Revision 1

Predicting olivine formation environments using machine learning and implications for magmatic sulfide prospecting

Word Count: 7020

Shengchao Xue[1], Yunyun Niu[2,*], Zhuosen Yao[3], Luyang Wang[1], Xiaohao Zhang[2], Qingfei Wang[1*]

[1] *State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences, Beijing 100083, China*

[2] *School of Information Engineering, China University of Geosciences, Beijing 100083, China*

[3] *School of Earth Resources, China University of Geosciences, Wuhan 430074, China*

Corresponding author: Yunyun Niu

86- 18911528973

yniu@cugb.edu.cn

ABSTRACT

Global volcanic and plutonic olivines record the compositional characteristics and physicochemical conditions of the parental magmas. Thus, they have significant potential for use as petrogenetic discriminators of the olivine formation environment and prospecting indicators for potential host rocks of magmatic sulfide deposits. Several data visualization approaches have been proposed by researchers to determine olivine origins. However, they can only discriminate specific olivine populations and require the incorporation of trace elements for which data are lacking globally. In this study, a machine-learning method consisting of the random forest algorithm and the synthetic minority oversampling technique (SMOTE) is used to discriminate the crystallization environments of olivine and predict the sulfide potential of olivine-bearing mafic-ultramafic intrusions. We employ a global dataset of 24341 olivine samples from twelve environments to determine the contents of MgO, FeO, Ni, Ca, Mn, and Cr and the Fo number (100×Mg/(Mg+Fe)). The results indicate that the proposed method can classify olivine into genetically distinct populations and distinguish olivine derived from mineralized intrusions from that derived from sulfide-barren intrusions with high accuracies (higher than 99% on average). We develop a dimensionality reduction algorithm to visualize the olivine classifications using low-dimensional vectors and an olivine classifier (accessible at http://cugb.online:8080/olivine_web/main.html). The model is used successfully to identify the contributions of distinct sources to regional magmatism using olivines from the late Permian picrite and basalt along the western margin of the Yangtze block (SW China) and to predict the sulfide potential of the newly-discovered Qixin mafic-ultramafic complex in the southern Central Asian Orogenic Belt (NW China). The findings suggest that the proposed approach enables the accurate identification of olivine origins in different formation environments and is a reliable indicator suitable for global Ni-Cu- platinum group element (PGE) exploration.

Keywords: Olivine chemistry, mineral formation environment, magmatic sulfide potential, formation environment classification, ore deposit prediction, machine learning

INTRODUCTION

Olivine is an important and ubiquitous mineral in global volcanic and plutonic ultramafic rocks and occurs in a wide range of ages and geological environments (Fig. 1; Lehnert et al., 2000; Sobolev et al., 2007; Li et al., 2012; Barnes et al., 2022). It is crucial for understanding the Earth's upper mantle and the genesis and evolution of mantle melts (Foley et al., 2013; Herzberg et al., 2016). The controlling factors of the olivine composition of mafic and ultramafic igneous rocks include the mantle melt composition, which is affected by different sources and potential temperatures (Herzberg, 2011), fractional crystallization with or without sulfide (Li and Naldrett, 1999; Li et al., 2007), magma replenishment and mixing, post-cumulus processes, including the reaction with interstitial silicate melts and/or sulfide liquids (Brenan, 2003; Barnes et al., 2013), and subsolidus re-equilibration (Cameron, 1976; Barnes, 1986; Mao et al., 2022). These controlling factors can be assessed by determining the percentage of magnesian end-member forsterite (Fo number, 100×Mg/(Mg+Fe)), which is the most important olivine parameter, and the contents of minor elements, such as Ni, Ca, Mn, and Cr, which are typically analyzed along with major elements. Therefore, the major-minor element compositions of olivine have been widely used as petrogenetic indicators to distinguish the sources of olivine populations (Green and Ringwood, 1967; Sobolev et al., 2005, 2007; Li and Ripley, 2010; Herzberg, 2011; Howarth and Harris, 2017). However, these studies mostly relied on one or several geochemical indicators (*e.g.*, Fo number, Ca, Ni, or Ti content, and Fe/Ni, Zn/Fe, Mn/Zn, or Ni/Co ratio) derived from traditional clustering techniques by naked eye, resulting in difficulties in distinguishing the olivine origin for a wide range of ages and tectonic settings.

More specially, olivine is an early-crystallized silicate phase of almost all mantle-derived mafic/ultramafic magmas that is parental to numerous important magmatic Ni-Cu-platinum group element (PGE) sulfide deposits worldwide (Li et al., 2007; Naldrett, 2011). The compatible character of Ni leads to a strong tendency to partition into olivine (Simkin and Smith, 1970), while its chalcophile character results in much stronger partitioning into sulfide liquid (Rajamani and Naldrett, 1978; Li and Ripley, 2010; Kiseeva and Wood, 2015; Yao et al., 2018). Hence, the olivine Ni content is often considered an indicator of magmatic sulfide deposits in potential host rocks (Li and Naldrett, 1999; Barnes et al., 2004, 2022; Le Vaillant et al., 2016). However, the use of Ni-olivine as a fertility tool requires significant a priori knowledge of the olivine formation processes and the exploration regions; thus, this indicator is of low effectiveness. In addition, a recent empirical data-driven approach has shown that no universal, clear discrimination can be made between sulfide-bearing and sulfide-barren intrusions using the olivine Ni content at a given Fo content (Barnes et al., 2022). The same results were observed when trace elements were considered (Mao et al., 2022). Hence, it is unclear whether systematic geochemical differences exist and whether they enable the reliable discrimination of sulfide-mineralized intrusions.

Decoding the hidden information of mineral chemistry in high-dimensional vectors is potentially an effective method and an alternative to traditional data visualization techniques. Machine learning is a branch of computer science that analyzes large data sets using complex algorithms. This method has been used for classification, prediction, and clustering (Kuwatani et al., 2015; Petrelli and Perugini, 2016; Ueki et al., 2018; Petrelli et al., 2020; Lösing and Ebbing, 2021; Wang et al., 2021; Cheng et al., 2022). It has also been used recently in mineral chemistry in Earth science and was successfully applied to distinguish barren sedimentary and pyrite ore deposits (Gregory et al., 2019), predict quartz-forming environments (Wang et al., 2021), and characterize magma fertility of porphyry copper deposits via zircon chemistry (Zou et al., 2022).

A wealth of data has been accumulated on major and minor elements in olivine, and some information has been collected systematically (Barnes et al., 2022; Cheng et al., 2022). This information can be used for petrological and geochemical interpretation. We utilize a global data set of major-minor elements of olivine from global volcanic and intrusive rocks and establish a random forest model to investigate the chemical features of olivine from different ages and environments. The objective is the geochemical discrimination of magmatism in different tectonic settings. We also use machine learning algorithms to evaluate the magmatic sulfide potential of mafic-ultramafic intrusions.

DATA AND CLASSIFICATION

The major and minor element data of global volcanic and plutonic olivines, including magmatic Ni-Cu sulfide orebodies, were obtained from publications listed in Supplemental Table 1 (Table S1) and from the compilation of published data on olivine mineral chemistry available at https://zenodo.org/record/5787901 (Barnes et al., 2022).

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8893. http://www.minsocam.org/

The sample set used in this study comprises 24341 analyses of olivine grains from 54 locations worldwide (Fig. 1), including the contents of (wt.%) SiO_2, FeO, MgO, Cr_2O_3, MnO, CaO, and (ppm) Ni components and the Fo number (the mole percentage MgO/(MgO+FeO)) of olivine. These element concentrations were obtained from electron microprobe analyses with detection limits of minor elements, such as Ni, Mn, Ca, and Cr, of ~10^2 ppm or lower. The locations, ages, tectonic settings, rock types, and data sources of these olivine samples are provided in Table S1.

Volcanic and plutonic olivines were classified based on the tectonic environments and ages, including greenstone belts from the Archean (A-GB) to the Proterozoic (P-GB), continental blocks (continental large igneous provinces (LIP) and rifts) from the Proterozoic (P-CB) to the Phanerozoic (Ph-CB), convergent margins/orogenic belts from the Proterozoic (P-OB) to the Phanerozoic (Ph-OB), Phanerozoic (Ph-MORB) mid-ocean ridge basalts, and Phanerozoic (Ph-OLIP) oceanic large igneous provinces. We also included the status of the host rocks in the classification: volcanics (V), small intrusions (SI), and large layered mafic intrusions (LMI). Olivines from komatiite and other mafic volcanics in oceanic LIPs were divided into two types (OLIP-VK and OLIP-V) because high-Ni olivines (up to 5,000 ppm) tend to be crystallized from komatiitic magmas, whereas the Ni content is lower in olivine crystallized from more fractionated magmas (Crocket, 2002; Arndt et al., 2005; Barnes and Lightfoot, 2005). Some extreme examples of Ni enrichment, such as the ultra-nickeliferous olivine of the Kevitsa Ni-Cu-PGE mineralized intrusion in northern Finland (Yang et al., 2013), were not included in the data. As a result, we classified the olivines into twelve genetic types corresponding to distinct formation environments: A-GB, P-GB, P-CB-LMI, P-CB-SI, P-OB-LMI, P-OB-SI, Ph-CB-SI, Ph-CB-V, Ph-MORB, Ph-OB-SI, Ph-OLIP-V, and Ph-OLIP-VK (Table S1). These genetic types were divided into two subgroups based on the mineralization level of the host mafic-ultramafic intrusion: mineralized (M) containing economic Ni-Cu-(PGE) sulfide and sub-economic disseminated Ni-Cu-(PGE) sulfide mineralization, and barren (B) containing minor or no sulfide mineralization. Figure 2 shows the correlations between the Fo values and the contents or ratios of critical elements for the twelve types. The different types of olivine cannot be distinguished when the Fo content exceeds 65 mol.%. The only distinction can be observed in fayalite-rich olivines (Fo < 60 mol.%) of PCB-LMI and Ph-CB-V types (Table S1).

METHODS

Random forest algorithm

The random forest algorithm (Breiman, 2001; Cutler, 2012) is an ensemble learning method that utilizes multiple decision trees. Features are randomly selected to train the decision trees. It does not use the same training set to train the base classifier but utilizes bootstrap resampling (Efron, 1992), a random sampling method with replacement. Therefore, some samples may be selected multiple times, whereas other samples (~ 37%) may never be sampled even if an infinite number of samples are selected. These samples are referred to as out-of-bag data and are often used in the validation set to evaluate the model's generalization performance (out-of-bag estimate) (Wolpert and Macready, 1999). In the Bootstrap sampling method, T sampling sets are selected as the training sets of T individual learners, which are trained separately.

The classification prediction results of the individual learners are combined by voting, which calculates the number of votes obtained for each prediction result and selects the highest number of votes as the final decision. The final prediction result is defined in Eqs. (1) and (2):

$$H(\boldsymbol{x}) = c_{\mathrm{N}} \tag{1}$$

$$N = \underset{j}{\operatorname{argmax}} \sum_{i=1}^{T} h_i^j(\boldsymbol{x}) \tag{2}$$

where x denotes the training data, h_i is the individual learner i. Each individual learner predicts a class label from the class label set $\{c_1, c_2, c_3, \ldots, c_N\}$; $h_i^j(\boldsymbol{x})$ is the output of class label c_j by the individual learner. It is an n-dimensional vector. The prediction results of all individual learners are combined, and the *argmax* function is used to obtain the maximum class index inx of the vector, which is the final predicted class c_{inx}.

Random forest is a type of classification and regression tree (CART) (Breiman, 2017) but it uses an ensemble of many trees and random feature selection for training. At each node of the tree, a random feature subset is selected containing n ($n < d$) features (with d features), and the optimal feature is chosen. The value of n determines the degree of randomness, where $n = log_2 d$ (Breiman, 2001). The Gini index is used to select the optimal feature and the optimal binary segmentation point of the feature. The flowchart of the random forest algorithm is shown in Figure 3.

Model construction processes

The olivine dataset was divided into a training set and a test set with a ratio of 9:1. The training set was used for model training, and the test set was to test the generalization ability of the model. The out-of-bag data were used for the validation. We standardized the data by subtracting the average value from each feature value and dividing it by the variance. Hence, the data were clustered around 0, and the variance was 1. The samples were unbalanced, i.e., there was a difference in the number of samples in different categories. Model training is sensitive to the number of samples, resulting in a low accuracy rate for categories with few samples. Therefore, we used the synthetic minority oversampling technique (SMOTE) to augment the data, which deals with unbalanced samples by interpolation. The labels of the sample data were one-hot encoded to facilitate model training.

A grid search strategy and five-fold cross-validation (Kohavi, 1995) were used to tune the hyperparameters of the random forest model. After setting the candidate hyperparameter value range, the model was trained, and the optimum hyperparameter combination was selected based on the predicted score of the model on the olivine data set. Five-fold cross-validation divides the training data set into five subsets with the same sample size: four training subsets and one validation subset. The training subset was used to train the model, and the validation subset was used for the evaluation. Each subset was used as a training set so that the model was trained and evaluated five times. The predicted score of the model is the average of the five predicted scores. The grid search strategy enables the model to select the optimal hyperparameters to obtain the optimal model. The model was then trained using the olivine data training data set and validated with the out-of-bag data.

We used four evaluation metrics: accuracy, precision, recall, and F1 score (Dice, 1945; Sorensen, 1948). Accuracy is the proportion of correctly predicted olivine samples to the total number of samples. Precision is the proportion of correctly predicted positive samples to all predicted positive samples, and recall is the proportion of correctly predicted positive samples to the number of true positive and false negative samples. The F1 score is the harmonic mean of precision and recall.

RESULTS AND DISCUSSION

Feature importance

The olivine dataset contains eight features, each of which has a different degree of influence on the classification results. Therefore, feature importance analysis is performed on the olivine data to measure the contribution of the input features to the model prediction result and determine the degree of correlation between the feature and the target. For a given feature, a higher importance score indicates that the feature has more importance in the classification. The influence of each feature on the prediction results is analyzed using four classification criteria, including the olivine formation environment, mineralization status, volcanic status, and petrography group. Only features with high importance are considered because after the model reaches the highest accuracy, adding more features to the model does not increase the accuracy but decreases it slightly (Zou et al., 2022).

The feature importance of the four classification criteria is shown in Figure 4. The ranking of the features for identifying the olivine formation environment is from high to low the contents of CaO, Ni, FeO, Cr_2O_3, Fo, MgO, MnO, and SiO_2 (Fig. 4a). The CaO and Ni contents have the highest importance scores, while the FeO, Fo, Cr_2O_3, and MgO contents have comparable feature importance scores for predicting the mineralization status (Fig. 4b). The CaO content has the top score for predicting the volcanic status and petrography group, and the Ni content ranks second for predicting the volcanic status and is less important for predicting the petrography group (Fig. 4c-d). In all cases, The SiO_2 content has the lowest feature importance scores for all four criteria. The importance scores of the olivine elements for predicting the olivine formation environment are plausible. Olivine consists of >99% MgO, FeO, and SiO_2 (Foley et al., 2013). Thus, the major element contents provide insufficient information on the formation environment, but the comparable feature importance scores of the MgO, FeO, and Fo contents suggest that the combination of these features contributes significantly to the classification results. Olivine has a simple crystal structure; the octahedral M1 and M2 sites have a similar size (73 and 76.7 pm) (Zanetti et al., 2004) and provide locations for trace elements that have similar cation radii and charges to substitute for Mg^{2+} (72 pm) and Fe^{2+} (78 pm). The divalent cations Ni^{2+} (69 pm) and Mn^{2+} (67 pm) readily enter these octahedral sites; thus, they have high contents in olivine, and their partitioning behavior is controlled by the temperature and the MgO concentrations of olivine and silicate melts (Matzen et al., 2013; 2017a; b). The Ni and Mn contents in the mantle-derived melt are sensitive to the tectonic setting (Yao et al., 2018; Chen et al., 2022), which may also hold for their contents in the olivine crystallized from the mantle-derived melt. However, the values and ranges of the partition coefficient of Mn between olivine and the melt are ~0.5-1 (Matzen et al., 2017b), limiting the variability of the Mn content in the olivine. Therefore, Mn has an insignificant role (Fig. 4). In contrast, the partition coefficient of Ni is highly variable (~3-90, Foley et al., 2013; Matzen et al., 2017a), indicating strong compositional variation. There is also competition in the partitioning of Ni between olivine and sulfide, explaining the high importance scores of Ni in the three classification results (Fig. 4a-c).

Calcium is the most complex elements in our dataset. The large radius of Ca^{2+} (100 pm) causes it to preferentially occupy the larger M2 sites by Ca-Fe substitution (Coogan et al., 2005), which is counterbalanced by a higher proportion of Mg^{2+} at the smaller M1 sites (Di Stefano et al., 2019). Therefore, the partitioning of Ca into olivine becomes more pronounced as the Fo number decreases (Libourel, 1999). In addition, the inverse relationship between pressure and the olivine Ca content has been widely used to distinguish between volcanic and mantle olivines (Simkin and Smith, 1970; Foley et al., 2013), whereas the temperature effect was also proposed to be critical for Ca partitioning

(O'Reilly et al., 1997). Systematic experiments have shown that the amount of Ca entering olivine is proportional to the number of network-modifying Ca cations available in the melts, which is highly sensitive to the alumina, alkali, and ferrous iron contents of the melt (Libourel, 1999). The magmatic H_2O content also affects the partitioning of CaO between olivine and the silicate melt. Therefore, a Ca-in-olivine geohygrometer was fabricated to detect olivine grains crystallized from hydrous subduction zone lavas (Gavrilenko et al., 2016). The CaO content of olivine contains hidden information on the alkali, aluminum, and water contents of the melt, the pressure-temperature conditions, and the melt structure (i.e., depolymerization) (Mysen, 2007). These parameters are critical for classifying the olivine formation environments but cannot be obtained from other major-minor elemental compositions of olivine. Therefore, the CaO content has the highest importance score for the four classification criteria (Fig. 4).

Classification results and visualization

The confusion matrix (Stehman, 1997) of the olivine classification is shown in Figures 5-7. The dark blue cells on the diagonal show the records that are in agreement. The normalized confusion matrices show that the accuracies of the random forest model for the formation environment (Fig. 5), mineralization status (Fig. 6), volcanic status (Fig. 7a), and petrography group (Fig. 7b) are very high (0.99, 0.99, 1.0, and 0.99, respectively) (Table 1). The proposed model has high evaluation indicators scores for different classification criteria, demonstrating its applicability for the discrimination of the olivine origins.

The Uniform Manifold Approximation and Projection (UMAP) algorithm (McInnes, 2018) is used for dimensionality reduction. The output was used to train a random forest model and create a two-dimensional decision boundary (Fig. 8). The UMAP algorithm provides a low-dimensional visualization of the proximity relationship between the sample points by calculating the similarity score between sample points. It finds the low-dimensional representation of the manifold by learning the manifold structure in the high-dimensional space. It should be noted that cross-correlations occur between the element concentrations of the olivine in this study (Fig. A1), and they have a negligible effect on dimensionality reduction. The classification accuracy of the model is listed in Table 2.

Model validation

Relationship between regional magmatism and olivine sources. The Emeishan LIP (ELIP) (~260 Ma) is a good example of a mantle plume-derived continental mafic LIP. It is located in the western part of the South China block and extends from southwestern China into northern Vietnam (Chung and Jahn, 1995; Xu et al., 2004; Zhang et al., 2006; Ali et al., 2010; Shellnutt, 2014). The ELIP covers an area of ~0.3×10^6 km^2 and is predominantly composed of basalts, with exposures of ultramafic and felsic volcanic rocks, layered mafic-ultramafic intrusions, and silicic plutonic rocks (Shellnutt, 2014). The South China block is composed of the Yangtze and Cathaysia blocks, which were amalgamated through the westward subduction and subsequent closure of the intervening ocean during the Neoproterozoic (Yao et al., 2019). Geological records suggest that the western margin of the Yangtze block was part of a Neoproterozoic arc system (Cawood et al., 2020). The late Permian plume and Neoproterozoic subduction-modified mantle interaction is reflected in the heterogeneity in isotopes of volcanic and plutonic rocks, such as the elevated $\delta^{18}O$ of olivine from some picritic rocks (Yu et al., 2017), the low $\delta^{18}O$ of zircon from the magmatic Fe-Ti-V oxide ore deposit (Tang et al., 2021), and the extremely enriched Nd isotopes ($\varepsilon_{Nd}(t) < -5$) of the late Permian mafic dykes exposed along the western margin of the Yangtze block (Wang et al., 2022).

Based on the plume-modified mantle interaction at the western margin, we infer that the olivine from the late Permian picrite and basalt along the margin should be a mixture of multiple olivine populations and may exhibit affinities to two distinct sources. However, the random forest model classified the olivine from these picritic and basaltic rocks (Hexi, Jizushan, and Huangcaoba areas) as the Ph-CB-V type (Fig. 9a; Table S2), which does not contain signatures of the orogenic belt. Hence, we investigated the Sr-Nd isotopes of the Ph-CB-V type-dominated picritic and basaltic rocks and found that these rocks had much higher $\varepsilon_{Nd}(t)$ values (-0.94 to 5.07) than those of the spatially related isotopically enriched mafic dykes (Fig. 9b; Table S3). The analytical methods and results are provided in Table S3. According to our classification results and the distinct groups of Sr-Nd isotopes (Wang et al., 2022), we suggest that the melting of a plume mantle component contributed the most picrite and basalt, whereas the mafic dykes at the western margin of the Yangtze block were generated by the interaction of two distinct sources, i.e., a mantle plume and a Neoproterozoic subduction-modified, Nd isotope-enriched lithospheric mantle. Therefore, the proposed machine learning model has significant potential for revealing the component contribution of heterogeneous sources to regional magmatism and constraining their geodynamic environments.

Mineral prospectivity of small intrusions in orogenic belt. The Central Asian Orogenic Belt (CAOB) is one of the largest accretionary orogens on the Earth and is situated between the Siberian, European, Tarim, and Sino-Korean cratons (Sengör et al., 1993; Jahn et al., 2000; Kröner et al., 2007; Windley et al., 2007 and references therein). Examples of magmatic sulfur-rich deposits generated in small mafic/ultramafic intrusions are commonly assigned to post-subduction magmatism in the CAOB, northern China (e.g., Kalatongke, Huangshandong, Huangshan, Tulaergen, Poyi, and Hongqiling-Piaohechuan; Song et al., 2009; Gao et al., 2013; Wei et al., 2013, 2019; Li et al., 2019; Xue et al., 2016, 2021, 2022). These large deposits and more than a hundred prospects are excellent targets for the localization of Cu-Ni mineralization in this accretionary orogenic belt. The Qixin mafic-ultramafic complex is a newly discovered mafic-ultramafic intrusion under active exploration at the southern margin of the CAOB. The complex is characterized by a large gabbroic body and several small ultramafic bodies intruding into older gabbroic and metamorphic rocks (Xue et al., 2019). The ultramafic bodies are composed predominantly of lherzolite, troctolite, and minor amounts of websterite.

As listed in Table S2, the random forest model predictions via olivine chemistry

DOI: https://doi.org/10.2138/am-2022-8893. http://www.minsocam.org/

demonstrate that a large proportion of the olivine in the Qixin ultramafic-troctolitic rocks is classified into the mineralization states Ph-OB-SI-M (Fig. 9c; Table S2). The prediction results are consistent with the long-held belief that olivine grains originated from a ~285 Ma orogeny-related small intrusion. Explorations have shown that the Qixin ultramafic intrusion hosts significant Ni-Cu sulfide mineralization zones based on an ongoing drilling program (Fig. 9d). As a result, we are optimistic that the proposed machine learning model can accurately identify the sulfide signals in the mafic and ultramafic magma systems and is a reliable tool for regional prospecting. More importantly, it can be used to detect intrusion-scale mineralization in the feeder system at depth.

IMPLICATIONS

The proposed random forest algorithm enabled the accurate discriminations of twelve genetic types of olivine. We could determine the component contributions of different mantle sources to the formation of olivine originating from complex sources, such as the interactions between a plume and arc, plume and mid-ocean ridge, or a mid-ocean ridge and continental margin. The contribution provides new opportunities for using olivine chemistry and machine learning to accurately and effectively evaluate the magmatic sulfide fertility. It is a simple and efficient approach for global Ni-Cu-PGE exploration of mafic/ultramafic systems. Our olivine classifier of olivine forming-environments and sulfide mineralization status can be accessed via http://cugb.online:8080/olivine_web/main.html.

Large amounts of data on major-minor-trace elemental and isotopic compositions has been accumulated for various types of whole rocks and minerals in the last several decades. Due to an increased focus on "Big Data," machine learning techniques provide geologists with new tools to tackle problems that are challenging to solve using traditional methods. Our study shows that this technique is powerful in uncovering hidden information from massive data in Earth sciences. Machine learning methods are required to improve our understanding of geosystems and develop low-cost and high-accuracy prediction models for mineral exploration.

ACKNOWLEDGMENTS

This work was funded by the National Key Research and Development Project of China (Grant 2022YFC2903302), the National Natural Science Foundation of China (Grant 42172076), and the Fundamental Research Funds for the Central Universities (Grant 2652022204, 265QZ2022008 and 2-9-2019-043). We thank Yanning Wang, Heng Wang, and Jun Li for their assistance in fieldwork and sampling and Lihui Jia and Hu Guo for guidance in chemical analysis.

REFERENCES

Ali, J.R., Fitton, J.G., and Herzberg, C. (2010) Emeishan large igneous province (SW China) and the mantle-plume up-doming hypothesis. Journal of the Geological Society, 167, 953–959.

Arndt, N., Lesher, M., and Czamanske, G. (2005) Mafic-ultramafic magmas and their relationship to ore formation. Economic Geology 100th Anniversary Volume, 5–24.

Barnes, S.J., and Lightfoot, P.C. (2005) Formation of magmatic nickel-sulfide ore deposits and processes affecting their copper and platinum-group element contents. Economic Geology 100th Anniversary Volume, 179–213.

Barnes, S.J. (1986) The effect of trapped liquid crystallization on cumulus mineral compositions in layered intrusions. Contributions to Mineralogy and Petrology, 93, 524–531.

Barnes, S.J., Godel, B., Gürer, D., Brenan, J.M., Robertson, J., and Paterson, D. (2013) Sulfide-olivine Fe-Ni exchange and the origin of anomalously Ni rich magmatic sulfides. Economic Geology, 108, 1971–1982.

Barnes, S.J., Hill, R.E.T., Perring, C.S., and Dowling, S.E. (2004) Lithogeochemical exploration for komatiite-associated Ni-sulfide deposits: strategies and limitations. Mineralogy and Petrology, 82, 259–293.

Barnes, S.J., Yao, Z.S., Mao, Y.J., Jesus, A.P., Yang, S.H., Taranovic, V., and Maier, W.D. (2022) Nickel in olivine as an exploration indicator for magmatic Ni-Cu sulfide deposits: a data review and re-evaluation. American Mineralogist.

Breiman, L. (2001) Random forests. Machine learning, 45, 5–32.

Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. (2017) Classification and regression trees. Routledge.

Brenan, J.M. (2003) Effects of of f_{O2}, f_{S2}, temperature, and melt composition on Fe-Ni exchange between olivine and sulfide liquid: Implications for natural olivine-sulfide assemblages. Geochimica et Cosmochimica Acta, 67, 2663–2681.

Cameron, E.N. (1976) Post cumulus and subsolidus equilibration of chromite and coexisting silicates in the Eastern Bushveld Complex. In: (Irvine T.N. (ed)) Chromium: its Physicochemical Behavior and Petrologic Significance. Pergamon, 1021–1033.

Cawood, P.A., Wang, W., Zhao, T., Xu, Y., Mulder, J.A., Pisarevsky, S.A., Zhang, L., Gan, C., He, H., Liu, H., Qi, L., Wang, Y., Yao, J., Zhao, G., Zhou, M.F., and Zi, J.W. (2020) Deconstructing South China and consequences for reconstructing Nuna and Rodinia. Earth-Science Reviews, 204.

Chen, C., Yao, Z.S., and Wang, C.Y. (2022) Partitioning behaviors of cobalt and manganese along diverse melting paths of peridotitic and MORB-like pyroxenitic mantle. Journal of Petrology, 63, egac021.

Cheng, L.L., Wang, Y., and Yang, Z.F. (2022) Olivine in picrites from continental flood basalt provinces classified using machine learning. American Mineralogist, 107, 1045–1052.

Coogan, L.A., Hain, A., Stahl, S., and Chakraborty, S. (2005) Experimental determination of the diffusion coefficient for calcium in olivine between 900 °C and 1500 °C.

Geochimica et Cosmochimica Acta, 69, 3683–3694.

Chung, S.L., and Jahn, B.M. (1995) Plume-lithosphere interaction in generation of the Emeishan flood basalts at the Permian-Triassic boundary. Geology, 23, 889–892.

Crocket, J.H. (2002) Platinum-group element geochemistry of mafic and ultramafic rocks. The Geology, Geochemistry and Mineral Beneficiation of Platinum-Group Elements, 54, 177–210.

Cutler, A., Cutler, D.R., and Stevens, J.R. (2012) Random forests. Ensemble machine learning, 157–175. Springer, Boston, MA.

Di Stefano, F., Mollo, S., Blundy, J., Scarlato, P., Nazzari, M., and Bachmann, O. (2019) The effect of CaO on the partitioning behavior of REE, Y and Sc between olivine and melt: Implications for basalt-carbonate interaction processes. Lithos, 326, 327–340.

Dice, L.R. (1945) Measures of the amount of ecologic association between species. Ecology, 26, 297–302.

Efron, B. (1992) Bootstrap methods: another look at the jackknife. Breakthroughs in statistics, 569–593. Springer, New York, NY.

Foley, S.F., Prelevic, D., Rehfeldt, T., and Jacob, D.E. (2013) Minor and trace elements in olivines as probes into early igneous and mantle melting processes. Earth and Planetary Science Letters, 363, 181–191.

Gao, J.F., Zhou, M.F., Lightfoot, P.C., Wang, C.Y., Qi, L., and Sun, M. (2013) Sulfide saturation and magma emplacement in the formation of the Permian Huangshandong Ni-Cu sulfide deposit, Xinjiang, Northwestern China. Economic Geology, 108, 1833–

1848.

Gavrilenko, M., Herzberg, C., Vidito, C., Carr, M.J., Tenner, T., and Ozerov, A. (2016) A calcium-in-olivine geohygrometer and its application to subduction zone magmatism. Journal of Petrology, 57, 1811–1832.

Green, D.H., and Ringwood, A.E. (1967) The genesis of basaltic magmas. Contributions to Mineralogy and Petrology, 15, 103–190.

Gregory, D.D., Cracknell, M.J., Large, R.R., McGoldrick, P., Kuhn, S., Maslennikov, V.V., Baker, M.J., Fox, N., Belousov, I., Figueroa, M.C., Steadman, J.A., Fabris, A.J., and Lyons T.W. (2019) 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. Economic Geology, 114, 771–786.

Herzberg, C. (2011) Identification of source lithology in the Hawaiian and Canary Islands: Implications for origins. Journal of Petrology, 52, 113–146.

Herzberg, C., Vidito, C., and Starkey, N.A. (2016) Nickel-cobalt contents of olivine record origins of mantle peridotite and related rocks. American Mineralogist, 101, 1952–1966.

Howarth, G.H., and Harris, C. (2017) Discriminating between pyroxenite and peridotite sources for continental flood basalts (CFB) in southern Africa using olivine chemistry. Earth and Planetary Science Letters, 475, 143–151.

Jahn, B.M., Wu, F.Y., and Chen, B. (2000) Massive granitoid generation in Central Asia: Nd isotope evidence and implication for continental growth in the Phanerozoic.

Episodes, 23, 82–92.

Kiseeva, E.S., and Wood, B.J. (2015) The effects of composition and temperature on chalcophile and lithophile element partitioning into magmatic sulphides. Earth and Planetary Science Letters, 424, 280–294.

Kohavi, R. (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. Ijcai, 14, 1137–1145.

Kröner, A., Windley, B.F., Badarch, G., Tomurtogoo, O., Hegner, E., Jahn, B.M., Gruschka, S., Khain, E.V., and Wingate, M.T.D. (2007) Accretionary growth and crust formation in the Central Asian Orogenic Belt and comparison with the Arabian-Nubian shield. Geological Society of America Memoirs, 200, 181–209.

Kuwatani, T., Nagata, K., Okada, M., Watanabe, T., Ogawa, Y., Komai, T., and Tsuchiya, N. (2015) Machine-learning techniques for geochemical discrimination of 2011 Tohoku tsunami deposits. Scientific Reports, 4, 7077.

Le Vaillant, M., Fiorentini, M.L., and Barnes, S.J. (2016) Review of lithogeochemical exploration tools for komatiite-hosted nickel sulphide deposits. Journal of Geochemical Exploration, 168, 1–19.

Lehnert, K., Su, Y., Langmuir, C.H., Sarbas, B., and Nohl, U. (2000) A global geochemical database structure for rocks. Geochemistry, Geophysics, Geosystems, 1, 1012.

Li, C., and Ripley, E.M. (2010) The relative effects of composition and temperature on olivine-liquid Ni partitioning: Statistical deconvolution and implications for petrologic modeling. Chemical Geology, 275, 99–104.

Li, C., Naldrett, A.J., and Ripley, E.M. (2007) Controls on the Fo and Ni contents of olivine in sulfide-bearing mafic/ultramafic intrusions: principles, modeling, and examples from Voisey's Bay. Earth Science Frontiers, 14, 177–183.

Li, C., Ripley, E.M., and Tao, Y. (2019) Magmatic Ni-Cu and Pt-Pd Sulfide Deposits in China. Society of Economic Geologists Special Publications, 22, 483–508.

Li, C., and Naldrett, A.J. (1999) Geology and petrology of the Voisey's Bay intrusion: reaction of olivine with sulfide and silicate liquids. Lithos, 47, 1−31.

Li, C., Thakurta, J., and Ripley, E.M. (2012) Low-Ca contents and kink-banded textures are not unique to mantle olivine: evidence from the Duke Island Complex, Alaska. Mineralogy and Petrology, 104, 147–153.

Libourel, G. (1999) Systematics of calcium partitioning between olivine and silicate melt: implications for melt structure and calcium content of magmatic olivines. Contributions to Mineralogy and Petrology, 136, 63–80.

Lösing, M., and Ebbing, J. (2021) Predicting geothermal heat flow in Antarctica with a machine learning approach. Journal of Geophysical Research: Solid Earth, 126, e2020JB021499.

Mao, Y.J., Schoneveld, L., Barnes, S.J., Williams, M.J., Su, B.X., Ruprecht, P., Evans, N.J., and Qin, K.Z. (2022) Coupled Li-P zoning and trace elements of olivine from magmatic Ni-Cu deposits: implications for postcumulus re-equilibration in olivine. Journal of Petrology, 63, egac018.

Matzen, A.K., Baker, M.B., Becket, J.R., and Stopler, E.M. (2013) The temperature and pressure dependence of nickel partitioning between olivine and silicate melt. Journal of Petrology, 54, 2521–2545.

Matzen, A.K., Baker, M.B., Beckett, J.R., Wood, B.J., and Stolper, E.M. (2017a) The effect of liquid composition on the partitioning of Ni between olivine and silicate melt. Contributions to Mineralogy and Petrology, 172, 3.

Matzen, A.K., Wood, B.J., Baker, M.B., and Stolper, E.M. (2017b) The roles of pyroxenite and peridotite in the mantle sources of oceanic basalts. Nature Geoscience, 10, 530–535.

McInnes, L., Healy, J. and Melville, J. (2018) Umap: Uniform manifold approximation and projection for dimension reduction. arXiv preprint arXiv, 1802.03426.

Mysen, B. (2007) Partitioning of calcium, magnesium, and transition metals between olivine and melt governed by the structure of the silicate melt at ambient pressure. American Mineralogist, 92, 844–862.

Naldrett, A.J. (2011) Fundamentals of magmatic sulfide deposits. Reviews in Economic Geology, 17, 1–50.

O'Reilly, S.Y., Chen, D., Griffin, W.L., and Ryan, C.G. (1997) Minor elements in olivine from spinel lherzolite xenoliths: implications for thermobarometry. Mineralogical Magazine, 61, 257–269.

Petrelli, M., and Perugini, D. (2016) Solving petrological problems through machine learning: the study case of tectonic discrimination using geochemical and isotopic data.

Contributions to Mineralogy and Petrology, 171, 1–15.

Petrelli, M., Caricchi, L., and Perugini, D. (2020) Machine learning thermo-barometry: Application to clinopyroxene-bearing magmas. Journal of Geophysical Research: Solid Earth, 125, e2020JB020130.

Rajamani, V., and Naldrett, A.J. (1978) Partitioning of Fe, Co, Ni, and Cu between sulfide liquid and basaltic melts and the composition of Ni-Cu sulfide deposits. Economic Geology, 73, 82–93.

Simkin, T., and Smith, J.V. (1970) Minor-element distribution in olivine. Journal of Geology, 78, 304–325.

Sengör, A.M.C., Natal'in, B.A., and Burtman, V.S. (1993) Evolution of the Altaid tectonic collage and Paleozoic crustal growth in Asia. Nature, 364, 299–307.

Shellnutt, J.G. (2014) The Emeishan large igneous province: a synthesis. Geoscience Frontiers, 5, 369–394.

Simkin, T., and Smith, J.V. (1970) Minor-element distribution in olivine. The Journal of Geology, 78, 304–325.

Sobolev, A.V., Hofmann, A.W., Kuzmin, D.V., Yaxley, G.M., Arndt, N.T., Chung, S.L., Danyushevsky, L.V., Elliott, T., Frey, F.A., Garcia, M.O., Gurenko, A.A., Kamenetsky, V.S., Kerr, A.C., Krivolutskaya, N.A., Matvienkov, V.V., Nikogosian, I.K., Rocholl, A., Sigurdsson, I. A., Sushchevskaya, N.M., and Teklay, M. (2007) The amount of recycled crust in sources of mantle-derived melts. Science, 316, 412–417.

Sobolev, A.V., Hofmann, A.W., Sobolev, S.V., and Nikogosian, I.K. (2005) An olivine-free mantle source of Hawaiian shield basalts. Nature, 434, 590–597.

Song, X.Y., and Li, X.R. (2009) Geochemistry of the Kalatongke Ni-Cu-(PGE) sulfide deposit, NW China: implications for the formation of magmatic sulfide mineralization in a postcollisional environment. Mineralium Deposita, 44, 303–327.

Sorensen, T.A. (1948) A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. Biol. Skar, 5, 1–34.

Stehman, S.V. (1997) Selecting and interpreting measures of thematic classification accuracy. Remote sensing of Environment, 62, 77–89.

Tang, Q., Li, C., Ripley, E.M., Bao, J., Su, T., and Xu, S. (2021) Sr-Nd-Hf-O isotope constraints on crustal contamination and mantle source variation of three Fe-Ti-V oxide ore deposits in the Emeishan large igneous province. Geochimca et Cosmochimica Acta, 292, 364–381.

Ueki, K., Hino, H., and Kuwatani, T. (2018) Geochemical discrimination and characteristics of magmatic tectonic settings: A machine learning-based approach. Geochemistry, Geophysics, Geosystems, 19, 1327–1347.

Wang, Y., Qiu, K.F., Müller, A., Hou, Z.L., Zhu, Z.H., and Yu, H.C. (2021) Machine Learning Prediction of Quartz Forming-Environments. Journal of Geophysical Research: Solid Earth, 126(8), e2021JB021925.

Wang, Y.N., Xue, S.C., Klemd, R., Yang, L., Zhao, F., and Wang, Q.F. (2022) Late Permian plume and Neoproterozoic subduction-modified mantle interaction: Insights from geochronology and Sr-Nd-O isotopes of mafic dikes of the western Emeishan Large Igneous Province. American Journal of Science, 322, 993–1018.

Wei, B., Wang, C.Y., Lahaye, Y., Xie, L.H. and Cao, Y.H. (2019) S and C isotope constraints for mantle-derived sulfur source and organic carbon-induced sulfide saturation of magmatic Ni-Cu sulfide deposits in the Central Asian Orogenic Belt, North China. Economic Geology, 114, 787–806.

Wei, B., Wang, C., Li, C., and Sun, Y. (2013) Origin of PGE-depleted Ni-Cu sulfide mineralization in the Triassic Hongqiling No. 7 orthopyroxenite intrusion, Central Asian Orogenic Belt, northeastern China. Economic Geology, 108, 1813–1831.

Wessel, P., Smith, W.H., Scharroo, R., Luis, J., and Wobbe, F. (2013) Generic mapping tools: improved version released. Eos, Transactions American Geophysical Union, 94, 409–410.

Windley, B.F., Alexeiev, D., Xiao, W.J., Kroner, A., and Badarch, G. (2007) Tectonic models for accretion of the Central Asian Orogenic Belt. Journal of the Geological Society of London, 164, 31–47.

Wolpert, D.H., and Macready, W.G. (1999) An efficient method to estimate bagging's generalization error. Machine Learning, 35, 41–55.

Xu, Y.G., He,B., Chung, S.L., Menzies, M.A., and Frey, F.A. (2004) Geologic, geochemical, and geophysical consequences of plume involvement in the Emeishan flood-basalt province. Geology, 32, 917–920.

Xue, S.C., Deng, J., Wang, Q.F., Xie, W., and Wang, Y.N. (2021) The redox conditions and C isotopes of magmatic Ni-Cu sulfide deposits in convergent tectonic settings: the role of reduction process in ore genesis. Geochimica et Cosmochimica Acta, 306, 210–225.

Xue, S.C., Wang, Q.F., Deng, J., Wang, Y.N., and Peng, T.P. (2022) Mechanism of organic matter assimilation and its role in sulfide saturation of oxidized magmatic ore-forming system: insights from C-S-Sr-Nd isotopes of the Tulaergen deposit in NW China. Mineralium Deposita, 1–19.

Xue, S.C., Li, C., Wang, Q.F., Ripley, E.M., and Yao, Z.S. (2019) Geochronology, petrology and Sr-Nd-Hf-S isotope geochemistry of the newly-discovered Qixin magmatic Ni-Cu sulfide prospect, southern Central Asian Orogenic Belt, NW China. Ore Geology Reviews, 111, 103002.

Xue, S.C., Qin, K.Z., Li, C., Tang, D.M., Mao, Y.J., Qi, L., and Ripley, E.M. (2016) Geochronological, petrological, and geochemical constraints on Ni-Cu sulfide mineralization in the Poyi ultramafic-troctolitic intrusion in the northeast rim of the Tarim craton, western China. Economic Geology, 111, 1465–1484.

Yang, S.H., Maier, W.D., Hanski, E.J., Lappalainen, M., Santaguida, F., and Määttä, S. (2013) Origin of ultra-nickeliferous olivine in the Kevitsa Ni-Cu-PGE-mineralized intrusion, northern Finland. Contributions to Mineralogy and Petrology, 166, 81–95.

Yao, J., Cawood, P.A., Shu, L.S., and Zhao, G. (2019) Jiangnan Orogen, South China: A ~970–820 Ma Rodinia margin accretionary belt. Earth-Science Reviews, 196.

Yao, Z., Qin, K.Z., and Mungall, E. (2018) Tectonic controls on Ni and Cu contents of primary mantle-derived magmas for the formation of magmatic sulfide deposits. American Mineralogist, 18–106392.

Yu, S.Y., Shen, N.P., Song, X.Y., Ripley, E.M., Li, C., and Chen, L.M. (2017) An integrated chemical and oxygen isotopic study of primitive olivine grains in picrites from the Emeishan large Igneous Province, SW China: evidence for oxygen isotope heterogeneity in mantle sources. Geochimca et Cosmochimica Acta, 215, 263–276.

Zanetti, A., Tiepolo, M., Oberti, R., and Vannucci, R. (2004) Trace-element partitioning in olivine: modelling of a complete data set from a synthetic hydrous basanite melt. Lithos, 75, 39–45.

Zhang, Z., Mahoney, J.J., Mao, J., and Wang, F. (2006) Geochemistry of picritic and associated basalt flows of the western Emeishan flood basalt province, China. Journal of Petrology, 47, 1997–2019.

Zou, S., Chen, X., Brzozowski, M.J., Leng, C.B., and Xu, D. (2022) Application of machine learning to characterizing magma fertility in porphyry Cu deposits. Journal of Geophysical Research: Solid Earth, 127, e2022JB024584.

FIGURE CAPTIONS

Figure 1. Distribution of samples of global olivine-bearing volcanic and plutonic rocks used in this study. This map was created using the Generic Mapping Tool (GMT) package developed by Wessel et al., 2013, https://www.generic-mapping-tools.org/)-derived topographic data.

Figure 2. Scatterplots of Fo values and abundances of Ni (a), Fe/Ni (b), Ca (c), and Mn (d) in global volcanic and plutonic olivines. Overlaps exist in the olivine populations of the twelve genetic types.

Figure 3. Flowchart of the random forest algorithm.

Figure 4. Relative feature importance of the geochemical features of olivine obtained from Random forest using training data.

Figure 5. Confusion matrix of the olivine classification results for the formation environment. The rows show the predicted label, and the columns show the true label. A darker color indicates higher accuracy.

Figure 6. Confusion matrix of olivine classification results for mineralization status.

Figure 7. Confusion matrix of olivine classification results for (a) volcanic and

intrusive status and (b) petrography group.

Figure 8. Two-dimensional decision boundaries derived from UMAP and random forest.

Figure 9. (a) The predicted mantle sources of late Permian picrites and basalts in the western margin of the Yangtze block derived from the random forest model. (b) Plot of initial $^{87}Sr/^{86}Sr$ isotopic ratios vs. $\varepsilon_{Nd}(t)$ values for the mafic volcanic rocks and dykes in the western Yangtze block with two different sources. The mantle array is from DePaolo and Wasserburg (1979). (c) The predicted magmatic sulfide potential of the Qixin mafic-ultramafic complex at the southern margin of the CAOB derived from the random forest model. (d) Photographs and polished section microphotographs of the disseminated sulfide ore in the Qixin ultramafic intrusion. Abbreviation: Sulf, sulfide; Po, pyrrhotite; Pn, pentlandite; Ccp, chalcopyrite.

Appendix Figure 1. Correlation coefficient matrix for features of olivine.

Table 1 Summary of model performance under different classification criteria.

Classification criteria	Precision	Recall	F1 score	Accuracy
Forming-environment	0.99	0.99	0.99	0.99
Mineralization Status	0.99	0.99	0.99	0.99
Volcanic and Intrusive Status	1	1	1	1
Petrography Group	0.99	0.99	0.99	0.99

Note: The Precision, Recall and F1 score in the table are all weighted averages, which represent the cumulative sum of the proportion of the class samples in the total samples and the product of the corresponding indicators.

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8893. http://www.minsocam.org/

Table 2 The classification accuracy of decision boundary based on UMAP under different classification criteria.

Classification criteria	Accuracy
Forming-environment	0.88
Mineralization Status	0.9
Volcanic and Intrusive Status	0.94
Petrography Group	0.96

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8893. http://www.minsocam.org/

Figure 1

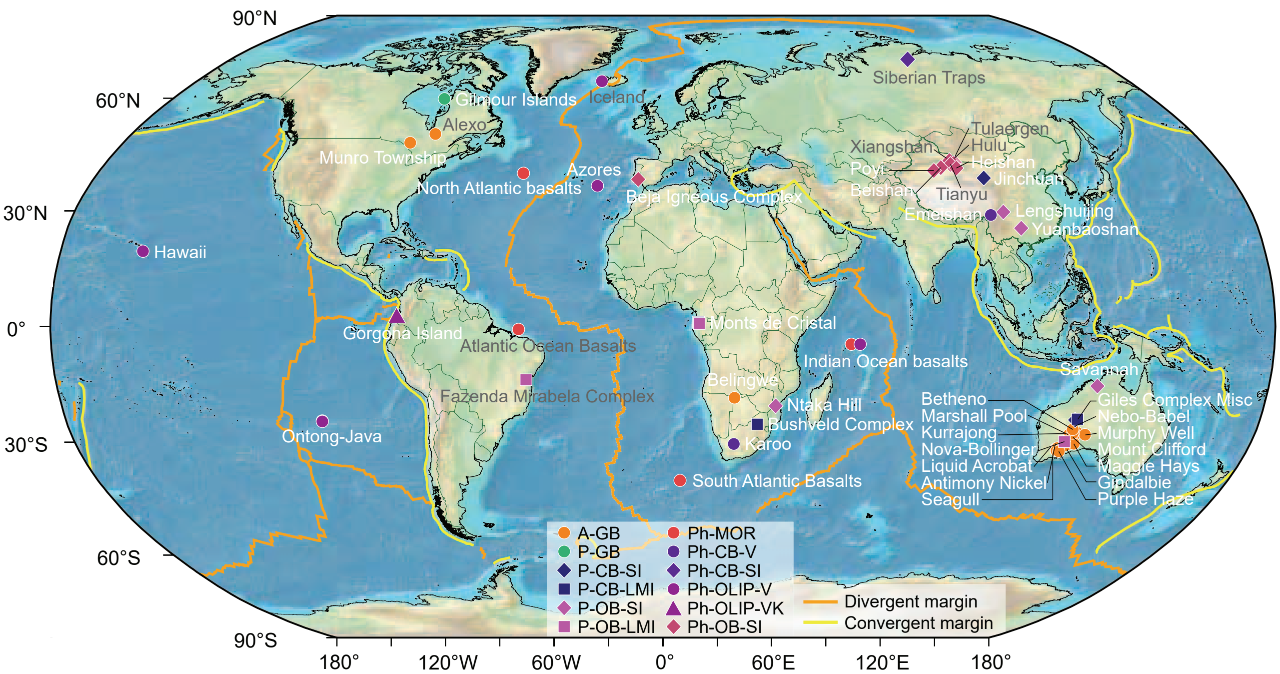
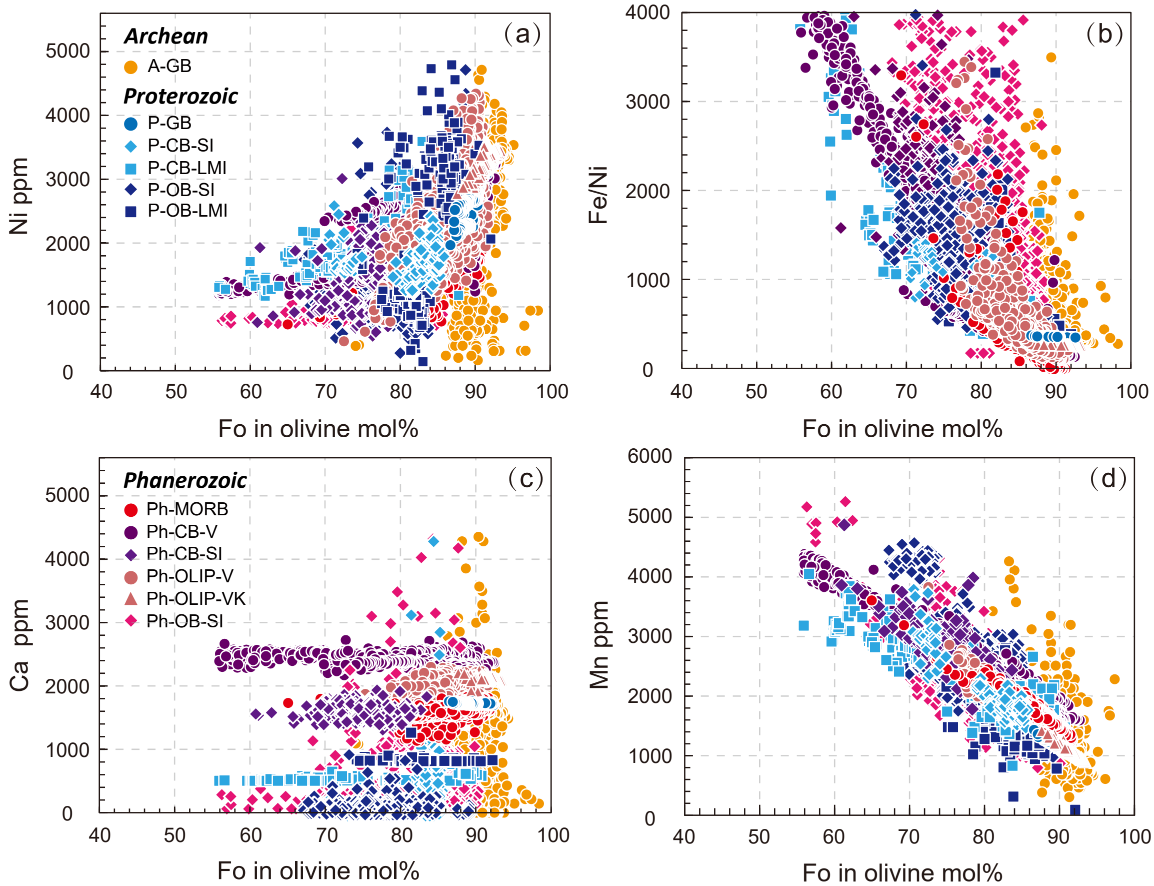
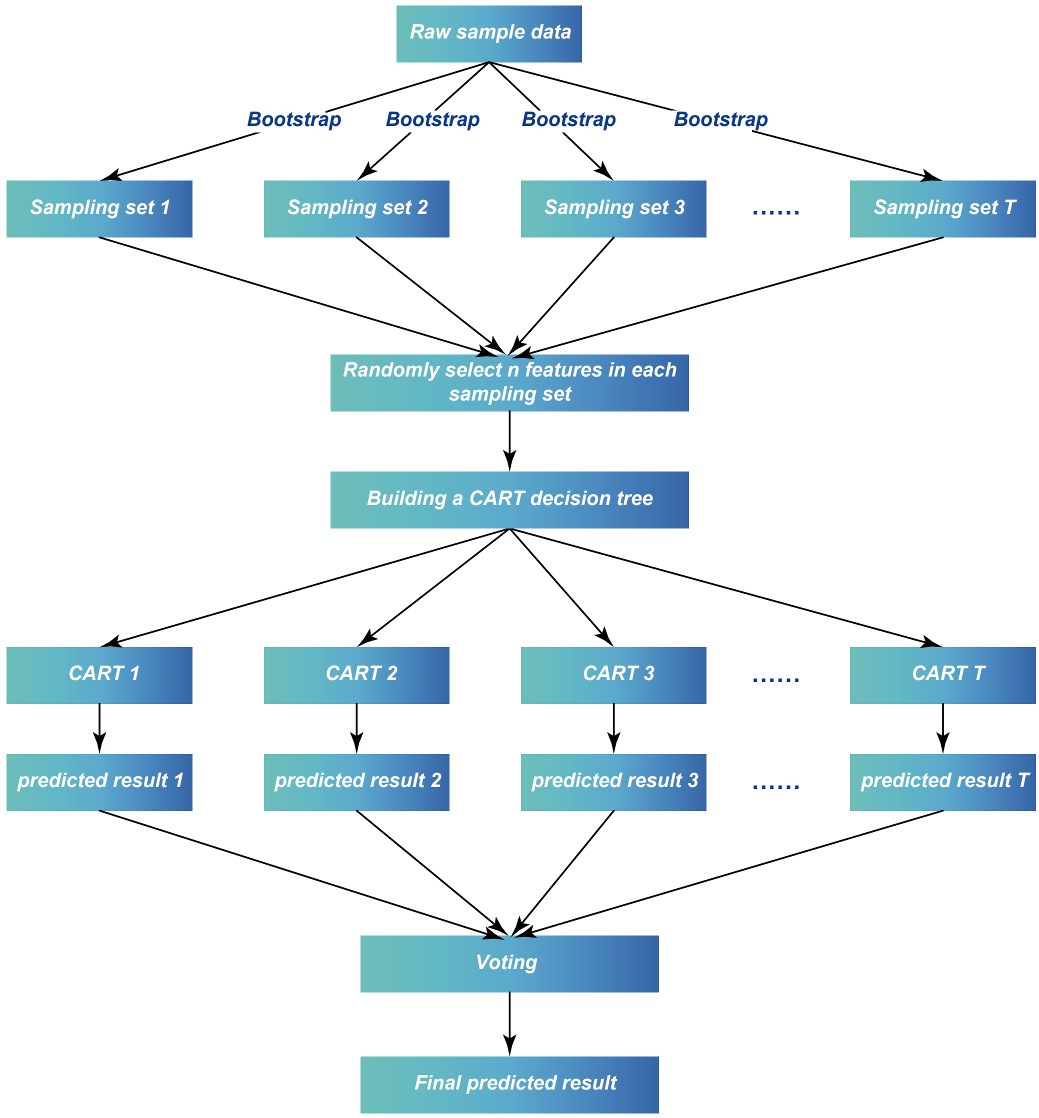


Figure 2


(a)
Archean
A-GB
Proterozoic
P-GB
P-CB-SI
P-CB-LMI
P-OB-SI
P-OB-LMI
Ni ppm
Fo in olivine mol%
(b)
Fe/Ni
Fo in olivine mol%
(c)
Phanerozoic
Ph-MORB
Ph-CB-V
Ph-CB-SI
Ph-OLIP-V
Ph-OLIP-VK
Ph-OB-SI
Ca ppm
Fo in olivine mol%
(d)
Mn ppm
Fo in olivine mol%

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8893. http://www.minsocam.org/

Figure 3



This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8893. http://www.minsocam.org/

Figure 4

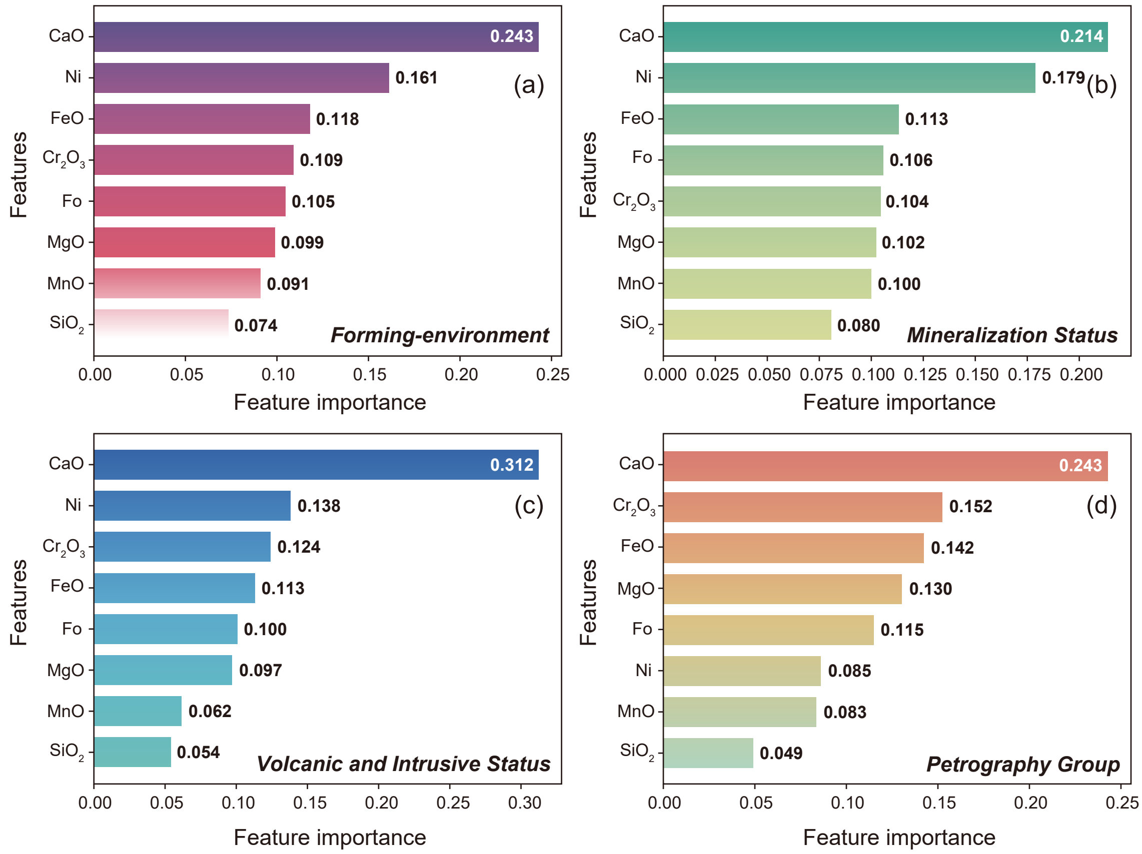
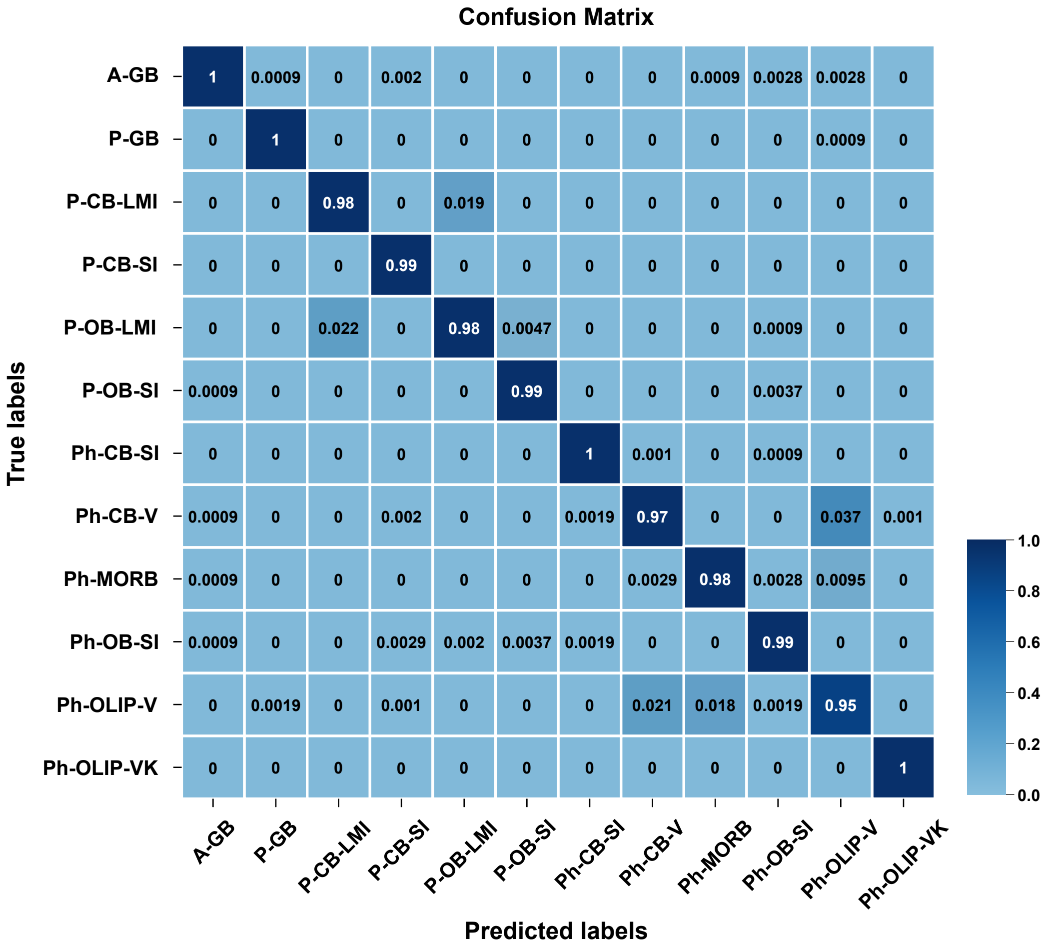


Figure 5


Confusion Matrix
True labels
Predicted labels
A-GB
P-GB
P-CB-LMI
P-CB-SI
P-OB-LMI
P-OB-SI
Ph-CB-SI
Ph-CB-V
Ph-MORB
Ph-OB-SI
Ph-OLIP-V
Ph-OLIP-VK
1 0.0009 0 0.002 0 0 0 0 0.0009 0.0028 0.0028 0
0 1 0 0 0 0 0 0 0 0 0.0009 0
0 0 0.98 0 0.019 0 0 0 0 0 0 0
0 0 0 0.99 0 0 0 0 0 0 0 0
0 0 0.022 0 0.98 0.0047 0 0 0 0.0009 0 0
0.0009 0 0 0 0 0.99 0 0 0 0.0037 0 0
0 0 0 0 0 0 1 0.001 0 0.0009 0 0
0.0009 0 0 0.002 0 0 0.0019 0.97 0 0 0.037 0.001
0.0009 0 0 0 0 0 0 0.0029 0.98 0.0028 0.0095 0
0.0009 0 0 0.0029 0.002 0.0037 0.0019 0 0 0.99 0 0
0 0.0019 0 0.001 0 0 0 0.021 0.018 0.0019 0.95 0
0 0 0 0 0 0 0 0 0 0 0 1
1.0
0.8
0.6
0.4
0.2
0.0

This is the peer-reviewed, final accepted version for American Mineralogist, published by the Mineralogical Society of America.
The published version is subject to change. Cite as Authors (Year) Title. American Mineralogist, in press.
DOI: https://doi.org/10.2138/am-2022-8893. http://www.minsocam.org/

Figure 6

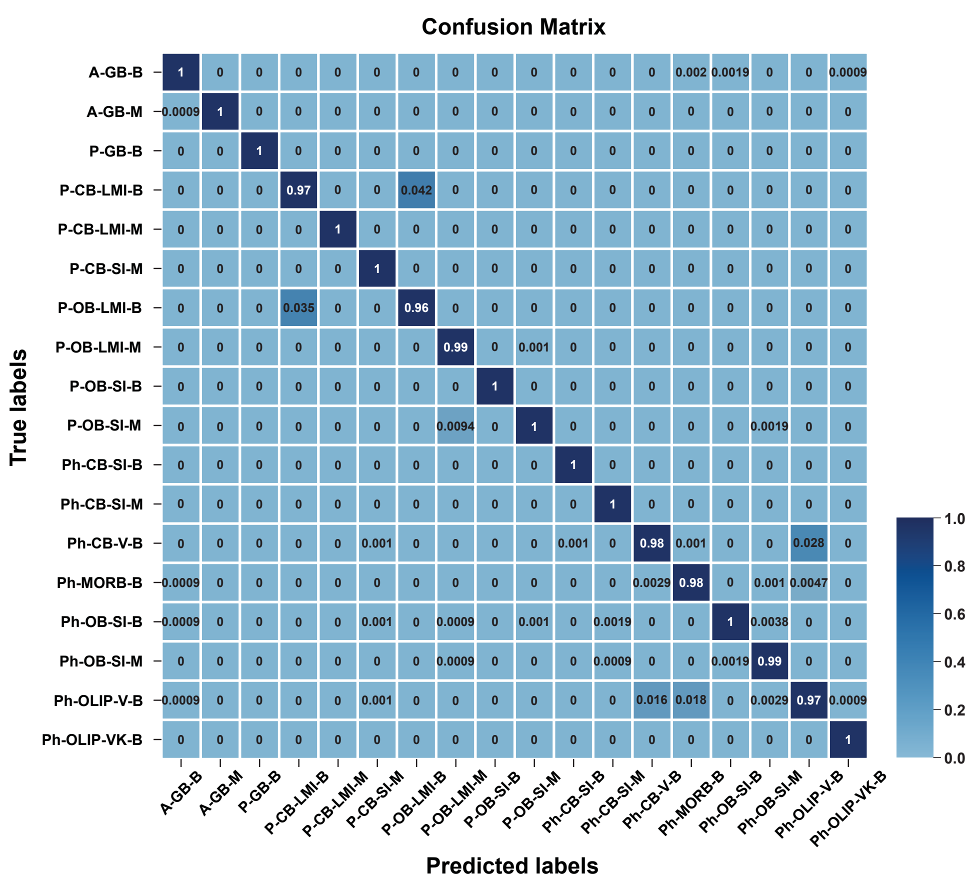


Figure 7

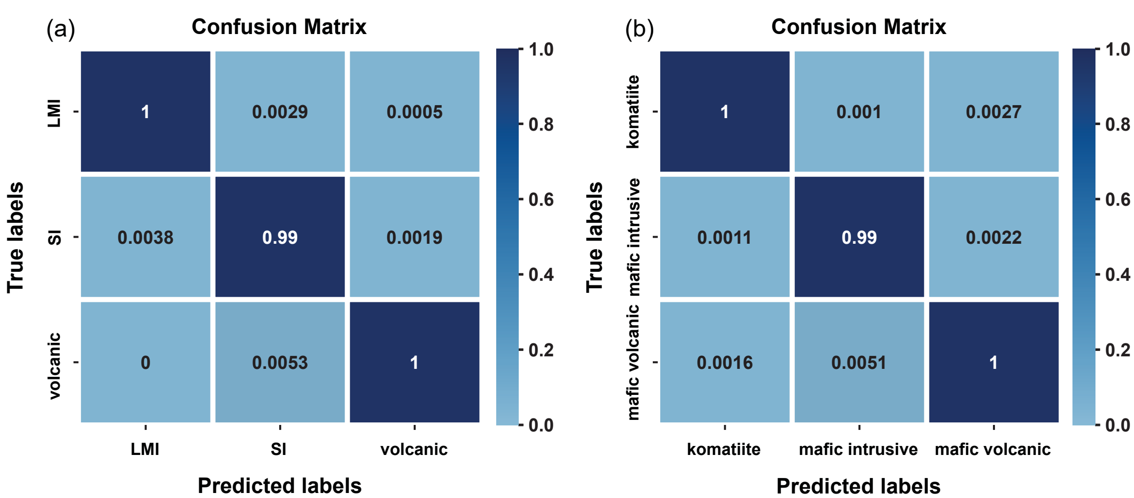
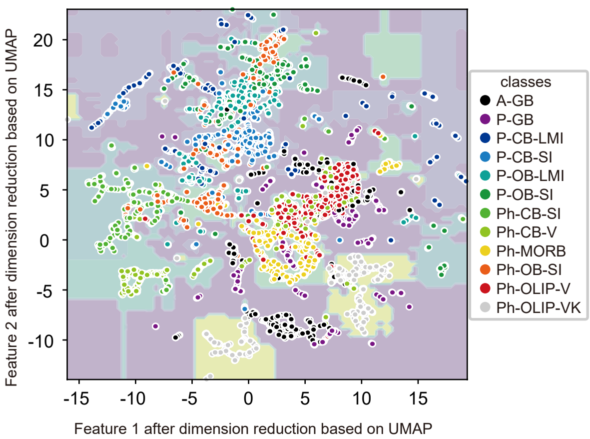
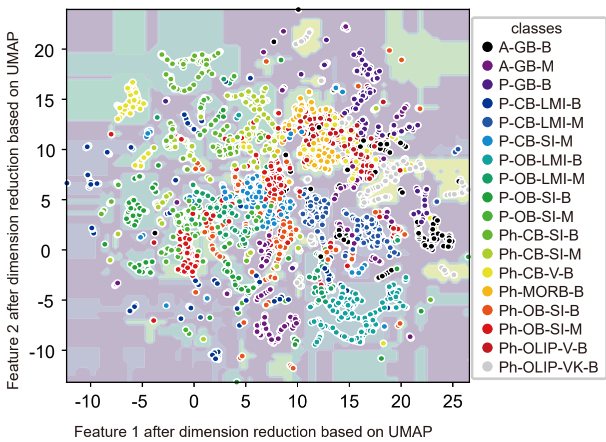

(a)
Confusion Matrix
True labels
LMI
SI
volcanic
1
0.0029
0.0005
0.0038
0.99
0.0019
0
0.0053
1
LMI
SI
volcanic
Predicted labels
1.0
0.8
0.6
0.4
0.2
0.0
(b)
Confusion Matrix
True labels
komatiite
mafic intrusive
mafic volcanic
1
0.001
0.0027
0.0011
0.99
0.0022
0.0016
0.0051
1
komatiite
mafic intrusive
mafic volcanic
Predicted labels
1.0
0.8
0.6
0.4
0.2
0.0

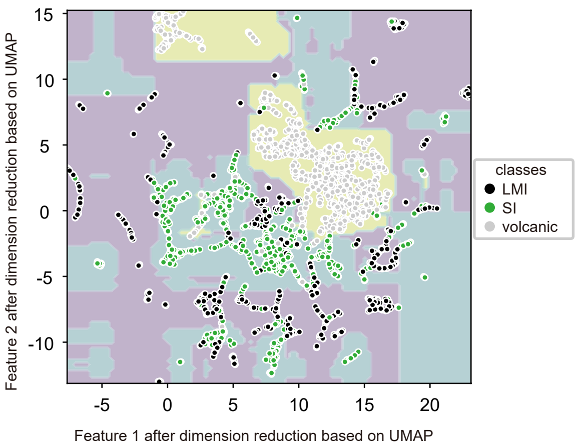
Figure 8



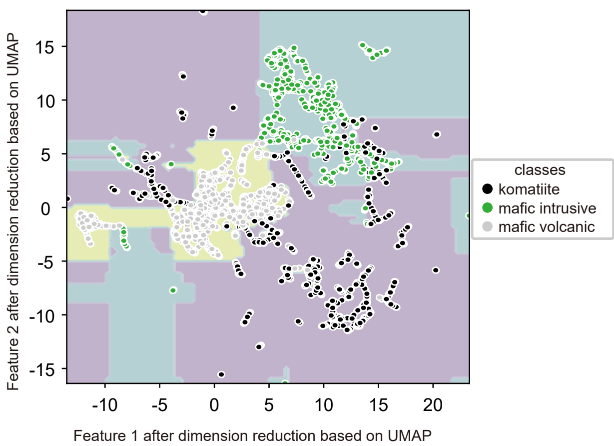
(a) Forming Environments



(b) Mineralization Status

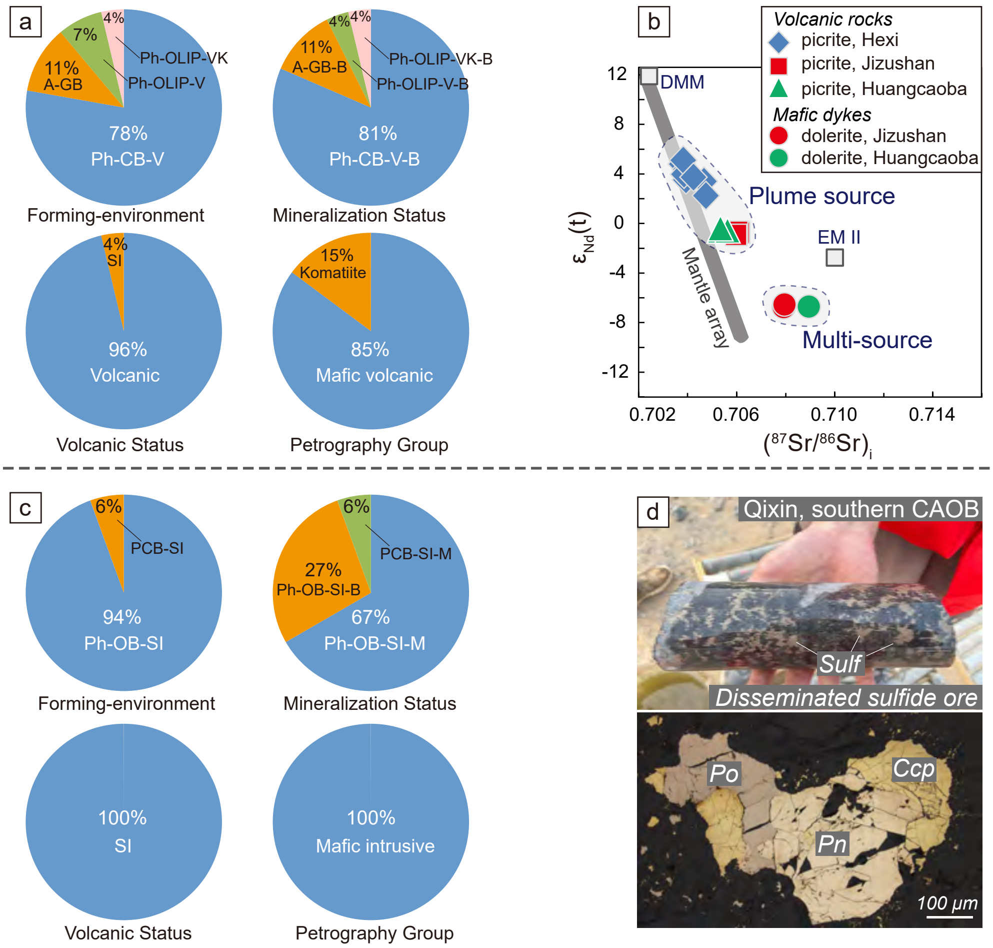


(c) Volcanic and Intrusion Status



(d) Petrography Groups

Figure 9


a
78%
Ph-CB-V
11%
A-GB
7%
4%
Ph-OLIP-VK
Ph-OLIP-V
Forming-environment
81%
Ph-CB-V-B
11%
A-GB-B
4%
4%
Ph-OLIP-VK-B
Ph-OLIP-V-B
Mineralization Status
96%
Volcanic
4%
SI
Volcanic Status
85%
Mafic volcanic
15%
Komatiite
Petrography Group
b
Volcanic rocks
picrite, Hexi
picrite, Jizushan
picrite, Huangcaoba
Mafic dykes
dolerite, Jizushan
dolerite, Huangcaoba
DMM
Plume source
EM II
Mantle array
Multi-source
εNd(t)
(87Sr/86Sr)i
c
94%
Ph-OB-SI
6%
PCB-SI
Forming-environment
67%
Ph-OB-SI-M
27%
Ph-OB-SI-B
6%
PCB-SI-M
Mineralization Status
100%
SI
Volcanic Status
100%
Mafic intrusive
Petrography Group
d
Qixin, southern CAOB
Sulf
Disseminated sulfide ore
Po
Ccp
Pn
100 μm