Applied Computing and Geosciences 29 (2026) 100325

Contents lists available at ScienceDirect

Applied Computing and Geosciences

journal homepage: www.sciencedirect.com/journal/applied-computing-and-geosciences

Geochemical anomaly delineation utilizing copula-based outlier detection method

Shahed Shahrestani [a], Emmanuel John M. Carranza [b,*], Ioan Sanislav [a]

[a] College of Science and Engineering, Economic Geology Research Centre (EGRU), James Cook University, Townsville, Australia
[b] Department of Geology, University of the Free State, Bloemfontein, South Africa

ARTICLE INFO

Keywords:
Copula-based outlier detection (COPOD)
Geochemical anomalies
Anomaly detection
Mineral occurrence

ABSTRACT

This study evaluates the effectiveness of Copula-Based Outlier Detection (COPOD) in identifying geochemical anomalies within the Toroud–Chah Shirin belt (TCSB) in Iran. The TCSB is a significant mineralized zone containing epithermal precious and base metal veins, skarn, gold placer, and Pb–Zn sedimentary-hosted deposits. Unlike proximity-based or learning-based models, COPOD is a fully deterministic and unsupervised statistical approach. It requires no hyperparameter tuning or assumptions regarding data distribution, making it ideal for the skewed, non-Gaussian nature of stream sediment datasets. By modeling multivariate dependencies through empirical cumulative distribution functions (ECDFs), COPOD captured complex element relationships, such as Ag–Pb and Bi–Au, which relate to sedimentary-hosted and epithermal gold deposits in the region. Comparative analysis using Receiver Operating Characteristic (ROC) curves demonstrates that COPOD outperforms both traditional uni-element mapping and the state-of-the-art Isolation Forest (IF) method. Using a 10% contamination threshold, the COPOD method identified 23 out of 32 known mineral occurrences, whereas the IF method captured 19. Furthermore, this study uses dimensional outlier graphs to provide transparent results, highlighting the influence of Co, Zn, Sb, and Pb on anomaly scores. Results from Lasso regression and random forest analysis further confirmed these elemental impacts. Comparison with the regional geological map shows that most anomalies occur within Paleogene volcanic units and the Cretaceous sedimentary unit that hosts Pb–Zn mineralization. However, some extend into surficial areas due to geochemical dispersion. Overall, COPOD offers a robust, efficient, and explainable alternative for multivariate geochemical anomaly delineation.

1. Introduction

The importance of geochemical anomalies in aiding the discovery of mineral deposits has long been acknowledged (e.g., Shahrestani et al., 2019; Zuo and Xu, 2024). Over the past decades, considerable research efforts have been directed towards devising innovative methods capable of reliably identifying geochemical anomalies while minimizing associated uncertainties (e.g., Shahrestani et al., 2020; Wang and Zuo, 2022; Chen et al., 2023). In this context, anomaly detection techniques have been closely aligned with the concept of geochemical anomalies, facilitating the differentiation between anomalous and non-anomalous zones. The anomaly detection methods applied to delineate geochemical anomalies can be categorized into unsupervised detection, supervised detection, ensemble methods, feature extraction, and hybrid learning. For instance, anomalies in geochemical exploration data can be identified using an unsupervised machine learning model, which employs an outlier detection process. Geochemical anomalies in exploration data can also be identified using a supervised machine learning model through a classification process. This model treats geochemical anomalies and geochemical background as two distinct classes. Subsequently, a decision function is determined to differentiate between these two classes, thereby separating geochemical anomalies from background (Chen et al., 2023). The fundamental concept underlying these methods aligns with the broad definition of geochemical anomalies, which denotes a departure from the typical geochemical baseline considered as normal. Nevertheless, the precise methodologies employed to ascertain such deviations vary across different anomaly detection approaches.

Methods that have been applied in the domains of geochemical anomaly detection and mineral prospectivity mapping include restricted Boltzmann machines (RBMs) (Chen, 2015), variational autoencoders (VAEs) (Xiong and Zuo, 2016; Zhao et al., 2023; Esmaeiloghli et al.,

* Corresponding author.
E-mail address: ejmcarranza@gmail.com (E.J.M. Carranza).

https://doi.org/10.1016/j.acags.2026.100325
Received 14 November 2025; Received in revised form 5 February 2026; Accepted 9 February 2026
Available online 10 February 2026

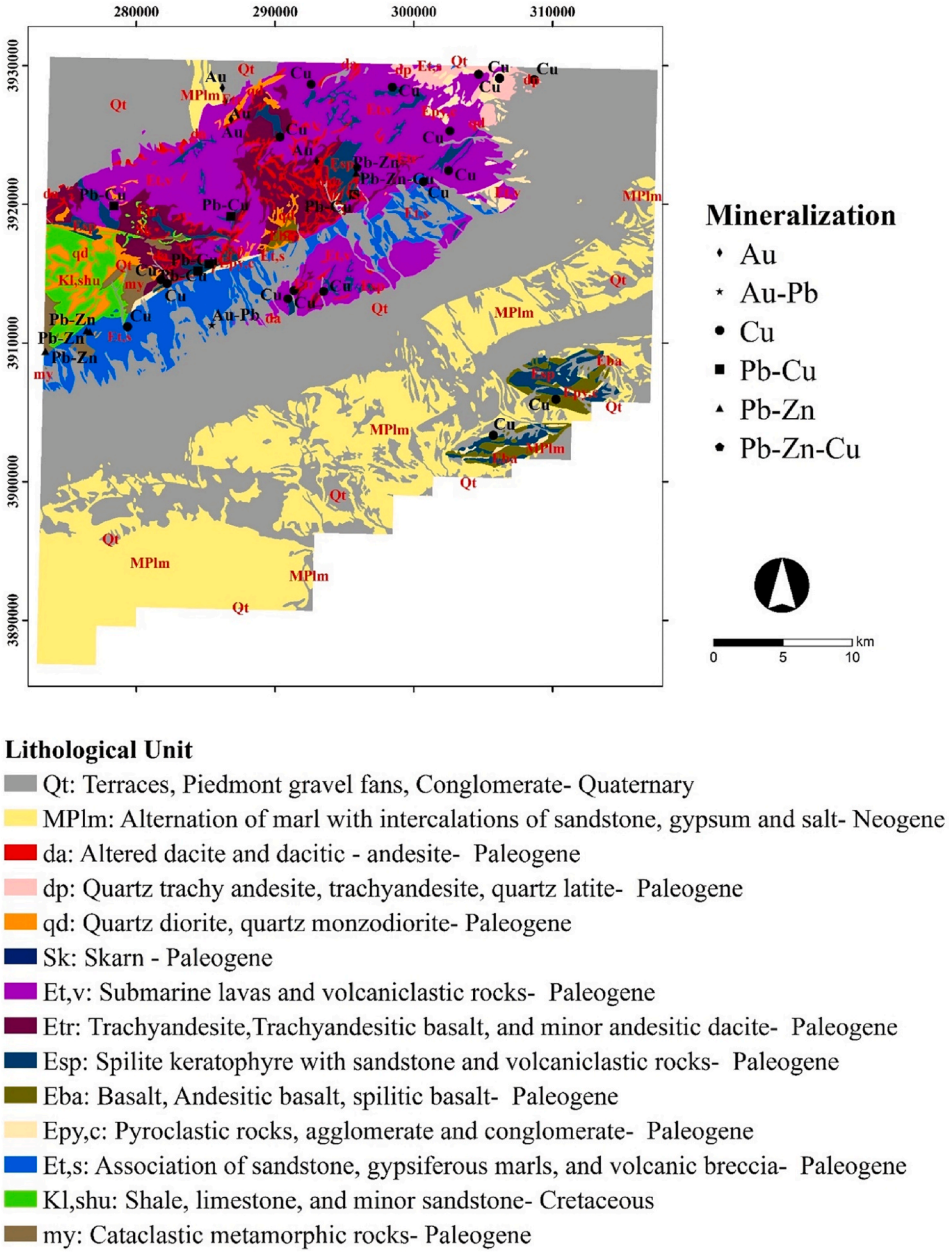

Fig. 1. Simplified geological map of the study area (after Eshraghi and Jalali, 2006).

2023), one-class support vector machines (OCSVMs) (Chen and Wu, 2017; Shahrestani and Carranza, 2024), deep autoencoders (Xiong et al., 2018), bat-optimized OCSVM model (Chen et al., 2019), isolation forest (iForest) (Chen and Wu, 2019), distance anomaly factors (Chen et al., 2021a), k-nearest neighbor (Chen et al., 2021b), generative adversarial network (GAN) (Zhang and Zuo, 2021), self-organizing map (Bigdeli et al., 2022), model averaging (Wang and Zuo, 2022), dictionary learning (Chen and Sui, 2022), local outlier factor (Puchhammer et al., 2024; Shahrestani and Carranza, 2024), and neural network based models (Wang et al., 2024). In recent years, significant progress has been made in the application of machine learning and, more recently, deep learning methods for the interpretation of geoscientific data. A wide range of approaches has been proposed, including LRR-Net (Li et al., 2023), SpectralGPT (Hong et al., 2023), VAE-BIRCH (Hoseinzade et al., 2025), 1DCNN combined with graph convolutional networks (Zuo and Xu, 2024), deep forest models (Liu et al., 2025), memory-augmented autoencoders (Luo et al., 2025), Bayesian deep learning approaches (Liu, 2025), and deep belief networks (e.g., Keykhay-Hosseinpoor et al., 2024).

Various anomaly detection methods produce distinct patterns of anomalies, revealing uncertainty in geochemical anomaly identification. The absence of a universally superior method suggests that there could be an optimal, yet unidentified, anomaly detector capable of effectively distinguishing significant anomalies while minimizing misclassification. The uncertainty arises from the variability in anomaly patterns identified by different detection techniques (Wang and Zuo, 2022). Several existing methods have certain limitations. Firstly, methods such as k-nearest neighbor, local outlier factor, one-class support vector machines, and isolation forest face performance challenges when dealing with high-dimensional data, leading to a significant reduction in data processing efficiency. Therefore, they are often restricted to low-dimensional datasets. Secondly, the majority of methods require the determination and tuning of hyperparameters, including the number of clusters for clustering-based models, network architecture for neural network-based models, and the selection of individual classifiers in ensemble models. In addition to these limitations,

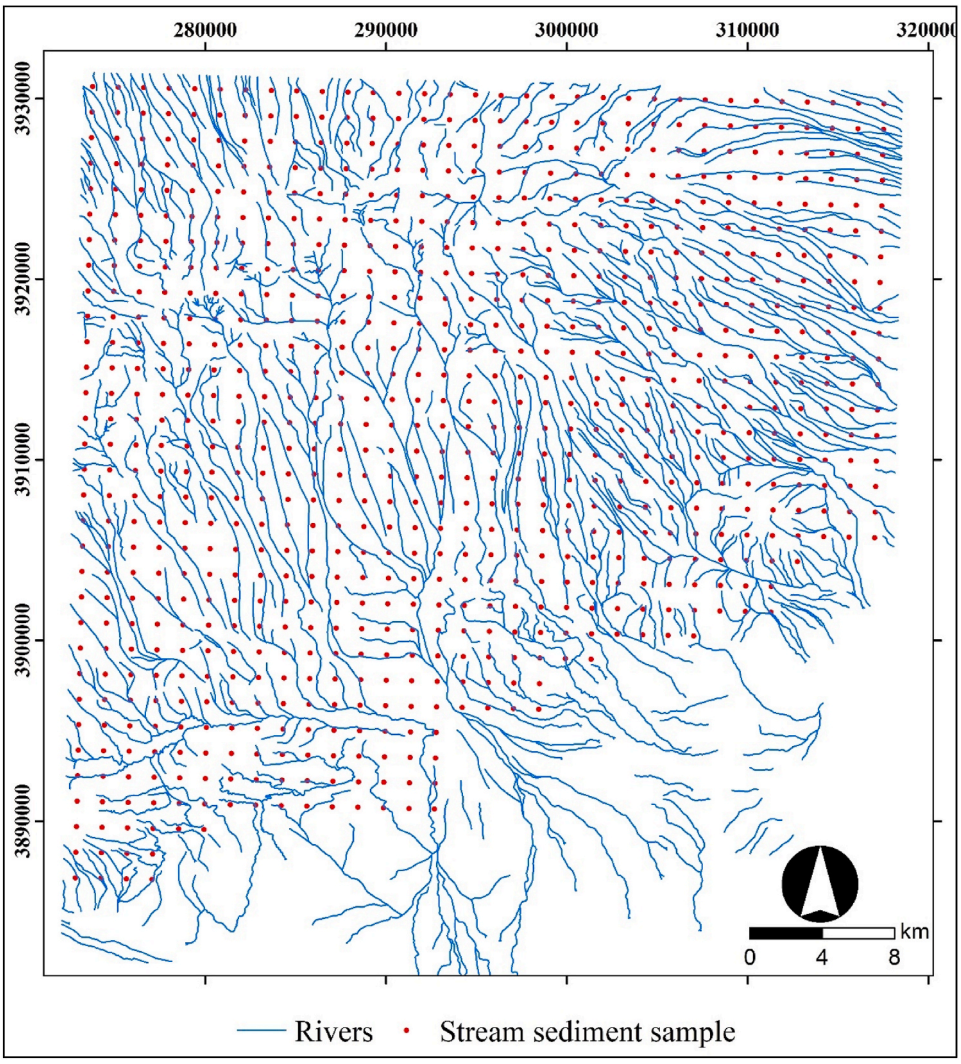

Fig. 2. Locations representing the center of each window where the stream sediment samples were collected in Moalleman area.

geochemical anomaly detection is further complicated by characteristics of geochemical data that are not adequately addressed by many existing approaches. Geochemical variables commonly exhibit non-Gaussian, skewed, and heavy-tailed distributions, which reduces the effectiveness of methods that rely on specific distributional assumptions. Moreover, reliable ground truth labels are generally unavailable in geochemical exploration datasets, making supervised learning and extensive model optimization impractical. Under such conditions, anomaly detection results become highly method-dependent, increasing uncertainty. Furthermore, many anomaly detection techniques provide limited interpretability, as they generate anomaly scores without clearly indicating the contribution of individual variables, which restricts their usefulness for geological interpretation and decision-making.

To address these challenges, this study applies copula-based anomaly detection (COPOD) (Li et al., 2020), which differs from many state-of-the-art anomaly detection techniques in several important aspects. Unlike proximity-based methods that rely on pairwise distance calculations or learning-based models that require training procedures, COPOD is a fully deterministic and unsupervised statistical approach. It is based on empirical cumulative distribution functions and involves no learning process or stochastic optimization. As a result, COPOD does not require model initialization or hyperparameter tuning, avoiding subjective parameter selection and potential training bias that commonly affect many existing methods. Another key advantage of COPOD is that it does not assume any predefined data distribution. This property is particularly important for geochemical datasets, which often exhibit skewed, heavy-tailed, and non-Gaussian distributions. By modelling multivariate dependencies through copula theory and ECDFs, COPOD is able to capture complex relationships among geochemical variables more effectively than methods that assume independence or specific distributional forms. In addition, COPOD does not require labeled data, making it well suited for geochemical exploration problems where reliable ground truth information is generally unavailable. COPOD also offers a high level of interpretability compared to many machine learning and deep learning approaches. By evaluating anomalous behaviour on a dimensional basis, COPOD quantifies the contribution of each variable to the overall anomaly score using dimensional outlier graphs. This allows domain experts to identify which geochemical elements are most responsible for anomaly formation, providing transparent and explainable results that support geological interpretation. Furthermore, COPOD is computationally efficient and scales well to high-dimensional datasets, as it avoids costly distance calculations and model training. These characteristics make COPOD a robust, interpretable, and efficient alternative to existing state-of-the-art anomaly detection techniques for multivariate geochemical data.

2. Geological background

The research area (Fig. 1), situated in northeastern Iran and depicted on the Moalleman geological map at a 1:100,000 scale (Eshraghi and

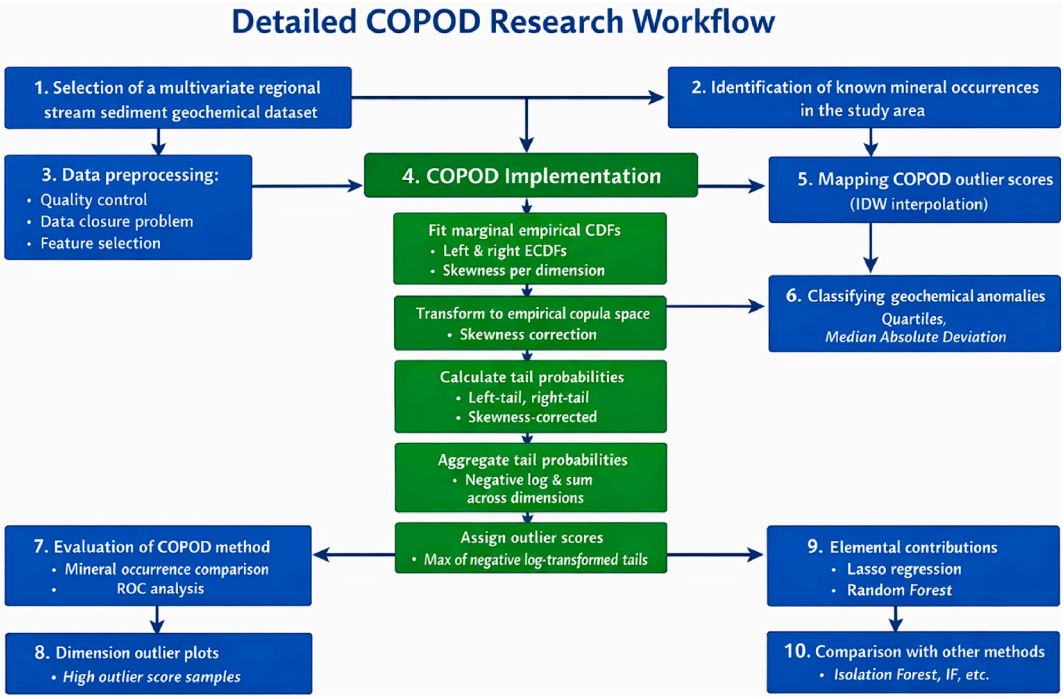

Fig. 3. Workflow diagram for delineating geochemical anomalies using the COPOD method.

Jalali, 2006), showcases a diverse array of geological formations. These formations span from Precambrian gneisses and schists to Tertiary volcanic rocks and granitic masses (Hushmandzadeh et al., 1978). Notably, the area encompasses segments of the Toroud–Chah Shirin belt (TCSB), renowned for its various mineral deposits, including copper and gold veins, gold placers, and quartz-base metal veins (Torshizian et al., 2021). The TCSB exhibits a rich variety of rock formations, such as green schist, metamorphosed dolomite, and limestone, with Eocene volcanic-pyroclastic assemblages overlaying them. Underneath these rock formations are Oligocene intrusive bodies, mainly composed of granodiorite. The predominant magmatic rocks within the TCB include andesite and basalt, evolving from acidic to more alkaline compositions over time, resulting in formations like lava flows, breccias, and andesite tuffs. The lithological units in the area also include granite, micro-granite, granodiorite, and various other compositions, forming magmatic series ranging from subalkaline to alkaline and categorized as I-type rocks (Mehrabi et al., 2015).

In the TCSB, there are numerous notable mineral deposits and abandoned mines. These include intrusion-related epithermal vein deposits such as Cheshmeh Hafez (Pb + Zn + Cu), Chah Messi (Cu ± Au), Gandy (Au–Ag ± Pb-Zn), Abolhassani (Pb–Zn ± Ag ± Au), Chalu (Cu ± Au), Darestan (Cu ± Au), Baghu (Au ± Cu) and its associated Au-bearing placers, and sedimentary hosted deposits such as Khanjar (Pb–Zn ± Ag) (e.g., TaleFazel et al., 2019; Ghezelbash et al., 2021; Shahrestani and Carranza, 2024). These mineral occurrences are meticulously mapped on the 100,000 scale geological map, categorized by their main elemental associations such as Au, Au–Pb, Cu, Pb–Cu, Pb–Zn, and Pb–Zn–Cu. For further insights into the geology and mineralization within the study area, readers are referred to Ghezelbash et al. (2021), Bigdeli et al. (2023), Esmaeiloghli et al. (2023), and Shahrestani and Carranza (2024).

3. Materials and methods

3.1. Geochemical dataset

The geochemical dataset used in this study covers approximately 1500 km^2, from where 819 stream sediment samples were collected systematically by the Jianxi Chinese Company (JCC) in 1993. To address the predominant N–S-trending high-density drainages in the Moalleman district, JCC implemented a systematic sampling grid to effectively gather stream sediment geochemical data. By employing 1412 m × 1412 m grid cells, 2–4 sub-samples were obtained from first- or second-order drainages within each cell (Hu, 1994). These sub-samples were then combined to produce a composite stream sediment sample per cell, ensuring a thorough dataset by merging multiple subsamples into representative composite samples, thus accurately reflecting sediment composition across the area (Fig. 2) (Ghezelbash et al., 2021; Azmi et al., 2021). Stream sediment sampling yields information relevant to upstream sources and may not indicate prospectivity at the sample locations. However, collecting samples from first- or second-order streams (and not higher-order streams) ensures that any detected geochemical anomaly is directly associated with its source (e.g., Carranza and Hale, 1997; Shahrestani et al., 2019). Stream sediment data are typically analyzed using catchment basin–based approaches (e.g., Carranza and Hale, 1997) because each sample commonly reflects material derived from an upstream drainage area rather than a single point location. In this study, however, the sampling strategy limits the spatial influence of individual samples, allowing them to be treated as quasi-point observations. As a result, deterministic Euclidean distance–based interpolation methods, such as inverse distance weighting, can be reasonably applied to interpolate anomaly scores between sampling locations. The samples were analyzed by AMDEL Co. (Australia) for trace elements (Ag, As, Au, Ba, Bi, Co, Cr, Cu, Hg, Mo, Ni, Pb, Sb, Sn, Sr, V, W, Zn) using ICP–MS (Ghezelbash et al., 2021). Quality assurance and control were carried out using Student's t-test and Fisher's test, which assessed the mean and variance of elemental data in duplicate samples. The findings demonstrated robust chemical analysis, with 95% confidence level (Hu, 1994; Koohzadi et al., 2021; Esmaeiloghli et al., 2023).

3.2. Copula-based outlier detection (COPOD)

The COPOD method comprises three steps (Fig. 3). First, it fits left and right empirical cumulative distribution functions on a t-dimensional multivariate input dataset and calculates the skewness vector of each dimension. Second, the empirical copula values of S_i corresponding to

the left and right empirical distribution functions of the previous stage are calculated, and the skewness of the corrected empirical copula values is determined. Third, the probability of each S_i being observed at least as extreme along each dimension is computed. The outlier score is determined by taking the maximum of the negative logarithm of the probability generated by the left tail empirical copula, the right tail empirical copula, and the skewness-corrected empirical copula. Intuitively, as the tail probability decreases, its negative logarithm increases. Therefore, a point is deemed an outlier if it exhibits a low left tail probability, a low right tail probability, or a low skewness-corrected tail probability.

The comprehensive mathematical formulation of the COPOD method can be found in Li et al., (2020). Here, we present the fundamental mathematical principles underlying the COPOD method. This method leverages copula models, which are designed to separate marginal distributions from the dependence structure of multivariate distributions. The copula function can be considered as a function that converts the $[0, 1]^t$ space into [0,1] using the cumulative distribution function (CDF) of a random vector (L_1, L_2, …, L_t) with uniform (0,1) marginals:

$$C_L(l) = \mathbb{P}\,(L_1 \le l_1, \ldots, L_t \le l_t) \tag{1}$$

where $P\,(L_j \le l_j) = l_j$ for j in 1, …t and $l_j \in [0,1]$ (Li et al., 2020). It benefits from the inverse sampling to convert a uniform distribution into a favorable distribution, thus:

$$S_j = F_j^{-1}(L_j) \sim F_j\Big) \tag{2}$$

For a random variable dataset (S_1, …, S_t) having joint distribution function F (s_1, …, s_t) with marginal distributions as F1, …, Ft, there is a copula (C) defined as (Sklar, 1959):

$$F(\mathbf{s}) = C(F_1(s_1), \ldots, F_t(s_t)) \tag{3}$$

In essence, a copula enables characterization of the combined distribution of (S_1, …, S_t) solely based on their individual marginal distributions. The copula equation, expressed in terms of the joint cumulative distribution function (CDF) and inverse CDFs, is obtained by substituting the inverse of Eq. (2) into Eq. (1):

$$\begin{aligned} C(\mathbf{l}) &= \mathbb{P}\,(F_{S1}(S_1) \le l_1, \ldots, F_{St}(S_t) \le l_t) \\ &= \mathbb{P}\,\left(S_1 \le F_{S1}^{-1}(S_1), \ldots, S_t \le F_{St}^{-1}(S_t)\right) \\ &= F_{\mathbf{S}}\left(F_{St}^{-1}(l_1), \ldots, F_{St}^{-1}(l_t)\right) \end{aligned} \tag{4}$$

Here, the empirical copula utilized by COPOD employs a nonparametric method that centers around fitting ECDFs. Considering $S_{j,i}=(S_{1,i}, S_{2,i}, \ldots, S_{t,i})$, i = 1, …, n as ith observation in jth dimension, the empirical CDF called $\widehat{F}(s)$ is:

$$\widehat{F}(s) = \mathbb{P}((-\infty, s]) = \frac{1}{n}\sum_{i=1}^{n} I(S_i \le s) \tag{5}$$

Using Eq. (2) the empirical copula observations called $\widehat{L}_i$ can be achieved as:

$$\left(\widehat{L}_{1,i}, \ldots, \widehat{L}_{t,i}\right) = \left(\widehat{F}_1(S_{1,i}), \ldots, \widehat{F}_t(S_{t,i})\right) \tag{6}$$

Substituting Eq. (6) in Eq. (4), we achieve:

$$\widehat{C}(l_1, \ldots, l_t) = \frac{1}{n}\sum_{i=1}^{n} I\left(\widetilde{L}_{1,i} \le l, \ldots, \widetilde{L}_{t,i} \le l_t\right) \tag{7}$$

$\widehat{C}(\boldsymbol{l})$ has discrete uniform marginals on $\{1/n, 2/n, \ldots, 1\}^t$ and based on central limit theorem will converge to $C(\mathbf{l})$ (Nelsen, 2006).

The probability of observing a point at least as extreme as each observation s_i is to be computed. Specifically, assuming s_i to be distributed according to some t-variate distribution function F_S, $F_S(s_i) = P\,(S \le s_i)$ (left tail probability) and $1-F_S(s_i) = P\,(S \ge s_i)$ (right tail probability) are to be calculated. If x_i is an outlier, rare occurrence is expected. Therefore, if either $F_S(s_i)$ or $1-F_S(s_i)$ is extremely small, evidence suggests that this point has rare occurrence and is, therefore, likely an outlier. The empirical computation of tail probabilities begins with the computation of the empirical CDF using Eq. (5). Subsequently, the empirical CDFs allow for the approximation of the inverse of Eq. (2), facilitating the derivation of estimated copula observations l_j by inputting each s_j into $\widehat{F}_j$. The estimated copula observations denote the probability of observing something as extreme as s along the jth dimension. Finally, the empirical estimation of the left tail probability is achieved by the multiplication of all l_j. The corresponding right tail probabilities are determined by considering $C(1\text{-}l) = \mathbb{P}\,(L_1 \ge l_1, \ldots, L_t \ge l_t)$. A simple substitution could be directly using $-S = [-S_{1,i}, -S_{2,i}, \ldots, -S_{t,i}]$ into Eq. (5) to achieve right tail ECDF.

In cases where outliers cluster at one extreme end of the distribution, skewness correction becomes necessary. For instance, in a dataset where all outliers concentrate at the left end, utilizing left tail probabilities proves effective, capturing smaller outliers accurately. Conversely, using right tail probabilities performs poorly due to the absence of significantly large outliers, leading to misidentification of relatively large points. If the dataset were reversed, with outliers at the top right corner, right tail ECDFs would work well while left tail ECDFs would struggle. Averaging tail probabilities is effective when both large and small outliers are present. Considering the skewness of the dataset is crucial, with positive skewness favoring the use of right tail ECDFs for better results. In the COPOD method, the skewness vector $b_i = [b_1, \ldots, b_t]$ is calculated using the standard equation, thus:

$$b_i = \frac{\frac{1}{n}\sum_{i=1}^{n}(-\overline{si})^3}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(s_i - \overline{si})^2}^{\,3}} \tag{8}$$

The "curse of dimensionality" refers to the emergence of unique challenges in high-dimensional spaces that are absent in low-dimensional contexts. This phenomenon poses significant obstacles for outlier detection algorithms (Li et al., 2020). Equation (7) exemplifies the difference between low and high-dimensional settings. As the dimensionality increases, the likelihood of encountering values less than or equal to l_j ($\widetilde{L}_{j,i} \le l_j$) for all dimensions diminishes exponentially. Li et al., (2020) prevented the reduction of tail probabilities, and leveraging the monotonicity property of the log () function, the sum of negative log probabilities is utilized instead, thus:

$$\begin{aligned} -\log(\widehat{C}(\boldsymbol{l})) &= -\log\left(\mathbb{P}(\widetilde{L}_{1,i} \le l_1) \times \ldots \times \mathbb{P}(\widetilde{L}_{t,i} \le l_t)\right) \\ &= -\sum_{j=1}^{t}\log\left(\mathbb{P}(\widetilde{L}_{j,i} \le l_j)\right) = -\sum_{j=1}^{t}\log(l_j) \end{aligned} \tag{9}$$

The validity of the last statement is upheld due to the uniform distribution of $\widetilde{L}_{1,i}$ on [0, 1].

In the employed anomaly detection methodology, copula-based techniques were utilized to capture variable dependencies inherently, circumventing the need for additional transformations like centered log-ratio (clr), additive log-ratio (alr), and isometric log-ratio (ilr) transformations. Through copula functions, multivariate dependencies are flexibly modeled while preserving the original data scale and distribution. Consequently, in the conducted analysis, the necessity for these transformation techniques is obviated, as copula-based anomaly detection provides a direct and interpretable approach to identifying anomalies in multivariate datasets. In terms of computational cost, previous benchmark studies show that COPOD requires about 0.23s on average to process standard test datasets, which is comparable to other commonly used methods, such as isolation forest (IF). COPOD avoids pairwise distance calculations and does not require model training, which

Table 1
Counts of extreme data values per element through analysis of probability plots.

Element	Ag	As	Au	Ba	Bi	Co	Cr	Cu	Hg
Number of extremes	1	1	4	4	1	0	0	4	2
Element	Mo	Ni	Pb	Sb	Sn	Sr	V	W	Zn
Number of extremes	1	1	13	2	0	3	3	2	9

Table 2
Summary statistics of uni-element concentrations in the stream sediment samples collected from the study area (values in ppm).

Variable	Mean	St.Dev	Minimum	Maximum	Skewness
Ag	0.16	0.56	0.03	7.95	11.90
As	13.31	5.85	4.80	66.80	3.01
Au	0.003	0.007	0.0004	0.100	9.160
Ba	544.07	277.73	143.48	2922.32	4.77
Bi	0.28	0.41	0.08	6.13	8.69
Co	15.32	2.93	6.46	28.52	0.46
Cr	73.47	19.56	28.75	198.12	1.77
Cu	53.51	40.50	16.74	401.81	4.66
Hg	0.02	0.07	0.01	1.17	12.60
Mo	1.40	0.54	0.33	6.40	2.30
Ni	33.56	6.46	13.02	83.64	1.36
Pb	44.53	45.52	14.54	416.19	5.16
Sb	1.00	1.21	0.30	17.09	8.74
Sn	1.28	0.32	0.66	3.12	1.10
Sr	669.30	730.10	210.60	6207.10	4.34
V	147.03	56.05	47.03	453.29	1.65
W	0.97	0.29	0.41	3.40	2.07
Zn	102.21	128.11	42.13	1879.94	9.93

reduces its computational cost. Tests on synthetic datasets also show that COPOD can handle datasets with up to 10,000 dimensions and 1,000,000 observations within reasonable computation times on a standard personal computer, making it suitable for large geochemical datasets (Li et al., 2020).

COPOD was implemented using the PyOD library with its default settings (https://github.com/winstonll/COPOD/blob/master/models/cod.py)(Zhao et al., 2019). The method is almost parameter-free and does not require model training or hyperparameter tuning. The only parameter is the contamination rate, which was set to 0.1 and used only to define the decision threshold, meaning that the top 10% of samples with the highest anomaly scores were labeled as anomalies. Empirical cumulative distribution functions (ECDFs) were calculated independently for each variable to estimate left- and right-tail probabilities. Skewness was used automatically to select the more relevant tail, and the final outlier score for each sample was obtained by summing the negative log-probabilities across all dimensions.

4. Results and discussion

Initially, extreme values (mean + 3* St.Dev) (Table 1) were discerned using probability plots and box plots. Among the elements Au, Ba, Hg, Sb, W, As, Bi, Ag, Ni, Mo, Cr, Co, and Sn, only a minimal fraction of extreme values were detected. These outliers were rectified by substituting them with the 95th percentile value derived from the remaining dataset. Conversely, for elements Zn, Pb, Cu, Sr, and V with a higher number of outliers, a regression model was employed to truncate and anticipate extreme values. Although replacing extreme values in geochemical datasets is generally discouraged due to the potential loss of important mineralization signals, in this study the number of extreme values was very small relative to the total dataset (819 samples). The COPOD method relies on marginal empirical distributions, and replacing a few extreme values with the 95th percentile helps to reduce the influence of potential measurement errors or noise on tail probability estimation. This replacement stabilizes the empirical cumulative distribution functions without removing true anomalies, as the adjusted values remain sufficiently high to be identified as outliers by COPOD. Table 2 shows the summary statistics of elemental values post-extreme removal, comprising the geochemical dataset under analysis. Additionally, concentration histograms are depicted in Fig. 4. Upon examination of the histograms and summary statistics, it becomes evident that

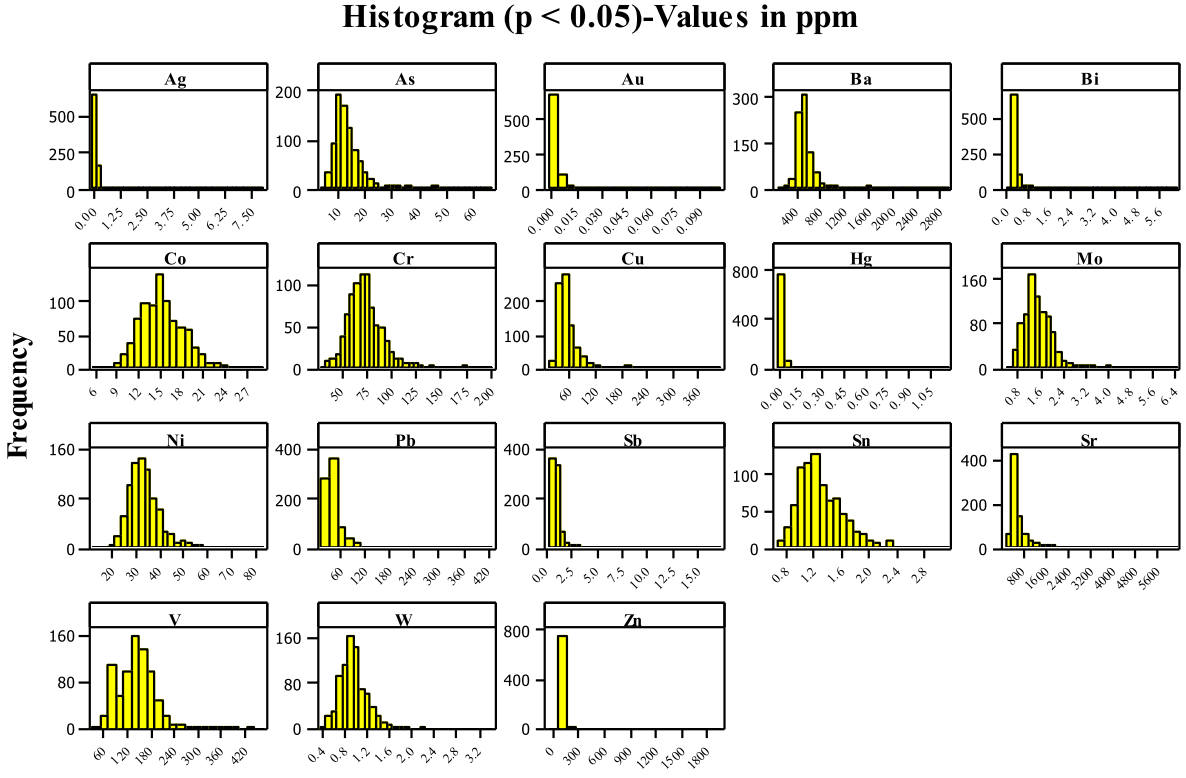

Fig. 4. Histograms of concentrations of 18 elements from the study area.

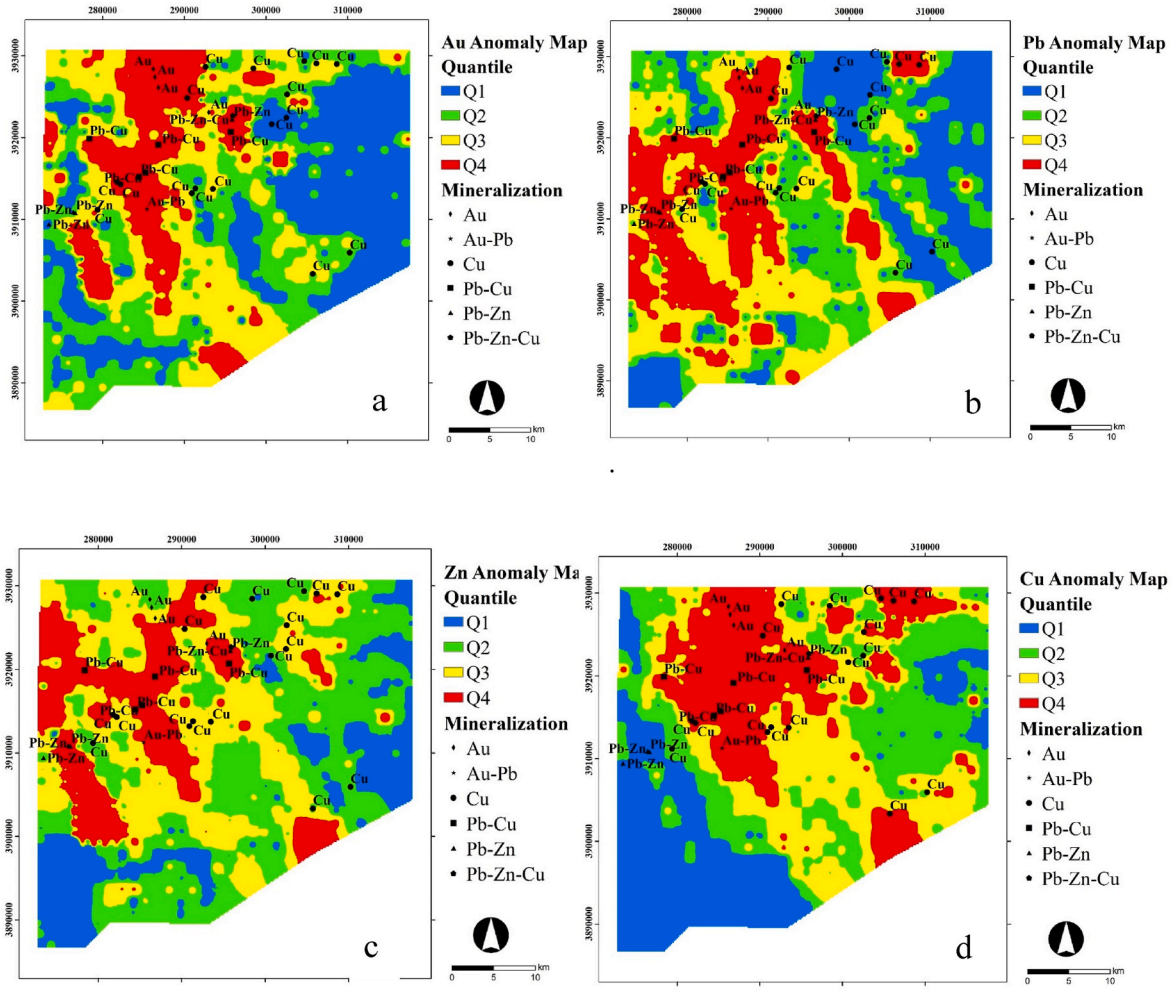

Fig. 5. Uni-element geochemical anomalies identified using the quantile classification for (a) Au, (b) Pb, (c) Zn, and (d) Cu.

elements such as Ag, Au, Bi, Hg, Sb, and Zn exhibit substantial skewness in their distributions. Conversely, Co, Cr, Ni, and V portray more symmetrical distributions, which is a prevalent trait observed in geochemical datasets, particularly observed in elements with very low concentrations, where skewed distributions are commonly encountered.

One way to delineate geochemical anomalies is mapping the main uni-element elemental concentrations such as Cu, Pb, Zn, and Au. Given that the distribution of these elements may deviate from normality, one straightforward approach to delineating geochemical anomalies is to utilize the quantile-based classification. Uni-element geochemical anomalies of Au, Zn, Cu, and Pb are presented in Fig. 5. The Cu anomaly mapping encountered challenges in accurately delineating some known Cu occurrences within the Cu anomaly class. Despite this difficulty, a distinct Cu Anomaly zone in the northern part of the study area clearly

Table 3
Spearman coefficients between empirical copula values of different elements.

	Ag	As	Au	Ba	Bi	Co	Cr	Cu	Hg	Mo	Ni	Pb	Sb	Sn	Sr	V	W	Zn
Ag	1																	
As	0.00	1.00																
Au	0.12	0.05	1.00															
Ba	0.18	−0.08	0.01	1.00														
Bi	0.07	0.16	0.25	−0.06	1.00													
Co	0.14	0.11	0.17	−0.09	0.19	1.00												
Cr	0.01	0.12	−0.01	0.07	0.03	0.27	1.00											
Cu	0.22	0.00	0.18	0.00	0.21	0.47	0.13	1.00										
Hg	0.21	0.08	0.10	0.18	0.09	0.18	0.10	0.22	1.00									
Mo	0.23	0.01	0.14	0.08	0.19	0.18	0.08	0.20	0.17	1.00								
Ni	−0.04	0.05	0.00	−0.02	0.09	0.14	0.32	0.11	0.05	0.04	1.00							
Pb	0.34	0.06	0.22	0.08	0.18	0.22	0.10	0.27	0.24	0.30	0.05	1.00						
Sb	0.17	0.20	0.09	0.06	0.11	0.23	0.14	0.06	0.15	0.15	0.05	0.25	1.00					
Sn	0.14	0.03	−0.01	0.07	0.06	0.15	0.11	0.13	0.05	0.05	0.10	0.07	0.11	1.00				
Sr	0.16	−0.21	−0.09	0.37	−0.12	0.01	0.03	0.07	0.17	0.11	0.03	0.04	−0.03	0.13	1.00			
V	0.09	0.02	−0.04	0.00	0.00	0.39	0.12	0.13	0.05	−0.04	−0.14	0.09	0.29	0.09	0.03	1.00		
W	0.09	0.13	0.23	−0.09	0.38	0.28	0.16	0.33	0.12	0.26	0.14	0.25	0.16	0.08	−0.09	0.04	1.00	
Zn	0.24	0.14	0.19	0.06	0.18	0.37	0.18	0.35	0.19	0.14	0.04	0.44	0.14	0.13	0.07	0.16	0.26	1.00

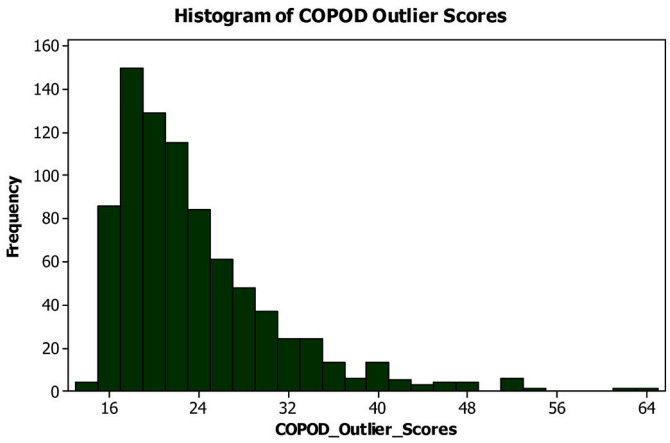

Fig. 6. Histogram of outlier scores generated by the COPOD method.

identifies known Pb–Cu and Au occurrences. Additionally, the Cu anomaly map shows spatial relation with a Cu occurrence in the southeast of the area. In contrast, the Pb anomaly maps highlighted all known Pb–Cu occurrences and some known Pb–Zn occurrences in the western region. Zn, with its greater mobility compared to Pb, tends to be more dispersed and diluted within stream sediments (Bouzekri et al., 2019; Drahota et al., 2024). However, despite this, the Zn anomaly map highlighted several known Pb–Zn and Pb–Cu occurrences. The Au anomaly map, similar to the Cu anomaly map, revealed a significant anomaly in the northern part, encompassing most of the known Au occurrences and several known Pb–Cu occurrences. The Au, Pb, and Zn uni-element anomaly maps demonstrated similarity in highlighting three known mineral occurrences in the northern part, including Pb–Cu, Pb–Zn, and Pb–Zn–Cu occurrences. However, distinct patterns emerged between the Au and Cu anomaly maps, indicating that varying lithological and geochemical factors influence the transportation and concentration of these elements within stream sediments in the study area.

In order to implement the COPOD method, the algorithm introduced by Li et al., 2020 was applied. This involved computing the ECDFs for both the left and right tails, along with skewness coefficients. Following this, empirical copula observations were computed. In the concluding stage, tail probabilities were estimated using Eq. (9), and the outlier scores were identified as the maximum values among the left, right, and skewness-corrected probabilities. Table 3 presents the Spearman coefficients derived from empirical copula values obtained through the COPOD method. The analysis revealed notable geochemically relevant relationships among the various elements. Significant correlations were observed among pairs such as Ag–Pb, Bi–Au, Pb–Au, Cr–Co, Cu–Co, Ni–Cr, Ba–Sr, Bi–W, V–Co, Pb–Cu, and Cu–Zn. These correlations can be linked to distinct mineral deposits within the study area. In the case of Bi–Au, trace minerals such as emplectite ($CuBiS_2$), wittichenite (Cu_3BiS_3), tellurobismuthite (Bi_2Te_3), tetradymite (Bi_2Te_2S), and native gold have been identified in the Darestan Cu ± Au epithermal deposit (TaleFazel et al., 2019). In a different example, the Khanjar carbonate-hosted deposit, the Ag–Pb association has been reported to show an enrichment of Ag ranging between 90 and 2400 g/t (Sabahi and Ebrahimi, 2015). Moreover, the Abolhassani and Gandy deposits are dominated by primary ore minerals such as chalcopyrite, pyrite, sphalerite, galena, and chalcocite. Additionally, trace minerals identified in these deposits include bornite, tennantite, emplectite ($CuBiS_2$), argentite, and native gold (TaleFazel et al., 2019).

Fig. 6 depicts the histogram of outlier scores generated by the COPOD method. Fig. 7 illustrates the classified anomaly map based on quantile classification of the anomaly map derived from the application of the COPOD method to the geochemical dataset from the study area. The classified anomaly map effectively identified all Au, Pb–Cu and Au–Pb, and Pb–Zn occurrences. Regarding Cu occurrences, while the COPOD method delineated several of these occurrences, similar to uni-element Cu anomaly map, some mineral occurrences were not delineated due to sampling scheme limitations or the presence of complex multivariate patterns in the study area that the COPOD method cannot capture. In order to compare the effectiveness of the multivariate COPOD method and the uni-element anomaly mapping method, two procedures were followed. In the first comparison, the number of mineral occurrences delineated in each quantile class (Q1 to Q4) of four uni-element and one COPOD-derived anomaly maps was contrasted (Table 4). The prospectivity maps in this study focus on Au, Zn, Cu, and Pb, as these are the main commodities of economic interest in the region

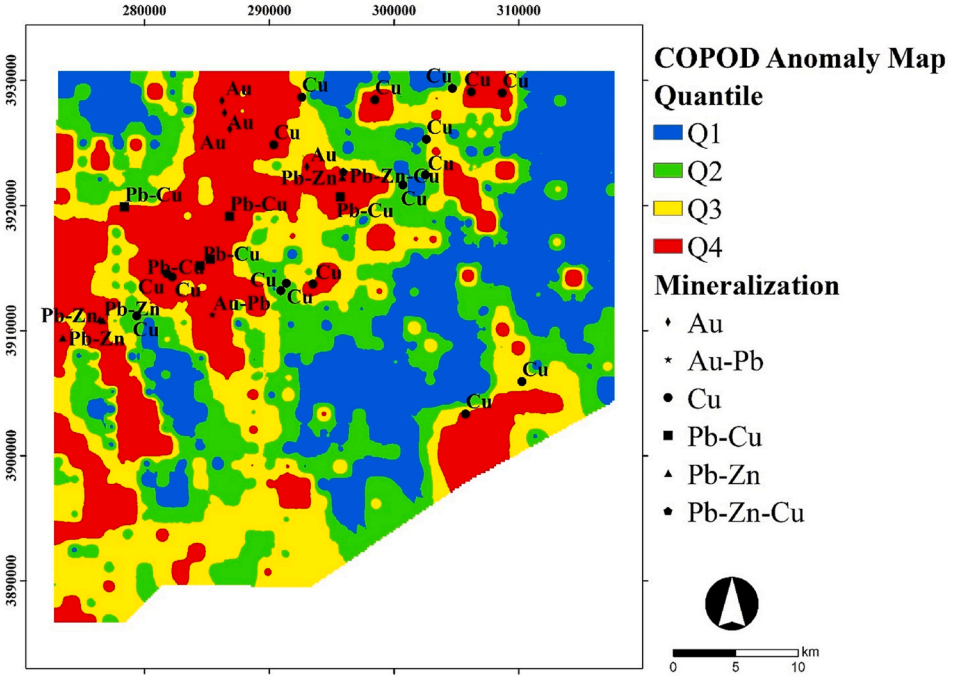

Fig. 7. Quantile-based anomaly map derived by the COPOD method.

Table 4
Number of known mineral occurrences in each anomaly class of the uni-element and COPOD-derived geochemical anomalies.

Anomaly map		Mineralization							
	Quantile	Cu	Pb-Zn	Au-Pb	Au	Pb-Cu	Pb-Zn-Cu	Total	Percentage
Au	Q1	2	2					4	13
	Q2	5	1					6	19
	Q3	7			1			8	25
	Q4	3	1	1	3	5	1	14	**44**
Zn	Q1	1						1	3
	Q2	5	1		2			8	25
	Q3	11			1			12	38
	Q4		3	1	1	5	1	11	**34**
Pb	Q1	4						4	13
	Q2	5	1					6	19
	Q3	4			1			5	16
	Q4	4	3	1	3	5	1	17	**53**
Cu	Q1		3					3	9
	Q2	3						3	9
	Q3	6						6	19
	Q4	8	1	1	4	5	1	20	**63**
COPOD	Q1							0	0
	Q2	7						6	18
	Q3	2					1	3	9
	Q4	8	4	1	4	5	0	23	**71**

and provide a clear basis for comparison between single-element anomaly mapping and the multivariate COPOD approach. Shahrestani and Carranza (2024) conducted a comprehensive principal component analysis (PCA) on the same geochemical dataset, which revealed significant elemental associations linked to distinct mineralization styles and lithological controls. Their results showed that the first seven principal components explain approximately 80% of the total elemental variability. Notably, PC2 highlights a strong association among Au, Bi, Cu, and W, indicative of epithermal gold mineralization characterized by the presence of chalcopyrite, pyrite, sphalerite, galena, and trace minerals such as bornite and tennantite in key deposits like Abolhassani and Gandy. PC5 and PC7 reflect hydrothermal and intrusive-related polymetallic mineralization, respectively, with elemental groups including Zn, Ag, Ni, Pb, Cu, Mo, V, and Ba. These PCA-derived elemental clusters provide robust justification for selecting Au, Zn, Cu, and Pb as pathfinder elements, as they capture the primary mineralization types and geochemical processes within the study area. Overall, the COPOD-derived and Cu anomaly maps showed the highest consistency between geochemical anomalies and known mineral occurrences. However, the most predominant feature of the Cu anomaly map was the high possibility of disregarding geochemical anomalies that originated from the Pb-Zn mineral deposits. In the case of COPOD-derived anomaly map, the contribution of geochemical anomalies related to Pb-Zn mineral deposits was also taken into account while similar number of Cu mineralization was delineated by the Cu and COPOD-derived anomaly maps in the highest anomaly rank (i.e., Q4). Moreover, regarding highest anomaly rank, the COPOD-derived anomaly map was more effective than either of the Pb and Zn anomaly maps in delineating Pb-Zn mineral occurrences. The same scenario was true when considering the COPOD-derived and Au anomaly maps in which all known Au mineralization fell into the highest anomaly class of the COPOD-derived anomaly map.

Another approach used to compare the anomaly maps was to utilize the receiver operating characteristic (ROC) curves whereby true positive rate is plotted against false positive rate at different threshold values (Fawcett, 2006). By incorporating labeled data representing known mineral deposits, ROC curves can be constructed to assess the performance of anomaly maps in which the area under the curve is utilized as a holistic performance metric across all threshold settings. In the ROC procedure, positive samples correspond to pixels that coincide with known mineral occurrences while negative samples are randomly selected pixels with no known mineral occurrences. Fig. 8 illustrates the ROC curve considering the four uni-element and one COPOD-derived

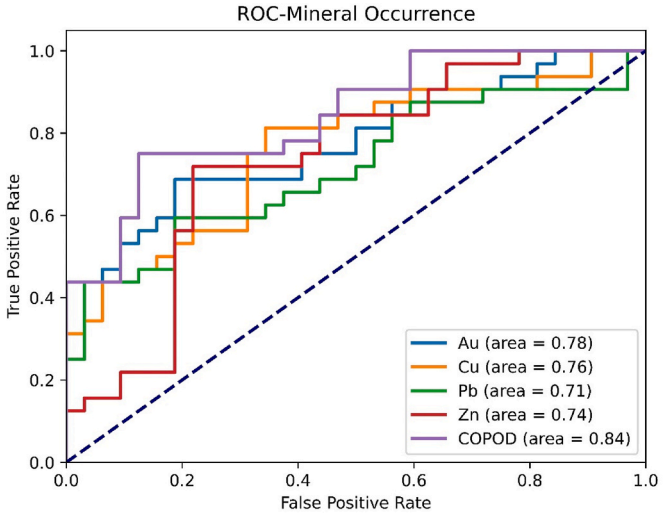

Fig. 8. ROC curves generated from the four univariate and one COPOD-derived anomaly maps, using the area under the curve as a proxy for efficiency.

anomaly maps. As can be seen, there was a considerable growth in area under curve value when the COPOD method was applied for geochemical anomaly delineation. We performed five repetitions of ROC analysis using different random selections of negative samples, and the AUC values for COPOD ranged from 78% to 87%. In all cases, COPOD anomaly maps outperformed the corresponding AUC values of the single-element maps.

The notable aspect of the COPOD method lies in its capacity to utilize the dimension outlier graph, which aids in comprehending the underlying causes of abnormality. To further illustrate this feature, dimension outlier graphs for the samples with the top five highest outlier scores (ID: 235, 339, 204, 793, and 72) are depicted in Fig. 8, along with their respective spatial locations. Fig. 9 also demonstrates that the COPOD-derived anomaly map, classified using a conservative 10% contamination fraction, effectively captures 21 out of 32 known mineral occurrences within the anomaly class. This indicates that despite the relatively small spatial extent of anomalous areas (approximately 10% of the study region), the COPOD method efficiently identifies geochemical anomalies associated with mineralization. There are both similarities and discrepancies observed in the trend of dimensional

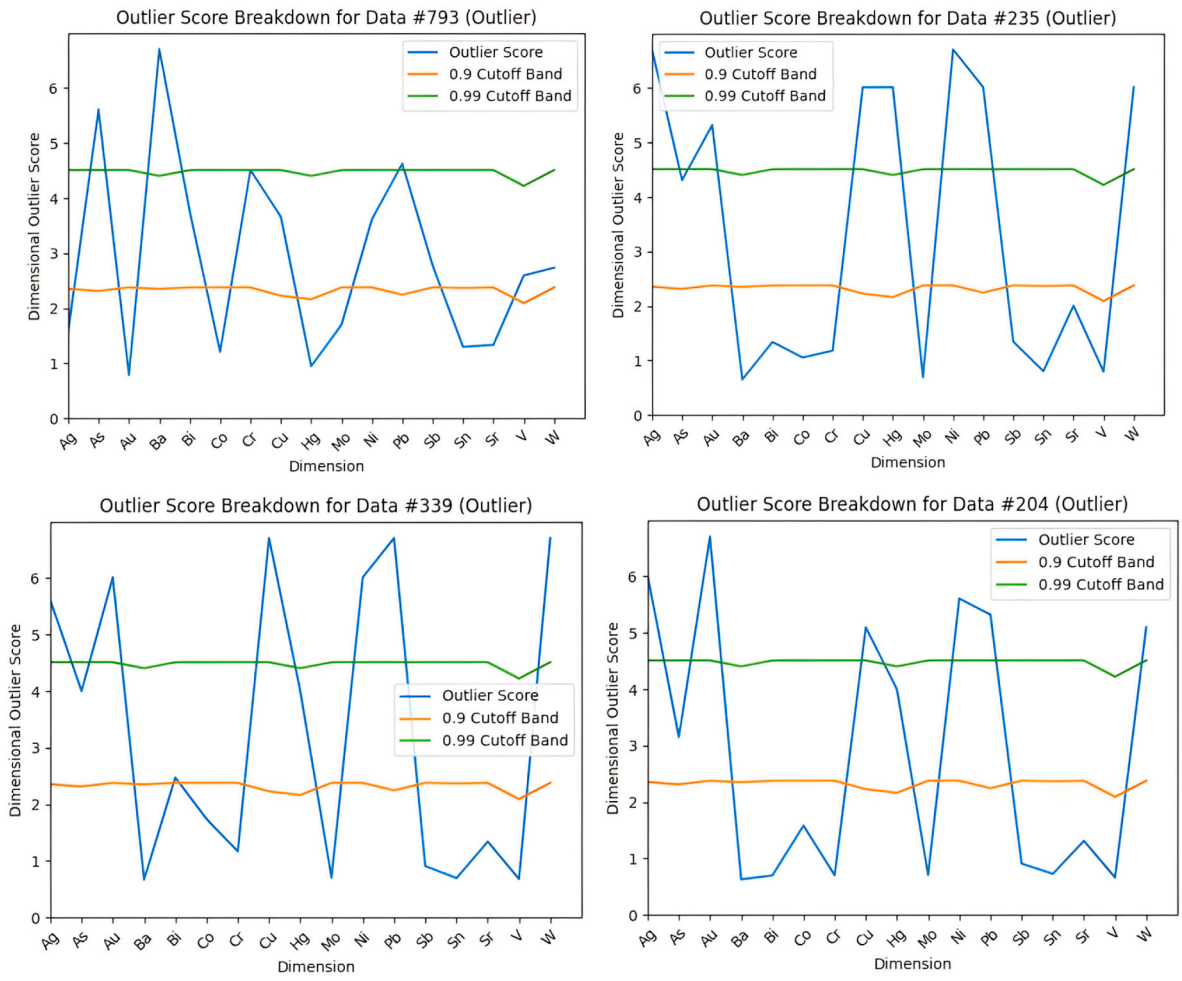

Fig. 9. Dimension outlier graphs for samples with the top five highest outlier scores (ID: 235, 339, 204, 793, and 72), along with their respective spatial locations based on the COPOD anomaly map encompassing 10-percent contamination fraction.

outlier scores among elements for the top high-scored samples. Specifically, samples with ID 204, 235, and 339, located downstream of two Pb–Zn occurrences in the western part of the study area, exhibit similar fluctuations in elemental concentrations. Notably, these samples show enrichment in Pb, Hg, Cu, and Au, suggesting the potential for placer-type mineralization originating from upstream anomalous sources. Distinct patterns were also evident in the case of samples 72 and 793, which demonstrate enrichment in Au pathfinder elements such as As and Bi. These samples are situated in close proximity to a zone favorable for Au mineralization. Additionally, among all five cases, elements such as Co, Sr, and Mo did not contribute to abnormality in these samples. It is important to note that, while these interpretations hold true for these specific samples, they may not be generalized across the entire study area.

In addition to visually representing the relative contributions of elements in outlier scores, this study utilized two methods for further analysis. Lasso regression, a technique within linear regression that incorporates L1 regularization, was employed for variable feature selection, exploring correlations among variables, and simplifying model complexity. Additionally, the random forest algorithm was utilized for tasks including classification, regression, and feature importance analysis. These methods were applied to the copula values of each dimension (element) and the outlier scores derived from COPOD. The random forest analysis utilized the following parameters: 100 trees in the forest, 2 splits for each tree, and a random state of 42 for reproducibility. Additionally, the dataset was partitioned into training (75%) and testing (25%) subsets for evaluation the performance of the model.

The order and coefficients obtained from Lasso regression are detailed in Table 5. From this regression analysis, the mean squared error (MSE) and R-squared values were computed as 14.54 and 0.69, respectively. The analysis suggests that elements such as Co, Zn, Sb, and Pb contributed significantly to the outlier scores derived by COPOD, while V, Ni, Cu, Sr, and Mo were considered less important.

In the random forest analysis, the MSE and R-squared values for the outlier scores of both the training and testing datasets were 5.51 and 0.89, respectively. Fig. 10 illustrates the relative impact of each feature (element) on the outlier scores derived by COPOD. Notably, elements like Sb, Co, Pb, Zn, and Au exhibited the highest impact on the COPOD-derived outlier scores, while V, As, Mo, Ni, and Sn had the least impact. Despite some variations, there was an overall similarity in the order derived from both methods.

The variability in the impact of elements on the anomaly detection process can be attributed to several factors. Firstly, the quality of available data, including the restricted variability observed in Mo among different samples, significantly influences the outcomes. Secondly, elements with known geochemical significance or associations with mineralization processes are more likely to influence anomaly scores. Conversely, lithological variation might primarily shape the distribution of geochemical anomalies for elements such as V, Ni, Cu, and Sr. Finally, elements that showcase more pronounced variability across samples typically exert a stronger influence on anomaly scores. This variability can be quantified using measures such as the skewness of

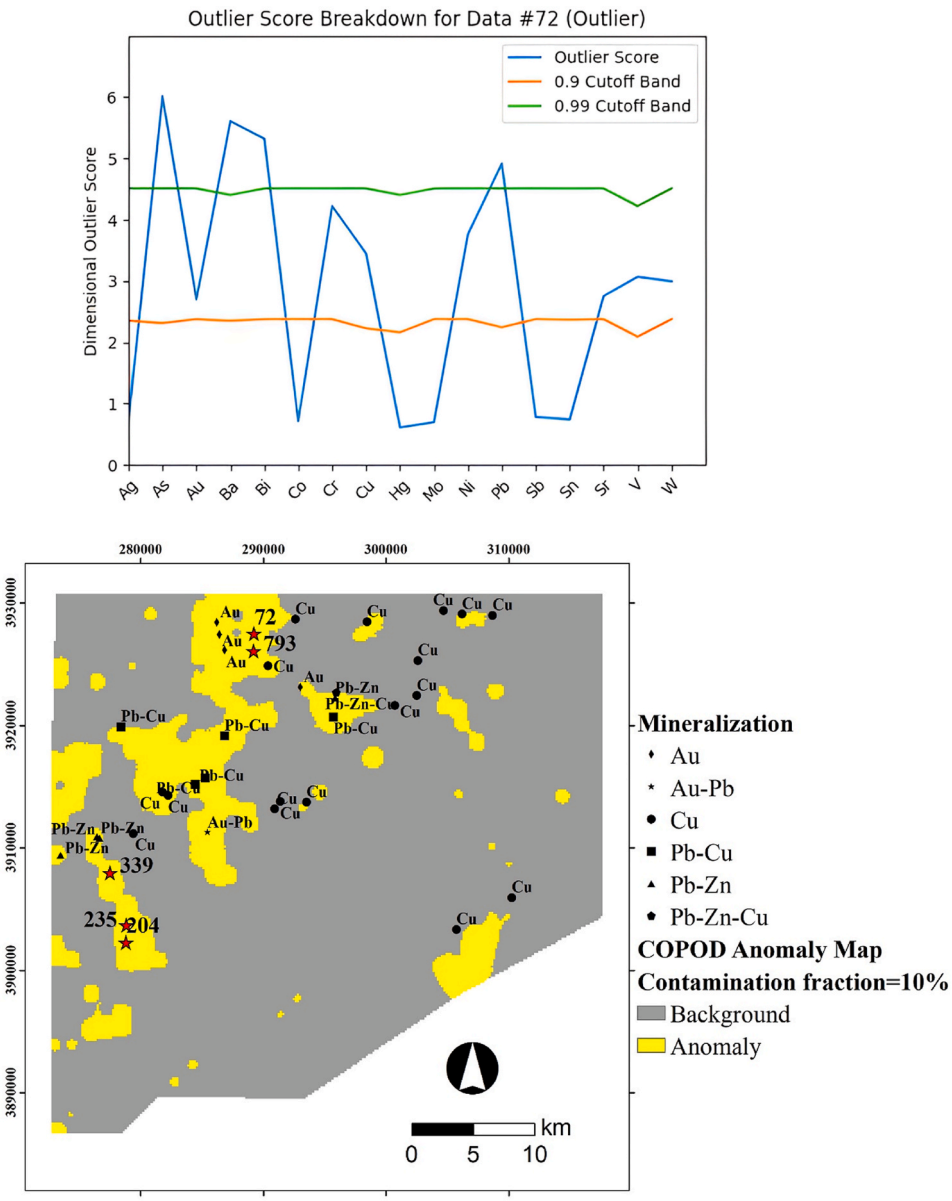

Fig. 9. (*continued*).

elemental values. However, there exists an exceptional case with Co. Despite its potentially lower skewness, Co significantly impacted the COPOD-derived outlier scores, likely due to other contributing factors such as its high correlations with Cu and Zn as primary elemental commodities (Table 3). Conversely, most elements that exhibited lower skewness values, including As, Cr, Ni, Mo, Sn, and V as indicated in Table 2, tended to have lesser impacts on COPOD-derived outlier scores.

Another way to assess the efficiency of the COPOD method in delineating geochemical anomalies is to compare its performance with other state-of-the-art techniques, such as the isolation forest (IF) method, which is robust to the curse of dimensionality (Shahrestani and Sanislav, 2025). Fig. 11 shows anomaly maps resulting from both COPOD and IF, classified using quartile thresholds. Overall, there is a general conformity between the spatial distribution of geochemical anomalies between the two methods; however, by focusing on the highest anomaly class (Q4), the superiority of the COPOD method over the IF anomaly map is evident, as it captures 23 mineral occurrences out of 32 compared to 19 localities by the IF method, however, some mineralization remain undetected in both anomaly maps.

Fig. 12 shows the overlap of the COPOD anomaly class with the regional geological map of the study area. Evidently, the anomaly class mainly covers the non-rocky zones of the area. However, since the outlier detection is based on multielement footprints derived from different lithological units, a complete agreement between the spatial distribution of anomalies and geological boundaries is not expected. Nevertheless, the anomaly class generally occurs within the Paleogene Etr unit, including trachyandesite, trachyandesitic basalt, and minor andesitic dacite, as well as the Et,v unit composed of submarine lavas and volcaniclastic rocks, and the Cretaceous Kl,shu unit consisting of shale, limestone, and minor sandstone, which host sedimentary Pb–Zn deposits.

Undoubtedly, samples containing elevated elemental values are more susceptible to being identified as outliers. However, when employing the COPOD method, a straightforward comparison of overall elemental concentrations in samples can be achieved through a scatter plot depicting the geometric means of the 18 elements against their anomaly scores. Geometric mean, being more robust to outliers than arithmetic mean, serves as a reliable metric. Based on the contamination factor and anomaly scores, samples with scores surpassing the threshold were categorized as the anomaly class, while the rest were labeled as

Table 5
Ranking of elements by Lasso regression based on outlier scores derived by COPOD.

Lasso Rank	Lasso Coefficient
Co	1.88
Zn	1.62
Sb	1.43
Hg	1.41
Pb	1.38
Ag	1.34
Ba	1.26
Au	1.12
Bi	1.09
Cr	0.98
Sn	0.87
As	0.76
W	0.73
Mo	0.67
Sr	0.66
Cu	0.59
Ni	0.58
V	0.49

background samples. As depicted in Fig. 13, there exists a noticeable disparity in the relationship between anomaly scores and geometric means of samples in the background class compared to those in the anomaly class. For background samples, the relationship appears less pronounced, indicating a lack of strong association between the two variables. On the other hand, a significant relationship exists between the anomaly score of each sample and the corresponding geometric mean of the concentrations of 18 elements in that sample within the anomaly class. This pattern is consistent with the typical behavior observed when employing outlier detection techniques in multi-element geochemical datasets.

However, implementation of the COPOD method in the current case study to detect geochemical anomalies faced several limitations. First, the reliance of the method on 18 multielement distributions may treat mineralization-related and unrelated elements similarly, potentially orienting geochemical anomalies. Second, the selection of copula models may influence the results, necessitating a comprehensive survey to understand accurately the efficacy of different copula models in capturing geochemical dependencies. Third, as in the current study, the tendency of the COPOD method to detect anomalies could be influenced more by lithological variations than by mineralization-related anomalies in certain study areas, complicating the interpretation of detected anomalies. Lastly, a lack of predefined criteria or pre-selection procedures in the COPOD method could hinder its ability to effectively filter out noise and focus on relevant anomalies, potentially leading to false positives or missed detections.

The results of this study also highlight directions for future research. Some mineral occurrences, particularly Cu mineralization, were not delineated by the COPOD method, which may be related to sampling density, lithological mixing, or complex multivariate geochemical patterns in stream sediment data. COPOD primarily focuses on marginal tail behaviour and may therefore have limited ability to identify outliers within mixed or overlapping geochemical populations, which are common in stream sediment samples. Future studies could evaluate the performance of COPOD after background correction or population separation to reduce lithological and catchment effects. In addition, integrating COPOD with other multivariate or ensemble anomaly detection methods, or incorporating geological and structural information, may improve the detection of subtle or masked mineralization signals. Further applications to datasets from different geological environments would also help to assess the robustness and broader applicability of the method.

5. Conclusions

1. The copula-based outlier detection (COPOD) method emerges as a highly effective tool for identifying geochemical anomalies, particularly in the context of mineral exploration. Through its application, COPOD accurately delineated the majority of known mineral occurrences in the study area.
2. The generation of uni-element anomaly maps proved to be a robust initial step in highlighting anomalous concentrations of key elements such as Cu, Pb, Zn, and Au. These maps provided valuable insights into the spatial distribution of mineralization and served as a foundation for further analysis.
3. Correlations between empirical copula values across various geochemical elements revealed significant relationships indicative of distinct mineralization types within the study area. These

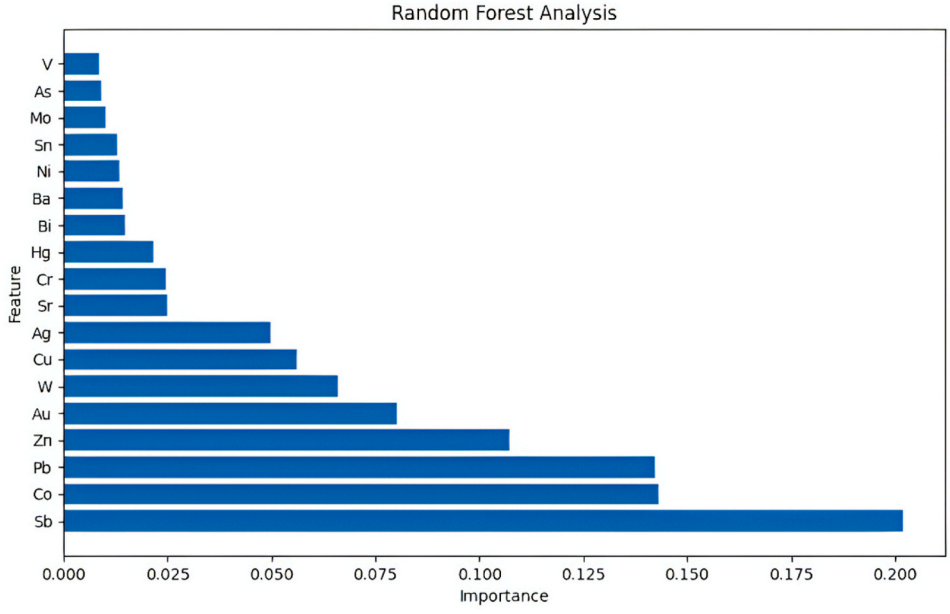

Fig. 10. Relative influence of each element determined by random forest analysis of the outlier scores generated by COPOD.

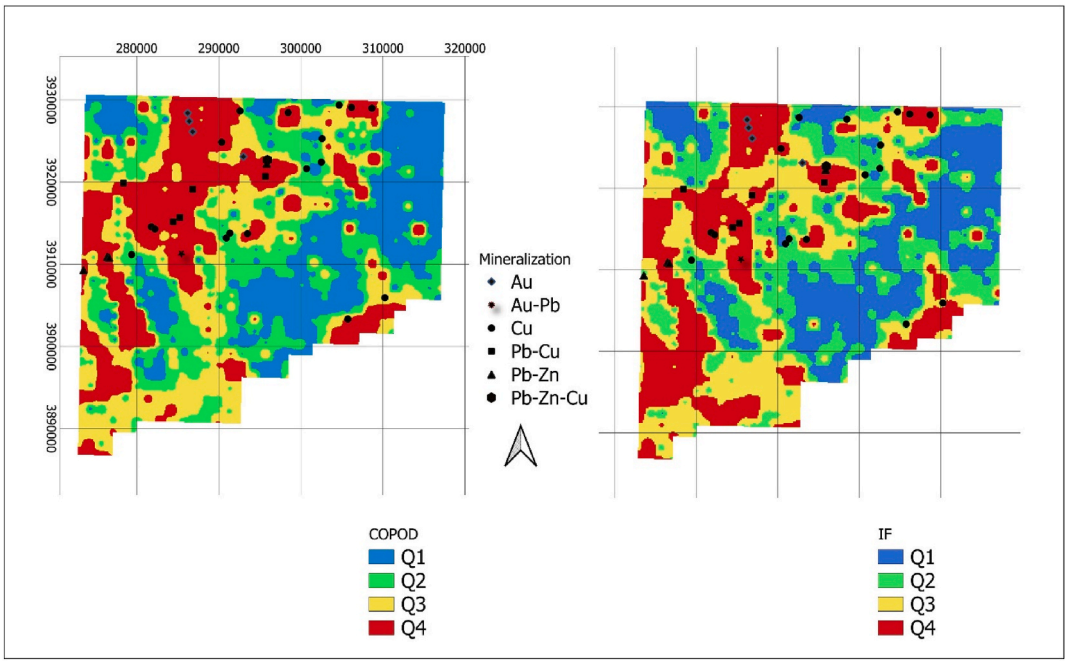

Fig. 11. Classified geochemical anomalies identified by COPOD and IF methods, shown alongside known mineral occurrences.

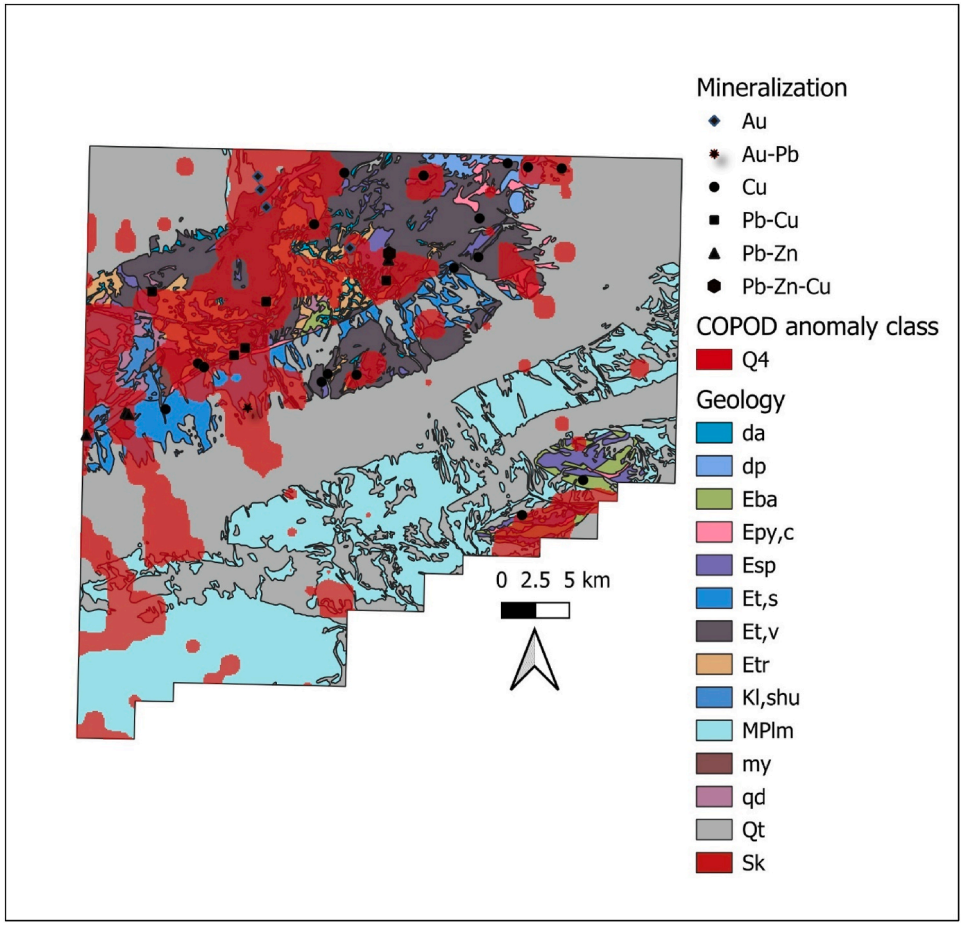

Fig. 12. Spatial overlap of the COPOD anomaly class with the regional geological map of the study area.

correlations offered valuable insights into the underlying geological processes driving anomaly formation.

4. Through interpretation of dimension outlier graphs, valuable insights were gained into the elemental contributions to anomaly

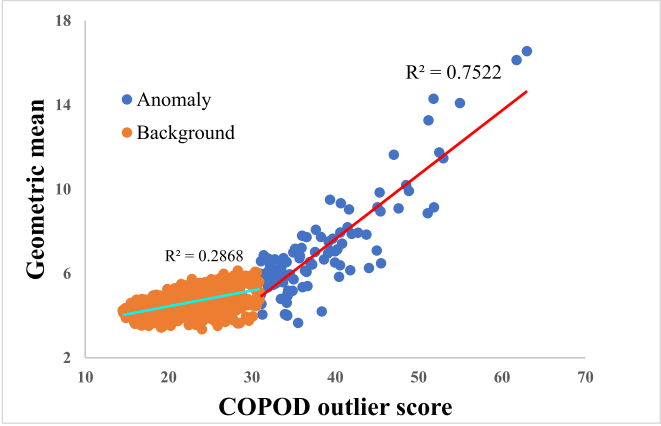

Fig. 13. Scatter plot depicting the relationship between the geometric means of concentrations of 18 elements per sample and the corresponding outlier scores derived from the COPOD method.

scores. Elements such as Co, Zn, Sb, and Pb were identified as having substantial influences on anomaly scores, highlighting their importance in anomaly detection.

5. The analysis of outlier scores between anomaly and background classes demonstrated distinct behaviors, with anomalies showing a strong correlation between outlier scores and geometric mean values, while background samples exhibited a weaker correlation.
6. The COPOD-derived anomaly map effectively delineated known Au, Pb–Cu, and Au–Pb occurrences. However, for Cu occurrences, some of these were not delineated by the COPOD method perhaps due to sampling limitations or complex patterns. Comparing multivariate and uni-element approaches, the COPOD-derived and Cu anomaly maps showed the highest consistency with known mineral occurrences. The COPOD method was effective in delineating Pb-Zn occurrences, outperforming Pb and Zn anomaly maps in the highest anomaly rank. Similarly, all known Au occurrences were captured in the highest anomaly class of the COPOD-derived anomaly map. The study compared the derived anomaly maps using ROC curves and found that the COPOD-derived map outperformed the uni-element maps in delineating geochemical anomalies, as evidenced by a substantial increase in the area under the curve of the ROC curve. This suggests that, compared to uni-element anomaly mapping, the COPOD method provides for more accurate and effective identification of deposit-related geochemical anomalies.
7. Further analysis using Lasso regression and random forest methods revealed the substantial impacts of individual elements on anomaly detection. Elements such as Co, Zn, Sb, and Pb were found to significantly influence anomaly scores, providing additional insights into the geochemical processes driving anomaly formation.

CRediT authorship contribution statement

Shahed Shahrestani: Writing – review & editing, Writing – original draft, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Emmanuel John M. Carranza:** Writing – review & editing, Supervision, Formal analysis. **Ioan Sanislav:** Writing – review & editing, Supervision.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

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.

Acknowledgments

The authors gratefully acknowledge the Geological Survey of Iran for generously providing the regional stream sediment geochemical data of the Moalleman area, which greatly contributed to this study.

Data availability

Data will be made available on request.

References

Azmi, H., Moarefvand, P., Maghsoudi, A., 2021. Gold anomaly ranking based on stream sediment geochemistry in the Fariman–Kashmar axis, NE Iran. Acta Geochimica 40, 135–149.
Bigdeli, A., Maghsoudi, A., Ghezelbash, R., 2022. Application of self-organizing map (SOM) and K-means clustering algorithms for portraying geochemical anomaly patterns in Moalleman district, NE Iran. J. Geochem. Explor. 233, 106923.
Bigdeli, A., Maghsoudi, A., Ghezelbash, R., 2023. Recognizing geochemical anomalies associated with mineral resources using singularity analysis and random forest models in the Torud-Chahshirin Belt, Northeast Iran. Minerals 13 (11), 1399.
Bouzekri, S., El Hachimi, M.L., Touach, N., El Fadili, H., El Mahi, M., 2019. The study of metal (As, Cd, Pb, Zn and Cu) contamination in superficial stream sediments around of Zaida mine (High Moulouya-Morocco). J. Afr. Earth Sci. 154, 49–58.
Carranza, E.J.M., Hale, M., 1997. A catchment basin approach to the analysis of reconnaissance geochemical-geological data from Albay Province, Philippines. J. Geochem. Explor. 60 (2), 157–171.
Chen, Y., 2015. Mineral potential mapping with a restricted Boltzmann machine. Ore Geol. Rev. 71, 749–760.
Chen, Y., Sui, Y., 2022. Dictionary learning for integration of evidential layers for mineral prospectivity modeling. Ore Geol. Rev. 141, 104649.
Chen, Y., Wu, W., 2017. Application of one-class support vector machine to quickly identify multivariate anomalies from geochemical exploration data. Geochem. Explor. Environ. Anal. 17 (3), 231–238.
Chen, Y., Wu, W., 2019. Isolation forest as an alternative data-driven mineral prospectivity mapping method with a higher data-processing efficiency. Nat. Resour. Res. 28 (1), 31–46.
Chen, Y., Wu, W., Zhao, Q., 2019. A bat-optimized one-class support vector machine for mineral prospectivity mapping. Minerals 9 (5), 317.
Chen, Y., Sun, G., Zhao, Q., 2021a. Detection of multivariate geochemical anomalies associated with gold deposits by using distance anomaly factors. J. Geochem. Explor. 221, 106704.
Chen, Y., Zheng, C., Sun, G., 2021b. Gold prospectivity modeling by combination of Laplacian eigenmaps and least angle regression. Nat. Resour. Res. 1–18.
Chen, Y., Du, X., Guo, M., 2023. Self-paced ensemble for constructing an efficient robust high-performance classification model for detecting mineralization anomalies from geochemical exploration data. Ore Geol. Rev., 105418
Drahota, P., Chatziantoniou, A., Mihaljevič, M., Culka, A., Melfou, M., Melfos, V., Voudouris, P., 2024. A mineralogical and geochemical assessment of the As-, Cu-, In-, Pb-, Sb-, and Zn-rich mine wastes at the Pefka epithermal deposit, Greece. J. Geochem. Explor. 256, 107336.
Eshraghi, S.A., Jalali, A., 2006. Geological Map of Moalleman, 1: 100000. Geological Survey of Iran (GSI).
Esmaeiloghli, S., Tabatabaei, S.H., Carranza, E.J.M., 2023. Infomax-based deep autoencoder network for recognition of multi-element geochemical anomalies linked to mineralization. Comput. Geosci. 175, 105341.
Fawcett, T., 2006. An introduction to ROC analysis. Pattern Recognit. Lett. 27 (8), 861–874.
Ghezelbash, R., Maghsoudi, A., Bigdeli, A., Carranza, E.J.M., 2021. Regional-scale mineral prospectivity mapping: support vector machines and an improved data-driven multi-criteria decision-making technique. Nat. Resour. Res. 30, 1977–2005.
Hu, D., 1994. Explanatory text for geological map of Torbat-E-Heydarieh, Stream Sediment Survey. Scale 1, 100000, Report No. 24.
Hushmandzadeh, A.R., Alavi Naini, M., Haghipour, A.A., 1978. Evolution of Geological Phenomenon in Totud Area. Geological Survey of Iran Report H5, p. 136.
Koohzadi, F., Afzal, P., Jahani, D., Pourkermani, M., 2021. Geochemical exploration for Li in regional scale utilizing Staged Factor Analysis (SFA) and Spectrum-Area (SA) fractal model in north central Iran. Iranian J. Earth Sci. 13 (4), 299–307.
Li, C., Zhang, B., Hong, D., Yao, J., Chanussot, J., 2023. LRR-Net: An interpretable deep unfolding network for hyperspectral anomaly detection. IEEE Transactions on Geoscience and Remote Sensing 61, 1–12.
Li, Z., Zhao, Y., Botta, N., Ionescu, C., Hu, X., 2020. COPOD: copula-based outlier detection. In: 2020 IEEE International Conference on Data Mining (ICDM). IEEE, pp. 1118–1123.

Liu, Y., 2025. Quantifying uncertainty of mineral prediction using a novel Bayesian deep learning framework. Artificial Intelligence in Geosciences 100164.
Liu, Y., Sun, T., Wu, K., Zhang, J., Zhang, H., Pu, W., Liao, B., 2025. Tungsten prospectivity mapping using multi-source geo-information and deep forest algorithm. Ore Geology Reviews 106452.
Mehrabi, B., Ghasemi, S.M., Tale, F.E., 2015. Structural control on epithermal mineralization in the Troud-Chah Shirin belt using point pattern and Fry analyses, north of Iran. Geotectonics 49, 320–331.
Nelsen, R.B., 2006. An Introduction to Copulas, second ed. Springer.
Puchhammer, P., Kalubowila, C., Braus, L., Pospiech, S., Sarala, P., Filzmoser, P., 2024. A performance study of local outlier detection methods for mineral exploration with geochemical compositional data. J. Geochem. Explor. 258, 107392.
Sabahi, F., Ebrahimi, S., 2015. Geology and mineralogy of Khanjar Rashm Pb-Zn deposit, SE Semnan. The 7th Conference of the Economic Geology Association of Iran.
Shahrestani, S., Carranza, E.J.M., 2024. Effectiveness of LOF, iForest, and OCSVM in detecting anomalies in stream sediment geochemical data. Geochem. Explor. Environ. Anal. geochem2024-009.
Shahrestani, S., Mokhtari, A.R., Carranza, E.J.M., Hosseini-Dinani, H., 2019. Comparison of efficiency of techniques for delineating uni-element anomalies from stream sediment geochemical landscapes. J. Geochem. Explor. 197, 184–198.
Shahrestani, S., Mokhtari, A.R., Fatehi, M., 2020. The use of unmixing technique in stream sediment geochemical exploration. J. Geochem. Explor. 208, 106339.
Shahrestani, S., Sanislav, I., 2025. How does dimensionality influence outlier detection effectiveness in multivariate geochemical data? insights from LOF and IF methods. Earth Science Informatics 18 (1), 27.
Sklar, M., 1959. Fonctions de répartition à n dimensions et leurs marges. Annales de l'Institut de Statistique de l'Université de Paris 8 (3), 229–231.
TaleFazel, E., Mehrabi, B., GhasemiSiani, M., 2019. Epithermal systems of the Torud–Chah Shirin district, northern Iran: Ore-fluid evolution and geodynamic setting. Ore Geol. Rev. 109, 253–275.
Torshizian, H., Afzal, P., Rahbar, K., Yasrebi, A.B., Wetherelt, A., Fyzollahhi, N., 2021. Application of modified wavelet and fractal modeling for detection of geochemical anomaly. Geochemistry 81 (4), 125800.
Wang, J., Zuo, R., 2022. Model averaging for identification of geochemical anomalies linked to mineralization. Ore Geol. Rev. 146, 104955.
Wang, J., Zuo, R., Liu, Q., 2024. Mapping Geochemical Anomalies by Accounting for the Uncertainty of mineralization-related Elemental Associations. EGUsphere, pp. 1–24, 2024.
Xiong, Y., Zuo, R., 2016. Recognition of geochemical anomalies using a deep autoencoder network. Comput. Geosci. 86, 75–82.
Xiong, Y., Zuo, R., Carranza, E.J.M., 2018. Mapping mineral prospectivity through big data analytics and a deep learning algorithm. Ore Geol. Rev. 102, 811–817.
Zhang, C., Zuo, R., 2021. Recognition of multivariate geochemical anomalies associated with mineralization using an improved generative adversarial network. Ore Geol. Rev. 136, 104264.
Zhao, Y., Nasrullah, Z., Li, Z., 2019. Pyod: A python toolbox for scalable outlier detection. Journal of machine learning research 20 (96), 1–7.
Zhao, B., Zhang, D., Tang, P., Luo, X., Wan, H., An, L., 2023. Recognition of multivariate geochemical anomalies using a geologically-constrained variational autoencoder network with spectrum separable module–A case study in Shangluo District, China. Appl. Geochem. 156, 105765.
Zuo, R., Xu, Y., 2024. A physically constrained hybrid deep learning model to mine a geochemical data cube in support of mineral exploration. Comput. Geosci. 182, 105490.