Trace Element Composition of Chalcopyrite from Volcanogenic Massive Sulfide Deposits: Variation and Implications for Provenance Recognition

Enzo Caraballo,[1,†] Georges Beaudoin,[1] Sarah Dare,[2] Dominique Genna,[2,3] Sven Petersen,[4]
Jorge M.R.S. Relvas,[5] and Stephen J. Piercey[6]

[1] *Département de géologie et génie géologique, Université Laval, 1065 avenue de la Médecine, Québec, Québec G1V 0A6, Canada*

[2] *Département des Sciences Appliquées, Université du Québec à Chicoutimi, 555, boul. de l'Université, Saguenay, Québec G7H 2B1, Canada*

[3] *CONSOREM, 555, boul. de l'Université, Saguenay, Québec G7H 2B1, Canada*

[4] *Helmholtz Centre for Ocean Research Kiel, Wischhofstr 1-3, 24148 Kiel, Germany*

[5] *Instituto Dom Luiz, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisbon, Portugal*

[6] *Memorial University of Newfoundland, 300 Prince Philip Drive, St. John's, Newfoundland and Labrador A1B 3X5, Canada*

Abstract

Chalcopyrite from 51 volcanogenic massive sulfide (VMS) and sea-floor massive sulfide (SMS) deposits from six lithostratigraphic settings was analyzed for trace elements by laser ablation-inductively coupled plasma-mass spectrometry (LA-ICP-MS) to evaluate its potential as an indicator mineral for exploration. Partial least squares discriminant analysis (PLS-DA) results reveal that chalcopyrite from different lithostratigraphic settings has different compositions reflecting host rock assemblages and fluid composition. Three random forest (RF) classifiers were developed to distinguish chalcopyrite from the six lithostratigraphic settings with a divisive approach. This method, which primarily classifies according to the major host-rock affinity and subsequently according to VMS settings, yielded an overall accuracy higher than 0.96 on test data. The model validation with literature data having the same elements required by the models yielded the highest accuracies (>0.90). In validation using published data with missing elements, the accuracy is moderate to high (0.60–1); however, the performances decrease significantly (<0.50) when the most important elements are missing. Similarly, RF regression models developed using all sets of analyzed elements to determine ccp/(ccp + sp) ratio (ccp = chalcopyrite; sp = sphalerite) in chalcopyrite within a single VMS setting reported high performances, thus showing a potential to predict the Cu/Zn ratio (Cu-rich vs. Zn-rich) of the mineralization based on chalcopyrite composition. This study demonstrates that trace element concentrations in chalcopyrite are primarily controlled by lithotectonic setting and can be used as predictors in an RF classifier to distinguish the different VMS subtypes.

Introduction

Chalcopyrite occurs in numerous mineral deposit types, from magmatic to hydrothermal systems. Its trace element composition is highly variable, as shown by several studies of magmatic Ni-Cu-platinum group element (PGE) deposits (Czamanske et al., 1992; Barnes et al., 2008; Godel and Barnes, 2008; Dare et al., 2010, 2011, 2014; Piña et al., 2012; Duran et al., 2016; Mansur et al., 2020; Caraballo et al., 2022) and hydrothermal environments (Hawley and Nichol, 1961; Rose, 1967; Harris et al., 1984; Cabri et al., 1985; Bajwah et al., 1987; Kase, 1987; Huston et al., 1995; Butler and Nesbitt, 1999; Moggi-Cecchi et al., 2002; Cook et al., 2011; Reich et al., 2013b; Wang et al., 2015; Wohlgemuth-Ueberwasser et al., 2015; Agangi et al., 2018; George et al., 2018b). The crystal structure, partitioning with cocrystallizing minerals, and magma/fluid composition are the principal factors controlling the trace element content of chalcopyrite (George et al., 2016; Caraballo et al., 2022). Thus, the trace element composition of chalcopyrite can be used for provenance studies (Hashmi et al., 2015; Duran et al., 2019; Caraballo et al., 2022).

An indicator mineral provides evidence on the presence of a type of mineralization, hydrothermal alteration, or lithology in bedrock; their morphology, textures, and/or geochemistry may be investigated in order to obtain information about transport and nature of mineralization (McClenaghan, 2005). The characteristics to consider a mineral as indicator in exploration are (1) a specific gravity relatively high (G>3.2) to be separated from sample media; (2) distinctive visual properties with minimal size of about 0.25 mm to be recuperated efficiently from sediments (e.g., till, soils, fluvial); (3) high resistance to weathering (e.g., oxidation) and/or mechanical transport; and (4) a wide range of trace element compositions that is sensitive to the environment of crystallization, including important genetic information about mineralization and host rocks (Averill, 2001, 2011; McClenaghan, 2005). With a specific gravity of 4.1 to 4.3, chalcopyrite is easily separated by density, and its physical properties enable recognition by optic microscope. In addition, chalcopyrite is among the sulfides most resistant to dissolution in the weathering environment, and several hypotheses have been proposed to explain the mechanisms leading to slow dissolution kinetics (Rimstidt et al., 1993; Crundwell et al., 2021). Averill (2011), based on stability in surficial sediments under temperate and humid tropical conditions, and Crundwell (2021), based on the electronic structure of the most common sulfides, proposed a relative scale of dissolution resistance/stability, which places chalcopyrite as one of the most resistant sulfides.

[†]Corresponding author: e-mail, enzo.caraballo-rojas.1@ulaval.ca

1

Moreover, the potential of chalcopyrite as an indicator mineral is highlighted by several studies on heavy mineral fractions in surficial sediments from glacial environments. Averill (2001 and 2011) showed that chalcopyrite can be used as an indicator mineral of volcanogenic massive sulfide (VMS), skarn, and Ni-Cu-PGE, since it is marginally stable in till sediments. Hashmi et al. (2015) demonstrated in till sediments from Mont Polley (British Columbia, Canada) that chalcopyrite indicates the presence of Cu-Au porphyry mineralization. Plouffe et al. (2016) showed that anomalies of chalcopyrite extend over 30 km² in till sediments associated with Cu-Mo porphyry mineralization in Gibraltar (British Columbia, Canada). Duran et al. (2019) showed that the 0.25- to 1-mm nonferromagnetic heavy mineral concentrates of Quaternary till and esker samples from the Churchill Province (Québec, Canada) contain thousands of chalcopyrite and pyrite grains and that their trace element compositions may be used to determine provenance. According to its physical properties, abundance and stability in surficial sediments, and ability to host a relatively wide variety of trace elements, chalcopyrite can record chemical variation associated with crystallization and composition of host rocks. Therefore, chalcopyrite can be used as an indicator for mineral exploration.

Multivariate statistical analysis, such as principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA), and machine learning methods such as random forest (RF) are useful to investigate the chemical variations of large data sets of trace elements in indicator minerals such as gold (Liu and Beaudoin, 2021; Liu et al., 2021), pyrite (Gregory et al., 2019), magnetite (Makvandi et al., 2016a, b; Duran et al., 2020; Bédard et al., 2022), scheelite (Manéglia et al., 2018; Grzela et al., 2019; Sciuba et al., 2020; Miranda et al., 2022), sphalerite (Frenzel et al., 2016; Zhang et al., 2022), rutile (Sciuba and Beaudoin, 2021), and tourmaline (Sciuba et al., 2021). However, studies on chalcopyrite using this approach are scarce, and previous work relied more on empirical discrimination diagrams (George et al., 2018b; Duran et al., 2019). Bédard et al. (2017) investigated trace element compositions of chalcopyrite from different ore deposit types using principal component analysis (PCA) and demonstrated its potential for recognizing chalcopyrite from different deposit types. They showed that trace element signatures discriminate chalcopyrite from volcanogenic massive sulfide (Ag, Sn, Zn, and Se), Ni-Cu-PGE (Ni, Te, and Pd), orogenic gold and red beds (Au and As), porphyry and skarn (Sb, In, and Se), and iron oxide-copper-gold (IOCG; Zn and Ba) deposits. Caraballo et al. (2022) demonstrated, using PLS-DA, that chalcopyrite from Ni-Cu sulfide and reef-type PGE deposits can be distinguished using Bi, In, Se, Sn, and Te. Furthermore, they showed that in Ni-Cu sulfide deposits, it was possible to discriminate chalcopyrite according to ore type (Cu-rich vs. Fe-rich) using the assemblage Sb, Se, Sn, Tl, and Zn. Based on cluster analysis and classification trees approach on trace elements in chalcopyrite from a single Ni-Cu-PGE deposit (Norilsk-Talnakh), Marfin et al. (2020) showed that there are chemical differences between massive vs. disseminated ore types. This study focuses on another single deposit type, VMS, to determine the intradeposit-scale variation of chalcopyrite composition according to both ore type and lithostratigraphic setting.

Volcanogenic massive sulfides occur as ancient accumulations of sulfide minerals (sphalerite, chalcopyrite, pyrite ± pyrrhotite, galena) that precipitated at or below the sea floor from convective hydrothermal fluids, in spatial relationship with volcanosedimentary sequences, and usually coeval with volcanic host rocks (Lydon, 1988; Barrie and Hannington, 1999; Franklin et al., 2005; Hannington, 2014). They are important sources of Zn, Cu, Pb, Ag, and Au, and of critical and strategic metal by-products such as Co, Sn, Se, Mn, Cd, In, Bi, Te, Ga, and Ge (Galley et al., 2007). The modern equivalent of VMS deposits are sea-floor massive sulfides (SMS), which form in active submarine hydrothermal systems, where high-temperature (≈350°C) black smoker vents are the most recognizable features and occur in extensional environments, principally at mid-ocean ridges, in back-arc basins, and along submarine volcanic arcs (Hannington et al., 2005; Fuchs et al., 2019).

This deposit type typically comprises a mound-shaped to tabular, stratabound body of massive sulfides (>60% sulfides) with altered silicates, with some underlain by a discordant stockwork or stringer veins composed of Cu-rich sulfides (Lydon, 1988; Barrie and Hannington, 1999; Galley et al., 2007; Hannington, 2014). In some VMS deposits, a zonation is present, with a systematic decrease of chalcopyrite/(sphalerite + galena) ratio from the inner, underlying stockwork zones to the outer margins (Lydon, 1988; Galley et al., 2007; Hannington, 2014). This zoning is principally a result of several high-temperature events, with replacement of early sulfides (Zn-Pb–rich) by Cu-rich ores as the system increased temperature in the lower parts and a reprecipitation of Zn-Pb minerals in the outer zones of the growing orebody (Eldridge et al., 1983; Ohmoto, 1996). The spatial distribution of these minerals could also be affected by the cooling path and the solubility of chalcopyrite and sphalerite, precipitating initially chalcopyrite and then sphalerite (Hannington et al., 1995; Hannington, 2014). Chalcopyrite precipitates at relatively high temperatures, between 270° and 360°C, at the base, whereas sphalerite is deposited at lower temperatures (175°–250°C) in the outer margins (Lydon, 1988; Hannington et al., 1995; Franklin et al., 2005; Hannington, 2014). In conductive cooling processes, pH variation in the fluid is limited, which could inhibit precipitation of sphalerite and trigger chalcopyrite precipitation at the base, whereas in a mixing cooling history, the pH variation is greater, leading to cocrystallization of chalcopyrite and sphalerite (Hannington et al., 1995). In hydrothermal systems, the behavior of elements is controlled principally by the stability of aqueous complexes of chlorine and sulfur (Hannington, 2014; Seward et al., 2014). Hannington (2014) indicates that there are two principal trace element associations according to temperature: (1) a low-temperature assemblage, comprising Zn, Pb, Ag, Cd, Sn, Sb, As, Hg, ±Tl, and ±W, in the outer margin; and (2) a higher-temperature association of Cu, Co, Bi, Se, In, ±Ni, and ±Mo in the deeper roots of the system. The composition of chalcopyrite is also determined by partitioning in cocrystallizing sulfides such as pyrite, sphalerite, and galena (Genna and Gaboury, 2015; George et al., 2016, 2018b).

Although several studies have analyzed chalcopyrite from VMS-SMS (sea-floor massive sulfide) deposits (e.g., Hawley and Nichol, 1961; George et al., 2018b; Wang et al., 2018;

Martin et al., 2019), only a few have been focused on variability of the trace element composition at deposit scale (Butler and Nesbitt, 1999; Revan et al., 2014; Maslennikov et al., 2017; Torró et al., 2022). Torró et al. (2022) showed that the increase of Se and Bi contents in chalcopyrite from Zn-rich to Cu-rich zones suggests an increase in temperature at the Sofía-D massive sulfide body in the María Teresa VMS deposit (Perú). Furthermore, at chimney scale, the same behavior is observed in the Broken Spur SMS deposit (Mid-Atlantic Ridge; Butler and Nesbitt, 1999), at Çayeli¸-Kutlular (Turkey; Revan et al., 2014), and in Urals deposits (Russia; Maslennikov et al., 2009, 2017), where chalcopyrite from inner zones is higher in Se than that from outer parts of the chimney, likewise supporting the model of temperature-controlled precipitation. However, more studies are needed to fully understand the chemical behavior of chalcopyrite in VMS deposits.

Based on the volcanosedimentary assemblage within a district, delimited by major time breaks, faults or major intrusions, and metal proportions and associated with specific tectonic settings, VMS deposits are classified into five lithostratigraphic types (Barrie and Hannington, 1999; Franklin et al., 2005). Furthermore, the sulfide mineralogy and abundance vary according to the composition of ore-forming hydrothermal fluids and the interaction with the underlying rocks (Hannington et al., 2005; Shanks et al., 2012; Hannington, 2014), such that mafic-hosted deposits are typically Cu rich, whereas bimodal-felsic deposits are higher in Zn, and siliciclastic-felsic deposits are higher in Pb (Galley et al., 2007; Hannington, 2014). The five VMS lithostratigraphic types are as follows: (1) mafic, associated with ophiolite assemblages in mature intraoceanic back-arc basins, dominated by mid-ocean ridge basalt (MORB) successions; (2) siliciclastic-mafic, related to mature back-arc settings in basalt-pelite or pelite-dominant sequences; (3) bimodal-mafic, associated with incipient-rifted bimodal volcanic arcs and basalt-dominated with up to 25% felsic volcanic strata; (4) bimodal-felsic, in continental margin arcs and related back-arcs, with felsic volcanics up to 35 to 75% and basalt 20 to 50%; and (5) siliciclastic-felsic, associated with mature epicontinental back-arcs and dominated by siliciclastic rocks (~80%). Patten et al. (2022) suggested there is also an ultramafic-hosted subgroup of VMS deposits, given their unique lithostratigraphic setting and metal endowment (e.g., Cu, Zn, Co, Ni). They form principally in slow- to ultra-slow–spreading mid-ocean ridges with limited effusive magmatic activity. Although chalcopyrite is present in all the subtypes of VMS deposits, there is not a systematic study that has compared the signatures between these different settings. Therefore, the ability to predict the VMS lithostratigraphic setting using chalcopyrite as an indicator mineral from till or river sediments at an early exploration stage could be an important tool to guide exploration strategies in covered terrains, in addition to geophysical techniques.

This study aims to evaluate the potential of chalcopyrite chemistry from ancient VMS deposits (Archean to Tertiary) and sea-floor massive sulfide (SMS) deposits to be used as an indicator mineral for exploration. We systematically investigated the variation of 36 trace elements in chalcopyrite from worldwide VMS-SMS deposits as a function of the deposit lithostratigraphic subtype (Franklin et al., 2005; Patten et al., 2022), ore type (massive/stockwork), and sulfide proportions.

We show that the chalcopyrite chemical composition varies as a function of lithostratigraphic setting and that, by using multivariate statistical analysis and an RF classification model combined with an understanding of the processes that control the trace element content of chalcopyrite, it is possible to distinguish chalcopyrite from different lithostratigraphic subtypes and to classify accurately unknown chalcopyrite samples by VMS-SMS subtype. In addition, we tested the models on literature data, demonstrating that internal variation at the deposit scale or in a single VMS subtype is considerably less important that the variation between VMS subtypes.

Methodology

Sample selection

A total of 100 chalcopyrite-bearing samples (n = 100) were selected from 51 VMS-SMS deposits from around the world (Fig. 1). The deposits belong to six lithostratigraphic types of settings according to the classification for VMS-SMS (Franklin et al., 2005; Patten et al., 2022), as follows (Table 1): (1) ultramafic (UM, n = 8); (2) mafic (M, n= 6); (3) siliciclastic-mafic (SM, n = 3); (4) bimodal-mafic (BM, n = 27); (5) bimodal-felsic (BF, n = 30), and (6) siliciclastic-felsic (SF, n = 22). The VMS samples (n = 89) comprise all of the lithostratigraphic types, except for the ultramafic type. The SMS samples (n = 11) are related to the ultramafic (n = 8; from the Logatchev and Irinovskoe vent fields) and bimodal-felsic types (n = 3; PacManus vent field). The sulfide proportions in each sample are reported as ccp/(ccp + sp), po/(po + sp + py), and ccp/(ccp + po) ratios (where ccp = chalcopyrite, sp = sphalerite, py = pyrite, and po = pyrrhotite). The sulfide proportion was calculated for each sample from percentage of sulfides in thin section and in macroscopic sample, and from information on position in deposit. Since this study focuses on chalcopyrite, samples from the Cu-rich areas of mineralization were targeted. However, chalcopyrite from the Matagami district comprises samples of both Cu-rich ($n = 7$) and Zn-rich ($n = 6$) ores and represents the most balanced sample set according to Cu/Zn ratio. Furthermore, according to microscopic and megascopic sample observations, we classified the samples into massive ore (>60% opaque minerals and massive structure) and stockwork ores (stringer/vein mineralization; App. 1, Table A1).

Most of the studied deposits are metamorphosed, from prehnite-pumpellyite to principally greenschist conditions, and only one has undergone amphibolite-grade metamorphism (Tétrault, Québec). Among the unmetamorphosed deposits are those from Semail (Lasail) and VMS-SMS deposits (Table 1; App. 1, Table A1). The siciclastic-felsic group presents the greater diversity of the metamorphic grades, whereas the siliciclastic-mafic subtype is at upper greenschist metamorphic facies.

The selected samples are representative of the different subtypes of VMS-SMS deposits. For 10 deposits we analyzed more than four samples, whereas we analyzed two to three samples in 10 other deposits and one in the remaining 31 deposits (Table 1). According to the compilation included in Franklin et al. (2005), among a total of 646 VMS deposits, only 63 (10%) correspond to mafic and 80 (12%) to siliciclastic-mafic settings (Fig. 3). An underlying class imbalance problem is present as a result of rarity of deposits be-
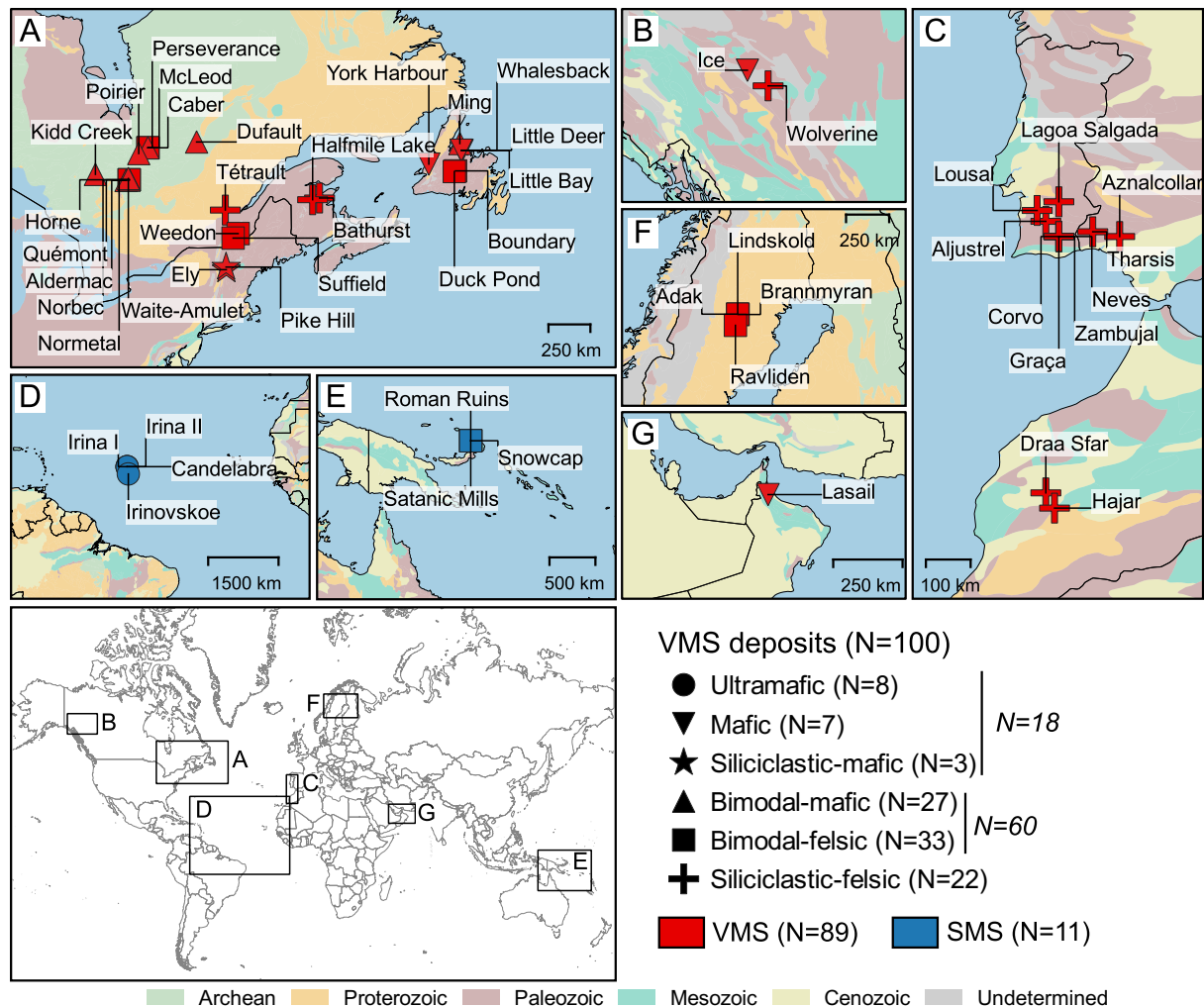
Fig. 1. Location of selected volcanogenic massive sulfide–sea-floor massive sulfide (VMS-SMS) deposits in this study. Classification according to Franklin et al. (2005) and Patten et al. (2022). Generalized geologic map of the world (Chorlton, 2007). *N* = number of samples.

longing to these settings. In our data, deposits from the mafic subtype represent about 12% (7 samples from six deposits; Fig. 1), belonging to four host tectonostratigraphic complexes with a total of 42 analyses (7.4%). Based on these sampling statistics, the mafic subtype would be considered representative, since its distribution is similar to the deposit distribution from Franklin et al. (2005). Concerning the siliciclastic-mafic subtype, according to Franklin et al. (2005), it represents 12% of compiled deposits. In our data, deposits from this setting represent about 4%, belonging to the same host tectonostratigraphic complex. The representative nature of the selected samples and deposits ensures that the analytical data will capture the more important characteristics of chalcopyrite composition in VMS-SMS deposits, but it is possible that some intradeposit variance or less common features will require detailed deposit-scale studies and investigation of other VMS-SMS deposits in the future.

Petrography

Mineralogical composition and detailed textural descriptions of each thin section, considered representative of the sample,

were carried out using a reflected-light microscope. From this information, paragenetic sequences were determined in each sample. Modal percentages were calculated using ImageJ 1.52p based on a mosaic of photomicrographs. To verify the presence of inclusions not detected under optical microscope, and to select regions without inclusions (<10 µm) for geochemical analysis, scanning electron microscopy (SEM) was used (model FEI Inspect F50, equipped with an energy-dispersive spectrometer [EDS] at Université Laval; App. 2, Fig. A1). Between five and seven grains of chalcopyrite were selected in each polished section for geochemical analysis, based on the amount and distribution of chalcopyrite, grain size, and mineralogical association.

Analytical methods

Major elements in 296 grains of chalcopyrite (*n* = 296) were analyzed by electron probe microanalysis (EPMA), using a CAMECA SX-100 equipped with five wavelength-dispersive spectrometers (WDS) at Université Laval, Québec. Beam size was 10 µm with a voltage of 15 kV and a current of 20 nA. The background was measured for 15 to 20 s on both sides, and a

peak counting time of 10 s was carried out. Standard Fe_2O_3 (Fe), FeS_2 (S), and $CuFeS_2$ (Cu) were used for the calibration. The results are given in Appendix 1 and Table A2 and show that the major element composition of all chalcopyrite is close to stoichiometric values.

Trace elements in 561 grains of chalcopyrite (n = 561) were analyzed by laser ablation inductively coupled plasma mass spectrometry (LA-ICP-MS) at LabMaTer (Université du Québec à Chicoutimi, UQAC) using an Excimer 193-nm RESOlution M-50 laser ablation system (Australian Scientific Instrument) equipped with a double volume cell S-155 (Laurin Technic) and coupled to an Agilent 7900 ICP-MS following the protocol described in Caraballo et al. (2022). Depending on grain size, a beam size of 55 μm and 33 μm, a line speed of 5 μm/s, a laser frequency of 15 Hz, and a fluence of 3 J/cm² were used to analyze chalcopyrite. Lines were made across the sulfide grains for a period of 20 to 60 s according to grain size, after measuring the gas blank for 30 s. The ablated material was carried into the ICP-MS by an argon-helium gas mix at a rate of 0.8 to 1 L/min for Ar and 350 mL/min for He, and 2 mL/min of nitrogen was also added to the mixture. Data reduction was carried out using the Iolite package for Igor Pro software (Paton et al., 2011). Full results are given in Appendix 1, Table A3.

Thirty-six isotopes were measured: ^{24}Mg, ^{29}Si, ^{34}S, ^{49}Ti, ^{51}V, ^{52}Cr, ^{55}Mn, ^{59}Co, ^{60}Ni, ^{65}Cu, ^{66}Zn, ^{71}Ga, ^{74}Ge, ^{75}As, ^{82}Se, ^{95}Mo, ^{101}Ru, ^{103}Rh, ^{107}Ag, ^{108}Pd, ^{111}Cd, ^{115}In, ^{118}Sn, ^{121}Sb, ^{128}Te, ^{137}Ba, ^{182}W, ^{185}Re, ^{189}Os, ^{193}Ir, ^{195}Pt, ^{197}Au, ^{202}Hg, ^{205}Tl, ^{208}Pb, and ^{209}Bi. The polyatomic interference of $^{40}Ar^{63}Cu$ on ^{103}Rh was corrected in chalcopyrite using a Cu blank. However, the ^{103}Rh values are not reported, as the interference corrections are too large in chalcopyrite (1% Cu produces about 0.1–0.2 ppm ^{103}Rh interference). Additionally, in most cases, PGE contents were below detection limit in chalcopyrite. Direct interference of ^{115}Sn on ^{115}In and ^{108}Cd on ^{108}Pd were corrected in chalcopyrite manually from ^{118}Sn and ^{111}Cd. We used ^{57}Fe as the internal standard with stoichiometric iron value of chalcopyrite (30.43%). Calibration was based on a combination of three international reference materials (RM): MASS-1 (Wilson et al., 2002), corresponding to a pressed powder pellet (ZnCuFeS) used to calibrate S, Cu, Zn, As, Se, Mo, Cd, Sn, Sb, Te, W, Hg, Tl, and Pb; GSE-1g (Jochum et al., 2005), an Fe-rich natural basaltic glass used to calibrate Si, Ti, V, Cr, Mn, Co, Ni, Ga, Ge, Ag, In, Ba, Re, and Bi; and Laflamme Po727, a synthetic FeS doped with PGEs, used to calibrate Au and PGEs. The reference materials used to monitor quality control of the analyses were as follows: CCu-1e, a copper concentrate (certified by CCRMP-Canadian Certified Reference Materials project, CANMET, Canada); UQAC-FeS1, a synthetic sulfide developed at LabMaTer (UQAC); and PTC-1b, a Cu-Ni-sulfide concentrate powder (Jochum et al., 2005). The results of analyses of reference materials and the ranges of detection limits are shown in Appendix 1, Table A4. Most elements (Ag, As, Au, Bi, Co, Ga, In, Ir, Mn, Mo, Ni, Os, Pb, Rh, Ru, Sb, Se, Sn, Te, Tl, and Zn) have a moderate to high accuracy relative to their reference values (relative difference [rel. diff.] <10–15%) and a moderate to high precision (relative standard deviation [RSD] <10–15%). Elements such as Ti, Ge, and Cd show lower accuracy (rel. diff. >15%)

and lower precision (RSD >15%; App. 1, Table A4). Differences between working values in a few of the reference materials and those measured in this study could be a result of heterogeneity and non-matrix-matched materials or due to the low level of confidence of reported working values for some elements in reference materials (provisional/informative values). As polyatomic interferences were detected on ^{49}Ti (from $^{33}S^{16}O^+$) and ^{74}Ge (from $^{40}Ar^{34}S^+$, $^{36}Ar^{38}Ar^+$, and/or $^{38}Ar^{36}S^+$), and Cd shows low accuracy and precision, these elements were not used for statistical analysis.

Elements such as Si (silicates), Zn-Cd (sphalerite), Pb (galena), Ni (pentlandite), and Ba (barite) were used to monitor for inclusions. On each time signal, only micrometric inclusion-free regions were selected for data reduction in order to have the most reliable composition of chalcopyrite (App. 2, Fig. A1).

Statistical methods

Data preprocessing: Before statistical analysis, the geochemical data were preprocessed for the values below detection limit (i.e., censored values). Since most imputation algorithms do not remain stable with higher degrees of censoring (Hron et al., 2010; Antweiler, 2015; Makvandi et al., 2016b), elements with more than 40% of censored values were excluded from the database. The log-ratio expectation-maximization (lrEM) algorithm was applied to impute censored values, using the zComposition R-package (Palarea-Albaladejo and Martín-Fernández, 2015). According to Palarea-Albaladejo and Martín-Fernández (2008), an alr-transformed component $y_{ij} = ln(x_{ij}/x_{iD})$ from a compositional vector x_i is considered a censored value if $y_{ij} \leq \psi_{ij}$, where ψ_{ij} is an alr-transformed threshold (detection limit). This adapted statistical iterative procedure ensures that censored values y_{ij} are replaced by values lower than threshold ψ_{ij} (Palarea-Albaladejo et al., 2007).

Compositional data describe quantitatively the parts of a whole, which sum to a constant (i.e., 100%), having thus a closed nature, leading to potential spurious correlations (Filzmoser et al., 2009; Egozcue and Pawlowsky-Glahn, 2011; Pawlowsky-Glahn and Buccianti, 2011). In order to avoid this problem, the data were transformed using centered log ratios following Aitchison (1986).

Exploratory data analysis: The exploratory analysis comprised univariate, bivariate, and multivariate methods to identify systematic variations in the data before building discrimination models. Univariate analysis (box plots) explores each variable in a data set separately. In this phase, to characterize different variables, the trimmed mean (μt) and the trimmed standard deviation (sdt), two robust measures of location and dispersion with the aim to reduce effects of outliers on estimator calculation, were used (Wilcox, 2011). These statistic estimators are presented for each variable (trace elements and mineralogical composition) in the format μt ± sdt plus the respective unit (ppm or vol %). The bivariate analysis was conducted to determine whether a statistical association exists between two variables and verify clusters in the data using, respectively, Pearson correlation coefficients and binary diagrams.

A PLS-DA based on clr-transformed and autoscaled data was performed. Previous studies have used PLS-DA on indicator minerals as a tool to classify unknown data (Makvandi et al., 2016b; Sciuba et al., 2020, 2021; Liu and Beaudoin, 2021;

CARABALLO ET AL.

Table 1. Summary of Studied VMS and SMS Deposits (details of sulfide textures and opaque mineralogy in App. 1, Table A1)

Setting[1]	Affinity[2]	Deposit name (no. of samples)	Host tectonostratigraphic complex (age)	Location
Ultramafic	Mafic-ultramafic	Irinovskoe (4)	Mid-Atlantic Ridge (?)	Irinovskoe field, Atlantic Ocean
Ultramafic	Mafic-ultramafic	Candelabra (2)	Mid-Atlantic Ridge (1–60 ka)	Logatchev-1 field, Atlantic Ocean
Ultramafic	Mafic-ultramafic	Irina I (1)	Mid-Atlantic Ridge (1–60 ka)	Logatchev-1 field, Atlantic Ocean
Ultramafic	Mafic-ultramafic	Irina II (1)	Mid-Atlantic Ridge (1–60 ka)	Logatchev-1 field, Atlantic Ocean
Mafic	Mafic-ultramafic	Lasail (2)	Semail ophiolite (95 Ma)	Al Batinah North, Oman
Mafic	Mafic-ultramafic	Little Deer (1)	Dunnage zone-Notre Dame subzone (0.48 Ga)	Lushs Bight, Newfoundland, Canada
Mafic	Mafic-ultramafic	Whalesback (1)	Dunnage zone-Notre Dame subzone (0.48 Ga)	Lushs Bight, Newfoundland, Canada
Mafic	Mafic-ultramafic	Little Bay (1)	Dunnage zone-Notre Dame subzone (0.48 Ga)	Newfoundland, Canada
Mafic	Mafic-ultramafic	York Harbour (1)	Humber zone-Bay of Islands ophiolite (0.48 Ga)	Newfoundland, Canada
Mafic	Mafic-ultramafic	Ice (1)	Slide Mountain terrane (0.28 Ga)	Yukon, Canada
Siliciclastic-mafic	Mafic-ultramafic	Ely (1)	Vermont copper belt (0.41 Ga)	Vermont, USA
Siliciclastic-mafic	Mafic-ultramafic	Pike Hill (2)	Vermont copper belt (0.41 Ga)	Vermont, USA
Bimodal-mafic	Bimodal	Poirier (1)	Abitibi greenstone belt (2.7 Ga)	Joutel, Quebec, Canada
Bimodal-mafic	Bimodal	Caber (3)	Abitibi greenstone belt (2.7 Ga)	Matagami, Quebec, Canada
Bimodal-mafic	Bimodal	Ming (5)	Dunnage zone-Notre Dame subzone (0.48 Ga)	Newfoundland, Canada
Bimodal-mafic	Bimodal	Aldermac (3)	Abitibi greenstone belt (2.7 Ga)	Noranda, Quebec, Canada
Bimodal-mafic	Bimodal	Dufault (3)	Abitibi greenstone belt (2.7 Ga)	Noranda, Quebec, Canada
Bimodal-mafic	Bimodal	Norbec (2)	Abitibi greenstone belt (2.7 Ga)	Noranda, Quebec, Canada
Bimodal-mafic	Bimodal	Normetal (1)	Abitibi greenstone belt (2.7 Ga)	Noranda, Quebec, Canada
Bimodal-mafic	Bimodal	Waite-Amulet (5)	Abitibi greenstone belt (2.7 Ga)	Noranda, Quebec, Canada
Bimodal-mafic	Bimodal	Kidd Creek (4)	Abitibi greenstone belt (2.7 Ga)	Timmins, Ontario, Canada
Bimodal-felsic	Bimodal	Suffield Mine (1)	Ascot-Weedon Complex (0.45 Ga)	Cantons de l'Est, Quebec, Canada
Bimodal-felsic	Bimodal	Weedon (1)	Ascot-Weedon Complex (0.45 Ga)	Cantons de l'Est, Quebec, Canada
Bimodal-felsic	Bimodal	McLeod (4)	Abitibi greenstone belt (2.7 Ga)	Matagami, Quebec, Canada
Bimodal-felsic	Bimodal	Perseverance (6)	Abitibi greenstone belt (2.7 Ga)	Matagami, Quebec, Canada
Bimodal-felsic	Bimodal	Boundary (1)	Dunnage zone-Exploits subzone (0.51 Ga)	Newfoundland, Canada
Bimodal-felsic	Bimodal	Duck Pond (1)	Dunnage zone-Exploits subzone (0.51 Ga)	Newfoundland, Canada

Table 1. (*Cont.*)

Host rock composition	Regional metamorphism	References
Pdt-Gabnor	Unmetamorphosed	Cherkashev et al. (2013); Petersen et al. (2009); Wohlgemuth-Ueberwasser et al. (2015)
Pdt-Gabnor	Unmetamorphosed	Cherkashev et al. (2013; Petersen et al. (2009); Wohlgemuth-Ueberwasser et al. (2015)
Pdt-Gabnor	Unmetamorphosed	Cherkashev et al. (2013; Petersen et al. (2009); Wohlgemuth-Ueberwasser et al. (2015)
Pdt-Gabnor	Unmetamorphosed	Cherkashev et al. (2013; Petersen et al. (2009); Wohlgemuth-Ueberwasser et al. (2015)
Bas and basaltic And	Unmetamorphosed	Gilgen et al. (2014); Stakes and Taylor (2003)
Pillow Bas	Lower greenschist	Kean et al. (1995); Piercey (2007)
Pillow Bas	Lower greenschist	Cloutier et al. (2015); Piercey (2007)
Pillow Bas	Lower greenschist	Kean et al. (1995); Piercey (2007)
Pillow Bas	Lower greenschist	Lode et al. (2015); Piercey (2007)
Bas	Lower greenschist	Piercey et al. (2008)
Meta-Bas/Meta-Sed/minor Meta-Maf-volc	Upper greenschist	McWilliams et al. (2010); Slack et al. (2001)
Meta-Bas/Meta-Sed/minor Meta-Maf-volc	Upper greenschist	McWilliams et al. (2010); Slack et al. (2001)
Rhy-Dac/And-Bas	Lower greenschist	Barrie et al. (1999); Dubé et al. (2007); Gaboury and Pearson (2008); Lafrance et al. (2000); Mercier-Langevin et al. (2007); Shriver and MacLean (1993); Taylor et al. (2014); Genna et al. (2014a); Genna et al. (2014b); Ioannou et al. (2007)
Rhy	Lower greenschist	Barrie et al. (1999); Dubé et al. (2007); Gaboury and Pearson (2008); Lafrance et al. (2000); Mercier-Langevin et al. (2007); Shriver and MacLean (1993); Taylor et al. (2014); Genna et al. (2014a); Genna et al. (2014b); Ioannou et al. (2007)
Bon/Bas/Rhy	Upper greenschist	Brueckner et al. (2014); Brueckner et al. (2015); Pilote et al. (2020); Pilote and Piercey (2013)
Rhy-And	Lower greenschist	Barrie et al. (1999); Dubé et al. (2007); Gaboury and Pearson (2008); Lafrance et al. (2000); Mercier-Langevin et al. (2007); Shriver and MacLean (1993); Taylor et al. (2014); Genna et al. (2014a); Genna et al. (2014b); Ioannou et al. (2007)
Rhy	Lower greenschist	Barrie et al. (1999); Dubé et al. (2007); Gaboury and Pearson (2008); Lafrance et al. (2000); Mercier-Langevin et al. (2007); Shriver and MacLean (1993); Taylor et al. (2014); Genna et al. (2014a); Genna et al. (2014b); Ioannou et al. (2007)
Rhy-And	Lower greenschist	Barrie et al. (1999); Dubé et al. (2007); Gaboury and Pearson (2008); Lafrance et al. (2000); Mercier-Langevin et al. (2007); Shriver and MacLean (1993); Taylor et al. (2014); Genna et al. (2014a); Genna et al. (2014b); Ioannou et al. (2007)
Rhy-And	Lower greenschist	Barrie et al. (1999); Dubé et al. (2007); Gaboury and Pearson (2008); Lafrance et al. (2000); Mercier-Langevin et al. (2007); Shriver and MacLean (1993); Taylor et al. (2014); Genna et al. (2014a); Genna et al. (2014b); Ioannou et al. (2007)
Rhy	Lower greenschist	Barrie et al. (1999); Dubé et al. (2007); Gaboury and Pearson (2008); Lafrance et al. (2000); Mercier-Langevin et al. (2007); Shriver and MacLean (1993); Taylor et al. (2014); Genna et al. (2014a); Genna et al. (2014b); Ioannou et al. (2007)
Rhy-Um	Lower greenschist	Barrie et al. (1999); Dubé et al. (2007); Gaboury and Pearson (2008); Lafrance et al. (2000); Mercier-Langevin et al. (2007); Shriver and MacLean (1993); Taylor et al. (2014); Genna et al. (2014a); Genna et al. (2014b); Ioannou et al. (2007)
Felsic-volc	Lower greenschist	Gauthier et al. (1994)
Felsic-volc	Lower greenschist	Gauthier et al. (1994)
Rhy-Rhydac	Lower greenschist	Barrie et al. (1999); Dubé et al. (2007); Gaboury and Pearson (2008); Lafrance et al. (2000); Mercier-Langevin et al. (2007); Shriver and MacLean (1993); Taylor et al. (2014); Genna et al. (2014a); Genna et al. (2014b); Ioannou et al. (2007)
Rhy-Rhydac	Lower greenschist	Barrie et al. (1999); Dubé et al. (2007); Gaboury and Pearson (2008); Lafrance et al. (2000); Mercier-Langevin et al. (2007); Shriver and MacLean (1993); Taylor et al. (2014); Genna et al. (2014a); Genna et al. (2014b); Ioannou et al. (2007)
Dac-Rhy/Bas	Lower greenschist	Buschette and Piercey (2016); Piercey (2007)
Dac-Rhy/Bas	Lower greenschist	McNicoll et al. (2010); Piercey (2007)

Table 1. Summary of Studied VMS and SMS Deposits (details of sulfide textures and opaque mineralogy in App. 1, Table A1)

Setting[1]	Affinity[2]	Deposit name (no. of samples)	Host tectonostratigraphic complex (age)	Location
Bimodal-felsic	Bimodal	Horne (4)	Abitibi greenstone belt (2.7 Ga)	Noranda, Quebec, Canada
Bimodal-felsic	Bimodal	Quémont (5)	Abitibi greenstone belt (2.7 Ga)	Noranda, Quebec, Canada
Bimodal-felsic	Bimodal	Roman Ruins (1)	Pual Ridge (20–50 ka)	Pacmanus field, Papua New Guinea
Bimodal-felsic	Bimodal	Satanic Mills (1)	Pual Ridge (20–50 ka)	Pacmanus field, Papua New Guinea
Bimodal-felsic	Bimodal	Snowcap (1)	Pual Ridge (20–50 ka)	Pacmanus field, Papua New Guinea
Bimodal-felsic	Bimodal	Adak (2)	Bothnia-Skelleftea unit (1.8 Ga)	Västerbotten, Sweden
Bimodal-felsic	Bimodal	Brannmyran (1)	Bothnia-Skelleftea unit (1.8 Ga)	Västerbotten, Sweden
Bimodal-felsic	Bimodal	Lindsköld (3)	Bothnia-Skelleftea unit (1.8 Ga)	Västerbotten, Sweden
Bimodal-felsic	Bimodal	Rävliden (1)	Bothnia-Skelleftea unit (1.8 Ga)	Västerbotten, Sweden
Siliciclastic-felsic	Siliciclastic-felsic	Aznalcollar (1)	Iberian Pyrite Belt (0.34–0.38 Ga)	Andalucia, Spain
Siliciclastic-felsic	Siliciclastic-felsic	Tharsis (1)	Iberian Pyrite Belt (0.34–0.38 Ga)	Andalucia, Spain
Siliciclastic-felsic	Siliciclastic-felsic	Aljustrel (1)	Iberian Pyrite Belt (0.34–0.38 Ga)	Beja, Portugal
Siliciclastic-felsic	Siliciclastic-felsic	Corvo (4)	Iberian Pyrite Belt (0.34–0.38 Ga)	Beja, Portugal
Siliciclastic-felsic	Siliciclastic-felsic	Graça (1)	Iberian Pyrite Belt (0.34–0.38 Ga)	Beja, Portugal
Siliciclastic-felsic	Siliciclastic-felsic	Neves (2)	Iberian Pyrite Belt (0.34–0.38 Ga)	Beja, Portugal
Siliciclastic-felsic	Siliciclastic-felsic	Zambujal (1)	Iberian Pyrite Belt (0.34–0.38 Ga)	Beja, Portugal
Siliciclastic-felsic	Siliciclastic-felsic	Hajar (1)	Guemassa Massif (0.34 Ga)	Marrakech-Safi, Morocco
Siliciclastic-felsic	Siliciclastic-felsic	Draa Sfar (1)	Jebilet Massif (0.34 Ga)	Marrakech-Safi, Morocco
Siliciclastic-felsic	Siliciclastic-felsic	Tétrault (1)	Greenville Province	Montauban, Quebec, Canada
Siliciclastic-felsic	Siliciclastic-felsic	Bathurst (4)	Bathurst Mining Camp (0.47 Ga)	New Brunswick, Canada
Siliciclastic-felsic	Siliciclastic-felsic	Halfmile Lake (1)	Bathurst Mining Camp (0.47 Ga)	New Brunswick, Canada
Siliciclastic-felsic	Siliciclastic-felsic	Lagoa Salgada (1)	Iberian Pyrite Belt (0.34–0.38 Ga)	Setúbal, Portugal
Siliciclastic-felsic	Siliciclastic-felsic	Lousal (1)	Iberian Pyrite Belt (0.34–0.38 Ga)	Setúbal, Portugal
Siliciclastic-felsic	Siliciclastic-felsic	Wolverine (1)	Yukon-Tanana terrane (0.34 Ga)	Yukon, Canada

Abbreviations: And = andesite, Bas = basalt, Bon = bonninite, Dac = dacite, Fel-volc = felsic volcanic rocks, Gabnor = gabbronorite, Maf-volc = mafic volcanic rocks, Nor = norite, Pdt = peridotite, Phy = phyllite, Qzt = quartzite, Rhy = rhyolite, Rhydac = rhyodacite, Sed = sedimentary rocks, SMS = sea-floor massive sulfide, Um = ultramafic rocks, VMS = volcanogenic massive sulfide, Volc = volcanic
[1]Classification according to Franklin et al. (2005)
[2]Classification according to major host rocks composition

Liu et al., 2021; Caraballo et al., 2022; Miranda et al., 2022). However, since in this study the PLS-DA recorded lower performance than the RF model, as presented below, this was used as a multivariate exploratory method. This method consists of a supervised linear classification, which models dependent variables Y from a set of predictor variables X by reducing the multidimensionality of the data (Wold et al., 2001; Ruiz-Perez et al., 2020). A PLS-DA model can be interpreted from two principal outputs: scores and loadings. After calculating latent variables, they can be used in a cartesian space (i.e., t_1-t_2) to project X-scores of the observations. The aim of these representations, called score plots, is to look for clusters indicating similar characteristics (Brereton and Lloyd, 2014). Likewise, the distance between clusters can suggest differences in attributes. On the other hand, loading plots show correlations between original variables. Superposing X-loadings (specifically X-weight, w^*) and Y-loadings in the same plot, the relationship between X-variables (elements) and Y-variables (classes) can be inferred (Wold et al., 2001; Dunn, 2019). Variables positively correlated are shown grouped, whereas those with a negative correlation plot diametrically opposed. Location of variables in a loading plot is dependent on contributions in discrimination; therefore, X-variables near the correlation circle (radius = 1) are interpreted to be highly variable between classes. In contrast, those located near the origin contribute weakly to classification, as variability is negligible. Since dummy variables are required to use PLS-DA on categorical data, the encoding, which consisted of binary integers where 1 is in class and 0 is out of class, was performed automatically by the CARET package in R.

Different estimators have been used to judge the importance of the X-variables on the classification model. Mehmood et al. (2012) described three filter methods (loading weights, regression coefficients, and variable importance on projection)

Table 1. (*Cont.*)

Host rock composition	Regional metamorphism	References
Rhy-And	Lower greenschist	Barrie et al. (1999); Dubé et al. (2007); Gaboury and Pearson (2008); Lafrance et al. (2000); Mercier-Langevin et al. (2007); Shriver and MacLean (1993); Taylor et al. (2014); Genna et al. (2014a); Genna et al. (2014b); Ioannou et al. (2007)
Rhy-And	Lower greenschist	Barrie et al. (1999); Dubé et al. (2007); Gaboury and Pearson (2008); Lafrance et al. (2000); Mercier-Langevin et al. (2007); Shriver and MacLean (1993); Taylor et al. (2014); Genna et al. (2014a); Genna et al. (2014b); Ioannou et al. (2007)
Dac-Rhy	Unmetamorphosed	Binns et al. (2007); Hannington et al. (2005); Reeves et al. (2011)
Dac-Rhy	Unmetamorphosed	Binns et al. (2007); Hannington et al. (2005); Reeves et al. (2011)
Dac-Rhy	Unmetamorphosed	Binns et al. (2007); Hannington et al. (2005); Reeves et al. (2011)
Felsic-volc	Lower greenschist	Allen et al. (1996); Årebäck et al. (2005); Johansson (2017); Skyttä et al. (2020)
Felsic-volc	Lower greenschist	Allen et al. (1996); Årebäck et al. (2005); Johansson (2017); Skyttä et al. (2020)
Felsic-volc	Lower greenschist	Allen et al. (1996); Årebäck et al. (2005); Johansson (2017); Skyttä et al. (2020)
Fel-volc to Maf-volc/argilites	Lower greenschist	Allen et al. (1996); Årebäck et al. (2005); Johansson (2017); Skyttä et al. (2020)
Rhy-Dac/phyllite-quartzite	Prehnite-pumpellyite	Almodóvar et al. (2019); Barrie et al. (2002); Inverno et al. (2015); Relvas et al. (2001); Relvas et al. (2006); Sáez et al. (1999); Tornos (2006)
Rhy-Dac/phyllite-quartzite	Prehnite-pumpellyite	Almodóvar et al. (2019); Barrie et al. (2002); Inverno et al. (2015); Relvas et al. (2001); Relvas et al. (2006); Sáez et al. (1999); Tornos (2006)
Rhy-Dac/phyllite-quartzite	Prehnite-pumpellyite	Almodóvar et al. (2019); Barrie et al. (2002); Inverno et al. (2015); Relvas et al. (2001); Relvas et al. (2006); Sáez et al. (1999); Tornos (2006)
Rhy-Dac/phyllite-quartzite	Prehnite-pumpellyite	Almodóvar et al. (2019); Barrie et al. (2002); Inverno et al. (2015); Relvas et al. (2001); Relvas et al. (2006); Sáez et al. (1999); Tornos (2006)
Rhy-Dac/phyllite-quartzite	Prehnite-pumpellyite	Almodóvar et al. (2019); Barrie et al. (2002); Inverno et al. (2015); Relvas et al. (2001); Relvas et al. (2006); Sáez et al. (1999); Tornos (2006)
Rhy-Dac/phyllite-quartzite	Prehnite-pumpellyite	Almodóvar et al. (2019); Barrie et al. (2002); Inverno et al. (2015); Relvas et al. (2001); Relvas et al. (2006); Sáez et al. (1999); Tornos (2006)
Rhy-Dac/phyllite-quartzite	Prehnite-pumpellyite	Almodóvar et al. (2019); Barrie et al. (2002); Inverno et al. (2015); Relvas et al. (2001); Relvas et al. (2006); Sáez et al. (1999); Tornos (2006)
Rhy-Rhydac	Lower greenschist	Admou et al. (2018); Marcoux et al. (2008); Moreno et al. (2008)
Rhy-Rhydac	Lower greenschist	Admou et al. (2018); Marcoux et al. (2008); Moreno et al. (2008)
Gneiss	Amphibolite	Tomkins (2007); Stamatelopoulou-Seymour and MacLean (1984)
Rhy-Rhydac	Upper greenschist	Goodfellow (2007); Mireku and Stanley (2006)
Rhy-Rhydac	Upper greenschist	Goodfellow (2007); Mireku and Stanley (2006)
Rhy-Dac/Phy-Qzt	Prehnite-pumpellyite	Almodóvar et al. (2019); Barrie et al. (2002); Inverno et al. (2015); Relvas et al. (2001); Relvas et al. (2006); Sáez et al. (1999); Tornos (2006)
Rhy-Dac/Phy-Qzt	Prehnite-pumpellyite	Almodóvar et al. (2019); Barrie et al. (2002); Inverno et al. (2015); Relvas et al. (2001); Relvas et al. (2006); Sáez et al. (1999); Tornos (2006)
Maf-volc and Fel-volc/dark Sed	Middle greenschist	Piercey et al. (2008)

used to determine variable importance. Variable importance on projection (VIP) quantifies the influence on the model by accumulating the importance of each variable *j* reflected by the weight *w* from each latent variable (Mehmood et al., 2012; Favilla et al., 2013). The average of VIP scores is 1, and the sum of squares of all VIP scores is equal to *M*. Thus, variables with VIP scores >1.0 are considered highly influential on the discrimination model, whereas VIP scores between 0.8 and 1.0 indicate a variable moderately influential, and VIP scores <0.8 represent less important variables (Chong and Jun, 2005; Gosselin et al., 2010; Mendez et al., 2020).

A second, widely used method to assess the importance of *X*-variables on the classification model is the vector or matrix of regression coefficients (Bpls), which measures the association between predictor and response variables (Mehmood et al., 2012). Partial least squares discriminant analysis is based on a regression algorithm, thus these coefficients are calculated for each modeled response or class (Ballabio and Consonni, 2013). Variables with a strong importance for a specific class have generally high absolute values of regression coefficients, whereas those with small absolute values do not contribute importantly (Gosselin et al., 2010; Mehmood et al., 2012; Ballabio and Consonni, 2013).

Random Forest: In order to test the trace element composition of chalcopyrite to discriminate different lithostratigraphic settings of VMS deposits, a supervised classification was performed using the RF algorithm. In addition to the classification model, a preliminary RF regression model was developed using trace elements in chalcopyrite to predict ccp/(ccp + sp) ratio. Random forest consists of an ensemble of methods using tree-type classifiers, which generates predictions based on class membership probabilities for individual predictions (Breiman, 2001). This algorithm is random in two ways: (1) each tree is based on a random subsample of the observations, and (2) each split within each tree is created based on a random subset of independent variables or predictors (Breiman, 2001; Gislason et al., 2006; Harris and Grunsky, 2015). The random subsamples are created from an internal resampling in the training set through a bootstrapping procedure, which provides a stable estimate of classification accuracy from out-of-bag (OOB) data (Harris and Grunsky, 2015). An RF algorithm is considered a nonparametric classifier; it

is insensitive to outliers and requires minimal input parameters—the number of decision trees and the number of random variables for each decision tree, called mtry (Harris and Grunsky, 2015). For classification problems, the dependent variable Y is categorical (classes), and the prediction is based on a majority class vote, whereas in regression, the outcome Y is numerical (e.g., sulfide ratio), and the final prediction is the average of the individual tree values (Svetnik et al., 2003; Grömping, 2009; Brownlee, 2018; Barker et al., 2021).

Before building the classification and regression models, a recursive feature elimination (RFE) incorporating resampling was performed to determine the optimal number of variables using an exploratory RF model. When the model is created, a measure of variable importance is calculated; thus, the least important predictor is iteratively eliminated prior to rebuilding the new model (Kuhn, 2008; Kuhn and Johnson, 2013). Thus, an optimal number of important variables is chosen and used to train the RF model.

The data set in this study presents a class imbalance problem, where the ultramafic, mafic, and particularly siliciclastic-mafic classes have the lowest number of samples. The class imbalance ratio between the overrepresented bimodal-felsic and underrepresented siliciclascic-mafic classes is ≈9. Class imbalance in a data set degrades the performance of machine learning models, leading to misclassifying the minority class samples (Abd Elrahman and Abraham, 2013). This problem has been discussed in several studies, and different approaches have been proposed to counteract its negative effect (e.g., Abd Elrahman and Abraham, 2013; Kuhn and Johnson, 2013; Weiss, 2013; Thabtah et al., 2020; Megahed et al., 2021). The sampling methods are mostly used to overcome the class imbalance and include preprocessing of data by creating a balanced training data set, either by undersampling the majority class and/or oversampling the minority class (Abd Elrahman and Abraham, 2013). However, Tarawneh et al. (2022) evaluated more than 70 oversampling methods, of which only six were developed to multiclass problems, and concluded that in the current forms, the sampling methods are unreliable for learning from imbalance data. Oversampling methods create fictitious samples, adding more samples to the minority class, and can result in an overfitting (Abd Elrahman and Abraham, 2013; Megahed et al., 2021). In contrast, undersampling methods randomly remove samples from the majority class, which can result in a loss of important information (Abd Elrahman and Abraham, 2013; Megahed et al., 2021). According to Weiss (2013), there is no standard on the degree of class imbalance in a data set to be considered unbalanced, indicating that overall, a class imbalance ratio (IR) ≤2 would be considered as marginally unbalanced, whereas an IR ≈10 would be considered modestly unbalanced, and an IR ≈1,000 would be considered extremely unbalanced. To avoid the problem of class imbalance, the problem was redefined with a divisive approach by classifying the samples according to major host rock affinity and training an RF model as a first step. In this method, samples are separated into three major classes, as follows (Table 1):

1. Mafic-ultramafic (M-UM) class: contains all samples from subtypes constituted principally by mafic-ultramafic host rocks. It comprises ultramafic, mafic, and siliciclastic-mafic subtypes.

2. Bimodal-mafic-felsic (BMF) class: contains all samples from bimodal-felsic and bimodal-mafic subtypes.
3. Siliciclastic-felsic (SF) class: corresponds to the original siliciclastic-felsic setting.

Once the samples are classified by the affinity model (RF model.01), two models will be used to reclassify between ultramafic, mafic, and siliciclastic-mafic subtypes (RF model.02-A) and between bimodal-mafic and bimodal-felsic subtypes (RF model.02-B).

For each RF model, the data set was split into training (70%) and test (30%) subsets using stratified sampling to obtain balanced classes. In order to improve the classification model, the optimal number of randomly selected variables in each split (mtry) was determined using a hyperparameter grid search function, implemented by the Caret package in R software. Performances were determined through 10-fold cross-validation with 10 repetitions during training, and different metrics were used to evaluate the model efficacy (accuracy, kappa, recall, specificity, precision, F1, and balanced accuracy). For the RF regression model, performance estimators were the root mean squared error (RMSE), the mean absolute error (MAE), and the coefficient of determination (R^2) in training and test data. Details of performance estimators and calculations are shown in Appendix 3. The variable importance in classification and regression models was measured with the mean decrease in accuracy (MDA) and mean decrease Gini (MDG) coefficients (Han et al., 2016). The index MDA uses permuting OOB samples to compute the importance of the variable, whereas the MDG index determines the impurity at each node (Harris and Grunsky, 2015; Gregory et al., 2019).

To verify the model classification accuracy, we used the remaining 30% of the total data set to compare the predicted classes with the actual classes (i.e., known classes) in a confusion matrix. Confusion matrices allow an evaluation of the classification performance, consisting of a square matrix $C \ x \ C$, where C is the number of classes and collects the outputs of the classification model. This allowed the calculation of multiple estimators, such as the accuracy, which measures the percentage of observations correctly classified (Ballabio and Consonni, 2013).

Results

Petrography

Mineralogical composition and textural features are summarized in Figure 2 and Appendix 1, Table A1. Most samples contain an assemblage of chalcopyrite, pyrrhotite, pyrite, and sphalerite in different proportions and textural relations, with minor and trace minerals (i.e., magnetite, cassiterite, tennantite). Chalcopyrite (66 ± 32 vol %) occurs in varying proportions, coexisting principally with anhedral pyrrhotite (4.2 ± 7.9 vol %; Fig. 2A) and commonly replacing pyrite (Fig. 2C) and/or sphalerite (Fig. 2C-E). Pyrite (8.7 ± 13 vol %) is present in most samples, occurring principally as subhedral to anhedral grains (Fig. 2C-E, G, I). Sphalerite (3.2 ± 4.1 vol %) is present as anhedral grains, often showing replacement textures by chalcopyrite in different proportions (Fig. 2C-F). In some cases, sphalerite occurs in fractured pyrite (Fig. 2D). Rarely,
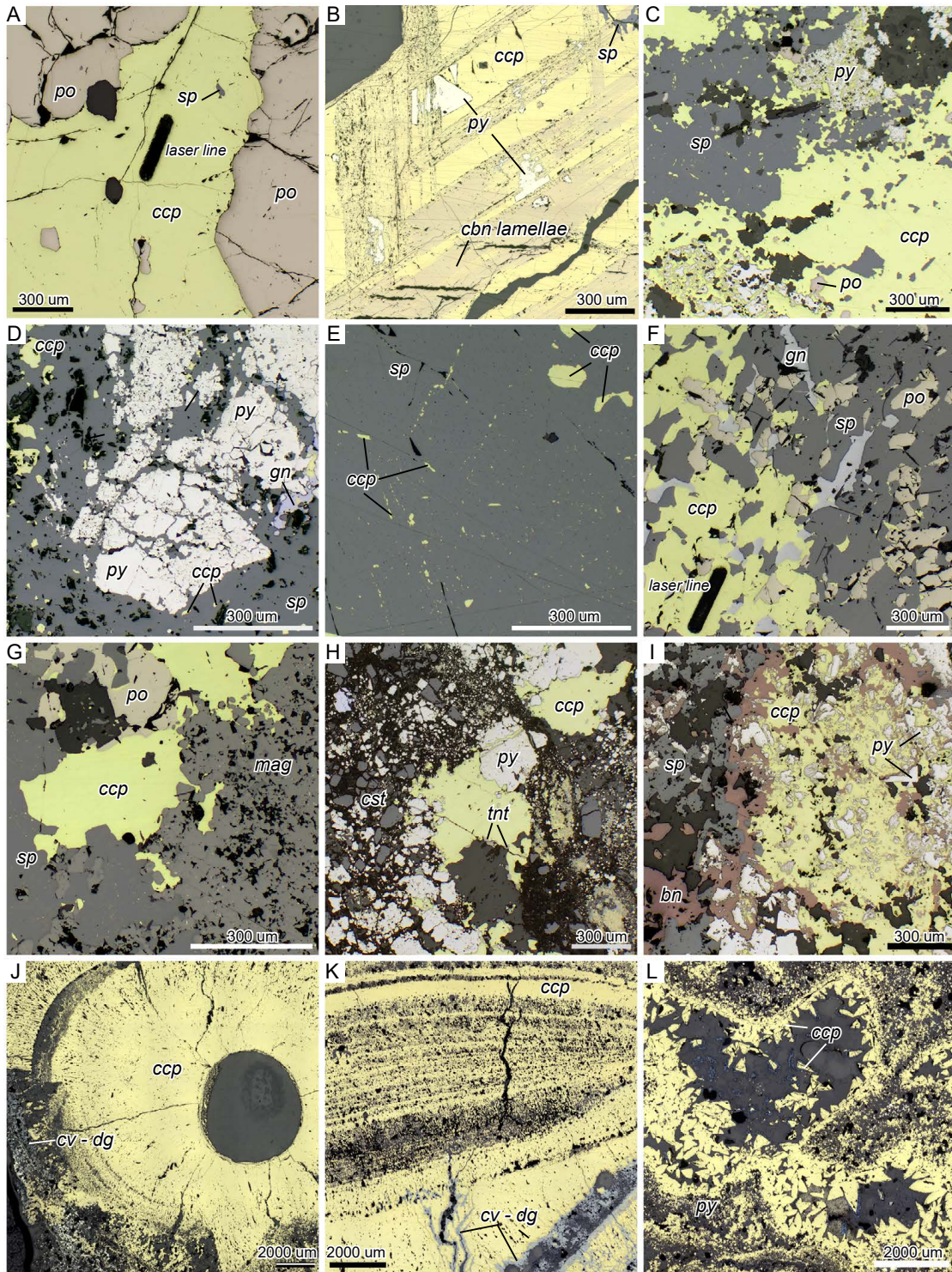
Fig. 2. Reflected light photomicrographs of chalcopyrite from volcanogenic massive sulfide–sea-floor massive sulfide (VMS-SMS) deposits from Cu-rich (A-C) and Zn-rich (D-F) samples. (A) pyrrhotite (po) coexisting with chalcopyrite (ccp) from the Aldermac deposit. (B) Star-like sphalerite (sp) and cubanite (cbn) lamellae exsolutions in chalcopyrite in association with pyrrhotite (Pike Hill deposit). (C) Pyrite (py) and sphalerite replaced by chalcopyrite (chalcopyrite disease) and (D) pyrite replaced by sphalerite; in addition, chalcopyrite replaces sphalerite and pyrite in samples from the Kidd Creek deposit. (E) Chalcopyrite replacing sphalerite (chalcopyrite disease) in sample from Waite-Amulet. (F) Galena (Gn), sphalerite, chalcopyrite, and pyrrhotite assemblage in sample from the Hajar deposit. (G) Magnetite (mag)-rich sample from the Quémont deposit, in association with chalcopyrite, pyrite, pyrrhotite, and sphalerite. (H) Anhedral cassiterite (cst) and euhedral tennantite (tnt) with pyrite and chalcopyrite in Sn-rich sample (Corvo deposit). (I) Sphalerite and pyrite replaced by chalcopyrite and bornite (Ice deposit). (J, K) Layered covellite-digenite (cv-dg) and chalcopyrite in sample from Roman Ruins (J) and Satanic Mills (PacManus field, K). (L) Subhedral chalcopyrite with minor pyrite from the Irinovskoe hydrothermal field.

exsolutions of pyrite in cubanite lamellae are present in chalcopyrite (Fig. 2B).

Magnetite is the most abundant iron oxide (1.9 ± 9.9 vol %), forming mainly anhedral to subhedral grains (Fig. 2G). Occasionally, magnetite is present in high concentrations, such as at the Quémont (≈81 vol %) and Caber (≈45 vol %) deposits (App. 1, Table A1). Cassiterite (<1.5 vol %) and tennantite (<1.5 vol %) are present in samples from the Iberian Pyrite Belt, principally in the stringer mineralization (Fig. 2H). Other Cu minerals such as bornite (<3 vol %) occur with chalcopyrite (Ice deposit; Fig. 2I), whereas covellite (<2.5 vol %) and digenite (<3 vol %) occur dominantly in SMS samples, where they are typically layered, intergrown with coarse- to fine-grained chalcopyrite and occasionally minor anhydrite and opal-A (i.e., amorphous silica; Fig. 2J-L).

Trace element composition of chalcopyrite

Among the analyzed trace elements, Mn, Co, Ni, Zn, Ga, Se, Ag, In, Sn, Sb, Te, Au, Tl, Pb, and Bi have <40% censored data. Arsenic, PGEs, Mo, and Re were below detection limits (<dl) in most samples. Of the elements used for statistical analysis, Zn and Se show the highest mean concentration values (312 and 260 ppm, respectively), whereas Ag, In, and Sn vary, typically between 40 and 100 ppm. Elements such as Pb, Ga, Mn, Bi, and Co have a range between 1 and 5 ppm, whereas the concentrations of Sb, Ni, Te, Au, and Tl are typically below 1 ppm. The data distribution is highly asymmetric with a large range of values and outliers (Fig. 3). Table 2 presents a statistical summary, and the complete database is given in Appendix 1, Table A3.

The box plots (Fig. 3) show that only a few elements have a significant difference amongst the VMS-SMS subtypes according to lithostratigraphy. In chalcopyrite from ultramafic deposits (*n* = 50), Co (275 ± 140 ppm), Te (26 ± 19 ppm), and to a lesser extent Se (718 ± 526 ppm) and Ni (38 ± 81 ppm) contents are higher than those in chalcopyrite from the other VMS-SMS subtypes. In contrast, chalcopyrite from ultramafic settings shows low values of In (4.1 ± 2.9 ppm), Sb (0.07 ± 0.14 ppm), Tl (<dl), Bi (0.19 ± 0.16 ppm), and the lowest content of Pb (0.13 ± 0.15 ppm). Chalcopyrite from mafic settings (*n* = 42) has a slightly higher concentration of Pb (9.3 ± 8.2 ppm) and the lowest Sn (1.5 ± 1.5 ppm) concentrations compared to chalcopyrite from other VMS-SMS subtypes. In bimodal-mafic deposits (*n* = 163), chalcopyrite is slightly enriched in Ag (95 ± 113 ppm); however, this value is highly biased by the values from samples from the Kidd Creek (sample 2369) and Normétal (sample 52) deposits, which have a trimmed mean of 1,491 and 940 ppm of Ag, respectively (Table 2; App. 1, Tables A1, A3). Chalcopyrite from the bimodal-felsic subtype (*n* = 168) has no significant differences compared to other VMS-SMS settings for most of the elements except for slightly higher Au concentrations (0.09 ± 0.14 ppm). However, the data shows a considerable dispersion. Chalcopyrite from the siliciclastic-mafic subtype (*n* = 18) shows higher content in Mn (92 ± 63 ppm), the lowest concentrations of Au (<dl), and the second lowest values in Pb (0.66 ± 0.47 ppm), after ultramafic settings. In siliciclastic-felsic deposits (*n* = 127), chalcopyrite is substantially higher in Sn (467 ± 414 ppm) than in the other subtypes, and to a lesser extent Bi (4.4 ± 5.5 ppm), In (140 ± 152 ppm), and

Sb (3.4 ± 3.8 ppm), whereas Ag (9.6 ± 11 ppm) and Te (<dl) have the lowest values.

Pearson correlation coefficients for chalcopyrite from the VMS-SMS settings show the positive correlation between Pb and Tl (r = 0.50) and to a lesser extent Te and Co (r = 0.41; Fig. 4A-B). Antimony-Co and Pb-Ni have dispersed, negative correlations (Fig. 4C-D). These binary diagrams confirm that chalcopyrite from ultramafic deposits is significantly different from the other settings. Although an imperfect grouping is shown between chalcopyrite in the remaining deposits, samples from siliciclastic-felsic deposits form a cluster in the Te-Co diagram with the lowest content of Te (Fig. 4B). Correlation matrices by deposit setting are shown in Appendix 2, Figure A2. Chalcopyrite from the bimodal-mafic deposit subtype shows a bimodal distribution of Te and Co, and to a lesser extent Ni (Figs. 3, 4). These samples correspond to different host tectonostratigraphic complexes (Abitibi belt vs. Dunnage zone-Notre Dame subzone). Chalcopyrite from the Dunnage zone-Notre Dame subzone complex (Newfoundland) has higher concentrations of Te than most of the chalcopyrite from the Abitibi belt. Furthermore, bimodality in Co and Ni is overall related to the deposit; however, it is less systematic.

Deposits showing differences as a function of the ccp/(ccp + sp) ratio and ore type (massive vs. stockwork) are presented in binary diagrams in Figure 5. Overall, chalcopyrite from samples with low ccp/(ccp + sp) ratio (i.e., Zn-rich samples) is higher in Ga, In, and Sn and lower in Bi and Co than that from samples with high ccp/(ccp + sp) ratios (i.e., Cu-rich samples; Fig. 5A-I). Gold, (except Ascot-Weedon; Fig. 5I), Zn (except Kidd Creek; Fig 5F), and Tl show no differences as a function of this ratio. Chalcopyrite from the Perseverance and McLeod deposits (Matagami; Fig. 5A-D) shows the same variations, except for Sn, whose concentrations are similar in chalcopyrite from Zn-rich and Cu-rich samples (Fig. 5C). From Matagami, only one analysis is from a stockwork sample; consequently, comparison between ore types is not possible.

In chalcopyrite from Kidd Creek (Fig. 5E-H), the differences are defined by ccp/(ccp + sp) ratio but not ore type. In particular, chalcopyrite from Zn-rich samples is considerably higher in Zn than in other deposits (Fig. 5F). The diagram of Tl-Co, however, shows no grouping as a function of ccp/(ccp + sp) (Fig. 5H) except for the Matagami and Ascot-Weedon samples. Furthermore, chalcopyrite from Ascot-Weedon (Suffield and Weedon deposits) forms separate groups according to ccp/(ccp + sp) ratios (Fig. 5I-L). Chalcopyrite from Cu-rich samples is depleted in Ga and Sn compared to chalcopyrite in Zn-rich samples (Fig. 5I). Other elements such as Ag and occasionally Se, Te, and Ni are enriched in chalcopyrite from Cu-rich samples, whereas Sb is principally higher in those from Zn-rich samples (App. 2, Fig. A3). Manganese and Pb do not show systematic variations. However, Pb has a high positive correlation with Sb at Kidd Creek, with a relative enrichment in Pb in Zn-rich samples (r = 0.83; App. 2, Fig. A3H). According to ore type, chalcopyrite from the Ming deposit shows significant differences between massive and stockwork ores, in which that from massive ores is higher in Au, Bi, and Co and lower in In, Sn, and Tl (Fig. 5M-P). Successive remobilization of these elements from sulfides previously formed at lower temperature (e.g., sphalerite,
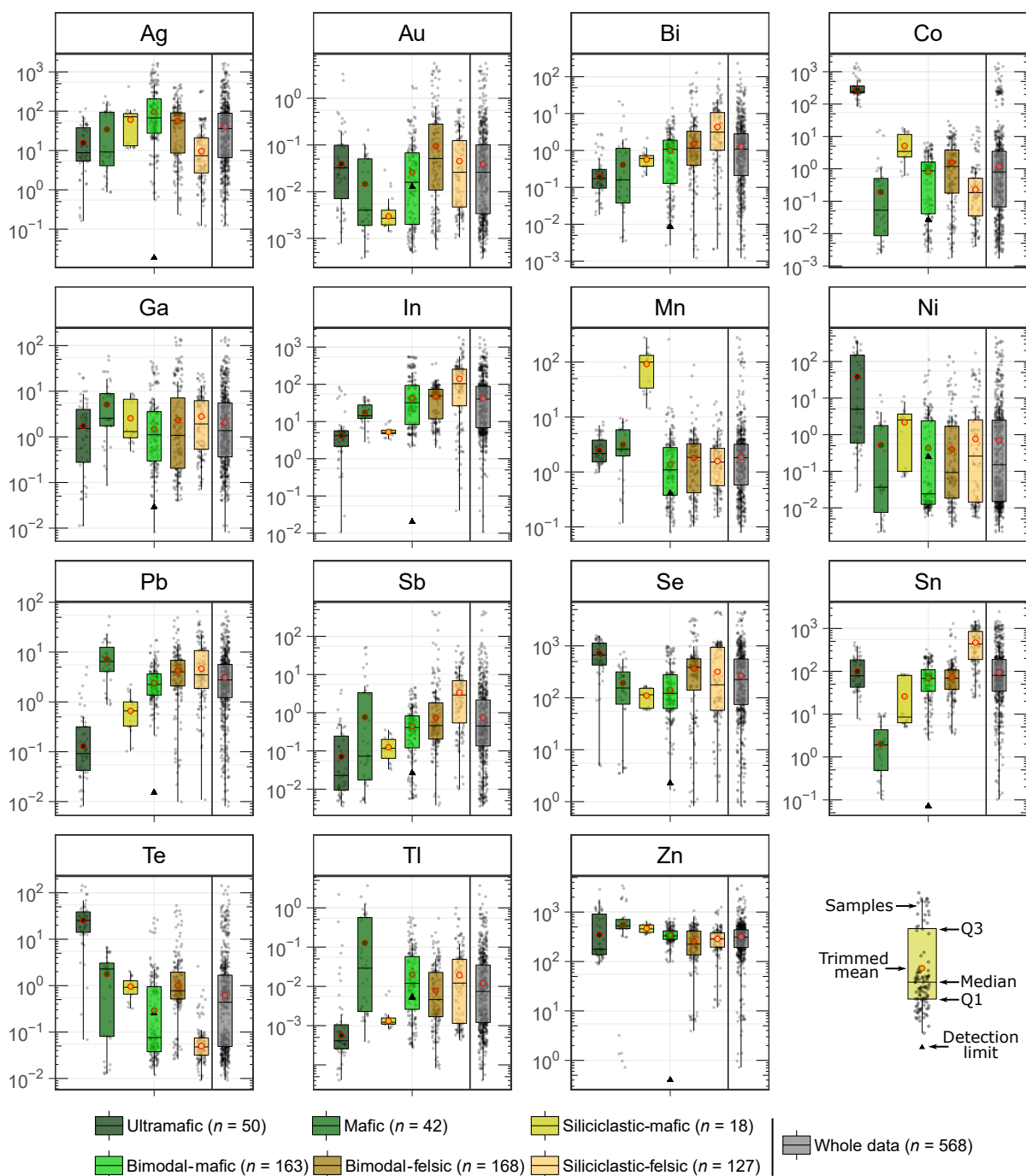
Fig. 3. Box plots showing the variation of selected trace elements (in ppm) in chalcopyrite according to the geologic setting of the deposit; n = number of analyses. Q1 and Q3 indicate first and third quartiles, respectively.

galena, or pyrite) suggests that the trace element composition of chalcopyrite is inherited from replaced assemblages.

Exploratory PLS-DA

The PLS-DA using all elements with less than 40% of censured values confirms that chalcopyrite can be grouped according to VMS-SMS settings (Fig. 6A). The first two latent variables explain about 36.5% of the variance of the 15 chemical variables ($t_1 = 24.7\%$; $t_2 = 11.8\%$; Fig. 6A), and at least six latent variables are necessary to explain more than 60%

of the data variance (App. 2, Fig. A5G). Estimators indicate a moderate to low performance of the model with an accuracy of 0.48 and kappa 0.29 (App. 2, Fig. A5G). Chalcopyrite from ultramafic, mafic, and to a lesser extent siliciclastic-felsic deposits shows significant differences from the other settings (Fig. 6A). Chalcopyrite from ultramafic deposits forms a group separated from the other subtypes, and there is a positive correlation with Co, Te ($\pm$Sn), and to a lesser extent Ni according to the score plot t_1-t_2 and loading plot qw*1-qw*2 (Fig. 6A, B), respectively. In contrast, Ag, Pb, Bi, and Sb are

Table 2. Summary of Selected Trace Elements (ppm) in Chalcopyrite from Volcanogenic Massive Sulfide Deposits
Used for Multivariate Statistical Analysis and Machine Learning

Setting	Name	Host tectonostratigraphic complex	n	Element	Mn	Co	Ni	Zn
				Isotope	55	59	60	66
		Median of detection limits		55 um	0.479	0.016	0.207	0.024
				33 um	1.03	0.028	0.410	0.051
Ultramafic								
	Candelabra	Mid-Atlantic Ridge	11	μ_t	3.41	678	266	755
				sd_t	2.78	706	132	723
	Irina I	Mid-Atlantic Ridge	6	μ_t	3.36	206	163	124
				sd_t	0.942	59.9	114	13.6
	Irina II	Mid-Atlantic Ridge	6	μ_t	3.83	829	0.635	1296
				sd_t	1.65	476	0.976	824
	Irinovskoe	Mid-Atlantic Ridge	27	μ_t	1.80	240	2.09	201
				sd_t	0.631	53.3	2.54	83.4
				μ_t	2.45	275	38.2	345
Summary ultramafic			50	sd_t	1.35	141	81.3	420
				med	2.17	245	4.95	178
Mafic								
	Little Bay	Dunnage zone-Notre Dame subzone	6	μ_t	2.38	0.039	0.766	646
				sd_t	0.345	0.016	1.28	127
	Little Deer	Dunnage zone-Notre Dame subzone	6	μ_t	6.00	0.315	1.84	406
				sd_t	1.21	0.155	2.18	153
	Whalesback	Dunnage zone-Notre Dame subzone	5	μ_t	7.55	1.05	0.596	493
				sd_t	2.38	0.454	1.29	51.2
	York Harbour	Humber zone-Bay of Islands ophiolite	6	μ_t	2.51	3.40	0.923	521
				sd_t	0.525	9.43	1.54	54.3
	Lasail	Semail ophiolite	14	μ_t	1.58	0.397	<dl	926
				sd_t	1.86	0.580	na	1414
	Ice	Slide Mountain terrane	5	μ_t	3.73	0.156	<dl	1.24
				sd_t	2.64	0.349	na	0.540
				μ_t	3.23	0.336	0.349	443
	Summary mafic		42	sd_t	2.70	0.479	0.815	412
				med	2.63	0.255	0.031	496
Bimodal-mafic								
	Aldermac	Abitibi greenstone belt	18	μ_t	4.04	2.18	1.80	421
				sd_t	1.89	1.81	2.08	84.4
	Caber	Abitibi greenstone belt	10	μ_t	<dl	<dl	<dl	176
				sd_t	na	na	na	95.1
	Dufault	Abitibi greenstone belt	20	μ_t	1.51	0.886	<dl	285
				sd_t	1.93	1.14	na	38.9
	Kidd Creek	Abitibi greenstone belt	28	μ_t	0.873	0.644	<dl	302
				sd_t	0.906	0.917	na	49.5
	Norbec	Abitibi greenstone belt	12	μ_t	0.800	0.383	2.03	331
				sd_t	0.607	0.407	1.34	24.8
	Normetal	Abitibi greenstone belt	5	μ_t	<dl	2.29	<dl	288
				sd_t	na	0.188	na	61.7
	Poirier	Abitibi greenstone belt	6	μ_t	2.52	0.033	6.78	161
				sd_t	0.217	0.035	2.83	11.1
	Waite-Amulet	Abitibi greenstone belt	33	μ_t	1.64	1.40	<dl	325
				sd_t	1.38	1.22	na	58.2
	Ming	Dunnage zone-Notre Dame subzone	31	μ_t	1.11	0.366	1.75	513
				sd_t	0.645	0.649	3.38	95.7
				μ_t	1.37	0.802	0.430	339
	Summary bimodal-mafic		163	sd_t	1.41	0.999	1.13	104
				med	1.09	0.866	0.024	330
Bimodal-felsic								
	Horne	Abitibi greenstone belt	24	μ_t	1.45	3.64	0.685	243
				sd_t	1.59	2.23	1.02	133
	McLeod	Abitibi greenstone belt	10	μ_t	2.89	0.212	<dl	232
				sd_t	4.51	0.255	na	43.0
	Perseverance	Abitibi greenstone belt	15	μ_t	<dl	0.235	<dl	235
				sd_t	na	0.584	na	42.9
	Quémont	Abitibi greenstone belt	32	μ_t	1.48	5.65	2.39	162
				sd_t	1.50	6.79	4.46	58.4
	Suffield mine	Ascot-Weedon Complex	6	μ_t	<dl	<dl	<dl	145
				sd_t	na	na	na	9.31
	Weedon	Ascot-Weedon Complex	7	μ_t	0.543	4.73	0.445	306
				sd_t	0.334	1.41	0.580	24.0

Table 2. (*Cont.*)

Ga	Se	Ag	In	Sn	Sb	Te	Au	Tl	Pb	Bi
71	82	107	115	118	121	128	197	205	208	209
0.079	0.029	0.011	0.070	0.017	0.221	0.007	0.004	0.006	0.004	0.006
0.149	0.030	0.026	0.147	0.027	0.501	0.012	0.007	0.011	0.007	0.004
0.235	1371	9.30	33.5	52.6	<dl	57.5	0.085	<dl	0.145	0.293
0.285	163	5.88	39.8	28.2	na	63.6	0.082	na	0.197	0.597
<dl	1195	2.94	5.78	44.9	<dl	24.9	0.006	<dl	0.048	0.040
na	129	3.92	1.35	9.02	0.344	11.8	0.012	na	0.020	0.033
1.22	375	1.51	1.75	216	<dl	19.3	0.031	<dl	0.138	0.159
1.56	496	1.66	2.54	42.8	na	31.4	0.023	na	0.139	0.085
3.77	503	32.3	3.36	123	<dl	23.3	0.045	<dl	0.155	0.282
2.90	222	25.7	2.30	101	na	14.8	0.048	na	0.163	0.229
1.74	719	15.6	4.16	100	0.070	25.5	0.039	0.001	0.129	0.190
2.19	526	19.6	2.93	91.1	0.136	19.3	0.047	0.001	0.150	0.160
1.53	669	8.82	4.18	78.4	0.023	25.4	0.033	0.000	0.092	0.172
2.50	110	1.95	12.4	2.12	<dl	5.64	0.024	0.030	7.60	7.96
0.135	59.4	1.34	1.40	0.393	na	0.902	0.057	0.028	1.89	11.0
1.89	378	94.8	36.3	0.171	3.88	2.08	<dl	0.758	7.14	1.12
0.432	54.6	9.64	3.49	0.087	1.26	2.33	na	0.211	3.09	0.366
8.65	549	70.2	7.78	3.28	34.2	2.27	<dl	0.830	17.4	<dl
0.389	265	48.0	6.28	3.26	24.0	0.173	na	0.551	2.53	na
0.484	264	120	17.8	1.81	<dl	2.63	0.031	<dl	5.74	0.252
0.322	7.15	141	14.2	0.544	na	0.326	0.043	na	2.16	0.550
5.43	40.0	13.2	10.2	4.69	<dl	0.712	0.031	0.008	15.2	0.223
7.08	54.8	10.2	6.22	6.72	na	1.02	0.037	0.009	22.3	0.227
17.8	112	8.30	29.3	0.443	0.401	0.083	<dl	0.024	11.4	0.162
15.0	29.5	7.45	18.2	0.168	0.350	0.009	na	0.036	14.4	0.191
3.84	160	29.5	14.9	1.48	0.483	1.81	0.013	0.084	9.26	0.391
4.74	154	42.5	11.1	1.55	1.18	1.89	0.023	0.192	8.19	0.511
2.41	130	20.5	13.9	1.32	0.051	2.22	0.004	0.019	7.57	0.287
1.93	149	124	35.7	42.9	0.344	0.536	0.021	0.009	1.98	1.10
2.06	32.7	86.1	43.6	51.5	0.390	0.519	0.022	0.012	1.23	1.08
0.285	461	9.01	16.4	12.4	0.995	13.9	0.013	0.013	2.64	0.086
0.221	460	25.3	16.7	4.07	0.690	13.3	0.018	0.018	1.93	0.078
2.96	271	88.4	179	88.9	0.421	0.048	0.045	0.048	2.37	1.94
6.16	399	57.3	140	26.9	0.240	0.013	0.064	0.058	1.08	1.60
2.55	114	170	211	93.3	<dl	0.053	<dl	0.022	3.39	0.770
1.82	98.9	283	216	83.6	na	0.045	na	0.019	2.95	0.904
0.349	259	222	61.1	140	<dl	0.023	0.010	<dl	0.925	0.881
0.190	309	327	90.1	110	na	0.006	0.013	na	0.498	1.32
4.03	139	1491	19.6	35.2	1.96	0.088	<dl	0.007	2.86	3.90
0.526	38.3	182	1.11	1.81	1.36	0.025	na	0.006	2.10	1.81
<dl	76.6	32.0	6.72	21.4	<dl	0.756	0.056	<dl	3.82	1.65
na	3.86	3.71	0.360	4.12	na	0.224	0.036	na	0.767	0.532
6.16	60.6	91.0	30.0	70.1	0.427	0.049	0.123	0.059	2.27	1.89
11.7	41.5	73.8	11.1	28.1	0.507	0.022	0.176	0.075	1.06	1.68
1.99	230	187	14.9	99.2	1.15	1.86	0.050	0.017	2.72	0.346
1.67	188	352	18.4	83.2	0.898	0.880	0.042	0.026	1.83	0.385
1.48	140	94.8	41.8	69.9	0.447	0.293	0.026	0.020	2.41	0.990
1.80	123	113	49.5	58.6	0.460	0.509	0.036	0.029	1.57	1.11
1.11	122	67.5	32.0	68.3	0.410	0.076	0.016	0.012	2.25	1.07
7.46	1349	61.5	57.4	82.2	0.357	5.43	0.285	0.016	4.52	6.36
9.83	1193	16.3	67.5	27.8	0.287	4.09	0.198	0.018	2.92	6.35
0.182	57.2	35.0	37.0	22.0	<dl	0.525	0.017	0.010	1.84	0.310
0.259	67.6	13.3	47.7	6.32	na	0.205	0.027	0.012	1.09	0.195
1.90	782	3.98	53.3	23.0	1.04	8.43	0.019	<dl	2.79	0.096
2.28	482	4.44	13.0	7.24	0.981	11.2	0.011	na	2.07	0.172
0.160	392	204	51.3	79.6	0.303	0.506	0.832	0.008	3.95	2.04
0.062	104	150	26.0	37.0	0.187	0.295	1.03	0.009	2.08	2.04
129	31.9	16.7	5.44	134	6.16	0.064	0.025	0.015	4.88	0.110
19.4	3.92	10.1	0.919	26.2	8.01	0.019	0.031	0.021	2.65	0.087
4.35	42.3	123	9.13	4.32	0.393	1.35	<dl	0.047	7.45	1.14
0.633	2.65	2.61	0.364	0.729	0.045	0.238	na	0.011	6.61	0.184

Table 2. (*Cont.*)

Setting	Name	Host tectonostratigraphic complex	n	Element	Mn	Co	Ni	Zn
Bimodal-felsic								
	Adak	Bothnia-Skelleftea unit	11	μ_t	3.20	1.30	0.574	454
				sd_t	0.812	0.566	1.27	93.8
	Brannmyran	Bothnia-Skelleftea unit	6	μ_t	3.70	2.02	0.359	576
				sd_t	0.393	0.817	0.832	51.2
	Lindsköld	Bothnia-Skelleftea unit	18	μ_t	3.75	2.73	0.775	483
				sd_t	1.47	2.82	1.20	87.8
	Rävliden	Bothnia-Skelleftea unit	6	μ_t	3.58	1.76	<dl	492
				sd_t	0.502	0.678	na	68.3
	Boundary	Dunnage zone-Exploits subzone	5	μ_t	2.34	13.3	<dl	486
				sd_t	0.421	2.31	na	64.9
	Duck Pond	Dunnage zone-Exploits subzone	7	μ_t	<dl	<dl	<dl	135
				sd_t	na	na	na	29.1
	Roman Ruins	Pual Ridge	8	μ_t	1.97	0.469	0.508	60.8
				sd_t	0.327	0.219	0.620	8.19
	Satanic Mills	Pual Ridge	7	μ_t	1.30	0.136	2.17	7.68
				sd_t	0.404	0.137	2.65	0.727
	Snowcap	Pual Ridge	6	μ_t	1.28	0.056	8.90	6.51
				sd_t	0.406	0.070	9.79	0.769
				μ_t	1.80	1.55	0.398	255
	Summary bimodal-felsic		168	sd_t	1.79	2.24	0.820	192
				med	1.84	1.19	0.093	225
Siliciclastic-mafic								
	Ely	Vermont copper belt	6	μ_t	27.5	3.08	2.90	375
				sd_t	10.5	1.01	2.25	28.2
	Pike Hill	Vermont copper belt	12	μ_t	119	7.44	1.79	519
				sd_t	26.7	8.31	2.16	62.7
				μ_t	92.0	5.12	2.16	466
	Summary siliciclastic-mafic		18	sd_t	62.7	5.42	2.35	107
				med	99.9	3.46	2.53	457
Siliciclastic-felsic								
	Bathurst	Bathurst Mining Camp	21	μ_t	1.94	0.309	<dl	324
				sd_t	0.576	0.142	na	30.9
	Halfmile Lake	Bathurst Mining Camp	5	μ_t	1.17	1.05	1.15	415
				sd_t	0.089	0.068	0.252	30.4
	Tétrault	Greenville Province	5	μ_t	2.57	14.6	3.03	159
				sd_t	0.308	6.28	3.69	32.9
	Aljustrel	Iberian Pyrite Belt	6	μ_t	4.23	0.169	0.348	367
				sd_t	5.79	0.316	0.892	25.5
	Aznalcollar	Iberian Pyrite Belt	6	μ_t	0.715	0.060	19.5	256
				sd_t	0.387	0.081	30.2	34.6
	Corvo	Iberian Pyrite Belt	28	μ_t	1.00	0.192	0.758	197
				sd_t	0.939	0.238	1.03	166
	Graça	Iberian Pyrite Belt	6	μ_t	0.618	0.074	0.599	314
				sd_t	0.653	0.100	1.59	74.9
	Lagoa Salgada	Iberian Pyrite Belt	7	μ_t	0.824	0.028	0.355	225
				sd_t	0.441	0.009	0.458	60.2
	Lousal	Iberian Pyrite Belt	7	μ_t	1.74	0.462	0.952	216
				sd_t	2.04	0.630	1.32	27.0
	Neves	Iberian Pyrite Belt	7	μ_t	2.39	0.050	2.73	192
				sd_t	0.487	0.039	2.93	126
	Tharsis	Iberian Pyrite Belt	4	μ_t	<dl	<dl	<dl	449
				sd_t	na	na	na	107
	Zambujal	Iberian Pyrite Belt	7	μ_t	3.42	0.231	20.4	178
				sd_t	3.72	0.268	23.7	19.2
	Draa Sfar	Jebilet Massif	6	μ_t	0.940	0.193	<dl	273
				sd_t	1.10	0.100	na	22.9
	Hajar	Jebilet Massif	6	μ_t	3.88	0.189	<dl	466
				sd_t	2.47	0.016	na	51.1
	Wolverine	Yukon-Tanana terrane	6	μ_t	2.68	5.42	43.4	1002
				sd_t	0.448	1.70	7.00	120
				μ_t	1.57	0.227	0.762	286
	Summary siliciclastic-felsic		127	sd_t	1.37	0.294	1.35	130
				med	1.51	0.184	0.260	286

Notes: μ_t = trimmed mean (0.2); sd_t = trimmed standard deviation; *med* = median; n = number of analyses; <dl = below detection limit; na = not applicable

Table 2. (*Cont.*)

| Ga	Se	Ag	In	Sn	Sb	Te	Au	Tl	Pb	Bi
0.739	509	84.7	85.8	81.6	0.379	0.483	0.043	<dl	8.06	1.08
0.455	47.5	6.39	18.7	21.4	0.317	0.165	0.054	na	3.58	0.549
1.46	428	87.3	68.4	179	1.05	0.462	0.068	<dl	4.50	2.26
0.140	35.8	1.97	9.60	26.1	0.876	0.065	0.068	na	1.98	1.18
1.17	329	85.4	47.9	104	1.91	0.823	0.017	0.008	4.83	0.779
1.46	156	53.2	31.8	45.0	1.91	0.281	0.020	0.009	3.17	0.433
0.523	599	79.9	73.5	73.3	0.339	0.770	0.094	<dl	7.73	1.00
0.107	23.2	2.37	13.5	8.24	0.117	0.198	0.078	na	2.14	0.532
0.257	271	1.11	6.78	45.2	8.33	1.35	0.021	0.024	16.0	0.799
0.117	18.4	0.736	0.939	3.78	5.71	0.588	0.023	0.028	16.1	0.396
47.4	33.8	6.34	48.3	41.0	4.51	0.270	<dl	0.028	8.54	6.44
6.61	11.5	1.99	2.99	1.76	3.30	0.353	na	0.026	6.66	4.55
1.62	305	5.87	10.4	26.4	<dl	1.43	0.051	<dl	0.049	0.338
0.695	31.9	3.77	2.06	20.1	na	1.08	0.058	na	0.023	0.233
21.2	4.87	2.96	49.9	734	29.7	1.52	0.120	<dl	0.287	21.7
3.78	2.04	0.595	21.0	176	27.8	0.885	0.139	na	0.272	20.8
29.5	1.02	4.76	68.7	148	302	3.74	2.67	0.009	5.19	29.7
3.55	0.181	1.41	18.8	52.1	245	2.96	2.63	0.007	4.92	23.4
2.34	361	56.7	47.1	72.1	0.735	1.01	0.094	0.008	4.25	1.52
4.61	309	58.0	40.1	51.8	0.936	0.948	0.142	0.011	3.50	1.73
1.08	385	58.2	49.2	69.0	0.460	0.775	0.051	0.005	4.01	1.16
0.810	158	12.1	4.05	5.89	<dl	1.55	<dl	<dl	1.07	0.385
0.311	9.18	0.682	0.921	1.25	na	0.351	na	na	0.829	0.317
4.17	85.2	83.3	5.90	45.0	<dl	0.696	<dl	<dl	0.524	0.618
4.41	36.5	18.3	0.970	55.6	na	0.320	na	na	0.429	0.212
2.55	110	60.5	5.21	25.9	0.125	0.958	0.003	0.001	0.660	0.562
3.16	55.0	48.8	1.17	41.4	0.090	0.559	0.001	0.000	0.474	0.273
1.31	114	72.5	5.06	8.55	0.118	0.920	0.003	0.001	0.660	0.608
3.06	936	25.2	78.9	165	0.221	0.057	0.247	<dl	2.55	7.04
3.57	184	10.0	22.1	70.6	0.121	0.013	0.199	na	1.51	7.48
0.304	56.4	63.5	22.4	543	1.22	0.027	0.114	0.024	5.25	7.09
0.038	2.92	3.64	6.53	175	1.02	0.003	0.160	0.029	1.65	4.98
0.091	43.3	35.5	2.01	141	<dl	0.344	0.050	<dl	3.11	0.787
0.021	1.90	28.4	1.48	94.6	na	0.049	0.063	na	0.765	0.517
4.36	51.7	4.60	77.5	321	4.56	0.049	0.023	0.209	2.29	0.846
3.10	9.05	2.02	9.30	198	6.10	0.031	0.025	0.334	2.97	0.880
0.614	84.0	0.322	238	375	3.03	0.026	<dl	0.039	2.63	0.185
0.077	5.22	0.331	5.59	27.2	1.95	0.009	na	0.040	2.52	0.197
7.35	407	3.84	278	771	5.35	0.069	0.023	0.040	7.89	8.45
5.96	484	3.21	124	349	4.58	0.038	0.029	0.071	7.26	7.96
3.33	267	5.36	320	1007	6.00	0.036	0.007	0.042	4.14	1.66
1.41	132	3.02	134	415	5.85	0.018	0.006	0.048	4.25	1.54
0.429	187	8.75	54.0	714	18.5	0.030	0.086	0.091	10.8	3.95
0.230	29.8	9.14	34.1	301	17.6	0.006	0.115	0.112	8.98	2.88
0.297	97.4	5.31	549	312	21.2	0.048	0.042	0.026	17.9	16.2
0.156	24.9	2.24	147	43.3	9.67	0.028	0.080	0.018	5.96	6.36
1.39	48.5	5.44	421	700	3.46	0.023	0.005	<dl	2.72	0.348
0.797	51.5	1.95	233	700	2.98	0.008	0.002	na	1.77	0.467
4.66	2.40	315	0.433	434	0.519	0.016	0.017	0.053	0.850	<dl
0.263	1.01	15.4	0.288	189	0.458	0.007	0.011	0.039	0.889	na
7.85	1097	0.733	235	1058	5.00	0.043	0.006	0.015	4.45	4.90
1.08	56.9	0.131	21.0	79.7	1.73	0.009	0.005	0.010	3.27	1.66
0.283	98.7	3.26	7.68	35.0	4.18	0.070	0.281	0.016	23.3	28.7
0.047	2.22	0.925	2.15	15.0	2.71	0.013	0.103	0.012	7.77	13.3
12.6	8.38	61.4	5.56	856	7.90	0.033	0.030	0.027	11.1	0.170
4.12	5.79	4.69	1.26	64.6	7.15	0.009	0.018	0.035	6.84	0.148
0.613	4310	10.1	1.61	101	2.81	0.084	0.015	0.024	0.360	1.85
0.070	236	2.27	0.182	13.3	2.06	0.011	0.011	0.032	0.322	0.655
2.79	316	9.62	140	467	3.39	0.050	0.045	0.019	4.65	4.36
3.43	466	10.9	153	415	3.81	0.029	0.067	0.026	4.80	5.49
1.92	178	7.38	104	437	2.92	0.048	0.026	0.012	3.50	3.13
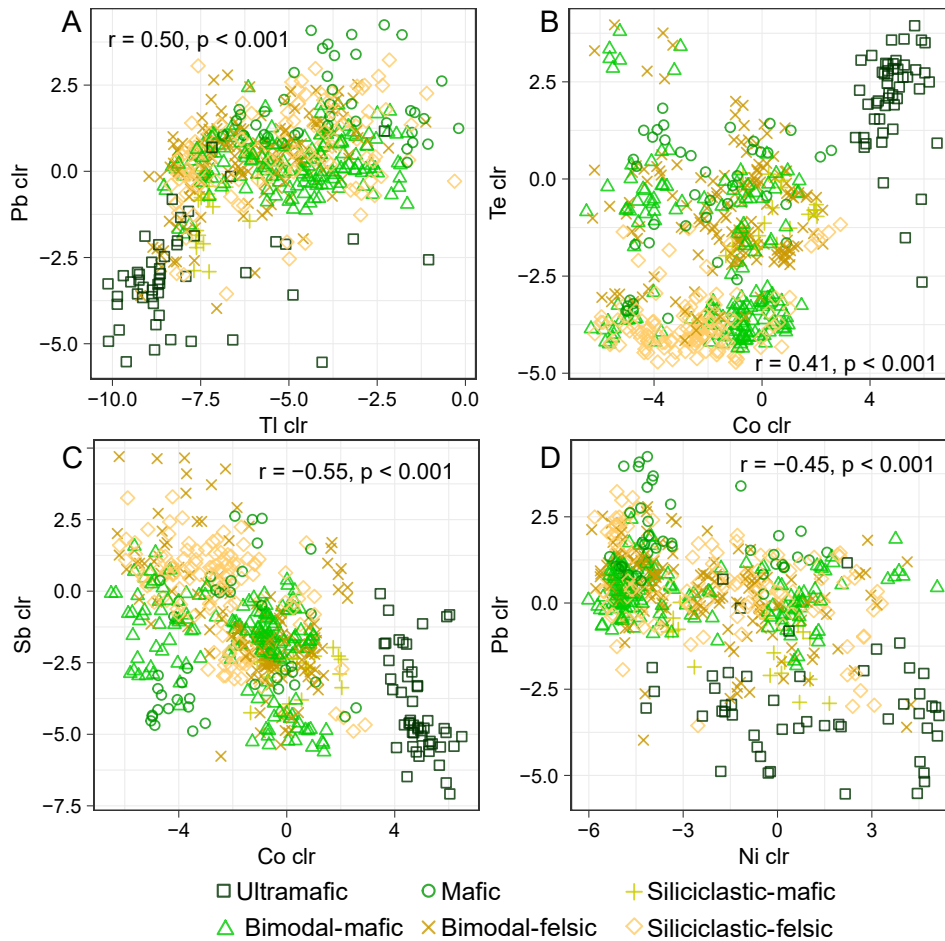
Fig. 4. Binary diagrams of elements with highest Pearson correlation coefficients showing grouping in function of the geologic setting of the deposit. Chalcopyrite from ultramafic deposits is separated from the remaining classes in most diagrams, to a lesser extent in mafic and siliciclastic-felsic subtypes. Abbreviations: clr = centered-log transformation, r = Pearson correlation coefficient, p = p-value.

situated opposed to the UM (ultramafic) pole, suggesting a negative correlation with chalcopyrite from the ultramafic subtype (Fig. 6B). Relatively high regression coefficients of Co, Te, Ag, Pb, Bi, and Sb confirm a high correlation with chalcopyrite from the ultramafic deposit subtype (Fig. 6C). Samples from mafic deposits form a group associated with positive scores on latent variable t_2 separated from the remaining subtypes (Fig. 6A) and are positively correlated with Pb and Ga and negatively correlated with Sn (Fig. 6B, C). The positive correlation between Pb and Ga and the mafic subtype is shown in loading plots qw*3-qw*4 (Pb; App. 2, Fig. A5A) as well as loading plots qw*7-qw*8 and qw*9-qw*10 (Ga; App. 2, Fig. A5C, D), and by their higher concentrations than in the other deposit subtypes (Table 2). Also, regression coefficients show that Ag, Bi, and Co have a moderate negative correlation with chalcopyrite from mafic deposits (Fig. 6C). These relations are evidenced by the loading plots in Appendix 2, Figure A5, where Ag, Bi, and Co tend to be opposed to the M (mafic) pole. Chalcopyrite from siliciclastic-mafic deposits forms a compact group in the upper left part of the score plot t_1-t_2, relatively near to the center (Fig. 6A). These samples are principally correlated with Mn, which records a high positive regression coefficient (Fig. 6C). Lead, Tl, Sn,

and Au have moderate negative regression coefficients, which is evidenced by loading plots in Appendix 2, Figure A5, which show a relative opposed position to the SM (siliciclastic-mafic) pole. Chalcopyrite from bimodal-mafic and bimodal-felsic deposits overlaps considerably, plotting principally at the center of the score plot t_1-t_2 and loading plot qw*1-qw*2 (Fig. 6A, B). However, according to the regression coefficients, chalcopyrite from the bimodal-mafic deposit subtype has a high positive correlation with Ag and moderate positive correlation with Tl and Zn, whereas Pb and to a lesser extent Co, Mn, Sb, and Te record a moderate negative correlation. Loading plots qw*3-qw*4, qw*7-qw*8, and qw*9-qw*10 illustrate these relations (App. 2, Fig. A5A, C and D, respectively). On the other hand, chalcopyrite from the bimodal-felsic subtype records high positive regression coefficients in Te and Pb, whereas Ag, Au, and Sb show moderate positive values. In contrast, Tl and Zn show negative regression coefficients (Fig. 6C). The association between Te, Pb, and chalcopyrite from bimodal-felsic settings is evidenced in the loading plot in Appendix 2, Figure A5. Chalcopyrite from the siliciclastic-felsic subtype is grouped in the lower right quadrant (Fig. 6A), positively correlated with Sn and to a lesser extent Se, and negatively correlated with Ag and Te (Fig. 6B, C).
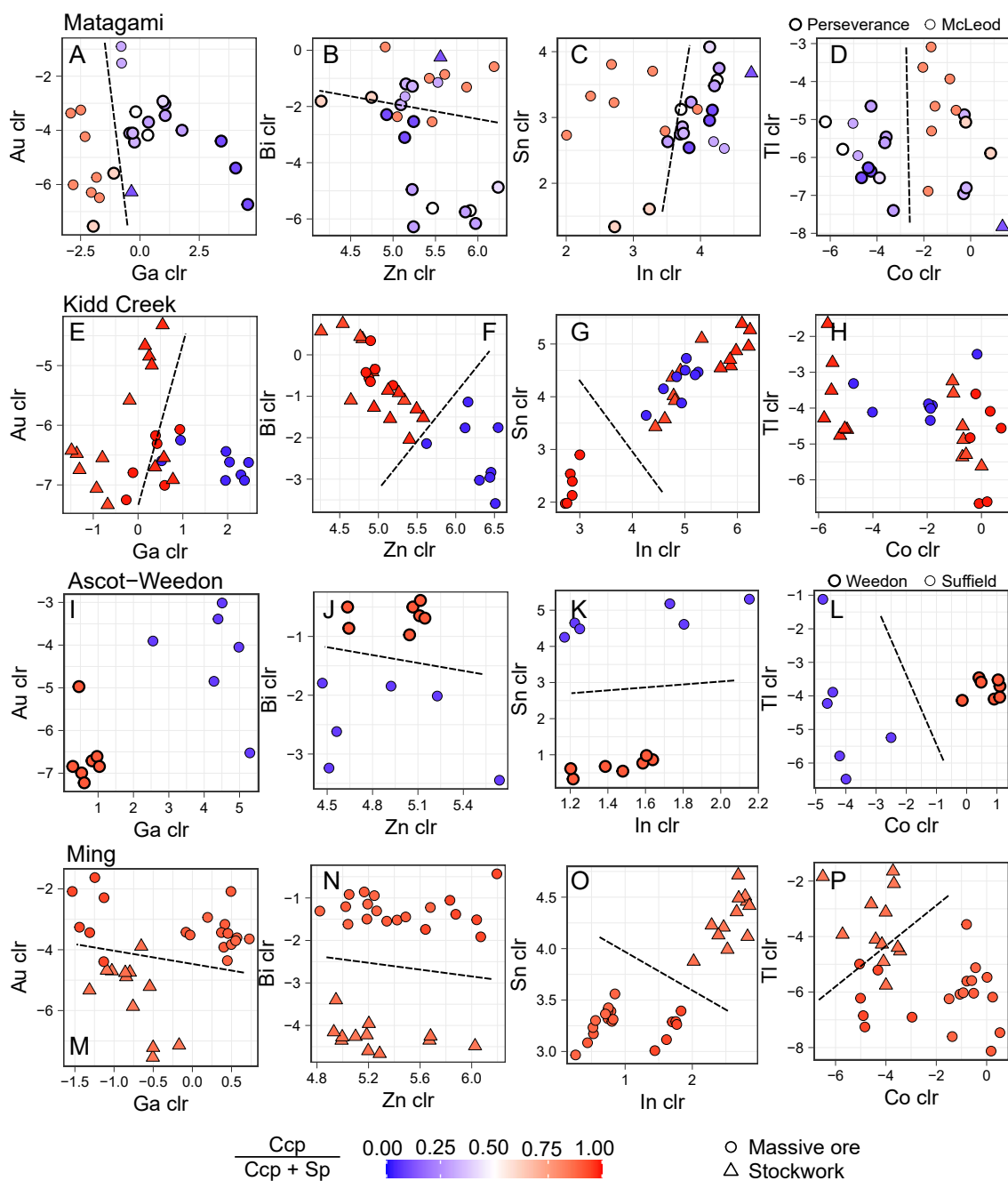
Fig. 5. Binary diagrams by deposit showing differences between chalcopyrite in function of ccp/(ccp + sp) ratio and ore type. Dashed lines show approximate limits between groups. Clr = centered-log transformation.

In order to evaluate the impact of metamorphism, the sulfide proportion, and the ore type on PLS-DA results by VMS setting (Fig. 6), each factor was projected on score plot t_1-t_2 (Figs. 7, 8; App. 2, Figs. A6-A8). Among the studied VMS subtypes, only mafic, bimodal-mafic, bimodal-felsic, and siliciclastic-felsic have samples from different metamorphic facies. Chalcopyrite records similar trace element composition within each VMS subtype, independently of the metamorphic facies, although a certain grouping is observed within the siliciclastic-felsic subtype (Fig. 7B-F). In the mafic subtype, chalcopyrite from unmetamorphosed deposits has no significant difference from that from deposits with lower

greenschist metamorphism (Fig. 7B). In the bimodal-mafic subtype, chalcopyrite from deposits with lower and upper greenschist metamorphism shows similar trace element composition, with a considerable overlapping at the center of score plot (Fig. 7D). In the bimodal-felsic subtype, chalcopyrite from unmetamorphosed deposits (SMS deposits from PACMANUS) forms a group plotting at the edge and outside of the 95% confidence ellipse (Fig. 7E), which encloses approximately 95% of the data points (Michael et al., 2013). Chalcopyrite in unmetamorphosed deposits from the bimodal-felsic subtype plotting outside of the confidence ellipse has similar composition to that from siliciclastic-felsic
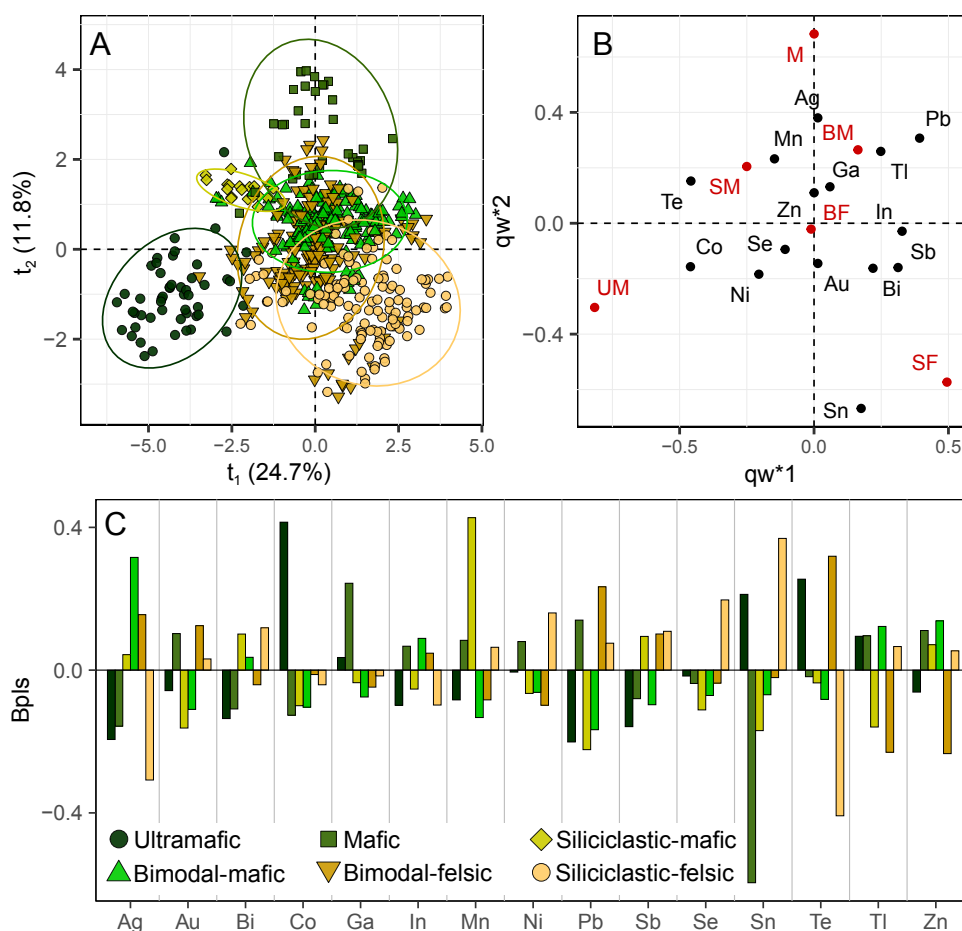
Fig. 6. Partial least squares discriminant analysis (PLS-DA) on trace elements in chalcopyrite by geologic setting of the deposit. Chalcopyrite from ultramafic, mafic, and siliciclastic-felsic settings shows an important separation from the other subtypes in score plot t₁-t₂ (A). According to loading plot (B) and regression coefficients (C), chalcopyrite from the ultramafic subtype is associated with a high content of Co (±Te), whereas chalcopyrite from mafic settings is correlated with high Ag and low Sn concentrations. In contrast, chalcopyrite from siliciclastic-felsic settings is associated with high Sn and low Ag concentrations. Abbreviations: BF = bimodal-felsic, BM = bimodal-mafic, Bpls = regression coefficients, M = mafic, SF = siliciclastic-felsic, SM = siliciclastic-mafic, UM = ultramafic. Confidence ellipses enclose approximately 95% of the data.

deposits with prehnite-pumpellyite metamorphism (Fig. 7E). In the siliciclastic-felsic subtype, chalcopyrite from deposits with upper and lower greenschist forms a group separate from those with prehnite-pumpellyite and, to a lesser extent, middle greenschist metamorphism, whereas chalcopyrite from deposits with amphibolite metamorphism plots outside of the confidence ellipse (Fig. 7F). Although metamorphism has an effect on the trace element contents in chalcopyrite, the modifications are smaller than the much larger variation in composition due to the different lithotectonic subtypes of VMS deposits, such that it does not significantly affect the trace element signature.

Similarly, trace element composition of chalcopyrite from samples with different sulfide proportions shows variation within each VMS subtype (Fig. 8; App. 2, Figs. A6-A8). Chalcopyrite from samples with different ccp/(ccp + sp) ratios in bimodal-mafic and bimodal-felsic VMS subtypes has similar trace element composition, with an important overlap at the center of the score plot (Fig. 8D, E). Although samples with different ccp/(ccp + sp), po/(po + py + py), and ccp/(ccp + po) ratios and different ore types show a slight grouping

within each VMS subtype (App. 2, Figs. A6-A8), the trace element composition of chalcopyrite remains constrained to the groups defined by the PLS-DA by VMS setting.

Since the ccp/(ccp + sp) ratio provides important information on location in the deposit and its distribution is skewed in whole data, a PLS-DA was carried out separately on chalcopyrite from bimodal-mafic (Fig. 9A, B), bimodal-felsic (Fig. 9C, D), and siliciclastic-felsic settings (Fig. 9E, F) as a preliminary approach. The PLS-DA results show that Ga, Sn, Tl, In, and to a lesser extent Sb and Zn are positively correlated with samples with a low ccp/(ccp + sp) ratio, whereas Bi, Co, Ni, Au, Se, Mn, and Te are associated with samples with a high ccp/(ccp + sp) ratio. Results for the bimodal-mafic setting show two clusters in the score plot t₁-t₂: (1) chalcopyrite from samples with a low ccp/(ccp + sp) ratio is associated with Ga, Sn, Tl, In, and to a lesser extent Sb and Zn, whereas (2) chalcopyrite from samples with a high ccp/(ccp + sp) ratio is principally associated with Te, Au, Se, Bi, Mn, Ni, and Co (Fig. 9A, B). Chalcopyrite from Cu-rich stockwork ores plots mainly with chalcopyrite from low ccp/(ccp + sp) ratio massive samples. Chalcopyrite from bimodal-felsic settings shows
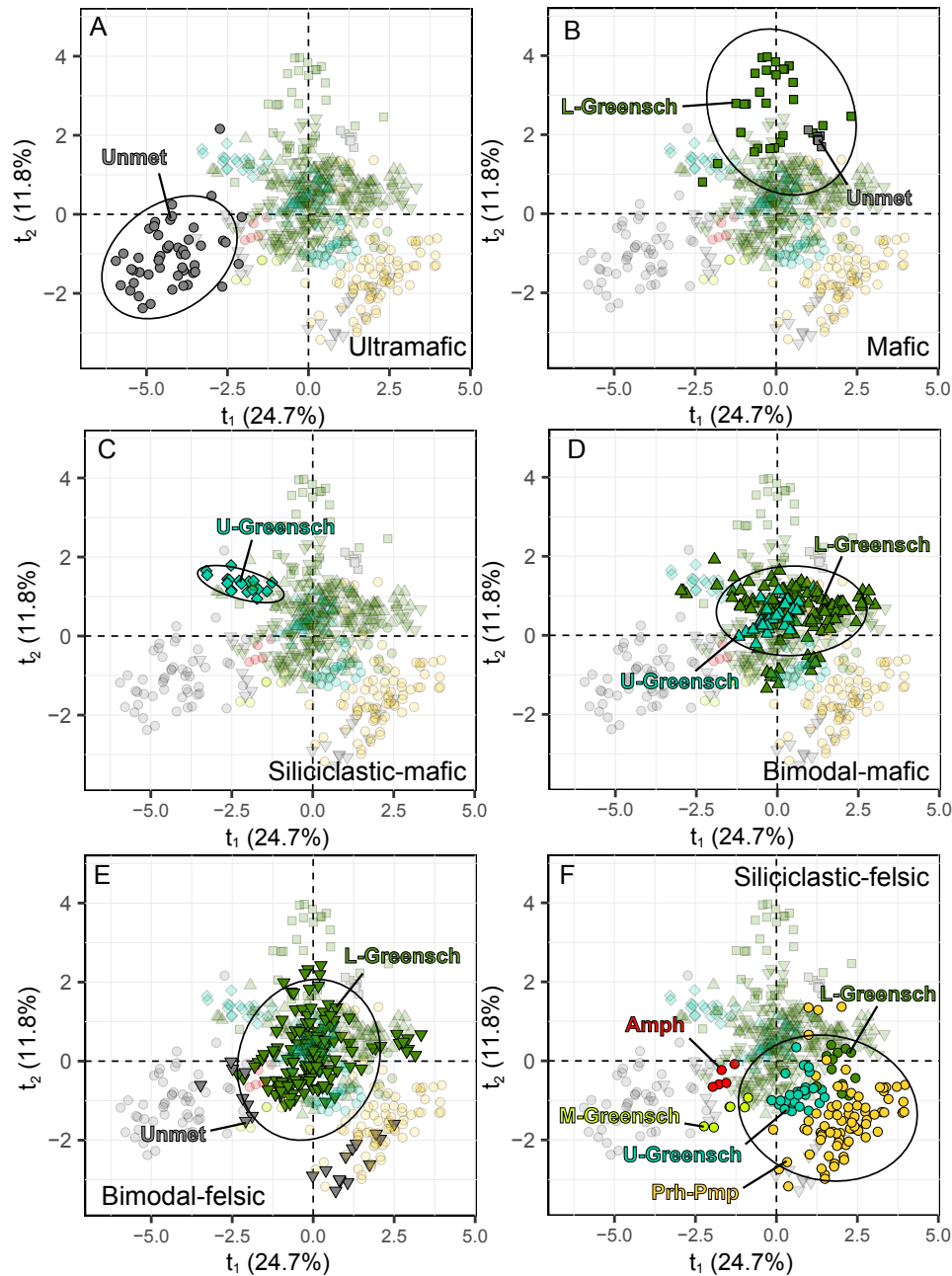
Fig. 7. Metamorphic grades of samples projected on partial least squares discriminant analysis (PLS-DA) results according to lithotectonic settings. (A) Samples from unmetamorphosed sea-floor massive sulfide (SMS) ultramafic-hosted deposits. (B) Mafic subtype group comprising samples from deposits with different metamorphic grades, showing similar trace element composition. (C) Samples from siliciclastic-mafic settings have upper greenschist metamorphism. (D, E) Chalcopyrite from bimodal subtypes showing similar trace element composition independent of metamorphic grade. (F) In siliciclastic-felsic settings, chalcopyrite from deposits with prehnite-pumpellyite, lower greenschist, upper greenschist, and to a lesser extent middle greenschist and amphibolite metamorphism has similar trace element composition. Abbreviations: Amp = amphibolite, L-Greensch = lower greenschist, M-Greensch = middle greenschist, Prh-Pmp = prehnite-pumpellyite, U-Greensch = upper greenschist, Unmet = unmetamorphosed.

similar relations according to ccp/(ccp + sp) ratios (Figs. 9C, D). Indium, Ga, Sn, and to a lesser extent Zn, Sb, and Pb are positively correlated with Zn-rich samples (Fig. 9D). To a lesser extent, in siliciclastic-felsic settings, chalcopyrite from Zn-rich samples (Fig. 9E) with relatively low ccp/(ccp + sp) ratios is associated with Ga, Mn, Zn, Sb, and marginally Pb and Ag (Fig. 9F).

Supervised classification and regression with Random Forest

Performance estimators in the optimized models, calculated from cross validation during training, are summarized in Table 3. Figure 10 shows results of recursive feature elimination, indicating that the optimal variables are as follows: Sn, Co, Te,
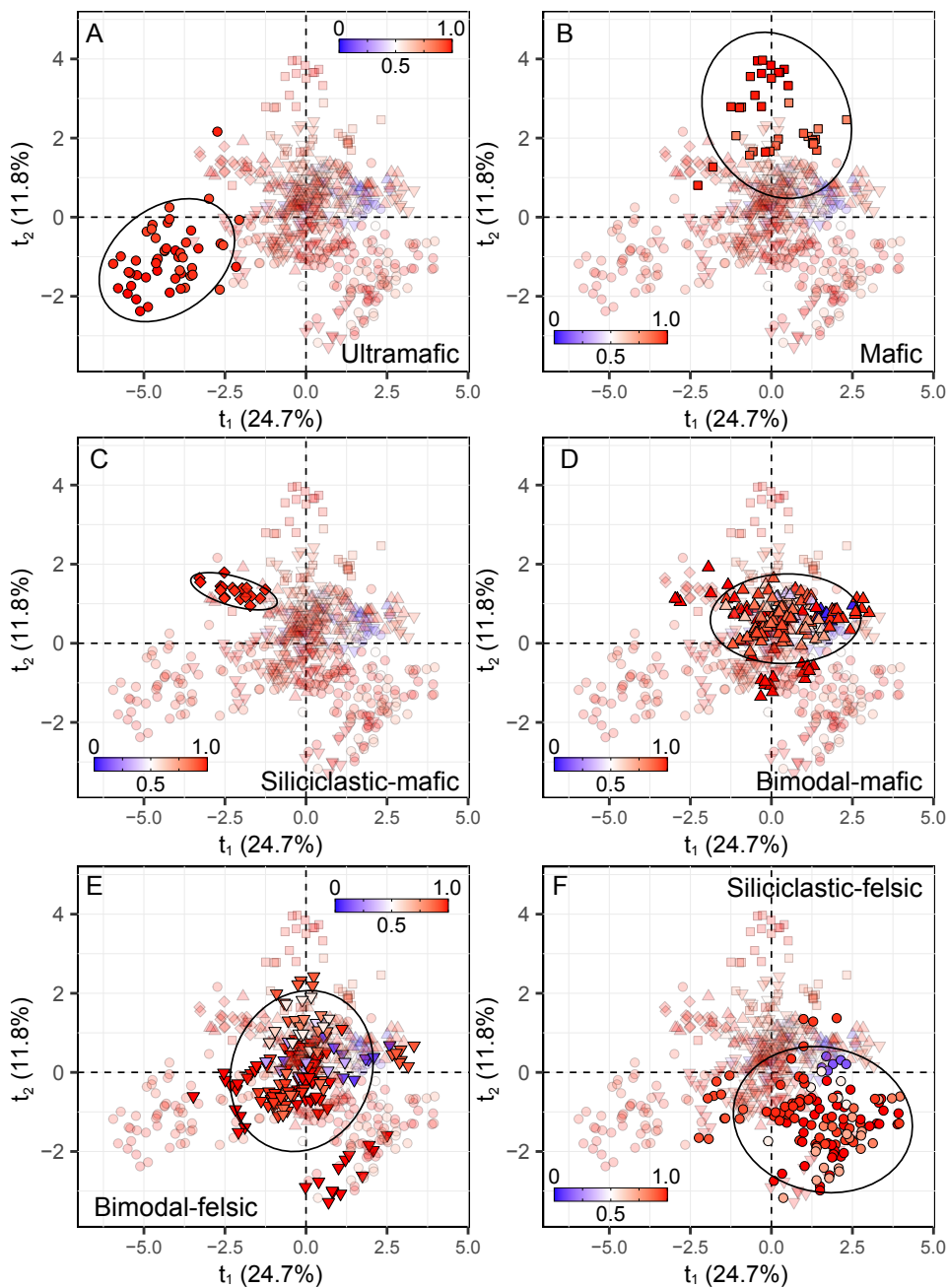
Fig. 8. The ccp/(ccp + sp) ratio (where ccp = chalcopyrite and sp = sphalerite) of the samples projected on partial least squares discriminant analysis (PLS-DA) results, showing that sulfide proportions have no impact on grouping according to lithotectonic setting.

In, Zn, Se, Ag, Sb, and Pb for RF model.01 (affinity); Co, Sn, Mn, Te, In, Pb, Se, Tl, Ni, Ag, Zn, and Bi for RF model.02-A (ultramafic vs. mafic vs. siliciclastic-mafic subtypes); and Te, Se, Zn, In, Ag, Sb, Co, Ga, Sn, and Ni for RF model.02-B (bimodal-mafic vs. bimodal-felsic subtypes). All models resulted in kappa values greater than 0.90. Minimum training accuracy was 0.95, recorded by RF model.02-B, whereas maximum training accuracy was 0.99, reported by RF model.02-A. Classification outcomes for the test data are presented in Tables 4, 5, and 6 for each model. Random Forest model.01 correctly identified the classes from chalcopyrite trace element composition with an overall accuracy of 0.98, whereas RF model.02-

A and RFmodel.02-B recorded accuracies of 0.97 and of 0.96, respectively.

In addition to classifying chalcopyrite by VMS setting using trace elements, an RF regression model was developed for bimodal-mafic, bimodal-felsic, and siliciclastic-felsic deposit subtypes to predict the ccp/(ccp + sp) ratio in samples. Figure 11 summarizes the performance and shows the variable importance for each model. The regression model performs better for chalcopyrite from bimodal-mafic VMS settings (Fig. 11A, B), reporting the lowest errors in training (root mean squared error [RMSE]tr: 0.09; mean absolute error [MAE] tr: 0.06) and test data (RMSEtest: 0.06; MAEtest: 0.06), and
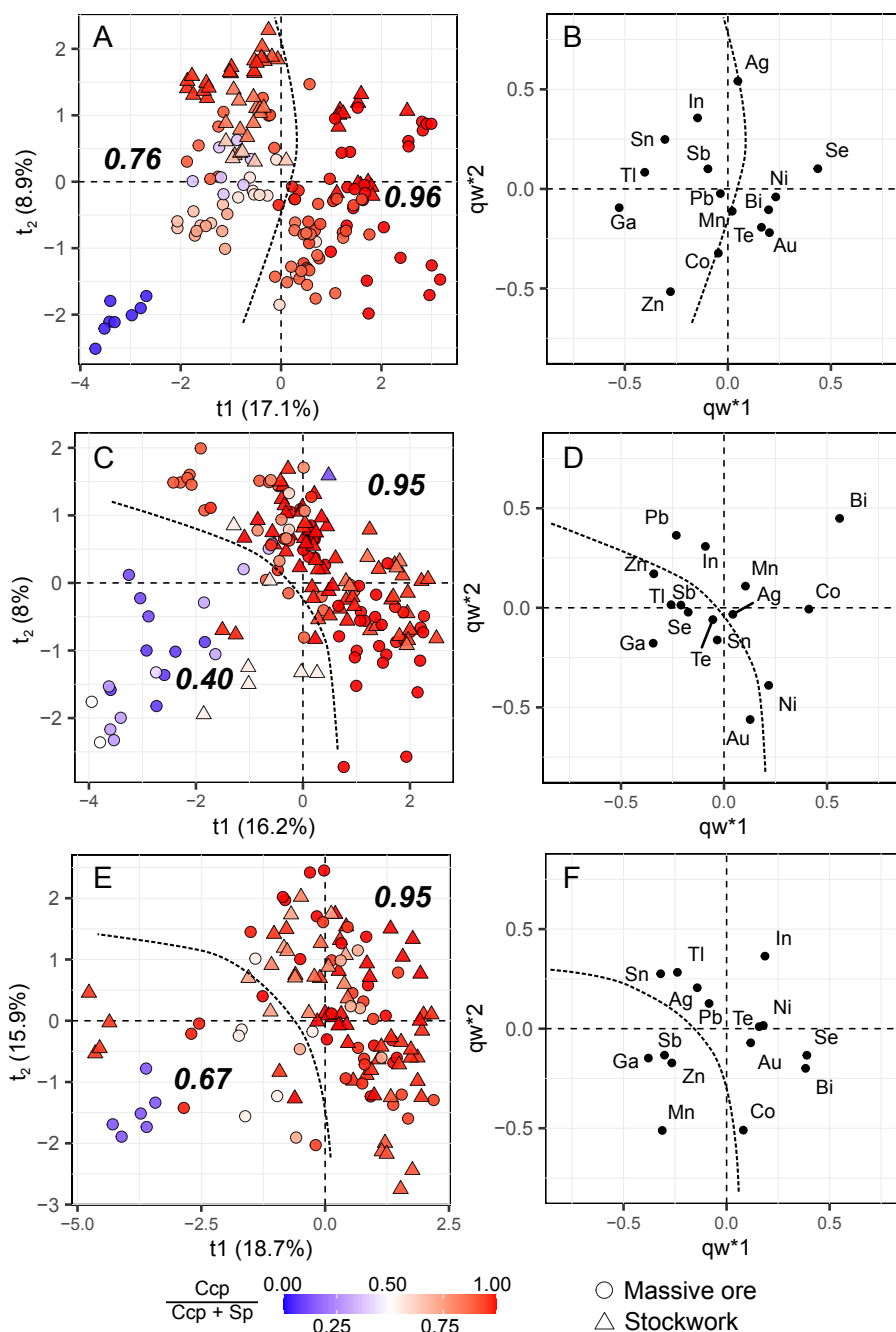
Fig. 9. Partial least squares discriminant analysis (PLS-DA) on trace elements in chalcopyrite from bimodal-mafic (A, B), bimodal-felsic (C, D), and siliciclastic-felsic deposits (E, F) according to ccp/(ccp + sp) ratio (where ccp = chalcopyrite and sp = sphalerite). In bold-italic: trimmed means of chalcopyrite-sphalerite proportion of each group delineated by the dotted lines.

a higher coefficient of determination (R^2tr: 0.84; R^2test: 0.93). In bimodal-felsic settings (Fig. 11C, D), RF regression model errors for training (RMSEtr: 0.12; MAEtr: 0.08) and test data (RMSEtest: 0.12; MAEtest: 0.07) are slightly higher, whereas the coefficient of determination is lower (R^2tr: 0.78; R^2test: 0.79), indicating a performance slightly lower than bimodal-mafic. Finally, in siliciclastic-felsic settings (Fig. 11E, F), the performance of the RF regression model is the lowest, reporting the highest errors in training (RMSEtr: 0.13; MAEtr: 0.09) and test data (RMSEtest: 0.14; MAEtest: 0.10) as well

as the lowest coefficient of determination (R^2tr: 0.50; R^2test: 0.55). The results of RF regression models confirm that there is potential to predict ccp/(ccp + sp) ratio and determine Cu-rich and Zn-rich zones in bimodal-mafic, bimodal-felsic, and siliciclastic-felsic VMS settings, as suggested by PLS-DA results.

Discussion

Numerous parameters affect the trace element composition of sulfides in VMS deposits, such as host rock composition,

Table 3. Classification Performance Estimators of the Random Forest Models

Estimator	RF model.01 (Affinity)		RF model.02-A (M-UM)		RF model.02-B (BMF)	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Accuracy (Acc)	0.97	0.03	0.99	0.04	0.95	0.04
Kappa (K)	0.95	0.05	0.98	0.06	0.91	0.07
F1-score (F1)	0.97	0.03	0.99	0.04	0.95	0.04
Specificity (Spe)	0.98	0.02	0.99	0.02	0.97	0.05
Precision (Pre)	0.98	0.02	0.99	0.02	0.97	0.05
Recall (Recall)	0.96	0.04	0.98	0.07	0.94	0.06
Balanced accuracy (Bacc)	0.97	0.03	0.99	0.04	0.95	0.04

Abbreviations: BMF = bimodal-mafic-felsic, M = mafic, UM = ultramafic

temperature, f_{O_2}, a_{S_2} and pH of the fluid in the reaction zone, volatile content, as well as cooling, mixing with seawater, and boiling (Barrie and Hannington, 1999; Franklin et al., 2005; Galley et al., 2007; Hannington, 2014; Fuchs et al., 2019). During precipitation, trace elements partition between chalcopyrite and coprecipitating sulfides, such as sphalerite, galena, and pyrite/pyrrhotite (Reich et al., 2013a; Lockington et al., 2014; Genna and Gaboury, 2015; George et al., 2015, 2016, 2018b; Carvalho et al., 2018; Torró et al., 2022; Xing et al., 2022). Although the Cu-Zn zonation model is consistent in most deposits and at a variety of scales, locally VMS and SMS are often paragenetically complex (Hannington, 2014). Deposits can be affected by different thermal histories involving continuous dissolution and precipitation processes (i.e., zone refining) and potentially trace element remobilization, as shown by replacement of preexisting sulfides (Eldridge et al., 1983; Ohmoto, 1996; Hannington, 2014).

Trace element composition of chalcopyrite as a tool to sulfide assemblage recognition

Based on our PLS-DA results, at the deposit scale, chalcopyrite from Cu-rich samples with high ccp/(ccp + sp) ratios is enriched in Bi, Co, Ni, Au, Ag, Te, and Se (Figs. 5, 9; App. 2, Figs. A3, A4). According to Hannington (2014), these elements correspond to a typical high-temperature polymetallic assemblage, including Ag transport under reducing conditions. Similar results are reported in pyrite from high-temperature Cu-rich VMS ore zones, which also are enriched in Se, Bi, and Co (Genna and Gaboury, 2015). Silver-bearing chalcopyrite has been reported in VMS deposits (Broken Spur vent field: Butler and Nesbitt, 1999; Izok Lake: Harris et al., 1984) and other hydrothermal systems (Cook and Chryssoulis, 1990; Cook et al., 2011; Reich et al., 2013b, 2020) incorporated in solid solution. In contrast, chalcopyrite from Zn-rich samples with low ccp/(ccp + sp) ratios is high in Ga, Sn, In, Sb, and to a lesser extent Tl (Figs. 5, 9; App. 2, Figs. A3, A4), typical of low-temperature assemblages (Hannington, 2014). In Zn-rich samples, chalcopyrite commonly replaces pyrite and sphalerite, with the latter showing typical replacement textures such as chalcopyrite disease (Fig. 2D, E; Barton and Bethke, 1987; Bortnikov et al., 1991; Nagase and Kojima, 1997; Hannington, 2014). Previous studies have demonstrated that sphalerite and galena are the principal hosts of elements such as Ga, In, Sb, and Tl in hydrothermal deposits (Lockington et al., 2014; George et al., 2015, 2016; Carvalho et al., 2018; Torró et al., 2022; Xing et al., 2022). Likewise, pyrite is also known to be an important sink of Tl and Sb, in particular in Zn-rich zones, among others (Genna and Gaboury, 2015; Dehnavi et
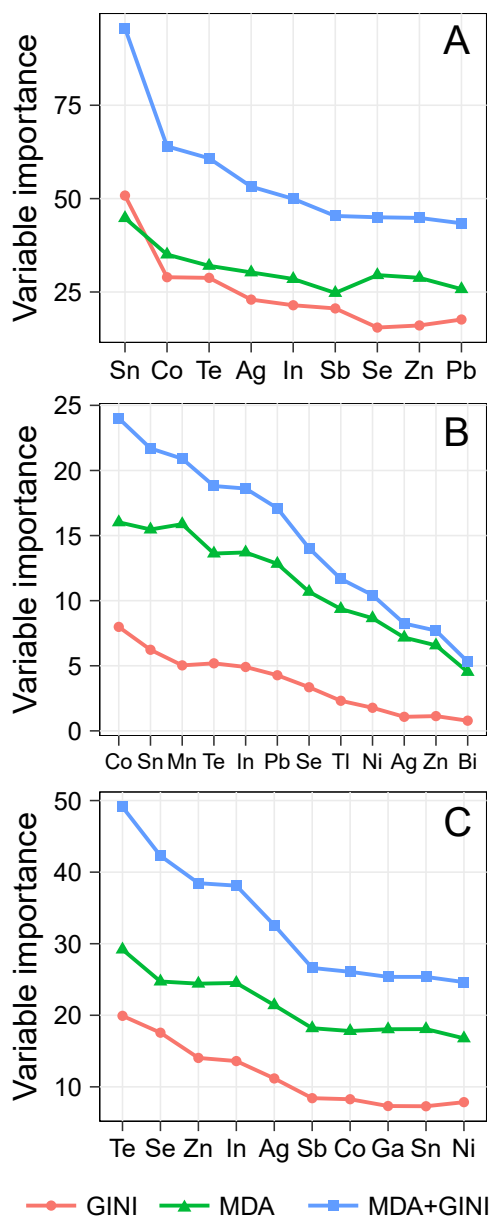

Fig. 10. Variable importance measures of the optimal elements after recursive feature elimination for (A) RF model.01 (affinity), (B) RF model.02-A (mafic-ultramafic-siliciclastic mafic), and (C) RF model.02-B (bimodal mafic-bimodal felsic). Note: GINI = Gini index, MDA = mean decrease in accuracy, MDA + GINI = total scores.

Table 4. Confusion Matrix and Performance Estimators for
RF Model.01 on Test Data

	Actual			
		Mafic-ultramafic	Bimodal	Felsic	Total
Predicted	Mafic-ultramafic	**31**	0	0	31
	Bimodal	0	**99**	3	102
	Felsic	0	0	**35**	35
	Total	31	99	38	**168**
	accuracy by class	1.00	1.00	0.92	
	overall accuracy	0.98			
	recall	0.97			
	kappa	0.97			

Note: Bold numbers indicate samples correctly classified in each class, allowing for calculation of overall accuracy

Table 5. Confusion Matrix and Performance Estimators for
RF Model.02-A on Test Data

	Actual			
		Ultramafic	Mafic	Siliciclastic-mafic	Total
Predicted	Ultramafic	**15**	1	0	10
	Mafic	0	**10**	0	1
	Siliciclastic-mafic	0	0	**5**	5
	Total	15	11	5	**31**
	accuracy by class	1.00	0.91	1.00	
	overall accuracy	0.97			
	recall	0.97			
	kappa	0.96			

Note: Bold numbers indicate samples correctly classified in each class, allowing for calculation of overall accuracy

Table 6. Confusion Matrix and Performance Estimators for
RF Model.02-B on Test Data

	Actual		
		Bimodal-mafic	Bimodal-felsic	Total
Predicted	Bimodal-mafic	**48**	3	47
	Bimodal-felsic	1	**47**	3
	Total	49	50	**99**
	accuracy by class	0.98	0.94	
	overall accuracy	0.96		
	recall	0.96		
	kappa	0.95		

Note: Bold numbers indicate samples correctly classified in each class, allowing for calculation of overall accuracy

al., 2018; Ren et al., 2021). Probably during the zone refining process, a successive remobilization of these elements from sulfides previously formed at lower temperature (e.g., sphalerite, galena, or pyrite) led to their incorporation into chalcopyrite at different replacement degrees. Chalcopyrite still forms at high temperature; however, it overprints the low-temperature assemblage. Differences between chalcopyrite composition from massive and stockwork ores in some deposits are not systematic, suggesting that conditions of precipitation are similar.

Trace element composition of chalcopyrite as a tool for VMS setting recognition

Partial least squares discriminant analysis results demonstrate that lithostratigraphic setting of the VMS-SMS deposits (e.g., Franklin et al., 2005; Patten et al., 2022) is the main control on the trace element content in chalcopyrite (Fig. 6). Moreover, PLS-DA allows for determination of the element association for each lithostratigraphic setting, information not provided by RF classification model. Our results show that chalcopyrite from ultramafic-hosted SMS deposits (Logatchev and Irinovskoe hydrothermal fields) is significantly enriched in Co, Ni, and Te when compared to other subtypes (Figs. 3, 4). This is in agreement with results previously published from Logatchev (Wohlgemuth-Ueberwasser et al., 2015). Previous studies on tourmaline reported unusually high Ni and Co concentrations in orogenic gold deposits that are hosted in mafic-ultramafic rocks (Hazarika et al., 2015; Sciuba et al., 2021), possibly indicating a common enrichment process related to the host rock lithology. Ultramafic rocks have abundant ferromagnesian minerals and minor Ni-Co–enriched magmatic sulfides (e.g., olivine; Vasyukova and Williams-Jones, 2022). During fluid-rock interaction, Ni and Co can be leached, transported, and precipitated in the massive sulfide deposits (Barrie and Hannington, 1999; Hannington, 2014). High values of Te in chalcopyrite from ultramafic-hosted deposits may also have been derived from sulfide droplets in ultramafic rocks (Hattori et al., 2002). In contrast, Pb is significantly depleted in chalcopyrite from ultramafic-hosted deposits relative to other settings (Figs. 3, 4, and 6). Chalcopyrite from Cu-rich samples shows negative correlation coefficients between Pb-Se (r = –0.41) and Pb-In (r = –0.39), whereas Pb is positively correlated with Tl (r = 0.35) (App. 2, Fig. A2B), suggesting that temperature could be a factor controlling concentration of Pb and Tl (Hannington, 2014). Selenium and In show strong correlation with Cu in most VMS deposits, which form stable H_2Se and $InCl_3$ complexes at high temperatures, respectively (Ishihara and Endo, 2007; Hannington, 2014; Carvalho et al., 2018). Additionally, ultramafic rocks are characterized by low Pb concentrations compared to the other rock types (Hofmann, 1988; Workman and Hart, 2005; Digis Team, 2022), which also influences the low Pb contents of the ultramafic VMS subtype.

Most trace elements in chalcopyrite from the mafic subtype of VMS deposits do not show significant differences compared to chalcopyrite from the other settings (Fig. 3). Although Pb has a low average concentration in mafic rocks (3.6 ppm in basalt; 4.2 ppm in gabbro; Digis Team [2022] database), it is the only element showing a trimmed mean in chalcopyrite higher than for the other subtypes (9.26 ppm; Table 2; Fig. 3). The association between Pb and mafic-hosted deposits is also evidenced by high positive regression coefficients from PLS-DA (Fig. 6C) and loading plot qw*3-qw*4 (App. 2, Fig. A5A). According to our results, chalcopyrite from the Whalesback (17.4 ppm, *n* = 5), Lasail (15.2 ppm, *n* = 14), and to a lesser extent Ice (11.4 ppm, *n* = 5) deposits has higher concentration of Pb than the remaining mafic VMS deposits (Table 2). Mafic VMS deposits are on average dominated by Cu-rich and Pb-Zn–poor sulfide assemblages with scarce to absent galena (Franklin et al., 2005; Galley et al., 2007; Hannington, 2014).
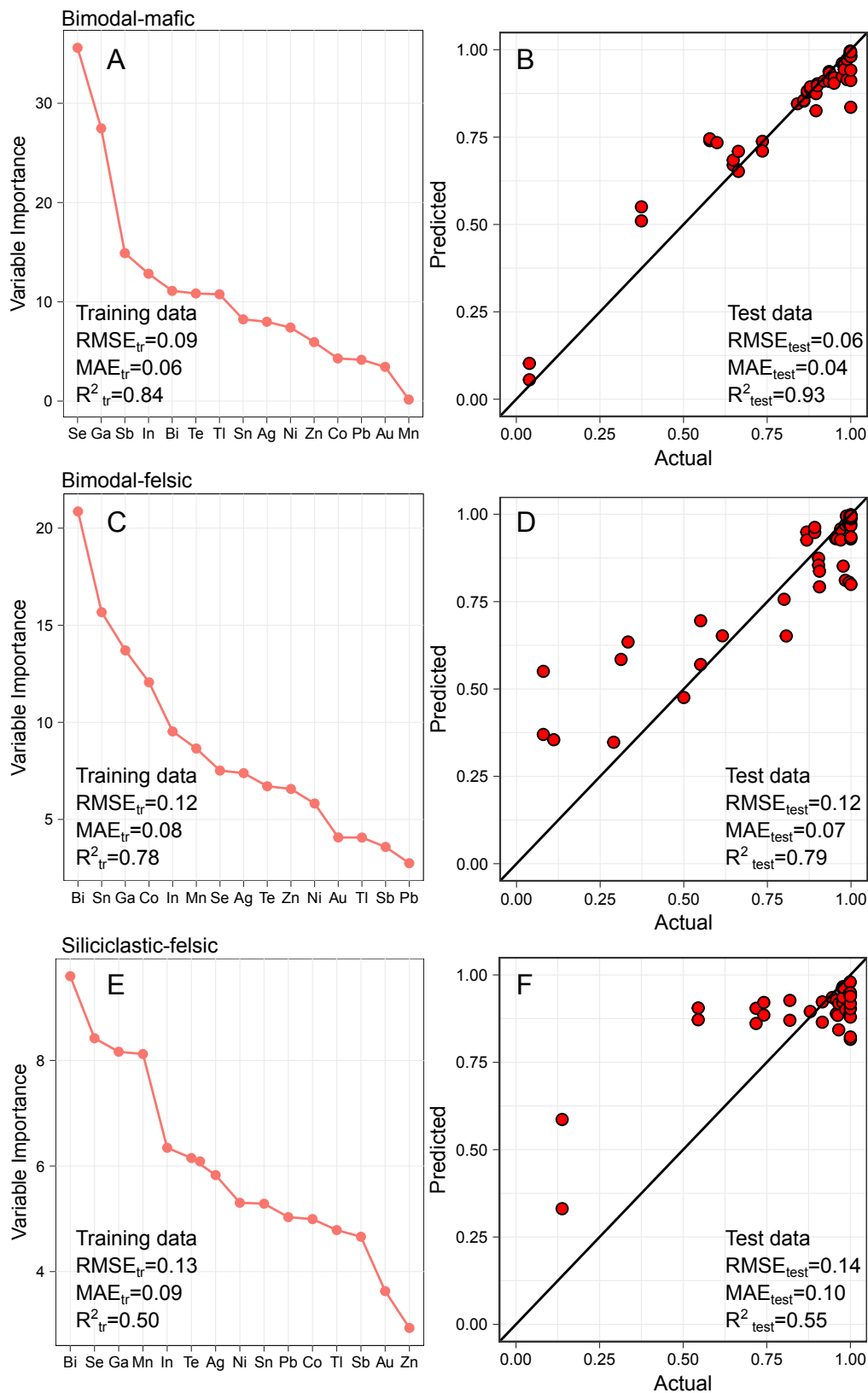
Fig. 11. Random forest regression model results for chalcopyrite from bimodal-mafic (A, B), bimodal-felsic (C, D), and siliciclastic-felsic (E, F) settings. Variable importance is shown at the left side with the performance metrics for training data. Actual and predicted ccp/(ccp + sp) ratio values for test data are shown at the right side with their corresponding performance metrics. Abbreviations: MAE = mean absolute error, RMSE = root mean squared error, R_2 = coefficient of determination.

According to George et al. (2016), chalcopyrite may be an important trace element host of Pb in the absence of sphalerite and/or galena. The high Pb content in chalcopyrite from the mafic subtype could be related to the near absence of galena, which led to more availability of Pb to partition into chalcopyrite. In addition, Pb can be efficiently mobilized not only from the sheeted dike zones (i.e., epidosite zones) but also from Sb- and Pb-enriched plagiogranites during hydrothermal alteration of the oceanic crust, such as that at the Troodos ophiolite, Cyprus (Patten et al., 2016, 2017). As such, Martin et al. (2019) reported high concentrations of Pb in chalcopyrite from the Troodos mafic-hosted VMS deposit. In the Lasail deposit of the Oman ophiolite, Stakes and Taylor (2003) demonstrated the existence of multiple superimposed magmatic and hydrothermal events, and that the large plagiogranite bodies and associated gabbroic rocks found with the contemporaneous late mafic dikes played an important role as a source of heat and metals (such as Pb) for massive sulfide formation.

In the bimodal-mafic subtype, chalcopyrite shows a slightly higher trimmed mean of Ag (94.8 ppm) than the other subtypes (Fig. 3) and a high positive regression coefficient from PLS-DA (Fig. 6C). The high concentration of Ag in chalcopyrite from bimodal-mafic settings is significantly biased by samples from Normétal (sample 51) and Kidd Creek (sample 2369), with maximum values of 1,640 and 1,031 ppm, respectively (App. 1, Table A3). The bias is also evidenced by the median of Ag in chalcopyrite from bimodal-mafic deposits (68 ppm Ag), which is significantly different from the mean value (95 ppm Ag; Table 2). Except for Au, Se, and Pb, chalcopyrite from bimodal-mafic and bimodal-felsic subtypes does not show significative differences from each other (Fig. 6). Bimodal deposits are characterized by a variable proportion of mafic and felsic volcanic rocks, representing a transition of host rock composition (Franklin et al., 2005) leading to similar trace element composition in chalcopyrite, as shown by important overlapping of chalcopyrite composition from these settings in our PLS-DA results (Fig. 6A).

Chalcopyrite from bimodal-felsic settings has higher concentrations of Au than the remaining subtypes. Chalcopyrite from the Quémont (0.83 ± 1.0 ppm) and Horne (0.29 ± 0.20 ppm) deposits is enriched in gold compared to the other deposits (<0.05 ppm), reaching a maximum of 5.6 and 1.1 ppm Au, respectively (Table 2; App. 1, Table A3). The Quémont and Horne deposits are considered two of the largest and richest gold-rich VMS deposits worldwide (Dubé et al., 2007; Mercier-Langevin et al., 2007, 2011), which is clearly recorded in the chalcopyrite composition. Chalcopyrite from the Snowcap bimodal-felsic–hosted SMS occurrence in the PACMANUS hydrothermal field also records a high concentration of Au (2.7 ± 2.6 ppm) with a maximum of 4.8 ppm. Studies have reported that chalcopyrite-rich chimneys in PACMANUS are enriched in Au (Binns et al., 2004, 2007; Hannington et al., 2005; Fuchs et al., 2019). In addition, chalcopyrite from bimodal-felsic compared to bimodal-mafic settings has a higher content of Pb, which may be explained by a greater proportion of felsic volcanic host rocks. Lead is principally derived from the destruction of feldspar, particularly abundant in felsic rocks from subduction-related volcanic arc systems floored by continental basement (Hannington et al., 1995; Hannington, 2014).

The trace element composition in chalcopyrite from siliciclastic-mafic settings has the highest values of Mn (Fig. 3), particularly those from Pike Hill mine in the Vermont copper belt (119 ppm; Table 2). The association between Mn and the siliciclastic-mafic subtype is supported by a high positive regression coefficient and the loading plots qw*1-qw*2 and qw*3-qw*4 (Fig. 6B, C; App. 2, Fig A5A). In metamorphosed hydrothermal systems, Mn-rich siliciclastic sedimentary rocks (i.e., coticule) that are usually intercalated with other meta-exhalites are typically found in rift-hosted volcanosedimentary base metal sulfide deposits, typical for the siliciclastic-mafic type (Spry et al., 1998). Slack et al. (2001) also reported the presence of Mn-rich garnet-quartz rocks in the Elizabeth and Pike Hill mines in the Vermont copper belt and suggest that they have formed as exhalative chemical precipitates on the sea floor. According to Hannington (2014), Mn is common at high concentrations in the most reduced fluids involved in the formation of VMS deposits. However, it rarely forms stable sulfides and is lost to distal hydrothermal facies, such as exhalites. The presence of coticules in Pike Hill mine suggests mobilization of Mn by hydrothermal fluids during VMS deposit formation, which was in part incorporated in chalcopyrite.

Finally, our results show that chalcopyrite from siliciclastic-felsic settings has higher concentrations of Sn, Sb, Bi, and In than the other subtypes (Figs. 3, 6). In particular, chalcopyrite from Neves-Corvo is enriched in Sn (1,058 ppm) and to a lesser extent In (Table 2), in agreement with the study of Carvalho et al. (2018). Likewise, chalcopyrite from the Halfmile Lake (5,643 ppm) and Hajar (856 ppm) deposits also shows high values of Sn (Table 2). Several studies have proposed a model for ore formation at Neves-Corvo involving direct magmatic contributions of Sn in addition to footwall metasediments (Relvas et al., 2001, 2006; Huston et al., 2011; Carvalho et al., 2018; Li et al., 2019; Marques et al., 2020). Similarly, Goodfellow (2007) reported Sn enrichment at Bathurst, also proposing a magmatic component for the metal source based on radiogenic Pb isotopes. The high concentration of Bi in chalcopyrite from siliciclastic-felsic settings may therefore also be explained by the presence of a magmatic component, since significant coenrichment of Bi and In have been used as evidence of direct contributions of metal from magmatic fluids (Hannington, 2014).

Predicting lithostratigraphic setting of VMS-SMS deposits with Random Forest (from affinity to VMS subtype)

The results of the predictions of literature data are summarized in Tables 7 and 8. All analyses were classified primarily by affinity with RF model.01; subsequent analyses correctly predicted were reclassified by subtype with RF model.02-A or RF model.02-B for those predicted as belonging to mafic-ultramafic or bimodal classes, respectively. Literature data with the complete set of elements analyzed by LA-ICP-MS required by the models correspond to George et al. (2018b), Wang et al. (2018), and Melekestseva et al. (2020a) (Table 7).

The highest rates of analyses correctly classified according to affinity (RF model.01) were recorded for the ultramafic-hosted SMS deposits Ashadze-2 hydrothermal field (Mid-Atlantic Ridge; Melekestseva et al., 2020a) and Kairei vent field (Central Indian Ridge; Wang et al., 2018), with 100%, followed by the mafic-hosted Vorta deposit, Romania (George

Table 7. Results of the Predictions of Literature Data with All Elements Required by the Random Forest Models

				Actual classes		RF model.01 (affinity) Sn-Co-Te-In-Zn-Se-Ag-Sb-Pb	Predicted classes			RF model.02-A (M-UM) Co-Mn-Sn-Te-Pb-In-Se-Ni-Tl-Ag-Zn-Bi	Predicted classes			RF model.02-B (BMF) Te-Zn-In-Se-Ag-Co-Ga-Sn-Sb-Ni	Predicted classes	
Author	Deposit	Metamorphism	n	Subtype	Affinity	Number correct[1]	M-UM	BMF	SF	Number correct	UM	M	SM	Number correct	BM	BF
Melekestseva et al. (2020a)	Ashadze-2 (SMS)	Unmetamorphosed	7	UM	M-UM	7	**100.0%**	0.0%	0.0%	7	**100.0%**	0.0%	0.0%			
Wang et al. (2018)	Kairei (SMS)	Unmetamorphosed	16	UM	M-UM	16	**100.0%**	0.0%	0.0%	15	**93.8%**	6.2%	0.0%			
George et al. (2018)	Vorta, Romania	Low grade	8	M	M-UM	7	**87.5%**	12.5%	0.0%	7	0.0%	**87.5%**	0.0%			
George et al. (2018)	Sulitjelma, Norway	Upper greenschist	111	BM[2]	M-UM	107	3.6%	**96.4%**	0.0%					67	**60.4%**	36.0%

Percentage in bold = percentage of analyses correctly classified on total of samples (n)
Abbreviations: BF = bimodal-felsic, BM = bimodal-mafic, BMF = bimodal-mafic-felsic, M = mafic, M-UM = mafic-ultramafic, SF = silicisclastic-felsic, SM = silicisclastic-mafic, UM = ultramafic
[1]Number of samples correctly classified
[2]Inconclusive classification in literature

et al., 2018b), with 87.5% (Table 7). Similarly, the Ashadze-2 and Kairei deposits reported the highest rates of analyses correctly classified according to VMS subtype (RF model.02-A), with 100% and 93.8%, respectively, followed by the Vorta deposits which registered an accuracy of 87.5%. The Sulitjelma deposit has an imprecise classification using literature data, and although several studies indicate that the host rocks show characteristics similar to the mafic subtype (Cook and Halls, 1990; Cook, 1996; Barrie et al., 2010a), Pedersen et al. (1991) proposed that Sulitjelma formed during the initial stages of back-arc spreading in the late Ordovician, more akin to the bimodal-mafic setting. Considering this uncertainty of the lithostratigraphic setting of the Sulitjelma deposit, the RF classification model yields results more consistent with a bimodal affinity (96.4%) and bimodal-mafic subtype (60.4%).

Predicting samples with missing elements: Published trace element data of chalcopyrite is scarce for VMS-SMS deposits. This is especially true for analyses of the elements required by our RF classification models. Machine learning algorithms provide some techniques to deal with missing predictors, such as bagged classification trees or k-Nearest Neighbor (kNN; (Provost and Saar-Tsechanski, 2007; Kuhn and Johnson, 2013). We tested bagged classification trees to impute missing predictors in studies where one or multiple elements were not analyzed. During our RF model training, bagged classification trees can be built for each predictor in order to use them for imputation of missing data (Provost and Saar-Tsechanski, 2007; Kuhn and Johnson, 2013).

Among studies with missing variables, those with a low number of nonanalyzed elements (<2) and of moderate to low variable importance in classification (e.g., Ga, In), recorded the highest rate of chalcopyrite correctly classified (Table 8). The highest proportion of correct prediction according to affinity (Table 8) were obtained for the bimodal-felsic María Teresa deposit, Peru (100%; Torró et al., 2022), the bimodal-felsic Çayeli-Kutlular deposit (82.1%; Revan et al., 2014) and the ultramafic-hosted Dergamysh deposit, Russia (77.3%; Melekestseva et al., 2020b). The remaining deposits, except for the bimodal-felsic SMS Roman Ruins, Papua New Guinea (74.1%; Wohlgemuth-Ueberwasser et al., 2015), which have at least half of the key elements missing, reported an overall accuracy lower than 65% with a minimum of 24.1% for the mafic-hosted Troodos deposits, Cyprus (Table 8; Martin et al., 2019). According to RF models by VMS settings, the highest percentage of analyses corrected classified were reported for bimodal-felsic María Teresa (89.6%), the bimodal-felsic Çayeli-Kutlular deposit (79.1%), and the ultramafic-hosted Dergamysh deposit (63.7%). Furthermore, the ultramafic-hosted SMS Logatchev deposit, the bimodal-felsic SMS Roman Ruins and Satanic Mills deposits, and the mafic-hosted Troodos deposits registered an overall accuracy lower than 60%. An increase of the number of nonanalyzed elements, particularly those with a moderate to high variable importance, degrades considerably the performance of the classifier, since less information is available to determine the provenance of the analyses. In addition to the number of missing elements, another important factor impacting the performance of the models is the percentage of censured values in the original data. This is the case of the data set from Martin et al. (2019), which has more than 30% of values below detection limit (in addi-

Table 8. Results of the Predictions of Literature Data with Missing Elements Using Random Forest (RF) Models

					Actual classes			RF model.01 (affinity) Sn-Co-Te-Ag-In-Sb-Se-Zn-Pb				RF model.02-A (M-UM) Co-Sn-Mn-Te-In-Pb-Se-Tl-Ni-Ag-Zn-Bi				RF model.02-B (BMF) Te-Se-Zn-In-Ag-Sb-Co-Ga-Sn-Ni		
									Predicted classes				Predicted classes				Predicted classes	
Author	Deposit	Metamorphism	Missing elements	n	Subtype	Affinity	Number correct[1]	M-UM	BMF	SF	Number correct	UM	M	SM	Number correct	BM	BF
Melekestseva et al. (2020b)	Dergamysh, Russia	Greenschist	Ga-In	22	UM	M-UM	17	**77.3%**	18.2%	4.5%	14	**63.7%**	13.6%	0.0%			
Wohlgemuth-Ueberwasser et al. (2015)	Logatchev (SMS)	Unmetamorphosed	Co-Ni-Zn-Ga-Ag-In-Tl-Pb-Bi	29	UM	M-UM	15	**51.7%**	48.3%	0.0%	15	**51.7%**	0.0%	0.0%			
Martin et al. (2019)	Troodos, Cyprus	Unmetamorphosed	Tl-Mn-Ga-In-Sn	41	M	M-UM	10	**24.1%**	59.3%	16.6%	1	20.3%	**1.9%**	1.9%			
Torró et al. (2022)	Cerro de Maimón, Dominican Republic	Greenschist/ blueschist	Te	11	M[2]	BMF	11	**100.0%**	0.0%	0.0%	11	0.0%	**100.0%**	0.0%			
Torró et al. (2022)	María Teresa, Peru	Contact metamorphism	Te	29	BF	BMF	29	0.0%	**100.0%**	0.0%					26	10.4%	**89.6%**
Revan et al. (2014)	Çayeli-Kutlular, Turkey	Unmetamorphosed	Ga-In	67	BF	BMF	55	17.9%	**82.1%**	0.0%					53	3.0%	**79.1%**
Wohlgemuth-Ueberwasser et al. (2015)	Roman Ruins (SMS)	Unmetamorphosed	Co-Ni-Zn-Ga-Ag-In-Tl-Pb-Bi	77	BF	BMF	55	0.0%	**71.4%**	28.6%					34	27.3%	**44.1%**
Wohlgemuth-Ueberwasser et al. (2015)	Satanic Mills (SMS)	Unmetamorphosed	Co-Ni-Zn-Ga-Ag-In-Tl-Pb-Bi	25	BF	BMF	16	0.0%	**64.0%**	36.0%					15	4.0%	**60.0%**

Percentage in bold = percentage of analyses correctly classified on total of samples (n)
Abbreviations: BF = bimodal-felsic, BM = bimodal-mafic, BMF = bimodal-mafic-felsic, M = mafic, M-UM = mafic-ultramafic, SF = siliciclastic-felsic, SM = siliciclastic-mafic, SMS = sea-floor massive sulfide, UM = ultramafic
[1] Number of samples correctly classified
[2] Inconclusive classification in literature

tion to the missing elements), decreasing the precision of the imputation, and consequently the performance of the prediction. On the other hand, although all samples from Cerro de Maimón deposit (Torró et al., 2022) were predicted as mafic subtype, its classification in the literature is inconclusive since it is considerably deformed. The RF classification model (RF model.02-B) yields results more consistent with a mafic subtype, with 100% of samples predicted as belonging to this setting.

Effects of metamorphism on classifier predictions

Several studies have documented the textural and chemical effects of deformation and remobilization on sulfides from VMS settings during metamorphism, principally on pyrite (Barrie et al., 2009, 2010b; George et al., 2018a; Conn et al., 2019; Zhao et al., 2021), sphalerite (George et al., 2016; Zhao et al., 2021), galena (George et al., 2015), and pyrrhotite (Steadman and Large, 2016). Solid mechanical transfer, liquid-state chemical transfer, and/or partial melting are the principal mechanisms involved in metamorphic remobilization of massive sulfide ores (Marshall et al., 1998; Tomkins, 2007; Tomkins et al., 2007; Andersson et al., 2016). Although little is known about the effect on chalcopyrite, some studies suggest that at high grades of metamorphism (upper amphibolite to granulite facies) Sn, In, Ga and Ag are preferentially incorporated into chalcopyrite, especially when it coexists with sphalerite, probably through partial melting, as well as a chemical remobilization mechanism combined with mechanical processes (George et al., 2016; Zhao et al., 2021).

In this study, samples come from a range of low metamorphic grades (Table 1). The majority of deposits underwent low-grade prehnite-pumpellyite (Iberian Pyrite Belt) and lower greenschist (Abitibi greenstone belt). In the Iberian Pyrite Belt deposits, studies have reported tectonic stacking of massive sulfide lenses, metal remobilization and local enrichment associated with Variscan deformation and metamorphism (Marignac et al., 2003; Tornos, 2006; Castroviejo et al., 2011;Almodóvar et al., 2019). Moreover, Almodóvar et al. (2019) suggested that postmineralization and pre-Variscan deformation sill emplacement in some Iberian Pyrite Belt deposits (e.g., Aguas Teñidas) conducted to a local increase of temperature leading to a Cu remobilization, probably by a liquid-state mechanism. Few studies have been conducted on remobilization mechanisms in sulfides at greenschist facies. For the McLeod deposit (Matagami, Québec), Genna and Gaboury (2019) suggested that semivolatile elements, principally hosted in pyrite (e.g., Sb, Tl, W, Sn), could be remobilized from the massive sulfide lens to the host rhyolite during prograde greenschist metamorphism by the transformation of pyrite to pyrrhotite. Concerning the deposits from the Bothnia-Skelleftea Unit, Skyttä et al. (2020) indicated that they have been affected by polyphase ductile deformation and metamorphism at variable temperature and low-pressure conditions, the later event under lower greenschist conditions. In these deposits, both solid-state remobilization, which includes fold hinge sulfide thickening and limb attenuation, and liquid-state remobilization has led to gold enrichment (Weihed et al., 2002). Samples from Ming, Bathurst, and the Vermont copper belt have undergone upper greenschist metamorphism (Table 1; App. 1 Table A1). At Ming, minor metal remobiliza-

tion by metamorphism and deformation has been recorded (Brueckner et al., 2015, 2016; Pilote et al., 2020). On the other hand, the Bathurst and Halfmile Lake deposits show textural characteristics typically related to remobilization of sulfides (Dehnavi et al., 2018), by solution transfer, principally in pyrite but also probably in pyrrhotite, sphalerite, and chalcopyrite (Roo and Staal, 2003). Samples from the Tétrault deposit of Montauban (Grenville Province, Québec) reached the lower amphibolite facies. In this deposit, Tomkins (2007) demonstrated that mechanical remobilization of sulfides was the most important mechanism for controlling the distribution of Zn-Pb mineralization, whereas a combination of partial melting and prograde hydrothermal mechanisms were responsible for Ag-Au remobilization.

Our PLS-DA results (Fig. 7) demonstrate that any modification of the trace element composition of chalcopyrite due to variable metamorphic grade is in fact smaller than the much larger variation in composition due to the different lithotectonic subtypes of VMS deposits. This is particularly well shown for the mafic subtype, where chalcopyrite from unmetamorphosed and lower greenschist grade deposits plot within the limits of the 95% confidence ellipse, indicating no significant differences in trace element composition (Fig. 7B). For the siliciclastic-felsic subtype, although chalcopyrite from deposits at greenschist and amphibolite facies form a group separated from those at prehnite-pumpellyite and to a lesser extent middle greenschist metamorphism, the signature of chalcopyrite from this subtype as a whole is still distinctive from all other VMS subtypes (Fig. 7F). As such, the RF model remains robust even when metamorphic grade varies, i.e., the prediction is not significantly affected. The literature data used to validate the RF models also includes samples from deposits that have undergone metamorphism at different conditions (Tables 7, 8). Yet, the RF models reported a high accuracy for prediction (>60%) in samples from the ultramafic-hosted Dergamysh deposit (Melekestseva et al., 2020b) affected by greenschist metamorphism (Table 8) in spite of the fact that the ultramafic class was trained exclusively with chalcopyrite from unmetamorphosed SMS deposits. Similarly, literature data from unmetamorphosed deposits, such as the bimodal-felsic Çayeli-Kutlular deposit (Revan et al., 2014), reported more than 79% of samples correctly classified (Table 8), although the training set was mostly from deposits at the lower greenschist facies.

Conclusions

This study shows that the trace element composition of chalcopyrite is principally controlled by the lithostratigraphic setting, which is affected by the host rock composition, but with a second-order variation due to the temperature of the sulfide assemblage (Cu-rich vs. Zn-rich) in which chalcopyrite is replacing/precipitating. In addition, results demonstrate that metamorphic grade does not significantly impact PLS-DA results nor the ability of RF models to identify VMS subtype. Chalcopyrite from ultramafic-hosted SMS settings is characterized by high content in Co, Ni, and Te most likely as a result of leaching of these elements by the fluids during the alteration of the ferromagnesian minerals and magmatic sulfides of the footwall ultramafic rocks. Chalcopyrite from mafic-hosted VMS, although it does not show significant

differences in most trace elements from the other subtypes, has higher concentrations of Pb, which could be derived from alteration of plagiogranites in the oceanic crust and/or result from the lack of galena coprecipitating in mafic-hosted deposits. Chalcopyrite from bimodal settings of VMS shows similar content for most elements since they are constituted by a variable proportion of mafic and felsic volcanic rocks, representing a transition of host rock composition. However, Au and Pb are enriched in chalcopyrite from the bimodal-felsic subtype, with Pb likely derived from the destruction of feldspar in felsic volcanic rocks. In siliciclastic-mafic deposits, chalcopyrite is particularly enriched in Mn. The presence of Mn-rich meta-exhalites in this deposit type suggests that Mn-rich hydrothermal fluids were involved in forming siliciclastic-mafic deposits, as recorded by chalcopyrite. Finally, in siliciclastic-felsic settings, chalcopyrite is characterized by high contents of Sn, Sb, Bi, and In, associated with a magmatic component in addition to footwall metasediments. On the other hand, chalcopyrite from Cu-rich samples is associated with high contents of high-temperature elements (Bi, Co, Ni, Au, Ag, Te, and Se), whereas that from Zn-rich samples is associated with high Ga, Sn, In, Sb, and Tl, a typical low-temperature assemblage, suggesting that during the zone refining process, a successive remobilization of these elements from sulfides previously formed at lower temperature (e.g., sphalerite, galena, or pyrite) led to their incorporation into chalcopyrite at different replacement degrees.

We developed a set of three RF models to classify chalcopyrite from VMS deposits according to the six lithostratigraphic subtypes proposed in the literature. Our approach, consisting of one RF model to first classify chalcopyrite according to major host rock affinity and two subsequent RF models to reclassify chalcopyrite according to VMS subtype, recorded an overall accuracy higher than 0.95. Further testing using data from the literature illustrates that the performance of the prediction of our classification model is determined by the deposit setting and the number and importance of variables used (quality of the data). Chalcopyrite from the ultramafic subtype shows the highest classification rates because this subtype is significantly different in trace element composition from all others. The prediction results show the highest performance in predictions with literature data containing all elements required by the RF models. In studies with missing elements, the classification accuracy decreases significantly, especially when a high number or the most important variables are missing. Finally, RF regression models indicate that there is a potential to predict ccp/(ccp + sp) ratio from trace element composition, and we suggest further studies involving a wider range of samples.

Acknowledgments

REFERENCES

Abd Elrahman, S.M., and Abraham, A., 2013, A review of class imbalance problem: Journal of Network and Innovative Computing, v. 1, p. 332–340.

Agangi, A., Reddy, S., Plavsa, D., Vieru, C., Selvaraja, V., LaFlamme, C., Jeon, H., Martin, L., Nozaki, T., and Takaya, Y., 2018, Subsurface deposition of Cu-rich massive sulphide underneath a Palaeoproterozoic seafloor hydrothermal system—the Red Bore prospect, Western Australia: Mineralium Deposita, v. 53, p. 1–18.

Aitchison, J., 1986, The statistical analysis of compositional data: London, Chapman and Hall, 416 p.

Almodóvar, G.R., Yesares, L., Sáez, R., Toscano, M., González, F., and Pons, J.M., 2019, Massive sulfide ores in the Iberian Pyrite Belt: Mineralogical and textural evolution: Minerals, v. 9, p. 653.

Andersson, S.S., Jonsson, E., and Högdahl, K., 2016, Metamorphism and deformation of a Palaeoproterozoic polymetallic sulphide-oxide mineralisation: Hornkullen, Bergslagen, Sweden: GFF, v. 138, p. 410–423.

Antweiler, R.C., 2015, Evaluation of statistical treatments of left-censored environmental data using coincident uncensored data sets. II. Group comparisons: Environmental Science & Technology, v. 49, p. 13,439–13,446.

Averill, S.A., 2001, The application of heavy indicator mineralogy in mineral exploration with emphasis on base metal indicators in glaciated metamorphic and plutonic terrains: Geological Society, London, Special Publications, v. 185, p. 69–81.

——2011, Viable indicator minerals in surficial sediments for two major base metal deposit types: Ni-Cu-PGE and porphyry Cu: Geochemistry: Exploration, Environment, Analysis, v. 11, p. 279–291.

Bajwah, Z.U., Seccombe, P.K., and Offler, R., 1987, Trace element distribution, Co: Ni ratios and genesis of the Big Cadia iron-copper deposit, New South Wales, Australia: Mineralium Deposita, v. 22, p. 292–300.

Ballabio, D., and Consonni, V., 2013, Classification tools in chemistry. Part 1: Linear models. PLS-DA: Analytical Methods, v. 5, p. 3790–3798.

Barker, R.D., Barker, S.L.L., Cracknell, M.J., Stock, E.D., and Holmes, G., 2021, Quantitative mineral mapping of drill core surfaces II: Long-wave infrared mineral characterization using μXRF and machine learning: Economic Geology, v. 116, p. 821–836.

Barnes, S.-J., Prichard, H.M., Cox, R.A., Fisher, P.C., and Godel, B., 2008, The location of the chalcophile and siderophile elements in platinum-group element ore deposits (a textural, microbeam and whole rock geochemical study): Implications for the formation of the deposits: Chemical Geology, v. 248, p. 295–317.

Barrie, C.D., Boyle, A.P., and Salter, M., 2009, How low can you go? - Extending downwards the limits of plastic deformation in pyrite: Mineralogical Magazine, v. 73, p. 895–913.

Barrie, C.D., Boyle, A.P., Cook, N.J., and Prior, D.J., 2010a, Pyrite deformation textures in the massive sulfide ore deposits of the Norwegian Caledonides: Tectonophysics, v. 483, p. 269–286.

——2010b, Pyrite deformation textures in the massive sulfide ore deposits of the Norwegian Caledonides: Tectonophysics, v. 483, p. 269–286.

Barrie, C.T., and Hannington, M.D., 1999, Classification of volcanic-associated massive sulfide deposits based on host-rock composition: Reviews in Economic Geology, v. 8, p. 1–11.

Barton, P.B., and Bethke, P.M., 1987, Chalcopyrite disease in sphalerite; pathology and epidemiology: American Mineralogist, v. 72, p. 451–467.

Bédard, É., Goulet, A., Sappin, A.-A., Makvandi, S., and Beaudoin, G., 2017, Chalcopyrite as an indicator mineral to fingerprint mineral deposit types: A preliminary study: SGA, Biennial Meeting, 14th, Quebec City, Canada, 2017, Proceedings, p. 1115–1118.

Bédard, É., De Bronac de Vazelhes, V., and Beaudoin, G., 2022, Performance of predictive supervised classification models of trace elements in

magnetite for mineral exploration: Journal of Geochemical Exploration, v. 236, article 106959.

Binns, R.A., McConachy, T., and McInnes, B., 2004, Eastern Manus basin, Papua New Guinea: Guides for volcanogenic massive sulphide exploration from a modern seafloor analogue: CSIRO Explores, v. 2, p. 59–80.

Binns, R.A., Barriga, F., and Miller, D.J., 2007, Leg 193 synthesis; anatomy of an active felsic hosted hydrothermal system, eastern Manus basin, Papua New Guinea: International Ocean Drilling Program, Proceedings of the Ocean Drilling Program, Scientific Results (CD ROM), v. 193, p. 71.

Bortnikov, N.S., Genkin, A.D., Dobrovol'Skaya, M.G., Muravitskaya, G.N., and Filimonova, A.A., 1991, The nature of chalcopyrite inclusions in sphalerite; exsolution, coprecipitation, or" disease"?: Economic Geology, v. 86, p. 1070–1082.

Breiman, L., 2001, Random Forests: Machine Learning, v. 45, p. 5–32.

Brereton, R.G., and Lloyd, G.R., 2014, Partial least squares discriminant analysis: Taking the magic away: Journal of Chemometrics, v. 28, p. 213–225.

Brownlee, J., 2018, Master machine learning algorithms: Discover how they work and implement them from scratch: Melbourne, Machine Learning Mastery, https://machinelearningmastery.com/master-machine-learning-algorithms/.

Brueckner, S.M., Piercey, S.J., Layne, G.D., Piercey, G., and Sylvester, P. J., 2015, Variations of sulphur isotope signatures in sulphides from the metamorphosed Ming Cu(-Au) volcanogenic massive sulphide deposit, Newfoundland Appalachians, Canada: Mineralium Deposita, v. 50, p. 619–640.

Brueckner, S.M., Piercey, S.J., Pilote, J.-L., Layne, G.D., and Sylvester, P.J., 2016, Mineralogy and mineral chemistry of the metamorphosed and precious metal-bearing Ming deposit, Canada: Ore Geology Reviews, v. 72, p. 914–939.

Butler, I.B., and Nesbitt, R.W., 1999, Trace element distributions in the chalcopyrite wall of a black smoker chimney: Insights from laser ablation inductively coupled plasma mass spectrometry (LA–ICP–MS): Earth and Planetary Science Letters, v. 167, p. 335–345.

Cabri, L.J., Campbell, J.L., Laflamme, J.H.G., Leigh, R.G., Maxwell, J.A., and Scott, J.D., 1985, Proton-microprobe analysis of trace elements in sulfides from some massive-sulfide deposits: The Canadian Mineralogist, v. 23, p. 133–148.

Caraballo, E., Dare, S., and Beaudoin, G., 2022, Variation of trace elements in chalcopyrite from worldwide Ni-Cu sulfide and Reef-type PGE deposits: Implications for mineral exploration: Mineralium Deposita, v. 57, p. 1293–1321.

Carvalho, J.R.S., Relvas, J., Pinto, A.M.M., Frenzel, M., Krause, J., Gutzmer, J., Pacheco, N., Fonseca, R., Santos, S., and Caetano, P., 2018, Indium and selenium distribution in the Neves-Corvo deposit, Iberian Pyrite Belt, Portugal: Mineralogical Magazine, v. 82, p. S5–S41.

Castroviejo, R., Quesada, C., and Soler, M., 2011, Post-depositional tectonic modification of VMS deposits in Iberia and its economic significance: Mineralium Deposita, v. 46, p. 615–637.

Chong, I.-G., and Jun, C.-H., 2005, Performance of some variable selection methods when multicollinearity is present: Chemometrics and Intelligent Laboratory Systems, v. 78, p. 103–112.

Chorlton, L.B., 2007, Generalized geology of the world: bedrock domains and major faults in GIS format: Geological Survey of Canada, Open File, v. 5529.

Conn, C.D., Spry, P.G., Layton-Matthews, D., Voinot, A., and Koenig, A., 2019, The effects of amphibolite facies metamorphism on the trace element composition of pyrite and pyrrhotite in the Cambrian Nairne Pyrite Member, Kanmantoo Group, South Australia: Ore Geology Reviews, v. 114, article 103128.

Cook, N.J., 1996, Mineralogy of the sulphide deposits at Sulitjelma, northern Norway: Ore Geology Reviews, v. 11, p. 303–338.

Cook, N.J., and Chryssoulis, S.L., 1990, Concentrations of invisible gold in the common sulfides: The Canadian Mineralogist, v. 28, p. 1–16.

Cook, N.J., and Halls, C., 1990, Coticules at Sulitjalma, Norway and their possible origin: Norsk Geologisk Tidsskrift, v. 70, p. 153–158.

Cook, N.J., Ciobanu, C.L., Danyushevsky, L.V., and Gilbert, S., 2011, Minor and trace elements in bornite and associated Cu–(Fe)-sulfides: A LA-ICP-MS studyBornite mineral chemistry: Geochimica et Cosmochimica Acta, v. 75, p. 6473–6496.

Crundwell, F.K., 2021, The impact of light on understanding the mechanism of dissolution and leaching of sphalerite (ZnS), pyrite (FeS$_2$) and chalcopyrite (CuFeS$_2$): Minerals Engineering, v. 161, article 106728.

Crundwell, F.K., Bryson, L.J., van Aswegen, A., and Knights, B.D.H., 2021, Effect of chopped light on the dissolution and leaching of chalcopyrite: Minerals Engineering, v. 160, article 106703.

Czamanske, G.K., Kunilov, V.E., Zientek, M.L., Cabri, L.J., Likhachev, A.P., Calk, L.C., and Oscarson, R.L., 1992, A proton microprobe study of magmatic sulfide ores from the Noril'sk-Talnakh District, Siberia: The Canadian Mineralogist, v. 30, p. 249–287.

Dare, S.A.S., Barnes, S.-J., and Prichard, H.M., 2010, The distribution of platinum group elements (PGE) and other chalcophile elements among sulfides from the Creighton Ni-Cu-PGE sulfide deposit, Sudbury, Canada, and the origin of palladium in pentlandite: Mineralium Deposita, v. 45, p. 765–793.

Dare, S.A.S., Barnes, S.-J., Prichard, H.M., and Fisher, P.C., 2011, Chalcophile and platinum-group element (PGE) concentrations in the sulfide minerals from the McCreedy East deposit, Sudbury, Canada, and the origin of PGE in pyrite: Mineralium Deposita, v. 46, p. 381–407.

Dare, S.A., Barnes, S.-J., Prichard, H.M., and Fisher, P.C., 2014, Mineralogy and geochemistry of Cu-rich ores from the McCreedy East Ni-Cu-PGE deposit (Sudbury, Canada): Implications for the behavior of platinum group and chalcophile elements at the end of crystallization of a sulfide liquid: Economic Geology, v. 109, p. 343–366.

Dehnavi, A.S., McFarlane, C.R.M., Lentz, D.R., and Walker, J.A., 2018, Assessment of pyrite composition by LA-ICP-MS techniques from massive sulfide deposits of the Bathurst Mining Camp, Canada: From textural and chemical evolution to its application as a vectoring tool for the exploration of VMS deposits: Ore Geology Reviews, v. 92, p. 656–671.

Digis Team, 2022, GEOROC Compilation: Rock types: Geoscience Centre Göttingen, https://georoc.eu/georoc/new-start.asp.

Dubé, B., Gosselin, P., Mercier-Langevin, P., Hannington, M., and Galley, A., 2007, Gold-rich volcanogenic massive sulphide deposits: Geological Association of Canada, Mineral Deposits Division, Special Publication no. 5, p. 75–94.

Dunn, K., 2019, Process improvement using data: Learnche.org, 421 p., https://learnche.org/pid/contents.

Duran, C.J., Barnes, S.-J., and Corkery, J.T., 2016, Trace element distribution in primary sulfides and Fe-Ti oxides from the sulfide-rich pods of the Lac des Iles Pd deposits, Western Ontario, Canada: Constraints on processes controlling the composition of the ore and the use of pentlandite compositions in exploration: Journal of Geochemical Exploration, v. 166, p. 45–63.

Duran, C.J., Dubé-Loubert, H., Pagé, P., Barnes, S.-J., Roy, M., Savard, D., Cave, B.J., Arguin, J.-P., and Mansur, E.T., 2019, Applications of trace element chemistry of pyrite and chalcopyrite in glacial sediments to mineral targeting: Example from the Churchill Province, northern Quebec, Canada: Journal of Geochemical Exploration, v. 196, p. 105–130.

Duran, C.J., Barnes, S.-J., Mansur, E.T., Dare, S.A.S., Bédard, L.P., and Sluzhenikin, S.F., 2020, Magnetite chemistry by LA-ICP-MS records sulfide fractional crystallization in massive nickel-copper-platinum group element ores from the Norilsk-Talnakh mining district (Siberia, Russia): Implications for trace element partitioning into magnetite: Economic Geology, v. 115, p. 1245–1266.

Egozcue, J.J., and Pawlowsky-Glahn, V., 2011, Análisis composicional de datos en Ciencias Geoambientales: Boletín Geológico y Minero, v. 122, p. 439–452.

Eldridge, C. S., Barton, P. B., Jr., and Ohmoto, H., 1983, Mineral textures and their bearing on formation of the Kuroko orebodies: Society of Economic Geologists, Economic Geology Monograph 5, p. 241–281.

Favilla, S., Durante, C., Vigni, M.L., and Cocchi, M., 2013, Assessing feature relevance in NPLS models by VIP: Chemometrics and Intelligent Laboratory Systems, v. 129, p. 76–86.

Filzmoser, P., Hron, K., and Reimann, C., 2009, Principal component analysis for compositional data with outliers: Environmetrics, v. 20, p. 621–632.

Franklin, J., Gibson, H., and Jonasson, G., 2005, Volcanogenic massive sulfide deposit: Economic Geology, 100th Anniversary Volume, p. 523–560.

Frenzel, M., Hirsch, T., and Gutzmer, J., 2016, Gallium, germanium, indium, and other trace and minor elements in sphalerite as a function of deposit type—A meta-analysis: Ore Geology Reviews, v. 76, p. 52–78.

Fuchs, S., Hannington, M.D., and Petersen, S., 2019, Divining gold in seafloor polymetallic massive sulfide systems: Mineralium Deposita, v. 54, p. 789–820.

Galley, A.G., Hannington, M.D., and Jonasson, I., 2007, Volcanogenic massive sulphide deposits: Mineral deposits of Canada: A synthesis of major deposit-types, district metallogeny, the evolution of geological provinces, and exploration methods: Geological Association of Canada, Mineral Deposits Division, Special Publication, v. 5, p. 141–161.

Genna, D., and Gaboury, D., 2015, 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: Economic Geology, v. 110, p. 2087–2108.

——2019, Use of semi-volatile metals as a new vectoring tool for VMS exploration: Example from the Zn-rich McLeod deposit, Abitibi, Canada: Journal of Geochemical Exploration, v. 207, article 106358.

George, L., Cook, N.J., Ciobanu, C.L., and Wade, B.P., 2015, Trace and minor elements in galena: A reconnaissance LA-ICP-MS study: American Mineralogist, v. 100, p. 548–569.

George, L.L., Cook, N.J., and Ciobanu, C.L., 2016, Partitioning of trace elements in co-crystallized sphalerite-galena-chalcopyrite hydrothermal ores: Ore Geology Reviews, v. 77, p. 97–116.

George, L.L., Biagioni, C., D'Orazio, M., and Cook, N.J., 2018a, 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 Geology Reviews, v. 102, p. 59–105.

George, L.L., Cook, N.J., Crowe, B.B., and Ciobanu, C.L., 2018b, Trace elements in hydrothermal chalcopyrite: Mineralogical Magazine, v. 82, p. 59–88.

Gislason, P.O., Benediktsson, J.A., and Sveinsson, J.R., 2006, Random forests for land cover classification: Pattern recognition letters, v. 27, p. 294–300.

Godel, B., and Barnes, S.-J., 2008, Platinum-group elements in sulfide minerals and the whole rocks of the JM Reef (Stillwater Complex): Implication for the formation of the reef: Chemical Geology, v. 248, p. 272–294.

Goodfellow, W.D., 2007, Metallogeny of the Bathurst mining camp, northern New Brunswick: Geological Association of Canada, Mineral Deposits Division, Special Publication, v. 5, p. 449–469.

Gosselin, R., Rodrigue, D., and Duchesne, C., 2010, A Bootstrap-VIP approach for selecting wavelength intervals in spectral imaging applications: Chemometrics and Intelligent Laboratory Systems, v. 100, p. 12–21.

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, v. 114, p. 771–786.

Grömping, U., 2009, Variable importance assessment in regression: Linear regression versus random forest: The American Statistician, v. 63, p. 308–319.

Grzela, D., Beaudoin, G., and Bédard, É., 2019, Tourmaline, scheelite, and magnetite compositions from orogenic gold deposits and glacial sediments of the Val-d'Or district (Quebec, Canada): Applications to mineral exploration: Journal of Geochemical Exploration, v. 206, article 106355.

Han, H., Guo, X., and Yu, H., 2016, Variable selection using mean decrease accuracy and mean decrease gini based on random forest: Institute of Electrical and Electronics Engineers (IEEE), International Conference on Software Engineering And Service Science (ICSESS), 7th, Beijing, China, 2016, Proceedings, p. 219–224.

Hannington, M.D., 2014, 13.18 - Volcanogenic massive sulfide deposits, *in* Holland, H.D., and Turekian, K.K., eds., Treatise on geochemistry, 2nd ed.: Oxford, Elsevier, p. 463–488.

Hannington, M.D., Jonasson, I.R., Herzig, P.M., and Petersen, S., 1995, Physical and chemical processes of seafloor mineralization at mid-ocean ridges: Seafloor hydrothermal systems: Physical, Chemical, Biological, and Geological Interactions, v. 91, p. 115–157.

Hannington, M.D., de Ronde, C.D., and Petersen, S., 2005, Sea-floor tectonics and submarine hydrothermal systems: Economic Geology, 100th Anniversary Volume, p. 111–141.

Harris, D.C., Cabri, L.J., and Nobiling, R., 1984, Silver-bearing chalcopyrite, a principal source of silver in the Izok Lake massive-sulfide deposit; confirmation by electron and proton-microprobe analyses: The Canadian Mineralogist, v. 22, p. 493–498.

Harris, J.R., and Grunsky, E.C., 2015, Predictive lithological mapping of Canada's North using Random Forest classification applied to geophysical and geochemical data: Computers & Geosciences, v. 80, p. 9–25.

Hashmi, S., Ward, B., Plouffe, A., Leybourne, M., and Ferbey, T., 2015, Geochemical and mineralogical dispersal in till from the Mount Polley Cu-Au porphyry deposit, central British Columbia, Canada: Geochemistry: Exploration, Environment, Analysis, v. 15, p. 234–249.

Hattori, K.H., Arai, S., and Clarke, D.B., 2002, Selenium, tellurium, arsenic and antimony contents of primary mantle sulfides: The Canadian Mineralogist, v. 40, p. 637–650.

Hawley, J.E., and Nichol, I., 1961, Trace elements in pyrite, pyrrhotite and chalcopyrite of different ores: Economic Geology, v. 56, p. 467–487.

Hazarika, P., Mishra, B., and Pruseth, K.L., 2015, Diverse tourmaline compositions from orogenic gold deposits in the Hutti-Maski greenstone belt, India: Implications for sources of ore-forming fluids: Economic Geology, v. 110, p. 337–353.

Hofmann, A. W., 1988, Chemical differentiation of the Earth: The relationship between mantle, continental crust, and oceanic crust: Earth and Planetary Science Letters, v. 90, p. 297–314.

Hron, K., Templ, M., and Filzmoser, P., 2010, Imputation of missing values for compositional data using classical and robust methods: Computational Statistics & Data Analysis, v. 54, p. 3095–3107.

Huston, D.L., Sie, S.H., Suter, G.F., Cooke, D.R., and Both, R.A., 1995, Trace elements in sulfide minerals from eastern Australian volcanic-hosted massive sulfide deposits; Part I, Proton microprobe analyses of pyrite, chalcopyrite, and sphalerite, and Part II, Selenium levels in pyrite; comparison with delta 34S values and implications for the source of sulfur in volcanogenic hydrothermal systems: Economic Geology, v. 90, p. 1167–1196.

Huston, D.L., Relvas, J.M.R.S., Gemmell, J.B., and Drieberg, S., 2011, The role of granites in volcanic-hosted massive sulphide ore-forming systems: An assessment of magmatic-hydrothermal contributions: Mineralium Deposita, v. 46, p. 473–507.

Ishihara, S., and Endo, Y., 2007, Indium and other trace elements in volcanogenic massive sulfide ores from the Kuroko, Besshi and other types in Japan: Bulletin of the Geological Survey of Japan, v. 58, p. 7–22.

Jochum, K.P., Nohl, U., Herwig, K., Lammel, E., Stoll, B., and Hofmann, A.W., 2005, GeoReM: A new geochemical database for reference materials and isotopic standards: Geostandards and Geoanalytical Research, v. 29, p. 333–338.

Kase, K., 1987, Tin-bearing chalcopyrite from the Izumo vein, Toyoha mine, Hokkaido, Japan: The Canadian Mineralogist, v. 25, p. 9–13.

Kuhn, M., 2008, Building predictive models in R using the caret package: Journal of Statistical Software, v. 28, p. 1–26.

Kuhn, M., and Johnson, K., 2013, Applied predictive modeling: New York, Springer, 600 p.

Li, X., Zhao, K.-D., Jiang, S.-Y., and Palmer, M.R., 2019, In-situ U-Pb geochronology and sulfur isotopes constrain the metallogenesis of the giant Neves Corvo deposit, Iberian Pyrite Belt: Ore Geology Reviews, v. 105, p. 223–235.

Liu, H., and Beaudoin, G., 2021, Geochemical signatures in native gold derived from Au-bearing ore deposits: Ore Geology Reviews, v. 132, article 104066.

Liu, H., Beaudoin, G., Makvandi, S., Jackson, S.E., and Huang, X., 2021, Multivariate statistical analysis of trace element compositions of native gold from orogenic gold deposits: Implication for mineral exploration: Ore Geology Reviews, v. 131, article 104061.

Lockington, J.A., Cook, N.J., and Ciobanu, C.L., 2014, Trace and minor elements in sphalerite from metamorphosed sulphide deposits: Mineralogy and Petrology, v. 108, p. 873–890.

Lydon, J., 1988, Ore deposit models# 14. Volcanogenic massive sulphide deposits Part 2: Genetic models: Geoscience Canada, v. 15, p. 43–65.

Makvandi, S., Ghasemzadeh-Barvarz, M., Beaudoin, G., Grunsky, E.C., McClenaghan, M.B., and Duchesne, C., 2016a, Principal component analysis of magnetite composition from volcanogenic massive sulfide deposits: Case studies from the Izok Lake (Nunavut, Canada) and Half-mile Lake (New Brunswick, Canada) deposits: Ore Geology Reviews, v. 72, p. 60–85.

Makvandi, S., Ghasemzadeh-Barvarz, M., Beaudoin, G., Grunsky, E.C., McClenaghan, M.B., Duchesne, C., and Boutroy, E., 2016b, Partial least squares-discriminant analysis of trace element compositions of magnetite from various VMS deposit subtypes: Application to mineral exploration: Ore Geology Reviews, v. 78, p. 388–408.

Manéglia, N., Beaudoin, G., and Simard, M., 2018, Indicator minerals of the Meliadine orogenic gold deposits, Nunavut (Canada), and application to till surveys: Geochemistry: Exploration, Environment, Analysis, v. 18, p. 241–251.

Mansur, E.T., Barnes, S.-J., Duran, C.J., and Sluzhenikin, S.F., 2020, Distribution of chalcophile and platinum-group elements among pyrrhotite, pentlandite, chalcopyrite and cubanite from the Noril'sk-Talnakh ores: Implications for the formation of platinum-group minerals: Mineralium Deposita, v. 55, p. 1215–1232.

Marfin, A.E., Ivanov, A.V., Abramova, V.D., Anziferova, T.N., Radomskaya, T.A., Yakich, T.Y., and Bestemianova, K.V., 2020, A trace element

classification tree for chalcopyrite from Oktyabrsk deposit, Norilsk-Talnakh Ore District, Russia: LA-ICPMS study: Minerals, v. 10, p. 716.

Marignac, C., Diagana, B., Cathelineau, M., Boiron, M.-C., Banks, D., Fourcade, S., and Vallance, J., 2003, Remobilisation of base metals and gold by Variscan metamorphic fluids in the south Iberian pyrite belt: Evidence from the Tharsis VMS deposit: Chemical Geology, v. 194, p. 143–165.

Marques, A.F.A., Relvas, J.M.R.S., Scott, S.D., Rosa, C., and Guillong, M., 2020, Melt inclusions in quartz from felsic volcanic rocks of the Iberian Pyrite Belt: Clues for magmatic ore metal transfer towards VMS-forming systems: Ore Geology Reviews, v. 126, article 103743.

Marshall, B., Vokes, F.M., and Larocque, A.C.L., 1998, Regional metamorphic remobilization: Upgrading and formation of ore deposits: Reviews in Economic Geology, v. 11, p. 19–38.

Martin, A.J., Keith, M., McDonald, I., Haase, K.M., McFall, K.A., Klemd, R., and MacLeod, C.J., 2019, Trace element systematics and ore-forming processes in mafic VMS deposits: Evidence from the Troodos ophiolite, Cyprus: Ore Geology Reviews, v. 106, p. 205–225.

Maslennikov, V.V., Maslennikova, S.P., Large, R.R., and Danyushevsky, L.V., 2009, Study of trace element zonation in vent chimneys from the Silurian Yaman-Kasy volcanic-hosted massive sulfide deposit (Southern Urals, Russia) using laser ablation-inductively coupled plasma mass spectrometry (LA-ICPMS): Economic Geology, v. 104, p. 1111–1141.

Maslennikov, V.V., Maslennikova, S.P., Large, R.R., Danyushevsky, L.V., Herrington, R.J., Ayupova, N.R., Zaykov, V.V., Lein, A.Y., Tseluyko, A.S., Melekestseva, I.Y., and Tessalina, S.G., 2017, Chimneys in Paleozoic massive sulfide mounds of the Urals VMS deposits: Mineral and trace element comparison with modern black, grey, white and clear smokers: Ore Geology Reviews, v. 85, p. 64–106.

McClenaghan, M.B., 2005, Indicator mineral methods in mineral exploration: Geochemistry: Exploration, Environment, Analysis, v. 5, p. 233–245.

Megahed, F. M., Chen, Y.-J., Megahed, A., Ong, Y., Altman, N., and Krzywinski, M., 2021, The class imbalance problem: Nature Methods, v. 18, p. 1270–1272.

Mehmood, T., Liland, K.H., Snipen, L., and Sæbø, S., 2012, A review of variable selection methods in partial least squares regression: Chemometrics and Intelligent Laboratory Systems, v. 118, p. 62–69.

Melekestseva, I., Maslennikov, V., Tret'yakov, G., Maslennikova, S., Danyushevsky, L., Kotlyarov, V., Large, R., Beltenev, V., and Khvorov, P., 2020a, Trace element geochemistry of sulfides from the Ashadze-2 hydrothermal field (12°58′ N, Mid-Atlantic Ridge): Influence of host rocks, formation conditions or seawater?: Minerals, v. 10, p. 743, doi: 10.3390/min10090743.

Melekestseva, I.Y., Maslennikov, V.V., and Maslennikova, S.P., 2020b, Trace-elements in sulfides of the Dergamysh cobalt-bearing massive sulfide deposit, the Southern Urals: Mode of occurrence and matter sources: Lithosphere (Russia), v. 20, p. 499–516.

Mendez, K.M., Broadhurst, D.I., and Reinke, S.N., 2020, Migrating from partial least squares discriminant analysis to artificial neural networks: a comparison of functionally equivalent visualisation and feature contribution tools using jupyter notebooks: Metabolomics, v. 16, p. 17.

Mercier-Langevin, P., Dubé, B., Hannington, M.D., Davis, D.W., Lafrance, B., and Gosselin, G., 2007, The LaRonde Penna Au-rich volcanogenic massive sulfide deposit, Abitibi greenstone belt, Quebec: Part I. Geology and Geochronology°: Economic Geology, v. 102, p. 585–609.

Mercier-Langevin, P., Hannington, M.D., Dubé, B., and Bécu, V., 2011, The gold content of volcanogenic massive sulfide deposits: Mineralium Deposita, v. 46, p. 509–539.

Michael, F., Georges, M., and John, F., 2013, Elliptical Insights: Understanding statistical methods through elliptical geometry: Statistical Science, v. 28, p. 1–39.

Miranda, A.C.R., Beaudoin, G., and Rottier, B., 2022, Scheelite chemistry from skarn systems: Implications for ore-forming processes and mineral exploration: Mineralium Deposita, v. 57, p. 1469–1497.

Moggi-Cecchi, V., Cipriani, C., Rossi, P., Ceccato, D., Rudello, V., and Somacal, H., 2002, Trace element contents and distribution maps of chalcopyrite: A micro-PIXE study: PERIODICO di MINERALOGIA, v. 71, p. 101–109.

Nagase, T., and Kojima, S., 1997, An SEM examination of the chalcopyrite disease texture and its genetic implications: Mineralogical Magazine, v. 61, p. 89–97.

Ohmoto, H., 1996, Formation of volcanogenic massive sulfide deposits: The Kuroko perspective: Ore Geology Reviews, v. 10, p. 135–177.

Palarea-Albaladejo, J., and Martín-Fernández, J.A., 2008, A modified EM alralgorithm for replacing rounded zeros in compositional data sets: Computers & Geosciences, v. 34, p. 902–917.

——2015, zCompositions—R package for multivariate imputation of left-censored data under a compositional approach: Chemometrics and Intelligent Laboratory Systems, v. 143, p. 85–96.

Palarea-Albaladejo, J., Martín-Fernández, J.A., and Gómez-García, J., 2007, A parametric approach for dealing with compositional rounded zeros: Mathematical Geology, v. 39, p. 625–645.

Paton, C., Hellstrom, J., Paul, B., Woodhead, J., and Hergt, J., 2011, Iolite: Freeware for the visualisation and processing of mass spectrometric data: Journal of Analytical Atomic Spectrometry, v. 26, p. 2508–2518.

Patten, C.G.C., Pitcairn, I.K., Teagle, D.A.H., and Harris, M., 2016, Mobility of Au and related elements during the hydrothermal alteration of the oceanic crust: implications for the sources of metals in VMS deposits: Mineralium Deposita, v. 51, p. 179–200.

Patten, C.G.C., Pitcairn, I.K., and Teagle, D.A.H., 2017, Hydrothermal mobilisation of Au and other metals in supra-subduction oceanic crust: Insights from the Troodos ophiolite: Ore Geology Reviews, v. 86, p. 487–508.

Patten, C.G.C., Coltat, R., Junge, M., Peillod, A., Ulrich, M., Manatschal, G., and Kolb, J., 2022, Ultramafic-hosted volcanogenic massive sulfide deposits: an overlooked sub-class of VMS deposit forming in complex tectonic environments: Earth-Science Reviews, v. 224, article 103891.

Pawlowsky-Glahn, V., and Buccianti, A., 2011, Compositional data analysis: Theory and applications: John Wiley & Sons, 378 p.

Pedersen, R.-B., Furnes, H., and Dunning, G., 1991, A U/Pb age for the Sulitjelma Gabbro, North Norway: Further evidence for the development of a Caledonian marginal basin in Ashgill-Llandovery time: Geological Magazine, v. 128, p. 141–153.

Pilote, J.-L., Piercey, S.J., and Mercier-Langevin, P., 2020, Evolution of the subseafloor hydrothermal system associated with the Ming VMS deposit, Newfoundland Appalachians, and its controls on base and precious metal distribution: Mineralium Deposita, v. 55, p. 913–936.

Piña, R., Gervilla, F., Barnes, S.-J., Ortega, L., and Lunar, R., 2012, Distribution of platinum-group and chalcophile elements in the Aguablanca Ni-Cu sulfide deposit (SW Spain): Evidence from a LA-ICP-MS study: Chemical Geology, v. 302, p. 61–75.

Plouffe, A., Ferbey, T., Hashmi, S., and Ward, B., 2016, Till geochemistry and mineralogy: Vectoring towards Cu porphyry deposits in British Columbia, Canada: Geochemistry: Exploration, Environment, Analysis, v. 16, p. 213–232.

Provost, F., and Saar-Tsechanski, M., 2007, Handling missing values when applying classification models: Journal of Machine Learning Research, v. 8, p. 1625–1657.

Reich, M., Deditius, A., Chryssoulis, S., Li, J.-W., Ma, C.-Q., Parada, M.A., Barra, F., and Mittermayr, F., 2013a, Pyrite as a record of hydrothermal fluid evolution in a porphyry copper system: A SIMS/EMPA trace element study: Geochimica et Cosmochimica Acta, v. 104, p. 42–62.

Reich, M., Palacios, C., Barra, F., and Chryssoulis, S., 2013b, "Invisible" silver in chalcopyrite and bornite from the Mantos Blancos Cu deposit, northern Chile: European Journal of Mineralogy, v. 25, p. 453–460.

Reich, M., Román, N., Barra, F., and Morata, D., 2020, Silver-rich chalcopyrite from the active Cerro Pabellón geothermal system, Northern Chile: Minerals, v. 10, p. 113.

Relvas, J.M., Tassinari, C.C., Munhá, J., and Barriga, F.J., 2001, Multiple sources for ore-forming fluids in the Neves Corvo VHMS Deposit of the Iberian Pyrite Belt (Portugal): Strontium, neodymium and lead isotope evidence: Mineralium Deposita, v. 36, p. 416–427.

Relvas, J.M.R.S., Barriga, F.J.A.S., Ferreira, A., Noiva, P.C., Pacheco, N., and Barriga, G., 2006, Hydrothermal alteration and mineralization in the Neves-Corvo volcanic-hosted massive sulfide deposit, Portugal. I. Geology, mineralogy, and geochemistry: Economic Geology, v. 101, p. 753–790.

Ren, Y., Wohlgemuth-Ueberwasser, C.C., Huang, F., Shi, X., Li, B., Oelze, M., Schreiber, A., and Wirth, R., 2021, Distribution of trace elements in sulfides from Deyin hydrothermal field, Mid-Atlantic Ridge – Implications for its mineralizing processes: Ore Geology Reviews, v. 128, article 103911.

Revan, M.K., Genç, Y., Maslennikov, V.V., Maslennikova, S.P., Large, R.R., and Danyushevsky, L.V., 2014, Mineralogy and trace-element geochemistry of sulfide minerals in hydrothermal chimneys from the Upper-Cretaceous VMS deposits of the eastern Pontide orogenic belt (NE Turkey): Ore Geology Reviews, v. 63, p. 129–149.

Rimstidt, J.D., Chermak, J.A., and Gagen, P.M., 1993, Rates of reaction of galena, sphalerite, chalcopyrite, and arsenopyrite with Fe(III) in acidic solutions: American Chemical Society, ACS Symposium Series, v. 550, p. 2–13.

de Roo, J.A., and van Staal, C.R., 2003, Sulfide remobilization and sulfide breccias in the Heath Steele and Brunswick deposits, Bathurst Mining

Camp, New Brunswick: Society of Economic Geologists, Economic Geology Monograph 11, p. 479–496.

Rose, A.W., 1967, Trace elements in sulfide minerals from the Central district, New Mexico and the Bingham district, Utah: Geochimica et Cosmochimica Acta, v. 31, p. 547–585.

Ruiz-Perez, D., Guan, H., Madhivanan, P., Mathee, K., and Narasimhan, G., 2020, So you think you can PLS-DA?: BMC Bioinformatics, v. 21, article 207225.

Sciuba, M., and Beaudoin, G., 2021, Texture and trace element composition of rutile in orogenic gold deposits: Economic Geology, v. 116, p. 1865–1892.

Sciuba, M., Beaudoin, G., Grzela, D., and Makvandi, S., 2020, Trace element composition of scheelite in orogenic gold deposits: Mineralium Deposita, v. 55, p. 1149–1172.

Sciuba, M., Beaudoin, G., and Makvandi, S., 2021, Chemical composition of tourmaline in orogenic gold deposits: Mineralium Deposita, v. 56, p. 537–560.

Seward, T.M., Williams-Jones, A.E., and Migdisov, A.A., 2014, The chemistry of metal transport and deposition by ore-forming hydrothermal fluids, *in* Holland, H.D., and Tukerian, K.K., eds., Treatise on geochemistry: Elsevier, Oxford, p. 29–57.

Shanks, W.C., Koski, R.A., Mosier, D.L., Schulz, K.J., Morgan, L.A., Slack, J.F., Ridley, W.I., Dusel-Bacon, C., Seal Ii, R.R., and Piatak, N.M., 2012, Volcanogenic massive sulfide occurrence model: Chapter C in Mineral deposit models for resource assessment, in Shanks, W.C., and Thurston, R., eds., Scientific Investigations Report: U.S. Department of the Interior, U.S. Geological Survey, doi 10.3133/sir20105070C.

Skyttä, P., Weihed, P., Högdahl, K., Bergman, S., and Stephens, M.B., 2020, Paleoproterozoic (2.0–1.8 Ga) syn-orogenic sedimentation, magmatism and mineralization in the Bothnia-Skellefteå lithotectonic unit, Svecokarelian orogen, Sweden: Geological Society of London, Memoir 50, p. 83–130.

Slack, J.F., Offield, T.W., Woodruff, L.G., and Shanks, W.C., 2001, Geology and geochemistry of Besshi-type massive sulfide deposits of the Vermont copper belt: Society of Economic Geologists, Guidebook Series, v. 35, p. 193–211.

Spry, P.G., Peter, J.M., and Slack, J.F., 1998, Meta-exhalites as exploration guides to ore: Reviews in Economic Geology, v. 11, p. 163–201.

Stakes, D.S., and Taylor, H.P., 2003, Oxygen isotope and chemical studies on the origin of large plagiogranite bodies in northern Oman, and their relationship to the overlying massive sulphide deposits: Geological Society, London, Special Publications, v. 218, p. 315.

Steadman, J.A., and Large, R.R., 2016, Synsedimentary, diagenetic, and metamorphic pyrite, pyrrhotite, and marcasite at the Homestake BIF-hosted gold deposit, South Dakota, USA: Insights on Au-As ore genesis from textural and LA-ICP-MS trace element studies: Economic Geology, v. 111, p. 1731–1752.

Svetnik, V., Liaw, A., Tong, C., Culberson, J. C., Sheridan, R.P., and Feuston, B.P., 2003, Random Forest: A classification and regression tool for compound classification and QSAR modeling: Journal of Chemical Information and Computer Sciences, v. 43, p. 1947–1958.

Tarawneh, A.S., Hassanat, A.B., Altarawneh, G.A., and Almuhaimeed, A., 2022, Stop oversampling for class imbalance learning: A review: IEEE Access, v. 10, p. 47,643–47,660.

Thabtah, F., Hammoud, S., Kamalov, F., and Gonsalves, A., 2020, Data imbalance in classification: Experimental evaluation: Information Sciences, v. 513, p. 429–441.

Tomkins, A.G., 2007, Three mechanisms of ore re-mobilisation during amphibolite facies metamorphism at the Montauban Zn-Pb-Au-Ag deposit: Mineralium Deposita, v. 42, p. 627–637.

Tomkins, A.G., Pattison, D.R.M., and Frost, B.R., 2007, On the initiation of metamorphic sulfide anatexis: Journal of Petrology, v. 48, p. 511–535.

Tornos, F., 2006, Environment of formation and styles of volcanogenic massive sulfides: The Iberian Pyrite Belt: Ore Geology Reviews, v. 28, p. 259–307.

Torró, L., Benites, D., Vallance, J., Laurent, O., Ortiz-Benavente, B.A., Chelle-Michou, C., Proenza, J.A., and Fontboté, L., 2022, Trace element geochemistry of sphalerite and chalcopyrite in arc-hosted VMS deposits: Journal of Geochemical Exploration, v. 232, article 106882.

Vasyukova, O.V., and Williams-Jones, A.E., 2022, Constraints on the genesis of cobalt deposits: Part II. Applications to natural systems: Economic Geology, v. 117, p. 529–544.

Wang, Y., Han, X., Petersen, S., Frische, M., Qiu, Z., Cai, Y., and Zhou, P., 2018, Trace metal distribution in sulfide minerals from ultramafic-hosted hydrothermal systems: Examples from the Kairei vent field, Central Indian Ridge: Minerals, v. 8, p. 526.

Wang, Z., Xu, D., Zhang, Z., Zou, F., Wang, L., Yu, L., and Hu, M., 2015, Mineralogy and trace element geochemistry of the Co-and Cu-bearing sulfides from the Shilu Fe-Co-Cu ore district in Hainan Province of South China: Journal of Asian Earth Sciences, v. 113, p. 980–997.

Weihed, P., Weihed, J.B., Sorjonen-Ward, P., and Matsson, B., 2002, Post-deformation, sulphide-quartz vein hosted gold ore in the footwall alteration zone of the Palaeoproterozoic Långdal VHMS deposit, Skellefte District, northern Sweden: GFF, v. 124, p. 201–210.

Weiss, G.M., 2013, Foundations of imbalanced learning, *in* He, H., and Ma, Y., eds., Imbalanced learning: Foundations, algorithms, and applications: Wiley Online Library, p. 13–41.

Wilcox, R.R., 2011, Introduction to robust estimation and hypothesis testing: Academic press, 608 p.

Wilson, S.A., Ridley, W.I., and Koenig, A.E., 2002, Development of sulfide calibration standards for the laser ablation inductively-coupled plasma mass spectrometry technique: Journal of Analytical Atomic Spectrometry, v. 17, p. 406–409.

Wohlgemuth-Ueberwasser, C.C., Viljoen, F., Petersen, S., and Vorster, C., 2015, Distribution and solubility limits of trace elements in hydrothermal black smoker sulfides: An in-situ LA-ICP-MS study: Geochimica et Cosmochimica Acta, v. 159, p. 16–41.

Wold, S., Sjöström, M., and Eriksson, L., 2001, PLS-regression: A basic tool of chemometrics: Chemometrics and Intelligent Laboratory Systems, v. 58, p. 109–130.

Workman, R.K., and Hart, S.R., 2005, Major and trace element composition of the depleted MORB mantle (DMM): Earth and Planetary Science Letters, v. 231, p. 53–72.

Xing, B., Mao, J., Xiao, X., Liu, H., Zhang, C., Guo, S., Li, H., Huang, W., and Lai, C., 2022, Metallogenic discrimination by sphalerite trace element geochemistry: An example from the Fengyan Zn-Pb deposit in central Fujian, SE China: Ore Geology Reviews, v. 141, article 104651.

Zhang, J., Shao, Y., Liu, Z., and Chen, K., 2022, Sphalerite as a record of metallogenic information using multivariate statistical analysis: Constraints from trace element geochemistry: Journal of Geochemical Exploration, v. 232, article 106883.

Zhao, Z., Wei, J., Liang, S., and Gao, T., 2021, Sulfide remobilization and trace element redistribution during metamorphism and deformation at the Xitieshan Pb-Zn deposit, NW China: Ore Geology Reviews, v. 136, article 104170.

Enzo Caraballo is currently a Ph.D. student at the Université Laval in Québec (Canada), working on trace element composition of chalcopyrite and its potential as an indicator mineral for exploration. He graduated as an engineering geologist from Universidad Central de Venezuela and obtained a Master 2 degree at the Université Savoie-Mont Blanc (France). His research interest involves ore mineral chemistry, ore deposit genesis, multivariate statistical analysis, and machine learning applied to mineral exploration. He worked more than seven years as a lecturer in mineralogy and petrology at the Universidad Central de Venezuela.