Comput Geosci (2016) 20:1013–1028
DOI 10.1007/s10596-016-9574-3

ORIGINAL PAPER

Quantifying the uncertainty in the spatial layout of rock type domains in an iron ore deposit

Mohammad Maleki[1,2,3] · Xavier Emery[1,2] · Alejandro Cáceres[1,2] · Diniz Ribeiro[4] · Evandro Cunha[4]

Received: 22 May 2015 / Accepted: 9 May 2016 / Published online: 28 May 2016

Abstract This work addresses the problem of delineating the spatial layout of ten rock type domains in an iron ore body and of assessing the uncertainty in the domain boundaries. A stochastic approach is proposed to this end, based on truncated Gaussian simulation, which consists in defining successive partitions of the space that comply with the desired spatial continuity and contact relationships between rock type domains. The sequencing of these domains is driven by their position (surficial vs. underlying rocks), granulometry (compact vs. friable rocks), and grades (rich vs. poor) of iron, alumina, manganese and loss on ignition. A total of 100 realizations are produced, conditioned to available drill hole data, and used to quantify the uncertainty in the occurrence of each rock type domain, at both global and local scales.

✉ Xavier Emery
xemery@ing.uchile.cl

1 Department of Mining Engineering, University of Chile, Santiago, Chile

2 Advanced Mining Technology Center, University of Chile, Santiago, Chile

3 CSIRO-Chile International Center of Excellence in Mining and Mineral Processing, Santiago, Chile

4 Gerência de Recursos Minerais Ferrosos, VALE, Belo Horizonte, Brazil

Keywords Geological uncertainty · Geostatistical simulation · Truncated Gaussian model

1 Introduction

Because of the relationships between geological properties, such as rock types, alteration or mineral zones, and metallurgical properties, modeling the spatial layout of geological domains has become an essential stage prior to the evaluation of mineral resources and ore reserves. Moreover, since the mining industry is increasingly aware of the importance of assessing geological uncertainty and its impact in production schedules ([12] and references therein), there is a growing interest in developing stochastic models of geological domains instead of deterministic models.

To date, different stochastic modeling approaches have been proposed for such a purpose, in which the geological domains are represented through spatial random fields known at a set of sampling locations (e.g., exploration drill holes) but unknown elsewhere. Among these approaches, truncated Gaussian and plurigaussian simulation [3] has received attention in the past decade, especially in the fields of mining [5, 16, 28–32], petroleum [1, 4, 19, 27], and hydrogeology [6, 7, 26]. However, its practical use is often restricted, especially in what refers to the number of geological domains that can be simulated and the contact relationships that can be reproduced [3, 7]. This paper aims to present a hierarchical approach that overcomes these practical restrictions (Section 2), with an application to an iron ore deposit with ten rock types of interest (Section 3).

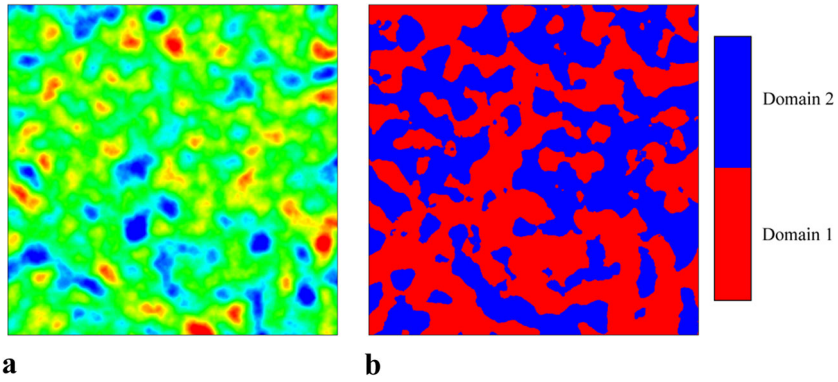

Fig. 1 **a** Simulated Gaussian random field and **b** corresponding domains

2 Methodology

2.1 Truncated Gaussian simulation

The truncated Gaussian model aims to simulate a categorical random field, in which each category represents a specific geological domain, by truncating a Gaussian random field. The number of categories and their proportions directly depend on the number of truncation thresholds and on their values [3]. For instance, in order to simulate two geological domains, each of which covering 50 % of the study area, one needs a single truncation threshold set to 0 (median of the standard Gaussian distribution) (Fig. 1).

The design of the truncated Gaussian model is motivated by its simplicity of use and by the existence of numerous algorithms for simulating Gaussian random fields (at least, in the case of stationary fields) and for conditioning to pre-existing categorical data [21]. Along with the truncation threshold(s), one needs to determine the auto-covariance function of the underlying Gaussian random field in order to complete the determination of the model. This covariance function can be inferred from the covariance functions of the domain indicators, since there is a one-to-one relationship between Gaussian and indicator covariance functions [3]. In practice, instead of covariance functions, one usually works with variograms. If one denotes by y the truncation threshold, by G the cumulative distribution function of the standard normal distribution, and by γ_y and γ the indicator and Gaussian variograms, respectively, then the following relationship holds [9]:

$$\gamma_y = G(y)\,[1 - G(y)] - \frac{1}{2\pi}\int_0^{\arcsin(1-\gamma)} \exp\left\{-\frac{y^2}{1+\sin(\theta)}\right\} d\theta \qquad (1)$$

In practice, the right-hand side of this equation can be calculated by quadrature or by using expansions into Hermite polynomials [13]. Given a truncation threshold y, this allows numerically determining the value of the indicator

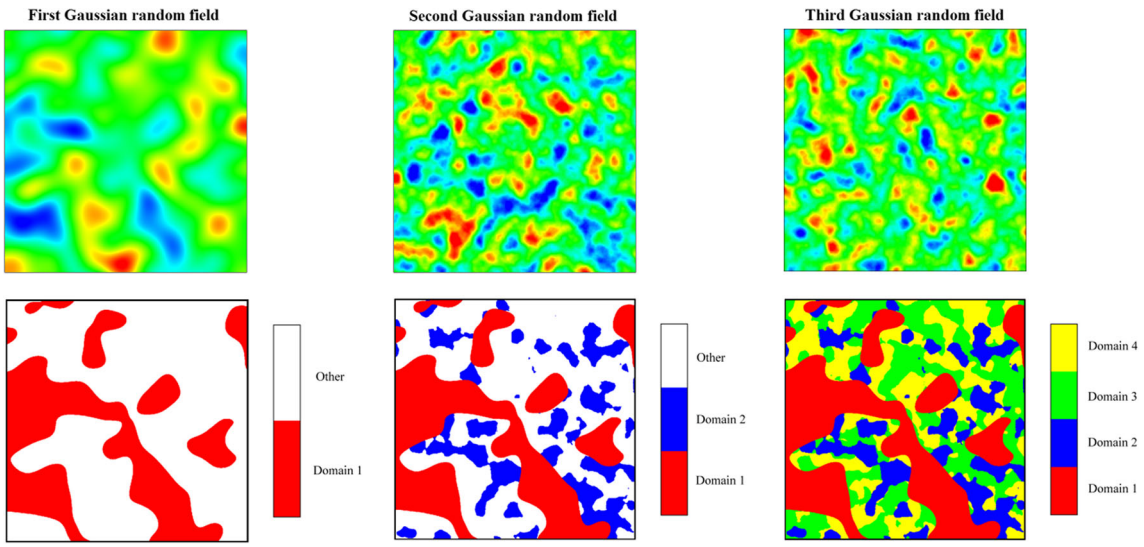

Fig. 2 Simulated Gaussian random fields (*top*) and corresponding geological domains (*bottom*) obtained with the proposed hierarchical approach

Table 1 Codification of categorical data into indicator data

	Indicator of first Gaussian random field	Indicator of second Gaussian random field	Indicator of third Gaussian random field
Domain 1	1	No information	No information
Domain 2	0	1	No information
Domain 3	0	0	1
Domain 4	0	0	0

variogram (γ_y) knowing that of the Gaussian variogram (γ) and vice versa.

When more than two geological domains are considered and there is no natural sequence in the domains, the truncated Gaussian approach is often too restrictive, since it does not allow all the domains being mutually in contact [3]. To overcome such a restriction, the idea is to perform successive applications of this model, as explained next.

2.2 Hierarchical simulation

In order to simulate more than two geological domains, one option is to iteratively apply truncated Gaussian simulation, so that the domains are simulated through a hierarchical process. At each stage, by truncating a new synthetic Gaussian random field (independent of the previous ones), one can split the space into two or more regions, depending on whether one or more truncation thresholds are considered. The regions so obtained can be combined in order to form the domains of interest. For instance (Fig. 2),

- At the first stage (first layer), one defines two regions, corresponding to domain 1 and its complement (Fig. 2 left).
- At the second stage (second layer), one defines, in the complement of domain 1, two regions corresponding to domain 2 and its complement (Fig. 2 middle).
- ...
- The last layer defines two regions, corresponding to domain $(n-1)$ and domain n (Fig. 2 right).

2.3 Inference of model parameters

2.3.1 Sequencing of geological domains

The proposed hierarchical approach is a particular case of plurigaussian simulation, which relies on the truncation of several Gaussian random fields [3]. One of its advantages is that one can control the sequencing of geological domains. In particular, the domain defined in the first layer (Fig. 2 left) appears in the foreground and cross-cuts all the other domains, while the domains defined in the last layer (Fig. 2 right) appear in the background. Based on this observation, the partitioning of space can be done according to geological criteria such as age or origin of the domains to be simulated [23]. In the case study presented in Section 3, the sequencing of the rock types will be driven by considerations on the associated variables (granulometry and grades of elements of interest).

2.3.2 Truncation thresholds

Once the sequencing of geological domains has been defined, the truncation thresholds can be determined in accordance with the proportion of space covered by each domain, in a similar way as in the traditional truncated Gaussian model [3]. For instance, consider the case of four domains with equal proportions (0.25), defined as in Fig. 2. In such a case, the truncation threshold to apply to the first Gaussian random field (Fig. 2 left) is $G^{-1}(1/4) = -0.6745$;

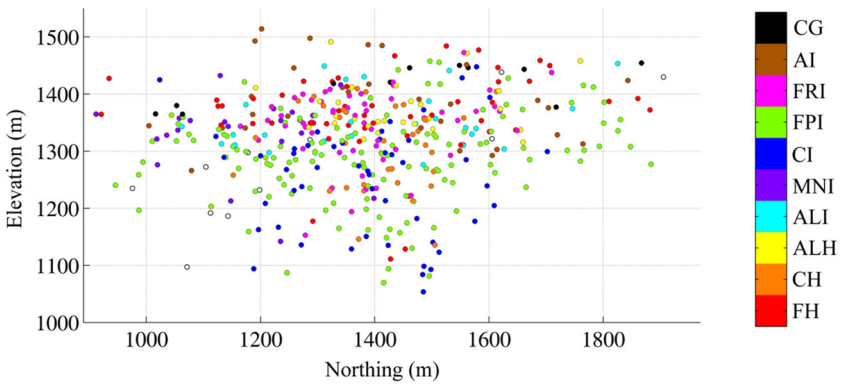

Fig. 3 Cross section showing locations of drill hole data with prevailing rock type

Table 2 Statistics of available data

		Global	FH	CH	ALH	ALI	MNI	CI	FPI	FRI	AI	CG
Number of data		4465	455	308	239	314	204	486	1332	600	239	150
Fe	Mean value	51.76	64.77	66.12	63.80	51.64	49.04	42.64	42.92	56.98	55.41	61.08
	Variance	133.90	10.66	17.40	3.99	68.75	68.11	76.57	48.21	18.88	41.63	15.25
	Minimum	9.92	37.13	40.03	51.10	29.67	26.92	25.56	24.32	41.80	29.07	48.60
	Maximum	69.70	69.35	69.53	68.78	66.29	66.53	68.32	69.31	68.32	65.20	68.65
Al	Mean value	1.31	1.24	0.73	2.21	2.22	1.96	0.47	0.81	1.09	2.62	3.97
	Variance	1.66	0.57	0.42	1.42	1.89	1.87	0.30	0.29	0.61	3.81	5.63
	Minimum	0.10	0.13	0.11	0.21	0.12	0.16	0.10	0.10	0.13	0.11	0.55
	Maximum	12.60	5.54	4.31	9.62	9.35	6.74	8.37	3.31	6.47	12.60	12.30
Granulometric	Mean value	33.10	23.77	72.18	33.42	24.61	25.39	60.52	24.87	22.42	35.04	45.50
fraction	Variance	552.27	294.92	477.09	492.75	311.39	309.53	255.25	242.49	216.89	396.58	304.04
	Minimum	0.12	0.97	1.77	0.63	0.61	0.69	0.79	0.30	0.85	2.27	3.76
	Maximum	98.18	94.79	98.18	92.70	90.41	87.62	96.52	95.75	91.23	88.97	86.41
Mn	Mean value	0.30	0.13	0.04	0.23	0.14	3.54	0.10	0.09	0.09	0.13	0.07
	Variance	1.94	1.36	0.03	0.40	0.05	20.61	0.62	0.04	0.05	0.04	0.04
	Minimum	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	Maximum	32.06	24.73	2.88	4.54	2.17	32.06	17.12	3.10	3.69	1.30	1.66
LOI	Mean value	2.25	1.78	1.28	4.01	3.63	3.01	1.11	1.35	1.56	6.62	5.92
	Variance	4.06	1.14	1.11	3.37	1.63	3.33	0.94	0.81	1.17	3.28	8.61
	Minimum	0.06	0.20	0.13	0.62	0.18	0.22	0.06	0.10	0.10	1.87	0.56
	Maximum	13.84	7.08	8.02	11.04	10.37	9.34	9.46	9.02	8.77	13.84	12.87

the truncation threshold to apply to the second Gaussian random field (Fig. 2 middle) is $G^{-1}(1/3) = -0.4307$, while the truncation threshold to apply to the third Gaussian random field (Fig. 2 right) is $G^{-1}(1/2) = 0$.

2.3.3 Variogram analysis

The first stage to variogram analysis is to convert each categorical data, which indicates the domain prevailing at a sampling location, into indicator data associated with each Gaussian random field and truncation threshold, depending on whether or not this field is below the threshold. An example with the three Gaussian random fields of Fig. 2 is given in Table 1, in which the original categorical data are codified into three possible indicator values: 0, 1, or unknown.

The second stage to variogram analysis is to calculate, for each Gaussian random field, the experimental variogram of the associated indicator data and to transform such an indicator variogram into a variogram for the Gaussian random field itself, using the one-to-one relationship linking both variograms (1). When several truncation thresholds apply to the same Gaussian random field, several indicators can be defined, for each of which one can calculate an experimental variogram for the Gaussian random field;

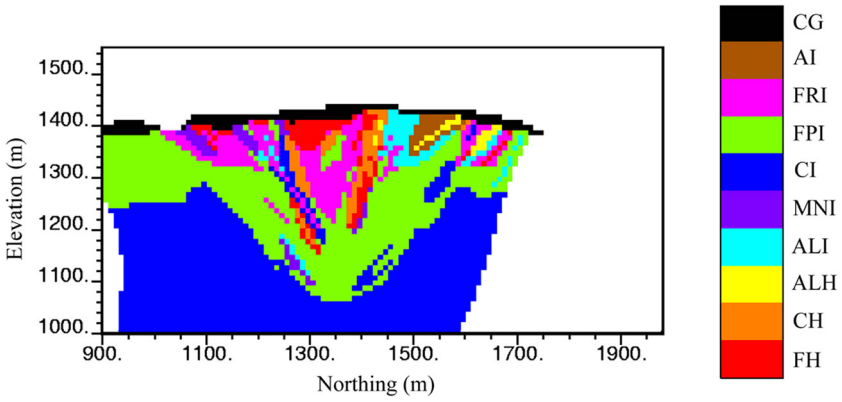

Fig. 4 Cross section showing interpreted rock type model

Table 3 Splitting into canga vs. other ferruginous rocks

Rock group	Rock codes	Relative proportions
Canga	CG	0.035
Other ferruginous rocks	HF, HAL, HC, IF, IFR, IC, IA, IAL, IMN	0.965

Table 4 Splitting into compact vs. friable rocks

Rock group	Rock codes	Relative proportions
Compact	CH, CI	0.19
Friable	FH, ALH, ALI, MNI, FPI, FRI, AI	0.81

these experimental variograms are then averaged in order to stay with a single experimental variogram for the Gaussian random field.

Once the experimental variograms of all the Gaussian random fields are obtained, they have to be fitted with permissible variogram models with a unit sill, which completes the variogram analysis stage. This can be done manually or by use of automatic fitting algorithms under a constraint on the total sill value [14, 22].

2.4 Simulation algorithm

The simulation of the geological domains can be performed through the following steps [3, 21]:

1. Simulate the underlying Gaussian random fields at the data locations, conditionally to the categorical data, which provide inequality constraints on the Gaussian values to be simulated. This step can be realized by a Markov chain Monte Carlo algorithm known as the Gibbs sampler. For each Gaussian random field, the sampler is initialized by assigning values that are consistent with the inequality constraints at the data locations. Then, each location is visited successively and its value is updated conditionally to the inequality constraints and to the values at the other locations. It can be shown [18] that, as the number of updates increases, the simulated values converge in distribution to the desired Gaussian random fields conditioned to the inequality constraints.
2. Simulate the Gaussian random fields at the target locations, which can be realized by any multivariate Gaussian simulation algorithm. In this work, a turning bands algorithm is used, for which the simulated fields are mixtures of piecewise linear functions, piecewise quadratic functions, or weighted cosine waves associated with random frequencies and random phases [15, 17, 21]. These simulated fields are subsequently conditioned to the Gaussian data obtained in the previous stage by means of conditioning kriging, a procedure also known as substitution of residuals [9, 11].
3. Truncate the conditioned Gaussian random fields to obtain the simulated geological domains.

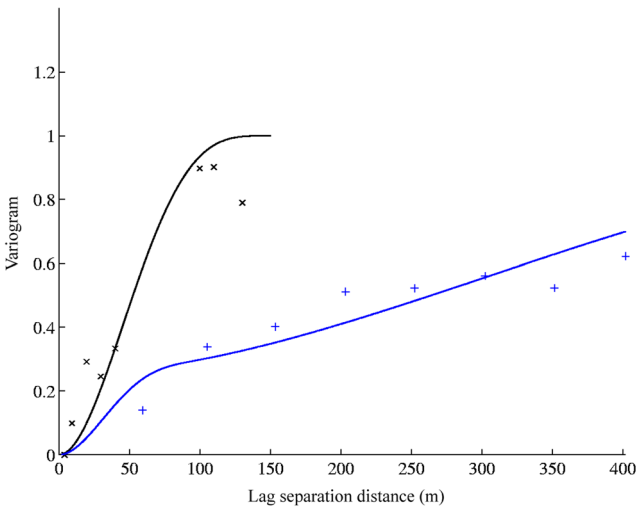

Fig. 5 Experimental (*crosses*) and modeled (*solid line*) variograms for the first Gaussian random field, along vertical (*black*) and horizontal (*blue*) directions

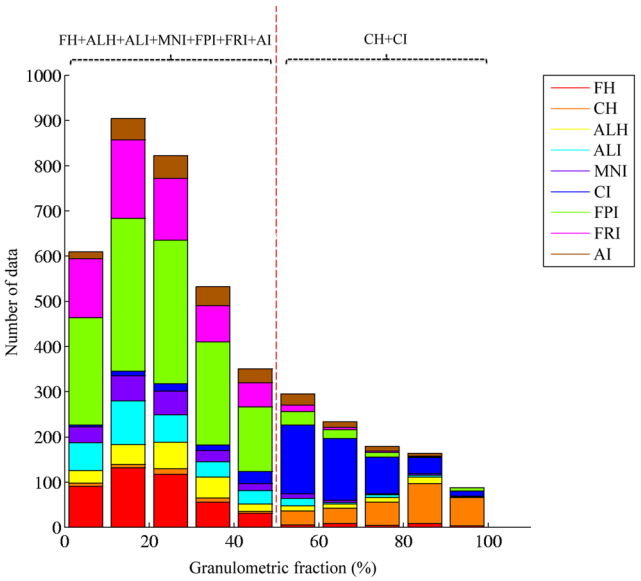

Fig. 6 Histogram of granulometric fraction for rock types considered in the second layer

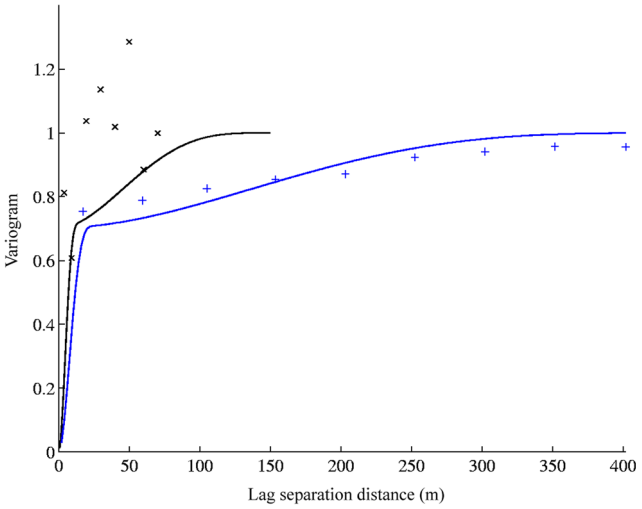

Fig. 7 Experimental (*crosses*) and modeled (*solid line*) variograms for the second Gaussian random field, along vertical (*black*) and horizontal (*blue*) directions

3 Case study

3.1 Presentation of the data

The previous hierarchical approach is now applied to an iron deposit recognized by a set of exploration drill holes (Fig. 3). For confidentiality reasons, the name and location of the deposit are not disclosed. The available data consists of 4465 samples with information on the grades (expressed in percent) of four elements of interest (iron, alumina, manganese and loss on ignition), the granulometric fraction (between 0 and 100 %) above 6.3 mm, and the dominant rock type, which is classified into 11 categories:

- Friable hematite (FH)
- Compact hematite (CH)
- Alumina-rich hematite (ALH)
- Alumina-rich itabirite (ALI)
- Manganese-rich itabirite (MNI)
- Compact itabirite (CI)
- Friable poor itabirite (FPI)
- Friable rich itabirite (FRI)
- Amphibolitic itabirite (AI)
- Canga ore (CG)
- Waste.

Hematite is an oxide-facies formation, while itabirite is a laminated, metamorphosed oxide-facies formation in which the iron is present as thin layers of hematite, magnetite or martite. The itabirites were weathered in different scales, resulting in rocks with a variety of grain size partition and iron enrichment zones. The definitions of the friable/compact or rich/poor subgroups are related to the observed granulometric fraction and/or the grades of the main elements, as it will be explained later. On the other hand, the surficial canga ore occurs as a layer of concretional material composed mainly of goethite derived from the weathering of iron formation. The main statistics of the grade variables are indicated in Table 2.

Also, a lithological model with 219 × 108 × 55 grid nodes spaced 10 m × 10 m × 10 m was provided, with information on the rock type interpreted by geologists (Fig. 4). Although this model has been constructed in agreement with the drill hole data and the geological knowledge of the deposit, it is of interest to assess the uncertainty in the spatial layout of the ten ore rock types within the domain under study (colored area in Fig. 4). Outside this domain, the material is assumed to be waste and of no interest.

3.2 Truncated Gaussian modeling

To simulate the ten ore rock types, a hierarchical model is proposed, in which the space is successively partitioned into subdomains, based on the truncation of synthetic Gaussian random fields that relate to the granulometry or to the grade of an element of interest. In the following, the waste domain is assumed perfectly known and will not be simulated. The next subsections provide details on the model. The relative proportions presented in these subsections were calculated from the interpreted geological model.

3.2.1 First layer (surficial/underlying rocks)

The ore domain can be split into two subgroups according to the rock type origin: on the one hand, an overlying conglomerate or canga (CG) and, on the other hand, the remaining ferruginous rock types (Table 3). The truncation

Table 5 Splitting into hematite vs. itabirite

	Rock group	Rock codes	Relative proportions
Compact	Compact hematite	CH	0.388
	Compact itabirite	CI	0.612
Friable	Friable poor itabirite	FPI	0.3937
	Friable rich itabirite	ALI, MNI, FRI, AI	0.4012
	Friable hematite	FH, ALH	0.2051

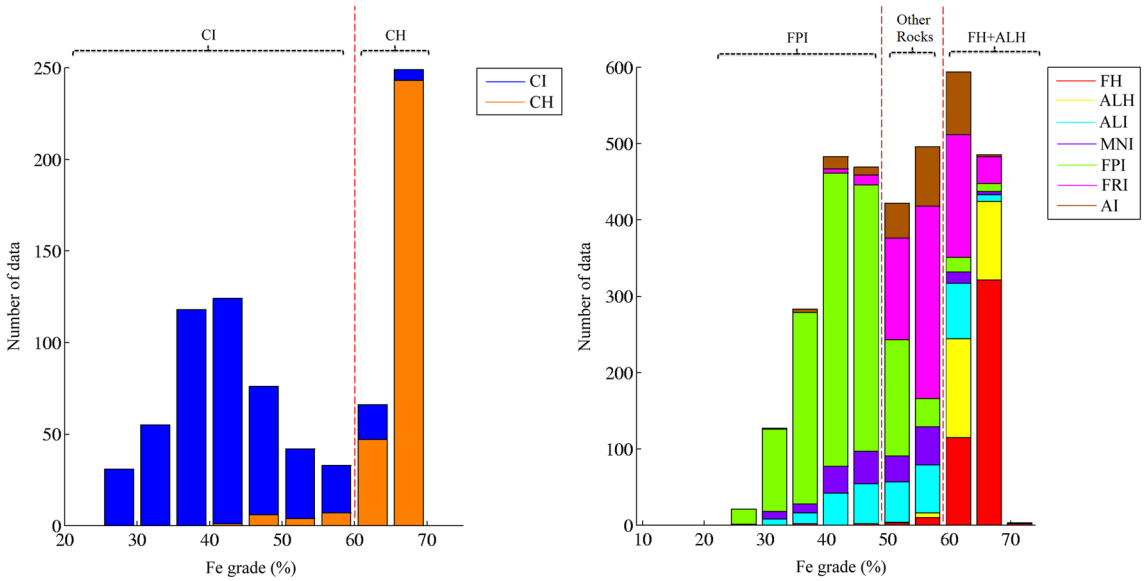

Fig. 8 Histograms of iron grade for rock types considered in the third layer

threshold for this first layer is $G^{-1}(0.035) = -1.81$. The experimental variogram of the associated Gaussian random field, calculated along the vertical and horizontal directions as explained in Section 2.3.3, is displayed in Fig. 5, together with the model fitted with a least-square algorithm [14].

3.2.2 Second layer (granulometry)

The second layer consists in splitting the ferruginous rock group into compact vs. friable rocks (Table 4). Although not strictly identical, the underlying Gaussian random field to truncate is related to the granulometric fraction above 6.3 mm (Fig. 6). Based on the relative proportions of compact and friable rocks, the truncation threshold is set to $G^{-1}(0.19) = -0.723$. As for the first layer, variogram analysis has been held along the horizontal and vertical directions (Fig. 7).

3.2.3 Third layer (iron grade)

Each of the previous groups (compact and friable) is now split according to iron mineralization (Table 5). In the case of compact rocks, one can distinguish hematite (iron grade mostly above 62 %) vs. itabirite (iron grade below 62 %). In the case of friable rocks, two subgroups of itabirite can be identified: iron-rich itabirite (iron grade between 52 and 62 %) and iron-poor itabirite (iron grade below 52 %). All these subgroups are defined by the truncation of a single Gaussian random field, which is related (although not identical) to

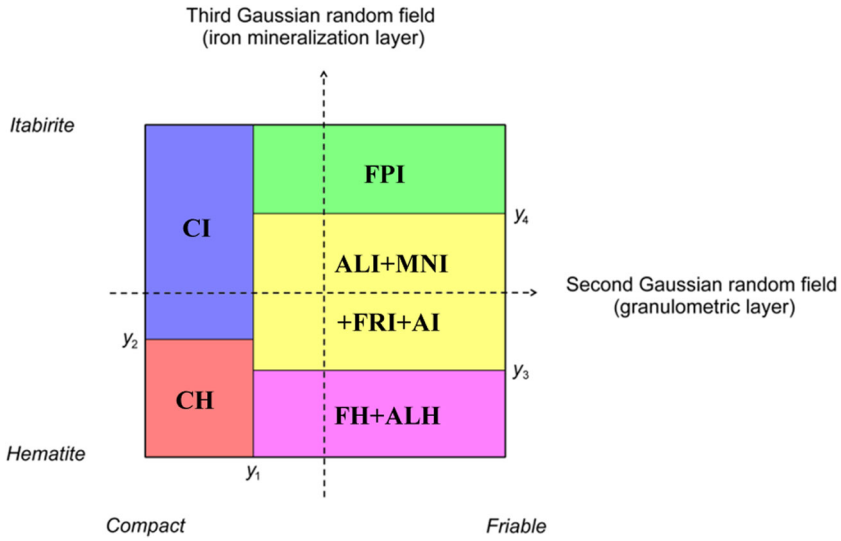

Fig. 9 Contact relationships between rock types for the second and third layers

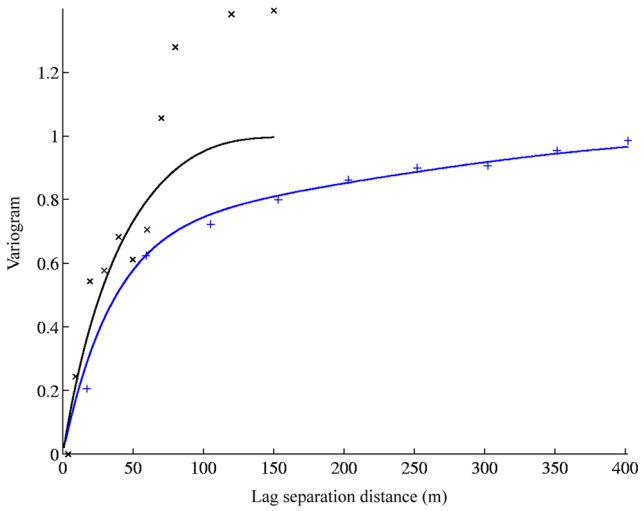

Fig. 10 Experimental (*crosses*) and modeled (*solid line*) variograms for the third Gaussian random field, along vertical (*black*) and horizontal (*blue*) directions

the iron grade (Fig. 8). Note that, with such truncations, the friable poor itabirite has no contact with the friable hematite. Also, using a single Gaussian random field in the third layer for splitting both the friable and compact groups implies that compact itabirite is likely to be in contact with friable itabirite, while compact hematite is likely to be in contact with friable hematite, thus avoiding contacts between compact itabirite and friable hematite or between friable itabirite and compact hematite (Fig. 9).

According to the relative proportions indicated in Table 5, three truncation thresholds have to be defined: the first one, related to compact rocks, is $G^{-1}(0.388) = 0.3418$, while the second and third ones, related to friable rocks, are $G^{-1}(0.2051) = -0.8234$ and $G^{-1}(0.5988) = 0.2504$. The variogram of the underlying Gaussian random field is displayed in Fig. 10.

3.2.4 Fourth layer (alumina grade)

The alumina content is used as the next criterion for splitting the previously defined friable hematite and friable rich itabirite groups (Fig. 11). Again, a single Gaussian random field (related to the alumina grade) is used to define the subgroups (Fig. 12). Based on the relative proportions given in Table 6, the truncation thresholds for this layer are $G^{-1}(0.3444) = -0.4$ and $G^{-1}(0.4422) = -0.1455$, respectively. The variogram of the Gaussian random field to truncate is displayed in Fig. 13.

3.2.5 Fifth layer (manganese grade)

In the fifth layer, the other friable itabirite is divided into two subgroups according to the manganese content, defining a manganese-rich itabirite (with manganese grade above 1 %) and manganese-poor itabirite (with manganese grade less than 1 %) (Fig. 14). Based on the relative proportions given in Table 7, the truncation threshold is $G^{-1}(0.2695) = -0.6144$. The variogram of the Gaussian random field to truncate is displayed in Fig. 15.

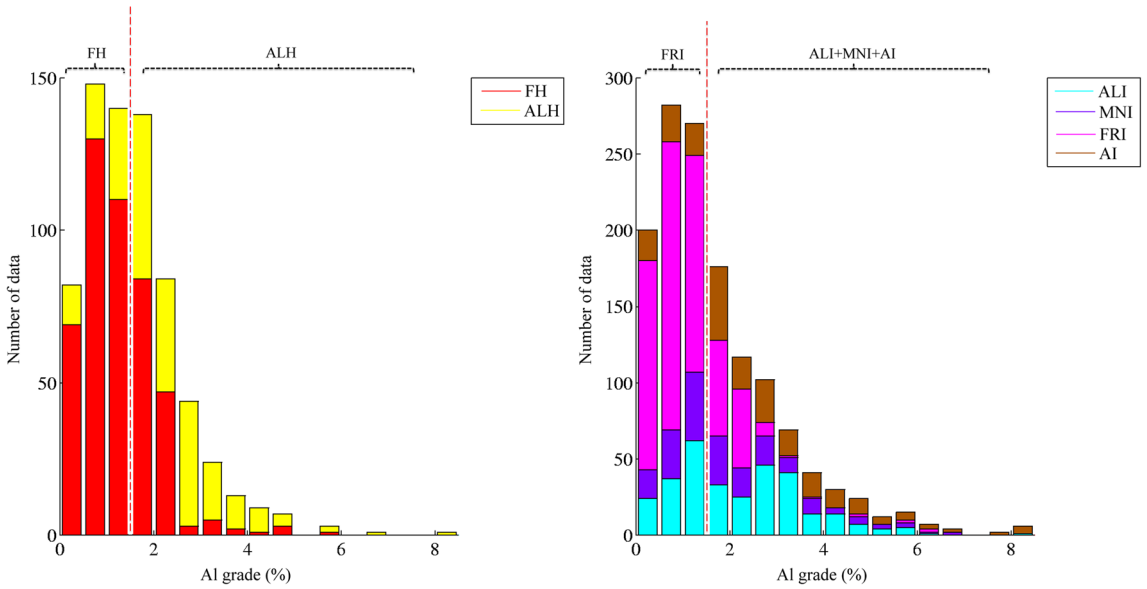

Fig. 11 Histograms of alumina grade for rock types considered in the fourth layer

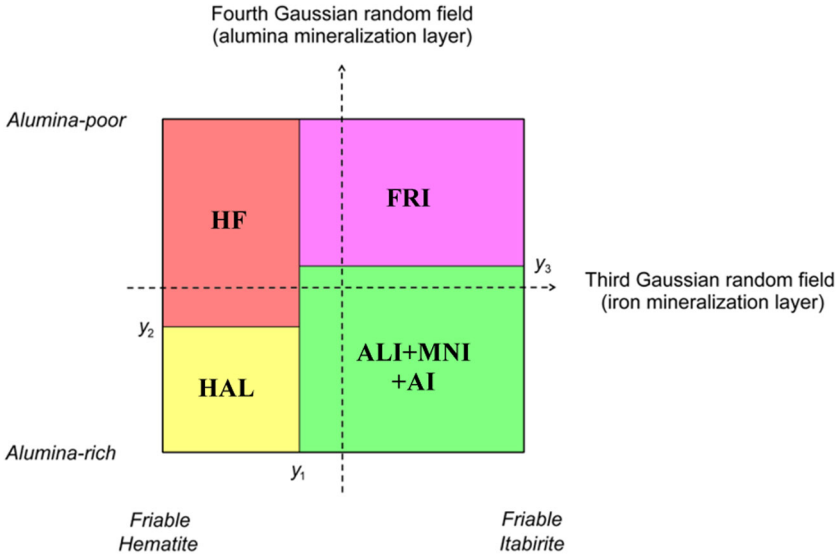

Fig. 12 Contact relationships between rock types for the third and fourth layers

Table 6 Splitting friable hematite and friable itabirite

	Rock group	Rock codes	Relative proportions
Hematite	Alumina-rich hematite	ALH	0.3444
	Friable hematite	FH	0.6556
Itabirite	Friable rich itabirite	FRI	0.4422
	Other friable itabirite	ALI, MNI, AI	0.5578

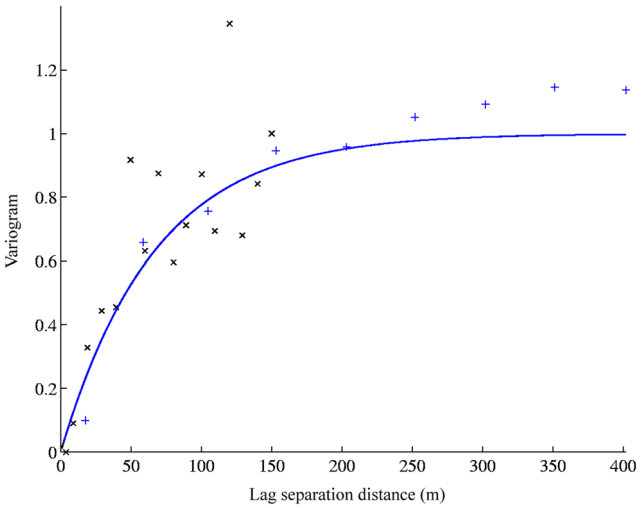

Fig. 13 Experimental (*crosses*) and modeled (*solid line*) variograms for the fourth Gaussian random field

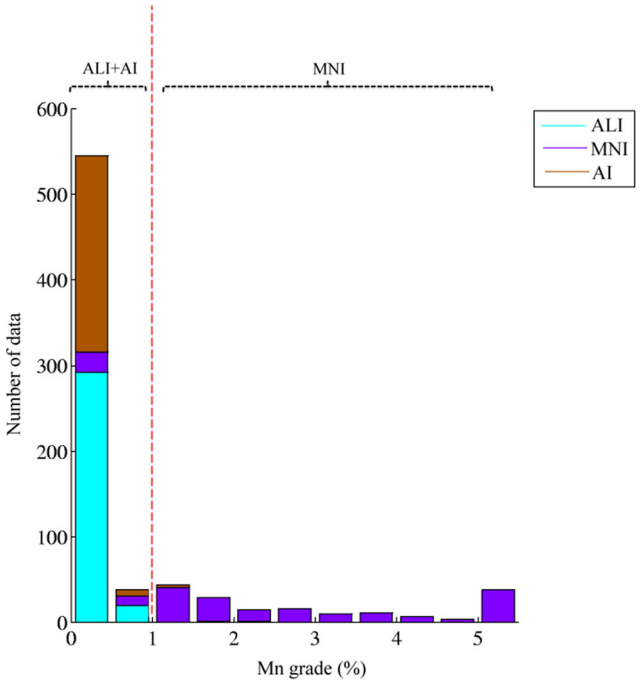

Fig. 14 Histogram of manganese grade for rock types considered in the fifth layer

Table 7 Splitting into manganese-rich vs. manganese-poor itabirite

Rock group	Rock codes	Relative proportions
Manganese-rich itabirite	MNI	0.2695
Manganese-poor itabirite	ALI, AI	0.7305

3.2.6 Sixth layer (loss on ignition)

The loss on ignition is the last criterion for splitting the manganese-poor itabirite defined in the previous layer into two subgroups: amphibolitic itabirite with a loss on ignition mostly greater than 5 %, and alumina-rich itabirite, with a loss on ignition mostly less than 5 % (Table 8 and Fig. 16). Based on the relative proportions given in Table 8, the truncation threshold for this layer is $G^{-1}(0.4322) = -0.1708$. The variogram of the Gaussian random field to truncate is displayed in Fig. 17.

3.2.7 Summary

The hierarchical model and Gaussian variograms are summarized in Fig. 18 and Table 9, respectively.

3.3 Hierarchical simulation

After determining the truncation thresholds and variogram models for the six underlying Gaussian random fields, the rock types are simulated in the study area conditionally to the drill hole data. The hierarchical simulation algorithm has been implemented by using an adaptation of a publicly available code [13], and 100 realizations of the rock types are constructed. Figure 19 shows two different realizations in the same cross section as Fig. 4. Each realization provides a structurally accurate map of the rock types, which reproduces the conditioning data, as well as the rock type contact relationships, proportions, and spatial continuity.

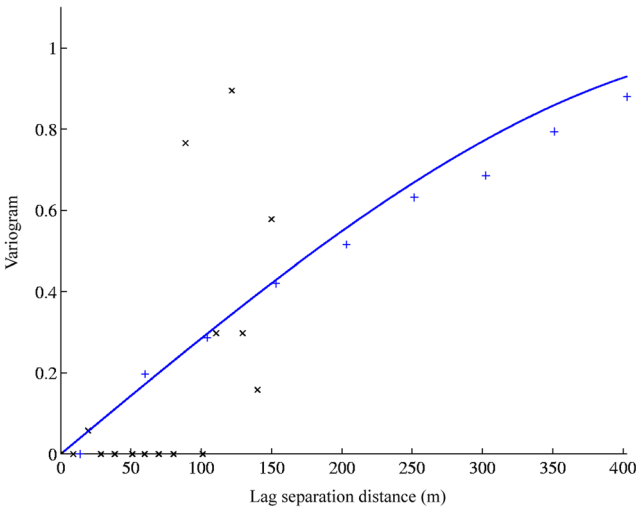

Fig. 15 Experimental (*crosses*) and modeled (*solid line*) variograms for the fifth Gaussian random field

Table 8 Splitting manganese-poor itabirite into amphibolitic vs. alumina-rich itabirite

Rock group	Rock codes	Relative proportions
Amphibolitic itabirite	AI	0.4322
Alumina-rich itabirite	ALI	0.5678

3.4 Post-processing the realizations

3.4.1 Probability maps

The first and maybe most useful and expedient output that can be constructed from the simulation results are maps showing the probability of occurrence of each rock type in the study area, which help to measure the uncertainty in the layout of each rock type. Figure 20 shows such maps in a cross section of the study area, calculated by using 100 realizations.

3.4.2 Most probable rock type

One can also construct a lithological model of the study area by selecting the most probable rock type at each location and calculate the associated probability of occurrence (Fig. 21). A low probability indicates high uncertainty, therefore a risk of finding a rock type that differs from what is expected.

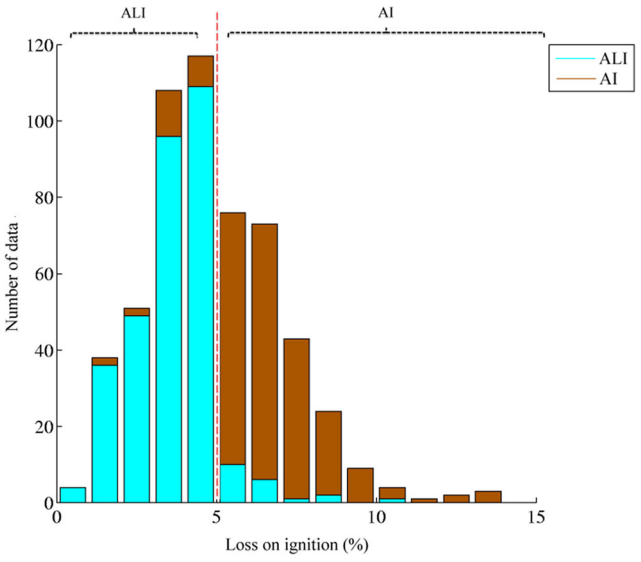

Fig. 16 Histogram of loss of ignition for rock types considered in the sixth layer

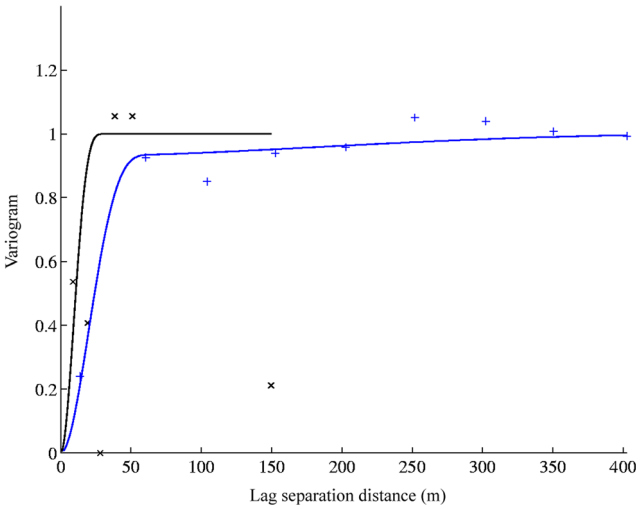

Fig. 17 Experimental (*crosses*) and modeled (*solid line*) variograms for the sixth Gaussian random field, along vertical (*black*) and horizontal (*blue*) directions

Table 9 Summary of variogram models

Gaussian random field	Basic nested structure	Sill	Horizontal range (m)	Vertical range (m)
1	Cubic	0.253	100	150
	Cubic	0.747	1000	150
2	Cubic	0.703	25	15
	Cubic	0.297	450	150
3	Exponential	0.767	110	85
	Cubic	0.233	700	165
4	Exponential	0.998	200	200
	Cubic	0.002	350	350
5	Spherical	1.00	520	520
6	Cubic	0.930	65	30
	Cubic	0.070	600	35

3.4.3 Matches and mismatches with the interpreted model

Another useful parameter that can be calculated relates to the match and mismatch between the most probable rock type (Fig. 21) and the interpreted lithological model (Fig. 4). This parameter is an indicator of the consistency between the simulated and interpreted models and can be used to identify the sectors with greater geological uncertainty (those for which the interpreted lithological model is likely to be mistaken).

The numbers of matches and mismatches in the study area are shown in Table 10, respectively. For instance, as can be seen in the first row of Table 10, there are 9198 blocks flagged as friable hematite (FH) in both the interpreted lithological model and the most probable model (matches). However, there are 529 blocks that are flagged as FH in the interpreted model but as CH in the most probable model

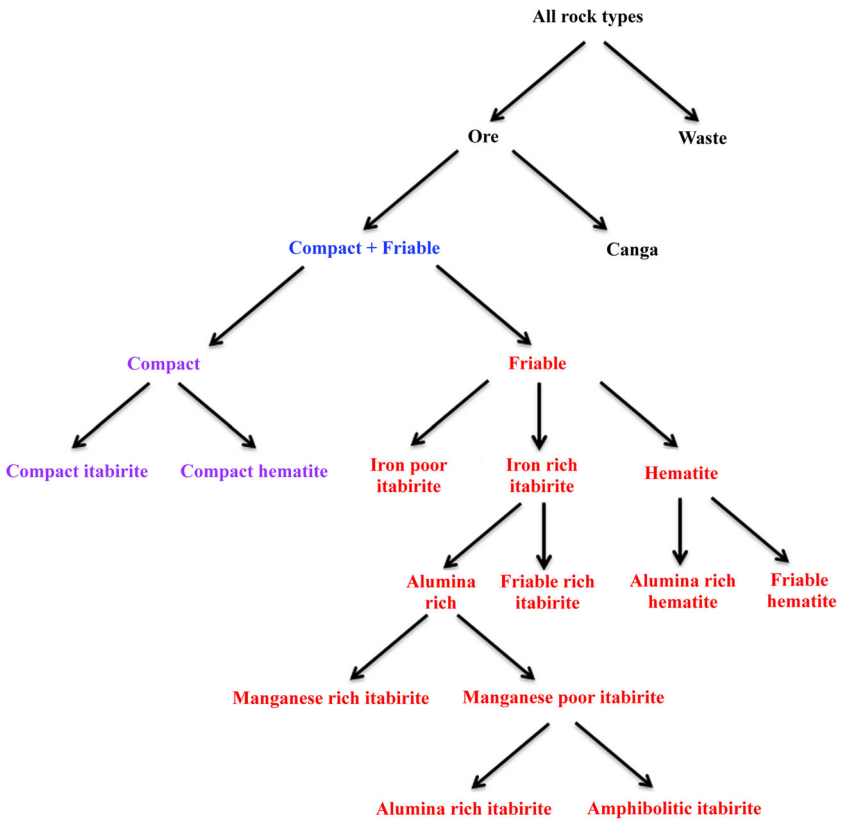

Fig. 18 Hierarchical modeling of rock types

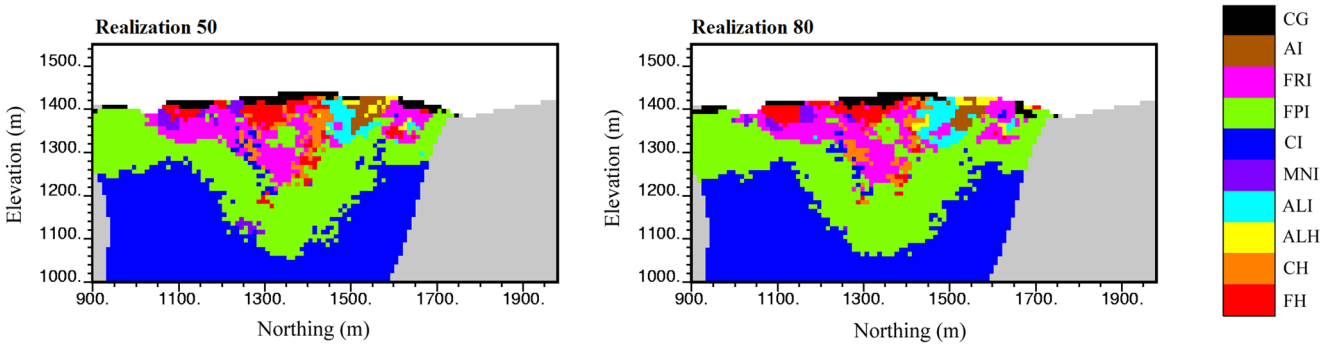

Fig. 19 Rock type realizations in a cross section of the study area

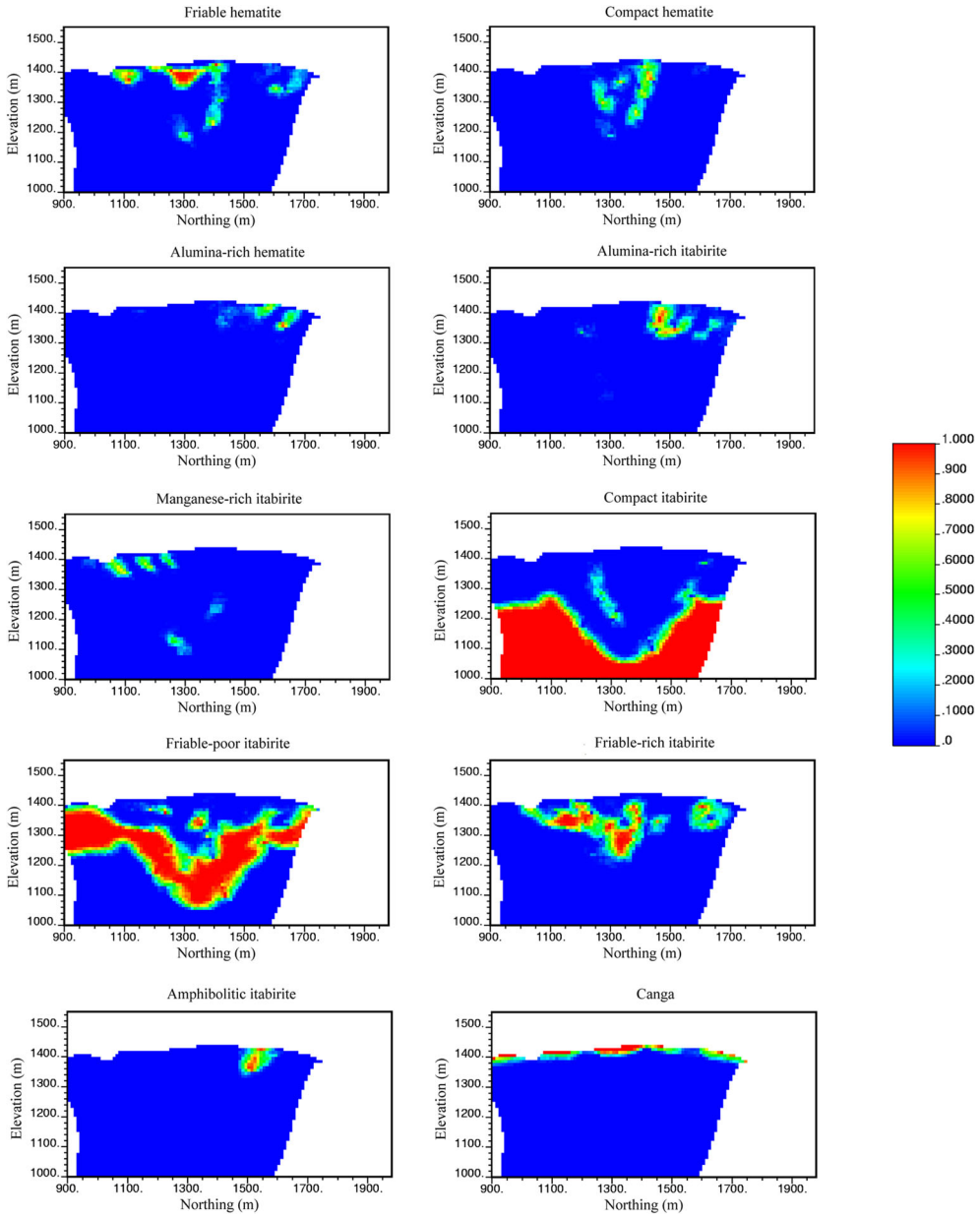

Fig. 20 Probability map of each rock type in a cross section of the study area

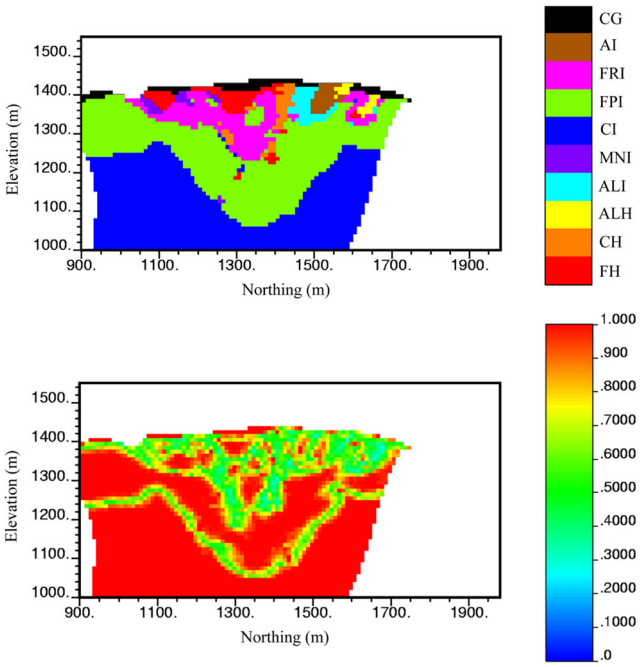

Fig. 21 Map of the most probable rock type and its associated probability of occurrence in a cross section of the study area

(mismatches). For the simulation all over the study area, one observes 602,927 matches over a total of 651,952 ore blocks. In other words, the interpreted model matches the most probable model for 92.5 % of the ore blocks.

3.4.4 Uncertainty in rock type proportions

For each realization, one can calculate the proportion of blocks that belong to each rock type and use this parameter for determining the uncertainty in the proportion of space covered by each rock type and for comparing with the proportions in the interpreted model. Table 11 indicates the minimum, maximum, and average proportions of rock types over 100 realizations, which are globally consistent with the proportions in the interpreted model.

Table 11 Statistics of rock type proportions for 100 realizations and interpreted model

	Minimum over realizations (%)	Maximum over realizations (%)	Average over realizations (%)	Interpreted model (%)
FH	0.97	1.27	1.11	0.97
CH	0.63	0.79	0.70	0.67
ALH	0.47	0.60	0.53	0.44
ALI	1.32	1.58	1.44	1.28
MNI	0.51	0.66	0.59	0.49
CI	26.96	27.78	27.31	27.41
FPI	13.37	14.26	13.87	14.27
FRI	1.61	1.80	1.70	1.56
AI	0.93	1.17	1.03	0.94
CG	1.81	2.31	2.07	2.32

3.5 Model validation

When inferring the model parameters and simulating a given spatial attribute (in the present case, a rock type), there is interest to know about the accuracy and quality of the simulated values. Accordingly, validation is an essential step of the simulation process. To this end, the original data are divided into two subsets randomly. Then, one of the subsets (training subset) is used as conditioning data for simulating the rock types at the data locations of the other subset (testing subset). Using the realizations, the probability of occurrence of each rock type can be calculated for each data of the testing subset. Then, for each rock type, the data of the testing subset are sorted by increasing probability of occurrence: for a given probability p (up to some calculation tolerances), one expects that a proportion p of the testing

Table 10 Matches between interpreted (rows) and most probable (columns) rock types

	FH	CH	ALH	ALI	MNI	CI	FPI	FRI	AI	CG	Total
FH	9198	529	489	127	116	66	615	771	52	581	12,544
CH	642	6324	142	32	22	143	739	389	16	241	8690
ALH	408	155	4130	89	84	16	127	161	163	304	5637
ALI	252	54	190	12,093	164	140	1926	355	904	532	16,610
MNI	184	2	58	96	4815	85	771	184	36	87	6318
CI	99	531	18	208	106	345,449	7856	411	77	134	354,889
FPI	858	406	241	2038	755	5060	171,298	2393	868	868	184,785
FRI	1207	395	298	501	362	368	1984	14,486	169	470	20,240
AI	93	41	270	707	10	77	501	95	9832	549	12,175
CG	761	59	358	571	223	176	1375	578	661	25,302	30,064
Total	13,702	8496	6194	16,462	6657	351,580	187,192	19,823	12,778	29,068	651,952

Table 12 Results of split-sample validation

Rock type	Probability interval	Average probability	Number of validation data	Expected number of validation data	Effective number of validation data	Difference
FH	0.0–0.2	0.072	350	25.31	28	−2.69
	0.2–0.4	0.289	148	42.81	41	1.81
	0.4–0.6	0.491	98	48.11	33	15.11
	0.6–0.8	0.689	70	48.22	42	6.22
	0.8–1.0	0.902	60	54.14	47	7.14
CH	0.0–0.2	0.069	275	19.01	30	−10.99
	0.2–0.4	0.287	99	28.37	30	−1.63
	0.4–0.6	0.481	64	30.78	42	−11.22
	0.6–0.8	0.702	47	33.01	37	−3.99
	0.8–1.0	0.886	27	23.91	22	1.91
ALH	0.0–0.2	0.070	207	14.43	27	−12.57
	0.2–0.4	0.290	67	19.44	22	−2.56
	0.4–0.6	0.505	34	17.18	10	7.18
	0.6–0.8	0.691	27	18.67	17	1.67
	0.8–1.0	0.918	37	33.96	34	−0.04
ALI	0.0–0.2	0.064	314	20.23	40	−19.77
	0.2–0.4	0.269	109	29.36	30	−0.64
	0.4–0.6	0.496	66	32.75	32	0.75
	0.6–0.8	0.697	53	36.93	34	2.93
	0.8–1.0	0.901	28	25.23	25	0.23
MNI	0.0–0.2	0.065	157	10.26	22	−11.74
	0.2–0.4	0.292	46	13.45	12	1.45
	0.4–0.6	0.512	28	14.34	8	6.34
	0.6–0.8	0.715	31	22.15	24	−1.85
	0.8–1.0	0.898	30	26.93	27	−0.07
CI	0.0–0.2	0.070	370	26.03	27	−0.97
	0.2–0.4	0.283	176	49.87	51	−1.13
	0.4–0.6	0.488	96	46.86	53	−6.14
	0.6–0.8	0.688	67	46.07	49	−2.93
	0.8–1.0	0.903	52	46.98	46	0.98
FPI	0.0–0.2	0.069	550	38.13	26	12.13
	0.2–0.4	0.289	241	69.53	53	16.53
	0.4–0.6	0.495	189	93.53	86	7.53
	0.6–0.8	0.698	202	141	139	2
	0.8–1.0	0.912	289	263.63	247	16.63
FRI	0.0–0.2	0.070	487	34.31	34	0.31
	0.2–0.4	0.287	198	56.81	57	−0.19
	0.4–0.6	0.486	137	66.65	69	−2.35
	0.6–0.8	0.689	90	62	69	−7
	0.8–1.0	0.901	59	53.18	49	4.18
AI	0.0–0.2	0.064	158	10.15	17	−6.85
	0.2–0.4	0.297	58	17.2	20	−2.8
	0.4–0.6	0.487	42	20.45	15	5.45
	0.6–0.8	0.695	34	23.64	31	−7.36
	0.8–1.0	0.908	31	28.15	25	3.15

Table 12 (continued)

Rock type	Probability interval	Average probability	Number of validation data	Expected number of validation data	Effective number of validation data	Difference
CG	0.0–0.2	0.072	96	6.91	7	−0.09
	0.2–0.4	0.298	66	19.7	19	0.7
	0.4–0.6	0.499	44	21.96	19	2.96
	0.6–0.8	0.683	14	9.56	12	−2.44
	0.8–1.0	0.924	14	12.93	13	−0.07

data actually matches this rock type. The results of this process (Table 12) show that the proposed model is accurate, insofar as the observed data proportions are close to the expected probabilities.

4 Discussion and conclusions

The stochastic modeling of geological domains has received an increasing attention over the past decades in the characterization of ore deposits, oil reservoirs, and aquifers, insofar as it allows better understanding the geological structure in the subsurface, mapping the uncertainty in the domain layout at locations where sampling information is not available and consequently identifying the locations where the geological interpretation is likely to be mistaken. The hierarchical application of the truncated Gaussian model, as proposed in this paper, turns out to be congenial and flexible, as it allows reproducing (1) the sequencing of the domains and their contact relationships, (2) the proportion and spatial continuity of each geological domain, via the definition of the truncation thresholds and Gaussian variograms, and (3) the available conditioning data at sampling locations. The case study demonstrated the straightforwardness of this approach to modeling geological domains and its capabilities, overcoming the limitations of the standard truncated Gaussian model that is often restricted to the simulation of few domains and to specific orderings of the domains, e.g., sedimentary sequences.

Despite these benefits, there are still some opportunities for further enhancements to get geologically more realistic images, in particular in what refers to the reproduction of connectivity constraints (e.g., when one desires a domain to connect different spatial locations without interruption), local anisotropies (when the main direction of continuity changes in space), and specific shapes, such as weathering features along the compositional bands that are observed in the interpreted lithological model of Fig. 4, or coniform shapes that mimic the geological geometries related to different genetic processes (e.g., deposition, infiltration by gravity, intrusive bodies, or metamorphic geological processes). Currently, these features can be reproduced by resorting to object-based [8], process-based [10], or multiple-point geostatistical models [25]. Such models are actually not incompatible with the proposed hierarchical approach, as they can be used in one or more layers instead of the truncated Gaussian model. The latter can also be extended so as to reproduce connectivity constraints [2] or meandering structures [20]; another alternative is to define truncations based not only on half-bounded intervals but also bounded intervals or unions of intervals, which may yield geological domains with greater spatial continuity and connectivity [24].

Acknowledgments This work was supported by VALE S.A., through the project entitled "Geostatistical cosimulation of grades and rock types for iron resource evaluation," by the COPEC-UC Foundation, through project 2014.J.057, and by the Chilean Commission for Scientific and Technological Research, through Projects CONICYT/FONDECYT/REGULAR/N°1130085 and CONICYT PIA Anillo ACT1407.

References

1. Albertao, G.A., Grell, A.P., Badolato, D., dos Santos, L.R.: 3D geological modeling in a turbidite system with complex stratigraphic-structural framework—an example from Campos Basin, Brazil. SPE Annual Technical Conference & Exhibition, Dallas, Texas, SPE 95612 (2005)
2. Allard, D.: Simulating a geological lithofacies with respect to connectivity information using the truncated Gaussian model. In: Armstrong, M., Dowd, P.A. (eds.) Geostatistical simulations, pp. 197–211. Kluwer, Dordrecht (1994)
3. Armstrong, M., Galli, A., Beucher, H., Le Loc'h, G., Renard, D., Doligez, B., Eschard, R., Geffroy, F.: Plurigaussian simulations in geosciences. Springer, Berlin (2011)
4. Barbier, M., Hamon, Y., Doligez, B., Callot, J.P., Floquet, M., Daniel, J.M.: Stochastic joint simulation of facies and diagenesis: a case study on early diagenesis of the Madison Formation (Wyoming, USA). Oil & Gas Science and Technology – Revue d'IFP Energies Nouvelles **67**(1), 123–145 (2012)

5. Carrasco, P., Ibarra, F., Rojas, R., Le Loc'h, G., Séguret, S.: Application of the truncated Gaussian simulation method to a porphyry copper deposit. In: Magri, E. (ed.) Proceedings of the 33rd international symposium on application of computers and operations research in the mineral industry, pp. 31–39 (2007)
6. Cherubini, C., Giasi, C.I., Musci, F., Pastore, N.: Application of truncated plurigaussian method for the reactive transport modeling of a contaminated aquifer. In: Proceedings of the 4th IASME/WSEAS International Conference on Water Resources, Hydraulics & Hydrology, pp. 119–124. WSEAS Press (2009)
7. Cherubini, C., Giasi, C.I., Musci, F., Pastore, N.: Checking simulations of a geolithological model obtained by means of nested truncated bigaussian method. Int. J. Math. Models Methods Appl. Sci. **3**(2), 152–161 (2009)
8. Chessa, A.G., Martinius, A.W.: Object based modelling of the spatial distribution of fluvial sandstone deposits. In: Christie, M.A., Da Silva, F.V., Farmer, C.L., Guillon, O., Heinemann, Z.E., Lemonnier, P., Regtien, J.M.M., van Spronsen, E. (eds.) Proceedings of the 3rd European conference on the mathematics of oil recovery, pp. 5–14. Delft University press (1992)
9. Chilès, J.P., Delfiner, P.: Geostatistics: modeling spatial uncertainty. Wiley, New York (2012)
10. Cojan, I., Fouche, O., Lopez, S., Rivoirard, J.: Process-based reservoir modelling in the example of meandering channel. In: Leuangthong, O., Deutsch, C.V. (eds.) Geostatistics Banff, vol. 2004, pp. 611–619. Springer, Dordrecht (2005)
11. De Fouquet, C.: Reminders on the conditioning kriging. In: Armstrong, M., Dowd, P.A. (eds.) Geostatistical simulations, pp. 131–145. Kluwer, Dordrecht (1994)
12. Dimitrakopoulos, R.: Advances in orebody modelling and strategic mine planning. Australasian Institute of Mining and Metallurgy, Melbourne (2011)
13. Emery, X.: Simulation of geological domains using the plurigaussian model: new developments and computer programs. Comp. Geosci. **32**(9), 1189–1201 (2007)
14. Emery, X.: Iterative algorithms for fitting a linear model of coregionalization. Comp. Geosci. **36**(9), 1189–1201 (2010)
15. Emery, X., Arroyo, D., Porcu, E.: An improved spectral turning-bands algorithm for simulating stationary vector Gaussian random fields. Stoch. Env. Res. Risk A. (2016). in press
16. Emery, X., González, K.E.: Incorporating the uncertainty in geological boundaries into mineral resources evaluation. J. Geol. Soc. India **69**(1), 29–38 (2007)
17. Emery, X., Lantuéjoul, C.: TBSIM: a computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. Comp. Geosci. **32**(10), 1615–1628 (2006)
18. Freulon, X., De Fouquet, C.: Conditioning a Gaussian model with inequalities. In: Soares, A. (ed.) Geostatistics Tróia'92, pp. 201–212. Kluwer, Dordrecht (1993)
19. Galli, A., Le Loc'h, G., Geffroy, F., Eschard, R.: An application of the truncated pluri-gaussian method for modeling geology. In: Coburn, T.C., Yarus, J.M., Chambers, R.L. (eds.) Stochastic modeling and geostatistics: Principles, methods, and case studies, volume II, vol. 5, pp. 109–122 (2006)
20. Langlais, V., Beucher, H., Renard, D.: In the shade of truncated Gaussian simulation. In: Ortiz, J.M., Emery, X. (eds.) Proceedings of the eighth international geostatistics congress, pp. 799–808. Gecamin Ltda, Santiago (2008)
21. Lantuéjoul, C.: Geostatistical simulation: models and algorithms. Springer, Berlin (2002)
22. Larrondo, P.F., Neufeld, C.T., Deutsch, C.V.: VARFIT: a program for semi-automatic variogram modelling. In: Deutsch, C.V. (ed.) Fifth annual report of the center for computational geostatistics, p. 17. University of Alberta (2003)
23. Madani, N., Emery, X.: Simulation of geo-domains accounting for chronology and contact relationships: application to the Río Blanco copper deposit. Stoch. Env. Res. Risk A. **29**(8), 2173–2191 (2015)
24. Marcotte, D., Gloaguen, E.: A class of spatial multivariate models based on copulas. In: Ortiz, J.M., Emery, X. (eds.) Proceedings of the eighth international geostatistics congress, pp. 177–186. Gecamin Ltda, Santiago (2008)
25. Mariethoz, G., Caers, J.: Multiple-point geostatistics: Stochastic modeling with training images. Wiley, New York (2014)
26. Mariethoz, G., Renard, P., Cornaton, F., Jaquet, O.: Truncated plurigaussian simulations of aquifer heterogeneity. Ground Water **47**(1), 13–24 (2009)
27. Remacre, A.Z., Zapparolli, L.H.: Application of the plurigaussian simulation technique in reproducing lithofacies with double anisotropy. Rev. Bras. Geocienc. **33**(2), 37–42 (2003)
28. Riquelme, R., Le Loc'h, G., Carrasco, P.: Truncated Gaussian & plurigaussian simulations of lithological units in Mansa Mina Deposit. In: Ortiz, J.M., Emery, X. (eds.) Proceedings of the 8th international geostatistics congress, pp. 819–828, Santiago (2008)
29. Talebi, H., Asghari, O., Emery, X.: Application of plurigaussian simulation to delineate the layout of alteration domains in Sungun copper deposit. Cent. Eur. J. Geosci. **5**(4), 514–522 (2013)
30. Talebi, H., Asghari, O., Emery, X.: Simulation of the lately injected dykes in an Iranian porphyry copper deposit using the plurigaussian model. Arab. J. Geosci. **7**(7), 2771–2780 (2014)
31. Skvortsova, T., Armstrong, M., Beucher, H., Forkes, J., Thwaites, A., Turner, R.: Simulating the geometry of a granite-hosted uranium orebody. In: Armstrong, M., Bettini, C., Champigny, N., Galli, A., Remacre, A. (eds.) Geostatistics Rio 2000, pp. 85–100. Kluwer, Dordrecht (2002)
32. Yunsel, T., Ersoy, A.: Geological modeling of rock type domains in the Balya (Turkey) lead-zinc deposit using plurigaussian simulation. Cent. Eur. J. Geosci. **5**(1), 77–89 (2013)