

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>

Perma-Link to Publication Repository of HZDR:

<https://www.hzdr.de/publications/Publ-27863>

Release of the secondary publication
on the basis of the German Copyright Law § 38 Section 4.

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

4 Hassan Talebi¹, Ute Mueller¹, Raimon Tolosana-Delgado², K. Gerald van den Boogaart²

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

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

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

10
11
12 **Abstract**

13
14 An accurate prediction of benefit in ore deposits with heterogeneous spatial variations requires
15 the definition of geological domains that differentiate the types of mineralogy, alteration, and
16 lithology as well as the prediction of full mineral and geochemical compositions within each
17 modelled domain and across boundaries between different domains. This paper proposes and
18 compares various approaches (different combination of log-ratio transformation, gaussian and
19 flow anamorphosis, and deterministic or probabilistic geological models) for geostatistical
20 simulation of geochemical compositions in the presence of several geological domains.

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

29

30

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

33

34 **1 Introduction**

35 Ore deposits usually consist of ore materials with different characteristics. In order to maximize
36 revenue in a mining project, a decision must be made regarding what processing plants are
37 needed and the processing destinations of ore materials with different characteristics. For a
38 better classification of ore materials, features such as rock type, alteration, microstructure,
39 geochemical and mineral composition must be measured at sample locations and be predicted
40 at the block model of the target deposit. These features normally have complex statistical and
41 spatial relationships which should be reproduced in the predicted models (Boisvert et al. 2013;
42 Maleki et al. 2016; Mery et al. 2017; Montoya et al. 2012; Mueller et al. 2014; Talebi et al.
43 2017; Tolosana-Delgado et al. 2014; Tolosana-Delgado et al. 2015; van den Boogaart et al.

44 2014). Among the different features of ore materials, mineral and geochemical compositions
45 have a great impact on the final destination of materials and performance of processing plants.
46 However, the compositional nature of these data induces several challenges for multivariate
47 geostatistical techniques to predict them at the block support (Tolosana-Delgado et al. 2014;
48 Tolosana-Delgado et al. 2015). Compositional data are multivariate, non-negative values
49 which represent the importance of some parts of a whole. In such data, the constant sum
50 constraint forces at least one covariance to be negative and induces spurious correlations.
51 Furthermore, they carry just relative information (Aitchison 1986). To transform compositional
52 data into unbounded space and to increase mathematical tractability, different log-ratio
53 transformations can be applied prior to using standard (geo)statistical techniques (McKinley et
54 al. 2016; Pawlowsky-Glahn and Egozcue 2016; Pawlowsky-Glahn et al. 2015; Pawlowsky-
55 Glahn and Olea 2004; Tolosana-Delgado and van den Boogaart 2013; van den Boogaart and
56 Tolosana-Delgado 2013). Most of the multivariate geostatistical simulation techniques are
57 based on the assumption of multivariate multigaussianity of the data. In real case studies, log-
58 ratio transformed data alone do not ensure this assumption, so the log-ratio transform has to be
59 combined with a normal score transform prior to using geostatistical simulation techniques in
60 order to not violate the assumptions of multigaussianity (Chilès and Delfiner 2012; Mueller et
61 al. 2014). However, compositional data do not have a unique, canonical representation and
62 several log-ratio transformations are available, making invariance of the simulated results
63 under the choice of log-ratio transform desirable. Normal score transformations based on
64 quantile matching do not have the invariance property and in addition the transformed data
65 might not be multivariate normal. Flow anamorphosis (FA) has been designed to address these
66 challenges, which means, it is capable of transforming original multivariate data to multivariate
67 normal space and at the same time being invariant under the choice of log-ratio transform
68 (Mueller et al. 2017; van den Boogaart et al. 2017).

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

84 The objective of this contribution is to compare different approaches for geostatistical
85 simulation of geochemical compositions to assess mineral resources in a nickel-cobalt laterite
86 deposit. To evaluate the effect of geological domaining on the accuracy of the predicted
87 geochemical compositions, three approaches are analysed: geological controls are ignored, a
88 deterministic geological model is applied, and a probabilistic geological model is used (Talebi
89 et al. 2016). The probabilistic geological model is calculated based on a plurigaussian (PGS)
90 model (Armstrong et al. 2011; Emery 2007). An isometric log-ratio transformation (ilr) is used
91 to transform compositional data from the simplex to real space (Egozcue et al. 2003). The ilr-
92 transformed data are transformed to normal scores and subsequently simulation is used to
93 generate realisations at unsampled locations. Classical gaussian anamorphosis (GA) and flow

94 anamorphosis are compared based on the accuracy of the predicted compositions, as well as
95 their capability to reproduce complex statistical and spatial patterns present in the input data.
96 The various approaches considered are illustrated through an application to a nickel-cobalt
97 laterite deposit and their performances are evaluated against a set of validation boreholes.

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

109

110 **2 Methodology**

111 **2.1 Compositional Data Analysis**

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

$$S^D = \left\{ \vec{Z}(u_\alpha) = [z_1(u_\alpha), z_2(u_\alpha), \dots, z_D(u_\alpha)] \mid z_i(u_\alpha) \geq 0; i = 1, 2, \dots, D; u_\alpha \in \mathbf{A}; \sum_{i=1}^D z_i(u_\alpha) = m \right\}, \quad (1)$$

117

118 where $z_i(u_\alpha)$ represents the i^{th} component measured at location u_α within the study area \mathbf{A} .
 119 The number m is the constant sum and common values are 1 (proportions), 100 (percentages),
 120 10^6 (ppm), and 10^9 (ppb). The constant sum constraint is known to induce problems of spurious
 121 correlation (see Aitchison (1982), for a detailed report). Compositional data carry just relative
 122 information, which is appropriately represented by taking log-ratio transformations (Aitchison,
 123 1986). Pairwise log-ratio transformation (Aitchison 1986), centred log-ratio transformation
 124 (Aitchison 1986), and isometric log-ratio transformation (Egozcue et al. 2003) are utilized in
 125 this study. Independently of which transform used, the resulting log-ratio scores happen to be
 126 free of the constraints of positivity and constant sum or of the spurious correlation problem,
 127 which make log-ratio scores more amenable to (geo)statistical treatment. The pairwise log-
 128 ratio transformation (pwlr) is defined as follows

129

$$pwlr(\vec{Z}(u_\alpha)) = \begin{bmatrix} 0 & \ln\left(\frac{z_1(u_\alpha)}{z_2(u_\alpha)}\right) & \dots & \ln\left(\frac{z_1(u_\alpha)}{z_D(u_\alpha)}\right) \\ \ln\left(\frac{z_2(u_\alpha)}{z_1(u_\alpha)}\right) & 0 & \dots & \ln\left(\frac{z_2(u_\alpha)}{z_D(u_\alpha)}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \ln\left(\frac{z_D(u_\alpha)}{z_1(u_\alpha)}\right) & \ln\left(\frac{z_D(u_\alpha)}{z_2(u_\alpha)}\right) & \dots & 0 \end{bmatrix}. \quad (2)$$

130

131

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

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

133

134

135 Finally, ilr transformation is defined as follows

136

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

137

138

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

147

148 **2.2 Flow Anamorphosis**

149 The flow anamorphosis introduced by van den Boogaart et al. (2017) is a multivariate form of
150 gaussian anamorphosis. In this method a kernel density estimate is deformed from the given
151 multivariate density of the observations into the density of a standard multivariate normal
152 distribution. Several statistical tests of multivariate normality can be applied to check the fit of
153 the transformed data to normality (Korkmaz et al. 2014; Mardia 1970; Székely and Rizzo 2005;
154 Székely and Rizzo 2013). The fit is highly dependent on the selection of the two parameters
155 of FA, σ_0 and σ_1 (initial and final spreads of the smoothing kernels of the kernel density

156 estimates). Deformation of the underlying space is controlled by σ_0 . The smaller the value, the
157 closer the transformed data are to multivariate normality. Selecting a proper σ_0 is dependent on
158 the number of variables (D), sample size, and complexity of the input data (Mueller et al. 2017).
159 On the other hand, σ_1 controls the ranges of the transformed distributions. In this study, in order
160 to force the marginal distributions of transformed data to have standard deviations close to
161 unity, σ_1 is selected as $\sigma_1 = \sigma_0 + 1$ (Mueller et al. 2017). Figure 1 shows the process of
162 simulation via FA.

163 [\[Fig. 1 about here.\]](#)

164
165 Figure 1(a) depicts the scatter plot and marginal kernel smoothing density estimates of the input
166 data with two simulated variables whose relationship is complex. The characteristics include
167 the presence of some outliers, nonlinear relationships, several populations, and
168 heteroscedasticity. Figure 1(b) and Fig. 1(c) show the co-deformation of the underlying space
169 and final distribution of the transformed data respectively. The transformed data in multivariate
170 normal space can be simulated (Fig. 1(d)) via many geostatistical algorithms (Alabert 1987a;
171 Deutsch and Journel 1998; Emery 2008; Emery et al. 2016; Emery and Lantuéjoul 2006).
172 Several experiences have shown that the FA-transformed data are not only multivariate normal
173 but often also exhibiting lack of spatial cross-correlation which make the geostatistical
174 simulation of such orthogonal factors, straightforward (Mueller et al. 2017; van den Boogaart
175 et al. 2017). Otherwise, multivariate simulation or spatial decorrelation technique followed by
176 univariate simulation could be implemented. The simulated results are back-transformed to the
177 original space via FA^{-1} . Figure 1(e) and Fig. 1(f) show the co-deformation back from simulated
178 multivariate normal space to the original space and final distribution of the simulated data
179 respectively. A visual comparison of these two plots shows that all the characteristics of the
180 input the data are very well reproduced.

181

182 **2.3 Geological Domaining**

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

204

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

205

206 Here $\vec{r}(u_\alpha)$ is the vector of perturbed proportions at location u_α , $R(u_\alpha) = \sum_{k=1}^K b_k q_k(u_\alpha)$,
 207 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
 208 $(r_1(u_\alpha), r_2(u_\alpha), \dots, r_K(u_\alpha))$, form the new probabilistic model of geology and can be used to
 209 determine the adjusted most probable map of geological domains.

210

211 **2.4 Approaches to Geostatistical Simulation of Compositional Data**

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

219 In the first (\mathbf{M}_0) and second (\mathbf{M}_1) proposed methods there is no geological control (Table 1).
 220 In these two scenarios all input compositions (without considering a geological domain
 221 partition of the deposit) are transformed to real space via an ilr transformation and subsequently
 222 to normal space via GA and FA, respectively. If the normal scores are spatially correlated, they
 223 are transformed to spatial orthogonal factors via Min/Max autocorrelation factors (MAF),
 224 (Bandarian et al. 2008; Desbarats and Dimitrakopoulos 2000; Rondon 2012; Switzer and Green
 225 1984). This orthogonalization makes the simulation step more straightforward. However,
 226 normal scores obtained via FA are normally spatially orthogonal and independent (Mueller et

227 al. 2017; van den Boogaart et al. 2017). The normal scores (or MAFs) are simulated
228 independently at the simulation grid (in this case all the locations of the validation data) via
229 turning bands (TB) technique (Emery and Lantuéjoul 2006). The simulated results are back-
230 transformed to the original space afterwards.

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

239 The last proposed method (\mathbf{M}_3) is geostatistical simulation of geochemical compositions using
240 a probabilistic geological model. In this case input compositions are divided into subsets based
241 on their associated geological domain. Compositions in each subset are transformed to ilr space
242 and subsequently transformed to multivariate normal space via FA. Normal scores from each
243 subset are simulated on the entire simulation grid and back-transformed to the original space
244 independently. This process provides several sets of simulated geochemical compositions
245 associated with the geological domains. Final simulated compositions can be obtained via
246 weighting the simulated compositions associated with the different domains by the local
247 probabilities of occurrence of each domain (Emery and González 2007a; Emery and González
248 2007b; Talebi et al. 2015)

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

249

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

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

254

255 Finally methods based on a deterministic model of geology (\mathbf{M}_2) and a probabilistic model
256 (\mathbf{M}_3) are repeated with the corrected deterministic (\mathbf{M}_{2c}) and probabilistic (\mathbf{M}_{3c}) models. The
257 correction is based on the method described in Sect 2.3. Table 1 provides a summary of the
258 proposed methods.

259

[\[Table 1 about here.\]](#)

260

261

262

263 **3 Case Study: Murrin Murrin Nickel-Cobalt Laterite Deposit**

264 Murrin Murrin East (MME) is a nickel-cobalt laterite deposit located in Western Australia, at
265 about 60 km southeast of Laverton and at an average elevation of approximately 420 metres
266 above mean sea level. The orebody is approximately 1,500 meters long, 600 meters wide, and
267 30 meters thick. Figure 2 shows a satellite image of the orebody together with the location of
268 the boreholes.

269

270

[\[Fig. 2 about here.\]](#)

271

272

273 **3.1 Geological Description**

274 Laterite deposits are formed during chemical weathering of ultramafic rocks near the surface
275 of the earth. At MME, nickel laterite deposits occur as laterally extensive, undulating blankets
276 of mineralisation with strong vertical trends covering basement ultramafic rocks (Murphy
277 2003). Proximity of the deposit to a major salt lake probably influenced the development of the
278 regolith (Markwell 2001). Based on the geochemical interpretation and the logging information
279 from geologists, lateritic weathering of the ultramafic rocks at MME has produced a profile
280 that may be broadly divided into four geological units (Camuti and Riel 1996; Markwell 2001;
281 Monti and Fazakerley 1996). The sequence of these units from the bottom to the top is as
282 follows: (i) ultramafic rocks (UM) at the base of weathering overlain by (ii) a Saprolite zone
283 (SA) overlain by (iii) a Smectite zone (SM) and finally capped by (iv) a Ferruginous zone (FZ,
284 Fig. 3(b)). **UM** occurs as a thin layer of unweathered ultramafic rock at the base of the deposit.
285 Due to the undulating nature of different layers in this deposit, some outcrops of UM can be
286 recognized at the surface of the deposit. **SA** consists mainly of Lizardite and Smectite. SA is
287 enriched in Mg, but low amounts of Fe are present in this zone. A strong contrast between Mg
288 in SM and SA allows easier domaining of the two units (Markwell 2001). **SM** consists mainly
289 of Smectite and is confined to the shoulders of Saprolite domes. SM is enriched in Ni and Co
290 (Fig. 3) and depleted in Mg. **FZ** is composed predominantly of Goethite, Kaolin, and
291 Maghemite. FZ is less enriched in Ni, but it can host significant amounts of Co, especially at
292 the transition to the Smectite zone (Fig. 3).

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

295 [\[Table 2 about here.\]](#)

296

297

298 **3.2 Dataset**

299 In total, 17,512 samples (of 1 meter length) from 926 RC holes (Fig. 2) make up the database
300 for this study. Four rock types (FZ, SA, SM, and UM) are considered to define compositionally
301 homogeneous domains. Three major (Fe, Al, and Mg) and two target (Ni and Co) elements are
302 the variables of interest in this study. Since the data are compositional, a filler variable is
303 introduced to achieve closure and to retain the intuitive relationship between each component
304 and the mass of its associated element. The data set was subdivided into two subsets: 453 holes
305 (8,694 samples) are considered for validation and 473 holes (8,818 samples) for geostatistical
306 modelling, called input data from hereon. The input data are used for constructing the various
307 geostatistical models. Locations of the validation and input boreholes can be seen in a cross
308 section of the deposit for northing 300m (Fig. 3(a)). Figure 4 depicts the histograms of different
309 geochemical components in the validation and input sets. Figure 5 shows centred ternary
310 diagrams (Buccianti et al. 1999; Pawlowsky-Glahn et al. 2015) of the sub-compositions. In this
311 case study geochemical compositions are composed of different populations. The multi-
312 population character is consistent with the presence of four geological units (Fig. 4 and Fig. 5).
313 These geological units control the statistical and spatial distribution of the geochemical
314 compositions. Figure 6 shows the vertical curves of the rock type proportions and clr-
315 transformed of the geochemical components.

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

329

330 [\[Fig. 3 about here.\]](#)

331 [\[Fig. 4 about here.\]](#)

332 [\[Fig. 5 about here.\]](#)

333 [\[Fig. 6 about here.\]](#)

334 [\[Fig. 7 about here.\]](#)

335

336

337 3.3 Compositional Contact Analysis

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

352

353

[\[Fig. 8 about here.\]](#)

354

355

356 A comparison of the contact diagrams of the raw components with the associated clr-
357 transformed ones shows that the behaviour of the means at the transition zone has been
358 exaggerated unrealistically in the raw contact analysis. Except for Filler which shows different
359 behaviour in the clr plot, other components show similar behaviour in raw and clr analysis. Al,
360 Fe, and Mg show abrupt transitions (variation of means as a function of distance) while the

361 local variations are stable (standard deviations as a function of distance). All the diagrams of
362 the Co (raw, clr, and pwlr) show an increase of the mean and the standard deviation
363 (proportionality effect) at the transition zone. The reason for this behaviour would be the
364 presence of Smectite units at the transition from Ferruginous to Saprolite zone, not recognised
365 during logging (Smectite units are enriched in Co). Pairwise log-ratio of the Co/Ni and Al/Fe
366 means and standard deviations are stable across geological units. The results suggest that these
367 two geological domains (FZ and SA) can be considered as stationary for some subcompositions
368 such as Al and Fe, but domaining would be necessary for other subcompositions or ratios, for
369 example Mg and Ni. Similar assessments were conducted on the other contacts between the
370 four geological domains and results supported the necessity of domaining. Given the results of
371 compositional contact analysis and statistical analysis of compositional data in Sect 3.2, it is
372 reasonable to partition the study area to four homogeneous geological units (four rock types)
373 prior to geostatistical modelling.

374

375 **3.4 Deterministic and Probabilistic Geological Models**

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

383 The last column in Table 3 shows the adjusted proportions in the most probable rock type map,
384 demonstrating an improvement in the proportion of the SM units. Figure 9(g) to Fig. 9(j) show

385 cross-sections of the validation boreholes, coloured by the adjusted probabilities, for northing
386 300m. A visual comparison of Fig. 9(c) and Fig. 9(h) indicates a clear improvement in the
387 proportion of the SM units, more representative of reality (Fig. 9(a)).

388

389 [\[Table 3 about here.\]](#)

390

391

392 [\[Fig. 9 about here.\]](#)

393

394

395

396 **4 Results and Discussion**

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

400 Transformation of ilr transformed data into normal space in \mathbf{M}_0 was based on GA. Figure 11
401 shows the scatterplots of the normal scores obtained by GA. Although the marginal
402 distributions are normal, the scores are not multivariate normal. An analysis of the
403 omnidirectional experimental semivariograms and cross-variograms further revealed that these
404 normal scores are spatially correlated, with Tercan's $\bar{\tau}$ and $\bar{\kappa}$ (Tercan 1999) equal to 0.1973
405 and 0.8147, respectively. To make the simulation step more straightforward, normal scores
406 data were transformed to spatially orthogonal factors via MAF. MAF transformation improved
407 spatial orthogonality with $\bar{\tau}$ and $\bar{\kappa}$ equal to 0.0782 and 0.9689, respectively. Orthogonal factors

408 were simulated at the entire validation holes via the TB algorithm. The simulated results were
409 back-transformed to the simplex afterwards to recover outputs in the original scale as
410 percentages.

411 On the other hand, \mathbf{M}_1 transforms the ilr-transformed data to multivariate normal space via FA.
412 Due to the complexity of the data and the number of the observations and variables,
413 multivariate normality was not achieved by a single FA. Two successive FA with the same
414 parameters ($\sigma_0 = 0.1$ and $\sigma_1 = 1.1$) were required to achieve multivariate normality. Figure 12
415 shows the normal scores are obviously close to multivariate normal and uncorrelated, so
416 statistically independent. Spatial structural analysis (variography) showed further that the
417 normal scores are spatially orthogonal, with Tercan's $\bar{\tau}$ and $\bar{\kappa}$ equal to 0.0656 and 0.9873,
418 respectively, so they could be simulated independently. The scores were simulated at the
419 validation holes independently via TB algorithm and back-transformed to composition
420 afterward.

421 [\[Fig. 10 about here.\]](#)

422 [\[Fig. 11 about here.\]](#)

423 [\[Fig. 12 about here.\]](#)

424

425

426 For simulation based on a deterministic geological model input compositions were divided into
427 four subsets based on their associated rock types (FZ, SM, SA, and UM). Compositions in each
428 subset were transformed to ilr space and subsequently transformed to multivariate normal space
429 via FA independently. The validation holes were also divided into four mutually exclusive and
430 exhaustive parts based on the most probable rock types (Fig. 9(f)) and the corrected most
431 probable rock types (Fig. 9(k)), provided the simulation grids for \mathbf{M}_2 and \mathbf{M}_{2c} respectively.

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

436 Unlike the deterministic approaches (\mathbf{M}_2 and \mathbf{M}_{2c}), in the probabilistic approach normal scores
437 from each subset of input compositions were simulated at the entire simulation grid. This
438 process provided four sets of simulated geochemical compositions associated with the four
439 geological domains (FZ, SM, SA, UM). Final simulated compositions obtained by weighting
440 the simulated compositions associated with the different domains by the probabilities of
441 occurrence of each domain (Eq. 6). Two sets of probabilities are available, the raw probabilities
442 (Fig. 9(b) to Fig. 9(e)) obtained by the PGS model (\mathbf{M}_3) and adjusted probabilities (\mathbf{M}_{3c})
443 obtained by the proposed correction method in Sect 2.3 (Fig. 9(g) to Fig. 9(j)).

444 The simulated compositional models based on the six proposed approaches were validated
445 against reality available at the validation boreholes (Fig. 3(a)). Figure 13 shows the global
446 histogram reproduction of the six methods. Best realisations were achieved by \mathbf{M}_1 followed by
447 \mathbf{M}_{2c} . Visual comparison of histogram reproduction of \mathbf{M}_2 and \mathbf{M}_{2c} reveals that adjusting
448 proportions (Sect 2.3) has improved the histogram reproduction especially for Co component.
449 Worst reproductions were achieved by \mathbf{M}_0 .

450 [\[Fig. 13 about here.\]](#)

451

452

453 To check the ability of the methods to reproduce the complex patterns in the simplex space,
454 ternary diagrams were plotted and compared to each other (Fig. 14). Due to the averaging
455 nature of \mathbf{M}_3 and \mathbf{M}_{3c} , different populations have been mixed up, which is not acceptable. On

456 the other hand, \mathbf{M}_0 was not able to reproduce different populations and complex patterns inside
457 the simplex space. Here, \mathbf{M}_1 generated the best results followed by \mathbf{M}_2 and \mathbf{M}_{2c} .

458 The smoothing effect of the \mathbf{M}_3 and \mathbf{M}_{3c} is more obvious from the experimental variograms of
459 the simulated models (Fig. 15). In Fig. 15 continuous black lines are input data, dashed black
460 lines are validation data and grey lines are realisations. The sills of the variograms for one of
461 the simulated component (Ni) are systematically less than the sills for input and validation data,
462 representing a systematic reduction of the spatial variability. Adjusting probabilities of
463 geological domains has improved the variogram reproduction for the deterministic approaches
464 (compare \mathbf{M}_2 and \mathbf{M}_{2c}).

465 Risk quantification in grade-tonnage curves (GTC) and predicting true curves are critical for
466 feasibility studies and capital investment in mining projects. The accuracy of the six methods
467 was investigated based on their capability for reproducing input GTCs and predicting true
468 (validation data in this study) GTCs. Figure 16 and Fig. 17 show the GTCs for Ni and Co
469 components respectively. The smoothing effect of the \mathbf{M}_3 and \mathbf{M}_{3c} is again clear from the GTCs
470 of both Ni and Co. Comparing GTCs of \mathbf{M}_2 and \mathbf{M}_{2c} , an improvement can be recognised due
471 to the implementation of the proposed technique for adjusting underrepresented domains.
472 GTCs for Co component generated by \mathbf{M}_0 is not acceptable. \mathbf{M}_1 generated the most satisfactory
473 results followed by \mathbf{M}_{2c} .

474 Finally, the six proposed methods were compared to check the presence of any systematic bias
475 in the predicted proportions of samples above cut-offs (Fig. 18). In multi-element deposits,
476 there are more than one target elements (in this study two target elements: Ni and Co). In Fig.
477 18, each cell of the maps is associated with two cut-offs (one for Ni and one for Co). For each
478 cell the expected proportion of samples above the two cut-offs were calculated from all
479 realisations and subsequently subtracted from the real associated proportion calculated from

480 validation data. These maps can be compared to show the capability of the proposed methods
481 for avoiding any serious bias (over or underestimations of ore tonnages). M_0 , M_3 , and M_{3c}
482 show clear systematic over and underestimations in some range of cut-offs. In this regard, M_1
483 generated the best (least bias) results, followed by M_{2c} .

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

489

490 [\[Fig. 14 about here.\]](#)

491 [\[Fig. 15 about here.\]](#)

492 [\[Fig. 16 about here.\]](#)

493 [\[Fig. 17 about here.\]](#)

494 [\[Fig. 18 about here.\]](#)

495

496

497 **5 Conclusion**

498 This paper compares various geostatistical approaches for simulation of geochemical
499 compositions and their application to mineral resource evaluation. Results and several
500 validation tests showed that the classical transformation to normal space (gaussian
501 anamorphosis) is not capable of reproducing complex statistical patterns inside data and should
502 be replaced with more advanced transformations. The method for modelling geochemical

503 compositions based on a probabilistic geological model exhibits smoothing effects due the
504 averaging nature of the algorithm. Although this method generates satisfactory results for
505 kriged maps (Emery and González 2007a; Emery and González 2007b; Talebi et al. 2015), it
506 should be avoided for simulation purposes. The proposed technique for adjusting
507 underrepresented domains improved the result of simulation and should be used in the cases
508 where there are important geological domains with small proportions such as SM in this case
509 study. Flow anamorphosis is a vital element for geostatistical modelling of geochemical
510 composition due to its invariance properties and capability for reproducing complex patterns
511 in data such as: outliers, multi-population, nonlinearity, and heteroscedasticity. In the case
512 study presented, a simulation involving a global flow anamorphosis without domaining was
513 best capable of reproducing all performance targets (histograms, variograms, grade and
514 tonnage curves). In the authors' opinion, this remarkable property might occur again in other
515 settings in which domains emerge as the effect of chemical processes mostly involving the
516 composition modelled.

517

518 **Acknowledgments**

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

521

522

523 **References**

524 Aitchison J (1982) The Statistical Analysis of Compositional Data Journal of the Royal
525 Statistical Society Series B (Methodological) 44:139-177
526 Aitchison J (1986) The statistical analysis of compositional data. Chapman & Hall Ltd,

527 Alabert F (1987a) The practice of fast conditional simulations through the LU decomposition
528 of the covariance matrix *Mathematical Geology* 19:369-386 doi:10.1007/bf00897191
529 Alabert F (1987b) Stochastic imaging of spatial distributions using hard and soft information.
530 Master's thesis. Department of Applied Earth Sciences, Stanford University.
531 Armstrong M et al. (2011) *Plurigaussian Simulations in Geosciences*. Springer Berlin
532 Heidelberg, Berlin, Heidelberg. doi:10.1007/978-3-642-19607-2_3
533 Bandarian EM, Bloom LM, Mueller UA (2008) Direct minimum/maximum autocorrelation
534 factors within the framework of a two structure linear model of coregionalisation
535 *Computers & Geosciences* 34:190-200
536 doi:<https://doi.org/10.1016/j.cageo.2007.03.015>
537 Boisvert JB, Rossi ME, Ehrig K, Deutsch CV (2013) Geometallurgical Modeling at Olympic
538 Dam Mine, South Australia *Mathematical Geosciences* 45:901-925
539 doi:10.1007/s11004-013-9462-5
540 Buccianti A, Pawlowsky-Glahn V, Barceló-Vidal C, Jarauta-Bragulat E (1999) Visualization
541 and modeling of natural trends in ternary diagrams: A Geochemical Case Study vols I
542 and II, Tapir, Trondheim (N), pp 139–144
543 Camuti K, Riel R (1996) Mineralogy of the Murrin Murrin nickel laterites Australasian
544 Institute of Mining and Metallurgy 6/96
545 Chilès J-P, Delfiner P (2012). In: *Geostatistics: Modeling Spatial Uncertainty*. John Wiley &
546 Sons, Inc., p 699. doi:10.1002/9780470316993.fmatter
547 Desbarats AJ, Dimitrakopoulos R (2000) Geostatistical Simulation of Regionalized Pore-Size
548 Distributions Using Min/Max Autocorrelation Factors *Mathematical Geology* 32:919-
549 942 doi:10.1023/A:1007570402430
550 Deutsch CV, Journel AG (1998) *GSLIB: Geostatistical Software Library and User's Guide*.
551 Oxford University Press. New York.
552 Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, Barceló-Vidal C (2003) Isometric
553 Logratio Transformations for Compositional Data Analysis *Mathematical Geology*
554 35:279-300 doi:10.1023/a:1023818214614
555 Emery X (2007) Simulation of geological domains using the plurigaussian model: New
556 developments and computer programs *Computers & Geosciences* 33:1189-1201
557 doi:<https://doi.org/10.1016/j.cageo.2007.01.006>
558 Emery X (2008) A turning bands program for conditional co-simulation of cross-correlated
559 Gaussian random fields *Computers & Geosciences* 34:1850-1862
560 doi:<https://doi.org/10.1016/j.cageo.2007.10.007>
561 Emery X, Arroyo D, Porcu E (2016) An improved spectral turning-bands algorithm for
562 simulating stationary vector Gaussian random fields *Stochastic Environmental*
563 *Research and Risk Assessment* 30:1863-1873 doi:10.1007/s00477-015-1151-0
564 Emery X, González KE (2007a) Incorporating the uncertainty in geological boundaries into
565 mineral resources evaluation. *J. Geol. Soc. India* 69(1), 29–38
566 Emery X, González KE (2007b) Probabilistic modelling of lithological domains and its
567 application to resource evaluation *Journal of the Southern African Institute of Mining*
568 *and Metallurgy* 107:803-809
569 Emery X, Lantuéjoul C (2006) TBSIM: A computer program for conditional simulation of
570 three-dimensional Gaussian random fields via the turning bands method *Computers &*
571 *Geosciences* 32:1615-1628 doi:<https://doi.org/10.1016/j.cageo.2006.03.001>
572 Korkmaz S, Goksuluk D, Zararsiz G (2014) MVN: an R package for assessing multivariate
573 normality *R J* 6:151–162
574 Maleki M, Emery X, Cáceres A, Ribeiro D, Cunha E (2016) Quantifying the uncertainty in the
575 spatial layout of rock type domains in an iron ore deposit *Computational Geosciences*
576 20:1013-1028 doi:10.1007/s10596-016-9574-3

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

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

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

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

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

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

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

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

597 Mueller U, van den Boogaart KG, Tolosana-Delgado R (2017) A Truly Multivariate Normal
598 Score Transform Based on Lagrangian Flow. In: Gómez-Hernández JJ, Rodrigo-Ilarri
599 J, Rodrigo-Clavero ME, Cassiraga E, Vargas-Guzmán JA (eds) *Geostatistics Valencia*
600 2016. Springer International Publishing, Cham, pp 107-118. doi:10.1007/978-3-319-
601 46819-8_7

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

604 Ortiz JM, Emery X (2006) Geostatistical estimation of mineral resources with soft geological
605 boundaries: a comparative study The South African Institute of Mining and Metallurgy
606 106:577–584

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

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

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

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

616 Rondon O (2012) Teaching Aid: Minimum/Maximum Autocorrelation Factors for Joint
617 Simulation of Attributes *Mathematical Geosciences* 44:469-504 doi:10.1007/s11004-
618 011-9329-6

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

621 Switzer P, Green AA (1984) Min/max autocorrelation factors for multivariate spatial imaging.
622 Technical Report No. 6, Department of Statistics, Stanford University, Stanford, CA,
623 14pp.

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

- 626 Székely GJ, Rizzo ML (2013) Energy statistics: A class of statistics based on distances *Journal*
627 *of Statistical Planning and Inference* 143:1249-1272
628 doi:<https://doi.org/10.1016/j.jspi.2013.03.018>
- 629 Talebi H, Asghari O, Emery X (2013) Application of plurigaussian simulation to delineate the
630 layout of alteration domains in Sungun copper deposit *Central European Journal of*
631 *Geosciences* 5:514-522 doi:10.2478/s13533-012-0146-3
- 632 Talebi H, Asghari O, Emery X (2014) Simulation of the lately injected dykes in an Iranian
633 porphyry copper deposit using the plurigaussian model *Arabian Journal of Geosciences*
634 7:2771-2780 doi:10.1007/s12517-013-0911-8
- 635 Talebi H, Asghari O, Emery X (2015) Stochastic rock type modeling in a porphyry copper
636 deposit and its application to copper grade evaluation *Journal of Geochemical*
637 *Exploration* 157:162-168 doi:<https://doi.org/10.1016/j.gexplo.2015.06.010>
- 638 Talebi H, Lo J, Mueller U (2017) A Hybrid Model for Joint Simulation of High-Dimensional
639 Continuous and Categorical Variables. In: Gómez-Hernández JJ, Rodrigo-Ilarri J,
640 Rodrigo-Clavero ME, Cassiraga E, Vargas-Guzmán JA (eds) *Geostatistics Valencia*
641 2016. Springer International Publishing, Cham, pp 415-430. doi:10.1007/978-3-319-
642 46819-8_28
- 643 Talebi H, Sabeti EH, Azadi M, Emery X (2016) Risk quantification with combined use of
644 lithological and grade simulations: Application to a porphyry copper deposit *Ore*
645 *Geology Reviews* 75:42-51 doi:<https://doi.org/10.1016/j.oregeorev.2015.12.007>
- 646 Tercan AE (1999) Importance of Orthogonalization Algorithm in Modeling Conditional
647 Distributions by Orthogonal Transformed Indicator Methods *Mathematical Geology*
648 31:155-173 doi:10.1023/a:1007557701073
- 649 Tolosana-Delgado R (2006) *Geostatistics for constrained variables: positive data, compositions*
650 *and probabilities. Application to environmental hazard monitoring.* PhD thesis,
651 University of Girona, Girona (Spain).
- 652 Tolosana-Delgado R, Mueller U, van den Boogaart KG (2016) Compositionally Compliant
653 Contact Analysis. In: Raju NJ (ed) *Geostatistical and Geospatial Approaches for the*
654 *Characterization of Natural Resources in the Environment: Challenges, Processes and*
655 *Strategies.* Springer International Publishing, Cham, pp 11-14. doi:10.1007/978-3-319-
656 18663-4_2
- 657 Tolosana-Delgado R, Mueller U, van den Boogaart KG, Ward C (2014) Compositional Block
658 Cokriging. In: Pardo-Igúzquiza E, Guardiola-Albert C, Heredia J, Moreno-Merino L,
659 Durán JJ, Vargas-Guzmán JA (eds) *Mathematics of Planet Earth: Proceedings of the*
660 *15th Annual Conference of the International Association for Mathematical*
661 *Geosciences.* Springer Berlin Heidelberg, Berlin, Heidelberg, pp 713-716.
662 doi:10.1007/978-3-642-32408-6_154
- 663 Tolosana-Delgado R, Mueller U, van den Boogaart KG, Ward C, Gutzmer J (2015) Improving
664 processing by adaption to conditional geostatistical simulation of block compositions
665 *Journal of the Southern African Institute of Mining and Metallurgy* 115:13-26
- 666 Tolosana-Delgado R, van den Boogaart KG (2013) Joint Consistent Mapping of High-
667 Dimensional Geochemical Surveys *Mathematical Geosciences* 45:983-1004
668 doi:10.1007/s11004-013-9485-y
- 669 van den Boogaart KG, Mueller U, Tolosana-Delgado R (2017) An Affine Equivariant
670 Multivariate Normal Score Transform for Compositional Data *Mathematical*
671 *Geosciences* 49:231-251 doi:10.1007/s11004-016-9645-y
- 672 van den Boogaart KG, Tolosana-Delgado R (2013) *Analyzing Compositional Data with R.*
673 Springer-Verlag Berlin Heidelberg. doi:10.1007/978-3-642-36809-7
- 674 van den Boogaart KG, Tolosana-Delgado R, Lehmann M, Mueller U (2014) On the joint
675 multipoint simulation of discrete and continuous geometallurgical parameters. Paper

676 presented at the Proceedings Orebody Modelling and Strategic Mine Planning
 677 Symposium 2014, Melbourne,
 678 Wackernagel H (2003) Gaussian Anamorphosis with Hermite Polynomials. In: Multivariate
 679 Geostatistics: An Introduction with Applications. Springer Berlin Heidelberg, Berlin,
 680 Heidelberg, pp 238-249. doi:10.1007/978-3-662-05294-5_33

681

682

683

684

685

686 **List of Figures**

687 **Fig. 1** Geostatistical simulation via Flow Anamorphosis 29

688 **Fig. 2** Borehole location map of the Murrin Murrin East (MME)..... 30

689 **Fig. 3** Cross sections of boreholes for northing 300m and 100m thickness: locations of input and
 690 validation boreholes (a), spatial distributions of different rock types (b) Ni grade (c) and Co grade (d)
 691 31

692 **Fig. 4** Histogram of the geochemical components of input and validation sets, coloured by the
 693 proportion of rock types in each bin 32

694 **Fig. 5** Ternary diagrams of the geochemical compositions of input and validation sets, coloured by the
 695 rock types (large triangles) and kernel density (small triangles) 33

696 **Fig. 6** Vertical proportion curves of different rock types and clr-transformed geochemical components
 697 34

698 **Fig. 7** Scatterplots of clr-transformed geochemical components (upper triangle is input and lower
 699 triangle is validation set)..... 35

700 **Fig. 8** Compositional contact analysis for two dominant geological domains (FZ and SA). Mean values
 701 and standard deviations are represented by continuous and dashed lines respectively (black for input set
 702 and red for validation set) 36

703	Fig. 9 Cross sections of validation boreholes for northing 300m and 50m thickness. a) true rock types.	
704	b) to e) probability of FZ, SM, SA, and UM respectively. f) most probable rock types. g) to j) adjusted	
705	probability of FZ, SM, SA, and UM respectively. k) adjusted most probable rock types.....	37
706	Fig. 10 Histograms and scatterplots of ilr-transformed input data (coloured by kernel density estimate)	
707	38
708	Fig. 11 Histograms and scatterplots of the transformed data to normal space via a GA (coloured by	
709	kernel density estimate)	39
710	Fig. 12 Histograms and scatterplots of the transformed data to normal space via FA (coloured by kernel	
711	density estimate)	40
712	Fig. 13 Histogram reproduction of the six proposed methods for simulation of geochemical	
713	compositions. Continuous black lines are input data, dashed black lines are validation data, and grey	
714	lines are realisations	41
715	Fig. 14 Ternary diagrams of input and validation data (three components: Ni, Co, Fe), and one	
716	realisation (randomly selected) from each method	42
717	Fig. 15 Experimental variogram reproduction (Ni component) of the six proposed methods in vertical	
718	(short range) and horizontal (long range) directions.....	43
719	Fig. 16 Grade-tonnage curves (for Ni component) of the six proposed methods. Continuous black lines	
720	are the proportion of samples above Ni cut-offs while continuous red lines are the average grades for	
721	input data. Dashed lines are for validation data while grey lines are different realisations.	44
722	Fig. 17 Grade-tonnage curves (for Co component) of the six proposed methods Continuous black lines	
723	are the proportion of samples above Co cut-offs while continuous red lines are the average grades for	
724	input data. Dashed lines are for validation data while grey lines are different realisations.	45
725	Fig. 18 The difference between the expected proportions above cut-offs (Ni and Co), calculated from	
726	realisations, and real proportions above cut-offs, calculated from the validation data	46
727		
728		

729

730

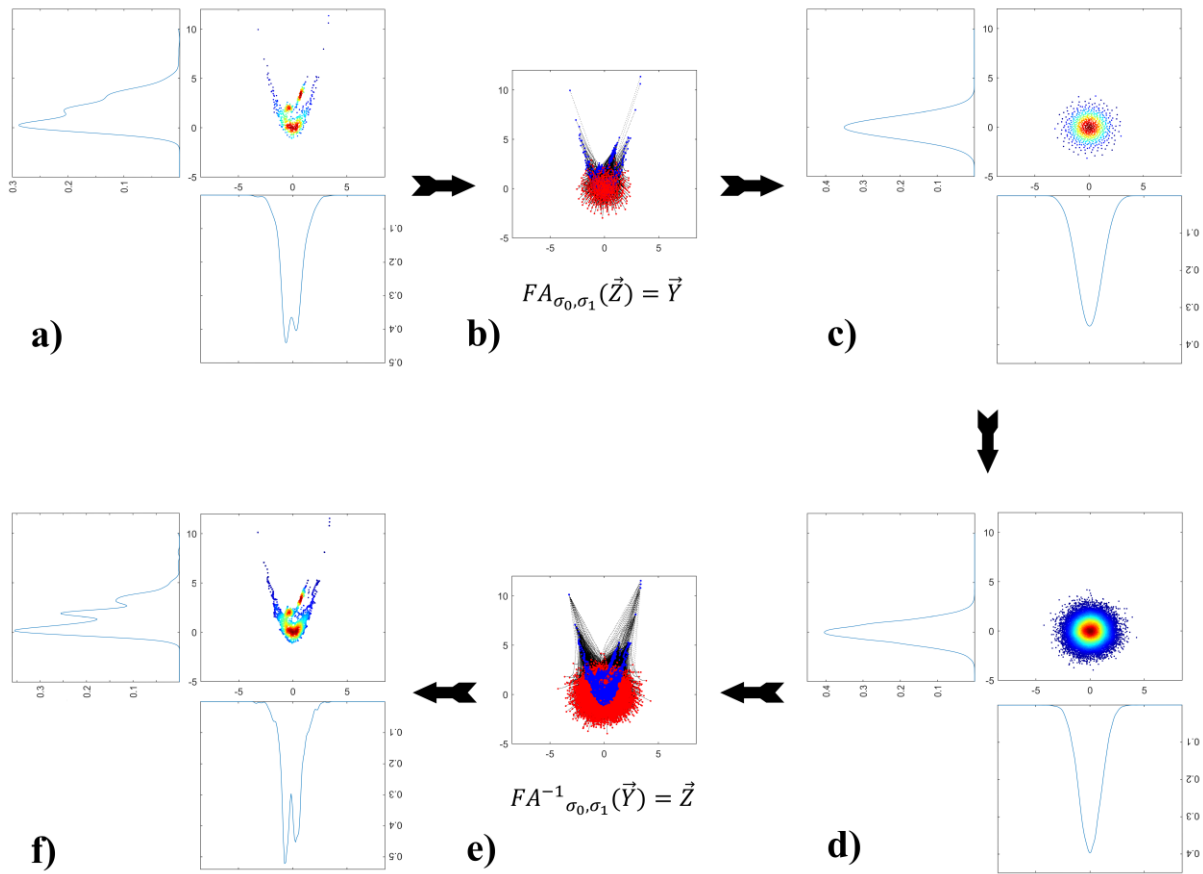
731

732

733

734

735



736

737

Fig. 1 Geostatistical simulation via Flow Anamorphosis

738

739

740

741

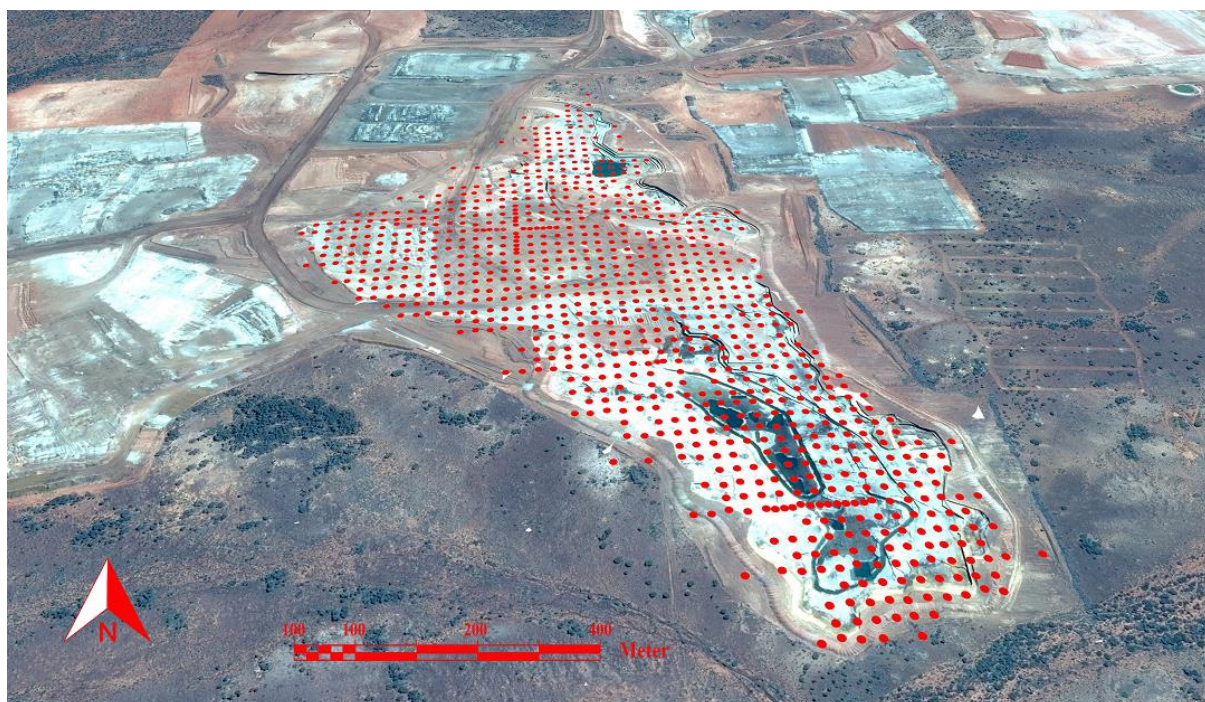
742

743

744

745

746



747

748

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

749

750

751

752

753

754

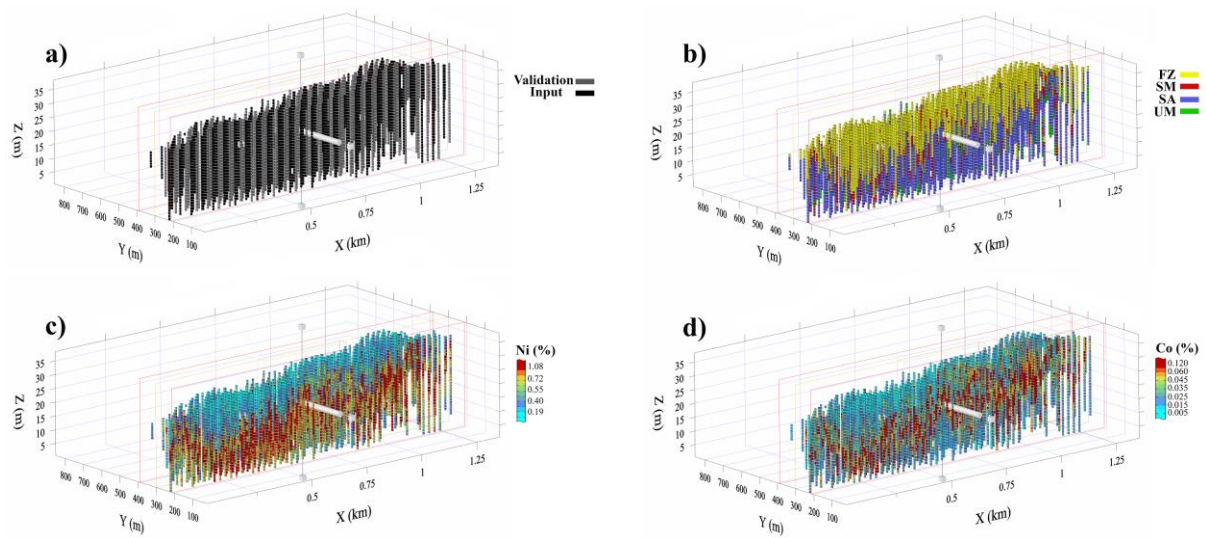
755

756

757

758

759



760

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

763

764

765

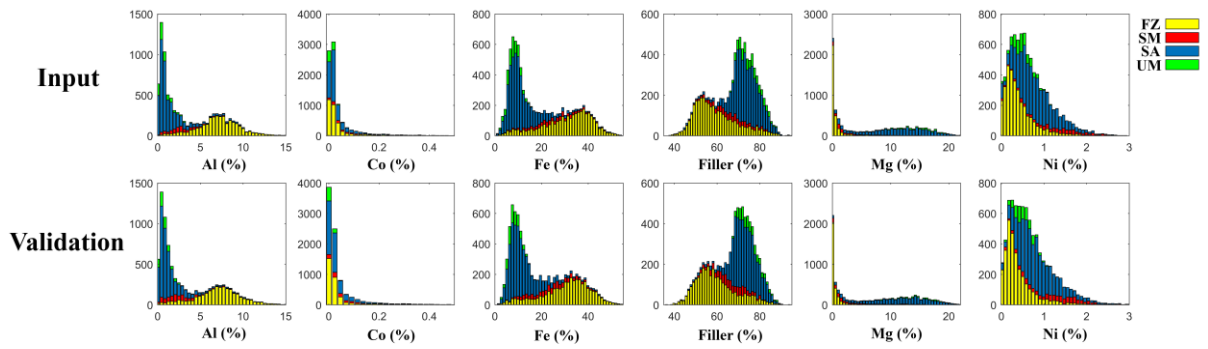
766

767

768

769

770



771

772 **Fig. 4** Histogram of the geochemical components of input and validation sets, coloured by the proportion of rock

773 types in each bin

774

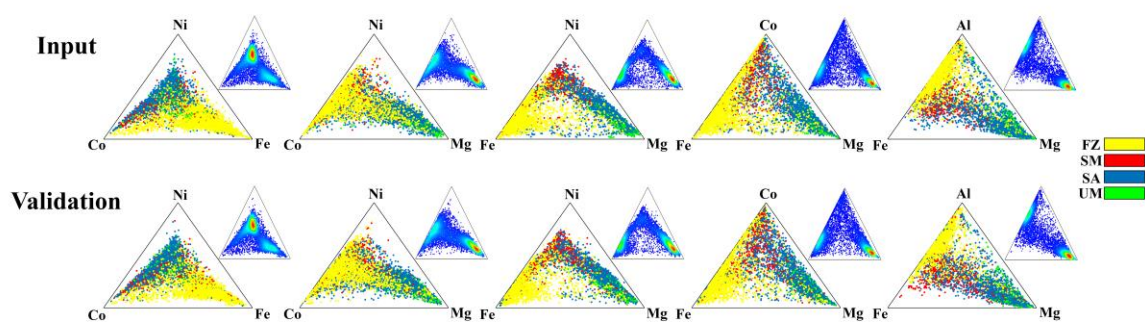
775

776

777

778

779



780

781 **Fig. 5** Ternary diagrams of the geochemical compositions of input and validation sets, coloured by the rock types

782 (large triangles) and kernel density (small triangles)

783

784

785

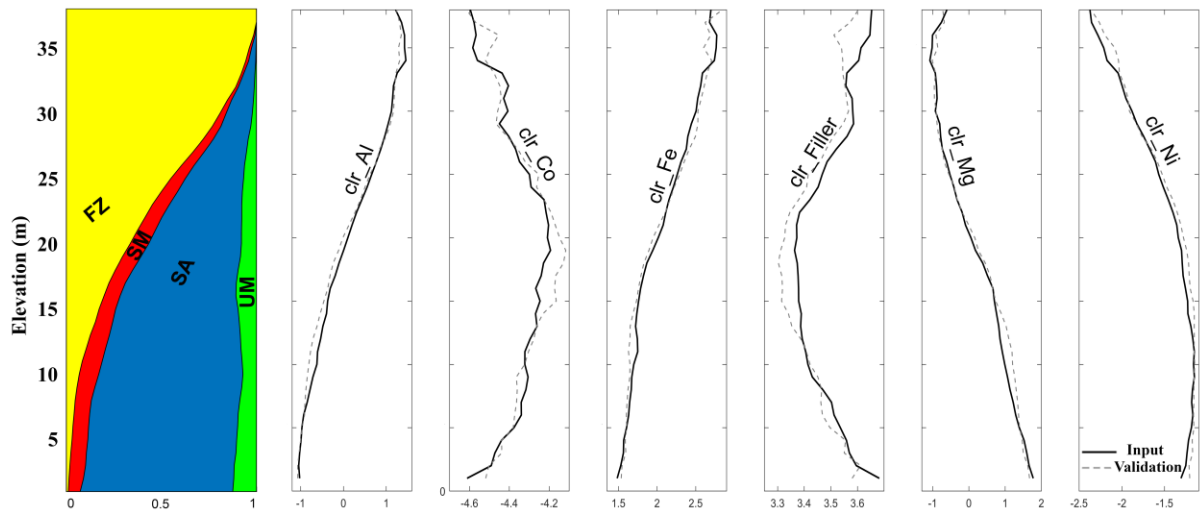
786

787

788

789

790



791

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

793

794

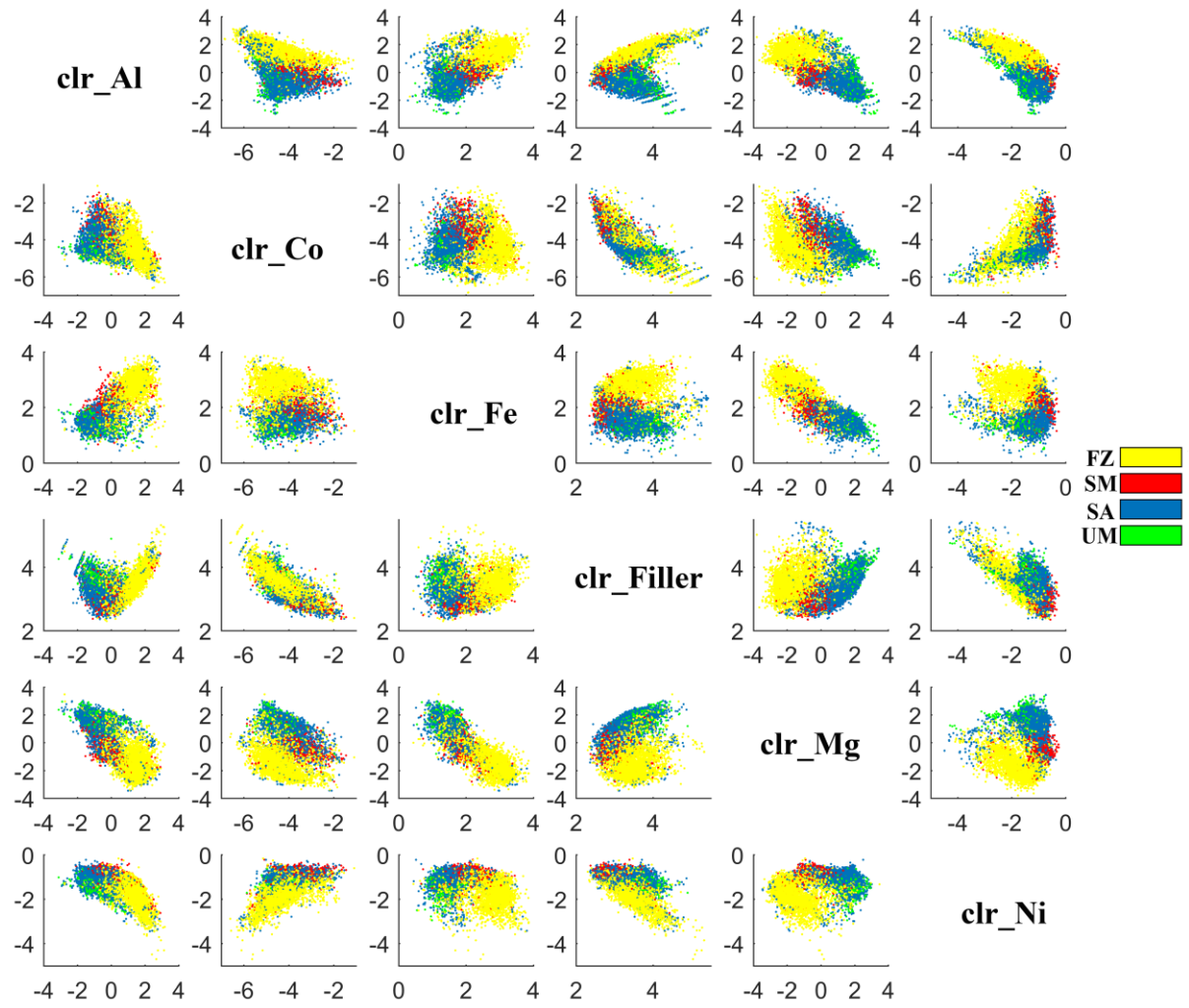
795

796

797

798

799



800

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

803

804

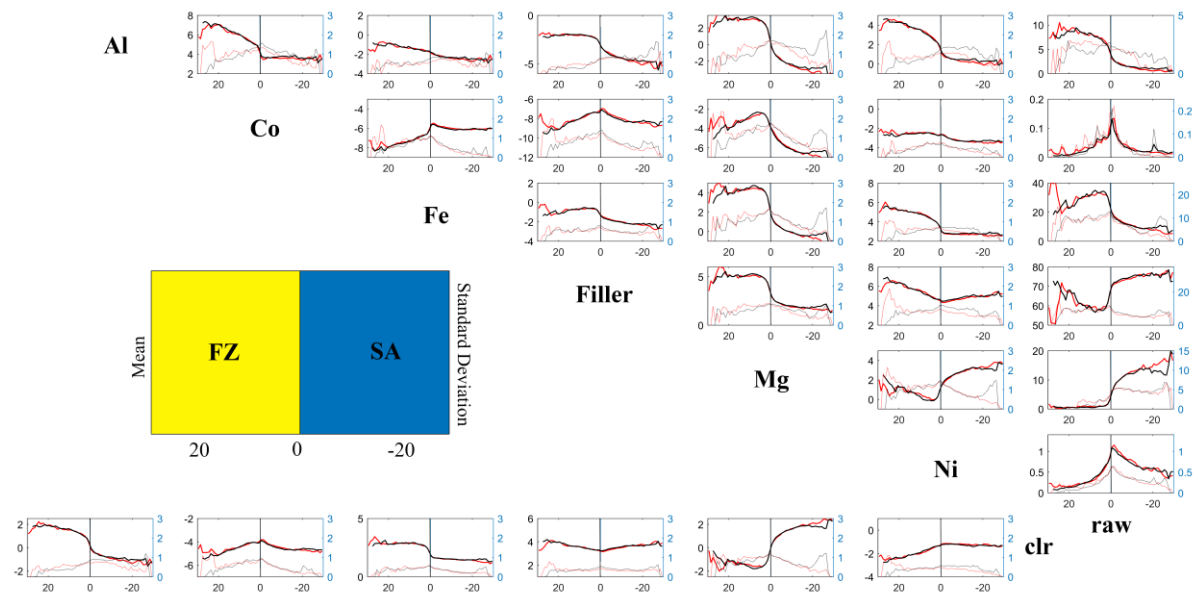
805

806

807

808

809



810

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

814

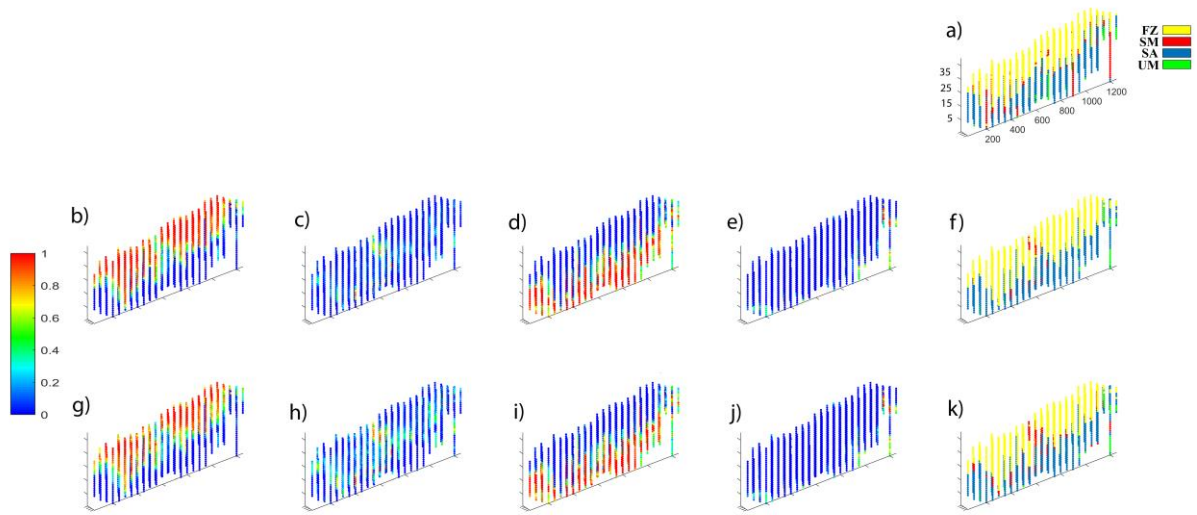
815

816

817

818

819



820

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

824

825

826

827

828

829

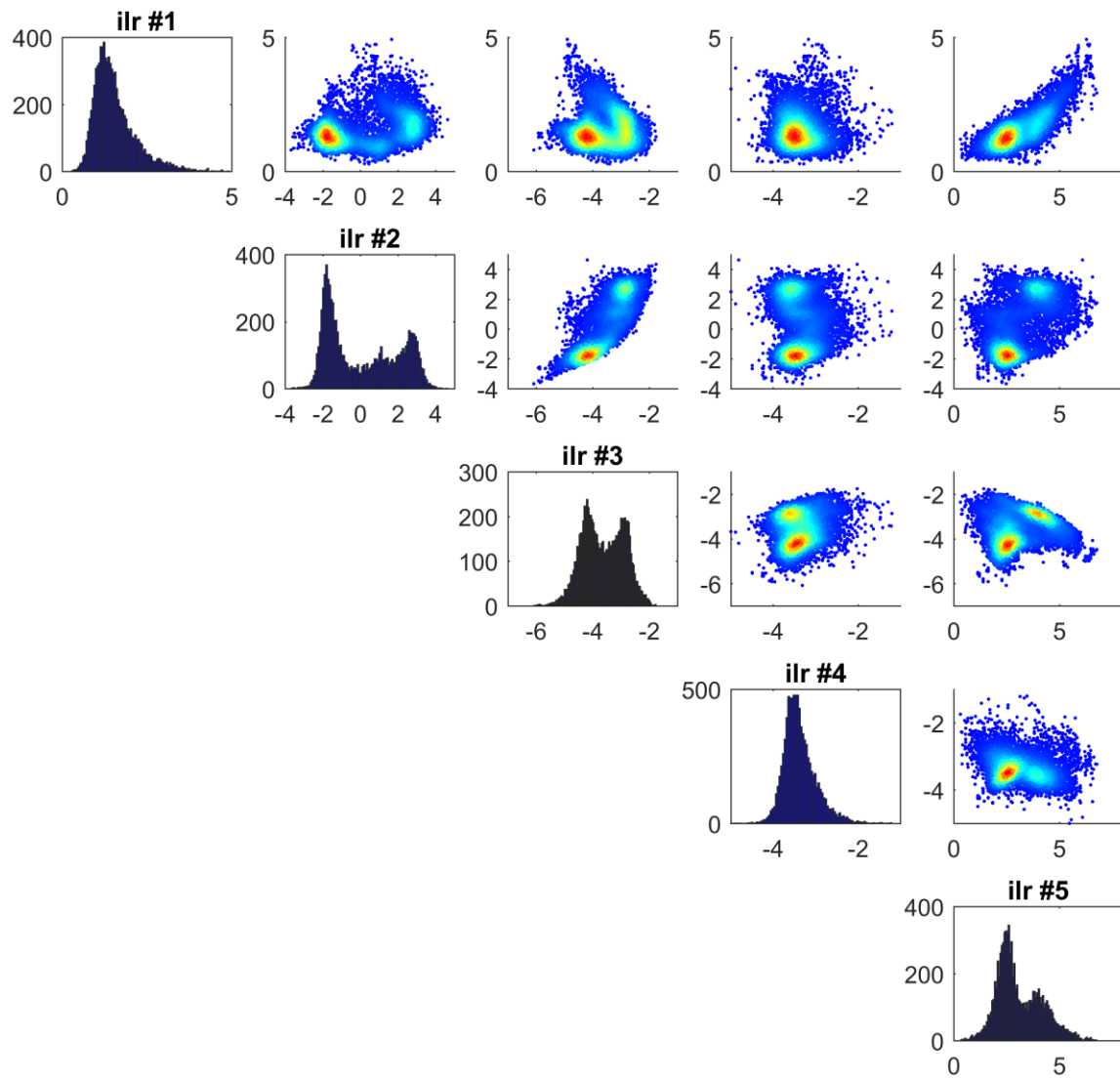
830

831

832

833

834



835

836

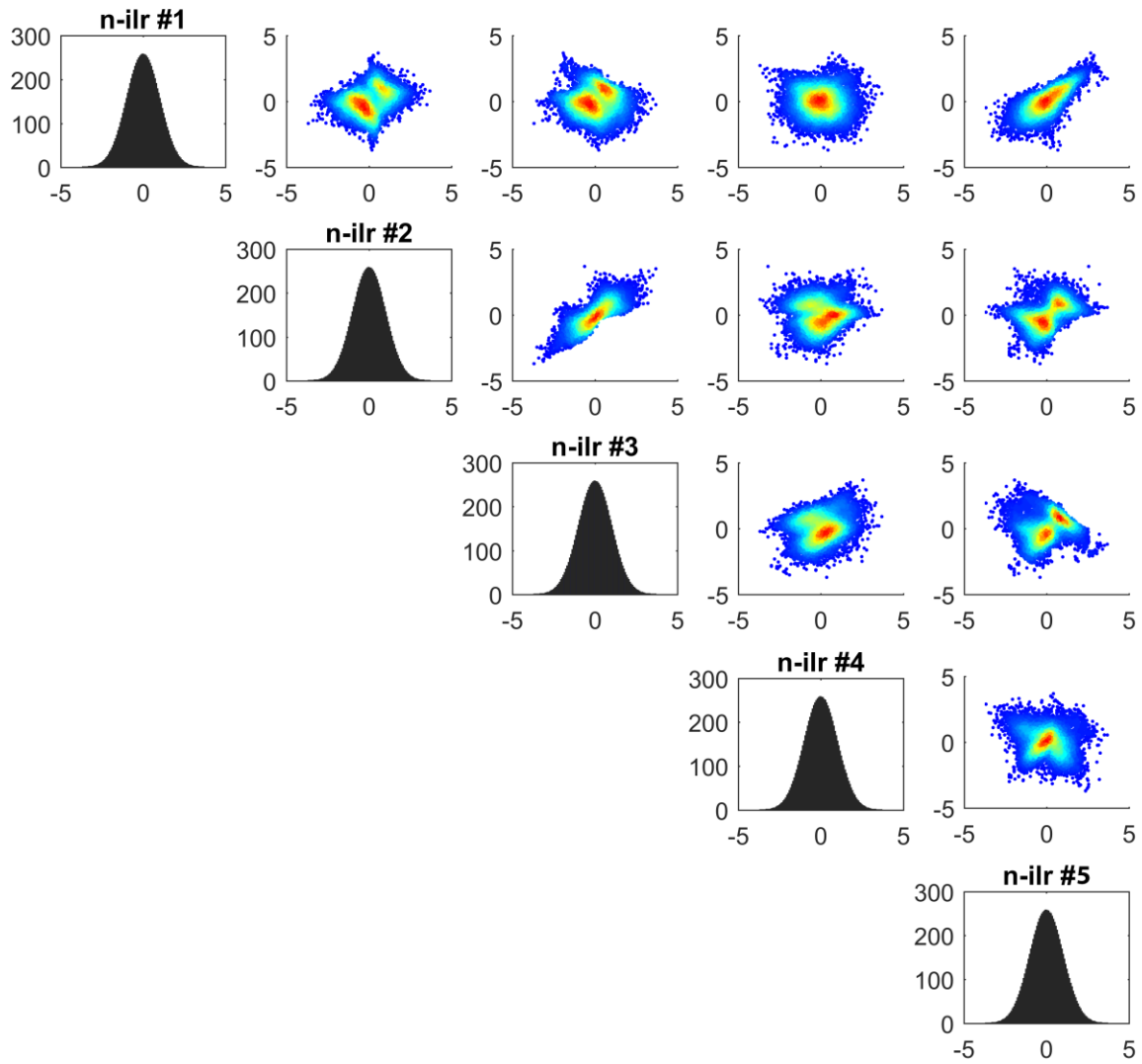
Fig. 10 Histograms and scatterplots of ilr-transformed input data (coloured by kernel density estimate)

837

838

839

840

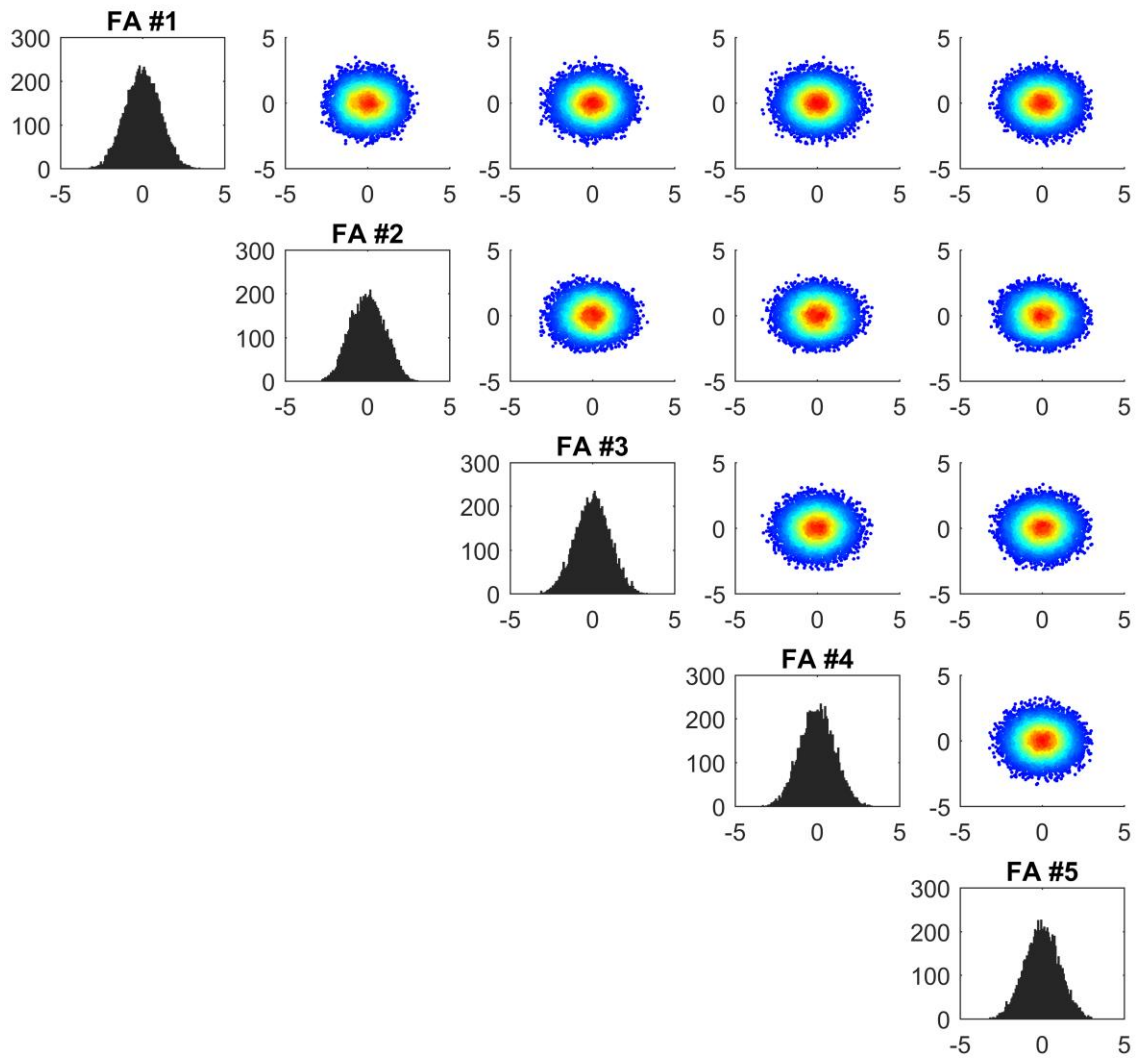


841

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

844

845



846

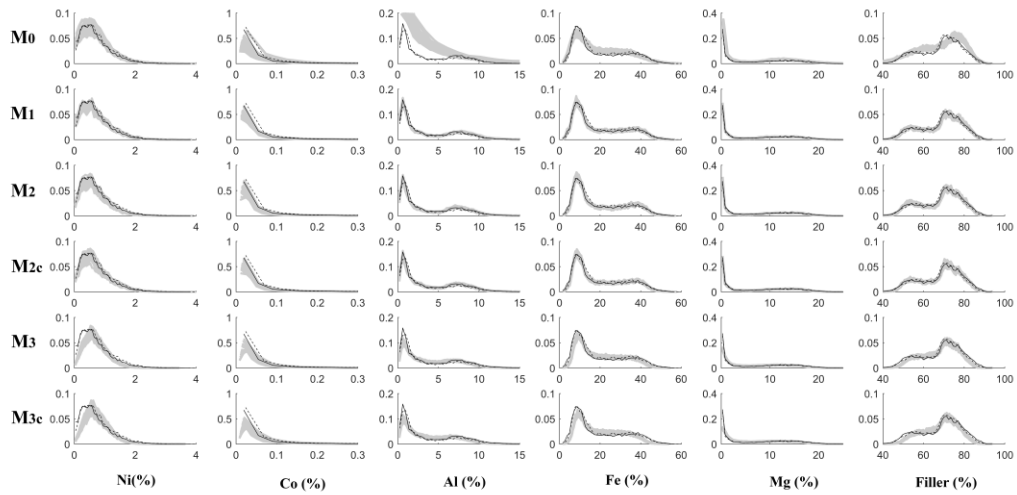
847 **Fig. 12** Histograms and scatterplots of the transformed data to normal space via FA (coloured by kernel density

848 estimate)

849

850

851



852

853 **Fig. 13** Histogram reproduction of the six proposed methods for simulation of geochemical compositions.

854 Continuous black lines are input data, dashed black lines are validation data, and grey lines are realisations

855

856

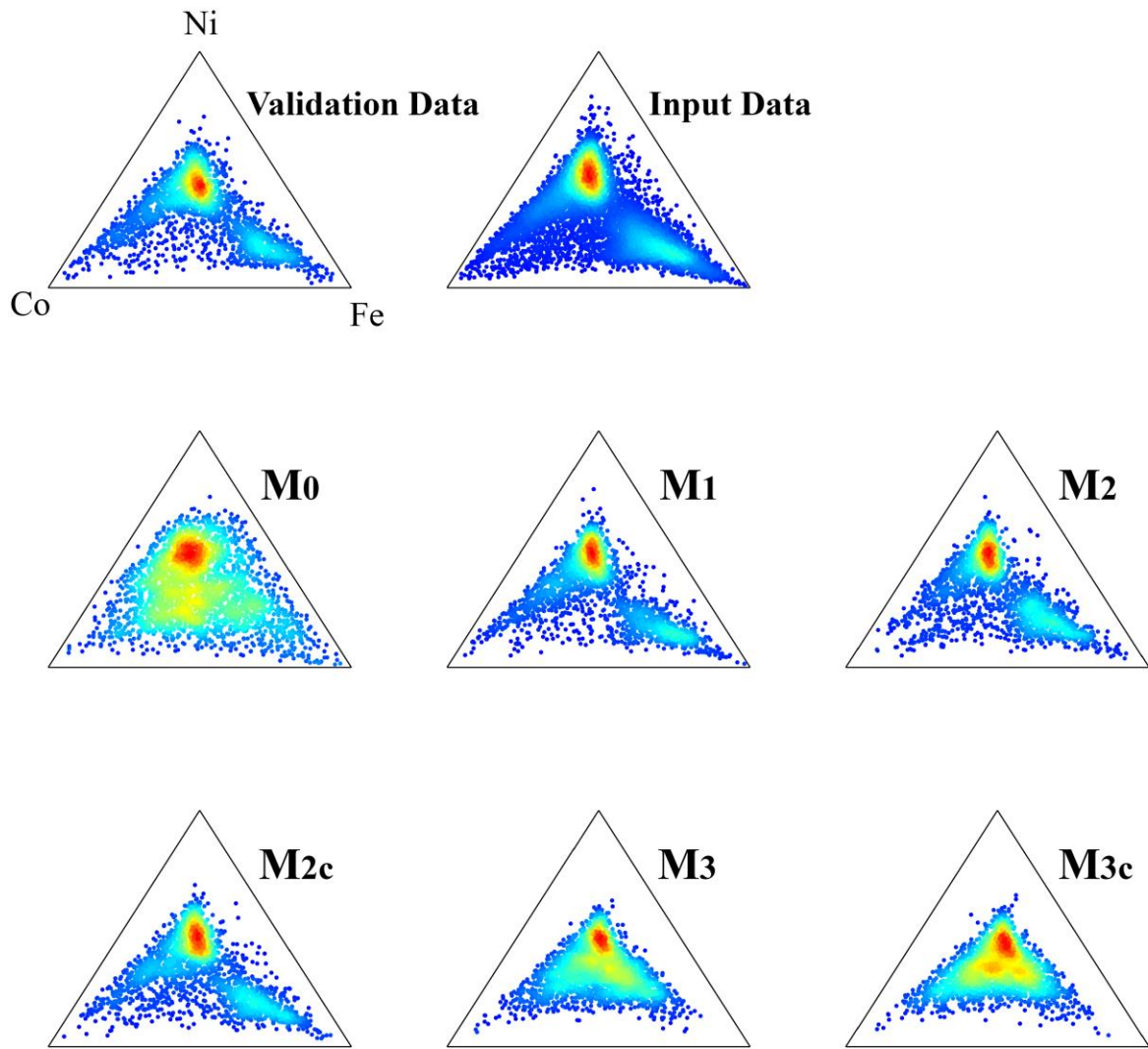
857

858

859

860

861



862

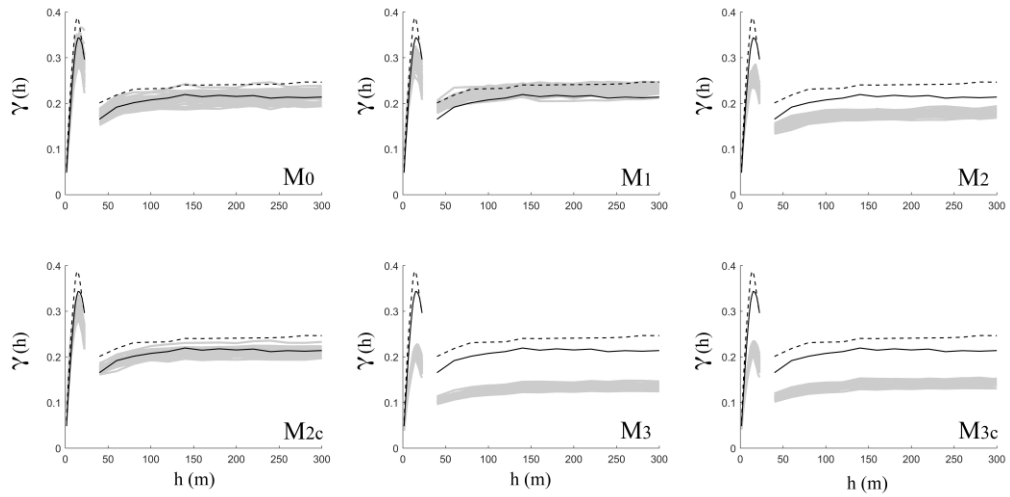
863 **Fig. 14** Ternary diagrams of input and validation data (three components: Ni, Co, Fe), and one realisation
 864 (randomly selected) from each method

865

866

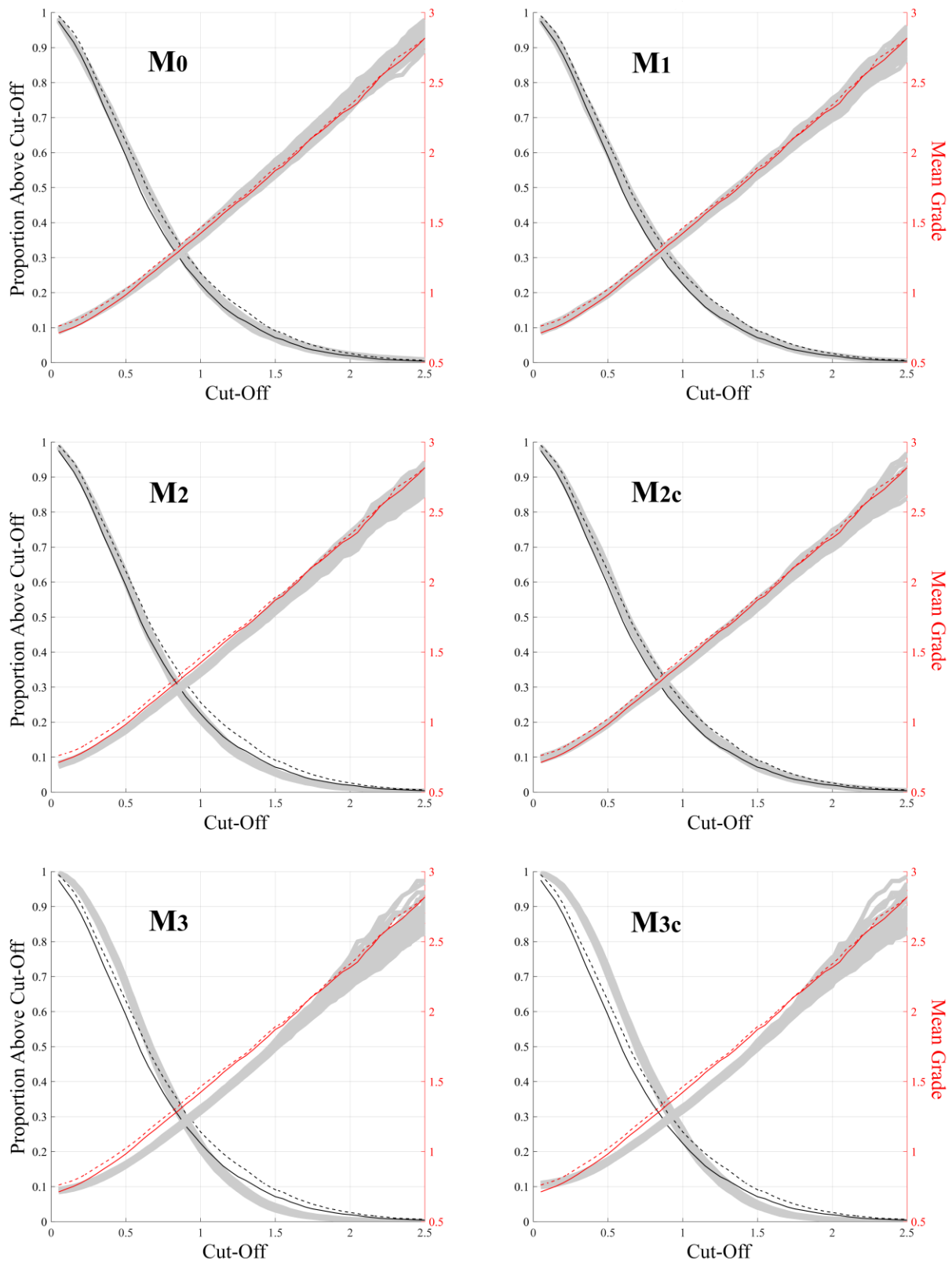
867

868



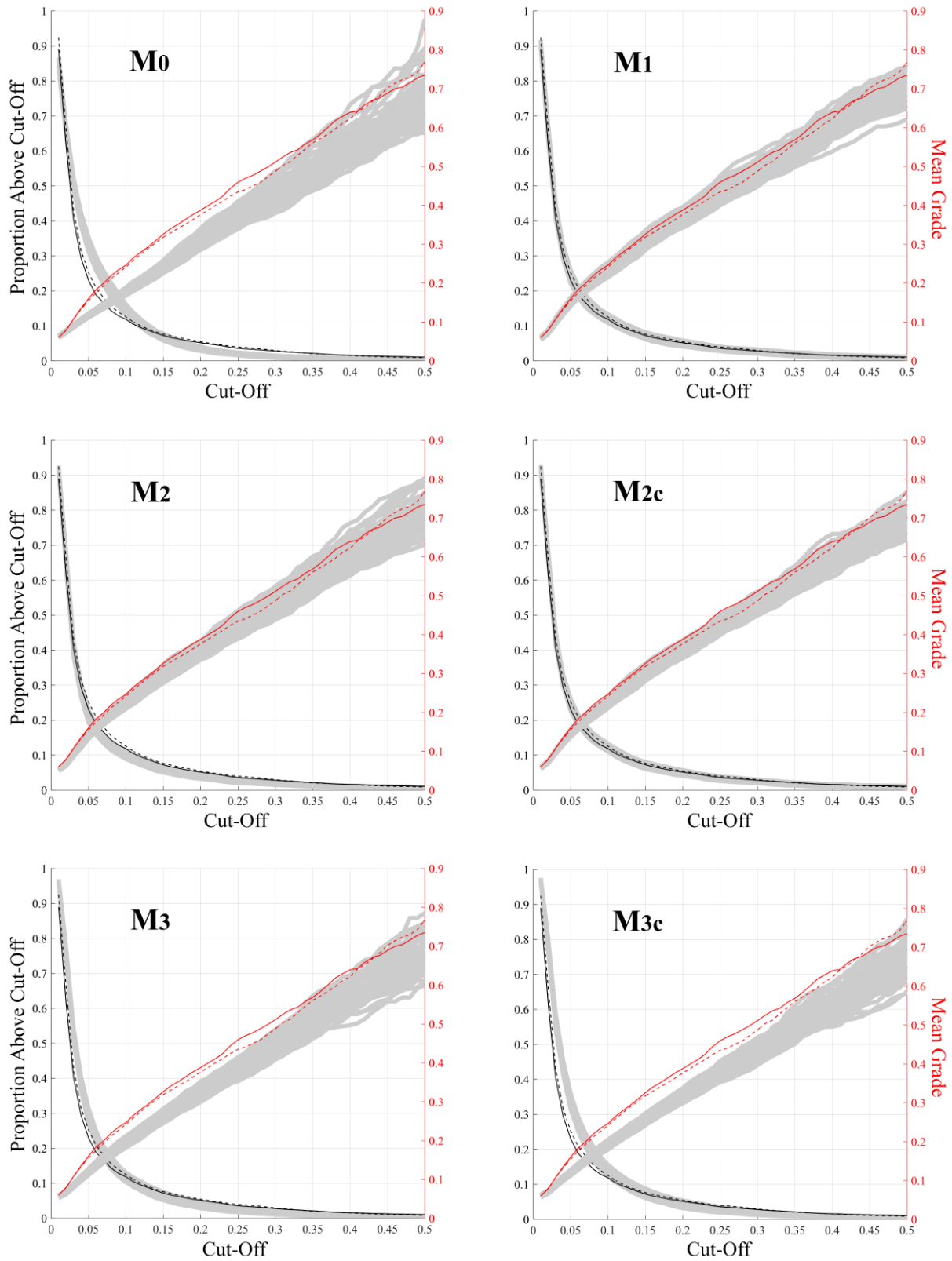
869

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



872

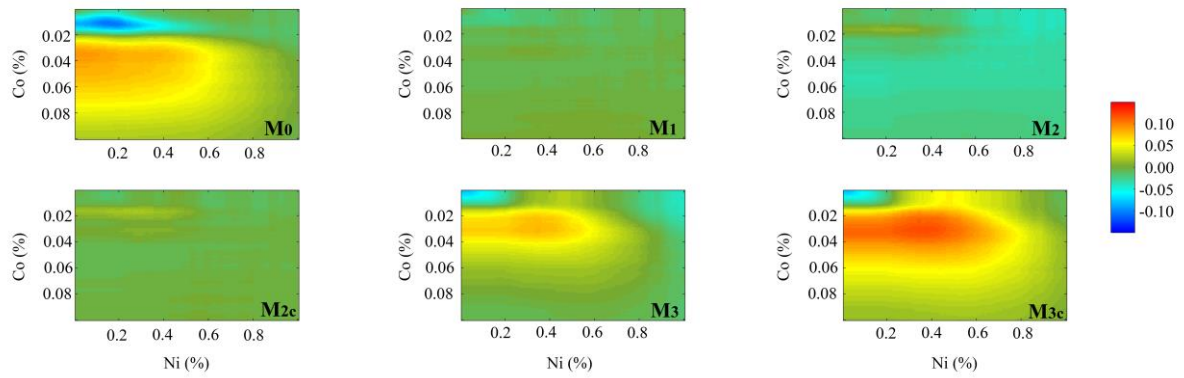
873 **Fig. 16** Grade-tonnage curves (for Ni component) of the six proposed methods. Continuous black lines are the
 874 proportion of samples above Ni cut-offs while continuous red lines are the average grades for input data. Dashed
 875 lines are for validation data while grey lines are different realisations.



876

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

880



881

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

884

885

886

887

888

889

890

891

892

893

894

895

896

897 **List of Tables**

898 Table 1 Proposed methods and the related features 48

899 Table 2 Ore mineralogy of different geological units at MME 49

900 Table 3 Proportions of rock types 50

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

Table 1 Proposed methods and the related features

Code	Geological control	Adjust proportions	Transform to real space	Transform to normal	Factorization	Spatial simulation
M₀	No geological control	-	ilr	GA	MAF	TB
M₁	No geological control	-	ilr	FA	-	TB
M₂	Deterministic geological model	-	ilr	FA	-	TB
M_{2c}	Deterministic geological model	✓	ilr	FA	-	TB
M₃	Probabilistic geological model	-	ilr	FA	-	TB
M_{3c}	Probabilistic geological model	✓	ilr	FA	-	TB

917

918

919

920

921

922

923

924

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	-	-	-	-	-	-	-

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

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

940

941

942