ASTROBIOLOGY
Volume 24, Number 11, 2024
© Mary Ann Liebert, Inc.
DOI: 10.1089/ast.2024.0019

A Machine-Learning Approach to Biosignature Exploration on Early Earth and Mars Using Sulfur Isotope and Trace Element Data in Pyrite

Maria C. Figueroa,[1] Daniel D. Gregory,[2] Kenneth H. Williford,[3] David J. Fike,[4] and Timothy W. Lyons[1]

Abstract

We propose a novel approach to identify the origin of pyrite grains and distinguish biologically influenced sedimentary pyrite using combined *in situ* sulfur isotope (δ^{34}S) and trace element (TE) analyses. To classify and predict the origin of individual pyrite grains, we applied multiple machine-learning algorithms to coupled δ^{34}S and TE data from pyrite grains formed from diverse sedimentary, hydrothermal, and metasomatic processes across geologic time. Our unsupervised classification algorithm, K-means++ cluster analysis, yielded six classes based on the formation environment of the pyrite: sedimentary, low temperature hydrothermal, medium temperature, polymetallic hydrothermal, high temperature, and large euhedral. We tested three supervised models (random forest [RF], Naïve Bayes, k-nearest neighbors), and RF outperformed the others in predicting pyrite formation type, achieving a precision (area under the ROC curve) of 0.979 ± 0.005 and an overall average class accuracy of 0.878 ± 0.005. Moreover, we found that coupling TE and δ^{34}S data significantly improved the performance of the RF model compared with using either TE or δ^{34}S data alone. Our data provide a novel framework for exploring sedimentary rocks that have undergone multiple hydrothermal, magmatic, and metamorphic alterations. Most significant, however, is the demonstrated potential for distinguishing between biogenic and abiotic pyrite in samples from early Earth. This approach could also be applied to the search for potential biosignatures in samples returned from Mars. Key Words: pyrite—biosignatures—machine learning—trace elements—sulfur isotopes. Astrobiology 24, 1110–1127.

1. Introduction

Current methods for identifying signatures of past life are complicated by the rarity of unambiguous fossils and difficulties associated with preservation of ancient organic materials (*e.g.*, Brasier et al., 2005). However, minerals that form through direct or indirect biological processes, while incorporating trace elements (TEs) in concentrations that are proportional to those in the surrounding fluid, can provide details about past environmental conditions and about the life present at the time of formation. In addition, advances in spatially resolved geochemical techniques for analyzing discrete mineral grains can significantly minimize the overprinting effects potentially present in bulk-sample techniques.

Pyrite (FeS_2), the most common metal sulfide mineral at Earth's surface (Rickard and Luther, 2007), can form through multiple pathways. These include abiotic chemical precipitation within magmatic and volcanic/hydrothermal systems as well as through biologically mediated pathways in sedimentary systems. In this context, biogenic pyrite is defined as containing sulfur that experienced recent and local metabolic processing, in contrast to abiotic pyrite that incorporates sulfur without contemporaneous metabolic involvement.

---

[1]Earth and Planetary Sciences, University of California, Riverside, California, USA.
[2]Department of Earth Sciences, University of Toronto, Toronto, Canada.
[3]Blue Marble Space Institute of Science, Seattle, Washington, USA.
[4]Earth and Planetary Sciences, Washington University, St. Louis, Missouri, USA.

A portion of this paper was previously published as the dissertation for Maria Cristina Figueroa submitted as part of the requirements for the degree of Doctor of Philosophy in Earth and Planetary Sciences, June 2023. A copy of the dissertation can be found at https://escholarship.org/content/qt9wh1v37s/qt9wh1v37s.pdf.

Typically, pyrite formation in sedimentary deposits initiates with sulfur-metabolizing microorganisms. Generally, pyrite forms when reactive iron reacts with hydrogen sulfide (H_2S) produced primarily from microbial sulfate reduction (MSR). In this process, sulfate-reducing bacteria reduce sulfate (SO_4^{2-}) to H_2S. This biological process can be represented by the general equation: $SO_4^{2-} + 2CH_2O \rightarrow H_2S + 2HCO_3^-$ (Canfield, 2001). Once H_2S is produced, it can react with iron (Fe^{2+}) in the environment to form pyrite. The simplified pyrite formation can be represented by the general equation: $Fe^{2+} + 2H_2S \rightarrow FeS_2 + 4H^+$ (Berner, 1984; Canfield et al., 1992; Rickard, 2012; Rickard and Luther, 2007).

In addition, other sulfide-producing microorganisms can also contribute to pyrite formation. Sulfur-disproportionating bacteria, for example, can produce H_2S from intermediate sulfur species, such as elemental sulfur (S^0), which can contribute to the formation of pyrite (Picard et al., 2016; Thamdrup et al., 1993; Fike et al., 2015). The metabolic processes of sulfur-metabolizing microbes result in large shifts toward lighter sulfur isotopes (^{32}S relative to ^{34}S), resulting in comparatively low $\delta^{34}S_{py}$ values that are often considered indicative of life (Canfield and Teske, 1996; Canfield and Thamdrup, 1994; Fike et al., 2015; Jørgensen et al., 2019).

Although biogenic processes are the primary contributors to sedimentary pyrite formation, abiotic processes, such as thermochemical sulfate reduction (TSR), can also occur during late-stage diagenesis via sulfate reduction linked to organic matter and hydrothermal heating (>110°C) (Jiang et al., 2018, 2015; Machel, 2001; Machel et al., 1995). Pyrite formed via TSR generally results in ^{34}S enrichment relative to microbially mediated pyrite (Kiyosu and Krouse, 1990; Machel, 2001; Machel et al., 1995). However, distinguishing biogenicity solely based on $\delta^{34}S$ values becomes challenging due to overlapping signatures influenced by various factors such as the source and availability of sulfur, system temperature, and disproportionation mechanisms (*e.g.*, Eldridge et al., 2016; Marini et al., 2011; Tino et al., 2023).

As pyrite forms, it incorporates a wide variety of TE in concentrations that can be characteristic of its environment. Recent work using *in situ* micro- and nanoscale measurements from sedimentary pyrite deposited on the seafloor has shown that Mo and Ni content in pyrite is linked to changes in local and global oxygen availability, whereas Co, Cu, and Zn are associated with variations in nutrient availability and productivity (Johnson et al., 2015; Large et al., 2014; Mukherjee and Large, 2020; Sweere et al., 2023). Thus, TE concentrations in pyrite serve as an indicator of paleoenvironmental conditions and, in particular, allow for the distinction between low- and high-temperature origins (Gregory et al., 2019b). Our results and related statistical modeling suggest that this combination of $\delta^{34}S$ and TE content could enable the recognition of pyrite as a biosignature on early Earth and Mars.

1.1. Sulfur isotopes in pyrite

Through geologic time, the burial of oxidized and reduced S species in marine sediments has been closely tied to the redox state of Earth's atmosphere. For instance, the lack of oxygen in the Archean atmosphere led to low sulfate concentrations in the oceans and a small $^{34}S/^{32}S$ fractionation effect in the resulting pyrite, relative to the fractionations after the Great Oxidation Event ~2.3–2.4 billion years ago (*e.g.*, Habicht et al., 2002; Crowe et al., 2014). Since isotopic signatures of S metabolisms are among the earliest evidence of life (ca. 3.45Ga; *e.g.*, Bontognali et al., 2012; Ohmoto et al., 1993; Shen et al., 2001), identifying S isotope fractionations unambiguously associated with microbial cycling is of immense interest in studies of early life.

Pyrite S isotopes have been used to reconstruct the early evolution of sulfur metabolisms and ocean redox chemistry (Canfield and Teske, 1996; Fike et al., 2015; Ohmoto et al., 1993; Shen et al., 2009; Johnson et al., 2015). However, these isotope records alone are not conclusive biosignatures, as large ranges in $\delta^{34}S_{py}$, including low $\delta^{34}S_{py}$ values, can be explained by multiple nonbiological (abiotic) processes (Wernette et al., 2018).

Biogenic S isotope fractionations in sedimentary pyrite are affected by multiple factors, including the source of sulfur, the rate and extent of MSR, and the formation and preservation of pyrite. These factors can track evolving redox chemistry of the oceans and the life within those waters (Leavitt et al., 2013; Sim et al., 2019, 2011). Moreover, recent studies are in agreement that sedimentary pyrite formation is affected by global and local processes (Atienza et al., 2023; Fike et al., 2015; Gomes et al., 2018; Gomes and Hurtgen, 2015; Gorjan et al., 2012; Pasquier et al., 2021a, 2021b). Global processes include the evolution of seawater sulfate concentrations and isotopic composition, which reflect the input and output of sulfur over long timescales related to weathering, volcanism, burial, and subduction (Jørgensen et al., 2019; Zhu et al., 2021). Local processes include the availability of organic matter, iron, and oxygen, which affect the biological community and the associated fractionations. Also important is the degree of basin restriction, which affects the sulfate inventory and the isotopic exchange (Fike et al., 2015; Pasquier et al., 2021a, 2021b). These processes, although informative about the history of the sedimentary environment, are major drivers of variability in the $\delta^{34}S_{py}$ record.

Pyrite that reflects purely abiotic processes typically has $\delta^{34}S$ values characteristic of its formation processes. For example, pyrite that forms from magmatic-hydrothermal or metamorphic processes typically has $\delta^{34}S$ values close to the ~0 ‰ typical of mantle-derived sulfur (Hutchison et al., 2020; Labidi et al., 2015). However, in certain hydrothermal systems, the formation of pyrite can involve the reaction of sulfate minerals (*i.e.*, anhydrite or barite) with hydrothermal fluids, which can result in significant S isotope fractionation between the reactants and products. The extent of this fractionation depends on the temperature, pressure, redox conditions, and the kinetics of the reactions. The typical $\delta^{34}S$ values for pyrite formed through these processes can range from approximately −10‰ to +10‰, depending on the fractionation effect and the initial $\delta^{34}S$ of the sulfate (Hutchison et al., 2020; Labidi et al., 2015; LaFlamme et al., 2018).

Moreover, TSR can have fractionation effects similar to the generally interpreted $\delta^{34}S$ fractionation (>21‰) indicative of biological origin (Machel, 2001; Tino et al., 2023). The fractionation effect for TSR depends on the temperature (>150°C), the equilibrium state, and the kinetic effect of the reaction, which can cause the fractionation to be either higher or lower than the theoretical equilibrium value (Cai et al., 2003; Hu et al., 2022; Meshoulam et al., 2016). Consequently, TSR has the potential to produce pyrite with δ^{34}S values similar to biogenic pyrite; however, TE enrichment patterns and morphologies of TSR-derived pyrite differ from those of biogenic pyrite (Hu et al., 2022; Machel, 2001; Tino et al., 2023). For instance, TSR results in a more pronounced enrichment of TEs reflective of the composition of hydrothermal fluids, whereas MSR-related enrichment is generally less pronounced and more selective (*e.g.*, Sun et al., 2024). In addition, the higher temperature in TSR leads to larger and well-defined crystals compared with the frequently observed framboidal morphology in microbially mediated pyrite, reflecting biological processes and organic matter interactions (*e.g.*, Liu et al., 2019).

1.2. Pyrite trace elements

During formation, pyrite incorporates a variety of TE in concentrations that can be characteristic of its environment. TE sources are variable and can come from the water column, hydrothermal fluids, and detrital material, including metals adsorbed onto Fe(hydr)oxides and organic matter. For sedimentary pyrite, the TE content is expected to scale with the concentrations, and thus bioavailability, of dissolved TEs within the surrounding water column or pore fluids. Redox-sensitive TEs such as Mo, Cu, and Zn are useful in revealing redox conditions in the water column and during diagenesis, and varying concentrations over long timescales can align with marine inventories and thus prevailing redox conditions in the biosphere (Gregory, 2020; Gregory et al., 2017; Large et al., 2017, 2014). More specifically, a combination of TEs (U, Mo, Cr, and Co) can be used to determine whether the overlying water column was oxic, suboxic, anoxic, or euxinic (Algeo and Lyons, 2006; Anbar et al., 2007; Gregory et al., 2015; Lyons et al., 2003; Morford et al., 2001; Tribovillard et al., 2006). An advantage of using TE concentrations lies in their strong retention in pyrite during subsequent hydrothermal and metamorphic events—up to greenschist facies metamorphism (Genna and Gaboury, 2015; George et al., 2018; Large et al., 2007).

There can be disagreement about the role of hydrothermal fluids for syngenetic metal enrichments in the rock record; however, established patterns, including those presented in the present study, help us make that distinction. For example, pyrite formed in high-temperature hydrothermal environments (*i.e.*, iron oxide copper-gold [IOCG] deposits) contains greater amounts of Co, Ni, Te, and Se and lesser amounts of Tl, Mn, Sb, and Hg compared with pyrite formed in lower-temperature environments (*i.e.*, sedimentary deposits; Reich et al., 2013; Keith et al., 2016; Sykora et al., 2018; Gregory et al., 2019b). There is, nonetheless, significant overlap in TE concentrations between biologically mediated sedimentary pyrite and abiotically precipitated forms, which elevates the relevance of this study.

1.3 Machine learning

Recent advances in machine learning (ML) have the potential to advance our understanding of complex geochemical patterns and the processes they reflect. ML algorithms can improve pattern identification from large and diverse datasets, improving interpretations, predictions, and decision making. This value is elevated by the significant overlap that exists for $\delta^{34}S_{py}$ and TE data from different pyrite formation types, specifically between sedimentary and hydrothermal processes. These relationships are not easily resolved by traditional statistical techniques based on empirical methods that subjectively group data. An important distinction from commonly used statistical models is that ML models can find patterns in data without requiring assumptions, which offers greater prediction accuracy for poorly understood associations. ML techniques have been used successfully in previous work to predict ore deposit type (Gregory et al., 2019b). Building on these datasets and associated successes, we couple new and previously published data of TE and $\delta^{34}S_{py}$ from pyrite grains from diverse sedimentary and hydrothermal deposits across geologic time. Beyond value added to the TE approach, we specifically explore whether the inclusion of S isotope data improves our ability to identify pyrite origins with confidence.

2. Methods

Data preprocessing was primarily executed in Python, with random forest (RF) classifier training and prediction evaluation conducted in the open-source data mining software platform Orange (Demšar et al., 2013). This project arose from multiple published pyrite analyses, including from the Geological Survey of Western Australia (Belousov et al., 2015) and the Geological Survey of South Australia (Gregory et al., 2015). Additional data from various ore deposits have been generated, leading to the current dataset of 865 pyrite analyses. The data are represented in Figure 1. Details follow, and a thorough statistical summary is available in the Supplementary Data S1.

2.1. Trace elements

LA-ICP-MS data are from a number of different sources and represent those in previously published peer-reviewed articles (Gilbert et al., 2014; Gregory et al., 2019a, 2019b, 2015). For the present and previous studies, LA-ICP-MS facilities at the University of Tasmania, Australia, and the University of Toronto, Canada, were used to analyze pyrite TE concentrations. Both facilities followed the same protocol with spot size and standards. Detailed analytical procedures are available in the work of Gregory et al. (2019a). All samples were analyzed for Co, Ni, Cu, Zn, As, Mo, Ag, Sb, Te, Au, Tl, and Pb, and these are the elements emphasized here. Beam size varied from 10 to 100 μm, depending on the size of the pyrite analyzed and the goals of the primary study. For each analysis, background was measured for 30 s before laser ablation periods of 40 to 60 s. The analyses were conducted in a pure He atmosphere, and Ar was added to the gas stream before injection into the ICP-MS to improve aerosol transport. Standards were analyzed at the start and end of each sample change and approximately every 25 analyses in between. The standards STDGL3 (Belousov et al., 2015) and GSD-1G (Guillong et al., 2005) were used. The software package SILLS (Guillong et al., 2008) was used for data reduction and concentration calculations.

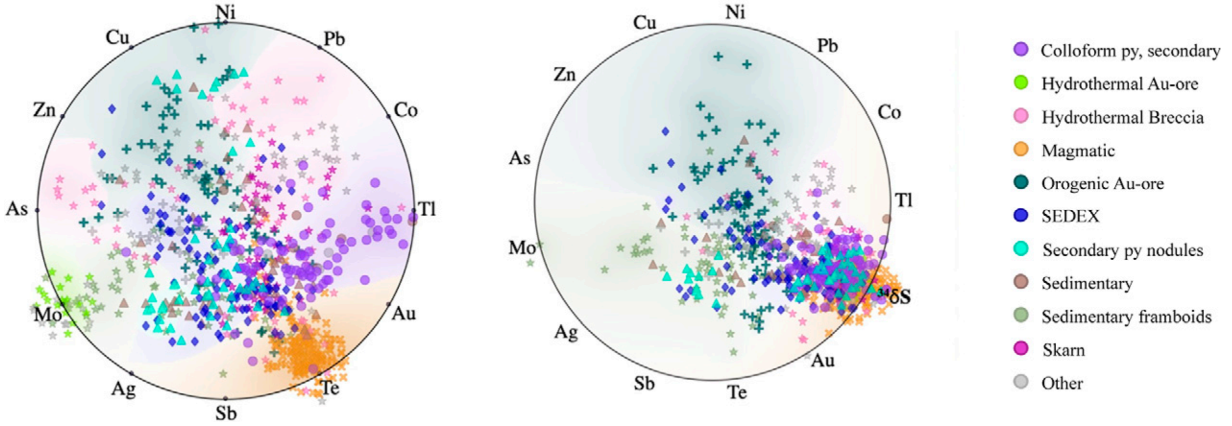


FIG. 1. Radial visualization of the dataset. Left: Distribution of the dataset across all 12 trace elements (TEs) delineated in terms of pyrite deposit type. All samples used in this study are included. Right: Distribution of the dataset across all 12 TE and δ^{34}S data points. Each axis represents one of the 12 TEs. Data points are plotted based on their values for these elements, with their positions reflecting a balance of all TE concentrations. Proximity of a point to an axis indicates a stronger influence of the corresponding TE.

2.2. Sulfur isotope analysis

In situ S isotope analyses were completed using the sensitive high-resolution microprobe (SHRIMP SI) at the Australian National University, with a spot size of 20 μ and 5–13 analyses conducted as a linear transect across the samples. The instrumental uncertainties for the SHRIMP SI measurements were below 0.3‰ (2σ). The detailed SHRIMP SI analytical procedures, including the standards used, are available in the work of Gregory et al. (2019b). SIMS measurements were collected on the Cameca IMS 7f-GEO at Washington University. Replicate analyses were conducted on an in house standard ($\delta^{34}S = 0.13 \pm 0.30$‰; Meyer et al., 2017) and a Balmat standard ($\delta^{34}S = 15.1$‰). The 2σ standard deviation was 0.3‰. The isotope values are expressed in standard delta notation (δ^{34}S) as per mil (‰) deviation from the Vienna-Canyon Diablo Troilite reference standard.

2.3. Data processing

To mitigate the influence of TEs from micro-inclusions of other minerals within the pyrite that might be included during the ablation, the data were screened to ensure that no analyses had > 1% Zn, 2% As, 1% Cu, 1% Ni, and 2% Co, as suggested by Gregory et al. (2015). For measurements that were below the detection limit, we imputed half the detection limit for those elements. This approach acknowledges the presence of the element at a low concentration while avoiding introducing artificial zeros that could skew the data. Zeros and null values represent true non-detections by the instrument and were left as null in the dataset.

Pyrite groups were defined by data that correspond to the same deposit type (*e.g.*, carbonaceous shales) as classified through previous studies of each. In instances where > 60% of values for a specific element within a pyrite group were below detection limit or missing, we excluded that group from the master dataset. This cutoff was selected to maintain a balance between preserving a substantial sample size and ensuring the integrity of data quality. Outliers, identified by using the local outlier factor method, a Euclidean distance technique (Alghushairy et al., 2021), were also systematically removed to enhance the robustness of the dataset for subsequent analyses.

2.4. Data clustering

In the present study, we used data clustering to categorize data points with similar attributes and determine the optimal number of classes for ML models. To accomplish this, we first adjusted the data by normalizing the features (the 12 TEs of interest and $\delta^{34}S_{py}$) to have values between 0 and 1. This allows for an even evaluation and comparison of data values per feature and improves the clustering analysis outcomes. We then applied the K-means++ clustering algorithm (Arthur and Vassilvitskii, 2007) to the normalized data using the software package scikit-learn in Python. K-means++ is an unsupervised ML technique for grouping (clustering) data based on the similarity of their features. This method groups data by initiating a randomized smart centroid and selecting the rest of the centroids based on the maximum square distances.

To determine the optimal number of clusters, we used the silhouette coefficient method (Rousseeuw, 1987), which measures the similarity of a data point to its own cluster compared with other clusters and ranges from −1 to 1. A higher silhouette coefficient indicates a better-defined cluster, whereas a negative coefficient indicates that the data point may be better placed in a different cluster.

2.5. Model selection

We used Orange, an open source ML and data visualization package in Python (Demšar et al., 2013), for the development of ML models. Three supervised machine learning algorithms were optimized for the dataset: RF, Naïve Bayes, and k-nearest neighbors (K-NN). These algorithms were selected given their capabilities to perform multivariable classifications. In addition, these algorithms have been used for similar Earth science applications (*e.g.*, Shelton et al., 2021; Lui et al., 2022).

The RF algorithm uses a collection of relatively uncorrelated decision trees in which each tree predicts the classification of an unknown sample. The class that most of the decision trees produce is ultimately the final prediction (*e.g*., Breiman, 2001). The naïve Bayes classification method is a simple probabilistic classifier that uses Bayes' theorem and assumes each of the features is independent and contributes equally to the model outcome (Dai et al., 2017; Zhang and Sakhanenko, 2019). It classifies new samples based on the maximum a posteriori decision rule, which assigns a data point to a class by the highest posterior probability from the feature values and assumes that features are independent from each other. The K-NN algorithm uses labeled input data, in this case by pyrite formation environment, and produces a distance matrix to calculate the distance between the training data and the new data (Campos et al., 2016; Cover and Hart, 1967). Any new data are classified based on the class of the nearest known data points. The code used in the present study is available at https://github.com/mcfig/pyrite-ML.

2.6. Model training and validation

A 70/30 split was made on the dataset to divide it into a training set and a validation set. The sampling split was made following the replicable (deterministic) method and the stratified method to ensure that they follow the compositional proportions of the original dataset. Confusion matrixes were produced to compare the performance of the model for each class. Overall model performance was determined by using statistical results for false positives and false negatives for each algorithm (*i.e*., accuracy, recall, and precision). In addition, the receiver operator characteristic (ROC) curve and the area under the ROC curve (AUC) (Mandrekar, 2010) were used as evaluation metrics. The ROC curve plots true positive rates against false-positive rates; subsequently, the AUC summarizes the ROC curve, representing the probability of correctly identifying pyrite types with a single value (0–1), where 1 indicates perfect discrimination.

3. Results

3.1. Cluster analysis

The K-means++ clustering algorithm was applied with the number of clusters varying from 2 to 10. The optimal number of clusters was determined to be five based on the silhouette coefficient method, as it gave the highest silhouette score (0.847). Cluster 1 (C1) contained a larger portion of the data points with 44.6%; C2 contained 20.5%, whereas C3, C4, and C5 contained 14.5%, 12.6%, 7.8%, respectively. To visualize the clusters, the data points are plotted in Figure 2 as a heatmap that highlights the distinctive feature values associated with each cluster. Observations within each cluster show that they differ slightly in their feature values, yet data points are effectively separated into temperature-dependent groups. For instance, C1 included all the sedimentary samples as well as pyrite from orogenic gold deposits (with some exceptions, see Section 3.2); C2 included all the homogenous large euhedral pyrite. C3 included the pyrite that formed in low-temperature hydrothermal fluid deposits (Carlin type and Irish-type deposits), whereas C4 included all the pyrite from polymetallic hydrothermal deposits, and C5 included all the pyrite from high-temperature skarn deposits.

3.2. Determining classes

With the exception of C1, all the clusters grouped the pyrite samples according to their temperature of formation and mineral deposit type. For example, C2 consisted of all homogenous large euhedral pyrite, whereas C3 contained pyrite samples formed from low-temperature hydrothermal fluids (Carlin type and Irish type deposits). C4 included pyrite samples from polymetallic hydrothermal deposits, and C5 consisted of pyrite from high-temperature skarn deposits. However, C1 included all the sedimentary samples, including modern marine framboids, pyrites from shales, and synsedimentary (pre-ore) pyrite samples from ore deposits, such as SEDEX and gold deposits, as well as pyrite data from orogenic gold deposits and SEDEX deposits that formed at higher temperatures (300–450°C; Tan et al., 2022; Zhang et al., 2022).

Based on these results, we divided the C1 cluster into two classes: (1) sedimentary and (2) medium temperature. This approach allowed us to define six distinct classes of pyrite from the five clusters: sedimentary and medium temperature (both from C1), low-temperature hydrothermal (C3), polymetallic hydrothermal (C4), high-temperature hydrothermal (C5), and large euhedral (C2). Table 1 summarizes the classification labels of the pyrite samples based on their temperature of formation and mineral deposit type. The proportion of each class was generally well balanced ($13-22\%$ of the total data), except for the high-temperature class, which comprised only 8% of the total dataset. Despite this observation, we retained the high-temperature pyrite deposits as a singular class due to their geochemical significance and relevance in understanding the chemical fingerprint of magmatic and metasomatic processes in pyrite. In addition, the decision to label one class as "large euhedral," diverging from the temperature-dependent labels assigned to other groups, is underpinned by the unique characteristics and formation conditions of these particular homogeneous crystals of approximately 1 cm^3 in size.

Large euhedral and homogeneous pyrite crystals are predominantly found in hydrothermal vein deposits typically under temperatures of approximately $200-300$°C, although this temperature range can vary depending on multiple factors such a geothermal gradient and fluid chemistry (Ridley, 2013). However, the key determinants for this pyrite crystal structure are the slow cooling rate and ample space provided by the vein systems for crystal growth. The slow growth process allows for the systematic exclusion of TE from the crystal structure during growth. By contrast, rapid crystal growth can trap more TE within the crystal lattice and/or nanoinclusions. As these crystals continue to grow, their lattices become increasingly ordered, further reducing the potential for TE incorporation. The inclusion of these analyses in our dataset is essential, as they provide an end member for pyrite depleted in TE that can potentially be found in other planetary bodies, thus contributing to our understanding of pyrite formation under a broad range of conditions.

Feature heterogeneity within each of the six defined classes was observed and is presented in Figure 2. In this context, heterogeneity refers to the variation in TE concentrations and δ^{34}S values within each of the six classes and indicates that the data points within a class exhibit a range of values. The polymetallic hydrothermal class displayed the highest overall

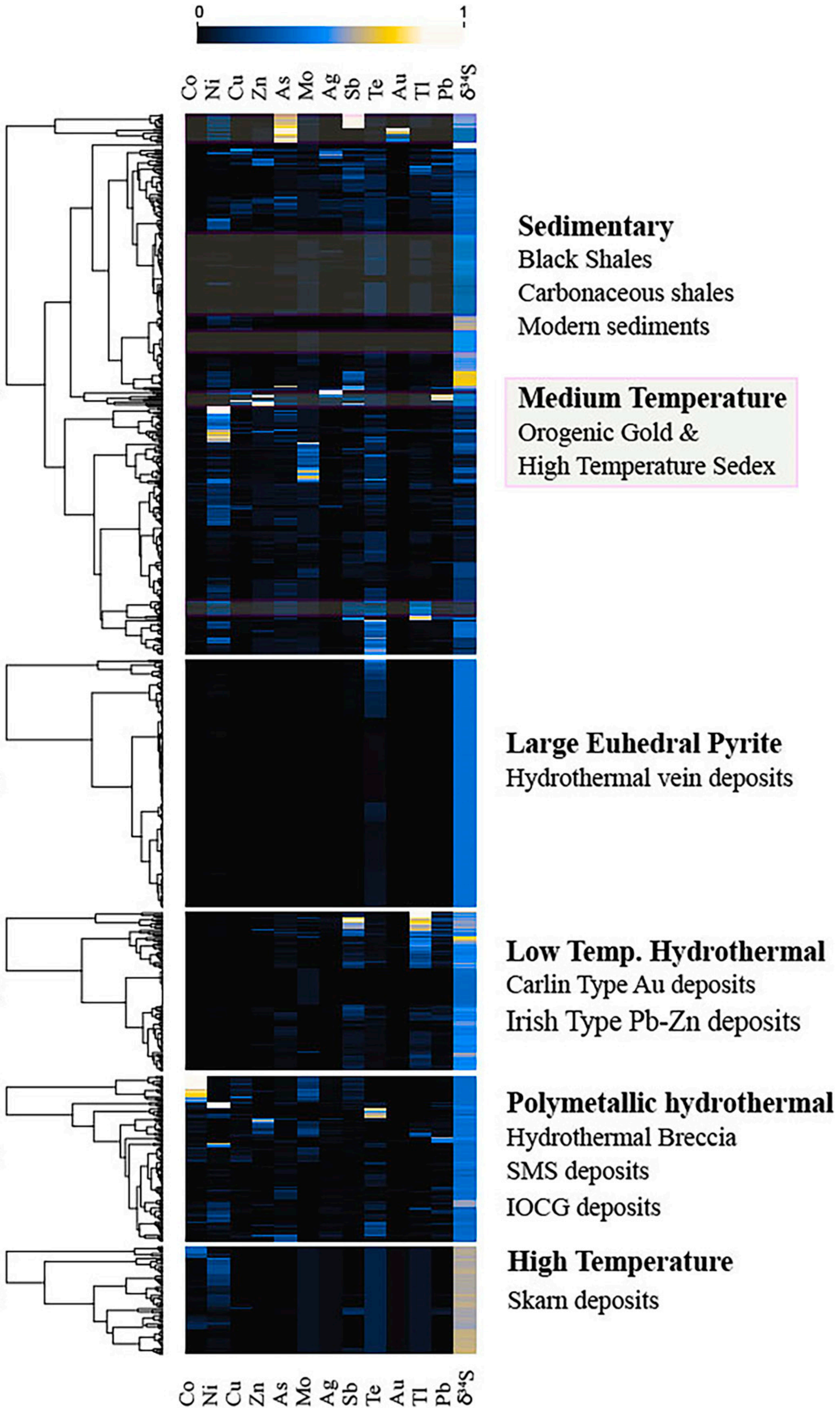


FIG. 2. Heatmap illustrating clusters. The heatmap represents the five clusters generated by K-means++, showcasing the variability in features shaded and normalized within the range of 0 and 1. Clusters are grouped, with descriptions provided on the right size. Notably, C1, the largest cluster, encompasses sedimentary pyrite and medium temperature pyrite (shaded and outlined in pink).

Table 1. Classification of Pyrite Samples by Temperature and Deposit Type

Class	*Deposit Types*	*Temperature Range (°C)*	*No. data Analysis*	*Studies*
Sedimentary	Black shales Carbonaceous shales Modern sediments	< 100	272	Gregory et al. (2019a) D. Li et al. (2019) J. Li et al. (2019) Zhang et al. (2022)
Low-temperature hydrothermal	Carlin-type gold Irish-type Pb-Zn	100 < *T* < 300	103	D. Li et al. (2019) J. Li et al. (2019)
Polymetallic hydrothermal	Hydrothermal Breccia SMS IOCG	250 < *T* < 450	119	This study Meng et al. (2020) Ding et al. (2021)
Medium temperature	Orogenic gold High-temperature SEDEX	300 ≤ *T* ≤ 450	177	Tan et al. (2022) Zhang et al. (2022)
High-temperature	Skarn	> 450	58	Ding et al. (2021)
Large euhedral	Hydrothermal veins	approx. 200 < *T* < 300	168	Gilbert et al. (2014)

heterogeneity among all features. The sedimentary class exhibited high heterogeneity in specific features, particularly in δ^{34}S, along with low concentrations of Co, Zn, and Au compared with the other classes. The medium-temperature class showed homogeneity in Mo and Co and heterogeneity in the remaining features. The low-temperature hydrothermal class had overall low concentrations in Co, Ni, Cu, Zn, and Au and showed heterogeneity in the remaining features. The high-temperature class displayed strong homogeneity in all features. Finally, the large euhedral class exhibited overall low concentrations in all TE, except for Te, and showed homogeneity in δ^{34}S values.

3.3. Distribution between pyrite classes

Box plots in figures 3 and 4 illustrate the distribution of several features (δ^{34}S and TEs) among pyrite classes. The δ^{34}S plot reveals that the sedimentary class has the largest distribution, with values ranging from −34.1‰ to +44.3‰, including both the lowest and highest values in the dataset. In contrast, large euhedral has the narrowest distribution, with an average value of +5.3 ± 0.1‰ (Fig. 3). The TEs box plots (Fig. 4) yielded several noteworthy observations. The sedimentary class, which is central to our investigation, exhibits a high degree of variability in all TE concentrations, as indicated by the extended whiskers in the box plot. Furthermore, the sedimentary pyrite and polymetallic hydrothermal pyrite classes demonstrate similarities in the concentrations of Cu, As, and Pb. This is evidenced by comparable median values (represented by the line inside the box) and elongated lower whiskers, indicating a degree of heterogeneity within these classes. In contrast, the large euhedral pyrite class emerges as the most distinctive, characterized by generally lower concentrations and a distribution skewed toward the upper quartile.

Figure 5 depicts feature associations among pyrite classes, revealing overall scattering and weak correlations. Specifically, the sedimentary class exhibits significant scattering that overlaps with the medium temperature, polymetallic hydrothermal, and low-temperature hydrothermal classes. In contrast, the high-temperature class exhibits lower scattering, associated with overall lower TE abundance and a strong differentiation by δ^{34}S (see Fig. 2), as indicated by the clustering of data points around the δ^{34}S axis. The large euhedral class demonstrates the least scattering and is strongly influenced by their Te abundance and δ^{34}S. Although variations in TE concentrations among the classes are discernible, the

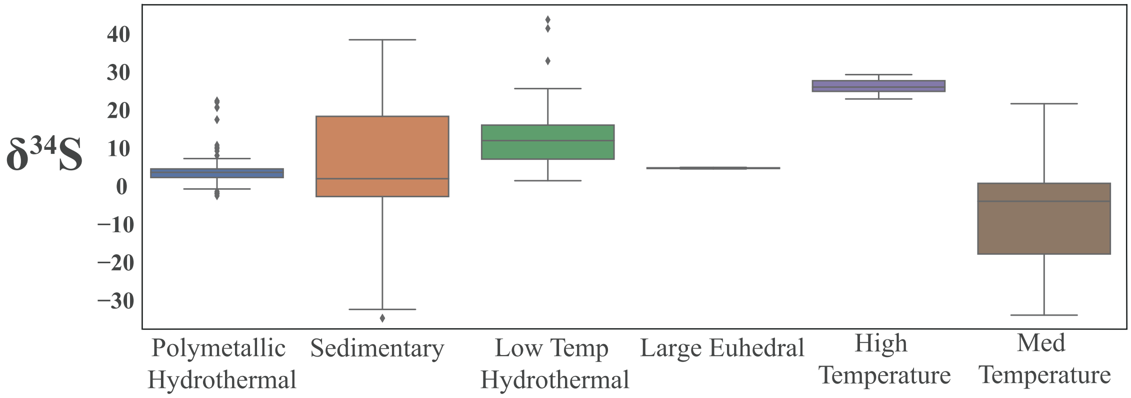


FIG. 3. Box plot showing the distribution of pyrite classes in relation to δ^{34}S values. The box plot captures the distribution of the six pyrite classes with respect to their δ^{34}S values. Each box represents the interquartile range (IQR) of δ^{34}S for a specific pyrite class, with the median indicated by the central line. Whiskers extend to the minimum and maximum values within 1.5 times the IQR. Outliers, represented as individual points, are displayed.

similarities in element concentrations make it challenging to identify clear trends or patterns solely based on their distribution. A pairwise scatterplot of the data, color coded by class, is provided in the supplementary information (Supplementary Fig. S2) and further emphasizes the scattering and complexity of the dataset. These observations underscore the intricate behavior of TEs in pyrite and the need for ML techniques in classifying pyrite formation types.

3.4. Model evaluation and selection

We evaluated three classification algorithms—RF, naïve Bayes, and K-NN—for their ability to predict pyrite formation type. Table 2 summarizes the performance of each algorithm in terms of average area under the ROC curve (AUC), class accuracy (CA), precision, and recall over 10-fold cross validation.

In the context of our multi-class classification evaluation, the AUC provides an aggregate measure of how well the model distinguishes between the various pyrite formation types. It represents the probability that the model correctly ranks instances from different classes. RF demonstrated superior performance with an average AUC of 0.979 ± 0.005 and an overall average CA of 0.878 ± 0.005.

To put this into perspective, RF outperformed the other models by exhibiting 10.9% higher accuracy than Naïve Bayes (CA = 0.792 ± 0.005) and 35.5% higher accuracy than K-NN (CA = 0.648 ± 0.008). In addition, higher precision (0.883 ± 0.005) and recall (0.878 ± 0.005) scores for each class indicate a better performance in classifying both positive and negative instances for RF when compared with the other models. Moreover, the low standard deviation of the RF model ($\sigma = 0.005$, $n = 10$) indicates good reproducibility and consistency in its performance. Based on these results, we selected RF as the algorithm of choice for the following stages of model development.

3.4.1. Evaluation of features. To test whether pyrite δ^{34}S values are useful features for pyrite type identification, we created three models for comparison. One model had only the 12 TEs as features (variables), another used only the δ^{34}S values, and the third coupled TE and $\delta^{34}S_{py.}$

In Figure 6, we present the results of cross-validation tests, with the training set comprising 70% of the data and the test set containing the remaining 30%, depicted as confusion matrixes. The model utilizing only the TEs as features achieved an overall CA of 0.943. In contrast, the model relying solely on δ^{34}S as the feature attained a lower overall CA of 0.724. However, when coupling TE with δ^{34}S, the model demonstrated remarkable improvement, achieving an overall CA of 0.991. Specifically for the sedimentary pyrite class, the TE + δ^{34}S model achieved perfect AUC, recall, and precision scores (1.000), whereas the TE model showed lower scores (AUC = 0.985, recall = 0.867, precision = 0.951), and the δ^{34}S-only model scored poorly (AUC = 0.803, recall = 0.587, precision = 0.628). Our results show that the addition of δ^{34}S to the TE dataset significantly enhances the classification performance, particularly for sedimentary pyrite, which is the pyrite formation type of interest in the search for biogenicity.

3.4. 2 RF prediction scores. To further analyze the performance of the RF classifier in distinguishing pyrite types, we present a focused examination of prediction scores from a blind test utilizing pyrite data from seafloor massive sulfide (SMS) not used during the model training phase (data from Meng et al., 2020); this effort is depicted in Figure 7. The model exhibited a robust ability to accurately classify euhedral pyrite (Py-II) from the interior wall and massive pyrite (Py-III) from the innermost interior wall. Both pyrite types were confidently assigned to the polymetallic hydrothermal class, garnering high prediction scores of 0.88 ± 0.09 and 0.97 ± 0.02, respectively.

However, the predictive capability diminished for anhedral pyrite (Py-I) from the outermost chimney zone, with a confidence score of 0.53 ± 0.19 for polymetallic hydrothermal. Interestingly, there was a subsequent uptick in the probability assigned to sedimentary pyrite, reaching 0.22 ± 0.06. Notably, one sample was predicted as sedimentary with a low probability score of 0.27, followed closely by polymetallic hydrothermal prediction with 0.25.

4. Discussion

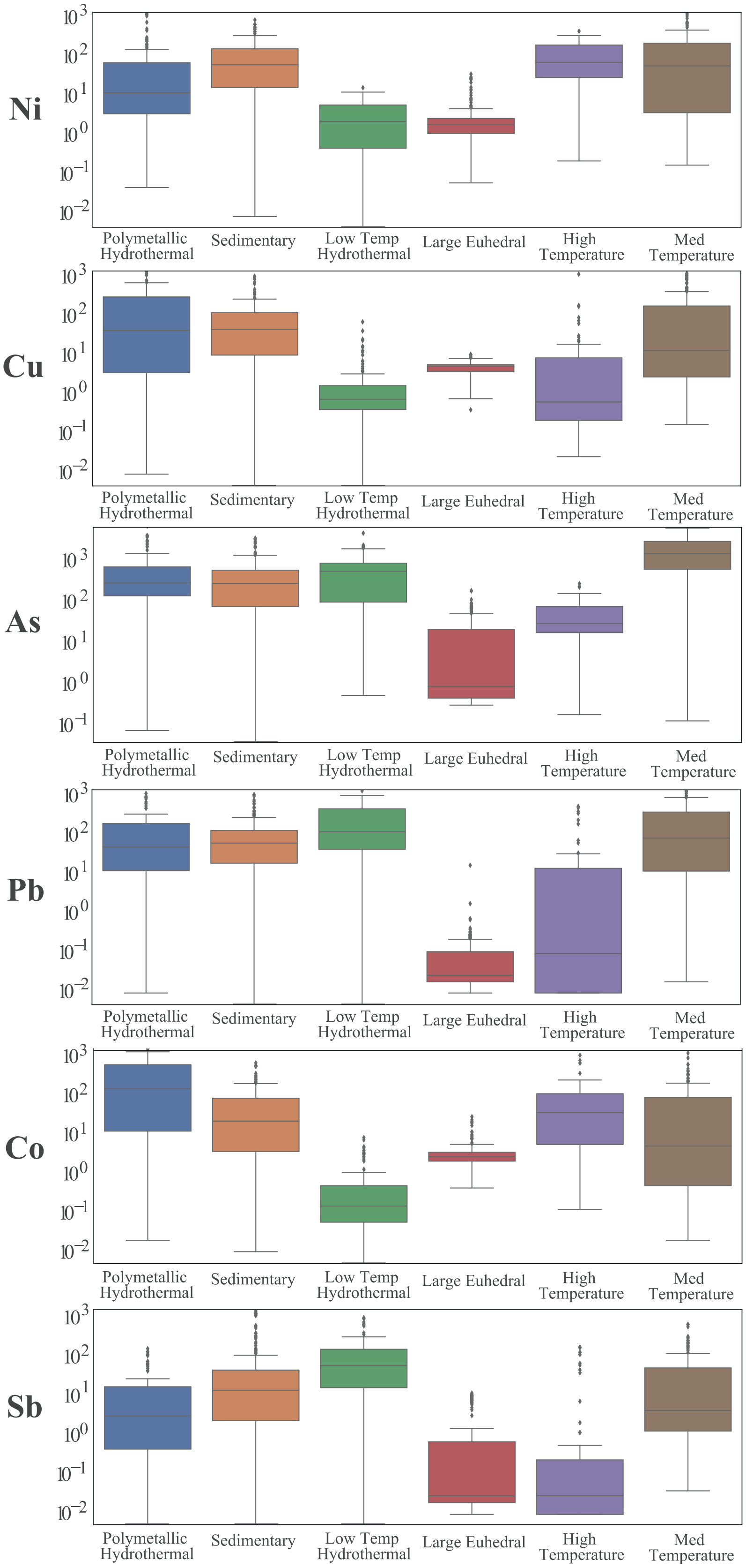
4.1. Machine learning for improved pyrite classification

The results of this study demonstrate the potential of ML to accurately classify pyrite types based on temperature of formation and deposit type by using TE and δ^{34}S data. Traditional statistical methods have inherent limitations due to the nonlinear relationship between geochemical variables. Our chosen method for determining classes, K-means++ clustering, avoids assumptions about data distribution and linear relationships. This method iteratively minimizes the distance between data points and centroid, which results in mechanically meaningful clusters compared with traditional methods, such as hierarchical clustering or principal component analysis. These traditional statistical methods assume that the relationships among the variables are linear. This assumption does not hold for our geochemical dataset, where the relationship among variables can be nonlinear. Pairwise scatterplots of the data that further reinforce this argument are available in the SI (Supplementary Fig. S1–S2). In contrast, our method can better capture the underlying structure of the data, resulting in groupings with elevated mechanistic meaning compared with traditional statistical methods.

The observations from the data distributions (Fig. 3–5) further highlight the greater degree of variability in TE concentrations, specifically in sedimentary environments. However, our use of cluster analysis facilitated the grouping of

→

FIG. 4. Box plot illustrating the distribution of pyrite classes in relation to trace elements. These box plots show the distributions of pyrite classes in terms of six trace elements—Ni, Cu, As, Pb, Co, and Sb in ppm. Each box represents the interquartile range (IQR) of trace element values for a specific pyrite class, with the median indicated by the central line. Whiskers extend to the minimum and maximum values within 1.5 times the IQR, and outliers are depicted as individual points.


Ni
Cu
As
Pb
Co
Sb
10^3
10^2
10^1
10^0
10^{-1}
10^{-2}
Polymetallic Hydrothermal
Sedimentary
Low Temp Hydrothermal
Large Euhedral
High Temperature
Med Temperature

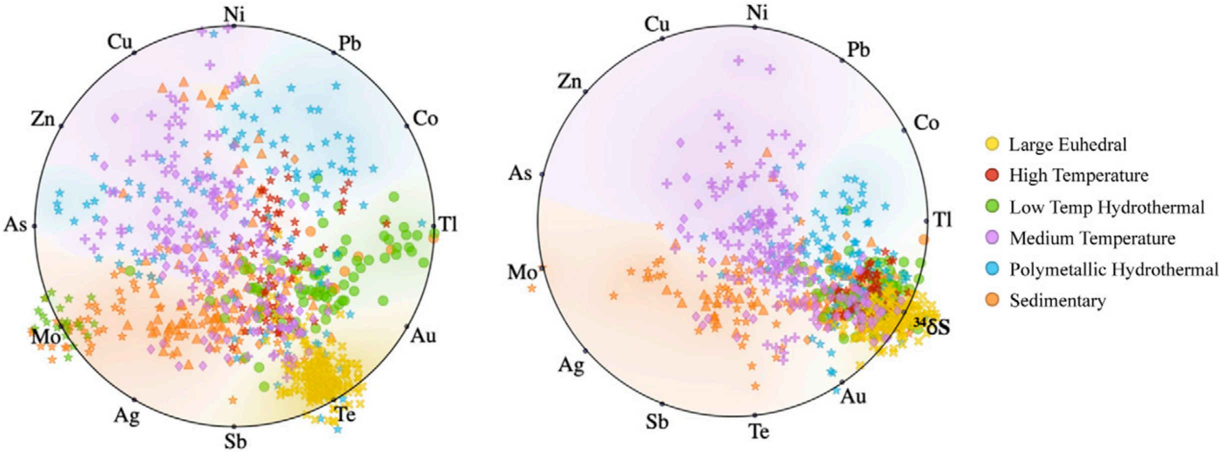


FIG. 5. Radial visualization by pyrite classes. Left: Distribution across data for all 12 trace elements (TEs). Right: Distribution across data for all 12 TEs and δ^{34}S. Data points are color-coded into pyrite label classes.

the pyrite data into six distinct classes that further provided insights into their geochemical variability and the relationships among multiple features.

4.2. Distinguishing diverse pyrite types

4.2.1. Cluster analysis. From the clustering results, we identified that grouping pyrite by temperature of formation is an effective way to create robust classes for predictive models. The subsequent classification method is beneficial for predictive models in differentiating sedimentary pyrite from hydrothermal/magmatic-sourced pyrite, as it can give important context about its origin and the geological setting in which it formed.

Traditional geochemical methods for deciphering ancient high- and low-temperature settings are typically a multistep process that includes a variety of techniques, including geochemical and isotopic analyses of pyrite and other minerals, host rocks, and fluids, and followed by implementing quantitative analysis for interpretation—particularly for large datasets across multiparameter space (*e.g.*, George et al., 2018; Gregory et al., 2019a, 2017; Johnson et al., 2015; Large et al., 2007; J. Li et al., 2019). Here, we provide a novel approach that requires fewer steps compared with traditional geochemical protocols. For example, our classifier was able to accurately distinguish the multiple generations of pyrite from a gold deposit, as described in the work of J. Li et al. (2019). It distinguished the syn-sedimentary pyrite data from the post-depositional pyrite and correctly identified them as sedimentary and low-temperature hydrothermal, respectively.

4.2.2. Random Forest. RF can accurately classify pyrite based on temperature of formation and mineral deposit type. The superior performance of the RF algorithm over Naïve Bayes and K-NN may be attributed to its superior ability to manage nonlinear relationships. Specifically, the construction of multiple decision trees using a random subset of features allows RF to capture nonlinear relationships and interactions among the various features, which can be difficult to capture using Naïve Bayes and K-NN. The combination of multiple decision trees also helps reduce the impact of outliers and noise on the final prediction. Another advantage of RF is that it is less prone to overfitting compared with Naïve Bayes and K-NN. This difference is because bootstrap aggregation (bagging) and random feature selection reduce variance in the model and prevent overfitting. Naïve Bayes, on the other contrary, can overfit if the features are strongly correlated, and K-NN is prone to overfitting if the number of neighbors is too small.

The initial classification of pyrite based only on TE data resulted in a misclassification rate of 11% for sedimentary pyrite as low-temperature hydrothermal and 2.2% as polymetallic hydrothermal (Fig. 3). However, when δ^{34}S data were added to the classification model, the accuracy for sedimentary pyrite improved significantly, which resulted in a perfect classification rate. These results further indicate that δ^{34}S values alone are not a strong biological signature; see the work of Moreras-Marti et al. (2022) for an exhaustive discussion about S isotopes as potential biosignatures for Mars exploration. Despite the limitations that S isotopes alone hold, the sulfur data significantly improve the detection of sedimentary pyrite when coupled with TE results.

Table 2. Model Evaluations. These Are the Average Values and Standard Deviation ($n = 10$) for the Evaluation Metrics for Each Model

Model	*AUC*	*CA*	*Precision*	*Recall*
Random forest	0.979 ± 0.005	0.878 ± 0.005	0.883 ± 0.005	0.878 ± 0.005
Naïve Bayes	0.964 ± 0.005	0.792 ± 0.005	0.814 ± 0.004	0.792 ± 0.003
K-NN	0.887 ± 0.008	0.648 ± 0.008	0.654 ± 0.010	0.648 ± 0.010

AUC is the area under the ROC curve, CA is class accuracy, precision quantifies the number of positive class predictions that belong to the positive class. Recall quantifies how many of the true positives in the dataset were classified correctly.

Our results highlight that combining δ^{34}S with TEs results in improved distinction of sedimentary pyrite from low-temperature hydrothermal and polymetallic depositional systems, which was a limitation in the model when using only TEs. Therefore, the incorporation of δ^{34}S data into the pyrite classification model and categorization of the classes by temperature of formation improves the model's accuracy, meaning it makes more correct predictions of pyrite type. In addition, this coupling improves the model's robustness, meaning the model is able to provide reliable predictions despite noise in data distribution.

4.3. Unraveling biological influences in pyrite

The results presented highlight the potential to accurately distinguish sedimentary pyrite from diverse thermally impacted types. Although the model shows promise in delineating pyrite formation types, a more extensive understanding of biogenic pyrite is necessary. It is also important to acknowledge that sedimentary pyrite alone is not exclusively a biosignature, as debates continue with regard to the biological and abiotic involvement in its mineralization process. Recent work on microbial involvement in the formation of pyrite spherules shows ambiguity in deciphering between biologically controlled or induced mineralization (Berg et al., 2020; Duverger et al., 2021; Truong et al., 2023). However, these findings emphasize that pyrite formation in the presence of microbial activity often results in distinct morphologies that are not replicated by abiotic processes. Therefore, while pyrite itself may not be a definitive biosignature, its formation pathway, coupled with detailed geochemical and isotopic analyses, can provide strong evidence for biogenic processes. This underscores the importance of using a multifaceted approach in biosignature identification, where pyrite's formation environment and associated signatures play a crucial role. Additionally, experimental studies on abiotic low-temperature sedimentary pyrite formation and TE incorporation (e.g., Baya et al., 2022, 2021; Le Pape et al., 2017; Morin et al., 2017) may elucidate the abiotic controls on TE incorporation and abundance, potentially serving as a valuable end member for abiotic sedimentary pyrite.

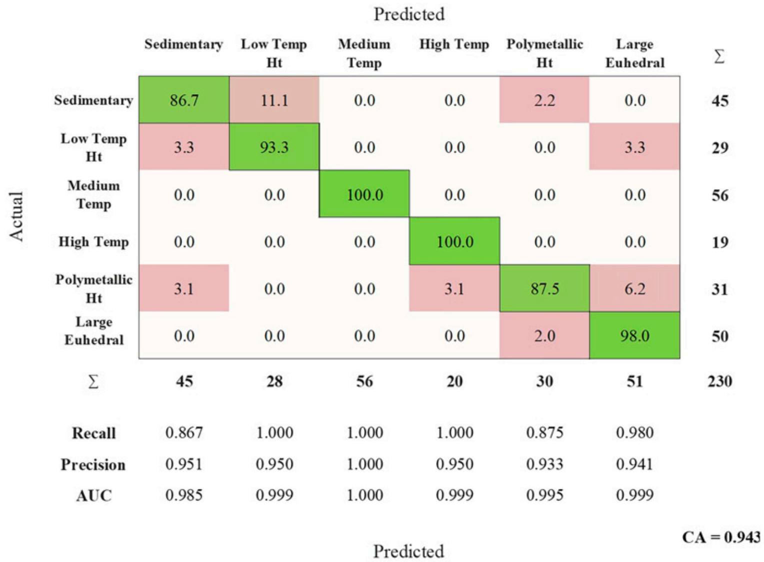
Further, the role of biology cannot be entirely ruled out for some low- to moderate-temperature hydrothermal deposits. Modern SMS deposits, for instance, are intricately linked to seawater and chemosynthesis-based influences. Our investigation reveals that the influence of biology on SMS samples can be inferred from the RF prediction scores, detailed in subsection 3.4.2 and depicted in Figure 7. The pyrite samples from the innermost parts of the chimney were classified as polymetallic hydrothermal with overall high confidence (predictive) scores. Further, the confidence score for pyrite from the outermost chimney zone diminished, accompanied by an increased probability of sedimentary origin. This shift in prediction scores for samples from the outermost zone suggests that the model can capture influences associated with sedimentary processes even within a hydrothermal system. These diverse influences, in this context, are primarily driven by exposure to seawater and microbially mediated sulfate reduction. The associated prediction scores highlight the importance of considering the complex interplay of processes within a given system and how they may map out spatially and temporally. Similarly, other low-temperature hydrothermal deposits such as VMS (the ancient analog of SMS deposits) and SEDEX systems can also exhibit biologically derived influences. During ore formation, "cooler" zones can lead to the precipitation of pyrite influenced by seawater and other biologically mediated processes. In addition, the model's capability to classify pyrite based on temperature-dependent conditions carries broader implications concerning geological processes, particularly in relation to tectonism, with focus on the polymetallic hydrothermal and medium temperature classes. For example, IOCG, and orogenic gold deposits are primarily associated with active tectonic zones such as orogenic belts and continental margins. Both deposit types require fracture and shear zones created by tectonic forces to allow for the movement of mineralizing fluids (Goldfarb et al., 2005; Ridley, 2013; Skirrow, 2022). The ability of our classification tool to identify pyrite based on temperature-dependent conditions thus offers a unique lens for unraveling the temporal dynamics of Martian tectonism and hydrothermal activity. This possibility not only contributes to constraining potential habitable periods on Mars but also aids in validating and interpreting other biosignatures on Earth and potential signs of life beyond.

4.4. Addressing dataset limitations

One limitation of this study is the relatively small size of the dataset used for training and analysis. More data should help to further validate the results and refine the approach. This consideration is particularly relevant for the large euhedral group, which currently exhibits strong association with Te, likely due to the data originating from a single pyrite specimen. The limited sample size introduces the risk of overfitting, whereby the model learns specific patterns instead of generalizable features. Thus, expanding the dataset to include a wider range of pyrite samples from diverse environments and deposit types is crucial for several reasons. First, it serves to mitigate the risk of overfitting by ensuring that the model is not tailored to localized characteristics, thereby fostering a more generalizable and robust model. Second, the addition of more data provides an opportunity to subdivide the sedimentary class into syndepositional pyrite and pyrite formed, altered, or overprints by abiotic processes. In addition, increasing the amount of hydrothermally

→

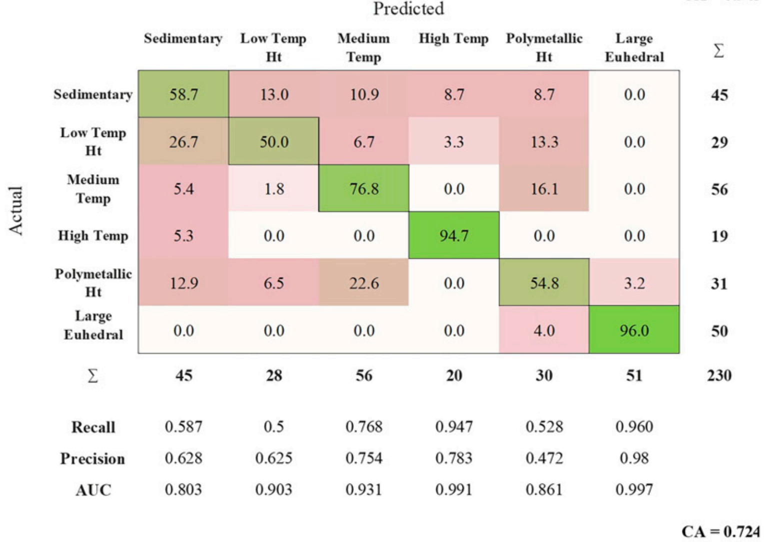
FIG. 6. Confusion matrices. *Top:* Confusion matrix for the model using only TEs. *Middle:* Confusion matrix for the model using only δ^{34}S. *Bottom:* Confusion matrix for the model using both TEs and δ^{34}S. The matrices display the percentage of instances, with the actual classes on the y-axis and the predicted classes on the x-axis. Color shades represent the percentage of predicted instances per class, ranging from 0% (white) to pink-green hues in between, and deep green representing 100%. The sum on the far right and bottom indicates the total instances per class. For each confusion matrix, recall, precision, and area under the ROC curve (AUC) scores are displayed along the bottom. Also shown are the overall classification accuracy (CA).



Predicted

Actual	Sedimentary	Low Temp Ht	Medium Temp	High Temp	Polymetallic Ht	Large Euhedral	Σ
Sedimentary	86.7	11.1	0.0	0.0	2.2	0.0	**45**
Low Temp Ht	3.3	93.3	0.0	0.0	0.0	3.3	**29**
Medium Temp	0.0	0.0	100.0	0.0	0.0	0.0	**56**
High Temp	0.0	0.0	0.0	100.0	0.0	0.0	**19**
Polymetallic Ht	3.1	0.0	0.0	3.1	87.5	6.2	**31**
Large Euhedral	0.0	0.0	0.0	0.0	2.0	98.0	**50**
Σ	**45**	**28**	**56**	**20**	**30**	**51**	**230**
Recall	0.867	1.000	1.000	1.000	0.875	0.980	
Precision	0.951	0.950	1.000	0.950	0.933	0.941	
AUC	0.985	0.999	1.000	0.999	0.995	0.999	

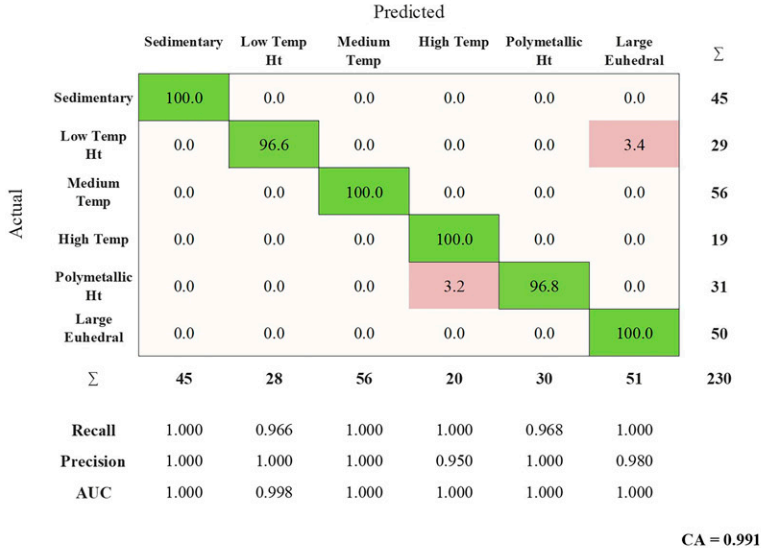
CA = 0.943



Predicted

Actual	Sedimentary	Low Temp Ht	Medium Temp	High Temp	Polymetallic Ht	Large Euhedral	Σ
Sedimentary	58.7	13.0	10.9	8.7	8.7	0.0	**45**
Low Temp Ht	26.7	50.0	6.7	3.3	13.3	0.0	**29**
Medium Temp	5.4	1.8	76.8	0.0	16.1	0.0	**56**
High Temp	5.3	0.0	0.0	94.7	0.0	0.0	**19**
Polymetallic Ht	12.9	6.5	22.6	0.0	54.8	3.2	**31**
Large Euhedral	0.0	0.0	0.0	0.0	4.0	96.0	**50**
Σ	**45**	**28**	**56**	**20**	**30**	**51**	**230**
Recall	0.587	0.5	0.768	0.947	0.528	0.960	
Precision	0.628	0.625	0.754	0.783	0.472	0.98	
AUC	0.803	0.903	0.931	0.991	0.861	0.997	

CA = 0.724



Predicted

Actual	Sedimentary	Low Temp Ht	Medium Temp	High Temp	Polymetallic Ht	Large Euhedral	Σ
Sedimentary	100.0	0.0	0.0	0.0	0.0	0.0	**45**
Low Temp Ht	0.0	96.6	0.0	0.0	0.0	3.4	**29**
Medium Temp	0.0	0.0	100.0	0.0	0.0	0.0	**56**
High Temp	0.0	0.0	0.0	100.0	0.0	0.0	**19**
Polymetallic Ht	0.0	0.0	0.0	3.2	96.8	0.0	**31**
Large Euhedral	0.0	0.0	0.0	0.0	0.0	100.0	**50**
Σ	**45**	**28**	**56**	**20**	**30**	**51**	**230**
Recall	1.000	0.966	1.000	1.000	0.968	1.000	
Precision	1.000	1.000	1.000	0.950	1.000	0.980	
AUC	1.000	0.998	1.000	1.000	1.000	1.000	

CA = 0.991

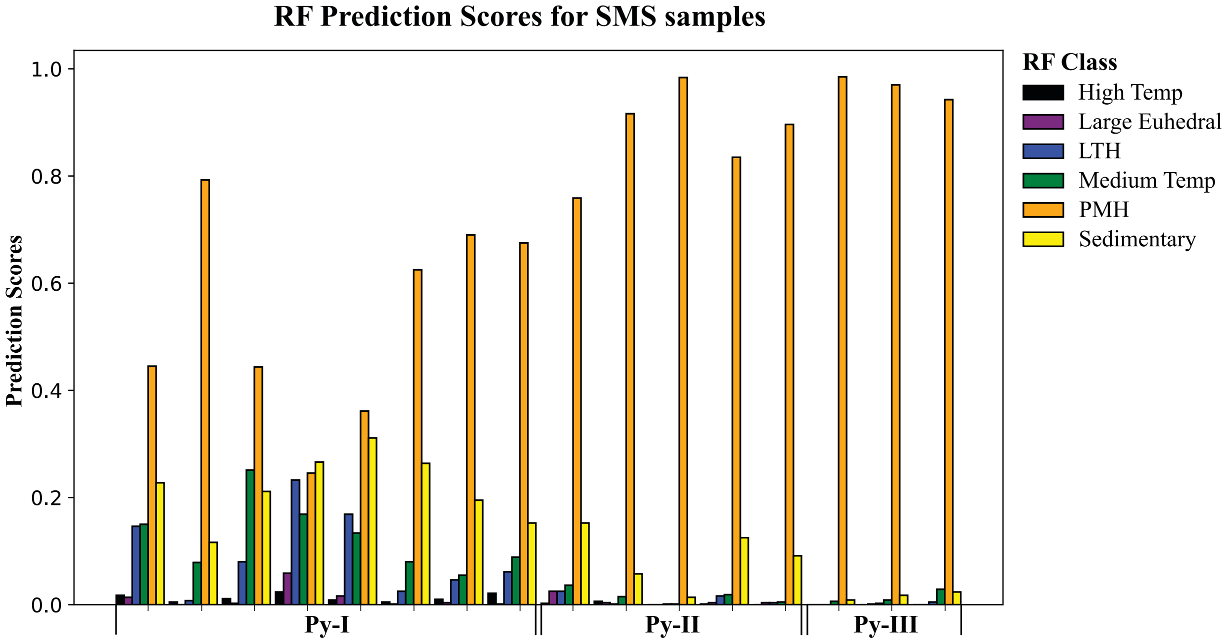


FIG. 7. Random forest (RF) prediction scores for SMS samples. The pyrite samples are from an active pyrite-dominated chimney from the East Pacific Rise hydrothermal field (data from Meng et al., 2020). Three pyrite types were defined: Py-I refers to anhedral pyrite from the outermost chimney zone; Py-II is euhedral pyrite from the interior wall; and Py-III massive pyrite from the innermost interior wall. LTH, low-temperature hydrothermal; PMH, polymetallic hydrothermal.

influenced pyrite would help in subdividing low-temperature and polymetallic hydrothermal classes and further refine the definition of a representative biogenic end member. Third, a more extensive dataset enhances the applicability of the classification tool, broadening its scope to include pyrite samples from a wider array of geologic contexts. These motivations, combined with our promising initial results, underscore the importance of additional studies that include *in situ* pyrite TE and δ^{34}S data.

5. Conclusions and Implications

This study leverages a combination of TE and $\delta^{34}S_{py}$ data to identify the origin of pyrite from various geological deposits. Using cluster analysis, we grouped pyrite samples based on their temperature of formation and mineral deposit type, resulting in six distinct classes: sedimentary, low temperature hydrothermal, medium temperature, polymetallic hydrothermal, high temperature, and large euhedral. Among the three classification algorithms evaluated—RF, Naïve Bayes, and K-NN—RF proved most effective in predicting pyrite formation type. The incorporation of δ^{34}S values as a feature for pyrite classification, when combined with TE data, yielded superior accuracy compared with using TE and δ^{34}S values alone.

The implications of this study are significant for determining the origin of pyrite in diverse geological settings, particularly between biogenic and abiotic formation. The six pyrite classes identified serve as a foundation for future research related to biogeochemical processes in pyrite. More broadly, our investigation highlights the value of using ML techniques alongside geochemical data to unravel complex geochemical systems in ways that can facilitate mineral exploration and biosignature identification.

The classification tool developed in this study has the potential to identify specific markers of biological activity, including those retained after alteration of pyrite. Such insight could guide exploration for signs of early life on Earth and potentially on Mars. Given the abundance of iron and sulfur in the Martian crust (Gellert et al., 2015; Toulmin et al., 1977), their interactions and distributions are closely linked to the formation and weathering of Fe-S minerals.

Pyrite has been identified in certain Martian meteorites, such as the NWA 7533 impact breccia, where pyrite is the major sulfur host (Lorand et al., 2020). Although detection of pyrite on Mars remains limited, rover-based instrumentation has observed other Fe-S minerals, such as pyrrhotite (Vaniman et al., 2014; MacArthur et al., 2019). In addition, sulfates, which can be a product of oxidative weathering of pyrite, have been detected through remote sensing and direct observation at various locations on the Martian surface (Baccolo et al., 2021; Ehlmann and Edwards, 2014; Squyres et al., 2004). These observations suggest that Martian pyrite deposits may indeed exist, particularly under the reducing conditions below the oxidized surface layers.

The geochemical history of Mars has been significantly influenced by its sulfur cycle, analogous to the carbon cycle dominating surficial processes on Earth (Franz et al., 2017; Gaillard et al., 2013; Halevy et al., 2007). This raises important considerations regarding the potential differences in the sulfur cycle systems between Mars and Earth that could have implications for the S isotope signatures of Martian pyrite. Although it is essential to recognize and investigate these variations, our approach, which leverages terrestrial pyrite S isotopes and TE as biosignatures, can still provide valuable information. Terrestrial studies provide a robust framework for identifying biogenic influences through well-established S isotope fractionation patterns and TE distributions. By applying this framework to Martian samples, we can develop a baseline understanding of potential biosignatures.

Future missions, particularly those with sample return as a primary focus, offer promising possibilities in the search for evidence of past life on Mars. Future studies should further investigate how the unique conditions and sulfur cycle dynamics of early Mars might confound the results when applying interpretations to Martian samples. This would enhance our ability to accurately identify and understand sulfur-related biosignatures on Mars while ensuring our methodology remains adaptable and relevant across different planetary contexts.

Acknowledgments

Special thanks to Dr. Bahram Mobasher (UCR) and the NASA MIRO FIELDS Fellowship for their generous support and encouragement. The authors also thank Dr. Clive Jones (WashU) for his invaluable assistances with SIMS analyses and Dr. Michael Tuite (JPL) for the generous assistance in sharing of laboratory space. Additional thanks to the NSERC Discovery grant provided to D.D.G. Funding at UCR was provided to T.W.L. by the NASA Astrobiology Institute under the Cooperative Agreement NNA15BB03A and the Interdisciplinary Consortia for Astrobiology Research (ICAR) Program issued through the Science Mission Directorate.

Authors' Contributions

M.C.F., D.D.G., and T.W.L. developed the project. M.C.F. and D.D.G. carried out the sample preparation and analyses with support from K.H.W. and D.J.F.; M.C.F. wrote the article with contributions from all the authors.

Author Disclosure Statement

The authors declare no conflicts of interests.

Funding Information

This work was supported by the NASA Astrobiology Institute provided to T.W.L. under Cooperative Agreement No. NNA15BB03A issued through the Science Mission Directorate and the NASA Interdisciplinary Consortia for Astrobiology Research (ICAR).

Supplementary Material

Supplementary Data S1
Supplementary Figure S1
Supplementary Figure S2

References

Algeo TJ, Lyons TW. Mo–Total organic carbon covariation in modern anoxic marine environments: Implications for analysis of paleoredox and paleohydrographic conditions. *Paleoceanography* 2006;21(1); doi: 10.1029/2004PA001112

Alghushairy O, Alsini R, Soule T, et al. A review of local outlier factor algorithms for outlier detection in big data streams. *BDCC* 2021;5(1):1; doi: 10.3390/bdcc5010001

Anbar AD, Duan Y, Lyons TW, et al. A Whiff of oxygen before the great oxidation event? *Science* 2007;317(5846):1903–1906; doi: 10.1126/science.1140325

Arthur D, Vassilvitskii S. K-Means++: The Advantages of Careful Seeding. Proc Eighteenth Annu ACM-SIAM Symp Discrete Algorithms Soc Ind Appl Math 2007;Philadelphia, PA, USA:1027–1035.

Atienza NMM, Gregory DD, Taylor SD, et al. Refined views of ancient ocean chemistry: Tracking trace element incorporation in pyrite framboids using atom probe tomography. *Geochim Cosmochim Acta* 2023;357:1–12; doi: 10.1016/j.gca.2023.07.013

Baccolo G, Delmonte B, Niles PB, et al. Jarosite formation in deep antarctic ice provides a window into acidic, water-limited weathering on mars. *Nat Commun* 2021;12(1):436; doi: 10.1038/s41467-020-20705-z

Baya C, Le Pape P, Baptiste B, et al. Influence of trace level as or ni on pyrite formation kinetics at low temperature. *Geochim Cosmochim Acta* 2021;300:333–353; doi: 10.1016/j.gca.2021.01.042

Baya C, Le Pape P, Baptiste B, et al. A methodological framework to study the behavior and kinetic influence of V, Mn, Co, Ni, Cu, Zn, As, Se and Mo during pyrite formation *via* the polysulfide pathway at ambient temperature. *Chem Geol* 2022;613:121139; doi: 10.1016/j.chemgeo.2022.121139

Belousov I, Danyushevsky L, Olin P, et al. STDGL3—a New Calibration Standard for Sulphide Analysis by LA-. Goldschmidt 2015.

Berg JS, Duverger A, Cordier L, et al. Rapid Pyritization in the presence of a sulfur/sulfate-reducing bacterial consortium. *Sci Rep* 2020;10(1):8264; doi: 10.1038/s41598-020-64990-6

Berner RA. Sedimentary pyrite formation: An update. *Geochim Cosmochim Acta* 1984;48(4):605–615; doi: 10.1016/0016-7037(84)90089-9

Bontognali TRR, Sessions AL, Allwood AC, et al. Sulfur isotopes of organic matter preserved in 3.45-billion-year-old stromatolites reveal microbial metabolism. *Proc Natl Acad Sci U S A* 2012;109(38):15146–15151; doi: 10.1073/pnas.1207491109

Brasier MD, Green OR, Lindsay JF, et al. Critical testing of earth's oldest putative fossil assemblage from the ∼3.5Ga apex chert, chinaman creek, western Australia. *Precambrian Res* 2005;140(1–2):55–102; doi: 10.1016/j.precamres.2005.06.008

Breiman L. Random Forests. *Mach Learn* 2001;45(1):5–32; doi: 10.1023/A:1010933404324

Cai C, Worden RH, Bottrell SH, et al. Thermochemical sulphate reduction and the generation of hydrogen sulphide and thiols (mercaptans) in triassic carbonate reservoirs from the sichuan basin, china. *Chem Geol* 2003;202(1–2):39–57; doi: 10.1016/S0009-2541(03)00209-2

Campos GO, Zimek A, Sander J, et al. On the evaluation of unsupervised outlier detection: Measures, datasets, and an empirical study. *Data Min Knowl Disc* 2016;30(4):891–927; doi: 10.1007/s10618-015-0444-8

Canfield DE. 12. biogeochemistry of sulfur isotopes. In: Stable Isotope Geochemistry. (Valley JW and Cole DR. eds) De Gruyter: Berlin, Boston; 2001; pp. 607–636; doi: 10.1515/9781501508745-015

Canfield DE, Raiswell R, Bottrell SH. The reactivity of sedimentary iron minerals toward sulfide. *Am J Sci* 1992;292(9):659–683; doi: 10.2475/ajs.292.9.659

Canfield DE, Teske A. Late proterozoic rise in atmospheric oxygen concentration inferred from phylogenetic and sulphur-isotope studies. *Nature* 1996;382(6587):127–132.

Canfield DE, Thamdrup B. The production of 34s-depleted sulfide during bacterial disproportionation of elemental sulfur. *Science* 1994;266(5193):1973–1975; doi: 10.1126/science.11540246

Cover T, Hart P. Nearest neighbor pattern classification. *IEEE Trans Inform Theory* 1967;13(1):21–27; doi: 10.1109/TIT.1967.1053964

Crowe SA, Paris G, Katsev S, et al. Sulfate was a trace constituent of archean seawater. *Science* 2014;346(6210):735–739; doi: 10.1126/science.1258966

Dai X, Müller H-G, Yao F. Optimal bayes classifiers for functional data and density ratios. *Biometrika* 2017;104(3): 545–560; doi: 10.1093/biomet/asx024

Demšar J, Curk T, Erjavec A, et al. Orange: Data mining toolbox in python. *J Mach Learn Res* 2013;14(71):2349–2353.

Ding T, Wang J, Tan T, et al. Sulfur isotopic and trace-element compositions of pyrite from the zankan iron deposit, west kunlun orogenic belt, China: Possible early cambrian banded iron formations. *Geol J* 2021;56(11):5738–5754; doi: 10.1002/gj.4270

Duverger A, Bernard S, Viennet J-C, et al. Formation of pyrite spherules from mixtures of biogenic fes and organic compounds during experimental diagenesis. *Geochem Geophys Geosyst* 2021;22(10):e2021GC010056; doi: 10.1029/2021GC010056

Ehlmann BL, Edwards CS. Mineralogy of the martian surface. *Annu Rev Earth Planet Sci* 2014;42(1):291–315; doi: 10.1146/annurev-earth-060313-055024

Eldridge DL, Guo W, Farquhar J. Theoretical estimates of equilibrium sulfur isotope effects in aqueous sulfur systems: Highlighting the role of isomers in the sulfite and sulfoxylate systems. *Geochim Cosmochim Acta* 2016;195:171–200; doi: 10.1016/j.gca.2016.09.021

Fike DA, Bradley AS, Rose CV. Rethinking the ancient sulfur cycle. *Annu Rev Earth Planet Sci* 2015;43(1):593–622; doi: 10.1146/annurev-earth-060313-054802

Franz HB, McAdam AC, Ming DW, et al. Large sulfur isotope fractionations in martian sediments at Gale Crater. *Nature Geosci* 2017;10(9):658–662; doi: 10.1038/ngeo3002

Gaillard F, Michalski J, Berger G, et al. Geochemical reservoirs and timing of sulfur cycling on mars. *Space Sci Rev* 2013; 174(1–4):251–300; doi: 10.1007/s11214-012-9947-4

Gellert R, Clark BC. , III, and MSL and MER Science Teams. In situ compositional measurements of rocks and soils with the alpha particle x-ray spectrometer on NASA's mars rovers. *Elements* 2015;11(1):39–44; doi: 10.2113/gselements.11.1.39

Genna D, Gaboury D. Deciphering the hydrothermal evolution of a VMS System by LA-ICP-MS using trace elements in pyrite: An example from the bracemac-mcleod deposits, abitibi, canada, and implications for exploration*. *Econ Geol* 2015;110(8):2087–2108; doi: 10.2113/econgeo.110.8.2087

George LL, Biagioni C, D'Orazio M, et al. Textural and trace element evolution of pyrite during greenschist facies metamorphic recrystallization in the southern apuan alps (Tuscany, Italy): Influence on the formation of Tl-Rich Sulfosalt Melt. *Ore Geol Rev* 2018;102:59–105; doi: 10.1016/j.oregeorev.2018.08.032

Gilbert SE, Danyushevsky LV, Goemann K, et al. Fractionation of sulphur relative to iron during laser Ablation-ICP-MS analyses of sulphide minerals: Implications for quantification. *J Anal At Spectrom* 2014;29(6):1024–1033; doi: 10.1039/C4JA00012A

Goldfarb RJ, Baker T, Dubé B, et al. Distribution, character, and genesis of gold deposits in metamorphic terran. In: One Hundredth Anniversary Volume. (Hedenquist JW, Thompson JFH, Goldfarb RJ eds) Society of Economic Geologists; 2005; p. 0; doi: 10.5382/AV100.14

Gomes ML, Fike DA, Bergmann KD, et al. Environmental insights from high-resolution (SIMS) Sulfur isotope analyses of sulfides in proterozoic microbialites with diverse mat textures. *Geobiology* 2018;16(1):17–34; doi: 10.1111/gbi.12265

Gomes ML, Hurtgen MT. Sulfur isotope fractionation in modern euxinic systems: Implications for paleoenvironmental reconstructions of paired sulfate–sulfide isotope records. *Geochim Cosmochim Acta* 2015;157:39–55; doi: 10.1016/j.gca.2015.02.031

Gorjan P, Kaiho K, Fike DA, et al. Carbon- and sulfur-isotope geochemistry of the hirnantian (late ordovician) wangjiawan (riverside) section, south china: Global correlation and environmental event interpretation. *Palaeogeogr Palaeoclimatol Palaeoecol* 2012;337–338:14–22; doi: 10.1016/j.palaeo.2012.03.021

Gregory D, Mukherjee I, Olson SL, et al. The formation mechanisms of sedimentary pyrite nodules determined by trace element and sulfur isotope microanalysis. *Geochim Cosmochim Acta* 2019a;259:53–68; doi: 10.1016/j.gca.2019.05.035

Gregory DD. The Pyrite Trace Element Paleo-Ocean Chemistry Proxy. In: Elements in Geochemical Tracers in Earth System Science Cambridge University Press; 2020.

Gregory DD, Cracknell MJ, Large RR, et al. Distinguishing ore deposit type and barren sedimentary pyrite using laser ablation-inductively coupled plasma-mass spectrometry trace element data and statistical analysis of large data sets. *Econ Geol* 2019b;114(4):771–786; doi: 10.5382/econgeo.4654

Gregory DD, Large RR, Halpin JA, et al. The chemical conditions of the late archean hamersley basin inferred from whole rock and pyrite geochemistry with Δ33S and *δ*34S isotope analyses. *Geochim Cosmochim Acta* 2015;149:223–250; doi: 10.1016/j.gca.2014.10.023

Gregory DD, Lyons TW, Large RR, et al. Whole rock and discrete pyrite geochemistry as complementary tracers of ancient ocean chemistry: An example from the neoproterozoic doushantuo formation, China. *Geochim Cosmochim Acta* 2017;216:201–220; doi: 10.1016/j.gca.2017.05.042

Guillong M, Hametner K, Reusser E, et al. Preliminary characterisation of new glass reference materials (GSA-1G, GSC-1G, GSD-1G and GSE-1G) by laser ablation-inductively coupled plasma-mass spectrometry using 193 Nm, 213 Nm and 266 Nm wavelengths. *Geostand Geoanalyt Res* 2005; 29(3):315–331; doi: 10.1111/j.1751-908X.2005.tb00903.x

Guillong M, Meier DM, Allan MM, et al. Appendix A6: SILLS: A MATLAB-Based program for the reduction of laser ablation icp-ms data of homogeneous materials and inclusions. *Mineral Assoc Can Short Course* 2008;40:328–333.

Habicht KS, Gade M, Thamdrup B, et al. Calibration of sulfate levels in the archean ocean. *Science* 2002;298(5602):2372–2374; doi: 10.1126/science.1078265

Halevy I, Zuber MT, Schrag DP. A sulfur dioxide climate feedback on early mars. *Science* 2007;318(5858):1903–1907; doi: 10.1126/science.1147039

Hu G, Guo J, Tian L, et al. The effect of thermochemical sulphate reduction on the carbon isotope ratio of individual light hydrocarbons associated with natural gas. *Front Earth Sci* 2022;10.

Hutchison W, Finch AA, Boyce AJ. The sulfur isotope evolution of magmatic-hydrothermal fluids: Insights into ore-forming processes. *Geochim Cosmochim Acta* 2020;288: 176–198; doi: 10.1016/j.gca.2020.07.042

Jiang L, Worden RH, Cai C. Generation of isotopically and compositionally distinct water during thermochemical sulfate reduction (tsr) in carbonate reservoirs: Triassic feixianguan formation, sichuan basin, China. *Geochim Cosmochim Acta* 2015;165:249–262; doi: 10.1016/j.gca.2015.05.033

Jiang L, Worden RH, Yang C. Thermochemical sulphate reduction can improve carbonate petroleum reservoir quality. *Geochim Cosmochim Acta* 2018;223:127–140; doi: 10.1016/j.gca.2017.11.032

Johnson SC, Large RR, Meffre S. Primary Metal Enrichment and Metamorphism at the Talvivaara Ni-Zn-Cu-Co Deposit, Finland. Proc SGA Bienn Meet 2015;(Society of Geology Applied to Mineral Deposits: Geneva, Switzerland).

Jørgensen BB, Findlay AJ, Pellerin A. The biogeochemical sulfur cycle of marine sediments. *Front Microbiol* 2019;10:849; doi: 10.3389/fmicb.2019.00849

Keith M, Häckel F, Haase KM, et al. Trace element systematics of pyrite from submarine hydrothermal vents. *Ore Geol Rev* 2016;72:728–745; doi: 10.1016/j.oregeorev.2015.07.012

Kiyosu Y, Krouse HR. The role of organic acid in the abiogenic reduction of sulfate and the sulfur isotope effect. *Geochem J* 1990;24(1):21–27; doi: 10.2343/geochemj.24.21

Labidi J, Cartigny P, Jackson MG. Multiple sulfur isotope composition of oxidized samoan melts and the implications of a sulfur isotope ‘mantle array’ in chemical geodynamics. *Earth Planet Sci Lett* 2015;417:28–39; doi: 10.1016/j.epsl.2015.02.004

LaFlamme C, Hollis SP, Jamieson JW, et al. Three-Dimensional spatially constrained sulfur isotopes highlight processes controlling sulfur cycling in the near surface of the iheya north hydrothermal system, okinawa trough. *Geochem Geophys Geosyst* 2018;19(8):2798–2812; doi: 10.1029/2018GC007499

Large RR, Halpin JA, Danyushevsky LV, et al. Trace element content of sedimentary pyrite as a new proxy for deep-time ocean–atmosphere evolution. *Earth Planet Sci Lett* 2014;389:209–220; doi: 10.1016/j.epsl.2013.12.020

Large RR, Maslennikov VV, Robert F, et al. Multistage sedimentary and metamorphic origin of pyrite and gold in the giant sukhoi log deposit, lena gold province, Russia. *Econ Geol* 2007;102(7):1233–1267; doi: 10.2113/gsecongeo.102.7.1233

Large RR, Mukherjee I, Gregory DD, et al. Ocean and atmosphere geochemical proxies derived from trace elements in marine pyrite: Implications for ore genesis in sedimentary basins. *Econ Geol* 2017;112(2):423–450; doi: 10.2113/econgeo.112.2.423

Le Pape P, Blanchard M, Brest J, et al. Arsenic incorporation in pyrite at ambient temperature at both tetrahedral s-i and octahedral feii sites: Evidence from EXAFS-DFT analysis. *Environ Sci Technol* 2017;51(1):150–158; doi: 10.1021/acs.est.6b03502

Leavitt WD, Halevy I, Bradley AS, et al. Influence of sulfate reduction rates on the phanerozoic sulfur isotope record. *Proc Natl Acad Sci U S A* 2013;110(28):11244–11249; doi: 10.1073/pnas.1218874110

Li D, Chen H, Sun X, et al. Coupled trace element and sims sulfur isotope geochemistry of sedimentary pyrite: Implications on pyrite growth of caixiashan Pb–Zn Deposit. *Geosci Front* 2019;10(6):2177–2188; doi: 10.1016/j.gsf.2019.05.001

Li J, Zhao C, Huang Y, et al. In-Situ sulfur isotope and trace element of pyrite constraints on the formation and evolution of the nibao carlin-type gold deposit in SW China. *Acta Geochim* 2019;38(4):555–575; doi: 10.1007/s11631-019-00342-0

Liu Z, Chen D, Zhang J, et al. Pyrite morphology as an indicator of paleoredox conditions and shale gas content of the longmaxi and wufeng shales in the middle yangtze Area, South China. *Minerals* 2019;9(7):428; doi: 10.3390/min9070428

Lorand J-P, Labidi J, Rollion-Bard C, et al. The sulfur budget and sulfur isotopic composition of martian regolith breccia NWA 7533. *Meteorit & Planetary Scien* 2020;55(9):2097–2116; doi: 10.1111/maps.13564

Lui TCC, Gregory DD, Anderson M, et al. Applying machine learning methods to predict geology using soil sample geochemistry. *Appl Comput Geosci* 2022;16:100094; doi: 10.1016/j.acags.2022.100094

Lyons TW, Werne JP, Hollander DJ, et al. Contrasting sulfur geochemistry and Fe/Al and Mo/Al ratios across the last oxic-to-anoxic transition in the cariaco basin, Venezuela. *Chem Geol* 2003;195(1–4):131–157; doi: 10.1016/S0009-2541(02)00392-3

MacArthur JL, Bridges JC, Hicks LJ, et al. Mineralogical constraints on the thermal history of martian regolith breccia northwest Africa 8114. *Geochim Cosmochim Acta* 2019;246:267–298; doi: 10.1016/j.gca.2018.11.026

Machel HG. Bacterial and thermochemical sulfate reduction in diagenetic settings—old and new insights. *Sediment Geol* 2001;140(1–2):143–175; doi: 10.1016/S0037-0738(00)00176-7

Machel HG, Krouse HR, Sassen R. Products and distinguishing criteria of bacterial and thermochemical sulfate reduction. *Appl Geochem* 1995;10(4):373–389; doi: 10.1016/0883-2927(95)00008-8

Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol* 2010;5(9):1315–1316; doi: 10.1097/JTO.0b013e3181ec173d

Marini L, Moretti R, Accornero M. Sulfur isotopes in magmatic-hydrothermal systems, melts, and magmas. *Rev Mineral Geochem* 2011;73(1):423–492; doi: 10.2138/rmg.2011.73.14

Meng X, Li X, Chu F, et al. Trace element and sulfur isotope compositions for pyrite across the mineralization zones of a sulfide chimney from the east pacific rise (1-2°S). *Ore Geol Rev* 2020;116:103209; doi: 10.1016/j.oregeorev.2019.103209

Meshoulam A, Ellis GS, Said Ahmad W, et al. Study of thermochemical sulfate reduction mechanism using compound specific sulfur isotope analysis. *Geochim Cosmochim Acta* 2016;188:73–92; doi: 10.1016/j.gca.2016.05.026

Meyer NR, Zerkle AL, Fike DA. Sulphur cycling in a neoarchaean microbial mat. *Geobiology* 2017;15(3):353–365; doi: 10.1111/gbi.12227

Moreras-Marti A, Fox-Powell M, Cousins CR, et al. Sulfur isotopes as biosignatures for mars and europa exploration. *JGS* 2022;179(6):jgs2021–j134; doi: 10.1144/jgs2021-134

Morford JL, Russell AD, Emerson S. Trace metal evidence for changes in the redox environment associated with the transition from terrigenous clay to diatomaceous sediment, saanich inlet, BC. *Mar Geol* 2001;174(1–4):355–369; doi: 10.1016/S0025-3227(00)00160-2

Morin G, Noël V, Menguy N, et al. Nickel accelerates pyrite nucleation at ambient temperature. *Geochem Persp Let* 2017:6–11; doi: 10.7185/geochemlet.1738

Mukherjee I, Large RR. Co-Evolution of trace elements and life in precambrian oceans: The pyrite edition. *Geology* 2020;48(10):1018–1022; doi: 10.1130/G47890.1

Ohmoto H, Kakegawa T, Lowe DR. 3.4-Billion-Year-Old biogenic pyrites from barberton, south africa: Sulfur isotope evidence. *Science* 1993;262(5133):555–557; doi: 10.1126/science.11539502

Pasquier V, Bryant RN, Fike DA, et al. strong local, not global, controls on marine pyrite sulfur isotopes. *Sci Adv* 2021a;7(9): eabb7403; doi: 10.1126/sciadv.abb7403

Pasquier V, Fike DA, Halevy I. Sedimentary pyrite sulfur isotopes track the local dynamics of the peruvian oxygen minimum zone. *Nat Commun* 2021b;12(1):4403; doi: 10.1038/s41467-021-24753-x

Picard A, Gartman A, Girguis PR. What do we really know about the role of microorganisms in iron sulfide mineral formation? *Front Earth Sci* 2016;4; doi: 10.3389/feart.2016.00068

Reich M, Deditius A, Chryssoulis S, et al. Pyrite as a record of hydrothermal fluid evolution in a porphyry copper system: A SIMS/EMPA trace element study. *Geochim Cosmochim Acta* 2013;104:42–62; doi: 10.1016/j.gca.2012.11.006

Rickard D. Sedimentary Pyrite. In: Developments in Sedimentology Elsevier; 2012; pp. 233–285; doi: 10.1016/B978-0-444-52989-3.00006-4

Rickard D, Luther GW. Chemistry of iron sulfides. *Chem Rev* 2007;107(2):514–562; doi: 10.1021/cr0503658

Ridley J. Hydrothermal Ore Deposits I: Magmatic and Orogenic Environments. In: Ore Deposit Geology. 1st ed. Cambridge University Press; 2013; doi: 10.1017/CBO9781139135528

Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math* 1987; 20:53–65; doi: 10.1016/0377-0427(87)90125-7

Shelton JL, Jubb AM, Saxe SW, et al. Machine learning can assign geologic basin to produced water samples using major ion geochemistry. *Nat Resour Res* 2021;30(6):4147–4163; doi: 10.1007/s11053-021-09949-8

Shen Y, Buick R, Canfield DE. Isotopic evidence for microbial sulphate reduction in the early archaean era. *Nature* 2001; 410(6824):77–81; doi: 10.1038/35065071

Shen Y, Farquhar J, Masterson A, et al. Evaluating the role of microbial sulfate reduction in the early archean using quadruple isotope systematics. *Earth Planet Sci Lett* 2009;279(3–4): 383–391; doi: 10.1016/j.epsl.2009.01.018

Sim MS, Bosak T, Ono S. Large sulfur isotope fractionation does not require disproportionation. *Science* 2011;333(6038): 74–77; doi: 10.1126/science.1205103

Sim MS, Ogata H, Lubitz W, et al. Role of APS reductase in biogeochemical sulfur isotope fractionation. *Nat Commun* 2019;10(1):44; doi: 10.1038/s41467-018-07878-4

Skirrow RG. Iron Oxide Copper-Gold (IOCG) Deposits – A Review (Part 1): Settings, mineralogy, ore geochemistry and classification. *Ore Geol Rev* 2022;140:104569; doi: 10.1016/j.oregeorev.2021.104569

Squyres SW, Grotzinger JP, Arvidson RE, et al. In situ evidence for an ancient aqueous environment at meridiani planum, mars. *Science* 2004;306(5702):1709–1714; doi: 10.1126/science.1104559

Sun Y, Jiao Y, Cuney M, et al. Sulfur isotope and trace element constraints on the conditions of pyrite formation from the diantou-shuanglong sandstone-hosted uranium deposit, ordos basin, china: Implications for uranium mineralization. *Ore Geol Rev* 2024;165:105921; doi: 10.1016/j.oregeorev.2024.105921

Sweere TC, Dickson AJ, Vance D. Nickel and zinc micronutrient availability in phanerozoic oceans. *Geobiology* 2023; 21(3):310–322; doi: 10.1111/gbi.12541

Sykora S, Cooke DR, Meffre S, et al. Evolution of pyrite trace element compositions from porphyry-style and epithermal conditions at the lihir gold deposit: Implications for ore genesis and mineral processing. *Econ Geol* 2018;113(1):193–208; doi: 10.5382/econgeo.2018.4548

Tan H, Shao Y, Liu Q, et al. Textures, trace element geochemistry and in-situ sulfur isotopes of pyrite from the xiaojiashan gold deposit, jiangnan orogen: Implications for ore genesis. *Ore Geol Rev* 2022;144:104843; doi: 10.1016/j.oregeorev.2022.104843

Thamdrup B, Finster K, Hansen JW, et al. Bacterial disproportionation of elemental sulfur coupled to chemical reduction of iron or manganese. *Appl Environ Microbiol* 1993;59(1): 101–108; doi: 10.1128/aem.59.1.101-108.1993

Tino CJ, Stüeken EE, Arp G, et al. Are large sulfur isotope variations biosignatures in an ancient, impact-induced hydrothermal mars analog? *Astrobiology* 2023;23(10):1027–1044; doi: 10.1089/ast.2022.0114

Toulmin , III P, Baird AK, Clark BC, et al. Geochemical and mineralogical interpretation of the viking inorganic chemical results. *J Geophys Res* 1977;82(28):4625–4634; doi: 10.1029/JS082i028p04625

Tribovillard N, Algeo TJ, Lyons T, et al. Trace metals as paleoredox and paleoproductivity proxies: An update. *Chem Geol* 2006;232(1–2):12–32; doi: 10.1016/j.chemgeo.2006.02.012

Truong C, Bernard S, Le Pape P, et al. Production of carbon-containing pyrite spherules induced by hyperthermophilic thermococcales: A biosignature? *Front Microbiol* 2023;14: 1145781; doi: 10.3389/fmicb.2023.1145781

Vaniman DT, Bish DL, Ming DW, et al. MSL Science Team. Mineralogy of a mudstone at yellowknife bay, gale crater, mars. *Science* 2014;343(6169):1243480; doi: 10.1126/science.1243480

Wernette BW, Ripley EM, Li C, et al. Strongly negative δ34s values associated with secondary pyrite above and below the j-m reef, stillwater complex, montana. *Chem Geol* 2018;493: 58–66; doi: 10.1016/j.chemgeo.2018.05.023

Zhang W-D, Li B, Lu A-H, et al. In-situ pyrite trace element and sulfur isotope characteristics and metallogenic implications of the qixiashan Pb-Zn-Ag polymetallic deposit, eastern China. *Ore Geol Rev* 2022;144:104849; doi: 10.1016/j.oregeorev.2022.104849

Zhang Y-C, Sakhanenko L. The naive bayes classifier for functional data. *Stat Probab Lett* 2019;152:137–146; doi: 10.1016/j.spl.2019.04.017

Zhu G, Li T, Huang T, et al. Quantifying the seawater sulfate concentration in the cambrian ocean. *Front Earth Sci* 2021;9.

Address correspondence to:
M.C. Figueroa Pacific
Coastal and Marine Science Center
U.S. Geological Survey
Santa Cruz, CA
USA

E-mail: mfigueroa@usgs.gov

Submitted February 15, 2024
Accepted October 7, 2024
Associate Editor: Michael Storrie-Lombardi

Abbreviations Used

AUC = area under the ROC curve
CA = class accuracy
IOCG = iron oxide copper-gold
IQR = interquartile range
K-NN = K-nearest neighbors
(LA-)ICP-MS = (laser ablation) inductively coupled plasma mass spectrometry
LTH = low-temperature hydrothermal
ML = machine learning
MSR = microbial sulfate reduction
PMH = polymetallic hydrothermal
RF = Random Forest
ROC = receiver operator characteristic
SHRIMP-SI = sensitive high-resolution microprobe – stable isotope
SMS = seafloor massive sulfide
TE = trace element
TSR = thermochemical sulfate reduction