*Desalination and Water Treatment*
www.deswater.com

doi: 10.1080/19443994.2016.1180474

57 (2016) 26993–27002
December

Taylor & Francis
Taylor & Francis Group

# Investigation of groundwater hydrochemical characteristics using the multivariate statistical analysis in Ain Djacer area, Eastern Algeria

Dalila Ziani[a], Abderrahmane Boudoukha[a,*], Abderrahmane Boumazbeur[b], Lahcen Benaabidate[c], Chemseddine Fehdi[d]

[a]*Laboratory of Applied Research in Hydraulics, Hadj Lakhdar University of Batna, National Road n°3. Fesdis Pole 05 078, Batna 05000, Algeria, Tel./Fax: +213 33 86 97 24; emails: dalila_ziani@yahoo.fr (D. Ziani), boudoukha_abderrahmane@yahoo.fr (A. Boudoukha)*

[b]*Research Laboratory Sedimentary Environment, Mineral and Water Resources of Eastern Algeria, University of Tebessa, Tebessa 12000, Algeria, Tel. +213 0 560476250; email: f080263@hotmail.com*

[c]*Laboratory of Georesources and Environment, Faculty of Sciences and Technology, University of Sidi Mohammed Ben Abdellah, Immouzer Road, P.O. Box 2202, Fez 30000, Morocco, Tel. +212 35 60 29 53; Fax: +212 35 60 82 14; email: benaabidate@yahoo.fr*

[d]*Laboratory Water and Environment, Department of Geology, Tébessa University, Tébessa 12002, Algeria, Tel. +213 0 772293267; email: fehdi@yahoo.fr*

Received 6 April 2015; Accepted 17 March 2016

## ABSTRACT

The hydrochemical and multivariate statistical techniques such as the factor analysis (FA), the principal component analysis (PCA), and the cluster analysis (CA) were used to determine the main factors and mechanisms controlling groundwater chemistry of Ain Djacer Mio-Plio-Quaternary aquifer, Eastern Algeria. Twenty-one groundwater samples of this aquifer were monitored, during March 2013, with the objective of identifying the geochemical processes and their relation with groundwater quality as well as to get an insight into the hydrochemical evaluation of groundwater. The Piper diagram showed that water facies are of a Ca-$HCO_3$-type near limestone limits and $SO_4$–Na in the center part of the plain. The FA and the PCA revealed two factors that explained 65.1% of the total variance in water quality dataset. The first factor is salinization, which shows strong associations between total dissolved solids (TDS), $SO_4^{2-}$, $Cl^-$, $Ca^{2+}$, $Mg^{2+}$, and $Na^+$. The second factor is mostly associated with chemical fertilizers and represents the high positive load of $NO_3^-$ and $K^+$ and the high negative load of pH and $HCO_3^-$. However, the CA, based on major ion contents, defined three main groups, reflecting different hydrochemical processes. The first group has low salinity (TDS < 1,000 mg/l) and a Ca-$HCO_3$ water type. Samples of this group are mostly located in recharge areas and suggest dissolution of carbonate rocks. The third group has high salinity (TDS > 2,000 mg/l) and a $SO_4$–Na water type. Samples from this group are mostly located in the center part of the plain which constitutes the discharge area. The second group has intermediate salinity (1,000 < TDS < 2,000 mg/l) and Ca-Na water type. These three groups present a salinity increasing with groundwater flow direction. The chemical monitoring allowed relating the presence of nitrates in the study area to agricultural activity. This study highlights that reactions responsible for the hydrochemical evolution in this area are of three categories:

---

*Corresponding author.

(1) dissolution of evaporite minerals, (2) precipitation of carbonate minerals, and (3) anthropogenic action.

*Keywords:* Hydrochemical; Groundwater; Multivariate statistical; Ain Djacer; Algeria

## 1. Introduction

Groundwater quality is influenced by several factors, such as the infiltrating water chemistry, the geological nature of reservoir rock, the decomposition of organic matter, and anthropogenic factors [1]. The dissolution of evaporate and carbonate minerals from rocks surrounding water is the dominant factor controlling the chemical composition [2]. This dissolution brings significant change in water chemical composition, while others, such as silicate minerals, dissolve more slowly and therefore have less effect on this composition [3]. The temperature also has an essential role for groundwater to acquire a given chemistry [4,5]. Besides, temperature plays as a catalyst in chemical reactions and its increasing allows the acceleration of dissolution of minerals present in the geologic environment. Most studies [6–8] showed that groundwater chemistry is mainly a function of the water–rock interaction processes. Understanding the hydrogeochemistry particularities of this aquifer represents a great opportunity to further apply the multivariate statistical analysis which is a quantitative approach allowing to classify groundwater samples, to study correlations between chemical parameters, and to evaluate the similarity between groundwater sampling sites. Multivariate analysis, such as principal component analysis (PCA), factor analysis (FA), and cluster analysis (CA), aiming to interpret the governing processes through data reduction and classification are recognized as powerful tools to deal with the increasing number of hydrochemical parameters. [9,11] used FA to discuss the geochemical evolution and mineralization and groundwater contamination. Additionally, the CA was also used to interpret the hydrochemical data based on factor scores [12–14]. Water samples are classified into three groups. The first, group 1, has low salinity, with increased water–rock interaction, waters in groups 2 and 3 become more saline, changing composition along the direction of groundwater flow. This technique was used to investigate water chemical evolution along the groundwater flow [15–18]. Given the relatively complex setting and geological history of the study area, PCA, FA, and CA could help to distinguish the role of geological and hydrogeological contexts on the chemical evolution.

The use of thermodynamic equilibrium by PHREEQC code allows calculating the saturation indices of some carbonate and evaporates minerals [19]. This expresses the degree of chemical equilibrium between water and the solid matrix in the aquifer and can be considered a parameter to measure dissolution and/or precipitation in water–rock interaction investigation [20,21].

The study area contains more than 300,000 inhabitants distributed across the plain of Ain Djacer. Agriculture is the main activity as well as market gardening cereals (barley and corn). This farming requires the use of fertilizers such as ammonium nitrate, urea, phosphorus and potassium, superphosphate, potassium chloride and to a lesser extent, ammonium sulfate, sodium nitrate, calcium, and potassium sulfate [22]. Surface water resources are very limited and then agricultural and industrial activities depend on groundwater. In Ain Djacer region, the absence of wastewater treatment station induces the discharge of polluted water in nature provoking groundwater contamination. This contamination is increased by leaching of fertilizer in excess uses in agricultural activity.

In this study, the PCA, FA, and CA are applied and used to categorize the spatial variation of groundwater chemistry and to determine the origin of chemical elements existing in Ain Djacer groundwater.

## 2. Materials and methods

### 2.1. Geographical setting

Ain Djacer region is located in Eastern Algeria and centered on latitude 35°49´50´´N and longitude 6°00´ 16´´E. The Ain Djacer catchment area is of 210 km, it is a relatively flat terrain with an average altitude of 850 m above sea level and it is surrounded by hills and mountains culminating up to 1,230 m. This morphology permits the setup of a temporary endorheic drainage system reflecting a semi-arid climate with an annual rainfall of about 260 mm and an annual mean temperature of 15.9°C [23]. The Cretaceous rocks dominated by dolomitic limestone and sandstone constitute the main geological formation in the study area and its surrounding, sometimes with thick marl layers (Fig. 1). Triassic formations are also present and consist mainly of evaporatic rocks. These Triassic formations are made of a mixture of gypsum, halite, clay, deformed dolomite, and sometimes conglomerates.

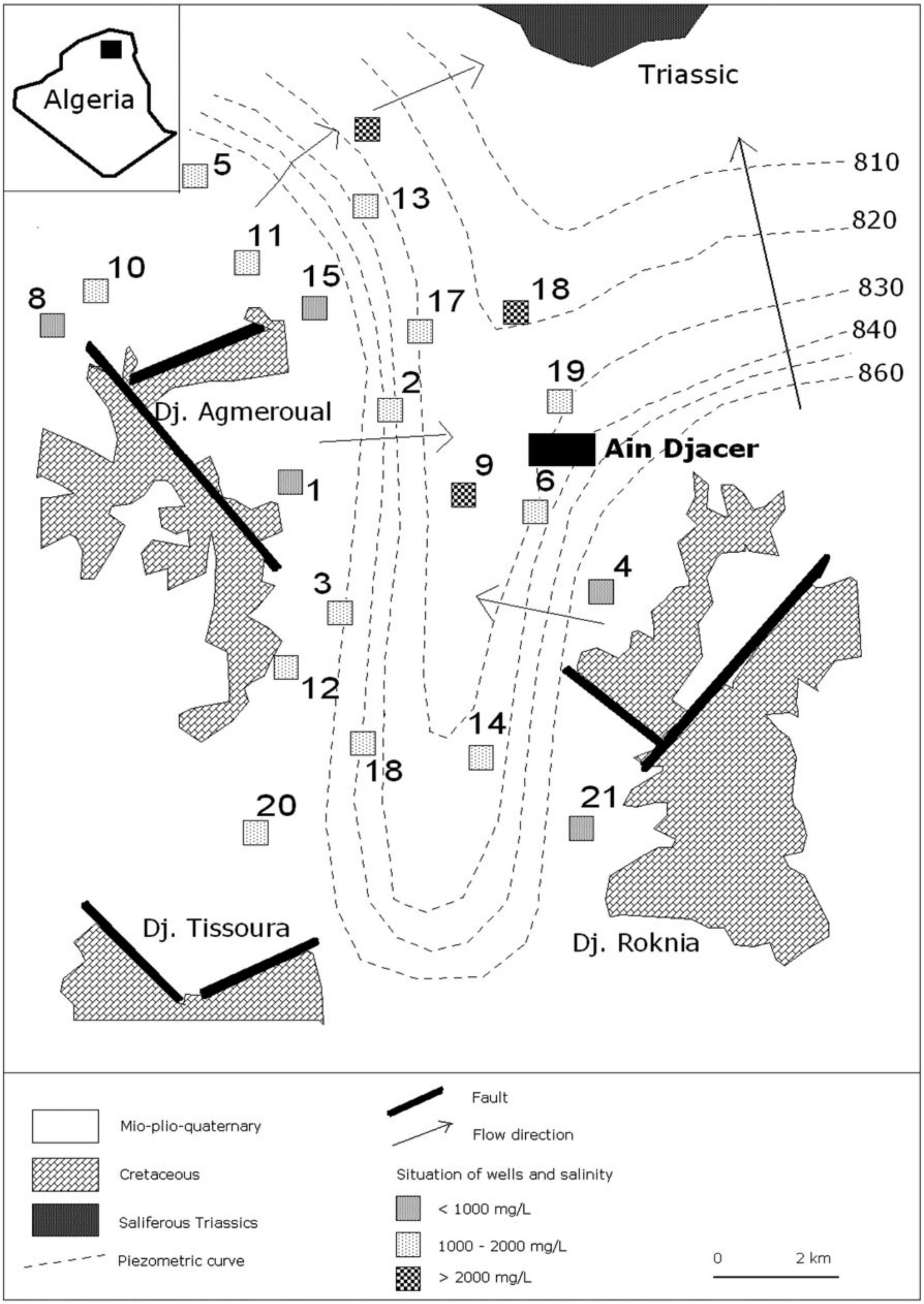

Fig. 1. Map of naturals conditions of Ain Djacer area.

The cover formation is of Mio-Plio-Quaternary age corresponding to red clay, sometimes gypsum overled by lacustrine limestones, conglomerates, and alluvium [24]. The structural setting shows that the study area is located in a syncline structure, filled with Mio-Plio-Quaternary alluvial material. The Triassic formations outcrop by a fracturation network of SW-NE and NW-SE directions.

Previous hydrogeological studies in the area [25–27] highlighted the existence of a shallow aquifer located in the Mio-Plio-Quaternary alluvial formations and resting on clay and gypsiferous marl. The groundwater is recharged by discharge occurring through fractured carbonate formations of the borders, as well as by vertical infiltration of meteoric water arriving into the basin by streams from the bordering hills and mountains surrounding Ain Djacer depression. It is an unconfined aquifer which thickness varies between 10 m at the borders and 100 m in the center of the plain. These alluvial deposits have an average permeability of $10^{-4}$ m/s [28]. Since 1997 and up to 2013, water table has remained at the same level. Within the aquifer, groundwater flows in a west–east direction in the west and east–west direction in the eastern part imposed by inclination of the bedrock. These two flow directions converge toward the center of the plain and then they together get drained to the north. This groundwater situation suggests a recharge from cretaceous limestone and a discharge to the north.

### 2.2. Sample collection and analysis

Twenty-one groundwater water samples were collected in the study area during March 2013 (Fig. 1). These samples were taken from drinking water wells after 15 min of pumping, time sufficient enough for water temperature to stabilize and become representative of the temperature of the whole aquifer. These samples were taken using two polypropylene clean bottles washed with acid, and each sample was immediately filtered *in situ* through 0.45 μm filters on cellulose acetate. Filtrate for cation analyses was poured into 100 $cm^3$ polyethylene bottles and immediately acidified to a $pH < 2$ by the addition of Merck™ ultrapure nitric acid (5 ml 6 $NHNO_3$).

Samples for anion analysis were collected into 250 $cm^3$ polyethylene bottles and were not acidified. All the samples were stored in an ice jacket at a temperature of <4℃ and later transferred to the laboratory of the National Agency for Water Resources of Constantine and stored in a refrigerator at a temperature below 4℃ until they have been analyzed (within 1 week). Temperatures ($T_w$) and pH are measured *in situ* using a multi-parameter WTW (P3 MultiLine pH/LF-SET). Chemical elements that have been analyzed are calcium ($Ca^{2+}$), magnesium ($Mg^{2+}$), sodium ($Na^+$), potassium ($K^+$), chloride ($Cl^-$), bicarbonate ($HCO_3^-$), sulfate ($SO_4^{2-}$), nitrate ($NO_3^-$), and TDS. This was achieved using standard methods as suggested by the American Public Health Association [28]. The ionic balance is generally around 5%.

### 2.3. Methods

Thermodynamic study was carried out using the PHREEQC software to calculate the saturation indices (SI) and ionic speciation of related minerals. The calculation of the SI was conducted according to Eq. (1):

$$\mathrm{IS} = \log\,(\mathrm{IAP}/K) \tag{1}$$

where IAP is the ion activity and $K$ is the equilibrium constant.

The SI of each mineral is used to evaluate the state of saturation of water. This SI also controls the chemistry and equilibrium with the solid phase [21]. In general, the water–rock equilibrium is reached when SI = 0. If it is above 0, water is supersaturated and then precipitation of minerals is necessary to attain equilibrium. However, if it is below 0, water is undersaturated and so dissolution of minerals is necessary to reach the equilibrium [29].

As mentioned, each of the twenty-one sampling site is characterized by ten chemical and physical variables and therefore, 210 data to be synthesized, making the regional hydrogeochemical study a multivariate problem. The multivariate statistical analysis is a quantitative and independent approach of groundwater classification allowing categorizing groundwater samples and making correlations between chemical parameters and groundwater samples [30]. Multivariate statistical techniques permit to simplify and organize large datasets to provide meaningful insight [31]. In this study, three multivariate methods were applied using Excel 2010 (Microsoft Office®) and Statistica version 6.1 (StatSoft Inc., 2004): the CA, the principal components analysis (PCA), and FA. The description of PCA, FA, and CA techniques and the methodology used for their application are detailed in [31]. Multivariate statistical analysis requires that the observations are conformed to normal distribution, and data transforms, such as log-transform [29], rank ordered [32], and normalization [33], are carried out before applying the multivariate statistical analysis. Indeed, before conducting PCA and FA, the Kaiser–Meyer–Olkin (KMO) [34] and Bartlett's sphericity [35] tests were performed on the parameter correlation matrix to examine the validity of PCA and FA. Therefore, these two techniques were conducted for all samples physicochemical data, and the results were 0.705 for the KMO and 891.9 ($p < 0.0001$) for Bartlett's sphericity, indicating that PCA and FA may be useful in providing significant reductions in dimensionality.

PCA and FA were also applied in the treatment of these chemical data. Although PCA and FA are exploratory and descriptive methods, this investigation aims to identify the main factors that control the chemistry of the groundwater [36,37]. This multivariate statistical method has been widely applied to determine the phenomena of the environment around the world [2,9,14,30,31,38] and in Algeria [9,39] and was used successfully to study the hydrogeochemical processes [36,37] and to account for the degree of mutually shared variability between individual pairs of water quality variables. This study is a very helpful tool in promoting research and opening new frontiers of knowledge [40]. Furthermore, Pearson's correlation analysis method is applied to describe the relationship between two hydrogeochemical parameters at a significant level of $p < 0.05$.

CA is also another data reduction method used to classify entities with similar properties. CA comprises a series of multivariate methods which are used to highlight true groups of data [9]. In clustering, the objects are grouped such that similar objects fall into the same class [38]. Besides, CA has the advantage of not requesting any prior knowledge of the number of clusters, which the nonhierarchical method does. A review by [9] suggests Ward's clustering procedure to be the best, because it yields a larger proportion of correct classified observations than any other method. Hence, Ward's clustering procedure is applied in this study.

## 3. Results and discussion

### 3.1. General hydrochemistry

The statistical characteristics of chemical analysis (mg/l) are showed in Table 1. The average of pH is about 7.5 ± 0.25 indicating a low alkalinity of groundwater. This value is in the normal range (6.5–8.0) for freshwater ecosystems [41]. The average temperature is about 18 ± 1.1°C. Higher values of $Ca^{2+}$, $Mg^{2+}$, and $HCO_3^-$ are localized in eastern and the western parts of the plain. The fractured and karstified limestone formations are indeed the recharge boundaries from where water acquires the concentration in $Ca^{2+}$ and $HCO_3^-$ and the low salinity (TDS < 1,000 mg/l) close to the former boundaries. However, higher values in TDS, $Na^+$, $K^+$, $Cl^-$, and $SO_4^{2-}$ are recorded in the center of plain. This indicates that in the part of area, and hydrochemical is highly influenced by the presence of evaporate minerals. This in accordance with groundwater flow allows the concentration in chemical elements and the presence of Mio-Pliocene-Quaternary evaporite rocks as shown in previous studies [42]. In this aquifer, nitrate ($NO_3^-$) concentration varies from 1 to 111 mg/l with a mean value of 30.8 mg/l, which suggests local pollution of water induced by agricultural activity and domestic wastewater. The absence of a WWTP generates the discharge of wastewater in the nature and then increasing groundwater contamination by simple infiltration. Leaching of fertilizer used excessively and without any control in agricultural harms the groundwater quality in the study area.

In the plain boundaries, the order of abundance of the major cations in groundwater is $Ca^{2+} > Mg^{2+} > Na^+ > K^+$ and the abundance of the major anions is $HCO_3^- > Cl^- > SO_4^{2-}$. However, in the center of plain, the order of abundance of the major cations is $Na^+ > Ca^{2+} > Mg^{2+} > K^+$ and $Cl^- > SO_4^{2-} > HCO_3^-$ for the major anions. This is in connection with water–rock interaction and groundwater flow.

### 3.2. Statistical analysis

#### 3.2.1. Correlation analysis

The correlation coefficient is commonly used to measure and establish the relationship between two variables. The Pearson correlation coefficient matrix of ten variables of Ain Djacer area measured samples is given in Table 2. The pH shows significant negative correlation between bicarbonates ($HCO_3^-$, $Ca^{2+}$, and $Mg^{2+}$). This deals with the calc-carbonic equilibrium where pH influences the dissolution of carbonate rocks. TDS values exhibit high positive correlation with $HCO_3^-$, $SO_4^{2-}$, $Cl^-$, $Ca^{2+}$, $Mg^{2+}$, and $Na^+$. This indicates that these elements are the main component of salinity [43]. $Ca^{2+}$, $Mg^{2+}$, and $HCO_3^-$ presented a strong positive correlation indicating a common source. $Cl^-$, $Na^+$, $K^+$, $SO_4^{2-}$, and $Ca^{2+}$ possess a moderate-to-strong positive correlation between each other. This is attributed to the effect of leaching and dissolution of Triassic soluble salts [44]. This dissolution is verified by the indices stability. Thus, the saturation indices indicate an undersaturation state of groundwater with respect to halite, gypsum, and anhydrite and a supersaturation state with respect to all groups of

Table 1
Statistical summary of groundwater hydrochemical parameters

| | pH | TDS | Tw | $Ca^{2+}$ | $Mg^{2+}$ | $Na^{+}$ | $K^{+}$ | $Cl^{-}$ | $SO_4^{2-}$ | $HCO_3^{-}$ | $NO_3^{-}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Min | 7 | 450 | 16 | 90 | 41 | 86 | 1 | 214 | 156 | 140 | 1 |
| Mean | 7.5 | 1,574.3 | 18.0 | 167.8 | 95.5 | 265.0 | 8.0 | 456.8 | 388.1 | 339.7 | 30.8 |
| Max | 8 | 2,600 | 20 | 370 | 166 | 690 | 15 | 995 | 685 | 860 | 111 |
| SD | 0.25 | 483.5 | 1.1 | 73.4 | 35.8 | 171.9 | 3.3 | 238.4 | 117.1 | 118.1 | 30.9 |

Table 2
Pearson's coefficients correlation for physicochemical parameters

| | pH | TDS | $Ca^{2+}$ | $Mg^{2+}$ | $Na^{+}$ | $K^{+}$ | $Cl^{-}$ | $SO_4^{2-}$ | $HCO_3^{-}$ | $NO_3^{-}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| pH | 1 | | | | | | | | | |
| TDS | −0.008 | 1 | | | | | | | | |
| $Ca^{2+}$ | **−0.668** | **0.687** | 1 | | | | | | | |
| $Mg^{2+}$ | **−0.770** | **0.764** | **0.846** | 1 | | | | | | |
| $Na^{+}$ | 0.354 | **0.921** | 0.387 | 0.462 | 1 | | | | | |
| $K^{+}$ | 0.235 | 0.314 | 0.257 | 0.333 | **0.965** | 1 | | | | |
| $Cl^{-}$ | 0.350 | **0.742** | 0.453 | 0.420 | **0.799** | **0.822** | 1 | | | |
| $SO_4^{2-}$ | −0.021 | **0.888** | **0.655** | 0.414 | **0.754** | **0.651** | **0.654** | 1 | | |
| $HCO_3^{-}$ | **−0.630** | **0.942** | **0.812** | **0.635** | 0.414 | 0.397 | 0.345 | −0.041 | 1 | |
| $NO_3^{-}$ | −0.160 | −0.335 | −0.388 | −0.234 | −0.282 | **0.534** | **0.592** | **0.696** | −0.219 | 1 |

Note: The bold values indicate the correlated variables at the 0.05 level.

the main carbonate minerals: calcite, aragonite, and dolomite (Table 3). The $NO_3^{-}$ shows a moderate correlation with $SO_4^{2-}$, $Cl^{-}$, and $K^{+}$ indicating an effect of the anthropogenic activities such as agricultural activities.

### 3.2.2. FA/PCA

The analytical results of FA were performed for the 21 samples and 10 variables (Table 4). The use of FA to water quality assessment has increased, mainly according to the need of obtaining appreciable data reduction for analysis and decision [39]. [45] Proposed to use only the factors with eigenvalues exceeding one. Under this criterion, only those factors with eigenvalues greater or equal to 1 will be accepted as possible sources of variance in data, with the highest priority ascribed to the factor that has the highest eigenvector sum. The reason for choosing one is that a factor must have a variance as large as that of a single standardized original variable to be acceptable [9]. Two factors which explain 75.1% of the total variance are obtained based on the criteria for factor selection (eigenvalues > 1.0). The parameter weights for the two components from the FA of the dataset are given in Table (4). The factor loading is classified as "strong", "moderate", and "weak" corresponding to absolute loading values of 0.75, 0.75–0.50, and 0.5, respectively [46]. [47] reported that factor scores can be related to the intensity of a particular described process. Extreme negative values (close to –1) indicate areas essentially unaffected by the process, and the positive scores (close to + 1) are the areas the most affected.

Factor 1 explains the greatest amount of the variance (57%) and is characterized by highly positive loadings in TDS, $Ca^{2+}$, $Mg^{2+}$, $Na^{+}$, $Cl^{-}$, $SO_4^{2-}$, and $HCO_3^{-}$, which were 0.735, 0.732, 0.654, 0.853, 0.877, 0.742, and 0.934, respectively. The TDS association with these elements is due to the different hydrochemical processes that increase the water salinity. Due to the association of Ca-Mg-$HCO_3$, Ca-$SO_4$, and Na-Cl, factor 1 is defined as the "salinity" factor in reference to $CaCO_3$, $MgCa(CO_3)_2$, $CaSO_4$, and NaCl salts. Similar observations have also been reported by [48].

Factor 2 explains 18.1% of the total variance of the dataset and shows significant highly positive loadings in $K^{+}$ and $NO_3^{-}$, which predicts the association of this factor with chemical fertilizers. Large amounts of fertilizers, such as urea and commercial composites, have been applied for long time in agriculture in the study area which remains a cereal vocation. The great

Table 3
Statistical summary of thermodynamic speciation (saturation indices of some minerals and the partial pressure of $CO_2$) calculations using PHREEQC

| | Anhydrite | Gypsum | Halite | Calcite | Aragonite | Dolomite | Log ($P_{CO2}$) |
|---|---|---|---|---|---|---|---|
| Min | −1.44 | −1.94 | −6.09 | −0.02 | −0.07 | −0.49 | −2.17 |
| Mean | −1.03 | −1.05 | −4.59 | 0.14 | 0.12 | 0.42 | −2.28 |
| Max | −0.83 | −0.50 | −2.8 | 0.74 | 0.59 | 1.45 | −1.41 |
| SD | 0.2 | 0.13 | 0.27 | 0.20 | 0.19 | 0.22 | 0.14 |

plowing before the rain period permits the application of chemical fertilizers used by plants during its development. The autumn precipitations favorite the leaching of some of these fertilizers which will be transported to groundwater. The association of these two parameters reflects the influence of fertilizers on groundwater pollution and can thus be termed “the agriculture contamination factor” [49].

### 3.2.3. Cluster analysis

There are two types of CA analysis: R and Q-modes. The advantage of using the hierarchical method of CA is due to the fact that it does not request any prior knowledge of the number of cluster [9]. In the present paper, CA was applied using the Euclidian distance as a measure distance between samples and Ward’s method as a linkage rule for the classification of hydrogeochemical data of the Ain Djacer area. The dendrogram of the nine physicochemical parameters based on the CA can be divided into three main clusters (Fig. 2). First cluster group shows narrow association between $Ca^{2+}$ and $HCO_3^-$. Second cluster group shows narrow association between $Na^+$, $K^+$, and $Ca^{2+}$ indexes. These indexes have high values in most of the sampling sites. The third group shows the closest association between TDS, $SO_4^{2-}$, $Mg^{2+}$, and $NO_3^-$. These findings corroborate with the results of correlation analysis. The dendrogram of sampling sites produced four major groupings (Fig. 3). Three main groups can be verified in this figure. The first cluster group (G1) shows close association between sites 1, 4, 8, 15, and 21, with low salinity (TDS = 714 mg/l) characterized by $Ca^{2+}$ and $HCO_3^-$. It is located near the eastern and western outcrops. The third cluster group (G3) shows close association between sites 7, 9, and 16, with high salinity (TDS = 2,580 mg/l) characterized by Na-Ca and Cl–$SO_4$. This group is influenced by the Triassic saliferous outcrops located to the North and by groundwater flow. The second cluster group (G2) shows close association between sites 2, 3, 5, 6, 10, 11, 12, 13, 14, 17, 18, 19, and 20, with intermediate salinity (TDS = 1,370 mg/l) characterized by $Na^+$ and $Cl^-$. This logic of succession (G1-G2-G3) is in a good agreement with the flow path that allows the concentration of elements during its movement within the aquifer. This succession can be reviewed with the F1-F2 diagram. Fig. 4 shows the plot of the two factor scores (F1 vs. F2) for datasets. The distribution of factor scores for datasets suggests a continuous variation of chemical

Table 4
Variance and component matrixes

| Variable | Factor 1 | Factor 2 |
|---|---|---|
| pH | −0.453 | 0.153 |
| TDS | **0.735** | 0.346 |
| $Ca^{2+}$ | **0.732** | 0.301 |
| $Mg^{2+}$ | **0.654** | 0.265 |
| $Na^+$ | **0.853** | 0.028 |
| $K^+$ | 0.180 | **0.652** |
| $Cl^-$ | **0.877** | 0.007 |
| $SO_4^{2-}$ | **0.742** | 0.060 |
| $HCO_3^-$ | **0.934** | 0.180 |
| $NO_3^-$ | 0.260 | **0.650** |
| Eigenvalue | 3.89 | 1.40 |
| Variance (%) | 57.0 | 18.1 |
| Cumulative (%) | 57.0 | 75.1 |

Note: The bold values indicate the moderate-to-high loadings factors (>0.65).

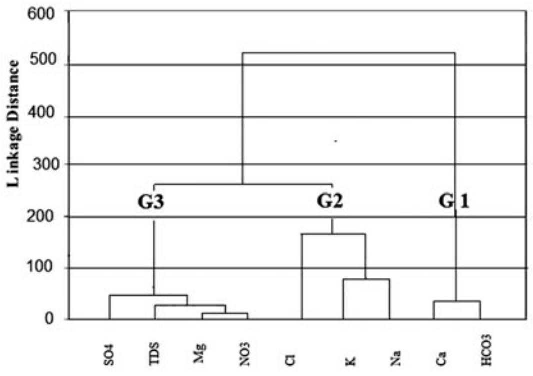

Fig. 2. Dendrogram of CA of variables.

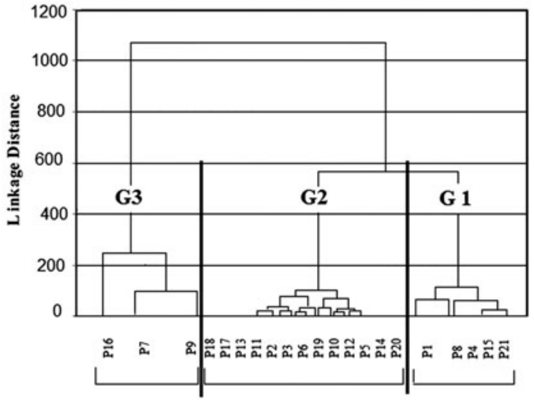

Fig. 3. Dendrogram of CA of samples hydrochemistry.

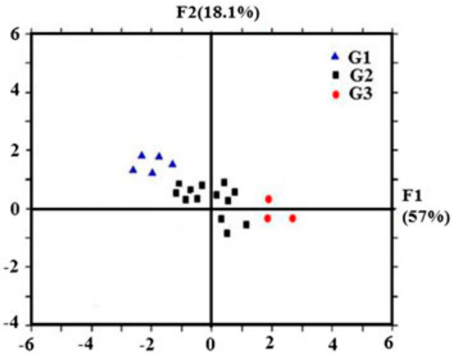

Fig. 4. Plots of PCA scores for F1 vs. F2.

and physical properties of some samples for datasets for G1 to G2 and to G3. This is in connection with the water–rock interaction and groundwater flow showing the groundwater chemistry acquiring process.

## 4. Conclusion

This paper is a contribution to the study of hydrochemical functioning of the Ain Djacer shallow aquifer using multivariate statistical techniques including FA, PCA, and CA. Interpretation of analytical data showed that the abundance of major ions in the boundaries tends as follows $Ca^{2+} > Mg^{2+} > Na^{+} > K^{+}$ and $HCO_3^{-} > Cl^{-} > SO_4^{2-}$. However, in the center of the plain, the order of abundance of the major ion is $Na^{+} > Ca^{2+} > Mg^{2+} > K^{+}$ and $Cl^{-} > SO_4^{2-} > HCO_3^{-}$. The TDS values exhibit high positive correlation with $HCO_3^{-}$, $SO_4^{2-}$, $Cl^{-}$, $Ca^{2+}$, $Mg^{2+}$, and $Na^{+}$. This is in agreement with the saturation indices which show saturation in mineral carbonates and under-saturation in evaporite minerals. The presence of $NO_3^{-}$ in groundwater is related to agricultural activity. The FA and PCA identified two factors accounting for 75.1% of the total variance in the dataset. Their loadings allowing the interpretation of the hydrochemical processes that take place in the area. Factor 1 is defined as the "salinity" factor, whereas factor 2 is associated with chemical fertilizers. Three major water types are suggested by the Q-mode CA analysis following salinity increase simultaneously to flow direction. These techniques can successfully be used to derive information from dataset about the possible influences of environment on groundwater quality and also to identify natural groupings in these data.

## Acknowledgements

The authors would like to thank Dr Miriam Balaban Editor in Chief at Desalination and Water Treatment Journal for his patience and greatly appreciate the constructive and thoughtful comments of the anonymous(s) reviewer(s).

## References

[1] S.K. Frape, P. Fritz, R.H. McNutt, Water-rock interaction and chemistry of groundwaters from the Canadian Shield, Geochim. Cosmochim. Acta 48 (1984) 1617–1627.

[2] M. Ghodbane, A. Boudoukha, L. Benaabidate, Hydrochemical and statistical characterization of groundwater in the Chemora area, Northeastern Algeria, Desalin. Water Treat. 57(32) (2016) 14858–14868.

[3] R.M. Garrels, F.T. MacKenzie, Origin of the chemical compositions of some springs and lakes, in: Equilibrium Concepts in Natural Waters, Am. Cancer Soc., Washington, DC, 1967.

[4] J.D. Hem, Study and interpretation of the chemical characteristics of natural water, third ed., US Geol. Surv. Water Supply Paper 2254, 1989.

[5] A. Boudoukha, M. Athamena, Characterization of the thermal waters of all South Sétif, Algeria J. Water Sci. 25 (2012) 103–119.

[6] J. Hartmann, Z. Berner, D. Stüben, N. Henze, A statistical procedure for the analysis of seismotectonically induced hydrochemical signals: A case study from the Eastern Carpathians, Romania, Tectonophysics 405 (2005) 77–98.

[7] S. Adams, R. Titus, K. Pietersen, G. Tredoux, C. Harris, Hydrochemical characteristics of aquifers near Sutherland in western Karoo, South Afr. J. Hydrol. 241 (2001) 91–103.

[8] W.D. Alberto, D.M. Del Pilar, A.M. Valeria, P.S. Fabiana, H.A. Cecilia, B.M. De Los Angeles, Pattern recognition techniques for the evaluation of spatial and temporalvariations in water quality. A case study: Suquia River Basin (Cordoba-Argentina), Water Res. 35 (2001) 2881–2894

[9] L. Belkhiri, A. Boudoukha, L. Mouni, T. Baouz, Application of multivariate statistical methods and inverse geochemical modeling for characterization of

groundwater—A case study: Ain Azel plain (Algeria), Geoderma 159 (2010) 390–398.
[10] S.X. Meng, J.B. Maynard, Use of statistical analysis to formulate conceptual models of geochemical behavior: Water chemical data from the Botucatu aquifer in São Paulo state, Brazil, J. Hydrol. 250 (2001) 78–97.
[11] I.M. Farnham, K.J. Stetzenbach, A.K. Singh, K.H. Johannesson, Deciphering groundwater flow systems in Oasis Valley, Nevada, using trace element chemistry, multivariate statistics, and geographical information system, Math. Geol. 32 (2003) 943–968.
[12] H.J. Suk, K.K. Lee, Characterization of a ground water hydrochemical system through multivariate analysis: Clustering into ground water zones, Ground Water 37 (1999) 358–366.
[13] R. Reghunath, T.R.S. Murthy, B.R. Raghavan, The utility of multivariate statistical techniques in hydrogeochemical studies: An example from Karnataka, India, Water Res. 36 (2002) 2437–2442.
[14] J.H. Kim, R.H. Kim, J.H. Lee, T.J. Cheong, B.W. Yum, H.W. Chang, Multivariate statistical analysis to identify the major factors governing groundwater quality in the coastal area of Kimje, South Korea, Hydrol. Processes 19 (2005) 1261–1276.
[15] L.N. Plummer, J.F. Busby, R.W. Lee, B.B. Hanshaw, Geochemical modeling of the madison aquifer in parts of Montana, Wyoming, and South Dakota, Water Resour. Res. 26 (1990) 1981–2014.
[16] E. Rosenthal, B.F. Jones, G. Weinberger, The chemical evolution of Kurnub. Group paleowater in Sinai-Negev-a mass balance approach, Appl. Geochem. 27 (1998) 1–17.
[17] L.E. Eary, D.D. Runnels, K.J. Esposito, Geochemical controls on groundwater composition at the Cripple Creek mining district, Colorado, Appl. Geochem. 18 (2002) 1–24.
[18] C. Güler, G.D. Thyne, Hydrologic and geologic factors controlling surface and groundwater chemistry in Indian Wells-Owens Valley area, southeastern California, USA, J. Hydrol. 285 (2004) 177–198.
[19] D.L. Parkhurst, C.A.J. Appelo, PHREEQC for Windows Version 1.4.07. A Hydrogeochemical Transport Model, US Geological Survey Software, Washington, DC, 1999.
[20] R.J. Gibbs, Mechanisms controlling world water chemistry, Science 170 (1970) 1088–1090.
[21] J.I. Drever, The Geochemistry of Natural Waters, thired, Prentice Hall, New Jersey, NJ, 1997.
[22] A. Boudoukha, M. Boulaarak, Water pollution Grouz dam by nutriments (Eastern Algeria), Bull. Serv. Geol. Nat. 24 (2013) 139–149.
[23] A. Mebarki, Water Resources and Management in Algeria. The Watersheds of the East, Office of University Publications, Algeria, 2009.
[24] J.M. Vila, The Alpine chain Eastern Algeria and the Algerian–Tunisian border, PhD Thesis in, Natural Sci. Univ. Pierre and Marie Curie, Paris VI, 1980.
[25] N. Chabour, Hydrogeological Study of the Plain of Ain Djasser (Boughzel), Magister memory. University of Constantine, Algeria, 1997.
[26] S. Bencer, State of Water Resources in the Region of Ain Djacer and Estimating the Chemical Quality of Groundwater, Batna. Magister memory. University of Batna, Algeria, 2005.
[27] D. Ziani, Study of Nitrogen flow in the plain of Ain Djacer, Magister memory. University of Batna, Algeria, 2009.
[28] American Water Works Association and Water Environment Federation, Standard Methods for the Examination of Water and Wastewater, twenty-first ed., American Public Health Association, Washington, DC, 2005, 20001–3710.
[29] D.L. Parkhurst, C.A.J. Appelo, PHREEQC for Windows Version 1.4.07. A Hydrogeochemical Transport Model, US Geological Survey Software, Washington, DC, 1999.
[30] S.M. Yidana, D. Ophori, B. Banoeng-Yakubo, A multivariate statistical analysis of surface water chemistry data—The Ankobra basin, J. Environ. Manage. 88 (2008) 697–707.
[31] V. Cloutier, R. Lefebvre, R. Therrien, M.M. Savard, Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system, J. Hydrol. 353(3–4) (2008) 294–313.
[32] M. Laaksoharju, C. Skårman, E. Skårman, Multivariate mixing and mass balance (M3) calculations, a new tool for decoding hydrogeochemical information, Appl. Geochem. 14 (1999) 861–871.
[33] V. Cloutier, Origin and geochemical evolution of groundwater in the Paleozoic Basses-Laurentides sedimentary rock aquifer system, St. Lawrence Lowlands, Quebec, Canada. PhD Thesis, INRS-Eau, Terre & Environnement, Quebec, Canada, 2004 (in French and English).
[34] C. Miller, J. Denis, S. Ator, J. Brakebill, Nutrients in streams during baseflow in selected environmental sevfings of the potomac river basin, J. Am. Water Resour. Assoc. 33 (1997) 1155–1171.
[35] S. Ravichandrana, R. Ramanibai, N.V. Pundarikanthan, Ecoregions for describing water quality patterns in Tamiraparani basin, South India, J. Hydrol. 178 (1996) 257–276.
[36] H.F. Kaiser, A Second Generation Little Jiffy, Psychometrika, 35 (1970) 401–415.
[37] M.S.A. Bartlett, A note on multiplying factors for various chi-squared approximations, J. R. Stat. Soc. Ser. B, 16 (1954) 296–298
[38] P. Dagnélie, Theoretical and Applied Statistics, Vol. 2, Inferences One- and Two-dimensional. Edis. Boeck & Larcier, Bruxelles, 2006.
[39] M. Laaksoharju, I. Gurban, C. Skarman, E. Skarman, Multivariate mixing and mass balance (M3) calculations, a new tool for decoding hydrogeochemical information, Appl. Geochem. 14 (1999) 861–871.
[40] A. Tiri, N. Lahbari, A. Boudoukha, Multivariate statistical analysis and geochemical modeling to characterize the surface water of oued Chemora basin, Algeria, Nat. Resour. Res. 23 (2014) 379–391.
[41] D.M. Joshi, N.S. Bhandari, A. Kumar, N. Agrawal, Statistical analysis of physicochemical parameters of water of river Ganga in Haridwar district, Rasayan J. Chem. 2 (2009) 579–587.
[42] S.A.A. Rashid, M.B. Gasim, M.E. Toriman, H. Juahir, M.K.A. Kamarudin, A. Azid, N.A. Abd Aziz, Water quality deterioration of Jinjang River, Kuala Lumpur: Urban risk case water pollution. Arab World Geogr., 16 (2013) 349–362

[43] A. Boudoukha, Hydrogeological Study and Modeling of the Aquifer System of El Eulma-Aın Lahdjar (Eastern Region of Setif. Algeria), PhD Thesis, Lab. Structural Geology. Univ. Franche Comté, Besancon, 1988.

[44] P. Olive, The System $CO_2/H_2O*CaCO_3$ and Sulfate-Sulfide System, practical Handbook, C. R. G. Thonon les Bains, 1976.

[45] R.A. Freeze, J.A. Cherry, Groundwater, Prentice Hall Inc., Englewood Cliffs, New Jersey, NJ, 1979.

[46] H.F. Kaiser, A second generation little jiffy, Psychometrika 35 (1970) 401–415.

[47] C.P. Unmesh, K.S. Sanjay, R. Prasant, B.N. Binod, B. Dinabandhu, Application of factor and cluster analysis for characterization of river and estuarine water systems—A case study: Mahanadi River (India), J Hydrol. 331 (2006) 434–445.

[48] M.G. Dalton, S.B. Upchurch, Interpretation of hydrochemical facies by factor analysis, Ground Water 16 (1978) 228–233.

[49] N.S. Rao, D.J. Devadas, K.V.S. Rao, Interpretation of groundwater quality using principal component analysis from Anantapur district, Andhra Pradesh, India Environ. Geosci. 13 (2006) 239–259.

[50] A. Zghibi, A. Merzougui, L. Zouhri, J. Tarhouni, Understanding groundwater chemistry using multivariate statistics techniques to the study of contamination in the Korba unconfined aquifer system of Cap-Bon (North-east of Tunisia), J. Afr. Earth. Sci. 89 (2014) 1–15.