2002.
[34] H. Wackernagel, Multivariate Geostatistics: An Introduction with Applications, Springer, Berlin, 2003.
[35] J. Rivoirard, Cours De Géostatistique Multivariable, Ecole des Mines de Paris, Paris, 2003.
[36] D.E. Myers, Multivariable geostatistical analysis for environmental monitoring, Sci. Terre 27 (1988) 411–427.
[37] G. Buttafuoco, A. Castrignano, A.C. Busoni E, Dimase Studying the spatial structure evolution of soil water content using multivariate geostatistics, J. Hydrol. 311 (2005) 202–218.
[38] E. Polus, N. Flipo, C. de Fouquet, M. Poulin, Geostatistics for assessing the efficiency of a distributed physically-based water quality model: Application to nitrate in the Seine River, Hydrol. Process. 25 (2011) 217–233.
[39] A.M. Subyani, A.M. Al-Dakheel, Multivariate geostatistical methods of mean annual and seasonal rainfall in southwest Saudi Arabia, Arab. J. Geosci. 2 (2009) 19–27.
[40] R. Webster, M.A. Oliver, Geostatistics For Environmental Scientists, Wiley, Chichester, 2007.
[41] D. Sollitto, M. Romic, A. Castrignanò, D. Romic, H. Bakic, Assessing heavy metal contamination in soils of the Zagreb region (Northwest Croatia) using multivariate geostatistics, Catena 80 (2010) 182–194.
[42] Tunisian Water Code, Official Journal of Tunisian Republic, 22, April 1975.