



# Groundwater salinity in a semi-arid region of central-eastern Tunisia: insights from multivariate statistical techniques and geostatistical modelling

S. M'nassri<sup>1</sup> · L. Dridi<sup>1</sup> · G. Schäfer<sup>2</sup> · M. Hachicha<sup>3</sup> · R. Majdoub<sup>1</sup>

Received: 5 July 2018 / Accepted: 12 April 2019 / Published online: 6 May 2019  
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

## Abstract

This study aims to investigate the hydrochemical behaviour of a shallow aquifer of a semi-arid environment in central-eastern Tunisia. It is based on a combined approach of multivariate statistical tools and a geostatistical analysis. The analysis of 49 groundwater samples underlined the heterogeneity between water classes. The factor analysis and the analysis of principal components highlighted that the deterioration of the groundwater quality results essentially from natural factors, such as rock weathering, and marginally from associated anthropogenic factors, in most cases, with artificial groundwater recharge from water irrigation at the soil surface. The natural factors are clearly associated with chloride, sodium and sulphate, and they are responsible for the water quality class with a salinity higher than 6 g/L. The anthropogenic factors are characterized by calcium and potassium ions. Based on the variograms obtained for the vector of both factors, their spatial interpolation by kriging showed that the high values of vector 1 (the natural factors) were recorded near the sabkha Sidi El Hani, indicating a rock-water interaction. However, vector 2 is high in the central part of the Ouled Chamekh plain, likely due to the water irrigation of agricultural soils.

**Keywords** Groundwater quality · Hydrochemical behaviour · Multivariate statistical methods · Variograms · Kriging · Tunisia

## Introduction

Over the last two decades, many regions worldwide have been severely affected by periods of extended drought, especially the African continent due to changes in climatic conditions. According to Nyenjie and Batelaan (2009), groundwater resources are the first to be affected by global warming. This environmental problem is evidenced through

the decline in groundwater level and the continuous deterioration of its quality (Kamel et al. 2010).

Located on the southern bank of the Mediterranean, Tunisia is a country with majority of the territory dominated by arid and semi-arid climate, where the water resources are scarce and heterogeneous in time and space (Majdoub et al. 2012). These characteristics seem to limit the socioeconomic development of the country, especially in the agricultural sector (Hamzaoui-Azaza et al. 2013; Hassen et al. 2016; Halimi et al. 2016). Furthermore, salinization has constituted a major factor responsible for the deterioration of water quality for several decades (Farid et al. 2013). It constitutes one of the negative effects of climate change on the availability of groundwater resources and severely threatens their quality. This has become a major concern for Tunisia in safeguarding these resources for future generations.

Various quantitative methods of hydrochemical groundwater data evaluation have recently been developed to provide decision-making tools that would aid the agencies responsible for the management of water resources (Farid

✉ G. Schäfer  
schafer@unistra.fr

<sup>1</sup> Département du Génie des Systèmes Horticoles et du Milieu Naturel, Université de Sousse, Institut Supérieur Agronomique de Chott Meriem, BP 42, 4042 Chott Meriem, Sousse, Tunisia

<sup>2</sup> Université de Strasbourg, CNRS, Laboratoire d'Hydrologie et de Géochimie de Strasbourg (LHyGeS) UMR 7517, 67000 Strasbourg, France

<sup>3</sup> Institut National de Recherches en Génie Rural, Eaux et Forêts, BP 10, 2080 Ariana, Tunisia

et al. 2013; Hajd Ammar et al. 2014; Bouzourra et al. 2015; Alssane et al. 2015; Fadili et al. 2016).

Multivariate methods are generally used for the hydrochemical analysis of groundwater samples. They provide a quantitative insight that can highlight relationships that are difficult to elucidate by the direct analysis of the data. Moreover, these methods have the advantage of providing a concise description of the main information resulting from the input data.

Several research studies have demonstrated the relevance of multivariate statistical and geostatistical approaches in groundwater data evaluation. For example, Jiang et al. (2009) and Belkhiri et al. (2010) used multivariate methods to analyse the variation in the hydrochemical composition of groundwater. In addition, Cloutier et al. (2008) combined statistical tools with hydrochemical methods to study independent factors controlling the quality of groundwater in Quebec. Furthermore, Yidana et al. (2012) used factor analysis to identify the origin of fluorides in a volcanic aquifer in Ghana. The results showed that the high concentrations of fluorides are from the dissolution of silicates. They highlighted a close relationship between the geology of the aquifer and water chemistry. Moreover, Kim et al. (2012) combined analyses of chemometrics and kriging for identifying groundwater contamination sources and origins at the Masan coastal area in Korea. They found that the main sources of contamination are seawater, nitrate and iron. Gbolo and Gerla (2013) used multivariate analysis to characterize nutrients from an abandoned feedlot and Mikhailov et al. (2007) used multivariate analysis as a tool for the ecological assessment and characterization of landfills. Venkatramanan et al. (2016) also used geostatistical techniques to evaluate groundwater contamination and its sources in Miryang City, Korea.

The unconfined aquifer of Sidi El Hani is situated between the sabkha Sidi El Hani and the sabkha Cherita in central-eastern Tunisia. It experiences multiple constraints due to the scarcity of surface water, which is limited only to the sabkhas and the wadi Cherita. Population increase and the increase in agriculture land uses in the area have had a negative impact on the groundwater quality (M'nassri et al. 2016). This present paper focuses on the hydrochemical investigation of the groundwater of Sidi El Hani aquifer to identify the main factors currently affecting water quality. Multivariate statistical methods including a hierarchical component analysis (HCA), a factor analysis (FA) and a principal component analysis (PCA) were used. The HCA was used to group the hydrochemical data, subsequently revealing what makes them homogeneous and thus leading to a better understanding of the salinity distribution. The FA and PCA enabled an examination of the independent factors responsible for the hydrochemical modification of the groundwater quality. Then, geostatistical modelling was

used to provide a comprehensive view of the spatial distribution of the salinity and to adequately summarize any structural information related to the zone of influence of a regionalized chemical variable.

## Materials and methods

### Study area

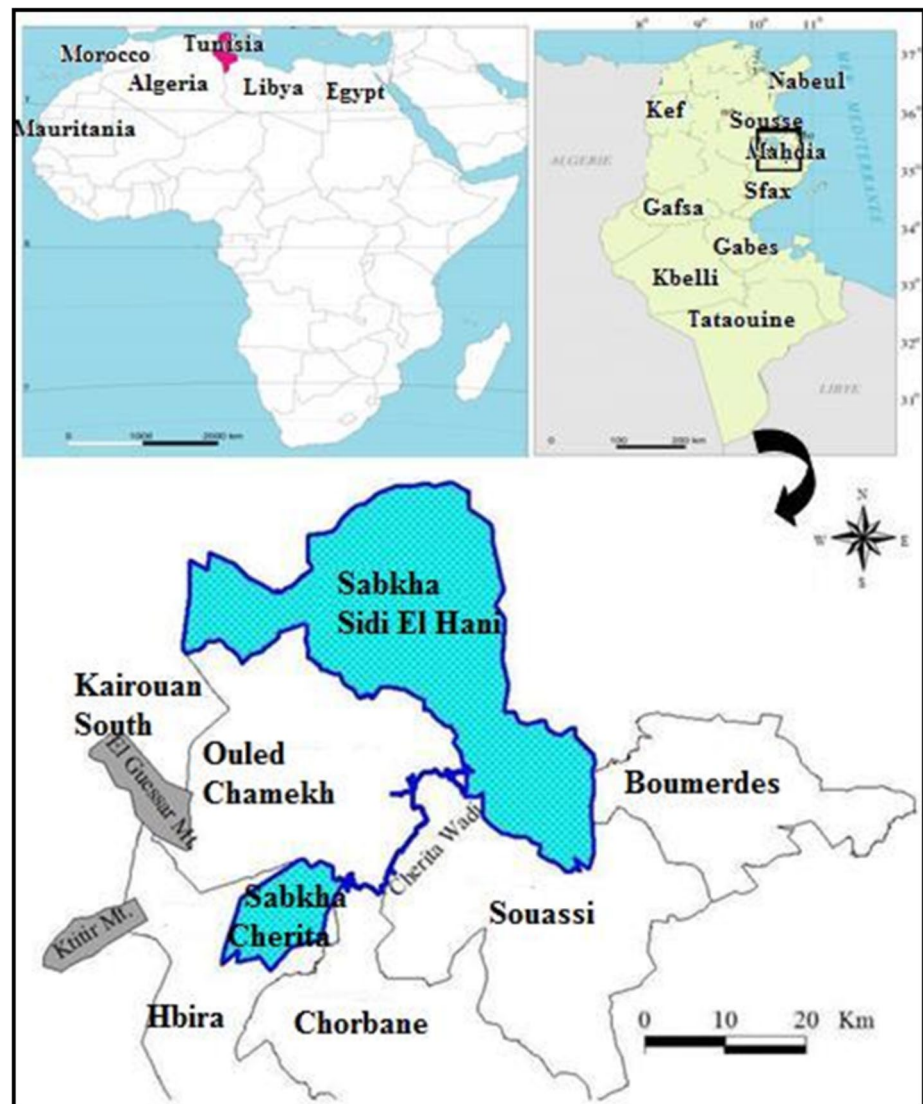
The unconfined aquifer of Sidi El Hani is located in the Ouled Chamekh Plain in central-eastern Tunisia (Fig. 1). This plain covers an area of 346 km<sup>2</sup> and belongs to the Mahdia region. The Ouled Chamekh plain is bordered by the sabkha of Sidi El Hani in the north and the northeast, by the wadi Cherita in the east, by the sabkha of Cherita and the Ktitir Mountain to the south, and by the El Guessat Mountain to the west. The sabkha Sidi El Hani is a north-west-southeast elongated depression with an average altitude of 33 m, covering an area of approximately 370 km<sup>2</sup>. This saline lake drains the sabkha Cherita, which covers an area of 70 km<sup>2</sup> through the hydrographic network of the wadi Cherita. The altitude of the Cherita depression is 60 m (Essefi et al. 2013). The climate of the area is semi-arid, characterized by an irregular and relatively weak annual rainfall of approximately 270 mm. The annual average temperature is 9 °C and 30 °C for the winter and summer, respectively. The potential evapotranspiration is estimated to be approximately 1500 mm/year.

### Geological and hydrogeological setting

According to Castany (1948) and Maamri and Rabhi (2003), the studied aquifer is formed by a basin filled with sediments of Mio-Plio-Quaternary age comprising the Holocene (Qh), Upper Pleistocene (QP3), Middle Pleistocene (Qp2) and Lower Pleistocene (Qp1), highlighted in Fig. 2. These formations are primarily composed of continental deposits of clay, marl, clayey sand, gypsum, silt and sandy clay. The study area is strongly controlled by the Ktitir, Hajeb Layoun and Sidi El Hani faults which control the Mio-Plio-Quaternary deposits (Khomsi et al. 2006). These faults have NW-SE and NNE-SSW to E-W orientations (Essefi et al. 2013).

The unconfined aquifer of Sidi El Hani mainly comprises sands, sandy clays and clayey sands. The direction of the groundwater flow is towards the sabkhas of Sidi El Hani and Cherita and the wadi Cherita (Fig. 3). The potentiometric surface map was interpolated using a Geographic Information System (GIS) software. The interpolation was performed using groundwater elevation data from 15 wells. This was done using the kriging method which allows a linear estimation with the minimum of variance. The water table

**Fig. 1** Location of the study area



is located at a depth of 5–30 m. The transmissivity varies between  $2 \times 10^{-4}$  and  $5 \times 10^{-3}$  m<sup>2</sup>/s, and the hydraulic gradients calculated in a regional flow model are on the order of 0.3–0.45%. These values were obtained from the previous numerical studies of the regional groundwater flow (Dridi et al. 2013).

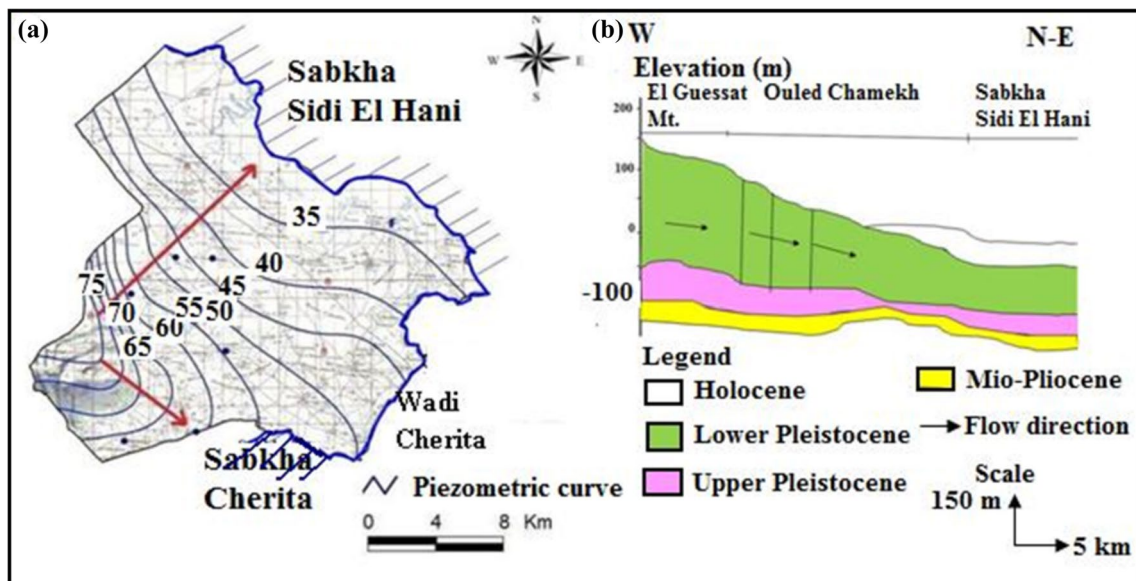
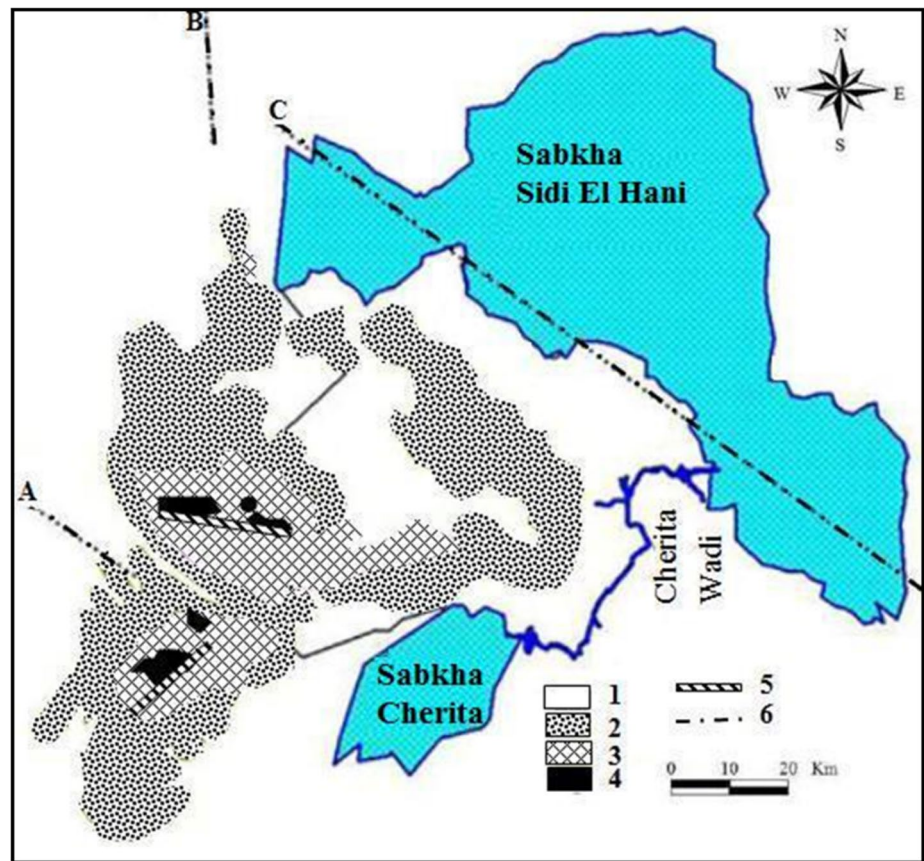
### Sample collection and analytical procedure

Groundwater sampling was performed during March and April 2015. Forty-nine samples were collected from shallow wells that have depths varying from 30 to 50 m (Fig. 4). Following the standard procedures introduced by Eaton et al. (1995), we carried out sampling and preservation of water samples as follows: All water samples were taken after pumping of the corresponding wells for 15–20 min to obtain representative values under ambient aquifer conditions. Each sample was collected in a new 1 L polyethylene bottle and

then stored at a temperature below 4 °C before analysis in the laboratory. The temperature ( $T$ ), electrical conductivity (EC) and hydrogen ion activity (pH) of each groundwater sample were measured in the field during the sampling activities. Concerning the pH measurements, the electrode was calibrated with a reference buffer solution of pH 4, 7 and 8 at each location.

The chemical analyses shown in Table 1 were accomplished in accordance with the standard methods of the American Association of Public Health (APHA 1995). Potassium ( $K^+$ ) and sodium ( $Na^+$ ) were measured by flame photometry (Rodier 1996). Sulfate ( $SO_4^{2-}$ ) concentrations were measured by the colorimetric spectrophotometry using  $BaCl_2$ . Chlorides ( $Cl^-$ ) were determined by standard  $AgNO_3$  titration (Rodier 1996). Bicarbonates ( $HCO_3^-$ ) were determined by titration with  $H_2SO_4$  (Rodier 1996), and calcium ( $Ca^{2+}$ ) and magnesium ( $Mg^{2+}$ ) concentrations were determined titrimetrically using standard ethylenediaminetetraacetic acid (Rodier 1996).

**Fig. 2** Geological summary map of the study area: 1 holocene; 2 upper pleistocene; 3 lower pleistocene; 4 mio-pliocene; 5 anticline axis; 6 fault; A, Ktitir Fault; B, Hajeb Layoun Fault; C, Sidi El Hani Fault (M'nassri et al. 2018)

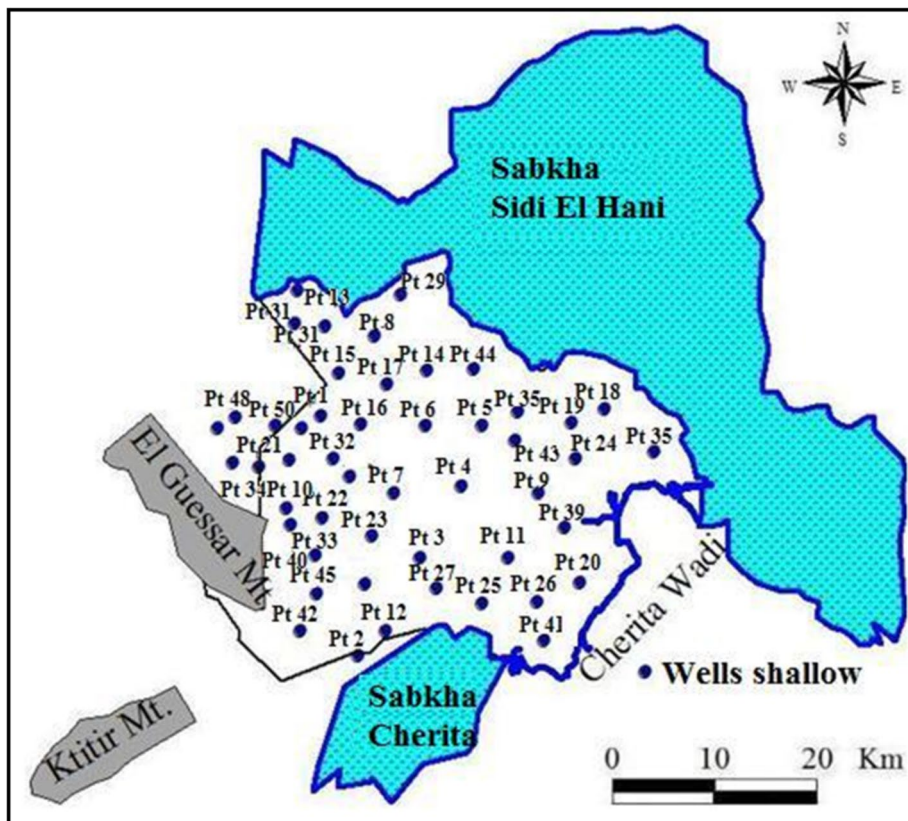


**Fig. 3** a Potentiometric surface map from (Dridi et al. 2013) (groundwater levels are in meters) and b hydrological cross-section of the Sidi El Hani aquifer from (M'nassri et al. 2019)

The accuracy of chemical analysis was carefully checked by the repeated analyses of samples and then by calculating the

percent charge balance error (CBE) (%) as defined by (Freeze and Cherry 1979):

**Fig. 4** Location of water samples



**Table 1** Analytical methods used for the analysis of chemical elements

Chemical elements	Analytical methods
Na <sup>+</sup> and K <sup>+</sup>	Flame photometry
SO <sub>4</sub> <sup>2-</sup>	Colorimetric photometry
Cl <sup>-</sup> and HCO <sub>3</sub> <sup>-</sup>	Titration
Ca <sup>2+</sup> and Mg <sup>2+</sup>	Titration
TDS	Drying and weighing

$$CBE = \left[ \frac{(\sum m_c - \sum m_a)}{(\sum m_c + \sum m_a)} \right] \times 100, \tag{1}$$

where  $m_c$  is the molality of cationic species and  $m_a$  is the molality of the anionic species.

The results of the analysis were judged to be acceptable when the CBE was less than or equal to  $\pm 5\%$ . No samples in the database had a CBE greater than this value.

TDS were measured by evaporating a pre-filtered sample to dryness. The accuracy of the chemical analysis was carefully studied using repeated analyses of three samples. At the end of the measurements, we chose the mean value of the repeated analyses.

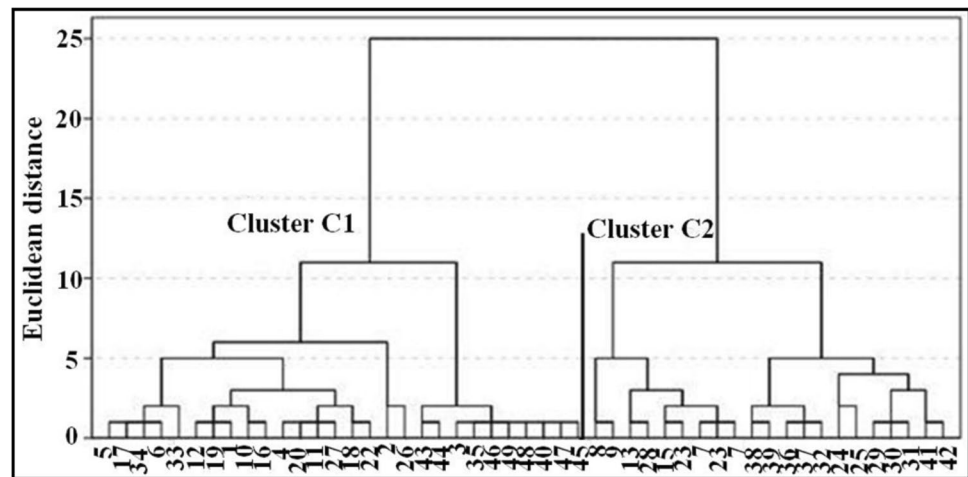
The analyses of total dissolved solids (TDS) and major elements such as Ca<sup>2+</sup>, Mg<sup>2+</sup>, Na<sup>+</sup>, K<sup>+</sup>, Cl<sup>-</sup>, SO<sub>4</sub><sup>2-</sup> and HCO<sub>3</sub><sup>-</sup> were conducted in the hydrological laboratory of the Higher Institute of Agronomy of Chott Mariem and in the laboratory of water analysis at the National Institute for Research in Rural Engineering, Water and Forest (Tunisia).

**Multivariate statistical analysis and geostatistical analysis**

In our study, the multivariate statistical analysis was based on HCA, FA and PCA. It was performed with the Statistical Package for the Social Sciences (SPSS, software version 18) and applied to a subgroup of the whole hydrochemical dataset of 49 samples and 11 parameters ( $T$ , pH, TDS, EC, Ca<sup>2+</sup>, Mg<sup>2+</sup>, Na<sup>+</sup>, K<sup>+</sup>, Cl<sup>-</sup>, HCO<sub>3</sub><sup>-</sup>, SO<sub>4</sub><sup>2-</sup>) using their similarities (Belkhiri et al. 2010; Hamzaoui-Azaza et al. 2009; Belkhiri and Mouni 2014; Bouzourra et al. 2015). This stage was preceded by a normalization of the data according to their  $z$  scores to achieve a normal distribution and homogeneity and to ensure that all of the parameters were close in terms of their variances (Guler et al. 2002; Cloutier et al. 2008; Kolsi-Hajji et al. 2013).

The hierarchical component analysis was performed to divide the dataset into hierarchical groups based on the

**Fig. 5** Hierarchical component analysis of the groundwater of Sidi El Hani



similarity between the samples and to explain the causes of the hydrochemical variation from one location to another. According to Alberto et al. (2001) and Yidana et al. (2010), this classification is known as Q-mode classification, in which the classification plot using Euclidean distances (straight line distance between two points in  $c$ -dimensional space defined by  $c$  variables). The Ward agglomeration method was used as a linkage criterion in this analysis (Ward 1963). This combination was selected to reveal the most unique sample association. Indeed, the most similar parameters were clustered together in the same group (Fig. 5).

A factor analysis was applied to study the variance observed in the dataset. It derives a subset of uncorrelated variables called factors (Cloutier et al. 2008; Kolsi-Hajji et al. 2013). The adequacy of the factor solution was verified by the Kaiser–Meyer–Olkin (KMO) index and the Bartlett sphericity test (Kaiser 1960). If the KMO is greater than 0.5, the factor analysis is appropriate to provide a significant reduction in data size. The total numbers of factors generated from the factor analysis that have an eigenvalue greater than 1 can establish a pattern of variation among the variables and reduce the large dataset into factors for easy interpretation (Jiang et al. 2009; Yidana et al. 2012). The first factor, which has the highest eigenvalue, represents the most important source of variation in the dataset, whereas the last factor represents the least important process that contributes to the chemical variation (Maoui et al. 2009; Hachicha et al. 2008).

The principal component analysis (PCA) is far more commonly used than factor analysis (FA) and permits also to obtain the component which reflects both common and unique variance of the variables. PCA may be seen as a variance-focused approach that reproduces both the total variable variance with all components as well as the correlations.

It was applied in our study to load on the factor and to interpret the correlation coefficients between the variables and the factor using the maximum variance of the factor after rotation (varimax rotation) to maximize the variation

among the variables for each factor. Note that “varimax rotation” is an application of an orthogonal matrix to the factor matrix (Sachez-Martoz et al. 2001; Kim et al. 2009). As a convention, variables with significant contributions to the final factor model should not be less than 50% (Belkhiri et al. 2010). The factors selected in the previous stage are then laid out in vectors. They are the image of the principal processes responsible for the variation of the hydrochemical composition of the groundwater. Each vector is then modelled in the geostatistical stage by a characteristic vector (Sachez-Martoz et al. 2001; Kim et al. 2009) using the program GS+ version 7 (Yidana et al. 2010). The geostatistical analysis establishes the spatial distribution of the selected vectors by computing their experimental isotropic variograms (Louati et al. 2015). The experimental variograms  $\gamma(h)$  were fitted by the Gaussian model expressed by Eq. 2. Then, a cross-validation test based on the method of ordinary kriging (OK) (Goovaerts 1997) was performed to detect both the reliability of the adopted model and the reliability of the spatially interpolated data. In the last stage, kriging maps were produced to clarify the possible distribution factors:

$$\gamma(h) = C_0 + C \left[ 1 - e^{-\left(\frac{h}{A}\right)^2} \right] \quad (2)$$

where  $C_0$  is the nugget effect ( $m^2$ ),  $(C_0 + C)$  ( $m^2$ ) is the sill of the variogram,  $A$  ( $m$ ) is the range and  $h$  ( $m$ ) is the distance between two sampling points.

## Results and discussion

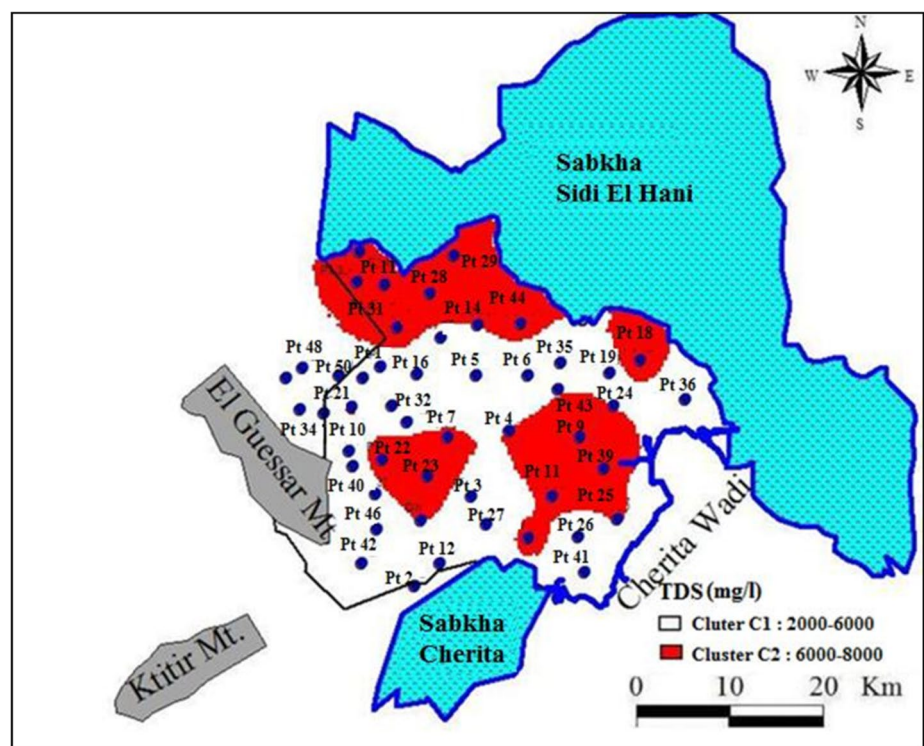
### Hierarchical component analysis

The dendrogram obtained from the HCA revealed the existence of two classes of groundwater in the unconfined aquifer of Sidi El Hani (Fig. 5). The first class, C1, encompassed

**Table 2** Statistical analysis of hydrochemical parameters of the clusters

Parameters	C1				C2			
	Min	Max	Mean	SD	Min	Max	Mean	SD
T (°C)	21.0	29.0	24.3	1.8	18.0	28.0	23.9	2.5
TDS (mg/L)	2400.0	6620.0	3880.7	976.0	2500.0	8413.3	5809.2	1353.9
EC (mS/cm)	3.9	8.2	5.4	1.2	3.8	12.0	7.8	1.9
pH	6.9	8.2	7.6	0.4	7.0	8.5	7.8	0.3
Ca (mg/L)	105.0	540.0	346.6	101.9	180.0	700.0	368.0	125.3
Mg (mg/L)	79.0	303.8	173.9	57.9	60.8	461.7	302.6	105.1
Na (mg/L)	410.9	1152.8	700.0	169.4	589.7	1461.9	1078.9	235.6
K (mg/L)	14.0	70.4	27.7	13.7	39.4	19.6	28.8	5.0
HCO <sub>3</sub> (mg/L)	457.5	1220.0	300.4	116.0	488.0	152.5	294.8	104.0
Cl (mg/L)	726.3	2130.0	1194.5	296.0	1088.1	2666.8	1852.2	365.3
SO <sub>4</sub> (mg/L)	3404.2	1275.0	2325.0	536.0	1920.0	5472.0	3526.3	848.3

**Fig. 6** Spatial interpolation of measured salinity

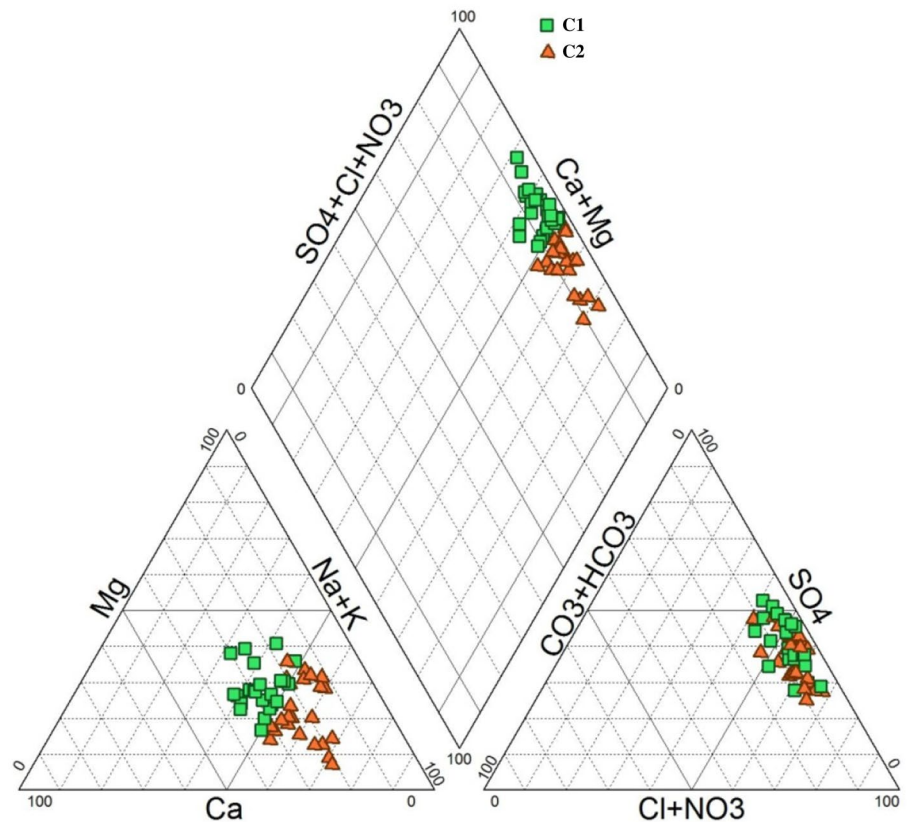


53% of the water samples. It was characterized by a fairly high salinity ranging, between 2400 and 6620 mg/L, with an average and a standard deviation of approximately 3881 and 976 mg/L, respectively (Table 2). The results of the chemical analyses of the major elements show the predominance of chloride, sodium and sulphate. They varied from 726 to 2130 mg/L, from 411 to 1153 mg/L and from 3404 to 1275 mg/L, respectively, whereas potassium was almost non-existent in the groundwater. The second class, C2, represented 47% of the total samples with a high salinity that exceeded 6000 mg/L. As noted for the C1 class, C2 was marked by a predominance of chloride, sodium and sulphate but with higher concentrations. The sodium concentrations

varied between 590 and 1462 mg/L with an average of 1079 mg/L, and the sulphate contents ranged between 1920 and 5472 mg/L. However, the chloride concentrations varied between 1088 and 2667 mg/L with an average and a standard deviation of approximately 1852 and 365 mg/L, respectively.

According to the classes identified by HCA, the spatial distribution of the salinity shows that high salinity was recorded in the neighbourhoods of the sabkha Sidi El Hani and the wadi Cherita (Fig. 6). The presence of evaporites in this area, including gypsum, anhydrite and halite, may be directly linked to the higher salinity in these samples compared to the samples located in the western part. This can be explained by the presence of permeable limestone that

**Fig. 7** Piper diagram of Sidi El Hani groundwater



**Table 3** KMO index and Bartlett's test of the factor analysis

KMO index	0.63
Bartlett's sphericity test	
Approximate Chi square	382.3
Test definition	55
Significance level	0.00

allows a high infiltration rate of rainwater into the groundwater, and thus contributes to the dilution of dissolved salts (Bekkoussa et al. 2013).

Furthermore, it is noteworthy that the water table near the mountain of El Guessat is rather deep, which may explain why the groundwater is protected against anthropogenic contamination. Represented on the Piper diagram (Piper 1944), the water of class C1 has mixed facies of Cl–SO<sub>4</sub>–Na–Ca. However, C2 is characterized by a sodium chloride facies (Fig. 7). The two types suggest the presence of process of dissolving minerals, particularly halides and sulphated rocks and/or contamination of the water by anthropogenic activities (Amadou et al. 2014; Najib et al. 2017).

### Factor analysis and principal component analysis

The FA was applied to 49 samples and 11 variables, including T, pH, EC, TDS, Ca, Mg, Na, K, Cl, HCO<sub>3</sub> and SO<sub>4</sub>. It should deform the least possible initial configuration of the individuals and variables. This was confirmed by an index of KMO higher than 0.5 and by a Bartlett sphericity test with a value lower than 0.001 (Table 3). The eigenvalue is the variance explained by a factor, which indicates the variance it captured; the higher the value, the more variance it has captured (Awadallah and Yousry 2012).

As a result, two factors with eigenvalues greater than 1 were extracted from the “varimax-rotation”. The first factor, F1, expressed more than 42.6% of the total variance, whereas the second factor, F2, had a variance of 15.9% (Table 4). In this study, the factor loading was classified as moderate because the value of cumulative variance varied between 0.5 and 0.75% (Belkhiri and Narany 2015).

The first factor, shown in Table 5, represents the most important factor controlling the evolution of the groundwater. It was strongly correlated with TDS, EC, Na, Cl, SO<sub>4</sub> and Mg as follows: TDS (0.91), EC (0.89), Cl (0.86), Na (0.84), SO<sub>4</sub> (0.69) and Mg (0.67). This confirms the regrouping of these elements with the electric conductivity on the axis F1 (Fig. 8). On the other hand, the second factor (F2) was also correlated with Ca (0.81) and K (0.62).



**Table 4** Percentage of variance of selected factors (using the principal component analysis extraction method)

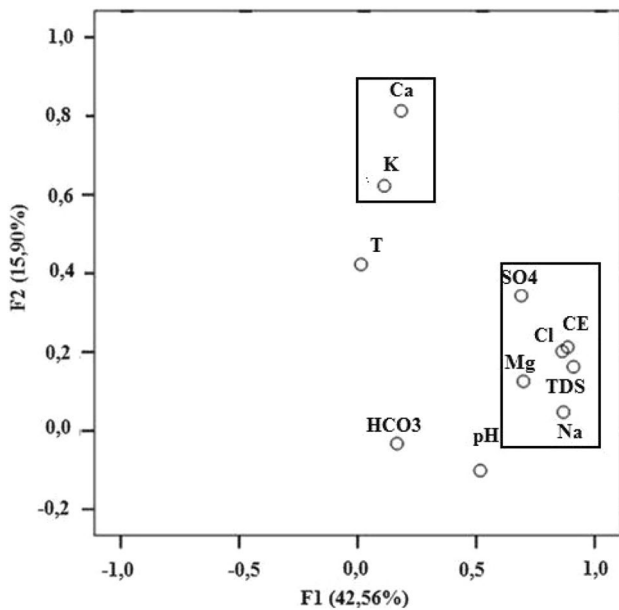
Selected factor	Without rotation			Rotation of factor		
	Eigen value	% Total of variance	% Of cumulative variance	Eigen value	% Total of variance	% Of cumulative variance
F1	4.68	42.6	42.6	4.33	39.38	39.38
F2	1.75	15.9	58.47	1.74	15.87	55.26

**Table 5** Factor analysis of the physicochemical parameters

Parameters	Component	
	F1	F2
T	0.014	0.423
TDS	0.910	0.162
EC	0.886	0.212
pH	0.517	-0.102
Ca <sup>2+</sup>	0.183	0.813
Mg <sup>2+</sup>	0.671	0.125
Na <sup>+</sup>	0.841	0.062
K <sup>+</sup>	0.112	0.622
HCO <sub>3</sub> <sup>-</sup>	0.165	-0.033
Cl <sup>-</sup>	0.863	0.202
SO <sub>4</sub> <sup>2-</sup>	0.690	0.343

of the phenomena responsible for the release of these ions in groundwater. On the other hand, the dissolved salts are related primarily to sodium (0.774), chloride (0.756) and sulphate (0.621). These elements are thus the most relevant variables defining the salinization of the groundwater (Kharroubi et al. 2012). In fact, the correlation between Ca and SO<sub>4</sub> may be related to the dissolution of gypsum. However, the high correlation between the Na and Cl may be derived from the dissolution of halite.

Factor F1 can be considered as a possible path for natural mineralization (geology and evaporation). It appears to be the most important mechanisms in the formation process of ions in water: the majority of the variables were regrouped with the electric conductivity around axis F1 (see Fig. 8), indicating the influence of the water–rock interaction on natural mineralization. F1 thus represents the natural factor derived by evaporite and carbonate dissolution. It suggests that the natural process is the most important mechanism controlling the ions in the given groundwater. These elements originate from the long-lasting contact of water within the aquifer, either in the water-saturated zone or during the crossing of the unsaturated zone. This mechanism is highlighted by both the rock weathering over time and the precipitation of ions resulting from the dissolution of evaporites, particularly halite, gypsum and anhydrite, and carbonates, such as calcite and dolomite (Trabelsi et al. 2007; Fadili et al. 2016). As shown in the study of M'nassri et al. (2019), the contamination of the Sidi El Hani groundwater is due to dissolution of halite, cation exchange, and precipitation of carbonate minerals such as calcite and dolomite coupled with the dissolution of gypsum, and evaporation. The correlation of F2 with calcium and potassium suggests that F2 is mainly related to anthropogenic activities. The calcium and potassium likely originate from the infiltration of surface waters such as rain and irrigation water, which dissolve and leach the salts concentrated in the soil or in the unsaturated zone (Yermani et al. 2003; Bekkoussa et al. 2013). In addition, the Gibbs diagram (Fig. 9) underlines that evaporation is the primary mechanism responsible for the increase of concentration of all species present in the groundwater. Furthermore, evaporation leads to precipitation of the least soluble mineral. Hence, an amount of salt could accumulate on the topsoil. This salt may then be

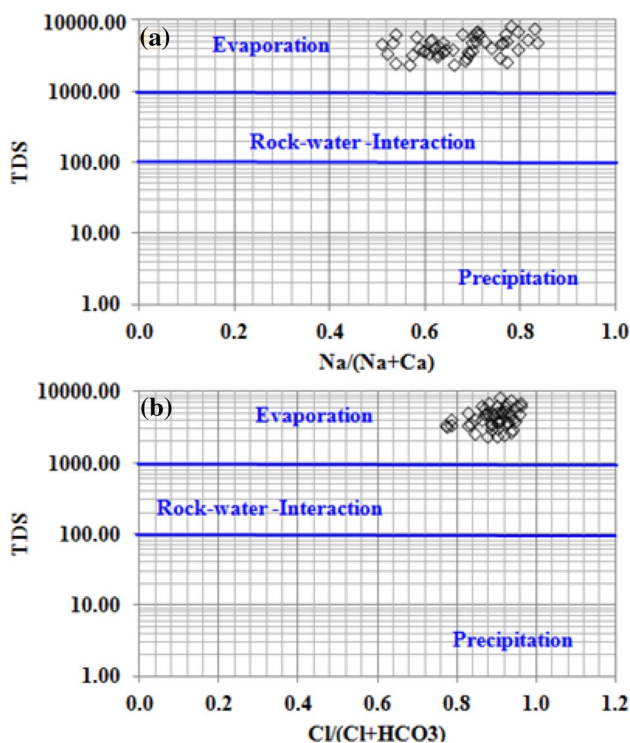


**Fig. 8** Principal component analysis: scatter plot of F1 versus F2

The correlation matrix given in Table 6 shows significant correlation between the major elements. The chlorides are strongly correlated with sodium (0.799) and magnesium (0.665). The sulphates have a correlation with calcium (0.577). These correlations indicate the similarity

**Table 6** Correlation matrix of the physicochemical parameters

	T	TDS	EC	pH	Ca <sup>2+</sup>	Mg <sup>2+</sup>	Na <sup>+</sup>	K <sup>+</sup>	HCO <sub>3</sub> <sup>-</sup>	Cl <sup>-</sup>	SO <sub>4</sub> <sup>2-</sup>
T	1.000										
TDS	0.173	1.000									
EC	0.218	0.932	1.000								
pH	0.126	0.389	0.329	1.000							
Ca <sup>2+</sup>	0.181	0.238	0.289	0.092	1.000						
Mg <sup>2+</sup>	-0.141	0.566	0.478	0.264	0.250	1.000					
Na <sup>+</sup>	0.138	0.729	0.774	0.362	0.156	0.286	1.000				
K <sup>+</sup>	0.212	0.208	0.215	0.126	0.217	0.081	0.207	1.000			
HCO <sub>3</sub> <sup>-</sup>	0.383	0.158	0.210	0.297	0.090	-0.146	0.262	0.023	1.000		
Cl <sup>-</sup>	0.106	0.745	0.756	0.335	0.254	0.665	0.797	0.330	0.117	1.000	
SO <sub>4</sub> <sup>2-</sup>	-0.142	0.654	0.621	0.215	0.537	0.597	0.545	0.056	-0.207	0.505	1.000



**Fig. 9** Gibbs plot of water samples of the Sidi El Hani basin: total dissolved solids (TDS) as a function of weight ratios of **a** Na/(Na + Ca) and **b** (Cl/Cl + HCO<sub>3</sub>) (M'nassri et al. 2018)

**Table 7** Variogram model parameters obtained for V1 and V2 and the correlation coefficients of the validity of the Gaussian model

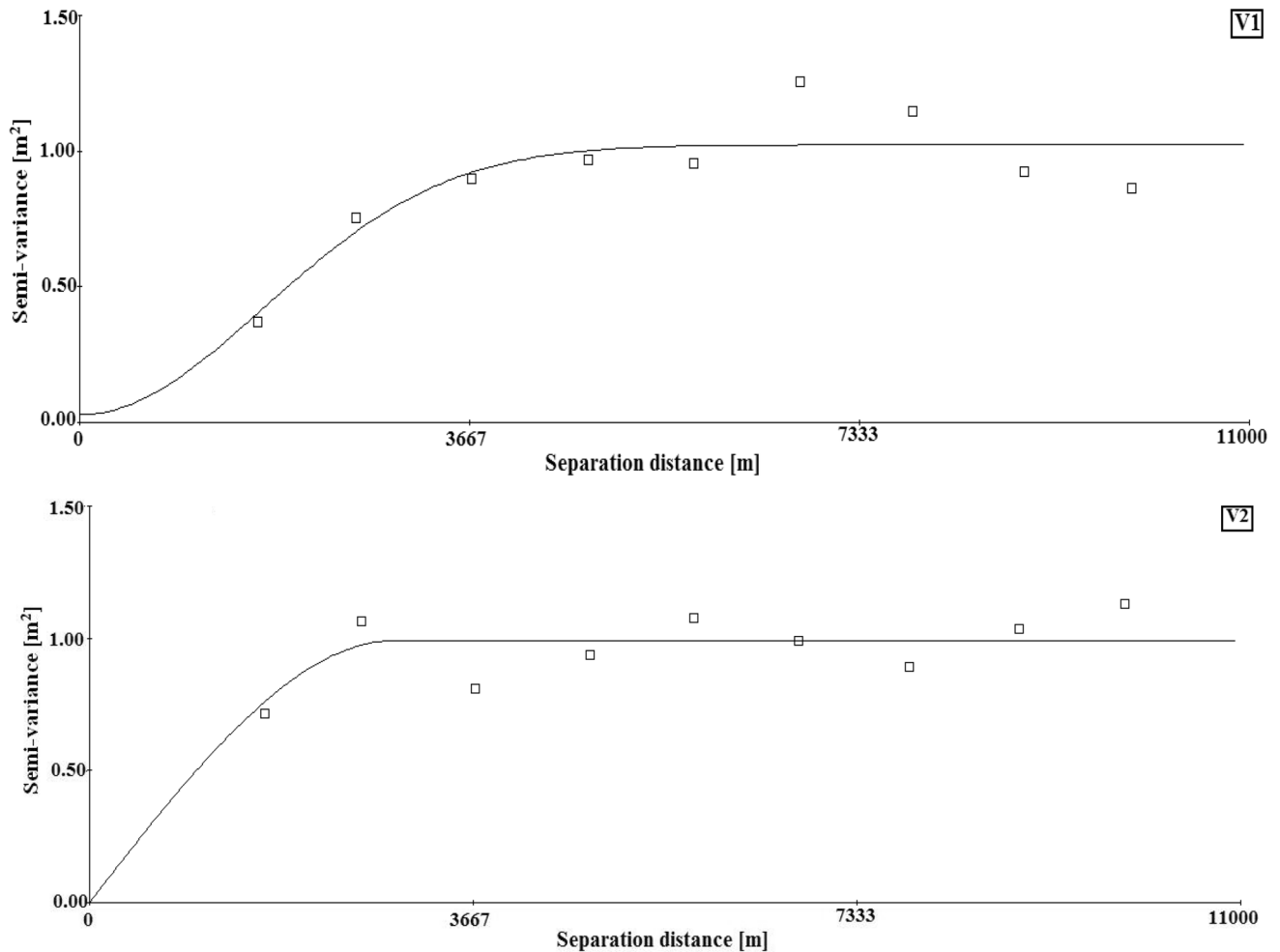
Vector	Nugget effect (C <sub>0</sub> ) (m <sup>2</sup> )	Sill (C <sub>0</sub> +C) (m <sup>2</sup> )	(C <sub>0</sub> +C)/C (-)	Coefficient of determination r <sup>2</sup>
V1	0.036	1.003	0.495	0.75
V2	0.220	0.880	0.600	0.62

leached under rainfall or irrigation and then transferred to deep soil layers (M'nassri et al. 2018).

### Geostatistical modelling of factors F1 and F2

Semivariograms computed from the OK model illustrated the spatial distribution vector V1 and vector V2 associated with F1 and F2, respectively. The best-fit semivariogram model parameters are shown in Table 7. They are characterized by a nugget effect, in which the model intersects the y-axis at 0.036 and 0.22 for V1 and V2, respectively. Moreover, coefficients of determination  $r_1^2 = 0.75$  and  $r_2^2 = 0.62$  were obtained for V1 and V2, respectively. Additionally, the normalized sill [(C<sub>0</sub> + C)/C] presents 49.5% and 60% for V1 and V2, respectively. These values are lower than 75%, indicating a moderate spacing between the sampling points (Belkhiri and Narany 2015). Therefore, the semivariogram of the studied parameters is highlighted in Fig. 10, in which the black lines represent the empirical semivariances and the scatter point symbolize the experimental semivariances.

Subsequently, the mapping of V1 and V2 was applied by means of the point kriging tools. The probability map of V1, which correlates with the concentrations of Na, Cl and SO<sub>4</sub>, reveals that the highest concentrations were recorded near the sabkha Sidi El Hani. This may be due to the dissolution of salt rocks (such as halite and gypsum) combined with the up-coning of salted waters trapped in the sedimentary series in the catchment area of the wells. In addition, this zone was characterized by a rather low hydraulic gradient corresponding to a very slow groundwater flow that may cause an increase in the concentration of dissolved salts. However, the obtained map of V2 resulting from the presence of dissolved calcium and potassium in the groundwater shows that the highest value was recorded in the central plain of Ouled Chamekh and in the neighbourhood of the sabkha Cherita. Thus, the V2



**Fig. 10** Experimental isotropic variogram of vectors V1 and V2

vector is, mainly, associated with activities of the agricultural sector, such as the use of fertilizers.

## Conclusions

A combined approach of multivariate statistical methods (HCA, FA, and PCA) and geostatistical modelling was used to characterize the hydrochemical state of the groundwater of Sidi El Hani and to identify the main factors responsible for its evolution. The results based on HCA highlight two classes of water samples: C1 and C2. The C1 class, composed of 53% of the samples, is characterized by a fairly high salinity and a Cl–Ca–Mg–SO<sub>4</sub> facies. However, the C2 class gathers 47% of water points where the salinity exceeds 6 g/L. It represents sodic chlorinated facies.

The results of the FA and PCA factorial analysis (AF) demonstrated that the hydrochemical evolution of the groundwater resulted from two independent factors. The first factor (F1) was represented primarily by chloride (Cl<sup>-</sup>),

sodium (Na<sup>+</sup>), sulphate (SO<sub>4</sub><sup>2-</sup>) and magnesium (Mg<sup>2+</sup>). However, the second factor (F2) was characterized by calcium (Ca<sup>2+</sup>) and potassium (K<sup>+</sup>). The chart in the factorial space showed that F1 was associated with the water samples having a strong electric conductivity for which the ionic acquisition was controlled by natural factors such as rock weathering. On the other hand, F2 gathered the water samples for which the salinity was controlled by anthropogenic factors such as the return of irrigation water charged by the fertilizers used in agriculture.

The geostatistical modelling gave a comprehensive view of the influence of vectors V1 and V2 associated with the two identified factors (F1) and (F2), respectively. The mapping of the variables shows that two major areas can be defined on a regional level. It can be concluded that the chemical evolution of the sampling points close to the sabkha of the Sidi El Hani is strongly controlled by natural factors (originating from evaporite materials, mainly from the Holocene). The rest of the plain is under the influence of anthropogenic factors related mainly to the intensive use

of fertilizers. However, the heterogeneity of salinization observed in the groundwater of Sidi El Hani underlines the difficulty of identifying the origin of the major elements present.

The combination of statistical and geostatistical tools seems to produce a better picture of the main factors influencing the groundwater quality than former studies, but it does not provide detailed information on the physicochemical processes involved. To further specify the origin of each chemical element and the current state of its flux in the aquifer, numerical studies of reactive transport should be performed to supplement the knowledge of the functioning of the aquifer and to provide answers concerning the transport processes involved. Thus, the statistical and geostatistical study of groundwater samples contributes to preventing the salinization of groundwater and improving water quality.

**Acknowledgements** This work is part of a research project entitled “salinization of groundwater of the Sidi El Hani aquifer.” It was funded by Institution of Research and Higher agricultural Education (IRESA).

## References

- Alberto WD, Del Pilar DM, Valeria AM, Fabiana PS, Cecilia HA, Los Angeles BM (2001) Pattern Recognition techniques for the evaluation of spatial and temporal variation in water quality A case study: Suquia river basin (Cordoba-Argentina). *Water Resour J* 35(12):2881–2894
- Allsane A, Trabelsi R, Dovonon L, Odeloui D, Boukari M, Zouari K, Mama D (2015) Chemical evolution of the continental shallow aquifer in the South of coastal sedimentary basin of Benin (West-Africa) using multivariate factor analysis. *J Water Resour Prot* 7:496–515
- Amadou H, Louati M, Manzola AS (2014) Caractérisation hydrochimique des eaux souterraines de la région de Tahoua (Niger). *J Appl Biosci* 80:7161–7172
- APHA (1995) Standard method for the examination of water and wastewater, 19th edn. American Public Health Association, Washington, DC, p 500
- Awadallah A, Yousry M (2012) Identifying homogeneous water quality regions in the Nile River using multivariate statistical analysis. *Water Resour Manag* 26(7):2039–2055
- Bekkoussa B, Jourde H, Batiot Ch, Meddi M, Khaldi A, Azzaz H (2013) Origine de la salinité et des principaux éléments majeurs des eaux de la nappe phréatique de la plaine de Ghriss, Nord-Ouest algérien. *Hydrol Sci J* 58(5):1111–1127
- Belkhir L, Mouni L (2014) Geochemical characterization of surface water and groundwater in Soummam Basin, Algeria. *Nat Resour Res* 23(4):393–407
- Belkhir L, Narany TS (2015) Using multivariate statistical analysis, geostatistical techniques and structural equation modeling to identify spatial variability of groundwater quality. *Water Resour Manag J* 29(6):2073–2089
- Belkhir L, Boudouhka A, Mouni L, Baouz T (2010) Application of multivariate statistical methods and inverse geochemical modeling for characterization of groundwater of groundwater—a case study: Ain Azel plain (Algeria). *Geoderma* 159:390–398
- Bouzourra H, Bouhlila R, Elango L, Salma F, Ouslati N (2015) Characterization of mechanisms and processes of groundwater salinization in irrigated coastal area using statistics, GIS, and hydrochemical investigations. *Environ Sci Pollut Res J* 22(4):2643–2660
- Castany G (1948) Les fosses d’effondrement de Tunisie – Géologie et Hydrogéologie. Plaine de Grombalia et cuvettes de la Tunisie Orientale. *Annales des mines et de la géologie* 3:126
- Cloutier C, Lefebvre R, Therrien R, Savad M (2008) Multivariate statistical analysis geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. *J Hydrol* 353:294–313
- Dridi L, Majdoub R, Hachicha M (2013) Modélisation hydrogéologique des écoulements souterrains au niveau de la région d’Ouled Chamekh. Actes des 17<sup>èmes</sup> Journées Scientifiques de l’INRGREF, Hammet, Tunisie
- Eaton AD, Clesceri LS, Greenberg AE (1995) Standard methods for the examination of water and wastewater, 19th edn. American Public Health Association, Washington, DC
- Essefi H, Tagorti MA, Touri J, Yachi Ch (2013) Hydrocarbons migration through groundwater convergence towards saline depression A case study Sidi El Hani, discharge playa, Tunisian Sahel. *ISRN Environ Chem*. <https://doi.org/10.1155/2013/709190>
- Fadili A, Najib S, Mehdi K, Riss J, Makan A, Boutayeb K, Guessir H (2016) Hydrochemical features and mineralization processes in coastal groundwater of Oualidia, Morocco. *J Afr Earth Sci* 116:233–247
- Farid I, Trabelsi R, Zouari K, Abid K, Ayachi M (2013) Deciphering the interaction between quaternary and continental Sabkhas aquifers in Central Tunisia using hydrochemical and isotopic tools. *Environ Earth Sci* 70:3289–3309
- Freeze RA, Cherry JA (1979) *Groundwater*. Prentice-Hall Inc., Englewood Cliffs, p 604
- Gbolo P, Gerla P (2013) Statistical analysis to characterize transport of nutrients in groundwater near an abandoned feedlot. *Hydrol Earth Syst Sci* 17:4897–4906
- Goovaerts P (1997) *Geostatistics for natural resource evolution*. Oxford University Press, New York
- Guler C, Thyne GD, McCray JE, Tumer AK (2002) Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeol J* 10:455–474
- Hachicha W, Masmoudi F, Haddar M (2008) Formation of machine groups and part families in cellular manufacturing systems using a correlation analysis approach. *Int J Adv Manuf Technol* 36(11):1157–1169
- Hajd Ammar F, Chakir N, Zouari K, Deschamps P (2014) Hydro-geochemical processes in the complex terminal aquifer of southern Tunisia: an integrated investigation based on geochemical and multivariate statistical methods. *J Afr Earth Sci* 100:81–95
- Halimi S, Baali S, Kherici N, Zairi M, Bouhsina S (2016) Irrigation and risk of saline pollution. Example: groundwater of Annaba Plain. *Eur Sci J* 12(6):1857–7431
- Hamzaoui-Azaza F, Bouhlila R, Gueddari M (2009) Geochemistry of fluoride and major ion in the groundwater samples of Triassic aquifer (South Eastern Tunisia), through multivariate and hydrochemical techniques. *J Appl Sci Res* 5(11):1941–1951
- Hamzaoui-Azaza F, Tlili-Zrelli B, Bouhlila R, Gueddari M (2013) An integrated statistical methods and modeling minerals-water interaction to identifying hydrochemical processes in groundwater in southern Tunisia. *Chem Speciat Bioavailab* 25(3):165–178
- Hassen I, Hamzaoui-Azaza F, Bouhlila R (2016) Application of multivariate statistical analysis and hydrochemical and isotopic investigations for evaluation of groundwater quality and its suitability for drinking and agriculture purpose: case of Oum Ali-Thelepte aquifer, Central Tunisia. *Environ Monit Assess J* 188(3):135
- Jiang Y, Wu Y, Groves C, Yuan D, Kambesis P (2009) Natural and anthropogenic factors affecting the groundwater quality in the

- Nandong Karst underground river system in Yunan, China. *J Contam Hydrol* 109:49–61
- Kaiser HF (1960) The application of electronic computers to factor analysis. *Educ Psychol Meas* 20:141–151
- Kamel S, Dassi L, Zouari K (2010) Approche hydrogéologique et hydrochimique des échanges hydrochimiques entre aquifères profond et superficiel du bassin du Djérid, Tunisie. *J Sci Hydrol* 51(4):713–730
- Kharroubi A, Tkahigue F, Agoubi B, Azri Ch, Bouri S (2012) Hydrochemical and statistical studies of the groundwater salinization in Mediterranean arid zones: case of the Jerba coastal aquifer in southeast Tunisia. *Environ Earth Sci J* 67:2089–2100
- Khomsis S, Bédir M, Soussi M, Ben Jemia MG, Ben Ismail-Lattrache K (2006) Mise en évidence en subsurface d'évènements compressifs Eocène moyen-supérieur en Tunisie orientale (Sahel): généralité de la phase atlasique en Afrique du Nord. *C R Geosci* 338:41–49
- Kim KO, Yun ST, Choi BY, Chae GT, Kim K, Kim HS (2009) Hydrochemical and multivariate statistical interpretations of spatial control of nitrate concentrations in a shallow alluvial aquifer around oxbow lakes (Osong area, central Korea). *J Contam Hydrol* 107:114–127
- Kim TH, Chung SY, Park N, Hamm SY, Lee SY, Kim BW (2012) Combined analyses of chemometrics and kriging for identifying groundwater contamination sources and origins at the Masan coastal area in Korea. *Environ Earth Sci* 67(5):1373–1388
- Kolsi-Hajji S, Bouri S, Hachicha W, Ben Dhia H (2013) Implementation and evaluation of multivariate analysis for groundwater hydrochemistry assessment in arid environments: a case study of Hajeb Elyoun-Jelma, Central Tunisia. *Environ Earth Sci J* 70(5):2215–2224
- Louati D, Majdoub R, Rigane H, Abida H (2015) Spatial estimation of soil salinity with ordinary Kriging (Eastern Tunisia). *Int J Adv Res* 3(5):274–283
- M'nassri S, Dridi L, El Amri A, Hachicha M, Majdoub R (2016) Aptitude des eaux de la nappe phréatique de la cuvette de Sidi El Hani à l'irrigation (Centre-Est de la Tunisie). *J Mater Environ Sci* 7(12):4742–4753
- M'nassri S, Dridi L, Lucas Y, Schäfer G, Hachicha M, Majdoub R (2018) Identifying the origin of groundwater salinisation in the Sidi El Hani basin in central-eastern, Tunisia. *J Afr Earth Sci* 147:443–449
- M'nassri S, Lucas Y, Dridi L, Schäfer G, Majdoub R (2019) Coupled hydrogeochemical modeling using KIRMAT to assess water-rock interaction in a saline aquifer in central-eastern Tunisia. *Appl Geochem J* 102:229–242
- Maamri R, Rabhi M (2003) Carte géologique de la Tunisie. Feuille Oued Chérta. Office National des Mines, Tunisie
- Majdoub R, Hachicha M, El Amri A, Melki M (2012) Etude de la dynamique de l'eau et du transfert des sels un sol sablo-limoneux du Sahel Tunisien. *Eur J Sci Res* 80(4):499–507
- Maoui A, Kherici N, Derradji F (2009) Hydrochemistry of an Albian sandstone aquifer in a semi arid region, Ain oussera, Algeria. *Environ Geol* 60(4):689–701
- Mikhailov EV, Tupicina OV, Bykov DE, Chertes KL, Rodionova OY, Pomerantsev AL (2007) Ecological assessment of landfills with multivariate analysis—a feasibility study. *Chemom Intell Lab Syst* 87:147–154
- Najib S, Fadili A, Mehdi K, Riss J, Makan A (2017) Contribution of hydrochemical and geoelectrical approaches to investigate salinization process and seawater intrusion in the coastal aquifer of Chaouia, Morocco. *J Contam Hydrol* 198:34–36
- Nyenjie P, Batelaan O (2009) Estimating the effects of climate change on groundwater recharge and baseflow in the upper Ssezibwa catchment, Uganda. *Hydrol Sci J* 54(4):713–726
- Piper AM (1944) A graphic procedure in the geochemical interpretation for water analyses. American geophysical Union: Papers Hydrology
- Rodier J (1996) L'analyse de l'eau naturelle: eaux résiduaires, eau de mer. p 1384
- Sachez-Martoz F, Jimenez ER, Pulido BA (2001) Mapping groundwater quality variables using PCA and geostatistics: a case study of Bajo Andarax, southeastern Spain. *Hydrol Sci J* 46(2):227–242
- Trabelsi R, Zairi M, Ben Dhia H (2007) Groundwater salinization of the Sfax superficiel aquifer, Tunisia. *Hydrogeol J* 15(7):1341–1355
- Venkatramanan S, Chung SY, Kim TH, Kim BW, Selvam S (2016) Geostatistical techniques to evaluate groundwater contamination and its sources in Miryang City, Korea. *Environ Earth Sci* 75:994. <https://doi.org/10.1007/s12665-016-5813-0>
- Ward JH (1963) Hierarchical grouping to optimize an objective function. *J Am Stat Assoc* 58:236–244
- Yermani M, Zouari K, Michelot JL, Mamou A, Moumni L (2003) Approche géochimique du fonctionnement de la nappe profonde Gafsa Nord (Tunisie Centrale). *J Sci Hydrol* 48(1):95–108
- Yidana MS, Yakubo-Banoeng B, Akabzaa TM (2010) Analysis of groundwater quality using multivariate and spatial analyses in the Keta basin, Ghana. *J Afr Earth Sci* 58:220–234
- Yidana MS, Ophori D, Yakubo-Banoeng B, Samed A (2012) A factor Model to explain the hydrochemistry and causes of fluorides enrichment in groundwater from the Middle voltain sedimentary aquifers in the northern region, Ghana. *ARPN J Eng Appl Sci* 7(1):1918–6608

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.