Computers & Geosciences 140 (2020) 104484

Contents lists available at ScienceDirect

Computers and Geosciences

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

Recognizing multivariate geochemical anomalies for mineral exploration by combining deep learning and one-class support vector machine

Yihui Xiong, Renguang Zuo *

State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan, 430074, China

ARTICLE INFO

Keywords:
Multivariate geochemical data
Deep learning
One class support vector machine
Deep belief network

ABSTRACT

The recognition of multivariate geochemical anomalies is important for mineral exploration. Big data analytics, which involves the whole data and variables, is an alternative manner to delineate multivariate geochemical anomalies in support of machine learning algorithms due to their strong ability to capture the complex intrinsic and diverse links between geochemical characteristics and mineralization. However, this method faces the issue of data redundancy and calculation complexity, and high-dimensional problems raise great challenges for anomaly detection. This is the curse of dimensionality problem, which hinders the development of a variety of techniques for anomaly detection. In this study, a hybrid model that combines unsupervised deep belief networks (DBNs) and one-class support vector machine (OCSVM) is adopted to address the high-dimensional geochemical anomalies detection problem. In this model, the relevant features first extracted through the DBN are used as the input of the OCSVM. The decision function values of the hybrid method are employed to map the geochemical patterns related to iron mineralization. The comparative results on the performance of the hybrid model and the other three anomaly detection models (deep autoencoder model, OCSVM, and hybrid model with principal component analysis and OCSVM) in terms of the area under curve(AUC) values, suggest that the hybrid method of the DBN and OCSVM can efficiently recognize the geochemical anomalies related to iron mineralization. The DBN can extract the geochemical information, reduce the redundant features, and further enhance the scalability of the OCSVM for processing high-dimensional geochemical data. The extracted geochemical anomalies, which show a close spatial relationship with the Yanshanian intrusions, can provide significant guidance for the next round of mineral exploration.

1. Introduction

Over the past several decades, geochemical anomalies have shown an increasingly important role in the fields of both exploration (Hawkes and Webb, 1962; Carranza, 2008) and environmental geochemistry (Reimann and Garrett, 2005). The distribution of the geochemical patterns is attributable to the abundant primary and secondary geological processes, such as mineralization processes, which usually occur at different spatiotemporal scales, and interact on each other with extraordinarily complicated ways (Cheng, 2012). In this context, the exploration of the geochemical patterns, such as geochemical background and anomalies, are challenging regarding their characteristics of complexity, nonlinearity, and variability (Cheng et al., 1994, 2000; Goovaerts, 1997; Agterberg, 2001; Cheng, 2007, 2012). Thus, have attracted widespread attention in the field of exploration geochemistry (Cohen et al., 2010; Zuo and Wang, 2016; Zuo et al., 2016, 2019).

Methods for the recognition of geochemical anomalies range from traditional methods, such as mean ± 2 standard deviations (Hawkes and Webb, 1962), probability graphs (Sinclair, 1974), exploratory data analysis, e.g. histograms, boxplots (Tukey, 1977; Kürzl, 1988), and multivariate statistics (Filzmoser et al., 2005), to geostatistics (Matheron, 1962) and the fractal/multifractal models (Cheng et al., 1994, 2000; Cheng, 2007). Recently, machine learning algorithms are gaining more attention regarding geochemical anomaly detection (e.g., Chen et al., 2014; Xiong and Zuo, 2016; Moeini and Torab, 2017; Zhang et al., 2019; Zuo et al., 2019).

Owing to their ability to learn complex and high-level features from data (LeCun et al., 2015), machine learning and deep learning algorithms have shown immense potential for modelling complex and nonlinear data, such as multivariate geochemical data resulting from various complex geological evolutions (Xiong et al., 2016; Zuo et al., 2019). Some supervised and unsupervised machine learning algorithms

* Corresponding author.
E-mail address: zrguang@cug.edu.cn (R. Zuo).

https://doi.org/10.1016/j.cageo.2020.104484
Received 20 July 2019; Received in revised form 30 March 2020; Accepted 31 March 2020
Available online 8 April 2020

have been successfully introduced to model the geochemical patterns related to mineralization. For example, the supervised learning methods, including neural networks (Ziaii et al., 2009, 2012; Zhao et al., 2016), metric learning (Wang et al., 2019, Wang et al., 2019), support vector machines (Gonbadi et al., 2015), and ensemble learning methods, such as AdaBoost and random forest (Gonbadi et al., 2015), can model labeled geochemical data and differentiate the geochemical anomaly from the background. Additionally, the unsupervised anomaly detection methods, such as one-class SVM (OCSVM, Chen and Wu, 2017), continuous restricted Boltzmann machines (CRBM, Chen et al., 2014) and isolation forest (Wu and Chen, 2018), have been adopted for the detection of geochemical anomalies in mineral exploration.

The recent development of data acquisition technologies, such as satellites and sensors, have enabled great progress in the collection of big data. Such datasets often hold large numbers of samples (inputs), large varieties of classes (outputs), and high dimensionality (features) (Chen and Lin, 2014). Big data brings great opportunities and prospects for a variety of fields, including geochemical data processing (Zuo et al., 2019). In this field, all the geochemical variables are considered based on the idea of big data analytics to reveal both positive and negative geochemical anomalies, as well as their statistical correlations with known mineralization (Zuo and Xiong, 2018). However, large numbers of input features may lead to the curse of dimensionality, which refers to the condition where an exponential increase of training data with the dimensionality is needed to support a statistically sound and reliable result (Bengio and LeCun, 2007; Erfani et al., 2016). Under these circumstances, the generalization error of the shallow architectures enumerated above would increase with the number of redundant features for their difficulties in learning efficient representations (Bengio and LeCun, 2007). For example, SVMs (including both supervised SVM for classification and unsupervised OCSVM for anomaly detection) comprise a layer of kernel functions for the input together with a single processing layer. The SVMs are best suited for small datasets with several features, but they are inefficient for large-scale datasets with high-dimensional feature space as the data complexity increases exponentially with the number of features (Vempati et al., 2010; Erfani et al., 2015, 2016). Thus, the kernel function based SVMs suffer limitations in providing complex features, but they can provide robust decision boundaries with the available well-behaved feature vectors (Abe, 2005). To overcome the limitations of the OCSVM models with complex and high-dimensional data, dimensionality reduction is necessary to obtain a set of relevant and valuable variables (Bengio et al., 2005; Erfani et al., 2015, 2016). The core process of dimension reduction is to extract meaningful representations from high-dimensional data. Shallow architectures are inefficient for high-level representation extraction due to their limitations of transforming the original data into a shallow feature space with only one or two hidden layers. In contrast to shallow learning, deep learning provides architectures for automatically learning a compact high-level representation from high-dimensional data by transforming the representation at a lower level into a higher level (Erhan et al., 2009; Schmidhuber, 2015). Recently, the deep learning framework, deep autoencoder, was successfully introduced into the field of multivariate geochemical anomaly detection based on its strong ability to automatically identify efficient geochemical features (Xiong and Zuo, 2016; Moeini and Torab, 2017; Chen et al., 2019; Zhang et al., 2019; Zuo and Xiong, 2020).

This paper presents a hybrid model for geochemical anomaly recognition with high-dimensional data through a case study conducted in the southwestern Fujian District (China). This architecture is different from that of the single nonlinear method (e.g., OCSVM) for multivariate geochemical anomaly detection, it provides the advantages of both DBNs (Hinton et al., 2006) and OCSVMs. This new detection process involves three main steps. First, regarding big data analytics, instead of using a partial sample, all the available geochemical samples are included in the training data to investigate the statistical correlations to replace the causal relations among the available data. Then, an unsupervised DBN is adopted to extract the compact high-level features that are robust against irrelevant variations in the input. Finally, an OCSVM is trained on well-behaved features provided by the DBN to detect the multivariate geochemical anomalies.

2. Methods

A new unsupervised hybrid model, combining DBN and OCSVM (Fig. 1), is adopted to recognize geochemical anomalies with the benefits of both DBN and OCSVM. The DBN is first trained to learn the representative features of the geochemical data for non-linear dimensionality reduction. Then, the features learned from the original data are used to train the OCSVM to effectively separate the multivariate geochemical anomaly from the geochemical background. In the following sections, the implementation of the DBN for dimensionality reduction will be described first, and then the output of the DBN used as input of OCSVM is elaborated.

2.1. Deep belief network

Deep belief network (DBN) is a type of deep neural network involving multiple layers of latent variables with interconnections between but not within each hidden layer. The DBN can be used for both unsupervised training on a set of examples to learn representational features and perform non-linear dimensionality reduction, as well as supervised classification based on the extracted features (Hinton et al., 2006). A DBN comprises an undirected, generative energy-based restricted Boltzmann machine (RBM), which can provide a fast layer-by-layer unsupervised pre-training procedure (Fig. 2). As a variant of the Boltzmann machines, RBMs are restricted to having no connections between the units within each layer and having symmetric connections between the visible and hidden layers defined by a matrix of weights w_{ij} (Fig. 3a). This restriction establishes more superior training strategies for the training of Boltzmann machines, particularly contrastive divergence (CD) (Fig. 3b), originally proposed for training the products of expert models (Hinton, 2002). The CD algorithm can update the connection weight, w_{ij}, with the gradient descent procedure as follows:

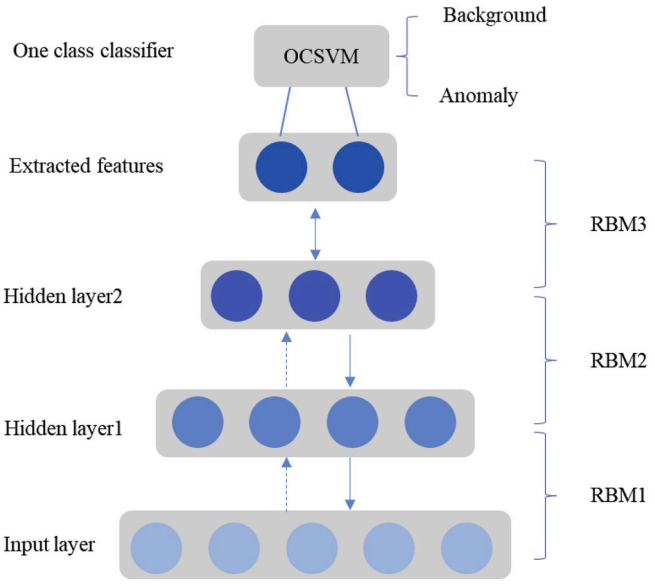

Fig. 1. Model architecture of the hybrid DBN + OCSVM, which contains DBN, top down, for feature extraction, and OCSVM, bottom up, for anomaly detection.

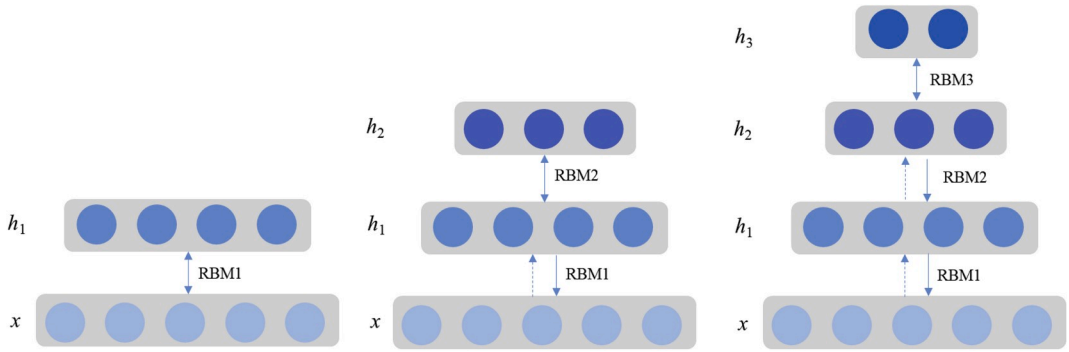

Fig. 2. Constructing process of DBN containing three hidden layers h_1, h_2, h_3, one input layer x. The top two layers have undirected connections and form an associative memory. The layers below have directed, top-down, generative connections that can be used to map a state of the associative memory to an image. There are also directed, bottom-up, recognition connections that are used to infer a factorial representation in one layer from the binary activities in the layer below.

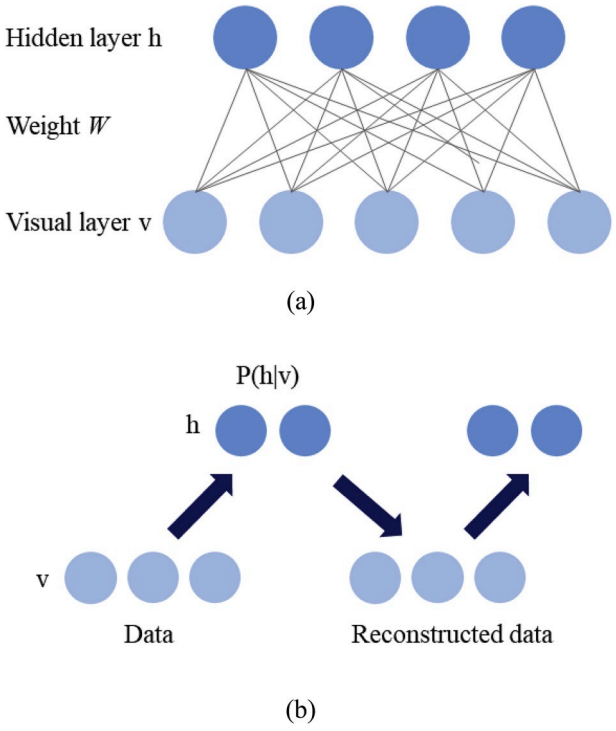

Fig. 3. (a) An architecture of restricted Boltzmann machines containing hidden layer, visual layer and connection weight; (b) Contrastive divergence for restrict Boltzmann machine.

$$w_{ij}\left(t+1\right) = w_{ij}\left(t\right) + \eta_w \frac{\partial log(p(v))}{\partial w_{ij}}, \tag{1}$$

The second part on the right of the equation ($\frac{\partial log(p(v))}{\partial w_{ij}}$) is the gradient, the simple form of which is $\eta_w(\langle v_i h_j\rangle - \langle \widehat{v}_i \widehat{h}_j\rangle)$. Here, $p(v)$ represents the probability of the visible vector, $\langle\cdot\rangle$ represents the expected value with respect to the corresponding probability distribution, and η_w denotes the learning rate for the weights. The relationships between the states of the visible units (v_i for original data and $\widehat{v}_i$ for the reconstructed data) and the states of hidden units (h_j for original data and $\widehat{h}_j$ for the reconstructed data) can be determined as follows:

1. Update the hidden units in parallel given the visible units: $\widehat{h}_j = f\left(b_j + \sum_i w_{ij} v_i\right)$
2. Update the visible units in parallel given the hidden units: $\widehat{v}_i = f\left(b_i + \sum_j w_{ij} h_j\right)$. The first two steps are called "reconstruction".
3. Perform the weight update: $\Delta w_{ij} = w_{ij}(t+1) - w_{ij}(t) = \eta_w(\langle v_i h_j\rangle - \langle \widehat{v}_i \widehat{h}_j\rangle)$

Here, f is the sigmoid function, b_i and b_j represent the biases of the visible unit i and the hidden unit j, respectively.

Once the training of an RBM is completed, another RBM is "stacked" on the top of it, taking its input from the output of the previous RBM. The new RBM is then trained with the procedure outlined above until the desired stopping criterion is reached. The stacking RBMs of exhibits a strong ability to generate non-linear feature detectors at different levels denoting increasingly and more representative features of the data. Thus, the weights of the multi-layer neural network are initialized by the weights of the stacked RBMs instead of the traditional method of random initialization weights. This network can be adopted for dimensionality reduction, fine-tuning by back-propagation or classification by the addition of a logistic regression layer.

2.2. One-class support vector machine

For the second step of the DBN + OCSVM architecture, the detected features of the DBN are fed to the OCSVM for geochemical anomaly detection. An SVM is a promising supervised learning algorithm that is highly effective for both regression and classification tasks. The SVM solves the issues of linear inseparability by constructing a discriminant function in a high-dimensional domain with nonlinear inner product kernel functions, and aims to find a hyperplane with the maximum boundary from a different class dataset (Cortes and Vapnik, 1995). As an important branch of pattern recognition, one-class classification is a special two-class issue, whose training data contain only one class of information. A one class classifier, such as an OCSVM (Schölkopf et al., 2001), can solve the one class classification and anomaly detection issues in an unsupervised way. Regarding the extracted features from the DBN, $X = \{x_1, x_2, ..., x_m\}, x_i \in R^d$ is the training sample set of the OCSVM, with R^d and d representing the input space and its dimension, respectively. The main objective of the OCSVM is to find a hyper-plane or hyper-sphere boundary to separate the geochemical background from the anomalies by maximizing the distance between them (Fig. 4). This decision boundary can be defined as:

$$f(x) = sign(w \cdot \Phi(x) - \rho), \tag{2}$$

Here $\Phi(x)$ is the map function, whose inner product constructs the kernel functions $K(x_i, x) = \Phi(x_i)^T \Phi(x)$ in the SVM solutions (Hastie et al., 2004). Common choices for the kernel functions include: linear, radial basis (RBF), polynomial, and sigmoid (Min and Lee, 2005). To

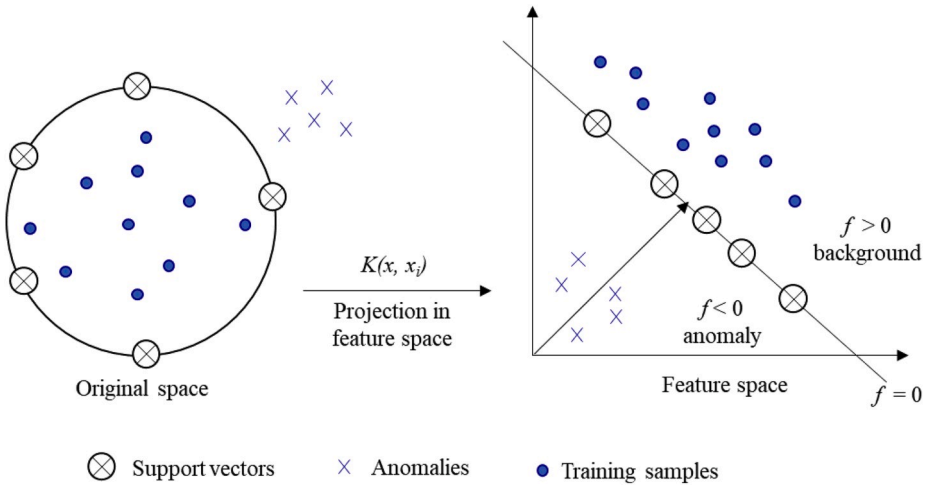

Fig. 4. OCSVM for anomaly detection through non-linear transformation from input space to high-dimensional feature space. The meaning of symbols is presented in Section 2.2.

obtain a weight vector, w, and an offset, ρ, for separating the anomalies records with a maximum margin, the following quadratic optimization issue must be addressed:

$$min_{w,\xi,\rho}\frac{1}{2}w^{T}w+\frac{1}{vm}\sum_{i=1}^{m}\xi_i \quad - \quad \rho, \tag{3}$$

$$\text{subject to } w^{T}\Phi(x)\geq -\rho-\xi_i, \xi_i\geq 0, i=1,2,...,m$$

where ξ_i represents the slack variables penalized in the objective function and $\nu\in(0,1)$ is a parameter corresponding to the ratio of "anomalies" in the training samples.

The Lagrange multipliers changed the quadratic optimization prob-

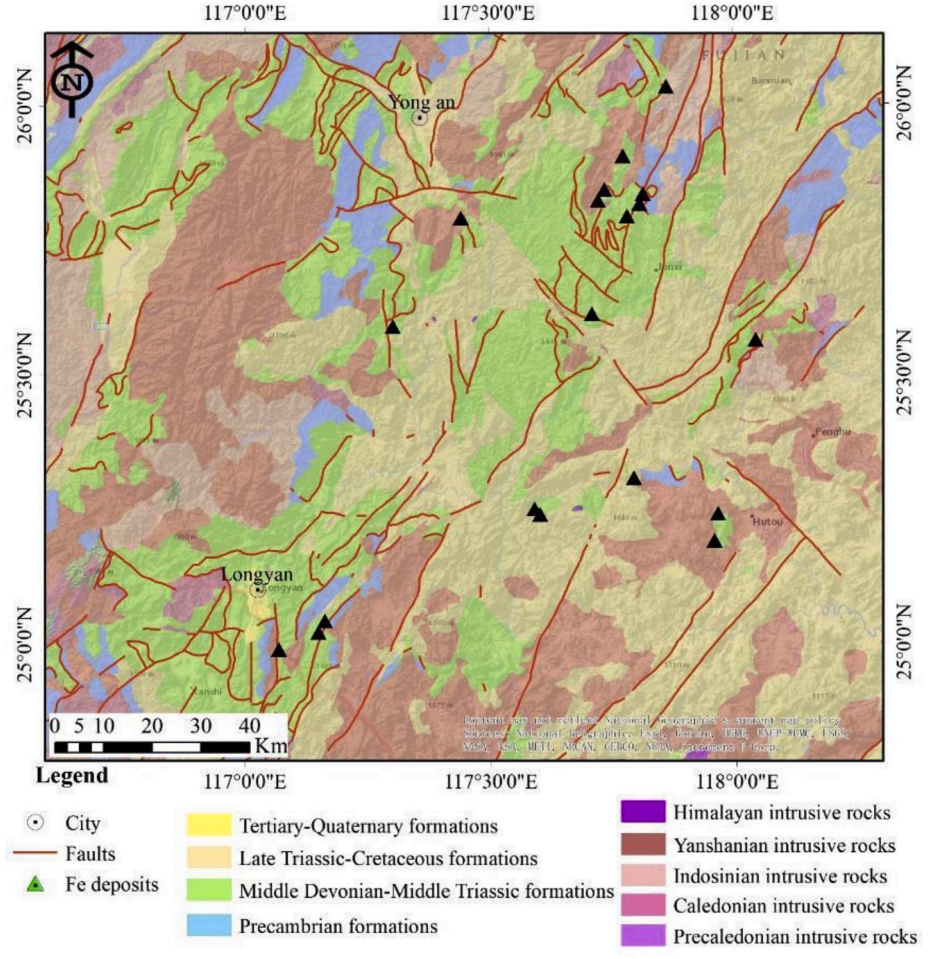

Fig. 5. Simplified geological map of southwest Fujian district (revised from Xiong and Zuo, 2016).

lem to the dual problem which is expressed as follows:

$$w = \sum_{i=1}^{m} a_i \Phi\left(\mathrm{x}_i\right) \quad s.t.\ 0 \le a_i \le \frac{1}{vm}, \sum_{i=1}^{m} a_i = 1, \tag{4}$$

Thus, the decision function $f(\mathrm{x})$ is expressed as:

$$f\left(x\right) = sign\left(\sum_{i=1}^{m} a_i K(x_i, x) - \rho\right), \tag{5}$$

In multivariate geochemical anomaly recognition, this decision function can be regarded as an index to determine whether a sample, x, belongs to anomaly via the following rules:

$$f(x) > 0, \quad if \quad x \in background;$$

$$f(x) < 0, \quad if \quad x \in anomaly; \tag{6}$$

3. Study area and data

The southwestern Fujian district (SFD) (Fig. 5), located in Yangzte Block and Cathaysia Block, is renowned for the Mesozoic extensive tectonic magmatism and numerous discovered iron deposits (e.g. Makeng, Luoyang and Dapai iron deposit) distributed along the early Hercynian Yong'an–Meixian fold belt above the Caledonian substrate (Ge et al., 1981; Han and Ge, 1983). The geological evolution occurred in this district are multi-stage which consist of clastic and carbonate sedimentation, Indosinian W–E orogenic movement, the conversion of structure stress field of early Yanshanian, and extensive magmatism, and their corresponding evolution timing are late Devonian–Permian (D_3–P), Early–Middle Permian (T_1–T_2), Late Permian–Middle Jurassic (T_3–J_2) and Late Jurassic–Early Cretaceous (J_3–K_1), respectively (Zhou and Li, 2000; Sun et al., 2007). The zircon U–Pb dating of granitoids in Zhongjia, Dayang, Juzhou, Luoyang and Dapai are 99 Ma (Yang et al., 2008), 125–145 Ma (Zhang et al., 2012a; Wang et al., 2015), 132 Ma (Wang et al., 2015), 131–132 Ma (Zhang et al., 2012b) and 134 Ma (Yuan et al., 2013), respectively. These ages indicate that the regional Fe-related polymetallic mineralization exhibit a close temporal and spatial relationships with Yanshanian granites which act as the heat and fluid source of mineralization. The three dominant fault structures, Zhenghe-Dapu, Nanping-Ninghua, and Shanghang-Yunxiao, combined with the widespread secondary faults are favorable channels for fluids flow (Zhang and Zuo, 2014; Zhang et al., 2016). The dominant lithologies and ore-hosting formations are the late Paleozoic marine sedimentary rocks and the middle−lower Carboniferous carbonate and clastic rock, respectively. The ore-hosting formations, discontinuously distributed along Zhenghe-Dapu Fault, can provide depositional space for mineralization because their suitable physical/chemical condition for deposition (Zhang and Zuo, 2014).

In addition, skarn alterations, which are direct evidence for skarn-type deposits, can be reflected by geochemical anomalies detected from exploration geochemical data. These data were collected at a 2 km × 2 km grid (Xie et al., 1997), and analyzed for 39 geochemical element concentrations. The analytical system of these 39 elements is introduced in detailed in Xie et al. (2008). The general methods for geochemical anomalies detection only contain parts of elements associated with mineralization, and neglect several elements for mapping geochemical signatures associated with mineralization. In this context, the identified geochemical patterns could not report all geochemical properties associated with mineralization (Zuo and Xiong, 2018). A novel idea based on big data analytics for geochemical data processing considers statistical correlations instead of causal relations based on all the available geochemical data (Zuo and Xiong, 2018;Xiong et al., 2018). Thus, all the 39 elements of geochemical data in this district are adopted for mining statistical correlations between geochemical anomalies and known deposits in this study. The geochemical data are composition data and should be processed prior to data analysis due to the closure problem (Aitchison, 1986; Egozcue et al., 2003; Filzmoser and Hron, 2008). The isometric logratio transformation (*ilr*) (Egozcue et al., 2003) was applied to process the raw geochemical data to address the closure problem.

4. Results and discussion

4.1. Choice of parameters

The selection of appropriate model parameters is critical to the performance of the hybrid method of DBN + OCSVM. According to the big data idea of Zuo and Xiong (2018), all elements of the stream sediment data in the study area were selected as the input for the training of the DBN, Thus, the number of input layer units and hidden layer units of the DBN were assigned as 39, 128, 64, 32 and 16, respectively. The detailed optimal parameters selection process of the DBN have been drawn in Xiong and Zuo (2016).

The performance of OCSVM is directly determined by the appropriate parameters: v and the kernel function, K. The receiver operating characteristic (ROC) (Fawcett, 2006) and the area under the curve (AUC) were used as the performance measures for selecting v and K. An AUC ranging from 0.5 to 1, can be employed as the performance metrics of the DBN and OCSVM. The closer the AUC is to 1, the better the performance of the model. Chen and Wu (2017) discussed the relationship between the AUC of an identified anomaly and the varied v value. With the increase of v, the AUC value first increased and then decreased, reaching the maximum at 0.25. This idea was adopted here to set parameter v here. The decision function values based on the OCSVM, the hybrid DBN + OCSVM, as well as the comparative method PCA + OCSVM (a hybrid model consisting of principle component analysis (PCA) and OCSVM) and were calculated. Note that PCA is used as a reference dimensionality reduction algorithm as it is one of the most prevalent approaches for feature extraction. All three methods can obtain the maximum AUC values at 0.3; thus, the parameter v for controlling the anomaly ratio was set to 0.3 (Fig. 6).

Another critical step for modelling SVMs is to select a suitable kernel function from the four widely used kernel functions, such as linear, polynomial, RBF and sigmoid. The multivariate geochemical anomaly maps extracted by RBF kernels have the largest AUC value, followed by the polynomial kernels (Fig. 7). This conclusion is consistent with that in Zuo and Carranza (2011), where the accuracies of the RBF and polynomial kernels were higher compared with those of the linear and sigmoid kernels.

According to the selected appropriate model parameters, the decision function values corresponding to each cell in the study area could

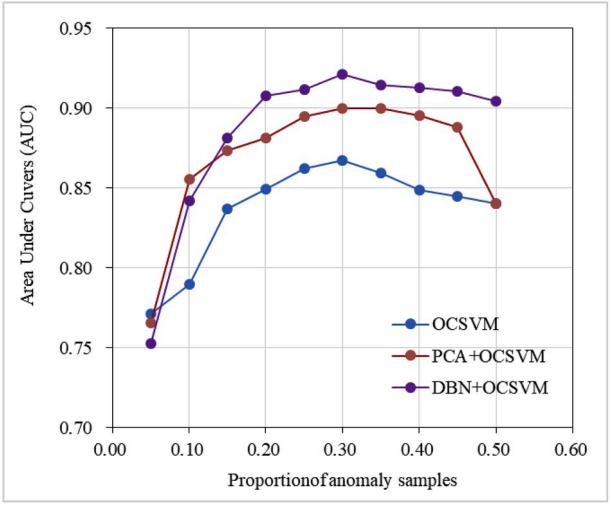

Fig. 6. Diagram of AUCs with varying parameter v.

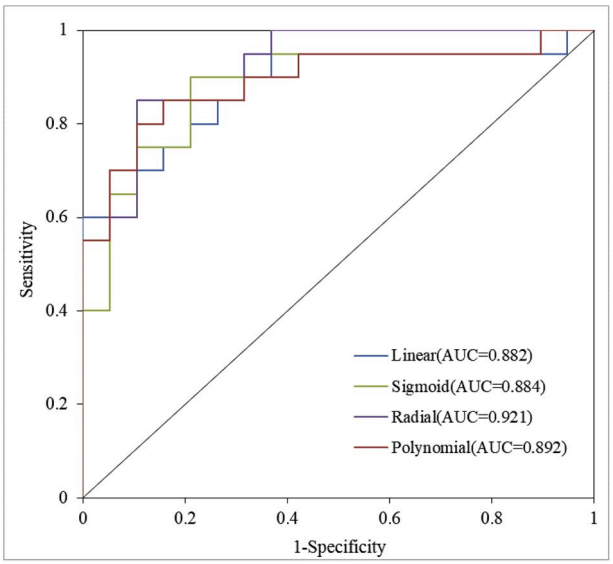

Fig. 7. Comparison of ROC of the hybrid DBN-OCSVM based on different kernel functions.

be calculated by the DBN + OCSVM model. The geochemical samples on each cell with positive decision function values are distinguished as background; while those with negative decision function values are distinguished as anomalies. The extracted geochemical anomalies show a close spatial correlation with the discovered iron deposits (Fig. 8). ROC curve and AUC values were adopted to further quantify the effectiveness of the extracted geochemical anomalies. The AUC value of the hybrid method for the identified geochemical anomalies is 0.921, indicating that the anomalies extracted by the hybrid model show close spatial correlation with known iron deposits, and can effectively delineate the multivariate geochemical anomalies reflecting iron mineralization in SFD.

4.2. *Influence of dimensionality reduction*

To further evaluate the validity of the model for high-dimensional geochemical anomaly recognition, a comparative study of the four multivariate geochemical anomaly detection models was conducted involving a deep autoencoder network (DAE), OCSVM, DBN + OCSVM, and PCA + OCSVM. Among these four models, DAE, as a variation of DBN, comprises of an encoder and a decoder that is trained by minimizing the difference between the input and reconstructed output (Hinton and Salakhutdinov, 2006). The DAE has proved to be a powerful tool for multivariate geochemical anomaly detection by considering the reconstruction error as the anomaly recognition index, whose high and low values correspond with the geochemical anomalies and background, respectively (Xiong and Zuo, 2016). The anomaly recognition index of the other three models are the decision function values of the OCSVM, which are positive for the geochemical background samples and negative for the geochemical anomaly samples. The extracted geochemical anomalies of these four architectures show close spatial distribution patterns, thus exhibiting a close spatial relationship with the discovered iron deposits (Figs. 8 and 9).

The AUC values of the anomalies extracted by the different models

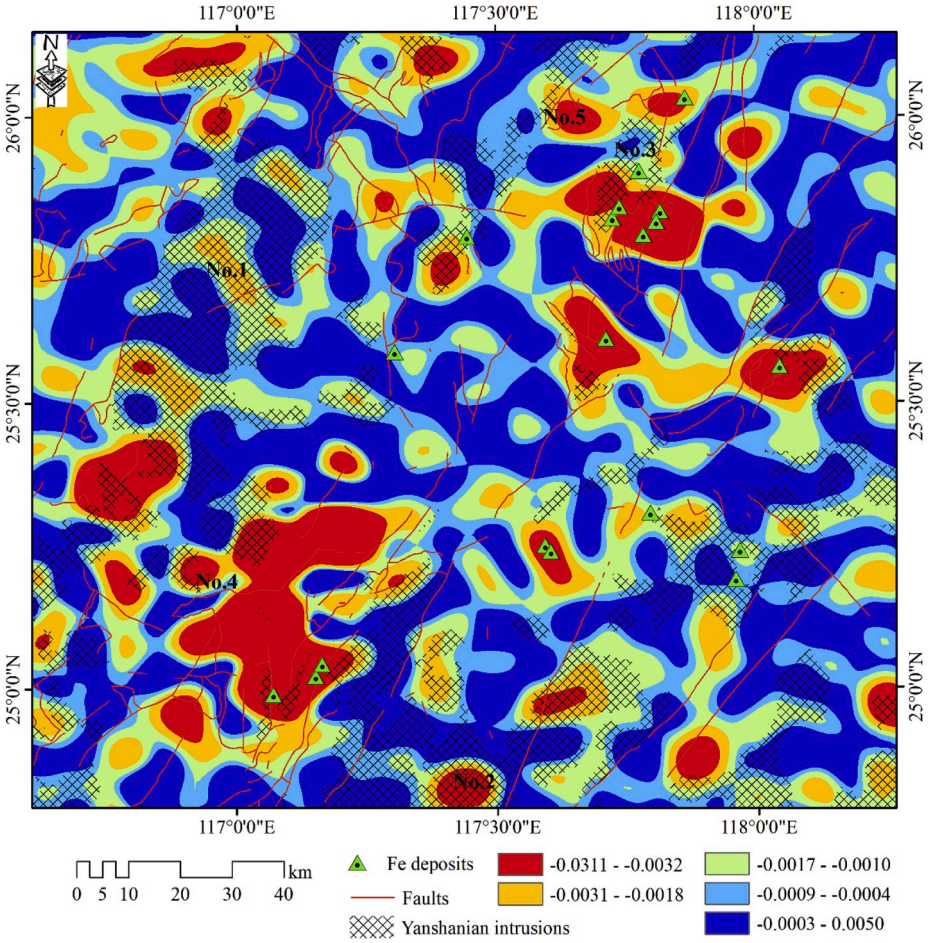

Fig. 8. Geochemical anomaly maps based on hybrid DBN + OCSVM.

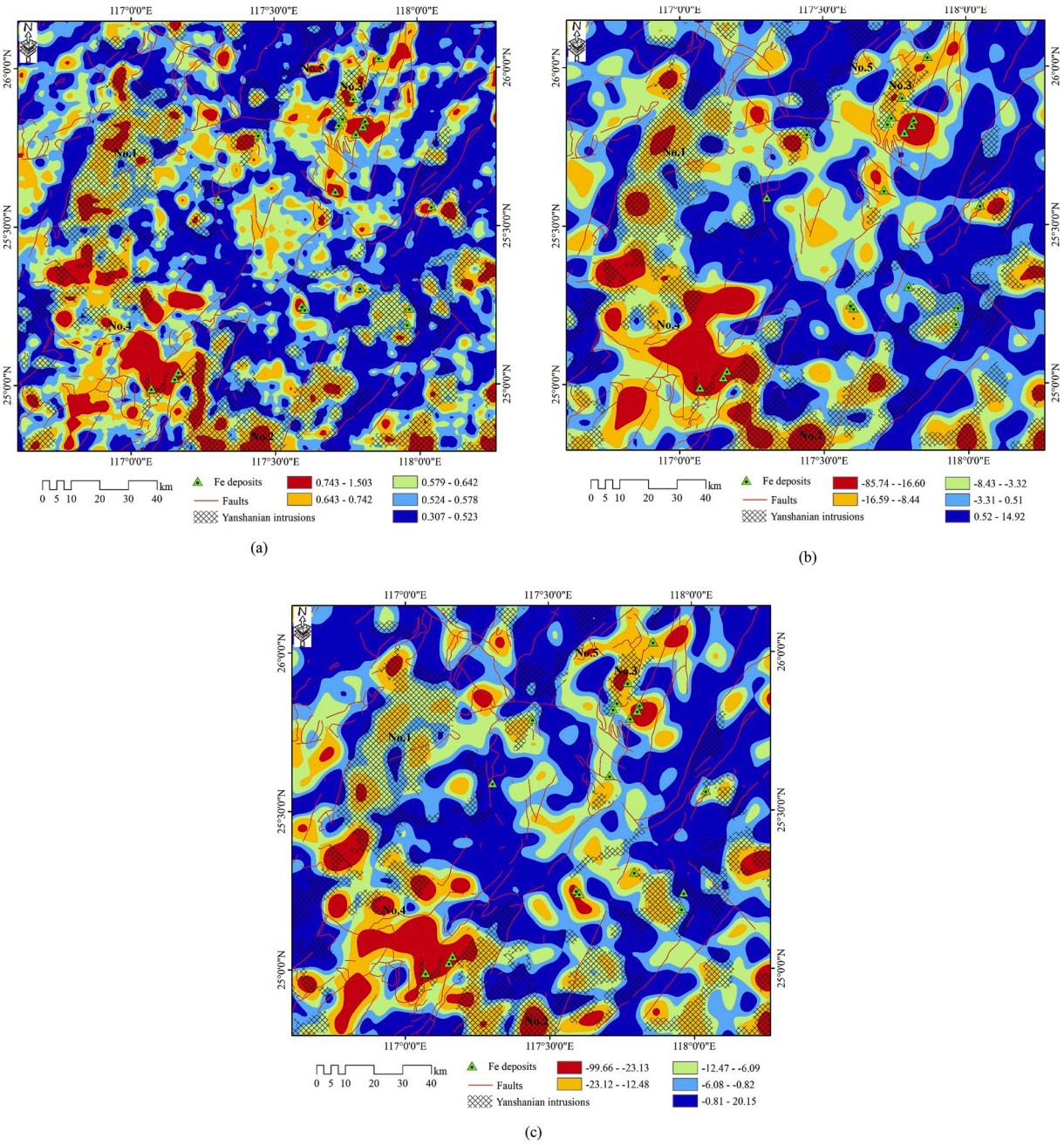

Fig. 9. Geochemical anomaly maps based on (a) deep autoencoder; (b) One class SVM; (c) hybrid method with PCA and one class SVM.

are 0.921 (DBN + OCSVM), 0.900 (DAE), 0.900 (PCA + OCSVM), and 0.867 (OCSVM), respectively (Fig. 10). Note that the number of principal components was set to 16 to retain 95% of the variance. As expected, the mapped geochemical anomalies obtained by the OCSVM show a lower capability for geochemical datasets with all 39 elements. In contrast, the geochemical anomalies detected by the hybrid OCSVM suggest that both linear feature extraction method (e.g., PCA) and nonlinear feature detection technique (e.g., DBN) can improve the accuracy of the OCSVM technique. Among them, the DBN achieved the best accuracy. In terms of their AUC values, the PCA-based OCSVM shows an increase of approximately 3.3%, while the DBN-based OCSVM shows an increase of 5.4%. Compared with linear dimensional reduction methods (e.g. PCA), deep learning methods (such as DBN) are better at characterizing complex unknown high-dimensional geochemical data caused by the geological processes. This can be further supported by the results of the AUC values with different values of ν, where the AUC values of the hybrid OCSVM (PCA + OCSVM and DBN + OCSVM) are generally higher than the single OCSVM, and DBN + OCSVM are generally higher than PCA + OCSVM.

From a geological point, most geochemical anomalies obtained in this study are distributed in the Yanshanian intrusions, such as the area numbered No.1, No.2 and No.3, or around the boundaries of the Yanshanian intrusions, such as the area numbered No.4 and No.5. This observation is consistent with the conclusions that Yanshaian intrusions are the heat and fluid sources for the formation of Fe mineralization, and the contact zones of the granites are the appropriate positions for the Fe mineral deposits in this district (Zhang et al., 2016), indicating a credible result was derived from this study. The detected anomalies can also provide a significant assistance for the next round of mineral exploration in this district.

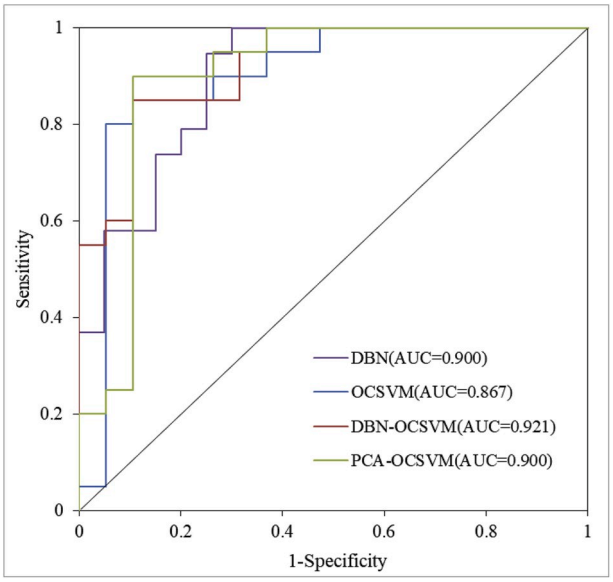

Fig. 10. Diagram of AUCs derived from different methods.

5. Conclusions

In this study, an unsupervised anomaly detection algorithm for high-dimensional data was adopted for identifying geochemical anomalies that might indicate iron polymetallic mineralization in SFD. The algorithm is a hybrid architecture of DBN and OCSVM, where the DBN was used for dimensionality reduction and feature extraction to transform the geochemical data into a lower dimensional set of features, which were further used as input for the training of OCSVM. The hybrid DBN + OCSVM model employs the strong abilities of the DBN in extracting data representations, and effectively solves the problems that limit OCSVM when dealing with complex high-dimensional data. The other classical anomaly detection methods, besides the hybrid DBN + OCSVM, such as the DAE, OCSVM, and hybrid PCA + OCSVM models were adopted for recognizing the multivariate geochemical anomalies in SFD. The extracted geochemical anomalies of these four architectures exhibited similar spatial patterns, and a close spatial relationship with the known deposits. The comparative results of the hybrid and stand along methods (DAE, OCSVM, and hybrid PCA + OCSVM) suggest that the DBN can extract geochemical information from deeper levels and better reduce the number of redundant features. This enhances the scalability of the OCSVM for processing the geochemical data containing high-dimensional records.

Author contribution

The contribution of each author: Yihui Xiong and Renguang Zuo collaborated to finish the design of the hybrid model. Yihui Xiong wrote the code to carry out the model and wrote the paper. Renguang Zuo revised the paper.

Computer code availability

The codes based on MATLAB platform for the hybrid of DBN and OCSVM are accessible at https://github.com/CUG-MG-GROUP/DBN-OCSVM.

Declaration of competing interest

The authors have declared that no conflict of interest exists.

Acknowledgement

Thanks due to two anonymous reviewers' comments and suggestions which help us improve this study. This research was jointed supported by the National Natural Science Foundation of China (No. 41772344), and MOST Special Fund from the State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences (MSFGPMR25).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cageo.2020.104484.

References

Abe, S., 2005. Support Vector Machines for Pattern Classification. Springer, London.
Agterberg, F.P., 2001. Multifractal simulation of geochemical map patterns. In: Geologic Modeling and Simulation. Springer, Boston, MA, pp. 327–346.
Aitchison, J., 1986. The Statistical Analysis of Compositional Data. Chapman & Hall, London, 416pp.
Bengio, Y., LeCun, Y., 2007. Scaling learning algorithms towards AI. Large-scale kernel machines 34, 1–41.
Bengio, Y., Delalleau, O., Le Roux, N., 2005. The curse of dimensionality for local kernel machines. Techn. Rep 1258.
Carranza, E.J.M., 2008. Geochemical Anomaly and Mineral Prospectivity Mapping in GIS, 11. Elsevier.
Chen, X., Lin, X., 2014. Big data deep learning: challenges and perspectives. IEEE Access 2, 514–525.
Chen, Y., Wu, W., 2017. Application of one-class support vector machine to quickly identify multivariate anomalies from geochemical exploration data. Geochem. Explor. Environ. Anal. 17, 231–238.
Chen, Y., Lu, L., Li, X., 2014. Application of continuous restricted Boltzmann machine to identify multivariate geochemical anomaly. J. Geochem. Explor. 140, 56–63.
Chen, L., Guan, Q., Xiong, Y., Liang, J., Wang, Y., Xu, Y., 2019. A Spatially Constrained Multi-Autoencoder approach for multivariate geochemical anomaly recognition. Comput. Geosci. 125, 43–54.
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 for predicting undiscovered mineral deposits in covered areas. J. Geochem. Explor. 122, 55–70.
Cheng, Q., Agterberg, F.P., Ballantyne, S.B., 1994. The separation of geochemical anomalies from background by fractal methods. J. Geochem. Explor. 51, 109–130.
Cheng, Q., Xu, Y., Grunsky, E., 2000. Integrated spatial and spectrum method for geochemical anomaly separation. Nat. Resour. Res. 9, 43–52.
Cohen, D.R., Kelley, D.L., Anand, R., Coker, W.B., 2010. Major advances in exploration geochemistry, 1998-2007 Geochem. Explor. Environ. Anal. 10, 3–16.
Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach. Learn. 20, 273–297.
Egozcue, J.J., Pawlowsky-Glahn, V., Mateu-Figueras, G., Barcelo-Vidal, C., 2003. Isometric logratio transformations for compositional data analysis. Math. Geol. 35, 279–300.
Erfani, S., Baktashmotlagh, M., Rajasegarar, S., Karunasekera, S., Leckie, C., 2015. R1SVM: a randomised nonlinear approach to large-scale anomaly detection. In: AAAI Conference on Artificial Intelligence, pp. 432–438.
Erfani, S.M., Rajasegarar, S., Karunasekera, S., Leckie, C., 2016. High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning. Pattern Recogn. 58, 121–134.
Erhan, D., Manzagol, P.A., Bengio, Y., Bengio, S., Vincent, P., 2009. The difficulty of training deep architectures and the effect of unsupervised pre-training. In: Artificial Intelligence and Statistics, pp. 153–160.
Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recogn. Lett. 27, 861–874.
Filzmoser, P., Hron, K., 2008. Outlier detection for compositional data using robust methods. Math. Geosci. 40, 233–248.
Filzmoser, P., Garrett, R.G., Reimann, C., 2005. Multivariate outlier detection in exploration geochemistry. Comput. Geosci. 31, 579–587.
Ge, C., Han, F., Zhou, T., Chen, D., 1981. Geological characteristics of the Makeng iron deposit of marine volcano-sedimentary origin. Acta Geosicientia Sinica 3, 47–69 (in Chinese with English Abstract).
Gonbadi, A.M., Tabatabaei, S.H., Carranza, E.J.M., 2015. Supervised geochemical anomaly detection by pattern recognition. J. Geochem. Explor. 157, 81–91.
Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, 482pp.
Han, F., Ge, C., 1983. Geological and Geochemical Features of Submarine Volcanic Hydrothermal-Sedimentary Mineralization of Makeng Iron Deposit, Fujian Province. Bulletin of the Institute of Mineral Deposits Chinese Academy of Geological Science7, pp. 1–118 (in Chinese with English abstract).
Hastie, T., Rosset, S., Tibshirani, R., Zhu, J., 2004. The entire regularization path for the support vector machine. J. Mach. Learn. Res. 5, 1391–1415.

Hawkes, H.E., Webb, J.S., 1962. Geochemistry in Mineral Exploration. Harper and Row, New York, NY.
Hinton, G.E., 2002. Training products of experts by minimizing contrastive divergence. Neural Comput. 14, 1771–1800.
Hinton, G.E., Osindero, S., Teh, Y.W., 2006. A fast learning algorithm for deep belief nets. Neural Comput. 18, 1527–1554.
Kürzl, H., 1988. Exploratory data analysis: recent advances for the interpretation of geochemical data. J. Geochem. Explor. 30, 309–322.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436.
Matheron, G., 1962. Traité de géostatistique appliquée. Editions Technip.
Min, J.H., Lee, Y.C., 2005. Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. Expert Syst. Appl. 28, 603–614.
Moeini, H., Torab, F.M., 2017. Comparing compositional multivariate outliers with autoencoder networks in anomaly detection at hamich exploration area, east of Iran. J. Geochem. Explor. 180, 15–23.
Reimann, C., Garrett, R.G., 2005. Geochemical background-concept and reality. Sci. Total Environ. 350, 12–27.
Schmidhuber, J., 2015. Deep learning in neural networks: an overview. Neural Network. 61, 85–117.
Schölkopf, B., Platt, J.C., Shawe-Taylor, J., Smola, A.J., Williamson, R.C., 2001. Estimating the support of a high-dimensional distribution. Neural Comput. 13, 1443–1471.
Sinclair, A.J., 1974. Selection of threshold values in geochemical data using probability graphs. J. Geochem. Explor. 3, 129–149.
Sun, W., Ding, X., Hu, Y., Li, X., 2007. The golden transformation of the Cretaceous plate subduction in the west Pacific. Earth Planet Sci. Lett. 262, 533–542.
Tukey, J.W., 1977. Exploratory Data Analysis. Addison-Wesley, Reading, USA.
Vempati, S., Vedaldi, A., Zisserman, A., Jawahar, C.V., 2010. Generalized RBF feature maps for efficient detection. In: British Machine Vision Conference, 1–11.
Wang, Z., Dong, Y., Zuo, R., 2019a. Mapping geochemical anomalies related to Fe–polymetallic mineralization using the maximum margin metric learning method. Ore Geol. Rev. 107, 258–265.
Wang, S., Zhang, D., Vatuva, A., Yan, P., Ma, S., Feng, H., Yu, T., Bai, Y., Di, Y., 2015. Zircon U-Pb geochronology, geochemistry and Hf isotope compositions of the Dayang and Juzhou granites in Longyan, Fujian and their geological implications, 44, 450–468 (In Chinese with English abstract).
Wang, Z., Zuo, R., Dong, Y., 2019b. Mapping geochemical anomalies through integrating random forest and metric learning methods. Nat. Resour. Res. 28, 1285–1298.
Wu, W., Chen, Y., 2018. Application of isolation forest to extract multivariate anomalies from geochemical exploration data. Glob. Geol. 21, 36–47.
Xie, X., Mu, X., Ren, T., 1997. Geochemical mapping in China. J. Geochem. Explor. 60, 99–113.
Xie, X., Wang, X., Zhang, Q., Zhou, G., Cheng, H., Liu, D., Cheng, Z., Xu, S., 2008. Multiscale geochemical mapping in China. Geochem. Explor. Environ. Anal. 8, 333–341.
Xiong, Y., Zuo, R., 2016. Recognition of geochemical anomalies using a deep autoencoder network. Comput. Geosci. 86, 75–82.
Xiong, Y., Zuo, R., Carranza, E.J.M., 2018. Mapping mineral prospectivity through big data analytics and a deep learning algorithm. Ore Geol. Rev. 102, 811–817.
Yang, Z., Zhang, D., Feng, C., She, H., Li, J., 2008. SHRIMP zircon U–Pb dating of quartz porphyry from Zhongjia tin–polymetallic deposit in Longyan area, Fujian Province, and its geological significance. Miner. Deposits 27, 329–335 (in Chinese with English Abstract).
Yuan, Y., Feng, H., Zhang, D., Di, Y., Wang, C., Ni, J., 2013. Geochronology of Dapai iron-polymetallic deposit in Yongding city, Fujian province and its geological significance. Acta Mineral. Sin. (Suppl. l), 73–75 (in Chinese with english abstract).
Zhang, Z., Zuo, R., 2014. Sr–Nd–Pb isotope systematics of magnetite: implications for the genesis of Makeng Fe deposit, southern China. Ore Geol. Rev. 57, 53–60.
Zhang, C., Li, L., Zhang, C., Wang, J., 2012a. LA-ICP-MS zircon U-Pb ages and Hf isotopic compositions of Dayang granite from Longyan, Fujian province. Geoscience 26, 434–444 (in Chinese with English abstract).
Zhang, D., Wu, G., Di, Y., Wang, C., Yao, J., Zhang, Y., Lv, L., Yuan, Y., Shi, J., 2012b. Geochronology of diagenesis and mineralization of the Luoyang iron deposit in Zhangping city, Fujian province and its geological significance. Earth Science. J. China Univ. Geosci. 37, 1217–1231.
Zhang, Z., Zuo, R., Xiong, Y., 2016. A comparative study of fuzzy weights of evidence and random forests for mapping mineral prospectivity for skarn-type Fe deposits in the southwestern Fujian metallogenic belt, China. Sci. China Earth Sci. 59, 556–572.
Zhang, S., Xiao, K., Carranza, E.J.M., Yang, F., Zhao, Z., 2019. Integration of auto-encoder network with density-based spatial clustering for geochemical anomaly detection for mineral exploration. Comput. Geosci. 130, 43–56.
Zhao, J., Chen, S., Zuo, R., 2016. Identifying geochemical anomalies associated with Au–Cu mineralization using multifractal and artificial neural network models in the Ningqiang district, Shaanxi, China. J. Geochem. Explor. 164, 54–64.
Zhou, X., Li, W., 2000. Origin of Late Mesozoic igneous rocks in Southeastern China: implications for lithosphere subduction and underplating of mafic magmas. Tectonophysics 326, 269–287.
Ziaii, M., Pouyan, A.A., Ziaei, M., 2009. Neuro-fuzzy modelling in mining geochemistry: identification of geochemical anomalies. J. Geochem. Explor. 100, 25–36.
Ziaii, M., Ardejani, F.D., Ziaei, M., Soleymani, A.A., 2012. Neuro-fuzzy modeling based genetic algorithms for identification of geochemical anomalies in mining geochemistry. Appl. Geochem. 27, 663–676.
Zuo, R., Carranza, E.J.M., 2011. Support vector machine: a tool for mapping mineral prospectivity. Comput. Geosci. 37, 1967–1975.
Zuo, R., Wang, J., 2016. Fractal/multifractal modeling of geochemical data: a review. J. Geochem. Explor. 164, 33–41.
Zuo, R., Xiong, Y., 2018. Big data analytics of identifying geochemical anomalies supported by machine learning methods. Nat. Resour. Res. 27, 5–13.
Zuo, R., Xiong, Y., 2020. Geodata science and geochemical mapping. J. Geochem. Explor. 209, 106431.
Zuo, R., Carranza, E.J.M., Wang, J., 2016. Spatial analysis and visualization of exploration geochemical data. Earth Sci. Rev. 158, 9–18.
Zuo, R., Xiong, Y., Wang, J., Carranza, E.J.M., 2019. Deep learning and its application in geochemical mapping. Earth Sci. Rev. 192, 1–14.