Journal of Geochemical Exploration 141 (2014) 6–14

Contents lists available at ScienceDirect

Journal of Geochemical Exploration

journal homepage: www.elsevier.com/locate/jgeoexp

Mapping of Fe mineralization-associated geochemical signatures using logratio transformed stream sediment geochemical data in eastern Tianshan, China

Wenlei Wang [a,b], Jie Zhao [b,c,*], Qiuming Cheng [b,c]

[a] MLR Key Laboratory of Metallogeny and Mineral Assessment, Institute of Mineral Resources, Chinese Academy of Geological Sciences, Beijing 100037, China
[b] Department of Earth and Space Science and Engineering, Department of Geography, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada
[c] State Key Lab of Geological Processes and Mineral Resources, China University of Geosciences, Beijing 100083, China

ARTICLE INFO

Article history:
Received 1 July 2013
Accepted 21 November 2013
Available online 1 December 2013

Keywords:
Desert coverage
Compositional data
Closure effect
Multivariate outlier detection
Principal component analysis

ABSTRACT

Geochemical data used in this paper to identify geochemical signatures associated with Fe mineralization in eastern Tianshan mineral district, China are compositional data. In order to eliminate the spurious relationships between compositions, isometric logratio (ilr) transformation is currently employed to deal with the closure effects. The opened geochemical data are further analyzed by principal component analysis. By back-transforming to the clr space, PCA results of ilr transformed data with geological meanings of their counterparts assist in recognition of spatial distributions of both intermediate-felsic igneous rocks and fault systems in the study area. In comparison with log-transformed geochemical data, the ilr transformed ones are theoretically reasonable to apply standard statistics. In addition, multivariate outliers are identified to investigate their associations with the formation of these two geological features. At the end, this paper suggests that the ilr transformation is necessary to be applied routinely before applying statistical treatments to raw geochemical data.

1. Introduction

Eastern Tianshan mineral district is a famous polymetallic mineralized zone in China (Han and Zhao, 2003; Mao et al., 2005; Zhang et al., 2008). Caused by a prolonged history of tectonic evolution, magmatism (e.g., volcanic eruption, magmatic emplacement, and hydrothermal activities) was widely spread across the space, which benefited metallic mineralization in this area (Cui et al., 2008). Among various types of mineralization, volcanic-sedimentary Fe deposits with great profits have greatly interested geologists in China. Fe deposits hosted by volcanic strata are mostly located in the Jueluotag Carboniferous volcanic belt where the volcanic edifices are distributed in high density. Most Fe deposits were hydrothermally altered by the Late Carboniferous-–Middle Permian intermediate-felsic intrusions (Ding, 1990; Han et al., 2002). Furthermore, indicated by existing literature, the volcanic strata were genetically and spatially dominated by regional fault systems (BGEDXP, 2009). Therefore, better understandings to spatial distributions of both intermediate-felsic igneous rocks and fault systems will benefit future Fe exploration in eastern Tianshan district, China.

With development of computer sciences in past decades, exploratory data sets were frequently employed to recognize geo-anomalies (Cheng, 2007) or outliers in a statistical sense (Filzmoser et al., 2005) for mapping of mineral potential. As an important source of geo-information, geochemical data recording multiple element concentration have been successfully processed by advanced multivariate analytical methods (e.g., factor analysis, cluster analysis, etc.) to identify mineralization-associated geological bodies and delineate mineralization-favored spaces (Bogoch et al., 1993; Brantley and White, 2009; Cheng, 2007; Cheng et al., 2011; Hao et al., 2007; Wang et al., 2011, 2012, 2013; Zhao et al., 2012, 2013, 2014). On the other hand, outlier detection as one of the main tasks in multivariate data analysis has been implemented in many case studies since outliers with different distributions to the regular observations may indicate atypical phenomena in space (Barceló et al., 1996; Filzmoser and Hron, 2008; Filzmoser et al., 2005, 2009b, 2012; Reimann et al., 2005; Rose et al., 1979). According to the detection, relations between secondary geo-processes (e.g., mineralization, sedimentation, volcanism, etc.) and the origin of outliers can be investigated and further benefit mineral exploration.

Commonly used exploratory data sets (e.g., mineral contents of igneous rocks, elemental concentration of whole-rock geochemical samples, rock types of sediment samples, etc.) are often compositional data which have long been of concern in the geological field (Aitchison, 1982, 1984; Buccianti, 2011; Buccianti et al., 2006; Carranza, 2011; Chayes and Trochimczyk, 1978; Rollinson, 1992). The closure effect frequently results in misleading conclusions, especially when standard statistical treatments are applied to those data sets. Standing for relative information of different parts to a whole, compositional data can be always represented as summing to a unit constant sum constraint (e.g., 1 for the case if an observed physical quantity is in parts per unit, 100 for the case if the physical quantity is in percentage, etc.) (Filzmoser and Hron, 2009; Pawlowsky-Glahn and Buccianti, 2011; Pawlowsky-Glahn

* Corresponding author. Tel.: +86 13520709895.
E-mail address: zhaojie.ca@gmail.com (J. Zhao).

0375-6742/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.gexplo.2013.11.008

and Egozcue, 2006). Unlike unconstrained variables whose values are free to vary from $-\infty$ to $+\infty$ in Euclidean space independently, compositional data (e.g., exploratory data by geological observations) lie in the simplex (Aitchison, 1986) and are often with positive values ranging from 0 to 1 (or a constant sum) (Pawlowsky-Glahn and Egozcue, 2006). Restricted by the force of a constant sum, geo-information carried by compositions is trading off each other. Taking one geochemical sample as an example, an increase in the proportion of one element will cause a decrease of the other(s) to some degree. As a result, correlation coefficients of compositions are not free to vary from -1 to $+1$ independently, and there must be at least one negative correlation coefficient; furthermore, correlation coefficients are negatively biased (Pawlowsky-Glahn and Egozcue, 2006; Thomas and Aitchison, 2006). Consequently, correlations among these compositions are often spurious, misleading and/or meaningless in a statistical sense (Rollinson, 1993). Standard statistical methods (e.g., principal component analysis) employed to examine relations among open variables might be inappropriate for the analysis of untransformed compositional data, since these methods are mostly designed for Euclidean space (i.e., open systems) (Aitchison, 1983).

In practice, logratio transformations are commonly employed in geochemical data processing to open closed systems for better understandings of realistic relationships among compositions (Carranza, 2011; Egozcue et al., 2003; Filzmoser et al., 2012; Gallo and Buccianti, 2013; Verma et al., 2006). Logratio transformations process compositional data by two treatments: defining ratios of compositional parts and taking logarithm on the ratios. The former is to decompose the closure effect by selecting proper divisors, while the latter is to make the transformed compositional data lognormally distributed (Aitchison, 1982, 1986; Filzmoser et al., 2009a; Zhou, 1998). In general, three main logratio transformations are frequently applied to compositional data: (1) additive logratio (alr) transformation (Aitchison, 1982, 1983, 1986); (2) centered logratio (clr) transformation (Aitchison, 1982, 1983, 1986); and (3) isometric logratio (ilr) transformations (Egozcue et al., 2003), advantages and drawbacks of which will be reviewed in later sections.

The current research demonstrates a geochemical exploration model which employs the logratio transformed stream sediment geochemical data to investigate geochemical signatures associated with Fe mineralization in eastern Tianshan mineral district, China. The modeling process consists of spatial delineation and outlier detection. First, principal component analysis (PCA) is applied to ilr transformed geochemical data for mapping of spatial distributions of mineralization-associated intermediate-felsic igneous rocks and fault systems. Second, outliers resulted from tectonic–magmatic activities in the study area are recognized by a multivariate outlier detection method, which are beneficial to identify the geochemical signatures of these two geological features.

2. Geological background

Eastern Tianshan mineral district is located in the eastern Xinjiang province, China, and situated in the junction zone of the Siberian, Kazakhstan–Junggar, and Tarim plates. As a part of Jueluotag mineralization zone (Han and Zhao, 2003; Hou et al., 2006), it is bound in the north by the Turpan–Hami basin (part of the Junggar plate), in the south by the Aqikekuduke–Shaquanzi fault zone, in the west by the Xiaorequanzi area, and in the east by the Late Paleozoic Beishan rift (Mao et al., 2005) (Fig. 1).

Influenced by the Early Paleozoic subduction of the Junggar crust (in the north) under the Tarim crust (in the south) (Ma et al., 1993; Wang et al., 1994; Zhang et al., 2005), fault systems are well developed in this area. Three E–W trending faults including the Kanggurtag–Huangshan, Yamansu and Aqikekuduke–Shaquanzi from the north to the south make up the tectonic framework of the study area. They are the south boundaries of the Kanggurtag–Harlik (A), the Qiugemingtashi–Huangshan (B), and the Aqishan–Yamansu subareas (C), respectively (Li et al., 2002; Mao et al., 2005; Yang et al., 1996). Stratigraphically, the Qiugemingtashi–Huangshan subarea clamped in the middle is a series of disordered strata which were highly deformed and metamorphosed. On the contrary, strata in the Kanggurtag–Harlik and the Aqishan–Yamansu subareas are ordered volcanic–sedimentary strata deposited from the Middle Ordovician to Upper Carboniferous and the Lower Carboniferous to Middle Permian, respectively (Li et al., 2002; Mao et al., 2002, 2005; Wang et al., 2006; Zhang et al., 2004; Zhao et al., 2012).

Tectonism dominated the magmatic activities and consequent Fe mineralization in eastern Tianshan district (BGEDXP, 2009). E–W trending faults restricting the extent of stratigraphical subareas confine occurrences of magmatic activities at a regional scale (e.g., volcanism, sedimentation, igneous intrusion, etc.) (Mao et al., 2002; Qin et al., 2002; Wang et al., 1994; Xia et al., 2005). Furthermore, intersections of faults striking along N-E and N-W orientations controlled the local allocation of volcanic edifices (BGEDXP, 2009; Ma et al., 1997; Qin et al., 2002). Located in volcanic arc, the Yamansu Formation (C_1y) consisting of terrigenous clastics, carbonate rocks, and mafic–felsic bimodal volcanic rocks in the Aqishan–Yamansu area is the main host strata of Fe deposits. (BGEDXP, 2009; Han et al., 2002; Jiang et al., 2002; Ma et al., 1997; Wang et al., 2006; Yang et al., 1996). Mechanically, ore materials migrated along magma towards ground surface and precipitated in lower basins after volcanic eruption (Han et al., 2002). As a result, preliminary Fe ore bodies were formed around fault-controlled volcanic edifices during flow, cooling, and solidification of magma. As an important characteristic, most of these ore bodies were further concentrated and enriched by hydrothermal fluids differentiated from post-mineralization granitoid intrusions (Chen, 1999; Ma et al., 1993, 1997; Wang, 2005; Zhang and Xie, 2001). Therefore, tectonically controlled magmatism played significant roles in Fe mineralization. Better understandings of those igneous rocks can improve metallogeny studies of the Fe deposits. Fortunately, outcrops of these igneous rocks (i.e., intrusions and extrusions) through prolonged history of surface denudation provide important clues for Fe exploration in eastern Tianshan mineral district, since the Fe ores are apt to form within contact zones of intrusions and volcanic sedimentary rocks (BGEDXP, 2009; Jiang et al., 2002; Li et al., 2002; Wang et al., 2007). Therefore, identification of both intermediate-felsic igneous rocks and fault systems are beneficial and necessary to Fe exploration.

Because the eastern Tianshan mineral district is located in the Gobi Desert area, mineral exploration in this district is greatly impeded by natural terrain features (e.g., eolian sand, caliche, and regolith). Recording geochemical signatures inherited from bedrocks, geochemical data are commonly used to identify geochemical anomalies associated with various geological bodies and to interpret geological phenomena (Bogoch et al., 1993; Brantley and White, 2009; Cheng, 2007; Hao et al., 2007; Wang et al., 2011, 2012; Zhao et al., 2012, 2013). The stream sediment geochemical data currently used to identify the spatial information of geological bodies associated with Fe mineralization were collected and analyzed by Chinese National Geochemical Mapping Project as part of the "Regional Geochemistry National Reconnaissance (RGNR) Project". Samples were collected within drainage basins, in which 39 elements/compounds were mainly analyzed by means of X-ray Fluorescence (Xie et al., 1997; Zhuang et al., 2003). The concentration of all elements/compounds is smoothed by averaging all samples within each 2×2 km^2 cell. Thus, over eight thousand samples are employed in current research. Detailed information about the RGNR can be found in Xie et al. (1997).

3. Methodology

3.1. Logratio transformation

Confined by the constant sum representation, compositional data (e.g., raw geochemical data) lie in a restricted space and carry only

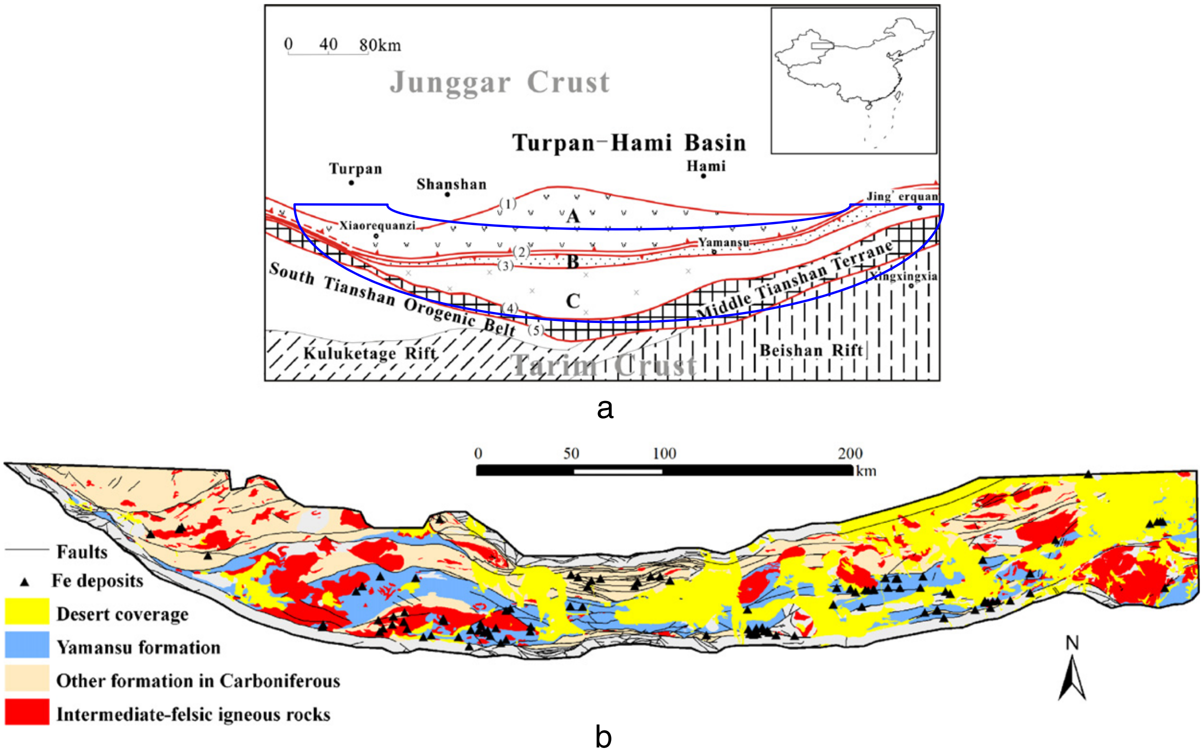

Fig. 1. Geological maps of the study area. a: The study area and its tectonic setting (modified from Mao et al., 2005). A = Kanggurtag–Harlik area. B = Qiugemingtashi–Huangshan ductile shear zone. C = Aqishan–Yamansu island arc. (1) = Dacaotan fault. (2) = Kanggurtag–Huangshan fault. (3) = Yamansu fault. (4) = Aqikekuduke–Shaquanzi fault. (5) = South margin fault of Middle Tianshan. The study area is outlined in blue. b: Geological map of the study area.

relative information (contained in ratios between parts) rather than the absolute one (Pawlowsky-Glahn and Egozcue, 2006). Being relative proportions, none of the compositional parts can vary between $(-\infty, +\infty)$ independently. For example, if SiO_2 contained in a sample of igneous rock accounts for 69% of the total weight, then the content of another constituent like MgO can only take values less than 31%. In the restricted space or so-called simplex (Aitchison, 1986), data are following Aitchison geometry where D-part compositions only contain D-1 dimensional information (Egozcue and Pawlowsky-Glahn, 2006). In other words, the correlation and/or covariance matrices of compositional data are singular (Pawlowsky-Glahn and Egozcue, 2006). Since most standard statistical methods are designed for Euclidean geometry, application of these treatments to raw geochemical data would result in misleading conclusions even though the log-transformed data are normally distributed (Filzmoser et al., 2012). Therefore, logratio transformation which can convert compositional data to Euclidean space is necessary to deal with the closure effect, and logratio transformed data can then be analyzed by unconstrained multivariate statistics appropriately (Aitchison and Egozcue, 2005; Carranza, 2011). The additive logratio (alr), the centered logratio (clr), and the isometric logratio (ilr) transformations are three main approaches to convert compositional data to an open system (Carranza, 2011; Pawlowsky-Glahn and Egozcue, 2006).

Constant sum representation of D-part compositional data $x = (x_1, \ldots, x_D)^T$ form a simplex (denoted as S^D), where the compositional parts x_i ($i = 1, 2, \ldots, D$) (e.g., 39 elements/oxides recorded in currently used stream sediment geochemical samples) are strictly positive components summing to a constant (e.g., 100%). Geochemical data as a closed system can be opened by alr (Eq. (1)), clr (Eq. (2)) and ilr (Eq. (3)) transformations, respectively (Egozcue et al., 2003; Filzmoser et al., 2009a):

$$y_i = \log\frac{x_i}{x_D} \quad (i = 1,\ 2,\ \ldots, D-1) \tag{1}$$

$$y_i = \log\frac{x_i}{\sqrt[D]{\prod_{i=1}^{D} x_i}} \quad (i = 1,\ 2,\ \ldots, D) \tag{2}$$

$$y_i = \sqrt{\frac{D-i}{D-i+1}}\log\frac{x_i}{\sqrt[D-i]{\prod_{j=i+1}^{D} x_j}} \quad (i = 1,\ 2,\ \ldots, D-1). \tag{3}$$

Some general and specific properties of these three forms of logratio transformations derived from literatures are reviewed as follow:

(1) Formulas for the alr and clr transformations are relatively simple. By alr transformation, one compositional part (i.e., denominator or divisor) is selected to divide the remaining parts (i.e., numerators) and then log-transformation is taken on the ratios (Aitchison, 1986). By clr, the transformation utilizes the geometric mean of all parts as the divisor and then log-transformation is applied to the ratios (Aitchison, 1986). Different from alr and clr, by ilr transformation the stepwise elimination of compositional parts is taken into calculation of geometric mean which performs as the divisor. It provides one-to-one conversion of compositions from the Aitchison geometry into a vector in the Euclidean real space (Filzmoser et al., 2012; Hron et al., 2010). Therefore, compared with clr and ilr transformations, results of alr transformation are subjective regarding the divisor (Filzmoser et al., 2009b).

(2) Both alr and ilr transformations reduce the number of resulting variables (i.e., from S^D to R^{D-1}). It means that one variable will be sacrificed during alr and ilr transformations (Reimann et al., 2012). On the contrary, the clr transformation preserves D real components (i.e., R^D), transformed data are adding up to zero. Consequently, clr transformed vectors are collinear and still constrained in a sub-space (Egozcue et al., 2003; Filzmoser et al., 2012; Pawlowsky-Glahn and Egozcue, 2006). However, clr transformation which treats all components symmetrically

enables an easier interpretation of single clr variables in sense of the original compositional parts (Filzmoser et al., 2009b; Reimann et al., 2012).

(3) Only ilr transformed vectors lie in orthogonal systems, and standard statistics designed for Euclidean space are consequently applicable for the ilr transformed variables (Buccianti, 2013; Carranza, 2011; Filzmoser et al., 2009b, 2010; Pawlowsky-Glahn and Egozcue, 2006). However, by non-linear functions, ilr transformed variables do no longer possess the sense of their original counterparts. The insufficiency of direct connections to original variables results difficulties in interpretation of statistical results (Filzmoser and Hron, 2009; Filzmoser et al., 2009a; Pawlowsky-Glahn and Egozcue, 2006). Attention should be paid during the interpretation of statistical results. Specifically, for PCA procedure in a current study, this problem can be calibrated by back-transforming statistical results (e.g., loadings and scores from PCA) based on ilr transformed variables to clr space (Filzmoser et al., 2009b). Through this approach, interpretable results can then be derived. A more detailed introduction regarding back-transformation of PCA results to clr space can be referred to Filzmoser et al. (2009b).

Overcoming drawbacks of subjectivity of alr transformation and collinearity caused by clr transformation, ilr transformation generating correct equivalent of compositions in real Euclidean space has been broadly practiced to decompose closed number systems (Buccianti, 2013; Carranza, 2011; Filzmoser et al., 2010, 2012). Furthermore, analysis from clr space makes the PCA results of ilr transformed compositions much more interpretable.

3.2. PCA

PCA is a classic multivariate analysis technique which has been commonly used to examine relationships among variables. By matrix transformation (i.e., orthogonal transformation), multiple related variables can be converted into uncorrelated principal components (PCs) based on a covariance or correlation coefficient matrix (Cheng et al., 2011; Horel, 1984; Jolliffe, 2002; Loughlin, 1991). Since only the first few PCs posses most of variances of input data sets which are retained for further interpretation, PCA is an efficient tool in reducing dimensionality of multi-variable data sets.

For a $n \times p$ data matrix X with p variables x_i ($i = 1, \dots, n$), PCs are frequently derived from its covariance matrix $C(X)$ (Filzmoser et al., 2005). Based on the covariance matrix, the eigenvalues and eigenvectors can be calculated:

$$\det[C(X) - \lambda I] = 0 \tag{4}$$

$$[C(X) - \lambda I]U = 0 \tag{5}$$

where, I is the $p \times p$ identity matrix, and "det" is the determinant of the matrix formed by $C(X) - \lambda I$. λ_j ($j = 1, 2, \dots, p$) calculated from the characteristic equation of $C(X)$ is the eigenvalue, and $U = [a_{j1}, a_{j2}, \dots, a_{jp}]$ is the eigenvector matrix. Each PC$_j$ can be expressed as a linear combination of the p variables (i.e., $X_1, X_2, \dots, X_p$) as:

$$\mathrm{PC}_j = a_{j1}X_1 + a_{j2}X_2 + \dots + a_{jp}X_p \tag{6}$$

where PC$_j$ is the scores of the jth PC ($j = 1, \dots, p$).

In practice, PCA has been commonly used to geochemical data analysis. According to the loadings of each geochemical variable, geo-information is interpreted in support of geological exploration.

3.3. Multivariate outlier detection

Belonging to other distributions, outliers resulting from one or more different secondary geo-processes are of primary interest of statistical analysis of geochemical data (Filzmoser et al., 2005). For a p-dimensional multivariate sample x_i ($x_1, \dots, x_n$), rather than extreme values which are defined as extremely high or low values along single coordinates, multivariate outliers (Barnett and Lewis, 1994) are detected based on Mahalanobis distance (MD):

$$MD_i = \left((x_i - t)^T C^{-1} (x_i - t)\right)^{1/2} \quad (i = 1, \ \dots, n) \tag{7}$$

where t and C are estimations of location (i.e., the multivariate arithmetic mean or centroid) and scatter (i.e., covariance matrix), respectively (Barceló et al., 1996; Filzmoser and Hron, 2008; Filzmoser et al., 2012). For the multivariate normally distributed data, the Mahalanobis distance is approximately chi-square distributed with p degrees of freedom (χ^2_p). Outliers are defined as observations with great values of MD_i (Filzmoser et al, 2005). Selection of t and C are significant to multivariate outlier detection. However, the normally used t and C mentioned above may often mislead the statistical results since these estimators are sensitive to the distribution of outliers (Filzmoser et al., 2012; Hampel et al., 1986).

A robust method, the minimum covariance determinant (MCD) estimator, proposed by Rousseeuw (1985) is one of the commonly used techniques to recognize reliable outliers, which provides affine transformations of t and C in calculation (Filzmoser et al., 2012). The Mahalanobis distance estimated from Eq. (7) by using the MCD estimator is so-called as robust distances (RDs), and outliers determined by using these RDs account for approximately 25% of total samples. However, the main drawback of this method is the difficulties in differentiating true outliers from extreme values because of using a fixed threshold value (e.g., $\chi^2_{p;0.98}$). To overcome the drawback, an adaptive threshold (i.e., adjusted quantile) was proposed by Filzmoser et al. (2005) to calculate robust Mahalanobis distances. As introduced in Filzmoser et al. (2005), $G_n(u)$ denotes the empirical distribution function of the squared robust distances RD^2, and $G(u)$ is the distribution function of χ^2_p. Therefore, multivariate outliers can be detected by comparing the tails of $G_n(u)$ and $G(u)$ (i.e., departures of the empirical from the theoretical distribution):

$$p_n(\delta) = \sup_{u \geq \delta} (G(u) - G_n(u))^+ \tag{8}$$

where $\delta = \chi^2_{p;1-\alpha}$ defines the tails; "+" means the positive differences. In addition, Filzmoser et al. (2005) suggested two approximate formulas to calculate the critical value p_{crit} for distinguishing between outliers and extremes:

$$p_{crit}(\delta, n, p) = \frac{0.24 - 0.003p}{\sqrt{n}} \quad \text{for} \ \ p \leq 10. \tag{9}$$

$$p_{crit}(\delta, n, p) = \frac{0.252 - 0.0018p}{\sqrt{n}} \quad \text{for} \ \ p > 10. \tag{10}$$

where, n is the sample size and p is the dimensions. A more detailed information of the robust and adaptive outlier detection methods can be referred to Filzmoser et al. (2005, 2012).

4. Data processing and results

Benefitting from the development of computer sciences and spatial analysis techniques, numbers of algorithms have been greatly improved and applied to geo-data sets. Ratios of geological variables have been extensively used to deal with geological issues. Their applications can be found in geochronological analysis (e.g., Rb–Sr isochron diagrams where $^{87}Sr/^{86}Sr$ is plotted against $^{87}Rb/^{86}Sr$) (Rollinson, 1993), ore genesis analysis (e.g., Co/Ni in pyrite to reveal ore types) (Bralia et al., 1979), remote sensing interpretation (e.g., the band ratio of TM bands 5/7 to map hydrothermal alteration) (Sabins, 1999), geophysical analysis (e.g., velocity ratio of V_p/V_s to examine physical properties of sedimentary rocks) (Wilkens et al., 1984), etc. The logratio transformations dealing with the closure effect of compositional exploratory data (Aitchison, 1986; Egozcue and Pawlowsky-Glahn, 2006; Filzmoser et al., 2012; Pawlowsky-Glahn and Egozcue, 2006) can be used to investigate geological issues as well. For example, the concept of balances (Egozcue and Pawlowsky-Glahn, 2005) taking care of both statistical sense and geological meaning of the ratio was successfully applied in many cases. Since balances describe relative behaviors of groups of compositional parts in an orthogonal geometry, standard statistics can be applied to the transformed variables, appropriately (Buccianti, 2013; Carranza, 2011; Filzmoser and Hron, 2009; Hron et al., 2010; Pawlowsky-Glahn and Egozcue, 2006; Tolosana-Delgado et al., 2005).

Delineations of spatial distributions of intermediate-felsic igneous rocks and fault systems are significant to Fe mineral exploration in eastern Tianshan mineral district, China (Cheng, 2012; Zhao et al., 2012, 2013). As an important source of geo-information, stream sediment geochemical data were frequently used to fulfill these objectives by geological interpreters due to their advantages in providing clues to the presence of geological bodies on/near the surface (Rose et al., 1979). However, currently employed geochemical data are typical compositional data restrained in a closed system, and spurious relationships between concentration values of 39 elements/oxides may yield misleading results if standard statistics are applied. Therefore, the ilr transformation calibrating problems of the closure effect is currently involved in pre-processing of geochemical data; after that, to recognize intermediate-felsic igneous rocks and fault systems in the study area, PCA is further employed to integrate the ilr transformed variables.

In order to depict the spatial distribution of intermediate-felsic igneous rocks, geochemical distributions of seven major rock-forming components (i.e., SiO_2, Na_2O_3, MgO, Fe_2O_3, K_2O, CaO, and Al_2O_3) are used, concentration values of which are recorded in wt.%. In addition, igneous rocks in this area can be characterized by enrichment of Ba and Be in intermediate-felsic igneous rocks and enrichment Li in mafic igneous rocks (BGEDXP, 2009). Concentration values of these three elements are recorded in ppm. In addition, being geo-pressure relief zones, fault systems are ideal places for hydrothermal fluid flow and ore materials precipitation. Fault zones are often accompanied with prominent geochemical anomalies of certain elements (Qian, 2009). In other words, these geochemical anomalies are indicative to the spatial distribution of fault systems. Different from the strategy of igneous rock identification, geochemical elements used for fault system recognition are trace elements Au, As, Hg and Sb which are recorded in ppm in geochemical data. These elements are extremely sensitive to changes of surrounding circumstances and readily dissolve into or precipitate from hydrothermal fluids. Specifically, As, Hg, and Sb performing as mineralizers can benefit Au mineralization within fault systems. In this study, geochemical anomalies of these elements are chosen as indicators to faults (BGEDXP, 2009; He and Chen, 2002; Yuan et al., 1979). Therefore, in order to recognize spatial distributions of intermediate-felsic igneous rocks and fault systems in eastern Tianshan mineral district, China, 14 geochemical variables (i.e., SiO_2, Na_2O_3, MgO, Fe_2O_3, K_2O, CaO, Al_2O_3, Ba, Be, Li, Au, As, Hg and Sb) from 8768 stream sedimentary samples (i.e., a matrix of 8768×14) are chosen to investigate objective geological features.

In this paper, two experiments are demonstrated as follows:

(1) PCA applied to 14 log-transformed geochemical variables. This experiment demonstrates a traditional treatment to geochemical data, which cannot open the closed system.
(2) PCA applied to 14 ilr transformed geochemical variables. Rather than the first case, input variables for PCA in this experiment are opened geochemical data. Since ilr transformed variables are short of direct interpretability, PCA results (e.g., loadings, scores, etc.) are back-transformed to the clr space to fulfill recognitions of the objective geological features (Filzmoser et al., 2009b; Zuo et al., 2013).

Comparing biplots of these two experiments (Fig. 2), the closure effect is distinct in the biplot of the PCA result based on log-transformed data (Fig. 2a). Most variables are plotting oppositely to felsic oxides (i.e., SiO_2, K_2O, Na_2O, and Al_2O_3). It corresponds to a primary environmental context of the study area that dominating surface features are rich of felsic oxides. Since these felsic oxides are prevalent in either Gobi-desert coverage or igneous rocks (i.e., two main lithological units in the study area), concentrations of other compositional parts in these two ground features must be greatly declined due to the closure effect. In the biplot, these variables are plotting towards negative loadings of PC1. In addition, variables (i.e., Au, As, Hg, Sb, MgO, Fe_2O_3, CaO and Li) are ungrouped in the biplot (Fig. 2a) that illustrates spurious relationships among compositional parts in the closed system. In other words, the real paragenetic characteristics of element associations

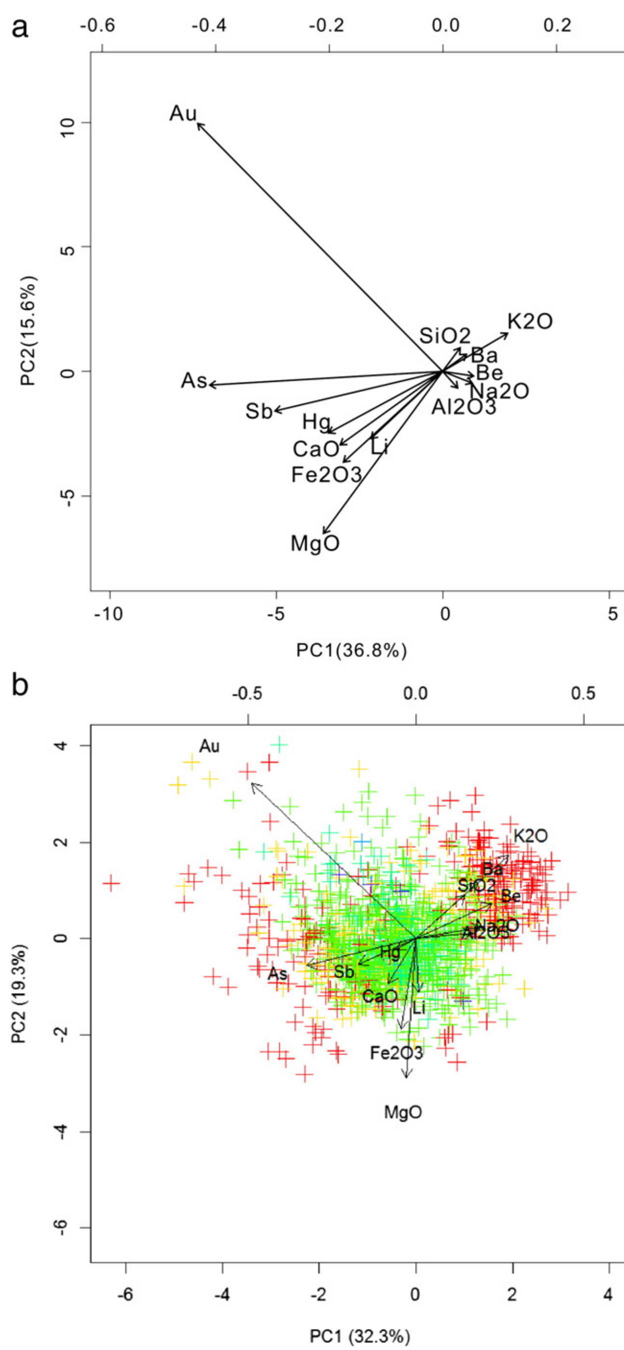

Fig. 2. Biplots of the first two PCs for a: the log-transformed and b: ilr transformed geochemical data.

indicative to specific geological processes or geochemical signatures of geological bodies cannot be reflected sufficiently.

On the contrary, the biplot of the PCA result based on ilr transformed data (Fig. 2b) demonstrates a nicely grouped plotting of variables that implies characteristics of the opened system in real space. Lying in an orthogonal space, PCA applied to ilr transformed data will be reasonable from the viewpoint of statistics. By back-transforming to clr space for a better interpretation, rays of ilr transformed variables in the biplot (Fig. 2b) are grouped and pointing towards different directions. These features indicate that the PCA result of ilr transformed data is achieved in a right geometry (Filzmoser et al., 2010). Rays of SiO_2, K_2O, Na_2O, Al_2O_3, Ba, and Be in the fourth quadrant are indicative to the most prevalent ground features (i.e., Gobi-desert and intermediate-felsic igneous rocks) in the study area. Two noticeable sub-groups in this quadrant are SiO_2-K_2O-Ba and Al_2O_3-Na_2O-Be, respectively. The former implies occupation of quartz and feldspars in either intermediate-felsic igneous rocks or sandy covers; whereas the latter implies clay minerals produced by extravagant eolation. The downward-pointed association of Fe_2O_3, MgO, CaO, and Li corresponds to geochemical signatures of mafic igneous rocks (e.g., the Yamansu Formation which is mainly composed of intermediate-mafic volcanic rocks, volcaniclastics, carbonate rocks, tuff, etc.) which are the primary hosts of volcanic-sedimentary Fe deposits (BGEDXP, 2009). The group of As, Sb, and Hg lying in the third quadrant corresponds to geochemical signatures of tectonic processes and/or their end products (i.e., fault systems). The separation of Au from this group might imply extraordinary mobility of Au in the natural environment. It may not paragenetically exist with As, Sb and Hg, although these elements are common mineralizers of Au.

From PCA score maps of these two experiments, geo-information of grouped element associations descriptive and/or indicative to intermediate-felsic igneous rocks and fault systems can be derived from Figs. 3 and 4, respectively. According to the two biplots (Fig. 2), geochemical signatures of the felsic group (i.e., SiO_2, K_2O, Na_2O, Al_2O_3, Ba, and Be) indicative to spatial distributions of intermediate-felsic igneous rocks can be represented by high PC1 scores of the log-transformed (Fig. 3a) and the ilr transformed variables (Fig. 3b). In these maps, reddish patterns are coincident with outcrops of intermediate-felsic igneous rocks and (SiO_2-rich) sandy covers in this area. On the contrary, the element association descriptive of mafic rocks and fault systems is shown in opposite directions of felsic concentrations. For better visualization, a reversed color scheme of Fig. 3 is applied (Fig. 4). The low PC1 scores of log-transformed (Fig. 4a) and the ilr transformed variables (Fig. 4b) are indicating mafic rocks and fault systems. In these maps, reddish patterns coincide with fault traces and outcrops of volcanic strata, especially the Yamansu Formation.

Comparing these two experiments from a statistical sense, only PCA result of ilr transformed geochemical data processed in right geometry can be accepted to investigate real relationships among geochemical variables. In the current study, PC1 scores of ilr transformed variables not only display patterns coincident with outcrops of objective geological features, but also follow both statistical and geological guidance.

Detection of outliers is an important step in multivariate data analysis, since the outliers with different distributions to regular observations may indicate atypical phenomena in space. In eastern Tianshan mineral district, tectono-magmatic activities as singular geo-processes (Wang et al., 2012, 2013; Zhao et al., 2012) often result in enrichment or depletion of certain elements/compounds. Consequently, outliers of geochemical data resulted from extraneous processes (Filzmoser et al., 2005) rather than background are necessary to be detected to assist in recognition of intermediate-felsic igneous rocks and fault systems in the study area. In this paper, a multivariate outlier detection technique (Filzmoser et al., 2009b, 2012) is implemented by employing the *R* package *mvoutlier* (Filzmoser et al., 2012). Multivariate outliers are detected using the adaptive approach of Filzmoser et al. (2005). Shown in the biplot (Fig. 2b), the outliers are labeled as symbol "+".

For each observation, distances to the medians of the univariate ilr variables are measured. After that, the median of all these distances determines the color of the symbol. A high value coincident with red symbol "+" indicates that most univariate parts of the corresponding observation have values greater than the average; whereas a low value coincident with the blue or green symbol "+" indicates that most univariate parts of the corresponding observation are mainly low values (Filzmoser et al., 2012). In the biplot (Fig. 2b), multivariate outliers in red color indicating high "extremeness" are clustered in both felsic and faults directions. It can be inferred that these two groups of multivariate outliers are derived from the tectono-magmatic activities. The spatial information of these outliers is probably corresponding to the distributions of end products of the geo-processes that are intermediate-felsic igneous rocks and fault systems. In addition, it is not surprising to see that the aggregation of multivariate outliers

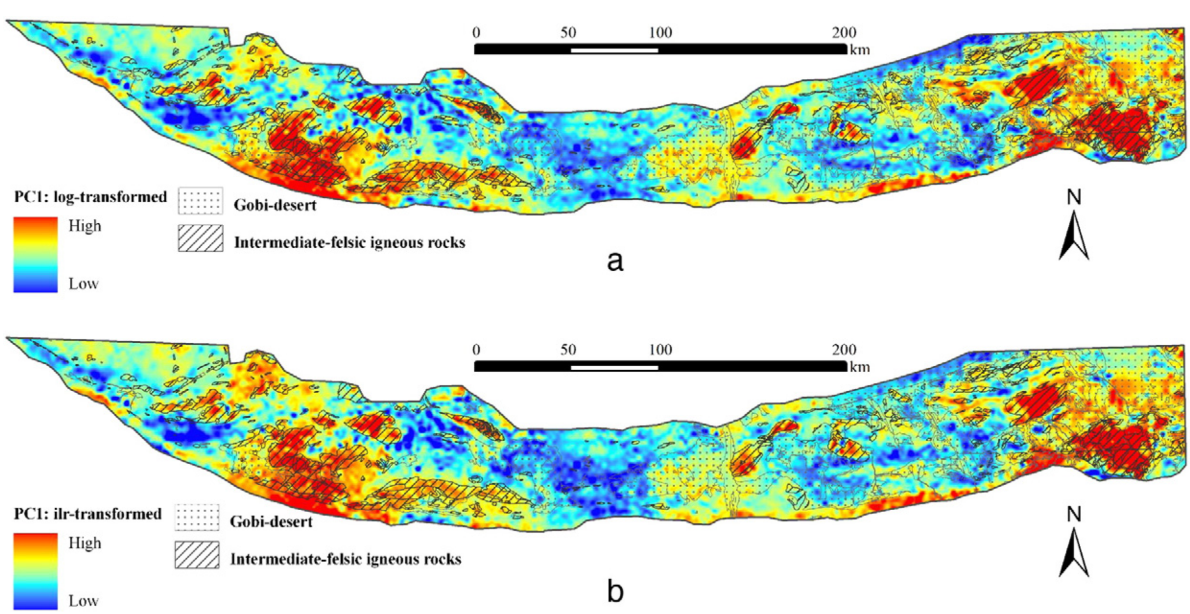

Fig. 3. PCA scores (reddish patterns) of two experiments indicating the spatial distributions of intermediate-felsic igneous rocks in eastern Tianshan mineral district, China. a: PC1 scores of log-transformed variables; b: PC1 scores based on ilr transformed variables.

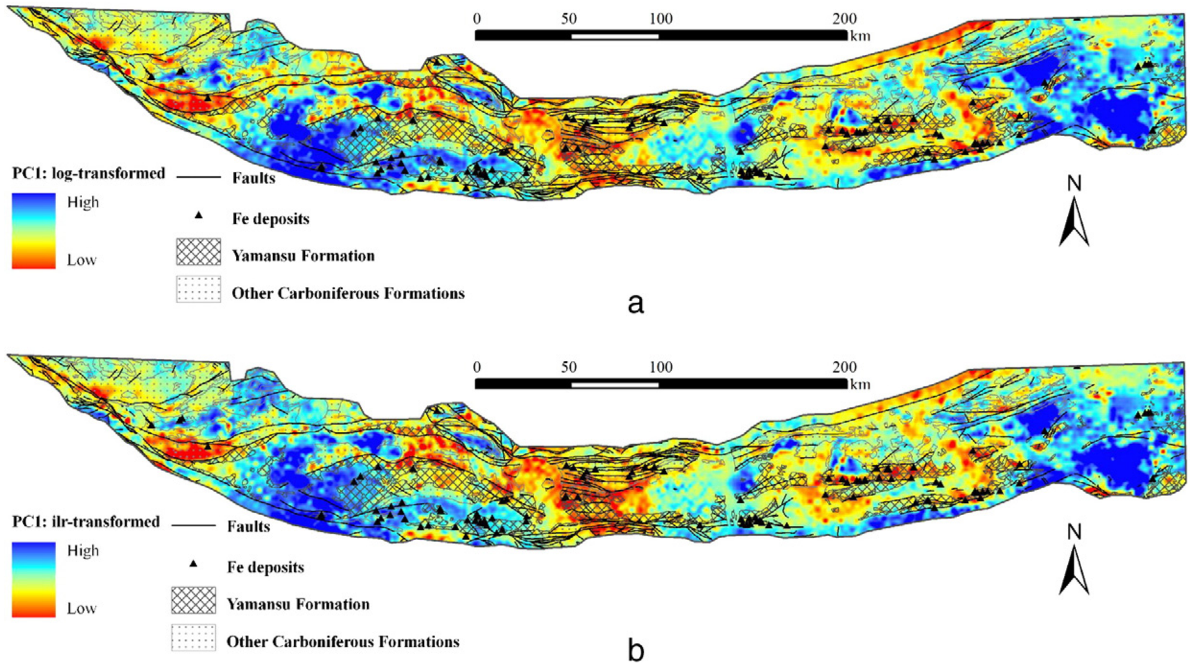

Fig. 4. A reversed color scheme of Fig. 3. PCA scores (reddish patterns) of two experiments indicating spatial distribution of fault systems and volcanic strata in eastern Tianshan mineral district, China. a: PC1 scores of log-transformed variables; b: PC1 scores based on ilr transformed variables.

along the felsic direction is much denser than the fault direction, since intermediate-felsic igneous rocks occupy more areas than fault systems (i.e., line features) do in a 2-dimensional scenario. When cross-referencing to the outliers in univariate variables (Fig. 5a), it is easy to find that outliers with high values of ilr transformed fault elements (i.e., Au, As, Hg, and Sb) are more sparsely distributed; whereas the outliers with high values of felsic elements (i.e., SiO_2, K_2O, Na_2O, Al_2O_3, Ba, and Be) are plotted densely. From the spatial distributions of these multivariate outliers (Fig. 5b), clusters of multivariate outliers in red are generally observed at the locations of intermediate-felsic igneous rocks, whereas, the sparsely distributed outliers in red are prone to appear along the fault traces.

5. Summary and discussion

This paper applies PCA to geochemical data for mapping of spatial distributions of intermediate-felsic igneous rocks and fault systems in eastern Tianshan mineral district, China. As an important source of geo-information, geochemical data have been widely practiced in mapping of mineral exploration targets and various geological features. However, special concerns to the closure effect of geochemical data were not sufficiently considered in many cases during the procedure of data processing. Specifically, constrained in simplex, the geochemical data processed by standard statistics designed for Euclidean space are theoretically inappropriate. Currently used stream sediment geochemical data as typical compositional data are restrained within a closed system; therefore, in order to explore real relationships among these geochemical variables associated with the two geological features, the ilr transformation is employed to reduce the closure effect. Two experiments of PCA demonstrated in this paper suggest that the ilr transformation generating correct statistical outputs and releasing realistic insights into the structure of the compositional data is significant to the utilization of statistical approaches to compositional data.

In addition, multivariate outliers that could be of interest to geologists are detected as well. By the adaptive approach of multivariate outlier detection, the outliers caused by the formation of geological features or associated with the intermediate-felsic igneous rocks and fault systems are distinguished from extreme values and marked in a spatial scenario. In this paper, both PCA and multivariate outlier detection can be supportive to inspect the locations or spatial distributions of these two geological features. However, the differences between these two methods are: PCA results identify spatial distributions of geological features relying on covariance or correlation of geochemical variables; whereas, the multivariate outlier detection focuses on data structure to recognize samples with different distributions (e.g., for this study area, they are caused by the tectono-magmatism) to regular observations. Based on ilr transformed geochemical data, the two objective geological features are interpreted by PCA and multivariate outlier detection, appropriately. Achieved results are not only beneficial to future Fe exploration in the study area, but also provide another meaningful geological study to the community of compositional data analysis. Therefore, ilr transformation is necessary and suggested to be used routinely when standard statistical treatments, especially the PCA and/or factor analysis are applied to raw geochemical data.

In this paper, geological guidance to recognize geological features is dependent on the selection of geochemical variables. The concept of Balances mentioned in Section 3 which can construct the coordinates with more flexibility to investigate geological features is suggested to be attempted in further studies.

Acknowledgments

Thanks are due to two anonymous reviewers, Dr. Antonella Buccianti and Dr. Eric Grunsky as the guest editors for their critical reviews and constructive comments that improved the manuscript. This research has been jointly supported by a NSERC Discovery Research Grant (ERC-OGP0183993), by a research project on "New methods for mineral resource potential assessment" (1212010733811), and by a research project on "Quantitative models for prediction of strategic mineral resources in China" (201211022) by the China Geological Survey.

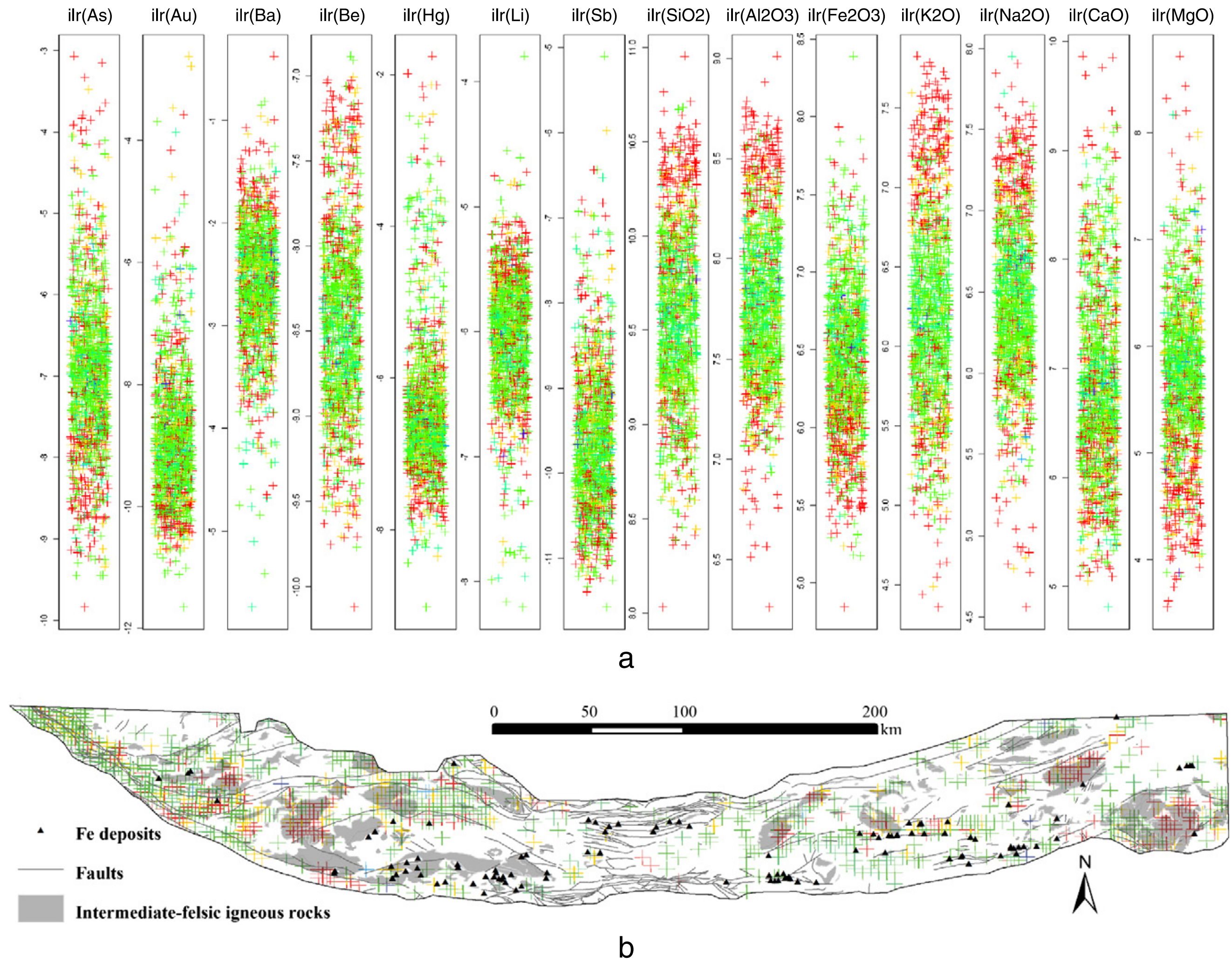

Fig. 5. Outliers detection results based on ilr transformed geochemical data. a: univariate scatterplots for 14 selected elements/compounds of the geochemical data. Only outliers are shown using the symbol color; b: a map to show spatial distributions of detected outliers with symbol color indicated.

References

Aitchison, J., 1982. The statistical analysis of compositional data. J. R. Stat. Soc. Ser. B Methodol. 44, 139–177.

Aitchison, J., 1983. Principal component analysis of compositional data. Biometrika 70, 57–65.

Aitchison, J., 1984. Reducing the dimensionality of compositional data sets. Math. Geol. 16, 617–635.

Aitchison, J., 1986. The Statistical Analysis of Compositional Data, first ed. Chapman and Hall, London.

Aitchison, J., Egozcue, J., 2005. Compositional data analysis: where are we and where should we be heading? Math. Geol. 37, 829–850.

Barceló, C., Pawlowsky, V., Grunsky, E., 1996. Some aspects of transformations of compositional data and the identification of outliers. Math. Geol. 28, 501–518.

Barnett, V., Lewis, T., 1994. Outliers in Statistical Data, third ed. John Wiley & Sons, Chichester (584 pp.).

BGEDXP, 2009. Report of mineralization exploration in Eastern Tianshan, unpublished report (in Chinese). Bureau of Geological Exploration & Development of Xinjiang Province (BGEDXP)..

Bogoch, R., Shirav, M., Beyth, M., Halicz, L., 1993. Geochemistry of ephemeral stream sediments in the Precambrian mountainous arid terrain of southern Israel. J. Geochem. Explor. 46, 349–364.

Bralia, A., Sabatini, G., Troja, F., 1979. A revaluation of the Co/Ni ratio in pyrite as geochemical tool in ore genesis problems. Mineral. Deposita 14, 353–374.

Brantley, S.L., White, A.F., 2009. Approaches to modeling weathered regolith. Rev. Mineral. Geochem. 70, 435–484.

Buccianti, A., 2011. Natural laws governing the distribution of the elements in geochemistry: the role of the log-ratio approach. In: Pawlowsky-Glahn, V., Buccianti, A. (Eds.), Compositional Data Analysis: Theory and Applications, Chapter 18, pp. 255–266.

Buccianti, A., 2013. Is compositional data analysis a way to see beyond the illusion? Comput. Geosci. 50, 165–173.

Buccianti, A., Mateu-Figueras, G., Pawlowsky-Glahn, V., 2006. Compositional data analysis in the geosciences: from theory to practice. Geol. Soc. Lond. Spec. Publ. 264, 212.

Carranza, E.J.M., 2011. Analysis and mapping of geochemical anomalies using logratio-transformed stream sediment data with censored values. J. Geochem. Explor. 110, 167–185.

Chayes, F., Trochimczyk, J., 1978. An effect of closure on the structure of principal components. Math. Geol. 10, 323–333.

Chen, Y., 1999. Geochemistry of Granitoids from the Eastern Tianshan Mountains and Northern Qinling Belt, first ed. Geological Publish House, Beijing (in Chinese).

Cheng, Q., 2007. Mapping singularities with stream sediment geochemical data for prediction of undiscovered mineral deposits in Gejiu, Yunnan Province, China. Ore Geol. Rev. 32, 314–324.

Cheng, Q., 2012. Singularity theory and methods for mapping geochemical anomalies caused by buried sources and fore predicting undiscovered mineral deposits in covered areas. J. Geochem. Explor. 122, 55–70.

Cheng, Q., Bonham-Carter, G., Wang, W., Zhang, S., Li, W., Xia, Q., 2011. A spatially weighted principal component analysis for multi-element geochemical data for mapping locations of felsic intrusions in the Gejiu mineral district of Yunnan, China. Comput. Geosci. 37, 662–669.

Cui, B., He, Z., Zhao, L., Dong, L., Liu, T., Qu, J., 2008. Research on metallogenic systems of gold and copper deposits in the middle section of Eastern Tianshan, Xinjiang, China. Earth Sci. Frontiers 15, 13–17 (in Chinese with English abstract).

Ding, T., 1990. The geological characteristics of stratabound iron deposits in the Yamansu formation in Xinjiang. Regional Geology of China 1990. 269–272 (in Chinese with English abstract).

Egozcue, J., Pawlowsky-Glahn, V., 2005. Groups of parts and their balances in compositional data analysis. Math. Geol. 37, 795–828.

Egozcue, J., Pawlowsky-Glahn, V., 2006. Simplicial geometry for compositional data. Compositional data analysis in the geosciences: from theory to practice. Geol. Soc. Lond. Spec. Publ. 264, 145–160.

Egozcue, J., Pawlowsky-Glahn, V., Mateu-Figueras, G., Barceló-Vidal, C., 2003. Isometric logratio transformations for compositional data analysis. Math. Geol. 35, 279–300.

Filzmoser, P., Hron, K., 2008. Outlier detection for compositional data using robust methods. Math. Geosci. 40, 233–248.

Filzmoser, P., Hron, K., 2009. Correlation analysis for compositional data. Math. Geosci. 41, 905–919.

Filzmoser, P., Hron, K., Reimann, C., 2005. Principal component analysis for compositional data with outliers. Environmetrics 20, 621–632.

Filzmoser, P., Hron, K., Reimann, C., 2009a. Univariate statistical analysis of environmental (compositional) data: problems and possibilities. Sci. Total Environ. 407, 6100–6108.

Filzmoser, P., Hron, K., Reimann, C., 2009b. Principal component analysis for compositional data with outliers. Environmetrics 20, 621–632.

Filzmoser, P., Hron, K., Reimann, C., 2010. The bivariate statistical analysis of environmental (compositional) data. Sci. Total Environ. 408, 4230–4238.

Filzmoser, P., Hron, K., Reimann, C., 2012. Interpretation of multivariate outliers for compositional data. Comput. Geosci. 39, 77–85.

Gallo, M., Buccianti, A., 2013. Weighted principal component analysis for compositional data: application example for the water chemistry of the Arno river (Tuscany, central Italy). Environmetrics 24, 269–277.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J., Stahel, W., 1986. Robust Statistics: The Approach Based on Influence Functions, first ed. Wiley, New York (311 pp.).

Han, C., Zhao, G., 2003. Major types and characteristics of Late Paleozoic ore deposit, East Tianshan, Northwest China. Int. Geol. Rev. 45, 798–813.

Han, C., Mao, J., Yang, J., Wang, Z., Cui, B., 2002. Types of Late Palaeozoic endogenetic metal deposits and related geodynamical evolution in the East Tianshan. Acta Geol. Sin. 76, 222–233 (in Chinese with English abstract).

Hao, L., Lu, J., Li, L., Mo, G., Yan, G., Shi, Y., Zhao, Y., 2007. Method of using regional geochemical data in geological mapping in shallow overburden areas. Geol. China 34, 710–715 (in Chinese with English abstract).

He, S., Chen, Z., 2002. The zoning of surface tectonic geochemistry in Guizhou and its significance. Guizhou Geology 19, 148–155 (in Chinese with English abstract).

Horel, J., 1984. Complex principal component analysis: theory and examples. J. Clim. Appl. Meteorol. 23, 1660–1673.

Hou, G., Tang, H., Liu, C., 2006. Geochemical characteristics of the Late Paleozoic volcanics in Jueluotage tectonic belt, eastern Tianshan and its implications. Acta Petrol. Sin. 22, 1167–1177 (in Chinese with English abstract).

Hron, K., Templ, M., Filzmoser, P., 2010. Imputation of missing values for compositional data using classical and robust methods. Comput. Stat. Data Anal. 54, 3095–3107.

Jiang, F., Qin, K., Fang, T., Wang, S., 2002. Types, geological characteristics, metallogenic regularity and exploration target of iron deposits in eastern Tianshan Mountains. Xinjiang Geol. 20, 379–384 (in Chinese with English abstract).

Jolliffe, I.T., 2002. Principal Component Analysis, second ed. Springer, New York.

Li, W., Ren, B., Yang, X., Li, Y., Chen, Q., 2002. The intermediate-acid intrusive magmatism and its geodynamic significance in Eastern Tianshan region. Northwest. Geol. 35, 41–64 (in Chinese with English abstract).

Loughlin, W., 1991. Principal component analysis for alteration mapping. Photogramm. Eng. Remote Sens. 57, 1163–1169.

Ma, R., Wang, C., Ye, S., 1993. Tectonic Framework and Crustal Evolution of Eastern Tianshan Mountains, first ed. Nanjing University Press, Nanjing (in Chinese).

Ma, R., Shu, L., Sun, J., 1997. The Tectonic Deformation, Evolution and Metallization in the Eastern Tianshan Belt, Northwest China, first ed. Geological Publish House, Beijing (in Chinese).

Mao, J., Yang, J., Han, C., Wang, Z., 2002. Metallogenic systems of polymetallic copper and gold deposits and related metallogenic geodynamic model in Eastern Tianshan, Xinjiang. Earth Science-Journal of China University of Geosciences, 27 413–424 (in Chinese with English abstract).

Mao, J., Goldfarb, R.J., Wang, Y., Hart, C.J., Wang, Z., Yang, J., 2005. Late Paleozoic base and precious metal deposits, East Tianshan, Xinjiang, China: characteristics and geodynamic setting. Episodes: Journal of International Geoscience, 18 23–36.

Pawlowsky-Glahn, V., Buccianti, A., 2011. Compositional Data Analysis: Theory and Applications, first ed. John Wiley & Sons, Chichester (378 pp.).

Pawlowsky-Glahn, V., Egozcue, J., 2006. Compositional data and their analysis: an introduction. Compositional data analysis in the geosciences: from theory to practice. Geol. Soc. Lond. Spec. Publ. 264, 1–10.

Qian, J., 2009. Tectono-geochemical prospecting method and its application in searching for sediment-hosted, disseminated gold deposits. Geol. Explor. 45, 60–67 (in Chinese with English abstract).

Qin, K., Fang, T., Wang, S., Zhu, B., Feng, Y., Yu, H., Xiu, Q., 2002. Plate tectonics division, evolution and metallogenic settings in eastern Tianshan mountains, NW-China. Xinjiang Geol. 20, 302–308 (in Chinese with English abstract).

Reimann, C., Filzmoser, P., Garrett, R.G., 2005. Background and threshold: critical comparison of methods of determination. Sci. Total Environ. 346, 1–16.

Reimann, C., Filzmoser, P., Fabian, K., Hron, K., Birke, M., Demetriades, A., Dinelli, E., Ladenberger, A., the GEMAS Project Team, 2012. The concept of compositional data analysis in practice—total major element concentrations in agricultural and grazing land soils of Europe. Sci. Total Environ. 426, 196–210.

Rollinson, H., 1992. Another look at the constant sum problem in geochemistry. Mineral. Mag. 56, 469–475.

Rollinson, H., 1993. Using Geochemical Data: Evaluation, Presentation, Interpretation, first ed. Longman, London.

Rose, A.W., Hawkes, H.W., Webb, J.S., 1979. Geochemistry in Mineral Exploration, second ed. Academic, London.

Rousseeuw, P.J., 1985. Multivariate estimation with high breakdown point. In: Grossmann, W., Pflug, G., Vincze, I., Wertz, W. (Eds.), Mathematical Statistics and Applications, vol. B, pp. 283–297.

Sabins, F., 1999. Remote sensing for mineral exploration. Ore Geol. Rev. 14, 157–183.

Thomas, C.W., Aitchison, J., 2006. Log-ratios and geochemical discrimination of Scottish Dalradian limestones: a case study. Compositional data analysis in the geosciences: from theory to practice. Geol. Soc. Lond. Spec. Publ. 264, 25–41.

Tolosana-Delgado, R., Otero, N., Pawlowsky-Glahn, V., Soler, A., 2005. Latent compositional factors in the Llobregat river basin (Spain) hydrogeochemistry. Math. Geol. 37, 681–702.

Verma, S., Guevara, M., Agrawal, S., 2006. Discriminating four tectonic settings: five new geochemical diagrams for basic and ultrabasic volcanic rocks based on log-ratio transformation of major-element data. J. Earth Syst. Sci. 115, 485–528.

Wang, X., 2005. Analysis of the geology and genesis of Yamansu Fe deposit. Contrib. Geol. Miner. Resour. Res. 20, 125–128 (in Chinese with English abstract).

Wang, C., Ma, R., Shu, L., Zhu, W., 1994. Study on the regional metamorphism and the tectonic settings in the eastern Tianshan orogenic belt. Journal of Nanjing University (Natural Science Edition), 30 494–503 (in Chinese with English abstract).

Wang, J., Wang, Y., He, Z., 2006. Ore deposits as a guide to the tectonic evolution in the East Tianshan Mountains, NW China. Geol. China 33, 461–469 (in Chinese with English abstract).

Wang, H., Liu, T., Wang, Q., Zheng, Q., 2007. Geochemical character in ore-forming belt of Tianshan–Beishan mineralization zone, Xinjiang. J. Earth Sci. Environ. 29, 141–144 (in Chinese with English abstract).

Wang, W., Zhao, J., Cheng, Q., 2011. Analysis and integration of geo-information to identify granitic intrusions as exploration targets in southeastern Yunnan District, China. Comp. Geosci. 37, 1946–1957.

Wang, W., Zhao, J., Cheng, Q., Liu, J., 2012. Tectonic-geochemical exploration modeling for characterizing geo-anomalies in southeastern Yunnan district, China. J. Geochem. Explor. 122, 71–80.

Wang, W., Zhao, J., Cheng, Q., 2013. Fault trace-oriented singularity mapping technique to characterize anisotropic geochemical signatures in Gejiu mineral district, China. J. Geochem. Explor. 134, 27–37.

Wilkens, R., Simmons, G., Caruso, L., 1984. The ratio V_p/V_s as a discriminant of composition for siliceous limestones. Geophysics 49, 1850–1860.

Xia, Z., Xu, X., Xia, L., Li, X., Ma, Z., Wang, L., 2005. Geochemistry of the Carboniferous–Permian post-collisional granitic rocks from Tianshan. Northwest. Geol. 38, 1–14 (in Chinese with English abstract).

Xie, X., Mu, X., Ren, T., 1997. Geochemical mapping in China. J. Geochem. Explor. 60, 99–113.

Yang, X., Tao, H., Luo, G., Ji, J., 1996. Basic features of plate tectonics in east Tianshan of China. Xinjiang Geol. 14, 221–227 (in Chinese with English abstract).

Yuan, J., Zhu, S., Zhai, Y., 1979. Mineral Deposits, first ed. Geological Publishing House, Beijing (in Chinese).

Zhang, H., Xie, L., 2001. New views on origin of Yamansu iron deposit in Xinjiang autonomous region. Journal of Changchun Institute of Technology (Natural Science Edition), 2 26–29 (in Chinese with English abstract).

Zhang, L., Xiao, W., Qin, K., Ji, J., Yang, X., 2004. Types, geological features and geodynamic significances of gold–copper deposits in the Kanggurtag metallogenic belt, eastern Tianshan, NW China. Int. J. Earth Sci. 93, 224–240.

Zhang, L., Xiao, W., Qin, K., Qu, W., Du, A., 2005. Re–Os isotopic dating of molybdenite and pyrite in the Baishan Mo–Re deposit, eastern Tianshan, NW China, and its geological significance. Miner. Deposita 39, 960–969.

Zhang, L., Qin, K., Xiao, W., 2008. Multiple mineralization events in the eastern Tianshan district, NW China: isotopic geochronology and geological significance. J. Asian Earth Sci. 32, 236–246.

Zhao, J., Wang, W., Dong, L., Yang, W., Cheng, Q., 2012. Application of geochemical anomaly identification methods in mapping of intermediate and felsic igneous rocks in eastern Tianshan, China. J. Geochem. Explor. 122, 81–89.

Zhao, J., Wang, W., Cheng, Q., 2013. Investigation of spatially non-stationary influences of tectono-magmatic processes on Fe mineralization in eastern Tianshan, China with geographically weighted regression. J. Geochem. Explor. 134, 38–50.

Zhao, J., Wang, W., Cheng, Q., 2014. Application of geographically weighted regression to identify spatially non-stationary relationships between Fe mineralization and its controlling factors in eastern Tianshan, China. Ore Geol. Rev. 57, 628–638.

Zhou, D., 1998. Geological compositional data analysis: difficulties and solutions. Earth Science-Journal of China University of Geosciences, 23 147–152 (in Chinese with English abstract).

Zhuang, D., Liu, T., Hu, J., Wang, X., 2003. The review and prospect of regional geochemical exploration in Xinjiang. Geophys. Geochem. Explor. 27, 425–427 (in Chinese with English abstract).

Zuo, R., Xia, Q., Wang, H., 2013. Compositional data analysis in the study of integrated geochemical anomalies associated with mineralization. Appl. Geochem. 28, 202–211.