

Identifying key hydrochemical processes in a confined aquifer of an arid basin using multivariate statistical analysis and inverse modeling

Changchang Fu · Wenjing Zhang ·
Shengyu Zhang · Xiaosi Su · Xueyu Lin

Received: 10 August 2013 / Accepted: 15 April 2014 / Published online: 7 May 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract Tertiary fractured permeable confined aquifer, which covered about 70 % of the studying area, played an important role in alleviating drinking water shortages. However, about 58 and 79 % of the groundwater samples exceeded the desirable limits for fluoride (1.5 mg/L) and TDS (1,000 mg/L). Two multivariate statistical methods, hierarchical cluster analysis (HCA) and principal components analysis (PCA), were applied to a subgroup of the dataset in terms of their usefulness for groundwater classification, as well as to identify the key processes controlling groundwater geochemistry. In the PCA, two principal factors have been extracted, which could explain 73 % of the total data variability. Among them, factor 1 revealed the source of groundwater salinity and factor 2 explained the elevated fluoride. Two major groups were classified by HCA and Group 1 was near the groundwater recharge zone and Group 2 was mainly distributed over the groundwater discharge zone. Inverse modeling (NET-PATH) results indicated that the hydrochemical evolution

was primarily controlled by (1) the dissolution of mirabilite, gypsum and halite for the source of groundwater salinity; (2) the release of the adsorbed fluoride through desorption or through competition with HCO_3^- under alkalinity condition for the elevated fluoride in the groundwater.

Keywords Xiji Basin · Tertiary confined groundwater · Hydrochemistry evolution · Multivariate statistical analysis · Inverse modeling

Introduction

Sustainable management of scarce freshwater resources is a major challenge for arid regions, where there are also pressures related to population growth, economic development, climate change, pollution and groundwater over-abstraction (Du et al. 2013; Su et al. 2013). Groundwater is commonly the dominant water resource in these arid areas, so it is important to understand the hydrochemical processes that govern the groundwater quality for sustainable management of the water resource.

The study area, Xiji County, is a region in Northwest China with serious water shortages and the water available per capita is only 95 m³ each year (Gao 2012). Some researches have been done to investigate hydrochemical evolution in the southern Ningxia Autonomous Region, including Xiji County (Yang et al. 2009; Zhang et al. 2006; Chen et al. 2009; Li et al. 2006). The groundwater system of Xiji Basin was firstly isolated by the northeast boundary of Moon Mountain and the southwest boundary of loess hilly-gully geomorphology that extended into Gansu province (Yang et al. 2009). Two reasons were assumed for the saline groundwater of the southern Ningxia region: (1)

Electronic supplementary material The online version of this article (doi:10.1007/s12665-014-3290-x) contains supplementary material, which is available to authorized users.

C. Fu · W. Zhang (✉) · X. Su · X. Lin
Key Laboratory of Groundwater Resources and Environment,
Ministry of Education, Jilin University, Changchun 130021,
China
e-mail: zhangwenjing80@126.com

C. Fu · W. Zhang · X. Su · X. Lin
Institute of Environment and Resources,
Jilin University, Changchun 130021, China

S. Zhang
Key Laboratory of Songliao Aquatic Environment, Ministry of
Education, Jilin Architectural and Civil Engineering Institute,
Changchun 130118, China

the dissolution of great amounts of gypsum, mirabilite and other minerals (primary source); (2) the obturated water reserving structure (Zhang et al. 2006). Further, the hydrochemical evolution of the other two basins adjacent to Xiji Basin (Haiyuan Basin and Qingshuihe Basin, locate north and east of Xiji Basin, respectively) has been investigated (Chen et al. 2009; Li et al. 2006). However, the hydrochemical processes for the salinity and the elevated fluoride in the groundwater of Xiji Basin are still not clear.

Multivariate statistical analyses have been applied to conduct the present investigations. As a part of these analyses, hierarchical cluster analysis (HCA) and principal component analysis (PCA) have been used with success to classify groundwater and identify major mechanisms influencing groundwater chemistry (Usunoff and Guzmán-Guzmán 1989; Farnham et al. 2003; Ghrefat et al. 2012; Huang et al. 2013; Zumlot et al. 2013). When the hydrochemical interpretation is combined with the inverse modeling, models with more confidence level will be found to quantitatively identify the hydrochemical processes along the flow path (Atkinson 2011; Lorite-Herrera et al. 2008).

This paper will attempt to identify the geological factors that presently affect the water chemistry in the region. Inverse modeling techniques and two well-proven multivariate statistical methods (HCA and PCA) are used to assist in the analysis.

Study area

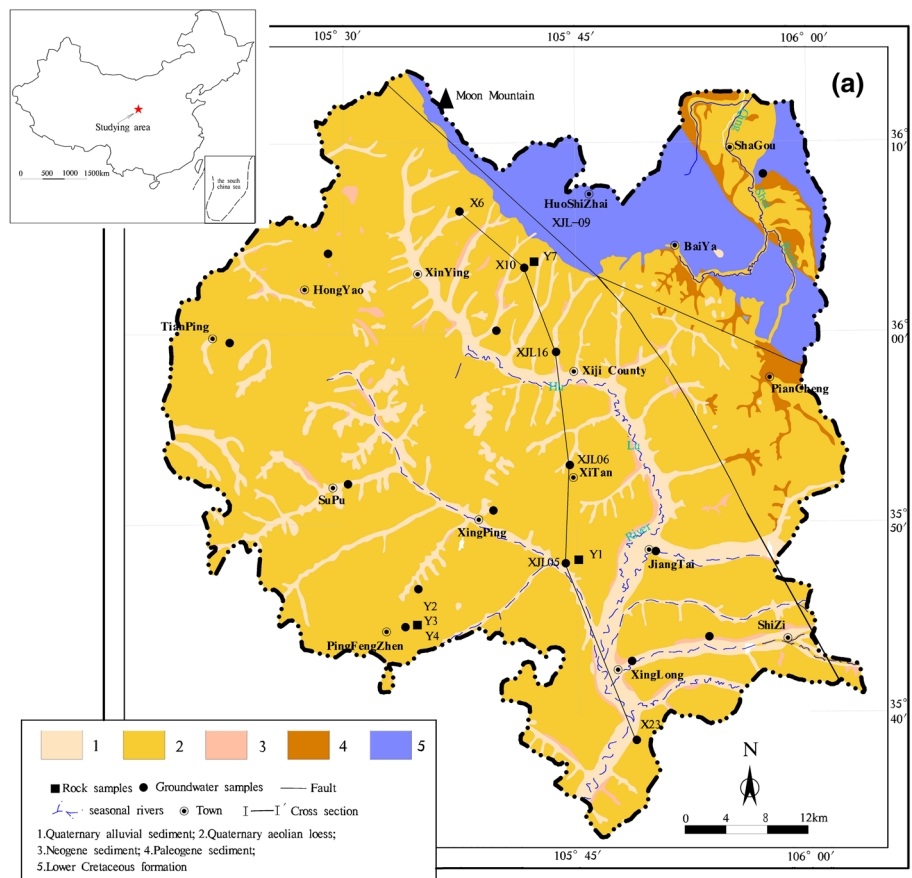
Xiji Basin is located in the southwest of Ningxia Autonomous Region with an area of 3,144 km². The annual average precipitation is approximately 410 mm and approximately 56 % of the annual rainfall falls between July and September. The mean monthly air temperature ranges between a minimum of −7 °C in December–February and a maximum of 17 °C in June–August, with a mean annual temperature of 5.6 °C. The mean annual potential evapotranspiration ranges from 1,400 to 1,600 mm. The main surface water body in the study area is Hulu River and its tributaries.

Geological formations of Xiji Basin have been severely eroded, with remains of Proterozoic, Cretaceous, Tertiary and Quaternary formations (Fig. 1b). Tertiary formation is overlain by Quaternary loess and underlain by Cretaceous formation. A compact clayey soil layer exists between Quaternary and Tertiary formation with thickness from 3 to 10 m. Quaternary gravel layers are mainly distributed along the river banks. Lower Cretaceous formation (K_{1s}) is exposed in the northeastern part, which consists of feldspar sandstone and is weathered severely with well-

developed fissures. Based on lithology and field relationships, the Tertiary formation has generally been divided into Paleogene (E_{2s} and E_{3q}) and Neocene (N_{1h} and N_{2g}) formation. The deposits of E_{2s} and E_{3q} primarily consisted of conglomerate and sandy mudstone, respectively. N_{1h} and N_{2g} consisted of argillaceous sandstone and sandstone with conglomerate interlayer, respectively. The hydraulic conductivity (*K*) of E_{2s}, E_{3q}, N_{1h} and N_{2g} is 0.0165 m/d (0.00253–0.0305, *n* = 2), 0.0013 m/d (*n* = 1), 0.667 m/d (*n* = 1) and 0.0725 m/d (0.0274–0.1384, *n* = 3), respectively. There is an abnormal *K* (6.0 m/d) in E_{2s} formation appeared at X23 well, which could locate on the fault zone. During the period of E_{2s}, E_{3q} and N_{1h}, salts have been accumulated in the deposits of Xiji Basin, including mirabilite, gypsum and magnesite (Zhang et al. 2006). Deposits during Pliocene (N_{2g}) period, however, were generally variegated mudstone, sandstone and conglomerate with low salinity. By the end of Pliocene, Xiji Basin was uplift under the influence of strong fault fold caused by the Himalayan movement (phase III). Furthermore, the eastern part experienced much more upliftment than the western part, which resulting in all the northeastern part formations were mostly eroded while only N_{2g} formation was eroded in the southeast part. In the joint of the two parts, namely the Moon Mountain-Shizi fault (Fig. 1a), parts of Pliocene coarse grain layers were distributed with lower salt content, micropores, bigger hydraulic gradient, as well as stronger water exchange capacity (Huo et al. 1989; Yang et al. 2009).

There are three types of aquifer, which includes Cretaceous fractured permeable unconfined aquifer, Tertiary fractured permeable confined aquifer (E_{2s}, E_{3q}, N_{1h} and N_{2g}) and Quaternary unconsolidated pore aquifer. Tertiary fractured permeable confined aquifer was the focus of our study. It covered about 70 % of the study area, and its thickness varied from more than 400 m in the northern part to 200 m in the southern part. The groundwater level mapping of Tertiary fractured permeable confined aquifer (Fig. 1, ESM only) was made in July 2012 from 17 wells. This map shows that the main groundwater flow direction is N–S from the recharge zone situated in the foot hill zone in the north of the basin through the central parts towards the south area. The hydraulic gradient was approximately 0.007. Groundwater recharge was mainly via a blind fault from the Cretaceous aquifer. In addition, the leakage of pore water from overlying Quaternary unconsolidated formation could be neglected because of a confining bed between the two aquifers. The major discharge of the studying aquifer included the flow to the adjacent regions and artificial abstraction.

Fig. 1 **a** Simplified geological map of Xiji County; **b** hydrogeological cross section along the I–I' line



Methodology

Sampling and analytical methods

A regional geological survey of Xiji Basin was carried out in 2012, and groundwater was sampled from private,

municipal and observation wells, which were distributed over the whole region (Fig. 1a). Field parameters of temperature, pH and TDS were measured. Other chemical (Ca, Mg, Na, K, Cl, SO₄, HCO₃, CO₃ and F) analyses were carried out in the laboratory using atomic absorption spectrometer (Shimadzu AA-6300CF) and ion

Table 1 Mineral composition of the Tertiary confined aquifer

Samples	Relative quantity $\omega(B)/10^{-2}$									
	quart	K-feldspar	Plagioclase	Calcite	Dolomite	Amphibole	Halite and mirabilite	Zeolite	Illite	Chlorite
Y1	56	3	9	11	–	1	2	–	12	6
Y2	42	2	6	3	–	–	1	8	26	12
Y3	62	3	6	15	–	–	1	–	12	1
Y4	20	1	3	2	–	–	2	–	58	5
Y5	31	6	18	2	1	–	1	7	23	11

chromatography (ICS-2100). Then, ionic balance was calculated by cations minus anions over total ion concentration (Appelo and Postma 2005). Of the 24 samples, 2 groundwater samples with ionic balance error above 10 % were rejected.

For an accurate determination of the potential water–rock reactions occurring in the hydrochemical evolution, five core samples of the target aquifer were analyzed using X-ray diffraction in the test center of Jilin University. Table 1 shows that the main minerals in the target aquifer are quartz, illite, plagioclase, chlorite, calcite, zeolite and K-feldspar. Although the amount of these minerals exceeds more than 97 %, they are, however, either unreactive minerals or with much lower reaction rate. So these minerals are probably not governing factors in the hydrochemical evolution. On the contrary, although the content percentage of halite and mirabilite is 1–2 %, they are easier to dissolve into water. As the groundwater flow is very slow, the dissolution of mirabilite and halite is probably the main source of the groundwater salinity.

Data preparation for multivariate statistical analysis

Multivariate statistical analysis is a quantitative and independent approach of groundwater classification allowing the grouping of groundwater samples and the making of correlations between chemical parameters and groundwater samples (Cloutier et al. 2008). The aim of the factor analysis for hydrogeochemical data is to explain the observed relations in a simpler term that was expressed as a new set of factor (Matalas and Reicher 1967). In this study, two multivariate methods were applied using SPSS 17.0: hierarchical cluster analysis (HCA) and principal component analysis (PCA).

Optimal results in multivariate statistical analyses require normal distribution and homoscedasticity. In this respect, the data for each parameter were tested for normal distribution using statistics from SPSS 17.0. The parameter is considered following normal distribution when the coefficients of skewness and kurtosis are all smaller than 1.96 (at 0.05 confidence level). The skewness coefficient of Na^+ , K^+ , Mg^{2+} and SO_4^{2-} is all larger than

1.96 (Table 2), indicating that these four parameters do not follow normal distribution and need to be log normally transformed. Positively skewed chemical data are commonly log normally transformed and standardized for multivariate statistical analyses (Steinhorst and Williams 1985; Güler et al. 2002; Cloutier et al. 2008; Yidana et al. 2012). Then, standardization of the data (X_i) results in new values (Z_i) that have zero mean and are measured in units of standard deviation (s). The standardized data are obtained by subtracting the mean of the distribution from each data and dividing by the standard deviation of the distribution, $Z_i = (X_i - \text{mean})/s$ (Davis 1986). Inter-correlation matrix of the variables is computed from the standardized variables. The correlation coefficient matrix quantifies the linear relationship existing between pairs of variables present therein. The percentages of eigenvalues are computed since the eigenvalues quantify the contribution of a factor to the total variance (i.e., the sum of the eigenvalues). The factor extraction is done using a minimum acceptable eigenvalue that is >1 (Kaiser 1960). The factor-loading matrix is rotated to an orthogonal simple structure, according to varimax rotation, which results in the maximization of the variance of the factor loading of the variables. This procedure renders a new rotated factor matrix in which each factor is described in terms of only those variables and affords a greater ease for interpretation. Factor loading is the measure of the degree of closeness between the variables and the factor (Dalton and Upchurch 1978). The lower part of Fig. 2 is a summary of the methodology used for hydrochemical data collection and preparing hydrochemical data for the analysis.

Inverse modeling

Inverse modeling (Plummer 1992) has been widely used in interpreting the geochemical processes that account for the hydrochemical and isotopic evolution of groundwaters. One of the inverse modeling codes, NETPATH, is used in the present study.

NETPATH 2.0 (Plummer et al. 1994) is an interactive Fortran 77 computer program that can be used to calculate

Table 2 Statistical summaries of hydrochemical variables

Parameter	Range	Median	Mean	Standard deviation	Coefficient of skewness	Coefficient of kurtosis	Health limits of China
Ca ²⁺	28.1–429.1	208.1	184.4	118.2	0.76	−0.51	–
Mg ²⁺	34.6–332.5	112.1	126.8	87.2	2.55	0.76	–
Na ⁺	80.8–2,400	717.5	797.5	610.9	2.36	1.44	–
K ⁺	1.4–17.2	5.8	6.2	4.0	2.13	1.33	–
HCO ₃ [−]	67.1–399.3	220.2	218.2	96.1	0.61	−0.86	–
Cl [−]	47.5–1,567	533.1	584.3	431.7	1.52	−0.01	250
SO ₄ ^{2−}	111.3–4,911.5	1,457.5	1,655.4	1,252.2	2.23	1.29	250
F [−]	0.4–2.4	1.5	1.5	0.6	−0.06	−0.97	1.0
pH	7.4–8.4	8.0	7.9	0.2	−0.51	−0.1	6.5–8.5
TDS	189.0–9,114.9	2,975.8	3,246.4	2,352.3	2.12	1.34	1,000

All values are in mg/L except pH

the mass transfers in every possible combination of the selected phases that account for the observed changes in the selected chemical and (or) isotopic compositions observed along the flow path. NETPATH contains two Fortran 77 codes: DB.FOR and NETPATH.FOR. In the present study, we first use DB.FOR to input and edit the chemical data and then run WATEQF included in DB to calculate mineral saturation indices. NETPATH.FOR is then run to calculate every possible hydrochemical mass balance reaction.

Results and discussion

General hydrochemistry

A summary of the major hydrochemical variables of Tertiary fractured permeable confined aquifer is presented in Table 2. It indicated that SO₄^{2−}, Cl[−] and Na⁺ were the dominant ions in the study area. The pH in groundwater samples ranged from 7.4 to 8.4, indicating that the groundwater environment was slightly alkaline. The TDS concentration ranged from 189.0 to 9,114.9 mg/L with a mean concentration of 3,246.4 mg/L. The fluoride concentration ranged from 0.4 to 2.4 mg/L, and the mean concentration was 1.5 mg/L. About 58 % of the samples exceeded the desirable limit of fluoride for drinking water (1.5 mg/L), while 79 % exceeded that of TDS (1,000 mg/L) (WHO 2006).

Results of multivariate statistical analysis

Principal component analysis (PCA)

Tables 3 and 4 present the factor loadings matrix. Only the first two components extracted had eigenvalues >1, and accounted for 73 % of the total variance in the dataset. Varimax normalized rotation was applied to

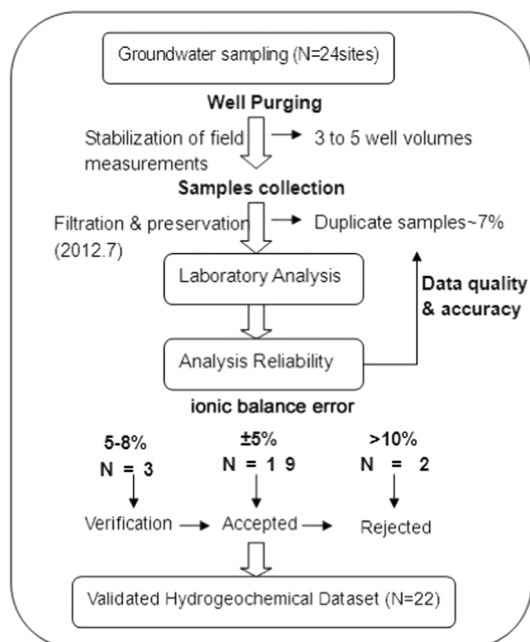
maximize the variance of the first two principal axes (Lorite-Herrera et al. 2008).

Factor 1 explained 55 % of the total variance and had high positive loadings for most of the major ions (Ca²⁺, Mg²⁺, Na⁺, K⁺, Cl[−] and SO₄^{2−}), which probably reflects the source of the groundwater salinity as these elements are the main components of TDS. Further, it was assumed that the salinity was mainly produced by the dissolution of mirabilite and halite, as the Tertiary formations contained appreciable quantities of halite and mirabilite (Table 1). Ca²⁺ and Mg²⁺ also had high positive factor loadings, which suggest that these two components contributed remarkably to the groundwater salinity. The most likely process was cation exchange as most groundwater samples' saturation index (SI) of calcite and dolomite was above 0 (Fig. 2, ESM only), which means that the groundwater has reached thermodynamic equilibrium with these mentioned minerals; therefore, the dissolution of these two carbonate minerals attributed scarcely to Ca²⁺ and Mg²⁺ in the groundwater.

Figure 3 (ESM only) shows that SO₄^{2−} and Na⁺ were the dominant anion and cation, respectively, especially when the TDS was above 1,000 mg/L. Then, it was crucial to identify the source of Na⁺ and SO₄^{2−}.

Samples were all distributed near the mirabilite and halite dissolution line (1:1 line), which indicated that Na⁺ was mainly from the dissolution of mirabilite and halite (Fig. 3a). However, some samples were on the above of the dissolution line, indicating that some Na⁺ was removed from groundwater by cation exchange process according to the reaction Ca(Mg)X(s) + 2Na⁺ → Na₂X(s) + Ca²⁺ (Mg²⁺) (where X is a solid substrate such as a clay mineral). During this process, Na⁺ in the solution was exchanged with Ca²⁺ and Mg²⁺ in the sediments. Then, it can be concluded that Na⁺ was mainly from the dissolution of mirabilite and halite. A part of Na⁺ was further modified by cation exchange. Moreover, the referred conclusion can be further confirmed through the relationship illustrated in

1. Hydrochemical Data Collection & Validation



2. Multivariate Statistical Analysis

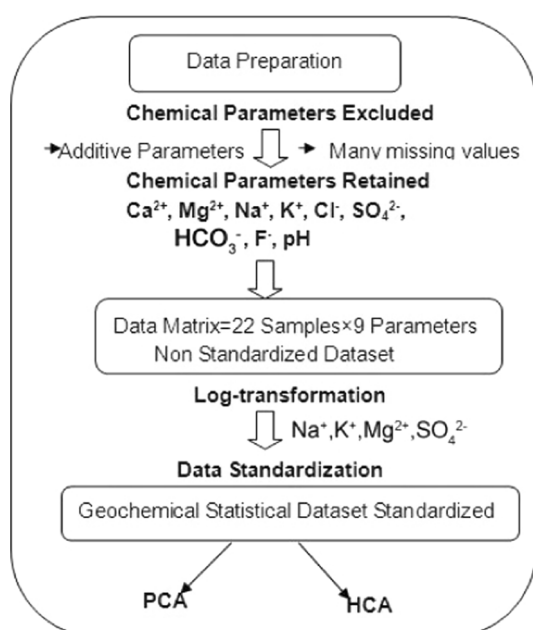


Fig. 2 Methodological flow chart, from groundwater sampling to data analysis

Fig. 3b, in which $(\text{Na} + \text{K} - \text{Cl})$ and $[(\text{SO}_4 + \text{HCO}_3) - (\text{Ca} + \text{Mg})]$ varied proportionally (Garcia et al. 2001; Carol et al. 2009; Moussa et al. 2010).

The positive correlation between Na^+ and SO_4^{2-} (Fig. 4a) suggested that the dissolution of mirabilite was the major source of SO_4^{2-} . However, a considerable

Table 3 Factor loadings matrix from R-mode factor analysis

Variable	Factor 1	Factor 2
Ca^{2+}	0.809	-0.033
HCO_3^-	-0.194	-0.755
Cl^-	0.787	0.221
F^-	0.254	0.746
pH	-0.441	0.593
Mg^{2+}	0.893	-0.229
Na^+	0.828	0.483
K^+	0.828	0.370
SO_4^{2-}	0.919	0.273

number of samples were under the 1:1 line, suggesting that SO_4^{2-} might have other sources. Indeed, the saturation indexes showed that some samples were undersaturated with respect to gypsum (Fig. 4b), indicating that gypsum dissolution was another source of SO_4^{2-} .

Factor 2, which accounted for 19 % of the total variance, contained a high positive loading of F^- and probably reflected the source of fluoride in the groundwater. It was found that HCO_3^- and pH also had high factor loadings, which means that the enrichment of F^- had a close relationship with HCO_3^- and pH (Table 3).

The values of pH for the majority of the high-fluoride groundwater samples were within the range of 7.4–8.4, indicating that the high-fluoride groundwater commonly belonged to alkaline water (Fig. 4a, ESM only). High-fluoride samples, however, displayed a negative correlation with HCO_3^- (Fig. 4b, ESM only), suggesting that HCO_3^- was against the release of fluoride to groundwater.

It has been proven that the dissolution of F^- -bearing minerals (mainly was fluorite, Speer 1984) would be enhanced by calcite precipitation and Ca^{2+} scavenging by cation exchange as well as desorption from Fe-oxides, all of that are the main mechanisms that could control fluoride concentration in natural waters. Comparing both mechanisms, dissolution always predominated at lower pH (primary source), whereas desorption always occurred at higher pH. In general, F^- was preferentially adsorbed to mineral surfaces under neutral to acidic conditions (Hiemstra and Van Riemsdijk 2000; Tang et al. 2009). However, increasing alkalinity in the aquatic reservoirs leads to the release of the adsorbed fluoride through desorption or through competition with HCO_3^- (Borgnino et al. 2013). Then, it can be assumed that some fluoride was released to groundwater by a competitive sorption with HCO_3^- (secondary source) (Currell et al. 2011; Borgnino et al. 2013). But Ca^{2+} had a really small factor loading with factor 2 (Table 3), which means that fluorite dissolution was not shown by factor 2 or most of Ca^{2+} was not from fluorite dissolution. Thus, the results obtained here

Table 4 Variances explained by the various factors

Factor	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of variance	Cumulative (%)	Total	% of variance	Cumulative (%)	Total	% of variance	Cumulative (%)
1	4.91	54.59	54.59	4.91	54.59	54.59	4.58	50.94	50.94
2	1.70	18.84	73.43	1.70	18.84	73.43	2.03	22.50	73.43
3	0.91	10.10	83.53						
4	0.65	7.18	90.71						
5	0.38	4.24	94.95						
6	0.28	3.16	98.11						
7	0.12	1.30	99.41						
8	0.05	0.52	99.93						
9	0.01	0.07	100						

Extraction method: principal component analysis

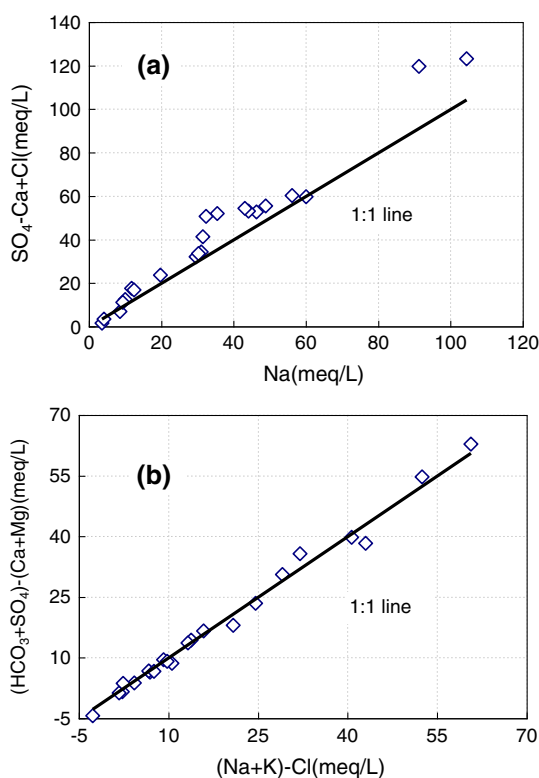


Fig. 3 a Na/(SO₄-Ca) + Cl relationship; b (Na + K-Cl)/(SO₄ + HCO₃)-(Ca + Mg) relationship

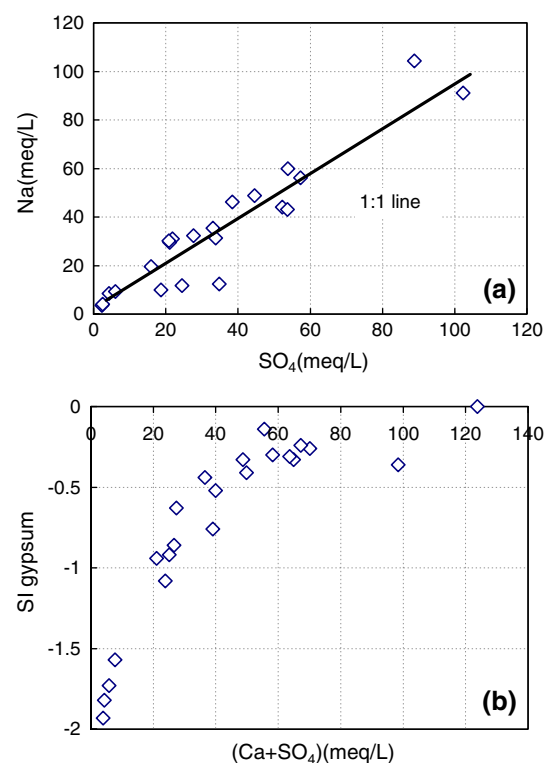


Fig. 4 a SO₄/Na relationship; b Geochemical relationships of (Ca + SO₄)/SI gypsum

indicated that desorption was the main contribution within the study area under alkaline conditions.

As the groundwater salinity was mainly due to the dissolution of sulfate minerals, it was necessary to identify the relationships between the increase of SO₄²⁻ and that of fluoride for these waters.

The activity coefficient of fluoride decreased from above 0.9 to 0.7 (Fig. 5a), indicating that with SO₄²⁻ added into the groundwater, the reduction of fluorine activity favors

further dissolution of F⁻-bearing minerals. Meanwhile, almost all samples were undersaturated with respect to fluorite (Fig. 5b), which suggested that, where present, they might be dissolved. However, the saturation index increased from -2.1 to 0 in coincidence with SO₄²⁻ increasing, indicating that groundwater was gradually saturated with respect to fluorite.

So, there are two major reasons for the observed elevation of fluoride content: one is the direct contribution of

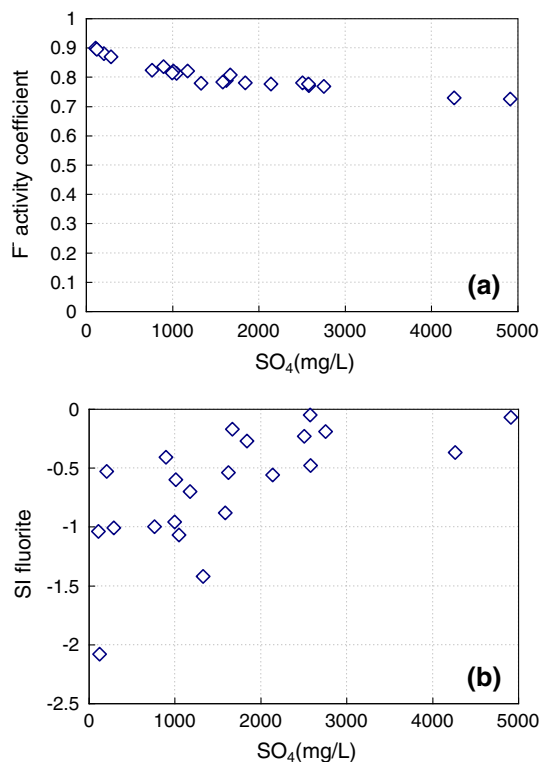


Fig. 5 Fluoride activity coefficient and Saturation Index of fluorite versus SO_4 plots of groundwater samples from Xiji Basin

fluorite dissolution, the other is the release of the adsorbed fluoride through desorption or through competition with HCO_3^- which appears to be more important.

Hierarchical cluster analysis (HCA)

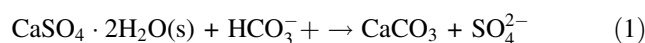
HCA has widely been used to classify hydrochemical data (Hussein 2004; Helstrup et al. 2007; Papaioannou et al. 2010; El-Hames et al. 2013). It has been proven that the high degree of spatial and statistical coherence could be used to support a model of hydrochemical evolution where the changes in water chemistry were a result of increasing water–rock interactions along the flow paths (Güler et al. 2002).

The main result of the HCA performed on the 22 groundwater samples was the dendrogram (Fig. 6). For this project, the Euclidean distance was used to measure the degree of similarity/dissimilarity between sampling sites. The initial cluster was formed by the linkage of the two samples with the greatest similarity, and then the steps were repeated until all observations have been classified. With this hydrochemical dataset, Ward's method was used to combine the clusters. This method was commonly applied in water chemistry investigations (Thyne et al. 2004; Papaioannou et al. 2010; El-Hames et al. 2013). Ward's method used an analysis of variance approach to

evaluate the distances between clusters, attempting to minimize the sum of squares of any two (hypothetical) clusters that can be formed at each step. Therefore, using the Euclidean distance as a distance measure and Ward's method as a linkage rule produced the most distinctive groups (Güler et al. 2002).

The classification of the samples into clusters was based on a visual observation of the dendrogram. In this study, the phenon line was drawn across the dendrogram at a linkage distance of 10. Thus, samples with a linkage distance lower than 10 were grouped into the same cluster. This position of the phenon line allowed a division of the dendrogram into two clusters of the groundwater samples, named Group 1 (include two subgroup, named G.1a and G.1b) and Group 2 and Stiff diagrams were generated for the arithmetic averages of the measures of the chemical variables used in this study (Fig. 6). Piper diagram was also generated (Fig. 7).

Each cluster (group) represented the position of its members in the groundwater flow regime and/or the key processes influencing groundwater hydrochemistry in the areas where the membership of these clusters was presented. For instance, G.1a samples were distributed mainly in the north area (Fig. 8), near the groundwater recharge zone and accepted the northern Cretaceous groundwater recharge. G.1a represented a 'Na–Mg– HCO_3 – SO_4 –Cl' groundwater type, and had low TDS (mean value = 318 mg/L). A significant feature of G.1a samples was that all major ions were in relatively small quantities and had a large proportion of HCO_3^- among the anions (Fig. 7). Group 2 samples were spread over the most studying areas (Fig. 8), and had the characteristics of groundwater from the drainage zone. Group 2 represented a 'Na– SO_4 –Cl' groundwater type with high TDS (mean value = 4,054 mg/L). Na^+ , Cl^- and SO_4^{2-} were significantly increased from G.1a to Group 2 (Fig. 7). The groundwater type of G.1b was, however, 'Ca–Na–Mg– SO_4 ' with moderate TDS (mean value = 1,867 mg/L) though they were also located in the middle and south area (Fig. 8). This moderate evolution groundwater type was probably because they were along the Moon Mountain–Shizi fault. The fault connected the shallow Quaternary 'Ca– HCO_3 ' groundwater with low TDS and deep confined Tertiary 'Na– SO_4 –Cl' groundwater with high TDS. Then, gypsum of the aquifer would dissolve and formed calcium carbonate deposits, leading to a 'Ca–Mg– SO_4 ' or 'Mg–Ca– SO_4 ' mixed type (Eq. 1).



These results further confirmed the assumption that the evolution of Tertiary fractured permeable confined groundwater was influenced by water–rock interaction flowing from G.1a to Group 2.

Fig. 6 Dendrogram for the groundwater samples, showing the division into two clusters and the mean concentrations and the mean concentrations Stiff diagram of each cluster

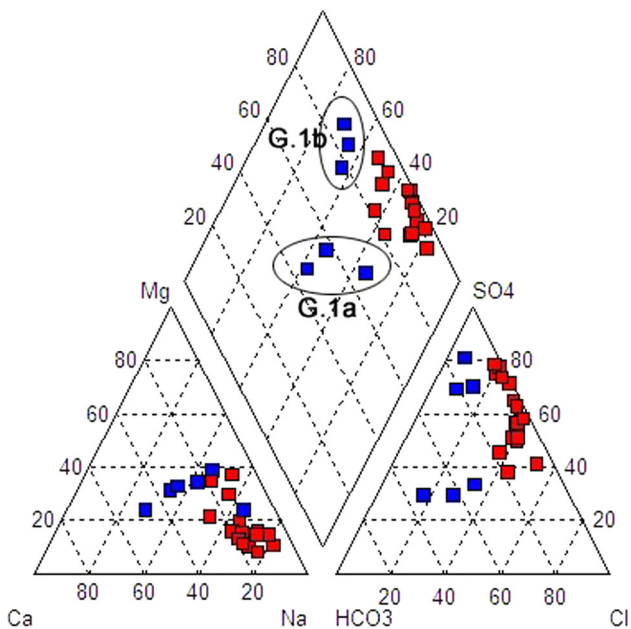
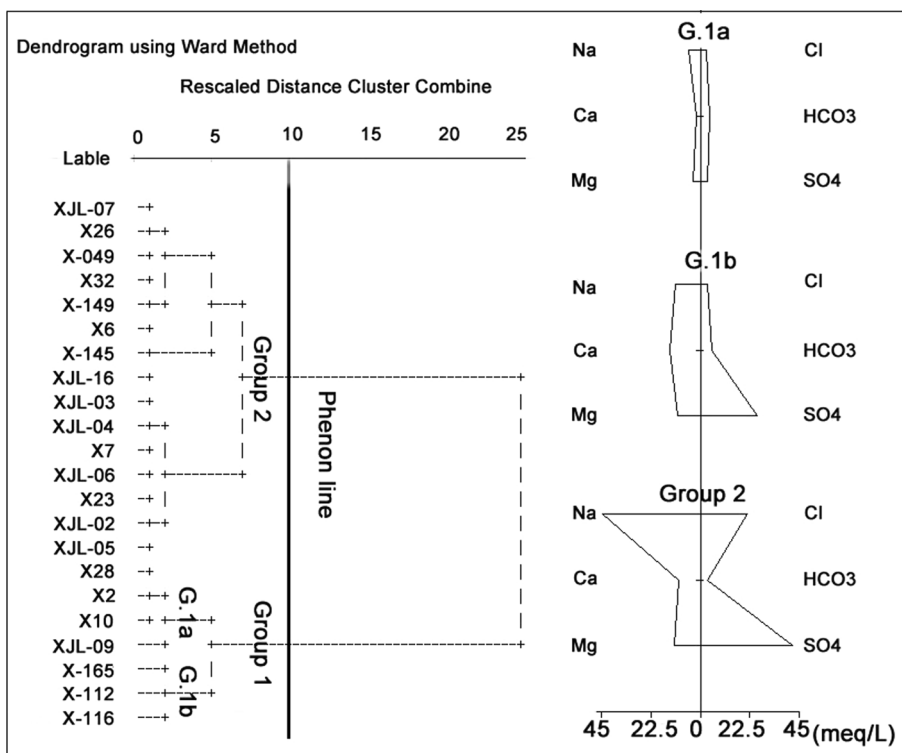


Fig. 7 Piper plot of groundwater chemistry

Inverse modeling

In an attempt to model the geochemical evolution of Xiji Tertiary confined groundwater systems, one representative flow path was evaluated using inverse modeling methods.

The I → I' line was selected as the representative flow path (Fig. 1a) because: (1) the final solute compositions were evolved; (2) it originated from groundwater recharge zone and ended in groundwater discharge zone. Based on the initial and the final water composition and the flow path mineralogy, inverse modeling was performed. NETPATH (Plummer et al. 1994) was used to evaluate the plausible mineral–water reactions along the flow path. NETPATH performed mass balance calculations for selected chemical compositions with minerals and gas phases. Considering NETPATH cannot validate results against SI, only modeling results that were consistent with the calculated SI of modeled mineral phases were considered plausible.

Well-characterized mineral assemblages for both reactant and product phases were critical for inverse reaction modeling (Bowser and Jones 2002). In NETPATH modeling, reactants and product phases were described in terms of constraints and phases. The major elements such as Ca, Mg, Na, C, Cl, and S were selected as constraints in this inverse modeling.

Samples given in Table 5 were chosen for the inverse modeling, with X10 sample as “initial” water and XJL05 sample as “final” water along the groundwater flow path. The information of the lithology, general hydrochemical evolution patterns and saturation indices was used to constrain the models. Mirabilite, magnesite, halite, gypsum and two common carbonate minerals (calcite and dolomite) were selected. Cation exchange was also chosen because it

Fig. 8 Regional distribution of groundwater clusters

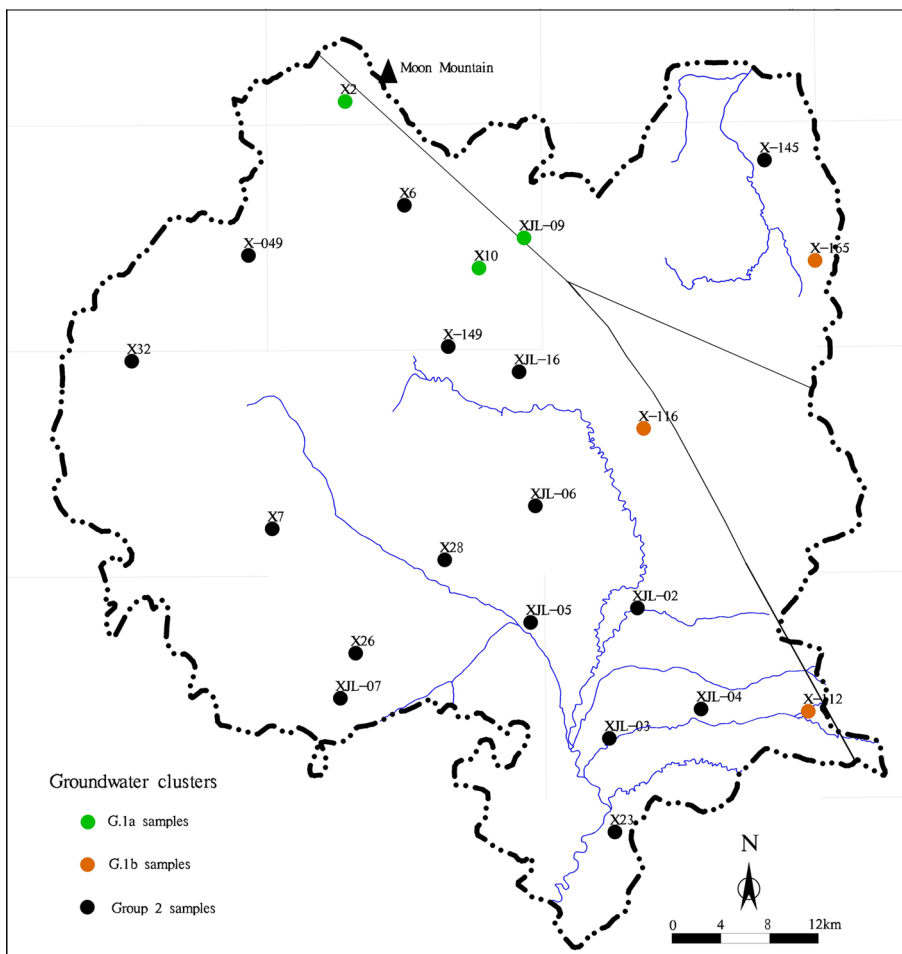


Table 5 Parameter values for the groundwater samples used in the inverse modeling calculations

Name	Ca ²⁺	Mg ²⁺	Na ⁺	K ⁺	HCO ₃ ⁻	Cl ⁻	SO ₄ ²⁻	F ⁻	pH	Temperature (°C)
X10	28.1	41.9	94.2	1.7	225.8	86.2	122.5	0.4	8.1	9.0
XJL05	227.8	85.4	1,013.2	6.8	125.1	435	2,508.7	2.1	8.0	12.0

was ubiquitous in aquifers with clay and fine particles. The aquifer was considered to be a closed system with respect to CO₂.

The detailed results of the modeling are presented in Table 6 and overall 5 models were present here.

It has been known that SO₄²⁻ was mainly from the dissolution of mirabilite (primary source) and gypsum (second sources). Na⁺ was mainly from the dissolution of mirabilite and halite. Cation exchange (Ca/Mg → Na) was probably existed and scavenging some Na⁺ of the groundwater. Based on these, model 4 was considered as the most reasonable one.

The bar graph of Fig. 9 presenting the results of model 4 was used to examine the influence of lithology on the dissolution and precipitation of mineral phases. Figure 9 shows the moles transferred from each phase, dissolved or

precipitated. In the area, the dominant processes were the dissolution of mirabilite, gypsum and halite along with the precipitation of calcite. Elevated calcium and magnesium concentration was likely related to the cation exchange (Na replacing Ca/Mg in fine-grained sediments such as clays found in the area).

Conclusion

About 58 and 79 % of the samples from Tertiary fractured permeable confined aquifer exceeded the desirable limits for fluoride (1.5 mg/L) and TDS (1,000 mg/L) in drinking water.

In the PCA, two principal factors were extracted from the dataset, which explained 73 % of the total data

Table 6 Summary of mass transfer for selected inverse models

Phases	Stoichiometry	Model 1	Model 2	Model 3	Model 4	Model 5
Halite	NaCl	9.96	9.96	9.96	9.96	9.96
Plagioclase	Ca _{0.38} Na _{0.62} Al _{1.38} Si _{2.62} O ₈	1.54	1.54	1.54	1.54	1.54
Calcite	CaCO ₃	-17.49	-6.30	-9.81	-2.12	-4.21
Dolomite	CaMg(CO ₃) ₂	7.68	2.09	-	-	-
Gypsum	CaSO ₄ ·2H ₂ O	23.18	8.85	23.18	3.50	8.85
Magnesite	MgCO ₃	-	-	7.68	-	2.09
Mirabilite	Na ₂ SO ₄ ·10H ₂ O	-	14.33	-	19.68	14.33
EXCHANGE	Ca ²⁺ /Mg ²⁺ → Na ⁺	14.33	-	14.33	-5.35	-

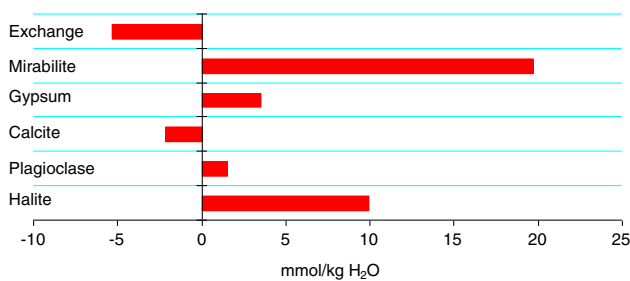


Fig. 9 NETPATH inverse modeling results: mmol/kg H₂O of different transferred phases in model. *Positive values* are for dissolved phases, and *negative values* are for precipitated ones

variability. Among them, factor 1 explained 55 % of the total variance and had high positive loadings for most of the major ions (Ca²⁺, Mg²⁺, Na⁺, K⁺, Cl⁻ and SO₄²⁻). In combination with the analyses of ion ratio, it revealed that the dissolution of mirabilite, halite and gypsum was the main source for the groundwater salinity. Factor 2, which accounted for 19 % of the total variance, contained a high positive loading for F⁻, pH and HCO₃⁻. It reflected that there were two major reasons for the observed elevation of fluoride content: one was the release of the adsorbed fluoride through desorption or through competition with HCO₃⁻ under alkaline conditions (primary source), the other was the direct contribution of fluorite dissolution (secondary source). All groundwater samples could be classified into two hydrochemical types by HCA: Group 1 (can be further divided into two subgroups, including G.1a and G.1b) and Group 2. G.1a located in the northwest region and represented a ‘Na–Mg–HCO₃–SO₄–Cl’ groundwater type with low TDS (mean value = 318 mg/L), reflecting the hydrochemistry of groundwater in recharge area; Group 2 distributed mostly in the south area and represented ‘Na–SO₄–Cl’ type with high TDS (mean value = 4,587 mg/L), reflecting the characteristics of groundwater fully evolved.

The dominant processes, which were revealed by inverse modeling, for the hydrochemistry evolution (mainly the increasing salinity) along the groundwater flow

regime were (1) the dissolution of mirabilite, gypsum and halite; (2) cation exchange, releasing Ca²⁺ and Mg²⁺ to groundwater; (3) the precipitation of calcite. The information provided in this study would be useful for current and future groundwater management in the region.

Acknowledgments This work was supported by the National Natural Science Foundation of China (41103045). The authors are grateful for the support provided by the “985 Project” and the “Key Laboratory of Groundwater Resources and Environment, Ministry of Education” of Jilin University.

References

Appelo CAJ, Postma D (2005) Geochemistry, groundwater and pollution, 2nd edn. Balkema, Netherlands

Atkinson JC (2011) Geochemistry analysis and evolution of a bolson aquifer, basin and range province in the southwestern United States. *Environ Earth Sci* 64:37–46

Borgnino L, Garcia MG, Bia G, Stupar YV, Coustumer PL, Depetris PJ (2013) Mechanisms of fluoride release in sediments of Argentina’s central region. *Sci Total Environ* 443:245–255

Bowser CJ, Jones BF (2002) Mineralogic controls on the composition of natural waters dominated by silicate hydrolysis. *Am J Sci* 302:582–662

Carol E, Kruse E, Mas-Pla J (2009) Hydrochemical and isotopical evidence of ground water salinization processes on the coastal plain of Samborombón Bay, Argentina. *J Hydrol* 365:335–345

Chen L, Zhang FW, Cheng YP, Lin WJ, Chen J, Zhang L (2009) Hydrogeochemical characteristics and evolution laws of groundwater in Haiyuan Basin, Ningxia. *Geoscience* 23(1):9–14 (in Chinese)

Cloutier V, Lefebvre R, Therrien R, Savard MM (2008) Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. *J Hydrol* 353:294–313

Currell M, Cartwright I, Raveggi M, Han DM (2011) Controls on elevated fluoride and arsenic concentrations in groundwater from the Yuncheng Basin, China. *Appl Geochem* 26:540–552

Dalton MG, Upchurch SB (1978) Interpretation of hydrochemical facies by factor analysis. *Groundwater* 16:228–233

Davis JC (1986) Statistics and data analysis in geology. Wiley, New York

Du SH, Su XS, Zhang WJ (2013) Effective storage rates analysis of groundwater reservoir with surplus local and transferred water used in Shijiazhuang City, China. *Water Environ J* 27:157–169

- El-Hames AS, Hannachi A, Al-Ahmadi M, Al-Amri N (2013) Groundwater quality zonation assessment using GIS, EOFs and hierarchical clustering. *Water Resour Manage* 27(7):2465–2481
- Farnham IM, Johannesson KH, Singh AK, Hodge VF, Stetzenbach KJ (2003) Factor analytical approaches for evaluating groundwater trace element chemistry data. *Anal Chim Acta* 490:123–138
- Gao H (2012) The water volume and water characteristics of Xiji County. *Value Eng* 28:75–76 (in Chinese)
- Garcia MG, Del Hidalgo M, Blesa MA (2001) Geochemistry of groundwater in the alluvial plain of Tucuman province, Argentina. *J Hydrol* 9:597–610
- Ghrefat HA, Yusuf N, Jamarh A, Nazzal J (2012) Fractionation and risk assessment of heavy metals in soil samples collected along Zerqa River, Jordan. *Environ Earth Sci* 66:199–208
- Güler C, Thyne G, McCray J, Turner K (2002) Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeol J* 10:455–474
- Helstrup T, Jorgensen NO, Banoeng-Yakubo B (2007) Investigation of hydrochemical characteristics of groundwater from the Cretaceous–Eocene limestone aquifer in southern Ghana and southern Togo using hierarchical cluster analysis. *Hydrogeol J* 15(5):977–989
- Hiemstra T, Van Riemsdijk WH (2000) Fluoride adsorption on goethite in relation to different types of surface sites. *J Colloid Interface Sci* 225:94–104
- Huang LM, Deng CB, Huang N, Huang XJ (2013) Multivariate statistical approach to identify heavy metal sources in agricultural soil around an abandoned Pb–Zn mine in Guangxi Zhuang Autonomous Region, China. *Environ Earth Sci* 68:1331–1348
- Huo FC, Pan XS, You GL (1989) Geological outline of Ningxia [M]. Science Press, Beijing, pp 45–72
- Hussein MT (2004) Hydrochemical evaluation of groundwater in the Blue Nile Basin, eastern Sudan, using conventional and multivariate techniques. *Hydrogeol J* 12(2):144–158
- Kaiser HF (1960) The application of electronic computers to factor analysis. *Educ Psychol Meas* 20:141–151
- Li XQ, Yu QS, Hou XW, Zhang L (2006) A study of the characteristics of groundwater circulation and the formation of bitter and saline groundwater in the Qingshuihe Basin in the southern Ningxia. *Hydrogeol Eng Geol* 1:46–51 (in Chinese)
- Lorite-Herrera M, Jimenez-Espinosa R, Jimenez-Millan J, Hiscock KM (2008) Integrated hydrochemical assessment of the Quaternary alluvial aquifer of the Guadalquivir River, southern Spain. *Appl Geochem* 23(8):2040–2054
- Matalas NC, Reicher BJ (1967) Some comments on the use of factor analysis. *Water Resour Res* 3:213–223
- Moussa AB, Zouari K, Marc V (2010) Hydrochemical and isotope evidence of groundwater salinization processes on the coastal plain of Hammamet-Nabeul, north-eastern Tunisia. *Phys Chem Earth* 36:167–178
- Papaoannou A, Dovriki E, Rigas N, Pleras P, Rigas I, Kokkora M, Papastergiou P (2010) Assessment and modeling of groundwater quality data by environmetric methods in the context of public health. *Water Resour Manage* 24(12):3257–3278
- Plummer LN (1992) Geochemical modeling of water-rock interaction: Past, present, future. Dissertation, Amsterdam
- Plummer LN, Prestemnn EC, Parkhurst DL (1994) An interactive code (NETPATH) for modeling net geochemical reactions along flow path Version 2.0. US Geological Survey Water Resources Investigations Report 94-4169
- Speer JA (1984) Micas in igneous rocks. In: Bailey SW (ed) *Reviews in mineralogy: micas*. Mineralogical Society of America, vol 13, pp 299–349
- Steinhorst RK, Williams RE (1985) Discrimination of groundwater sources using cluster analysis, MANOVA, canonical analysis and discriminant analysis. *Water Resour Res* 21:1149–1156
- Su XS, Xu W, Du SH (2013) Responses of groundwater vulnerability to artificial recharge under extreme weather conditions in Shijiazhuang City, China. *J Water Supply: Res Technol-Aqua*
- Tang Y, Guan X, Wang J, Gao N, McPhail MR, Chusuei CC (2009) Fluoride adsorption onto granular ferric hydroxide: effects of ionic strength, pH, surface loading, and major co-existing anions. *J Hazard Mater* 171:774–779
- Thyne G, Güler C, Poeter E (2004) Sequential analysis of hydrochemical data for watershed characterization. *Groundwater* 42(5):711–723
- Usunoff EJ, Guzmán-Guzmán A (1989) Multivariate analysis in hydrochemistry: an example of the use of factor and correspondence analyses. *Ground Water* 27:27–34
- WHO (2006) Guidelines for drinking-water quality. Recommendations, 4th edn. World Health Organization, Geneva
- Yang HF, Zhang FW, Wang GL (2009) Controlling structure and dividing result of groundwater system in Southern Ningxia. *Arid Land Geogr* 32(4):558–565 (in Chinese)
- Yidana SM, Banoeng-Yakubo B, Sakyi PA (2012) Identifying key processes in the hydrochemistry of a basin through the combined use of factor and regression models. *J Earth Syst Sci* 121(2):491–507
- Zhang FW, Chen L, Yu QS, Xue ZQ (2006) Research on the forming mechanism of the ground water quality in the Southern Ningxia Autonomous Region. *Geol Rev* 52(6):810–814 (in Chinese)
- Zumlot T, Batayneh A, Nazal Y, Ghrefat H, Mogren S, Elawadi E, Laboun A, Qaisy S (2013) Using multivariate statistical analyses to evaluate groundwater contamination in the northwestern part of Saudi Arabia. *Environ Earth Sci* 70:3277–3287