



A spatial-compositional feature fusion convolutional autoencoder for multivariate geochemical anomaly recognition

Qingfeng Guan^{a,b}, Shuliang Ren^{a,b}, Lirong Chen^c, Bin Feng^d, Yao Yao^{a,e,*}

^a School of Geography and Information Engineering, China University of Geosciences, Wuhan, Hubei, China

^b National Engineering Research Center of GIS, China University of Geosciences, Wuhan, Hubei, China

^c Development and Research Center, China Geological Survey, Xicheng District, Beijing, China

^d Institute of Geophysical and Geochemical Exploration, Chinese Academy of Geological Sciences, Langfang, Hebei, China

^e Alibaba Group, Hangzhou, 311121, Zhejiang Province, China

ARTICLE INFO

Keywords:

Geochemical anomaly recognition
Spatial structural feature
Compositional relationships
Feature fusion
Autoencoder

ABSTRACT

The spatial structural features and compositional relationships of multivariate geochemicals are influenced by complex geological processes (e.g., diagenesis and mineralization), and can help identify geochemical anomalies and provide key references for mineral resource exploration. However, previous machine-learning-based models often treat spatial structural features or compositional relationships separately. Based on the multitask stack autoencoder structure, this study proposes a feature fusion convolutional autoencoder (FCAE) to extract and fuse the spatial structural features and compositional relationships of multivariate geochemicals for identifying geochemical anomalies. In addition, a three-stage training (3ST) strategy combining greedy layerwise pretraining and overall fine-tuning is adopted to calibrate the FCAE. To assess the performance, the proposed FCAE was used to identify the anomalies related to the Cu ore in the southwest area of the Wuyishan polymetallic metallogenic belt in China. The results showed that fusing both spatial structural features and compositional relationships effectively improved the accuracy of the anomaly identification. The FCAE outperformed several existing models by achieving an AUC of 0.863, a recall of 0.909, and the highest intersection point of the P-A plot in the experiments. In addition, the FCAE is less sensitive to the size of the convolution window, which makes the method more applicable and reliable for mineral resource exploration.

1. Introduction

Geochemical analysis has been widely used in mineral resource exploration due to its low cost and fast implementation (Beus and Grigorian 1977; Xie and Ren 1993; Leybourne and Cameron 2010). Recognizing geochemical anomalies is a key task of geochemical exploration (Hawkes and Webb 1963; Zuo 2017). These anomalies are often treated as important indicators of mineralization and are of essential value for mineral resource exploration (Cheng et al., 1994; Zuo et al., 2009).

Many methods have been developed to identify geochemical anomalies (Harris et al., 2000), such as the mean \pm k standard deviations (Hawkes and Webb 1963), exploratory data analysis (Kürzl 1988), moving average technique (Zuo et al., 2016) and spatial fractal methods (Cheng and Agterberg 1996; Cheng 2007). These methods can identify

geochemical anomalies by using the spatial structural patterns and statistical laws of geochemistry. However, they are mainly based on idealized assumptions regarding the frequency distribution (e.g., normal distribution) and linear, low-order features, which cannot fully accommodate the complexity of geochemical patterns (Xiong and Zuo 2016; Zuo 2017).

In recent years, many machine learning methods have been adopted in geochemical exploration and geochemical anomaly identification (Zuo et al., 2019). With the ability to model nonlinear systems, machine learning methods can capture complex and multistage geological events without any known data distribution (Cracknell and Reading 2014). Machine learning has been demonstrated to effectively integrate multivariate geochemical data and extract geochemical anomalies related to mineralization. For example, Twarakavi et al. (2006) predicted the arsenic concentration in sediments using the robust

* Corresponding author. YAO School of Geography and Information Engineering, China University of Geosciences, Wuhan 430078, Hubei province, China.
E-mail addresses: guanqf@cug.edu.cn (Q. Guan), renshuliang@cug.edu.cn (S. Ren), chenlirong@mail.cgs.gov.cn (L. Chen), fengbin@igge.cn (B. Feng), yaoy@cug.edu.cn (Y. Yao).

<https://doi.org/10.1016/j.cageo.2021.104890>

Received 9 November 2020; Received in revised form 10 July 2021; Accepted 13 July 2021

Available online 16 July 2021

0098-3004/© 2021 Elsevier Ltd. All rights reserved.

least-square support vector machines (LS-SVMs) based on the Au concentration distribution in Circle City, Alaska. In the Ningqiang mineral area, China, Zhao et al. (2016) identified the Au-Cu-related geochemical anomalies using artificial neural networks (ANNs) and spectrum-area (S-A) multifractals. Li et al. (2019) constructed a deep convolutional neural network (DCNN) classification model based on AlexNet and effectively extracted the relationships between Mn deposits and 21 geochemicals.

The above machine learning methods do not require any assumptions regarding the data distribution and can effectively integrate multivariate geochemical data and extract information related to mineralization. Nevertheless, a key issue with supervised learning methods is the requirement of a large number of labeled samples. In cases of geochemical anomaly detection, supervised learning methods require many known mineral sites as the labeled samples to train the model. Such a requirement makes supervised learning methods inapplicable in areas without sufficient known mineral sites (Nickerson et al., 2001; Krawczyk 2016). In contrast, unsupervised or self-supervised learning methods do not rely on labeled samples for the training process. As a typical type of self-supervised learning method, the autoencoder (AE) neural network includes one or more encoder layers and a corresponding number of decoder layers (Fig. 1) (Rumelhart et al., 1986). Although training using unlabeled samples, the encoder layers learn the general patterns of the samples, such that the decoder layers can use the learned patterns to reconstruct the samples as close as possible to the original ones (Hinton and Salakhutdinov 2006). As geochemical anomalous areas often cover 1.5%–5% of the total area (Chen et al. 2009, 2014), the majority of geochemical samples represent the general background in a region. Through training, an AE mainly learns the geochemical background patterns from the majority of the samples and can reconstruct the background, while anomalous samples with much fewer occurrences have little impact on the learned patterns. Therefore, anomalous samples can be easily distinguished because of the reconstruction errors are higher than those of background samples (Valentine and Trampert 2012; Sun et al., 2014). The greater the reconstruction error is, the more anomalous the sample.

A number of studies have used AEs for geochemical anomaly recognition. For example, Chen et al. (2014a) used a continuously restricted Boltzmann machine to reconstruct the geochemical background and compare it with the measured values to identify geochemical anomalies without the need for ore-forming samples. Xiong and Zuo (2016) proposed a geochemical anomaly recognition method based on a deep autoencoder (DAN). The results showed that extracting and learning the compositional relationships among multiple geochemical elements were helpful for geochemical anomaly recognition. Chen et al. (2019a) constructed a multiconvolutional autoencoder (MCAE), which can extract the spatial structural features of geochemical elements via

convolution to identify geochemical anomalies.

The spatial structural feature of a certain geochemical element represents the pattern of the spatial distribution of this element within a certain area. The compositional features of a group of geochemical elements represent the pattern of the compositional relationships among these elements within geochemical samples. Previous studies have suggested that the complex spatial and compositional features of geochemicals are the result of various geological events and complex geological processes (Debaille et al., 2006; Zuo 2011; Xiong and Zuo 2016; Zuo et al., 2016). Therefore, the analysis of spatial structural features and compositional relationships of geochemical elements can facilitate the detection and understanding of the primary and secondary geological processes (e.g., mineralization processes), hence providing key information for identifying anomalies related to mineralization. Existing unsupervised machine learning methods can learn these features from large amounts of unlabeled or unclassified geochemical data. However, they only consider either the spatial structural features (such as Chen et al. (2019a)) or the compositional relationships (such as Xiong, Zuo (2016) and Chen et al. (2019b)) of geochemicals and cannot effectively integrate these two aspects to identify geochemical anomalies.

Therefore, the purpose of this study is to construct a self-supervised deep learning model to combine the spatial structural features and compositional relationships of multiple geochemical elements to identify geochemical anomalies related to mineralization. This study proposes a feature fusion convolutional autoencoder (FCAE) based on the stack autoencoder and a multitask learning mechanism. To assess the effectiveness of the FCAE, this study used it to detect copper-related anomalies in the southwestern Wuyishan polymetallic metallogenic belt in China. The receiver operating characteristic (ROC) curve, correlation C-value of weights-of-evidence and prediction-area (P-A) plot were used as performance indicators to evaluate the FCAE, in comparison with several existing methods. In addition, the model's sensitivity to the size of the convolution window was also analyzed.

2. Methods

A new self-supervised model, FCAE, is proposed in this study. The FCAE can extract and fuse the spatial structural features and compositional relationships of geochemical elements for multivariate geochemical anomaly recognition. As shown in Fig. 2, the principle of anomaly detection using the FCAE is the same as for other AE-based methods. The patterns of high-probability samples (i.e., geochemical background) can be better learned and reconstructed by the autoencoder, and those samples with much smaller probabilities (i.e., geochemical anomalies) have larger errors after reconstruction and can be identified (Xiong and Zuo 2016; Zhou and Paffenroth 2017; Chen

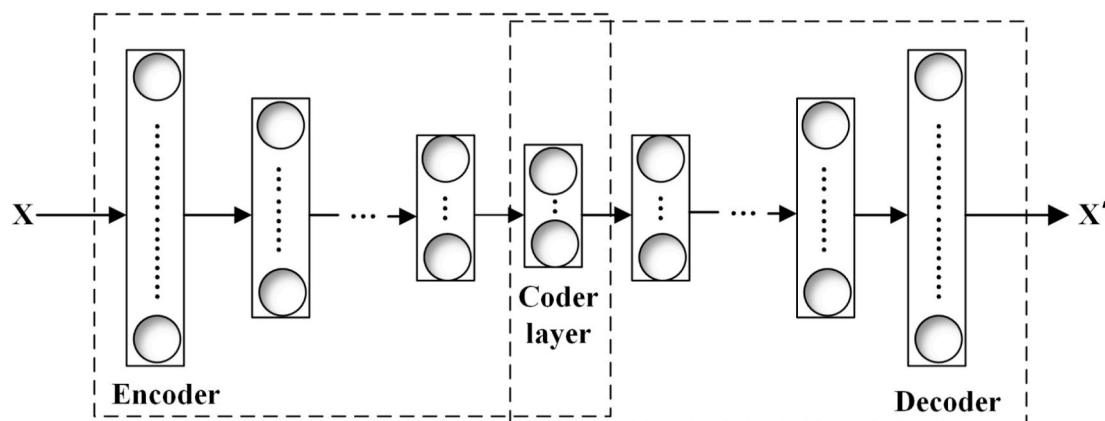


Fig. 1. Schematic diagram of the autoencoder (AE) neural network. X represents the original sample (often a multivariate sample), and X' represents the reconstructed sample.

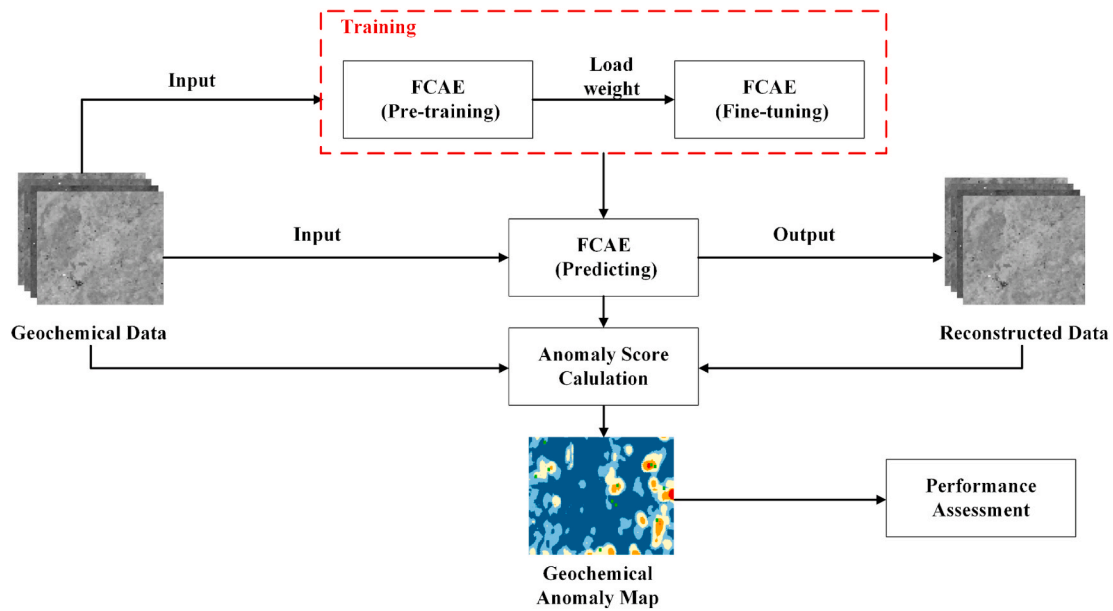


Fig. 2. Flowchart of multivariate geochemical anomaly detection using the FCAE.

et al. 2019a, 2019b). By measuring the differences between the reconstructed samples and the original samples, the anomaly score at each sampling location can be calculated and an anomaly map can be generated. In addition, this study adopts a three-stage training (3ST) strategy that combines greedy layerwise pretraining and overall fine-tuning to calibrate the FCAE.

2.1. Structure of the FCAE

As shown in Fig. 3, the FCAE is composed of two subautoencoders, namely AE-s and AE-c. Specifically, AE-s is designed to extract the spatial structural features of the input geochemical elements, and AE-c extracts the compositional relationships among these elements. The encoder and decoder of the FCAE are made up of the encoders and decoders of the AE-s and AE-c, respectively. The FCAE takes the concentration maps of a group of geochemical elements, and reconstructs the

concentration maps.

2.1.1. Subautoencoder for the spatial structural features (AE-s)

As shown in Fig. 3, the AE-s is a subautoencoder located on the outer layers of the FCAE and is responsible for learning the spatial structural features of all input geochemical elements. Multiple convolutional autoencoders (CAEs) constitute the AE-s, each CAE corresponding to one element. The strengths of the convolutional neural networks (CNNs) and AE neural networks are combined in the CAEs (Masci et al., 2011; Chen et al., 2019a). CNNs can effectively extract the features of images, such as color, texture and density (LeCun et al., 2015). AEs are often used to capture the information embedded in a large number of samples, and reconstruct the samples as close as possible to the original samples.

As shown in Fig. 4, each CAE in the AE-s is composed of spatial information encoder components and decoder components. The encoder consisting of convolutional layers and max-pooling layers can learn the

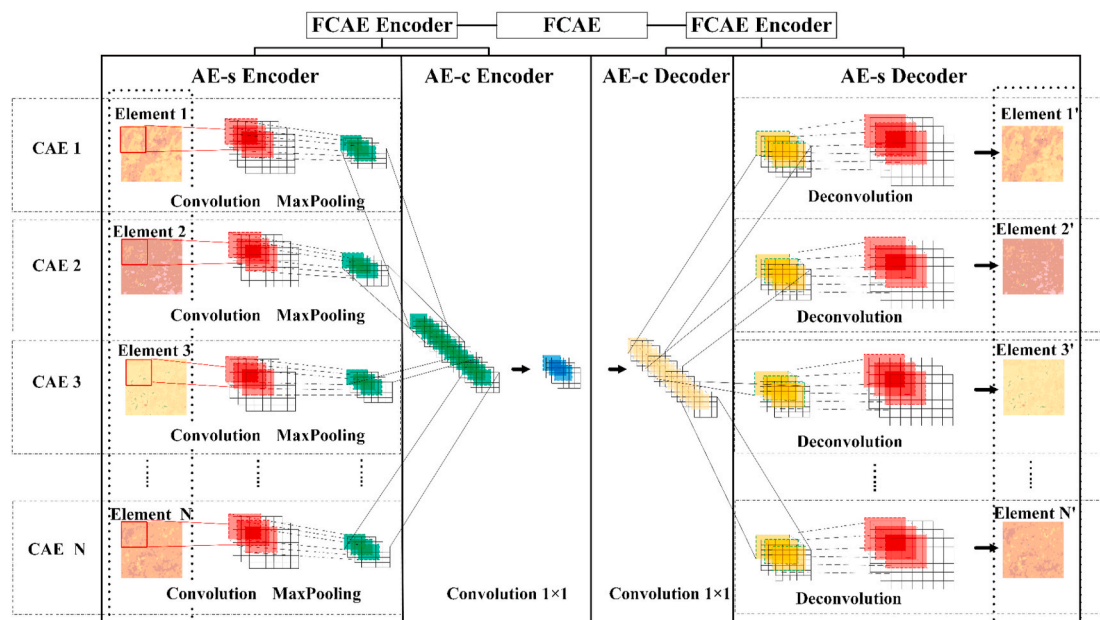


Fig. 3. Structure of the feature fusion convolutional autoencoder (FCAE).

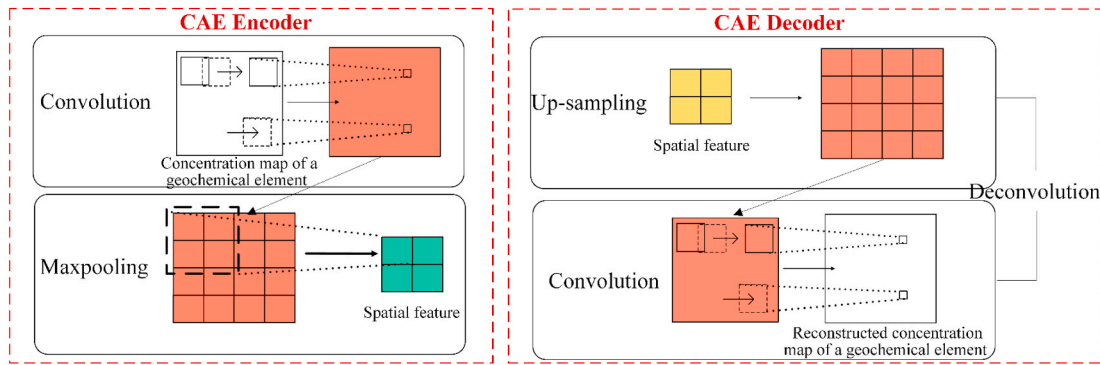


Fig. 4. Encoder and decoder of a CAE in AE-s and the red rectangle represents the convolution kernel. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

spatial structural features of the geochemical background. The decoder, comprising the unpooling layers and convolutional layers, is used to reconstruct the geochemical background with the features learned by the encoder (Ribeiro et al., 2018). Note that deconvolution is realized by using the upsampling layer and the convolutional layer (Chollet et al., 2015; Noh et al., 2015).

2.1.2. Subautoencoder for compositional relationships (AE-c)

Although the spatial features of each geochemical element can be obtained via the AE-s encoder, each element is treated independently. The compositional relationships among multiple geochemical elements also contain key information and can help identify anomalies related to mineralization, as the source or process that has generated the anomalies is commonly associated with a suite of elements (Xiong and Zuo 2016, 2020). Therefore, the FCAE includes a subautoencoder (i.e., the AE-c) to fuse the spatial features of multiple geochemical elements extracted by the AE-s and further obtain the compositional relationships among them.

As shown in Fig. 5, the encoder component and the decoder component of the AE-c are both composed of convolutional layers. Compared with the AE-s, the window size of the convolutional layer is 1×1 in the AE-c. During training, AE-c learns the correlation among multiple channels (feature layers) at the same location. Therefore, the AE-c can learn complex nonlinear relationships among multiple elements.

2.1.3. Activation functions and the cost function

In FCAE, the rectified linear Unit (ReLU) and sigmoid (Nair and Hinton 2010) are used as the activation functions. Specifically, the sigmoid is only used in the output layer. Using different activation functions in different layers can balance the advantages and disadvantages of these two activation functions (Chen et al., 2017). Compared with the sigmoid, the ReLU is more computationally efficient, and does not suffer from the vanishing gradient problem. However, when the weight update is triggered by a significant gradient, it produces “Dying ReLU” (Lu et al., 2019). Therefore, we normalize the input data and apply the sigmoid function to the output layer. The output and input are normalized to the

range of [0,1], which can help prevent the “Dying ReLU problem” during backpropagation (Chen et al., 2017). The cost function of the FCAE is the mean square error (MSE). The adaptive learning rate (AdaDelta) method (Zeiler 2012) is used to train the FCAE.

2.2. Three-stage training (3ST) strategy

The complex structure of the FCAE is likely to lead to the problem of vanishing gradients during model training (Bengio et al., 2007; Bengio 2009). That is, the weights are modified in the hidden layers near the input layer, while the weights in the hidden layers near the output layer are minimally updated or not at all (Larochelle et al., 2009). It is challenging to obtain the optimal weights for the two subautoencoders in the FCAE, especially for the multiple CAEs in the AE-s, which affects the effectiveness of feature extraction and fusion. To address these problems, this study adopts a three-stage training (3ST) strategy that combines greedy layerwise pretraining and overall fine-tuning (Fig. 6). In the first two stages, the AE-s and AE-c are pre-trained separately. In the third stage, the pretrained AE-s and AE-c are combined together as the FCAE, and the entire model is trained as a whole.

First, the AE-s are pretrained using multivariate geochemical samples. Specifically, each CAE of the AE-s is trained independently using the concentration map of the corresponding geochemical element as input, and the output of each CAE is the reconstructed concentration map of the corresponding element. By observing the change in the cost value during training, the suboptimal weights (w_1 and w'_1) of the AE-s and the spatial features of all elements are obtained.

Second, the AE-c is pretrained independently using the spatial features of elements extracted by the AE-s as the input, and the output is the reconstructed spatial features of elements. After training, the suboptimal weights (w_2 and w'_2) of the AE-c are obtained.

Third, the pretrained AE-s and AE-c are combined into the FCAE, and all the obtained suboptimal weights (i.e. w_1 , w_2 , w'_1 and w'_2) of the AE-s and AE-c are loaded into the FCAE. The FCAE is then trained as a whole (fine-tuned) until the entire model converges (Hinton and Salakhutdinov 2006). For the training, the concentration maps of the

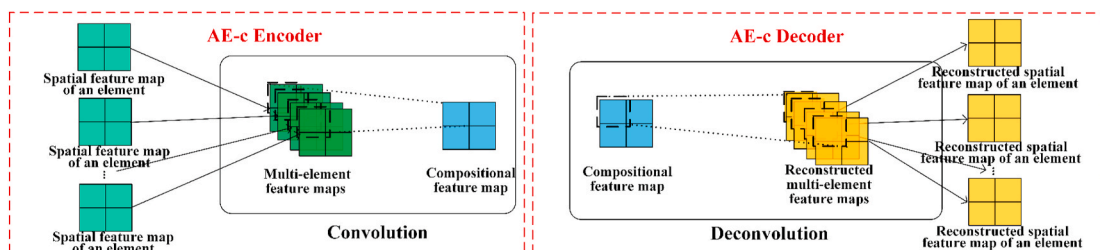


Fig. 5. Encoder and decoder of the AE-c.

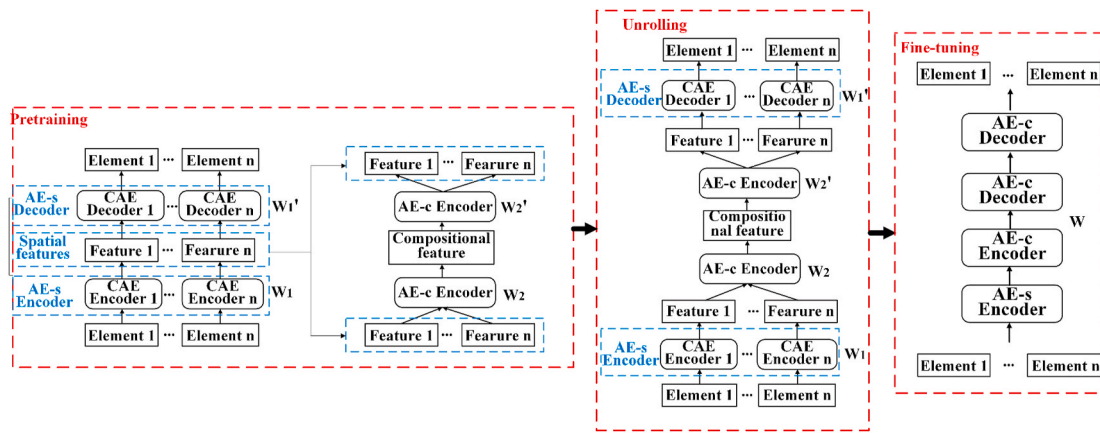


Fig. 6. Schematic diagram of the three-stage training (3ST) strategy of the FCAE.

geochemical elements are used as the input, and the output is the reconstructed concentration maps.

Such a 3ST strategy can help mitigate the vanishing gradient problem for complex neural network structures, as the hidden layers close to the input layer can be effectively updated during the pretraining of individual components and the entire model is then fine-tuned to ensure that all components to work together as a whole (Hinton and Salakhutdinov 2006; Erhan et al., 2010; Mohammadi and Kain 2014).

2.3. Anomaly score calculation

Reconstruction errors are often used to detect anomalies in many AE-based studies (Valentine and Trampert 2012; An and Cho 2015; Xiong and Zuo 2016). In this study, the multivariate Euclidean distance between the original sample and reconstructed sample is calculated as the reconstruction error (also termed anomaly score) for each sampling location (Chen et al., 2019a). The higher is the anomaly score is, the more anomalous the sample is.

2.4. Performance assessment

2.4.1. Receiver operating characteristic curve

To evaluate the performance of a geochemical anomaly detection method, the ratio of the detected mineral deposits to the total mineral deposits must be considered, as well as the area of the detected anomalous areas. Therefore, this study adopts the receiver operating characteristic (ROC) curve for the quantitative evaluation of the model (Fawcett 2006; Zuo 2017; Chen et al., 2019a). Specifically, a set of discerningly ordered thresholds of anomaly scores were determined according to the range of the anomaly scores. At a certain threshold, the study region is divided into anomalous regions and background regions. Subsequently, using the known deposits as the reference, at every threshold of anomaly score, the true positive rate (TPR) and the false positive rate (FPR) are calculated as follows:

$$TPR = \frac{\text{Number of mineral deposits in the anomalous region}}{\text{Number of known mineral deposits in the study region}} \quad (2)$$

$$FPR = \frac{\text{Nonmineral area within the anomalous region}}{\text{Total nonmineral area in study region}} \quad (3)$$

In geochemical anomaly detection, TPR refers to the proportion of the known deposits covered by anomalous regions to all known deposits in the entire study region. FPR refers to the proportion of nonmineral area in the anomaly region to the total nonmineral area of the entire study region. The larger the TPR value and the smaller the FPR value, the higher the overlap between the anomaly area circled by the model and the known deposits. To evaluate the performance of the model at all

thresholds, the ROC curve is plotted with the FRPs and TRPs along the X horizontal and Y vertical axes, respectively. The closer the area under the curve (AUC) is to 1, the greater the accuracy of the model (Bradley 1997; Fawcett 2006). The closest point on the ROC curve to the upper left corner ($x = 0, y = 1$), which represents the largest TPR and the smallest FPR, is used as the best segmentation threshold for dividing anomalies and backgrounds (Chen 2019). The ROC curve has been widely used in the evaluation of geochemical anomaly detection results (Zuo 2017; Zuo et al., 2019; Chen et al., 2019a).

2.4.2. Weights-of-evidence

The Weights-of-evidence (WofE) is a geostatistical method based on Bayesian probability theory, that evaluates the correlation between the anomaly maps and mineral deposits (Agterberg et al., 1993). This is calculated as follows:

$$W^+ = \ln \frac{\text{Area}(D \cap B) / \text{Area}(D)}{\text{Area}(D \cap B^c) / \text{Area}(D^c)} \quad (3)$$

$$W^- = \ln \frac{\text{Area}(D \cap B^c) / \text{Area}(D)}{\text{Area}(D^c \cap B^c) / \text{Area}(D^c)} \quad (4)$$

$$C = W^+ - W^- \quad (5)$$

where B and D indicate the anomalous region and known mine region, respectively. W^+ and W^- refer to the positive and negative weights calculated, respectively, from the anomaly map and known occurrences. The difference between W^+ and W^- is denoted as C , which represents the correlation between the anomaly map and the mine deposit (Zhang et al., 2013). The larger the value of C is, the more significant the correlation between the anomaly map and the mine deposit.

2.4.3. Prediction-area (P-A) plot

In addition to the AUC, a common machine learning metric, this study also uses the prediction-area (P-A) plot, a traditional metric of geochemical anomaly identification, to evaluate the performance of the model in terms of both anomaly area and prediction rate of mineral deposits (Yousefi and Carranza 2015a, 2015b; Li et al., 2020). In the P-A plot of an anomaly map, there are two curves, the curve of the prediction rate of known mineral deposits (P) and the curve of the percentage of the area of the anomaly region (A). The intersection point that is a common point on both the abovementioned curves is a criterion to evaluate and weight the whole anomaly map. The sum of the prediction rate and occupied area for the intersection point is equal to 100. If the intersection point is higher on the vertical axis, it indicates that the model identifies a smaller anomaly area and a higher accuracy rate. In other

words, the performance of the model that obtained this anomaly map is better (Agterberg and Bonham-Carter 2005; Yousefi and Carranza 2015a). This study normalizes the anomaly scores obtained by each model as (horizontal axis), and evaluates the performances by comparing the intersection locations.

3. Experiment and evaluation

3.1. Study area and data

To assess the performance, the FCAE was applied to the southwest area of the Wuyishan metallogenic belt to identify the geochemical anomalies related to copper ore. The Wuyishan metallogenic belt has long experienced the convergence of global supercontinents and the breakup of the northern and southern continents, and is endowed with preferential mineralization backgrounds and conditions (Ding et al., 2016). This great belt is located in the Cathaysia block. Several copper deposits have been found in the southwestern part of the metallogenic belt (i.e., the case study area, as shown in Fig. 7). The main types of copper deposits are continental volcanic copper-lead-zinc polymetallic deposits, skarn copper polymetallic deposits, and porphyry copper deposits (Qiu et al., 2010).

The stream sediment geochemical data were collected from 3484 sampling points in the study area at a scale of 1:200,000 by the Institute of Geophysical and Geochemical Exploration, Chinese Academy of Geological Sciences (IGGE). The dataset includes 32 elements and 6 oxides. All concentration values were normalized into the range [0, 1] using min-max normalization.

3.2. Selection of related elements

To identify the geochemical anomalies related to copper, the geochemical elements and oxides that are strongly related to copper were chosen as indicators of anomaly identification (Chen et al., 2014a). The Pearson correlation coefficient between copper and other

geochemical variables is calculated as follows Benesty et al. (2009). According to the correlation coefficient test table, two elements are significantly correlated if their correlation coefficient is greater than 0.104, when the significance level is 0.001 (Chen et al., 2014a). The results showed that silver, gold, zinc, arsenic, cadmium, antimony, tin, lead, phosphorus, and sodium oxide had high correlations with copper (Table 1). Therefore, these 11 geochemical variables were selected as the geochemical indicators of copper ore.

3.3. Application of the FCAE

3.3.1. Data preprocessing

Because the FCAE uses a convolutional layer to extract the spatial features, the input data of the model should be in the form of a matrix. However, the original multivariate geochemical data are points, and the concentration values are missing at some locations. To solve this problem, we interpolated the geochemical values of the sampling locations using the inverse distance weighting (IDW) method (Watson and Philip 1985) to generate concentration maps with spatial dimensions of 57×92 , at a spatial resolution of $1634.5 \text{ m} \times 1492.8 \text{ m}$ ($0.0147^\circ \times 0.0147^\circ$) (Fig. 8). Note that all concentration values were normalized into the range of [0, 1].

3.3.2. Parameter setting

In addition to the structure of the model, the setting of the hyperparameters is also important. The hyperparameters of the FCAE primarily include the size and number of convolution windows, which can directly affect the accuracy (Gu et al., 2018). The size of the convolution window (also termed the recognition domain in Chen et al. (2019a)) is related to the spatial structure of the data (Bellehumeur et al., 1994). To better learn the spatial structural features of the geochemical background, the number of geochemical background samples in the recognition domain needs to be much greater than the number of anomalous samples. Generally, the stability of the spatial structure can be seen as an indicator for the selection of the recognition domain. We selected a size

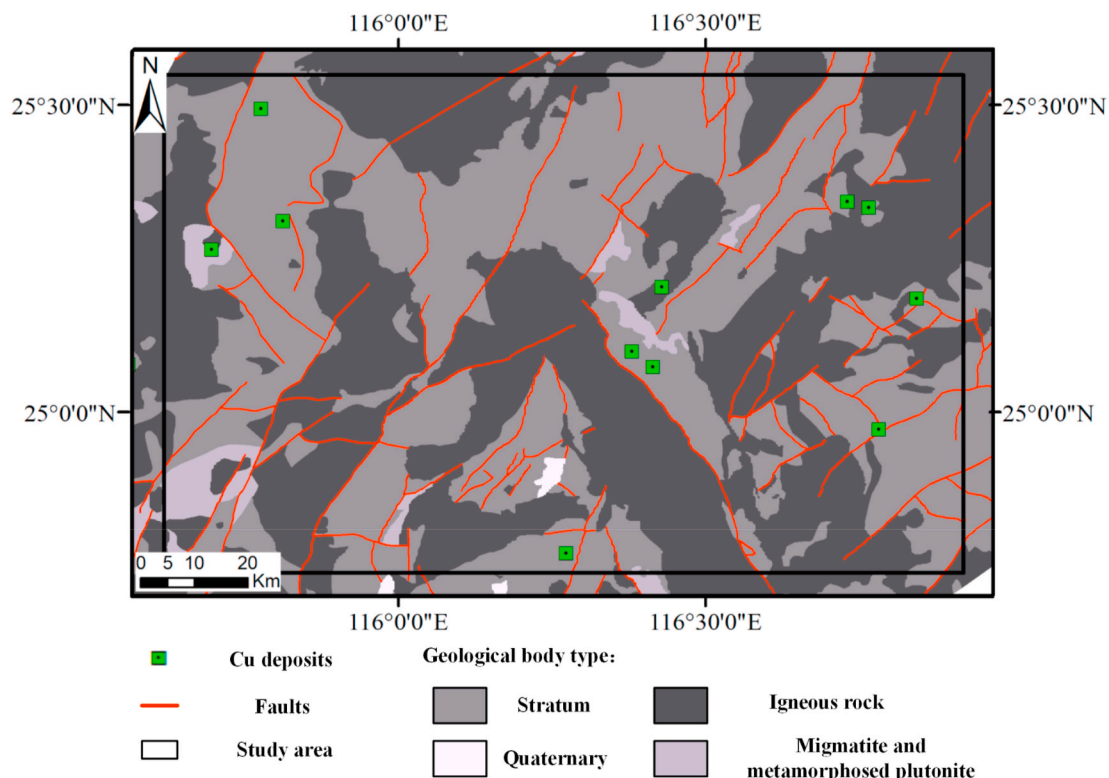


Fig. 7. Geological map of the study area (data from the Institute of Geophysical and Geochemical Exploration, Chinese Academy of Geological Sciences (IGGE)).

Table 1
Correlation coefficients between copper and the other 9 elements and compound.

	Ag	Au	Zn	As	Cd	Sb	Ti	Pb	Na ₂ O	P
Cu	0.742	0.531	0.525	0.434	0.379	0.353	0.244	0.242	-0.156	0.142

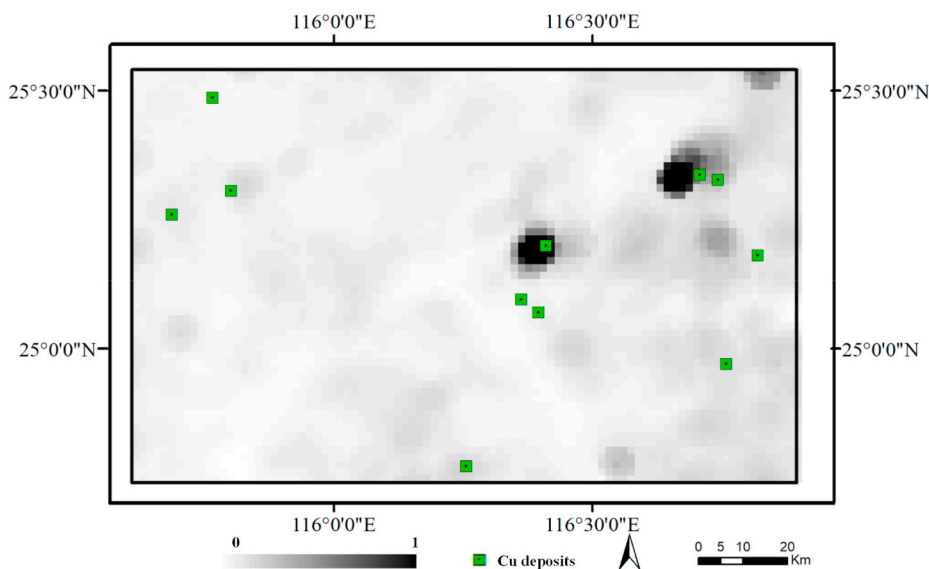


Fig. 8. A normalized concentration map after interpolation (taking Ag as an example).

of the domain for which the structure did not change significantly as the appropriate size of the convolution window (Chen et al., 2019a). In geochemical data analysis, global Moran's I, can be used to measure the spatial distribution patterns of the regional geochemical variables (Bin et al., 2017). The size of the convolution window can be determined using global Moran's I of the geochemical data.

As shown in Fig. 9, this study calculated global Moran's I values for different window sizes (i.e., the size of the window measured by the numbers of rows and columns of the grids). When the size of the window was greater than 17, the Moran's I value slowly decreased as the increasing size of the window. This indicates that the spatial structure in the recognition domain is relatively stable and that geochemical background samples predominate when the size of the domain is greater than 17. However, when the size of the window became larger (>29), Moran's I value started to approach 0. This indicates that the pattern of

variation in the background space (spatial heterogeneity) is difficult to capture, when the size of the window is larger than 29. Finally, we set the convolution size as 26.

The setting of the number of hidden layers is often based on experience and previous studies. Chen et al. (2019a) and Steck (2019) demonstrated that as the number of hidden layers increased, the reconstruction error of the model decreased gradually. However, the reduction of reconstruction error does not represent the accuracy of the background reconstruction. Conversely, overly deep structures learned unwanted features (i.e., the pattern of geochemical anomalies), which made the background reconstruction less accurate. Thus, the hidden layers of the decoder and encoder in the FCAE are set to 2. Furthermore, as suggested by previous studies, the number of convolution kernels is often a power of 2 (8, 16, 32, etc.) (Simonyan and Zisserman 2014; Chollet et al., 2015; Gu et al., 2018). To extract enough features without

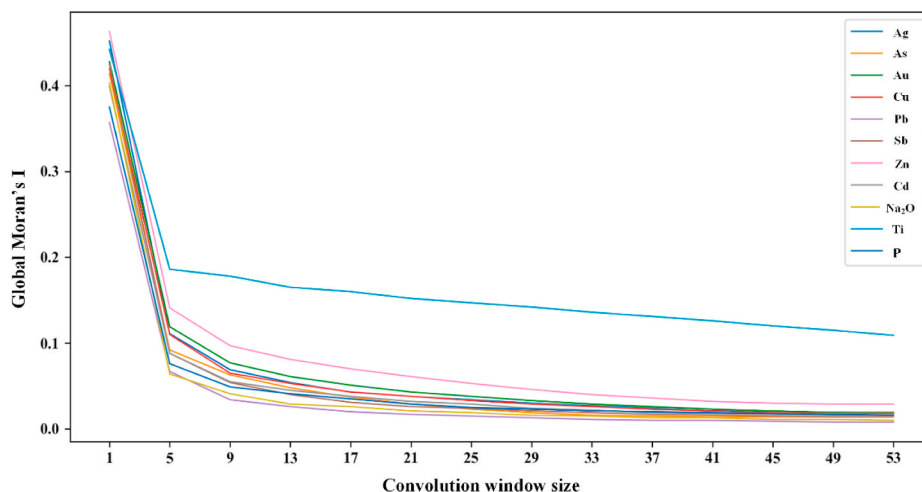


Fig. 9. Variation in global Moran's I index with different convolution window sizes.

overfitting, the number of convolution kernels was set to 16.

3.3.3. FCAE training

The FCAE was trained using the 3ST strategy that combines greedy layerwise pretraining and overall fine-tuning, with the input data of 11 geochemical indicators. Each element corresponds to a separate input layer. A stable cost value meant that the model was considered relatively well-trained, and then we loaded the weights of pretrained models into the overall fine-tuning. Fig. 10 shows the effects of two training strategies, i.e., 3ST and direct training (without pretraining), on the FCAE model. The model trained by 3ST has faster convergence and less fluctuation than the model without pretraining. The cost value of the 3ST is also lower, which indicates that the model has a better capability of extracting and reconstructing the multivariate geochemical background.

3.3.4. Anomalies map generation

After training, the FCAE generated the reconstructed concentration maps of geochemical elements. As mentioned before, such reconstructed maps can be seen as the backgrounds of elements, because the FCAE had learned the spatial structural features and compositional relationships of elements embedded in the majority of the samples (i.e., background samples), while the anomalous samples had little impact on the reconstructed maps. Subsequently, the multivariate Euclidean distance (i.e., the anomaly score) between the original concentrations of the elements and the reconstructed concentrations was calculated for each sample. Finally, the anomaly map was generated based on the anomaly scores (Fig. 11). However, limited by the lack of ore point labels during training, the FCAE can only cause each sample to obtain a continuous anomalous score, and does not directly predict that it is a mineral deposit or background. Such a limitation results in the mineral and non-mineral areas not being directly distinguished, but rather causing a tendency from high to low scores in the anomaly map. As shown in Fig. 11, anomalous high values also appear in the area around the resultant Cu deposits. This is a common problem for geochemical anomaly identification using unsupervised methods.

3.4. Performance evaluation and uncertainty analysis

3.4.1. Performance evaluation

To quantitatively evaluate the performance of the FCAE, the AUC, P-A plot and WofE were calculated using the known Cu deposit sites in the study area. In addition, considering the difference in the problem solved by unsupervised and supervised models, three unsupervised models, the Mahalanobis distance (MAHAL) (Chen et al., 2014b), deep autoencoder

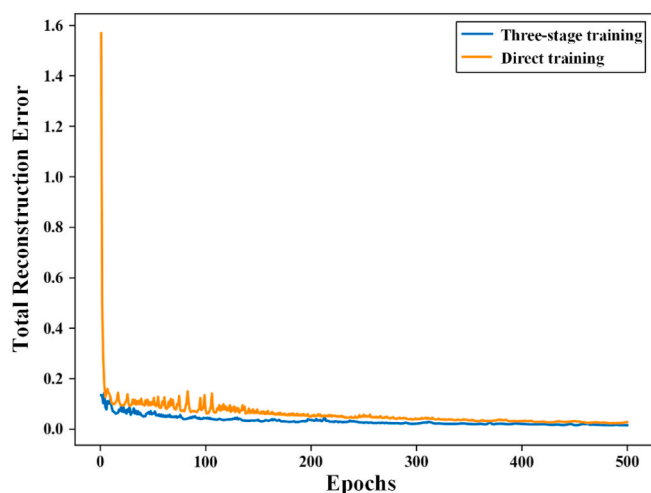


Fig. 10. Variation in the loss function (cost function, i.e., MSE) with the number of training epochs.

network (DAN) (Xiong and Zuo 2016), and multiconvolutational autoencoder (MCAE) (Chen et al., 2019a), were used as comparison models. The MAHAL is the baseline model in geochemical anomaly identification, which is widely used in comparison experiments. The DAN and MCAE are the latest proposed deep learning models, and have been shown to be superior to shallow machine learning models for anomaly detection (Zuo 2017; Zuo et al., 2019; Chen 2019). Note, that in this experiment, both the MAHAL based on copper elements (MAHAL-Cu) and the MAHAL based on the 11 related geochemical elements (MAHAL-M) were calculated.

As shown in Fig. 12, the AUC of the FCAE is significantly higher than that of the other methods, which indicates that the FCAE performed the best for geochemical anomaly detection for Cu mineralization. In particular, when the prediction rate of all models is up to 80%, only the FCAE percentage of mineral-free with anomaly regions to study mineral-free area is less than 20%. Additionally, based on the vertical values of the left and right axes corresponding to the intersection points in the P-A plot, the FCAE model obtains a higher position of the intersection points, which means that the anomalous area obtained by the FCAE model not only covers more known mineral deposits, but also has a smaller anomaly area.

To further evaluate the performance of the model, we selected the best threshold based on the ROC (i.e., the closest point on the ROC curve to the upper left corner, see section 2.4.1) to segment the anomaly and background regions (Fig. 13). The recall rate (the ratio of copper deposits located in the anomalous region to all known copper deposits) and the correlation C-value WofE from each model are shown in Table 2. The FCAE achieved the highest values (recall = 0.909, C = 3.586) among all methods.

The recall rate of the MAHAL-M is higher (recall = 0.818) than that of the MAHAL-Cu, and the anomaly region identified by the MAHAL-M is closer to the known Cu deposit (Fig. 14B). The main reason for this result is that multivariate data contain not only more information related to mineralization, but also redundant information. Simple statistical methods can extract only linear features, and cannot extract key details from multivariate data. This causes the area of the anomalous region identified based on the 11 related geochemical elements to be larger than that identified by a single Cu element (Timme et al., 2014).

By learning either the compositional relationships among elements or the spatial structural features of elements, the DAN and MCAE are able to identify some geochemical anomalies related to copper deposits. The DAN had a better recognition result for the copper deposits in the northwest corner (Fig. 13C), while the MCAE had a better recognition result for the copper deposits in the central area (Fig. 13D). By combining both the spatial structural features and compositional relationships, the FCAE outperformed other methods, and identified most of the known Cu deposits (10 out of 11) in the study area (Fig. 13E).

3.4.2. Sensitivity analysis on convolution window size

The convolutional layers of the FCAE can extract the spatial structural features of geochemical exploration data and improve anomaly recognition related to mineralization. The size of the convolution window is a key parameter that may greatly affect the performance (Gu et al., 2018). In many studies that use convolution neural networks, the size of the convolution window was set by empirical experience. Therefore, the sensitivity of the model to the convolution window size must be considered in real applications. In this study, the MCAE model and FCAE model with different sizes of convolution windows were divided into eleven groups to compare and analyze the sensitivity. To better simulate the real applications, each group model was trained 10 times with random initial weights.

When the size of the convolution window was the same, the highest AUC value of the FCAE was higher than that of the MCAE (Fig. 14 A and B). Compared with MCAE, the fluctuation range of the AUC values of the FCAE of each group was smaller. Furthermore, the mean AUC and mean recall rate of FCAE in each group were also greater than those of the

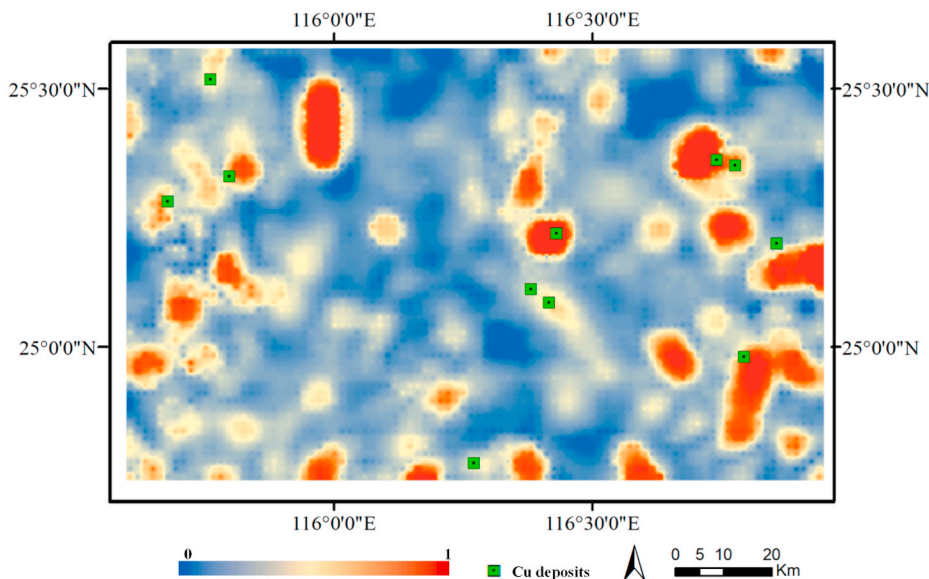


Fig. 11. Anomaly map generated by the FCAE. For better visualization, the anomaly scores are normalized into the range of [0,1].

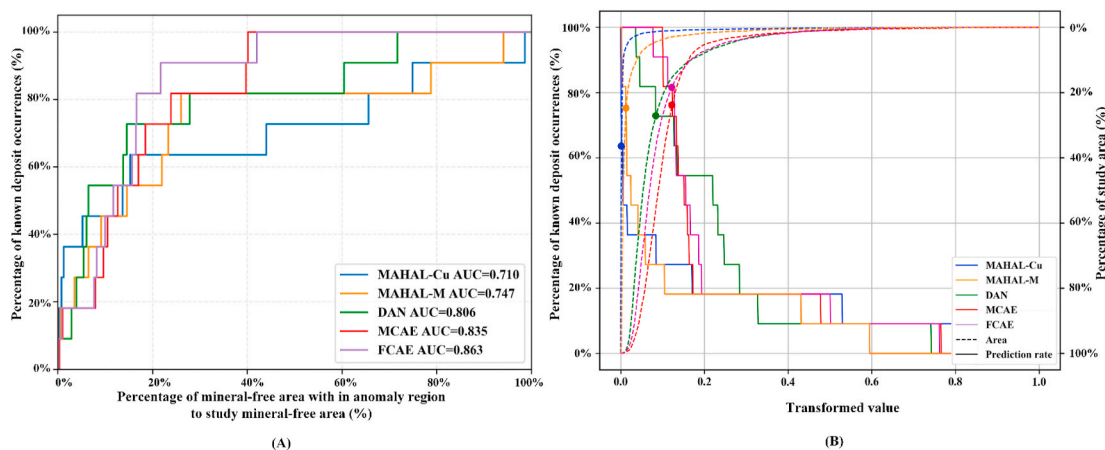


Fig. 12. Receiver operating characteristics (ROC) curves and prediction-area (P-A) plots of various approaches. MAHAL-Cu: Mahalanobis distance with Cu; MAHAL-M: Mahalanobis distance with multiple related elements; DAN: deep autoencoder network; MCAE: multiconvolutinal autoencoder; FCAE: feature fusion convolutional autoencoder.

MCAE in each group (Fig. 14 C and D). These results indicate that the FCAE is less sensitive to the size of the convolution window than the MCAE, which is more conducive for real-world applications.

4. Discussion

In contrast with the existing deep learning models, the FCAE solves the problem of separation of spatial structural features and composition relationships of geochemical variables. The FCAE can extract and fuse both spatial structural features and composition relationships to identify geochemical anomalies without labeled geochemical data. In the case study, the FCAE achieved the highest accuracy in all evaluation indicators, compared with other methods. The maximum AUC of the FCAE reached 0.863, with the recall rate of known Cu deposits in the geochemical anomaly region identified by the FCAE reaching 0.909. Such a result shows that the FCAE model has better anomaly detection capabilities.

It must be noted that unsupervised learning methods, including the FCAE, are based on the premise that geochemical anomalies only take a small proportion of the whole study area and that the majority of samples represent the geochemical background. Since previous studies have

suggested that geochemical anomalies often cover 1.5%–5% of the total area (Chen et al. 2009, 2014a), the FCAE and other unsupervised methods can be applied in most cases. However, we still need to clarify that the AE is counterproductive in areas where mineral deposits are particularly rich and occupy the majority of the area.

This study demonstrates that it is advantageous to integrate spatial structural features and component relationships for geochemical anomaly recognition. In future studies on geochemical anomaly identification, more methods of multifeature fusion should be considered. In this study, we use the stack and multitask mechanism for feature fusion, which not only increases the generalization ability of the model but also enables nonlinear fusion between different features (Evgeniou and Pontil 2004; Zhou and Paffenroth 2017). However, there are many ways to integrate models and features in machine learning (e.g., multimodel averaging, staking, dynamic classifier selection and sequential combination) (Zhao et al. 2019, 2020). How to better fuse multiple features based on the characteristics of geochemical data is also worthy of further study.

This study also analyzed the model’s sensitivity to parameters (i.e., the convolution window size for the CAE). The results suggest that the FCAE is less sensitive to the size of the convolution window than the

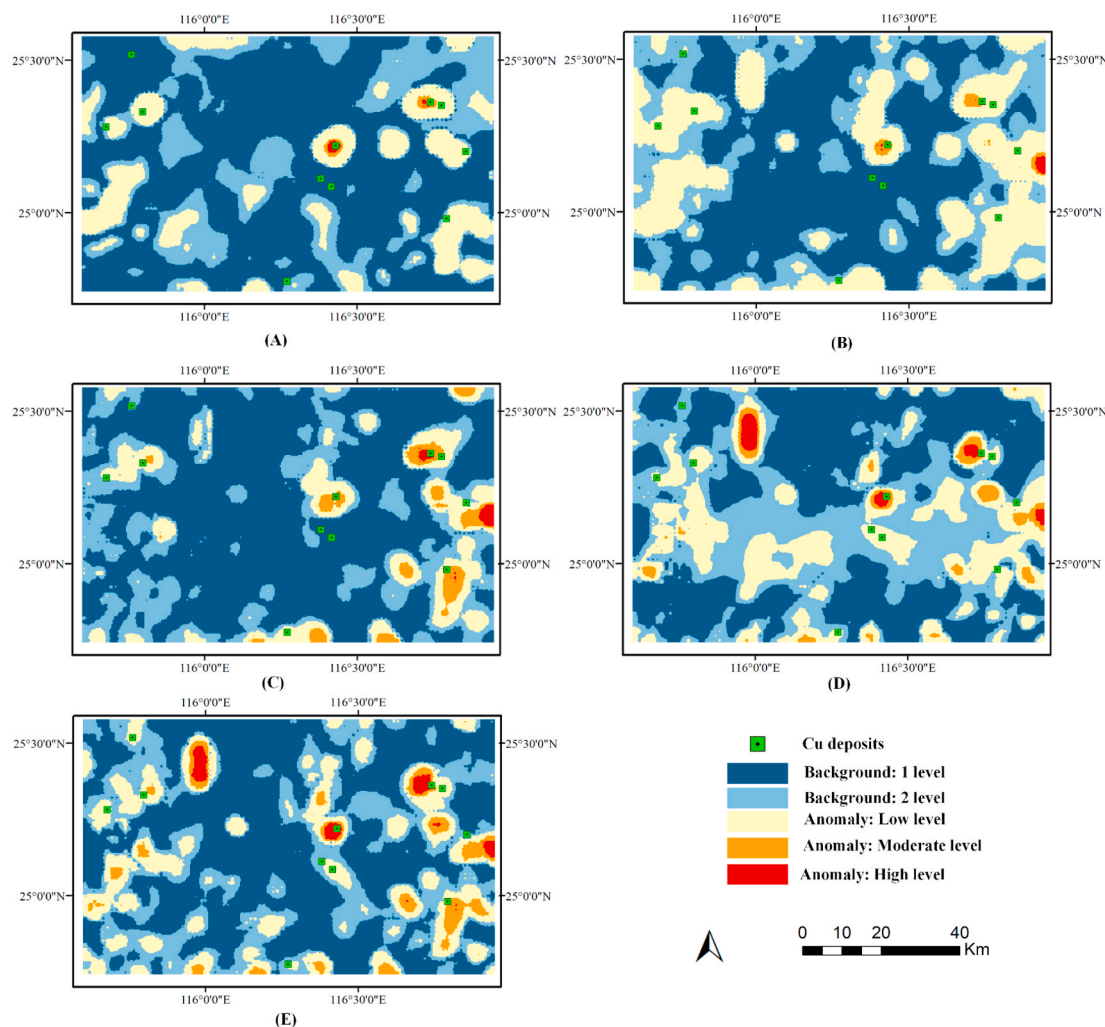


Fig. 13. Anomaly recognition results of each model based on the optimal threshold obtained by the ROC curve (A) MAHAL-Cu: Mahalanobis distance with Cu; (B) MAHAL-M: Mahalanobis Distance with multiple related elements; (C) DAN: deep autoencoder network; (D) MCAE: multiconvolutinal autoencoder; (E) FCAE: feature fusion convolutional autoencoder.

Table 2

Performance indicators of the various methods. MAHAL-Cu: Mahalanobis distance with Cu; MAHAL-M: Mahalanobis distance with multiple related elements; DAN: deep autoencoder network; MCAE: multiconvolutinal autoencoder; FCAE: feature fusion convolutional autoencoder.

Model	AUC	Recall	C
MAHAL-Cu	0.710	0.636	2.276
MAHAL-M	0.747	0.818	2.549
DAN	0.806	0.727	2.753
MCAE	0.835	0.818	2.666
FCAE	0.863	0.909	3.586

MCAE, which is more conducive in real applications. The MCAE only extracts the spatial structural features to identify geochemical anomalies, and is more sensitive to the size of the convolution window. In contrast, the FCAE fuses the spatial structural features and composition relationships for multivariate geochemical anomaly recognition. The compositional relationships are extracted from multiple geochemical variables at the sampling locations without considering the feature extraction window size (Xiong and Zuo 2016). Therefore, the FCAE is less sensitive to the size of the convolution window. Previous studies have shown that the combined use of multi-sized convolution kernels is beneficial to spatial structural feature extraction and reduces the parameter sensitivity. In the future, we will refer to the structure of

ResNet (He et al., 2016) and use multiple repetitive small convolution kernels instead of a large convolution kernel. The features obtained by each convolutional layer will be fused to further reduce the model's sensitivity to the size of the convolution window.

This study still has many limitations and opportunities for future studies. First, the FCAE uses convolutional layers for extracting spatial features, which can be used only when the input data are in regular grids (Zhou et al., 2018). However, most study areas in the natural environment have irregular shapes, which limits the application of the model (Zhu et al., 2007). Second, this study did not consider other data sources that may also provide information for anomaly recognition (e.g., remote sensing data and geophysical data). All these limitations require further study.

5. Conclusions

The spatial structural features and composition relationships of geochemical elements is the synthesis of various geological processes. This study proposes a feature fusion convolutional autoencoder (FCAE) for multivariate geochemical anomaly recognition. FCAE extracts and fuses both the spatial structural features and composition relationships of geochemical variables to identify the geochemical anomalies accurately. It provides a new method for the identification of geochemical anomalies through deep learning. We designed a case study to

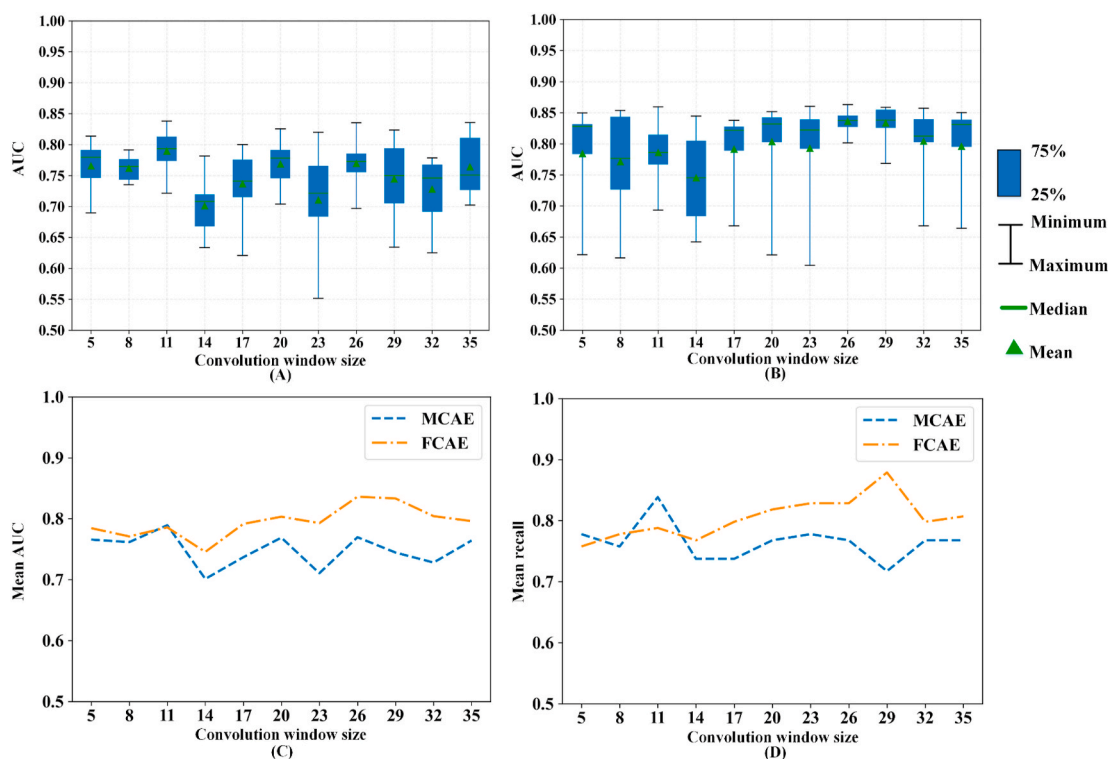


Fig. 14. Accuracies of the MCAE and FCAE with different convolution window sizes. (A) and (B) are the AUC box plots of the MCAE and FCAE with different convolution window sizes, respectively; (C) and (D) are the mean AUC and mean recall rate of the MCAE and FCAE with different convolution window sizes, respectively.

demonstrate the performance of the FCAE, in the southwest area of the Wuyishan Cu-Pb-Zn polymetallic metallogenic belt. The FCAE achieved the highest accuracy in all indicators (AUC = 0.863 and recall = 0.909) compared with a few existing methods, proving that integrating both spatial structural characteristics and composition relationships helps identify geochemical anomalies. Compared to the MCAE, the FCAE is less sensitive to the size of the convolutional window, which improves its ease of use in real-world applications.

Funding

This work was supported by the National Key Research and Development Program of China [2019YFB2102903, 2016YFC0600508], the National Natural Science Foundation of China [41801306, U1711267].

Authorship statement

Yao Yao provided the idea and designed the experimental procedure. Qingfeng Guan and Shuliang Ren completed the experiment and wrote the paper. Lirong Chen revised the paper. Lirong Chen and Bin Feng provided the data needed for the paper.

Computer code availability

The codes based on Python for the FCAE are available at <https://github.com/whuyao/FCAE>.

Declaration of competing interest

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

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.cageo.2021.104890>.

References

- Agterberg, F.P., Bonham-Carter, G.F., 2005. Measuring the performance of mineral-potential maps. *Nat. Resour. Res.* 14 (1), 1–17.
- Agterberg, F.P., Bonham-Carter, G.F., Cheng, Q., Wright, D.F., 1993. Weights of evidence modeling and weighted logistic regression for mineral potential mapping. *Comput. Geol.* 25, 13–32.
- An, J., Cho, S., 2015. Variational autoencoder based anomaly detection using reconstruction probability. *Spl. Lect. IE 2* (1), 1–18.
- Bellehumeur, C., Marcotte, D., J E Brak, M., 1994. Multi-element relationships and spatial structures of regional geochemical data from stream sediments, southwestern Quebec, Canada. *J. Geochem. Explor.* 51 (1), 11–35.
- Benesty, J., Chen, J., Huang, Y., Cohen, I., 2009. Pearson correlation coefficient. In: *Noise Reduction in Speech Processing*. Springer, Berlin, Heidelberg, pp. 1–4.
- Bengio, Y., 2009. Learning deep architectures for AI. *Found. Trends Mach. Learn.* 2 (1), 1–127.
- Bengio, Y., Lamblin, P., Popovici, D., Larochelle, H., 2007. Greedy layer-wise training of deep networks. *Adv. Neural Inf. Process. Syst.* 19, 153–160.
- Beus, A.A., Grigorian, S.V., 1977. *Geochemical Exploration Methods for Mineral Deposits*. Applied Pub. 1977.
- Bin, J.L., Zhou, T., Yuan, F., Zhang, D., Liu, L., Liu, G., 2017. A method for identifying geochemical anomalies based on spatial autocorrelation. *Sci. Surv. Mapp.* 42, 24–27.
- Bradley, A.P., 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recogn.* 30 (7), 1145–1159.
- Chen, J., Sathe, S., Aggarwal, C., Turaga, D., 2017. Outlier detection with autoencoder ensembles. In: *Proceedings of the 2017 SIAM International Conference on Data Mining Society for Industrial and Applied Mathematics*, pp. 90–98.
- Chen, L., 2019. *Multivariate Geochemical Anomaly Recognition Using Spatial Constrained Autoencoders*. China University of Geosciences.
- Chen, L., Guan, Q., Feng, B., Yue, H., Wang, J., Zhang, F., 2019a. A multi-convolutional autoencoder approach to multivariate geochemical anomaly recognition. *Minerals* 9 (5), 270.
- Chen, L., Guan, Q., Xiong, Y., Liang, J., Wang, Y., Xu, Y., 2019b. A Spatially Constrained Multi-Autoencoder approach for multivariate geochemical anomaly recognition. *Comput. Geosci.* 125, 43–54.
- Chen, Y., Lu, L., Li, X., 2014a. Application of continuous restricted Boltzmann machine to identify multivariate geochemical anomaly. *J. Geochem. Explor.* 140, 56–63.

- Chen, Y.L., Lu, L.J., Li, X.B., 2014b. Kernel Mahalanobis distance for multivariate geochemical anomaly recognition. *J. Jilin Univ.: Earth Sci. Ed.* 44, 396–408.
- Chen, Z.J., Cheng, Q.M., Chen, J.G., 2009. Comparison of different models for anomaly recognition of geochemical data by using eample ranking method. *Earth Sci. J. China Univ. Geosci.* 34 (2), 353–364, 2009.
- 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 (1–2), 314–324.
- Cheng, Q., Agterberg, F.P., 1996. Multifractal modeling and spatial statistics. *Math. Geol.* 28 (1), 1–16.
- Cheng, Q., Agterberg, F.P., Ballantyne, S.B., 1994. The separation of geochemical anomalies from background by fractal methods. *J. Geochem. Explor.* 51 (2), 109–130.
- Chollet, F.C.C.O., Others, 2015. *Keras*. GitHub.
- Cracknell, M.J., Reading, A.M., 2014. Geological mapping using remote sensing data: a comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. *Comput. Geosci.* 63, 22–33.
- Debaille, V., Blichert-Toft, J., Agraniar, A., Doucelance, R.E.G., Schiano, P., Albarede, F., 2006. Geochemical component relationships in MORB from the mid-Atlantic ridge, 22–35 N. *Earth Planet Sci. Lett.* 241 (3–4), 844–862.
- Ding, J., Fan, J., Yin, J., Liu, Y., 2016. Geological characteristics and mineral resource potential of the Wuyishan Cu-Pb-Zn polymetallic metallogenic belt. *Acta Geol. Sin.* 7, 1537–1555.
- Erhan, D., Courville, A., Bengio, Y., Vincent, P., 2010. Why does unsupervised pre-training help deep learning?. In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, pp. 201–208.
- Evgeniou, T., Pontil, M., 2004. Regularized multi-task learning. In: *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 109–117.
- Fawcett, T., 2006. An introduction to ROC analysis. *Pattern Recogn. Lett.* 27 (8), 861–874.
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., Others, 2018. Recent advances in convolutional neural networks. *Pattern Recogn.* 77, 354–377.
- Harris, J.R., Wilkinson, L., Grunsky, E.C., 2000. Effective use and interpretation of lithochemical data in regional mineral exploration programs: application of Geographic Information Systems (GIS) technology. *Ore Geol. Rev.* 16 (3–4), 107–143.
- Hawkes, H.E., Webb, J.S., 1963. Geochemistry in mineral exploration. *Soil Sci.* 95 (4), 283.
- He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778.
- Hinton, G.E., Salakhutdinov, R.R., 2006. Reducing the dimensionality of data with neural networks. *Science* 313 (5786), 504–507.
- Kürzl, H., 1988. Exploratory data analysis: recent advances for the interpretation of geochemical data. *J. Geochem. Explor.* 30 (1–3), 309–322.
- Krawczyk, B., 2016. Learning from imbalanced data: open challenges and future directions. *Progr. Artif. Intell.* 5 (4), 221–232.
- Larochelle, H., Bengio, Y., Louradour, J.E.R.O., Lamblin, P., 2009. Exploring strategies for training deep neural networks. *J. Mach. Learn. Res.* 10 (Jan), 1–40.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521 (7553), 436–444.
- Leybourne, M.I., Cameron, E.M., 2010. Groundwater in Geochemical Exploration.
- Li, S., Chen, J., Xiang, J., 2019. Applications of deep convolutional neural networks in prospecting prediction based on two-dimensional geological big data. *Neural Comput. Appl.* 1–17.
- Li, T., Xia, Q., Zhao, M., Gui, Z., Leng, S., 2020. Prospectivity mapping for Tungsten polymetallic mineral resources, Nanling metallogenic belt, South China: use of random forest algorithm from a perspective of data imbalance. *Nat. Resour. Res.* 29 (1), 203–227.
- Lu, L., Shin, Y., Su, Y., Karniadakis, G.E., 2019. Dying Relu and Initialization: Theory and Numerical Examples arXiv preprint arXiv:1903.06733.
- Masci, J., Meier, U., Cire C S An, D., Schmidhuber, J.U.R., 2011. Stacked convolutional auto-encoders for hierarchical feature extraction. In: *International Conference on Artificial Neural Networks*. Springer, pp. 52–59.
- Mohammadi, S.H., Kain, A., 2014. Voice conversion using deep neural networks with speaker-independent pre-training. In: *2014. IEEE Spoken Language Technology Workshop (SLT)*, pp. 19–23.
- Nair, V., Hinton, G.E., 2010. Rectified linear units improve restricted Boltzmann machines. In: *International Conference on International Conference on Machine Learning*. Omnipress, pp. 807–814.
- Nickerson, A., Japkowicz, N., Milios, E.E., 2001. In: *Using Unsupervised Learning to Guide Resampling in Imbalanced Data Sets*. AISTATS.
- Noh, H., Hong, S., Han, B., 2015. Learning deconvolution network for semantic segmentation. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1520–1528.
- Qiu, X., Lan, Y., Liu, Y., 2010. The key to the study of deep mineralization and the evaluation of ore-prospecting potential in the Zijinshan gold and copper deposit. *Diqiu Xuebao* 31 (2), 209–215.
- Ribeiro, M.E.S., Lazzaretti, A.E.E.E., Lopes, H.S.E.R., 2018. A study of deep convolutional auto-encoders for anomaly detection in videos. *Pattern Recogn. Lett.* 105, 13–22.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. *Nature* 323 (6088), 533–536.
- Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *Comput. Sci.* arXiv preprint arXiv:1409.1556.
- Steck, H., 2019. Embarrassingly shallow autoencoders for sparse data. In: *The World Wide Web Conference*, pp. 3251–3257.
- Sun, J., Steinecker, A., Glocker, P., 2014. Application of deep belief networks for precision mechanism quality inspection. In: *International Precision Assembly Seminar*. Springer, pp. 87–93.
- Timme, N., Alford, W., Flecker, B., Beggs, J.M., 2014. Synergy, redundancy, and multivariate information measures: an experimentalist's perspective. *J. Comput. Neurosci.* 36 (2), 119–140.
- Twarakavi, N.K., Misra, D., Bandopadhyay, S., 2006. Prediction of arsenic in bedrock derived stream sediments at a gold mine site under conditions of sparse data. *Nat. Resour. Res.* 15 (1), 15–26.
- Valentini, A.P., Trampert, J., 2012. Data space reduction, quality assessment and searching of seismograms: autoencoder networks for waveform data. *Geophys. J. Int.* 189 (2), 1183–1202.
- Watson, D.F., Philip, G.M., 1985. A refinement of inverse distance weighted interpolation. *Geo Process.* 2 (4), 315–327.
- Xie, X., Ren, T., 1993. National geochemical mapping and environmental geochemistry—progress in China. *J. Geochem. Explor.* 49 (1–2), 15–34.
- Xiong, Y., Zuo, R., 2016. Recognition of geochemical anomalies using a deep autoencoder network. *Comput. Geosci.* 86, 75–82.
- Xiong, Y., Zuo, R., 2020. Recognizing multivariate geochemical anomalies for mineral exploration by combining deep learning and one-class support vector machine. *Comput. Geosci.* 140, 104484.
- Yousefi, M., Carranza, E., 2015a. Prediction-area (P-A) plot and C-A fractal analysis to classify and evaluate evidential maps for mineral prospectivity modeling. *Comput. Geosci.* 79, 69–81.
- Yousefi, M., Carranza, E.J.M., 2015b. Fuzzification of continuous-value spatial evidence for mineral prospectivity mapping. *Comput. Geosci.* 74, 97–109.
- Zeiler, M.D., 2012. Adadelata: An Adaptive Learning Rate Method arXiv preprint arXiv: 1212.5701.
- Zhang, D., Cheng, Q., Zuo, R., 2013. A comparison of two different unit division methods in weights of evidence. *J. Jilin Univ.* 43, 1040–1051.
- 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.
- Zhao, Y., Nasrullah, Z., Hryniewicki, M.K., Li, Z., 2019. LSCP: locally selective combination in parallel outlier ensembles. In: *Proceedings of the 2019 SIAM International Conference on Data Mining*. Society for Industrial and Applied Mathematics, pp. 585–593.
- Zhao, Y., Wang, X., Cheng, C., Ding, X., 2020. Combining machine learning models and scores using combo library. In: *Proceedings of the AAAI Conference on Artificial Intelligence*, New York, USA, vol. 34, pp. 13648–13649, 9.
- Zhou, C., Paffenroth, R.C., 2017. Anomaly detection with robust deep autoencoders. In: *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 665–674.
- Zhou, J., Cui, G., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., Sun, M., 2018. Graph Neural Networks: A Review of Methods and Applications arXiv preprint arXiv: 1812.08434.
- Zhu, Y., Xiao, K., Song, G., Yan, S., Chen, Z., Mei, Y., Li, C., Liu, Y., Wang, M., 2007. Geological Feature and Metallogenic Pedigree of Ore Deposits in the Major Metallogenic Regions (Belts) in China. Geological Publishing House, Beijing.
- Zuo, R., 2011. Identifying geochemical anomalies associated with Cu and Pb-Zn skarn mineralization using principal component analysis and spectrum-area fractal modeling in the Gangdese Belt, Tibet (China). *J. Geochem. Explor.* 111 (1–2), 13–22.
- Zuo, R., 2017. Machine learning of mineralization-related geochemical anomalies: a review of potential methods. *Nat. Resour. Res.* 26 (4), 457–464.
- 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., Cheng, Q., Agterberg, F.P., Xia, Q., 2009. Application of singularity mapping technique to identify local anomalies using stream sediment geochemical data, a case study from Gangdese, Tibet, western China. *J. Geochem. Explor.* 101 (3), 225–235.
- 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.