Multivariate geochemical classification of chromitite seams in the Bushveld Complex, South Africa

Bachmann, K.; Menzel, P.; Tolosana-Delgado, R.; Schmidt, C.; Hill, M.; Gutzmer, J.;

Originally published:

February 2019

Applied Geochemistry 130(2019), 106-117

DOI: https://doi.org/10.1016/j.apgeochem.2019.02.009

Perma-Link to Publication Repository of HZDR:

https://www.hzdr.de/publications/Publ-28821

Multivariate geochemical classification of chromitite seams in the Bushveld Complex, South Africa

Kai Bachmann[1,2]*, Peter Menzel[1], Raimon Tolosana-Delgado[1], Christopher Schmidt[3], Moritz Hill[3], Jens Gutzmer[1]

[1]Helmholtz-Zentrum Dresden-Rossendorf, Helmholtz Institute Freiberg for Resource Technology, Chemnitzer Str. 40, D-09599 Freiberg/Sachsen, Germany

[2]Department of Mineralogy, TU Bergakademie Freiberg, Brennhausgasse 14, D-09596 Freiberg/Sachsen, Germany

[3]Cronimet Mining Group, Building C Willow Wood Office Park, Cnr 3rd & Cedar Roads, Broadacres 2021, Johannesburg, South Africa

*Tel: +49 351 260 4426

k.bachmann@hzdr.de

Abstract

The Bushveld Complex, the largest layered mafic-ultramafic intrusion worldwide, is host of numerous, laterally continuous and chemically similar chromitite seams. Based on their stratigraphic position the seams are subdivided into a lower, middle and upper group (LG, MG and UG). Within these groups the seams are numbered successively – from the base to the top of each group. Attempts of discriminating between single seams based on their composition have failed – mainly due to the significant overlap of compositional fields, e.g. of chromitite mineral assemblages and chromite mineral chemistry between (neighboured) seams. In this contribution a tailored and easy to use multivariate classification scheme for the chromitite seams is proposed, based on a comprehensive classification routine for the LG and MG chromitites. This routine allows a clear attribution with known uncertainty of eight distinct chromitite seams. The study was carried out at the Thaba Mine, a chromite mine located on the western limb of the Bushveld Complex. The classification is based on a large geochemical database (N = 1205) from Thaba Mine. It comprises of a hierarchical discrimination approach relying on linear discriminant analysis and involves five distinct steps. Using default homogeneous prior probabilities, classification results are excellent for the first discrimination steps (LGs vs. MGs, 97 %; LG-6 vs. LG-6A, 94 %) and very good for the following steps (MG-1/2 vs. MG-3/4, 86 %; MG-1 vs. MG-2, 92 %; MG-3 vs. MG-4, 93 %; MG-4 vs. MG-4Z, 97 %; MG-4 vs. MG-4A, 88 %). The classification scheme was tested using the same sample set as a training set with unknown composition. Overall classification results for unknown samples belonging to one of the seams are 81 %. Hence, the classification scheme is at least valid for the Thaba mine. The approach may, however, be extended across the entire Bushveld, provided that an appropriate geochemical data set is available.

Keywords

Linear discriminant analysis, PGE, Thaba Mine

Introduction

The Bushveld Complex, the world's largest layered mafic-ultramafic intrusion, accounts currently for 45 % of global annual chromite ore production and contains 40 % of global chromite reserves (USGS 2018). Besides vast resources of chromite, the Bushveld Complex also contains the largest global resource of platinum-group elements (PGE) as well as significant resources of Ni and Cu (e.g., von Gruenewaldt 1977; Maier et al. 2013). The chromite resource of the Bushveld Complex is hosted by a series of laterally continuous and chemically similar chromitite seams (e.g., Fourie 1959; Naldrett et al. 2012) in the Critical Zone of the ultramafic and mafic rocks of the Rustenburg Layered Suite (RLS) (Maier et al. 2013 and references therein, Cawthorn 2015 and references therein). The chromitite seams are numbered according to their stratigraphic position within the Critical Zone (CZ) from the base upwards and are generally subdivided into lower, middle and upper group (LG, MG and UG) chromitites. In this way, the Lower Group ranges from the LG-1 to LG-7; the Middle Group comprises of MG-1, MG2a,b,c, MG-3, MG-3a, MG-4, MG-4a and the Upper Group consists of the UG-1, UG-2 and, in some areas, the UG-3 and UG-3a seams (Schürmann et al. 1998). In addition, the CZ can be further subdivided into the Lower Critical Zone (LCZ) and the Upper Critical Zone (UCZ), which are separated by the first appearance of cumulus plagioclase between the middle group chromitites MG-2 and MG-3 (Veksler et al. 2015 and references therein).

The present contribution provides a novel multivariate geochemical classification scheme to correctly identify chromitite seams of the Bushveld Complex. This approach is introduced and tested on a test case that includes eight distinct chromitite seams at the Thaba Chromite Mine. It uses a hierarchical approach relying on linear discriminant analysis and weighting prior probabilities.

This approach augments the current practice in exploration or exploitation of chromite layers (the terms chromitite "layer" and "seam" are used interchangeably in this study) identifcation by their stratigraphic position, the nature of interlayer silicates and partly on the thickness of the chromitite (Kinnaird et al. 2002).

Despite its apparent simplicity, the subdivision of the seams into different groups obscures a somewhat greater complexity. For example, the MG-1 may locally comprise of multiple layers of chromitite, as at Tweefontein mine (Kinnaird et al., 2002), and in the eastern Bushveld, the MG-1 may be underlain by a chromitite layer referred to as MG-0 (Kinnaird et al., 2002). Furthermore, possible bifurcations of individual chromitite layers, such as the UG-1 in the Dwars River area (e.g., Voordouw et al. 2009), leads to variability in the number of chromitite seams observed at any given locality. Tracing seams over large distances within the Bushveld can also be challenging. Although single seams can be followed for tens of km in the both eastern and western lobes of the Bushveld Complex (Cawthorn and Webb (2001) implied continuity of individual layers over more than 300 km), in some areas, chromitite horizons are missing or correlation between two chromitite seams may be tenuous or even impossible, depending largely on the outcrop situation (Kinnaird et al. 2002). Furthermore, it is well-known that not all chromitite seams occur across the entire Bushveld Complex; indeed many are restricted to certain “compartments” (Naldrett et al. 2012).

Still, there are some gradual trends in (i) chemistry, (ii) mineral assemblage and (iii) mineral chemistry that have been identified in previous studies of the chromitite layers of the CZ. These include: (i) Chromite seams show a systematic chemical variation from the bottom to the top chromitite layer in terms of the Cr:Fe ratio and the abundance and proportion of PGE. Amongst others, Hatton and Von Gruenewaldt (1987) as well as Schürmann et al. (1998) showed a general decrease in chromium concentrations, reflected in a decreasing Cr/Fe ratio stratigraphically upwards from 1.6 in the LG-6 down to <1.3 in the UG-2. PPGE (Pt, Pd, Rh) contents display a progressive and substantial increase upwards in the CZ. The IPGE (Os, Ir, Ru) values, in contrast, remain broadly constant or rise only slightly (e.g., Naldrett and von Gruenewaldt 1989; Scoon and Teigler 1994). Naldrett et al. 2009 noted that the Pt/Ru (and Pd/Ru) ratio is highly variable in all major chromitite seams, but that Ru/Ir and Ru/Rh ratios are relatively consistent, suggesting that Pt and Pd are controlled by different concentration mechanisms than the IPGE and Rh.

(ii) The mineralogy of the chromitite seams and their associated accessory minerals provide further information for the differentiation and recognition of different seams. The gradual systematic

variation of chromitite composition with stratigraphic elevation is reflected in mineral assemblages and chromite mineral chemistry. In general, chromite exceeds 50 wt% (up to 95 wt%) in all chromitite seams; associated minerals, however, change from predominant orthopyroxene in the lowest LG chromitites to orthopyroxene and plagioclase in the MG chromitites and finally to plagioclase with minor orthopyroxene in the uppermost UG chromitites (Bachmann et al. 2018, Kinnaird et al. 2002). Orthopyroxene oikocrysts and poikilitic texture of pyroxene enclosing euhedral to subhedral chromite chadacrysts are common features in the seams (Kaufmann et al. 2018). Accessory minerals include amphibole, clinopyroxene, biotite/phlogopite, chlorite, talc, serpentine, quartz, and carbonates, base metal sulphides (BMS) and platinum-group minerals (PGM) (e.g., Bachmann et al. 2018).

(iii) Naldrett et al. (2009) document systematic variation mineral chemistry of chromite across the RLS stratigraphy. They observed two trends: 1) increasing Cr/(Cr + Al) and decreasing Mg/(Mg + Fe^{2+}) from LG-1 to LG-4 (trend A), and 2) decreasing Cr/(Cr + Al) still associated with decreasing Mg/(Mg + Fe^{2+}) from LG-5 to MG-2 (trend B).

Despite these well recognized trends, all attempts of robust discrimination between different chromitite seams suffer from (i) considering less variables than available (e.g., Cr/Fe ratio, various PGE ratios), (ii) subjective correlation of seams largely based on interlayer silicates and thickness of chromitites, (iii) significant overlap of compositional fields, e.g. of chromitite mineral assemblages and chromite mineral chemistry between (neighbouring) seams, (iv) added complexity due to post magmatic processes, such as hydrothermal alteration or oxidation (Junge et al. 2015, Oppermann et al. 2017).

To overcome the complications listed above, an enhanced chromitite discrimination scheme should be built on a comprehensive database and appropriate multivariate statistics (Tolosana-Delgado et al. 2017). The resulting classification should be applicable either in restricted ‘compartments’ of the Bushveld or at least locally within a deposit. In this study we propose such a classification scheme based on linear discriminant analysis using (i) a large database covering a significant number of chromitite seams, and (ii) considering the compositional nature of whole rock

chemical analyses of main element oxides and PGE. Output data is in the form of classification probabilities, hence automatically assessing the uncertainty on the classification and the seam correlation. The model also effectively identifies the main separation variables between distinct seams within the dataset. Furthermore, the results are used to test the viability of the scheme and simplify it, whenever possible. Finally, possible implications for the Bushveld Complex in terms of genesis and/or post-magmatic modification are briefly considered. The developed classification tool might be useful for several purposes, such as exploration for PGE and/or chromite deposits, but also mine-scale mapping and metallogenetic studies (e.g. Kaufmann et al. 2018).

Analytical methods: sampling and QAQC

As a case study, we used a comprehensive data set comprising chromitite samples from the Thaba Mine, located on the western limb of the Bushveld Complex. The operation is run by Cronimet, S.A. (Pty) Ltd., predominantly exploiting the LG-6/-6A and the MG chromitite layers. The LG-6 and LG-6A occur in close vicinity to each other and generally comprise a single seam with minor pyroxenite partings. The LG-7 is not developed as a single seam but as chromite stringers and less persistent chromite-bearing intervals, which cannot be traced for longer distances along strike or down dip. The Middle Group succession present at the Thaba Mine includes five main seams, named from the base upwards MG-1, MG-2, MG-3, MG-4 and MG-4A. While the MG-1 comprises an individual seam, the MG-2 is split by a pyroxenite parting up to 1 m in thickness in roughly 30 % of the cases. The MG3 is usually a single seam that locally contains multiple thin pyroxenite partings. The MG-4 has a similar appearance, often with numerous pyroxenite partings, it is overlain by approximately 2 m of barren pyroxenite; the MG-4, in turn, is closely associated with the minor seams MG-4Z (at the base of the MG-4) and MG-4A (just above the MG-4). A couple of minor seams also appear in the MG at Thaba mine e.g. the MG-3A. Further detailed information about the microstructure and the mineral assemblage of the studied chromitites, especially LG-6, LG-6A, MG-1 and MG-2 are described in Bachmann et al. (2018, 2019).

Sampling and structure of the data set

The data set used in this study was provided by Cronimet SA (Pty) Ltd., comprising more than 1200 chromitite samples and covers 14 distinct chromitite seams (LG-1 – LG-7, MG-1 – MG-4, LG-6A, MG-4A, MG-4Z) from Thaba mine. The data shown are restricted to massive chromitite, excluding footwall or hanging wall seams, but analyses may contain fine (up to a couple of cm-thick) silicate intercalations, pyroxenite partings, pyroxene oikocrysts, etc. These remain in the data set in order to represent the whole range of possible chromitite compositions within the mine. Furthermore, the depths of the assayed samples range from ca. 3 m to ca. 620 m below present day land surface, i.e. the data set includes samples affected by modern weathering – marked by alteration of silicates and oxidation of trace sulphide minerals. In accordance to literature (Junge et al. 2015), the weathered zone – or zone affected by supergene alteration – at Thaba Mine is restricted to intersections <50 m below present day land surface. Using 50m depth as a threshold suggests that a large part of the sample set (N = 370) considered here may have been affected by weathering.

During exploration and exploitation at Thaba Mine the composition of the chromitite seams is routinely constrained by only six major element oxides (Cr_2O_3, FeO, Al_2O_3, MgO, SiO_2, CaO) and P. In addition, the company data base contained values for Au and PGE concentrations, including Pt, Pd, Rh, Ru and Ir. The database was filtered beforehand for any missing values (such as “not analysed” or “below the detection limit”) among the major element oxides and the PGE. Because of incompleteness of data 83 samples were discarded. Due to limited available information all seams represented by less than 30 samples with complete records were not further considered. This included the seams LG-1 – LG-5 and LG-7 (totalling N=12 cases). The remaining 1110 chemical analyses contain data for eight different chromitite seams, namely (stratigraphically upwards) the LG-6, LG-6A, MG-1, MG-2, MG-3, MG-4Zero (MG-4Z), MG-4 and MG-4A.

Because only massive chromitite was analysed, multiple analyses of the same chromitite seam occur within the dataset. For example, wherever lower and upper MG-2 seams were observed, a chemical assay was carried out for each of these chromitites. This routine was also applied to the MG-3 and MG-4 with respect to cm- to dm-thick pyroxenite parting intersections. This approach

results in up to eight chemical assays for some of the MG-3 chromitites with analysed intervals ranging from 5 cm up to 200 cm. Similar results are available for MG-4, with up to 11 analyses per seam, with sample intervals ranging from 4 to 150 cm. Detailed information about the analyses per seam is provided in Table 1. For improvement of classification results we combined all assays belonging to a single seam intersection for the MG-3 and MG-4/4A/4Zero weighted by the length of the analysed section (Dataset A). For comparison the full number of assays was used for the classification scheme (Dataset B; appendix A). Even though results for Dataset B might be more robust, as the effect of aggregating data for Dataset A is known to reduce the variability (*support effect*), which could falsely reduce the uncertainty on the classification (Chilès and Delfiner, 1999), Dataset A consists of comparable analyses of full chromitite analyses over a succession of eight seams. Therefore, it reduces bias caused by the sampling procedure and represents a realistic scenario, e.g. of a drilling program within an exploration campaign. Nevertheless, the implications on the discrimination scheme using Dataset B will be evaluated as well. Working with Dataset B is expected to permit the application of the classification scheme with only limited amount of chromitite seams, for example in field mapping and sampling campaigns.

Analytical Methodology

Chemical data including major element oxides analyses of four of the chromitite seams from Thaba were recently published in Bachmann et al. (2019). Therein, possible sources of errors and the quality of the dataset were discussed for the LG-6, LG-6A, MG-1 and MG-2 seams. Identical reasoning and discussion was extended to the comprehensive data set used in this study. It was shown that Cronimet SA (Pty) has conducted a comprehensive program of quality control and quality assurance (QAQC) for all core samples, well comparable to best practice approaches (e.g., Abzalov 2008 and references therein). The assay data was sourced by Cronimet from two commercial laboratories, Setpoint Laboratories (Johannesburg) and METCHEM Laboratories (Johannesburg), within a time frame of almost four years during the duration of exploration campaigns. In both laboratories, major element analysis was carried out by X-ray fluorescence spectrometry (XRF) on fused glass beads.

Randomly inserted un-mineralized leuconorite samples were used as blanks and display Cr_2O_3 and FeO concentrations below the set threshold limit (<1% Cr_2O_3; <10% FeO). Quarter core field duplicate samples were used as blind duplicates and show relative deviations of less than 10 %. Internationally certified standards of South African chromite ores (SARM8 and SARM9) were randomly inserted to monitor Cr_2O_3 and FeO accuracy at Setpoint Laboratories, using an acceptance criterion of ±3σ. Overall, all analyses were reported within the acceptance criterion but FeO tended to be assayed at a lower concentration than its certified mean, with one analyses close to the 3σ acceptance criterion. Inconsistencies in the data were related to sample batches that were sent to different commercial laboratories during an extended period of time. While these minor inconsistencies may cause inappropriate results in "intra-seam" observations, "inter-seam" investigations are not significantly affected.

All PGE analyses were carried out by SetPoint Laboratories. Table 2 displays a statistical summary for the data. The PGE measurements contain several outliers but the data set is consistent without conspicuous statistical anomalies. The detection limit for all PGE concentrations was reported as 0.01 ppm. The quality of the PGE measurements can be assessed by duplicated measurements of the same sample. At first, the differences X_D between the two results, X_1 and X_2, of a duplicated measurement are calculated by$X_D = X_1 - X_2$. The uncertainty $X_{uncertainty}$ is calculated by:

$$X_{uncertainty} = Var[X_D] + E[X_D]^2, \quad \text{(I)}$$

where $E[X_D]$ and $Var[X_D]$, respectively, denote the average and variance of these differences. The PGE data set used in this study contains two types of duplicated measurements. Firstly, a random sample is duplicated and measured a second time (called umpires). The complete data set contains 17 of these duplicated measurements. Secondly, certified standards provided by African Mineral Standards as well as and un-mineralized leuconorite, acting as blanks, were randomly inserted into the sample batch as unknown to the analytical laboratory. Ninety chemical analyses were carried out on four different certified standards (AMIS0074, AMIS0075, AMIS0132 and AMIS0207). The measured results are compared to the reference values for the standard samples given in Table 3.

Table 4 displays the obtained uncertainties for the PGE. Since the uncertainties are based on measurements of the same sample, in the ideal case these multiple measurements should obtain the same result and, thus, the uncertainties should be zero. Values larger than zero reflect the analytical uncertainties of the measurements. To be able to make statements about the quality of the PGE measurements used in this study, the calculated uncertainty values were compared to the data variances. A high data quality can be assumed if the uncertainty values are significantly lower than the variances in the data. For all elements, the uncertainty values are at least one order of magnitude lower than the variances in the data (table 2, column "Var."). The uncertainties based on the duplicate measurements are slightly lower (except for Rh) than the uncertainties based on the standard sample measurements. The quality of the PGE measurements is appropriate for the aim of this study, since both types of uncertainties show similar results in spite of being based on different types of duplicate measurements. The uncertainty based on the duplicate samples may be an indicator for a nugget effect if it is significantly larger than the uncertainty based on standards. This is not the case for the used data, offering higher certainty that the results of this study are based on representative samples of the chromitite volumes sampled.

Classification strategy

In order to define an appropriate classification strategy, a first evaluation of the dataset is carried out, based on a series of box- and whisker plots, spider plots and ternary diagrams. Figure 1 shows the variability of the major element concentrations for the eight chromitite seams considered. Most major element oxide concentrations follow general stratigraphic trends that are well-documented for the chromitite seams of the CZ, including a decrease of Cr_2O_3 that is matched by an increase of SiO_2, CaO and (moderately) Al_2O_3 in stratigraphically upward direction (e.g. Scoon and Teigler 1994). Whilst these gradual trends distinguish well between the lowermost and uppermost chromitite seams of the CZ they are not useful to neighbouring chromitite seams.

It should be further noted that, when compared with a normal distribution, the spread of each variable shows a significant number of outliers, displayed as circles in the boxplots. These outliers

might challenge a discrimination scheme but seem to be rather common in the Bushveld Complex and are commonly attributed to local heterogeneities in the chromitites, e.g. the presence of large pyroxene oikocrysts. Figure 2A displays the whole range of variability in the PGE distribution in a chondrite-normalized format and separated into LG and MG chromitites. Using solely the PGE concentration, differentiation between the two groups is not possible, although the LG data shows significantly lower variability than the MG data. In order to evaluate the PGE profiles in detail box- and whisker plots without outliers are shown (Figure 2B-F). This provides a systematic overview over the actual PGE distribution in the different seams, without getting side-tracked by extreme and uncommon concentrations. We propose this type of PGE-spider diagrams for further use, when plotting bigger data sets, to get a more realistic impression on the actual PGE distribution.

Accordingly, the total PGE content progressively rises from LG-6 to MG-3 but decreases in MG-4A, which is in good agreement with previous studies (e.g. Scoon and Teigler 1994, Naldrett et al. 2012). The Pt/Ru ratio (Table 5) reflects rather flat PGE profiles for the lower chromitites and MG-4A, compared to chromitites further up in the stratigraphy. In contrast, there is no consistent trend recognizable for the Pt/Pd ratio throughout the succession. Nevertheless, this ratio differs between the different MG-4 seams, where MG-4Zero represents a significantly higher Pt/Pd ratio than MG-4(A). In addition, Au contents are similar in all chromitite seams, with concentrations close to the detection limit and very small variability. Hence, Au was excluded from further consideration.

According to Figure 1 and Figure 2, two subsets of elements might be useful to distinguish the chromitite seams. Using a ternary plot of Cr_2O_3-SiO_2-CaO (Figure 3A), LGs can be discriminated from MG-3 and MG-4(A, Z), while MG-1 and MG-2 overlap with both groups. In contrast, the Ru-Pt-Pd ternary diagram discriminates LGs from MG-1 and MG-2, while MG-4(A) is overlaps with MG1 and 2. MG-3 is shifted towards low Pt/Pd ratios and MG-4Z towards higher Pt values.

These descriptive evaluations highlight that whole rock compositions - including PGE concentrations - have to be used in an attempt to attain a proper classification. The absolute concentration of PGE appears to be of relevance, not only their relative proportions. Also, the separation of chromitite

seams should be performed in consecutive steps, as a global one-step classification did produce very unsatisfactory results (not shown here to save space).

In the following, LGs will be discriminated from MGs first; then LG-6 and LG-6A will be distinguished (Figure 4). Afterwards, focus will be on the differentiation of the MGs; MG-1/-2 will be split from MG-3/-4(A, Z), and MG-1 from MG-2. Finally, MG-3 and MG-4(A, Z) will be classified and in a last step the MG-4s distinguished. Given the overlap of several of these groups – especially LG-6/LG-6A, MG-1/MG-2, and the MG-4s – the output of the classification scheme will be a set of probabilities that a given analysis will stem from an individual chromitite seam, instead of a specific group without considering any uncertainty. This will be achieved in the next section with a multivariate statistical approach.

Multivariate statistical discrimination

We use a linear discriminant analysis (LDA, Fahrmeir and Hammerle 1984) as a multivariate statistical method for building each discrimination step. The set of components [Cr_2O_3, FeO, SiO_2, MgO, Al_2O_3, CaO, Ir, Ru, Rh, Pt, Pd] were previously transformed with an isometric logratio transformation (ilr), to account for their compositional nature (Aitchison 1986, Aitchison et al. 2002, Egozcue et al 2003). A set of variables forms a composition if its size (or total sum) is irrelevant for the problem under consideration, or if it is an artefact of the sampling procedure (van den Boogaart and Tolosana-Delgado 2008). The R software environment v3.2.4 (R Core Team 2016) with the additional packages "MASS" (function lda; Venables and Ripley, 2002) and "compositions" (function ilr; van den Boogaart et al. 2014) were used for building a compositional discriminant analysis. Discriminant analysis constructs a discriminant function as a linear combination of the available variables that maximize the differences between groups while minimizing their internal variability (Tolosana-Delgado et al. 2017). Therefore, not only the average values of each component but also their variabilities and correlations are important. Following the classification strategy displayed in Figure 4, we have built five steps of discrimination, including seven linear discrimination analyses. Because the amount of

data within each seam differs significantly (e.g. LG-6 = 226 vs. LG-6A = 35), it is crucial to weight the groups beforehand to not artificially favour the ones with more observations in this data set.

Stepwise classification scheme

The first discrimination step is the classification of chromitites into the two main groups – LGs and MGs. This is achieved by using all ilr-transformed, weighted components. Figure 5 displays the probability density estimates and boxplots of the scores of linear discriminant function 1 (LD1) by group, showing excellent separation of the LG from the MG chromitites. A so-called cross-validation was applied to this LDA. Therefore, each sample is removed from the dataset and its probability of belonging to each seam is predicted by means of a linear discriminant function built without it. The highest probability obtained gives the prediction from which chromitite seam the particular sample should have derived. In this way, a true and a predicted chromitite seam origin is available for each sample. The results of the cross-validation are in an excellent agreement between true and predicted seams with an overall misclassification rate of only 2.3 %, i.e., the rate of correct classification is far above 90%. Very similar misclassification rates (2.7%) were achieved using a naive reclassification, i.e. without the cross-validation computational burden (Figure 5). This can be expected due to the large data set used in this study and led us to use this naive reclassification strategy instead of cross-validation results for the following classification steps.

The second classification step needs to discriminate either LGs or MGs from one another. On the one hand, LG-6 is discriminated from LG-6A and, on the other hand, MG-1/MG-2 is distinguished from MGs-3/4/4A/4Zero. Figure 6A displays the probability density estimates and boxplots of the scores of LD1 by group, showing good separation of LG-6 from LG-6A. Classification results are displayed in Table 6A, confirming the very good agreement between true and predicted chromitite origin with a correct classification rate significantly higher than 90 %. The LDA for the MGs does not deliver such a clear result (Figure 6B); this is also reflected in the classification results, with a rate of misclassification of about 14 % (Table 6B). The third classification step is again twofold, discriminating MG-1 from MG-2 and MG-3 from MG-4/4A/4Zero. MG-1 and MG-2 display well-

separated probability density plots and boxplots (Figure 6C) with misclassification rates below 10 % (Table 6C). Discrimination of MG-3 from MG-4/4A/4Zero is also successful - with more than 90 % correct classification (Figure 6D, Table 6D). The final two classification steps are used to discriminate between the different MG-4 seams, namely MG-4/4A from MG-4Zero and MG-4 from MG-4A. While separating MG-4Zero works very well (3 % misclassification; Figure 6 E, Table 6E), an overlap in probability density plots (Figure 6F) and significant misclassification is recorded between the MG-4 and MG-4A, i.e., 12 % (Table 6F). The discrimination of the MG-3 and MG-4 chromitites will be further evaluated and discussed in the following section. This section will also present a global reclassification exercise, as a validation strategy reasonably analogous to cross-validation, i.e. less prone to overfitting than the naive reclassification.

Discussion

In the first part of this section, we integrate the results into a comprehensive classification process, complemented by a discussion on the influence of supergene alteration on the data. In the second part, we evaluate the scaling of the single components used in the discrimination scheme to establish possible simplifications.

Implementing a comprehensive classification process

The results of the discrimination of neighbouring seams as reported in the previous sections showed excellent results for the majority of attempts. However, for chromitite samples of unknown stratigraphic position, the challenge differs from just separating two seams: rather, the goal is to obtain probabilities that a sample belongs to a certain seam. Thus, we have to predict the probability of an overall successful attribution of an unknown sample, combining all LDA outcomes. This has been done for all major subgroups of chromitites, e.g., if there is the prior knowledge that a sample is from the LGs or the MGs, or if it is from the LCZ or UCZ, respectively. We thus use our data set as a training set of samples with unknown composition, only making sure to deal with massive chromitites and that the samples belong to any of our target chromitite seams.

To estimate the overall success of the procedure we go through all the steps outlined in Figure 4, without using further prior knowledge, i.e., the actual seam denotation. After classification the result is compared with the actual geological position of the sample and the overall performance is given as a percentage of correctly classified samples. In general, the use of chromitite whole rock assays yields good result, with more than 80 % correctly classified samples. This result is owed to a small number of seams with higher off-diagonal values in the misclassification table (Table 6), e.g. the classification of MG-1 and MG-4 chromitites (Table 7A). On the contrary, classifications performed with samples belonging only to the LCZ yield excellent results (90.4 %), with the aforementioned MG-1 showing the lowest correct classification (70 %, not shown). In addition, correct classification of chromitites from the UCZ reaches more than 85 %, with highest misclassification rates in the aforementioned MG-4, which shows a considerable overlap with MG-3 and MG-4A (Table 7C). In contrast, results are getting worse when considering all MG chromitites (77.2 % correct classified samples).

Accounting for support effects

It has been mentioned that aggregating data reduces the variability, which in turn underestimates the uncertainty on the classification (the support effect as known in the geostatistical literature, see e.g. Chilès and Delfiner, 1999). To evaluate these effects, the analysis was repeated for Dataset B, which shows higher heterogeneity is expected to cause more challenges in the classification of the chromitites (see appendix A). Indeed, the percentage of correct classification decreased for all MG samples. While for $LDA_{LGs\ vs.\ MGs}$ and $LDA_{MG12\ vs.\ MG34}$ only slight decreases (1-2 %) occur, misclassifications significantly raise from 7 to 11 % and 3 to 12 % for $LDA_{MG3\ vs.\ MG4}$ and $LDA_{MG4\ vs.\ MG4Z}$, , respectively. Most challenging is the discrimination of $LDA_{MG4\ vs.\ MG4A}$, with a drop in correct classifications from to 88 to 70 %. The overall performance decreased considerably by using Dataset B, with misclassification increasing from around 19 % to 34 % for the entire sample set (Table 7B). As expected, in particular the UCZ causes problems, showing only 65 % correct classifications (Table 7D). A more detailed evaluation of Dataset B can be found in the electronic supplementary material A.

As stated earlier, the significant drop in correct classifications is caused by the higher variability of Dataset B. The higher variability can be caused by nugget effects; however, the presence of relevant nugget effects was previously discarded for this dataset by testing original, duplicate and umpire samples– as described above. In addition, using averaged values, as done in Dataset A, is known to lower the variability by a factor that can be derived from theoretical considerations, i.e. the central limit theorem of Statistics under the assumption that the samples being mixed are mutually uncorrelated. Indeed, one could compare this theoretical reduction with the actual one, discrepancies eventually reflecting lack of independence, or a systematic chemical variation within the seams. The results may then be linked to possible implications for the genesis of the chromitite seams. However, this is beyond the scope of this paper and is left for further research. Nevertheless, it can be concluded that both data sets (Datasets A and B) will keep their practical use. Dataset A is well applicable to drill core samples, e.g. during exploration campaigns or in mining operations – especially around faults, where a proper seam correlation can be difficult but one has access to material along the whole width of the seam. In these cases, assuring an appropriate sampling of the whole chromitite seam is crucial. In contrast, for hand specimens collected during field mapping or samples from mining dumps, Dataset A might overestimate the correct classification of chromitites because only parts of the chromitites are sampled. In these cases Dataset B should be used.

The influence of alteration

It is important to note that the data set contains a significant number of samples exposed to supergene weathering. In addition, some samples may be affected by structurally-related hydrothermal alteration processes. These processes affect not only the mineral assemblage but also the geochemistry of chromitite samples (e.g., Hey 1999, Becker et al. 2014, Bachmann et al. 2019). Becker et al. (2014) documented an enrichment of alteration silicates, and Fe-oxides/hydroxides as well as remobilization of Pd in weathered PGM ores. Remobilization of Pd might occur during hydrothermal alteration and or supergene alteration (Junge et al. 2015, Oppermann et al. 2017).

Indeed, the large variation of Pd concentrations in the data set is attributed to weathering processes (cf. Figure 2). In addition, Bachmann et al. (2019) showed generally lower Cr_2O_3 and increasing FeO, SiO_2 and CaO concentrations for supergene-altered chromitite samples. This is similar to comparable altered and oxidized PGE ores elsewhere (Oberthür et al. 2003). Because the degree of supergene alteration is difficult to judge without detailed petrographic studies (Bachmann et al. 2019), the success of a geochemical classification scheme should not be significantly affected by the degree of weathering. In order to test if the classification procedure is robust against alteration we classified the Dataset A samples through the stepwise classification rules previously presented particularly tracking those coming from the weathering zone. The overall classification success gave misclassification rates of 17 % for the pristine samples. In contrast, the sample population from a depth of less than 50 m below present day land surface – and thus possibly affected by weathering - display only marginally higher misclassification rates, around 22 %, and confirm that the classification scheme can be used for pristine and weathered chromitite samples.

Influence of each component and simplification of classification

In order to better understand the actual discrimination function and the separation into groups of data, it is crucial to evaluate the influence of each component on the linear discriminant functions, i.e. the coefficients of linear combinations of log-transformed components that give rise to the discriminant functions. The coefficients used throughout the study are displayed in Figure 7. The bar plot can be expressed in an equation as follows:

$$LD_{LGsvs.MGs} = log\left[\frac{\left(C_{FeO}^{\alpha_{FeO}} \times C_{SiO_2}^{\alpha_{SiO_2}} \times C_{Al_2O_3}^{\alpha_{Al_2O_3}} \times C_{CaO}^{\alpha_{CaO}} \times \left(\beta C_{Pt}\right)^{\alpha_{Pt}} \times \left(\beta C_{Pd}\right)^{\alpha_{Pd}} \times \left(\beta C_{Rh}\right)^{\alpha_{Rh}} \times \left(\beta C_{Rh}\right)^{\alpha_{Rh}}\right)}{\left(C_{Cr_2O_3}^{\alpha_{Cr_2O_3}} \times C_{MgO}^{\alpha_{MgO}} \times \left(\beta C_{P}\right)^{\alpha_{P}} \times \left(\beta C_{Ru}\right)^{\alpha_{Ru}}\right)}\right],$$

where LD is the linear discriminant, C_i is the concentration of element i, α_i are the coefficients (the values reported in Figure 7 are normalized by its log-standard deviation, to make them comparable), and β is a factor to bring all concentrations into the same scale, i.e., from ppm to wt%. Thus, bigger values will have larger influences on the probability of the analysis to be sorted into one or the other class: variables in the numerator favour MGs, variables in the denominator LGs. Additionally, small

variability (higher homogeneity) of components of the dataset within a seam might lead to a leverage effect, making that variable more influential. For example, due to the little variation of MgO in the LG-6, it is used as an important discriminator against the LG-6A, even though the median of MgO is slightly higher for the latter (cf. Figure 1).

It is also important to note that bigger values increase the discrimination efficiency in either ways. For example, the discrimination between the LGs and MGs or MG-4 and MG-4A is less robust than the separation of MG-1 and MG-2, even though the misclassification might be similar or higher in the latter. This relationship is also valid for single components, thus only MgO, Al_2O_3 and Pt are important contributors to LD in LG-6 vs. LG-6A. Hence an approximation to $LD_{LG6vs.LG6A}$ might be computed as:

$$LD_{LG6vs.LG6A} = log\left[\frac{(\beta C_{Pt})^2}{C_{MgO} \times C_{Al_2O_3}}\right].$$

In general, these reduced equations can be simplified again using:

$$LD_{LG6vs.LG6A} = 2log\beta + log\left[\frac{{C_{Pt}}^2}{C_{MgO} \times C_{Al_2O_3}}\right].$$

whereas the term $2log\beta$ will can be expressed as a constant k summarizing all used factors β. Additionally, k is not needed to discriminate between seams, because it just acts as a translation constant. Hence, for a discrimination procedure to separate LG-6 from LG-6A analyses, concentrations of just three elements are needed to justify an appropriate classification. Detailed results for all LDAs shown in Figure 7 are displayed as boxplots in Figure 8. Especially the separation of LG-6 vs. LG-6A can be approximated by using not more than three elemental concentrations, while the classification of MG-1/2 vs. MG-3/4 and MG-1 vs. MG-2 need a rather extensive data set.

Conclusion and outlook

In this study a reliable multivariate discrimination scheme for chromitite seams of the LG and MG layers collected at the Thaba Mine of the western Bushveld Complex – including a clear attribution with known uncertainty – is proposed. The classification scheme is built on a large database that has

been considered as original (Dataset B) and as aggregated (Dataset A) datasets. While Dataset A is well applicable to drill core samples, e.g. during exploration campaigns or in mining operations, Dataset B might be used for hand specimens collected during field mapping or samples from mining dumps. This general approach may well be applicable to chromitite seams across the entire Bushveld as long as a comparable set of data is available. Such a tool should not only be able to agnostically discriminate the chromitites but would also reveal geochemical differences between the different "compartments" and trends within them. In order to extend the presented approach further research is necessary to optimise the proposed chromitite discrimination scheme, including further enlargement and improvement of the database by adding different data sets from other regions in the Bushveld Complex.

Such sophisticated classification should also involve data from the eastern Bushveld, and also extend the database to the base of the LCZ, including LG-1 to LG-5, as well as the UCZ, including UG-1 to UG-3. We invite the community and the mining companies to share with us their experience, data, and problems with the new discrimination scheme in order to further improve this tool.

Acknowledgments

This is a contribution of the German/South African R&D project AMREP—Applied Mineralogy for Resource Efficiency of Platinum-Group Metals—funded by the German Ministry of Education and Research (BMBF; grant number BMBF 033R119E). We thank the Cronimet Mining Group for providing access to their core shed and drill core intersections from the Thaba mine, for the contribution of additional analytical data and information on the local geology and beneficiation. The comments of Cynthia Sanchez-Garrido to this manuscript are gratefully acknowledged.

References

Abzalov, M (2008) Quality control of assay data: a review of procedures for measuring and monitoring precision and accuracy. Exploration and Mining Geology, 17(3-4):131–144.

Aitchison J (1986) The statistical analysis of compositional data. Monographs on Statistics and Applied Probability. Chapman & Hall Ltd., London (UK). (Reprinted in 2003 with additional material by The Blackburn Press). 416p.

Aitchison J, Barceló-Vidal C, Egozcue JJ, Pawlowsky-Glahn V (2002) A concise guide for the algebraic-geometric structure of the simplex, the sample space for compositional data analysis. In Proceedings of IAMG Vol2:387–392.

Bachmann K, Osbahr I, Tolosana-Delgado R, Chetty D, Gutzmer J (2018) Variation in platinum group mineral and base metal sulfide assemblages in the Lower Group chromitites of the western Bushveld Complex, South Africa. Can Min, https://doi.org/10.3749/canmin.1700094.

Bachmann K, Menzel P, Tolosana-Delgado R, Schmidt C, Hill M, Gutzmer J (2019)The use of assay data as a foundation for a geometallurgical model – the case of the Thaba Chromite Mine, South Africa. J Geochem Explo, under review.

Becker M, Wiese J, Ramonotsi M (2014) Investigation into the mineralogy and flotation performance of oxidised PGM ore. Minerals Engineering 65:24–32.

Cawthorn RG, Webb SJ (2001) Connectivity between the western and eastern limbs of the Bushveld Complex. Tectonophys 330:195–209.

Cawthorn RG (2015) The Bushveld Complex, South Africa. In Layered Intrusions, Springer, Dordrecht, 517–587.

Chilès JP, Delfiner P (1999) Geostatistics: Modeling spatial uncertainty. John Wiley & Sons, New York. Geostatistics: Modeling spatial uncertainty. John Wiley & Sons, New York.

Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, Barcelo-Vidal C (2003) Isometric logratio transformations for compositional data analysis. Mathematical Geology, 35(3), 279-300.

Fahrmeir L, Hammerle A (1984) Multivariate Statistische Verfahren. Walter de Gruijter, Berlin:796p.

Fourie GP (1959) The chromite deposits in the Rustenburg area. South Africa Geological Survey Bulletin 27:45.

Hatton CJ, Von Gruenewaldt G (1987) The geological setting and petrogenesis of the Bushveld chromitite layers. In: Stowe, C.R. (Ed.), Evolution of Chromium Ore Fields. Van Nostrand Reinhold, New York, pp. 109–143.

Hey PV (1999) The effects of weathering on the UG2 chromitite reef of the Bushveld Complex, with special reference to the platinum-group minerals. South Afr J Geol 102:251–260.

Junge M, Oberthür T, Kraemer D, Melcher F (2015) Distribution of platinum-group elements in pristine and near-surface ores from the Platreef, northern Bushveld Complex, South Africa. In: André-Mayer AS, Cathelineau M, Muches P, Pirad E, Sindern S (Editors). Mineral Resources in a Sustainable World. 13th Biennial SGA Meeting, 955–958.

Kaufmann FE, Vukmanovic Z, Holness MB, Hecht L (2018) Orthopyroxene oikocrysts in the MG1 chromitite layer of the Bushveld Complex: implications for cumulate formation and recrystallisation. Contributions to Mineralogy and Petrology 173(2):17.

Kinnaird JA, Kruger FJ, Nex PAM, Cawthorn RG (2002) Chromitite formation—a key to understanding processes of platinum enrichment. Applied Earth Science 111(1):23–35.

Maier WD, Barnes S-J, Groves DI (2013) The Bushveld Complex, South Africa: formation of platinum–palladium, chrome- and vanadium-rich layers via hydrodynamic sorting of a mobilized cumulate slurry in a large, relatively slowly cooling, subsiding magma chamber. Mineral Deposita 48:1–56.

Naldrett A, von Gruenewaldt G (1989) Association of platinum-group elements with chromitite in layered intrusions and Ophiolite Complexes. Economic Geology 84:180–187.

Naldrett AJ, Kinnaird J, Wilson A, Yudovskaya M, McQuade S, Chunnett G, Stanley C (2009) Chromite composition and PGE content of Bushveld chromitites: part 1—the lower and middle groups. Trans Inst Min Metall B 118:131–161.

Naldrett AJ, Wilson A, Kinnaird J, Yudovskaya M, Chunnett G (2012) The origin of chromites and related PGE mineralization in the Bushveld Complex: new mineralogical and petrological constraints. Mineralium Deposita 47:209–232.

Oberthür T, Weiser TW, Gast L, Kojonen K (2003) Geochemistry and mineralogy of platinum-group elements at Hartley platinum mine, Zimbabwe: Part 2: Supergene redistribution in the oxidized main sulfide zone of the Great Dyke, and alluvial platinum-group minerals. Miner Deposita 38:344–355.

Oppermann L, Junge M, Schuth S, Holtz F, Schwarz-Schampera U, Sauheitl L (2017) Mobility and distribution of palladium and platinum in soils above Lower and Middle Group chromitites of the western Bushveld Complex, South Africa. South Afr J Geol 120(4):511–524, DOI:10.25131/gssajg.120.4.511.

R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Schürmann LW, Grabe P-J, Steenkamp CJ (1998) Chromium. In: Wilson MGC, Anhaeusser CR (eds) The mineral resources of South Africa: Handbook 16. Council for Geosciences, CTP Book Printers, Cape Town, pp 90–105.

Scoon RN, Teigler B (1994) Platinum-group element mineralization in the Critical Zone of the Western Bushveld Complex: I. Sulfide-poor chromitites below the UG-2. Economic Geology 89:1094–1121.

Tolosana-Delgado R, von Eynatten H, Krippner A, Meinhold G (2017) A multivariate discrimination scheme of detrital garnet chemistry for use in sedimentary provenance analysis. Sedimentary Geology, https://doi.org/10.1016/j.sedgeo.2017.11.003.

USGS (2018) Minerals Commodity Summaries, 2018: Chromium. U.S. Geological Survey, U.S. Department of the Interior.

Van den Boogaart KG, Tolosana-Delgado R (2008) “Compositions”: a unified R package to analyze compositional data. Computers & Geosciences 34(4):320–338.

Van den Boogaart KG, Tolosana R, Bren M (2014) compositions: Compositional Data Analysis. R package version 1.40-1. https://CRAN.R-project.org/package=compositions.

Veksler IV, Reid DL, Dulski P, Keiding JK, Schannor M, Hecht L, Trumbull RB (2015) Electrochemical Processes in a Crystal Mush: Cyclic Units in the Upper Critical Zone of the Bushveld Complex, South Africa. J of Petrology 56 (6): 1229–1250, https://doi.org/10.1093/petrology/egv036.

Venables WN, Ripley BD (2002) Modern Applied Statistics with S. Fourth Edition. Springer, New York. ISBN 0-387-95457-0.

Von Eynatten H, Pawlowsky-Glahn V, Egozcue, JJ (2002) Understanding perturbation on the simplex: a simple method to better visualise and interpret compositional data in ternary diagrams. Mathematical Geology 34 (3): 249–257.

Von Gruenewaldt G (1977) The mineral resources of the Bushveld Complex. Miner Sci Eng 9(2):83–95.

Voordouw R, Gutzmer J, Beukes NJ (2009) Intrusive origin for upper group (UG1, UG2) stratiform chromitite seams in the Dwars River area, Bushveld Complex, South Africa. Mineralogy and Petrology, 97(1-2), 75. https://doi.org/10.1007/s00710-009-0072-3.

Figures

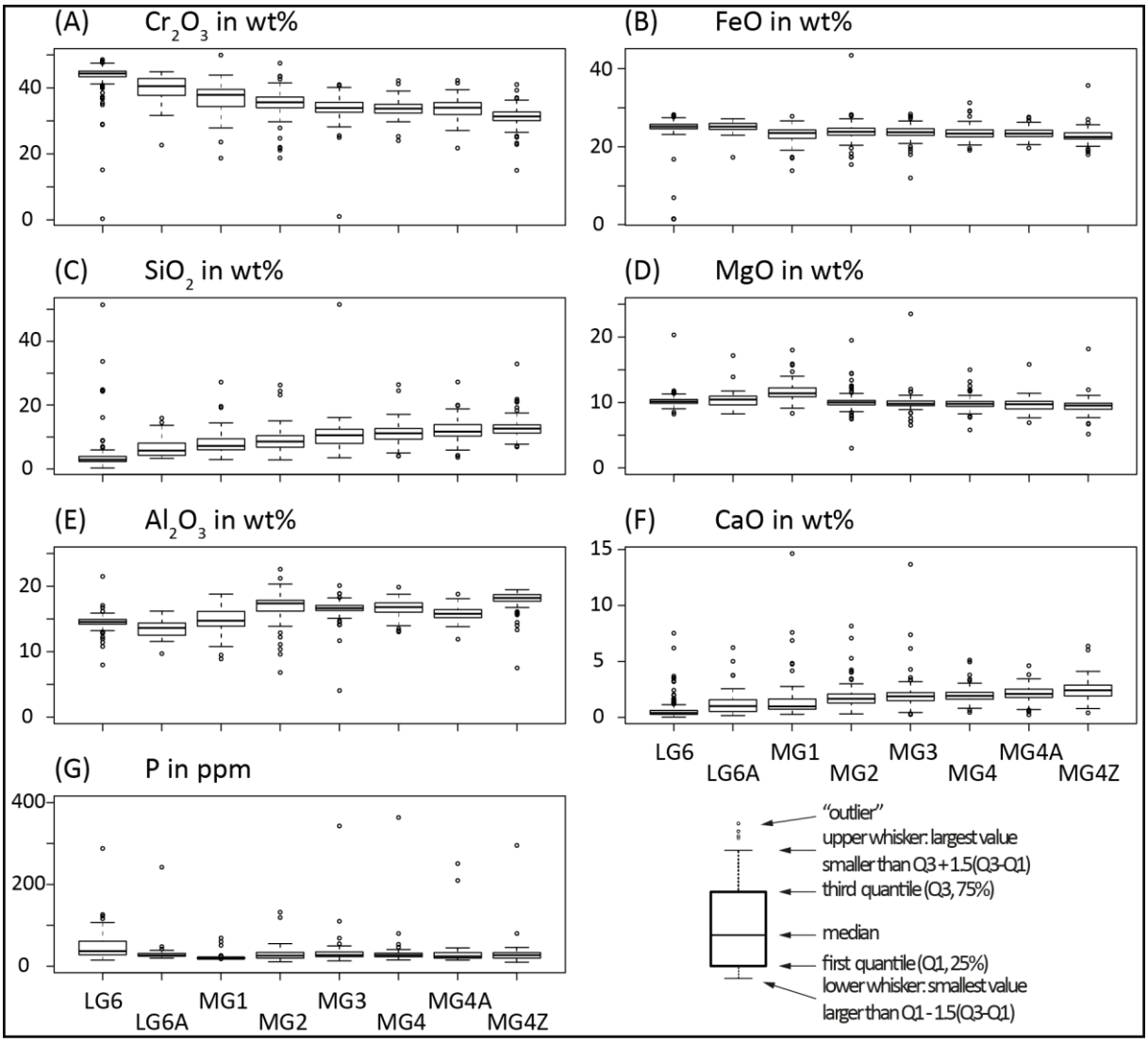


Figure 1 Major element concentrations in oxide wt% (except P – in ppm) for the LG-6 to MG-4A chromitites. The variability of individual seams is illustrated in box- and whisker-plots: (A) Cr_2O_3, (B) FeO, (C) SiO_2, (D) MgO, (E) Al_2O_3, (F) CaO, (G) P.

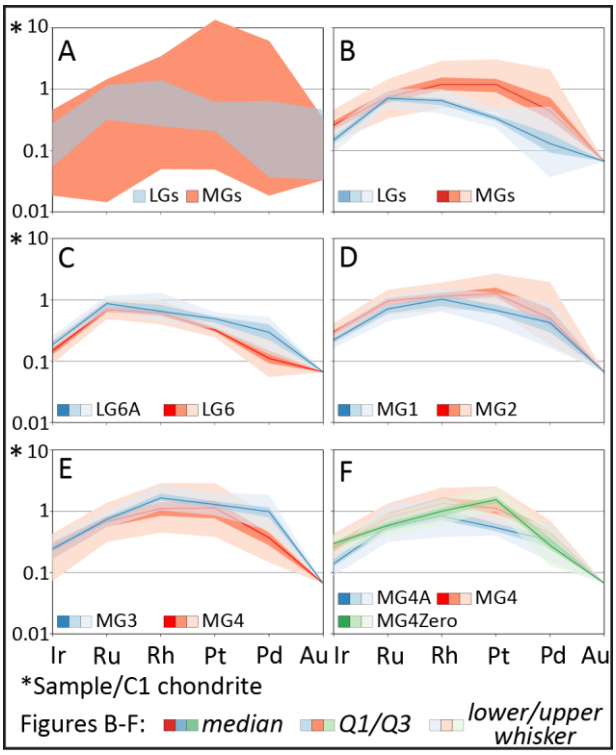


Figure 2 Chondrite-normalized PGE profiles for chromitite seams from LG-6 to MG-4A. (A) Whole dataset is expressed by an area that covers all the profiles for both LGs and MGs, respectively. (B-F) All PGE profiles are expressed as Box-and Whisker-Plots. For detailed explanation, please see Figure 1. MG-4 in (E) contains all PGE profiles of MG-4Zero, MG-4 and MG-4A. For details see (F). The chondritic concentrations used for the normalization are those of Naldrett and Duke (1980) (Ir=540; Ru=690; Rh=200; Pt=1,015; Pd=540; Au=150, all in parts per billion).

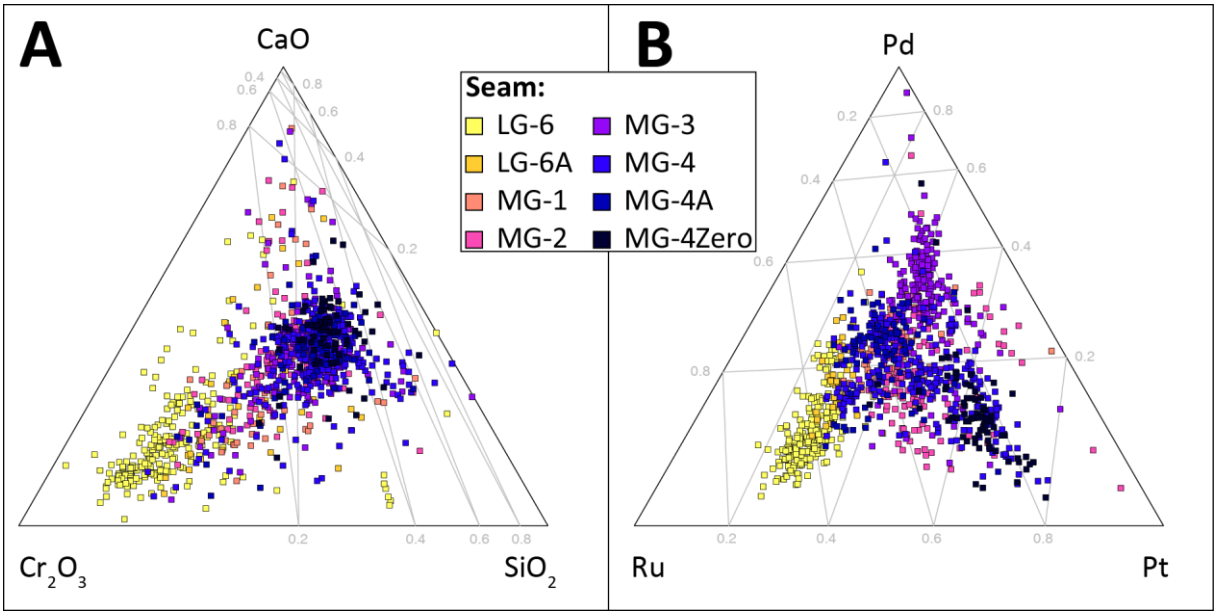


Figure 3 Ternary diagrams for (A) Cr_2O_3 – SiO_2 – CaO and (B) Ru – Pt – Pd centred according to von Eynatten et al. (2002) by division with their geometrical mean and split by chromitite seams.

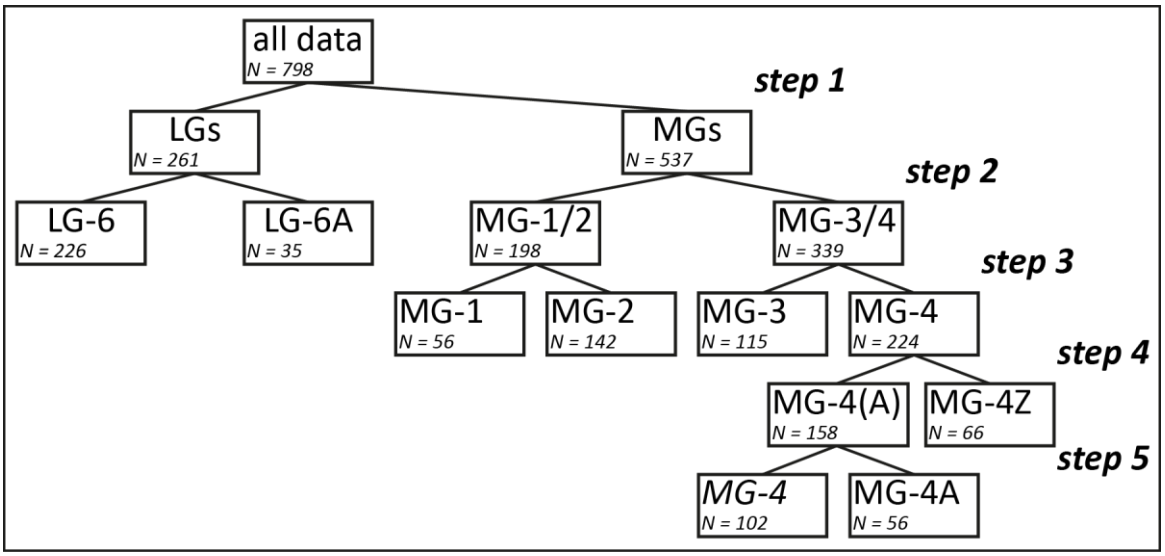
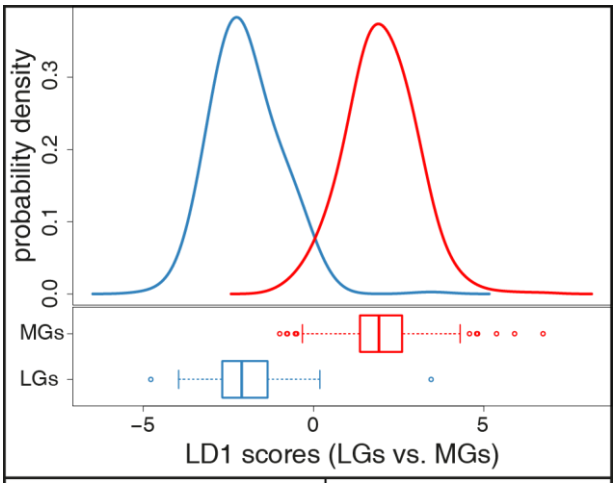


Figure 4 Hierarchical strategy of chromitite seam classification, with indication of sample size for each seam. MG-4: includes MG-4, MG-4Zero, MG-4A; MG-4(A): includes MG-4, MG-4A; MG-4: includes MG-4.



Missclassification 2.69%		True type LG	MG
Predicted type	LG	255	15
	MG	6	504
Correct Classification		97.70%	97.11%

Figure 5 Probability densities and boxplots of the scores of the linear discriminant function for lower group chromitites (LGs, blue) against middle group chromitites (MGs, red). Classification results for the two-way discrimination are shown in the lower part.

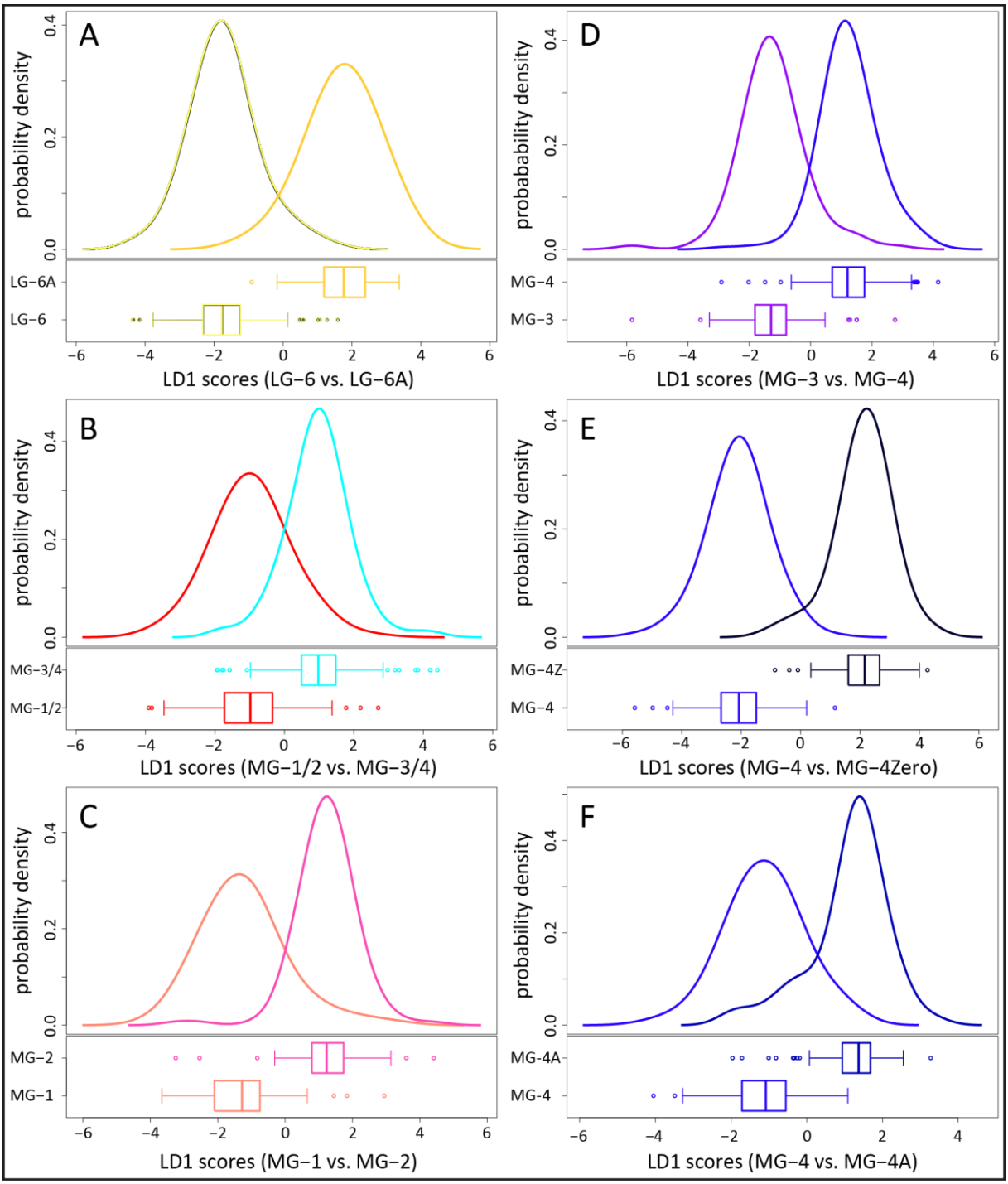


Figure 6 Probability densities and boxplots of the scores of the linear discriminant functions for lower and middle group chromitites following the classification scheme outlined in Figure 4. (A) LG-6 (yellow) vs. LG-6A (orange). (B) MG-1/2 (red) vs. MG-3/4 (turquoise). (C) MG-1 (pale red) vs. MG-2 (pink). (D) MG-3 (purple) vs. MG-4 (light blue). MG-4 contains data of MG-4/4A/4Zero. (E) MG-4 (light blue) vs. MG-4Zero (dark blue). MG-4 contains data of MG-4/4A. (F) MG-4 (light blue) vs. MG-4A (blue). MG-4Z and MG-4Zero are used synonymously.

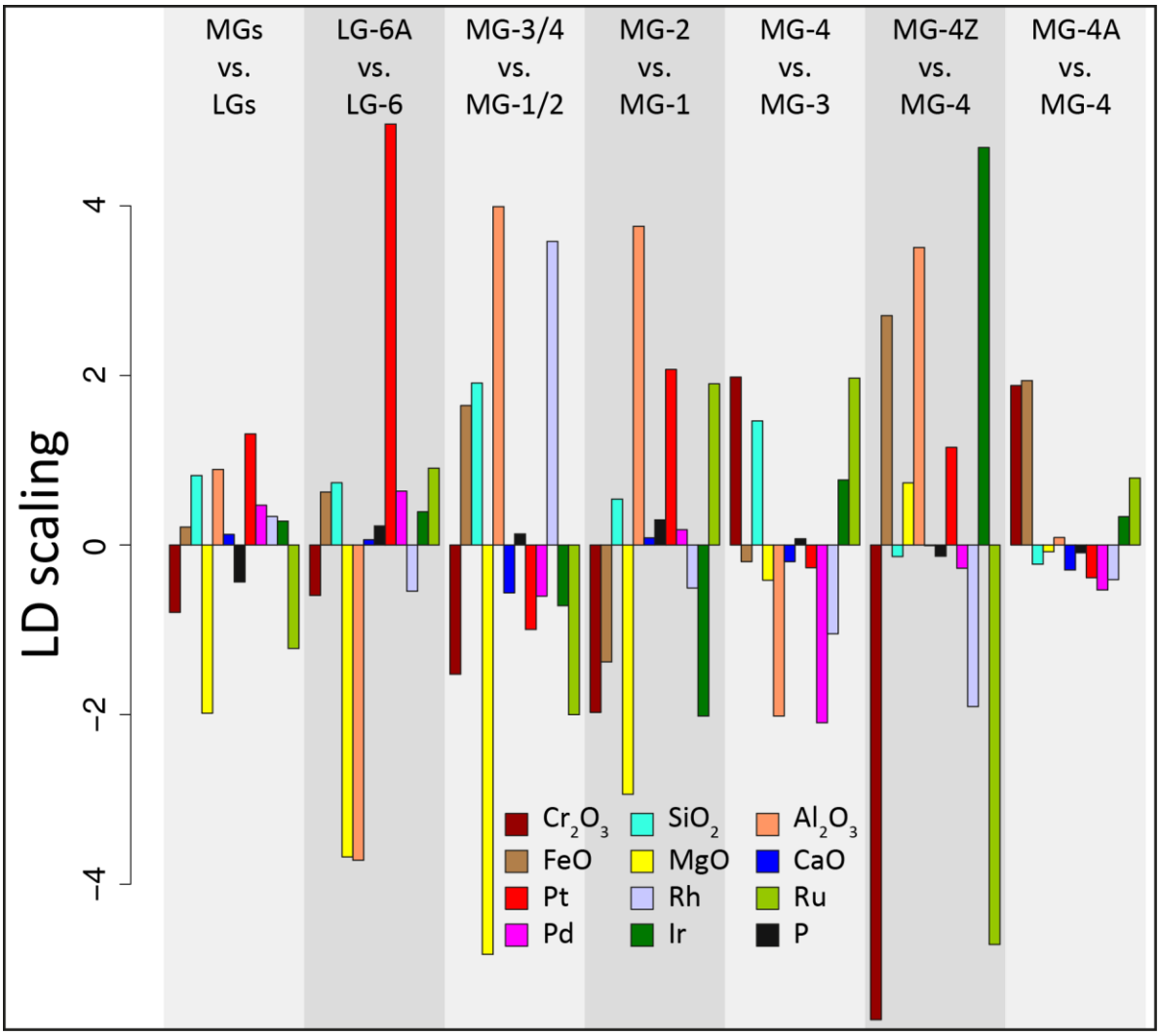


Figure 7 Normalized coefficients of the components on each linear discriminant function as a classification rule. For details see text.

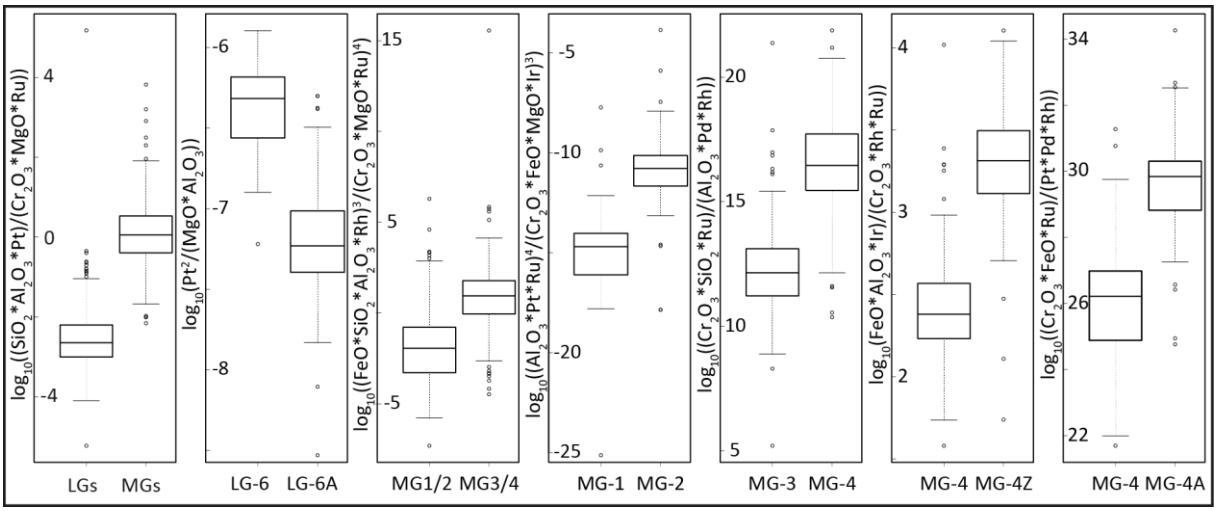


Figure 8 Box-and whisker plots showing simplified ratios to discriminate between chromitite seams along the classification path presented in Figure 4. For details see text. MG-4Z is MG-4Zero.

Tables

Table 1 Number of whole rock chemical analyses per seam.

Dataset	LG-6	LG-6A	MG-1	MG-2	MG-3	MG-4Zero	MG-4	MG-4A
A	226	35	56	142	115	66	102	56
B	226	35	56	142	202	77	289	83

Table 2 Statistical summary for available PGE data. All values (except variance) are given in [ppm], variance is given in [ppm]².

Element	Min.	1st Quantile	Median	3rd Quantile	Max.	Mean	Geom. Mean	Variance
Pt	0.04	0.41	0.93	1.45	13.7	1.057	0.820	0.783
Pd	0.01	0.12	0.19	0.32	6.17	0.274	0.190	0.107
Rh	0.01	0.14	0.21	0.30	1.38	0.236	0.208	0.017
Ir	0.01	0.08	0.12	0.16	0.46	0.124	0.114	0.003
Ru	0.01	0.41	0.49	0.61	1.98	0.517	0.488	0.030

Table 3 Reference PGE values for African Mineral Standards in [ppm].

Sample name	Pt	Pd	Rh	Ir	Ru
AMIS0074	1.1	0.7	0.2	-	0.4
AMIS0075	1.2	1.5	0.3	0.09	0.4
AMIS0132	0.5	0.2	0.1	0.05	0.2
AMIS0207	2.2	1.2	0.5	0.19	0.8

Table 4 Uncertainties for PGE measurements based on duplicated measurements (Duplicate) and measurements of standard samples (Standard). All values are given in [ppm]².

Element	Uncertainty (Duplicate)	Uncertainty (Standard)
Pt	0.066	0.091
Pd	0.006	0.020
Rh	0.005	0.004
Ir	0.001	0.001
Ru	0.008	0.016

Table 5 Relevant PGE weight ratios for all investigated seams.

Seam	Pt/Ru		Pt/Pd		PPGE/IPGE	
	Mean	Median	Mean	Median	Mean	Median
LG-6	0.5	0.5	2.9	2.9	1.3	1.3
LG-6A	0.6	0.6	1.7	1.7	1.5	1.4
MG-1	1.6	1.0	1.9	1.7	2.7	2.2
MG-2	2.0	1.2	3.2	2.5	3.0	2.3
MG-3	1.8	1.7	1.6	1.3	4.2	4.0
MG-4	1.5	1.4	2.8	2.6	2.9	2.8
MG-4A	1.0	0.9	2.1	1.7	2.5	2.4
MG-4Zero	2.7	2.6	5.7	5.2	3.3	3.2
MG-4_all	1.7		3.5		2.9	

Table 6 Classification results for the discrimination of chromitite seams.

(A) LG-6 vs. LG-6A				**(D) MG-3 vs. MG-4/4A/4Zero**			
Misclassification		True type		Misclassification		True type	
6.13%		LG-6	LG-6A	6.85%		MG-3	MG-4
Predicted type	LG-6	212	2	Predicted type	MG-3	98	12
	LG-6A	14	33		MG-4	10	201
Correct Classification		93.8%	94.3%	Correct Classification		90.7%	94.4%
(B) MG-1/2 vs. MG-3/4/4A/4Zero				**(E) MG-4/4A vs. MG-4Z**			
Misclassification		True type		Misclassification		True type	
14.26%		MG-1/2	MG-3/4	3.29%		MG-4	MG-4Z
Predicted type	MG-1/2	163	39	Predicted type	MG-4	146	3
	MG-3/4	35	282		MG-4Z	4	60
Correct Classification		82.3%	87.9%	Correct Classification		97.3%	95.2%
(C) MG-1 vs. MG-2				**(F) MG-4 vs. MG-4A**			
Misclassification		True type		Misclassification		True type	
7.58%		MG-1	MG-2	12.00%		MG-4	MG-4Z
Predicted type	MG-1	48	7	Predicted type	MG-4	201	22
	MG-2	8	135		MG-4Z	88	61
Correct Classification		85.7%	95.1%	Correct Classification		89.7%	84.9

Table 7 Classification results for the overall and UCZ discrimination of chromitite seams. Please note in (A) and (C) Dataset A, while in (B) and (D) Dataset B was used. For details see text.

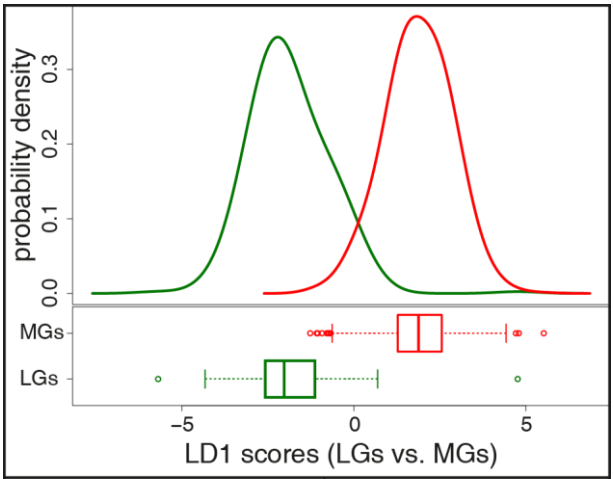
(A) Overall success Dataset A									
Misclassification		True Type							
19.4%		LG6	LG6A	MG1	MG2	MG3	MG4	MG4A	MG4Z
Predicted	LG6	211	2	1	0	0	0	0	0
Type	LG6A	13	29	10	1	0	1	2	0
	MG1	1	4	37	5	1	3	2	2
	MG2	0	0	7	103	9	14	3	4
	MG3	1	0	1	12	92	6	1	2
	MG4	0	0	0	10	5	65	6	2
	MG4A	0	0	0	7	0	6	39	0
	MG4Z	0	0	0	4	1	2	0	53
Correct Classification		93.4	82.9	66.1	72.5	85.2	67.0	73.6	84.1
(B) Overall success Dataset B									
Misclassification		True Type							
34.2%		LG6	LG6A	MG1	MG2	MG3	MG4	MG4A	MG4Z
Predicted	LG6	210	2	1	0	1	1	5	0
Type	LG6A	13	25	11	3	0	5	1	0
	MG1	2	8	35	5	3	15	9	2
	MG2	0	0	7	105	25	30	6	8
	MG3	0	0	1	9	157	30	1	5
	MG4	0	0	1	12	10	96	16	4
	MG4A	0	0	0	4	4	73	44	0
	MG4Z	1	0	0	4	2	39	1	58
Correct Classification		92.9	71.4	62.5	73.9	77.7	33.2	53.0	75.3

(C) UCZ success Dataset A						**(D) UCZ success Dataset B**				
Misclassification		True Type				Misclassification			True Type	
13.4%		MG3	MG4	MG4A	MG4Z	34.6%	MG3	MG4	MG4A	MG4Z
Predicted	MG3	98	8	1	3	MG3	177	36	3	7
Type	MG4	9	78	7	2	MG4	15	124	19	5
	MG4A	0	8	44	0	MG4A	7	83	60	0
	MG4Z	1	3	1	58	MG4Z	3	46	1	65
Correct Classification		90.7	80.4	83.0	92.1		87.6	42.9	72.3	84.4

Appendix A

Classification results of Dataset B

The results of the cross-validation are in an excellent agreement between true and predicted seams with an overall misclassification rate of 3.1 %, i.e., the rate of correct classification is far above 90%. Very similar misclassification rates (3.5%) were achieved using a naive reclassification, i.e. without the cross-validation computational burden (Figure A1).



Missclassification 3.51%		True type LG	MG
Predicted type	LG	250	28
	MG	11	821
Correct Classification		95.79%	96.70%

Figure A1 Probability densities and boxplots of the scores of the linear discriminant function for lower group chromitites (LGs, green) against middle group chromitites (MGs, red). Classification results for the two-way discrimination are shown in the lower part.

The second classification step now needs to discriminate either LGs or MGs from one another. Figure A2A displays the probability density estimates and boxplots of the scores of LD1 by group, showing very well separation of the LG-6 from the LG-6A chromitite. Classification results are displayed in Table A1A, confirming the very good agreement between true and predicted chromitite seams with a correct classification rate significantly higher than 90 %. Within this second step MG-1 and MG-2 were successfully discriminated from MGs-3/4/4A/4Zero (Figure A2B). However, the LDA does not deliver such a clear result as the previous steps, which is also reflected in the classification results, i.e., the rate of misclassification is about 16 % (Table A1B). The third classification step is again twofold, discriminating MG-1 from MG-2 and MG-3 from MG-4/4A/4Zero. MG-1 and MG-2 display well-separated probability density plots and boxplots (Figure A2C) with misclassification rates below 10 % (Table A1C). Discrimination of MG-3 from MG-4/4A/4Zero shows less clear results with below 90 % correct classification (Figure A2D, Table A1D). The final two classification steps are used to discriminate between MG-4 sub-seams, firstly MG-4/4A from MG-4Zero and secondly MG-4 from MG-4A. While separating MG-4Zero works quite well (12 % misclassification; Figure A2E, Table A1E),

a significant overlap in probability density plots (Figure A2F) and some significant misclassification was recorded between the MG-4 and MG-4A, i.e., almost 30 % (Table A1F).

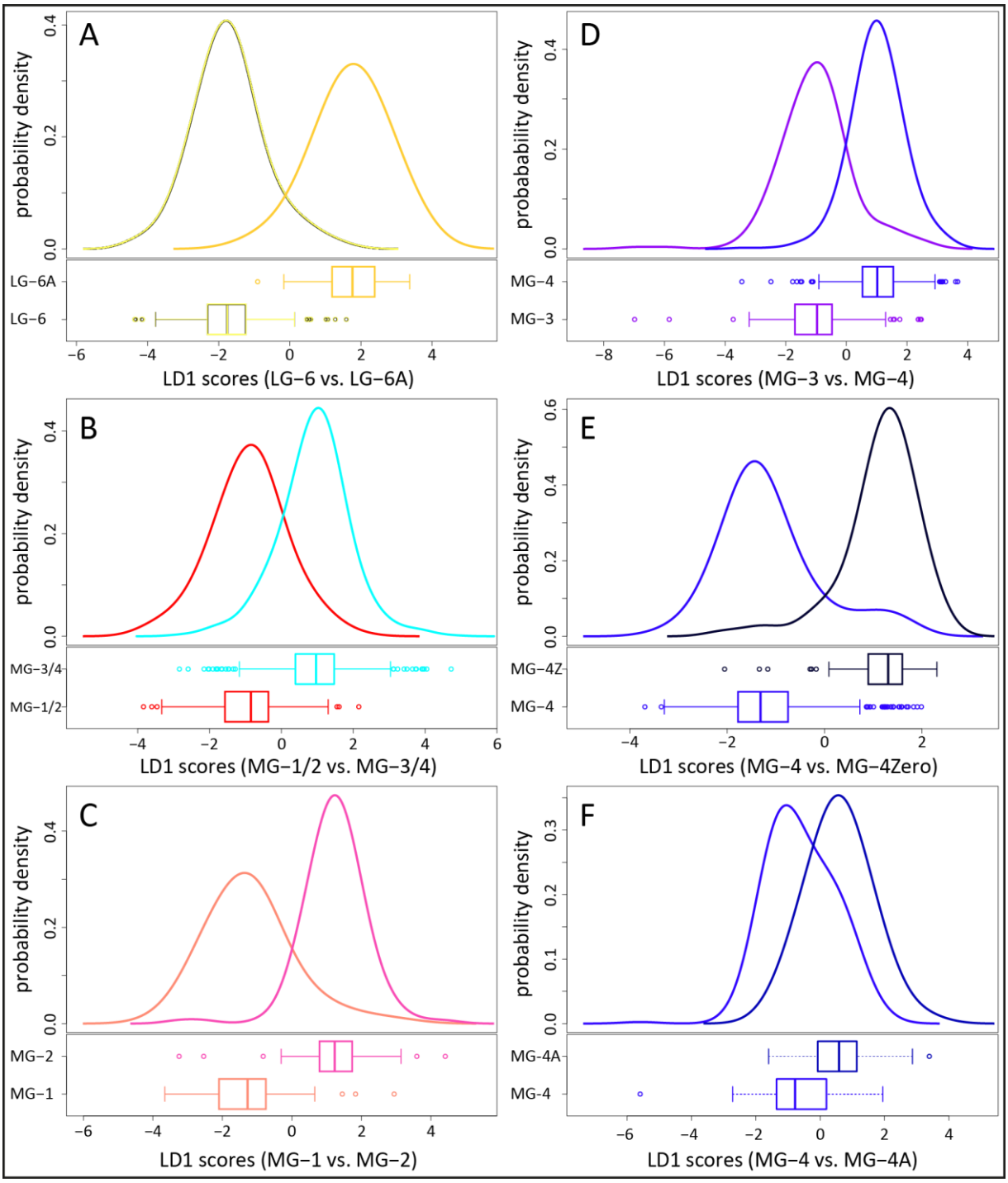


Figure A2 Probability densities and boxplots of the scores of the linear discriminant functions for lower and middle group chromitites following the classification scheme outlined in Figure 4. (A) LG-6 (yellow) vs. LG-6A (orange). (B) MG-1/2 (red) vs. MG-3/4 (turquoise). (C) MG-1 (pale red) vs. MG-2 (pink). (D) MG-3 (purple) vs. MG-4 (light blue). MG-4 contains data of MG-4/4A/4Zero. (E) MG-4 (light blue) vs. MG-4Zero (dark blue). MG-4 contains data of MG-4/4A. (F) MG-4 (light blue) vs. MG-4A (blue). MG-4Z and MG-4Zero are used synonymously.

Table A1 Classification results of Dataset B for the discrimination of chromitite seams. Please note in (D) MG-4 contains data of MG-4/4A/4Zero and in (E) MG-4 contains data of MG-4/4A.

(A) LG-6 vs. LG-6A				**(D) MG-3 vs. MG-4**			
Miss Classification		True type		Miss Classification		True type	
6.13%		LG-6	LG-6A	10.91%		MG-3	MG-4
Predicted type	LG-6	212	2	Predicted type	MG-3	177	46
	LG-6A	14	33		MG-4	25	403
Correct Classification		93.81%	94.29%	Correct Classification		87.62%	89.76%
(B) MG-1/2 vs. MG-3/4				**(E) MG-4 vs. MG-4Z**			
Miss Classification		True type		Miss Classification		True type	
16.02%		MG-1/2	MG-3/4	12.03%		MG-4	MG-4Z
Predicted type	MG-1/2	166	104	Predicted type	MG-4	324	6
	MG-3/4	32	547		MG-4Z	48	71
Correct Classification		83.84%	84.02%	Correct Classification		87.10%	92.21%
(C) MG-1 vs. MG-2				**(F) MG-4 vs. MG-4A**			
Miss Classification		True type		Miss Classification		True type	
7.58%		MG-1	MG-2	29.57%		MG-4	MG-4Z
Predicted type	MG-1	48	7	Predicted type	MG-4	201	22
	MG-2	8	135		MG-4Z	88	61
Correct Classification		85.71%	95.07%	Correct Classification		69.55%	73.49%