

Combined Use of Multivariate Statistical Analysis and Hydrochemical Analysis for Groundwater Quality Evolution: A Case Study in North China Plain

Rong Ma, Jiansheng Shi*, Jichao Liu, Chunlei Gui

Institute of Hydrogeology and Environmental Geology, Shijiazhuang 050061, China

ABSTRACT: Understanding the controlling factor of groundwater quality can enhance promoting sustainable development of groundwater resources. To this end, multivariate statistical analysis (MA) and hydrochemical analysis were introduced in this work. The results indicate that the canonical discriminant function with 7 parameters was established using the discriminant analysis (DA) method, which can afford 100% correct assignment according to the 3 different clusters (good water (GW), poor water (PW), and very poor water (VPW)) obtained from cluster analysis (CA). According to factor analysis (FA), 8 factors were extracted from 25 hydrochemical elements and account for 80.897% of the total data variance, suggesting that groundwater with higher concentrations of sodium, calcium, magnesium, chloride, and sulfate in southeastern study area are mainly affected by the natural process; the higher level of arsenic and chromium in groundwater extracted from northwestern part of study area are derived by industrial activities; domestic and agriculture sewage have important contribution to copper, iron, iodine, and phosphate in the northern study area. Therefore, this work can help identify the main controlling factor of groundwater quality in North China plain so as to make better and more informed decisions about how to achieve groundwater resources sustainable development.

KEY WORDS: factor, groundwater quality, hydrochemical variable, industrial activity, multivariate statistical analysis.

1 INTRODUCTION

Groundwater is one of the most important natural resources in North China Plain (NCP), which provides 56% of the drinking water supply for more than 100 million people; meanwhile, it also used for agriculture irrigation and industrial purposes (Yuan and Shen, 2013). Currently, groundwater contamination is recognized as one of the most serious environmental problems in NCP, which has attracted worldwide attention. “Previous investigation have indicated that in more than 40 cities groundwater is polluted in various degrees”, says Shi (2007), who was at the 34th International Association of Hydrogeologists (IAH) congress, “however, the problem of NCP is orders of magnitude bigger than anywhere else”. To achieve the sustainable groundwater resources development, it is necessary to characterize the situation of groundwater quality and detect the origin of each hydrochemical variable observed in groundwater. At present, various methods and models have been applied to evaluate the status of groundwater quality (Pang et al., 2013; Wang Y et al., 2013; Ma et al., 2012; Wang M et al., 2012; Xie et al., 2012; Weng et al., 2011), which have been accepted by the decision maker. However, the controlling factors of hydrochemical

composition, such as when dissolution, erosion, sedimentation, ion exchange, anthropogenic activities, and climate change, are difficult to identify exactly by the previous studies (Osman et al., 2012). The useful tools that are employed to detect the pollution sources of groundwater are multivariate statistical analysis (MA). The MA, such as cluster analysis (CA), factor analysis (FA), and discriminant analysis (DA), has been widely employed to characterize the spatial variation of groundwater quality and identify their influencing factors (Wang et al., 2013; Ujević Bošnjak et al., 2012; Zhao et al., 2012; Ariyibi et al., 2010; Dantu, 2010; Žibret and Šajn, 2010; Ravindra et al., 2008). On the other hand, there are many articles that have investigated the contribution of natural process (minerals dissolution and ion exchange) to hydrochemical composition using the hydrochemical analysis method (Li et al., 2013; Yang et al., 2012; Yidana and Yidana, 2010; Jalali, 2007; Stigter et al., 2006). For example, Xie et al. (2012) applied the molar Cl/Br ratio to detect the major origin of arsenic in Datong Basin, North China; Hamzaoui-Azaza et al. (2011) attempted to detect the origin of some elements in groundwater in southeastern Tunisia according to the Na/Cl, Ca/Mg, and Ca/SO₄ ratio.

In this article, the MA and hydrochemical analysis methods were used to characterize the groundwater quality evolution process, and to identify the controlling factors, which govern the chemical composition of groundwater in NCP.

Corresponding author: tiger7886@gmail.com

© China University of Geosciences and Springer-Verlag Berlin Heidelberg 2014

Manuscript received June 13, 2013.

Manuscript accepted October 25, 2013.

Ma, R., Shi, J. S., Liu, J. C., et al., 2014. Combined Use of Multivariate Statistical Analysis and Hydrochemical Analysis for Groundwater Quality Evolution: A Case Study in North China Plain. *Journal of Earth Science*, 25(3): 587–597, doi:10.1007/s12583-014-0446-2

2 MATERIALS AND METHODS

2.1 The Study Area

The study area is located in the west NCP, lying between longitudes 114°28'E and 115°26'E and latitudes 37°41'N and 38°12'N and is shown in Fig. 1. The overall area is 3 011 km², its terrains decline gently from northwest to southeast and the elevation ranges from 20 to 110 m. This region is characterized by

a semi-arid continental climate, with annual average air temperature range from 5.6 to 7.8 °C. During the period of 2001–2011, the average precipitation is approximately 400–800 mm; the annual potential evaporation far exceeds the annual precipitation with a mean value of 1 300 mm, and about 75% of the evaporation occurs from June to September (Ma et al., 2012).

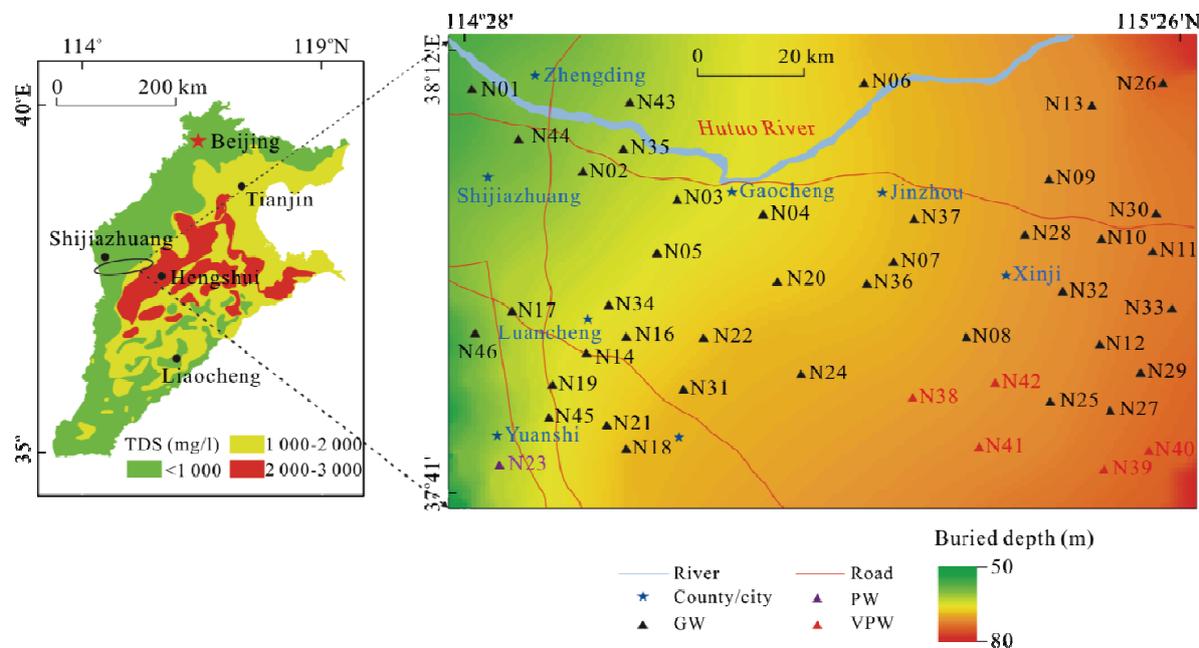


Figure 1. Study area and sample locations.

Regional geological studies revealed that the fluvial sediments from the Hutuo, Huaisha, and Xiao rivers are mainly deposited in the central and eastern parts of the study area, while alluvial and pluvial sediments in the western part in the front of the Taihang Mountains. From the hydrogeology point of view, four aquifers are divided according to their lithology. The lithology of the first aquifer is silty sand and medium fine sand, with a thickness of 15–40 m. The buried depth and TDS concentration of the second aquifer is described in Fig. 1, it can be seen from this figure that the bottom of the second aquifer is at depth of 50–80 m below ground surface showing a gentle slope from north to south. This aquifer is the current main exploitation sector for drinking, agriculture irrigation, and industrial uses due to the first aquifer has been depleted over the past decades. The grains become finer from west to east in this aquifer. The buried depth of the bottom of the third aquifer is 120–180 m, and the bottom of the fourth aquifer is 250–420 m, with consists of fine sand and medium sand. The regional geology and hydrogeology conditions are not described for the length of this article; their detailed information has been reported by the previous reports (Shi et al., 2011; Wang S Q et al., 2011; Hu et al., 2010).

2.2 Sample Collection

A total of 46 groundwater samples were collected in 250 mL clean polyethylene bottles between 2010 and 2011 in the second aquifer for analysis of 25 hydrochemical parameters. In

the case of boreholes and wells, the water samples were collected after pumping for 10–20 min. This was done to remove groundwater stored in the well. For all samples collected, parameters such as electrical conductivity (EC), temperature (T), total dissolved solids (TDS) and pH values were measured in situ using a multi parameter water quality meter. It is necessary to test these elements on-site due to they are likely to change during transport. Preservation and transportation of these samples to the laboratory followed standard methods. Rest of the variables was analyzed at the laboratory of the institute of Hydrogeological and Environmental Geology immediately after transportation to the laboratory. Hydrochemistry analysis was done in triplicates as per the standard recommended methods using double glass distilled water and analytical grade (AR) chemicals. Total hardness (TH), chloride, carbonate, and bicarbonate content were estimated by volumetric analyses; calcium, magnesium, sodium, potassium, sulfate, and phosphate were determined by photometric method; nitrate determination was performed using the brucine method; and inductively coupled mass spectrometry (ICP-MS) was employed to determine the concentrations of trace elements (GB/T 5750, 2006; WHO, 2006); heavy metals were measured using atomic fluorescence spectrometry (AFS). In this article, with the exception of EC and pH, all the hydrochemical variables were expressed in mg/L.

2.3 Methods

The MA approach, such as CA, FA, and DA, can be employed to analyze the hydrochemical variables without losing important information (Wang et al., 2013). It is widely used to identify the influencing factor of groundwater (Wang et al., 2013; Ariyibi et al., 2010; Dantu, 2010; Ravindra et al., 2008), providing a useful tool for sustainable groundwater resources management as well as contamination prevention. The relative concepts and detailed information of MA method are described as follows.

CA is one of the most popular tools for analyzing hydrochemical composition. In order to recognize the existence of different groups of groundwater, CA is employed to split the standardized hydrochemistry data into different clusters with similar hydrochemical variables characteristics so that each group represents a specific hydrochemistry process in the study area. There are two types of cluster analysis: R and Q modes, Q-mode CA is performed on the hydrochemistry data to group the samples in terms of hydrochemical variables in this work. There are several steps involves in CA method, first, the most similar samples are grouped after standardization based on Z-scores; subsequently, these initial clusters are merged according to their analogous behavior; finally, all subgroups are merged into a single group with the samples similarity decreases (Kamble and Vijay, 2011).

DA is a MA method that can be applied to classify the groundwater sample into a given groups using a few parameters. It performs on raw data and establishes a discriminant function for each cluster. The ratio between the correctly classified numbers and the total number of parameters is defined as the hit ratio, which is used to determine the relations between the measured parameters. In DA, the discriminate score of each sample can be established using the canonical discriminate function, which is defined as follows (Jang et al., 2012).

$$f_{ki} = C_i + \sum_{j=1}^n w_{ij} p_{kj} \quad (1)$$

where f_{ki} is the discriminate score of k th sample for i th canonical discriminate function, $k=1,2,\dots,m$, m is the number of groundwater samples; $i=1,2,\dots,l$, l is the number of canonical discriminate function; n denotes the number of parameters in canonical discriminates function; p_{kj} represents the concentration of j th hydrochemical variable of k th sample, $j=1,2,\dots,n$; and w_{ij} is the weight coefficient of p_{kj} in i th canonical discriminate function; C_i denotes the constant value in i th canonical discriminate function.

The primary purpose of FA is to analyze the interrelationships within a set of variables by creating one or more new factors each representing a new group. To reduce the data to an easily interpretable form, the first step is computation of a correlation coefficient matrix, which requires normal distribution of all variables. The eigenvalue are computed for all the new factors. The factor extraction has been done with a minimum acceptable eigenvalue as greater than 1. Factor loading can measure the degree of closeness between the variables and the factor. The largest loading, either positive or negative, demonstrates the meaning of the dimension; positive loading suggests that the contribution of the variables increases with the increasing load-

ing in a dimension; and negative loading indicates a decrease. Currently, FA method is widely used to detect the main process controlling the spatial variation of groundwater quality (Ujević Bošnjak et al., 2012; Aryafar and Doulati, 2011). In this paper, the R-mode FA was carried out using a set of data consisting of 46 groundwater samples to identify the controlling factors that influence the hydrochemical composition of NCP.

3 RESULTS AND DISCUSSIONS

3.1 Hydrochemical Analysis

Knowledge on hydrochemical variables (Fig. 2) is more important to characterize the chemical evolution of groundwater for identifying its controlling factor.

3.1.1 The influence of chloride salt

A 1 : 1 relationship between chloride and total cations (total cations= $\text{Ca}^{2+}+\text{Mg}^{2+}+\text{Na}^{+}+\text{K}^{+}$) would be expected from the dissolution of chloride salt. The relationship between chloride and total cations (Fig. 2a) shows all of the samples deviate from the 1 : 1 line, indicating chloride salt dissolution is not the major source of total cations. The upward deviations from the mixing line seem to be caused by another process, such as silicate and carbonate minerals dissolution and ion exchange.

3.1.2 Dissolution of silicate mineral and ion exchange

($\text{Ca}^{2+}+\text{Mg}^{2+}$)-(HCO₃⁻+SO₄²⁻) in meq/L was plotted against ($\text{Na}^{+}+\text{K}^{+}$)-Cl⁻ meq/L that was employed to evaluate the possible role of silicate mineral dissolution and ion exchange in the chemistry composition of groundwater from NCP (Fig. 2b). If this water-rock interaction is significant in hydrochemical composition controlling, the relation between these two parameters should be linear with a slope of -1.0 (Jalali, 2005; Kim et al., 2004). In Fig. 2b, data plots for all samples are close to a straight line ($R^2=0.913$) with a slope of -0.946. In addition, an increase in sodium concentration without an associated increase in chloride concentration was observed in Fig. 2c, which was regarded as an evidence of this process existing. Therefore, these two figures indicated that most of the samples were mainly influenced by silicate mineral dissolution and ion exchange.

3.1.3 The origin of sodium and potassium

The sodium-chloride relationship was widely used to identify the mechanisms for acquiring salinity and saline intrusions in semi-arid regions (Dixon and Chiswell, 1992). The dissolution of halite in water releases equal concentrations of sodium and chloride into the solution. 10.86% of samples fall around the equiline in Fig. 2c, the others deviate from the expected 1 : 1 relation indicating that the sodium may be derived from another process. There is a high correlation coefficient ($r=0.905$) between sodium and sulfate, reflecting that the excess of sodium in these samples mostly resulted from the dissolution of sodium sulfate minerals.

A molar $\text{Na}^{+}/\text{Cl}^{-}$ ratio greater than 1 is typically employed to reflect sodium released from silicate mineral dissolution (Meybeck, 1987). The molar $\text{Na}^{+}/\text{Cl}^{-}$ ratio in most groundwater samples (91.3%) are greater than 1, indicating silicate dissolution can be a probable origin of sodium in groundwater. For example, metamorphic rocks are widely distributed in

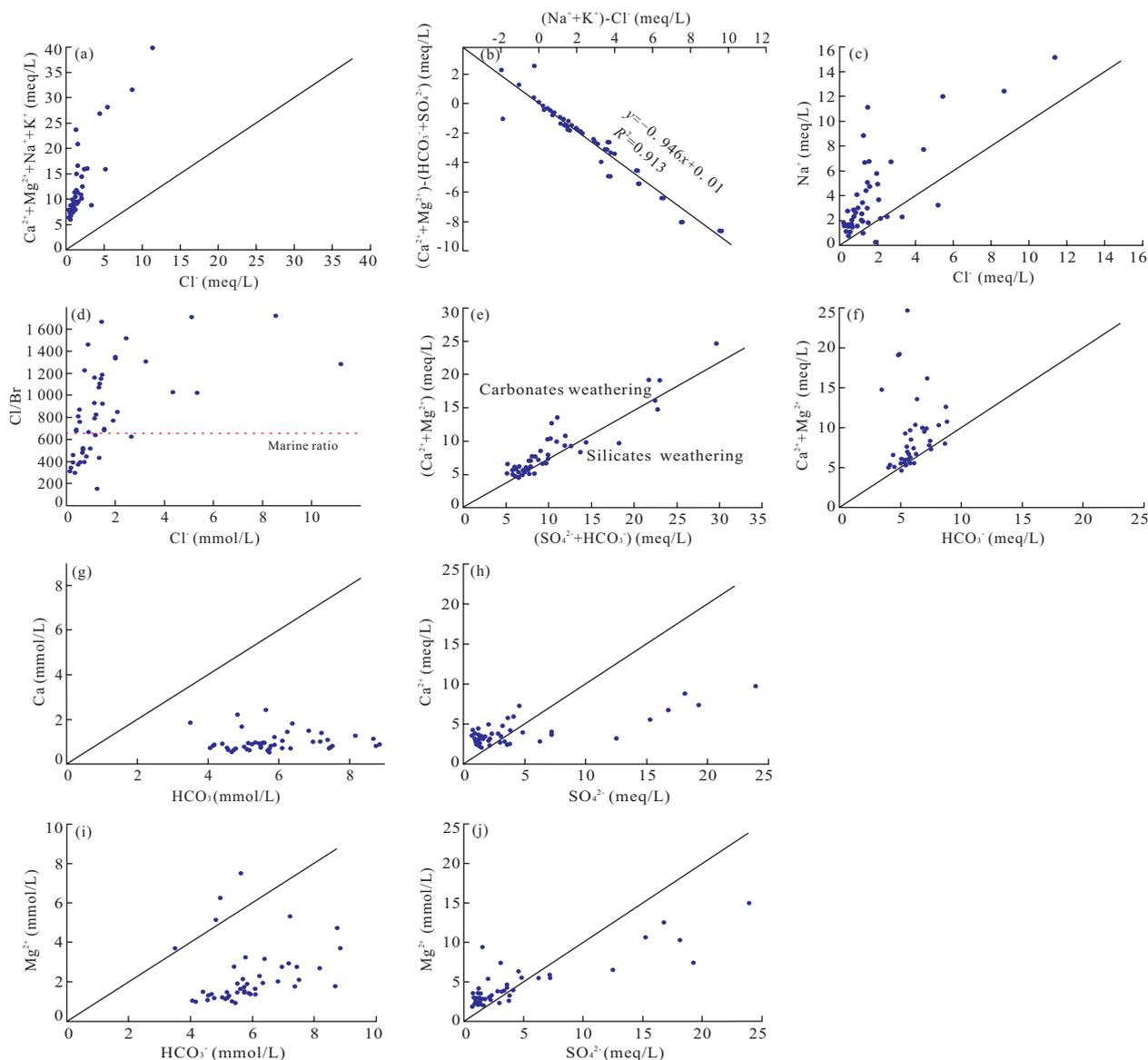


Figure 2. Relationship between major cations and anions. (a) Relationship between total cations and chloride in groundwater; (b) relationship between $(Ca^{2+}+Mg^{2+})-(HCO_3^-+SO_4^{2-})$ and $(Na^++K^+-Cl^-)$ in groundwater; (c) relationship between Na^+ and Cl^- in groundwater; (d) relationship between Cl^- concentration and mmolar Cl^-/Br^- ratios in groundwater; (e) relationship between $Ca^{2+}+Mg^{2+}$ and $HCO_3^-+SO_4^{2-}$ in groundwater; (f) relationship between $Ca^{2+}+Mg^{2+}$ and HCO_3^- in groundwater; (g) relationship between Ca^{2+} and HCO_3^- in groundwater; (h) relationship between Ca^{2+} and SO_4^{2-} in groundwater; (i) relationship between Mg^{2+} and HCO_3^- in groundwater; (j) relationship between Mg^{2+} and SO_4^{2-} in groundwater.

northwestern study area; dissolution of sodium plagioclase can be a potential source of sodium in groundwater. In addition, domestic sewage may be another origin of sodium and chloride in groundwater because it is always enriched in sodium relative to chloride (Jalali, 2007; Vengosh and Keren, 1996).

It is well known that the natural process, anthropogenic activities, and evaporation have stronger effects on the molar Cl^-/Br^- ratio, which is often applied to reflect the origin of chloride and bromide. Figure 2d demonstrates that 32 samples have higher molar Cl^-/Br^- ratio above 650. Out of 32 samples, a rapid increase in molar Cl^-/Br^- ratios with increasing chloride concentrations is observed in 6 samples, indicating that chloride in these samples are derived from halite dissolution; the rest of 26 samples may be associated with the fertilizers or salt used for

road de-icing (Cartwright et al., 2006). Groundwater samples were collected from the second aquifer, the evaporation does not affect the molar Cl^-/Br^- ratios in these samples (Xie et al., 2012; Cartwright et al., 2006). 14 samples have lower molar Cl^-/Br^- ratio fall within 650 in Fig. 2d, which are caused by Br-bearing compounds utilization, such as the ethyl dibromide and methyl bromide, rather than evaporation. Therefore, minerals (sodium plagioclase, sodium sulfate, and halite) dissolution and anthropogenic activities are the probable sources of sodium in groundwater samples.

According to correlation analysis, it is concluded that the potassium are released into the groundwater mostly through chemistry fertilizer and domestic sewage; water-rock interaction, such as K-feldspars, illite, and sylvite dissolution, is less fre-

quently.

3.1.4 The origin of calcium and magnesium

If calcium, magnesium, sulfate, and bicarbonate are derived from simple dissolution of calcite, dolomite and gypsum, then a charge balance should exist between the cations and anions (Jalali, 2005). As described in Fig. 2e, the relationship between $(\text{HCO}_3^- + \text{SO}_4^{2-})$ and $(\text{Ca}^{2+} + \text{Mg}^{2+})$ was employed to identify this water-rock interaction. It shows that about 89.13% of samples fall around and above the equiline, reflecting the effect of carbonate and sulfate minerals dissolution (Yu et al., 2012; Yidana and Yidana, 2010); 10.87% samples fall below the 1 : 1 line, indicating the influence of silicate dissolution and ion exchange.

Figure 2e has indicated that the concentrations of calcium and magnesium are affected by the dissolution of carbonate minerals, such as dolomite and calcite. Subsequently, the 1 : 1 relation between $\text{Ca}^{2+} + \text{Mg}^{2+}$ and bicarbonate (Fig. 2f) was applied to identify the impact of dolomite dissolution on these two elements. In Fig. 2f, 31.8% samples are well described by the 1 : 1 line, reflecting that dolomite dissolution contributes to calcium and magnesium ions in these samples. On the other hand, Fig. 2g describes the relation between $1/2\text{Ca}^{2+}$ and bicarbonate in groundwater. The correlation coefficient is 0.021 between these ions, indicating calcite is not the origin of calcium (Jalali, 2007,

2006). A plot of calcium and sulfate ion (Fig. 2h) shows that most of the samples deviate from the 1 : 1 line; only 4 samples fall around the equiline. This indicates that gypsum is not an important source of calcium.

Figure 2i illustrates that beside carbonate mineral dissolution; other sources of magnesium must be present in this work. Figure 2j shows 61.5% of samples are plotted along the 1 : 1 line (Fig. 2j), indicating that dissolution of magnesium sulfate minerals have significant impact on magnesium content in groundwater. In addition, Fig. 2e has indicated that magnesium may be resulted from the dissolution of silicate minerals. The aluminosilicates minerals, such as biotite and magnesium chlorite, are widely distributed in the western study area, which may be another source of magnesium in groundwater. Therefore, ion exchange, dissolution of aluminosilicates, magnesium sulfate mineral, and dolomite are the major sources of magnesium in groundwater.

3.2 Multivariate Statistical Analysis

3.2.1 Cluster analysis

In this article, CA, an unsupervised technique, was applied to discover similarity clusters between the sampling points. Results of clustering of groundwater samples obtained from the

Table 1 Statistic summary of concentrations of hydrochemical variables in three groups resulting from cluster analysis in the study area (unit: concentration in mg/L except pH)

	VPW			PW	GW		
	Min	Mean	Max		Min	Mean	Max
pH	7.5	7.70±0.16	7.9	7.5	7.1	7.43±0.18	7.9
TDS	1 537	1 843±368	2 474	309.12	327.2	555.2±203.1	1 226.8
TH	732.6	930±189.2	1 223.6	246.25	228.2	362.8±117.5	672.7
Na	177.5	257.7±68.1	347.6	23.33	2.9	68.32±48.61	255.25
K	0.48	3.49±3.87	10.2	0.75	0.473	3.46±5.69	29.325
Ca	109.8	152.2±33.3	193.7	58.48	42.22	70.32±20.96	144.90
Mg	88.4	133.6±33.8	179.8	24.36	21.54	44.59±20.03	113.19
Cl	44.5	218.4±136.6	397.9	19.25	6.98	45.08±33.47	181.96
Br	0.17	0.43±0.24	0.7	0.05	0.05	0.112±0.061	0.34
SO ₄	735.1	898.4±158.4	1 150.9	59.37	33.7	131.60±110.90	601
HCO ₃	214.2	319.2±82.3	440.2	248.64	254.5	366.77±72.81	539.9
Mn	0.004	0.13±0.11	0.28	0.003 8	0.031	0.21±0.15	0.65
Fe	0.000 2	0.053±0.08	0.19	0.009 6	0.004	0.25±0.25	0.97
Cu	0.001	0.013±0.017	0.043	0.009	0.001	0.034±0.041	0.144
Zn	0.13	0.24±0.15	0.50	0.061	0.057	0.30±0.44	2.10
Ba	0.01	0.042±0.031	0.09	0.197	0.059	0.26±0.12	0.55
Pb	0.001	0.015±0.021	0.05	0.05	0.013	0.09±0.12	0.59
As	0.005	0.01±0.004 4	0.015	0.313	0.008	0.018±0.009	0.048
Cr	0.02	0.032±0.008 8	0.04	0.3	0.018	0.024±0.004	0.038
NO ₃	19	43.79±20.13	63.2	12.2	0.01	1.36±1.37	5.33
F	0.67	0.91±0.24	1.29	0.84	0.24	0.72±0.36	1.54
I	0.01	0.033±0.032	0.075	0.01	0.005	0.078±0.057	0.26
NO ₂	0.001	0.16±0.22	0.54	0.001	0.001	0.014±0.027	0.12
PO ₄	0.02	0.11±0.11	0.27	0.06	0.02	0.10±0.13	0.56
COD	1.07	1.46±0.45	2.08	0.65	0.51	1.36±0.86	5.96

Bold values indicate strong loadings.

normalized data using Ward’s method divided the sampling sites into three groups, namely good water (GW), poor water (PW), and very poor water (VPW) based on their hydrochemical parameters (Table 1). The dendrogram resulting from CA was displayed in Fig. 3.

The VPW cluster includes 5 samples with the highest concentrations of TDS, TH, sodium, potassium, calcium, magnesium, chloride, sulfate, bicarbonate, and nitrate. Table 1 demonstrates that the TDS values of groundwater range from 1 537 to 2 874 mg/L with an average value of 1 843 mg/L in this group; the TH values of samples are range from 732.6 to 1 223.6 mg/L with an average of 930 mg/L; the sodium varied in the range from 177.5 to 347.6 mg/L with a mean value of 257.7 mg/L; chloride range from 44.5 to 397.9 mg/L with a mean value of 218.4 mg/L; sulfate is found in the concentration range of 735.1–1 150.9 mg/L with an average value of 898.4 mg/L for the studied samples. The calculation results reflect that groundwater in this group are unsuitable for human consumption due to the aforementioned elements content far exceeds their drinking water guideline values (WHO, 2006). Combined with the hydrochemical analysis results, it is found that higher concentrations of these elements are associated with natural process and anthropogenic activities.

The PW cluster is formed by the site N23 correspond to the poor groundwater quality site, where numerous domestically plastic plants are located. With the exception of chromium and arsenic, the concentrations of the others elements are lower, however, this sample is still unsuitable for drinking use. The concentrations of arsenic and chromium are about 20–30 times higher than their guideline values (WHO, 2006), which would yield health risk to human body, such as skin cancer, myocardial damage, and heart disease (Phan et al., 2010). Batayneh (2012) have reported that the chromium content of groundwater can characterize the extent of industrial activities to some extent. Therefore, the higher concentrations of chromium and arsenic for this sample may be attributed to the industrial effluent discharge (Wang L et al., 2011).

The GW group includes 40 samples, which distribute in the northern study area following the Hutuo River flow direction. These samples mainly receive untreated domestic sewage and agriculture non-point/point source pollutant. Fortunately, the contamination situation of these samples are not serious due to their hydrochemical elements content fall within the drinking water standard of WHO (2006), which are basically suitable for human consumption.

3.2.2 Discriminant analysis

The present study applied the stepwise DA to establish the discrimination model with 25 hydrochemical variables of groundwater between the given clusters. According to this method, 7 elements, sulfate, copper, barium, arsenic, chromium, nitrate, and iodine, were extracted to divide groundwater samples into the given clusters, and other variables were eliminated in this analysis. Two canonical discriminant functions were established to classify the groundwater contamination and their relative coefficient were list in Table 2.

The first function DF1 explains 93.2% of the total variance with high canonical correlation 0.997. This function is correlat-

ed with high arsenic (113.233) and chromium (185.681) and low barium (-7.638) and sulfate (-0.003). DF2 explains about 6.8% variance with canonical correlation 0.961, which is negatively

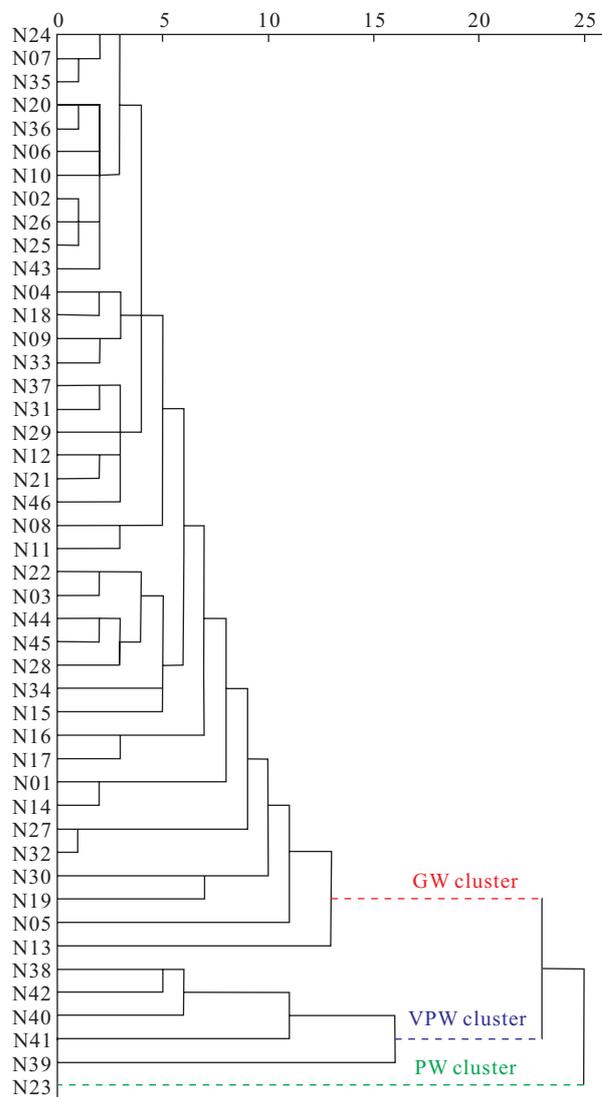


Figure 3. Dendrogram showing clustering of monitoring groundwater samples.

Table 2 Canonical discriminant function coefficients for groundwater quality assessment in study area for three clusters

Variables	DF1	DF2
SO ₄	-0.003	0.009
Cu	11.387	-0.34
Ba	-7.698	3.406
As	113.233	-20.316
Cr	185.681	22.08
NO ₃	0.021	0.093
I	10.148	-14.337
Canonical correlation	0.997	0.961
Constant	-7.085	-2.368

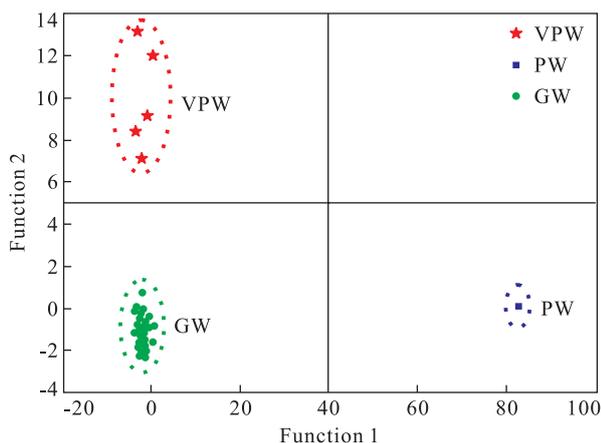


Figure 4. Scatter plot of two canonical discriminant functions showing three groundwater quality clusters.

correlated with iodine (-14.337) and arsenic (-20.316) and positively correlated with chromium (22.080) and barium (3.406).

The discriminant score of each sample were calculated using the canonical discriminant functions and 7 hydrochemical variables (Fig. 4). It can be seen from this figure that the DF1 discriminate the VPW and GW with PW cluster; and the

DF2 differentiate the GW and PW with VPW group (Rani et al., 2011). Therefore, only 7 hydrochemical variables were required to discriminate the 46 samples with 100% correct classification according to DA theory in this article.

3.2.3 Factor analysis

FA is a statistical method designed to analyze the interrelationships within a set of variables by reducing the complex information to an easily interpretable form. In this paper, the R-mode FA was carried out using a set of data consisting of 46 groundwater samples to identify and characterize the factors that affect the hydrochemical composition of the study area. The factor loadings and eigenvalues were described in Table 3. The first 8 factors were selected to represent the major hydrochemical reaction due to its eigenvalue greater than 1, which had helped in the formation of the present hydrochemistry system without losing significant information, and the factor loadings >0.5 were regarded as significant in the interpretation of the data (Dantu, 2010).

Table 3 shows the factor loadings of 8 selected factors, which represent about 80.897% of the total variance, and reflect various controlling factor of groundwater quality. Factor 1, which describes 32.97% of the variance, has a higher positive

Table 3 Factor analysis with rotated varimax scores for hydrochemical elements in groundwater

Variables	F1	F2	F3	F4	F5	F6	F7	F8
pH	0.210	-0.312	0.162	-0.016	-0.604	0.019	0.263	0.221
TDS	0.974	-0.028	-0.063	-0.132	0.117	0.033	0.034	0.027
TH	0.958	0.109	-0.052	-0.118	-0.129	0.058	-0.028	-0.008
Na	0.848	-0.077	-0.083	-0.136	0.358	-0.091	0.136	0.067
K	0.048	0.111	-0.045	-0.027	-0.073	0.091	0.886	0.003
Ca	0.822	0.015	-0.054	-0.104	-0.253	0.225	-0.205	-0.038
Mg	0.955	0.131	-0.039	-0.133	0.001	-0.084	0.084	0.031
Cl	0.926	0.059	-0.001	0.066	-0.077	-0.136	0.026	-0.007
Br	0.797	-0.096	-0.071	-0.018	0.018	0.336	0.094	-0.038
SO ₄	0.917	-0.169	-0.043	-0.163	0.191	0.136	0.014	0.013
HCO ₃	0.077	0.758	-0.162	-0.104	-0.218	-0.299	0.197	0.158
Mn	-0.081	0.181	-0.186	0.479	-0.054	0.501	0.316	-0.018
Fe	-0.183	0.331	-0.241	-0.349	0.514	-0.201	-0.214	-0.081
Cu	-0.146	-0.099	-0.128	0.775	-0.014	-0.045	-0.093	-0.101
Zn	-0.097	-0.474	-0.244	-0.069	-0.094	-0.202	0.388	-0.155
Ba	-0.406	0.706	0.037	0.088	-0.266	0.204	-0.087	-0.051
Pb	-0.163	0.057	-0.012	0.821	0.057	-0.033	-0.008	0.133
As	-0.123	-0.003	0.952	-0.084	-0.033	-0.015	-0.055	-0.077
Cr	-0.040	-0.123	0.949	-0.085	-0.009	-0.037	-0.036	-0.062
NO ₃	0.847	-0.238	0.157	-0.054	0.024	0.047	0.004	0.099
F	0.075	-0.518	-0.098	-0.306	-0.178	-0.456	-0.068	0.202
I	0.057	0.822	-0.142	-0.069	0.168	-0.064	0.084	-0.056
NO ₂	0.349	-0.168	-0.029	-0.254	0.196	0.722	-0.006	0.142
PO ₄	-0.058	0.078	-0.074	0.142	0.772	0.261	-0.237	-0.054
COD	0.004	0.011	-0.118	0.028	0.039	0.032	-0.018	0.927
Eigen value	8.572	2.691	2.129	1.962	1.732	1.498	1.358	1.090
% of Var	32.97	10.35	8.19	7.55	6.66	5.76	5.23	4.19
Cum. %	32.97	43.32	51.511	59.057	65.718	71.48	76.705	80.897

Bold values indicate strong loadings.

factor loading for TDS, TH, sodium, calcium, magnesium, chloride, sulfate, and nitrate. Kim et al. (2004) reported that if ammonium oxidation or aerobic decomposition of organic matter is the dominant source of nitrate, negative correlation is observed between nitrate and pH. In this paper, the correlation coefficient between nitrate and pH is 0.418, which illustrates that this ion is derived from another sources rather than ammonium oxidation or aerobic decomposition of organic matter. Combined with the hydrochemical analysis and CA results, factor 1 indicates that the origin of these elements may be expected from ion exchange and dissolution of silicate, dolomite, halite, and gypsum.

Factor 2 describes 10.35% of the variance has high positive factor loadings for bicarbonate, barium, and iodine; and a negative factor loading for fluoride. Therefore, this factor can be interpreted as the domestic sewage. Factor 3 accounts for 8.19% of the variance and indicates high loadings for arsenic and chromium. The chromium is widely distributed in earth crust, which has extremely lower concentration in rock and soil. Batayneh (2012) have proved that most of the chromium in water resource is associated with the industrial contamination. In this article, the higher concentrations of arsenic and chromium result from the industrial effluent discharge according to CA results, and therefore this factor is interpreted as the industrial activities.

Factor 4 exhibits 7.55% of the total variance and has significant relationships between copper and lead. The origins of these pollutants are mainly related to atmospheric deposition and waste disposal sites (Christensen et al., 1999). Both factors 5 and 6 account for 12.42% of the total variance have a higher loading for phosphate, manganese, and iron; for the element pH, a lower factor loading is observed. The strong negative loadings on pH and positive loading on iron and manganese reflect that these elements result from carbonate and silicate minerals dissolution (Ujević Bošnjak et al., 2012). In general, the higher level of phosphate in groundwater is associated with fertilizer, such as ammonium phosphates and ammonium sulfates (Kannel et al., 2011). Therefore, these two factors are attributed to fertilizer utilization and minerals dissolution.

Factor 7 exhibits 5.23% of the total variance has a higher loading for potassium, the hydrochemical analysis results have indicated that the potassium in groundwater are derived from anthropogenic activities, rather than such minerals dissolution as chloride salt, K-feldspars, and illite. Factor 8 describes 4.19% of the variance has high positive factor loading for chemical oxygen demand (COD). It is well known that COD element is closely related to the contamination of organic matters (Kannel et al., 2011), this factor is regarded as the domestic and agriculture sewage.

3.3 Discussions

Figure 2 demonstrates that there are many sources of sodium, potassium, calcium, and magnesium. With regard to carbonate minerals, the dissolution of dolomite rather than calcite has significant effects on the concentrations of calcium and magnesium. Therefore, the contribution of silicate and ion exchange, dolomite, and sulfate minerals to calcium are 59.5%, 31.8%, and 8.7%, respectively. And their contributions to

magnesium are 6.7%, 31.8%, and 61.5%, respectively. The sources of sodium are complex, which mainly include minerals dissolution (sodium sulfate, sodium plagioclase, and halite) and domestic sewage. The main origins of potassium are chemical fertilizer and domestic sewage, natural process has nothing influences on this element. In addition, according to the mMolar Cl/Br ratio, 69.57% of chloride is derived from halite dissolution, fertilizers, and salt used for road de-icing; and 30.43% of bromide is closely related to the pesticides, such as ethyl dibromide and methyl bromide.

Based on the results of MA, VPW cluster includes five samples (N38, N39, N40, N41, and N42) with the highest mineralized levels, which are located in the southeastern study area. The concentrations of TDS, TH, sodium, calcium, magnesium, and chloride far exceeded the drinking water standard of WHO (2006). The reason for this results is that groundwater as well as Hutuo River, flows through the middle of the area from west to east in the north, but flows southeastward in the south. Due to the water-rock interaction, a significant increase in the concentrations of TDS, TH, sodium, and magnesium were observed in the direction of northwest to southeast following the regional flow direction. In addition, combined with the hydrochemical analysis and FA results, it was found that nature process is mainly responsible for the higher concentrations of TDS, TH, magnesium, and sodium, in addition, domestic sewage is partly responsible for the high level of sodium in this group. There is only one sample (N23) in PW cluster; it has higher concentrations of arsenic and chromium. This result was attributed to the industrial effluent discharge from the nearby domestically plastic process plants. The concentrations of arsenic and chromium in N23 is about 20–30 times higher than their guideline values (WHO, 2006), causing various diseases to human body (Batayneh, 2012; Phan et al., 2010). GW cluster was formed by the rest of 40 samples, the potassium, bicarbonate, copper, lead, iodine, barium, and COD contents in these samples are less than the drinking water standard of WHO (2006), it is basically suitable for drinking uses. FA results demonstrate that the aforementioned elements are associated with domestic and agriculture sewage. These samples are located in the northern study area corresponding to the recharge zone of the aquifer. Accordingly, the GW cluster is influenced by the domestic sewage, leachate from waste sites, fertilizer, and atmospheric deposition, rather than water-rock interactions.

This study employed the MA method to identify the controlling factors of groundwater quality based on hydrochemical analysis of NCP. The findings presented here demonstrate that 25 parameters were required to characterize the groundwater quality without the MA and hydrochemical analysis in the past years. At present, only 7 hydrochemical variables can be applied to delineate groundwater quality evolution process and identify the controlling factors with a lower cost of computation and monitoring.

Therefore, the conclusions from the current work can be taken into account when developing policies for safe drinking water supplies. First, groundwater extracted from well or borehole can be divided into the different groups, VPW, PW, and GW, using the DA method according to its 7 hydrochemistry variables. Second, the treatment measures can be respectively

proposed according to its corresponding groundwater quality classification. If this sample belongs to VPW cluster, it should be treated to reduce the TDS, TH, sodium, and magnesium concentrations before human consumption. If it is divided into PW cluster, which is only suitable for use in industrial, rather than drinking and agriculture purposes, the excess of arsenic and chromium in irrigation water can accumulate in the edible part of crops, posing considerable health risk to humans and animals; if this groundwater belongs to GW group, indicating that it is currently suitable for drinking purposes, however, its quality tends to deterioration due to indiscriminate use of chemicals fertilizer in agriculture, and untreated effluents from domestic sectors. If no effective measures are taken, this situation will bring about serious harmful impacts on human health.

4 CONCLUSIONS

The hydrochemical compositions is affected by hydrogeology structure of NCP, a significant increase in the degree of TDS, TH, magnesium, and calcium is observed from northwest to southeast of study area, and therefore the highest mineralized groundwater is founded in southeastern part of study area corresponding to the discharge zone of the aquifer. Groundwater extracted from the northwestern study area, which distribute perpendicular to the groundwater flow direction, it is not influenced by the domestic sewage and chemistry fertilizer, and nature process, Therefore, the hydrochemical variables with the exception of arsenic and chromium contents are equal to background value at local field.

Due to growing population, agriculture expansion, and urbanization, groundwater resources sustainable development of NCP requires strategic planning and makes use of a wide range of groundwater pollution prevention measures. Anthropogenic activities are the major controlling factor of groundwater quality compared to nature process. Current strategies to address the groundwater pollution problem in the northern study area should focus on reducing the domestic sewage discharge and chemistry fertilizer utilization to prevent the groundwater quality consistent deteriorate; with respect to southwestern study area, the optimal measure is decreasing the three wastes contaminants discharge and reconfiguring the industrial structure. Groundwater quality in southeastern part of the study area is controlled by nature process, the prevention measures of groundwater contamination is reducing the total amount of groundwater exploitation to decrease the mean annual rates of groundwater renewal.

ACKNOWLEDGMENTS

This study was supported by the Major State Basic Research Development Program (No. 2010CB428800) and the Geological Survey Projects Foundation of Institute of Hydrogeology and Environmental Geology (No. SK201308).

REFERENCES CITED

- Ariyibi, E. A., Folami, S. L., Ako, B. D., et al., 2010. Application of the Principal Component Analysis on Geochemical Aata: A Case Study in the Basement Complex of Southern Ilesa Area, Nigeria. *Arabian Journal of Geosciences*, 4(1–2): 239–247, doi:10.1007/s12517-010-0175-5
- Aryafar, A., Doulati, A. F., 2011. R-Mod Factor Analysis, a Popular Multivariate Statistical Technique to Evaluate Water Quality in Khaf-Sangan Basin, Mashhad, Northeast of Iran. *Arabian Journal of Geosciences*, 6(3): 893–900, doi:10.1007/s12517-011-0367-7
- Batayneh, A. T., 2012. Toxic (Aluminum, Beryllium, Boron, Chromium and Zinc) in Groundwater: Health Risk Assessment. *International Journal of Environmental Science and Technology*, 9(1): 153–162, doi:10.1007/s13762-011-0009-3
- Cartwright, I., Weaver, T. R., Fifield, L. K., 2006. Cl/Br Ratios and Environmental Isotopes as Indicators of Recharge Variability and Groundwater Flow: An Example from the Southeast Murray Basin, Australia. *Chemical Geology*, 231(1–2): 38–56, doi:10.1016/j.chemgeo.2005.12.009
- Christensen, J. B., Botma, J. J., Christensen, T. H., 1999. Complexation of Cu and Pb by DOC in Polluted Groundwater: A Comparison of Experimental Data and Predictions by Computer Speciation Models (WHAM and MINTEQA2). *Water Research*, 33(15): 3231–3238, doi:10.1016/S0043-1354(99)00020-2
- Dantu, S., 2010. Factor Analysis Applied to a Geochemical Study of Soils from Parts of Medak and Sangareddy Areas, Medak District, Andhra Pradesh, India. *Environmental Monitoring and Assessment*, 162(1–4): 139–152, doi:10.1007/s10661-009-0782-3
- Dixon, W., Chiswell, B., 1992. The Use of Hydrochemical Sections to Identify Recharge Areas and Saline Intrusions in Alluvial Aquifers, Southeast Queensland, Australia. *Journal of Hydrology*, 135(1–4): 259–274, doi:10.1016/0022-1694(92)90091-9
- GB/T 5750, 2006. Standard Examination Methods for Drinking Water. Ministry of Public Health of the People's Republic of China, Beijing (in Chinese)
- Hamzaoui-Azaza, F., Ketata, M., Bouhlila, R., et al., 2011. Hydrogeochemical Characteristics and Assessment of Drinking Water Quality in Zeuss-Koutine Aquifer, Southeastern Tunisia. *Environmental Monitoring and Assessment*, 174(1–4): 283–298, doi:10.1007/s10661-010-1457-9
- Hu, Y., Moiwo, J. P., Yang, Y., et al., 2010. Agricultural Water-Saving and Sustainable Groundwater Management in Shijiazhuang Irrigation District, North China Plain. *Journal of Hydrology*, 393(3–4): 219–232, doi:10.1016/j.jhydrol.2010.08.017
- Jalali, M., 2005. Major Ion Chemistry of Groundwaters in the Bahar Area, Hamadan, Western Iran. *Environmental Geology*, 47(6): 763–772, doi:10.1007/s00254-004-1200-3
- Jalali, M., 2006. Chemical Characteristics of Groundwater in Parts of Mountainous Region, Alvand, Hamadan, Iran. *Environmental Geology*, 51(3): 433–446, doi:10.1007/s00254-006-0338-6
- Jalali, M., 2007. Hydrochemical Identification of Groundwater Resources and Their Changes under the Impacts of Human Activity in the Chah Basin in Western Iran. *Environmental Monitoring and Assessment*, 130(1–3): 347–364, doi:10.1007/s10661-006-9402-7

- Jang, C. S., Chen, J. S., Lin, Y. B., et al., 2012. Characterizing Hydrochemical Properties of Springs in Taiwan Based on Their Geological Origins. *Environmental Monitoring and Assessment*, 184(1): 63–75, doi:10.1007/s10661-011-1947-4
- Kamble, S. R., Vijay, R., 2011. Assessment of Water Quality Using Cluster Analysis in Coastal Region of Mumbai, India. *Environmental Monitoring and Assessment*, 178(1–4): 321–332, doi:10.1007/s10661-010-1692-0
- Kannel, P. R., Kanel, S. R., Lee, S., et al., 2011. Chemometrics in Assessment of Seasonal Variation of Water Quality in Fresh Water Systems. *Environmental Monitoring and Assessment*, 174(1–4): 529–545, doi:10.1007/s10661-010-1476-6
- Kim, K., Rajmohan, N., Kim, H., et al., 2004. Assessment of Groundwater Chemistry in a Coastal Region (Kunsan, Korea) Having Complex Contaminant Sources: A Stoichiometric Approach. *Environmental Geology*, 46(6–7): 763–774, doi:10.1007/s00254-004-1109-x
- Li, L., Wang, Y. X., Wu, Y., et al., 2013. Major Geochemical Controls on Fluoride Enrichment in Groundwater: A Case Study at Datong Basin, Northern China. *Journal of Earth Science*, 24(6): 976–986, doi:10.1007/s12583-013-0385-3
- Ma, R., Shi, J. S., Liu, J. C., 2012. Dealing with the Spatial Synthetic Heterogeneity of Aquifers in the North China Plain: A Case Study of Luancheng County in Hebei Province. *Acta Geologica Sinica (English Edition)*, 86(1): 226–245, doi:10.1111/j.1755-6724.2012.00624.x
- Meybeck, M., 1987. Global Chemical Weathering of Surficial Rocks Estimated from River Dissolved Loads. *American Journal of Science*, 287: 401–428, doi:10.2475/ajs.287.5.401
- Osman, R., Saim, N., Juahir, H., et al., 2012. Chemometric Application in Identifying Sources of Organic Contaminants in Langat River Basin. *Environmental Monitoring and Assessment*, 184(2): 1001–1014, doi:10.1007/s10661-011-2016-8
- Pang, Z. H., Yuan, L. J., Huang, T. M., et al., 2013. Impacts of Human Activities on the Occurrence of Groundwater Nitrate in an Alluvial Plain: A Multiple Isotopic Tracers Approach. *Journal of Earth Science*, 24(1): 111–124, doi:10.1007/s12583-013-0310-9
- Phan, K., Sthiannopkao, S., Kim, K. W., et al., 2010. Health Risk Assessment of Inorganic Arsenic Intake of Cambodia Residents through Groundwater Drinking Pathway. *Water Research*, 44(19): 5777–5788, doi:10.1016/j.watres.2010.06.021
- Rani, N., Sinha, R. K., Prasad, K., et al., 2011. Assessment of Temporal Variation in Water Quality of Some Important Rivers in Middle Gangetic Plains, India. *Environmental Monitoring and Assessment*, 174(1–4): 401–415, doi:10.1007/s10661-010-1465-9
- Ravindra, K., Stranger, M., Grieken, R. V., 2008. Chemical Characterization and Multivariate Analysis of Atmospheric PM_{2.5} Particles. *Journal of Atmospheric Chemistry*, 59(3): 199–218, doi:10.1007/s10874-008-9102-5
- Shi, J., 2007. The Developing Trend of Hydrogeology: A Review of the 34th IAH Congress. *Acta Geologica Sinica*, 28(6): 509–520, doi:10.3975/cagsb.2007.06.01 (in Chinese with English Abstract)
- Shi, J. S., Wang, Z., Zhang, Z. J., et al., 2011. Assessment of Deep Groundwater Over-Exploitation in the North China Plain. *Geoscience Frontiers*, 2(4): 593–598, doi:10.1016/j.gsf.2011.07.002
- Stigter, T. Y., Ribeiro, L., Carvalho Dill, A. M. M., 2006. Application of a Groundwater Quality Index as an Assessment and Communication Tool in Agro-Environmental Policies—Two Portuguese Case Studies. *Journal of Hydrology*, 327(3–4): 578–591, doi:10.1016/j.jhydrol.2005.12.001
- Ujević Bošnjak, M., Capak, K., Jazbec, A., et al., 2012. Hydrochemical Characterization of Arsenic Contaminated Alluvial Aquifers in Eastern Croatia Using Multivariate Statistical Techniques and Arsenic Risk Assessment. *Science of the Total Environment*, 420: 100–110, doi:10.1016/j.scitotenv.2012.01.021
- Vengosh, A., Keren, R., 1996. Chemical Modifications of Groundwater Contaminated by Recharge of Treated Sewage Effluent. *Journal of Contaminant Hydrology*, 23(4): 347–360, doi:10.1016/0169-7722(96)00019-8
- Wang, L., Wang, Y. P., Xu, C. X., et al., 2011. Analysis and Evaluation of the Source of Heavy Metals in Water of the River Changjiang. *Environmental Monitoring and Assessment*, 173(1–4): 301–313, doi:10.1007/s10661-010-1388-5
- Wang, M., Bai, Y. Y., Chen, W. P., et al., 2012. A GIS Technology Based Potential Eco-Risk Assessment of Metals in Urban Soils in Beijing, China. *Environmental Pollution*, 161: 235–242, doi:10.1016/j.envpol.2011.09.030
- Wang, S. Q., Song, X. F., Wang, Q. X., et al., 2011. Shallow Groundwater Dynamics and Origin of Salinity at Two Sites in Salinated and Water-Deficient Region of North China Plain, China. *Environmental Earth Sciences*, 66(3): 729–739, doi:10.1007/s12665-011-1280-9
- Wang, Y., Wang, P., Bai, Y., et al., 2013. Assessment of Surface Water Quality via Multivariate Statistical Techniques: A Case Study of the Songhua River Harbin Region, China. *Journal of Hydro-Environment Research*, 7(1): 30–40, doi:10.1016/j.jher.2012.10.003
- Weng, H. X., Ji, Z. Q., Chu, Y., et al., 2011. Benzene Series in Sewage Sludge from China and Its Release Characteristics during Drying Process. *Environmental Earth Sciences*, 65(3): 561–569, doi:10.1007/s12665-011-1100-2
- WHO, 2006. Guideline for Drinking Water Quality. World Health Organization, Geneva
- Xie, X., Wang, Y. X., Su, C., et al., 2012. Influence of Irrigation Practices on Arsenic Mobilization: Evidence from Isotope Composition and Cl/Br Ratios in Groundwater from Datong Basin, Northern China. *Journal of Hydrology*, 424: 37–47, doi:10.1016/j.jhydrol.2011.12.017
- Yang, M., Fei, Y. H., Ju, Y. W., et al., 2012. Health Risk Assessment of Groundwater Pollution—A Case Study of Typical City in North China Plain. *Journal of Earth Science*, 23(3): 335–348, doi:10.1007/s12583-012-0260-7

- Yidana, S. M., Yidana, A., 2010. Assessing Water Quality Using Water Quality Index and Multivariate Analysis. *Environmental Earth Sciences*, 59(7): 1461–1473, doi:10.1007/s12665-009-0132-3
- Yuan, Z. J., Shen, Y. J., 2013. Estimation of Agricultural Water Consumption from Meteorological and Yield Data: A Case Study of Hebei, North China. *Plos One*, 8(3): 1–9, doi:10.1371/journal.pone.0058685
- Yu, J. C., Mo, X. X., Yu, X. H., et al., 2012. Geochemical Characteristics and Petrogenesis of Permian Basaltic Rocks in Keping Area, Western Tarim Basin: A Record of Plume-Lithosphere Interaction. *Journal of Earth Science*, 23(4): 442–454, doi:10.1007/s12583-012-0267-0
- Zhao, Y., Xia, X. H., Yang, Z. F., et al., 2012. Assessment of Water Quality in Baiyangdian Lake Using Multivariate Statistical Techniques. *Procedia Environmental Sciences*, 13: 1213–1226, doi:10.1016/j.proenv.2012.01.115
- Žibret, G., Šajn, R., 2010. Hunting for Geochemical Associations of Elements: Factor Analysis and Self-Organising Maps. *Mathematical Geosciences*, 42(6): 681–703, doi:10.1007/s11004-010-9288-3