Desalination and Water Treatment
www.deswater.com

1944-3994 / 1944-3986 © 2010 Desalination Publications. All rights reserved.
doi: 10.5004/dwt.2010.1077

19 (2010) 307–317
July

Application of multivariate analysis to study water chemistry of groundwater in a semi-arid aquifer, Malayer, western Iran

Mohsen Jalali

Department of Soil Science, College of Agriculture, Bu-Ali Sina University, Hamadan, Iran
Tel. +98 (811) 4227090; Fax +98 (811) 4227012; email: Jalali@basu.ac.ir

Received 13 October 2009; Accepted in revised form 3 December 2009

ABSTRACT

Groundwater chemistry is controlled by many natural processes as well as anthropogenic processes. The identification of the influence of anthropogenic processes on groundwater chemistry is difficult particularly in cases of early phase of groundwater anthropogenic transformation. Multivariate analysis (principal component analysis and cluster analysis) was performed to identify a common source for groundwater chemistry and identification of anthropogenic processes affecting groundwater chemistry of a semi-arid aquifer, Malayer, western Iran. The principal component analysis (PCA) performed on groundwater identified three principal components controlling variability of groundwater chemistry. Magnesium, Na^+, K^+ and Cl^- content were associated in the same component (PC1) (salinity), most probably interpret as linked to anthropogenic activities. The Ca^{2+}, HCO_3^-, SO_4^{2-}, and Si (PC2) content were interpreted as related to the lithogenic. Finally, the NO_3^- and pH (PC3) was interpreted as reflecting long-term anthropogenic activity as a consequence of application of inorganic fertilizers, atmospheric deposition or industrial wastes. The results of this study demonstrate that multivariate analysis is probably a reliable method for differentiation of natural and anthropogenic processes affecting groundwater chemistry.

Keywords: Anthropogenic; Groundwater; Hydrochemistry; Multivariate analysis

1. Introduction

The monitoring of water quality is one of the important tools for sustainable development and provides important information for water management in arid and semi-arid regions. Intense agricultural and urban development has placed a high demand on groundwater resources in arid and semi-arid regions and also these resources at greater risk to contamination [1]. Complex processes control chemical composition of groundwater, which typically has a large range of chemical composition [2–4]. The composition of groundwater depends not only on natural factors such as composition of precipitation, geological structure and mineralogy of the watersheds and aquifers, geological processes within the aquifer [5], but also on human activities, which can alter these fragile groundwater system [6], either by polluting them or by changing the hydrological cycle [7].

Multivariate statistical techniques are widely used in environmental geochemistry [8]. These statistical methods, capable of distinguish complex relations among many variables, can be useful for source-identification problems. They also used to examine phenomena associated with water quality and to understand the behaviour and spatial patterns of water quality constituents. Principal component analysis is a multivariate statistical technique that can reveal the underlying structure of a dataset by means of distances between variables or samples in a multi-dimensional data space [9]. Principal component analysis is a powerful tool for determining which variables explain a large proportion of the variation and for identifying potential clusters among sites [10–12]. The first principal component (PC1) accounts for as much of the variation present as possible; it represents the line of maximum variance in the data. The second component (PC2) is uncorrelated with the first one and explains as much of the remaining variation as possible, and so on, until the total variance is accounted for [13].

Cluster analysis (CA) is a term for numerical techniques that separate data into constituent groups of individuals. Cluster analysis, also called segmentation analysis or taxonomy analysis, seeks to identify homogeneous subgroups of data in a population. That is, cluster analysis seeks to identify a set of groups which both minimize within-group variation and maximize between-group variation.

In some parts of Iran including Hamadan [14] and Bahar [15] area evidence of groundwater contamination due to anthropogenic contamination appear. Groundwater is a vital resource in this area, playing important roles as drinking water source, and as an ecosystem support. The investigation of groundwater contamination is complicated because changes of groundwater chemistry caused by contamination and natural hydrochemical processes overlap [12]. The present study shows the application of multivariate statistical technique for the identification of anthropogenic processes affecting groundwater chemistry of a semi-arid aquifer, Malayer, western Iran.

2. Material and methods

2.1. Study area

The study area is in the southern region of Malayer city in Hamadan province, western Iran, and located between longitudes 48°24′ and 49°8′ E and latitudes 34°00′ and 34°21′ N (Fig. 1). The area is characterized by hills in the western and southwest parts. The highest elevation (altitude) in the hilly area is 2232 m.a.s.l., whereas in the plains the lowest altitude is 1677 m.a.s.l. The area has a cold temperate climate. The climate of the study area is considered to be semi-arid, the annual precipitation being approximately 300 mm. Peak rain months include January, February, March, April, November, and December. Late spring and summer months are the driest, each with monthly rainfall below 5 mm. Another feature characterizing the precipitation in the study site is its irregular yearly distribution. The annual potential evaporation far exceeds the annual rainfall with a mean annual amount of approximately, 1945 mm estimated from 1975 to 2001. The most important economic activity of the area is agriculture, the chief crops are wheat (*Triticum aestivum* L.), barley (*Hordeum vulgare* L.), alfalfa (*Medicago sative* L.), grape (*Vitis vinifera* L.), sugar beet (*Beta vulgaris*), and vegetables with actual irrigation being lower than total theoretical demand, as there is considerable deficit in relation to the amount of irrigated land.

The aquifer is characterized by intensive exploitation as well as by a great variety of natural factors affecting groundwater chemistry. There is not any more water source in the studied area (surface streams like rivers or dammed water) that could be used for irrigation. The study area represents different types of land uses and most wells located in intensive agricultural sites. Pollution sources connected with agriculture are mainly livestock farms and poultry. Livestock and poultry manure is mostly spread on the land surface. The main risk concerning cultivated land is the excessive application of fertilizers.

Basement rocks in the southern of Malayer consist of limestone, sandstones, partly quartizic, some shale and conglomerate, granodiorite-diorite, hornfels and hornfels schist. Soils in the studied area are calcareous. Average electrical conductivity and pH for some of the soils in the studied area are 0.59 dS m^{-1} and 7.8, respectively. The

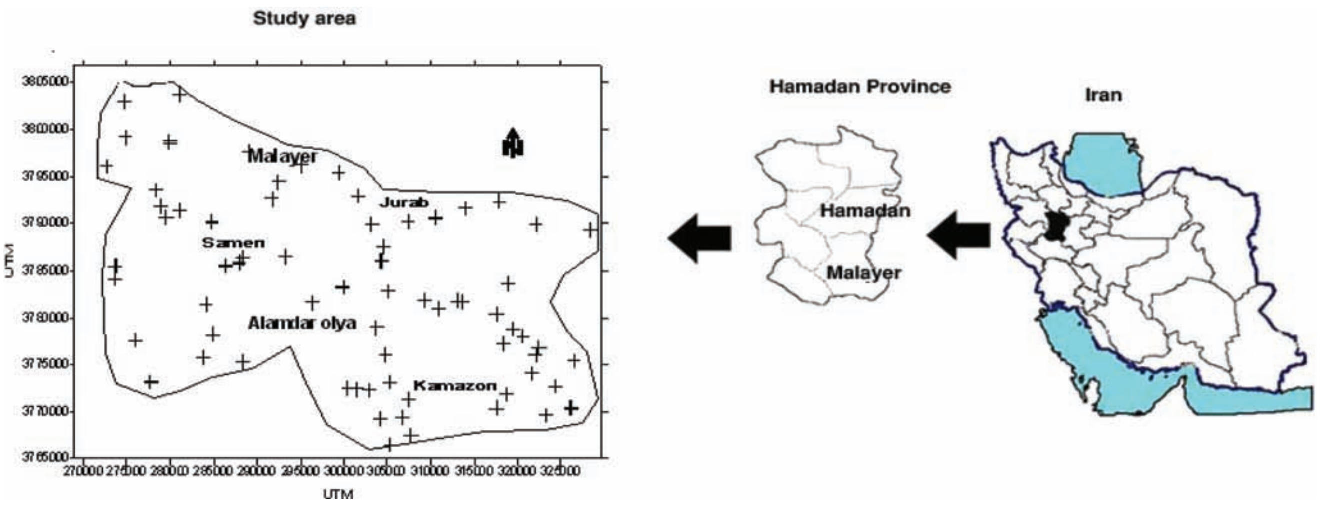


Fig. 1. Study area showing location of wells sampled for groundwater analysis.

amount of calcite in soils is about 14%. Average clay and sand contents are 200 and 440 g kg^{-1}, respectively. The soils are sandy loam, loam and clay loam in texture [16].

2.2. Sampling method and water analysis

Seventy one groundwater samples were collected from southern Malayer, during September, 2005 (Fig. 1). In order to avoid contamination, clean 500 ml plastic containers were used in drawing the water samples from wells. These containers had been rinsed with distilled water before sampling. The samples were stored in polyethylene containers, adequately labeled and preserved in the refrigerator until they were taken to the laboratory for measurement. Samples were analyzed in the laboratory for the major ions employing standard methods. The determinations were made within 48 h of collection. The pH and electrical conductivity (EC) were measured using pH and electrical conductivity meters. Calcium (Ca^{2+}) and magnesium (Mg^{2+}) were determined titrimetrically using standard EDTA [17]. Chloride was determined by standard $AgNO_3$ titration. Carbonate (CO_3^{2-}) and bicarbonate (HCO_3^-) were determined by titration with HCl. Sodium (Na^+) and potassium (K^+) were measured by flame photometry, sulfate (SO_4^{2-}) by spectrophotometric turbidimetry and $NO_3^-$by colorimetry with an UV-visible spectrophotometer [17]. Care was taken that the HCO_3^-, Ca^{2+}, and pH were analyzed within 12 h of sampling. Total dissolved solids (TDS) were computed by multiplying the EC (dS m^{-}1) by a factor of 640. Silicon was measured by ICP.

Saturation indices for calcite, gypsum, dolomite, anhydrite, aragonite and halite were computed through the computer programmers, PHREEQC [18]. The pH of groundwater can be changed from its true value in the groundwater during exposure at the surface to an environmental condition with different temperature and partial carbon dioxide pressure. Thus, the actual pH values could possibly be a tenth to a few tenths of a unit lower than the lab reported values and that the saturation indices for the carbonate minerals could be a little high as a result.

2.3. Data processing

The basic purpose of multivariate analysis is to generate groups of correlated elements from the initial data set. These groups are included into principal factors, which describes processes occurring in the investigate environment. Identification of the factors allows genetic interpretation of the environment. Principal components analysis and *k*-means clustering were done with the statistical software package MINITAB (version 13.1, Minitab Inc.). The general procedure of k-means clustering aims to minimize the variation within groups and maximize that between them based on a selected criterion [19]. The PCA was done on the correlation matrix.

3. Results and discussion

3.1. Major ion chemistry

The electrical conductivity indicates the amount of material dissolved in water, which range from 0.310 to 1.30, with an average value equal to 0.510 and a median of 0.460 dS m^{-1}. The recommended value of EC for potable water is 0.250 dS m^{-1}. Thus the EC of the all samples are above the maximum permissible limit. Total dissolved solids ranges from 193 to 807 with an average of 313 mg l^{-1}. The recommended value of TDS for potable water is 1000 mg l^{-1}. Thus the TDS of water samples are well within the maximum permissible limit. Waters are slightly alkaline, with a pH ranging between 5.9 and 8.4, with an average and median value equal to 7.9, indicating that the pH of water samples is well within the desirable limit (7.0–8.5).

Groundwater composition is dominated by Ca^{2+}, Na^+ and Mg^{2+} and HCO_3^-, Cl^- and SO_4^{2-}. Calcium and HCO_3^- can account for respectively 10.2–37.4% and 12–66.8% of the relative cationic and anionic content of the groundwater (in meq l^{-1}). These major ionic species in solution are positively correlated ($r = 0.62$, $P < 0.01$) and their contribution to the overall chemical composition of the groundwater is shown by the correlation between Ca^{2+} and the EC ($r = 0.50$, $P < 0.01$), HCO_3^- and EC ($r = 0.51$, $P < 0.01$). Sulfate can account for 4.5–44.5% of the relative anionic composition.

Chloride ion impacts a salty taste to water. The limit for domestic purposes is fixed at 250 mg l^{-1}. In the present study Cl^- content is well within the maximum permissible limit. The WHO guiding limit for SO_4^{2-} is 400 mg l^{-1} [20]. All of the water samples are within the maximum permissible limit. Most of the Ca^{2+}, Mg^{2+} and Na^+ contents in groundwater samples are well within the permissible limits.

Fig. 2 shows Piper diagrams corresponding to the data. It can be observed that the groundwater is basically a HCO_3^- water, with Ca^{2+} and Na^+ being prevailing cations. Thus, groundwater of the studied area is mostly classified into the Ca–HCO_3 and Na–HCO_3 water types.

The saturation indices (SI) describe quantitatively the deviation of water from equilibrium with respect to dissolved minerals. If the water is exactly saturated with the dissolved mineral, SI equals to zero. Positive values of SI indicate saturation, and negative one indicates undersaturated [3,4]. In Table 1 the SI for calcite, gypsum, dolomite, anhydrite, and halite are shown. Nearly 57% and 35% of water samples were oversaturated with respect to calcite and dolomite, respectively, and all water samples were undersaturated with respect to gypsum, anhydrite, and halite suggesting that these carbonate mineral phases may have influenced the chemical composition of the study area.

Table 1
Saturation indices for calcite, gypsum, dolomite, anhydrite, aragonite and halite

	$SI_{calcite}$	SI_{gypsum}	$SI_{dolomite}$	$SI_{anhydrite}$	SI_{halite}
Average	–0.02	–2.51	–0.16	–2.73	–7.78
Median	0.02	–2.58	–0.15	–2.8	–7.81
Max	0.55	–1.73	1.2	–1.95	–6.22
Min	–1.85	–2.99	–3.8	–3.21	–8.54
Skewness	–2.64	0.77	–2.07	0.77	0.95

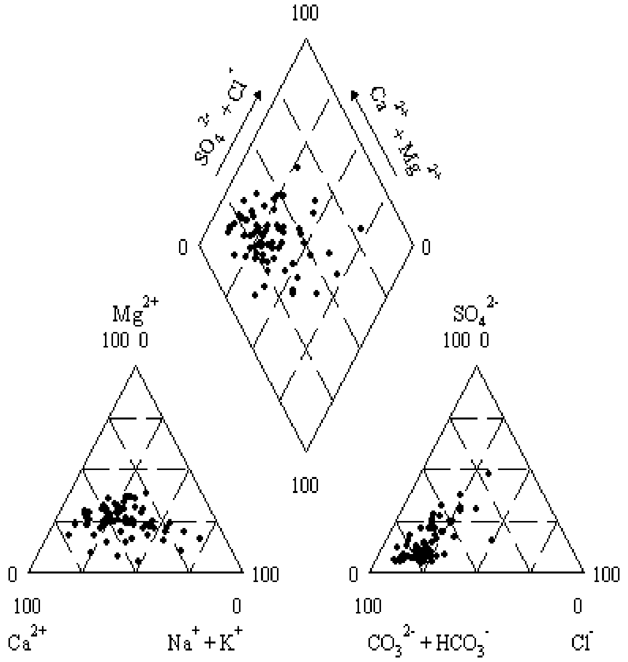


Fig. 2. Piper diagram for groundwater samples in the studied area.

3.2. Hydrogeochemical evaluation

Results from chemical analyses were used to identify the geochemical processes and mechanisms in the aquifer system. To study the relative importance of different weathering processes, binary plots of (Ca^{2+} + Mg^{2+}) vs. (HCO_3^- + SO_4^{2-}) was examined. If Ca^{2+}, Mg^{2+}, SO_4^{2-} and HCO_3^- are derived from simple dissolution of calcite, dolomite and gypsum, then a charge balance should exist between the cations and anions. Ion exchange tends to shift the points right because of the excess of SO_4^{2-} + HCO_3^- ions, which may be due to anthropogenic input in the groundwater. The plot (Fig. 3a) falls above the 1:1 line, indicating a portion of the HCO_3^- + SO_4^{2-} must be balanced by Na^+ and K^+. Thus, ion-exchange process in the groundwater system is dominated, indicating possible leaching of these anions from soil.

The source of Ca^{2+} and Mg^{2+} in the groundwater can be deduced from the Ca^{2+} + Mg^{2+}/HCO_3^- ratio. If Ca^{2+} and Mg^{2+} originate soley from dissolution of carbonate in the aquifer from the weathering, the ratio would be about 0.5 [3]. The average ratio of Ca^{2+} + Mg^{2+}/HCO_3^- in groundwaters (0.89) is higher than the suggested value of 0.50. The higher values indicate that excess Ca^{2+} and Mg^{2+} have been balanced by Cl^- and SO_4^{2-}. The fact that this ratio increases with salinity shows that Ca^{2+} and Mg^{2+} are released into the groundwater aquifer faster than biocarbonate, indicating some anthropogenic input in the aquifer system.

A plot of (Ca^{2+} + Mg^{2+}) vs. total cations (Fig. 3b) shows that most of the points lie above the 1:1 line, indicating an increasing contribution of alkalis to the major ions, which indicates anthropogenic input in the groundwater. The increase in alkalis with a simultaneous increase in Cl^- + SO_4^{2-} suggests a common source of these ions and presence of Na_2SO_4 in the soil which indicates leaching of these ions from soils (Fig. 3c).

High concentration of Na^+ in the groundwater can be attributed to the cation exchange and to the human activities. High concentrations of Na^+ in irrigated area are also the result of the repeated use of water. The Na^+–Cl^- relationship has often been used to identify the mechanisms for acquiring salinity and saline intrusions in semi-arid regions [21,22]. The high Na^+ and Cl^- contents detected in certain samples may suggest the dissolution of chloride salts. The dissolution of halite in water release equal concentrations of Na^+ and Cl^- into the solution. A parallel enrichment in both ions indicates dissolution of chloride salts or reconcentration processes by evaporation. There was a high correlation between Cl^- and Na^+ (r = 0.78, $P < 0.01$) in the groundwater samples. But analytical data deviate from the expected 1:1 relation (Fig. 3d), indicating that most of the Na^+ to be derived from other processes. The ionic relations suggest that Ca^{2+} and Mg^{2+} provided by calcite, dolomite and gypsum dissolution are replaced by Na^+ from a source other than NaCl.

A Na^+/Cl^- molar ratio greater than one is typically interpreted as reflecting Na^+ released from silicate weathering reactions [23]. The average molar ratio of Na^+/Cl^- (1.62) indicates much higher Na^+ values compared with Cl^-.

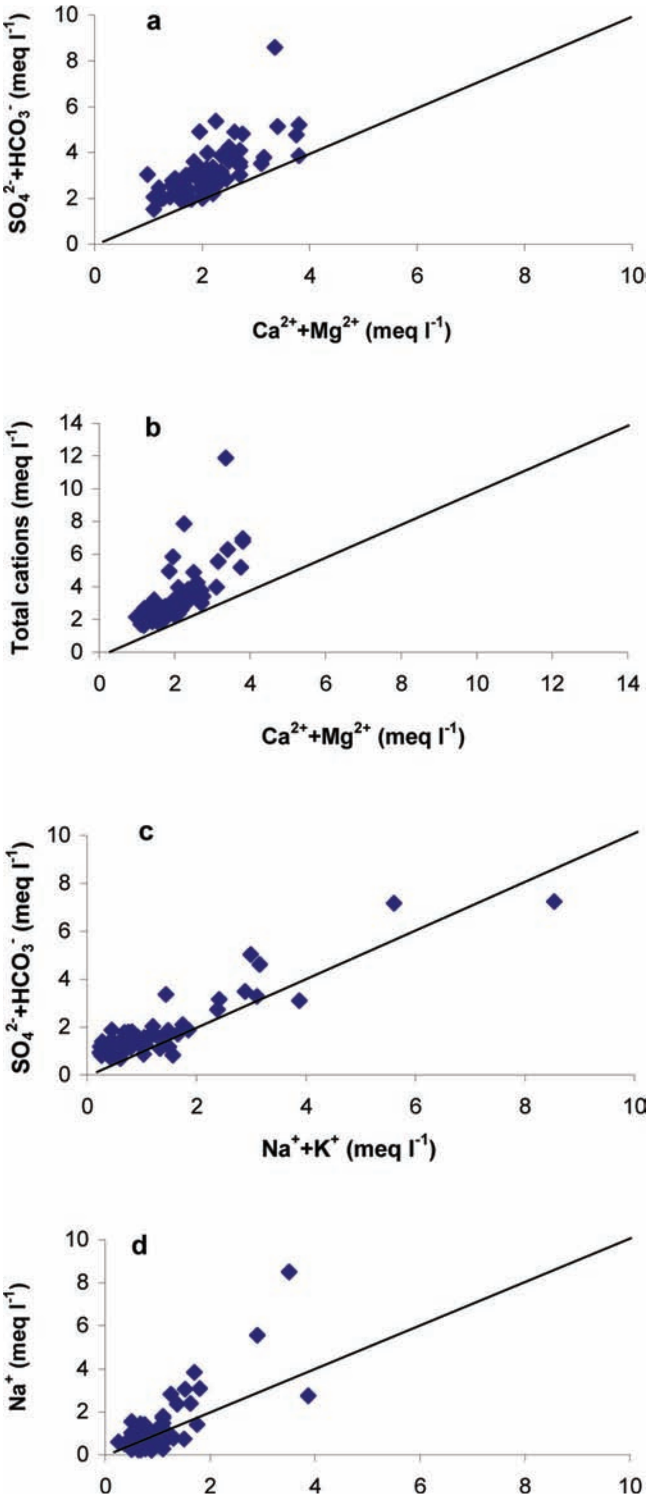


Fig. 3. Plots of diferent parameters.

The cation exchange between Ca^{2+} or Mg^{2+} and Na^+ may also explain the excess Na^+ concentration [24]. Those samples with a value of Na^+/Cl^- ratio higher than one also show a deficit in $Ca^{2+} + Mg^{2+}$ and this is consistent with a Ca^{2+}–Na^+ cation exchange process which leads to a softening of the water [25]. A Na–HCO_3 water type is usually an indication of cation exchange processes [26]. Such water type is referred to as exchange waters which imply that there is more HCO_3^- than the available alkaline-earth cations ($Ca^{2+} + Mg^{2+}$) in equivalent concentrations [26]. These excess bicarbonate ions then cause a release of the alkali ions (usually Na^+) into the solution by the exchange reaction with the exchange sites.

Agriculture has both indirect and direct effects on the quality of surface and groundwaters and is one of the key activities causing water quality degradation in many parts of Iran. Municipal and industrial wastewaters are not properly treated before discharged, which makes it easy for contaminants to reach groundwater. Arid conditions create irrigation-return flow to become concentrated in some chemical constituents because of a number of processes such as evapotranspiration, mineral dissolution, fertilizer application and agricultural residues [27]. The net loss in Ca^{2+} and HCO_3^- are common as a result of precipitation accompanied by a net gain of Na^+, K^+, Cl^-, SO_4^{2-} and NO_3^- [28].

3.3. Spatial distribution of EC and NO_3^-

Nitrate and electrical conductivity has been chosen as indicators of the degree of mineralization and pollution of groundwater, respectively. The common groundwater contaminant in the area was NO_3^- [14,15]. Nitrate is the most common form of nitrogen that occurs in surface and groundwater. Because of its anionic form, nitrate is very soluble and mobile in aqueous solution. It represents the oxidized end product in the nitrogen cycle of atmosphere, vegetation, upper soil and soil water zones. Concentrations of NO_3^- are the result of different pollution processes, municipal wastewaters, fertilizers application in agriculture.

Nitrate concentrations in the well samples varied from 4 to 73 with the average of 22.9 mg l^{-1}. In comparison with the WHO's drinking water guideline [20] of 50 mg l^{-1} for NO_3^-, a total of 2 wells (3%) showed higher concentrations (Fig. 4a). Results showed that of 71 wells, 18 (25.3%) had levels more than 30 mg l^{-1} NO_3^- (Fig. 4a). Groundwater with NO_3^- concentration exceeding the threshold of 13 mg l^{-1} NO_3^- is considered contaminated due to human activities (the so-called human affected value [29,30]). 69% of the water samples showed NO_3^- concentrations above the human affected value. The lines of equal nitrate concentrations are shown in Fig. 5b.

The frequency distribution chart (Fig. 4b) indicates that 84.5% of the samples contain EC in the range 0.40–0.60 dS m^{-1}. Fig. 5a shows distribution of EC in the studied area. Changes in groundwater salinity and chemical composition occur in the north east part of the studied area due to either natural and/or anthropogenic activity [28]. Evaporation of surface water and moisture in the

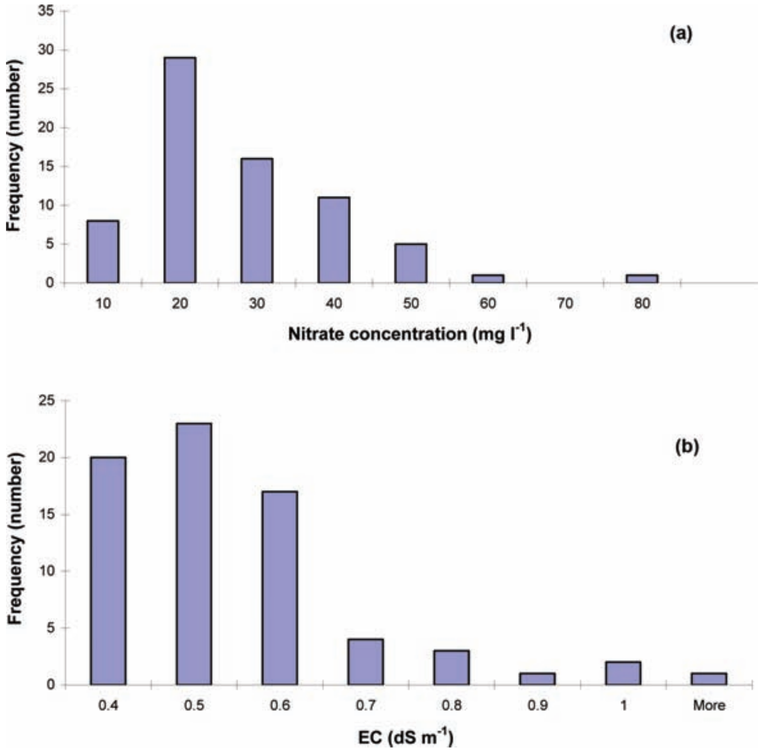


Fig. 4. Concentration distribution for (a) nitrate and (b) electerical conductivity.

unsaturated zone is the most important process affecting the chemical composition of shallow groundwater [28].

Increased water consumption has accelerated the transfer of water between surface and groundwater stores, leading to salinization of groundwater, especially in the north basin. The water resources in the studied area are not being used in an economically efficient manner and some water is wasted by unnecessary evaporation and traditional irrigation [14].

3.4. Statistical analysis

3.4.1. Correlation between variables

The Pearson correlation matrix was computed (Table 2) and the criterion values of probabilities ($P < 0.05$ and $P < 0.01$) for correlation significance are used. The correlations among the elements are similar to the conclusion drawn from the component loadings (following section), i.e. a strong positive correlation between EC and Cl^- (0.89 $P < 0.01$), EC and SO_4^{2-} (0.79 $P < 0.01$), EC and Mg^{2+} (0.75 $P < 0.01$), and EC and Na^+ (0.87 $P < 0.01$) was observed, which means that these parameters can be considered to be strongly correlated.

3.4.2. Principal component analysis

Principal component analysis was performed on the data of groundwater in order to better understand their interrelationships and to explore the reduction of the experimental variables. Table 3 shows the principal components, eigenvalues and associated variance explained in the physico-chemical data and communities. These components all had eigenvalues higher than one (the most significant one) [31]. The plot of factor loadings for the three components is presented in Fig. 6. Three extracted components explain 67.6% of data set variance. Component 1 explains 33.40% of the total variance, while component 2 explains 23.50%. Component 1 has the highest component loadings of Cl^-, Na^+, Mg^{2+} and, K^+ high component loadings of Ca^{2+} and HCO_3^- and low component loadings for the remaining parameters. This association strongly suggests that PC1 may represent a salinization (salinity) variable. Component 2 has the highest (negative) component loadings of Si, HCO_3^-, SO_4^{2-} and Ca^{2+}, high component loadings of pH and low component loadings for the remaining parameters. Component 3 has the highest (negative) component loadings of NO_3^-, high component loadings of pH and low component loadings for the remaining parameters. It can be assumed that some of the NO_3^- is derived from ammonium oxidation (nitrification). Nitrification is a two-step process in which NH_4^+ is converted first to NO_2^- and then to NO_3^- by specific micro-organisms [32]:

$$2NH_4^+ + 3O_2 \rightarrow 2NO_2^- + 2H_2O_2 + 4H^+$$

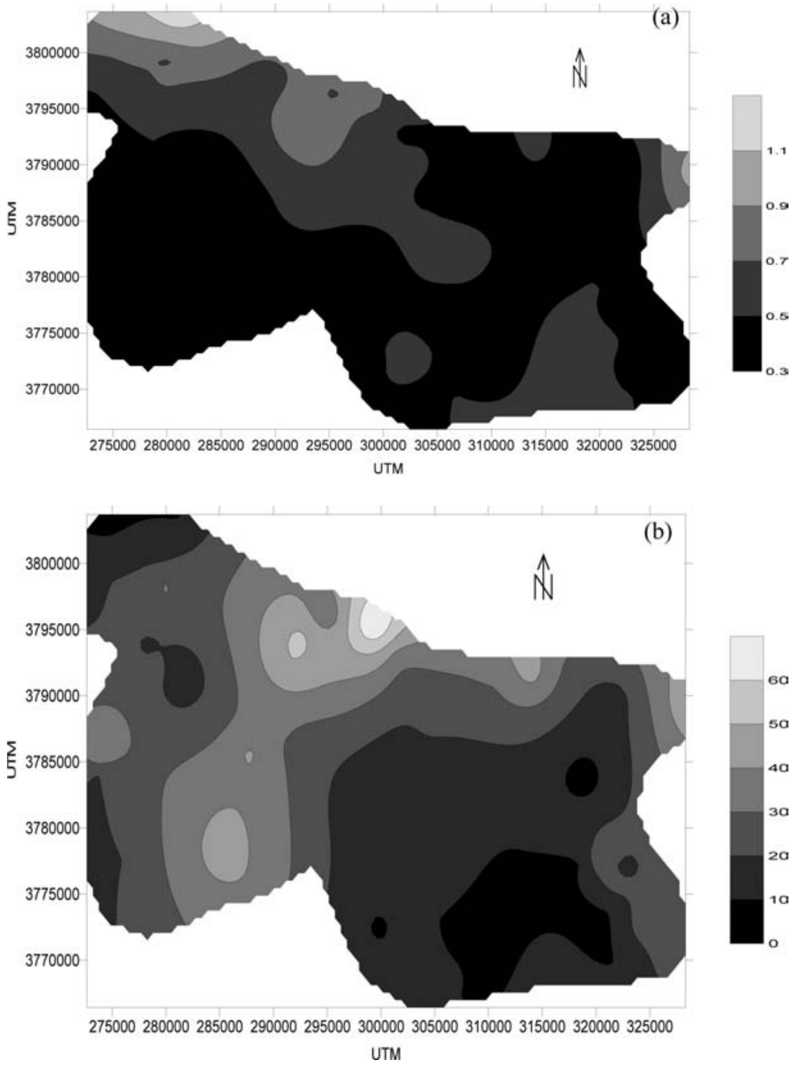


Fig. 5. Spatial distribution of (a) electerical conductivity (dS m^{-1}) and (b) nitrate (mg l^{-1}) in the water samples from the Malayer area.

Table 2
Pearson correlation coefficients for the 11 parameters

	pH	Ca^{2+}	Mg^{2+}	Na^{+}	K^{+}	Si	CO_3^{2-}	HCO_3^{-}	Cl^{-}	SO_3^{2-}	NO_3^{-}	EC
pH	1.000											
Ca^{2+}	-0.474^{**a}	1.000										
Mg^{2+}	–0.085	0.370^{**}	1.000									
Na^{+}	0.087	0.145	0.566^{**}	1.000								
K^{+}	–0.161	0.375^{**}	0.128	0.122	1.000							
Si	-0.241^{*}	0.323^{**}	0.139	0.236	0.236	1.000						
CO_3^{2-}	–0.153	0.118	–0.043	0.030	–0.093	0.130	1.000					
HCO_3^{-}	-0.266^{*}	0.617^{**}	0.366^{**}	0.249^{*}	0.349^{**}	0.629^{**}	0.157	1.000				
Cl^{-}	–0.037	0.302^{*}	0.698^{**}	0.780^{**}	0.274^{*}	0.013	0.223	0.223	1.000			
SO_4^{2-}	0.093	0.120	0.606^{**}	0.891^{**}	0.094	–0.189	–0.068	0.026	0.701^{**}	1.000		
NO_3^{-}	–0.105	0.092	0.165	0.132	–0.021	–0.011	-0.327^{**}	-0.293^{*}	0.136	0.242^{*}	1.000	
EC	–0.121	0.498^{**}	0.747^{**}	0.871^{**}	0.290^{*}	0.199	0.044	0.513^{**}	0.851^{**}	0.787^{**}	0.137	1.000

^{a*}, ** Significant at 5% and 1% level of significance respectively.

Table 3
Principal component loading

Variable	PC1	PC2	PC3
pH	–0.135	0.323	0.430
Ca^{2+}	0.298	–0.368	–0.162
Mg^{2+}	0.407	0.040	–0.144
Na^{+}	0.414	0.302	0.003
K^{+}	0.377	–0.050	0.131
Si	0.131	–0.427	–0.240
HCO_3^-	0.285	–0.475	0.126
Cl^-	0.428	0.160	0.154
SO_4^{2-}	0.363	0.394	0.117
NO_3^-	0.038	0.276	–0.800
Eigenvalues	3.344	2.352	1.065
% variance explained	33.40	23.50	10.70
% cumulative variance	33.40	57.00	67.60

$$2NO_2^- + O_2 \rightarrow 2NO_3^-$$

The source of NH_4^+ can be mineralization of organic N or N fertilizers containing or forming NH_4^+. The reactions show that nitrification of 1 mole of NH_4^+ produces 2 moles of H^+, resulting in increasing soil acidity. Nitrate is very mobile and subject to leaching losses.

The plot of factor scores for the first two factors are presented in Fig. 7. The data points with factor scores > 1 represent samples indicating anthropogenic transformations of groundwater chemistry (Fig. 7). The factor scores of factor 1 can be used as an indicator of groundwater anthropogenic contamination. Some outlier samples (2, 1, 43, and 47) partitioned in different clusters (following section), due to their variability in groundwater chemistry.

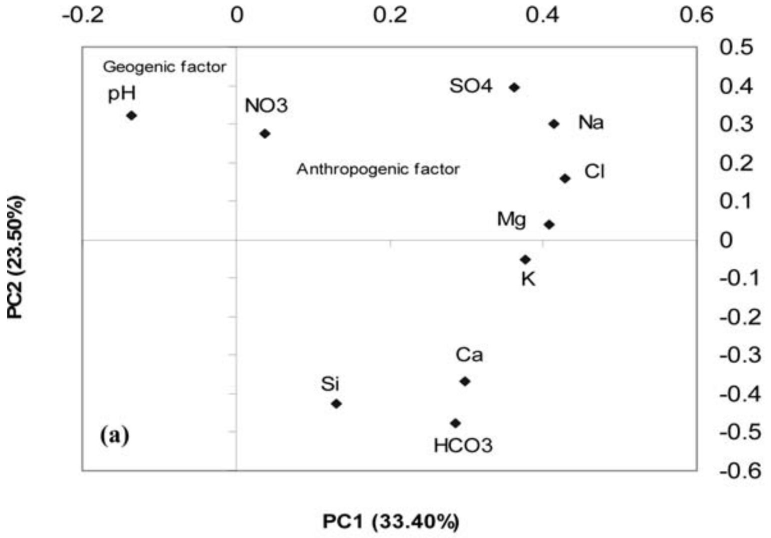
Thus, PC1 can be defined as an anthropogenic component due to the presence of high salinity levels in groundwater. The salinity component explains 33.4% of the total variance and is the most important component. High EC values can come from agrochemicals related to specific agronomic practices and wastewater used to irrigate some agricultural areas. PC2 can be considered to be a lithogenic component, as the variability of the elements seems to be controlled by parent rocks. The parent rock component explains 23.5% of the total variance. PC3 only includes pH and NO_3^- content. With respect to NO_3^-, the lack of relationship with the rest of the parameters may suggest that its origin is related to human activities. Nitrogen fertilizers are an important source of NO_3^- entering agricultural soils. Other sources of NO_3^- may include atmospheric deposition or anthropic waste such as sewage sludge, wastewater or waste materials. Nitrogen may become a highly problematic pollutant when it is added in large amounts to an ecosystem. Nitrate leached from soil are a major groundwater quality problem around the world [14]. A connection between agricultural and groundwater pollution is well established [33] and groundwater contamination by anthropogenic activities, such as urbanization and agricultural activities is a problem in arid and semi-arid regions [1,14]. Therefore, PC3 may also be an anthropogenic component.

In the present study, it seems reasonable to conclude that EC, Mg^{2+}, Na^+, SO_4^{2-}, Cl^-, and NO_3^- constitute an anthropogenic component related to human activities, whereas the remaining elements (Ca^{2+}, HCO_3^- and Si) appear to be associated with parent rocks.

3.4.3. Cluster analysis

In the absence of strong natural groups the use of different methods and clustering criteria can produce varying results when applied to the same data. Since *k*-means clustering (non-hierarchical cluster analysis) provides no direct means of determining the number of clusters, a hierarchical classification was produced to provide a starting point for the *k*-means clustering [19]. The hierarchical method suggested that the groundwater samples should be partitioned into five groups for the non-hierarchical cluster analysis. The resulting classification of the sites is shown in Table 4.

Group 2 only includes a well sample (sample 2) affected by high EC (0.960 dS m^{-1}), Na^+ (5.56 $mmol_c$ l^{-1}), and low NO_3^- (0.12 $mmol_c$ l^{-1}) content. Group 1 includes five well samples (samples 1, 24, 32, 33, and 43) having high content of Ca^{2+} (an average of 1.93 $mmol_c$ l^{-1}), Mg^{2+} (an average of 1.69 $mmol_c$ l^{-1}), and NO_3^- (an average of 0.62 $mmol_c$ l^{-1}). Group 3 includes eight well samples (3, 20, 23, 25, 26, 27, 38, and 40) with an average EC equal to 0.628 dS m^{-1}. Group 4 includes twenty five well samples (4, 6–13, 15, 17, 18, 21, 22, 29–31, 34–37, 39, 41–42, 70, and 71) with lowest EC, Ca^{2+}, and Cl^- (an average of 0.395 dS m^{-1}, 0.95 $mmol_c$ l^{-1} and 0.74 $mmol_c$ l^{-1}, respectively) . Finally,



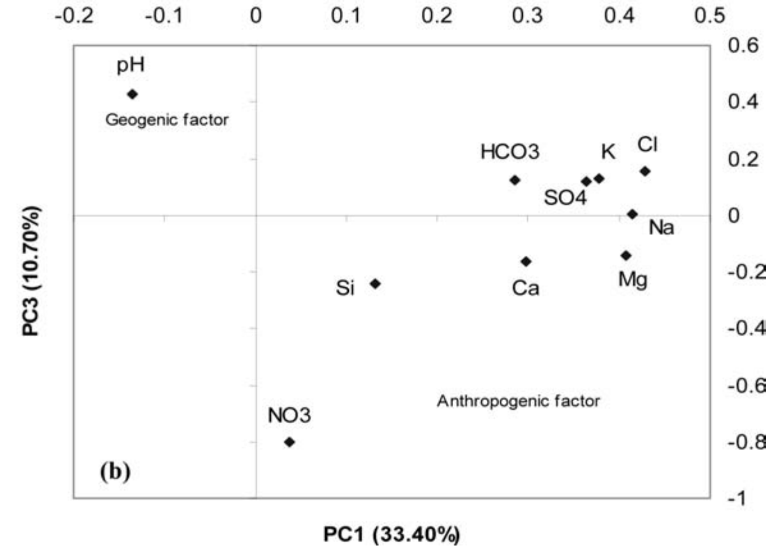


Fig. 6. Loading plots from PCA of chemical characteristics of groundwater in the studied area. (a) Principal component 1 vs. principal component 2; (b) principal component 1 vs. principal component 3.

Table 4
Class means of samples variables for the groundwater classification (concentrations are expressed in $mmol_c\ l^{-1}$ and EC in $dS\ m^{-1}$)

Variables	Group 1	Group 2	Group 3	Group 4	Group 5
pH	7.94	8.10	7.67	8.07	7.75
EC	0.931	0.960	0.628	0.395	0.480
Ca^{2+}	1.93	0.98	1.39	0.95	1.47
Mg^{2+}	1.69	1.27	0.95	0.61	0.71
Na^{+}	3.72	5.56	2.15	0.78	0.67
K^{+}	0.08	0.05	0.04	0.02	0.05
Si	2.46	1.37	2.32	2.14	2.41
HCO_3^-	3.20	1.10	2.57	1.80	2.82
Cl^-	2.43	2.90	1.27	0.74	0.87
SO_4^{2-}	2.32	4.27	1.21	0.59	0.38
NO_3^-	0.62	0.12	0.58	0.44	0.23

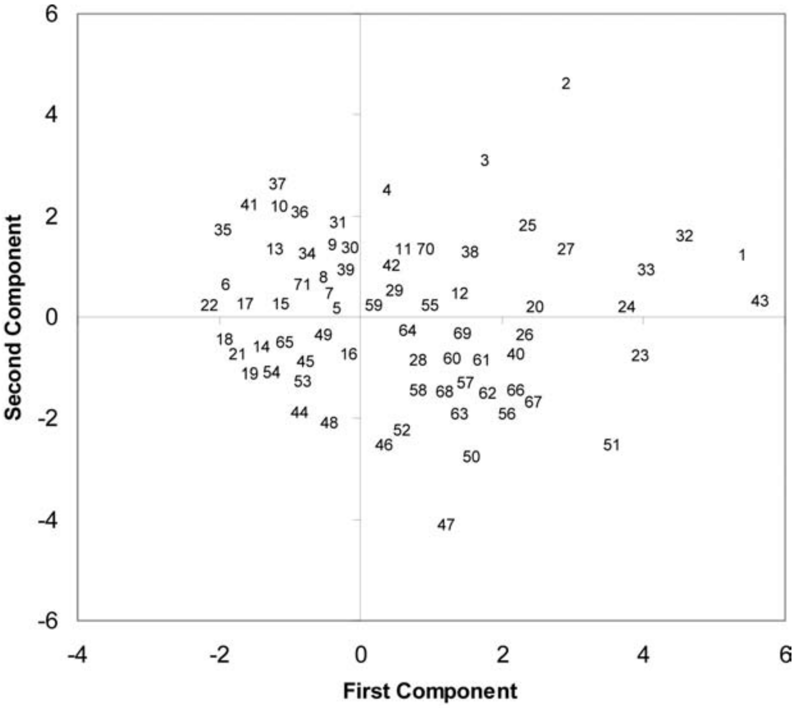


Fig. 7. Plot of component scores for the first two component.

wells in group 5, includes most of the water samples (32 water samples) (45%), with the lowest content of SO_4^{2-} and NO_3^- (an average of 0.230 dS m^{-1} and 0.38 mmol$_c$ l^{-1}).

4. Conclusion

The origin of groundwater chemistry of southern Malayer was identified using hydrogeochemical and multivariate analysis. The hydrochemical types of Ca–HCO_3 and Na–HCO_3 dominate the largest part of the studied area. Three components describe the main processes affecting groundwater chemistry: Component 1 (anthropogenic), marked by strong loading of, Mg^{2+}, Na^+, SO_4^{2-} and Cl^-. The concentrations of these elements have probably an anthropogenic origin and are associated with the influence of pollution from the land surface. Component 2 (geogenic) marked by strong loading of Ca^{2+}, HCO_3^- and Si. This component may reflect natural hydrogeochemical processes, connected mainly with the influence of this process on the solution of geochemical environment components. Component 3 (pH and NO_3^-) may also be an anthropogenic component. The results of multivariate analysis indicate that the influence of human activity is probably one of the most important factors controlling groundwater chemistry composition in the study area. This influence is reflected by the increased concentrations of the components, which in natural groundwater composition occur at low concentration. The results of multivariate analysis may be helpful for distinguished the vulnerability of an aquifer.

References

[1] M. Jalali, Chemical characteristics of groundwater in parts of mountainous region, Alvand, Hamadan, Iran. Environ. Geol., 51 (2006) 433–446.

[2] J.D. Hem, Study and Interpretation of Chemical Characteristics of Natural Water, 3 rd ed., USGS Water-Supply, Paper 2254, 1992.

[3] J.I. Drever, The Geochemistry of Natural Waters, 3rd ed., Prentice Hall, New Jersey, 1997, 436 p.

[4] C.A.J. Appelo and D. Postma, Geochemistry, Groundwater and Pollution. Rotterdam, A.A. Balkema, 1993, 536 p.

[5] L. Andre, M. Franceschi, P. Pouchan and O. Atteia, Using geochemical data and modelling to enhance the understanding of groundwater flow in a regional deep aquifer, Aquitaine Basin, south-west of France. J. Hydrol., 305 (2005) 40–62.

[6] K. Chen, J.J. Jiao, J. Huang and R. Huang, Multivariate statistical evaluation of trace elements in groundwater in a coastal area in Shenzhen, China. Environ. Pollut., 147 (2007) 771–780.

[7] B. Helena, R. Pardo, M. Vega, E. Barrado, J.M. Fernandez and L. Fernandez, Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. Water Res., 34(3) (2000) 807–816.

[8] T. Kikuchi, H.T. Hai and Sh. Tanaka, Characterization of heavy metal pollution in river sediment of Hanoi city and its downstream area by multivariate analyses. Desal. Wat. Treat., 4 (2009) 240–247.

[9] J.V. Cruz and C.S. Amaral, Major ion chemistry of groundwater from perched-water bodies of the Azores (Portugal) volcanic archipelago. Appl. Geochem., 19(3) (2004) 445–459.

[10] L.P. Gupta and V. Subramanian, Geochemical factors controlling the chemical nature of water and sediments in the Gomti River, India. Environ. Geol., 36 (1–2) (1998) 102–108.

[11] B. Škrbić and S. Čupić, Trace metal distribution in surface soils of Novi Sad and Bank Sediment of the Danube River. J. Environ. Sci. Health, Part A, 39 (2004) 1547–1558.

[12] K. Dragon, Application of factor analysis to study contamination of a semi-confined aquifer (Wielkopolska Buried Valley aquifer,

Poland). J. Hydrol., 331(1–2) (2006) 272–279.

[13] B. Škrbić and A. Onjia, Multivariate analyses of microelement contents in wheat cultivated in Serbia (2002). Food Control, 18 (2007) 338–345.

[14] M. Jalali, Nitrates leaching from agricultural land in Hamadan, western Iran. Agric Ecosyst Environ., 110 (2005) 210–218.

[15] M. Jalali and Z. Kolahchi, Nitrate concentration in groundwater of Bahar area, Hamadan. Iran. J. Soil Water Sci., 19(2) (2005) 194–202.

[16] M. Jalali, Phosphorous fixation and its effect on micronutrients availability (Zn, Fe, Mn, Cu) in semi-arid regions of Iran (Malayer and Takestan). MSc. thesis, Tehran University, 1991.

[17] D.L.Rowell, Soil Science: Methods and Applications. Longman Scientific and Technical, 1994.

[18] D.L. Parkhurst and C.A.J. Appelo, User's guide to PHREEQC (version 2) — A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations: US Geological Survey Water Resources Investigations, Report 99-4259, 1999, 310 p.

[19] M. Jalali, Site-specific potassium application based on the fertilizer potassium availability index of soil. Precision Agric., 8(3–4) (2007) 199–221.

[20] WHO Guidelines for Drinking Water Quality. 1. Recommendations, 2nd ed., World Health Organization, Geneva, 1993.

[21] W. Dixon and B. Chiswell, The use of hydrochemical sections to identify recharge areas and saline intrusions in alluvial aquifers, southeast Queensland, Australia. J. Hydrol., 130 (1992) 299–338.

[22] K. Sami, Recharge mechanisms and geochemical processes in a semi-arid sedimentary basin, Eastern Cape, South Africa. J. Hydrol., 139 (1992) 27–48.

[23] M. Meybeck, Global chemical weathering of surficial rocks estimated from river dissolved loads: Am. J. Sci., 287 (1987) 401–428.

[24] J. Stimson, S. Frape, R. Drimmie and D. Rudolph, Isotopic and geochemical evidence of regional-scale anisotropy and interconnetivity of an alluvial fan system, Cochabamba Valey, Bolivia. Appl. Geochem., 16 (2001) 1097–1114.

[25] M.C. Hidalgo and J. Cruz-Sanjulian, Groundwater composition, hydrochemical evolution and mass transfer in a regional detrital aquifer (Baza basin, southern Spain). Appl. Geochem., 16 (2001) 745–758.

[26] M.N. Tijani, Evolution of saline waters and brines in the Benue-Trough, Nigeria, Appl. Geochem., 19 (2004) 1355–1365.

[27] A. Cardona, J.J. Carrillo-Rivera, R. Huizer-Alvarez and E. Graniel-Castro, Salinization in coastal aquifers of arid zones: an example from Santo Domingo, Baja California Sur, Mexico. Environ. Geol., 45 (2004) 350–366.

[28] B.C. Ritcher and W.C. Kreitler, Geochemical Techniques for Identifying Sources of Groundwater Salinization, CRC Press, New York, 1993.

[29] M.R. Burkart and D.W. Kolpin, Hydrologic and land use factors associated with herbicides and nitrates in near-surface aquifers. J. Environ. Qual., 22 (1993) 646–656.

[30] D.A.V. Eckhardt and P.E. Stackelberg, Relation of groundwater quality to land use on Long Island, New York. Ground Water, 33 (1995) 1019–1033.

[31] N.J. Miller and J.C. Miller, Statistics and Chemometrics for Analytical Chemistry. 4th ed., Englewood Cliffs, NJ, Pearson Education, 2000.

[32] J.L. Havlin, J.D. Beaton, S.L. Tisdal and W.L. Nelson, Soil Fertility and Fertilizers: An Introduction to Nutrient Management. 6th ed., Prentice Hall, New Jersey, 1999.

[33] 33 P.A. Hamilton and D.R. Elsel, Effects of agriculture on groundwater quality in 5 regions of the united-states. Ground Water, 33(2) (1995) 217–226.