Geostatistical Simulation of Geochemical Compositions in the Presence of Multiple Geological Units: Application to Mineral Resource Evaluation

Talebi, H.; Mueller, U.; Tolosana Delgado, R.; van den Boogaart, K. G.;

Originally published:

August 2018

Mathematical Geosciences 51(2018)2, 129-153

DOI: https://doi.org/10.1007/s11004-018-9763-9

Geostatistical Simulation of Geochemical Compositions in the Presence of Multiple Geological Units - Application to Mineral Resource Evaluation

Hassan Talebi[1], Ute Mueller[1], Raimon Tolosana-Delgado[2], K. Gerald van den Boogaart[2]

[1]School of Science, Edith Cowan University, 270 Joondalup Drive, Joondalup, WA 6027,
Australia, htalebi@our.ecu.edu.au, u.mueller@ecu.edu.au

[2]Helmholtz Zentrum Dresden-Rossendorf, Helmholtz Institute Freiberg for Resources
Technology, Chemnitzerstrasse 40, D-09599 Freiberg (Saxony), Germany,

r.tolosana@hzdr.de, boogaart@hzdr.de

Abstract

An accurate prediction of benefit in ore deposits with heterogeneous spatial variations requires the definition of geological domains that differentiate the types of mineralogy, alteration, and lithology as well as the prediction of full mineral and geochemical compositions within each modelled domain and across boundaries between different domains. This paper proposes and compares various approaches (different combination of log-ratio transformation, gaussian and flow anamorphosis, and deterministic or probabilistic geological models) for geostatistical simulation of geochemical compositions in the presence of several geological domains.

Different approaches are illustrated through an application to a nickel-cobalt laterite deposit located in Western Australia. Four rock types (Ferruginous, Smectite, Saprolite, and Ultramafic) are considered to define compositionally homogeneous domains. Geochemical compositions are comprised of six different components of interest (Fe, Al, Mg, Ni, Co, Filler). Results suggest that the flow anamorphosis is a vital element for geostatistical modeling of geochemical composition due to its invariance properties and capability for reproducing complex patterns in input data including: presence of outliers, presence of several populations (due to the presence of several geological domains), nonlinearity, and heteroscedasticity.

Keywords: Compositional data, Log-ratio, Flow Anamorphosis, Geostatistical simulation, Geological domaining

1 Introduction

Ore deposits usually consist of ore materials with different characteristics. In order to maximize revenue in a mining project, a decision must be made regarding what processing plants are needed and the processing destinations of ore materials with different characteristics. For a better classification of ore materials, features such as rock type, alteration, microstructure, geochemical and mineral composition must be measured at sample locations and be predicted at the block model of the target deposit. These features normally have complex statistical and spatial relationships which should be reproduced in the predicted models (Boisvert et al. 2013; Maleki et al. 2016; Mery et al. 2017; Montoya et al. 2012; Mueller et al. 2014; Talebi et al. 2017; Tolosana-Delgado et al. 2014; Tolosana-Delgado et al. 2015; van den Boogaart et al.

2014). Among the different features of ore materials, mineral and geochemical compositions have a great impact on the final destination of materials and performance of processing plants. However, the compositional nature of these data induces several challenges for multivariate geostatistical techniques to predict them at the block support (Tolosana-Delgado et al. 2014; Tolosana-Delgado et al. 2015). Compositional data are multivariate, non-negative values which represent the importance of some parts of a whole. In such data, the constant sum constraint forces at least one covariance to be negative and induces spurious correlations. Furthermore, they carry just relative information (Aitchison 1986). To transform compositional data into unbounded space and to increase mathematical tractability, different log-ratio transformations can be applied prior to using standard (geo)statistical techniques (McKinley et al. 2016; Pawlowsky-Glahn and Egozcue 2016; Pawlowsky-Glahn et al. 2015; Pawlowsky-Glahn and Olea 2004; Tolosana-Delgado and van den Boogaart 2013; van den Boogaart and Tolosana-Delgado 2013). Most of the multivariate geostatistical simulation techniques are based on the assumption of multivariate multigaussianity of the data. In real case studies, log-ratio transformed data alone do not ensure this assumption, so the log-ratio transform has to be combined with a normal score transform prior to using geostatistical simulation techniques in order to not violate the assumptions of multigaussianity (Chilès and Delfiner 2012; Mueller et al. 2014). However, compositional data do not have a unique, canonical representation and several log-ratio transformations are available, making invariance of the simulated results under the choice of log-ratio transform desirable. Normal score transformations based on quantile matching do not have the invariance property and in addition the transformed data might not be multivariate normal. Flow anamorphosis (FA) has been designed to address these challenges, which means, it is capable of transforming original multivariate data to multivariate normal space and at the same time being invariant under the choice of log-ratio transform (Mueller et al. 2017; van den Boogaart et al. 2017).

On the other hand, the heterogeneity of geological units in an ore deposit requires defining domains that differentiate various characteristics such as: types of mineralogy, alteration, lithology, and microstructures. Indeed, these geological domains control the ore characteristics. One can think of deposits in which the mean grades and patterns of spatial continuity depend upon the rock and/or alteration type. For example a porphyry copper deposit with high grades in potassic alterations surrounded by low grades in argillic and propylitic alterations (Talebi et al. 2013), presence of lately injected barren dykes of different sizes and orientations (Talebi et al. 2014), and spatial modelling of geological units in an uranium roll-front deposit (Renard and Beucher 2012). Currently, the most prevalent approach to model the uncertainty in the spatial distribution of the elements of interest is to divide the study area into subdomains (geological units) based on geological interpretation and to predict the variables of interest within each domain separately. This approach defines just one interpretation of the geological domains and does not offer any measure of the uncertainty in the position of the domain boundaries. This uncertainty can be evaluated by use of geostatistical simulation methods for categorical variables (Alabert 1987b; Armstrong et al. 2011; Mariethoz and Caers 2015).

The objective of this contribution is to compare different approaches for geostatistical simulation of geochemical compositions to assess mineral resources in a nickel-cobalt laterite deposit. To evaluate the effect of geological domaining on the accuracy of the predicted geochemical compositions, three approaches are analysed: geological controls are ignored, a deterministic geological model is applied, and a probabilistic geological model is used (Talebi et al. 2016). The probabilistic geological model is calculated based on a plurigaussian (PGS) model (Armstrong et al. 2011; Emery 2007). An isometric log-ratio transformation (ilr) is used to transform compositional data from the simplex to real space (Egozcue et al. 2003). The ilr-transformed data are transformed to normal scores and subsequently simulation is used to generate realisations at unsampled locations. Classical gaussian anamorphosis (GA) and flow anamorphosis are compared based on the accuracy of the predicted compositions, as well as their capability to reproduce complex statistical and spatial patterns present in the input data. The various approaches considered are illustrated through an application to a nickel-cobalt laterite deposit and their performances are evaluated against a set of validation boreholes.

The paper is organised as follows: In Sect 2 the basics of the compositional data analysis are covered. Various log-ratio transformations used in this paper and transformation to multivariate normal space via flow anamorphosis are presented in this section. A new method for adjusting global proportion of geological domains is proposed in Sect 2.3. Various approaches for geostatistical simulation of regionalised compositions in the presence of several geological domains are presented in Sect 2.4. In Sect 3 the case study (Murrin Murrin nickel-cobalt laterite deposit) is introduced and the dataset is presented. A compositional contact analysis is implemented in Sect 3.3. Section 3.4 presents the process of generating deterministic and probabilistic geological models via a plurigaussian simulation approach. In Sect 4 the results are presented and proposed methods are compared to each other based on several criteria. Finally, some conclusions and final thoughts are presented in Sect 5.

2 Methodology

2.1 Compositional Data Analysis

Compositional data are multivariate data where the non-negative components are measured on the same scale and are constrained by a constant sum property, usually 100%. Geochemical and mineral compositions and proportions of various rock types or alteration types in a block are typical examples of compositional data in an ore deposit. The compositional space is a D-dimensional simplex

$$S^D = \left\{ \vec{Z}(u_\alpha) = [z_1(u_\alpha), z_2(u_\alpha), \cdots, z_D(u_\alpha)] \;\middle|\; z_i(u_\alpha) \geq 0;\ i = 1,2,\cdots,D;\ u_\alpha \in \boldsymbol{A}\,;\ \sum_{i=1}^{D} z_i(u_\alpha) = m \right\}, \tag{1}$$

where $z_i(u_\alpha)$ represents the i^{th} component measured at location u_α within the study area $\boldsymbol{A}$. The number m is the constant sum and common values are 1 (proportions), 100 (percentages), 10^6 (ppm), and 10^9 (ppb). The constant sum constraint is known to induce problems of spurious correlation (see Aitchison (1982), for a detailed report). Compositional data carry just relative information, which is appropriately represented by taking log-ratio transformations (Aitchison, 1986). Pairwise log-ratio transformation (Aitchison 1986), centred log-ratio transformation (Aitchison 1986), and isometric log-ratio transformation (Egozcue et al. 2003) are utilized in this study. Independently of which transform used, the resulting log-ratio scores happen to be free of the constraints of positivity and constant sum or of the spurious correlation problem, which make log-ratio scores more amenable to (geo)statistical treatment. The pairwise log-ratio transformation (pwlr) is defined as follows

$$pwlr\left(\vec{Z}(u_\alpha)\right) = \begin{bmatrix} 0 & ln\left(\dfrac{z_1(u_\alpha)}{z_2(u_\alpha)}\right) & \cdots & ln\left(\dfrac{z_1(u_\alpha)}{z_D(u_\alpha)}\right) \\ ln\left(\dfrac{z_2(u_\alpha)}{z_1(u_\alpha)}\right) & 0 & \ldots & ln\left(\dfrac{z_2(u_\alpha)}{z_D(u_\alpha)}\right) \\ \vdots & \vdots & \ddots & \\ ln\left(\dfrac{z_D(u_\alpha)}{z_1(u_\alpha)}\right) & ln\left(\dfrac{z_D(u_\alpha)}{z_2(u_\alpha)}\right) & & 0 \end{bmatrix}. \tag{2}$$

The centred log-ratio transformation (clr), is calculated via the following formula

$$clr\left(\vec{Z}(u_\alpha)\right) = ln\left(\frac{\vec{Z}(u_\alpha)}{\sqrt[D]{\prod_{i=1}^{D} z_i(u_\alpha)}}\right). \tag{3}$$

Finally, ilr transformation is defined as follows

$$ilr\left(\vec{Z}(u_\alpha)\right) = V \cdot clr\left(\vec{Z}(u_\alpha)\right), \tag{4}$$

where V is a $(D-1) \times D$ matrix whose columns are pairwise orthogonal vectors and sums to zero. Each matrix V satisfying these conditions give rise to an ilr transformation. There are infinitely many ilr transformations and often it is recommended to select one that increases the interpretability of the ilr scores. However, the aim of this study is to predict the geochemical compositions and model the associated uncertainty at unsampled locations accurately. Hence, a default orthonormal basis is used (Egozcue et al. 2003). The predicted compositions do not depend on the actual log-ratio transformation implemented for the computations (Tolosana-Delgado 2006).

2.2 Flow Anamorphosis

The flow anamorphosis introduced by van den Boogaart et al. (2017) is a multivariate form of gaussian anamorphosis. In this method a kernel density estimate is deformed from the given multivariate density of the observations into the density of a standard multivariate normal distribution. Several statistical tests of multivariate normality can be applied to check the fit of the transformed data to normality (Korkmaz et al. 2014; Mardia 1970; Székely and Rizzo 2005; Székely and Rizzo 2013). The fit is highly dependent on the selection of the two parameters of FA, σ_0 and σ_1 (initial and final spreads of the smoothing kernels of the kernel density estimates). Deformation of the underlying space is controlled by σ_0. The smaller the value, the closer the transformed data are to multivariate normality. Selecting a proper σ_0 is dependent on the number of variables (D), sample size, and complexity of the input data (Mueller et al. 2017). On the other hand, σ_1 controls the ranges of the transformed distributions. In this study, in order to force the marginal distributions of transformed data to have standard deviations close to unity, σ_1 is selected as $\sigma_1 = \sigma_0 + 1$ (Mueller et al. 2017). Figure 1 shows the process of simulation via FA.

[Fig. 1 about here.]

Figure 1(a) depicts the scatter plot and marginal kernel smoothing density estimates of the input data with two simulated variables whose relationship is complex. The characteristics include the presence of some outliers, nonlinear relationships, several populations, and heteroscedasticity. Figure 1(b) and Fig. 1(c) show the co-deformation of the underlying space and final distribution of the transformed data respectively. The transformed data in multivariate normal space can be simulated (Fig. 1(d)) via many geostatistical algorithms (Alabert 1987a; Deutsch and Journel 1998; Emery 2008; Emery et al. 2016; Emery and Lantuéjoul 2006). Several experiences have shown that the FA-transformed data are not only multivariate normal but often also exhibiting lack of spatial cross-correlation which make the geostatistical simulation of such orthogonal factors, straightforward (Mueller et al. 2017; van den Boogaart et al. 2017). Otherwise, multivariate simulation or spatial decorrelation technique followed by univariate simulation could be implemented. The simulated results are back-transformed to the original space via FA^{-1}. Figure 1(e) and Fig. 1(f) show the co-deformation back from simulated multivariate normal space to the original space and final distribution of the simulated data respectively. A visual comparison of these two plots shows that all the characteristics of the input the data are very well reproduced.

2.3 Geological Domaining

By constructing multiple spatial realisations of the geological domains, geostatistical simulation helps to improve the geological interpretation and to measure associated uncertainty. Several geostatistical methods for categorical variables can be used to this end. In the field of two-point geostatistics, PGS has gained popularity and proved to be suitable for reproducing complex configurations of geological domains without the need to define a training image. After simulating K geological domains at each location u_α in the study area $\boldsymbol{A}$, local proportions of the simulated domains, $(q_1(u_\alpha), q_2(u_\alpha), \dots, q_K(u_\alpha))$, can be calculated from the realisations. By assigning the geological domain $k, (k = 1,2, \cdots, K)$, with the highest proportion to each location u_α the most probable map is obtained. In this study, K spatial proportion maps (associated with K geological domains) and the most probable map are used as probabilistic and deterministic geological models respectively. It is known that the global proportions $\vec{Q} = (q_1, q_2, \dots, q_K)$ of simulated domains in the most probable map might be different from the global proportions $\vec{P} = (p_1, p_2, \dots, p_K)$ of the domains in the input data. Some economically important domains might be under-represented. To reduce under-representation of geological domains in the simulated geological model, it is common to post-process the proportions. One such method is the Soares correction (Soares 1998) which restores the global proportions of the input data, but is known to lead to artefacts and does not take account of the location of the datum to be adjusted. The correction method proposed here is also based on the proportions of domains in the input data, but takes location into account by borrowing a technique from compositional data analysis. The local proportions at each location u_α in the simulated geological model are perturbed by putting

$$\vec{r}(u_\alpha) = \big(r_1(u_\alpha), r_2(u_\alpha), \dots, r_K(u_\alpha)\big) = \left(\frac{b_1 q_1(u_\alpha)}{R(u_\alpha)}, \frac{b_2 q_2(u_\alpha)}{R(u_\alpha)}, \dots, \frac{b_K q_K(u_\alpha)}{R(u_\alpha)}\right). \qquad (5)$$

Here $\vec{r}(u_\alpha)$ is the vector of perturbed proportions at location u_α, $R(u_\alpha) = \sum_{k=1}^{K} b_k\, q_k(u_\alpha)$, and finally $(b_1, b_2, \dots, b_K) = \left(\frac{p_1}{q_1}, \frac{p_2}{q_2}, \dots, \frac{p_K}{q_K}\right)$. The resulting new local probabilities $\big(r_1(u_\alpha), r_2(u_\alpha), \dots, r_K(u_\alpha)\big)$, form the new probabilistic model of geology and can be used to determine the adjusted most probable map of geological domains.

2.4 Approaches to Geostatistical Simulation of Compositional Data

In this study several scenarios are investigated to assess the effects of geological models and transformation to normal space on the geostatistical simulation of geochemical compositions. Selected approaches to incorporate geological information are as follows: prediction without geological control, prediction by using a deterministic geological model, and prediction by using a probabilistic geological model. GA (Wackernagel 2003) and FA algorithms are compared to assess the effect of transformation to normal space and subsequent back-transformation on the spatial simulation of regionalised compositions.

In the first (**M_0**) and second (**M_1**) proposed methods there is no geological control (Table 1). In these two scenarios all input compositions (without considering a geological domain partition of the deposit) are transformed to real space via an ilr transformation and subsequently to normal space via GA and FA, respectively. If the normal scores are spatially correlated, they are transformed to spatial orthogonal factors via Min/Max autocorrelation factors (MAF), (Bandarian et al. 2008; Desbarats and Dimitrakopoulos 2000; Rondon 2012; Switzer and Green 1984). This orthogonalization makes the simulation step more straightforward. However, normal scores obtained via FA are normally spatially orthogonal and independent (Mueller et al. 2017; van den Boogaart et al. 2017). The normal scores (or MAFs) are simulated independently at the simulation grid (in this case all the locations of the validation data) via turning bands (TB) technique (Emery and Lantuéjoul 2006). The simulated results are back-transformed to the original space afterwards.

In the simulation via a deterministic geological model (**M**$_2$) input compositions are divided into several subsets based on their associated geological domains. Compositions in each subset are transformed to ilr space and subsequently transformed to multivariate normal space via FA. The simulation grid is also divided into mutually exclusive and exhaustive domains based on the most probable rock types (deterministic geological model) achieved by a PGS model. Normal scores from each subset are simulated at the associated part of the simulation grid. Finally simulated results at each part of the simulation grid are back-transformed to the original space independently.

The last proposed method (**M**$_3$) is geostatistical simulation of geochemical compositions using a probabilistic geological model. In this case input compositions are divided into subsets based on their associated geological domain. Compositions in each subset are transformed to ilr space and subsequently transformed to multivariate normal space via FA. Normal scores from each subset are simulated on the entire simulation grid and back-transformed to the original space independently. This process provides several sets of simulated geochemical compositions associated with the geological domains. Final simulated compositions can be obtained via weighting the simulated compositions associated with the different domains by the local probabilities of occurrence of each domain (Emery and González 2007a; Emery and González 2007b; Talebi et al. 2015)

$$\vec{Z}(u_\alpha) = \sum_{k=1}^{K} q_k(u_\alpha)\, \vec{Z}^k(u_\alpha), \tag{6}$$

where K is the number of classes (geological domains), $q_k(u_\alpha)$ is the probability of geological domain k at the location u_α calculated by a PGS model, and $\vec{Z}^k$ is the simulated composition associated with geological domain k. The final simulated vector $\vec{Z}(u_\alpha)$ is still a composition because

$$\sum_{i=1}^{D} z_i(u_\alpha) = \sum_{i=1}^{D}\sum_{k=1}^{K} q_k(u_\alpha)\, z_i^k(u_\alpha) = \sum_{k=1}^{K} q_k(u_\alpha) \sum_{i=1}^{D} z_i^k(u_\alpha) = \sum_{k=1}^{K} q_k(u_\alpha) m = m. \quad (7)$$

Finally methods based on a deterministic model of geology ($\mathbf{M_2}$) and a probabilistic model (M3) are repeated with the corrected deterministic ($\mathbf{M_{2c}}$) and probabilistic ($\mathbf{M_{3c}}$) models. The correction is based on the method described in Sect 2.3. Table 1 provides a summary of the proposed methods.

[Table 1 about here.]

3 Case Study: Murrin Murrin Nickel-Cobalt Laterite Deposit

Murrin Murrin East (MME) is a nickel-cobalt laterite deposit located in Western Australia, at about 60 km southeast of Laverton and at an average elevation of approximately 420 metres above mean sea level. The orebody is approximately 1,500 meters long, 600 meters wide, and 30 meters thick. Figure 2 shows a satellite image of the orebody together with the location of the boreholes.

[Fig. 2 about here.]

3.1 Geological Description

Laterite deposits are formed during chemical weathering of ultramafic rocks near the surface of the earth. At MME, nickel laterite deposits occur as laterally extensive, undulating blankets of mineralisation with strong vertical trends covering basement ultramafic rocks (Murphy 2003). Proximity of the deposit to a major salt lake probably influenced the development of the regolith (Markwell 2001). Based on the geochemical interpretation and the logging information from geologists, lateritic weathering of the ultramafic rocks at MME has produced a profile that may be broadly divided into four geological units (Camuti and Riel 1996; Markwell 2001; Monti and Fazakerley 1996). The sequence of these units from the bottom to the top is as follows: (i) ultramafic rocks (UM) at the base of weathering overlain by (ii) a Saprolite zone (SA) overlain by (iii) a Smectite zone (SM) and finally capped by (iv) a Ferruginous zone (FZ, Fig. 3(b)). **UM** occurs as a thin layer of unweathered ultramafic rock at the base of the deposit. Due to the undulating nature of different layers in this deposit, some outcrops of UM can be recognized at the surface of the deposit. **SA** consists mainly of Lizardite and Smectite. SA is enriched in Mg, but low amounts of Fe are present in this zone. A strong contrast between Mg in SM and SA allows easier domaining of the two units (Markwell 2001). **SM** consists mainly of Smectite and is confined to the shoulders of Saprolite domes. SM is enriched in Ni and Co (Fig. 3) and depleted in Mg. **FZ** is composed predominantly of Goethite, Kaolin, and Maghemite. FZ is less enriched in Ni, but it can host significant amounts of Co, especially at the transition to the Smectite zone (Fig. 3).

Table 2 shows the mineral assemblage in each geological unit at MME, obtained from several representative samples (Markwell 2001).

[Table 2 about here.]

3.2 Dataset

In total, 17,512 samples (of 1 meter length) from 926 RC holes (Fig. 2) make up the database for this study. Four rock types (FZ, SA, SM, and UM) are considered to define compositionally homogeneous domains. Three major (Fe, Al, and Mg) and two target (Ni and Co) elements are the variables of interest in this study. Since the data are compositional, a filler variable is introduced to achieve closure and to retain the intuitive relationship between each component and the mass of its associated element. The data set was subdivided into two subsets: 453 holes (8,694 samples) are considered for validation and 473 holes (8,818 samples) for geostatistical modelling, called input data from hereon. The input data are used for constructing the various geostatistical models. Locations of the validation and input boreholes can be seen in a cross section of the deposit for northing 300m (Fig. 3(a)). Figure 4 depicts the histograms of different geochemical components in the validation and input sets. Figure 5 shows centred ternary diagrams (Buccianti et al. 1999; Pawlowsky-Glahn et al. 2015) of the sub-compositions. In this case study geochemical compositions are composed of different populations. The multi-population character is consistent with the presence of four geological units (Fig. 4 and Fig. 5). These geological units control the statistical and spatial distribution of the geochemical compositions. Figure 6 shows the vertical curves of the rock type proportions and clr-transformed of the geochemical components.

There is a zoned weathering profile (Fig. 6), in which Mg as a mobile element effectively has been removed while Fe and Al, as less mobile elements, have been enriched residually in the Ferruginous zone. On the other hand, Co (one of the value elements), has been accumulated at the transition between Ferruginous and Saprolite zones (Talebi et al. 2017) where Smectite units are mainly located. Finally, the Smectite and Saprolite zones between fresh (ultramafic parent rock) and weathered zones (Ferruginous zones), have been enriched in Ni (the other value element). Figure 7 shows the scatterplots of the clr-transformed components. Complex relationships (nonlinearity, multi-population, and presence of outliers) can be recognised between different components. Based on the vertical curves (Fig. 6) and the scatterplots of the clr-transformed components (Fig. 7), Al-Fe, Ni-Co, and Mg-Ni are positively correlated. On the other hand, Al-Mg, Al-Ni, Co-Filler, Fe-Mg, Ni-Filler, are in terms of their clrs negatively correlated. From Fig. 4 to Fig. 7, it is clear that the validation and input sets are statistically and spatially similar.

[Fig. 3 about here.]

[Fig. 4 about here.]

[Fig. 5 about here.]

[Fig. 6 about here.]

[Fig. 7 about here.]

3.3 Compositional Contact Analysis

In order to evaluate the variation of the variables of interest at the transition zone between two geological domains, a contact analysis was conducted. To do this, statistics of the variables of interest such as means and standard deviations are plotted as functions of distance from the contact zone (Ortiz and Emery 2006). However, compositional data have spurious correlations and real variation in one component might cause apparent variations in others, so contact analysis of raw components should be avoided (Tolosana-Delgado et al. 2016). Although log-ratio transformed data can be treated as real data and compositional contact analysis can be implemented on the clr-transformed data, contact analysis of the pwlr-transformed data provides the geologists with an enriched view of the variations at the contact zone. Figure 8 shows the compositional contact analysis between the two dominant rock types, FZ and SA. Diagrams in the last row and last column show means and standard deviations as functions of the distance from the contact zone for the clr-transformed data and raw components respectively while the remaining diagrams show those of the pwlr-transformed data (row/column).

[Fig. 8 about here.]

A comparison of the contact diagrams of the raw components with the associated clr-transformed ones shows that the behaviour of the means at the transition zone has been exaggerated unrealistically in the raw contact analysis. Except for Filler which shows different behaviour in the clr plot, other components show similar behaviour in raw and clr analysis. Al, Fe, and Mg show abrupt transitions (variation of means as a function of distance) while the local variations are stable (standard deviations as a function of distance). All the diagrams of the Co (raw, clr, and pwlr) show an increase of the mean and the standard deviation (proportionality effect) at the transition zone. The reason for this behaviour would be the presence of Smectite units at the transition from Ferruginous to Saprolite zone, not recognised during logging (Smectite units are enriched in Co). Pairwise log-ratio of the Co/Ni and Al/Fe means and standard deviations are stable across geological units. The results suggest that these two geological domains (FZ and SA) can be considered as stationary for some subcompositions such as Al and Fe, but domaining would be necessary for other subcompositions or ratios, for example Mg and Ni. Similar assessments were conducted on the other contacts between the four geological domains and results supported the necessity of domaining. Given the results of compositional contact analysis and statistical analysis of compositional data in Sect 3.2, it is reasonable to partition the study area to four homogeneous geological units (four rock types) prior to geostatistical modelling.

3.4 Deterministic and Probabilistic Geological Models

The geological domains (rock types) for the validation data were simulated via a plurigaussian model (Armstrong et al. 2011). Overall 100 realisations were generated and used to determine the probability of occurrence for each rock type (Fig. 9(b) to Fig. 9(e)) and the most probable rock type (Fig. 9(f)). Proportions of the rock types in the most probable map (Table 3) show that, under-representation of the SM has occurred. Since SM units are highly mineralised in Co, prediction of these units with high accuracy is of great importance. The method described in Sect 2.3 was used to adjust the proportions of geological domains in the simulated models.

The last column in Table 3 shows the adjusted proportions in the most probable rock type map, demonstrating an improvement in the proportion of the SM units. Figure 9(g) to Fig. 9(j) show cross-sections of the validation boreholes, coloured by the adjusted probabilities, for northing 300m. A visual comparison of Fig. 9(c) and Fig. 9(h) indicates a clear improvement in the proportion of the SM units, more representative of reality (Fig. 9(a)).

[Table 3 about here.]

[Fig. 9 about here.]

4 Results and Discussion

Figure 10 shows the histograms and scatterplots (coloured by kernel density estimate) of the ilr-transformed input data. As the ilr-transformed data (in this study) are not multivariate normal (Fig. 10), a transformation to normal space is needed prior to geostatistical simulation.

Transformation of ilr transformed data into normal space in $\mathbf{M_0}$ was based on GA. Figure 11 shows the scatterplots of the normal scores obtained by GA. Although the marginal distributions are normal, the scores are not multivariate normal. An analysis of the omnidirectional experimental semivariograms and cross-variograms further revealed that these normal scores are spatially correlated, with Tercan's $\bar{\tau}$ and $\bar{\kappa}$ (Tercan 1999) equal to 0.1973 and 0.8147, respectively. To make the simulation step more straightforward, normal scores data were transformed to spatially orthogonal factors via MAF. MAF transformation improved spatial orthogonality with $\bar{\tau}$ and $\bar{\kappa}$ equal to 0.0782 and 0.9689, respectively. Orthogonal factors were simulated at the entire validation holes via the TB algorithm. The simulated results were back-transformed to the simplex afterwards to recover outputs in the original scale as percentages.

On the other hand, **M₁** transforms the ilr-transformed data to multivariate normal space via FA. Due to the complexity of the data and the number of the observations and variables, multivariate normality was not achieved by a single FA. Two successive FA with the same parameters ($\sigma_0 = 0.1$ and $\sigma_1 = 1.1$) were required to achieve multivariate normality. Figure 12 shows the normal scores are obviously close to multivariate normal and uncorrelated, so statistically independent. Spatial structural analysis (variography) showed further that the normal scores are spatially orthogonal, with Tercan's $\bar{\tau}$ and $\bar{\kappa}$ equal to 0.0656 and 0.9873, respectively, so they could be simulated independently. The scores were simulated at the validation holes independently via TB algorithm and back-transformed to composition afterward.

[Fig. 10 about here.]

[Fig. 11 about here.]

[Fig. 12 about here.]

For simulation based on a deterministic geological model input compositions were divided into four subsets based on their associated rock types (FZ, SM, SA, and UM). Compositions in each subset were transformed to ilr space and subsequently transformed to multivariate normal space via FA independently. The validation holes were also divided into four mutually exclusive and exhaustive parts based on the most probable rock types (Fig. 9(f)) and the corrected most probable rock types (Fig. 9(k)), provided the simulation grids for **M₂** and **M₂c** respectively.

Normal scores from each subset (for example, flow anamorphosed FZ data) were simulated at the associated part of the simulation grid (for example, locations in the validation holes with the most probable rock known as FZ). Finally simulated results for each domain were back-transformed to the original space independently.

Unlike the deterministic approaches ($\mathbf{M_2}$ and $\mathbf{M_{2c}}$), in the probabilistic approach normal scores from each subset of input compositions were simulated at the entire simulation grid. This process provided four sets of simulated geochemical compositions associated with the four geological domains (FZ, SM, SA, UM). Final simulated compositions obtained by weighting the simulated compositions associated with the different domains by the probabilities of occurrence of each domain (Eq. 6). Two sets of probabilities are available, the raw probabilities (Fig. 9(b) to Fig. 9(e)) obtained by the PGS model ($\mathbf{M_3}$) and adjusted probabilities ($\mathbf{M_{3c}}$) obtained by the proposed correction method in Sect 2.3 (Fig. 9(g) to Fig. 9(j)).

The simulated compositional models based on the six proposed approaches were validated against reality available at the validation boreholes (Fig. 3(a)). Figure 13 shows the global histogram reproduction of the six methods. Best realisations were achieved by $\mathbf{M_1}$ followed by $\mathbf{M_{2c}}$. Visual comparison of histogram reproduction of $\mathbf{M_2}$ and $\mathbf{M_{2c}}$ reveals that adjusting proportions (Sect 2.3) has improved the histogram reproduction especially for Co component. Worst reproductions were achieved by $\mathbf{M_0}$.

[Fig. 13 about here.]

To check the ability of the methods to reproduce the complex patterns in the simplex space, ternary diagrams were plotted and compared to each other (Fig. 14). Due to the averaging nature of $\mathbf{M_3}$ and $\mathbf{M_{3c}}$, different populations have been mixed up, which is not acceptable. On the other hand, $\mathbf{M_0}$ was not able to reproduce different populations and complex patterns inside the simplex space. Here, $\mathbf{M_1}$ generated the best results followed by $\mathbf{M_2}$ and $\mathbf{M_{2c}}$.

The smoothing effect of the $\mathbf{M_3}$ and $\mathbf{M_{3c}}$ is more obvious from the experimental variograms of the simulated models (Fig. 15). In Fig. 15 continuous black lines are input data, dashed black lines are validation data and grey lines are realisations. The sills of the variograms for one of the simulated component (Ni) are systematically less than the sills for input and validation data, representing a systematic reduction of the spatial variability. Adjusting probabilities of geological domains has improved the variogram reproduction for the deterministic approaches (compare $\mathbf{M_2}$ and $\mathbf{M_{2c}}$).

Risk quantification in grade-tonnage curves (GTC) and predicting true curves are critical for feasibility studies and capital investment in mining projects. The accuracy of the six methods was investigated based on their capability for reproducing input GTCs and predicting true (validation data in this study) GTCs. Figure 16 and Fig. 17 show the GTCs for Ni and Co components respectively. The smoothing effect of the $\mathbf{M_3}$ and $\mathbf{M_{3c}}$ is again clear from the GTCs of both Ni and Co. Comparing GTCs of $\mathbf{M_2}$ and $\mathbf{M_{2c}}$, an improvement can be recognised due to the implementation of the proposed technique for adjusting underrepresented domains. GTCs for Co component generated by $\mathbf{M_0}$ is not acceptable. $\mathbf{M_1}$ generated the most satisfactory results followed by $\mathbf{M_{2c}}$.

Finally, the six proposed methods were compared to check the presence of any systematic bias in the predicted proportions of samples above cut-offs (Fig. 18). In multi-element deposits, there are more than one target elements (in this study two target elements: Ni and Co). In Fig. 18, each cell of the maps is associated with two cut-offs (one for Ni and one for Co). For each cell the expected proportion of samples above the two cut-offs were calculated from all realisations and subsequently subtracted from the real associated proportion calculated from validation data. These maps can be compared to show the capability of the proposed methods for avoiding any serious bias (over or underestimations of ore tonnages). $\mathbf{M_0}$, $\mathbf{M_3}$, and $\mathbf{M_{3c}}$ show clear systematic over and underestimations in some range of cut-offs. In this regard, $\mathbf{M_1}$ generated the best (least bias) results, followed by $\mathbf{M_{2c}}$.

According to the aforementioned criteria, $\mathbf{M_1}$ outperformed other proposed techniques for geostatistical simulation of geochemical compositions. One reason is that $\mathbf{M_1}$ considers spatial correlations of geochemical components across geological boundaries. More important reason is the ability of geostatistical simulation via FA to reproduce multi-population characteristic of the input data which is consistent with the presence of several geological units.

[Fig. 14 about here.]

[Fig. 15 about here.]

[Fig. 16 about here.]

[Fig. 17 about here.]

[Fig. 18 about here.]

5 Conclusion

This paper compares various geostatistical approaches for simulation of geochemical compositions and their application to mineral resource evaluation. Results and several validation tests showed that the classical transformation to normal space (gaussian anamorphosis) is not capable of reproducing complex statistical patterns inside data and should be replaced with more advanced transformations. The method for modelling geochemical compositions based on a probabilistic geological model exhibits smoothing effects due the averaging nature of the algorithm. Although this method generates satisfactory results for kriged maps (Emery and González 2007a; Emery and González 2007b; Talebi et al. 2015), it should be avoided for simulation purposes. The proposed technique for adjusting underrepresented domains improved the result of simulation and should be used in the cases where there are important geological domains with small proportions such as SM in this case study. Flow anamorphosis is a vital element for geostatistical modelling of geochemical composition due to its invariance properties and capability for reproducing complex patterns in data such as: outliers, multi-population, nonlinearity, and heteroscedasticity. In the case study presented, a simulation involving a global flow anamorphosis without domaining was best capable of reproducing all performance targets (histograms, variograms, grade and tonnage curves). In the authors' opinion, this remarkable property might occur again in other settings in which domains emerge as the effect of chemical processes mostly involving the composition modelled.

Acknowledgments

The authors acknowledge financial support through DAAD-UA grant CodaBlockCoEstimation.

References

Aitchison J (1982) The Statistical Analysis of Compositional Data Journal of the Royal Statistical Society Series B (Methodological) 44:139-177

Aitchison J (1986) The statistical analysis of compositional data. Chapman & Hall Ltd, Alabert F (1987a) The practice of fast conditional simulations through the LU decomposition of the covariance matrix Mathematical Geology 19:369-386 doi:10.1007/bf00897191

Alabert F (1987b) Stochastic imaging of spatial distributions using hard and soft information. Master's thesis. Department of Applied Earth Sciences, Stanford University.

Armstrong M et al. (2011) Plurigaussian Simulations in Geosciences. Springer Berlin Heidelberg, Berlin, Heidelberg. doi:10.1007/978-3-642-19607-2_3

Bandarian EM, Bloom LM, Mueller UA (2008) Direct minimum/maximum autocorrelation factors within the framework of a two structure linear model of coregionalisation Computers & Geosciences 34:190-200 doi:https://doi.org/10.1016/j.cageo.2007.03.015

Boisvert JB, Rossi ME, Ehrig K, Deutsch CV (2013) Geometallurgical Modeling at Olympic Dam Mine, South Australia Mathematical Geosciences 45:901-925 doi:10.1007/s11004-013-9462-5

Buccianti A, Pawlowsky-Glahn V, Barceló-Vidal C, Jarauta-Bragulat E (1999) Visualization and modeling of natural trends in ternary diagrams: A Geochemical Case Study vols I and II, Tapir, Trondheim (N), pp 139–144

Camuti K, Riel R (1996) Mineralogy of the Murrin Murrin nickel laterites Australasian Institute of Mining and Metallurgy 6/96

Chilès J-P, Delfiner P (2012). In: Geostatistics: Modeling Spatial Uncertainty. John Wiley & Sons, Inc., p 699. doi:10.1002/9780470316993.fmatter

Desbarats AJ, Dimitrakopoulos R (2000) Geostatistical Simulation of Regionalized Pore-Size Distributions Using Min/Max Autocorrelation Factors Mathematical Geology 32:919-942 doi:10.1023/A:1007570402430

Deutsch CV, Journel AG (1998) GSLIB: Geostatistical Software Library and User's Guide. Oxford University Press. New York.

Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, Barceló-Vidal C (2003) Isometric Logratio Transformations for Compositional Data Analysis Mathematical Geology 35:279-300 doi:10.1023/a:1023818214614

Emery X (2007) Simulation of geological domains using the plurigaussian model: New developments and computer programs Computers & Geosciences 33:1189-1201 doi:https://doi.org/10.1016/j.cageo.2007.01.006

Emery X (2008) A turning bands program for conditional co-simulation of cross-correlated Gaussian random fields Computers & Geosciences 34:1850-1862 doi:https://doi.org/10.1016/j.cageo.2007.10.007

Emery X, Arroyo D, Porcu E (2016) An improved spectral turning-bands algorithm for simulating stationary vector Gaussian random fields Stochastic Environmental Research and Risk Assessment 30:1863-1873 doi:10.1007/s00477-015-1151-0

Emery X, González KE (2007a) Incorporating the uncertainty in geological boundaries into mineral resources evaluation. J. Geol. Soc. India 69(1), 29–38

Emery X, González KE (2007b) Probabilistic modelling of lithological domains and it application to resource evaluation Journal of the Southern African Institute of Mining and Metallurgy 107:803-809

Emery X, Lantuéjoul C (2006) TBSIM: A computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method Computers & Geosciences 32:1615-1628 doi:https://doi.org/10.1016/j.cageo.2006.03.001

Korkmaz S, Goksuluk D, Zararsiz G (2014) MVN: an R package for assessing multivariate normality R J 6:151–162

Maleki M, Emery X, Cáceres A, Ribeiro D, Cunha E (2016) Quantifying the uncertainty in the spatial layout of rock type domains in an iron ore deposit Computational Geosciences 20:1013-1028 doi:10.1007/s10596-016-9574-3

Mardia KV (1970) Measures of Multivariate Skewness and Kurtosis with Applications Biometrika 57:519-530 doi:10.2307/2334770

Mariethoz G, Caers J (2015) Multiple-Point Geostatistics: Stochastic Modeling with Training Images. John Wiley & Sons, Ltd. doi:DOI: 10.1002/9781118662953

Markwell T (2001) Murrin Murrin Ni/Co resource estimation: MME resource modelling. Anaconda Operations Pty Ltd. 18 pp.

McKinley JM et al. (2016) The single component geochemical map: Fact or fiction? Journal of Geochemical Exploration 162:16-28 doi:https://doi.org/10.1016/j.gexplo.2015.12.005

Mery N, Emery X, Cáceres A, Ribeiro D, Cunha E (2017) Geostatistical modeling of the geological uncertainty in an iron ore deposit Ore Geology Reviews 88:336-351 doi:https://doi.org/10.1016/j.oregeorev.2017.05.011

Monti R, Fazakerley VW (1996) The Murrin Murrin nickel cobalt project The Australasian Institute of Mining and Metallurgy:191-196

Montoya C, Emery X, Rubio E, Wiertz J (2012) Multivariate resource modelling for assessing uncertainty in mine design and mine planning Journal of the Southern African Institute of Mining and Metallurgy 112:353-363

Mueller U, Tolosana-Delgado R, van den Boogaart KG (2014) Approaches to the simulation of compositional data – a nickel-laterite comparative case study. Paper presented at the Orebody Modelling and Strategic Mine Planning Symposium 2014, Melbourne,

Mueller U, van den Boogaart KG, Tolosana-Delgado R (2017) A Truly Multivariate Normal Score Transform Based on Lagrangian Flow. In: Gómez-Hernández JJ, Rodrigo-Ilarri J, Rodrigo-Clavero ME, Cassiraga E, Vargas-Guzmán JA (eds) Geostatistics Valencia 2016. Springer International Publishing, Cham, pp 107-118. doi:10.1007/978-3-319-46819-8_7

Murphy M (2003) Geostatistical optimisation of sampling and estimation in a nickel laterite deposit. Edith Cowan University, MSc thesis (unpublished)

Ortiz JM, Emery X (2006) Geostatistical estimation of mineral resources with soft geological boundaries: a comparative study The South African Institute of Mining and Metallurgy 106:577–584

Pawlowsky-Glahn V, Egozcue JJ (2016) Spatial analysis of compositional data: A historical review Journal of Geochemical Exploration 164:28-32 doi:https://doi.org/10.1016/j.gexplo.2015.12.010

Pawlowsky-Glahn V, Egozcue JJ, Tolosana-Delgado R (2015) Modelling and Analysis of Compositional Data. John Wiley & Sons, Ltd. doi:10.1002/9781119003144.ch1

Pawlowsky-Glahn V, Olea RA (2004) Geostatistical Analysis of Compositional Data. Oxford University Press,

Renard D, Beucher H (2012) 3D representations of a uranium roll-front deposit Applied Earth Science 121:84-88 doi:10.1179/1743275812Y.0000000011

Rondon O (2012) Teaching Aid: Minimum/Maximum Autocorrelation Factors for Joint Simulation of Attributes Mathematical Geosciences 44:469-504 doi:10.1007/s11004-011-9329-6

Soares A (1998) Sequential Indicator Simulation with Correction for Local Probabilities Mathematical Geology 30:761-765 doi:10.1023/a:1022451504120

Switzer P, Green AA (1984) Min/max autocorrelation factors for multivariate spatial imaging. Technical Report No. 6, Department of Statistics, Stanford University, Stanford, CA, 14pp.

Székely GJ, Rizzo ML (2005) A new test for multivariate normality Journal of Multivariate Analysis 93:58-80 doi:https://doi.org/10.1016/j.jmva.2003.12.002

Székely GJ, Rizzo ML (2013) Energy statistics: A class of statistics based on distances Journal of Statistical Planning and Inference 143:1249-1272 doi:https://doi.org/10.1016/j.jspi.2013.03.018

Talebi H, Asghari O, Emery X (2013) Application of plurigaussian simulation to delineate the layout of alteration domains in Sungun copper deposit Central European Journal of Geosciences 5:514-522 doi:10.2478/s13533-012-0146-3

Talebi H, Asghari O, Emery X (2014) Simulation of the lately injected dykes in an Iranian porphyry copper deposit using the plurigaussian model Arabian Journal of Geosciences 7:2771-2780 doi:10.1007/s12517-013-0911-8

Talebi H, Asghari O, Emery X (2015) Stochastic rock type modeling in a porphyry copper deposit and its application to copper grade evaluation Journal of Geochemical Exploration 157:162-168 doi:https://doi.org/10.1016/j.gexplo.2015.06.010

Talebi H, Lo J, Mueller U (2017) A Hybrid Model for Joint Simulation of High-Dimensional Continuous and Categorical Variables. In: Gómez-Hernández JJ, Rodrigo-Ilarri J, Rodrigo-Clavero ME, Cassiraga E, Vargas-Guzmán JA (eds) Geostatistics Valencia 2016. Springer International Publishing, Cham, pp 415-430. doi:10.1007/978-3-319-46819-8_28

Talebi H, Sabeti EH, Azadi M, Emery X (2016) Risk quantification with combined use of lithological and grade simulations: Application to a porphyry copper deposit Ore Geology Reviews 75:42-51 doi:https://doi.org/10.1016/j.oregeorev.2015.12.007

Tercan AE (1999) Importance of Orthogonalization Algorithm in Modeling Conditional Distributions by Orthogonal Transformed Indicator Methods Mathematical Geology 31:155-173 doi:10.1023/a:1007557701073

Tolosana-Delgado R (2006) Geostatistics for constrained variables: positive data, compositions and probabilities. Application to environmental hazard monitoring. PhD thesis, University of Girona, Girona (Spain).

Tolosana-Delgado R, Mueller U, van den Boogaart KG (2016) Compositionally Compliant Contact Analysis. In: Raju NJ (ed) Geostatistical and Geospatial Approaches for the Characterization of Natural Resources in the Environment: Challenges, Processes and Strategies. Springer International Publishing, Cham, pp 11-14. doi:10.1007/978-3-319-18663-4_2

Tolosana-Delgado R, Mueller U, van den Boogaart KG, Ward C (2014) Compositional Block Cokriging. In: Pardo-Igúzquiza E, Guardiola-Albert C, Heredia J, Moreno-Merino L, Durán JJ, Vargas-Guzmán JA (eds) Mathematics of Planet Earth: Proceedings of the 15th Annual Conference of the International Association for Mathematical Geosciences. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 713-716. doi:10.1007/978-3-642-32408-6_154

Tolosana-Delgado R, Mueller U, van den Boogaart KG, Ward C, Gutzmer J (2015) Improving processing by adaption to conditional geostatistical simulation of block compositions Journal of the Southern African Institute of Mining and Metallurgy 115:13-26

Tolosana-Delgado R, van den Boogaart KG (2013) Joint Consistent Mapping of High-Dimensional Geochemical Surveys Mathematical Geosciences 45:983-1004 doi:10.1007/s11004-013-9485-y

van den Boogaart KG, Mueller U, Tolosana-Delgado R (2017) An Affine Equivariant Multivariate Normal Score Transform for Compositional Data Mathematical Geosciences 49:231-251 doi:10.1007/s11004-016-9645-y

van den Boogaart KG, Tolosana-Delgado R (2013) Analyzing Compositional Data with R. Springer-Verlag Berlin Heidelberg. doi:10.1007/978-3-642-36809-7

van den Boogaart KG, Tolosana-Delgado R, Lehmann M, Mueller U (2014) On the joint multipoint simulation of discrete and continuous geometallurgical parameters. Paper presented at the Proceedings Orebody Modelling and Strategic Mine Planning Symposium 2014, Melbourne,

Wackernagel H (2003) Gaussian Anamorphosis with Hermite Polynomials. In: Multivariate Geostatistics: An Introduction with Applications. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 238-249. doi:10.1007/978-3-662-05294-5_33

List of Figures

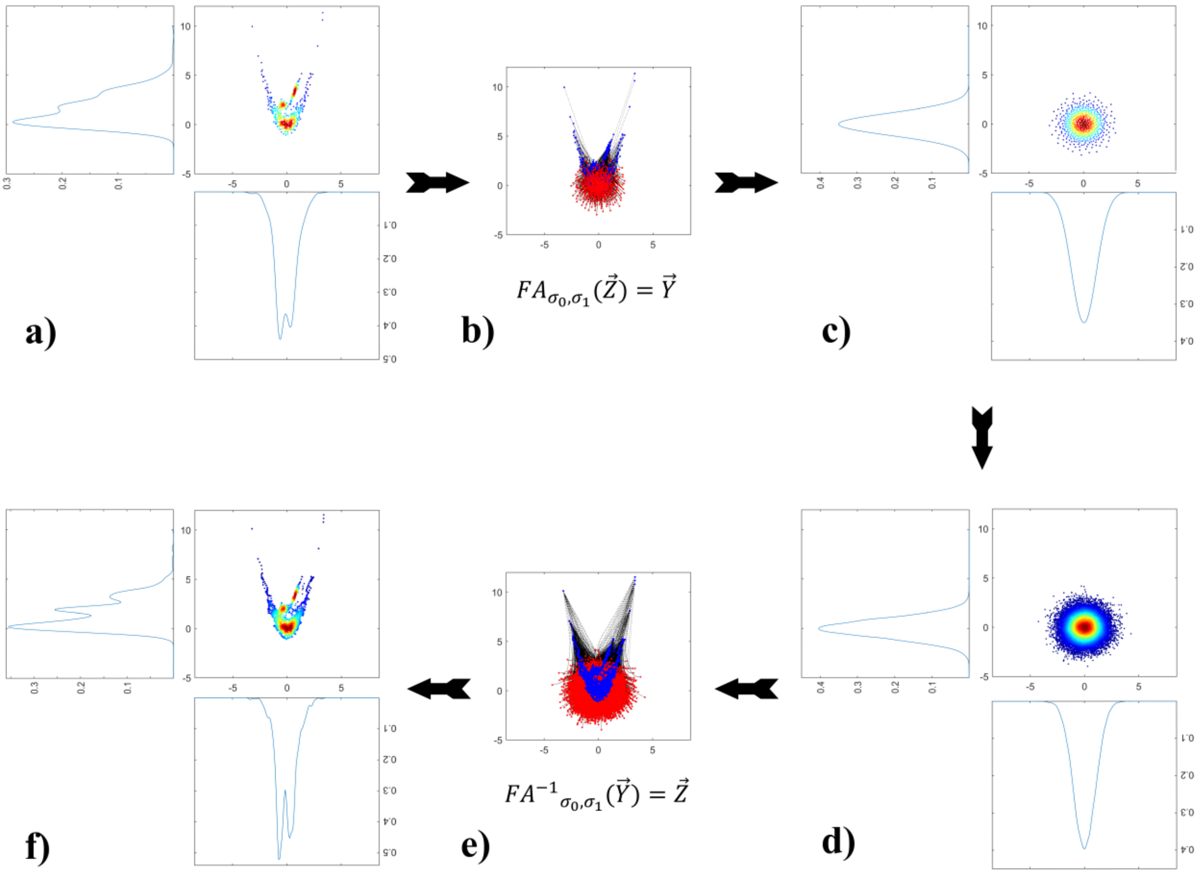

Fig. 1 Geostatistical simulation via Flow Anamorphosis

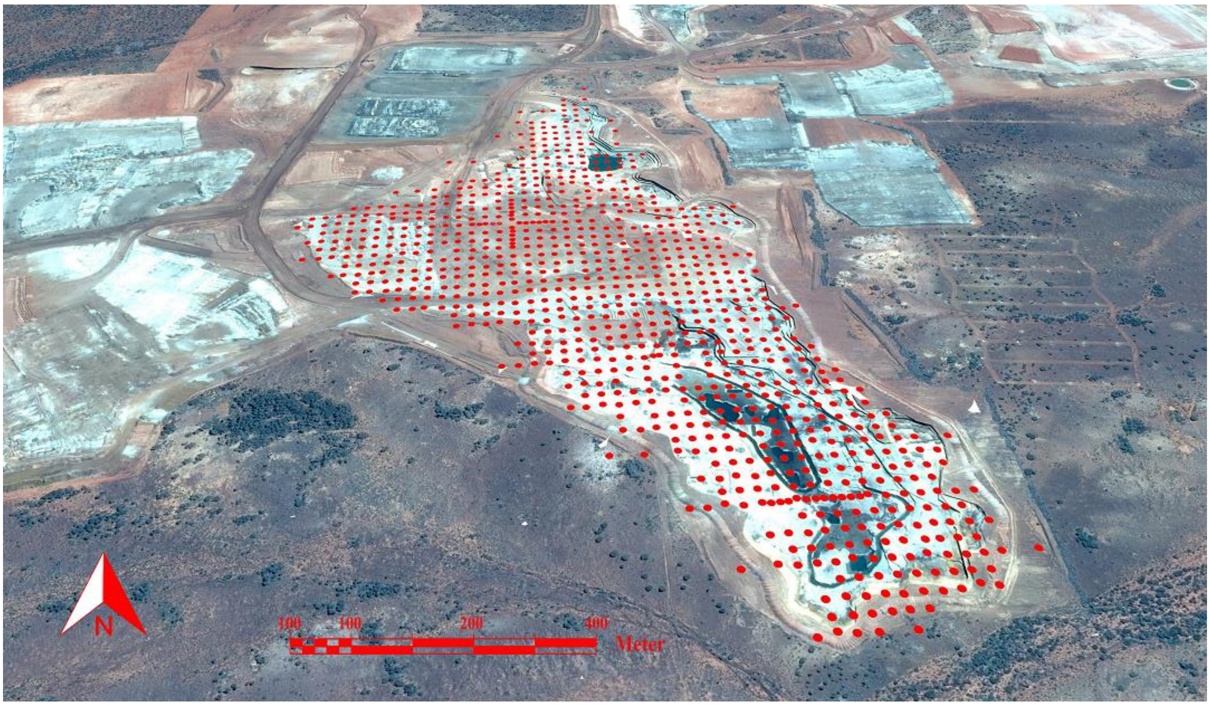

Fig. 2 Borehole location map of the Murrin Murrin East (MME)

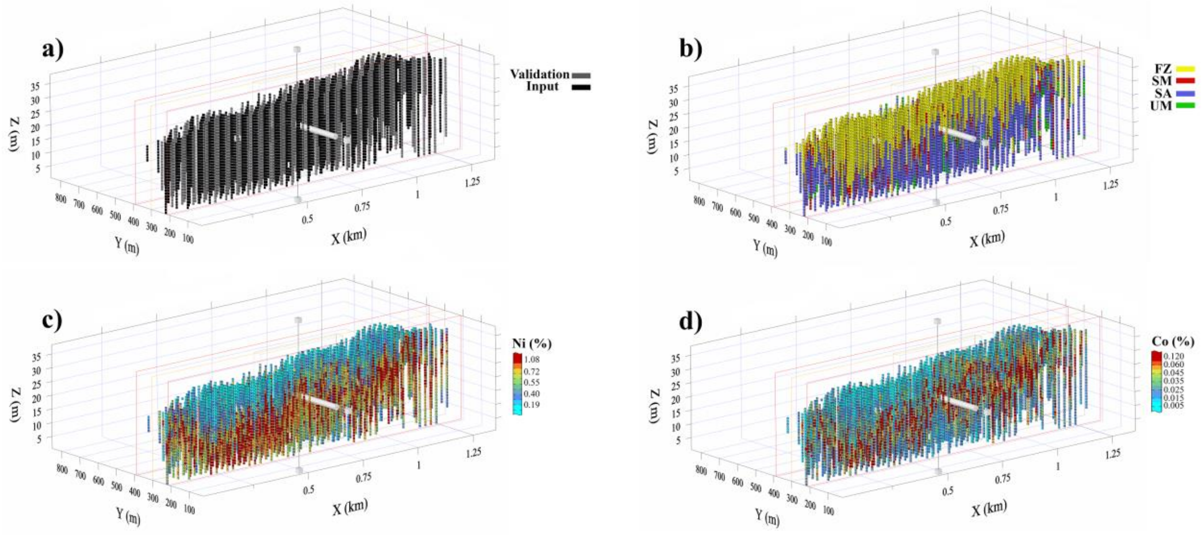

Fig. 3 Cross sections of boreholes for northing 300m and 100m thickness: locations of input and validation boreholes (a), spatial distributions of different rock types (b) Ni grade (c) and Co grade (d)

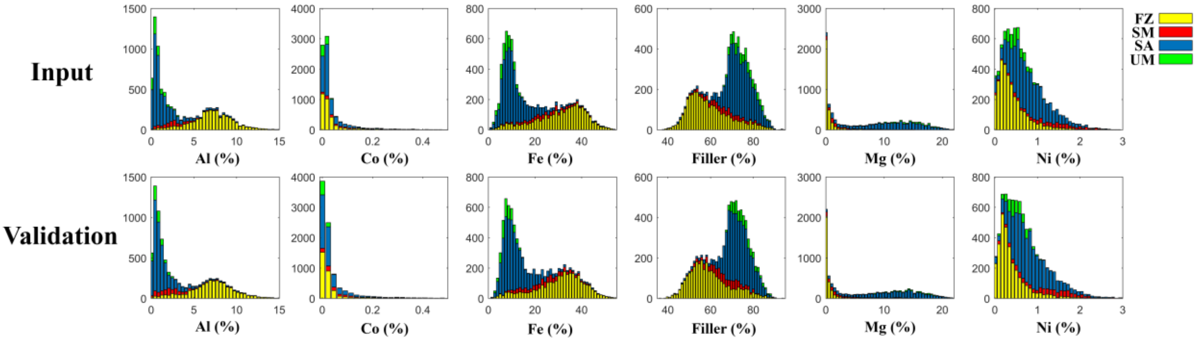

Fig. 4 Histogram of the geochemical components of input and validation sets, coloured by the proportion of rock types in each bin

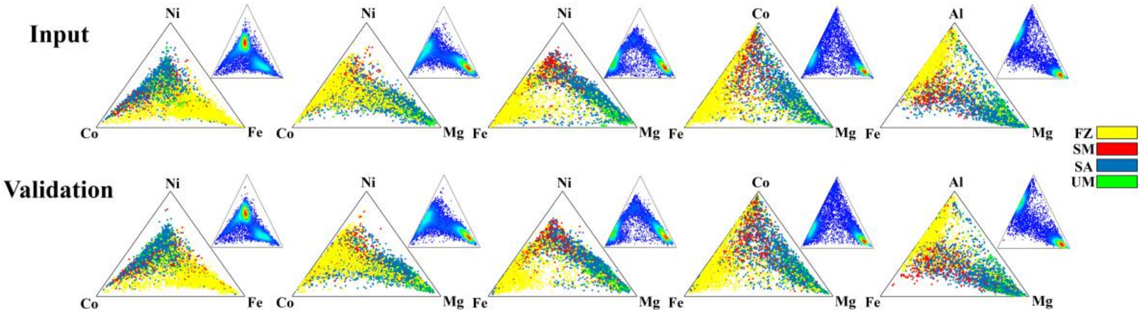

Fig. 5 Ternary diagrams of the geochemical compositions of input and validation sets, coloured by the rock types (large triangles) and kernel density (small triangles)

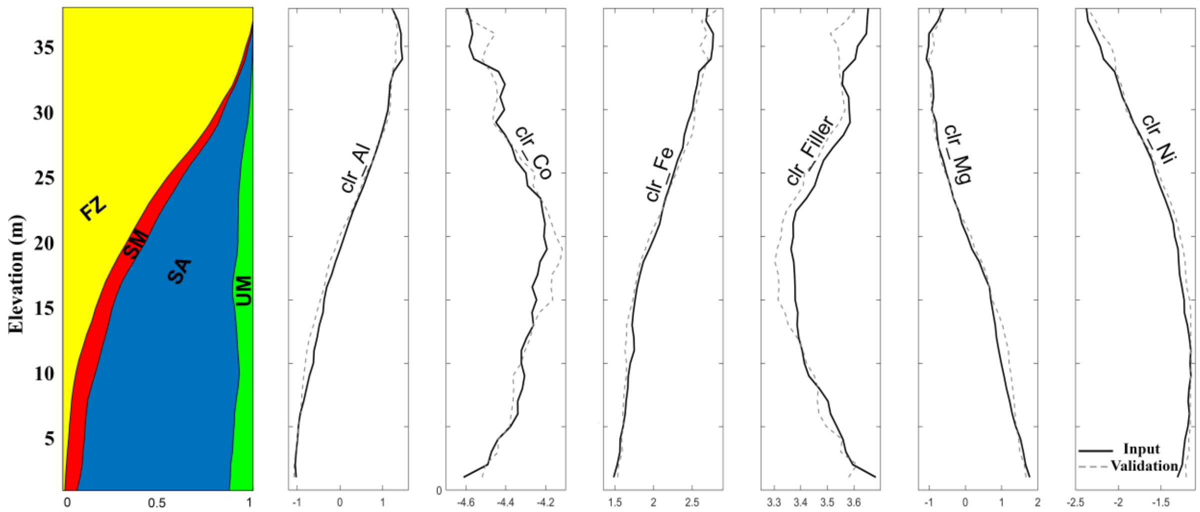

Fig. 6 Vertical proportion curves of different rock types and clr-transformed geochemical components

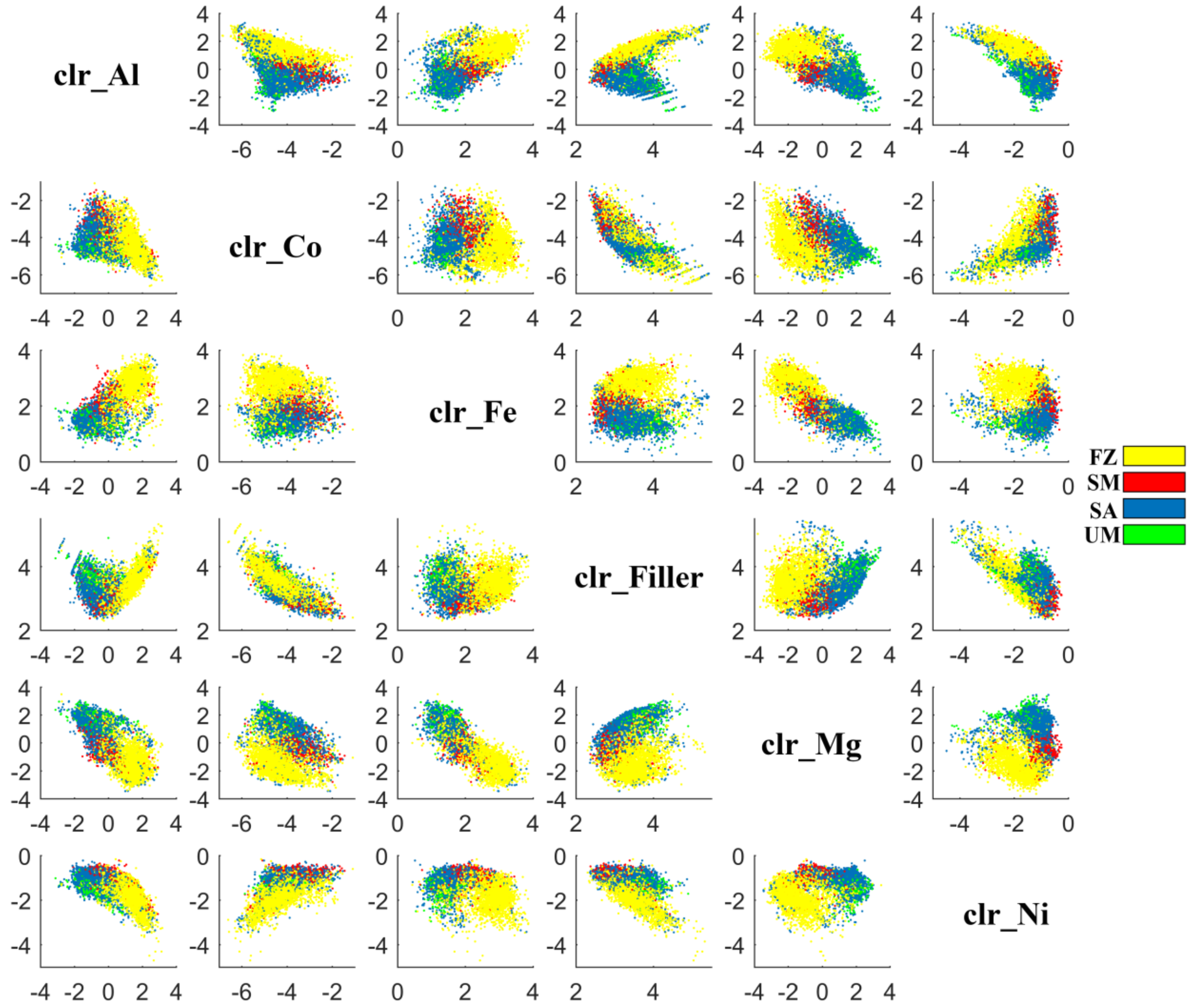

Fig. 7 Scatterplots of clr-transformed geochemical components (upper triangle is input and lower triangle is validation set)

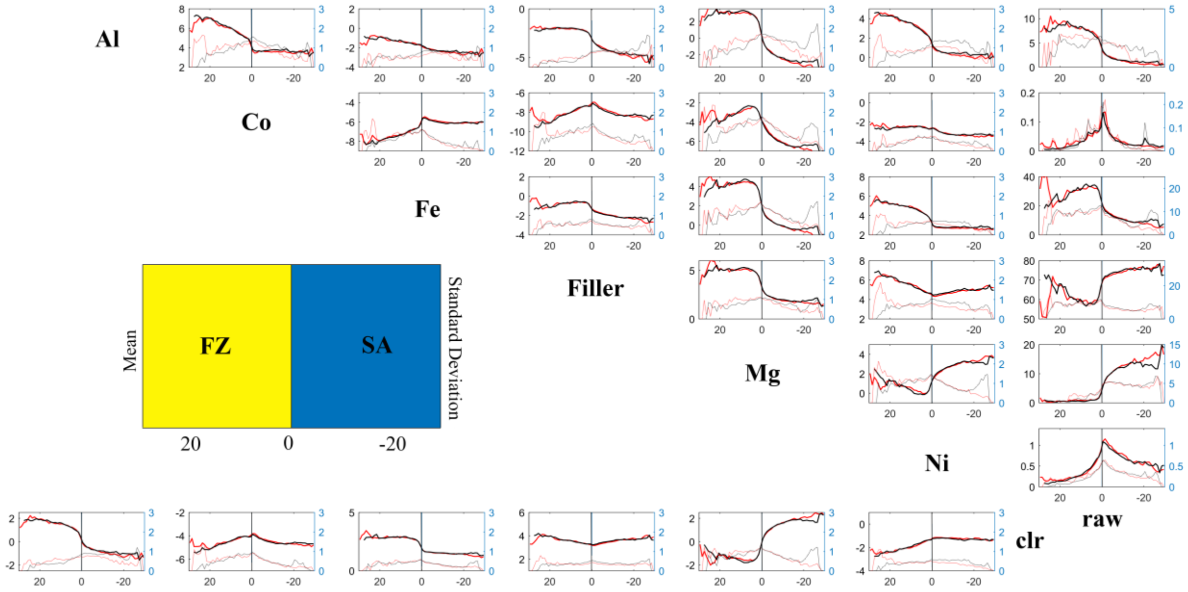

Fig. 8 Compositional contact analysis for two dominant geological domains (FZ and SA). Mean values and standard deviations are represented by continuous and dashed lines respectively (black for input set and red for validation set)

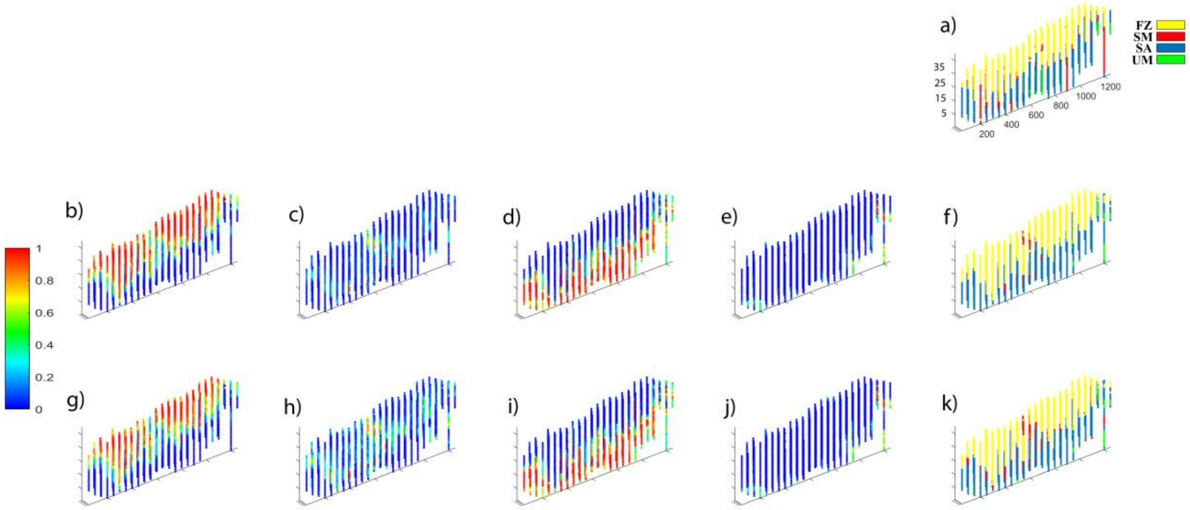

Fig. 9 Cross sections of validation boreholes for northing 300m and 50m thickness. a) true rock types. b) to e) probability of FZ, SM, SA, and UM respectively. f) most probable rock types. g) to j) adjusted probability of FZ, SM, SA, and UM respectively. k) adjusted most probable rock types

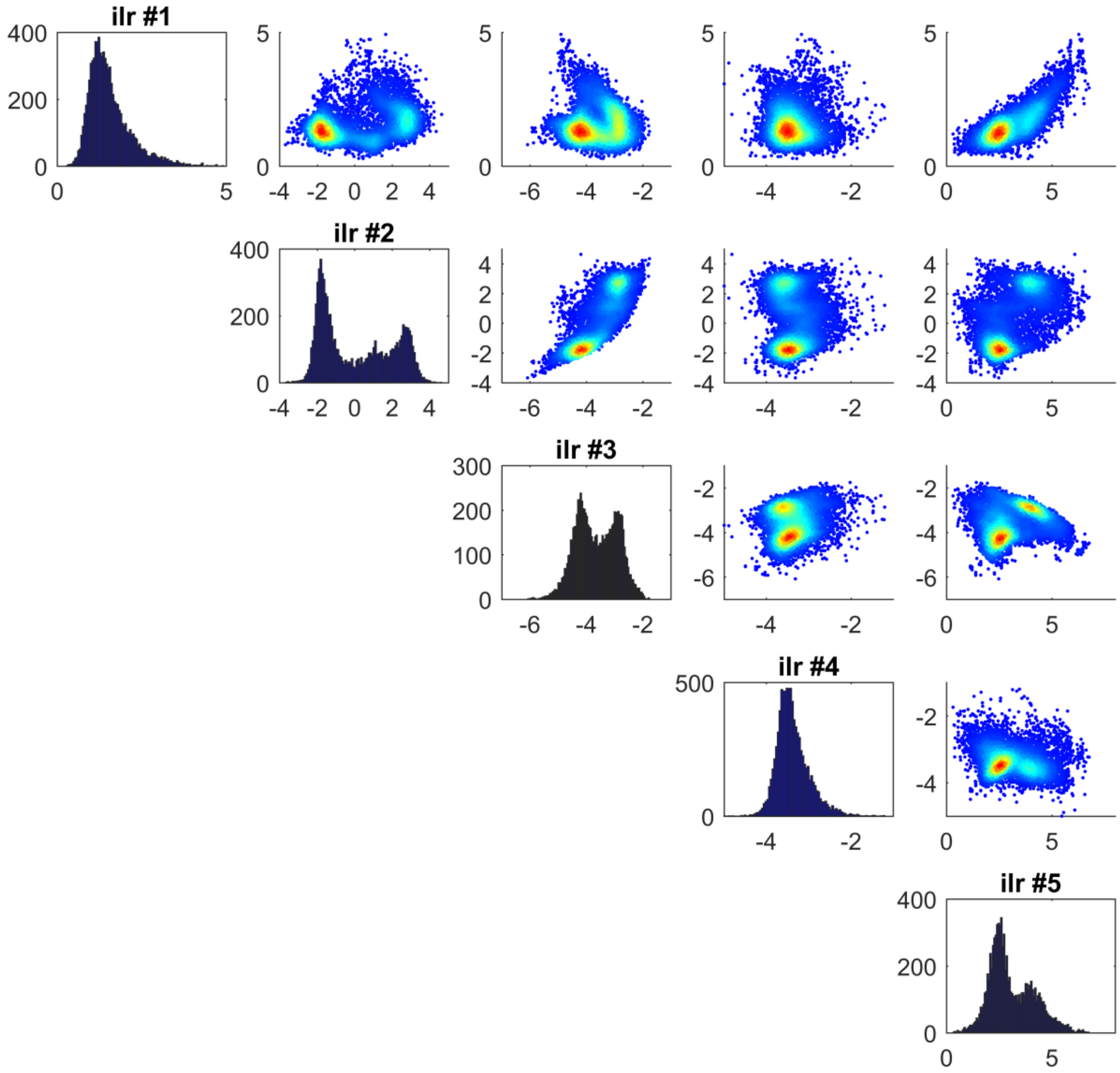

Fig. 10 Histograms and scatterplots of ilr-transformed input data (coloured by kernel density estimate)

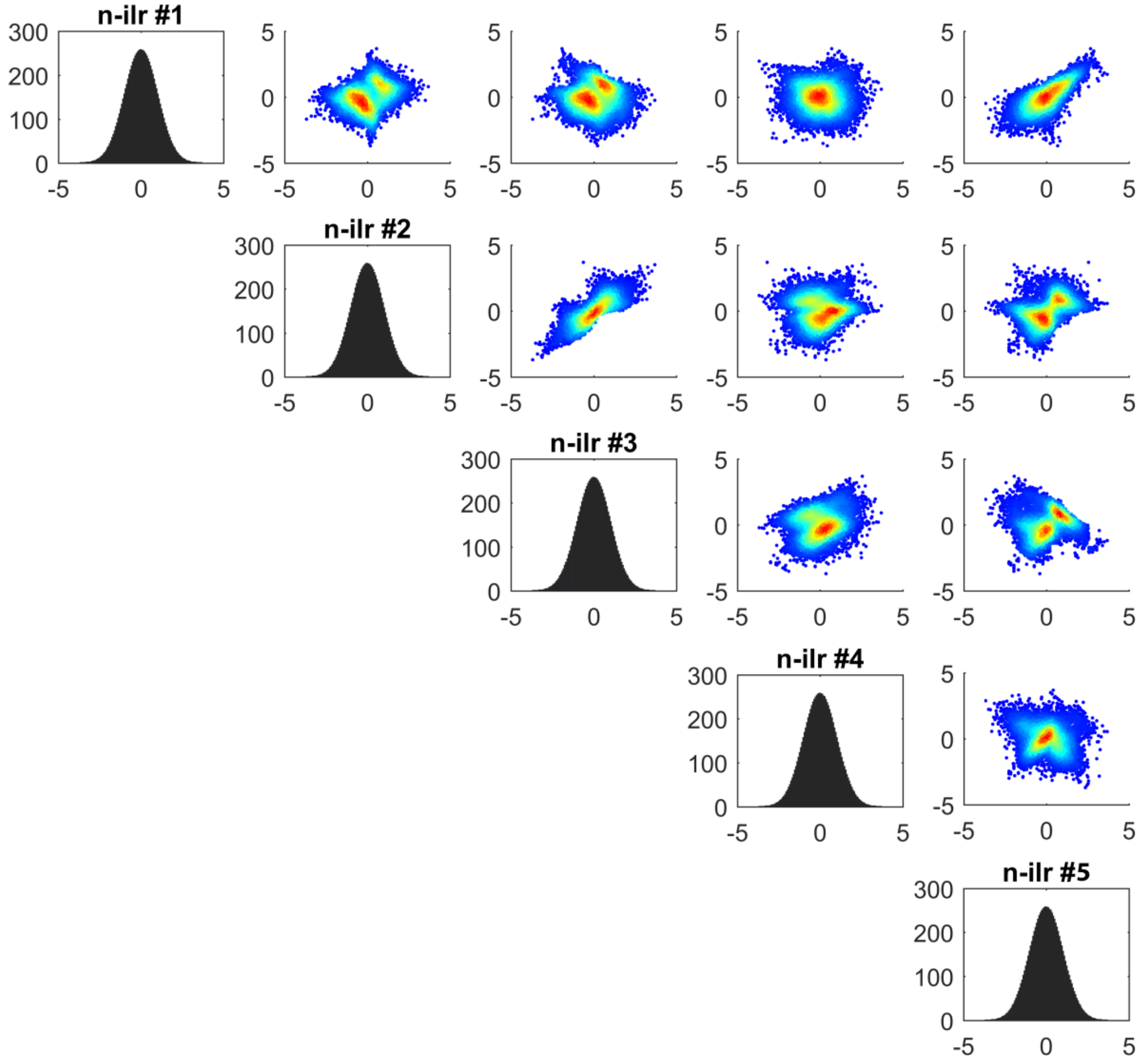

Fig. 11 Histograms and scatterplots of the transformed data to normal space via a GA (coloured by kernel density estimate)

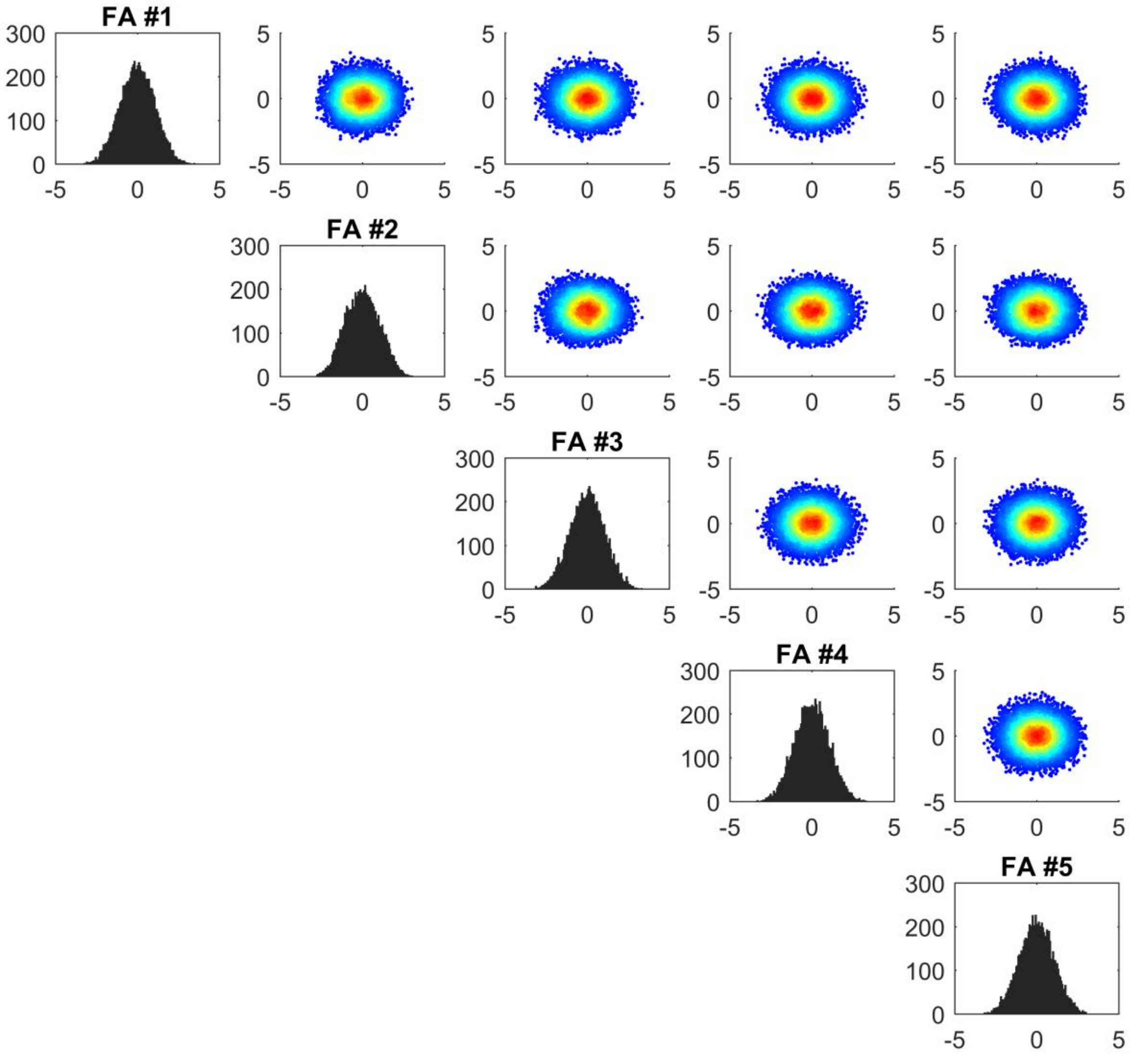

Fig. 12 Histograms and scatterplots of the transformed data to normal space via FA (coloured by kernel density estimate)

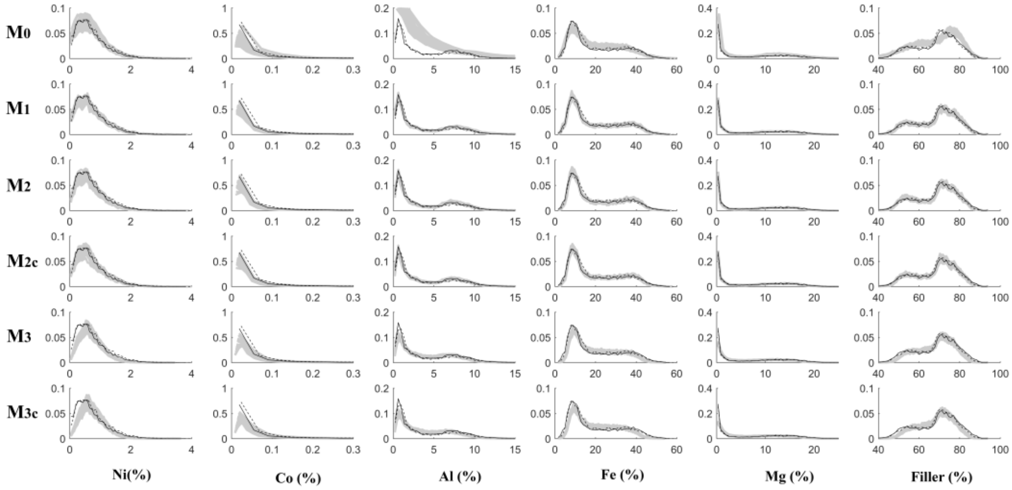

Fig. 13 Histogram reproduction of the six proposed methods for simulation of geochemical compositions. Continuous black lines are input data, dashed black lines are validation data, and grey lines are realisations

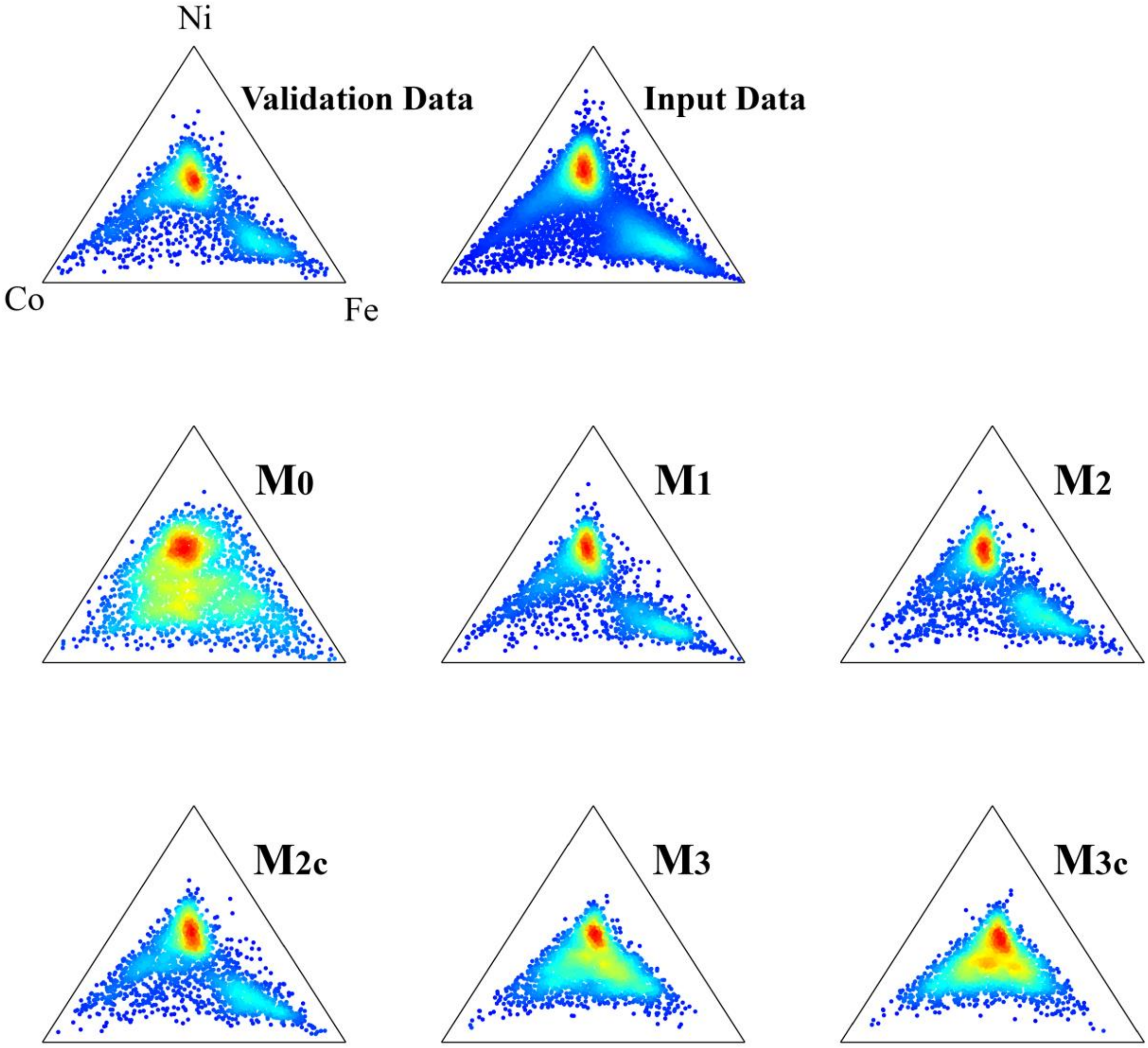

Fig. 14 Ternary diagrams of input and validation data (three components: Ni, Co, Fe), and one realisation (randomly selected) from each method

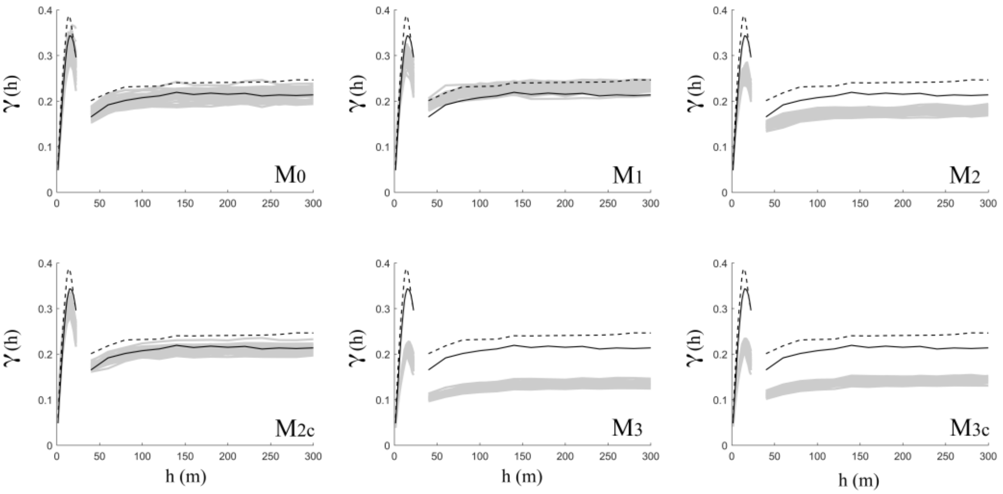

Fig. 15 Experimental variogram reproduction (Ni component) of the six proposed methods in vertical (short range) and horizontal (long range) directions.

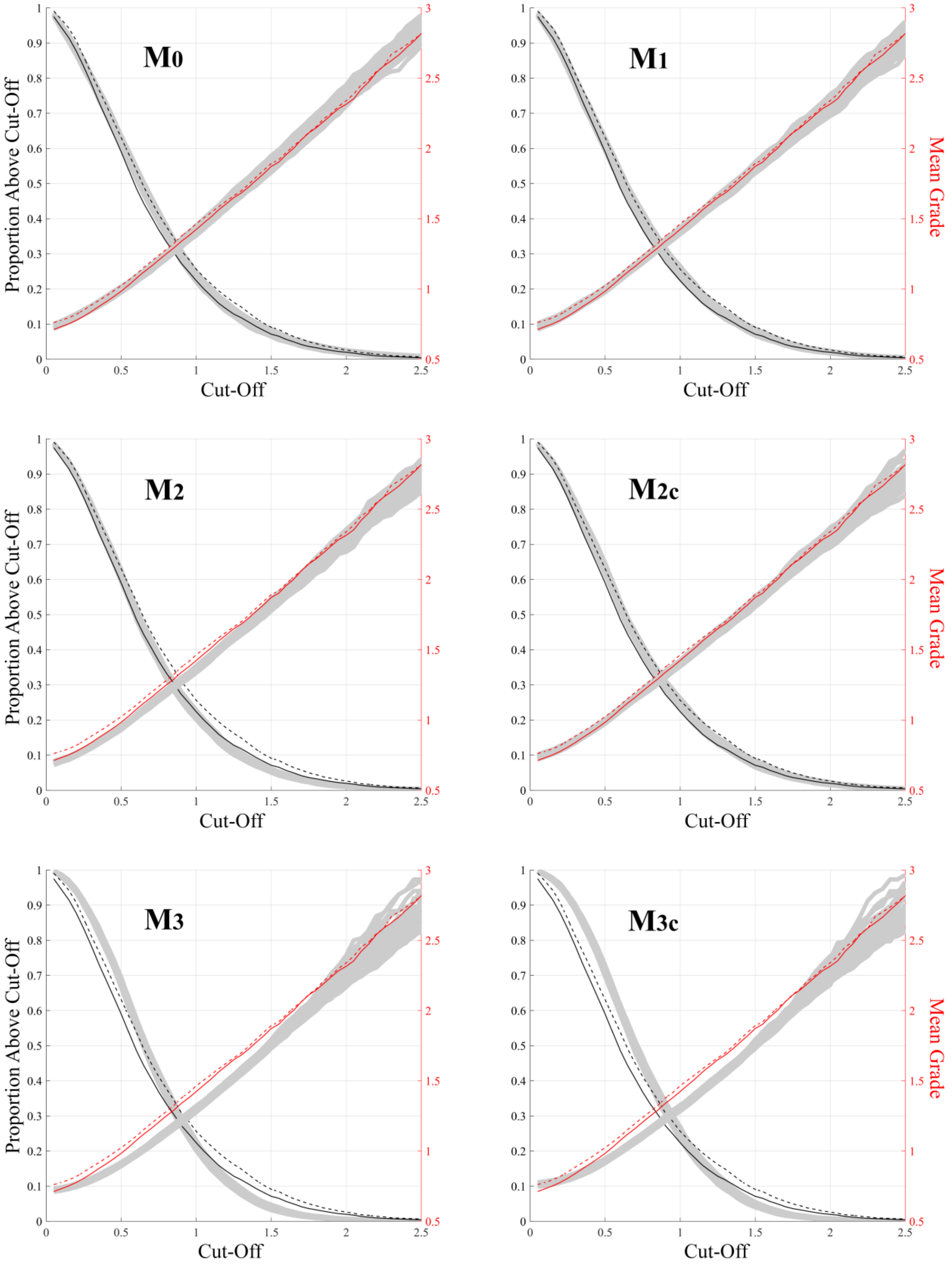

Fig. 16 Grade-tonnage curves (for Ni component) of the six proposed methods. Continuous black lines are the proportion of samples above Ni cut-offs while continuous red lines are the average grades for input data. Dashed lines are for validation data while grey lines are different realisations.

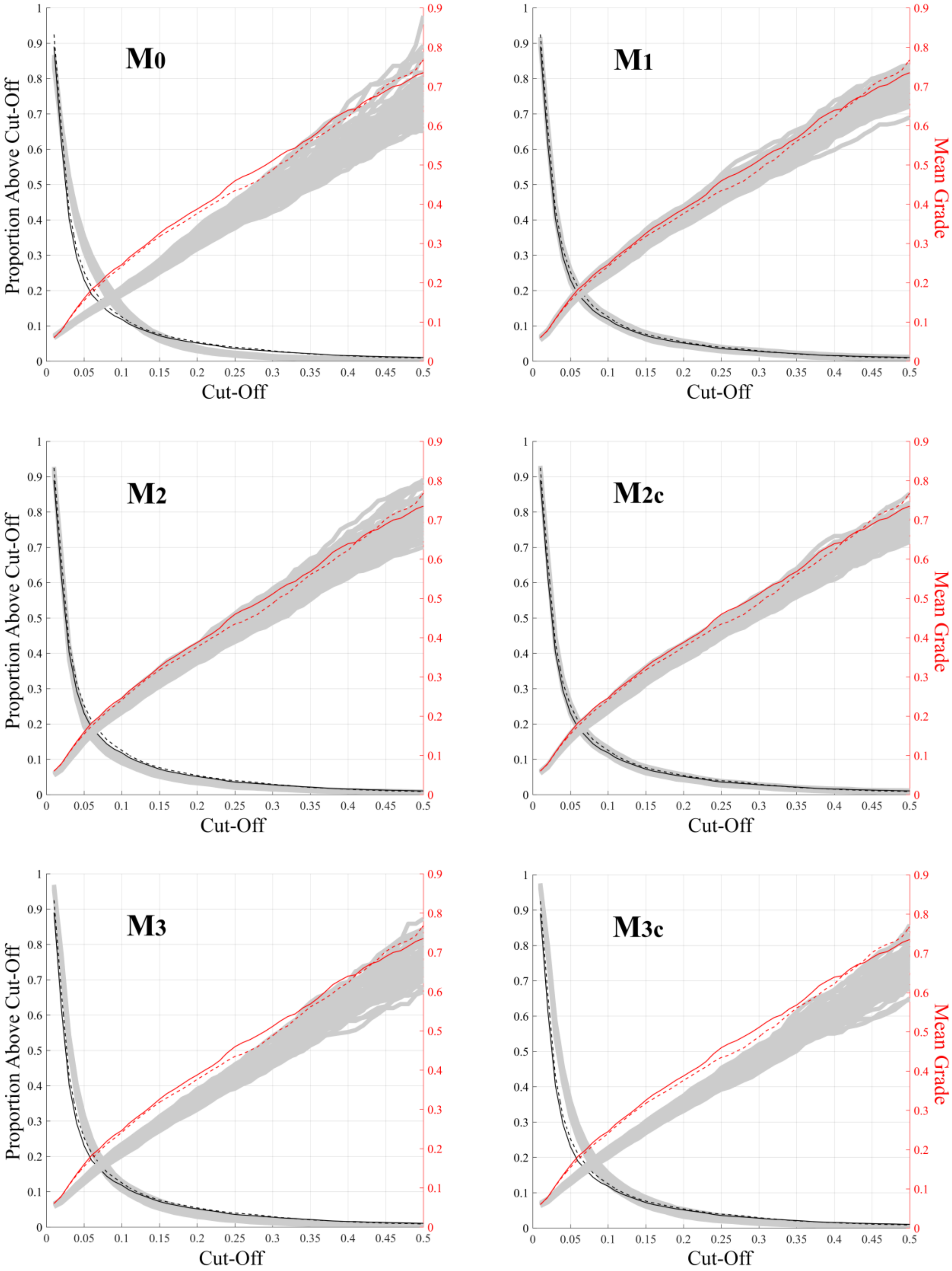

Fig. 17 Grade-tonnage curves (for Co component) of the six proposed methods Continuous black lines are the proportion of samples above Co cut-offs while continuous red lines are the average grades for input data. Dashed lines are for validation data while grey lines are different realisations.

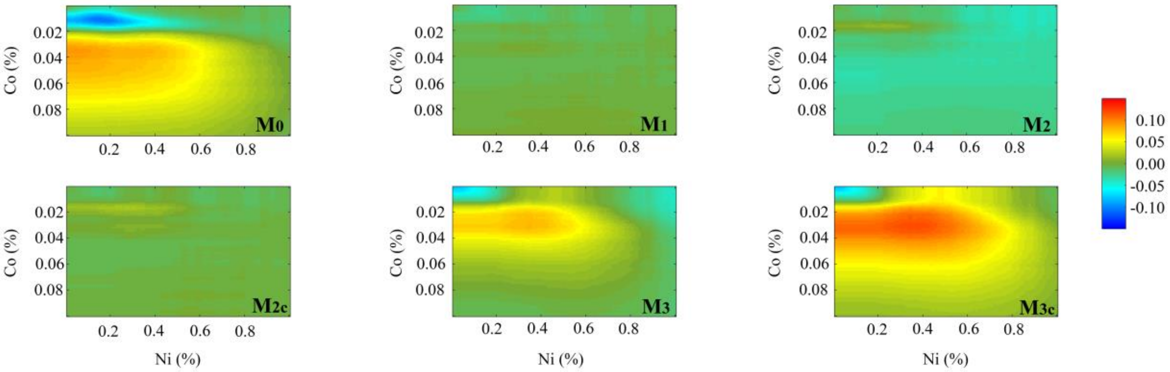

Fig. 18 The difference between the expected proportions above cut-offs (Ni and Co), calculated from realisations, and real proportions above cut-offs, calculated from the validation data

List of Tables

Table 1 Proposed methods and the related features

Code	**Geological control**	**Adjust proportions**	**Transform to real space**	**Transform to normal**	**Factorization**	**Spatial simulation**
M$_{0}$	No geological control	-	ilr	GA	MAF	TB
M$_{1}$	No geological control	-	ilr	FA	-	TB
M$_{2}$	Deterministic geological model	-	ilr	FA	-	TB
M$_{2c}$	Deterministic geological model	✓	ilr	FA	-	TB
M$_{3}$	Probabilistic geological model	-	ilr	FA	-	TB
M$_{3c}$	Probabilistic geological model	✓	ilr	FA	-	TB

Table 2 Ore mineralogy of different geological units at MME

Geological unit	**Kaolin**	**Goethite**	**Hematite**	**Maghemite**	**Chlorite**	**Smectite**	**Lizardite**
FZ	<50%	<50%	< 10%	<50%	-	-	-
SM	<2%	< 10%	<2%	<2%	< 10%	>75%	<2%
SA	-	< 10%	<2%	<2%	<2%	<50%	<50%
UM	-	-	-	-	-	-	-

Table 3 Proportions of rock types

	Validation	**Input**	**Most probable map**	**Most probable map (adjusted)**
FZ	0.352	0.367	0.417	0.376
SM	0.092	0.068	0.027	0.084
SA	0.485	0.486	0.479	0.462
UM	0.071	0.079	0.077	0.078