



Research papers

Using cluster analysis for understanding spatial and temporal patterns and controlling factors of groundwater geochemistry in a regional aquifer

Jing Yang^{a,b}, Ming Ye^{b,*}, Zhonghua Tang^{a,*}, Tian Jiao^a, Xiaoyu Song^c, Yongzhen Pei^d, Honghua Liu^e^a School of Environmental Studies, China University of Geosciences, Wuhan, Hubei 430074, China^b Department of Earth, Ocean, and Atmosphere Science, Florida State University, Tallahassee, FL 32306, USA^c Department of Civil and Coastal Engineering, University of Florida, Gainesville, FL 32611, USA^d School of Computer Science and Software Engineering, Tianjin Polytechnic University, Tianjin 300387, China^e Qingdao Geological Engineering Survey Institute and Key Laboratory of Urban Geology and Underground Space Resources, Shandong Provincial Bureau of Geology and Mineral Resources, Qingdao, Shandong 266100, China

ARTICLE INFO

This manuscript was handled by Corrado Corradini, Editor-in-Chief, with the assistance of Barbara Mahler, Associate Editor

Keywords:

Statistical analysis
Geochemical zones
Water-rock interactions
Jiangnan Plain
Three Gorges Dam

ABSTRACT

Understanding spatial and temporal patterns of groundwater geochemistry at the regional scale over a long time period is challenging, due to the lack of data and effective statistical approaches to characterize complex natural processes and anthropogenic activities. We applied a recently developed cluster analysis method to investigate spatial and temporal patterns and controlling factors of groundwater geochemistry in the confined aquifer of the Jiangnan Plain, China. The cluster analysis is applied to a dataset of 13,024 groundwater geochemical measurements for 11 geochemical parameters of 1,184 groundwater samples collected over 23 years from 29 monitoring wells distributed over the Jiangnan Plain. The cluster analysis yielded a classification of seven clusters, and the classification was confirmed by using principal component analysis and Stiff and Piper diagrams. Based on the spatial distribution of the seven clusters, the Jiangnan Plain is separated into four geochemical zones (i.e., recharge zone, transition zone, flow-through zone, and discharge-mixing zone) along the regional groundwater flow path, which has not been attempted in the past. The temporal changes of groundwater geochemistry are controlled by short- and long-term factors of water-rock interactions and anthropogenic activities. A particular finding of this study is that the Three Gorges Reservoir has a long-term impact on groundwater geochemistry, because the reservoir increased river discharge to groundwater after 2009. This study demonstrates that using the cluster analysis method together with hydrogeochemical analysis can identify spatiotemporal patterns and controlling factors of groundwater geochemistry at the regional scale over a long monitoring period.

1. Introduction

A thorough understanding of groundwater geochemistry is critical for protecting groundwater resources under the conditions of changing climate, growing population, and decreasing freshwater availability (Cloutier et al., 2008; Fendorf et al., 2010; Gorelick and Zheng, 2015; Han et al., 2016; Landon et al., 2011). Groundwater geochemistry is driven by various natural processes and anthropogenic activities (Güler and Thyne, 2004a; Güler et al., 2002; Reghunath et al., 2002; Tóth, 1999, 2009), and may vary both in space and time. Understanding spatial and temporal patterns of groundwater geochemistry not only requires designing a monitoring network to collect the right data (in terms of data constitutes, collection locations, collection times, etc.) for

tackling the problem of interest, but also requires using appropriate statistical methods to extract the spatial and temporal patterns embedded in measurements of groundwater geochemistry parameters. More importantly, hydrogeochemical analysis is needed to understand the natural and anthropogenic factors that control the spatial and temporal patterns of groundwater geochemistry. While obtaining a large amount of monitoring data for a region-scale study is not uncommon (e.g., Feyereisen et al., 2007; Güler and Thyne, 2004b; Jessen et al., 2017), it is still challenging to reveal the spatial and temporal patterns hidden in the data and to understand the controlling factors of the patterns at the regional scale. This study addresses these two challenges by using a recently developed method of cluster analysis together with hydrogeochemical analysis for a regional aquifer in

* Corresponding authors.

E-mail addresses: mye@fsu.edu (M. Ye), zhhtang@cug.edu.cn (Z. Tang).<https://doi.org/10.1016/j.jhydrol.2020.124594>

Received 11 December 2019; Accepted 14 January 2020

Available online 20 January 2020

0022-1694/ © 2020 Elsevier B.V. All rights reserved.

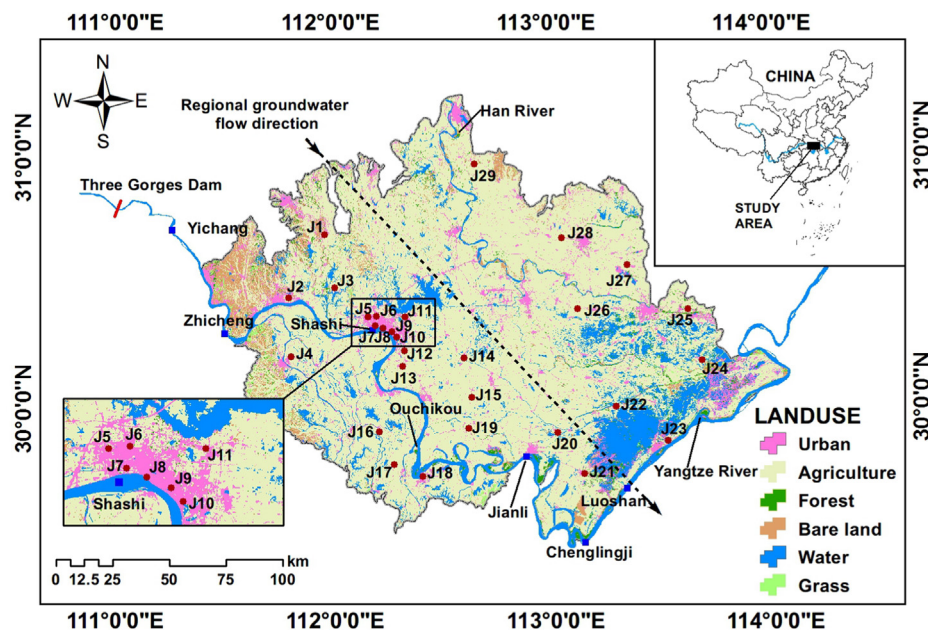


Fig. 1. Location of the study site (Jiangnan Plain), 29 monitoring wells of the aquifer (red dots), and six hydrological stations (blue squares) along the Yangtze River. Landsat Thematic Mapper satellite images of August 2014 (available at <https://www.gscloud.cn/> as of November 2019) were used to produce the land use map by using a supervised classification method with a maximum likelihood classification technique.

central China.

Using cluster analysis together with hydrogeochemical analysis has advanced our understanding on spatial and temporal patterns of groundwater geochemistry (Cloutier et al., 2008; Kim et al., 2003; Nguyen et al., 2015; Shrestha and Kazama, 2007; Simeonov et al., 2003; Wang et al., 2015). This is always done in two steps. In the first step, cluster analysis is conducted to classify groundwater geochemical data into a number of clusters, each of which reflects its own composition of groundwater geochemistry. Subsequently, hydrogeochemical analysis is conducted for the clusters to investigate spatial and temporal patterns in groundwater geochemistry. When dealing with spatiotemporal data obtained from a long-term monitoring network, many clustering methods have two limitations on how clusters are classified. One limitation is that cluster classification is conducted only for temporal means (i.e., the means over the entire sampling period) (e.g., Qian et al., 2007; Sayemuzzaman et al., 2018). While this kind of cluster analysis can help identify spatial patterns, it cannot be used to understand temporal patterns. The other limitation is that cluster classification is conducted separately to data of different sampling times or hydrological conditions (e.g., Hussain et al., 2008; Thyne et al., 2004). This kind of cluster analysis may be inadequate to simultaneously reveal spatial and temporal patterns when groundwater geochemistry substantially changes over time, because the number of clusters and the geochemical characteristics of the clusters can be dramatically different before and after the changes. Addressing the two limitations requires a cluster analysis method that deals with data in space and time simultaneously and can reveal spatial and temporal patterns.

Pacheco-Castro et al. (2018) recently developed a cluster classification method that has the potential to overcome the two limitations discussed above. Instead of conducting cluster classification for temporal means or data of different sampling times, the cluster analysis of Pacheco-Castro et al. (2018) is conducted for the entire monitoring data collected at all monitoring wells and sampling times. If groundwater samples are collected from m monitoring wells at s sampling times and p geochemical parameters are measured for each of $m \times s$ groundwater samples, the measurements form a matrix that has $m \times s$ rows and p columns. The cluster classification method of Pacheco-Castro et al. (2018) yields a column vector of $m \times s$ elements of clustering index for the combination of sampling wells and sampling times. This vector leads to s snapshots of clusters in space. Since the cluster classification is conducted for all monitoring data, the number of clusters and their

geochemical characteristics are the same for each snapshot. This makes it possible to jointly study the spatial and temporal patterns of groundwater geochemistry. Linking the snapshots to geological, hydrogeological, and geochemical knowledge of the site of interest can help understand the controlling factors of groundwater geochemistry.

This can be illustrated in the following two scenarios:

- (1) If the natural and anthropogenic factors controlling groundwater geochemistry are similar at several monitoring wells, the geochemistry at the wells should be similar, and the wells should belong to the same cluster. Studying the spatial distribution of the clusters can reveal large-scale patterns of groundwater geochemistry, which in turn helps identify the controlling factors of groundwater geochemistry in space.
- (2) If groundwater geochemistry changes at the wells, the wells may belong to another cluster. Therefore, studying the temporal changes of the clusters can reveal temporal patterns of groundwater geochemistry, which in turn helps identify the factors that control groundwater geochemistry in time.

Pacheco-Castro et al. (2018) applied their cluster classification method to 288 groundwater samples collected over three years (2009–2011) from a karst aquifer in Yucatan, Mexico. They found the following: (1) groundwater geochemistry at the west and in the coastal area of the site is controlled by seawater intrusion and sulfate-rich groundwater, (2) groundwater geochemistry at the middle and east part of the site is controlled by water-rock interactions and annual precipitation, respectively, and (3) groundwater geochemistry at two local areas of the site is controlled by anthropogenic activities. They also found that temporal variation of groundwater geochemistry at the site is caused by groundwater dilution due to changes in the amount and spatial distribution of precipitation. Since the cluster analysis of Pacheco-Castro et al. (2018) was for a small dataset collected over a short monitoring period, it is necessary to further evaluate their method for a large dataset over a long monitoring period.

The evaluation was conducted in this study by using a large amount of monitoring data collected from the confined aquifer of Jiangnan Plain, an alluvial plain located in central China with an area of approximately 27,400 km² (Fig. 1). The Yangtze River, the world's third longest river, flows through the plain, and the Three Gorges Dam, the world's largest hydroelectric project, is about 80 km upstream of the

plain (Fig. 1). Groundwater is mainly of Ca(Mg)-HCO₃ water type due to dissolution of carbonate minerals (Gan et al., 2014; Yu et al., 2017; Zhou et al., 2012). This area is undergoing extensive agricultural activities, and 72% of the land is for agricultural use (Fig. 1). Anthropogenic activities have affected groundwater geochemistry of the regional aquifer (Yang et al., 2018). To monitor groundwater quality, a monitoring network was established in 1990, and a large amount of monitoring data have been collected during the past three decades for a total of 21 groundwater geochemistry parameters specified by the National Quality Standard for Ground Water of China (Ministry of Environmental Protection of the People's Republic of China, 1994). The 21 parameters are pH, temperature, alkalinity, total dissolved solids (TDS), hardness, concentrations of seven major ions (Ca²⁺, Mg²⁺, K⁺, Na⁺, Cl⁻, SO₄²⁻, and HCO₃⁻), and concentrations of minor ions and trace constituents (NO₂⁻, NO₃⁻, NH₄⁺, Fe³⁺, Fe²⁺, Fe, Mn²⁺, F⁻, and total As). Groundwater contaminants reported in literature include arsenic (Duan et al., 2015, 2017; Gan et al., 2014) and wastewater infiltration (Niu et al., 2017) in the confined aquifer and nitrate (Yang et al., 2017, 2018) in the shallow aquifer of Jiangnan Plain. However, there are few published analyses of groundwater geochemistry at the plain scale, except the paper of Niu et al. (2017), who reported the temporal trends of multiple groundwater geochemistry (e.g., Cl⁻, SO₄²⁻, and NO₃⁻) but did not analyze spatial patterns of groundwater geochemistry. The overall spatial and temporal patterns of groundwater geochemistry and their controlling factors at the plain scale are still unknown.

This study used the groundwater geochemistry data collected over the period of 1992–2014. The dataset consists of a total of 13,024 geochemistry measurements for 1,184 groundwater samples, each of which has 11 selected groundwater geochemistry parameters (pH, Ca²⁺, Mg²⁺, K⁺, Na⁺, Cl⁻, SO₄²⁻, HCO₃⁻, NH₄⁺, F⁻, and Fe). This large dataset enables us to evaluate whether the cluster analysis method of Pacheco-Castro et al. (2018) can help better understand the spatial and temporal patterns of groundwater geochemistry and their controlling factors for a regional aquifer over a long period of 23 years. Based on the cluster analysis and hydrogeochemical analysis, we delineated for the first time four zones of groundwater geochemistry at the site which are related to aquifer recharge, regional groundwater flow, water-rock interactions, and anthropogenic activities. The four zones should be of value for future study of detailed groundwater geochemistry at the local scale. We also identified the long-term impacts of the Three Gorges Reservoir on groundwater geochemistry downstream of the reservoir. While the impacts of dam construction and operation on river hydrology have been well documented (Guo et al., 2018; Nilsson et al., 2005; Zhou et al., 2013), little is known about the impacts on groundwater geochemistry. This study provides a new insight on using cluster analysis for understanding long-term evolution of groundwater geochemistry at the regional scale.

2. Study area

Jiangnan Plain is a semi-closed basin with high elevation (~350 m) in the northwest and low elevation (~25 m) in the southeast portion of the plain, and the regional groundwater flow is from northwest to southeast. Based on precipitation data from the Jingzhou meteorological station (provided by the China Meteorological Data Service Center, available at <http://data.cma.cn/>), monthly average precipitation for the period of 1992–2014 ranges from 24 mm in December to 161 mm in July, with 41% of precipitation occurring from June to August. Mean annual rainfall increases from the northwest (~950 mm/yr) to the southeast (~1,350 mm/yr) (Wang, 2009). The Quaternary alluvial-lacustrine sediments deposited on the top of the bed rocks form the following three aquifers: the phreatic aquifer with Holocene and upper late Pleistocene deposits (0–20 m depth), the middle confined aquifer with late Pleistocene and middle Pleistocene deposits (20–100 m depth), and the deep confined aquifer with early Pleistocene (100–280 m depth). The main mineralogical compositions of the aquifer

sediments are clay minerals (about 29–50 wt%) and quartz (about 31–54 wt%). High contents of carbonate (up to 20 wt%) and albite (up to 21 wt%) are also detected in the sandy sediments (Duan et al., 2016, 2017). Discontinuous silty clay and embedded clay lenses (5–10 m in thickness) compose as the local aquitards between the phreatic and middle confined aquifer (Gan et al., 2018). The middle confined aquifer is the primary water supply for domestic, agricultural, and industrial water uses, and accounts for 64% of the total groundwater exploitation in Jiangnan Plain (Zhao, 2005). The studies of Duan (2016) and Gan et al. (2014) indicate that the middle confined aquifer is under reducing conditions with an average oxidation-reduction potential of -102 mV, and that the aquifer is enriched with organic matter with an average concentration of dissolved organic carbon of 6 mg/L. Water geochemistry of the middle-confined aquifer is the focus of this study, and all groundwater samples considered in this study were taken from this aquifer. The middle confined aquifer is referred to herein as the aquifer.

Analyses of hydrogen and oxygen stable isotopes suggest that rainfall is the main recharge source of the aquifer (Du et al., 2017). Groundwater recharge mainly occurs in the west part of the plain, where unconfined or semi-confined conditions prevail due to outcrops of the aquifer sediments. Leakage from the Yangtze River is another important recharge source as its thalweg elevation is lower than the elevation of the confining layer of the aquifer for most of the river reach downstream from the Three Gorges Dam (Fig. S1 in the Supplementary Information). Due to the direct interaction between the Yangtze River and the aquifer along the river, groundwater discharges to the river during dry seasons and receives recharge from the river during wet seasons (He and Tang, 2017). The invading river water affects groundwater geochemistry in the aquifer by mixing waters with different chemical compositions, especially after the quasi-normal stage of the Three Gorges Dam, as discussed in detail below.

3. Data and statistical procedures

Groundwater samples were collected from 29 monitoring wells (Fig. 1) twice a year during 1992–2014, one in the dry season from January to February and the other in the wet season from July to August. Since not all the wells were sampled regularly over the 23 years, the total number of samples is 1,190, less than 1,334 = 29 × 23 × 2. For the 21 groundwater geochemistry parameters, 11 parameters were selected for the cluster analysis, and they are pH, Ca²⁺, Mg²⁺, K⁺, Na⁺, Cl⁻, SO₄²⁻, HCO₃⁻, NH₄⁺, F⁻, and Fe. Temperature was excluded because it mainly reflects physical characteristic of groundwater but this study is focused on groundwater geochemistry. Hardness, TDS, and alkalinity were excluded because they can be obtained directly from the 11 parameters. For example, alkalinity can be approximated by the concentration of HCO₃⁻ given that pH values for most samples are less than 8.3 (Appelo and Postma, 2005). Concentrations of Mn²⁺, NO₂⁻, NO₃⁻, As, Fe²⁺, and Fe³⁺ were excluded because they were not continuously measured over the 23 years. Since NO₃⁻ concentrations are of growing concern to groundwater quality, we conducted another cluster analysis for a smaller dataset with NO₃⁻ included as a parameter, and the results are compared with those for the 11 parameters in Section 5.2.

For the 11 selected geochemical parameters, censored data (measured concentration values below detection limits) were reported for 335 groundwater samples for NH₄⁺, 35 samples for Fe, 14 samples for SO₄²⁻, 9 samples for Cl⁻, and 8 samples for F⁻. A common practice of processing censored data is to either exclude them from water quality analysis (Güler and Thyne, 2004b) or replace them with a value lower than detection limits (Cloutier et al., 2008; Sanford et al., 1993). To utilize the data to the extent possible and by following VanTrump and Miesch (1977), we replaced the censored data by 75% of the detection limits. Afterward, the charge balance error of the geochemical measurements was calculated. Following Güler et al. (2002) and Ghesquière et al. (2015), six samples with charge balance error above 10% were

excluded from further analysis. The final dataset used for the cluster analysis includes 13,024 groundwater geochemical measurements, forming a data matrix of 1,184 rows and 11 columns, corresponding to 1,184 groundwater samples and 11 geochemical parameters of each sample. Cluster analysis for this large amount of groundwater geochemical measurements is seldom reported in literature.

For the 13,024 groundwater geochemical measurements, hierarchical cluster analysis and principal component analysis were conducted by using Python with SciPy v0.19 library (Jones et al., 2001). Since the two statistical analyses have been well described in literature (Cloutier et al., 2008; Güler et al., 2002; Pant et al., 2018), we only describe here how the two analyses were implemented in this study. Following the literature (Güler et al., 2002; Ghesquière et al., 2015; Pacheco-Castro et al., 2018; Qian et al., 2007), we first applied the log-transformation to the geochemical data, and then standardized the transformed data by using the z-score calculation (subtracting sample mean and dividing the residuals by sample standard deviation) to remove the impacts of parameter units on the statistical analysis. The hierarchical cluster analysis of this study used the Ward's method with Euclidean distance (Ward, 1963), which calculates the distance between a pair of water samples in the space of the 11 groundwater geochemical parameters. The Ward's method uses an analysis of variance approach to evaluate the distances between clusters, and minimizes the sum of squares of distance between any two clusters that can be formed at each step of the cluster classification. This produces a dendrogram – a visual representation of the linkage distance during the history of cluster merging. The number of clusters is determined based on the phenon line, and changing the location of the phenon line on the dendrogram changes the number of clusters.

The principal component analysis was conducted independently from the hierarchical cluster analysis to verify whether the determined cluster number is reasonable. The principal component analysis is a data transformation technique for reducing the dimensionality of a large dataset by selecting a new set of uncorrelated variables that have lower dimensionality than the original data (Davis, 1990). For our dataset, the principal component analysis gave 11 eigenvalues (corresponding to the variance of the 11 geochemical parameters) and 11 eigenvectors (corresponding to 11 components that are linear combinations of the 11 geochemical parameters). The number of principal components is determined based on the Kaiser criterion, according to which only components with eigenvalues greater than one are selected (Kaiser, 1960). Using the principal components, the scores of the 1,184 groundwater samples are calculated and used to examine whether the determined cluster number is reasonable.

4. Classified seven clusters

This section presents the results of the hierarchical cluster analysis and the results of statistical analysis and geochemical analysis that were used to evaluate whether the classified seven clusters are reasonable. Fig. 2 shows the dendrogram of the hierarchical cluster classification for the 1,184 groundwater samples, and the phenon line drawn at the linkage distance of 25 led to seven clusters, denoted as Cluster C1–C7. To understand the geochemical characteristics of the seven clusters, the box plots of the groundwater geochemical parameters are shown in Fig. 3 for the seven clusters. Cluster C1 is characterized by low median concentration of F⁻ and high median concentrations of Ca²⁺, Mg²⁺, K⁺, HCO₃⁻, NH₄⁺ and Fe. Similar to Cluster C1, Cluster C2 also has low median concentration of F⁻ and high median concentrations of Ca²⁺, Mg²⁺ and HCO₃⁻. This is not surprising because of the short linkage distance between Clusters C1 and C2 (Fig. 2). Cluster C3 is characterized by substantially elevated concentrations of Na⁺ and Cl⁻. Cluster C4 is characterized by high concentrations of SO₄²⁻. For Clusters C5 and C6, their median concentrations of the majority of the geochemical parameters are close, except that median concentrations of SO₄²⁻ and F⁻ are significantly smaller in C5 than in C6. Cluster C7 is characterized by

low concentrations for all the parameters except K⁺, Cl⁻, SO₄²⁻, and NH₄⁺. The box plots allow a direct comparison of the main geochemical characteristics for the seven clusters.

Hydrogeochemical analysis (through Stiff and Piper diagrams) and principal component analysis were used to examine whether the classified seven clusters are reasonable. Fig. 2 plots the Stiff diagrams of the seven clusters based on the mean concentrations of the groundwater geochemical parameters. The Stiff diagrams display similarities and dissimilarities between the seven clusters. In particular, because of the high mean concentrations of Na⁺ and Cl⁻ (Fig. 3), the Stiff diagram of Cluster C3 is an upside-down triangle, which is substantially different from the shapes of the other six Stiff diagrams. It is noted that the Stiff diagrams of Clusters C2 and C4–C7 are similar. The five clusters are separated mainly because of the different concentrations of Ca²⁺, Mg²⁺, and HCO₃⁻, as shown in Fig. 3. For example, Fig. 3(h) shows that the HCO₃⁻ concentrations of Cluster C2 are significantly higher than those of Clusters C5–C7.

Fig. 4 plots the Piper diagram for the seven clusters of the 1,184 groundwater samples to further examine whether the cluster classification is reasonable. This figure indicates that the groundwater samples are mainly of the Ca(Mg)–HCO₃ water type, and this agrees with the findings of Gan et al. (2014, 2018) and Zhou et al. (2012). The groundwater samples of Cluster C3 are separated from those of the other six clusters, due to the high concentrations of Na⁺ and Cl⁻ of Cluster C3 samples. In the diamond plot, the distinction between the groundwater samples of Clusters C4–C7 is observed, and this confirms the classification of Clusters C4–C7, despite that the Stiff diagrams of the four clusters are similar. The groundwater samples of Clusters C1 and C2 overlap on each other due to similarly high concentrations of Ca²⁺, Mg²⁺ and HCO₃⁻ of the two clusters, as explained above. As a summary, the Stiff and Piper diagrams suggest that the classification of seven clusters is reasonable with respect to grouping the groundwater geochemical measurements and to revealing specific characteristics of groundwater geochemistry associated with each cluster.

The principal component analysis was conducted to further examine whether the classification of the seven clusters is reasonable. Table 1 lists the principal component loadings obtained after the Varimax normalized rotation for the first three principal components (with eigenvalues greater than one) as well as their explained variance. Following Liu et al. (2003), the strong loadings with absolute values larger than 0.75 are highlighted. Component 1 explains 35% of the total variance, and is characterized by the high loadings of Ca²⁺, Mg²⁺, and HCO₃⁻. Component 2 explains 18% of the total variances, and is dominated by the high loadings of Cl⁻ and SO₄²⁻. Component 3 only explains 10% of the total variance, and is thus less important than Components 1 and 2. Therefore, the analysis below is mainly based on Components 1 and 2. Based on the scores of the two components, Fig. 5 plots the groundwater samples to examine whether the classification of seven clusters is reasonable. To give a cloud view of the clustered groundwater geochemical measurements, seven ellipses are plotted for the seven clusters with the 80% confidence level, based on the eigenvalues and eigenvectors of the principal component analysis. The seven clusters are reasonably distinguished from each other in Fig. 5, despite marginal overlapping, indicating that the cluster classification is reasonable with respect to grouping the groundwater geochemical measurements.

It is worthwhile pointing out the difference between the results (cluster analysis and principal component analysis) and those of Gan et al. (2018). Based on 457 water samples (21 surface water samples, 91 phreatic groundwater samples, and 345 confined groundwater samples) collected in 2014 and 2015, Gan et al. (2018) applied cluster analysis and principal component analysis to the geochemical measurements for investigating the hydrogeological evolution in a small portion (1,800 km² in area) of Jiangnan Plain. Their principal component analysis gave three principal components corresponding to water-rock interactions, redox conditions, and anthropogenic activities. The

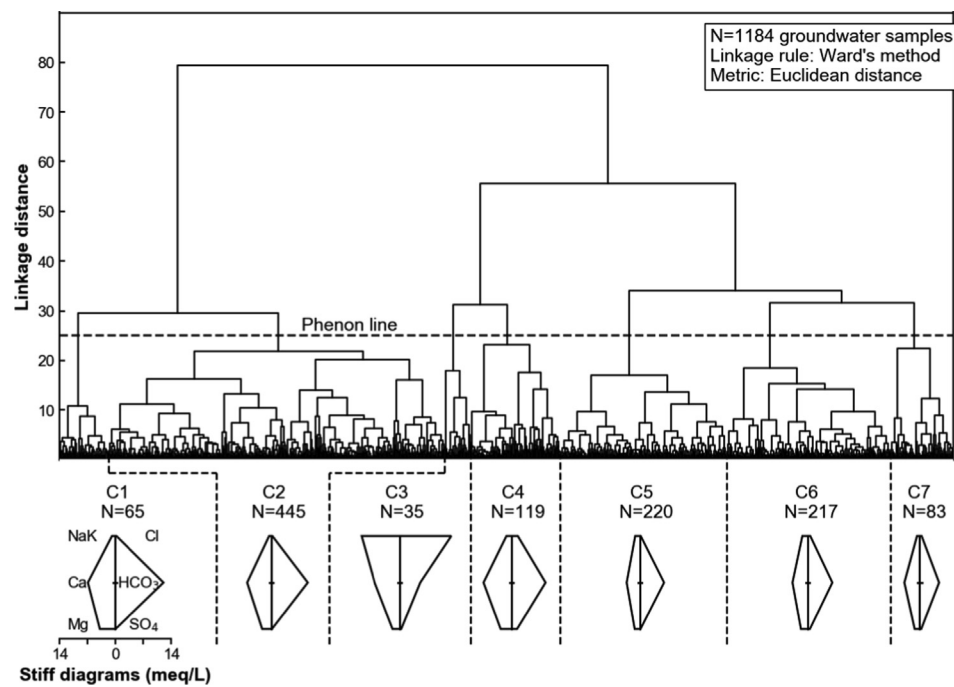


Fig. 2. Dendrogram of hierarchical cluster analysis for the 1,184 groundwater samples. Seven clusters are determined by the phenon line at linkage distance of 25. Stiff diagrams are plotted for the seven clusters based on the mean values of the geochemical parameters.

component of redox conditions was associated with the surface water sample and phreatic groundwater samples, because the redox conditions in surface water and phreatic groundwater are substantially different from those in confined groundwater. All our data are from the confined groundwater in a reducing environment. The reducing environment may affect some redox sensitive elements such as NO_3^- and Fe^{2+} in groundwater. For the NO_3^- concentrations measured for 1,067 groundwater samples, 582 measurements were censored, i.e., measured concentrations being lower than detection limit of 0.01 mg/L. This may be due to denitrification that occurs in reducing environments with enriched organic carbon as in the aquifer of this study. For the concentrations of Fe^{2+} and Fe^{3+} , after replacing the censored data by 75% of the detection limits, the mean concentrations of Fe^{2+} and Fe^{3+} are 2.37 mg/L and 0.42 mg/L, respectively; the higher concentrations of Fe^{2+} indicate a reducing environment in the aquifer, which is consistent with the finding of Niu et al. (2017) and Duan (2016). Note that, due to missing measurements of Fe^{2+} and Fe^{3+} concentrations, we only used total Fe concentrations in our study.

5. Spatial patterns and controlling factors

The hierarchical cluster analysis gave an array with 1,184 elements of cluster index (C1–C7) for the 1,184 groundwater samples. The cluster indices for each sampling time were extracted and plotted as a snapshot to show their spatial distributions at the well locations. The snapshots are the basis for investigating the spatial and temporal patterns of groundwater geochemistry.

5.1. Stable spatial patterns

Fig. 6 illustrates the snapshots of the seven clusters for the dry and wet seasons of the period of 2004–2014; the snapshots for the period of 1992–2003 are shown in Fig. S2 of the supplementary file. A visual comparison of the snapshots leads to a conclusion that, for each year, the spatial patterns of the dry and wet seasons are similar. For example, in the both dry and wet seasons, the wells of Cluster C2 are located in the central part, and the wells of Cluster C5 in the lower-left part of the study site. Over the 23 years, the spatial patterns of the clusters are

similar over time only with a few notable exceptions, indicating stable groundwater geochemistry.

Table 2 also suggests stability of the groundwater geochemistry at the study site. The table lists the number of samples of each cluster for the individual monitoring wells, and the last column of the table is the highest relative frequency. Taking well J1 as an example, this well belongs to Cluster C7 for 39 times, to C6 for 3 times, and to C4 for 2 times; the highest frequency is 39, and the highest relative frequency is $39/44 \approx 89\%$. A higher frequency indicates a more stable condition that groundwater geochemistry changes less frequently over time. For 22 out of the 29 monitoring wells, the highest frequency is larger than 80%, indicating that, for 80% of the sampling times, the spatial patterns of groundwater geochemistry do not change. Based on the stable spatial patterns and the controlling factors of groundwater geochemistry discussed below, four zones of groundwater geochemistry were delineated, and more details of the delineation are given in Section 6.

5.2. Controlling factors

The two components of principal component analysis (Table 1 and Fig. 5) reflect the controlling factors of the groundwater geochemistry. Component 1 is related to water-rock interactions because of its association with Ca^{2+} , Mg^{2+} and HCO_3^- in reference to natural weathering processes of carbonate minerals. Component 2 is related to the anthropogenic activities due to strong loadings of Cl^- and SO_4^{2-} , as the two ions are generally originated from anthropogenic activities at the study site (Gan et al., 2018; Zhou et al., 2012). This section provides detailed hydrogeochemical analysis to investigate the two controlling factors.

5.2.1. Water-rock interactions

Due to lacking detailed description of mineralogy of the aquifer, a full list of geochemical reactions (e.g., carbonate dissolution and silicate hydrolysis) are unknown. Based on the ratios of ion concentrations, we postulate the following five groundwater geochemical reactions to understand water-rock interactions in the aquifer,



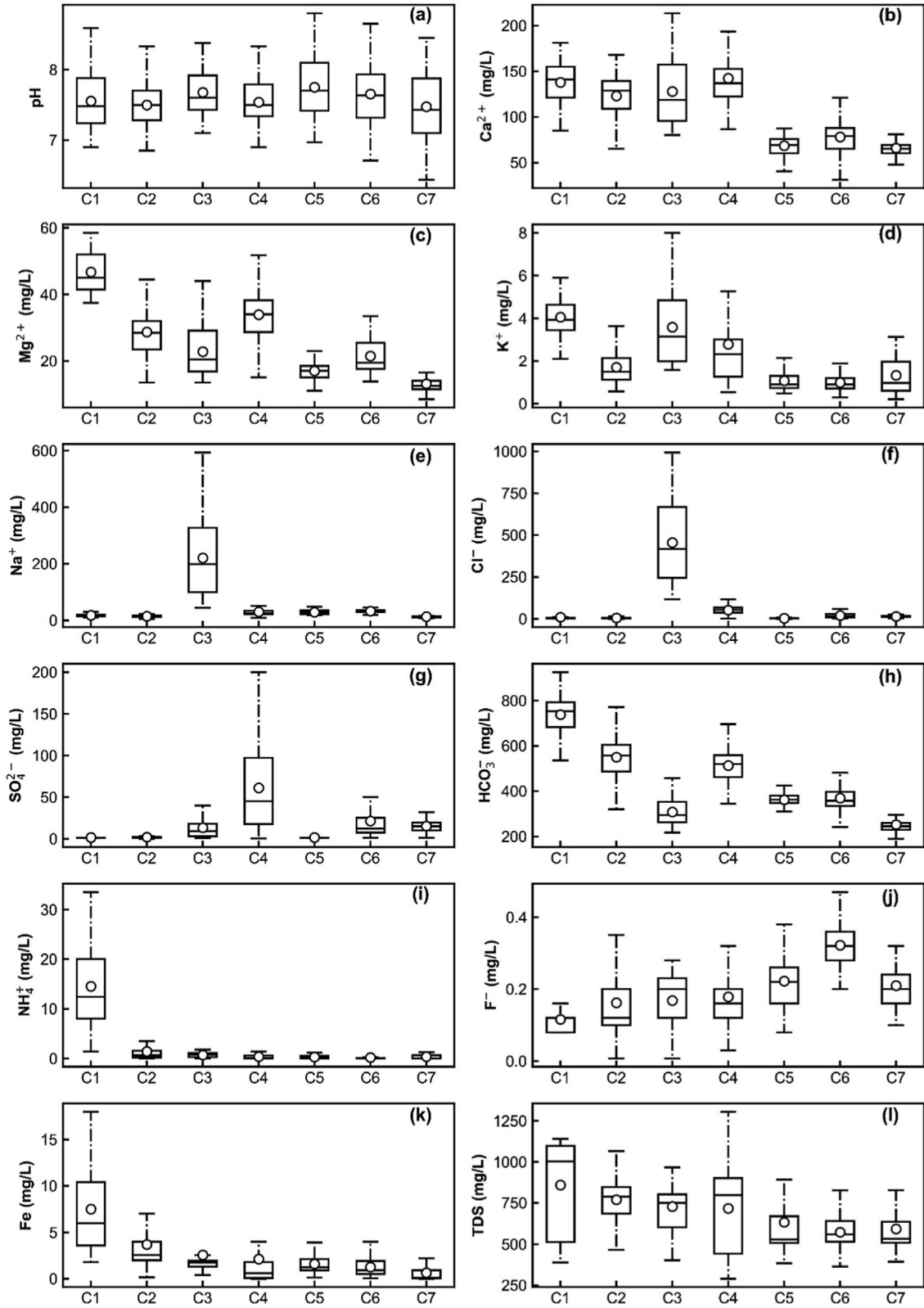
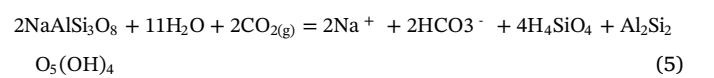
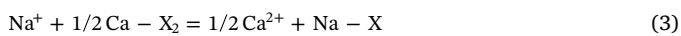
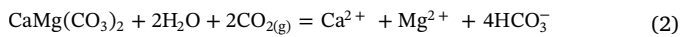


Fig. 3. Box plots of the eleven selected geochemical parameters for the seven clusters. Note TDS was not used for the statistical analysis and it was presented for comparison purpose. Open circles represent the mean values.



The first two reactions are for the dissolution of calcite and dolomite, the third reaction is for the cation exchange between Na^{+} and

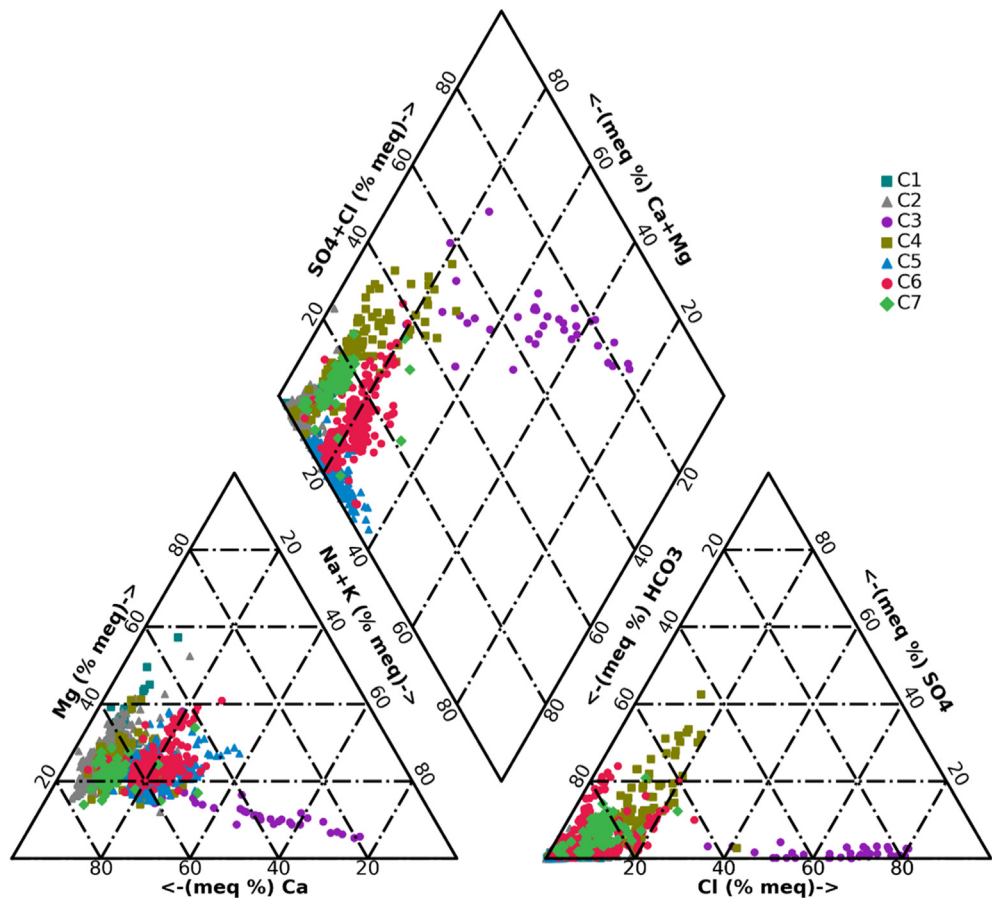


Fig. 4. Piper diagram for seven clusters of the 1,184 groundwater samples. The samples of one cluster are marked with the same color.

Table 1
Principal component loadings and explained variance for the first three components with Varimax normalized rotation. Loading values larger than 0.75 are in bold.

Parameter	Component 1	Component 2	Component 3
pH	−0.13	0.01	0.76
Ca ²⁺	0.82	0.09	−0.38
Mg ²⁺	0.82	0.02	−0.19
K ⁺	0.70	0.15	0.13
Na ⁺	−0.08	0.67	0.40
Cl [−]	0.06	0.87	0.05
SO ₄ ^{2−}	−0.14	0.80	−0.21
HCO ₃ [−]	0.82	−0.28	−0.32
NH ₄ ⁺	0.66	−0.36	0.15
Fe	0.58	−0.40	0.31
F [−]	−0.50	0.08	0.07
Eigenvalue	3.92	1.99	1.10
Explained variance (%)	35.68	18.12	10.00

Ca²⁺ (X denoting the exchanger resulting in increasing concentration of Ca²⁺), the fourth reaction is for the dissolution of halite, and the last reaction is for the albite hydrolysis.

Fig. 7 examines the above reactions at the study site by using the ion ratios. Fig. 7(a) plots the concentrations of Ca²⁺ and HCO₃[−], and the data (except those of Cluster C3) fall between the 1:2 and 1:4 lines. This is an evidence of calcite and dolomite dissolution (Wang et al., 2006), because the mole ratio of Ca²⁺/HCO₃[−] is 1:2 for calcite dissolution (Eq. (1)) and 1:4 for dolomite dissolution (Eq. (2)). Fig. 7(b) plots the concentrations of (Ca²⁺ + Mg²⁺) and HCO₃[−], and shows that the data (except those of Clusters C3 and C4) fall on the 1:1 line, suggesting that the concentrations of Ca²⁺, Mg²⁺, and HCO₃[−] can be largely explained by the dissolution of carbonate minerals. The data of Cluster C3 deviate

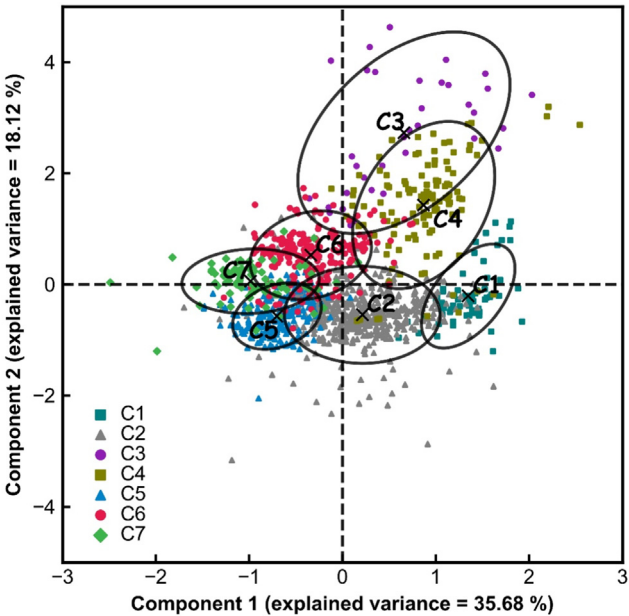


Fig. 5. Principal component scores for the first two components for the 1,184 groundwater samples. The samples of one cluster are marked with the same color, and the seven ellipses for the seven clusters are drawn with the 80% confidence level. Black crosses represent the mean scores for each cluster.

from the 1:1 line, and have larger Ca²⁺ concentrations. This can be explained by Eq. (3) of cation exchange (Appelo and Postma, 2005), considering that Na⁺ concentrations of Cluster C3 are significantly

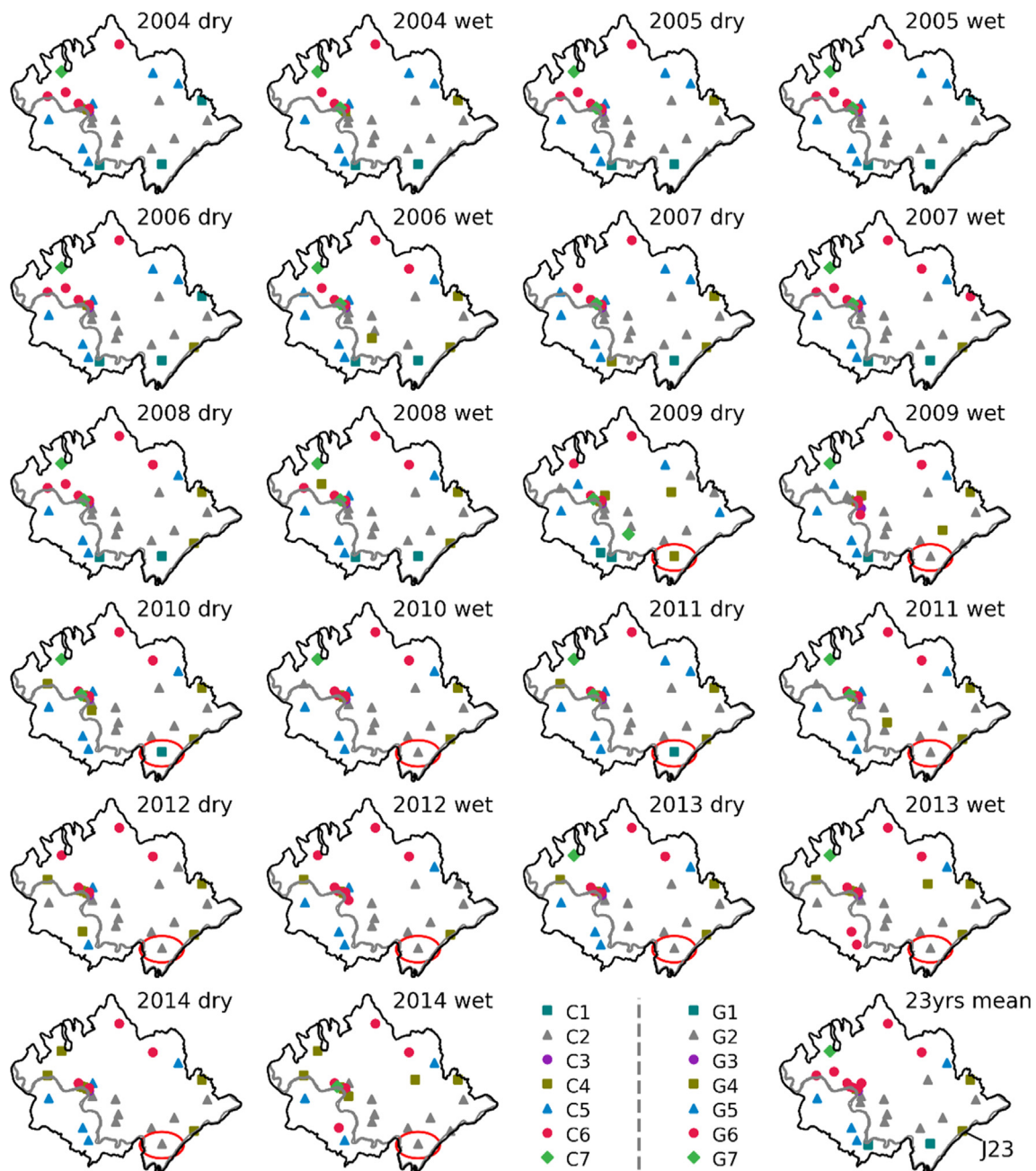


Fig. 6. Snapshots (except for the last one separated by dashed line) of the spatial distribution of the seven clusters C1–C7 for 2004–2014. The snapshots of the other years are shown in Fig. S1. The Yangtze River is represented by the grey line within the study area. Red open ellipses illustrate the temporal variation of the clusters for well J21 since 2009. The last snapshot shows the spatial distribution of clusters G1–G7 classified by using a conventional cluster analysis based for temporal averages of groundwater geochemistry parameters.

higher than Ca^{2+} concentration (Fig. 3) and that the aquifer is enriched in clay minerals and organic matter. The deviation of Cluster C4 data may be caused by the high SO_4^{2-} concentrations of this cluster (Fig. 7(c)) that result in low HCO_3^- concentration due to ionic balance. Fig. 7(d) shows that Na^+ concentrations are larger than Cl^- concentrations in Clusters C1, C2, C5, and C6 that are distributed over the study area. The larger Na^+ concentration is attributed to the dissolution of albite (Eq. (6)), because the Na^+ and Cl^- concentrations due to the dissolution of halite (Eq. (4)) should fall on the 1:1 line (Rajmohan and Elango, 2003; Rina et al., 2012). The albite dissolution may also explain the relatively high HCO_3^- concentrations shown in Fig. 7(b) for Clusters C1, C2, C5, and C6, as albite dissolution produces HCO_3^- .

Between calcite dissolution (Eqs. (1) and (2)) and albite hydrolysis (Eq. (5)), the former is more important than the latter for controlling

groundwater geochemistry in the study site for several reasons. First, silicate hydrolysis requires a weakly acid condition, but groundwater in the aquifer is slightly alkaline with average pH of 7.58, which is demonstrated in Fig. 3(a). Therefore, silicate hydrolysis may not be important for the alkaline groundwater. In addition, although the weight percentages of albite and carbonate are very close (21% vs. 20%), the concentration of Na^+ ranges between 0.1 and 2 mmol/L (Fig. 7(d)), but the concentration of Ca^{2+} ranges between 1 and 5 mmol/L (Fig. 7(a)). This indicates that albite hydrolysis is less important than carbonate dissolution. The last line of evidence is shown in Fig. 7(e) and (f), which plot the relation between the saturation index of calcite/dolomite and the concentration of $\text{Ca}^{2+}/\text{Mg}^{2+}$ in four zones of groundwater geochemistry from the recharge area to the discharge area of the aquifer (more details of the four zones are given in Section 6). The two figures

Table 2

Number of groundwater samples in each cluster for the 29 monitoring wells. Highest relative frequency is calculated by dividing the highest frequency (bold number) by the total sample number for each well.

Well	C1	C2	C3	C4	C5	C6	C7	Total	Highest relative frequency (%)
J1	0	0	0	2	0	3	39	44	89
J2	0	4	0	8	3	29	1	45	64
J3	0	1	0	0	0	45	0	46	98
J4	0	4	0	0	40	0	0	44	91
J5	0	1	0	0	0	29	1	31	94
J6	0	0	0	1	4	31	0	36	86
J7	0	16	0	4	1	3	5	29	55
J8	0	3	0	10	0	0	32	45	71
J9	0	0	0	43	0	0	1	44	98
J10	0	0	34	1	0	1	0	36	94
J11	0	3	0	2	39	0	0	44	89
J12	0	34	1	1	0	1	1	38	89
J13	0	35	0	1	0	1	0	37	95
J14	0	43	0	2	0	0	0	45	96
J15	0	44	0	1	0	0	0	45	98
J16	0	0	0	1	39	2	0	42	93
J17	1	0	0	0	41	2	0	44	93
J18	22	3	0	1	0	0	0	26	85
J19	0	38	0	1	0	0	1	40	95
J20	0	34	0	0	0	0	0	34	100
J21	31	13	0	1	0	0	0	45	69
J22	0	43	0	1	0	0	0	44	98
J23	1	25	0	16	0	0	1	43	58
J24	0	42	0	0	1	1	0	44	95
J25	10	16	0	18	0	1	0	45	40
J26	0	38	0	4	1	0	0	43	88
J27	0	3	0	0	39	0	0	42	93
J28	0	1	0	0	12	27	1	41	66
J29	0	1	0	0	0	41	0	42	98
Total	65	445	35	119	220	217	83	1184	—

show that the saturation index and the concentrations of Ca^{2+} and Mg^{2+} increase gradually along the regional groundwater flow path from Zone I (recharge area) to Zone IV (discharge area), indicating that carbonate dissolution occurs along the flow path.

It should be noted that the five geochemical reactions discussed above may be only a portion of water-rock interactions that can explain the ion ratios. A full list of geochemical reactions of water-rock interactions (including those related to the release of As and denitrification) is unavailable, due to the lack of mineralogical and biogeochemical data of the aquifer sediments. If the data were to become available, a more quantitative analysis (e.g., geochemical modeling using software such as PHREEQC) should be performed to better understand the water-rock interactions. This and collecting more information of the aquifer mineralogy are warranted in a future study.

5.2.2. Anthropogenic impacts

The phrase of anthropogenic impacts is a general term, and one specific focus of this study is on the pollution caused by extremely high concentrations of Na^+ , Cl^- , and SO_4^{2-} at several wells belonging to Clusters C3 and C4. Table 2 shows that 34 out of the 35 groundwater samples of Cluster C3 are from well J10, whose Na^+ and Cl^- concentrations are 20–100 times higher than those of other wells (Fig. 3(e) and (f)). Since well J10 is located at one of the largest pesticide factories in China, the high concentrations of Na^+ and Cl^- may be due to wastewater infiltration from the factory to the aquifer. Groundwater contamination and other environmental problems caused by the factory have been an environmental concern to the public (Ministry of Environmental Protection of the People's Republic of China, 2011). Elevated SO_4^{2-} concentration is a geochemical characteristic of Cluster C4 (Fig. 3(g)), and Table 2 shows that 43 samples of Cluster C4 are from well J9 located in a cotton mill and 18 samples from well J25 located in a water supply plant. Niu et al. (2017) stated that industrial activities

and potential leachates from the industrial areas probably accounted for high SO_4^{2-} concentrations, and a detailed field investigation is warranted in a future study to understand the exact reasons for the elevated SO_4^{2-} concentrations.

Another specific focus of this study is on nitrate concentrations that are always closely related to anthropogenic activities such as fertilizer use in agricultural lands, the major land use in the study site. To determine whether nitrate is a controlling factor of groundwater geochemistry, the cluster and principal component analyses were conducted twice for two different datasets without and with nitrate concentrations. The results discussed above are based on the dataset without nitrate concentrations. Nitrate was excluded from the analysis for the following two reasons: (1) nitrate concentrations were not continuously measured (among the 1,184 groundwater samples, nitrate concentrations were not measured for 117 samples, about 10% of the samples), and (2) nitrate concentrations measured for 582 samples were censored, i.e., measured concentrations being lower than detection limit of 0.01 mg/L. To include nitrate concentrations into the cluster and principal component analyses, the censored nitrate concentrations were replaced by 75% of the detection limit, resulting in a smaller dataset of 1,067 groundwater samples and 12 groundwater geochemical parameters (the original 11 ones plus nitrate).

For the dataset with nitrate, seven clusters were also classified based on the dendrogram plotted in Fig. S3. The seven clusters are denoted as CS1–CS7, and Table S1 lists the number of samples of each cluster for the individual monitoring wells. Clusters CS1–CS7 are equivalent to Clusters C1–C7 in terms of the wells included in the clusters. For example, Table 2 shows that wells J18 and J21 belong to Cluster C1, and the same two wells belong to Cluster CS1 as shown in Table S1. Comparing Tables 2 and S1 for all the clusters indicates that including nitrate concentrations in the cluster analysis did not change the spatial distribution of the clusters. This is further confirmed by comparing the snapshots of Clusters CS1–CS7 plotted in Figs. S4 and S5 with the snapshots of Clusters C1–C7 plotted in Figs. 6 and S2. Despite of small variation, the spatial distribution of the clusters does not change after including nitrate concentration in the cluster analysis. Therefore, the analysis below is still based on the seven clusters classified by using the 1,184 water samples with 11 groundwater geochemistry parameters with nitrate excluded.

The anthropogenic impacts discussed in this section are relevant to several findings of Niu et al. (2017). For example, Niu et al. (2017) found that industrial wastewater is responsible to increased Cl^- and SO_4^{2-} concentrations at several wells, and this is used in this study to understand why Cluster C3 is characterized with high Cl^- concentrations and Cluster C4 with high SO_4^{2-} concentrations. However, since our study differs from Niu et al. (2017) in terms of their purposes and scopes, certain findings of the two studies are not directly comparable. For example, Niu et al. (2017) found that NO_3^- concentrations has been increasing due to extensive fertilizer use, but we found that NO_3^- concentrations are not important to cluster classification because measured concentrations are low which may be due to denitrification given the reducing environment and high concentration of organic carbon in the aquifer as explained in Section 2. Therefore, we did not intensively compare the results of our study with those of Niu et al. (2017).

6. Four zones of groundwater geochemistry

Based on the highest frequency data listed in Table 2 and the understanding of the controlling factors of groundwater geochemistry discussed above, four groundwater geochemical zones were delineated and shown in Fig. 8. The figure also plots spatial distribution of sediment outcrops at land surface, which was also used for delineating the four zones. The delineation of the four zones is a major contribution of this study for understanding groundwater geochemistry at the study site. The zone delineation started with the wells belonging to the same

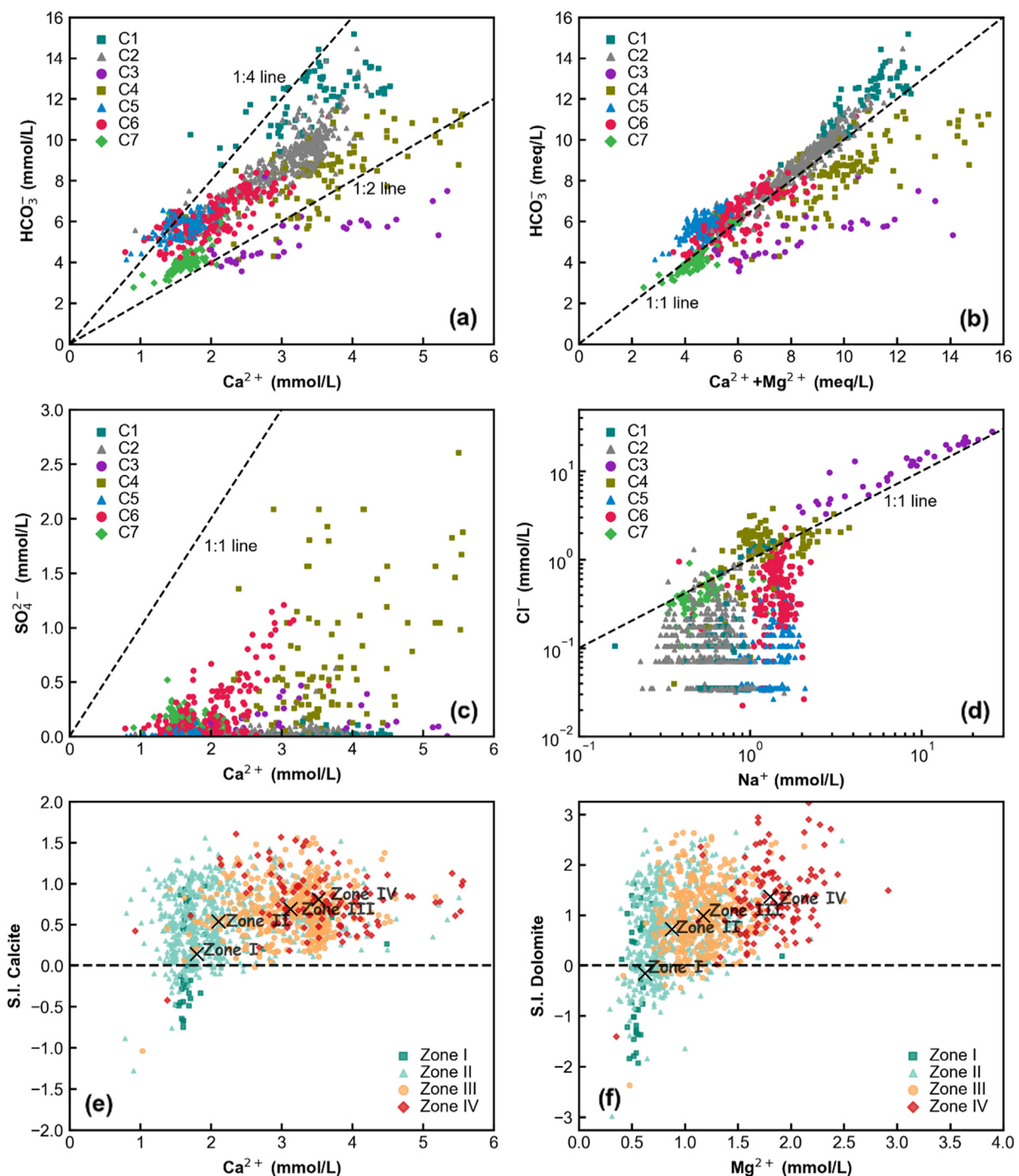


Fig. 7. Binary plot of (a) Ca^{2+} versus HCO_3^- concentrations, (b) $(\text{Ca}^{2+} + \text{Mg}^{2+})$ versus HCO_3^- concentrations, (c) Ca^{2+} versus SO_4^{2-} concentrations, (d) Na^+ versus Cl^- concentrations, (e) saturation index (S.I.) of calcite vs. Ca^{2+} concentrations, and (f) S.I. of dolomite vs. Mg^{2+} concentrations. In plots (a)–(d), data of different clusters are in different colors. In plots (e) and (f), data of different zones are in different colors, and the black crosses represent the mean values for each zone.

cluster. For example, Table 2 indicates that wells J18 and J21 belong to the same groundwater geochemistry zone, because 22 out of the 26 samples at well J18 belong to Cluster C1 and 31 out of 45 samples at well J21 belong to Cluster C1. Several clusters were combined into one groundwater geochemistry zones based on the hydrogeochemical analysis explained in detail below.

Zone I is the recharge zone of the aquifer, and the groundwater samples of this zone are from well J1 of Cluster C7 (Table 2). As shown in Fig. 8, the aquifer sediments (Q_3 and Q_2) are exposed at land surface, and an unconfined condition prevails in Zone I. Therefore, precipitation can directly recharge the aquifer, and groundwater geochemistry in this

zone is mainly of recharge water with weak water-rock interactions. As a result, groundwater geochemistry of Zone I is characterized by lowest concentrations of Ca^{2+} , Mg^{2+} and HCO_3^- (Fig. 3). This is also consistent with the results of principal component analysis (Fig. 5), as the groundwater samples of Cluster C7 have the lowest scores of Component 1, i.e., the water-rock interactions component. The boundary of Zone I was determined by the geological boundary between Q_2/Q_3 and Q_h sediments where the hydrogeological conditions may change from unconfined to confined conditions.

Zone II is the transition zone between the recharge zone (Zone I) and the flow-through zone (Zone III). Groundwater samples from this

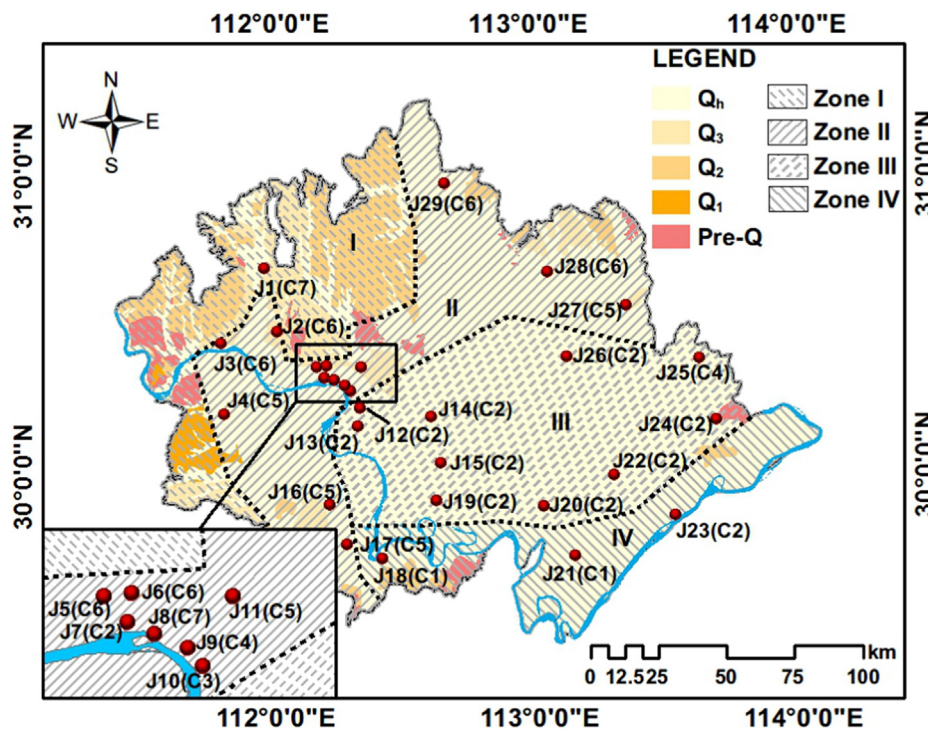


Fig. 8. Delineation of four zones of groundwater geochemistry. The figure background is spatial distribution of outcrops of Holocene (Q_h), Late Pleistocene (Q_3), Middle Pleistocene (Q_2), Early Pleistocene (Q_1), and Pre-Quaternary (Pre-Q) sediments at land surface (modified from Yang et al., 2017). Zone I is a recharge zone, Zone II is a transition zone, Zone III is a flow-through zone, and Zone IV is a discharge-mixing zone. The dotted lines are the approximate boundaries of the four zones. The labels are for the well index and the clusters to which the wells belong according to Table 2. For example J21(C1) means that well J21 belongs to C1 wells.

zone mainly belong to Clusters C5 and C6 (wells J4, J11, J16, J17, and J27 are C5 wells, and wells J2, J3, J5, J6, J28, and J29 are C6 wells). After precipitation infiltrates into the aquifer at the mountain area (Zone I) and flows downgradient to the flat area (Zone III) of the Jiangnan Plain, groundwater continuously reacts with the aquifer sediments. This explains the similarity of the ion concentrations between Clusters C5 and C6 (Fig. 2). Precipitation still affects groundwater geochemistry in the transition zone, because the depth of aquifer sediments is shallow (generally 5–10 m) and precipitation directly recharges groundwater due to discontinuous aquitard and semi-confined conditions in Zone II. Except the wells of Clusters C5 and C6, other wells located in Zone II were not used for the delineation of Zone II. For example, although well J10 of Cluster C3 and well J9 of Cluster C4 are located in Zone II, their groundwater geochemistry is affected mainly by anthropogenic activities, as explained above.

Zone III is the flow-through zone before groundwater is discharged out of the study site. Groundwater samples of this zone are all from C2 wells, except well J25 of Cluster C4. Although well J25 is identified as a C4 well, it is possible that this well belongs to Cluster C2 if the well is not affected by anthropogenic activities, because 18 samples of well J25 belong to Cluster C4 and 16 samples to Cluster C2 (Table 2). This flow-through zone is confined with the deepest (generally 20–40 m) aquifer sediments. Groundwater flows slowly in this zone due to low hydraulic conductivity and small hydraulic gradient (Zhao, 2005), which allows strong interaction between groundwater and aquifer sediments. Therefore, this zone is characterized by higher concentrations of Ca^{2+} , Mg^{2+} , and HCO_3^- in comparison with their relative concentrations in Zones I and II (Fig. 3).

Zone IV is the discharge-mixing zone before groundwater discharges into the Yangtze River. Groundwater samples from this zone are from wells J18 and J21 of Cluster C1 and well J23 of Cluster C2. The groundwater geochemistry of this zone is mainly characterized by the highest concentration of Ca^{2+} , Mg^{2+} , and HCO_3^- (Fig. 3), and this can be explained by the continuous water-rock interactions along the regional flow path. Groundwater and surface water interaction also plays an important role in groundwater geochemistry, recalling that the aquifer is hydraulically connected to the Yangtze River (Fig. S1). Groundwater discharges to the river when the river stage is lower than

the aquifer head, and receives recharge from the river otherwise. Well J23 is included in this zone due to its close distance (1.4 km) to the Yangtze River.

The four zones discussed above provide a conceptual model for understanding the spatial variation of groundwater geochemistry due to water-rock interactions. Taking the TDS concentration as an example, TDS increases along the regional groundwater flow direction from Zone I to Zone IV due to continuous water-rock interactions. As shown in Fig. 3(l), when groundwater moves from the northwest to the southeast, the average TDS concentration is 593 mg/L (for the wells of Cluster C7) in Zone I, increases to 603 mg/L (for the wells of Clusters C5 and C6) in Zone II, further increases to 770 mg/L (for the wells of Cluster C2) in Zone III, and finally reaches 859 mg/L (for the wells of Cluster C1) in Zone IV. The same pattern of spatial variation is also observed for the concentrations of Ca^{2+} , Mg^{2+} , and HCO_3^- shown in Figs. 3 and 7(a) and (b). As the four zones of groundwater geochemistry were delineated for the first time, more research may be needed to further examine the zone delineation. One way to do so is to first develop a regional-scale model of groundwater reactive transport based on existing geologic and hydrologic knowledge and data, and then use the model to simulate the concentrations of the groundwater geochemistry parameters in the aquifer. Since the simulated concentrations are for the entire aquifer (not limited to monitoring well locations), they may be used to examine the spatial distribution of the four zones. However, developing the groundwater reactive transport model for the regional aquifer is beyond the scope of this study.

7. Temporal variation and controlling factors

Although the groundwater geochemistry is generally stable over the 23 years, temporal variations are observed by examining the snapshots of the spatial patterns of the clusters (Fig. 6), recalling that changes of clusters indicate changes of groundwater geochemistry. A better way of examining the cluster changes is to plot the temporal variation of cluster index for each well, as shown in Fig. 9. The seven wells were selected because of their low relative frequency (less than 80%) of belonging to one cluster (Table 2). Taking well J8 as an example, Fig. 9(a) shows that this well belongs to Cluster C7 for 32 times, to

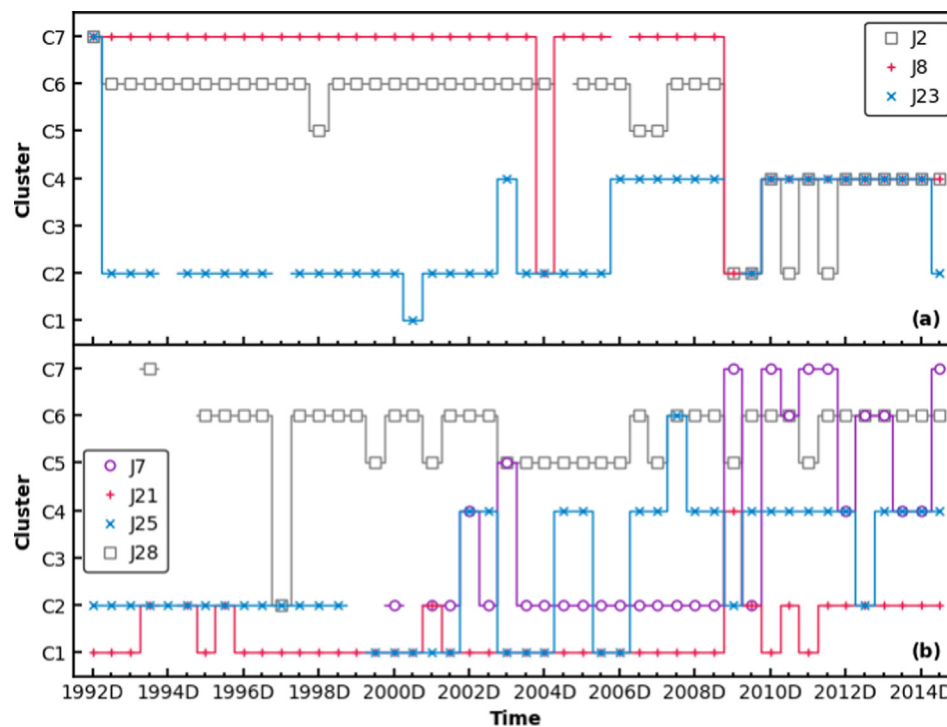


Fig. 9. Temporal changes of cluster index for (a) wells J2, J8 and J23, and (b) wells J7, J21, J25, and J28 with the highest relative frequency values less than 80% (Table 2) for the dry (D) and wet seasons of each year from 1992 to 2014.

Cluster C4 for 10 times, and to Cluster C2 for 3 times.

7.1. Temporal variation

Fig. 9 reveals both short- and long-term cluster changes. A short-term change means that a well belongs to a cluster for a short period (e.g., half a year or one year) and then changes to another cluster. A long-term cluster change means that, once a well changes to another cluster, the well stays in the cluster for a long period of time (e.g., several years). In Fig. 9, the temporal cluster changes are short term for wells J7 and J28. For J7, it occasionally belongs to Clusters C6 and C7. For J28, it stays in Cluster C6 for most of the time, but occasionally belongs to Cluster C5. This is not surprising, because the two clusters are located in the transition zone (Zone II) (Fig. 8) and have similar concentrations for the geochemical parameters (Fig. 3). The short-term changes at the two wells reflect the fluctuations of groundwater geochemistry in this zone. Short-term cluster changes are affected by multiple factors such as precipitation changes (Nguyen et al., 2015), fertilizer uses (Niu et al., 2017; Puckett, 1995), and seasonal pumping (Huang et al., 2018; Schaefer et al., 2016, 2017; Cheng et al. 2017). In addition, understanding the short-term changes requires collecting data at a short time scale (e.g., monthly). Since the monitoring data used in this study were collected at the seasonal scale, they are not adequate for explaining the short-term changes.

In Fig. 9, the temporal cluster changes are long term for wells J2, J8, J21, J23, and J25. For these five wells, J2 is located in an aquafarm, J8 about 300 m away from a cotton mill, J23 in a fertilizer plant, and J25 in a water supply plant. The four wells changed to Cluster C4 (characterized by high SO_4^{2-} concentrations) in the last several years, suggesting the long-term impacts of anthropogenic activities on groundwater geochemistry. The long-term cluster changes may reflect water type changes. Taking well J8 as an example, its water type changes from Ca-HCO_3 to $\text{Ca-HCO}_3\text{-SO}_4$ since 2010. This is consistent with the cluster change shown in Fig. 9(a), in that, since 2010, well J8 changes to Cluster C4 that is characterized with high SO_4^{2-} concentration (Fig. 3). However, since the anthropogenic activities are not

documented, detailed study of their impacts is difficult, if not impossible. The long-term change for well J21 is attributed to the impacts of the Three Gorges Reservoir, because the well changes from Cluster C1 to Cluster C2 after 2010 (Fig. 9(b)), when the reservoir operation started.

7.2. Impacts of Three Gorges Dam

Depending on the difference of water levels between the Yangtze River and the aquifer, the river-aquifer interactions change temporally and spatially, and the change substantially influences the recharge-discharge dynamics of the aquifer. In the upper reach of the Yangtze River upstream of the Ouchikou site (see its location in Fig. 1), the river stage is constantly higher than the groundwater level, and the aquifer receives recharge from the Yangtze River even in the dry seasons. This is illustrated in Fig. S6(a) that plots the river stage at the Shashi station (see its location in Fig. 1) and the nearby well J10 in the upstream reach of the river. In the lower reach of the Yangtze River downstream of the Ouchikou, the difference of water levels between the river and the aquifer changes in dry and wet seasons. During the dry seasons (from December to March of next year), the groundwater levels are higher than the river stages, and groundwater discharges to the river. During the wet seasons (from April to November), the groundwater levels are lower than the river stages, and groundwater is recharged by the river. This is illustrated in Fig. S6(b) that plots the river stage at the Luoshan station (see its location in Fig. 1) and the nearby well J23 in the downstream reach of the river. It is expected that the recharge-discharge dynamics have more substantial impacts on groundwater geochemistry in the downstream reach than in the upstream reach of the river. This is supported by the fact that there are essentially no cluster changes at upstream wells J9, J12, and J13. For wells J2, J8, J21, J23 and J25 with long-term cluster changes shown in Fig. 9, only well J21 is located near the downstream reach of the Yangtze River, and it is the focus of the analysis below. Since groundwater level at well J21 was not measured, we assume that the groundwater discharge-recharge dynamics at this well is similar to that of well J23 shown in Fig. S6(b).

Fig. 9(b) shows that well J21 mostly belonged to Cluster C1 from 1992 to 2008, but changed to Cluster C4 in the dry period of 2009, fluctuated between C1 and C2 from the wet period of 2009 to the wet period of 2011, and then remains as C2 afterward. These changes are also observed in the snapshots plotted in Fig. 6. We attribute these changes to the construction and operation of the Three Gorges Dam. The operation history of the Three Gorges Dam during 2003–2014 can be separated into four stages, i.e., initial stage, transitional stage, quasi-normal stage, and normal stage (Deng et al., 2016; Tang et al., 2016; Yang et al., 2014). The water level in the reservoir fluctuated seasonally, but increased gradually from 135 m to 175 m from June 2003 to October 2010 (175 m being the normal water level of the reservoir) (Fig. S7 of the supplementary file). Deng et al. (2016) reported that, while the reservoir construction and operation did not have remarkable impacts on water cycle of the Yangtze River for the first two stages, the river stages were significantly increased in the latter two stages in the dry seasons. According to Wang et al. (2013), the river stage increased about 0.33 m in the upper reach and 0.38 m in the lower reach of the Yangtze River after 2009. Considering that the Yangtze River penetrates into the aquifer and that the river stage is higher than the confining layer (Fig. S1), it is reasonable to expect that, during the quasi-normal and normal stages of the reservoir, the aquifer receives more recharge from the Yangtze River due to the river stage increase (He and Tang, 2017). In particular, the start of the quasi-normal stage of the reservoir in October 2008 appears to correspond to the change of well J21 from Cluster C1 to other clusters starting from the dry period of 2009. The normal stage of the reservoir starting from October 2010 appears to correspond to the well's permanent change from Cluster C1 to Cluster C2.

To quantify the mixing of groundwater and river water, the mixing fraction of river water was calculated by using a two-end member mixing model (Crandall et al., 1999),

$$f_{rw} = (C_{gw} - C_m) / (C_{gw} - C_{rw}) \quad (6)$$

where f_{rw} is the mixing factor of river water, i.e., the portion of river water in the mixed groundwater and river water, and C_m , C_{rw} , and C_{gw} denote the concentrations of a geochemical parameter in the mixed water, river water, and groundwater, respectively. The mixing factor of river water was calculated for Cl^- , Ca^{2+} , Mg^{2+} , and HCO_3^- . As a conservative tracer, Cl^- is commonly used for determining water mixing (Crandall et al., 1999; Kirchner et al., 2010); Ca^{2+} , Mg^{2+} , and HCO_3^- were chosen because they are the major ions produced by the water-rock interactions. The groundwater concentrations of the four parameters were set as the mean concentrations of the groundwater geochemistry measurements at well J21 before 2009. The corresponding river water concentrations were based on the data of Sample CJ23 list in Table 1 of Li et al. (2014), who studied water quality of the Yangtze River. This sample was selected because its sample location is close to well J21. It should be noted that, since the river sample was collected in the wet season of 2013, the calculated mixing factor of river water may not accurately reflect the temporal variation of the factor.

Fig. 10 shows the calculated mixing factors of river water for Cl^- , Ca^{2+} , Mg^{2+} , and HCO_3^- during the period of 2010–2014. Generally speaking, except that of Mg^{2+} in 2014, the mixing factors of the four geochemical parameters are similar, suggesting a consistent effect of river recharge on the groundwater geochemical parameters. In the dry periods of 2010 and 2011, the mixing factor is near zero. This explains why well J21 returned to Cluster C1 in the two periods (Fig. 9). Starting from the wet season of 2011, the mixing factor remains more or less stable (except the mixing factor of Mg^{2+} in 2014), and this corresponds to the normal stage of the Three Gorges Dam. It is thus concluded that the dilution of groundwater by Yangtze River water is the major reason for the temporal change of groundwater geochemistry. It is reasonable to expect that the impacts of the Three Gorges Dam on the groundwater geochemistry will continue after 2014. It is worth mentioning that the mixing factor was calculated only for well J21, not for well J18

(another C1 well), because geochemical measurements at well J18 after 2010 are not available. The deviation of the mixing fraction of Mg^{2+} after 2013 may be caused by the relatively large changes of Mg^{2+} concentration and/or the approximate values used for C_{rw} and C_{gw} . An improved understanding of the deviation may be gained in a future study.

Niu et al. (2017) also reported the impacts of the Three Gorges Reservoir on the groundwater geochemistry in the Jiangnan Plain. However, their rationale is different from ours, in that their conclusion is based on the trend analysis for the increase of pH, increase of NO_3^- -N concentrations, and decrease of NH_4^+ -N concentrations. In addition, their reported starting time of the impacts is 2003, whereas ours is 2009. Another difference is that their conclusion is based on the spatial mean values (i.e., the mean over all the 29 monitoring wells), whereas our conclusion is based on the data of well J21 near the Yangtze River. Using the data of individual wells may be more reasonable, because it is unlikely that the Three Gorges Reservoir affects all the 29 monitoring wells, especially those far from the Yangtze River. The impacts of the Three Gorges Reservoir on the groundwater geochemistry in the Jiangnan Plain is sophisticated, and to fully understand the impacts may require installing monitoring wells along a transect away from the river, conducting hydrodynamics modeling and/or stable isotope analysis, which however is beyond the scope of this study.

8. Comparison with conventional cluster analysis

The cluster analysis method used in this study enabled us to simultaneously identify the spatial and temporal patterns and controlling factors of the groundwater geochemistry. The simultaneous identification may not be achieved by using conventional methods of cluster analysis. This is illustrated by comparing the results of our cluster analysis with those obtained by using a conventional cluster analysis that uses temporal averages to identify spatial patterns. For each groundwater geochemistry parameter, its temporal average was calculated over the entire monitoring period, resulting in a data matrix of 29 rows and 11 columns corresponding to 29 wells and 11 groundwater geochemistry parameters. For this data matrix, the same procedure of hierarchical cluster analysis was performed, and the 29 wells were also classified into seven clusters (the phenon line was drawn at linkage distance of 5), denoted as Cluster G1–G7, based on the dendrogram plotted in Fig. S8.

The last snapshot of Fig. 6 plots the well locations of Clusters G1–G7. While this snapshot cannot be compared with any snapshot of Clusters C1–C7 shown in Figs. 6 and S2 for individual sampling times, roughly speaking, the spatial pattern shown in this snapshot is similar to that of Clusters C1–C7 shown in Figs. 6 and S2. This is not surprising because the spatial pattern is stable, as discussed in Section 5.1. However, the cluster analysis based on temporal averages cannot reveal temporal changes for the wells that belong to different clusters due to changes of groundwater geochemistry. Taking well J23 as an example, in the cluster analysis based on temporal averages, this well belongs to Cluster G4 characterized by high SO_4^{2-} concentrations in comparison with other G-clusters. In our cluster analysis, this well was classified into C2 for 2 times and into C4 for 16 times (Table 2). The snapshots of Clusters C1–C7 in Fig. 6 show that this well was classified into C4 since 2006. The reason is that SO_4^{2-} concentration changed dramatically after 2006. The mean SO_4^{2-} concentrations for the period of 1992–2005 is 3.76 mg/L, but increased dramatically to 104.40 mg/L for the period of 2006–2014. The change of SO_4^{2-} concentration over time was ignored in the cluster analysis based on temporal mean, but captured in our cluster analysis. This problem may become severe if groundwater geochemistry changes more frequent than in our study.

9. Conclusions

A recently developed cluster analysis method of Pacheco-Castro

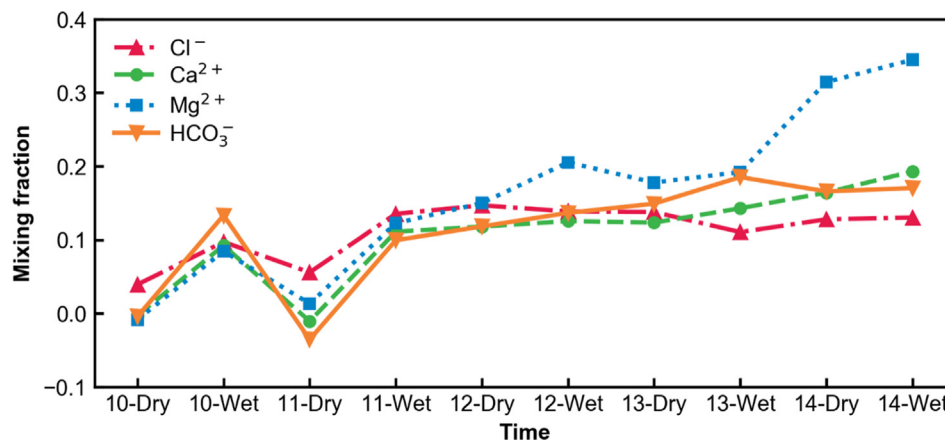


Fig. 10. Mixing fraction of Yangtze River calculated using the concentrations of Cl^- , Ca^{2+} , Mg^{2+} , and HCO_3^- of well J21 for dry (D) and wet (W) seasons of each year from 2010 to 2014.

et al. (2018) was used in this study for understanding the spatial and temporal patterns and controlling factors of groundwater geochemistry in the regional aquifer of the Jiangnan Plain, China. This cluster analysis method can handle the large dataset with 13,024 measurements of 1,184 groundwater samples collected over a long period of time (1992–2014) from 29 monitoring wells distributed over the Jiangnan Plain. The major conclusions related to understanding the regional-scale groundwater geochemistry in the confined aquifer of Jiangnan Plain are as follows:

- (1) The cluster analysis method of Pacheco-Castro et al. (2018) can be applied to a large dataset (e.g., the one used in this study) for investigating both spatial and temporal patterns of groundwater geochemistry. This is done by first classifying monitoring data of groundwater geochemistry into clusters and then examining spatial and temporal variations of the clusters to understand controlling factors of groundwater geochemistry based on hydrogeochemical analysis. The temporal changes of groundwater geochemistry identified by using the cluster analysis method of Pacheco-Castro et al. (2018) cannot be revealed by the conventional cluster analysis based on temporal averages.
- (2) In this study, the hierarchical cluster analysis yielded seven clusters, C1–C7, for the 1,184 groundwater samples. The principal component analysis of the 1,184 samples and hydrogeochemical analyses (through Stiff and Piper diagrams) indicate that the classified seven clusters are statistically and geochemically reasonable.
- (3) Based on spatial distribution of the seven clusters and hydrogeochemical analysis (mainly water-rock interactions) for the seven clusters, Jiangnan Plain was separated into four geochemical zones along the regional groundwater flow path, which is a major contribution of this study. For the water-rock interactions, the dissolution of calcite and dolomite is the most dominant controlling factor of groundwater geochemistry, followed by the albite hydrolysis.
- (4) Anthropogenic activities caused high concentrations of Cl^- , Na^+ , and SO_4^{2-} due to industrial production in a pesticides factory, a cotton mill, and a water supply plant. Nitrate concentrations are not an important factor to spatial patterns of groundwater geochemistry. The cluster analysis revealed the temporal variation of the seven clusters controlled by the Three Gorges Reservoir since 2009. The interactions between groundwater and the Yangtze River are confirmed by the calculated mixing fraction of the Yangtze River.

The findings above for understanding the spatial patterns of groundwater geochemistry at the regional scale are useful for water resource management such as designing future groundwater quality

monitoring networks. However, it should be noted that variations of groundwater geochemistry at small scales and their controlling factors are more complex, and more intensive groundwater samples are necessary for a detailed study in the future. The impacts of the Three Gorges Reservoir on groundwater geochemistry is of particular use, since the reservoir is the largest hydropower plant in the world and the Yangtze River is vulnerable to contamination. Further study of the impacts is warranted by other means such as calculating the mixing fraction using stable isotope data, which are currently unavailable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by China Geological Survey grants 1212011121142 and 1212011120084, and National Natural Science Foundation of China grant 51629901. The first author was supported by the Fundamental Research Funds for National Universities, China University of Geosciences (Wuhan) for his research at the Florida State University. The second author was supported by National Science Foundation grant EAR-1828827. We thank Barbara Mahler and two anonymous reviewers for their thoughtful comments that substantially improve the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jhydrol.2020.124594>.

References

- Appelo, C.A.J., Postma, D., 2005. *Geochemistry, Groundwater and Pollution*. Taylor and Francis, Great Britain.
- Cheng, Z., Zhang, Y., Su, C., Chen, Z., 2017. Chemical and isotopic response to intensive groundwater abstraction and its implications on aquifer sustainability in Shijiazhuang, China. *J. Earth Sci.* 28 (3), 523–534.
- Cloutier, V., Lefebvre, R., Therrien, R., Savard, M.M., 2008. Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. *J. Hydrol.* 353 (3–4), 294–313.
- Crandall, C.A., Katz, B.G., Hirtin, J.J., 1999. Hydrochemical evidence for mixing of river water and groundwater during high-flow conditions, lower Suwannee River basin, Florida, USA. *Hydrogeol. J.* 7 (5), 454–467.
- Davis, J.C., 1990. *Statistics and Data Analysis in Geology*. John Wiley & Sons Inc, New York.
- Deng, K., Yang, S., Lian, E., Li, C., Yang, C., Wei, H., 2016. Three Gorges Dam alters the Changjiang (Yangtze) river water cycle in the dry seasons: evidence from H-O

- isotopes. *Sci. Total Environ.* 562, 89–97.
- Du, Y., Ma, T., Deng, Y., Shen, S., Lu, Z., 2017. Sources and fate of high levels of ammonium in surface water and shallow groundwater of the Jiangnan Plain, Central China. *Environ. Sci. Process Impacts* 19 (2), 161–172.
- Duan, Y., 2016. *Seasonal variations of groundwater arsenic concentration in shallow aquifers at Jiangnan Plain*, (Doctoral dissertation). China University of Geosciences, Wuhan, China.
- Duan, Y., Gan, Y., Wang, Y., Deng, Y., Guo, X., Dong, C., 2015. Temporal variation of groundwater level and arsenic concentration at Jiangnan Plain, central China. *J. Geochem. Explor.* 149, 106–119.
- Duan, Y., Gan, Y., Wang, Y., Liu, C., Yu, K., Deng, Y., Zhao, K., Dong, C., 2017. Arsenic speciation in aquifer sediment under varying groundwater regime and redox conditions at Jiangnan Plain of Central China. *Sci. Total Environ.* 607, 992–1000.
- Fendorf, S., Michael, H.A., van Geen, A., 2010. Spatial and temporal variations of groundwater arsenic in south and southeast Asia. *Science* 328 (5982), 1123–1127.
- Feyereisen, G.W., Lowrance, R., Strickland, T.C., Sheridan, J.M., Hubbard, R.K., Bosch, D.D., 2007. Long-term water chemistry database, Little River Experimental Watershed, southeast Coastal Plain United States. *Water Resour. Res.* 43 (9).
- Gan, Y., Wang, Y., Duan, Y., Deng, Y., Guo, X., Ding, X., 2014. Hydrogeochemistry and arsenic contamination of groundwater in the Jiangnan Plain, central China. *J. Geochem. Explor.* 138, 81–93.
- Gan, Y., Zhao, K., Deng, Y., Liang, X., Ma, T., Wang, Y., 2018. Groundwater flow and hydrogeochemical evolution in the Jiangnan Plain, central China. *Hydrogeol. J.* 26 (5), 1609–1623.
- Ghesquière, O., Walter, J., Chesnaux, R., Rouleau, A., 2015. Scenarios of groundwater chemical evolution in a region of the Canadian Shield based on multivariate statistical analysis. *J. Hydrol.: Reg. Stud.* 4, 246–266.
- Gorelick, S.M., Zheng, C., 2015. Global change and the groundwater management challenge. *Water Resour. Res.* 51 (5), 3031–3051.
- Güler, C., Thyne, G.D., 2004a. Hydrologic and geologic factors controlling surface and groundwater chemistry in Indian Wells-Owens Valley area, southeastern California, USA. *J. Hydrol.* 285 (1–4), 177–198.
- Güler, C., Thyne, G.D., 2004b. Delineation of hydrochemical facies distribution in a regional groundwater system by means of fuzzy c-means clustering. *Water Resour. Res.* 40 (12).
- Güler, C., Thyne, G.D., McCray, J.E., Turner, K.A., 2002. Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeol. J.* 10 (4), 455–474.
- Guo, L.C., Su, N., Zhu, C.Y., He, Q., 2018. How have the river discharges and sediment loads changed in the Changjiang River basin downstream of the Three Gorges Dam? *J. Hydrol.* 560, 259–274.
- Han, D., Currell, M.J., Cao, G., 2016. Deep challenges for China's war on water pollution. *Environ. Pollut.* 218, 1222–1233.
- He, P., Tang, Z., 2017. Influence of the Three Gorges Reservoir operation on hydrological process and riparian groundwater dynamics in Jiangnan Plain section of the Yangtze River. *Saf. Environ. Eng.* 24 (6), 47–55 (In Chinese).
- Huang, K., Liu, Y.Y., Yang, C., Duan, Y.H., Yang, X.F., Liu, C.X., 2018. Identification of hydrobiogeochemical processes controlling seasonal variations in arsenic concentrations within a riverbank aquifer at Jiangnan Plain, China. *Water Resour. Res.* 54 (7), 4294–4308.
- Hussain, M., Ahmed, S.M., Abderrahman, W., 2008. Cluster analysis and quality assessment of logged water at an irrigation project, eastern Saudi Arabia. *J. Environ. Manage.* 86 (1), 297–307.
- Jessen, S., Postma, D., Thorling, L., Müller, S., Leskelä, J., Engesgaard, P., 2017. Decadal variations in groundwater quality: a legacy from nitrate leaching and denitrification by pyrite in a sandy aquifer. *Water Resour. Res.* 53 (1), 184–198.
- Jones, E., Oliphant, T., Peterson, P., 2001. In: *SciPy: Open Source Scientific Tools for Python*. <http://www.scipy.org/>.
- Kaiser, H.F., 1960. The application of electronic computers to factor analysis. *Educ. Psychol. Measur.* 20 (1), 141–151.
- Kim, J., Yum, B., Kim, R., Koh, D., Cheong, T., Lee, J., Chang, H., 2003. Application of cluster analysis for the hydrogeochemical factors of saline groundwater in Kimje, Korea. *Geosci. J.* 7 (4), 313–322.
- Kirchner, J.W., Tetzlaff, D., Soulsby, C., 2010. Comparing chloride and water isotopes as hydrological tracers in two Scottish catchments. *Hydrol. Process.* 24 (12), 1631–1645.
- Landon, M.K., Green, C.T., Belitz, K., Singleton, M.J., Esser, B.K., 2011. Relations of hydrogeologic factors, groundwater reduction-oxidation conditions, and temporal and spatial distributions of nitrate, Central-Eastside San Joaquin Valley, California, USA. *Hydrogeol. J.* 19 (6), 1203–1224.
- Li, X., Liu, Y., Zhou, A., Zhang, B., 2014. Sulfur and oxygen isotope compositions of dissolved sulfate in the Yangtze River during high water period and its sulfate source tracing. *Earth Sci. – J. China Univ. Geosci.* 39 (11), 1647–1654 (In Chinese).
- Liu, C., Lin, K., Kuo, Y., 2003. Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. *Sci. Total Environ.* 313 (1–3), 77–89.
- Ministry of Environmental Protection of the People's Republic of China, 1994. *National Quality Standard for Ground Water (GB/T 14848-93)* (in Chinese).
- Ministry of Environmental Protection of the People's Republic of China, 2011. *National Groundwater Pollution Prevention and Control Plan* (in Chinese).
- Nguyen, T.T., Kawamura, A., Tong, T.N., Nakagawa, N., Amaguchi, H., Gilbuena, R., 2015. Clustering spatio-seasonal hydrogeochemical data using self-organizing maps for groundwater quality assessment in the Red River Delta, Vietnam. *J. Hydrol.* 522, 661–673.
- Nilsson, C., Reidy, C.A., Dynesius, M., Revenga, C., 2005. Fragmentation and flow regulation of the world's large river systems. *Science* 308 (5720), 405–408.
- Niu, B., Wang, H., Loaiciga, H.A., Hong, S., Shao, W., 2017. Temporal variations of groundwater quality in the Western Jiangnan Plain, China. *Sci. Total Environ.* 578, 542–550.
- Pacheco-Castro, R., Pacheco Avila, J., Ye, M., Cabrera Sansores, A., 2018. Groundwater quality: analysis of its temporal and spatial variability in a karst aquifer. *Groundwater* 56 (1), 62–72.
- Pant, R.R., Zhang, F., Rehman, F.U., Wang, G., Ye, M., Zeng, C., Tang, H., 2018. Spatiotemporal variations of hydrogeochemistry and its controlling factors in the Gandaki River Basin, Central Himalaya Nepal. *Sci. Total Environ.* 622, 770–782.
- Puckett, L.J., 1995. Identifying the major sources of nutrient water pollution. *Environ. Sci. Technol.* 29, 408A–414A.
- Qian, Y., Migliaccio, K.W., Wan, Y., Li, Y., 2007. Surface water quality evaluation using multivariate methods and a new water quality index in the Indian River Lagoon, Florida. *Water Resour. Res.* 43 (8).
- Rajmohan, N., Elango, L., 2003. Identification and evolution of hydrogeochemical processes in the groundwater environment in an area of the Palar and Cheyyar River Basins, Southern India. *Environ. Geol.* 46 (1), 47–61.
- Reghunath, R., Murthy, T.R.S., Raghavan, B.R., 2002. The utility of multivariate statistical techniques in hydrogeochemical studies: an example from Karnataka, India. *Water Res.* 36 (10), 2437–2442.
- Rina, K., Datta, P.S., Singh, C.K., Mukherjee, S., 2012. Characterization and evaluation of processes governing the groundwater quality in parts of the Sabarmati basin, Gujarat using hydrochemistry integrated with GIS. *Hydrol. Process.* 26 (10), 1538–1551.
- Sanford, R.F., Pierson, C.T., Crovelli, R.A., 1993. An objective replacement method for censored geochemical data. *Math. Geol.* 25 (1), 59–80.
- Sayemuzzaman, M., Ye, M., Zhang, F., Zhu, M., 2018. Multivariate statistical and trend analyses of surface water quality in the central Indian River Lagoon area, Florida. *Environ. Earth Sci.* 77 (4).
- Schaefer, M.V., Guo, X., Gan, Y., Benner, S.G., Griffin, A.M., Gorski, C.A., Wang, Y., Fendorf, S., 2017. Redox controls on arsenic enrichment and release from aquifer sediments in central Yangtze River Basin. *Geochim. Cosmochim. Acta* 204, 104–119.
- Schaefer, M.V., Ying, S.C., Benner, S.G., Yanhua, D., Yanxin, W., Fendorf, S., 2016. aquifer arsenic cycling induced by seasonal hydrologic changes within the Yangtze River Basin. *Environ. Sci. Technol.* 50 (7), 3521–3529.
- Shrestha, S., Kazama, F., 2007. Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan. *Environ. Modell. Soft.* 22 (4), 464–475.
- Simeonov, V., Stratis, J.A., Samara, C., Zachariadis, G., Voutsas, D., Anthemidis, A., Sofoniou, M., Kouimtzi, T., 2003. Assessment of the surface water quality in Northern Greece. *Water Res.* 37 (17), 4119–4124.
- Tóth, J., 1999. Groundwater as a geologic agent: an overview of the causes, processes, and manifestations. *Hydrogeol. J.* 7 (1), 1–14.
- Tóth, J., 2009. *Gravitational Systems of Groundwater Flow: Theory, Evaluation, Utilization*. Cambridge University Press, Cambridge.
- Tang, Q., Bao, Y., He, X., Fu, B., Collins, A.L., Zhang, X., 2016. Flow regulation manipulates contemporary seasonal sedimentary dynamics in the reservoir fluctuation zone of the Three Gorges Reservoir, China. *Sci. Total Environ.* 548, 410–420.
- Thyne, G., Güler, C., Poeter, E., 2004. Sequential analysis of hydrochemical data for watershed characterization. *Groundwater* 42 (5), 711–723.
- VanTrump, G., Miesch, A.T., 1977. The U.S. geological survey RASS-STATPAC system for management and statistical reduction of geochemical data. *Comput. Geosci.* 3 (3), 475–488.
- Wang, G., 2009. *Analysis on climatic characteristics of temperature and precipitation in Hubei Province in recent 47 years*, (master's thesis). Lanzhou University, Lanzhou, China (in Chinese).
- Wang, H., Jiang, X., Wan, L., Han, G., Guo, H., 2015. Hydrogeochemical characterization of groundwater flow systems in the discharge area of a river basin. *J. Hydrol.* 527, 433–441.
- Wang, J., Sheng, Y., Gleason, C.J., Wada, Y., 2013. Downstream Yangtze River levels impacted by Three Gorges Dam. *Environ. Res. Lett.* 8 (4), 044012.
- Wang, Y., Guo, Q., Su, C., Ma, T., 2006. Strontium isotope characterization and major ion geochemistry of karst water flow Shentou, northern China. *J. Hydrol.* 328 (3–4), 592–603.
- Ward, J.H., 1963. Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* 58 (301), 236–244.
- Yang, J., Tang, Z., Jiao, T., Malik Muhammad, A., 2017. Combining AHP and genetic algorithms approaches to modify DRASTIC model to assess groundwater vulnerability: a case study from Jiangnan Plain, China. *Environ. Earth Sci.* 76 (12).
- Yang, J., Xiao, T., Li, H., Wang, Q., 2018. Spatial distribution and influencing factors of the NO₃-N concentration in groundwater in Jiangnan Plain, China. *Environ. Sci.* 38 (2), 710–717 (in Chinese).
- Yang, S., Milliman, J.D., Xu, K., Deng, B., Zhang, X., Luo, X., 2014. Downstream sedimentary and geomorphic impacts of the Three Gorges Dam on the Yangtze River. *Earth Sci. Rev.* 138, 469–486.
- Yu, H., Ma, T., Deng, Y., Du, Y., Shen, S., 2017. Hydrogeochemical characteristics of shallow groundwater in eastern Jiangnan Plain. *Earth Sci. – J. China Univ. Geosci.* 42 (5), 685–692 (in Chinese).
- Zhao, D., 2005. *The three-dimensional numerical simulation for groundwater system in Jiangnan Plain*, (master's thesis). China University of Geosciences, Wuhan, China (in Chinese).
- Zhou, J., Zhang, M., Lu, P., 2013. The effect of dams on phosphorus in the middle and lower Yangtze river. *Water Resour. Res.* 49 (6), 3659–3669.
- Zhou, Y., Wang, Y., Li, Y., Zwahlen, F., Boillat, J., 2012. Hydrogeochemical characteristics of central Jiangnan Plain, China. *Environ. Earth Sci.* 68 (3), 765–778.