Natural Resources Research (© 2022)
https://doi.org/10.1007/s11053-022-10088-x

Original Paper

Recognizing Multivariate Geochemical Anomalies Related to Mineralization by Using Deep Unsupervised Graph Learning

Qingfeng Guan,[1,2] Shuliang Ren,[1] Lirong Chen,[3] Yao Yao,[1,6,7] Ying Hu,[1] Ruifan Wang,[1] Bin Feng,[4] Landing Gu,[3] and Wenhui Chen[5]

Received 3 September 2021; accepted 6 June 2022

The spatial structure of geochemical patterns is influenced by various geological processes, one of which may be mineralization. Thus, analysis of spatial geochemical patterns facilitates understanding of regional metallogenic mechanisms and recognition of geochemical anomalies related to mineralization. Convolutional neural networks (CNNs) used in previous studies to extract spatial features require regular data (e.g., raster maps) as input. Due to the complex and diverse geological environment, geochemical samples are inevitably irregularly distributed and even partially missing in many spaces, leading to the inapplicability of CNN-based methods for geochemical anomaly identification. Also, interpolation from samples to regular grids often introduces uncertainties. To address these problems, this study innovatively transformed geochemical sampled point data into graphs and introduced graph learning to extract the geochemical patterns. Correspondingly, a novel framework of geochemical identification named GAUGE (recognition of Geochemical Anomalies Using Graph lEarning) is proposed. To assess the performance of the proposed method, this study recognized anomalies related to Au deposits in the Longyan area, the Wuyishan polymetallic metallogenic belt, China. For a set of regularly distributed samples, GAUGE achieved an accuracy similar to that of a traditional convolution autoencoder. More importantly, GAUGE achieved an area under the curve of 0.833, outperforming one-class support vector machine, isolation forest, autoencoder, and deep autoencoder network for a set of irregularly distributed samples by 10.6, 5.2, 4.8, and 2.5%, respectively. By introducing graph learning into geochemical anomaly recognition, this study provides a new perspective of extracting both spatial structure and compositional relationships of multivariate geochemical patterns, which can be applied directly to irregularly distributed samples in irregularly shaped regions without the need for interpolation. Such an improvement greatly enhances the applicability of machine learning methods in geochemical anomaly recognition, providing support for mineral resources evaluation and exploration.

KEY WORDS: Geochemical anomaly recognition, Graph learning, Unsupervised learning, Global Moran's I, Graph attention network, Autoencoder.

[1]School of Geography and Information Engineering, China University of Geosciences, Wuhan, Hubei, China.
[2]State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Wuhan, Hubei, China.
[3]Development and Research Center, China Geological Survey, Xicheng District, Beijing, China.
[4]Institute of Geophysical and Geochemical Exploration, Chinese Academy of Geological Sciences, Langfang, Hebei, China.
[5]College of Mathematics and Computer Science, Zhejiang A and F University, Hangzhou, Zhejiang, China.
[6]Alibaba Group, Hangzhou 311121, Zhejiang Province, China.
[7]To whom correspondence should be addressed; e-mail: yaoy@cug.edu.cn

Published online: 24 June 2022

INTRODUCTION

Mineral resources continue to support economic developments at the intersection of the industrial civilization age and the information civilization age (Brooks and Andrews, 1974; Christmann, 2018). As the contradiction between the supply and demand of mineral resources becomes acute, finding mineral deposits becomes crucial. Geochemical exploration methods are used widely due to their low costs and rapid implementation (Beus and Grigorian, 1977; Chao, 1984). Geochemical exploration aims to map geochemical patterns and geochemical anomalies by analyzing the distribution and variations in elements in soil, rocks, and stream sediments (Beus and Grigorian, 1977; Cameron, 2005). Mineral deposits are formed by element accumulation due to certain geological processes, and geochemical anomalies are often found at and around mineral deposits that are inconsistent with overall patterns (Zhao, 2002). Identifying these anomalies can narrow the target areas for mineral exploration and improve prospecting efficiency. The key to finding geochemical anomalies is distinguishing between geochemical background and anomalies. Geochemical background refers to samples that conform to a certain pattern, while anomalies are those that deviate significantly from this pattern (Matschullat et al., 2000). In general, affected by complex geological processes, geochemical patterns are reflected mainly by compositional relationships and spatial distribution of geochemical variables. Therefore, extracting geochemical background patterns from these two perspectives can provide the critical foundation for distinguishing between background and anomalies.

The methods for identification of anomalous geochemical patterns can be divided into traditional methods and machine learning (deep learning) methods. Considering the influence of geological processes on the compositional relationships among multiple geochemical variables, traditional methods often use multivariate statistical methods (e.g., principal component analysis (PCA) (Wold et al., 1987), factor analysis (FA), and cluster analysis (Fabrigar and Wegener, 2011)) to analyze relationships among geochemical variables, determine the optimal combination of variables, and identify anomalies through clustering. PCA and FA are the most commonly used multivariate statistical methods. Additionally, to extract the spatial structure of geochemical patterns, spatial statistics and geographical weighting are among the introduced traditional methods. For example, considering spatial constraints on geological characteristics on geochemical data, Cheng et al., (2011) used spatially weighted PCA to analyze multivariate geochemical variables; this type of PCA provides more information on anisotropic spatial patterns. Zuo and Xiong, (2020) used a spatial statistical method (i.e., Moran′s I) to quantify spatial patterns (including clusters of high values (high–high), clusters of low values (low–low), high outliers (high–low), or low outliers (low–high)) of samples, and to find anomalous samples that differed from the surrounding background. Fractal/multifractal models have also been applied in exploring spatial relationships in geochemical data. Considering the spatial variability, geometric attributes, and scale invariance of geochemical data, many fractal models (e.g., concentration-area (Cheng et al., 1994), spectrum-area (Cheng et al., 2000), concentration-distance (Li et al., 2003)) have been used widely in geochemical anomaly identification.

Many of these traditional methods (especially multifractal methods) can consider both compositional relationships and spatial features of geochemical variables. However, these methods rely mostly on certain prior assumptions and consider only linear, lower-order properties (Zuo et al., 2019). Due to complex geological and metallogenic processes, the distribution of geochemical patterns is more often than not multimodal and complex, which introduces great challenges to these traditional methods (Zuo, 2017).

In recent years, with the rapid advancement of information technology, scholars have increasingly adopted a variety of machine learning methods in geochemical analysis (Porwal et al., 2003; Twarakavi et al., 2006; Chen et al., 2019a, 2019b, 2019c; Li et al., 2019), due to the capabilities of these methods in modeling nonlinear systems and capturing complex multistage geological events (Carranza and Laborte, 2015; Zuo et al., 2019). Machine learning methods can be classified into reinforcement learning, supervised learning, and unsupervised/self-supervised learning (Jordan and Mitchell, 2015). Many machine-learning-based geochemical anomaly identification methods are suitable for supervised or unsupervised learning. In geochemical exploration, anomalous areas usually cover only 1.5–5% of a region of interest (Chen et al., 2009, 2014), thus making it necessary to spend considerable time and

labor finding mineralized samples. Due to the demand for labeled samples (i.e., known mineral deposits), supervised learning methods are unsuitable for areas without sufficient known mineral deposits. Unlike supervised learning methods, unsupervised learning methods (e.g., autoencoders and one-class support vector machines) do not rely on labeled samples (Barlow, 1989; Ahmad et al., 2017). Generally, unsupervised learning can be used for anomaly detection based on two essential premises: (1) there are considerably fewer anomalous samples than normal (background) samples; and (2) the characteristics of anomalous samples differ significantly from those of normal (background) samples. Specifically, unsupervised learning methods first characterize samples using dimension reduction or metrics. Samples that differ from the overall data distribution are considered anomalies. Mineralized samples are much fewer than non-mineralized (background) samples in geochemical exploration, thus making unsupervised learning strongly suitable for geochemical anomaly identification. Several unsupervised machine learning methods, including continuous restricted Boltzmann machines (Chen et al., 2014), one-class SVM (Chen and Wu, 2017), and isolation forest (Zhang et al., 2021b), have been applied to geochemical anomaly recognition.

Autoencoders (AE) are among the most widely used unsupervised machine learning methods in geochemical anomaly recognition (Chen et al., 2014, 2019a; Xiong and Zuo, 2016, 2020; Guan et al., 2021; Zhang et al., 2021a). An AE reconstructs data by learning the general pattern embedded in the original data, and then the difference between the original data and the reconstructed data is calculated as an indicator to detect anomalies (An and Cho, 2015; Zong et al., 2018; Tang et al., 2020). An AE consists of an encoder and decoder by stacking multiple layers. The encoder is responsible for learning the hidden features in input data and mapping these features into a low-dimensional space. The decoder can reconstruct the original data using latent representations (Hinton and Salakhutdinov, 2006). After reconstruction, samples with small probabilities (i.e., anomaly samples) have higher errors, while samples with large probabilities have more minor errors. Therefore, samples with small probabilities can be easily distinguished based on reconstruction errors. Shallow unsupervised machine learning algorithms, such as AEs, do not need to learn the nonlinear features of geochemistry under the premise of artificial hypotheses to identify anomalies. However, AEs cannot extract deeper features from complex geological systems (Zuo et al., 2019). More importantly, many machine learning methods cannot represent or extract the spatial structure of geochemical patterns.

Deep learning algorithms have been adopted to identify geochemical anomalies to solve the above problems. To address the insufficient feature recognition ability of AEs, Xiong and Zuo, (2016) proposed geochemical anomaly identification methods based on deep autoencoders and compared them with continuous restricted Boltzmann machines. The results showed that a deeper network could better learn the compositional relationships of geochemical variables and identify anomalies. Considering that background samples and anomalous samples may have the same mean value but different variations, Luo et al., (2020) used variational autoencoder to identify geochemical anomalies according to reconstruction probability rather than reconstruction error. Zhang and Zuo, (2021) also proposed an improved adversarial learned anomaly detection (ALAD) method, which combines the advantages of a deep variational autoencoder and a generative adversarial network, and significantly improves anomaly detection performance.

Parallel to the development of traditional methods, deep learning also needs to learn automatically geochemical spatial patterns. With the advantage of extracting spatial patterns from gridded data, many studies have introduced convolutional neural networks (CNNs) to qualify spatial patterns for geochemical exploration (Chen et al., 2019c; Li et al., 2020; Zhang et al., 2021c). Specifically, the procedure used by CNNs to learn geochemical spatial patterns can be summarized in the following three steps: (1) interpolate irregularly distributed geochemical samples into regular grids, such as raster maps; (2) use convolution kernels to scan and map grids to obtain a spatial feature map; and (3) repeat the operation of step (2) to obtain a deep feature map. Following these steps, Li et al., (2019) mined the composition relationship between the spatial distributions of geochemical data and manganese mineral deposits using a deep CNN. Considering that adjacent pixels are likely to belong to the same class, Zhang et al., (2021c) combined the pixel pair feature (PPF) and a deep CNN to solve the insufficiency of metallogenic (anomaly) labels.

In unsupervised learning, Chen et al., (2019c) proposed a convolutional autoencoder (CAE) model for geochemical anomaly identification. To

fuse the spatial pattern and compositional relationship of geochemical variables, Guan et al., (2021) proposed a spatial-compositional feature fusion convolutional autoencoder for multivariate geochemical anomaly recognition, which improved greatly the accuracy of geochemical anomaly recognition. In addition, acquiring and preprocessing geochemical data often leads to some geochemical variable data loss. Xiong and Zuo, (2021) used a stacking convolution denoising autoencoder (SCDAE) to extract the robust features of a model to reduce the sensitivity to some damaged data.

Although CNNs can provide the ability to extract spatial structures, they also have a few application limitations. For example, a convolution kernel can handle only regularly distributed and shaped grids (raster and images). However, due to the complex and diverse geological environment, geochemical samples collected in the field are often irregularly located in space and even partially missing (Cheng, 1999; Ge et al., 2005; Xiong and Zuo, 2021). The irregularly distributed geochemical samples must be interpolated into a regular distribution, such as raster and images, to be suitable inputs to CNNs. Such interpolation inevitably introduces uncertainties into the data (Wang and Zuo, 2019; Zuo et al., 2021) and degrades the pattern learning performance, hence lowering the ability for anomaly recognition. Therefore, an approach is urgently needed to learn spatial features and to identify anomalies directly from geochemical samples at random locations in arbitrary regions.

This study proposes a new unsupervised graph learning framework for geochemical anomaly recognition (GAUGE), which combines a graph neural network (GNN) with an AE to solve the problem. By constructing a topology graph of geochemical sample points according to their adjacency relations, GAUGE can extract both the spatial features and compositional relationships of geochemical variables collected at irregularly distributed locations and detect anomalies. The experiments showed that GAUGE achieved the highest anomaly detection accuracy compared with several existing methods. More importantly, GAUGE provides an end-to-end solution for deep-learning-based geochemical anomaly detection using the original geochemical samples. By enabling the extraction of spatial features of geochemical data from non-Euclidean data without error-introducing interpolation, GAUGE effectively broadens the applicability and feasibility of deep learning techniques in geochemical anomaly detection.

METHODS

A novel unsupervised framework for recognizing geochemical anomalies using graph learning, GAUGE, is proposed in this study. As the principle of AE in the Introduction above, GAUGE identifies anomalies by reconstructing errors. GAUGE can extract and fuse spatial structural features and compositional relationships of geochemical variables collected at irregular locations for multivariate geochemical anomaly identification.

The GAUGE architecture consists of three essential steps (Fig. 1): (1) geochemical topology graph construction, which aims to construct a graph of multivariate geochemical variables at a group of randomly located sampling points; (2) attributed graph autoencoder training: the AE comprises an attributed graph encoder and an attributed reconstruction decoder, which are responsible for modeling the spatial structure and compositional relationships of geochemical variables simultaneously and reconstructing the variables with the obtained node embeddings by a graph attention network (GAT), respectively; (3) anomaly detection: the Euclidean distance between pairs of original geochemical variable values and the reconstructed background values at each sampling point are calculated as the anomaly score, and an anomaly map is generated.

Constructing a Geochemical Topology Graph

A geochemical dataset is usually collected as a group of sampling points in an area. For geochemical anomaly detection, capturing the compositional features of geochemical variables alone may be insufficient because the spatial distribution of geochemical variables also reflects complex geological processes (e.g., mineralization). This study connects closely related sampling points to represent the spatial structure from point data and obtains an undirected graph $G = (X, A)$ as:

$$X = \left\{\overrightarrow{x_1}, \overrightarrow{x_2}, \ldots, \overrightarrow{x_N}\right\}, \overrightarrow{x_i} \in R^F \qquad (1)$$

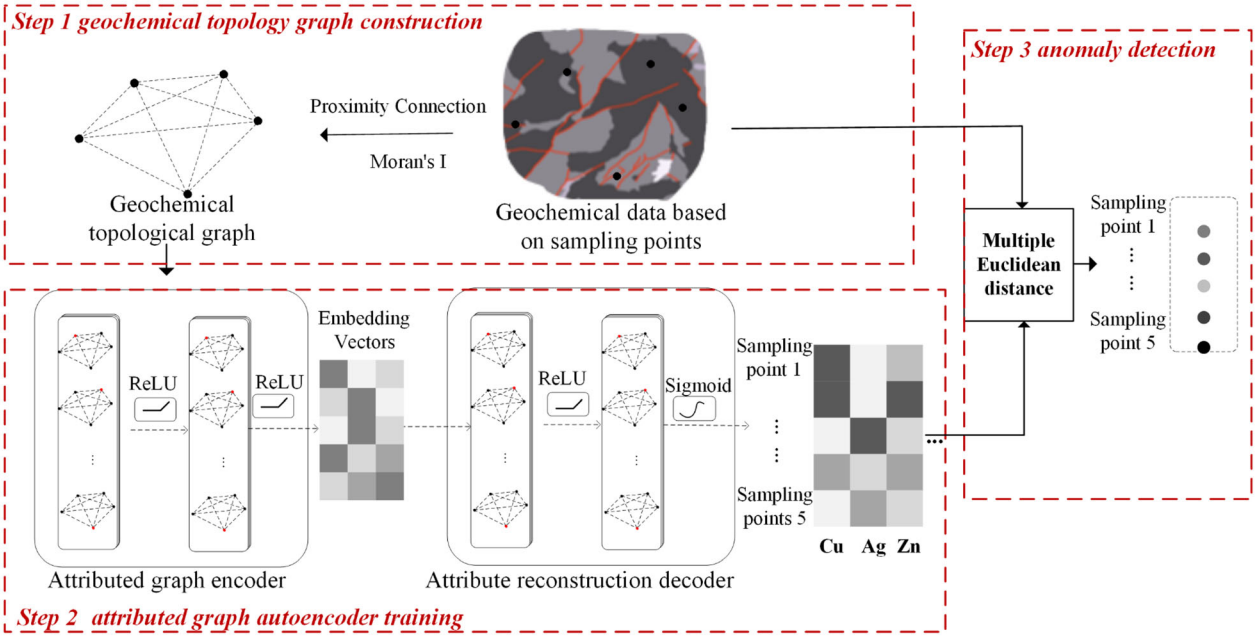

Figure 1. The overall GAUGE framework for geochemical anomaly identification.

$$A = \{A_{1,1}, A_{1,2}, \ldots, A_{N,N}\},\ A_{i,j} = \begin{cases} 1, & d_{i,j} \le K \\ 0, & d_{i,j} > K \end{cases} \tag{2}$$

where X and A represent the set of nodes (i.e., sampling points) and edges, respectively, N represents the number of nodes in the graph, F represents the number of features (i.e., geochemical variable) of each node, $d_{i,j}$ denotes the spatial distance between stations j and i, and K denotes the distance threshold when constructing the geochemical topology graph.

Generally, the smaller the distance between two points is, the more related are these two points are (Tobler, 2004). The sampling points are connected when the distance between two stations is less than K. Determining the appropriate distance threshold value (i.e., K) is critical. The connection here expresses only the geographical locations of sampled points. It does not represent the relationship between sampling points (spatial pattern of variables), and the exact relationship has to be calculated based on the attention coefficients of GAT as explained below.

As shown in Figure 2a, if the distance band (i.e., K) is too short, the nodes have too few neighbors, and this may lead to a situation in which background samples do not significantly outnumber anomalous samples in certain anomaly dense regions. The spatial relationship among background samples cannot be learned. The nodes have too many neighbors if the distance threshold is too large (as the blue area is in Fig. 2a); the spatial heterogeneity of the geochemical variables cannot be effectively represented.

To solve the above problem, global Moran′s I (Goodchild, 1986; see Appendix), a metric for measuring spatial structure, is introduced to GAUGE to measure the spatial distribution patterns of regional geochemical variables and to determine the appropriate distance threshold (Bin et al., 2017). As shown in Figure 2b, the global Moran′s I value with different distance bands was calculated for a set of sampling points in an area. If the Moran′s I value decreases rapidly as the window size increases, this indicates a strong relationship between the spatial structure and distance. To better learn the background features, the distance threshold should be larger than the current distance. If the Moran′s I value decreases slowly as the distance band increases, this indicates that the spatial structure is stable (Goodchild, 1986). This study selected the inflection point of the curve (i.e., the black point in Fig. 2b) as the optimal threshold (i.e., K) to balance partial heterogeneity and learning perfor-

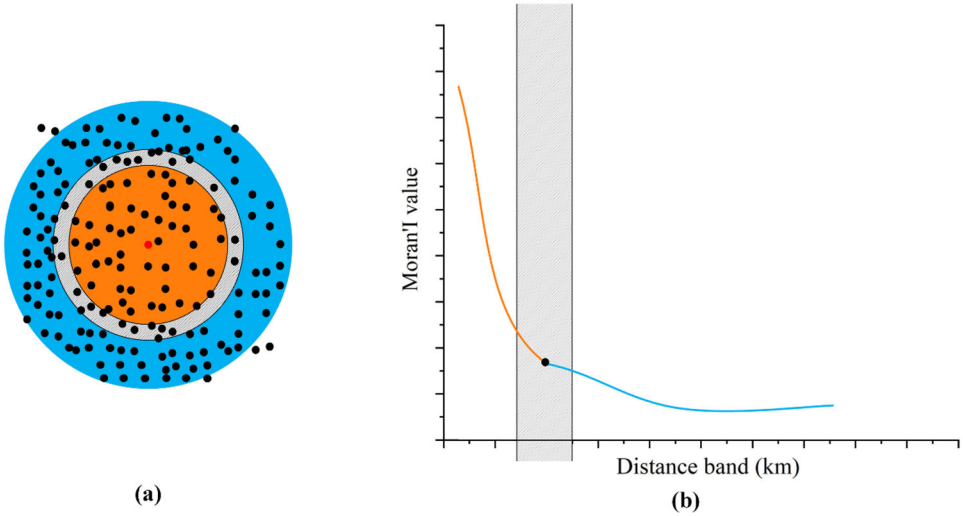

Figure 2. Schematic for finding the optimal K using Moran's I. (**a**) Neighborhood of sampled points at different thresholds (with red points as an example). (**b**) Schematic of global Moran's I.

mance. After the above steps, a geochemical graph was constructed to represent the spatial structure and the concentrations of geochemical variables for a group of samples in an area.

Attributed Graph Encoder

Like other encoders, the attributed graph encoder can learn features from data and map them in low-dimensional space. However, popular encoders based on convolutional layers can process only data with regular spatial arrangements and, therefore, cannot be applied to geochemical topology graphs. To solve this problem, we propose a new attributed graph encoder inspired by a graph attentional layer, namely GAT, proposed by Veli et al., (2017). Due to its ability to model network structure and nodal attributes seamlessly on an attributed graph, GAUGE can learn the spatial structure features and the geochemical characteristics at irregularly located sampling points. As shown in Figure 3, the graph attentional layer first calculates the weights (i.e., attention coefficients) of each graph edge according to the masked attention mechanism (Bahdanau et al., 2014) and then aggregates the neighboring node features by weight.

From a geological perspective, attention coefficients can, to some extent, indicate the anisotropy of geochemical patterns. The geochemical signatures of mineralization-favored spaces inherited from multiple geoprocesses are often anisotropic. This anisotropy can be characterized by the gradient variation or correlations of geochemical element concentrations at a sampling point with those at surrounding sampling points. To extract the nonlinear relationship between sampling points, we used a neural network approach with an attention mechanism to calculate weight e_{ij}. When extracting spatial patterns, the model can aggregate neighbor attributes (geochemical variables) based on this weight. In other words, when a sample point communicates with its neighbors with geochemical elements features, e_{ij} will be weighted on this feature. The larger the weighting coefficient is, the greater its effect on the central sampling point is. This process allows the model to consider the anisotropy of the geochemical signature of each sampling point when extracting geochemical patterns.

Mathematically, given an input graph $G(X, A)$ containing N nodes (sampling points), each node has a feature vector $\overrightarrow{x_i}$ and dimension F (the count of kinds of geochemical variables or the dimension of the extracted features). The attention coefficients e_{ij} between nodes i and j according to the attention mechanism is calculated as:

$$\begin{aligned} e_{ij} &= attention\left(W\overrightarrow{x_i}, W\overrightarrow{x_j}\right) \\ &= \text{LeakyReLU}\left(\vec{a}^T[W\overrightarrow{x_i}||W\overrightarrow{x_j}]\right) \end{aligned} \tag{3}$$

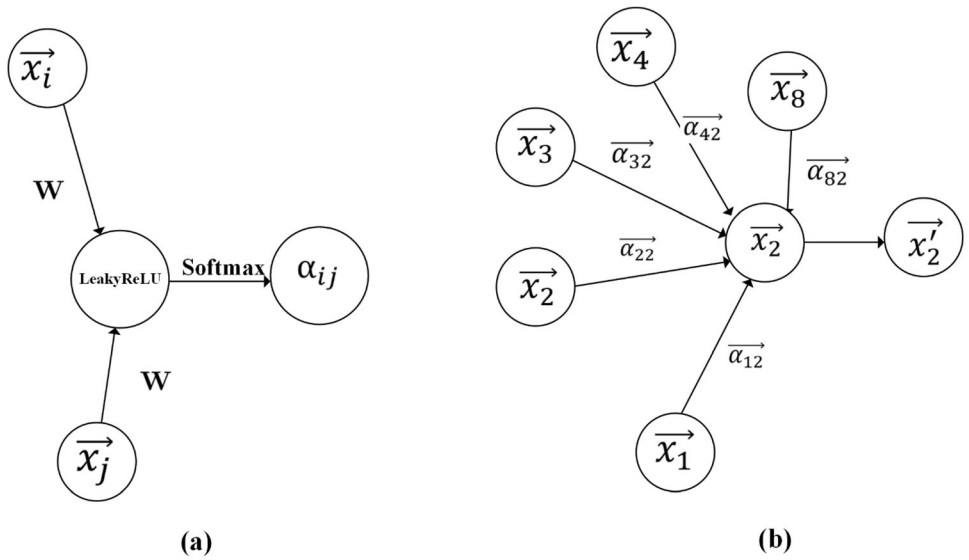

Figure 3. Illustration of the graph attention layer: (**a**) the attention mechanism; and (**b**) aggregating features of the node 2-based attention mechanism, with neighborhood {1,3,4,8}.

where *attention*() denotes a single-layer feedforward neural network with a weight vector $\overrightarrow{a}$, W denotes weight matrix for transforming the input features into higher-level features, and $.^T$ and $||$ represent transposition and concatenation operations, respectively. To make the coefficients compute between nodes efficiently, *softmax*() was used to normalize the neighbors of node i.

$$\alpha_{ij} = softmax(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{k\in N_i}\exp(e_{ik})} \tag{4}$$

where N_i represents the neighbors of node i and α_{ij} denotes the correlation between node i and node j variables. Details of LeakyReLU can be found in the Appendix. Fully expanded, these processes are illustrated in Figure 3a. As mentioned before, the anisotropy of sampling point 2 in terms of geochemical signature is characterized by the relationship weights of sampling point 2's neighborhood points {1,3,4,8} (Fig. 3b).

After obtaining the attention coefficients between nodes, GAT aggregates each neighboring node (Fig. 3b) and finally obtains the embedded features as:

$$\overrightarrow{x_i'} = \sigma\left(\sum_{k\in N_i}\alpha_{ij}W\overrightarrow{x_j}\right)\overrightarrow{x_i} \in N_i \tag{5}$$

where $\overrightarrow{x_i}$ and $\overrightarrow{x_i'}$ denote input data and embedded features, respectively, and σ is a nonlinear activation function. For a graph, all the above calculation processes can be expressed as:

$$X^{(k)} = \left(X^{(k-1)}, A|W(k)\right) \tag{6}$$

where $X^{(k-1)}$ is the input for the graph attention layer $k-1$ and $X^{(k)}$ is the output of the graph attention layer.

Any node can reach other nodes in a few steps in a small graph. Therefore, the perceptual field needs to have only a few layers to cover the entire graph and adding more layers will not be much help. Previous studies have shown that graph neural networks with two layers perform better (Kipf and Welling, 2016; Zhang et al., 2020). This study used two graph attention layers to construct the attributed graph encoder. The attributed graph encoder can be formulated as:

$$X^{(1)} = Relu\left(X^0, A|W(1)\right), X^0 \tag{7}$$

$$Z = X^{(2)} = Relu\left(X^1, A|W(2)\right) \tag{8}$$

where $W(1)$ is an input-hidden layer with input graph $X^{(1)}$ and $W(2)$ is a hidden-hidden layer with

hidden feature $X^{(1)}$. After applying two layers, the geochemical input graph can be learned and mapped to a low-dimensional vector space Z, which can seamlessly learn network structure and nodal attributes (i.e., spatial structural features and compositional relationships of geochemical variables).

Attribute Reconstruction Decoder

Similarly, to calculate reconstruction errors, we propose an attribute reconstruction decoder that reconstructs the nodal attributes (i.e., geochemical variables) from the encoded latent representations Z. Specifically, the attribute reconstruction decoder structure is symmetric to that of the encoder, and both use a graph attention layer to predict the original node attributes as:

$$X^{(3)} = Relu\left(X^{(2)}, A|W(3)\right) \tag{9}$$

$$\hat{X} = X^{(4)} = Sigmoid\left(X^{(3)}, A|W(4)\right) \tag{10}$$

Due to input data normalization, we applied the activation function $Sigmoid()$ to restrict the reconstruction values in the range of [0,1]. Details of *Relu* and *Sigmoid* can be found in Appendix.

Anomaly Detection

Geochemical anomaly scores are calculated for the sampling points. However, the widely used cost function in prediction models, i.e., the mean square error (MSE), focuses only on the reconstruction error of each variable while ignoring the differences in elemental contribution to mineralization among sampling points (Paszke et al., 2019). To address this issue, we replaced the MSE with the SPRE (sampling point reconstruction error), which is calculated as:

$$L_{SPRE} = \frac{1}{N}\sum_{i=1}^{N} A_i \tag{11}$$

$$A_i = \sqrt{\sum_{k=1}^{F}\left(x_i^k - x'^k_i\right)^2} \tag{12}$$

where x_i^k and x'^k_i denote the k th original and reconstructed features, respectively, of the i th node, N and F represent the number of nodes and the number of variable categories, respectively, and A_i represents the reconstruction error for each node. By minimizing the above cost function, our proposed model can continuously learn the background structure and compositional relationship from the input geochemical graph and then approximate iteratively the input attributed graph with encoded latent features until the cost function converges. Subsequently, we input the original geochemical variables into the completed training model to obtain the reconstructed values (background values) for each node. The multivariate Euclidean distance between the original geochemical values and the reconstructed background values is calculated as the reconstruction error. The anomaly map is generated based on the final reconstruction errors (i.e., A_i).

Performance Assessment

The primary purpose of geochemical exploration is to delimit areas of interest for further exploration; thus, geochemical anomaly detection should be evaluated from two perspectives: the percentage of discovered mineral deposits to the total mineral deposits and the size of the anomalous area. In this study, we used the prediction–area (P–A) plot (Yousefi and Carranza, 2015) and the receiver operating characteristic curve (ROC curve) (Fawcett, 2006) to evaluate GAUGE. Eight anomaly detection algorithms proposed in previous studies were chosen for comparison with GAUGE.

Comparison Methods

The selected algorithms can be applied to geochemical data collected at irregularly located points and can be classified into the following three categories.

(1) Linear Models, including minimum covariance determinant (MCD (Hardin and Rocke, 2004) and one-class support vector machine (OCSVM) (Schölkopf et al., 2001; Zuo and Carranza, 2011): embed the data in low-dimensional space, and the data that indicate poor results after projection in the low-dimensional space are considered outliers. Although these models have been applied successfully in geochemical anomaly detection,

they consider only nodal attributes (Xiong and Zuo, 2020).

(2) Proximity-based Models, including local outlier factor (LOF) (Breunig et al., 2000), connectivity outlier factor (COF) (Tang et al., 2002), cluster-based local outlier factor (CBLOF) (He et al., 2003) and isolation forest (IForest) (Liu et al., 2008): the distribution of anomalous and normal points is different in some indicators. Based on these indicators, proximity-based models compare each point distribution with the overall points to find the anomaly. LOF and COF measure the proximity of each point by calculating the local reachable density and the average connection distance ratio, respectively. IForest detects anomalies by comparing the number of spatial divisions required to isolate samples (Carranza and Laborte, 2015; Zhang et al., 2021b). Note that the distribution and density mentioned above describe the feature space. In other words, these models ignore the geographic locations of sampling points.

(3) Neural Network Models, including autoencoder (AE) (Kingma and Welling, 2013), deep autoencoder network (DAN) (Xiong and Zuo, 2016), convolutional autoencoder (CAE) (Chen et al., 2019c) and the proposed GAUGE: as mentioned in the Introduction and Method sections above, autoencoder networks are typically unsupervised anomaly detectors based on neural networks. These methods use the reconstruction error as the anomaly score. Compared with traditional methods, these methods have the advantages of nonlinear modeling systems and broader scalability and they have been used widely in geochemical anomaly detection (Zuo et al., 2019; Zhang et al., 2021a). It is worth mentioning that CAE can only be applied to an interpolated rectangular area because it considers the spatial characteristics of geochemical variables.

Prediction–Area (P–A) Plot

The prediction area (P–A) plot is a traditional indicator that evaluates models prospectively by combining the known mineral occurrence probability and the areas occupied. There are two y-axes in the P–A plot: the prediction rate of known mineral deposits (P) and the percentage of the anomaly region area (A). We can calculate and plot two curves (i.e., P–A). The intersection of these curves indicates that the sum of the predicted rate and the occupied area equals 100 and is a proper indicator to evaluate an anomaly map. If the intersection point is higher on the y-axis, the model identifies a smaller anomaly area and a higher prediction rate.

Receiver Operating Characteristic Curve

Although the P–A plot can evaluate model performance, it applies to the intersection only and cannot consider the full range of possible anomaly thresholds. To deal with these issue, the ROC curve, a commonly used evaluation method for machine learning, was also adopted in this study. Using the known deposits as references, at a certain anomaly score threshold all samples can be divided into four groups: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). Then, the true positive rate (TPR) and the false positive rate (FPR) at multiple thresholds can be calculated to generate a curve as follows:

$$\mathrm{TPR} = \frac{TP}{TP + FN} \tag{13}$$

$$\mathrm{FPR} = \frac{FP}{TN + FP} \tag{14}$$

In geochemical anomaly detection, TPR is the proportion of deposits labeled anomalous to all known deposits. FPR refers to the proportion of nonmineral points labeled anomalous to the total true nonmineral points. The larger the TPR value and the smaller the FPR value, the fewer the nonmineral points and the larger the mineral points among the anomalies delineated by a model. To evaluate the performance of the model comprehensively, the ROC curve is plotted with FPR and TPR as the x and y-axes, respectively. The area under the curve (AUC) indicates the comprehensive performance of a model considering all thresholds. The closer the AUC is to 1 (i.e., the maximum value of FPR $\times$ TPR), the greater the accuracy of the model. ROC curve is a popular indicator for evaluating geochemical anomaly results based on machine learning.

EXPERIMENT AND EVALUATION

Study Area and Dataset

Geology of the Study Area

To demonstrate the performance of the proposed method, the Longyan area in the Wuyishan polymetallic metallogenic belt in China was used as a case study area to identify geochemical anomalies related to Au deposits (Fig. 4a). The Wuyishan metallogenic belt, located within the active continental margin of southeastern China, is an important metallogenic area in the Cenozoic tectonic–magmatic belt of the Circum-Pacific (Mao et al., 2010; Lin et al., 2020). This region has long experienced the convergence of global supercontinents and the breakup of northern and southern continents (Jianhua et al., 2016). Such a special tectonic environment and the long-term complex tectonic–magmatic evolution history provided favorable geological conditions for mineralization. This region has many known deposits, and it is especially well known for the Zijinshan mega-gold mine, a typical case area for studying gold ore (So et al., 1998; Zhang et al., 2015; Li and Jiang, 2017).

The case study area (i.e., Longyan area, as shown in Fig. 4a) is located in the central part of the Wuyishan polymetallic metallogenic belt, southwest Fujian. The area is characterized by strong Indosinian–Yanshan periods of magmatic activity and the development of various stock, bedrock, and dike. In addition, the surrounding rocks in the area are strongly altered and the combination of various alteration material types has a distinct zoning pattern, which is a significant indicator of hydrothermal mineralization (Li et al., 2013; Mathieu, 2018). Although there are 12 Au deposits (including the Zijinshan mega-gold mine), except for Zijinshan, there are no published studies about most of the gold mines in Longyan.

According to previous geological studies, it can be determined that the gold mine here and the Zijinshan field belong to the same source and the same period of medium-acidic magmatic rock mineralization evolution. The main Au deposits in this region are ''Zijinshan type'' low-sulfidation epithermal Au deposits, porphyry Au deposits and medium–low temperature hydrothermal Au deposits (Zhang et al., 2015; Li and Jiang, 2017; Chen et al., 2019b). The geochemical anomaly model of these Au deposits in Longyan can be inferred from the confirmed Zijinshan gold mine.

Geochemical Characteristics of the Zijinshan Au Deposit

Recently, many studies analyzed the metallogenesis of Zijinshan Au deposits systematically by using geochemical data (Zhang et al., 2005; Zaw, 2007; Singer et al., 2008). The Au mineralization in the Zijinshan field is dominated by Cu–Au mineralization. The Au mine consists mainly of natural

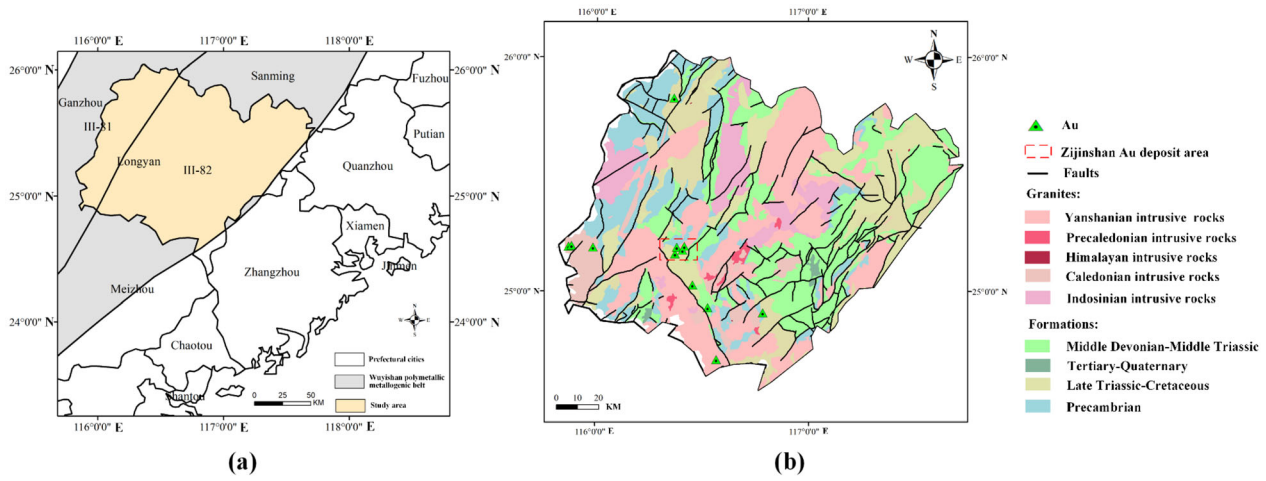

Figure 4. The case study area. (**a**) The Longyan area is located in the central part of the Wuyishan polymetallic metallogenic belt, southwest Fujian. (**b**) Simplified geological map of the Longyan area (data from the Institute of Geophysical and Geochemical Exploration, Chinese Academy of Geological Sciences (IGGE)).

Au endowment in limonite, accompanied by Ag, Cu, and Pb.

The regional geochemical elements of the Zijinshan mineral field are characterized by high concentrations and strong couplings, and there are obvious centers of high concentration aggregation. There is a clear boundary in the distribution of high and low concentrations of Cu, Pb, Au, Ag, Bi, and Mo, which can indicate the center and exterior of the field. The distributions of Cu, Pb, Zn, W, Bi, and Cd extend in the NE direction; those of Au, Bi, and Zn extend NW; those of Au and Mo extend E–W; and that of Sb spreads N–S. These elements' diversity and complex distributions reflect the interaction of different geological formations and multiple hydrothermal fluids. These geological effects are superimposed in the complex rock masses of Zijinshan, thus eventually leading to different element enrichment patterns in different geological sites (Dikang et al., 1997).

Specifically, the relationships between metallogenesis and accompanying elements can be categorized into the following three groups (Dikang et al., 1997; Huang et al., 1999):

(1) Au, Ag, Sb, Cu, Pb: these elements are associated with low-temperature volcanic-subvolcanic hydrothermal gold mineralization and low-temperature subvolcanic solution-type gold mineralization;
(2) Cu, W, Mo: these elements are associated with porphyry gold mineralization;
(3) Ti, V: spillover elements during gold mineralization.

Au Mineralization Indicator Element Dataset

Benefiting from unsupervised learning, GAUGE requires only geochemical variables for training. In total 4,812 sampling points at a scale of 1:200,000 were explored in the study area to collect stream sediment geochemical data, including 32 elements and five oxides (Xie et al., 1997). Due to lack of published studies on most Au deposits in Longyan, we cannot obtain detailed geochemical patterns of Au deposits in Longyan. However, the Au deposits here and in the Zijinshan field belong to the same source and the same period of medium-acidic magmatic rock mineralization evolution. The Zijinshan mega-gold mine has also been analyzed from a geochemical perspective. Thus, the geochemical anomaly model for the Au deposits in Longyan can be inferred from the confirmed Zijinshan mega-gold mine. In other words, we tried to recreate the actual application scenario, i.e., how to detect other gold-related anomalies in the area when the geochemical pattern of only one mine in the area is known. We selected 12 geochemical indicator variables (i.e., Ag, Au, Bi, Cu, Mo, Pb, Sb, Sr, Ti, V, W and Fe_2O_3). Because GAUGE enables direct learning of spatial features among nodes, we needed only to normalize the original sampled point data rather than interpolate or crop the data into rectangles. The normalized data were scaled to [0,1], thus maintaining the original elemental relationships. The data distribution was similar to that of the output data of the output layer after sigmoid mapping. Such an operation alleviates gradient disappearance and enables better loss calculation and model training.

Application of GAUGE

The global Moran's I value for different elements and oxides was calculated (Fig. 5a). There was a clear inflection point at approximately 20 km. As mentioned in the Methods section, we used 20 km as the distance threshold for constructing the topology graph. An edge connects a pair of sampling points if the distance between them is less than 20 km. Then, a topology graph of the geochemical sampling points was constructed for network training and anomaly detection.

To better train the autoencoder network in GAUGE, the learning rate decay strategy was introduced in this study. Figure 5b shows that the learning rate decreased to half of the original rate when the loss value did not decrease for 10 epochs. The encoder and reconstruction decoder learned and reconstructed the multivariate geochemical background better through this mechanism. Encoder and decoder training was completed when the loss stabilized at a low value (Fig. 5b). Then, we input the geochemical topology graph into the attribute encoder and attribute reconstruction decoder in turn and obtained the background values for each sampling point. Finally, the multivariate Euclidean distance (i.e., the anomaly score) between the original geochemical concentrations and the reconstructed concentrations was calculated per sample to generate the anomaly map.

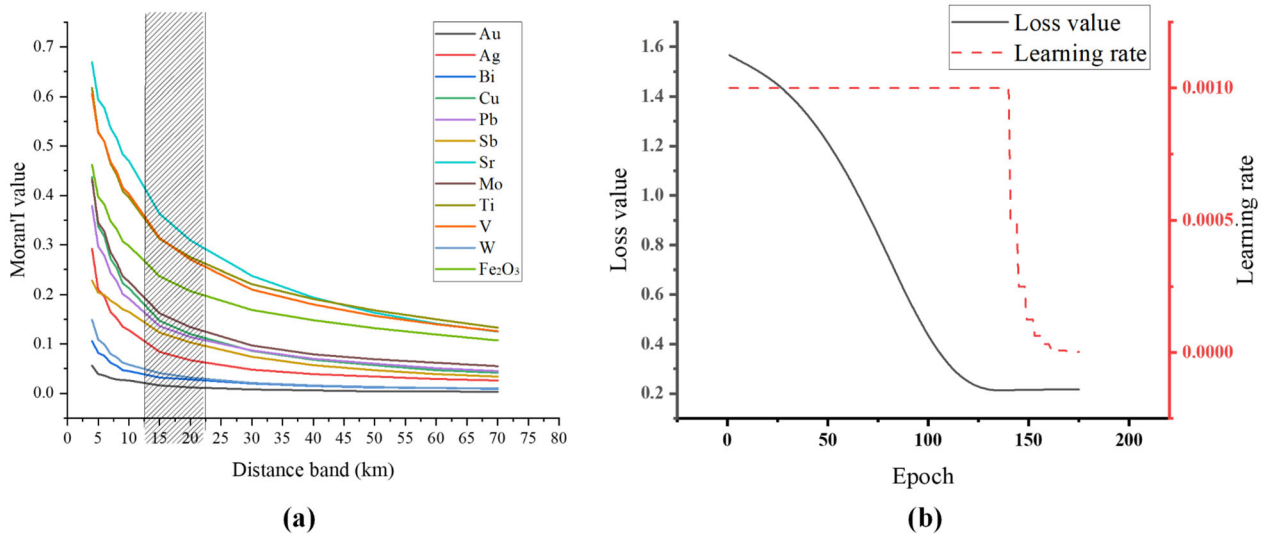

Figure 5. (**a**) Variation in global Moran's I index with different distance bands (i.e., *K*). (**b**) Variation in cost function and learning rate with number of training epochs.

Performance Evaluation

The anomaly values of the sampling points near the known Au deposits were higher than those farther ones (Fig. 6a). The anomaly score decayed with increasing distance from the deposits. More meaningfully, the anomaly distribution had a high coincidence with the Au deposits trend, and the anomaly can better reflect the mineral distribution. In addition, as the intersection of the two curves in the P-A plot (Fig. 6b) shows, GAUGE predicted more than 75% of the known Au deposits in less than 25% of the study area. This indicates the strong coincidence between the known mineral deposits and the areas with high anomaly scores; in other words, the anomaly map obtained by GAUGE can be used to guide mineral exploration.

The accuracy of autoencoder-based methods was higher than those of the other machine learning anomaly detection methods (Table 1). In particular, the AUC of GAUGE (0.833) with the graph attention layer was significantly higher than those of the other methods, thus indicating that GAUGE performed the best in geochemical anomaly detection of Au mineralization. To further demonstrate the advantage of GAUGE, we replaced its graph attention layer with a graph convolution layer (GCN) (i.e., GAUGE* in Table 1). Although the AUC obtained by GAUGE with the graph convolution layer did not reach 0.833, it was still more accurate compared to those of the other models. The reason for the difference between GAUGE and GAUGE* is that GAT can extract automatically and assign weights to the edges between nodes. Such a mechanism is similar to anisotropy in geochemical pattern. In geochemistry, the relationship between two sampling points is not just a binary one calculated based on geographic location (i.e., presence of an edge/absence of an edge). The relationship between a certain sampling point and a neighboring point varies at different directions. Various geological interactions cause the anisotropy and ultimately manifest themselves in the spatial distribution of geochemical elements. Therefore, the influence of geochemical variables at sampling sites on the relationship (i.e., continuous edge weight) should also be considered. This relationship among sampling points can be learned automatically and extracted through the GAT's attention mechanism, which helps GAUGE to reconstruct geochemical background and detect anomalies.

From a geological perspective, the study area has strongly magmatic history and it has many polymetallic deposits associated with intrusive and volcanic rocks. The granite in the region is one of the controls on mineralization (Qiu et al., 2010; Jianhua et al., 2016). The mineralization period in the Zijinshan region is divided into the Indosinian and Yanshanian periods, and Au deposits associated with extensional tectonics and mixed magmatism

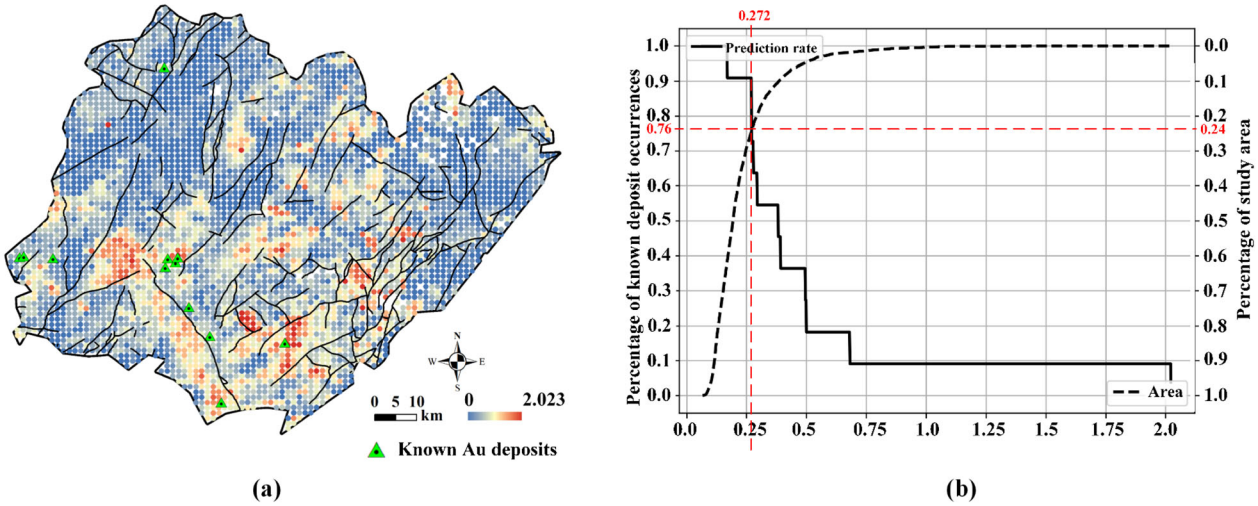

Figure 6. (**a**) Anomaly map obtained by GAUGE. (**b**) Prediction–area (P–A) plot.

Table 1. Performance indicators of various methods

Method	AUC
COF	0.712
OCSVM	0.753
LOF	0.767
IF	0.792
MCD	0.794
AE	0.795
CBLOF	0.809
DAN	0.813
GAUGE*	0.828
GAUGE	0.833

GAUGE* is the framework with GCN instead of GAT

also occurred in the late Yanshanian period (Development Research Center of Survey. and Fujian Institute of Geological Survey, 2014). The geochemical anomalies detected by GAUGE are located in or near granites (Figs. 4b, 6a). In addition, due to the indicators variables we obtained from the geochemical model of the Zijinshan area with a complex metallogenic environment, some mineralization dominated by tectonic precipitation (southeastern deposits) was also identified as high anomalies. For example, the Yongding area (southeastern part of Longyan) experienced slow subsidence and deposition from the Late Devonian onwards to the Early Triassic. The intrusive rocks in this area were not well developed; thus, the polymetallic deposit in this area is presumed to be volcanic hydrothermal–sedimentary types (Nai-Zheng et al., 2008; Development Research Center of Survey and Fujian Institute of Geological Survey, 2014). All the aforementioned results indicate that the results of GAUGE were credible. The detected geochemical anomalies can also provide significant indications for further exploration.

To better comprehend and highlight the advantages of GAUGE in considering spatial structure, we visualized the results obtained by the AE-based methods according to the optimal threshold determined by the AUC (Fig. 7). Compared with AE and DAN, GAUGE considers the neighborhood and spatial structure characteristics of the sampled points, thus making background or anomalies purer (e.g., region A). In other words, there is fewer ''salt-and-pepper noise'' anomalies and background in the region, and the anomaly map has better visual effect. In addition, we noted that, compared to the value of anomalies obtained by DAN, the value of anomalies obtained by GAUGE can better distinguish between anomalies and backgrounds. The high anomaly scores obtained by GAUGE are more significant and closer to the known Au deposits in region B, indicating that GAUGE results are more indicative of mineral resources.

To verify the ability of GAUGE to extract spatial features to identify anomalies, CAE and GAUGE were compared in this study. CAE cannot be applied to an irregular area for geochemical anomaly identification; hence, this study selected a rectangular area in the Longyan area to compare

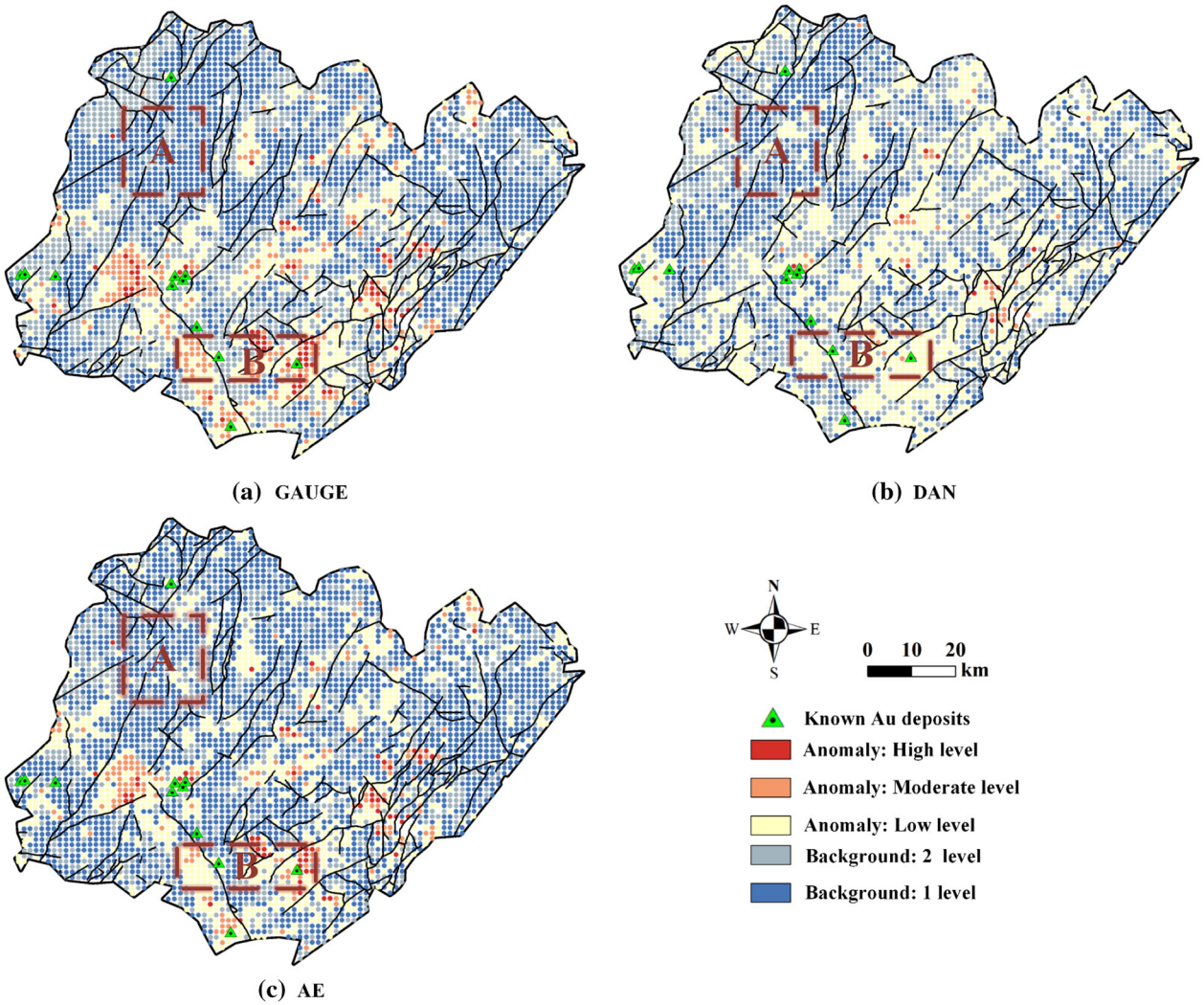

Figure 7. Anomaly recognition results of AE-based models using the optimal threshold obtained by the ROC curve. Detailed methods can be found in Chen et al., (2019a) and Guan et al., (2021).

CAE and GAUGE (Fig. 8a). The AUCs of CAE and GAUGE were 0.786 and 0.796, respectively (Fig. 8b). There is little difference in accuracy between the two models. Comparing Figure 8c and d shows that the distributions of different levels of background and anomalies obtained by the two models were also similar. The difference is that the background recognized by GAUGE was mostly level 1, thus making the anomaly score boundary between the background area and the anomaly area more obvious. In general, similar to the above, CAE and GAUGE showed the same advantages as DAN and AE because of the consideration of spatial structure feature or relationship among points besides attributes (i.e., element concentrations). Overall, the graph deep learning was feasible for geochemical anomaly identification. However, due to the advantage of graph neural networks, GAUGE can not only learn the spatial pattern of geochemical elements such as CAE but it can also be applied to irregular areas.

Uncertainty Analysis

The GAUGE can learn geochemical patterns and identify geochemical anomalies related to mineralization in irregular regions. In addition to model

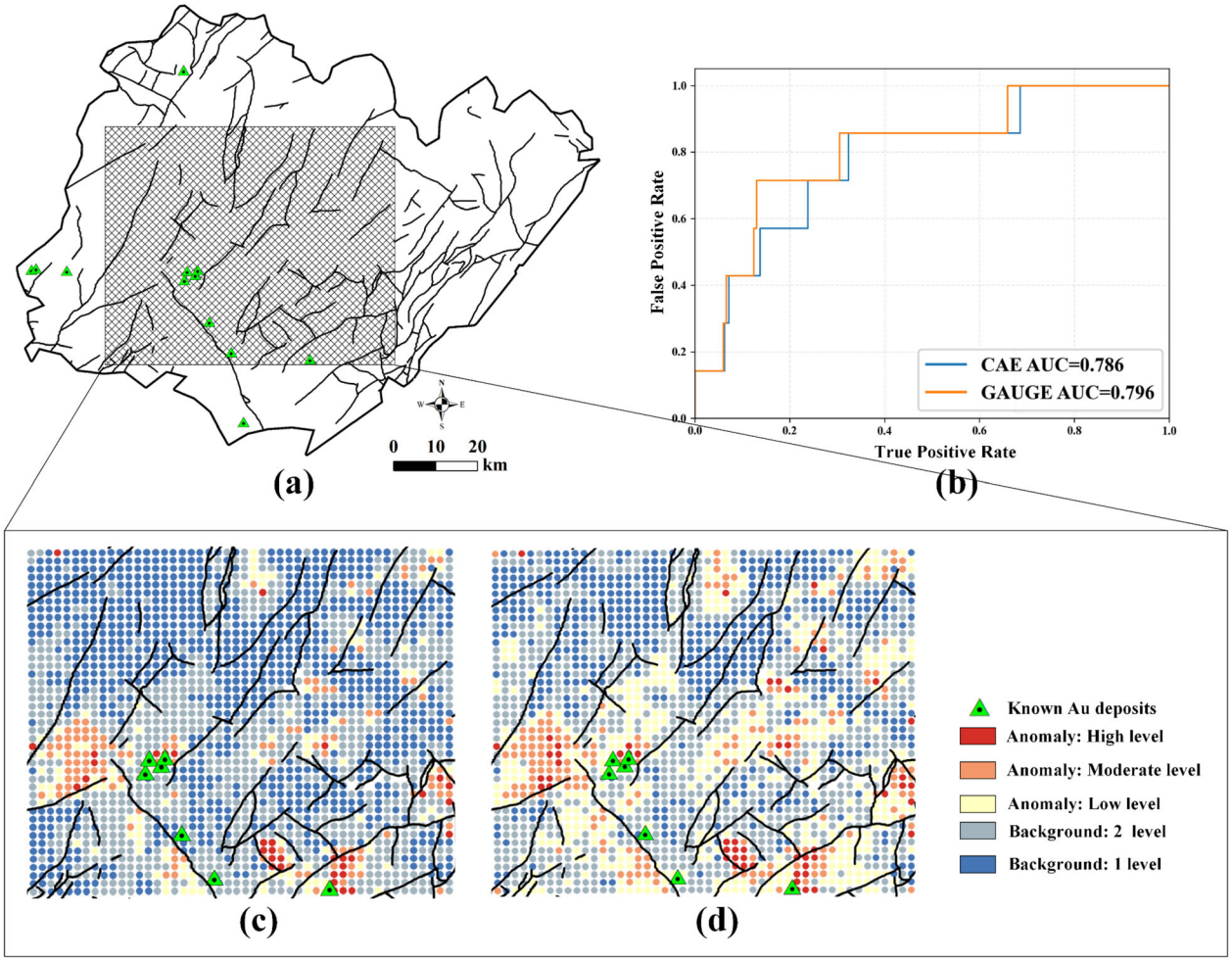

Figure 8. (**a**) Rectangular area selected for comparison of GAUGE and CAE. (**b**) AUCs of CAE and GAUGE. Anomalies detected by (**c**) GAUGE and (**d**) CAE.

structure, graph construction and choice of loss function produce many uncertainties in a model's performance. Different constructing methods and loss functions allow models to perform differently. In this study, comparative experiments were conducted separately to infer the optimal settings of the parameters by analyzing the accuracy variations to improve the application of GAUGE.

Comparison of Methods to Construct Topology Graph

In addition to distance thresholds, the influence of faults on geochemical spatial structure should be considered when constructing a geochemical graph. Similar to elemental concentrations, faults are readily available and basic in most geological datasets. From a data-driven perspective, it is natural to analyze whether the involvement of fault data in the graph affects the accuracy of a model. In general, faults are often the boundaries between two geological bodies, and the lithologic configuration of two sides of a geological body differs. To some extent, sampling points within the same geologic body are more strongly related than sampling points in different geologic bodies because of lithology. Therefore, this study compared the two methods (with fault truncation processing and without fault truncation processing) for constructing topology graphs. If a fault truncates the edge between two sampling points, this side is removed.

The relationship between faults and Au mineralization is also impacted by the timing of the formation of the faults. However, the purpose of the proposed GAUGE in this study was to identify geochemical anomalies more automatically and efficiently, and it was applied to unknown regions where detailed public information is lacking. Therefore, we discarded all a priori knowledge and used all faults for the composition graph and analysis (Fig. 9).

Compared with the AUCs of graphs without fault truncation processing, the AUCs of graphs with fault truncation processing were more stable because the fault truncation processing caused most of the edges in the graph to be truncated. Therefore, graph's and nodes' abilities to converge surrounding features were stable. However, due to fault truncation processing, some useful edges were also truncated thus causing the AUC of the graph to decrease after fault truncation processing (especially when the distance threshold was in the range from 0 to 50 km). Finding a reasonable threshold by using Moran's I was effective, thus ensuring that GAUGE obtained the highest accuracy. Additionally, we do not recommend truncating the edges when geological information is missing. The model accuracy is degraded because sampling points cannot obtain sufficient neighborhood information.

Comparison of Cost Functions

Unlike the MSE commonly used in many machine learning methods, the cost function in GAUGE improved (i.e., SPRE, as mentioned in the Methods section). To compare the performance of the two cost functions, we applied SPRE and MSE to various neural network methods. The AUCs of the methods using SPRE were significantly higher than those using MSE (Fig. 10). In particular, the AUC reached 0.833 when using GAUGE with GAT. These indicate that the performance of geochemical anomaly detection can be improved using the SPRE. This is because the MSE considers only the errors of each input and output variable equally. In other words, MSE treats different variables at different sampling points equally, thereby ignoring differences in the contribution of variables to mineralization among different sampling points. Therefore, the performance improved when we focused on the sampling point errors (i.e., by calculating each sampling point error first and then averaging the SPRE).

DISCUSSION AND CONCLUSIONS

The main task of geochemical exploration is to quantify spatial patterns and composition relationships of geochemical variables, and to identify geochemical anomalies related to mineralization. Geochemical patterns reflect complex geochemical processes. Better learning of these patterns has always been the main melody of developing geochemical methods. The emergence of deep learning methods (especially CNNs) provides approaches to learning complex and nonlinear geochemical patterns and to recognition of anomalies. Although combining CNNs and AE frameworks is the most popular and efficient unsupervised geochemical anomaly detection approach, there are still many problems. The deep learning model relies on a rectangular convolution operation in extracting spatial patterns (LeCun et al., 2015). Therefore, it is necessary to interpolate sampled point data into raster grids, which inevitably introduces uncertainties into the data. More seriously, existing deep learning methods that consider spatial patterns cannot be applied to irregular regions because geochemical data are inconsistent in resolution, partially missing and irregular. This study creatively introduces graph learning and constructs the GAUGE framework. This method can detect geochemical anomalies by extracting spatial structure and composition relationships directly from unlabeled geochemical sampling points. The Longyan area was selected as an irregular case area, and some geochemical sampling points are missing in this area. GAUGE was applied successfully to this study area. More than 75% of the known gold deposits were predicted in less than 25% of the study area, and the AUC obtained by GAUGE was 0.833. The proposed method solved the problem that the deep learning model cannot be applied to the geochemical anomaly recognition in irregular areas.

Unlike CNNs, which use convolution to extract spatial features, GAT and GCN are more similar to the improvements of DAN. Graph learning for extracting geochemical compositional relationships is the same as that of DAN. The difference is that GAT and GCN add a communication process between sampling points to learn spatial patterns (Veli

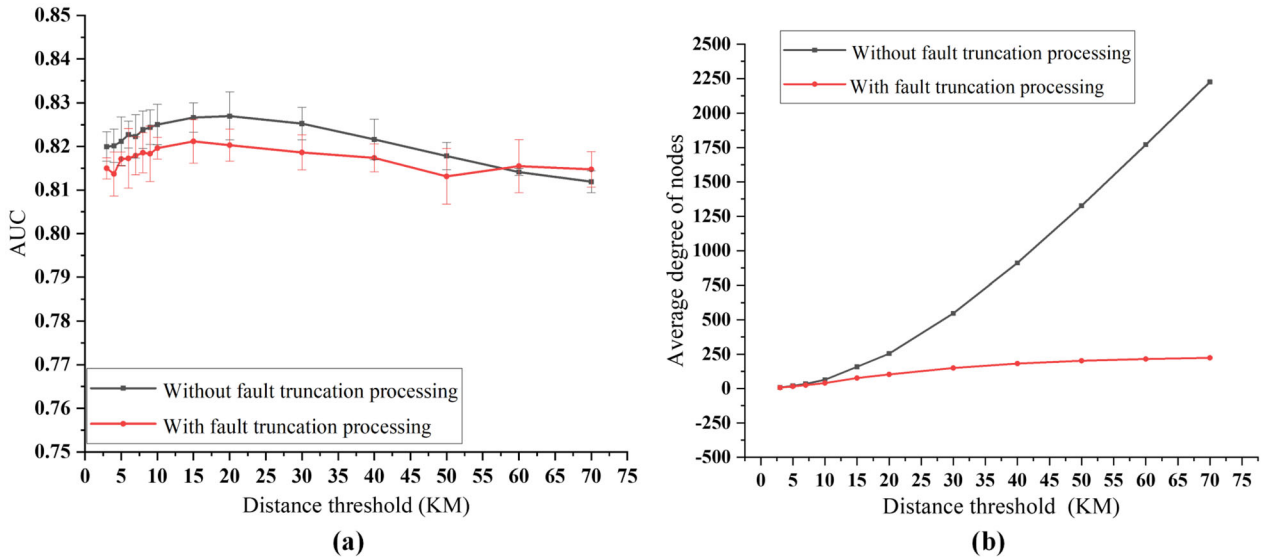

Figure 9. Effect of fault truncation processing on method accuracy. Comparisons of **(a)** accuracy (AUC) at different distance thresholds, and **(b)** average degree of graphs constructed at different distance thresholds.

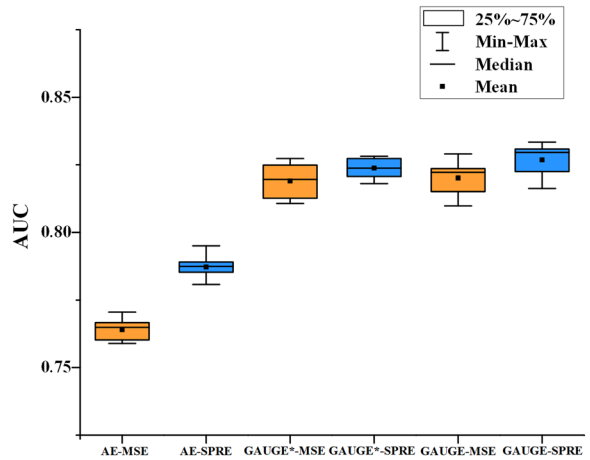

Figure 10. Performance of various methods with two cost functions (i.e., MSE and SPRE).

et al., 2017; Zhou et al., 2018). In other words, the features extracted at each sampling point are the relationships of its own collected attributes and the patterns of spatial relationships with surrounding sampling points. According to the Moran's I index, these patterns are similar to the aggregation patterns of sampling points quantified by Zuo and Xiong, (2020). In an irregular region, to compare with existing methods that consider only compositional relationships among geochemical variables (i.e., node attributes), the GAUGE framework (including GAT and GCN) achieved higher accuracy (an average improvement of 6.26%). Figure 8 illustrates that both the anomalous region and the background are purer when spatial pattern features are learned. Such a result is also more consistent with the geochemical distribution, i.e., the properties of adjacent samples are similar. Although the results and performances of GAUGE and CAE are similar, the former can be applied in irregular areas whereas that latter cannot. Therefore, graph deep learning methods (e.g., GAT and GCN) can replace CNNs in geochemical anomaly identification models. More deep learning models can be applied to identify geochemical anomalies in irregular areas. More importantly, GAUGE is helpful in geochemical anomaly identification and many other geoscience fields that need to consider spatial relations.

When extracting the spatial pattern of geochemical elements, graph convolution depends mainly on the interaction (edge) between nodes to represent the spatial relationship. How to convert geochemical exploration data into a graph is also a very important question. This study creatively proposes transforming geochemical sampled points into geochemical graph data. Referring to the method of determining the optimal convolution window size (Chen et al., 2019c; Guan et al., 2021), Moran's I was used to quantify the degree of spatial aggregation at

different thresholds and to determine the optimal distance threshold. GAUGE obtained the highest accuracy because the constructed graph balanced the spatial structure information and spatial heterogeneity. In addition, the influence of fault processing on accuracy of the model was analyzed. The results show that, when a geochemical graph without fault truncation was used, GAUGE performed better in anomaly detection. This result is consistent with Chen et al., (2019b). Hydrothermal mineralization is predominantly present in the Yongyan area (Chen et al., 2019b). Heat sources, fluid flow, activity pathways and chemical/physical traps are processes critical to hydrothermal mineralization. Faults/fractures likely provided favorable active pathways and physical traps for gold-bearing fluids (Mathieu, 2018). The difference in lithology is not highly critical in hydrothermal mineralization. The Moran's I index can detect the spatial distribution patterns of elements due to faults. More valuably, the comparison experiment demonstrated that we can manipulate the edges to give the graph some geological information, such as faults, thereby increasing the potential for GAUGE applications. Unlike a raster, edges in a graph can be carried or expressed more information. On the one hand, the importance of faults on mineralization is quantified by extracting the weights of edges of a trained model when the geological structure is unknown. On the other hand, when the influence of faults on mineralization is known, the effect can be added to the edge as a known weight to improve anomaly identification accuracy.

This study also analyzed the influence of a model's cost function (i.e., SPRE and MSE). Compared to the popular cost function MSE, the improved cost function SPRE also allowed GAUGE to maximize performance because it considers the differences in the contributions of variables to mineralization among the different sampling points. Such performance improvement is similarly effective for other autoencoder networks.

This study still has limitations and opportunities for future studies. First, the data used in this study are the original sampling data of geochemical exploration variables. However, due to the way geochemical data are measured, they carry the closure problem (Chayes, 1971). In the future, we will carry out suitable logratio transformation of geochemical exploration data to address the closure problem and thus further improve the accuracy of using GAUGE in extracting geochemical patterns for anomaly recognition. Second, compared to CNNs, graph learning is slower in extracting geochemical patterns for a larger quantity of data. Although GAT speeds up somewhat, its speed is intermediate in the range of comparative models. To the best of our knowledge, this is the first time that graph learning is introduced into the field of geochemical exploration. Graph deep learning can help a deep learning model applied to the recognition of geochemical anomalies in irregular areas. However, graph deep learning in other directions of geochemical exploration (e.g., prospectivity mapping) needs to be verified and further optimized. In the future, we believe that increasingly more research will be carried out on graph learning.

In general, this study introduces graph learning into geochemical feature extraction for the first time. A geochemical anomaly recognition framework (GAUGE) was constructed based on graph deep learning. Additionally, we innovatively proposed converting geochemical sampling point data into geochemical graphs. By constructing an unsupervised graph autoencoder, this study solved the problem that traditional deep learning models cannot extract geospatial patterns in irregular regions. When this framework recognizes anomalies in a regular area, GAUGE has the same advantages as the CAE. Compared to traditional methods, GAUGE considers the spatial pattern in irregular areas, and its model accuracy is higher compared to other models. This research firstly verified the adaptability and usability of graph learning in geochemical anomaly recognition, thereby greatly expanding the application of deep learning models in geochemical anomalies. We expect that graph learning will be broadly applied in future analysis of geochemical data for mineral exploration.

ACKNOWLEDGMENTS

Thanks are due to Dr. John Carranza's, Dr. Renguang Zuo's and two anonymous reviewers' comments and suggestions, which helped us improve this manuscript. This work was supported by the National Key Research and Development Program of China (Grant No. 2019YFB2102903), the National Natural Science Foundation of China (Grant No. 42171466, 41801306 and U1711267), the Scientific Research Program of the Department of Natural Resources of Hubei Province (Grant No.

ZRZY2021KJ02), the MOST Special Fund from the State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences (Grant No. MSFGPMR03-4), the ''CUG Scholar'' Scientific Research Funds at China University of Geosciences (Wuhan) (Grant No. 2022034), and the Zhejiang Provincial Natural Science Foundation (Grant No. LY18D010001).

DECLARATIONS

Conflict of Interest No conflict of interest exists in the submission of this manuscript.

APPENDIX

GLOBAL MORAN'S I METHOD FOR SPATIAL AUTOCORRELATION

Spatial autocorrelation indicates a significant spatial distribution pattern in space through the degree of correlation between spatial objects in a region. Global Moran's I is a common metric for quantitative representation of spatial autocorrelation (Goodchild, 1986). Its mathematical equation is:

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{i,j}(x_i - \overline{x})(x_j - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \quad (15)$$

where x_i, x_j are sampling values at sampling point i and j, respectively; $\overline{x}$ is the mean value; $w_{i,j}$ denotes weights representing the proximity relationship between sampling points i and j. Generally, $w_{i,j}$ is related to the distance band selection, and here we calculated the spatial weights for samples only within distance K. S_0 is the sum of all elements of the spatial weight matrix W. I is Global Moran's I value, which ranges from -1 to 1. The closer it is to 1, the stronger the spatial autocorrelation is.

ACTIVATION FUNCTIONS IN GAT

The activation function is a function that maps inputs to outputs in neurons. It is important for deep learning models to extract and understand complex and nonlinear patterns. Sigmoid, LeakyReLU, and ReLU are used in GAUGE, and their mathematical equations are as follows:

Sigmoid (Finney, 1952):

$$Sigmoid(x) = \frac{1}{1 + e^{-x}} \quad (16)$$

ReLU (Glorot et al., 2011):

$$Relu(x) = \max(0, x) \quad (17)$$

LeakyReLU (Maas et al., 2013):

$$LeakyReLU(x) = \max(ax, x) \quad (18)$$

where x indicates the input of activation function. $Sigmoid(x)$, $Relu(x)$, and $LeakyReLU(x)$ are the output of the respective functions. The $aofLeakyReLU$ defaults to 0.01.

REFERENCES

Ahmad, S., Lavin, A., Purdy, S., & Agha, Z. (2017). Unsupervised real-time anomaly detection for streaming data. *Neurocomputing, 262*, 134–147.

An, J., & Cho, S. (2015). Variational autoencoder based anomaly detection using reconstruction probability. *Special Lecture on IE, 2*(1), 1–18.

Bahdanau, D., Cho, K., & Bengio, Y. (2014). Neural machine translation by jointly learning to align and translate. *arXiv preprint* arXiv:1409.0473.

Barlow, H. B. (1989). Unsupervised learning. *Neural Computation, 1*(3), 295–311.

Beus, A. A., & Grigorian, S. V. (1977). Geochemical exploration methods for mineral deposits. *Earth Science Reviews, 14*(1), 67–69.

Bin, J. I., Zhou, T., Yuan, F., Zhang, D., Liu, L., & Liu, G. (2017). A method for identifying geochemical anomalies based on spatial autocorrelation. *Science Survery Mapping, 42*, 24–27.

Breunig, M.M., Kriegel, H., Ng, R.T., & Sander, J.O.R. (2000). LOF: identifying density-based local outliers. In *Proceedings of the 2000 ACM SIGMOD international conference on Management of data*, 93–104.

Brooks, D. B., & Andrews, P. W. (1974). Mineral Resources, Economic Growth, and World Populatic. *Science, 185*(4145), 13–19.

Cameron, E. M., et al. (2005). Geochemical Exploration. In R. C. Selley (Ed.), *Encyclopedia of Geology* (pp. 21–29). Oxford: Elsevier.

Carranza, E. J. M., & Laborte, A. G. (2015). Random forest predictive modeling of mineral prospectivity with small number of prospects and data with missing values in Abra (Philippines). *Computers & Geosciences, 74*, 60–70.

Chao, T. T. (1984). Use of partial dissolution techniques in geochemical exploration. *Journal of Geochemical Exploration, 20*(2), 101–135.

Chayes, F. (1971). *Ratio correlation: a manual for students of petrology and geochemistry*. University of Chicago Press.

Chen, Z.J., Cheng, Q.M., & Chen, J.G. (2009). Comparison of different models for anomaly recognition of geochemical data by using sample ranking method *Earth Science—Journal of China University of Geosciences*.

Chen, J., Cooke, D. R., Piquer, J. E., Selley, D., Zhang, L., & White, N. C. (2019a). Hydrothermal alteration, mineralization, and structural geology of the Zijinshan high-sulfidation Au-Cu deposit, Fujian Province. *Southeast China. Economic Geology, 114*(4), 639–666.

Chen, L., Guan, Q., Feng, B., Yue, H., Wang, J., & Zhang, F. (2019b). 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. (2019c). A spatially constrained multi-autoencoder approach for multivariate geochemical anomaly recognition. *Computers & Geosciences, 125*, 43–54.

Chen, Y., Lu, L., & Li, X. (2014). Application of continuous restricted Boltzmann machine to identify multivariate geochemical anomaly. *Journal of Geochemical Exploration, 140*, 56–63.

Chen, Y., & Wu, W. (2017). Application of one-class support vector machine to quickly identify multivariate anomalies from geochemical exploration data. *Geochemistry Exploration, Environment Analysis, 17*(3), 231–238.

Cheng, Q. (1999). Spatial and scaling modelling for geochemical anomaly separation. *Journal of Geochemical Exploration, 65*(3), 175–194.

Cheng, Q., Agterberg, F. P., & Ballantyne, S. B. (1994). The separation of geochemical anomalies from background by fractal methods. *Journal of Geochemical Exploration, 51*(2), 109–130.

Cheng, Q., Bonham-Carter, G., Wang, W., Zhang, S., Li, W., & Qinglin, X. (2011). A spatially weighted principal component analysis for multi-element geochemical data for mapping locations of felsic intrusions in the Gejiu mineral district of Yunnan. *China. Computers & Geosciences, 37*(5), 662–669.

Cheng, Q., Xu, Y., & Grunsky, E. (2000). Integrated spatial and spectrum method for geochemical anomaly separation. *Natural Resources Research, 9*(1), 43–52.

Christmann, P. (2018). Towards a more equitable use of mineral resources. *Natural Resources Research, 27*(2), 159–177.

Dikang, X., Zongxia, G., & Huimin, H. (1997). *Ore Formation Model and Mineral Search Model for Copper and Gold Deposits in Southeast China*. China University of Geosciences Press.

Fabrigar, L. R., & Wegener, D. T. (2011). *Exploratory factor analysis*. Oxford University Press.

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters, 27*(8), 861–874.

Finney, D. J. (1952). *Probit analysis: a statistical treatment of the sigmoid response curve*. Cambridge University Press.

Ge, Y., Cheng, Q., & Zhang, S. (2005). Reduction of edge effects in spatial information extraction from regional geochemical data: a case study based on multifractal filtering technique. *Computers & Geosciences, 31*(5), 545–554.

Glorot, X., Bordes, A., & Bengio, Y. (2011). Deep sparse rectifier neural networks. JMLR Workshop and Conference Proceedings, 315–323.

Goodchild, M.F. (1986). *Spatial autocorrelation*. Geo Books.

Guan, Q., Ren, S., Chen, L., Feng, B., & Yao, Y. (2021). A spatial-compositional feature fusion convolutional autoencoder for multivariate geochemical anomaly recognition. *Computers & Geosciences, 1*, 104890.

Hardin, J., & Rocke, D. M. (2004). Outlier detection in the multiple cluster setting using the minimum covariance determinant estimator. *Computational Statistics & Data Analysis, 44*(4), 625–638.

He, Z., Xu, X., & Deng, S. (2003). Discovering cluster-based local outliers. *Pattern Recognition Letters, 24*(9–10), 1641–1650.

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science, 313*(5786), 504–507.

Huang, C., Liu, Q., & Zhang, K. (1999). Geophysical and geochemical characters and ore-finding pattern of the zijinshan copper-gold orefield, in Shanghang County, Fujian Province. Geology of Fujian, 4(1999), 189–201.

Jianhua, D., Jianfu, F., Jiangning, Y., & Yaling, L. (2016). Geological Characteristics and mineral resource potential of the Wuyishan Cu-Pb-Zn polymetallic metallogenic belt. *Acta Geologica Sinica, 90*(7), 1537–1550.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: trends, perspectives, and prospects. *Science, 349*(6245), 255–260.

Kingma, D.P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint* arXiv:1312.6114.

Kipf, T.N., & Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *arXiv preprint* arXiv:1609.02907.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444.

Li, B., & Jiang, S. (2017). Genesis of the giant Zijinshan epithermal Cu-Au and Luoboling porphyry Cu–Mo deposits in the Zijinshan ore district, Fujian Province, SE China: A multi-isotope and trace element investigation. *Ore Geology Reviews, 88*, 753–767.

Li, C., Ma, T., & Shi, J. (2003). Application of a fractal method relating concentrations and distances for separation of geochemical anomalies from background. *Journal of Geochemical Exploration, 77*(2–3), 167–175.

Li, H., Li, X., Yuan, F., Jowitt, S. M., Zhang, M., Zhou, J., Zhou, T., Li, X., Ge, C., & Wu, B. (2020). Convolutional neural network and transfer learning based mineral prospectivity modeling for geochemical exploration of Au mineralization within the Guandian-Zhangbaling area, Anhui Province. *China. Applied Geochemistry, 122*, 104747.

Li, S., Chen, J., & Xiang, J. (2020). Applications of deep convolutional neural networks in prospecting prediction based on two-dimensional geological big data. *Neural computing and applications, 32*(7), 2037–2053.

Li, X., Fan, H., Santosh, M., Hu, F., Yang, K., & Lan, T. (2013). Hydrothermal alteration associated with Mesozoic granite-hosted gold mineralization at the Sanshandao deposit, Jiaodong Gold Province, China. *Ore Geology Reviews, 53*, 403–421.

Lin, J., Tang, G., Xu, T., Cai, H., Lu, Q., Bai, Z., Deng, Y., Huang, M., & Jin, X. (2020). P-wave velocity structure in upper crust and crystalline basement of the Qinhang and Wuyishan Metallogenic belts: constraint from the Wanzai-Hui'an deep seismic sounding profile. *Chinese Journal of Geophysics, 63*(12), 4396–4409.

Liu, F.T., Ting, K.M., & Zhou, Z. (2008). Isolation forest. In *2008 eighth IEEE international conference on data mining*. IEEE, 413–422.

Luo, Z., Xiong, Y., & Zuo, R. (2020). Recognition of geochemical anomalies using a deep variational autoencoder network. *Applied Geochemistry, 122*, 104710.

Maas, A.L., Hannun, A.Y., Ng, A.Y., & Others (2013). Rectifier nonlinearities improve neural network acoustic models. In *International Conference on Machine Learning*, 3.

Mao, J., Zhao, X., Ye, H., Hu, Q., Liu, K., & Yang, F. (2010). Tectono-magmatic mineralization and evolution in Wuyishan metallogenic belt. *Shanghai Geology, 31*(S1), 140–145.

Mathieu, L. (2018). Quantifying hydrothermal alteration: A review of methods. *Geosciences, 8*(7), 245.

Matschullat, J. O. R., Ottenstein, R., & Reimann, C. (2000). Geochemical background–can we calculate it? *Environmental Geology, 39*(9), 990–1000.

Nai-Zheng, X. U., Mao, J. R., Hai-Min, Y. E., Shen, M. T., Liu, Y. P., & Chen, L. Z. (2008). Geological characteristics and new ore-finding progress in the dapai lead and zinc deposit of yongding county, fujian province. *Geology and Prospecting, 44*(4), 20–23.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., & Others (2019). Pytorch: An imperative style, high-performance deep learning library. *arXiv preprint* arXiv:1912.01703.

Porwal, A., Carranza, E., & Hale, M. (2003). Artificial neural networks for mineral-potential mapping: a case study from Aravalli Province. *Western India. Natural Resources Research, 12*(3), 155–171.

Qiu, X.P., Lan, Y.Z., Fuzhou, Fujian, Beijing and Group, Z.M. (2010). The Key to the Study of Deep Mineralization and the Evaluation of Ore-prospecting Potential in the Zijinshan Gold and Copper Deposit. *Acta Geoscientica Sinica, 31*(2), 209–215.

Schölkopf, B., Platt, J. C., Shawe-Taylor, J., Smola, A. J., & Williamson, R. C. (2001). Estimating the support of a high-dimensional distribution. *Neural Computation, 13*(7), 1443–1471.

Singer, D. A., Berger, V. I., & Moring, B. C. (2008). *Porphyry copper deposits of the world: Database and grade and tonnage models: USGS Open-File Report 2008–1155*. USGS: Reston, VA, USA.

So, C., Dequan, Z., Yun, S., & Daxing, L. (1998). Alteration-mineralization zoning and fluid inclusions of the high sulfidation epithermal Cu-Au mineralization at Zijinshan, Fujian Province. *China. Economic Geology, 93*(7), 961–980.

Survey, D.R.C.O., & Survey, F.I.O.G. (2014). *Study on the geological background of mineralization and mineralization pattern of Wuyishan mineralization zone*. Geological Press.

Tang, J., Chen, Z., Fu, A.W., & Cheung, D.W. (2002). *Enhancing effectiveness of outlier detections for low density patterns*. Springer, 535–548.

Tang, Y., Zhao, L., Zhang, S., Gong, C., Li, G., & Yang, J. (2020). Integrating prediction and reconstruction for anomaly detection. *Pattern Recognition Letters, 129*, 123–130.

Tobler, W. (2004). On the first law of geography: a reply. *Annals of the Association of American Geographers, 94*(2), 304–310.

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. *Natural Resources Research, 15*(1), 15–26.

Veli, V.C., Kovi, C, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., & Bengio, Y. (2017). Graph attention networks. *arXiv preprint* arXiv:1710.10903.

Wang, J., & Zuo, R. (2019). Recognizing geochemical anomalies via stochastic simulation-based local singularity analysis. *Journal of Geochemical Exploration, 198*, 29–40.

Wold, S., Esbensen, K., & Geladi, P. (1987). Principal component analysis. *Chemometrics and Intelligent Laboratory Systems, 2*(1–3), 37–52.

Xie, X., Mu, X., & Ren, T. (1997). Geochemical mapping in China. *Journal of Geochemical Exploration, 60*(1), 99–113.

Xiong, Y., & Zuo, R. (2016). Recognition of geochemical anomalies using a deep autoencoder network. *Computers & Geosciences, 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. *Computers & Geosciences, 140*, 104484.

Xiong, Y., & Zuo, R. (2021). Robust feature extraction for geochemical anomaly recognition using a stacked convolutional denoising autoencoder. *Mathematical Geosciences*, 1–22.

Yousefi, M., & Carranza, E. J. M. (2015). Fuzzification of continuous-value spatial evidence for mineral prospectivity mapping. *Computers & Geosciences, 74*, 97–109.

Zaw, K. (2007). Mineral deposit types and metallogenic relations of South China and adjacent areas of Mainland SE Asia: implications for mineral exploration. *Geology.*.

Zhang, Z., Cui, P., & Zhu, W. (2020). Deep learning on graphs: A survey. *IEEE Transactions on Knowledge and Data Engineering*.

Zhang, S., Carranza, E.J.M., Xiao, K., Wei, H., Yang, F., Chen, Z., Li, N., & Xiang, J. (2021a). Mineral prospectivity mapping based on isolation forest and random forest: implication for the existence of spatial signature of mineralization in outliers. *Natural Resources Research*, 1–19.

Zhang, B., Wang, X., Ye, R., Zhou, J., Liu, H., Liu, D., Han, Z., Lin, X., & Wang, Z. (2015). Geochemical exploration for concealed deposits at the periphery of the Zijinshan copper–gold mine, southeastern China. *Journal of Geochemical Exploration, 157*, 184–193.

Zhang, C., & Zuo, R. (2021). Recognition of multivariate geochemical anomalies associated with mineralization using an improved generative adversarial network. *Ore Geology Reviews, 136*, 104264.

Zhang, C., Zuo, R., & Xiong, Y. (2021a). Detection of the multivariate geochemical anomalies associated with mineralization using a deep convolutional neural network and a pixel-pair feature method. *Applied Geochemistry, 130*, 104994.

Zhang, D. Q., Feng, C. Y., Li, D. X., She, H. Q., & Dong, Y. J. (2005). The evolution of ore-forming fluids in the porphyry-epithermal metallogenic system of Zijinshan area. *Acta Geoscientica Sinica, 26*(2), 127–136.

Zhang, S., Carranza, E. J. M., Wei, H., Xiao, K., Yang, F., Xiang, J., Zhang, S., & Xu, Y. (2021b). Data-driven mineral prospectivity mapping by joint application of unsupervised convolutional auto-encoder network and supervised convolutional neural network. *Natural Resources Research, 30*(2), 1011–1031.

Zhao, P. D. (2002). Three-component'' quantitative resource prediction and assessments: theory and practice of digital mineral prospecting. *Earth Science-Journal of China university of Geosciences, 27*(5), 482–489.

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.

Zong, B., Song, Q., Min, M.R., Cheng, W., Lumezanu, C., Cho, D., & Chen, H. (2018). Deep autoencoding gaussian mixture model for unsupervised anomaly detection. In *International conference on learning representations.*.

Zuo, R., Kreuzer, O.P., Wang, J., Xiong, Y., Zhang, Z., & Wang, Z. (2021). Uncertainties in GIS-based mineral prospectivity mapping: Key types, potential impacts and possible solutions. *Natural Resources Research*, 1–21.

Zuo, R. (2017). Machine learning of mineralization-related geochemical anomalies: a review of potential methods. *Natural Resources Research, 26*(4), 457–464.

Zuo, R., & Carranza, E. J. M. (2011). Support vector machine: a tool for mapping mineral prospectivity. *Computers & Geosciences, 37*(12), 1967–1975.

Zuo, R., & Xiong, Y. (2020). Geodata science and geochemical mapping. *Journal of Geochemical Exploration, 209*, 106431.

Zuo, R., Xiong, Y., Wang, J., & Carranza, E. J. M. (2019). Deep learning and its application in geochemical mapping. *Earth-Science Reviews, 192*, 1–14.