

Garnet major-element composition as an indicator of host-rock type: a machine learning approach using the random forest classifier

Jan Schönig^{1*}, Hilmar von Eynatten¹, Raimon Tolosana-Delgado², Guido Meinhold^{1,3,4}

¹*Geoscience Centre Göttingen, University of Göttingen, Goldschmidtstraße 3, 37077 Göttingen, Germany*

²*Helmholtz-Zentrum Dresden-Rossendorf, Helmholtz Institute Freiberg for Resources Technology, Chemnitz Straße 40, 09599 Freiberg, Germany*

³*School of Geography, Geology and the Environment, Keele University, Keele, Staffordshire, ST5 5BG, UK*

⁴*Institute for Geology, TU Bergakademie Freiberg, Bernhard-von-Cotta-Straße 2, 09599 Freiberg, Germany*

**corresponding author: jan.schoenig@uni-goettingen.de*

Supplementary Information 2

- **Extended information to *Model development* on pages 2–4**
- **Extended information to *Performance of the ‘setting and metamorphic facies’ model* on page 5–12**
 - **Performance of the discrimination scheme after Mange and Morton (2007) on pages 5–7**
 - **Performance of the discrimination scheme after Hardman et al. (2018) on pages 7–8**
 - **Performance of the discrimination scheme after Tolosana-Delgado et al. (2018) on pages 9–10**
 - **Effect of pressure and temperature on votes on pages 11–12**
- **Extended information to *Variable importance of the ‘setting and metamorphic facies’ model* on pages 13–14**
- **Extended information to *Variable importance of the ‘composition’ model* on pages 15–16**
- **Extended information to *Origin of main discriminators* on pages 17–23**
- **Extended *References* on pages 24–27**

Abbreviations

A – Alkaline

AM – Amphibolite facies

BS/GS – Blueschist/Greenschist facies

CS – Calcsilicate

EC/UHP – Eclogite/Ultrahigh-pressure facies

GR – Granulite facies

IF/S – Intermediate–felsic/metasedimentary

IG – Igneous rocks

M – Mafic

MA – Mantle rocks

MM – Metamorphic rocks

MS – Metasomatic rocks

UM – Ultramafic

Extended information to *Model development*

Random forest is a machine learning algorithm for supervised classification and regression developed by Breiman (2001). To understand the decisions made during the development of the garnet discrimination model (see *Model development* section), a description of the principle of the random forest classifier is given here. For a more detailed and mathematically based explanation, the reader is referred to the original work of Breiman (2001) and reviews treating this topic (Boulesteix et al. 2012; Ziegler and König 2014; Belgiu and Drăgut 2016; Biau and Scornet 2016).

The basis of the random forest algorithm are decision trees. Instead of growing a single tree trained on the classes of a dataset, though, an ensemble of trees is created, that is, a forest. Figure S-1 shows a flow chart that illustrates the principle of creating a random forest classification model on a simplified example. The example model is based on a database comprising 50 observations, each described by numerical values of ten variables. At first, the user has to split the dataset by defining classes that are aimed to be predicted by the model, here A, B, and C (step (a) in Fig. S-1). The number of observations in each class may be highly imbalanced as exemplified with a range from five (class “B”) to 30 observations (class “A”).

For a tree to be grown, the algorithm picks a random sample of observations (eventually, with repetition) from the defined classes of the dataset (step (b) in Fig. S-1). Afterwards, the sample is passed back to the dataset and a new random sample is taken for the next tree. Consequently, the random samples of the individual trees can resample observations included in the samples of other trees, like highlighted with the red colour-coded observations in the random samples 0001 and 1000 (Fig. S-1). As class “B” has five observations only, this holds for all trees in the example. This procedure, called bootstrapping, is repeated until the user defined number of trees to be grown (‘ntree’ parameter) is reached, here 1000. The number of observations taken for each sample from each of the classes can be the full number of observations available or can be defined by the ‘sampsize’ parameter, that is the size of samples to be drawn. In most cases, the overall classification success is higher when the total number of observations is considered for each sample. However, for an imbalanced dataset like the example, this high classification success may result from a preferred correct classification of the class with the highest number of observations. For instance, a classification success of 90 % corresponds to correctly predicted classes for 45 of the 50 observations, which could be generated by perfect prediction of “A” (30 of 30, i.e. 100 %), a fair prediction of “C” (13 of 15, i.e. 87 %), and poor prediction of “B” (2 of 5, i.e. 40 %). To avoid such high discrepancies, in the example the ‘sampsize’ parameter is chosen to be equal to the number of observations in the class with the lowest number of observations, that is class “B” with five observations. This approach usually results in a slightly lower classification success for the entire dataset, but more balanced success rates for the individual classes (Chen et al. 2004).

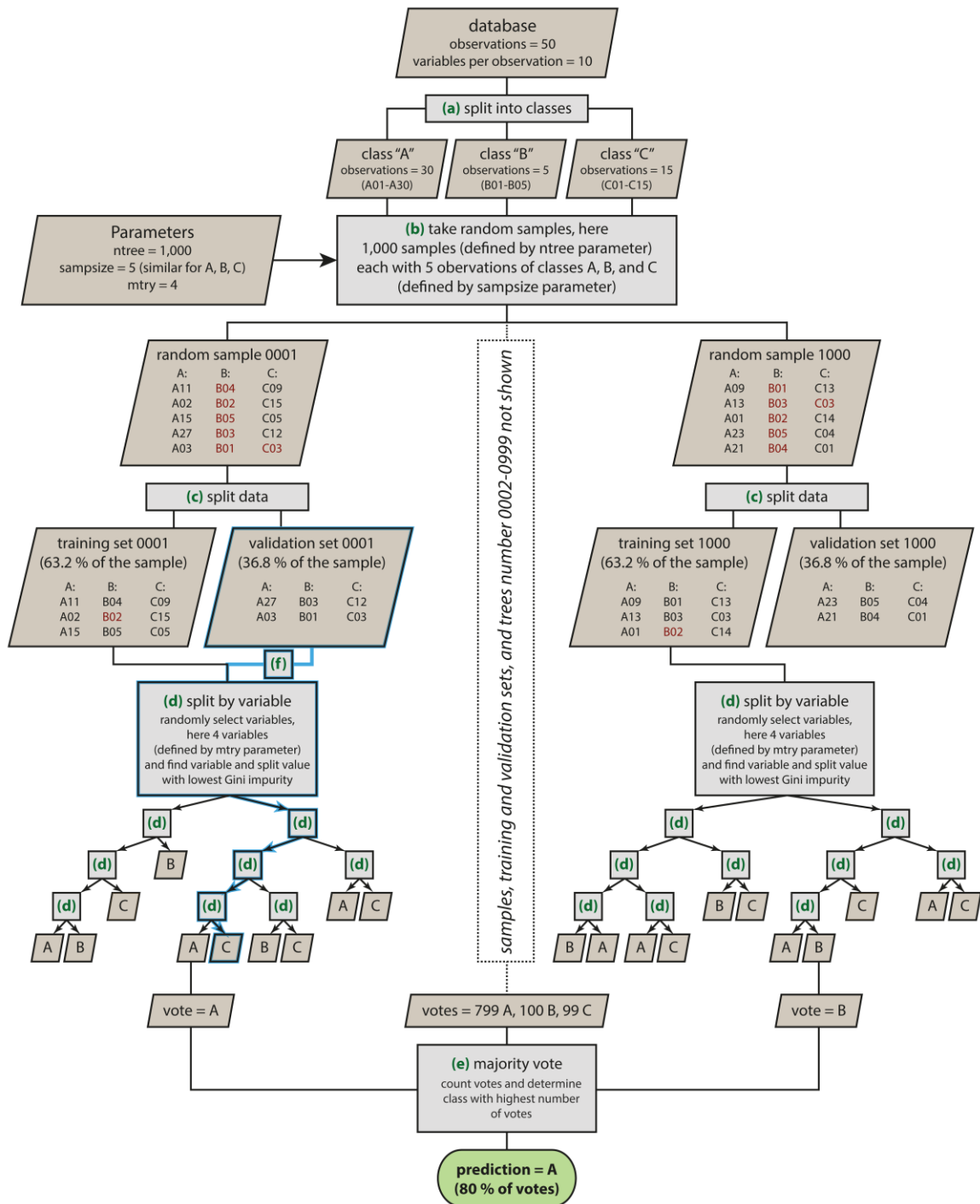


Figure S-1: Flow chart of a simplified example to create a random forest classification model, inspired by Belgiu and Drăgut (2016).

After a training sample is taken for a tree (step (b) in Fig. S-1), a validation set is constructed (step (c) in Fig. S-1), containing all observations excluded from the training set. The training set is used to grow the tree, while the validation set is used to evaluate its performance. At each node of the tree, the algorithm randomly picks a subset of variables from the dataset defined by the ‘mtry’ parameter, that is the number of variables randomly sampled and considered as candidates at each split. For classification, the default of ‘mtry’ is defined as the square root of the number of observations (Liaw

and Wiener 2002). In the example, 'mtry' was chosen by the user to be four. From these variable candidates, the algorithm selects one variable at a time from the variable subset and a threshold for this variable that shows the highest decrease in Gini impurity is computed (step (d) in Fig. S-1). Gini impurity measures the expected probability of incorrectly classifying a randomly chosen element in the dataset if it were randomly labelled according to the class distribution in the dataset (Breiman et al. 1984). Then, the algorithm picks the variable and the threshold that split the data with the highest reduction of the Gini impurity. This splitting produces two branches. The step is repeated branch by branch, creating new nodes in a hierarchical scheme, until all samples are assigned to their true class (with end branches called then leaves), resulting in a deep tree without pruning (optionally trees can be pruned to save computation time by increasing the 'nodesize' parameter, that is the minimum number of observations on each terminal node). Subsequently, an ensemble of trees, that is the forest, is created based on the training sets of all random samples taken defined by the 'ntree' parameter (step (b) in Fig. S-1).

The random selection of a subset of variables at each node makes the difference between trees grown using the random forest algorithm and trees grown using classification and regression trees (CART, Breiman et al. 1984). In CART, the best split variable is chosen from the entire set of variables. Thus, trees for the samples grown by CART can share high similarities, i.e. have a high correlation, because the best splitting variables are often the same. By contrast, trees grown by random forest show a much lower correlation, which becomes powerful in the last step of the random forest algorithm called bagging, or bootstrap aggregation. When predicting the class of an unknown observation, the sample will run through all of the trained trees. The prediction of a certain observation for a class made by a tree is called a 'vote'. Finally, the votes of all trees are counted and the class with the highest number of votes wins, that is the majority vote (step (e) in Fig. S-1, in this case 800 A, 101 B, 99 C). Similarly, the validation sets created are used to estimate the performance of the random forest model created by the training sets (step (f) in Fig. S-1, blue path). That is, a certain observation contributes either to the training or to the validation of each tree. The samples that have not contributed to the training of a certain tree are called out-of-bag (OOB) samples. The OOB error is defined by the sum of misclassified OOB observations divided by the total number of observations.

Extended information to *Performance of the ‘setting and metamorphic facies’ model*

Performance of the discrimination scheme after Mange and Morton (2007)

The discrimination scheme of Mange and Morton (2007) was originally developed to discriminate different source regions of garnet-bearing sedimentary rocks in the North Sea and Norwegian Sea (Morton 1985; Morton et al. 2004, 2005a, 2005b). The approach was to compare the composition of detrital garnet grains with that of potential surrounding source regions including Norway, Scotland, and Greenland. It was demonstrated that the major-element chemistry of detrital garnet is useful to trace back sediment from sink to source, with grains being derived from units dominantly consisting of amphibolite-facies rocks showing low Mg contents, those from mafic granulite- and eclogite-facies rocks showing high Mg and high Ca, and those from metasedimentary granulite-facies rocks having high Mg and low Ca. By considering several other garnet host rocks, Mange and Morton (2007) defined six different garnet types (A, Bi, Bii, Ci, Cii, and D) in a graphical ternary discrimination scheme based on the molar proportions of Fe+Mn, Mg, and Ca.

Figure S-2 shows the performance of this scheme based on the garnet database presented here (Supplementary Information 1). In Figure S-2a the performance of the classification scheme is evaluated solely for garnet compositions from host rocks which have been considered by Mange and Morton (2007). Results are represented splitting the data according to the classes defined in the present contribution: both probability density maps (showing the location and spread of the garnet composition belonging to each group) and performance barplots (showing the proportion of garnets from these group being classified on the classes of Mange and Morton (2007)) are displayed. From these barplots it is obvious that only mafic mantle garnets (MA-M) are not well represented by the respective compositional field. All other groups show a discrimination success of between 77 % and 97 %, with 88 % on average (i.e., the frequency of correctly classified garnets). Consequently, the scheme performs well for the application it was designed for.

The main limitation of the Mange and Morton (2007) scheme for provenance analysis refers to the usage as a petrogenetic indicator that arises from the wide range of host rocks represented by each defined garnet type and the fact that garnets from several specific host rocks are represented by more than one defined type. This is highlighted by the kernel density maps shown in Figure S-2a and even more clearly in Figure S-2b when reporting all subgroups included in the database into groups (Table 1). High compositional overlaps are given for granulite- and eclogite-facies garnet, in particular for mafic host-rock compositions, as well as for blueschist-/greenschist- and amphibolite-facies garnet. In addition, compositions of igneous garnet spread over a large part of the ternary plot and strongly overlap with compositions of blueschist-/greenschist- and amphibolite-facies garnet, but also overlap with higher-grade metamorphic garnet and metasomatic garnet.

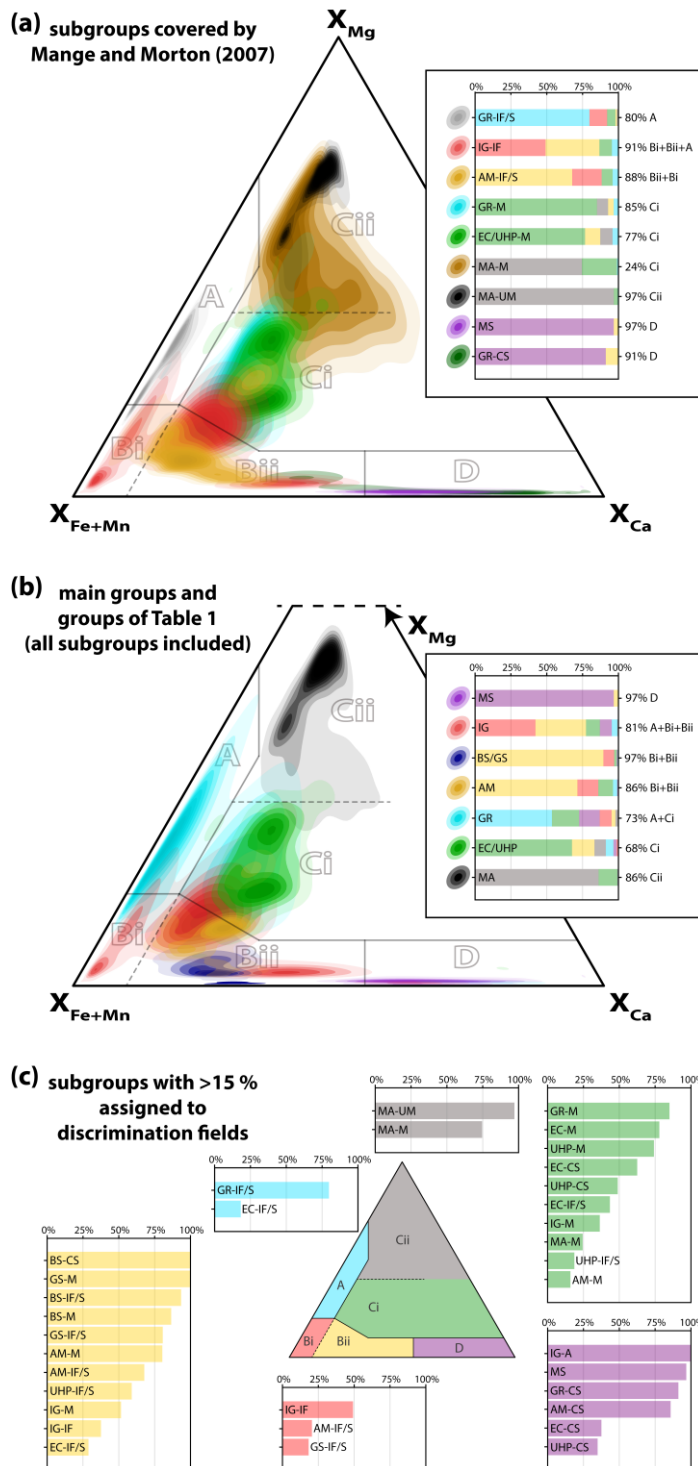


Figure S-2: Performance of the garnet discrimination scheme after Mange and Morton (2007). Discrimination fields correspond to type A – mainly derived from high-grade granulite-facies metasedimentary rocks or charnockites, but can also be supplied from intermediate–felsic igneous rocks; type B – mainly derived from amphibolite-facies metasedimentary rocks. Populations plotting exclusively in Bi suggest derivation from intermediate–felsic igneous rocks; Type C – derived mainly from high-grade metamafic rocks. High Mg contents (Type Cii) imply sourcing from ultramafic rocks; Type D – derived from metasomatic rocks, very low-grade metamafic rocks, or ultrahigh-temperature calc-silicate. (a) performance for all subgroups of Table 1 considered by Mange and Morton (2007); (b) performance for all groups of Table 1; (c) proportions of individual subgroups given in Table 1 that are assigned to the specific discrimination fields, solely amounts >15 % are shown. Compositions are given in molar proportions and shown as kernel density estimate maps in ternary diagrams, with bandwidth calculated after Venables and Ripley (2002). Assignments to garnet types are shown as barplots, colour coding corresponds to ternary diagram in (c), percentage to the right of the bars indicate assignment to the respective type(s) according to the scheme.

Implications of the compositional overlaps for provenance interpretations are evaluated in Figure S-2c, which shows how many subgroups of the database are represented in significant proportions >15 % by each of the six garnet types defined by Mange and Morton (2007). The best performance is given by type Cii garnet (mantle garnet), followed by type A garnet mostly indicative for intermediate–felsic/metasedimentary granulite-facies sources and lower but significant amounts of intermediate–felsic/metasedimentary eclogite-facies sources. Type Bi mostly represents garnet from

intermediate–felsic igneous host rocks, followed by significant proportions of greenschist- and amphibolite-facies intermediate–felsic/metasedimentary garnet. Most critical is type Bii garnet. Although most of the type Bii garnet belongs to lower grade metamorphic source rocks, i.e. greenschist, blueschist, and amphibolite facies, high amounts of intermediate–felsic/metasedimentary eclogite- and ultrahigh-pressure-facies garnets are also assigned to type Bii, as well as igneous garnet of intermediate–felsic to mafic composition. Conversely, most of Ci type garnet belongs to high-grade metamorphic source rocks, i.e. granulite, eclogite, and ultrahigh-pressure facies, but significant amounts of mafic igneous, mantle, and amphibolite-facies garnet are also represented by type Ci. Type D garnet represents metasomatic, alkaline igneous, and calcsilicate metamorphic garnet ranging from amphibolite to ultrahigh-pressure facies.

In conclusion, although changes in garnet composition observed in the $X_{\text{Fe+Mn}}-X_{\text{Mg}}-X_{\text{Ca}}$ ternary plot can imply a change in provenance, the large compositional overlaps from a wide range of garnet host rocks make it difficult to attribute these changes to the setting, geodynamic context, or protolith composition. For instance, only prominent changes like sediment containing dominantly type Bi and Bii garnets that is stratigraphically overlain by strata containing dominantly type A and Ci garnets may indicate the ongoing exhumation of deeper crustal units. Nevertheless, even such an interpretation is vulnerable considering that a similar trend of changing garnet composition could be produced by a change in protolith composition from intermediate–felsic/metasedimentary to mafic for high-grade metamorphic rocks. This becomes even more challenging given the strong overlap with garnet from igneous rocks.

Performance of the discrimination scheme after Hardman et al. (2018)

To reduce costs in diamond exploration campaigns, first surveys often focus on the identification of diamond/kimberlite indicator minerals like chromite, clinopyroxene, and garnet in modern sediments derived from potential target areas, which are by far more frequent than diamond (Nowicki et al. 2007). Garnet is of particular importance due to its occurrence in all dominant diamond-bearing lithologies. Thus, a robust geochemical tool to discriminate between mantle and crustal garnet is of economic interest. Ultramafic mantle garnet can be confidently discriminated from crustal garnet by the higher Cr_2O_3 component (e.g. Grütter et al. 2004) and its MgO enrichment relative to FeO (e.g. Schulze 2003; Tolosana-Delgado et al. 2018). In contrast, garnet of low-chromium (<1 wt% Cr_2O_3) mantle eclogite compositionally overlap with crustal garnet, especially with granulite- and eclogite-facies metamorphic garnet (Hardman et al. 2018). This leads to high misclassifications in the most applied graphical scheme for mantle versus crustal garnet after Schulze (2003), as demonstrated in Hardman et al. (2018; see their fig. 2).

Hardman et al. (2018) addressed the problem of discriminating low-chromium mantle and crustal garnet by a more robust database and a multivariate approach using linear discriminant analysis and logistic regression, resulting in a classification success of 93 to 94 %. The application requires knowledge of the Na₂O component, which is not covered by our database because Na₂O is often not reported, not measured or below detection limit, in particular for garnet of crustal origin. Thus, the graphical scheme of Hardman et al. (2018) is evaluated here, which is based on the natural logarithmic ratio of Ti to Si versus Mg to Fe, and reported to provide a high discrimination success of 90 %.

Figure S-3 shows the performance of the graphical scheme after Hardman et al. (2018) for main and subgroups of mantle garnet as well as groups of metamorphic garnet (Table 1). Mantle versus crustal garnet are correctly classified in 95 % of the cases (average). Even those groups presenting more difficulties to other schemes due to their lower wt% Cr₂O₃, that are garnet of mafic mantle rocks (98 % of observations <1 wt% Cr₂O₃) versus metamorphic eclogite- and ultrahigh-pressure garnet (100 % of observations <1 wt% Cr₂O₃), show a high classification success of on average 91 %. This also holds for low-chromium garnet from the group of ultramafic mantle rocks (36 % of observations <1 wt% Cr₂O₃), which are correctly classified in 95 % of the cases.

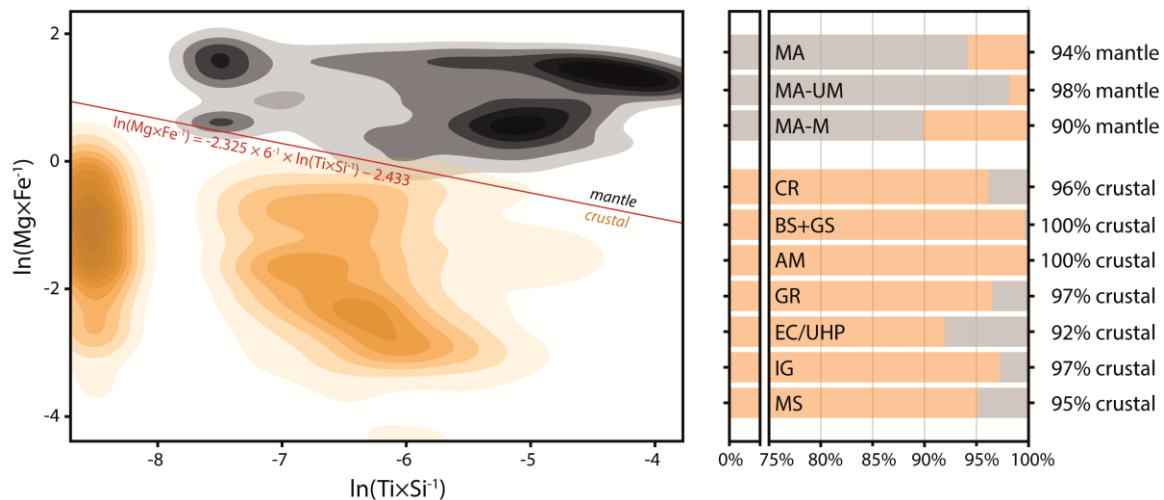


Figure S-3: Performance of the graphical mantle-versus-crustal garnet discrimination scheme after Hardman et al. (2018). Decision boundary with the highest overall classification success was chosen. As many observations of the database show TiO₂ values below detection limit (<0.03 wt%), TiO₂ values of crustal and mantle garnets have been replaced by 0.01 wt% and 0.03 wt%, respectively (most optimistic approach). Compositions are given as log-ratios of atomic proportions and are shown as kernel density estimate heatmaps, with bandwidth calculated after Venables and Ripley (2002). Assignments to crustal-versus-mantle sources are shown as barplots.

Performance of the discrimination scheme after Tolosana-Delgado et al. (2018)

To tackle the uncertainty derived from the large compositional overlaps, strict discrimination fields, and the usage of fewer variables than available, Tolosana-Delgado et al. (2018) applied a multivariate approach based on linear discriminant analysis to provide probabilities for detrital garnet of belonging to specific host rocks. In three hierarchical steps, the discrimination scheme first predicts probabilities for an ultramafic source, followed by an igneous source, and finally gives probabilities for metamorphic garnet of belonging to amphibolite-, granulite-, and eclogite-facies sources. Because Tolosana-Delgado et al. (2018) used a threshold of 0.02 wt% for Cr_2O_3 to apply the linear discriminant function for separating ultramafic garnet, which is not covered by the database detection limit of 0.03 wt%, and considered only the most abundant igneous garnet-bearing rock type, that is felsic plutonic, we focus here on the performance in discriminating metamorphic garnet sources into facies types, which is most challenging. The scheme requires a prior probability (see Tolosana-Delgado et al. 2018). Due to the unbalanced number of observations for the various groups included in the database (Table 1), a customized prior based on the number of observations turned out to favour the correct classification of eclogite-facies garnet (including ultrahigh-pressure) but strongly reduces the classification of granulite-facies garnet (only 10 % are correctly classified). To not promote a specific metamorphic group, the prior probability ‘equal-M’ was chosen (equal proportions of metamorphic host-rock types) for the following performance assessment, that is 30 % for A (eclogite facies), 30 % for B (amphibolite facies), 30 % for C (granulite facies), 5 % for D (ultramafic), and 5 % for E (felsic plutonic).

Figure S-4 shows the scheme performance based on the garnet database presented here (Supplementary Information 1), using density maps and barplots as in Figure S-2. In Figure S-4a solely garnet compositions from subgroups of Table 1 are tested which have been included in Tolosana-Delgado et al. (2018) with >33 observations. Barplots indicate that the majority of garnet from all subgroups show the highest probability for their respective host rock; on average 69 % classification success. Eclogite- and amphibolite-facies garnets are much better classified than granulite-facies garnet. By considering all groups of the database, the average success rate is 65 % (Fig. S-4b).

Although the classification success rates are lower compared to the classification after Mange and Morton (2007), the value of information in terms of the geodynamic context is higher. Giving probabilities instead of strict assignments is the major strength of the scheme. This makes better use of the available space in the ternary plot as shown by the kernel density maps (Fig. S-4a and S-4b). Notably, most misclassified grains show almost equal probabilities for two host-rock groups, and thus variations in detrital garnet composition can be more confidently related to metamorphic facies with

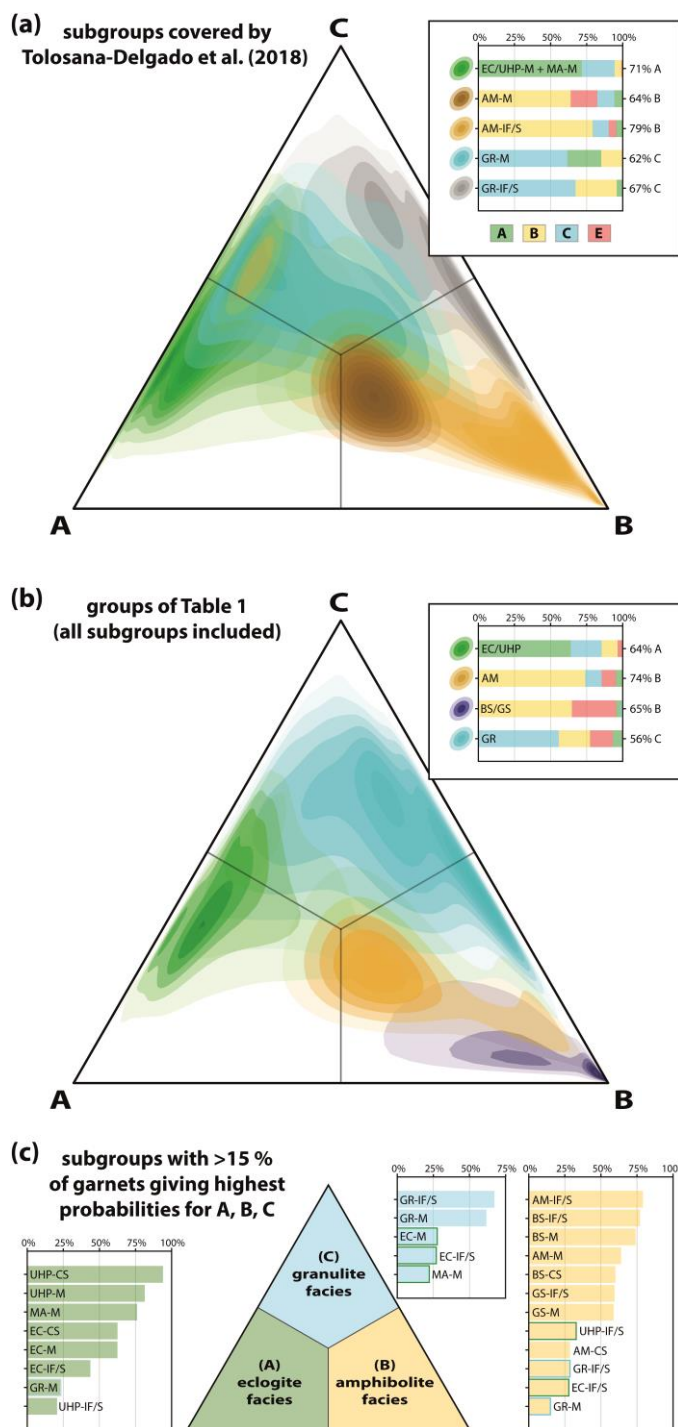


Figure S-4: Performance of the garnet discrimination scheme after Tolosana-Delgado et al. (2018) using the prior probability 'equal-M'. Host-rock types correspond to type A – eclogite facies; B – amphibolite facies; type C – granulite facies. (a) performance for all subgroups of Table 1 considered by Tolosana-Delgado et al. (2018); (b) performance for all groups of Table 1; (c) proportions of individual subgroups given in Table 1 that give the highest probability for individual host-rock types, solely amounts >15 % are shown. Probabilities are shown as kernel density estimate heatmaps in ternary diagrams, with bandwidth calculated after Venables and Ripley (2002). Assignments to host-rock types are shown as barplots, colour coding (frame: true class; filling: predicted class) corresponds to ternary diagram in (c) and type E (igneous) is shown in red.

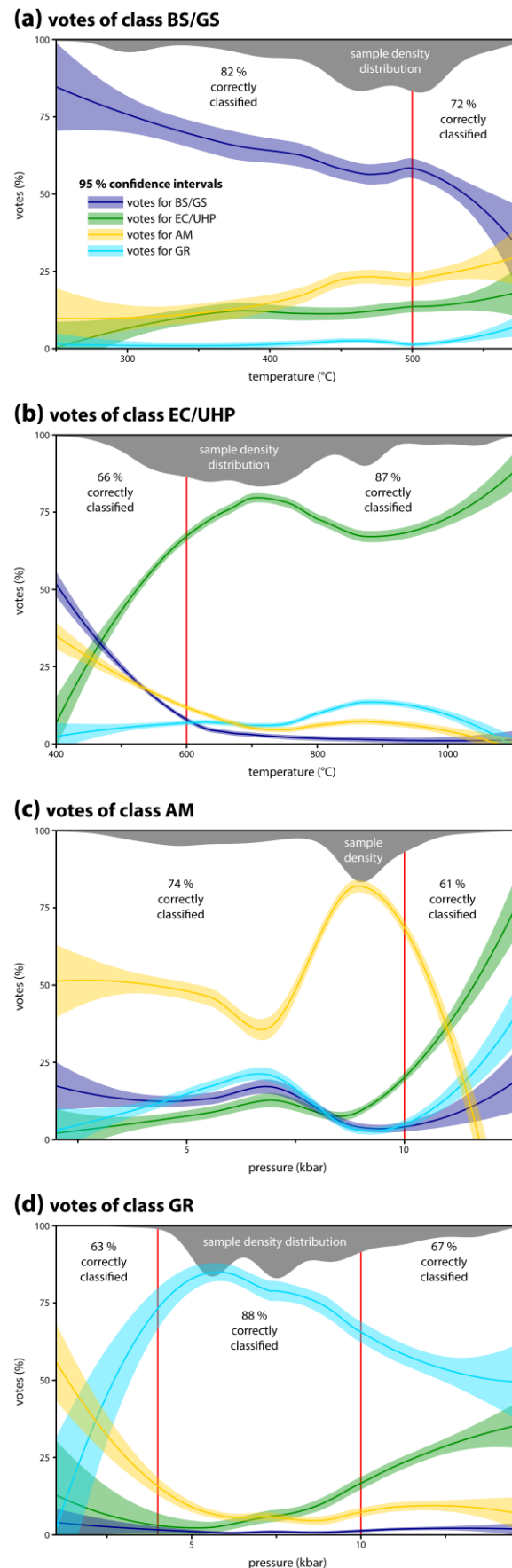
less influence of protolith composition on the results. However, significant amounts of eclogite-facies and mafic mantle garnet are incorrectly assigned to a granulite-facies source, those of mafic granulite facies to eclogite facies, and those of granulite, eclogite, and ultrahigh-pressure facies to amphibolite facies (see framed bars in Fig. S-4c). In addition, high proportions of metamorphic garnets are misclassified as felsic plutonic, in particular those of calcsilicate protolith composition and greenschist, blueschist, and amphibolite facies (not shown in Fig. S-4).

Effect of pressure and temperature on votes

A reflection of a continuous change in garnet composition with changing P – T conditions by voting results is highlighted in Figure S-5, which shows the internal distribution of votes within the individual metamorphic classes as a function of pressure and temperature for garnet in the database where P – T information is available. Classes BS/GS and EC/UHP are most sensitive to temperature. Garnets from class BS/GS show a continuous decrease in votes for BS/GS with increasing temperature, while votes for AM and, less clearly, EC/UHP increase (Fig. S-5a). This is exemplified by comparing BS/GS garnets that formed at $T \geq 500$ °C having a higher misclassification rate (28%) compared to $T < 500$ °C (18%). Contrarily, garnets from class EC/UHP show a significant decrease in votes for EC/UHP at $T < 600$ °C along with an increase for BS/GS and AM (Fig. S-5b).

Classes AM and GR are more sensitive to pressure. At $P > 10$ kbar a significant decrease in AM votes for garnets from class AM is

Figure S-5: Votes within the individual metamorphic classes as a function of temperature (a, b) or pressure (c, d). As trends are not obvious from the highly scattered data, vote regression curves are fitted and smoothed using local regression (locally estimated scatterplot smoothing) for classes (a) BS/GS ($n = 662$; 55 % of database observations), (b) EC/UHP ($n = 3,186$; 92 % of database observations), (c) AM ($n = 1,327$; 51 % of database observations), and (d) GR ($n = 856$; 85 % of database observations). Grey area indicates the distribution of observations given as kernel density estimate on a reverse relative scale (sample density distribution). Note that 95 % confidence intervals correspond to the estimated regression curve



observed, while votes for EC/UHP strongly increase simultaneously with a minor increase in GR and BS/GS votes (Fig. S-5c). Votes for garnets from class GR show a two-fold distribution with garnets crystallized at $P < 4$ kbar being more prone to be misclassified as AM and those crystallized at $P > 10$ kbar being more prone to be misclassified as EC/UHP (Fig. S-5d). The effect of pressure and temperature is highlighted for the individual metamorphic classes in the respective ternary plots by comparing the geometric means (Fig. S-6).

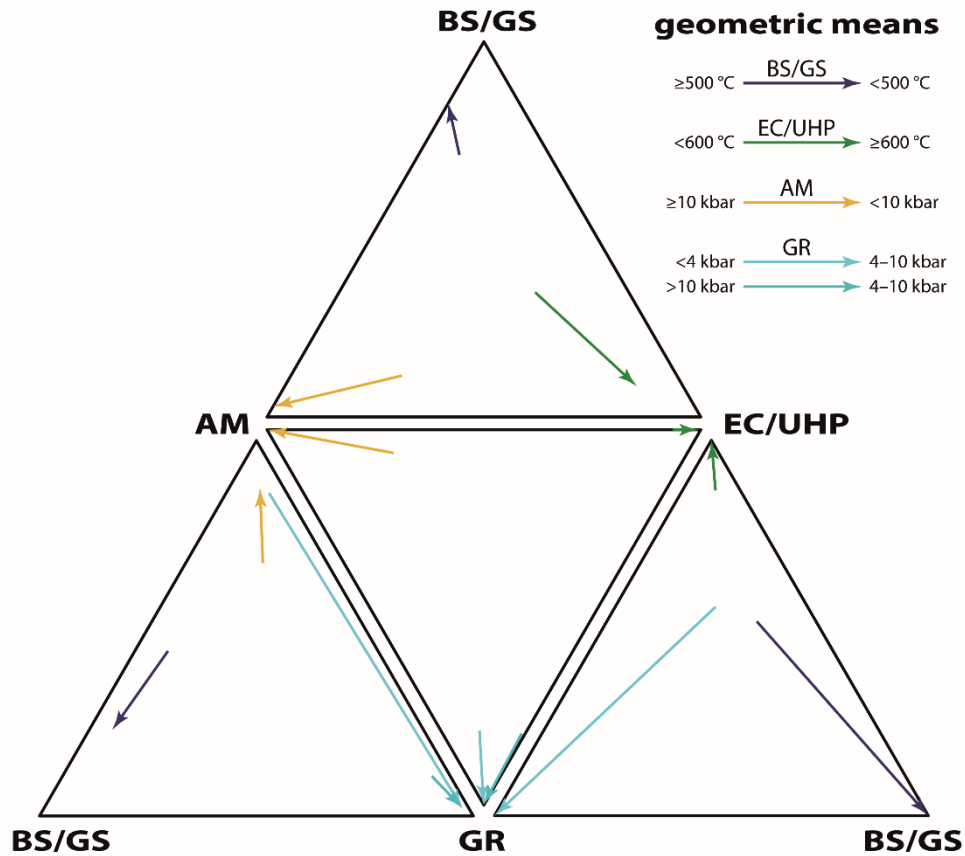


Figure S-6: Effect of pressure and temperature on geometric mean votes of metamorphic facies classes. Note that the AM–BS/GS–EC/UHP diagram is most important for classes BS/GS and AM, and the EC/UHP–GR–AM diagram is most important for classes EC/UHP and GR (cf. Fig. 1c).

Extended information to *Variable importance of the ‘setting and metamorphic facies’ model*

Consequences on misclassification rates for individual classes of the ‘setting and metamorphic facies’ model by excluding individual oxides

The highest influence on misclassification rates of class MA is observed by excluding TiO_2 , followed by MnO and Cr_2O_3 (Fig. 4, black line). In comparison to the full model (all eight oxides used), the vast majority of newly misclassified MA garnets are mafic (92 % for TiO_2 excluded, 81 % for MnO, and 75 % for Cr_2O_3) and now assigned to class EC/UHP (97 % for TiO_2 excluded, 93 % for MnO, and 91 % for Cr_2O_3). Thus, TiO_2 , MnO, and Cr_2O_3 are important variables for discriminating between mantle eclogites (subgroup MA-M) and crustal eclogites (subgroup EC/UHP-M).

For class MS, misclassification rates show the highest increase by excluding MgO (consistent with observations based on Gini impurity), followed by MnO and TiO_2 (Fig. 4, purple line). Excluding each of the three oxides mainly leads to newly misclassified MS garnets as GR (79 % for MgO, 54 % for MnO, and 76 % for TiO_2), followed by IG (31 % for MnO) and BS/GS (15 % for MnO, and 12 % for TiO_2). This indicates the importance of MgO, MnO, and TiO_2 for discriminating between metasomatic and low-pressure metamorphic as well as igneous garnet.

Variables TiO_2 and CaO have furthermore high importance for the correct classification of IG garnet, followed by MnO (Fig. 4, red line). The increase in misclassification rate by excluding TiO_2 is not only related to titanium-rich garnet of alkaline igneous rocks (16 % of newly misclassified garnets), but also to garnet of intermediate–felsic (77 %) and mafic (5 %) host-rock composition. In addition, although most of them are now assigned to AM (35 %), all other metamorphic classes are affected (22 % EC/UHP, 14 % GR, and 12 % BS/GS) as well as MS (17 %). Misclassifications by excluding CaO and MnO are mainly related to intermediate–felsic/metasedimentary host-rock compositions (97 % for CaO, and 93 % for MnO). CaO is most important for the discrimination to AM (51 %), followed by BS/GS (33 %). MnO is also most important for the discrimination to AM (63 %), followed by BS/GS and GR (both 11 %).

The highest influence on misclassification rates of class BS/GS are observed by excluding MgO, followed by CaO, and also $\text{FeO}_{\text{total}}$ and Al_2O_3 show considerable importance (Fig. 4, dark-blue line). In comparison to the full model, the newly misclassified BS/GS garnets correspond to both intermediate–felsic/metasedimentary as well as mafic host-rock compositions. All four oxides are of particular value for the discrimination of BS/GS to AM (54 % of newly misclassified garnets for MgO, 58 % for CaO, 47 % for $\text{FeO}_{\text{total}}$, and 84 % for Al_2O_3). For the discrimination to EC/UHP, MgO is most important (26 %), followed by $\text{FeO}_{\text{total}}$ (22 %), Al_2O_3 (16 %), and CaO (11 %). Conversely,

CaO is most important for the discrimination to IG (30 %), followed by FeO_{total} (25 %) and MgO (17 %), but Al₂O₃ does not affect the classification success.

For class AM, misclassification rates show the highest increase by excluding MnO and MgO, followed by TiO₂ (Fig. 4, yellow line). In comparison to the full model, newly misclassified AM garnets are mainly intermediate–felsic/metasedimentary (70 % for MnO excluded, 80 % for MgO, and 65 % for TiO₂) and mafic garnet constitute the subordinate portion. All three oxides show high importance for the discrimination of AM garnet to all other metamorphic classes. MnO is most important for the discrimination to EC/UHP, followed by BS/GS and GR, and MgO and TiO₂ are most important for BS/GS, followed by EC/UHP and GR. In addition, all three oxides affect the distinction of AM and IG, with TiO₂ being most important, followed by MnO and MgO.

The highest influence on misclassification rates of class GR are observed by excluding MnO, followed by CaO, and also MgO and TiO₂ show considerable importance (Fig. 4, sky-blue line). In comparison to the full model, the majority of newly misclassified GR garnets correspond to mafic host-rock compositions. All four oxides are of particular value for the discrimination of GR to EC/UHP (80 % of newly misclassified garnets for MnO, 51 % for CaO, 34 % for MgO, and 29 % for TiO₂). Considerable importance is also given for the discrimination of GR to AM (10 % for MnO, 36 % for CaO, 18 % for MgO, and 17 % for TiO₂). As observed for class MS, TiO₂ and MgO are important to distinguish GR garnet from MS garnet, with TiO₂ being also important for the discrimination to IG.

As for class GR, MnO, CaO, MgO, and TiO₂ are also important for the discrimination of EC/UHP versus all other metamorphic classes as well as MA (Fig. 4, green line). In addition, Cr₂O₃ shows considerable importance. To distinguish EC/UHP garnet from MA garnet, Cr₂O₃ is the most important oxide (49 % of newly misclassified garnets by excluding Cr₂O₃), followed by TiO₂ (40 %), MgO (19 %), and MnO (17 %). For the discrimination to GR garnet, CaO is most important (44 %), followed by MnO (26 %), Cr₂O₃ (21 %), TiO₂ (16 %), and MgO (13 %). MgO is most important for the discrimination to AM (39 %), followed by MnO (32 %), and CaO as well as TiO₂ (both 17 %). For the discrimination to BS/GS, CaO is most important (27 %), followed by Cr₂O₃ (22 %), MgO (21 %), TiO₂ (15 %), and MnO (14 %).

Extended information to *Variable importance of the ‘composition’ model*

Consequences on misclassification rates for individual classes of the ‘composition’ model by excluding individual oxides

The ln-ratios including TiO_2 are placed in a subordinate role for the mean decrease in Gini impurity and accuracy (Fig. 5a). In contrast, excluding TiO_2 when developing the ‘composition’ discrimination model leads to the highest increase in misclassification rates, followed by CaO , MgO , MnO , Cr_2O_3 , $\text{FeO}_{\text{total}}$, SiO_2 , and Al_2O_3 in decreasing order of importance (Fig. 6).

This is particularly related to the distinction of A and CS garnet (Fig. 5b), whereby the misclassification rate for A is 3.5 times higher when TiO_2 is excluded (Fig. 6, red line). In addition, TiO_2 is important for the classification success of IF/S and M (Fig. 6, yellow and green lines), although compositional differences are rather small (Fig. 5b). Excluding TiO_2 leads to misclassifications of IF/S as M, and M as IF/S for crustal garnet as well as M as UM for mantle garnet (see below).

Besides TiO_2 , the highest influence on misclassification rates is given by excluding CaO , followed by MgO , MnO , Cr_2O_3 , $\text{FeO}_{\text{total}}$, SiO_2 , and Al_2O_3 in decreasing order of importance (Fig. 6, bold grey line). Class UM is least affected by excluding oxides and solely the exclusion of MnO and $\text{FeO}_{\text{total}}$ leads to a considerable increase in the misclassification rate (Fig. 6, black line), whereby UM garnet being misclassified as M (92 % of newly misclassified garnets).

For class M, misclassification rates show the highest increase by excluding TiO_2 (see above), followed by CaO , Cr_2O_3 , MgO , and MnO (Fig. 6, green line). The exclusion of these oxides mainly leads to misclassification of crustal M garnets as IF/S (52 % of newly misclassified garnets for TiO_2 , 77 % for CaO , 24 % for Cr_2O_3 , 88 % for MgO , and 71 % for MnO), and of mantle M garnets as UM (45 % for TiO_2 , 23 % for CaO , 76 % for Cr_2O_3 , 12 % for MgO , and 25 % for MnO).

The highest influence on misclassification rates of class IF/S is observed by excluding MgO , followed by TiO_2 , MnO , Al_2O_3 , and CaO (Fig. 6, yellow line). By excluding these oxides, newly misclassified IF/S garnets belong to igneous and metamorphic host-rocks of all grades, except granulite facies. The vast majority of those garnets are now assigned to class M (93 % for MgO , 97 % for TiO_2 , 91 % for MnO , 89 % for Al_2O_3 , and 92 % for CaO).

For class A, misclassification rates strongly increase by excluding TiO_2 , followed by $\text{FeO}_{\text{total}}$, CaO , and Al_2O_3 , all leading to misclassifications as CS. Remarkably, the exclusion of MgO results in lower misclassifications, that are caused by A garnet previously assigned to CS. Conversely, the oxides CaO , MnO , and MgO have the highest influence on misclassification rates of CS garnet. The exclusion of these oxides results in both, the assignment of MS-CS garnet as A (17 % for CaO , 57 %

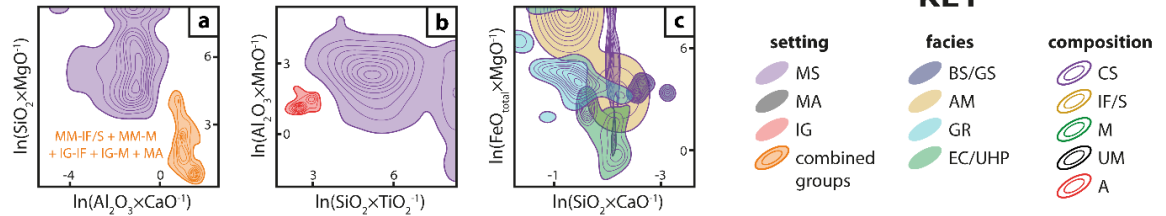
for MnO, and 33 % for FeO_{total}) and metamorphic CS garnets to all other composition classes (83 % for CaO, 43 % for MnO, and 67 % for FeO_{total}).

Extended information to *Origin of main discriminators*

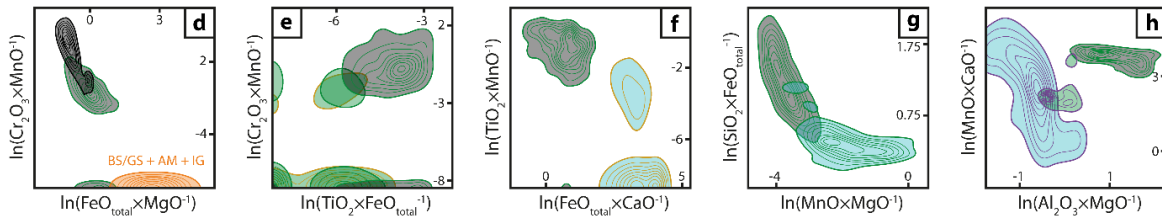
Detailed exploration of main discriminators for individual subgroups

Garnet of class MS in the ‘setting and metamorphic facies’ model shows the highest classification success (Table 3), reflecting its distinct chemical composition. MS garnet is very high in CaO and very low in MgO. In addition, garnet of class MS shows a broad range for $\ln(\text{SiO}_2 \times \text{Al}_2\text{O}_3^{-1})$ and higher values compared to all other classes. Consequently, a scatter plot of $\ln(\text{Al}_2\text{O}_3 \times \text{CaO}^{-1})$ versus $\ln(\text{SiO}_2 \times \text{MgO}^{-1})$ well discriminates MS garnet from most other classes (Fig. S-7a). For the remaining subgroups, IG-A can be discriminated based on the high TiO_2 content (Fig. S-7b), while the distinction from the various metamorphic facies subgroups of CS composition is more difficult. However, the tight range of $\ln(\text{SiO}_2 \times \text{CaO}^{-1})$ enables the separation to most CS garnets from classes

MS versus MA, IG, and MM



MA versus IG and MM



IG versus MM

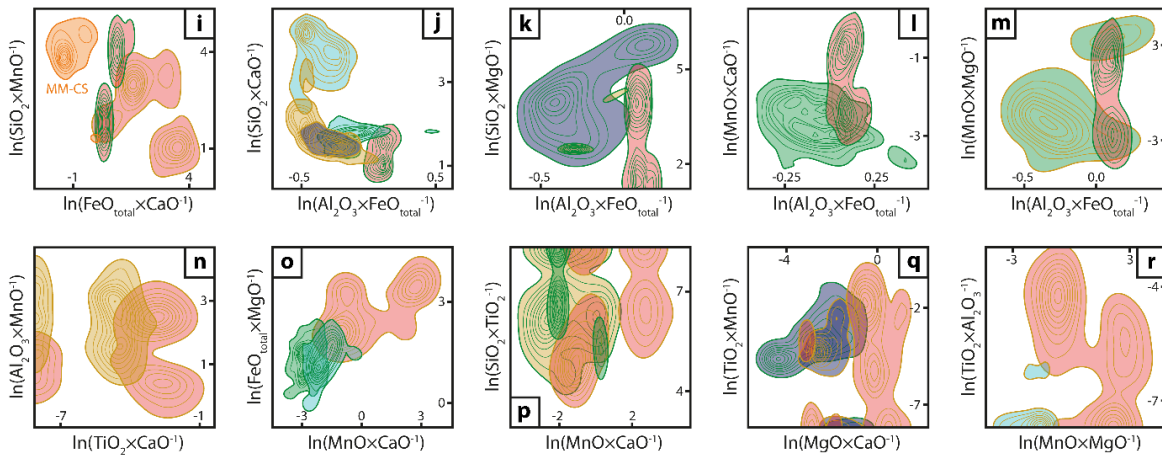


Figure S-7: Main discriminators to separate individual classes, groups, and subgroups of the ‘setting’ scheme. (a–c) separation of MS garnet from MA, IG, and MM garnet. (d–h) separation of MA garnet from IG and MM. (i–r) separation of IG garnet from MM. Note that individual fields are marked by the combination of colour fill (indicating setting/facies) and coloured contours (indicating composition).

BS/GS, AM, and GR. Considering additionally the low values for $\ln(\text{FeO}_{\text{total}} \times \text{MgO}^{-1})$ is useful to discriminate EC/UHP-CS garnet (Fig. S-7c).

The distinct composition of MS garnet is strongly related to the formation environment. Although skarn garnet can form in a wide range of settings and protolith compositions, by far the most are associated with igneous activity that leads to contact metamorphism of carbonates by heat supply and infiltrating metasomatic fluids at depth <12 km (e.g., Meinert 1992). Garnet mainly forms at the prograde anhydrous stage together with clinopyroxene, both being high in Ca^{2+} due to the availability given by protolith composition. Furthermore, high oxygen fugacity enables the stabilization of andradite (end member composition: $\text{Ca}_3\text{Fe}_2\text{Si}_3\text{O}_{12}$), that is the replacement of Al^{3+} by Fe^{3+} compared to grossular ($\text{Ca}_3\text{Al}_2\text{Si}_3\text{O}_{12}$), resulting in a garnet solid solution rich in grossular-andradite (e.g., Zhang and Saxena 1991). In contrast, clinopyroxenes crystallize mainly as solid solutions between diopside ($\text{CaMgSi}_2\text{O}_6$) and hedenbergite ($\text{CaFeSi}_2\text{O}_6$), and thus much of the available Mg^{2+} and Fe^{2+} are incorporated in clinopyroxene (Bin and Barton, 1988). An exceptional case are skarns with low oxygen fugacity like those forming tungsten and tin deposits, where some garnet populations are rich in Al^{3+} (e.g., Zhang and Saxena, 1991; Meinert, 1992). However, considering recently published compositions of garnets from these reduced environments (Duan et al. 2020; Im et al. 2020), >87 % of the garnets are correctly classified as MS, and the remaining are mainly classified as GR garnet of CS composition.

Garnet of class MA shows the second-highest classification success in the ‘setting and metamorphic facies’ model (Table 3). Due to the relatively high content of MgO and low content of $\text{FeO}_{\text{total}}$ in mantle rocks compared to crustal rocks, the $\ln(\text{FeO}_{\text{total}} \times \text{MgO}^{-1})$ represent the main discriminator. In combination with $\ln(\text{Cr}_2\text{O}_3 \times \text{MnO}^{-1})$ this enables almost perfect discrimination of garnet of subgroup MA-UM from those of crustal affinity (Fig. S-7d). The enrichment of Cr in MA garnet is mainly related to the lithophile behaviour of Cr, resulting in the accumulation of Cr in mantle mineral phases like chromium-spinel during partial melting of the upper mantle (e.g., Matrosova et al. 2020), while spinel is replaced by garnet at greater depths (e.g., Klemme et al. 2009). This also holds for garnet of subgroup MA-M versus most crustal garnets, except some overlap with garnet of subgroups EC/UHP-M, EC/UHP-CS, GR-M, and GR-IF/S. In addition, caution should be taken for some rare UHP-IF/S garnets that can have extremely high MgO contents (Chopin 1984) leading to misclassification as MA garnet.

The additional consideration of $\ln(\text{TiO}_2 \times \text{FeO}_{\text{total}}^{-1})$ helps in discriminating EC/UHP garnet (Fig. S-7e) and the consideration of $\ln(\text{FeO}_{\text{total}} \times \text{CaO}^{-1})$ combined with $\ln(\text{TiO}_2 \times \text{MnO}^{-1})$ is very useful to discriminate GR-IF/S garnet from MA-M garnet (Fig. S-7f). The importance of TiO_2 has also been observed by Hardman et al. (2018), but TiO_2 is not considered in most previous garnet provenance discrimination schemes. The higher TiO_2 content in MA-M compared to EC/UHP and GR-IF/S is

most likely related to higher temperatures experienced by MA-M, resulting in increasing solubility of Ti in garnet (Aulbach 2020). This can explain the increasing misclassification rates for MA garnet when TiO_2 is excluded (Fig. 4). Furthermore, variables $\ln(\text{MnO} \times \text{MgO}^{-1})$ versus $\ln(\text{SiO}_2 \times \text{FeO}_{\text{total}}^{-1})$ enable better discrimination of GR-M and MA-M garnet (Fig. S-2g), and $\ln(\text{Al}_2\text{O}_3 \times \text{MgO}^{-1})$ versus $\ln(\text{MnO} \times \text{CaO}^{-1})$ enable better discrimination of GR-CS and EC/UHP-CS from MA-M (Fig. S7-h).

Compared to MS and MA garnet, the discrimination of IG versus MM garnet is more challenging. Very distinct are only IG garnets of alkaline composition (IG-A), which can be separated by considering the low values for $\ln(\text{Al}_2\text{O}_3 \times \text{CaO}^{-1})$, $\ln(\text{SiO}_2 \times \text{MgO}^{-1})$, and $\ln(\text{SiO}_2 \times \text{TiO}_2^{-1})$ compared to MM garnet, given by the high andradite (Fe^{3+}) component resulting in lower amounts of Al_2O_3 along with high TiO_2 (Figs. S-7a and S-7b). This agrees with the increasing incorporation of Ti into garnet with decreasing silica activity and oxygen fugacity (e.g., Russell et al. 1999). Most important for discriminating IG garnet from MM garnet are $\ln(\text{MnO} \times \text{CaO}^{-1})$ and $\ln(\text{SiO}_2 \times \text{MgO}^{-1})$, showing that IG garnet is enriched in MnO but depleted in CaO and MgO compared to MM garnet of similar protolith composition (Fig. 7c). Probably the typically higher temperature–pressure gradient in igneous systems is the main driver. However, several subpopulations of IG garnet reflect a high chemical variability, reinforcing the need of considering multiple variables for a sufficient distinction to MM garnet.

Due to the distinct protolith composition, both IG-M and IG-IF garnet are decently separated from metamorphic garnet of calcsilicate composition (MM-CS) by $\ln(\text{FeO}_{\text{total}} \times \text{CaO}^{-1})$ and $\ln(\text{SiO}_2 \times \text{MnO}^{-1})$ (Fig. S-7i). To discriminate IG-M garnet from MM garnet of M and IF/S composition, $\ln(\text{Al}_2\text{O}_3 \times \text{FeO}_{\text{total}}^{-1})$ represents the most important variable (Figs. S-7j, S-7k, S-7l, and S-7m). This may be related to the high crystallization temperatures of IG-M garnet that leads to the relatively preferred incorporation of Fe into minerals co-existing with garnet like pyroxenes and biotite (e.g., Ferry and Spear 1978; Ganguly 1979; Dahl 1980). Combining $\ln(\text{Al}_2\text{O}_3 \times \text{FeO}_{\text{total}}^{-1})$ with $\ln(\text{SiO}_2 \times \text{CaO}^{-1})$ and $\ln(\text{SiO}_2 \times \text{MgO}^{-1})$ enables the discrimination of IG-M from most MM garnets (Figs. S-7j and S-7k), except those from EC/UHP rocks. The difficulties are mainly related to the higher CaO content of mafic host-rock compositions and the preferred partitioning of Mg into chlorite and chloritoid compared to garnet grown at low temperatures (e.g., Perchuk 1991). To discriminate IG-M from EC/UHP garnet, it is more useful to combine $\ln(\text{Al}_2\text{O}_3 \times \text{FeO}_{\text{total}}^{-1})$ with $\ln(\text{MnO} \times \text{CaO}^{-1})$ or $\ln(\text{MnO} \times \text{MgO}^{-1})$ but some overlaps remain (Figs. S-7l and S-7m), making their separation less pronounced (cf. Fig. 1d).

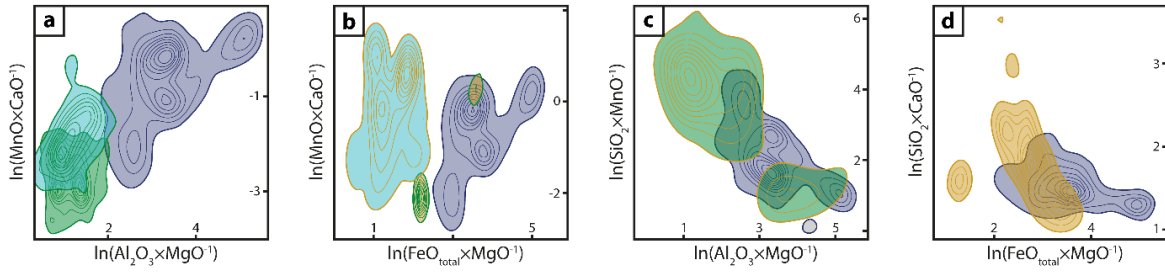
For the discrimination of IG-IF and MM garnet, $\ln$ -ratios including CaO, MnO, MgO, and TiO_2 are most important (Figs. S-7n, S-7o, S-7p, S-7q, and S-7r). CaO is particularly high in mafic MM garnet due to protolith composition, but it is also higher in the other MM classes. This relates to the higher-pressure conditions of MM garnet formation and Ca-rich garnet in igneous rocks only occurs when deeply emplaced (e.g., Anderson 2008). In contrast, MnO is often higher in IG-IF garnet, which

reflects the abundant crystallization from highly fractionated Al- and Mn-rich magmas (e.g., Dahlquist, 2007), enabling garnet growth at pressures as low as 3 kbar (e.g., Green 1977). In addition, at crystallization temperatures of IF melts, Mg is partitioning into the melt compared to garnet, resulting in low-Mg garnet (Green 1977). Most difficult is the separation of IG-IF garnet from AM garnet, reflecting the often more similar temperature and pressure conditions during garnet growth. The consideration of TiO_2 significantly enhances this discrimination, reflecting a higher average content of TiO_2 in IG-IF compared to AM (Fig. S-7n). Understanding the partitioning of Ti in garnet of igneous versus metamorphic systems is not straight forward, but the observed preferred incorporation of Ti into igneous biotite may contribute to this trend (Samadi et al. 2021).

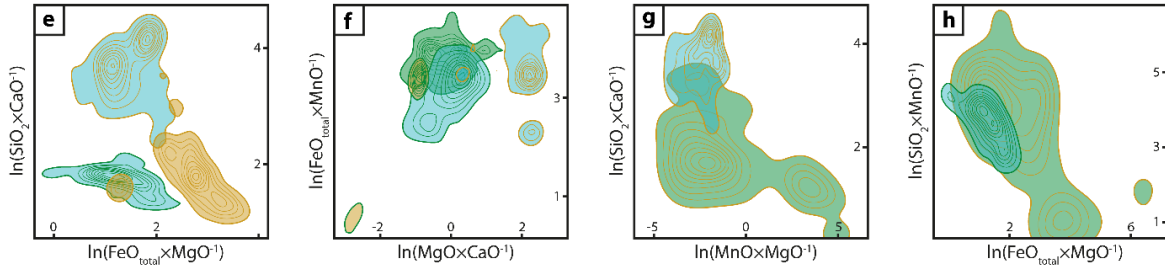
The most important variable combinations to discriminate different metamorphic classes and subgroups are shown in Figure S-8. The absence of ln-ratios that include TiO_2 shows its much lower importance for discrimination within the metamorphic classes (cf. Fig. 4). For class BS/GS, differences between IF/S and M composition are not highly pronounced and thus considered together. The most important variables to separate BS/GS from other metamorphic classes include MgO in the denominator with Al_2O_3 or $\text{FeO}_{\text{total}}$ in the nominator, whereby BS/GS garnet shows the highest values due to the low MgO content (Figs. S-8a, S-8b, S-8c, and S-8d). This agrees with the low-temperature conditions during garnet growth and the many exchange thermometers that imply increasing $\text{Fe} \times \text{Mg}^{-1}$ with decreasing temperature (Reverdatto et al., 2019, and references therein). In addition, high contents of MnO in BS/GS garnet improve the discrimination, in particular for AM-M, GR-M, EC-M, and EC-IF/S (Figs. S-8a, S-8b, and S-8c). This is consistent with the increasing stability field of garnet to lower pressure–temperature conditions with increasing MnO content of the protolith. Thus, earliest grown garnet typically shows highest MnO contents (e.g., Carlson 1989) and removes MnO from the effective bulk composition leading to an up-temperature shift of garnet stability and bell-shaped Mn-zoning patterns (Evans 2004). For BS/GS garnet versus other MM garnet of M composition, $\ln(\text{MnO} \times \text{CaO}^{-1})$ further increases the distance given by the higher CaO content of mafic protoliths (Figs. S-8a and S-8b). Remarkable overlaps remain for BS/GS and AM-IF/S as well as EC-IF/S (Figs. S-8c and S-8d).

In class GR, most distinct is the low CaO content of GR-IF/S garnet that enables almost perfect discrimination to AM-IF/S (Fig. S-8e), AM-M and EC-M (Fig. S-8f), as well as good discrimination to EC-IF/S (Fig. S-8g). This highlights the fundamentals of the garnet–aluminosilicate–plagioclase–quartz geothermobarometer based on the higher stability of anorthite at high-temperature/low-pressure compared to the higher stability of grossular + aluminosilicate + quartz at low-temperature/high-pressure (e.g., Ghent 1976; Koziol and Newton 1988). GR-M garnet can be discriminated from AM garnet by the higher MgO contents compared to $\text{FeO}_{\text{total}}$ or CaO, whereby the separation is higher for AM-IF/S (Fig. S-8e) than AM-M (Fig. S-8f). More difficult is the

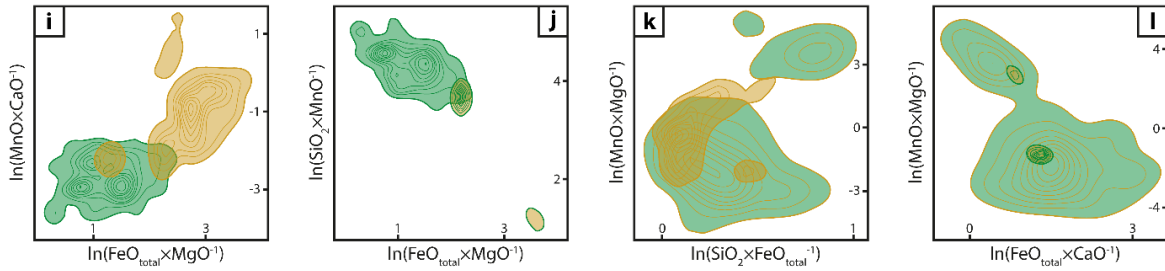
BS/GS versus AM, GR and EC/UHP



GR versus AM and EC/UHP



AM versus EC/UHP



KEY BS/GS AM GR EC/UHP IF/S M IF/S + M

Figure S-8: Main discriminators to separate individual classes, groups, and subgroups of the ‘metamorphic facies’ scheme. (a–d) separation of BS/GS garnet from AM, GR, and EC/UHP garnet. (e–h) separation of GR garnet from AM and EC/UHP. (i–l) separation AM garnet from EC/UHP. Note that individual fields are marked by the combination of colour fill (indicating setting/facies) and coloured contours (indicating composition).

discrimination of GR-M garnet from EC/UHP garnet (Figs. S-8g and S-8h), leading to higher misclassification rates (cf. Fig. 1d). Similarly, EC/UHP-M garnet can be separated from AM garnet by the lower $\ln(\text{FeO}_{\text{total}} \times \text{MgO}^{-1})$ and consideration of MnO and CaO (Figs. S-8i and S-8j), but the discrimination of EC-IF/S from AM garnet is difficult, at least by looking at two variables only (Figs. S-8k and S-8l).

Concerning composition, class CS shows the highest discrimination success, followed by A (Table 4). CS and A garnet are well separated from all other classes by their higher CaO content (Fig. S-9a), reflecting the high host-rock CaO content. The high TiO_2 content of A garnet allows the discrimination to CS (Fig. S-7b). Furthermore, UM garnet separates from IF/S and M garnet by the higher MgO and Cr_2O_3 content in combination with the lower $\text{FeO}_{\text{total}}$ content (Figs. S-9b and S-9c), in line with the element availability defined by the host-rock composition as well as the high-

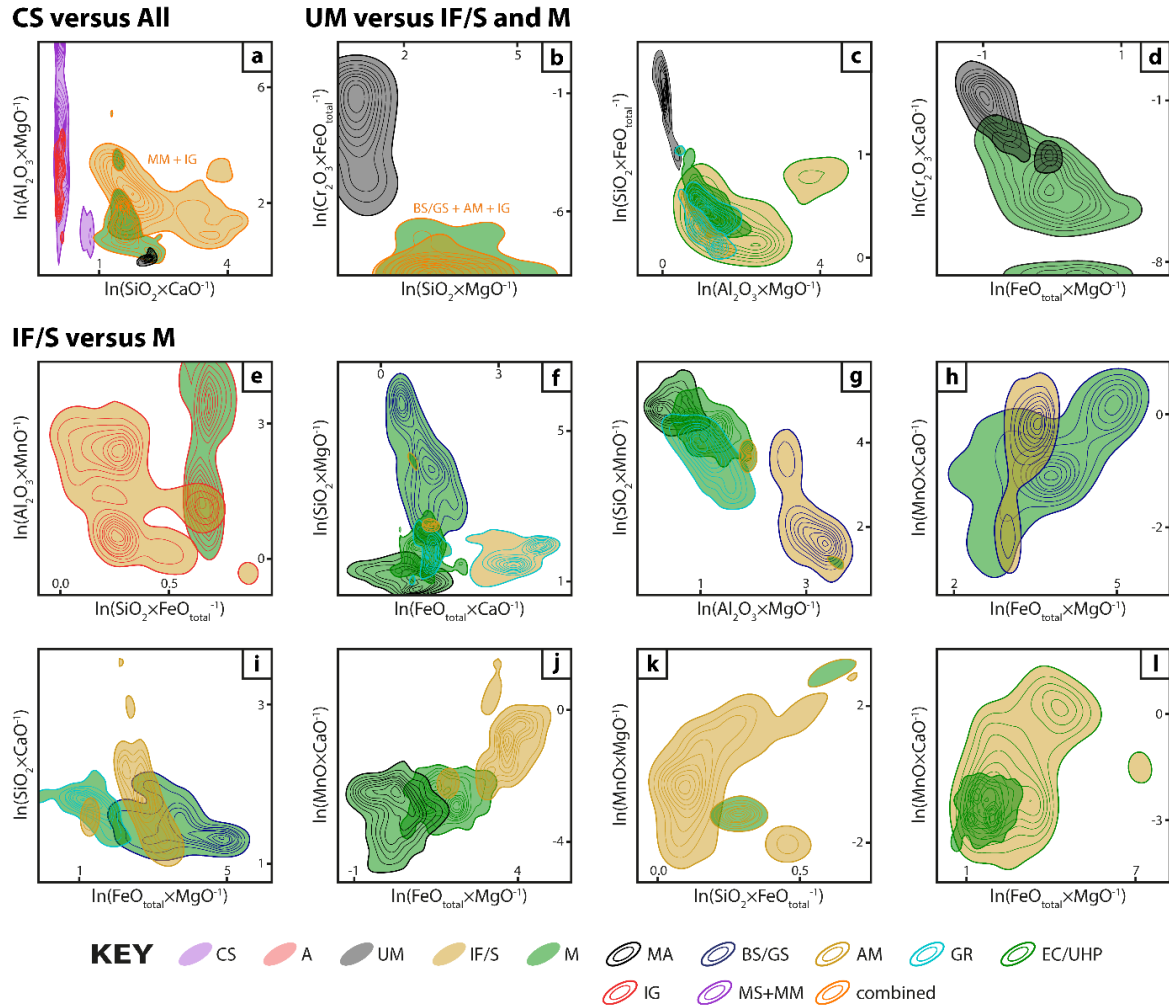


Figure S-9: Main discriminators to separate individual classes, groups, and subgroups of the 'composition' scheme. (a) separation of CS garnet from UM, IF/S, and M. (b–d) separation of UM garnet from IF/S and M. (e–l) separation IF/S garnet from M. Note that individual fields are marked by the combination of colour fill (indicating composition) and coloured contours (indicating setting/facies).

temperature formation conditions of mantle rocks. The similar formation conditions of MA-UM and MA-M make their discrimination most difficult, but particularly the higher CaO content of MA-M (basaltic) enables reasonable separation, although overlaps remain (Fig. S-9d).

For the discrimination of IF/S and M garnet, $\ln(\text{FeO}_{\text{total}} \times \text{MgO}^{-1})$ and $\ln(\text{MnO} \times \text{CaO}^{-1})$ are the most important variables, both being higher for IF/S compared to M garnet (Fig. 8d). Although these ratios represent the superordinate trend, they insufficiently split subpopulations, calling for a case-wise consideration. To separate IG-IF from MM-M garnet, Figures S-7o, S-7p, and S-7q show that the higher MnO and $\text{FeO}_{\text{total}}$ content as well as the lower CaO and MgO content are most important, agreeing with the superordinate trend. For the discrimination of IG-IF from IG-M, it is more useful to solely consider the higher $\text{FeO}_{\text{total}}$ content of IG-IF compared to IG-M (Fig. S-9e), in line with the lower crystallization temperatures.

The discrimination of metamorphic IF/S from M garnet is most distinct for GR-IF/S, which is well separated from all M subgroups by $\ln(\text{FeO}_{\text{total}} \times \text{CaO}^{-1})$ (Fig. S-9f). Although this is mainly a temperature effect as discussed before, it also enables the discrimination of GR-IF/S and GR-M, and thus reflects the higher CaO content of the mafic protolith. BS/GS-IF/S can be well separated from AM-M, GR-M, EC/UHP-M, and MA-M by the higher MnO and lower MgO content (Fig. S-9g), underlining the necessity of a higher MnO content of the protolith to stabilize garnet at low temperatures. As this is similar for IF/S and M protoliths, BS/GS-IF/S is difficult to be distinguished from BS/GS-M, where the superordinate trend of $\ln(\text{FeO}_{\text{total}} \times \text{MgO}^{-1})$ combined with $\ln(\text{MnO} \times \text{CaO}^{-1})$ is most important (Fig. S-9h). Variable $\ln(\text{FeO}_{\text{total}} \times \text{MgO}^{-1})$ also best separates AM-IF/S from M subgroups, whereby GR-M, EC/UHP-M and MA-M show higher values due to average higher temperatures, while BS-M shows lower values due to lower temperatures (Figs. S-9i and S-9j). The protolith effect is mainly given by higher CaO of M garnet, whereby the separation to EC/UHP-M and MA-M garnet is getting enhanced by their lower MnO content (Fig. S-9j). Best discrimination of AM-IF/S from AM-M garnet is given by the often lower $\text{FeO}_{\text{total}}$ and higher MnO of IF/S garnet (Fig. S-9k). Significantly higher overlaps are given for separating EC/UHP-IF/S from M garnet. Lower TiO_2 and Cr_2O_3 are most important to separate EC/UHP-IF/S from MA-M (Fig. S-7e). Higher MgO and lower MnO are most important to separate EC/UHP-IF/S from BS-M (Fig. S-8c) and GR-M (Fig. S-8h). A broader range in $\ln(\text{FeO}_{\text{total}} \times \text{CaO}^{-1})$ and $\ln(\text{MnO} \times \text{MgO}^{-1})$ is most important to separate EC/UHP-IF/S from AM-M (Fig. S-8l), and a higher $\ln(\text{MnO} \times \text{CaO}^{-1})$ value separates a subpopulation from EC/UHP-M garnet (Fig. S-9l).

References

- Anderson JL, Barth AP, Wooden JL, Mazdab F (2008) Thermometers and thermobarometers in granitic systems. *Reviews in Mineralogy and Geochemistry* 69:121–142. <https://doi.org/10.2138/rmg.2008.69.4>
- Aulbach S (2020) Temperature-dependent rutile solubility in garnet and clinopyroxene from mantle eclogite: Implications for continental crust formation and V-based oxybarometry. *Journal of Petrology* 61:egaa065. <https://doi.org/10.1093/petrology/egaa065>
- Belgiu M, Drăguț L (2016) Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing* 114:24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Biau G, Scornet E (2016) A random forest guided tour. *TEST* 25:197–227. <https://doi.org/10.1007/s11749-016-0481-7>
- Boulesteix AL, Janitza S, Kruppa J, König IR (2012) Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 2:493–507. <https://doi.org/10.1002/widm.1072>
- Breiman L, Friedman J, Stone CJ, Olshen RA (1984) *Classification and regression trees*. Chapman and Hall/CRC
- Breiman L (2001) Random forests. *Machine learning* 45:5–32. <https://doi.org/10.1023/A:1010933404324>
- Carlson WD (1989) The significance of intergranular diffusion to the mechanisms and kinetics of porphyroblast crystallization. *Contributions to Mineralogy and Petrology* 103:1–24. <https://doi.org/10.1007/BF00371361>
- Chen C, Liaw A, Breiman L (2004) Using random forest to learn imbalanced data. Technical report, # 666, Department of Statistics, University of California, Berkeley, CA. <http://www.stat.berkeley.edu/tech-reports/666.pdf>. Accessed 29th May 2021
- Chopin C (1984) Coesite and pure pyrope in high-grade blueschists of the Western Alps: a first record and some consequences. *Contributions to Mineralogy and Petrology* 86:107–118. <https://doi.org/10.1007/BF00381838>
- Dahl PS (1980) The thermal–compositional dependence of Fe²⁺–Mg distributions between coexisting garnet and pyroxene: applications to geothermometry. *American Mineralogist* 65:852–866
- Dahlquist JA, Galindo C, Pankhurst RJ, Rapela CW, Alasino PH, Saavedra J, Fanning CM (2007) Magmatic evolution of the Peñón Rosado granite: petrogenesis of garnet-bearing granitoids. *Lithos* 95:177–207. <https://doi.org/10.1016/j.lithos.2006.07.010>
- Duan XX, Ju YF, Chen B, Wang ZQ (2020) Garnet Geochemistry of Reduced Skarn System: Implications for Fluid Evolution and Skarn Formation of the Zhuxiling W (Mo) Deposit, China. *Minerals* 10:1024. <https://doi.org/10.3390/min10111024>

- Evans TP (2004) A method for calculating effective bulk composition modification due to crystal fractionation in garnet-bearing schist: implications for isopleth thermobarometry. *Journal of Metamorphic Geology* 22:547–557. <https://doi.org/10.1111/j.1525-1314.2004.00532.x>
- Ferry JT, Spear FS (1978) Experimental calibration of the partitioning of Fe and Mg between biotite and garnet. *Contributions to Mineralogy and Petrology* 66:113–117
- Ganguly J (1979) Garnet and clinopyroxene solid solutions, and geothermometry based on Fe-Mg distribution coefficient. *Geochimica et Cosmochimica Acta* 43:1021–1029. [https://doi.org/10.1016/0016-7037\(79\)90091-7](https://doi.org/10.1016/0016-7037(79)90091-7)
- Ghent ED (1976) Plagioclase–garnet–Al₂SiO₅–quartz: a potential geobarometer–geothermometer. *American Mineralogist* 61:710–714
- Green TH (1977) Garnet in silicic liquids and its possible use as a PT indicator. *Contributions to Mineralogy and Petrology* 65:59–67
- Grütter HS, Gurney JJ, Menzies AH, Winter F (2004) An updated classification scheme for mantle-derived garnet, for use by diamond explorers. *Lithos* 77:841–857. <https://doi.org/10.1016/j.lithos.2004.04.012>
- Hardman MF, Pearson DG, Stachel T, Sweeney RJ (2018) Statistical approaches to the discrimination of crust- and mantle-derived low-Cr garnet–major-element-based methods and their application in diamond exploration. *Journal of Geochemical Exploration* 186:24–35. <https://doi.org/10.1016/j.gexplo.2017.11.012>
- Im H, Jeong JY, Shin D, (2020) Genetic environment of W skarn and Pb-Zn vein mineralization associated with the Imog granite in the Taebaeksan Mineralized District, South Korea. *Ore Geology Reviews* 126:103721. <https://doi.org/10.1016/j.oregeorev.2020.103721>
- Klemme S, Ivanic TJ, Connolly JAD, Harte B (2009) Thermodynamic modelling of Cr-bearing garnets with implications for diamond inclusions and peridotite xenoliths. *Lithos* 112:986–991. <https://doi.org/10.1016/j.lithos.2009.05.007>
- Koziol AM, Newton RC (1988) Redetermination of the anorthite breakdown reaction and improvement of the plagioclase–garnet–Al₂SiO₅–quartz geobarometer. *American Mineralogist* 73:216–223
- Liaw A, Wiener M (2002) Classification and Regression by randomForest. *R News* 2:18–22
- Mange MA, Morton AC (2007) Geochemistry of heavy minerals. *Developments in Sedimentology* 58:345–391. [https://doi.org/10.1016/S0070-4571\(07\)58013-1](https://doi.org/10.1016/S0070-4571(07)58013-1)
- Matrosova EA, Bobrov AV, Bindi L (2020) Geochemistry of Chromium in the Earth's Mantle. Springer Nature Switzerland, Cham. <https://doi.org/10.1007/978-3-030-27018-6>
- Meinert LD (1992) Skarns and Skarn Deposits. *Geoscience Canada* 19:145–162. <https://journals.lib.unb.ca/index.php/GC/article/view/3773>

- Morton AC (1985) A new approach to provenance studies: electron microprobe analysis of detrital garnets from Middle Jurassic sandstones of the northern North Sea. *Sedimentology* 32:553–566.
<https://doi.org/10.1111/j.1365-3091.1985.tb00470.x>
- Morton AC, Hallsworth C, Chalton B (2004) Garnet compositions in Scottish and Norwegian basement terrains: a framework for interpretation of North Sea sandstone provenance. *Marine and Petroleum Geology* 21:393–410. <https://doi.org/10.1016/j.marpetgeo.2004.01.001>
- Morton AC, Whitham AG, Fanning CM (2005a) Provenance of Late Cretaceous to Paleocene submarine fan sandstones in the Norwegian Sea: integration of heavy mineral, mineral chemical and zircon age data. *Sedimentary Geology* 182:3–28. <https://doi.org/10.1016/j.sedgeo.2005.08.007>
- Morton AC, Whitham AG, Fanning CM, Claoué-Long J (2005b) The role of East Greenland as a source of sediment to the Vøring Basin during the Late Cretaceous. In: Wandås BTG, Nystuen JP, Eide E, Gradstein F (eds) *Onshore-Offshore Relationships on the North Atlantic Margin*. Norwegian Petroleum Society Special Publications 12:83–110. [https://doi.org/10.1016/S0928-8937\(05\)80045-1](https://doi.org/10.1016/S0928-8937(05)80045-1)
- Nowicki TE, Moore RO, Gurney JJ, Baumgartner MC (2007) Diamonds and associated heavy minerals in kimberlite: a review of key concepts and applications. *Developments in Sedimentology* 58:1235–1267.
[https://doi.org/10.1016/S0070-4571\(07\)58046-5](https://doi.org/10.1016/S0070-4571(07)58046-5)
- Perchuk LL (1991) Derivation of a thermodynamically consistent set of geothermometers and geobarometers for metamorphic and magmatic rocks. In: Perchuk LL (ed) *Progress in metamorphic and magmatic petrology*. Cambridge University Press, Cambridge, pp. 93–112.
<https://doi.org/10.1017/CBO9780511564444>
- Reverdatto VV, Likhonov II, Polyansky OP, Sheplev VS, Kolobov VY (2019) Mineral Geothermobarometry. In: Reverdatto VV, Likhonov II, Polyansky OP, Sheplev VS, Kolobov VY (eds) *The Nature and Models of Metamorphism*. Springer Nature Switzerland, Cham, pp. 55–82. https://doi.org/10.1007/978-3-030-03029-2_2
- Russell JK, Dipple GM, Lang JR, Lueck B (1999) Major-element discrimination of titanium andradite from magmatic and hydrothermal environments; an example from the Canadian Cordillera. *European Journal of Mineralogy* 11:919–935
- Samadi R, Torabi G, Kawabata H, Miller NR (2021) Biotite as a petrogenetic discriminator: Chemical insights from igneous, meta-igneous and meta-sedimentary rocks in Iran. *Lithos* 386:106016.
<https://doi.org/10.1016/j.lithos.2021.106016>
- Schulze DJ (2003) A classification scheme for mantle-derived garnets in kimberlite: a tool for investigating the mantle and exploring for diamonds. *Lithos* 71:195–213. [https://doi.org/10.1016/S0024-4937\(03\)00113-0](https://doi.org/10.1016/S0024-4937(03)00113-0)
- Tolosana-Delgado R, von Eynatten H, Krippner A, Meinhold G (2018) A multivariate discrimination scheme of detrital garnet chemistry for use in sedimentary provenance analysis. *Sedimentary Geology* 375:14–26.
<https://doi.org/10.1016/j.sedgeo.2017.11.003>

- Venables WN, Ripley BD (2002) Modern Applied Statistics with S. Springer-Verlag, New York.
<https://doi.org/10.1007/978-0-387-21706-2>
- Zhang Z, Saxena SK (1991) Thermodynamic properties of andradite and application to skarn with coexisting andradite and hedenbergite. *Contributions to Mineralogy and Petrology* 107:255–263
- Ziegler A, König IR (2014) Mining data with random forests: current options for real-world applications. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 4:55–63.
<https://doi.org/10.1002/widm.1114>