Geoscience Frontiers 17 (2026) 102248

Contents lists available at ScienceDirect

Geoscience Frontiers

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

Research Paper

A causal-aware artificial intelligence framework for mineral prospectivity mapping

Zitong Zhang [a,b], Jianping Chen [c,*], Fujie Jiang [a,b]

[a] State Key Laboratory of Petroleum Resources and Engineering, China University of Petroleum (Beijing), Beijing 102249, China
[b] College of Geoscience, China University of Petroleum (Beijing), Beijing 102249, China
[c] School of Earth Sciences and Resources, China University of Geosciences, Beijing 100083, China

ARTICLE INFO

Article history:
Received 20 June 2025
Revised 10 November 2025
Accepted 3 January 2026
Available online 5 January 2026

Keywords:
Mineral prospectivity mapping
Artificial intelligence
Causal feature
Machine learning
Granger causality
Geostatistics

ABSTRACT

Mineral prospectivity mapping (MPM) plays a vital role in locating potential mineral resources and informing exploration planning. Although data-driven machine learning (ML) techniques have demonstrated effectiveness in handling complex geological patterns, they often depend purely on statistical dependencies, ignoring the underlying causal mechanisms responsible for mineralization. This shortcoming can introduce spurious links and reduce the robustness of the model. In this study, we introduce a new causal-aware artificial intelligence (CausalAI) framework tailored for MPM, which incorporates both traditional metallogenic features and their causal interdependencies as input data. Firstly, we propose the Spatial Granger Causality (SGC) model to evaluate the causal relations between geochemical elements, drawing inspiration from the principles of geostatistics and the Granger causality. Secondly, we design a causal-aware feature extractor (Causactor) to mitigate spurious correlations between predictor variables and predicted responses, facilitating causal feature extraction. Finally, causal-guided ML methods are developed to predict mineralization and classify mineral resource types. A case study in the northern Cu-polymetallic metallogenic belt of the Sanjiang region, southwestern China, demonstrates the high performance of our framework for MPM. Overall, the development of the proposed CausalAI framework effectively integrates causal relations throughout the modeling process, promoting the synergy between causal inference and Earth sciences.

1. Introduction

Mineral prospectivity mapping (MPM) is an indispensable geospatial tool for delineating and prioritizing areas with a high potential for mineral deposits (Li et al., 2023b). As such, the development of innovative approaches capable of reliably forecasting and analyzing mineralization behaviour in regions exhibiting strong potential is of paramount importance (Wang et al., 2024b).

The methodologies underpinning MPM are derived from two principal strands of geoscience: knowledge- and data-driven approaches (Carranza et al., 2008a,b; Zuo and Xu, 2023). Knowledge-driven methods, which emphasize the metallogenic background and mineralization rules, are particularly effective in greenfield or district-scale regions. Expert knowledge is vital in these contexts, guiding exploration efforts with a nuanced understanding of geological backgrounds (Sun et al., 2019; Wang et al., 2024b). However, these methods often rely heavily on expert judgment and prior knowledge, which can introduce a degree of subjectivity and may not fully capture the geological structures and mineralization rules in poorly understood or underexplored regions.

The advancement of data science has ushered in what is known as the fourth paradigm of data-intensive science (Peters-Lidard et al., 2017). Emerging data-driven methods have become a significant trend, yielding numerous breakthroughs in MPM (Agterberg, 2011). Machine learning (ML) is at the core of these advancements, a branch of artificial intelligence (AI) that equips computers to model complex, hidden relationships. These methods, including logistic regression (LR), support machine vector (SVM),random forest (RF), artificial neural network (ANN), and other deep learning methods, are adept at analyzing the complex interplay between known mineral deposits and prospecting data (Li et al., 2023b; Liu et al., 2023; Lauzon and Gloaguen, 2024; Yang et al., 2024). Among these methods, ensemble learning algorithms such as RF,

* Corresponding author.
E-mail addresses: cugb_zzt@163.com (Z. Zhang), 3s@cugb.edu.cn (J. Chen), jfjhtb@163.com (F. Jiang).

https://doi.org/10.1016/j.gsf.2026.102248

extreme gradient boosting (XGBoost) (Yu et al., 2024), and LightGBM (LGBM) (Li et al., 2023a), alongside deep learning (DL) methods, are popular. Ensemble learning algorithms are favoured for their strong stability and generalization capabilities, achieved by harnessing the wisdom of crowds. Deep learning methods excel in modelling abstract, non-linear spatial associations between geological factors and mineral deposits.

Recent research has introduced DL methods into MPM, benefiting from their superior feature extraction capabilities, such as convolutional neural networks (CNNs), long short-term memory networks (LSTMs), and transformers. However, the black box nature of DL often results in a lack of confidence among exploration geoscientists regarding the applicability of these methods, and the extensive data requirements limit their widespread use in MPM (Hronsky and Kreuzer, 2019; Mou et al., 2023). Fundamentally, data-driven methods aim to model the inter-correlations between geological factors (predictor variables) and known mineral deposits (predicted responses), converting prospecting big data into mappable criteria and integrating these into comprehensive mineral potential maps (Zuo, 2020). Although the correlations here may suggest an underlying causal relationship, they are insufficient for establishing causality. Here, a causal relationship means that certain geological factors actively contribute to the mineralization process rather than merely co-varying with it. Correlation indicates a synchronous increase or decrease trend in variable changes, whereas a causal relationship is more stringent, requiring directionality and strength (Altman and Krzywinski, 2015; Yao et al., 2021). In other words, the core distinction between causal and correlational inference lies in that the former examines how effect variables respond to a change in the cause (Pearl, 2009; Wang et al., 2024a).

Recent reviews of causal inference frameworks and their applications in Earth system sciences underscore their transformative potential (Ebert-Uphoff and Deng, 2017; Pérez-Suay and Camps-Valls, 2018; Runge et al., 2019). Furthermore, Delforge et al. (2022) demonstrated the efficacy of causal inference methods (CIMs) in identifying hydrological connections from time series data, showcasing that multivariate nonlinear CIMs are particularly potent in discerning complex geoscientific relationships. Specific applications of causal inference also include attributing specific climatic events such as volcanic eruptions (Wilson and Russell, 2020), determining the causes of extreme weather events like hurricanes, droughts, and floods (Perkins-Kirkpatrick et al., 2022), and providing new insights into the selection of geochemical pathfinder elements via causal discovery algorithm (Luo and Zuo, 2025), among others.

In MPM, geoscientists diligently explore the relationships between mineralization and various geological factors, using geological reasoning supported by observation and experimentation, driven by the quest for causal relationships. Moving from mining correlations to uncovering causal relationships marks a pivotal advancement in mineral exploration. Rapid progress in causal inference and AI communities has sparked significant growth, merging these fields into a robust and expanding domain (Cui and Athey, 2022; Brand et al., 2023). The geological prospecting data used in MPM typically consist of observational data. Causal inferences are often drawn from such data in the absence of experimental control and without reliance on underlying mechanistic explanations. In such cases, Granger (1969) introduced a definition of causality in which the past values of a time series y_t are evaluated based on how effectively they can predict the future values of another series x_t. In his seminal paper, Granger defined y to be causal for x. This framework is founded on the notion that predictability implies causality, a principle now referred to as Granger causality (Shojaie and Fox, 2022). Today, Granger causality is widely used to analyze time series data across various fields, including economics, genomics, neuroscience, and particularly geoscience. For example, Kodra et al. (2011) employed a reversed-form Granger causality approach on the paired time series composed of globally averaged land surface temperatures and atmospheric CO_2 concentrations. Similarly, Mosedale et al. (2006) employed the Granger causality modeling approach to quantitatively assess how daily variations in sea surface temperature affect the short-term variability of the North Atlantic Oscillation, as simulated through a fully coupled general circulation system. Clearly, causal inference approaches offer notable potential in advancing methodological innovation within the MPM domain.

In order to close the gap between geoscientists and engineers working on causal inference techniques, we introduce a pioneering causal-aware AI framework, termed CausalAI. This framework is designed to enhance the power of ML models in MPM by embedding causal insights, which ensures the features extracted from the prospecting data are significant not only in terms of correlation but also causality. The main contributions can be summarized as follows.

(1) We propose the CausalAI framework, embedding explicit causal insights derived from spatial causality analysis. To the best of our knowledge, this is the first integration of a spatially explicit causal inference method into MPM, enabling conventional ML models to leverage geologically meaningful causal features rather than relying solely on correlations.
(2) Based on geostatistics principles, we develop the spatial Granger causality (SGC) model, a novel method grounded in geostatistics that extend traditional Granger causality to spatial data. This methodological advance allows the identification of causal relationships among geochemical elements under spatial dependence, providing a robust theoretical tool for mineralization studies.
(3) A meticulously designed causal-aware feature extractor (Causactor) is proposed to emphasize causality over mere correlation in prospecting data. By transforming SGC results into causal features, it enhances ML predictions and improves interpretability by linking outcomes to metallogenic processes.
(4) The effectiveness of the CausalAI framework is demonstrated in the northern Cu-polymetallic belt of the Sanjiang region in southwestern China.

The outline of this paper is as follows: Section 2 describes the geological background of the study area and data source. Section 3 provides detailed descriptions of the proposed CausalAI framework. Section 4 presents the experimental results. Section 5 concludes this paper and points out future lines.

2. Study area and data source

2.1. Geological background

The Sanjiang region, encompassing the three rivers of Lancangjiang, Nujiang, and Jinshajiang, is positioned in the eastern sector of the Qinghai-Tibetan Plateau (Fig. 1a). This region, where the Eurasian and Indian Plates converge, belongs to the eastern segment of the Tethys tectonic realm, exhibiting a north-northwest to nearly north–south orientation (Deng et al., 2014a). Its formation witnessed a series of tectonic evolutions and provided key information on the tectonic dynamics, such as the closure of the Paleo-Tethys Ocean and the amalgamation of the peri-Gondwanan micro-continental blocks and Paleozoic are terranes (Cocks and Torsvik, 2013; Deng et al., 2014b).

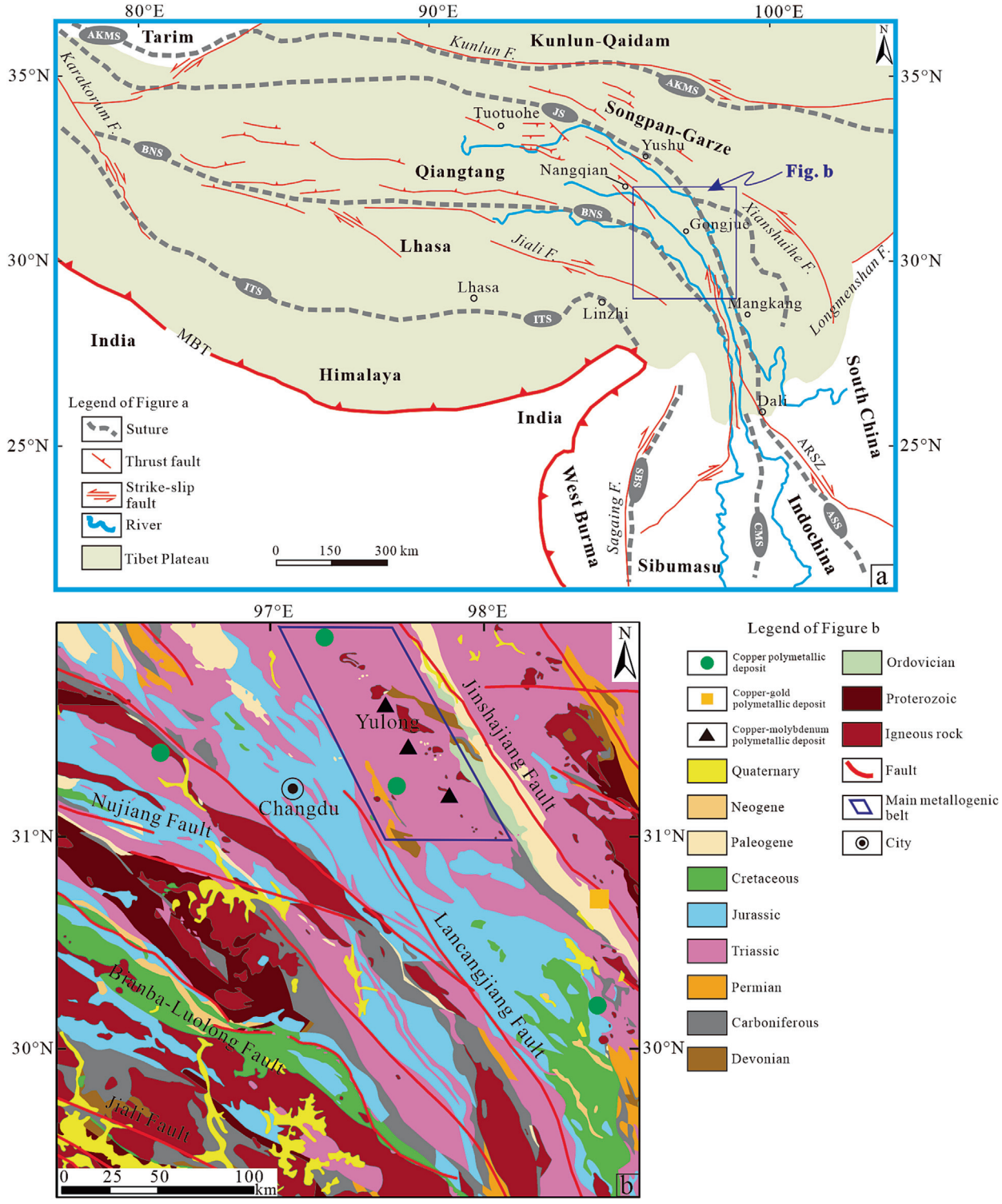

Fig. 1. Simplified geological maps of the study area. (a) Simplified tectonic map of the Tibetan plateau (adapted from Deng et al. (2014b)); (b) Sanjiang region and geological map of the study area. Full name of abbreviations: AKMS, Anyimaqin-Kunlun-Muztagh suture; JS, Jinshajiang suture; BNS, Bangong-Nujiang suture; ITS, Indus-Tsangpo suture; MBT, Main Boundary thrust; ARSZ, Ailaoshan-Red River shear zone.

The Changdu intracontinental rift-valley fold belt, along with its associated strike-slip thrust tectonic zone, is located in the northern area of the Sanjiang region. This area is characterized by substantial Cu-polymetallic ore deposits (Hou et al., 2004). The distribution of Cu-polymetallic ore deposits congruence implies a close correlation between the formation of these ore deposits and the intracontinental orogeny (Tang et al., 2006). Most porphyry copper ores are concentrated in the Yulong metallogenic belt (Fig. 1b), which is positioned west of the Jinshajiang Fault and east of the Lancangjiang Fault. Tectonic control is the main feature of mineralization in the area. The Jinshajiang and Lancangjiang Faults are significant ore-conducting faults, and the tensional thrust

nappe faults that intersect directly with the main tectonic structures play a crucial role in mineralization as ore-bearing faults (Hou et al., 2003; Tang et al., 2006).

Notably, the Yulong ore belt represents China's most extensive Cenozoic porphyry copper mineralization zone (Shen, 2021). Accordingly, this study focuses on the Yulong mineralization zone and its nearby regions. The dominant metamorphic crystalline basement within the study region is composed of the Middle Proterozoic Jitang Group and Ningduo Group, consisting mainly of schist and gneiss (Zhang et al., 2022b). The stratigraphic sequence is relatively comprehensive, spanning from the Lower Paleozoic to the Quaternary. Remarkably, the principal mineralized stratum is the Triassic carbonate rocks (Hou et al., 2003). The folds in the area are generally NW-SE oriented and compound dipping (Zhang et al., 2022b), and Triassic strata can be observed to be exposed on both sides of the dips (Fig. 1b). The largest intrusive rocks in this area were formed during the Indosinian and Yanshanian periods and mainly consist of granodiorite and diorite, which are S-type granite (Tao et al., 2011). The most relevant to mineralization are the Xishanian intrusive rocks, which are mainly dominated by dioritic granodiorite and orthogneiss (Hou et al., 2003, 2004; Tang et al., 2006).

2.2. Data source

Geochemical survey data serve as a crucial information source for MPM. The dataset applied in this study was derived from the 1:200,000-scale National Geochemical Mapping Program of China (Xie et al., 1997). In the targeted region, the sampling strategy involved one point per 4 km^2. Each collected sample contains concentration data for 39 chemical components, including both major and trace elements. To mitigate the impact of high dimensionality commonly encountered in AI-based modeling, we selected 18 representative elemental indicators for analysis, covering Au, Ag, As, Bi, Cu, Co, Cd, Cr, Hg, Mo, Mn, Ni, Pb, Sb, Sn, Ti, W, and Zn. This selection is guided by the regional metallogenic setting of the study area, where porphyry and hydrothermal deposits predominate (Hou et al., 2003), and classic prospecting indicators for such systems, with redundant elements further excluded via Pearson correlation coefficient analysis (Zhang et al., 2025) to ensure the retained set captures key processes including magmatism, fluid alteration, and mineralization enrichment.

To visualize the spatial distribution of these selected indicators, we adopted the inverse distance interpolation approach to produce raster maps, each constructed with a spatial resolution of 1 km by 1 km. Fig. 2 illustrates four thematic evidence layers derived from selected geochemical variables, such as Au, Ag, Cu, and Mo, which are broadly acknowledged to exhibit strong associations with Cu-polymetallic mineralization systems. In total, 18 spatial evidence layers, composed of 22,072 individual grid cells, were generated from the geochemical dataset used for this research.

Meanwhile, a total of 153 labeled positive instances (denoted as "1") are used to construct the supervised dataset for AI model training, obtained from the "GeoCloud" platform, which is a national-level geological information service platform led by China Geological Survey. These positive instances comprise 90 mineralized spots, 35 mineral occurrences, and 28 mineral deposits. In this context, mineralized spots refer to localized zones with direct traces of mineralization, mineral occurrences are confirmed locations of useful minerals with a broader scope than spots; and mineral deposits denote concentrated mineral zones that meet industrial mining standards. Negative instances (each labeled as "0") are drawn according to the selection criteria by Carranza et al. (2008c): (a) locations must be spatially random, (b) locations must be distal to all known deposits, (c) the number of negative instances should be equal to that of positive instances. Point pattern analysis is employed to determine a distance sufficiently far from positive instances in this study (Zuo and Carranza, 2011). As illustrated in Fig. 3a, we calculate and statistically plot the distances between each positive instance and its nearest neighbouring positive instance. The results show that the distance between 95% of positive instance pairs is within 25.05 km. This means that if one starts from any given positive instance, the probability of another positive instance appearing within this range is 95%, so this area is suitable for choosing negative instance locations. Therefore, 153 negative instances are randomly selected beyond this buffer distance from positive instance locations, as shown in Fig. 3b.

In this study, we perform hierarchical supervised classification tasks utilizing the constructed supervised dataset, including the classification of positive versus negative samples, as well as the classification of mineral resource types (i.e. mineralized spots, mineral occurrences, and mineral deposits).

3. Methods

The proposed CausalAI framework for MPM as proposed is presented in Fig. 4. It is structured three primary components: the acquisition and preprocessing of experimental datasets, identification of geochemical variables with causal correlation using the SGC model, and the extraction of causal features through a specialized causal-aware feature extractor, named Causactor.

3.1. Data preprocessing

Prior to the model construction, it is imperative to preprocess the raw data collected as described in Section 2.2. Geochemical data are inherently compositional, leading to a closure problem that can induce spurious correlations among variables. To facilitate accurate staticstical analysis and support subsequent experimental endeavors, the centered log-ratio (clr) transformation (Aitchison, 1982) is employed to address the closure problem inherent in the original geochemical data, thus enabling more reliable causal statistical analyses and feature extraction. The clr transformation is defined as follows:

$$w_i = ln \frac{g_i}{\prod_{j=1}^{P} g_j^{1/P}}, i = 1, 2, \cdots, P \tag{1}$$

where g_i and w_i represent the original and clr-transformed variables, respectively, and P denotes the number of geochemical variables.

Furthermore, to evaluate model performance robustly, a stratified hold-out validation is adopted (Yadav and Shukla, 2016). In each iteration, a fixed proportion of samples is randomly selected (with stratification to maintain the class ratio) for training, and the remaining samples are used for testing. This process is repeated five times with different random splits, and the average performance across the five runs is reported to ensure result stability and robustness. Compared with conventional k-fold cross-validation (Wong and Yeh, 2019; Nagasingha et al., 2024), the stratified hold-out validation offers greater flexibility in adjusting the train-test ratio. After extensive testing, the train-test ratios are set to 20%/80% for the binary classification task and 50%/50% for the multi-class mineral resource type classification. Such a procedure not only reduces the risk of overfitting but also provides a more reliable and stable estimate of the predictive capability of the models.

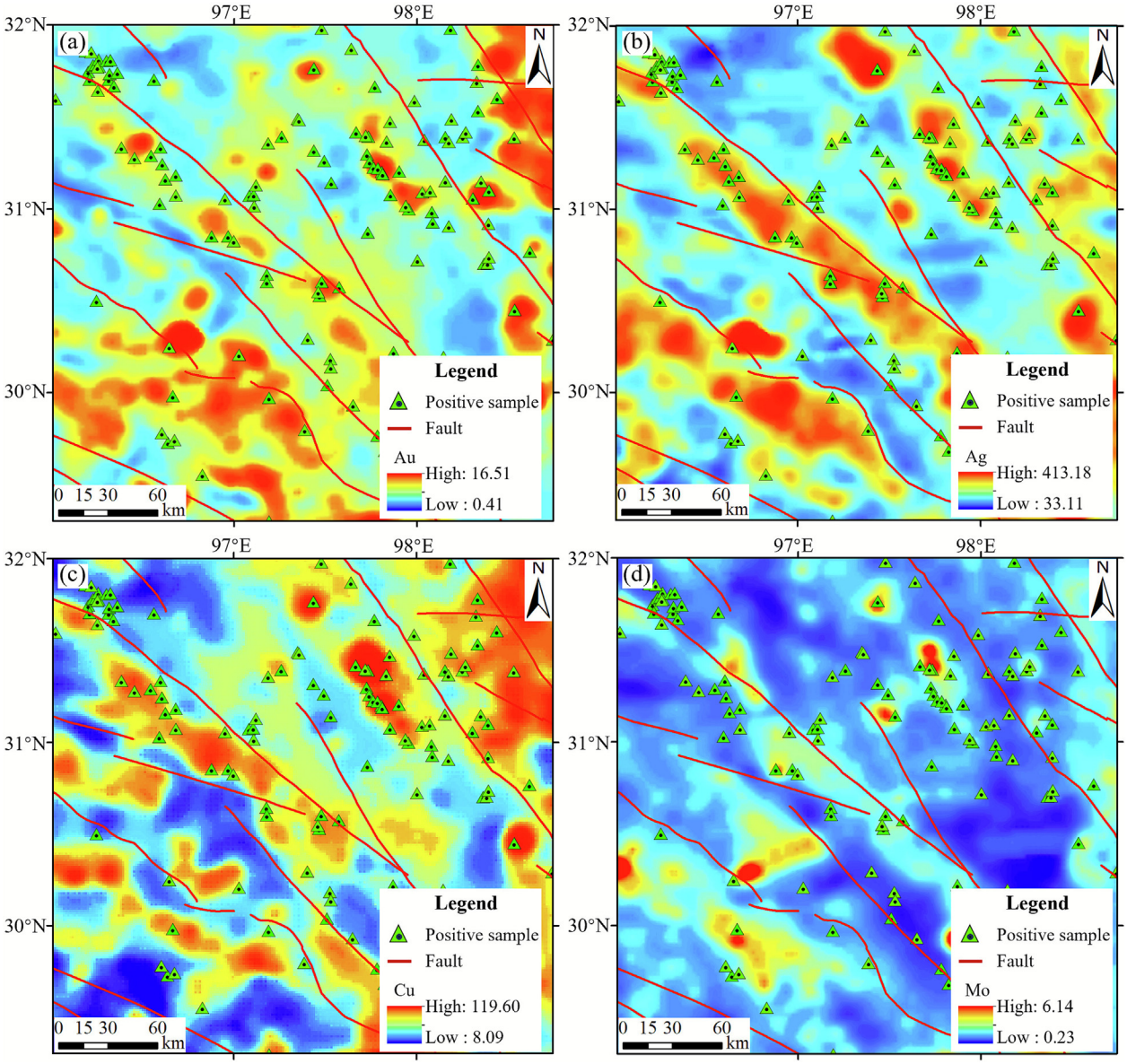

Fig. 2. Geochemical evidence layers represented by (a) Au; (b) Ag; (c) Cu; (d) Mo.

3.2. Preliminaries

3.2.1. Granger causality

Granger (1969) formulated a definition of causality grounded in the idea that past observations of a time series $y_{<t}$ may improve the ability to predict subsequent values of another series x_t. Denote by $H_{<t}$ the complete set of available information up to time $t-1$, and let $P(x_t|H_{<t})$ represent the best forecast of x_t conditioned on $H_{<t}$. Granger defined y as causal for x if the following inequality holds:

$$var[x_t - P(x_t|H_{<t})] < var[x_t - P(x_t|H_{<t}\{y_{<t}\})] \quad (2)$$

where $H_{<t}\{y_{<t}$ is defined as the set $H_{<t}$ with the past values of $y_{<t}$ removed. In other words, y is causal for x if incorporating the historical data of y reduces the variance of the best prediction error for x. Informally, if the past values of y improve the prediction of x, then y is defined as causal for x. Granger further hypothesized that, under certain assumptions, if y can predict x, an underlying causal mechanism must exist, implying that predictability is tantamount to causality. This concept is now widely recognized as Granger causality (Shojaie and Fox, 2022).

Although the above definition may appear somewhat broad and is not tied to any particular modeling assumption, Granger's initial rationale was based on the identifiability of an individual linear model. Let $x_t = (x_{1t}, x_{2t}, \cdots, x_{pt})^T$ denote a vector of variables at time t. Consider the following vector autoregressive (VAR) model:

$$A^0 x_t = \sum_{k=1}^{l} A^k x_{t-k} + \varepsilon_t \quad (3)$$

where $A^0, A^1, A^2, \cdots, A^l$ denote the $p \times p$ matrices corresponding to different lag terms, with l specifying the maximum lag, and ε_t indicating the vector of residuals.

Initial approaches to detecting Granger causality primarily focused on bivariate settings, often neglecting the effects of additional influencing factors. Considering the following models for x_t,

$$x_t = \sum_{k=1}^{l} a_k x_{t-k} + \varepsilon_{x,t} \quad (4)$$

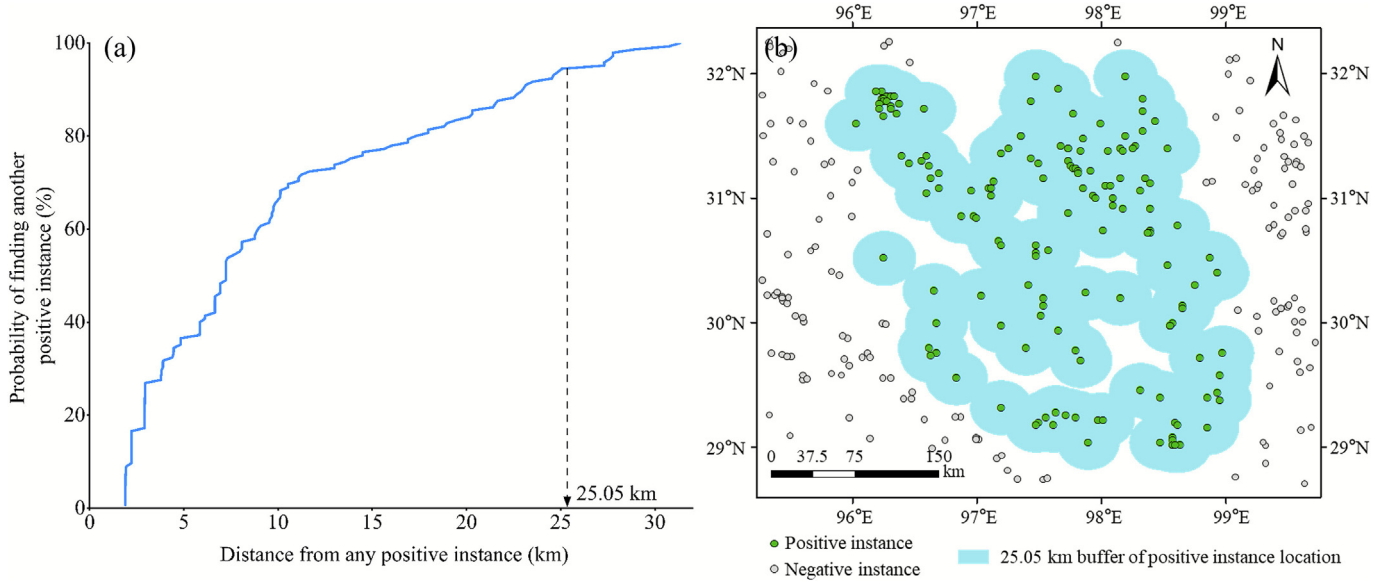

Fig. 3. (a) Result of point patter analysis showing neighboring distance and probability of finding another positive instance from any given positive instance. (b) Locations of positive instances and negative instances.

$$x_t = \sum_{k=1}^{l} b_k x_{t-k} + \sum_{k=1}^{l} c_k y_{t-k} + \varepsilon_{x|y,t} \tag{5}$$

where a_k, b_k, and c_k are the coefficients, $\varepsilon_{x,t}$ and $\varepsilon_{x|y,t}$ are the respective error terms. In this framework, y is said to Granger cause x if and only if $c_k \neq 0$ for some $1 \leq k \leq l$.

Later, Granger causality has been shown to be tested by the results of the F-test applied to two models (Sims, 1972; Chamberlain, 1982): the reduced model (see Eq. (4)) that includes past values of x, and the full model (see Eq. (5)) that includes past values of both x and y. Formally, the F-test is defined as:

$$F = \frac{(RSS_{\text{reduced}} - RSS_{\text{full}})/(r-s)}{RSS_{\text{full}}/(T-r)} \tag{6}$$

where RSS_{reduced} and RSS_{full} refer to the sum of squared residuals from the reduced and full models, respectively, involving r and s parameters (Tiku, 1967). In the F-test, Granger causality from y to x is established when the test statistic F surpasses the threshold value obtained from the F-distribution with degrees of freedom $(r-s)$ and $(T-r)$ under a given significance level (typically α).

3.2.2. Basic geostatistics theories

Regionalized variables are a fundamental concept in geostatistics, primarily used to describe and model the spatial variability of natural phenomena, which refer to random variables that exhibit spatial continuity or spatial correlation within a specific geographic region (Cressie and Chan, 1989). Mathematically, a regionalized variable $z(x)$ can be represented as a single instance of a spatially correlated stochastic process $Z(x)$ with spatial correlation, where x refers to spatial coordinates. It is noteworthy that geochemical survey data inherently exhibit the characteristics of regionalized variables, indicating strong spatial correlations. Assume a geographic region Ω, within which there are d spatial points $\{x_1, x_2, \ldots, x_d\}$ in a neighborhood ω surrounding the estimation point x_0. Let $Z_0 = \left(z_0^{(1)}, z_0^{(2)}, \ldots, z_0^{(p)}\right) \in R^p$ represent the observed regionalized variables at the estimation point.

Cokriging is an extension of the traditional kriging method, which is used for spatial interpolation of a single variable. In cokriging, we deal with two or more spatially correlated variables: one primary variable (usually the target of interest) and one or more secondary variables that provide additional information. The secondary variables are assumed to exhibit some form of spatial dependence with the primary variable. The core idea behind cokriging is that spatially correlated data from a secondary variable can be used to enhance the estimation of the primary variable. The spatial variation between values at locations $z(x)$ and $z(x+h)$ is described by the following formulation:

$$\gamma(h) = \frac{1}{2} var[z(x) - z(x+h)] = \frac{1}{2} E\left[\{z(x) - z(x+h)\}^2\right] \tag{7}$$

where h refers to the lag, representing a specific distance and direction from x, while $\gamma(h)$ denotes the variogram, which is a widely accepted tool for quantifying spatial autocorrelation (Wu and Murray, 2005). Furthermore, if two or more variables are needed, a cross-variogram is defined as follows:

$$\gamma_{uv}(h) = \frac{1}{2} E[\{z_u(x) - z_u(x+h)\}\{z_v(x) - z_v(x+h)\}] \tag{8}$$

Cokriging provides an unbiased estimator while maintaining both minimal and known variance (Curran, 2001). When estimating a target variable u over a block domain Ω, using data from sampled locations of u and another auxiliary variable v, our estimate will be:

$$\hat{z}_u(\Omega) = \sum_{i=1}^{N_u} \lambda_{ui} z_u(x_{ui}) + \sum_{j=1}^{N_v} \lambda_{vj} z_v(x_{vj}) \tag{9}$$

where N_u and N_v denote the number of sampling points corresponding to variables u and v, x_{ui} and x_{vj} mean their spatial sampling positions, and λ_{ui} and λ_{vj} represent the weights to be determined. There are also constraints need to be satisfied to ensure unbiasedness, which is detailed in Wu and Murray (2005). Cokriging constructs an optimal linear unbiased estimator by minimizing the estimation variance:

$$\sigma_u^2(\Omega) = E\left[\{z_u(\Omega) - \hat{z}_u(\Omega)\}^2\right] \tag{10}$$

Eq. (10) formulates an optimization task where λ_{ui} and λ_{vj} serve as decision variables, and the objective function is represented by $\sigma_u^2(\Omega)$. Classical Lagrangian methods are applicable for deriving the solution, as detailed in Myers (1982) and Atkinson et al. (1992).

3.3. Spatial Granger causality model

Traditional Granger causality model is designed for static time series data and cannot be directly applied to dynamic spatial data,

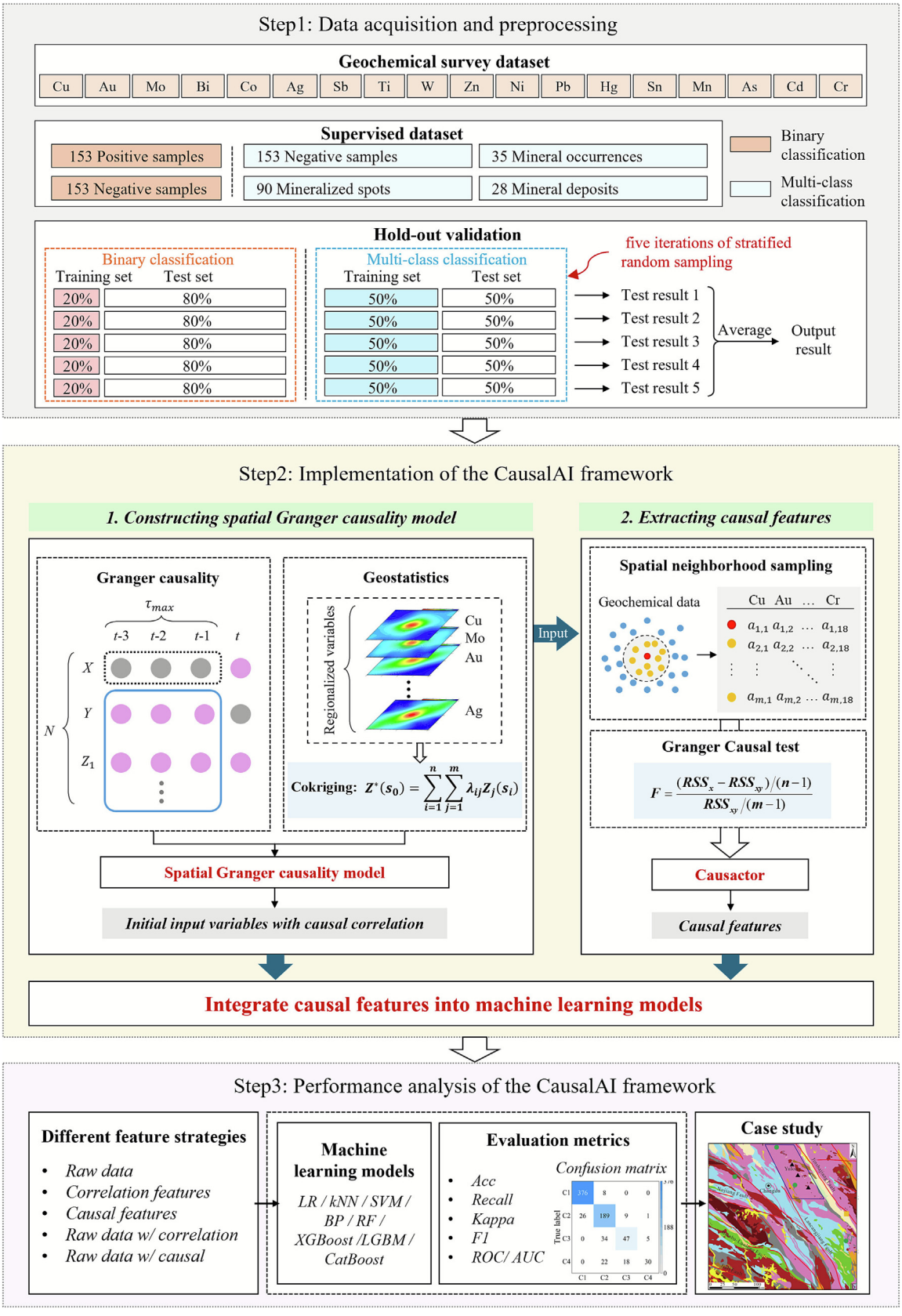

Fig. 4. Schematic illustration of the CausalAI framework for mineral prospectivity mapping (MPM).

such as geochemical survey data. Inspired by cokriging in geostatistics, we extend the Granger causality to spatial data, proposing the SGC model, which provides a more appropriate form for describing and quantifying the causal relationships between regionalized variables.

Specifically, the primary variable u at the estimated location can be predicted using the co-variable $z_k^v(v \neq u, k = 1, 2, \ldots, d)$ at all locations within the neighborhood ω, excluding the estimated point. Let $H_{z_k^u|k\neq 0}$ represent the primary variable u values collected from d different positions around the estimation site x inside the neighborhood region ω. Let $P\left(z_0^u|H_{z_k^u|k\neq 0}\right)$ represent the optimal prediction of z_0^u given $H_{z_k^u|k\neq 0}$. In the SGC model, the co-variable z_k^v is defined as causally related to the primary variable z_0^u if it satisfies the following relationship:

$$var\left[z_0^u - P\left(z_0^u|H_{z_k^u|k\neq 0}\right)\right] < var\left[z_0^u - P\left(z_0^u|\left(H_{z_k^u|k\neq 0} \cup H_{z_k^v|k\neq 0}\right)\right)\right] \tag{11}$$

where $H_{z_k^v|k\neq 0}$ represents the values of the co-variable v at d spatial locations within the neighborhood ω surrounding the estimated location x_0. This implies that by incorporating spatial data of the co-variable z_k^u, the variance of the optimal prediction error for the primary variable z_0^u is reduced, thereby identifying the spatial Granger causality from z_k^v to z_0^u.

Further, by analogy with Eqs. (4)–(5), we can build a regression model of spatial data based on the idea of cokriging method, as follows:

$$z_0^u = \sum_{k=1}^{d} a_k z_k^u + \varepsilon_u \tag{12}$$

$$z_0^u = \sum_{k=1}^{d} b_k z_k^u + \sum_{v\neq u}\sum_{k=1}^{d} c_k z_k^v + \varepsilon_{u|v} \tag{13}$$

where a_k, b_k, and c_k are the coefficients of the regionalized variables, d denotes the order of the regression model, that is, the number of spatial samples considered. ε_u and $\varepsilon_{u|v}$ are the error terms. The co-variable $z_v(x)$ is spatial Granger causal for the primary variable $z_u(x)$ if and only if $c_k \neq 0$ for some $1 \leq k \leq d$.

Similarly to Section 3.2.2, the spatial Granger causality can be obtained by comparing the F-test results about the two models of Eqs. (12) and (13). In this context, the reduced model derived from Eq. (12) represents predictions based solely on the primary variable z_0^u across d spatial locations within neighborhood ω, utilizing the regionalized variable values. Conversely, the full model incorporates both the primary variable z_0^u and the co-variable z_k^v, that captures all regionalized variable values at the same d spatial locations within ω. Regression analysis of these two models allows for the calculation of the corresponding sums of squares of regression residuals, denoted as RSS_{reduced} for the reduced model and RSS_{full} for the full model.

In the hypothesis testing, the null hypothesis is stated as $H_0 : c_1 = c_2 = \cdots = c_d = 0$, which asserts that the co-variable z_k^v does not belong to this regression. The F-test is employed to test this hypothesis, following a formula that adheres to an F-distribution characterized by degrees of freedom d and $n - m$:

$$F = \frac{(RSS_{\text{reduced}} - RSS_{\text{full}})/d}{RSS_{\text{full}}/(n-m)} \tag{14}$$

where n refers to how many samples have been estimated throughout the study region, d represents the count of regionalized variable values related to the primary variable, with respect to how many parameters are involved in RSS_{reduced} refers to the total count of parameters used in RSS_{full}. Utilizing this test, if the F-statistic exceeds the critical value F_α at a chosen significance level α, corresponding to the quantile from an F-distribution defined by degrees of freedom d and $n - m$, the null hypothesis is rejected. This implies that the co-variable does belong to this regression problem, indicating that z_k^v is a spatial Granger causality of z_0^u.

Similarly, in order to test whether other covariables are the spatial Granger causes when other geochemical elements are the primary variables, we can swap the positions of the principal variable and co-variable in the SGC model, and repeat the above steps.

It should be noted that the SGC model inherits the assumptions of the classical Granger framework, namely linearity and (second-order) stationarity of the underlying processes (Granger, 1988). These assumptions may not always strictly hold in geochemical datasets. To address this, clr transformation is applied to mitigate compositional nonstationarity, and the geostatistical concept of regionalized variables is adopted, assuming local rather than global stationarity. Linearity is approximated through regression-based formulations consistent with cokriging, which explicitly capture spatial dependencies. By integrating geostatistical principles into the causality framework, the SGC model accounts for spatial autocorrelation and provides a geologically interpretable extension of Granger causality. While these strategies cannot completely overcome the inherent limitations, they ensure that the SGC analysis remains statistically valid and practically applicable to mineral prospectivity mapping.

3.4. Causal-aware feature extractor

Based on the SGC model, we customized a causal-aware feature extractor (Causactor) specifically for the geochemical survey data. The Causactor converts SGC-derived causal relationships into ML-readable features, shifting the analysis focus from the mere statistical correlations captured by traditional feature engineering to geologically meaningful causal links between elements. This ensures model inputs are aligned with authentic metallogenic processes rather than spurious associations.

Given a sample point $P_0 = \left(P_0^1, P_0^2, \ldots, P_0^\xi\right) \in R^{1\times\xi}$ at spatial coordinate (x, y), where P_0^k represents the value at (x, y) in the k-th variable $(k = 1, 2, \ldots, \xi)$, and ξ corresponds to the total number of elements with causal relationships. Denote $P = \left[P^1, P^2, \ldots, P^\xi\right] \in R^{(\delta+1)\times\xi}$ by the ξ sample points $P_1, P_2, \ldots, P_\delta$ of the spatial nearest neighbors of P_0. The matrix P is used to represent extracted causal features between pairs of variables by the SGC model, specifically to determine whether the variable P^j (the j-th column vector) contributes to the prediction of the variable P^i (the i-th column vector).

Referring to Eqs. (12)–(14), we identify geochemical elements with spatial causal relationships and calculate the F-statistic for these elements to assess their mutual associations. Geochemical data often exhibit dynamic spatial variability, and the F-test is a robust tool for evaluating the significance of lags in the model, so as to determine whether specific elements play an important role in the causal chain. Moreover, the F-statistic provides a succinct and effective measure of causality, with a straightforward testing procedure. By analyzing the lags, the dynamic interactions among variables can be captured and quantified. Consequently, in our framework, the F-statistic is adopted as a key representation of causal features in the Causactor.

In this way, we can construct the causal matrix between the two variables. Unlike correlation, causal matrix is asymmetric, and all calculated causality values are 0 when $i = j$. Therefore, we select all off-diagonal elements in the causality matrix to obtain $\xi \times (\xi - 1)$ features, which represent the spatial causal relationships within a certain neighborhood of the sampling point P_0.

3.5. Machine learning models

The causal features derived in Section 3.4 are integrated with the raw data and then utilized as inputs to the ML models for MPM. We select eight representative ML classifiers to cover the principal families effective for geochemical data and to provide probabilistic outputs required for our analyses: LR as a linear and interpretable baseline; k-nearest neighbour's (kNN) as a non-parametric local method; SVM to capture non-linear decision boundaries via kernels; a back-propagation neural network (BP) as a shallow neural baseline; and four tree-based ensemble methods, namely RF, XGBoost, LGBM, and CatBoost, which are well known for their robustness to heterogeneous features and strong performance on small-to-moderate-sized datasets. All models are tuned with RandomizedSearchCV (scikit-learn 1.5) using five-fold cross-validation to ensure comparable optimization and mitigate overfitting.

3.6. Evaluation metrics

To assess the effectiveness of integrating causal features using machine learning models, evaluation metrics such as accuracy (Acc), recall, kappa coefficient (Kappa), and F1 are calculated as follows (Yin et al., 2023):

$$\mathrm{Acc} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{FP}+\mathrm{TN}+\mathrm{FN}} \tag{15}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{16}$$

$$\mathrm{Kappa} = \frac{\mathrm{ACC}-P_e}{1-P_e}; \quad P_e = \frac{\frac{n}{2}(\mathrm{TP}+\mathrm{FP})+\frac{n}{2}(\mathrm{FN}+\mathrm{TN})}{N\times N} \tag{17}$$

$$\mathrm{F1} = \frac{2\times\mathrm{TP}}{2\times\mathrm{TP}+\mathrm{FP}+\mathrm{FN}} \tag{18}$$

where true positive (TP), true negative (TN), false positive (FP), and false negative (FN) represent the specific outcomes of classification: TP means the actual label is Positive and the model correctly predicts Positive; TN means the actual label is Negative and the model correctly predicts Negative; FP means the actual label is Negative but the model incorrectly predicts Positive; and FN means the actual label is Positive but the model incorrectly predicts Negative. Here, N indicates the total number of samples. To assess model performance, we adopt the receiver operating characteristic (ROC) curve and its corresponding area under the curve (AUC), typically visualized by plotting the true positive rate against the false positive rate(Fawcett, 2006).

4. Experimental results

4.1. Spatial Granger causality between geochemical survey data

In the case study, the SGC model is applied to the northern Cu-polymetallic metallogenic belt of the Sanjiang region in southwestern China to investigate the causal relationships between geochemical elements. Table 1 summarizes the elements in the study area that exhibit causal relationships with respect to Cu. Due to the consistency of results across multiple lags, only the results for lags 1 and 2 are presented. As shown in Table 1, the elements exhibiting spatial causality with Cu include Ag, Au, Co, Mo, Ni, Sb, Zn, As, Cd, W, Nb, Pb, Ti, and Cr. Among these, Ag, Au, Pb, Zn, and Mo, which are the main metallogenic and associated elements in the copper polymetallic deposits of the study area, are accurately identified by the SGC model. Furthermore, this causal relationship is statistically significant, providing both qualitative and quantitative insights into the geochemical elements that are causally linked to mineralization. Based on these findings, the 14 geochemical elements listed above are selected as input data for feature extraction.

Table 1
Summary of spatial Granger causality geochemical elements conforming to Cu.

Lag	Causal relationship	*F*-test	
	causes Cu	*F*-statistic	*p*-value (%)
1	Ag	6.85	0.89
1	Au	9.94	0.16
1	Co	4.63	3.16
1	Mo	8.62	0.34
1	Ni	3.92	2.01
1	Sb	3.21	2.21
1	Zn	5.77	0.32
1	As	4.12	1.64
1	Cd	2.99	2.99
1	W	4.13	0.62
1	Nb	2.86	3.57
1	Pb	5.16	0.15
2	Ag	4.21	0.21
2	Au	3.28	1.08
2	Pb	2.80	2.49
2	Co	2.75	2.68
2	Mo	5.16	0.15
2	W	3.52	0.72
2	Zn	4.08	1.71
2	Ti	4.12	1.64
2	Cd	2.69	1.98
2	Nb	2.24	4.81
2	As	2.17	4.35
2	Sb	2.59	1.68
2	Ni	3.62	0.14
2	Cr	3.82	0.19

4.2. CausalAI framework performance comparison

Eight ML methods and five feature strategies are utilized to predict mineralization. To clarify each feature strategy in Table 2, the core definitions are supplemented as follows: "Raw data" refers to the original geochemical data processed by clr transformation; "Correlation features" are constructed via Pearson correlation coefficients between any two geochemical element groups to characterize synchronous concentration correlation, reflecting only statistical correlation; "Causal features" are extracted via the proposed Causactor, quantifying causal dependence between geochemical elements using the SGC model; "Raw data w/ correlation" combines raw data with the above correlation features, inputting to the model to integrate original and statistical correlation information; "Raw data w/ causal" integrates raw data with the above causal features, providing model inputs more consistent with mineralization mechanisms by integrating original concentration information and inter-element causal relationships.

The comparison results are summarized in Table 2. The results of "Raw data w/ causal" outperform other feature strategies with Acc values ranging from 0.16% to 37.63% higher than the other strategies. The Acc values for each model remain above 92.82%, with the RF achieving exceptional performance at 96.73%, which suggests that integrating causal features introduces additional discriminative information, uncovering deeper causal mechanisms within the data and improving the quality of feature representation.

Fig. 5 illustrates the ROC curves for the binary classification task under different feature strategies across eight ML methods. Overall, most cases exhibit satisfactory predictive performance, with the proposed "Raw data w/ causal" strategy achieving the highest AUC values in both the RF and CatBoost methods. Although the "Raw data w/ correlation" feature strategy also attains AUC values close to or exceeding 0.990 in the tree-based ensemble models

Table 2
Average test set results of various feature strategies utilizing ML models evaluated by five random stratified 20%/80% (train/test) hold-out validation for the binary classification (best results highlighted in bold).

Metrics	Input features	LR	kNN	SVM	BP	RF	XGBoost	LGBM	CatBoost
Acc (%)	Raw data	88.90	81.88	87.51	87.84	95.84	93.06	95.43	96.00
	Correlation features	68.24	70.45	71.76	57.47	69.88	65.47	66.29	69.80
	Causal features	77.88	74.04	76.65	65.80	87.18	83.35	85.55	86.53
	Raw data w/ correlation	91.67	85.55	92.82	92.57	96.33	93.14	95.92	96.16
	Raw data w/ causal	**93.71**	**92.82**	**94.45**	**95.10**	**96.73**	**93.31**	**96.08**	**96.57**
Recall (%)	Raw data	80.59	82.54	86.30	87.93	96.41	89.89	95.76	95.92
	Correlation features	72.92	88.42	82.06	67.54	77.00	65.25	66.23	82.22
	Causal features	89.40	88.91	84.01	73.74	89.56	87.28	89.07	90.38
	Raw data w/ correlation	84.83	84.34	91.84	92.50	95.11	**89.89**	95.27	94.45
	Raw data w/ causal	**93.31**	**89.72**	**92.99**	**95.60**	**96.08**	**89.89**	**96.90**	**96.08**
Kappa (×100)	Raw data	55.75	63.75	75.02	75.67	91.67	86.12	90.86	92.00
	Correlation features	36.48	40.88	43.50	14.92	39.75	30.94	32.57	39.58
	Causal features	77.80	48.07	53.30	31.58	74.37	66.69	71.10	73.06
	Raw data w/ correlation	83.35	71.10	85.63	85.14	92.65	86.29	91.84	92.33
	Raw data w/ causal	**87.43**	**85.63**	**88.90**	**90.20**	**93.47**	**86.61**	**92.16**	**93.14**
F1 (×100)	Raw data	78.47	82.01	87.37	87.86	95.86	92.84	95.45	96.00
	Correlation features	69.68	74.97	74.41	61.38	71.90	65.41	66.29	73.15
	Causal features	88.96	77.41	78.27	68.33	87.49	83.99	86.05	87.04
	Raw data w/ correlation	91.07	85.38	92.75	92.57	96.28	92.92	95.89	96.10
	Raw data w/ causal	**93.69**	**92.59**	**94.37**	**95.13**	**96.72**	**93.07**	**96.12**	**96.56**

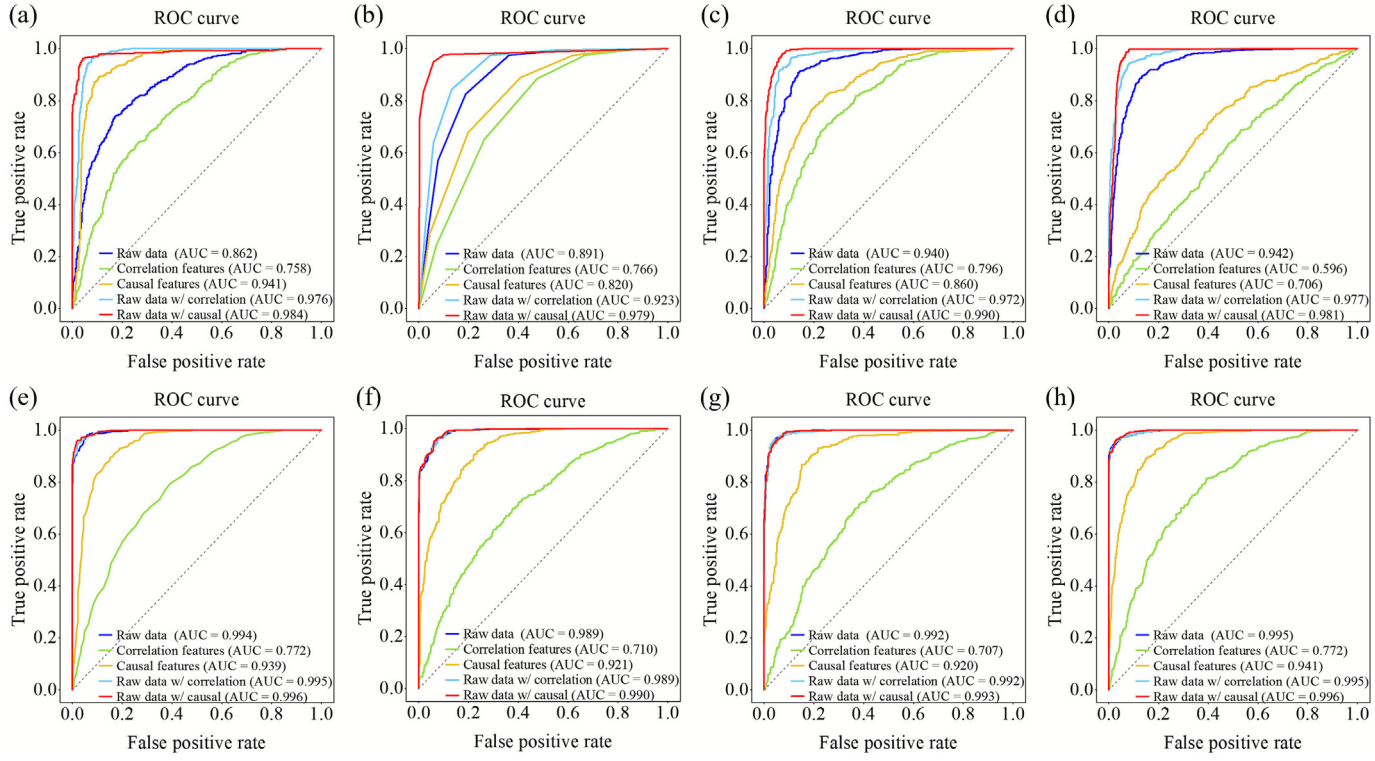

Fig. 5. ROC curves of various ML methods with different feature strategies. (a) LR; (b) kNN; (c) SVM; (d) BP; (e) RF; (f) XGBoost; (g) LGBM; (h) CatBoost.

(e.g., RF, XGBoost), such exceptional results are partly attributed to the inherent strength and high tolerance to overfitting of these methods. When considering the performance across all eight ML methods, the improvements achieved by the "Raw data w/ causal" strategy are more consistent and pronounced, further demonstrating the effectiveness and robustness of the proposed method. Overall, by comparing the performance of all ML methods, RF is identified as the most recommended model for the binary classification task.

Fig. 6 presents the confusion matrices of the binary classification task obtained using the RF method under different feature strategies. As shown, the "Raw data w/ causal" strategy yields the most accurate classification performance, with the smallest number of misclassified samples among all cases. This clearly demonstrates that incorporating causal features enhances the model's discriminative ability and reduces classification errors, further confirming the effectiveness and robustness of the proposed CausalAI framework.

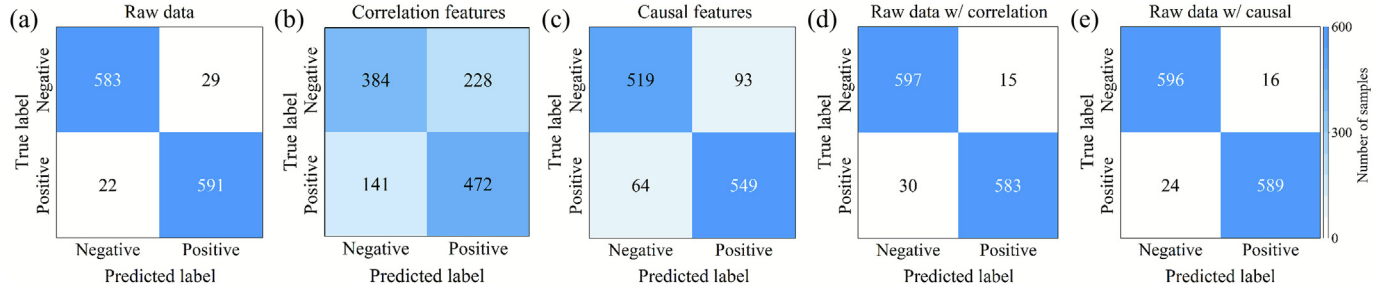

Fig. 6. Confusion matrices of the binary classification using the RF method under different feature strategies. (a) Raw data; (b) Correlation features; (c) Causal features; (d) Raw data w/ correlation; (e) Raw data w/ causal.

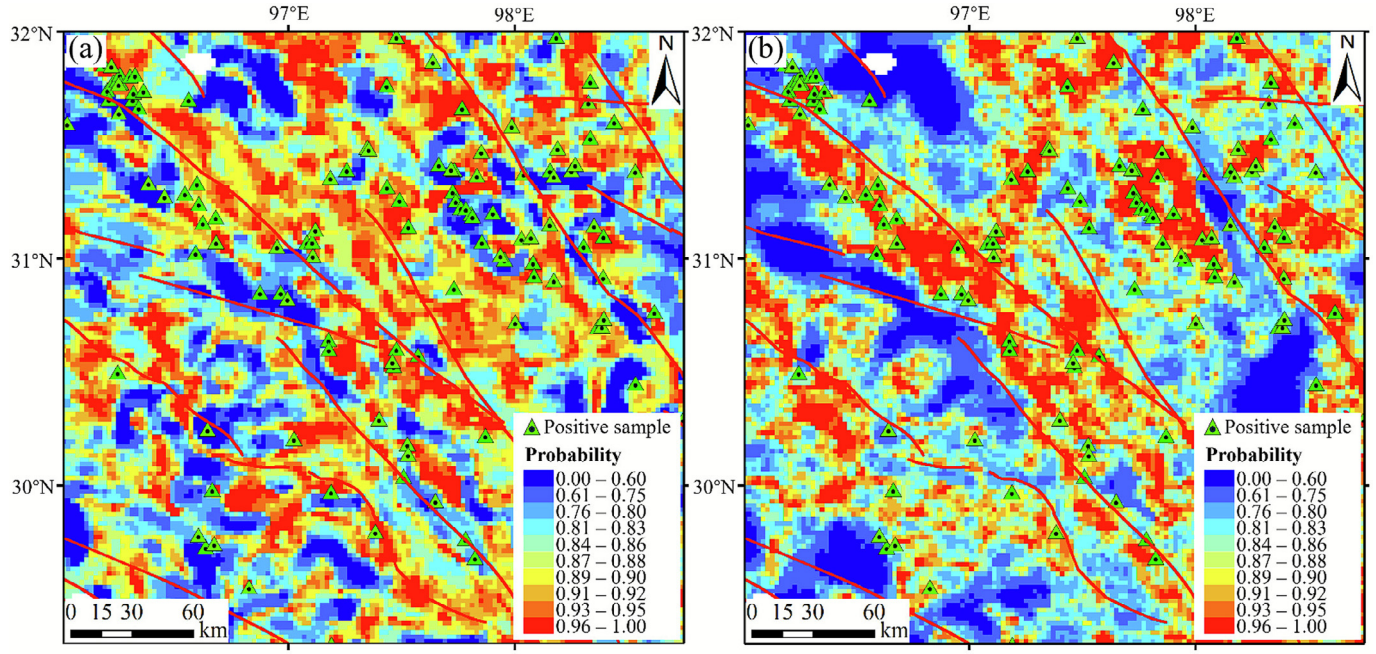

Fig. 7. Mineral prospectivity maps based on the RF method using (a) raw data with correlation features and (b) raw data with causal features.

Fig. 7 illustrates the mineral prospectivity maps generated using the RF model with two different feature strategies. The prospectivity map based on correlation features exhibits dispersed anomaly patterns with poorly defined zoning boundaries, indication a lower spatial coherence, as shown in Fig. 7a. In contrast, according to Fig. 7b, the prospectivity map derived from causal features demonstrates significantly enhanced spatial clustering, more apparent high potential zoning, and a more effective distribution of mineralized areas. Furthermore, there is a strong spatial correlation between the mapped anomalies and the distribution of known positive samples, highlighting the capability and efficiency of the CausalAI framework in mineral prospectivity analysis. Also, the success-rate curves (Rodriguez-Galiano et al., 2015) demonstrate an excellent prediction ability of using causal features in MPM, as displayed in Fig. 8a.

We conduct both quantitative and qualitative comparisons of our CausalAI framework (see Fig. 9a) against three widely-used deep learning methods: CNN, LSTM, and MLP-Mixer (see Fig. 9b–d) (LeCun et al., 2015; Zhang et al., 2022a). Here, the RF method is used to predict with the "Raw data w/ causal" feature strategy (RF_{causal}), which serves as the representative method of the proposed CausalAI framework. For all DL methods, we adopt consistent hyperparameter settings, including a learning rate of 1e−4, the root mean square propagation (RMSprop) optimizer, an input window size of 21 (considering the nearest 20 samples for distance calculation), and a batch size of 64. As shown in Fig. 8b, all five methods successfully identify approximately 55% of the known positive samples, covering about 40% of the study area. Notably, the RF_{causal} method demonstrates superior performance by capturing 80% of the known positive samples within 60% of the study area. Additionally, Fig. 9 further corroborates the exceptional efficacy of our proposed CausalAI framework.

We further categorize the positive samples into three types of mineral resources: mineralized spots, mineral occurrences, and mineral deposits, and utilize the CausalAI framework for multi-class classification, as detailed in Table 3. Consistent with the results observed in the binary classification task, the classification performance under the "Raw data w/ causal" strategy surpasses that of other feature strategies, further verifying the effectiveness of incorporating causal features. Quantitatively, Table 3 shows that the "Raw data w/ causal" strategy achieves the highest Acc (86.27%), representing notable improvements of 0.39%–10.07% compared with the raw data with the "Raw data w/ correlation" strategy. Among the ML methods evaluated, the CatBoost model exhibits the most stable and superior performance across all metrics, confirming it as the most suitable classifier for this scenario.

The confusion matrices shown in Fig. 10 provide further insight into these results. Due to the relatively large number of negative

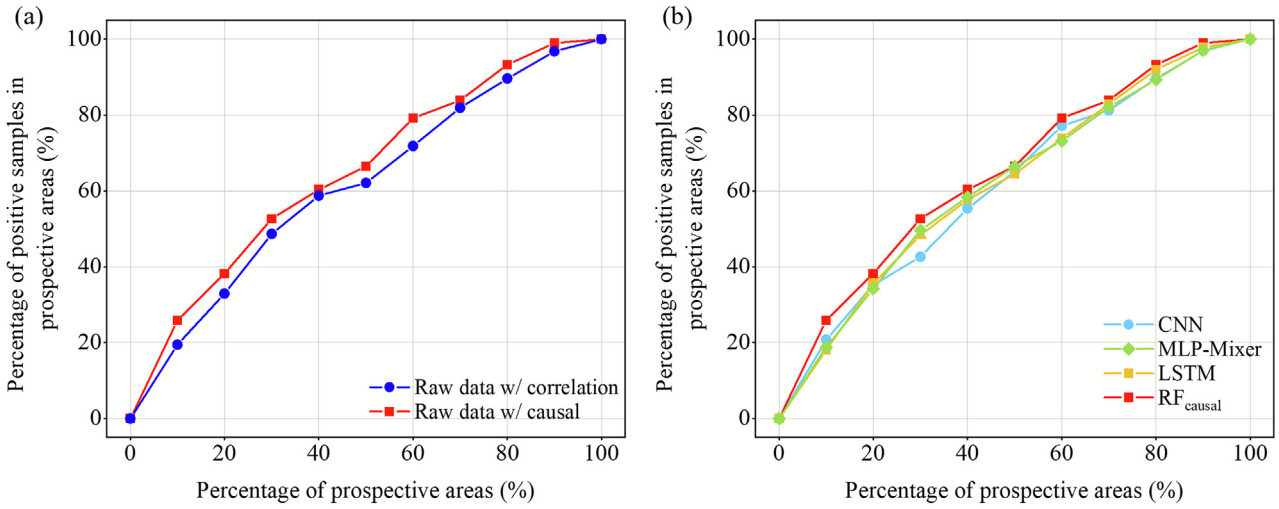

Fig. 8. Success-rate curves obtained by (a) the RF method using different feature strategies, and (b) a comparison between the RF method using the "Raw data w/ causal" feature strategy (RF_{causal}) and convolutional neural network (CNN), MLP-Mixer, and long short-term memory network (LSTM).

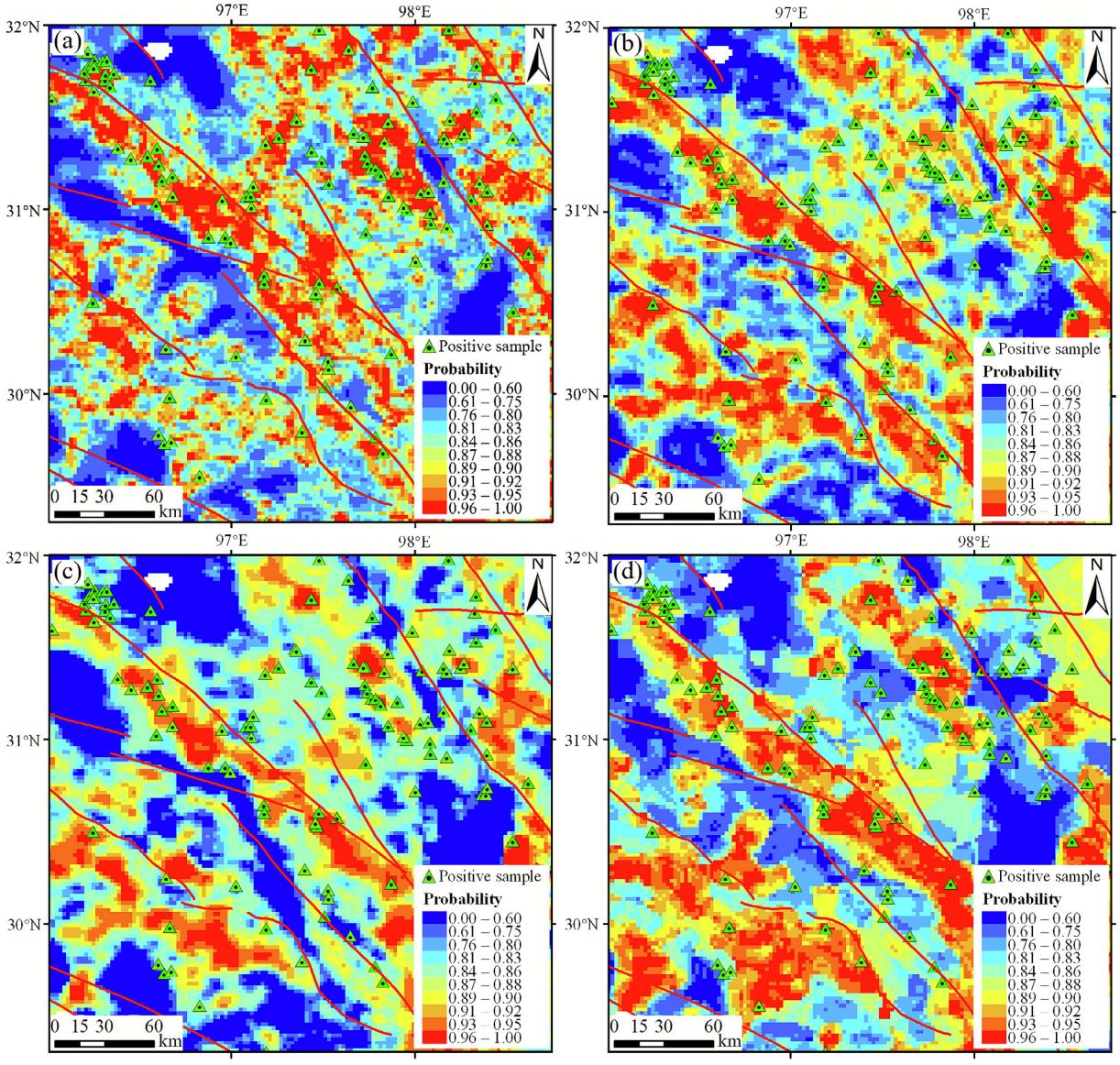

Fig. 9. Mineral prospectivity maps obtained by (a) RF causal method, (b) CNN, (c) LSTM, and (d) MLP-Mixer.

Table 3
Average test set results of various feature strategies utilizing ML models evaluated by five random stratified 50%/50% (train/test) hold-out validation for the classification of mineral resource types (best results highlighted in bold).

Metrics	Input features	LR	kNN	SVM	BP	RF	XGBoost	LGBM	CatBoost
Acc (%)	Raw data	74.51	72.42	74.90	74.12	76.73	83.01	78.30	83.92
	Correlation features	51.24	57.25	56.08	43.01	56.21	55.95	55.42	58.43
	Causal features	58.82	59.74	60.52	50.59	76.60	76.86	76.73	77.52
	Raw data w/ correlation	76.34	74.90	79.08	79.61	84.44	84.58	82.88	85.88
	Raw data w/ causal	**80.13**	**84.97**	**85.23**	**83.01**	**85.75**	**84.97**	**83.27**	**86.27**
Recall (%)	Raw data	57.89	54.31	54.20	61.17	60.05	71.42	64.20	69.86
	Correlation features	32.51	37.40	31.93	26.79	31.31	32.91	33.38	33.98
	Causal features	38.43	39.21	35.02	33.75	51.26	62.88	58.29	62.57
	Raw data w/ correlation	57.89	58.23	59.53	67.93	70.00	73.50	70.65	73.82
	Raw data w/ causal	**65.92**	**72.86**	**71.44**	**68.92**	**73.70**	**74.64**	**71.27**	**74.70**
Kappa (×100)	Raw data	59.03	55.05	58.76	59.54	62.57	73.02	65.05	74.22
	Correlation features	19.45	30.35	20.60	6.59	19.05	22.62	23.37	25.59
	Causal features	31.99	34.56	29.74	19.23	61.45	63.54	63.36	64.37
	Raw data w/ correlation	60.56	59.35	65.86	67.93	75.15	75.54	72.97	77.53
	Raw data w/ causal	**68.14**	**76.09**	**76.42**	**73.35**	**77.40**	**76.18**	**73.52**	**78.31**
F1 (×100)	Raw data	73.42	70.49	71.86	74.09	75.54	82.59	76.57	83.07
	Correlation features	48.54	53.76	48.25	40.91	47.33	49.74	50.27	50.89
	Causal features	56.01	56.09	52.77	48.72	72.21	76.47	76.54	76.95
	Raw data w/ correlation	74.00	73.53	76.91	79.40	83.62	84.17	82.50	85.34
	Raw data w/ causal	**79.29**	**84.43**	**84.38**	**82.75**	**85.25**	**84.61**	**82.86**	**85.86**

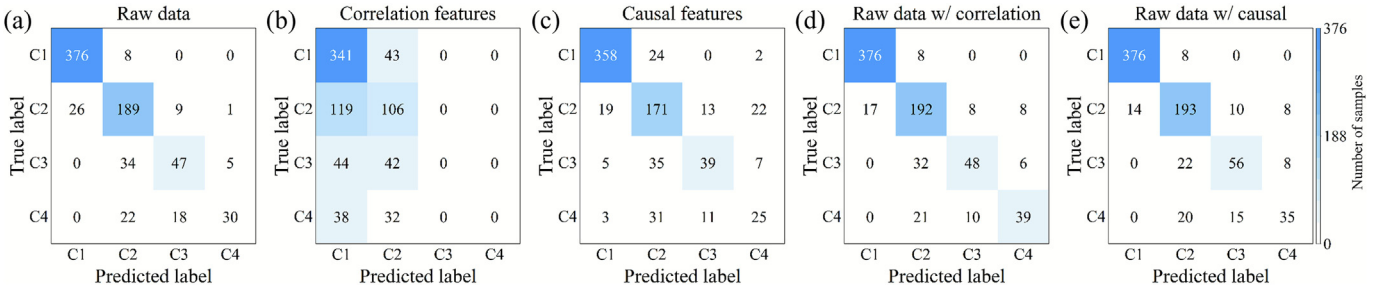

Fig. 10. Confusion matrices of the mineral resource type classification using the CatBoost method under different feature strategies. In this scenario, C1 represents the negative samples, C2 denotes the mineralized spots, C3 represents the mineral occurrences, and C4 denotes the mineral deposits. (a) Raw data; (b) Correlation features; (c) Causal features; (d) Raw data w/ correlation; (e) Raw data w/ causal.

samples (C1) and mineralized spots (C2) in the training set, these two classes are recognized with higher recall. In contrast, even though mineral occurrences (C3) and mineral deposits (C4) contain fewer training samples, the "Raw data w/ causal" feature strategy still achieves the highest overall number of correctly identified samples for these classes. These results collectively indicate that causal feature extraction not only enhances recognition for dominant classes but also improves the discriminative capability for minority classes, thereby strengthening the robustness of the proposed CausalAI framework in multi-class mineral resource type

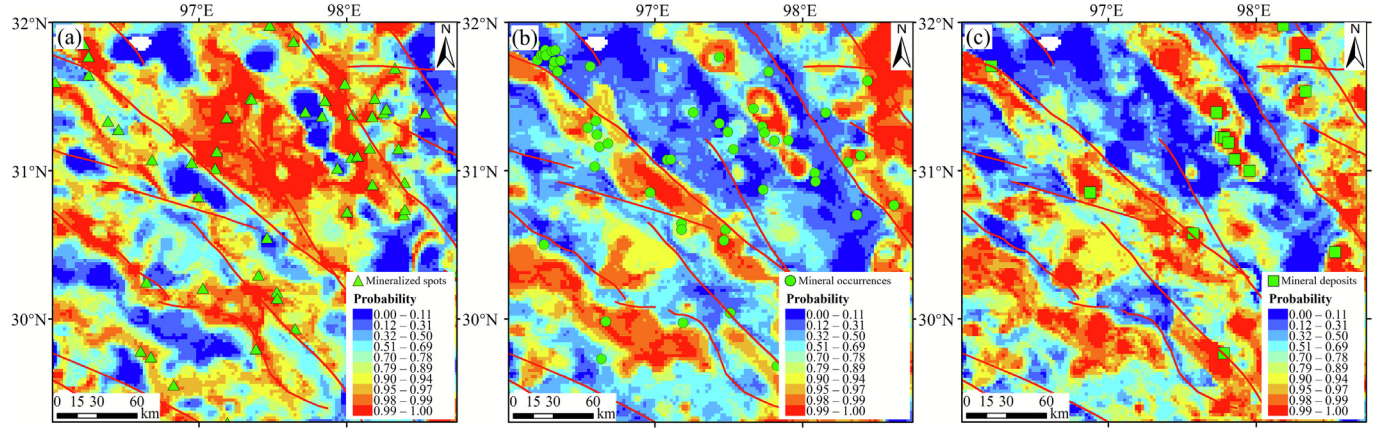

Fig. 11. Mineral prospectivity maps of different mineral resource types obtained by CatBoost causal method. (a) Mineralized spots; (b) Mineral occurrences; (c) Mineral deposits.

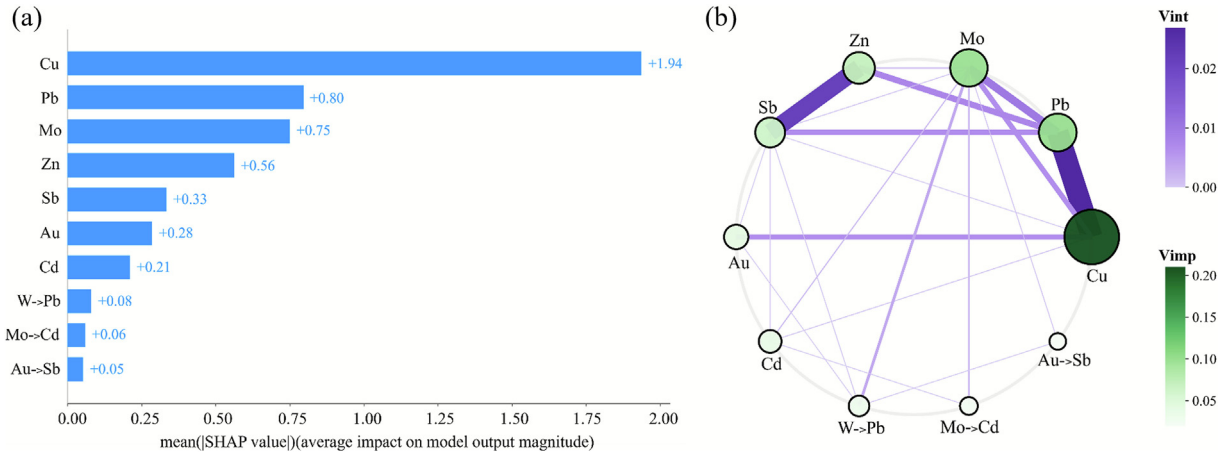

Fig. 12. SHAP analysis of binary classification using raw data with causal features. (a) Average absolute SHAP value, revealing the average impact on model output magnitude. (b) Variable importance and interaction effects (Li et al., 2024), higher variable importance (Vimp) is indicated with larger and darker nodes (green). Higher variable interaction strength (Vint) is plotted with thicker and darker lines (purple).

classification. As illustrated in Fig. 11, when predicting mineral resource types, the high-probability areas of the mineral prospectivity maps generated by the CausalAI framework exhibit a strong nesting relationship with the positive samples of different resource types. This result indicates that the causal feature representation derived from the CausalAI framework enhances the prediction accuracy of MPM in multi-class classification tasks.

4.3. Interpretability analysis

The interpretability of the proposed CausalAI framework is evaluated using Shapley Additive Explanations (SHAP) theory (Lundberg et al., 2020). SHAP quantifies the contribution of each feature to the model's output based on Shapley values from game theory, providing a measure of feature importance. We apply SHAP to the RF model that is trained using raw data and causal features in the binary classification task.

Fig. 12 presents SHAP summary plots from a global interpretability perspective. Specifically, Fig. 12a demonstrates that the model's predictions are dominantly driven by key individual metallogenic elements (e.g., Cu, Pb, Mo, Zn). Notably, several causal feature pairs (e.g., W->Pb, Mo->Cd) are also ranked among the most influential features. This validates that the causal features extracted by our framework provide additional discriminative information to the classifier, beyond the raw geochemical concentrations. Fig. 12b visualizes the network of variable importance (Vimp) and interaction effects (Vint). The analysis reveals that Cu and Pb not only rank as the top two most important variables (indicated by large dark green nodes) but also exhibit strong interactions with each other and with most other features. Furthermore, despite their lower importance rankings, Au and the W->Pb still exhibit a relatively significant interaction strength.

5. Discussion

5.1. Comparison with previous studies

MPM research has long evolved along two dominant paradigms, knowledge-driven and data-driven approaches, each with notable advancements yet inherent limitations that constrain their reliability in complex geological scenarios. Knowledge-driven MPM methods emphasize integrating expert-derived metallogenic rules with multi-source geo-data via algorithmic optimization. Early innovations such as multi-criteria decision-making (MCDM) techniques (e.g., fuzzy outranking) (Abedi et al., 2013) and uncertainty quantification frameworks (Daviran et al., 2022) have improved their practicality. However, these methods remain constrained by subjective expert judgment, propagating biases into final prospectivity maps. Traditional data-driven MPM methods, by contrast, leverages ML models to automatically capture non-linear relationships between predictor variables and mineralization, avoiding expert subjectivity (Mou et al., 2025). These models have demonstrated high predictive accuracy in data-rich settings. Yet they suffer from susceptibility to spurious correlations. This leads to reduced performance, as demonstrated by the relatively low accuracy of ML models when trained on raw geochemical data combined with correlation features than that combined with causal features (as shown in Table 2).

Our proposed CausalAI framework addresses these limitations by embedding causal inference into the MPM workflow. First, to resolve knowledge-driven subjectivity, we developed the SGC, which quantifies spatial causal relationships between geochemical elements. For example, the SGC model identified 14 elements (e.g., Au, Mo, Pb) with significant causal effects on Cu mineralization (see Table 1). These results align with established metallogenic theory for the Sanjiang region, validating the model's ability to objectively capture geologically meaningful causal signals. Second, to mitigate spurious correlations in data-driven methods, we propose the Causactor to extract causal features rather than statistical correlation features. Experimental results in Table 2 confirm that integrating these causal features with raw geochemical data ("Raw data w/ causal" strategy) outperforms other feature strategies across all 8 ML models.

In summary, the CausalAI framework addresses core limitations of previous MPM approaches to a certain extent. Its performance in the Sanjiang region demonstrates its potential to advance MPM from correlative mapping to causal mechanism-driven prediction, providing a more robust tool for mineral exploration.

5.2. Limitations of the CausalAI framework

The proposed CausalAI framework, despite its effectiveness in enhancing MPM accuracy via causal inference, still exhibits key limitations. A primary limitation is its exclusive reliance on geochemical data, failing to integrate multi-dimensional geological factors that jointly control mineralization. It is well recognized that mineralization is controlled by a combination of geological factors, including structure, lithology, and geophysics, in addition to geo-

chemistry. Although geochemical data are selected for their high spatial resolution consistent regional coverage, and direct linkage to ore-forming processes, their effectiveness in this study is context-dependent. The Sanjiang region's minimal anthropogenic disturbance preserves robust geochemical signatures of porphyry copper systems (closely tied to volcanic activity), and single-mode geochemical data have shown proven utility in porphyry copper exploration. However, they alone cannot capture the full spectrum of ore-controlling mechanisms. This single-data-source constraint may introduce biases in regions where non-geochemical factors dominate ore formation, limiting the framework's generalization to more complex geological settings.

Second, the framework demonstrates limited accuracy in identifying weakly mineralized targets, specifically mineralized spots and mineral occurrences. As reflected in the prediction results (Figs. 8 and 9), most of the unpredicted positive samples belong to these two types of targets rather than large mineral deposits. Geologically, mineralized spots and mineral occurrences typically correspond to early-stage or weakly developed mineralization events, characterized by limited ore-forming materials and low-intensity hydrothermal alteration. This leads to their geochemical anomalies being less pronounced and spatially discontinuous features that are difficult for CausalAI framework to capture. In contrast, mineral deposits, associated with stronger hydrothermal systems and large-scale fluid–rock interactions, generate robust and coherent geochemical footprints (e.g., high and continuous enrichment of Au, Mo, and Cu) that the framework can effectively recognize. Although the CausalAI framework can identify some weakly mineralized targets, its recall and precision for these targets remain notably lower than those for deposits, indicating a need for further optimization of causal feature extraction to capture subtle geochemical signals from weak mineralization.

5.3. Broader implications and future vision

The primary significance of this study extends beyond the numerical improvements in prediction accuracy. By introducing the CausalAI framework, we provide a methodological blueprint for incorporating causal reasoning into spatial predictive modeling. While DL models may offer greater capacity for modeling nonlinearities, their black-box nature often obscures the underlying decision-making process. Our work demonstrates that explicit causal feature engineering can significantly enhance model interpretability and geological plausibility without sacrificing performance. This establishes a crucial theoretical and practical foundation for the next generation of AI in geoscience, where causal understanding and predictive power are synergized. Our CausalAI framework should therefore be viewed not as a replacement for complex DL models, but as a necessary and foundational step toward building more trustworthy, robust, and geologically-aware exploration tools.

6. Conclusion

In this study, we propose a novel causal-aware artificial intelligence (CausalAI) framework for mineral prospectivity mapping (MPM), which integrates geostatistics with Granger causality theory to explicitly model the causal relationships among geochemical variables. The CausalAI framework comprises three core components: the Spatial Granger Causality (SGC) model for identifying causally relevant geochemical elements, a tailored causal-aware feature extractor (Causactor) for extracting causal features, and a suite of machine learning classifiers that combine raw geochemical data with causal features for mineralization prediction and mineral resource type classification. Through an empirical investigation conducted in the northern Cu-polymetallic mineralization zone of the Sanjiang region in southwestern China, we demonstrate that the CausalAI framework markedly enhances the reliability and precision of mineral prospectivity assessments. The predictive results exhibit strong spatial alignment with known mineral deposit distributions, with high-potential zones closely corresponding to areas of known mineralization. Notably, elements such as Au, Pb, Ag, Mo, and Zn recognized as key metallogenic and associated elements in the study area, are identified as causal factors by the SGC model, reinforcing the geological plausibility of the predictions. Meanwhile, Random Forest (RF) and CatBoost emerge as the most suitable classifiers for identifying both positive samples and mineral resource types. This study bridges the gap between data-driven modelling and geological reasoning by incorporating causal discovery into the mineral prospectivity workflow. The CausalAI framework demonstrates strong potential as a decision-support tool for mineral exploration, offering not only high predictive accuracy but also interpretability grounded in domain knowledge. Future work will focus on extending the framework to multi-source geological data, refining causal feature extraction methods, and developing causal-guided deep learning frameworks to model complex and nonlinear relationships in large-scale data. This will more fully unleash the potential of causal inference in mineral exploration. Furthermore, the framework will be applied to various metallogenic systems and underexplored regions to evaluate its generalizability and practical utility.

Declaration of Generative AI and AI-assisted technologies in the writing process

No applicable.

CRediT authorship contribution statement

Zitong Zhang: Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation. **Jianping Chen:** Validation, Supervision, Investigation, Funding acquisition, Conceptualization. **Fujie Jiang:** Validation, Investigation, Formal analysis.

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.

Acknowledgements

This work was supported by the Postdoctoral Fellowship Program and China Postdoctoral Science Foundation under Grant BX20250435 and the National Key Research and Development Program of China under Grant 2023YFC2906404-01, and in part by the Theory of Hydrocarbon Enrichment under Multi-Spheric Interactions of the Earth under Grant THEMSIE04010101.

References

Abedi, M., Norouzi, G.H., Fathianpour, N., 2013. Fuzzy outranking approach: a knowledge-driven method for mineral prospectivity mapping. Int. J. Appl. Earth Obs. 21, 556–567.

Agterberg, F., 2011. A modified weights-of-evidence method for regional mineral resource estimation. Nat. Resour. Res. 20 (2), 95–101.

Aitchison, J., 1982. The statistical analysis of compositional data. J. R. Stat. Soc. B. 44 (2), 139–160.

Altman, N., Krzywinski, M., 2015. Points of significance: association, correlation and causation. Nat. Methods. 12 (10), 899.
Atkinson, P.M., Webster, R., Curran, P.J., 1992. Cokriging with ground-based radiometry. Remote Sens. Environ. 41 (1), 45–60.
Brand, J.E., Zhou, X., Xie, Y., 2023. Recent developments in causal inference and machine learning. Annu. Rev. Sociol. 49 (1), 81–110.
Carranza, E.J.M., Hale, M., Faassen, C., 2008a. Selection of coherent deposit-type locations and their application in data-driven mineral prospectivity mapping. Ore. Geol. Rev. 33 (3–4), 536–558.
Carranza, E.J.M., Van Ruitenbeek, F.J.A., Hecker, C., van der Meijde, M., van der Meer, F.D., 2008b. Knowledge-guided data-driven evidential belief modeling of mineral prospectivity in Cabo de Gata, SE Spain. Int. J. Appl. Earth Obs. 10 (3), 374–387.
Chamberlain, G., 1982. The general equivalence of Granger and Sims causality. Econometrica 50 (3), 569–581.
Cocks, L.R.M., Torsvik, T.H., 2013. The dynamic evolution of the Palaeozoic geography of eastern Asia. Earth-Sci. Rev. 117, 40–79.
Cressie, N., Chan, N.H., 1989. Spatial modeling of regional variables. J. Am. Stat. Assoc. 84 (406), 393–401.
Cui, P., Athey, S., 2022. Stable learning establishes some common ground between causal inference and machine learning. Nat. Mach. Intell. 4 (2), 110–115.
Curran, P.J., 2001. Remote sensing: using the spatial domain. Environ. Ecol. Stat. 8, 331–344.
Carranza, E.J.M., Hale, M., Faassen, C., 2008c. Selection of coherent deposit-type locations and their application in data-driven mineral prospectivity mapping. Ore. Geol. Rev. 33, 536–558.
Daviran, M., Parsa, M., Maghsoudi, A., Ghezelbash, R., 2022. Quantifying uncertainties linked to the diversity of mathematical frameworks in knowledge-driven mineral prospectivity mapping. Nat. Resour. Res. 31 (5), 2271–2287.
Delforge, D., de Viron, O., Vanclooster, M., Van Camp, M., Watlet, A., 2022. Detecting hydrological connectivity using causal inference from time series: synthetic and real karstic case studies. Hydrol. Earth Syst. Sc. 26 (8), 2181–2199.
Deng, J., Wang, Q., Li, G., Li, C., Wang, C., 2014a. Tethys tectonic evolution and its bearing on the distribution of important mineral deposits in the Sanjiang region, SW China. Gondwana Res. 26 (2), 419–437.
Deng, J., Wang, Q., Li, G., Santosh, M., 2014b. Cenozoic tectono-magmatic and metallogenic processes in the Sanjiang region, southwestern China. Earth-Sci. Rev. 138, 268–299.
Ebert-Uphoff, I., Deng, Y., 2017. Causal discovery in the geosciences-using synthetic data to learn how to interpret results. Comput. Geosci. 99, 50–60.
Fawcett, T., 2006. ROC graphs with instance-varying costs. Pattern Recognit. Lett. 27 (8), 882–891.
Granger, C.W., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37 (3), 424–438.
Granger, C.W., 1988. Some recent development in a concept of causality. J. Econometrics. 39 (1–2), 199–211.
Hou, Z.Q., Zhong, D.L., Deng, W.M., 2004. A tectonic model for porphyry copper-molybdenum-gold metallogenic belts on the eastern margin of the Qinghai-Tibet Plateau. Geol. China 31 (1), 1–14.
Hou, Z.Q., Ma, H.W., Khin, Z., Zhang, Y.Q., Wang, M.J., Wang, Z., Pan, G.T., Tang, R.L., 2003. The Himalayan Yulong porphyry copper belt: product of large-scale strike-slip faulting in eastern Tibet. Econ. Geol. 98 (1), 125–145.
Hronsky, J.M., Kreuzer, O.P., 2019. Applying spatial prospectivity mapping to exploration targeting: fundamental practical issues and suggested solutions for the future. Ore. Geol. Rev. 107, 647–653.
Kodra, E., Chatterjee, S., Ganguly, A.R., 2011. Exploring Granger causality between global average observed time series of carbon dioxide and temperature. Theor. Appl. Climatol. 104, 325–335.
Lauzon, D., Gloaguen, E., 2024. Quantifying uncertainty and improving prospectivity mapping in mineral belts using transfer learning and Random Forest: a case study of copper mineralization in the Superior Craton Province, Quebec, Canada. Ore. Geol. Rev. 166, 105918.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521 (7553), 436–444.
Li, H., Li, X.H., Yuan, F., Zhang, M.M., Li, X.L., Ge, C., Wang, Z.Q., Guo, D., Lan, X.Y., Tang, M.H., Lu, S.M., 2023a. Genetic algorithm optimized light gradient boosting machine for 3D mineral prospectivity modeling of Cu polymetallic skarn-type mineralization, Xuancheng Area, Anhui Province, Eastern China. Nat. Resour. Res. 32 (5), 1897–1916.
Li, Q.K., Chen, G.X., Luo, L., 2023b. Mineral prospectivity mapping using attention-based convolutional neural network. Ore. Geol. Rev. 156, 105381.
Li, C., Xiao, K.Y., Sun, L., Tang, R., Dong, X.C., Qiao, B.C., Xu, D.H., 2024a. CNN-Transformers for mineral prospectivity mapping in the Maodeng–Baiyinchagan area, Southern Great Xing'an Range. Ore. Geol. Rev. 167, 106007.
Li, X.W., Su, X.X., Li, J., Anwar, S., Zhu, X.Q., Ma, Q., Wang, W.H., Liu, J.K., 2024b. Coupling image-fusion techniques with machine learning to enhance dynamic monitoring of Nitrogen content in winter wheat from UAV multi-source. Agriculture 14 (10), 1797.
Liu, H., Harris, J., Sherlock, R., Behnia, P., Grunsky, E., Naghizadeh, M., Rubingh, K., Tuba, G., Roots, E., Hill, G., 2023. Mineral prospectivity mapping using machine learning techniques for gold exploration in the Larder Lake area, Ontario, Canada. J. Geochem. Explor. 253, 107279.
Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S., 2020. From local explanations to global understanding with explainable AI for trees. Nat. Mach. Intell. 2 (1), 56–67.
Luo, Z.J., Zuo, R.G., 2025. Causal discovery and deep learning algorithms for detecting geochemical patterns associated with gold-polymetallic mineralization: a case study of the Edongnan Region. Math. Geosci. 57 (1), 193–220.
Mosedale, T.J., Stephenson, D.B., Collins, M., Mills, T.C., 2006. Granger causality of coupled climate processes: Ocean feedback on the North Atlantic Oscillation. J. Clim. 19 (7), 1182–1194.
Mou, N.N., Carranza, E.J.M., Wang, G.W., Sun, X., 2023. A framework for data-driven mineral prospectivity mapping with interpretable machine learning and modulated predictive modeling. Nat. Resour. Res. 32 (6), 2439–2462.
Mou, N.N., Carranza, E.J.M., Xue, J.L., Zhang, S., Wang, G.W., Song, H., Chen, Y.H., Ren, X.N., 2025. Interpretable machine learning for mineral prospectivity mapping in the Qulong-Jiama district, Tibet, China. Ore. Geol. Rev. 182, 106659.
Myers, D.E., 1982. Matrix formulation of co-kriging. J. Int. Assoc. Math. Geol. 14 (3), 249–257.
Nagasingha, L.M.A., Bérubé, C.L., Lawley, C.J.M., 2024. A balanced mineral prospectivity model of Canadian magmatic Ni (±Cu±Co±PGE) sulphide mineral systems using conditional variational autoencoders. Ore. Geol. Rev. 175, 106329.
Pearl, J., 2009. Causal inference in statistics: an overview. Stat. Surv. 3, 96–146.
Pérez-Suay, A., Camps-Valls, G., 2018. Causal inference in geoscience and remote sensing from observational data. IEEE Trans. Geosci. Remote. 57 (3), 1502–1513.
Perkins-Kirkpatrick, S.E., Stone, D.A., Mitchell, D.M., Rosier, S., King, A.D., Lo, Y.T.E., Pastor-Paz, J., Frame, D., Wehner, M., 2022. On the attribution of the impacts of extreme weather events to anthropogenic climate change. Environ. Res. Lett. 17 (2), 024009.
Peters-Lidard, C.D., Clark, M., Samaniego, L., Verhoest, N.E.C., Emmerik, T.V., Uijlenhoet, R., Achieng, K., Franz, T.E., Woods, R., 2017. Scaling, similarity, and the fourth paradigm for hydrology. Hydrol. Earth Syst. Sci. 21 (7), 3701–3713.
Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., Chica-Rivas, M., 2015. Machine learning predictive models for mineral prospectivity: an evaluation of neural networks, random forest, regression trees and support vector machines. Ore. Geol. Rev. 71, 804–818.
Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., Glymour, C., Kretschmer, M., Mahecha, M.D., Muñoz-Marí, J., van Nes, E.H., Peters, J., Quax, R., Reichstein, M., Scheffer, M., Schölkopf, B., Spirtes, P., Sugihara, G., Sun, J., Zhang, K., Zscheischler, J., 2019. Inferring causation from time series in Earth system sciences. Nat. Commun. 10 (1), 2553.
Shen, C., 2021. Research on prospecting prediction for the Yulong Cu metallogenic belt and its peripheral areas based on the model for spatial structures of regional geochemical multi-element anomalies, Tibet. Acta Geol. Sin-Engl. 95 (11), 3163–3177.
Shojaie, A., Fox, E.B., 2022. Granger causality: a review and recent advances. Annu. Rev. Stat. Appl. 9 (1), 289–319.
Sims, C.A., 1972. Money, income, and causality. Am. Econ. Rev. 62 (4), 540–552.
Sun, T., Chen, F., Zhong, L.X., Liu, W.M., Wang, Y., 2019. GIS-based mineral prospectivity mapping using machine learning methods: a case study from Tongling ore district, eastern China. Ore. Geol. Rev. 109, 26–49.
Tang, J.X., Zhong, K.H., Liu, Z.C., Li, Z.J., Dong, S.Y., Zhang, L., 2006. Intracontinent orogen and metallogenesis in Himalayan epoch: Changdu large composite basin, eastern Tibet. Acta Geol. Sin-Engl. 80 (9), 1364–1376.
Tao, Y., Bi, X.W., Li, J.G., Zhu, F.L., Liao, M.Y., Li, Y.B., 2011. Geochemistry and petrogenesis of the Jitang granitoids in Tibet, SW China. Acta Geol. Sin-Engl. 27 (9), 2763–2774.
Tiku, M.L., 1967. Tables of the power of the F-test. J. Am. Stat. Assoc. 62 (318), 525–539.
Wang, F.M., Li, J.L., Peng, D.L., Yi, Q.X., Zhang, X.Y., Zheng, J.Y., Chen, S.T., 2024a. Estimating soybean yields using causal inference and deep learning approaches with satellite remote sensing data. IEEE J.-Stars 17, 14161–14178.
Wang, L.Q., Yang, J., Wu, S.S., Hu, L.S., Ge, Y.Z., Du, Z.H., 2024b. Enhancing mineral prospectivity mapping with geospatial artificial intelligence: a geographically neural network-weighted logistic regression approach. Int. J. Appl. Earth Obs. 128, 103746.
Wilson, A.M., Russell, J.K., 2020. Glacial pumping of a magma-charged lithosphere: a model for glaciovolcanic causality in magmatic arcs. Earth Planet. Sci. Lett. 548, 116500.
Wong, T.T., Yeh, P.Y., 2019. Reliable accuracy estimates from k-fold cross validation. IEEE Trans. Knowl. Data Eng. 32 (8), 1586–1594.
Wu, C., Murray, A.T., 2005. A cokriging method for estimating population density in urban areas. Comput. Environ. Urban. 29 (5), 558–579.
Xie, X.J., Mu, X.Z., Ren, T.X., 1997. Geochemical mapping in China. J. Geochem. Explor. 60 (1), 99–113.
Zhang, Z.T., Zhang, X., Zhang, C.L., Feng, H.L., 2025. Category-wise sequence correlation filter for lithology identification with well logs. IEEE Trans. Geosci. Remote. 63, 1–14.
Yadav, S., Shukla, S., 2016. Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. In: 2016 IEEE 6th International Conference on Advanced Computing (IACC), pp. 78–83.
Yang, F.F., Zuo, R.G., Kreuzer, O.P., 2024. Artificial intelligence for mineral exploration: a review and perspectives on future directions from data science. Earth-Sci. Rev. 258, 104941.

Yao, L.Y., Chu, Z.X., Li, S., Li, Y.L., Gao, J., Zhang, A.D., 2021. A survey on causal inference. ACM t. Knowl. Discov. D. 15 (5), 1–46.
Yin, B.J., Zuo, R.G., Sun, S.Q., 2023. Mineral prospectivity mapping using deep self-attention model. Nat. Resour. Res. 32 (1), 37–56.
Yu, Z.B., Li, B.B., Wang, X.J., 2024. Mineral prospectivity mapping susceptibility evaluation based on interpretable ensemble learning. Ore. Geol. Rev. 173, 106248.
Zhang, H., Dong, Z.X., Li, B., He, S., 2022. Multi-Scale MLP-Mixer for image classification. Knowl-Based. Syst. 258, 109792.
Zhang, Z.B., Xu, Y., Miao, Y.J., Wang, W.F., Zhao, D.F., Chen, D.L., 2022b. Provenance and sedimentary environment of Paleogene Gongjue Formation in Qamdo Basin. Acta Sedimentol. Sin. 40(6), 1561-1581.
Zuo, R.G., Carranza, E.J.M., 2011. Support vector machine: a tool for mapping mineral prospectivity. Comput. Geosci. 37 (12), 1967–1975.
Zuo, R.G., 2020. Geodata science-based mineral prospectivity mapping: a review. Nat. Resour. Res. 29 (6), 3415–3424.
Zuo, R.G., Xu, Y., 2023. Graph deep learning model for mapping mineral prospectivity. Math. Geosci. 55 (1), 1–21.