Stoch Environ Res Risk Assess (2017) 31:893–913
DOI 10.1007/s00477-016-1365-9

ORIGINAL PAPER

Plurigaussian modeling of geological domains based on the truncation of non-stationary Gaussian random fields

Nasser Madani[1,2] · **Xavier Emery**[1,2]

Published online: 2 December 2016

Abstract The plurigaussian model is used in mining engineering, oil reservoir characterization, hydrology and environmental sciences to simulate the layout of geological domains in the subsurface, while reproducing their spatial continuity and dependence relationships. However, this model is well-established only in the stationary case, when the spatial distribution of the domains is homogeneous in space, and suffers from theoretical and practical impediments in the non-stationary case. To overcome these limitations, this paper proposes extending the model to the truncation of intrinsic random fields of order k with Gaussian generalized increments, which allows reproducing spatial trends in the distribution of the geological domains. Methodological tools and algorithms are presented to infer the model parameters and to construct realizations of the geological domains conditioned to existing data. The proposal is illustrated with the simulation of rock type domains in an ore deposit in order to demonstrate its applicability. Despite the limited number of conditioning data, the results show a remarkable agreement between the simulated domains and the lithological model interpreted by geologists, while the conventional stationary plurigaussian model turns out to be unsuccessful.

✉ Xavier Emery
xemery@ing.uchile.cl

Nasser Madani
naser.madani@yahoo.com

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

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

Keywords Geological domaining · Subsurface heterogeneity · Intrinsic random fields of order k · Generalized covariance function

1 Introduction

The primary motivation of this paper is the spatial modeling of geological domains in ore deposits, defined on the basis of lithology, mineralogy or alteration, but the tools and algorithms developed hereafter can be applied to other application fields, such as petroleum engineering, subsurface hydrology and environmental science to model lithofacies in oil reservoirs, aquifers or contaminated sites. In the mining industry, geological domaining is useful to define which material to extract, when and how it should be extracted, to determine the requirements for mining operations such as drilling, blasting, hauling and crushing, and to forecast the destination of the extracted material (dump, heap leaching, flotation plant, stock pile, etc.). It is also critical for modeling the mineral resources and ore reserves, as the spatial behavior of geological, geotechnical and geo-metallurgical variables, such as the grades of elements of interest and contaminants, rock density, fracturation intensity, work index, solubility ratio, acid consumption or metal recovery, are often controlled by the lithology, mineralogy or alteration. In such a case, a two-stage approach is frequently used to model the resources and reserves (Chilès and Delfiner 2012):

(1) Modeling the spatial layout of the geological domains based on the characteristics that control the mineralization, such as rock types, mineral types or alterations.

(2) Modeling the geological, geotechnical and geo-metallurgical variables of interest within each domain.

The second stage can be achieved through spatial interpolation methods (such as inverse distance weighting, kriging, simulation, etc.) and is not of direct interest in this paper. With respect to the first stage, the problem is two-fold. On the one hand, one aims at predicting which geological domains prevail at locations where direct information is not available. One the other hand, it is of interest to determine whether or not the layout of the geological domains is accurately predicted and what is the risk of misclassification.

Geological domaining can be done with deterministic or stochastic models. The former provide a unique description of the geological domain layout within the deposit and are based on a geological interpretation or on a model interpolated from the available data on the deposit, typically exploration drill hole samples. The latter consist in constructing multiple outcomes or realizations of the domains, represented by random sets (indicators) or categorical random fields, i.e., random fields that take nominal values or categories. These approaches contribute to uncertainty quantification and an improved modeling of the quantitative geological, geotechnical and geo-metallurgical variables of interest when these variables are homogeneously distributed in each geological domain but the layout of the domain boundaries is uncertain.

Among simulation approaches, truncated Gaussian simulation (Galli et al. 1994) is suited to the cases when the geological domains follow an ordered sequence, such as strata in sedimentary deposits, while plurigaussian simulation (Le Loc'h and Galli 1997) allows reproducing more complex transitions between the geological domains. The latter has found wide acceptance in the past decades for modeling ore deposits and petroleum reservoirs (Armstrong et al. 2011, and references therein).

However, as most of the geostatistical approaches used in the mining and petroleum industries, the current design of the plurigaussian model is based on an assumption of stationarity, through the use of Gaussian random fields whose finite-dimensional distributions are invariant by a translation in space. In a nutshell, such an assumption is applicable for regionalized variables that behave homogeneously throughout the region of interest (Matheron 1971). Specifically, first-order stationarity assumes that the mean of the random field is constant over space, while second-order stationarity assumes that, in addition to the existence of a constant mean, the covariance or the variogram between the values observed at two locations only depends on the geographical separation of these locations. Nonetheless, the geological domains, represented by indicator variables, often have a spatially varying mean and spatial continuity (covariance). For instance, mineralogical domains exhibit a vertical zonation, where some domains are likely to be found near the surface (gossan cap, leached zone, oxidized zone), while others are present only in depth (supergene sulfide enrichment zone and hypogene or primary sulfide-bearing zone) (Guilbert and Park 1986). Lateral variations are also often observed in the distributions of rock types or alterations. As an example, porphyry copper deposits generally present inner potassic and phyllic alteration domains, surrounded by a sericitic alteration domain, in turn surrounded by an argillic alteration domain and an outer propylitic alteration domain (Lowell and Guilbert 1970). An assumption of stationarity, implying that the probability of occurrence of every alteration domain is constant in space, is clearly inadequate in such a case. Stationary models may also be ill-suited to describe phenomena for which a spatial trend is present due to physical reasons, e.g. folded structures (domes, anticlines or synclines) in geological formations such as petroleum reservoirs and ore deposits.

Plurigaussian simulation can be adapted to non-stationary cases by considering spatially varying truncation thresholds (Beucher et al. 1993; Ravenne et al. 2002; Armstrong et al. 2011). However, in general, this procedure does not allow managing the uncertainty in the true parameters, as one considers the domain proportions as if they were perfectly known, which is not the case in reality (Biver et al. 2002), therefore an important source of uncertainty may be omitted. Also, handling spatial changes in the truncation thresholds becomes cumbersome in what refers to the inference of spatial continuity (covariance or variogram analysis) and, so far, the formalism is not clearly laid out (Armstrong et al. 2011).

To overcome these difficulties, this paper proposes to extend plurigaussian simulation to the non-stationary case by considering categorical random fields obtained through the truncation of intrinsic random fields of order k (shortly, IRF-k) with Gaussian generalized increments (Matheron 1973; Christakos 1992; Chilès and Delfiner 2012). The rationale is that the use of IRF-k will allow the reproduction of spatial trends (such as mineral domains that are present near the surface and absent at depth, or alteration domains that are present in the inner part of an ore body and absent in outer parts). To illustrate this idea, Fig. 1 displays isopleth maps of realizations of intrinsic random fields on the two-dimensional plane, with increasing orders (left side, from top to bottom). As the order increases, the intrinsic random field shows a greater short-scale regularity and, above all, clearer large-scale trends. In each case, two geological domains can be defined by truncating the intrinsic random field at a given level curve (right side): the

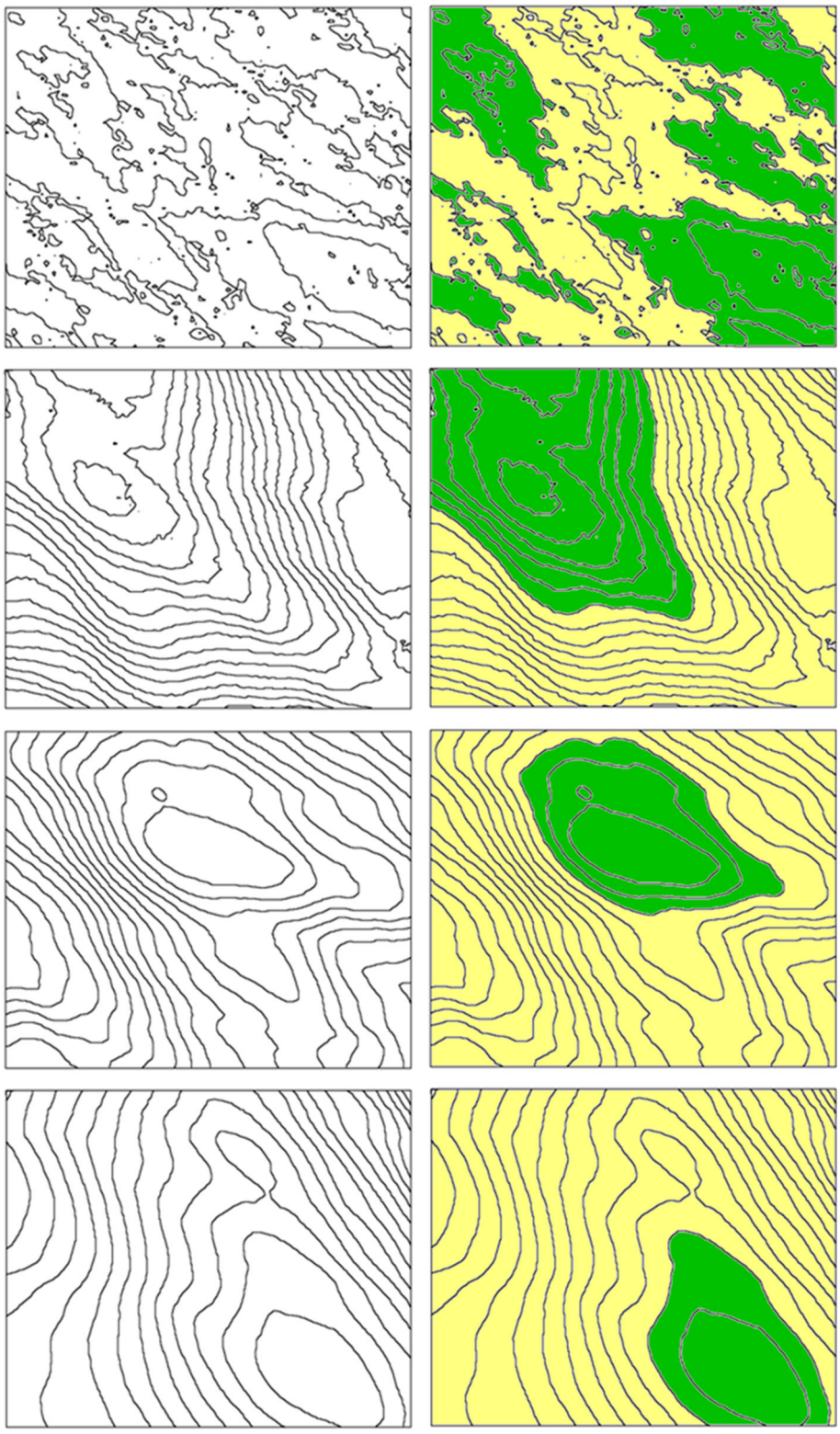

Fig. 1 Isopleth maps of intrinsic random fields of order k (*left*) and two geological domains (*green* and *yellow*) obtained by truncation at a specific threshold

trending behavior in the distribution of the domains so obtained gets more pronounced as the order increases.

The outline of this paper is as follows. Section 2 recalls the principles of plurigaussian modeling and the theory of intrinsic random fields of order k, then presents tools and algorithms for inferring the parameters of the proposed non-stationary plurigaussian model (in particular, in what refers to the identification of the spatial correlation

structure of the underlying intrinsic random fields) and for simulating the geological domains conditionally to existing data. In Sect. 3, an application to a case study in ore body modeling is presented. A general discussion and conclusions follow in Sects. 4 and 5, respectively.

2 Methodology

2.1 Stationary plurigaussian modeling

The plurigaussian model (Armstrong et al. 2011) relies on the simulation of one or more Gaussian random fields and their truncation to convert the simulated Gaussian values into categorical values representing a partition of space into domains (Fig. 2). The practical implementation of this model requires defining the following parameters:

(1) A truncation rule, which defines the contact relationships between domains. Such a rule can also account for their chronological relationships, as illustrated in Fig. 2c: the first geological domain cross-cuts the other two and therefore stands for a younger domain.
(2) Truncation thresholds, which are related to the proportion of space covered by each domain.
(3) The covariance functions of the underlying Gaussian random fields, which are related to the spatial continuity of the domains obtained by truncation.

The conditional simulation of the geological domains can be performed in three main steps (Lantuéjoul 2002):

(1) Simulate the underlying Gaussian random fields at the data locations, conditionally to the categorical data (presence or absence of a domain at sampling locations). This step can be realized by an iterative algorithm known as the Gibbs sampler (Geman and Geman 1984).
(2) Simulate the Gaussian random fields at the target locations conditionally to the values obtained in the previous step. This step can be achieved in two

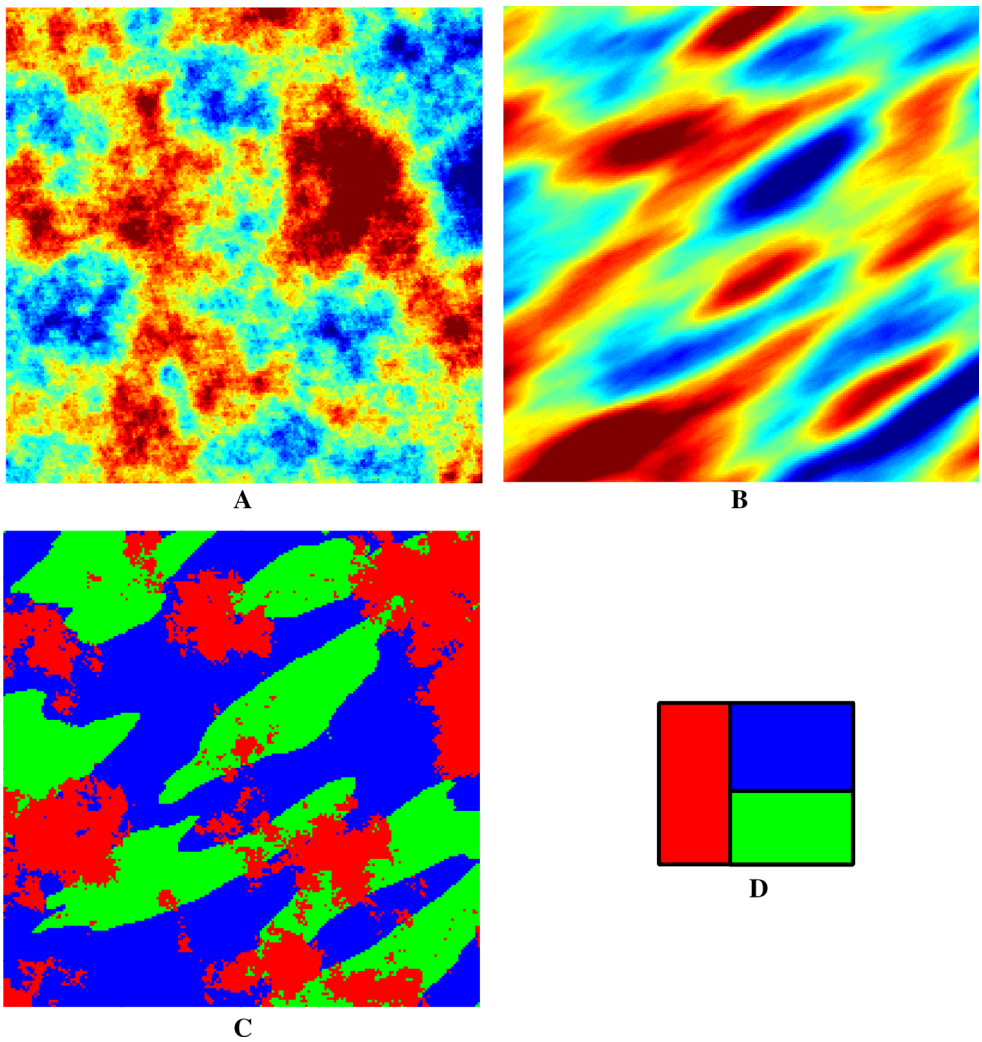

Fig. 2 Realizations of domains **c** obtained by truncating two independent Gaussian random fields (**a**, **b**). The truncation rule is represented by a two-dimensional flag (**d**), in which the abscissa axis represents the first Gaussian random field, the ordinate axis represents the second Gaussian random field, and the horizontal and vertical lines represent the truncation thresholds that define the partition of the bi-Gaussian space into domains

stages: first, a non-conditional simulation at the data and target locations, for which many algorithms are available; second, a post-conditioning by kriging that converts the non-conditional simulation into a conditional one (Chilès and Delfiner 2012).

(3) Apply the truncation rule to obtain the simulated domains.

In practice, to restrict the number of parameters and to ease their inference, one often considers two independent Gaussian random fields for the truncation rule. This entails a limitation in the number of domains that can be modeled, especially when all these domains are mutually in contact, a situation that often arises in practice. To overcome this limitation, Xu et al. (2006) and Emery (2007) propose to increase the number of Gaussian random fields, so as to increase the dimensionality of the truncation rule and to allow the domains to be mutually in contact.

The inference of the model parameters and the simulation often rely on an assumption of stationarity. It implies that the truncation thresholds are constant over the space and the Gaussian random fields are second-order stationary. To alleviate this assumption, one option is to use spatially varying thresholds, while the Gaussian random fields remain second-order stationary. In this case, the inference of truncation thresholds relies on the concept of vertical or regionalized proportion curves that give the proportion of each domain for each spatial location (Beucher et al. 1993; Ravenne et al. 2002). However, such an approach is challenging not only for the calculation of local domain proportions, which are assumed to be known without any uncertainty, but also for variogram analysis, for which the theoretical and practical frameworks are not clearly laid out (Armstrong et al. 2011).

2.2 Proposal for non-stationary plurigaussian modeling

The main contribution of this research is a generalization of the plurigaussian model to the non-stationary framework. Instead of working with stationary Gaussian random fields and spatially varying truncation thresholds, the key idea is to consider the truncation of intrinsic random fields of order k with generalized Gaussian increments. Hence, the non-stationarity is no longer modeled through spatially varying truncation thresholds, but through the underlying random fields to be truncated.

2.2.1 *Intrinsic random fields of order k*

A random field $Z = \{Z(\mathbf{x}): \mathbf{x} \in \mathbb{R}^d\}$ defined in the d-dimensional Euclidean space is an intrinsic random field of order 0 with no drift if its increments have zero mean and are second-order stationary (Chilès and Delfiner 2012), i.e.

$$\forall \mathbf{x},\ \mathbf{x}+\mathbf{h} \in \mathbb{R}^d,\ \begin{cases} E\{Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x})\} = 0 \\ \mathrm{var}\{Z(\mathbf{x}+\mathbf{h}) - Z(\mathbf{x})\} = 2\gamma(\mathbf{h}) \end{cases} \tag{1}$$

Accordingly, the spatial correlation structure of such a random field is represented by its variogram $\gamma(\mathbf{h})$, a function that only depends on the separation vector $\mathbf{h}$ between the chosen locations $\mathbf{x}$ and $\mathbf{x}+\mathbf{h}$.

To generalize this idea to higher order increments, let us denote by $\{f^\ell(\mathbf{x}) : \ell = 1\ldots L\}$ the set of monomials of the coordinates of $\mathbf{x}$ with degree less than or equal to some positive integer k. A generalized increment of order k of the random field $Z = \{Z(\mathbf{x}): \mathbf{x} \in \mathbb{R}^d\}$ is a linear combination

$$\mathsf{Z}(\lambda) = \sum_{i=1}^{I} \lambda_i Z(\mathbf{x}_i) \tag{2}$$

for a set of weights and locations $\lambda = \{(\lambda_i, \mathbf{x}_i) : i = 1\ldots I\}$ that filter out the polynomials of the location coordinates of degree less than or equal to k, i.e.

$$\forall \ell \in \{1,\ldots,L\}, \sum_{i=1}^{I} \lambda_i f^\ell(\mathbf{x}_i) = 0 \tag{3}$$

In the following, let Λ_k denote the set of such generalized increments $\lambda = \{(\lambda_i, \mathbf{x}_i) : i = 1\ldots I\}$ fulfilling Eq. (3) for some positive integer I.

In this context, the random field Z is an intrinsic random field of order k (IRF-k) if its generalized increments of order k are second-order stationary and have a zero mean, i.e.

$$\forall \lambda = \{(\lambda_i, \mathbf{x}_i) : i = 1\ldots I\} \in \Lambda_k, \quad \begin{cases} E\{\mathsf{Z}(\lambda)\} = 0 \\ \mathrm{var}\{\mathsf{Z}(\lambda)\} = \sum_{i=1}^{I}\sum_{j=1}^{I} \lambda_i \lambda_j K(\mathbf{x}_i - \mathbf{x}_j) \end{cases} \tag{4}$$

for some function $K(\mathbf{h})$ called a generalized covariance (Matheron 1973; Christakos 1992; Chilès and Delfiner 2012). Under these conditions, the covariance between any two generalized increments exists and can be calculated as

$$\begin{aligned} &\forall \lambda = \{(\lambda_i, \mathbf{x}_i) : i = 1\ldots I\} \in \Lambda_k, \\ &\forall \mu = \{(\mu_j, \mathbf{x}'_j) : j = 1\ldots J\} \in \Lambda_k, \\ &\mathrm{cov}\{\mathsf{Z}(\lambda), \mathsf{Z}(\mu)\} = \sum_{i=1}^{I}\sum_{j=1}^{J} \lambda_i \mu_j K(\mathbf{x}_i - \mathbf{x}'_j) \end{aligned} \tag{5}$$

Equation (5) shows that the covariance between generalized increments of Z can be calculated as if Z had a stationary covariance function, by replacing this hypothetical covariance by the generalized covariance.

Because any random field that differs from Z by a polynomial of degree less than or equal to k (that is, a linear combination of the monomials $f^\ell(\mathbf{x})$ with deterministic or random coefficients) possesses exactly the same generalized increments $\mathsf{Z}(\lambda)$ as Z, the definition of an intrinsic random field of order k can actually be generalized to an equivalence class Z of random fields that differ by polynomials of degree less than or equal to k and satisfy Eqs. (4) and (5). Any member of this equivalence class is called a representation of the IRF-k (Chilès and Delfiner 2012).

The intrinsic random field Z has Gaussian generalized increments if, furthermore, the finite-dimensional distributions of its generalized increments (Eq. 2) are multivariate normal. This property will be assumed in the next sections for plurigaussian modeling.

2.2.2 *Inference of a non-stationary plurigaussian model*

Let us revise the changes that the use of intrinsic random fields of order k implies in the inference of the plurigaussian model.

2.2.2.1 Truncation rule As the truncation rule establishes the contact relationships between the domains obtained by truncation, there is no change in this concept with respect to the stationary model.

2.2.2.2 Truncation thresholds and structural analysis In the stationary case, one can establish a one-to-one relationship between the covariance function of a standard Gaussian random field (ρ) and the covariance of the indicator random field (C_{I_y}) obtained by truncation at a given threshold y (Emery 2007):

$$C_{I_y}(\mathbf{h}) = g(y)^2 \sum_{p=1}^{+\infty} \frac{1}{p} H_{p-1}^2(y)\,[\rho(\mathbf{h})]^p \tag{6}$$

where $\{H_p\colon p \in \mathbb{N}\}$ are Hermite polynomials and g is the standard Gaussian probability density function. In the case of an IRF-k, the covariance may not exist and the spatial correlation is described by a generalized covariance (Eq. 5), which has no clear relation with the indicator random field obtained by truncation. To circumvent this difficulty, an idea is to consider a representation of the intrinsic random field that possesses an ordinary covariance function, which is called an internal representation (Matheron 1973; Chilès and Delfiner 2012).

Specifically, let $Z = \{Z(\mathbf{x})\colon \mathbf{x} \in \mathbb{R}^d\}$ be a representation of an IRF-k with generalized covariance $K(\mathbf{h})$ and Gaussian generalized increments, and truncate it at a given threshold z:

$$I_z(\mathbf{x}) = \begin{cases} 1 & \text{if } Z(\mathbf{x}) > \text{ z} \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

An internal representation $Y = \{Y(\mathbf{x})\colon \mathbf{x} \in \mathbb{R}^d\}$ of the IRF-k can be obtained by putting:

$$Y(\mathbf{x}) = Z(\mathbf{x}) - \sum_{i=1}^{n} \lambda_i(\mathbf{x})\, Z(\mathbf{x}_i) \tag{8}$$

where $\{\mathbf{x}_i\colon i = 1\ldots n\}$ are arbitrarily chosen locations and $\lambda_i(\mathbf{x})$ is the universal kriging weight assigned to location $\mathbf{x}_i$ when predicting $\mathbf{x}$ with a pure nugget effect model:

$$\begin{pmatrix} \mathbf{I}_n & \mathbf{F}_0^T \\ \mathbf{F}_0 & \mathbf{0}_{L,L} \end{pmatrix} \begin{pmatrix} \mathbf{\Lambda}(\mathbf{x}) \\ \mathbf{M}(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \mathbf{0}_{n,1} \\ \mathbf{F}(\mathbf{x}) \end{pmatrix} \tag{9}$$

where $\mathbf{I}_n$ is the identity matrix with size $n \times n$, $\mathbf{0}_{L,L}$ is the zero matrix with size $L \times L$, $\mathbf{0}_{n,1}$ is the zero vector with size $n \times 1$, $\mathbf{F}_0$ is the $L \times n$ matrix with $f^\ell(\mathbf{x}_i)$ as the entry at row ℓ and column i, $\mathbf{F}(\mathbf{x})$ is the $L \times 1$ vector with entry $f^\ell(\mathbf{x})$ at row ℓ, $\mathbf{\Lambda}(\mathbf{x})$ is the $n \times 1$ vector of kriging weights and $\mathbf{M}(\mathbf{x})$ is a $L \times 1$ vector of Lagrange multipliers. If $\mathbf{x}$ coincides with some $\mathbf{x}_i$, then the nugget effect has to be filtered out of the second member of the kriging system, so that Eq. (9) still holds. This equation is actually equivalent to a linear regression of n data (stored in the $n \times 1$ vector $\mathbf{Z}_0$ with entry $Z(\mathbf{x}_i)$ at row i) upon L variables (stored in matrix $\mathbf{F}_0$). The solution exists and is unique provided that the spatial configuration of the locations $\{\mathbf{x}_i\colon i = 1\ldots n\}$ used to construct the internal representation Y is such that matrix $\mathbf{F}_0\mathbf{F}_0^T$ is invertible. In particular, n must be greater than or equal to the number L of basic drift functions, which depends on the order k and space dimension d (Chilès and Delfiner 2012): $n \geq L = \frac{(k+d)!}{k!d!}$ (Table 1).

Under these conditions, by using the formula of the matrix inversion in block form (Bernstein 2009), the kriging weights are found to be:

$$\mathbf{\Lambda}(\mathbf{x}) = \mathbf{F}_0^T(\mathbf{F}_0\mathbf{F}_0^T)^{-1}\mathbf{F}(\mathbf{x}) \tag{10}$$

One therefore has (Eq. 8):

$$Y(\mathbf{x}) = Z(\mathbf{x}) - \mathbf{Z}_0^T\mathbf{F}_0^T(\mathbf{F}_0\mathbf{F}_0^T)^{-1}\mathbf{F}(\mathbf{x}) \tag{11}$$

On the one hand, $Y(\mathbf{x})$ differs from $Z(\mathbf{x})$ by a combination of the components of $\mathbf{F}(\mathbf{x})$, i.e., from a polynomial function of the coordinates of $\mathbf{x}$, so that Y is a representation of the same IRF-k as Z. On the other hand, $Y(\mathbf{x})$ is a generalized increment of order k, as it coincides with a universal kriging error at location $\mathbf{x}$, so that $\{(\lambda_i, \mathbf{x}_i) : i = 1\ldots n\} \in \Lambda_k$ [Eq. (3) holds because of the

Table 1 Number L of locations needed to construct an internal representation, as a function of the space dimension (d) and order of the intrinsic random field (k)

	$k = 0$	$k = 1$	$k = 2$
$d = 2$	1	3	6
$d = 3$	1	4	10

unbiasedness constraint of universal kriging]. The latter implies that the internal representation Y has a zero mean, a finite variance and an ordinary covariance function $C(\mathbf{x},\mathbf{x}')$, which can be expressed as a function of the generalized covariance $K(\mathbf{h})$:

$$C(\mathbf{x},\mathbf{x}') = K(\mathbf{x}-\mathbf{x}') - \sum_{i=1}^{n} \lambda_i K(\mathbf{x}_i-\mathbf{x}) - \sum_{i=1}^{n} \lambda_i K(\mathbf{x}_i-\mathbf{x}') + \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j K(\mathbf{x}_i-\mathbf{x}_j) \quad (12)$$

This internal representation can be standardized to a Gaussian random field with a unit variance by putting:

$$\forall \mathbf{x} \in \mathbb{R}^d, \tilde{Y}(\mathbf{x}) = \frac{Y(\mathbf{x})}{\sqrt{C(\mathbf{x},\mathbf{x})}} \quad (13)$$

The covariance function of $\tilde{Y}$ is

$$\rho(\mathbf{x},\mathbf{x}') = \frac{C(\mathbf{x},\mathbf{x}')}{\sqrt{C(\mathbf{x},\mathbf{x})\,C(\mathbf{x}',\mathbf{x}')}} \quad (14)$$

The indicator and truncation threshold (Eq. 7) can now be rewritten as:

$$I_z(\mathbf{x}) = \begin{cases} 1 & \text{if } \tilde{Y}(\mathbf{x}) > y(\mathbf{x}) \\ 0 & \text{otherwise} \end{cases} \quad (15)$$

with

$$y(\mathbf{x}) = \frac{z - p(\mathbf{x})}{\sqrt{C(\mathbf{x},\mathbf{x})}} \quad (16)$$

where $p(\mathbf{x}) = \mathbf{Z}_0^T \mathbf{F}_0^T (\mathbf{F}_0 \mathbf{F}_0^T)^{-1} \mathbf{F}(\mathbf{x})$ is a polynomial of the coordinates of $\mathbf{x}$, the coefficients of which depend on the components of vector $\mathbf{Z}_0$. Recalling the relationship between the covariance functions of a standard Gaussian random field and its indicator (Eq. 6), one obtains:

$$C_{I_z}(\mathbf{x},\mathbf{x}') = g(y(\mathbf{x}))\, g(y(\mathbf{x}')) \times \sum_{p=1}^{+\infty} H_{p-1}(y(\mathbf{x})) H_{p-1}(y(\mathbf{x}'))[\rho(\mathbf{x},\mathbf{x}')]^p \quad (17)$$

In practice, the above summation has to be truncated at a finite order, which must be high enough (say, $p = 200$) for the truncation error to be negligible. Finally, the non-centered indicator covariance between any two locations $\mathbf{x}$ and $\mathbf{x}'$ can be expressed as:

$$E\{I_z(\mathbf{x}) I_z(\mathbf{x}')\} = C_{I_z}(\mathbf{x},\mathbf{x}') + G(y(\mathbf{x}))\, G(y(\mathbf{x}')) \quad (18)$$

where G is the standard Gaussian cumulative distribution function. The right-hand side of this equation depends on the generalized covariance $K(\mathbf{h})$, the chosen locations $\{\mathbf{x}_i: i = 1\ldots n\}$ (known), the truncation threshold z and the components of vector $\mathbf{Z}_0$ [through the difference z–$p(\mathbf{x})$ that defines the numerator of $y(\mathbf{x})$ in Eq. (16)], while the left-hand side can be calculated experimentally for each pair of data locations, for which the indicator (0 or 1) is known. Accordingly, if the generalized covariance can be parameterized through a few parameters (e.g., by considering a set of known basic models weighted by unknown nonnegative coefficients), one can jointly estimate these parameters and the coefficients of z–$p(\mathbf{x})$ by least squares fitting.

Having determined z–$p(\mathbf{x})$ may not directly provide information on the truncation threshold z, which agrees with the formalism of intrinsic random fields: any random field that differs from Z by a constant is a representation of the same IRF-k and could be chosen as the random field to be truncated. Accordingly, without loss of generality, z can be set to any arbitrary value (for instance, 0).

In case of a complex truncation rule where several thresholds, say $z_1 \ldots z_q$, are applied to the same representation Z, then a similar equation can be derived to express the indicator direct and cross covariances as a function of the parameters of the generalized covariance $K(\mathbf{h})$, truncation thresholds and coefficients of the polynomial $p(\mathbf{x})$. In such a case, the least squares fitting allows jointly estimating the parameters of the generalized covariance together with the coefficients of z_j–$p(\mathbf{x})$ for $j = 1 \ldots q$: the relative values of $z_1 \ldots z_q$ are then found and it remains to fix one of these thresholds to get a solution, e.g., $z_1 = 0$.

2.2.2.3 Summary The steps for inferring the truncation thresholds and spatial correlation structure of the IRF-k are therefore the following:

(1) Choose an order k for the intrinsic random field to truncate.
(2) Choose a set of n locations $\{\mathbf{x}_i: i = 1\ldots n\}$ to define the internal representation (Eq. 11).
(3) Choose the basic nested structures that will compose the generalized covariance $K(\mathbf{h})$. Some parameters of these structures (scale factors, sills or multiplicative coefficients, shape parameters, exponents) may be assumed known, while others are unknown and will be determined through the following steps.
(4) Calculate the experimental covariance matrix of the indicator data (the matrix entries are 1 or 0).
(5) Calculate the theoretical covariance matrix of the indicator data by using Eq. (18). The entries of this matrix depend on the covariance parameters to be fitted and on the (unknown) values of the intrinsic random field (vector $\mathbf{Z}_0$) at locations $\{\mathbf{x}_i: i = 1\ldots n\}$.
(6) Find the covariance parameters and vector $\mathbf{Z}_0$ that minimize the total sum of squared errors between the experimental and theoretical covariance matrices. Since this least squares optimization problem is nonlinear in the parameters, it is advisable to keep

Table 2 Main rock types in the Río Blanco deposit

Code	Rock type	Abbreviation	Age
1	Pipe	PIP	Early Miocene
2	Porphyry	POR	Middle Miocene
3	Magmatic breccia	MAB	Late Miocene
4	Monolithic breccia	MOB	Early Pliocene
5	Tourmaline breccia	TOB	Early Pliocene
6	Granitoids	GD	Early Pliocene
7	Andesite	AND	Late Cretaceous to Early Pliocene

the generalized covariance model as parsimonious as possible, by avoiding choosing an excessive number n of locations on which to construct the internal representation (step 2) or an excessive number of nested structures or unknown parameters (step 3).

The above fitting procedure has to be repeated for each underlying IRF-k used in the plurigaussian model. The outputs are estimates of the values of vector $\mathbf{Z}_0$ and of the unknown parameters of the generalized covariance $K(\mathbf{h})$. Note that the order (step 1), set of locations used to construct the internal representation (step 2) and set of basic nested structures (step 3) are chosen by the practitioner, but they need not be the same for all the IRF-k. Also note that $\mathbf{Z}_0$ depends on the initial representation of the intrinsic random field at the chosen locations $\{\mathbf{x}_i: i = 1 \ldots n\}$ and will change in other realizations of this random field, while the generalized covariance parameters are the same for all the realizations of the intrinsic random field. Therefore, one only needs to keep the latter parameters for the subsequent simulation stage and should "forget" the value of $\mathbf{Z}_0$ delivered by the least-squares fitting algorithm.

As for most of the structural analysis approaches for an IRF-k (Chilès and Delfiner 2012), the identification of the spatial structure turns out to be semi-automatic and does not rest on a graphical fitting of some experimental covariance or variogram. This is the price to pay for dealing with non-stationary random fields.

2.2.3 *Conditional simulation of a non-stationary plurigaussian model*

The process of conditional simulation is similar to the stationary case exposed in Sect. 2.1, except for the following:

(1) The non-conditional simulation of an IRF-k with Gaussian generalized increments can be efficiently done by discrete or continuous spectral approaches (Stein 2002; Emery and Lantuéjoul 2006, 2008; Arroyo and Emery 2016) or by Gibbs sampling (Arroyo and Emery 2015). Most of the other Gaussian simulation algorithms fail or are approximate when applied to non-stationary random fields.

(2) In both the Gibbs sampler (step 1) and the post-conditioning of the realizations to data (step 2), one has to replace simple kriging by intrinsic kriging when determining the conditional distributions (Delfiner 1976; Emery 2008; Chilès and Delfiner 2012). Also, care must be taken for the Gibbs sampler, as the convergence of this iterative algorithm is ensured only when using a unique neighborhood implementation (Emery et al. 2014).

3 Case study

The plurigaussian model is now applied to a dataset from the Río Blanco copper deposit, located in the Central Chilean Andes. A geological description of this deposit can be found in Stambuk et al. (1988), Skewes and Stern (1995) and Serrano et al. (1996).

3.1 Available data

In the following, we will focus on a reduced region covering an area of 1.87 km^2 in the northern sector of the Río Blanco deposit, for which two types of data are available: exploration drill holes and an interpreted lithological model. The drill holes are in an irregular design along the northwest-southeast direction; the data consist of information on the prevailing rock types logged at 125 drill hole cores composited at 16 meters and codified into seven main categories (from the oldest to the youngest: andesite, granitoids, tourmaline breccia, monolithic breccia, magmatic breccia, porphyry and pipe) (Table 2). The lithological model consists of a discretization of the region into 8500 blocks with a volumetric support of 15 m $\times$ 15 m $\times$ 16 m, for which the (assumed) prevailing rock type domain is assigned to each block (Fig. 3). Along the vertical direction, the lithological model is limited to a single elevation (3348 m), while the 125 drill hole data are located between elevations 3340 and 3356 m.

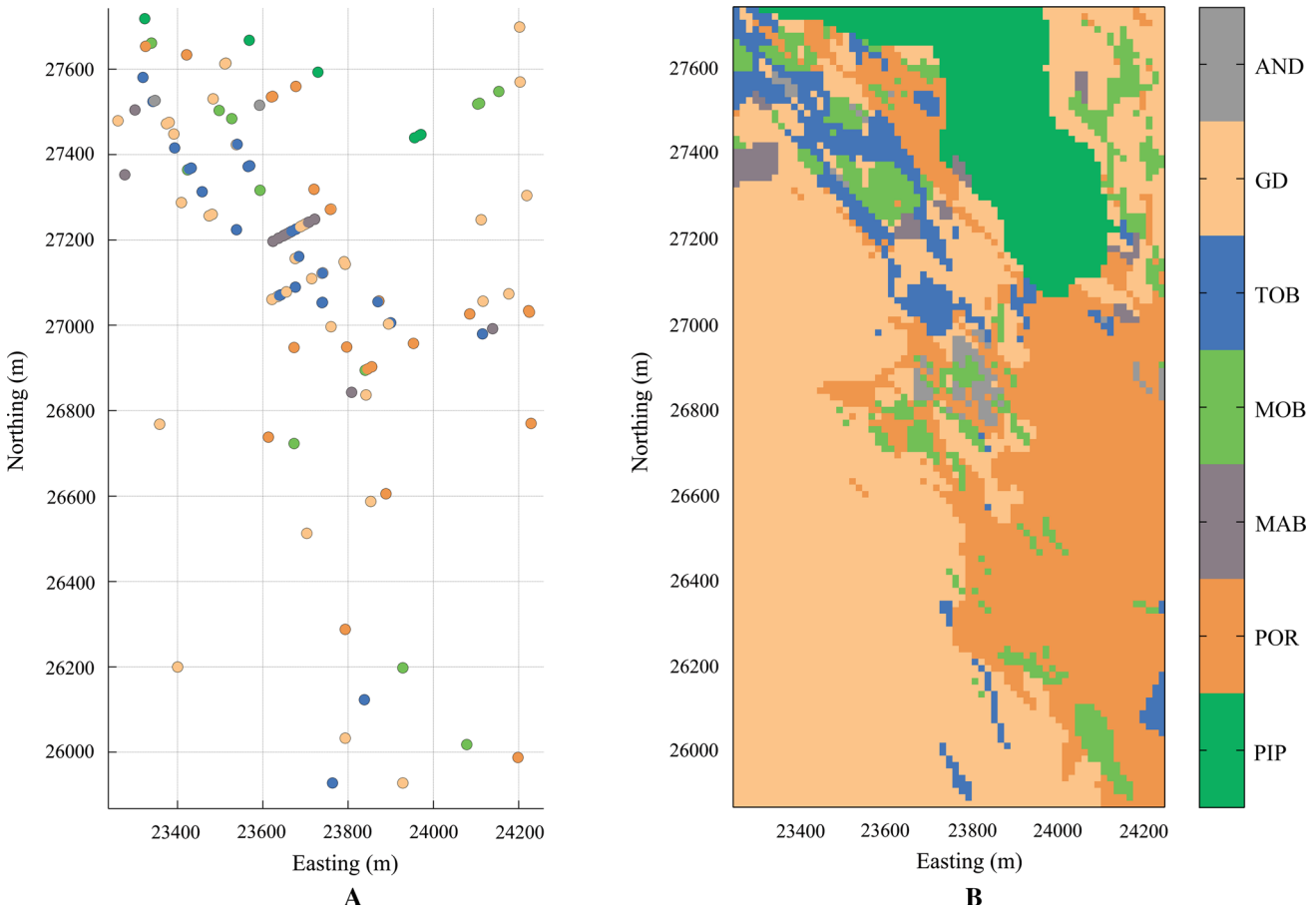

Fig. 3 Plan view showing **a** the drill hole data and **b** the interpreted lithological model for elevation 3348 m

Table 3 Estimated proportions of rock type domains

Rock type	Proportion in drill holes data (%)	Declustered proportions in drill hole data (%)	Proportions in interpreted lithological model (%)
PIP	4.8	7.2	10.5
POR	16.8	24.3	26.2
MAB	11.2	3.4	1.2
MOB	11.2	11.9	6.6
TOB	20.0	9.4	5.4
GD	32.8	42.6	49.1
AND	3.2	1.2	1.1

The proportions of rock type domains calculated from the interpreted lithological model and from the 125 drill hole data, before and after declustering, are presented in Table 3. The declustered proportions are obtained by interpolating the region of interest with the nearest neighbor technique, i.e., by assigning to each grid node the rock type observed on the closest drill hole data. For a given rock type, the deviations between the three calculated proportions are significant and reflect the uncertainty in the true unknown proportion.

3.2 Stationary plurigaussian model

In this subsection, a plurigaussian model will be elaborated based on the proportions of rock type domains calculated from the interpreted lithological model (Table 3), which here are assumed to be constant in space and perfectly known.

3.2.1 Truncation rule

In order to be consistent with the rock type chronology (Table 2), a model based on six independent Gaussian random fields $\{Z_1, Z_2, Z_3, Z_4, Z_5, Z_6\}$ and six truncation thresholds $\{z_1, z_2, z_3, z_4, z_5, z_6\}$ is proposed (Madani and Emery 2015), in which the rock type present at a given location $\mathbf{x}$ is defined as

$$R(\mathbf{x}) = \begin{cases} 1 & \text{if } Z_1(\mathbf{x}) \le z_1 \\ 2 & \text{if } Z_1(\mathbf{x}) > z_1 \text{ and } Z_2(\mathbf{x}) \le z_2 \\ 3 & \text{if } Z_1(\mathbf{x}) > z_1, Z_2(\mathbf{x}) > z_2 \text{ and } Z_3(\mathbf{x}) \le z_3 \\ 4 & \text{if } Z_1(\mathbf{x}) > z_1, Z_2(\mathbf{x}) > z_2, Z_3(\mathbf{x}) > z_3 \text{ and } Z_4(\mathbf{x}) \le z_4 \\ 5 & \text{if } Z_1(\mathbf{x}) > z_1, Z_2(\mathbf{x}) > z_2, Z_3(\mathbf{x}) > z_3, Z_4(\mathbf{x}) > z_4 \text{ and } Z_5(\mathbf{x}) \le z_5 \\ 6 & \text{if } Z_1(\mathbf{x}) > z_1, Z_2(\mathbf{x}) > z_2, Z_3(\mathbf{x}) > z_3, Z_4(\mathbf{x}) > z_4, Z_5(\mathbf{x}) > z_5 \text{ and } Z_6(\mathbf{x}) \le z_6 \\ 7 & \text{if } Z_1(\mathbf{x}) > z_1, Z_2(\mathbf{x}) > z_2, Z_3(\mathbf{x}) > z_3, Z_4(\mathbf{x}) > z_4, Z_5(\mathbf{x}) > z_5 \text{ and } Z_6(\mathbf{x}) > z_6 \end{cases} \tag{19}$$

With this truncation rule, all transitions between rock types are allowed, i.e., every rock type domain is likely to be in contact with any other.

3.2.2 Truncation thresholds

Accounting for the hierarchical nature of the model, the truncation thresholds $\{z_1, z_2, z_3, z_4, z_5, z_6\}$ are related to the proportions $\{p_1, p_2, p_3, p_4, p_5, p_6, p_7\}$ of the rock type domains in the following fashion:

$$\begin{aligned} p_1 &= G^{-1}(z_1) \\ p_2 &= [1 - G^{-1}(z_1)]G^{-1}(z_2) \\ p_3 &= [1 - G^{-1}(z_1)][1 - G^{-1}(z_2)]G^{-1}(z_3) \\ p_4 &= [1 - G^{-1}(z_1)][1 - G^{-1}(z_2)][1 - G^{-1}(z_3)]G^{-1}(z_4) \\ p_5 &= [1 - G^{-1}(z_1)][1 - G^{-1}(z_2)][1 - G^{-1}(z_3)] \\ &\quad [1 - G^{-1}(z_4)]G^{-1}(z_5) \\ p_6 &= [1 - G^{-1}(z_1)][1 - G^{-1}(z_2)][1 - G^{-1}(z_3)] \\ &\quad [1 - G^{-1}(z_4)][1 - G^{-1}(z_5)]G^{-1}(z_6) \\ p_7 &= [1 - G^{-1}(z_1)][1 - G^{-1}(z_2)][1 - G^{-1}(z_3)] \\ &\quad [1 - G^{-1}(z_4)][1 - G^{-1}(z_5)][1 - G^{-1}(z_6)] \end{aligned} \tag{20}$$

The last equation is actually redundant with the first six ones and only makes p_7 be the complement of the sum of proportions p_1–p_6. Knowing the proportions (p_1–p_7) of rock type domains, one can derive the truncation thresholds (z_1–z_6) to be applied.

3.2.3 Variogram analysis

At each data location, the information on the prevailing rock type can be codified into indicators, with one indicator associated with each Gaussian random field. The value of an indicator is 1 when the Gaussian random field is less than the associated threshold, 0 when it is greater than the threshold, or unknown when the Gaussian random field is not involved in the definition of the rock type domain (Table 4).

Variogram analysis relies on the calculation of the experimental variograms of the indicator data. To get more robust experimental variograms, the data from a wider region of the deposit, totalizing 12,006 drill hole data, have been used, insofar as the experimental variograms were not interpretable when considering only the 125 data in the region of interest. The three main anisotropy directions were identified as N20°W, N70°E and vertical, which agrees with the interpreted lithological model that exhibits a greater spatial continuity along the direction N20°W in the horizontal plane.

From the indicator variograms, one can derive the variograms of the underlying Gaussian random fields (Emery and Cornejo 2010) and fit a model to each experimental variogram. To ensure that the Gaussian random fields have a unit variance, the fitting is performed with a semi-automated algorithm (Emery 2010) that minimizes the mean squared error between experimental and modeled variograms under the constraint of a unit sill (Table 5). The reader is referred to Madani and Emery (2015) for details on the variogram calculation and modeling.

3.2.4 Conditional simulation and post-processing of the realizations

Provided with the truncation rule, truncation thresholds and variograms of the underlying Gaussian random fields, a total of one hundred realizations are constructed in the region of interest; the parameters used for simulation are indicated in Table 6.

With the realizations so obtained, one can assess the uncertainty in the rock type domains at a local (node-by-node) scale, by means of probability maps. The maps are constructed by calculating, for each grid node, the frequency of occurrence of each rock type domain over the 100 conditional realizations (Fig. 4). The probability map obtained for the first rock type domain (PIP) relatively agrees with the interpreted lithological model (Fig. 3b). In contrast, the probability map obtained for the sixth rock type domain (GD) bears little resemblance to the interpreted model, while that of the second rock

Table 4 Codification of rock type domains into indicators

Rock type	Rock code	$Z_1 < z_1$?	$Z_2 < z_2$?	$Z_3 < z_3$?	$Z_4 < z_4$?	$Z_5 < z_5$?	$Z_6 < z_6$?
PIP	1	1	Unknown	Unknown	Unknown	Unknown	Unknown
POR	2	0	1	Unknown	Unknown	Unknown	Unknown
MAB	3	0	0	1	Unknown	Unknown	Unknown
MOB	4	0	0	0	1	Unknown	Unknown
TOB	5	0	0	0	0	1	Unknown
GD	6	0	0	0	0	0	1
AND	7	0	0	0	0	0	0

Table 5 Fitted variogram models for the stationary case

Stationary Gaussian random field	Basic nested structure	Sill	Scale factor along N20°W (m)	Scale factor along N70°E (m)	Scale factor along vertical (m)
1	Cubic	1.000	1000	450	900
2	Cubic	0.713	70	40	80
	Cubic	0.287	500	60	Infinite
3	Cubic	1.000	1200	280	600
4	Cubic	0.084	20	30	20
	Cubic	0.403	300	250	400
	Cubic	0.513	800	250	Infinite
5	Cubic	0.013	10	60	60
	Cubic	0.117	60	60	60
	Cubic	0.314	300	200	220
	Cubic	0.556	Infinite	500	2500
6	Cubic	1.000	700	280	500

Table 6 Parameters for plurigaussian simulation

Simulation parameter	Value
Number of iterations (sweeps over the data set) for Gibbs sampler	3000
Number of lines for spectral-turning bands simulation	1000
Number of neighboring data for conditioning	All the data (125)
Radius (m) of search ellipsoid along direction N20°W	Infinite
Radius (m) of search ellipsoid along direction N70°E	Infinite
Radius (m) of search ellipsoid along vertical direction	Infinite

type domain (POR) is patchy and does not match at all. For these last two rock type domains, the probability is close to zero or to one near the conditioning data locations, but intermediate over the rest of the region, which means that the layout of the domain boundaries is highly uncertain.

These features were foreseeable as the result of assuming a stationary model, in which the probability of occurrence of each rock type domain is constant in space, reflecting a phenomenon that tends to repeat itself in space without any trend. The conditioning to the drill hole data improves the results only locally, but its effect is modest due to the small number of such data.

3.3 Plurigaussian model with spatially varying proportions

The stationary model developed in the previous subsection faces two limitations:

(1) The inability to reproduce spatial trends in the distribution of rock type domains.
(2) Possible misspecifications in the proportions of rock type domains, which may bias the results.

A non-stationary model may be able to overcome these limitations. One option is to assume that the proportions of rock type domains vary locally and to infer the truncation

thresholds accordingly, while considering that the underlying Gaussian random fields remain stationary. The calculation of local proportions of rock type domains can be done on the basis of the drill hole data or of the interpreted lithological model, by means of moving windows statistics or geographically-weighted statistics (Armstrong et al. 2011). A challenge is the choice of the moving window size, or of the kernel function in the case of geographically weighted statistics, which greatly influences the results. This choice is actually a trade-off between our uncertainty and our reliance on the interpreted model. With a very large window, the calculated proportions will tend to the global proportions, meaning that the spatial variations of rock type domain proportions that are seen in the interpreted lithological model tend to be ignored. On the contrary, with a very small window, the local proportions will be close to 0 or 1, depending on whether or not the rock type domain is present in the interpreted model at the target location; in such a case, one is extremely confident in the interpreted lithological model and the simulated rock type domains will not depart from this model too much.

As an illustration, an application of this approach with four different moving window sizes ($50 \times 50 \times 16$, $400 \times 400 \times 16$ and $1500 \times 1500 \times 16$ m^3) is presented in Fig. 5, where the local rock type proportions have been calculated from the interpreted lithological model shown in Fig. 3b, while the variograms of the Gaussian random fields are the same as those fitted in Sect. 3.2. In the first case (window of size $50 \times 50 \times 16$ m^3), the results indicate very little uncertainty, as the probability of finding a rock type is almost always close to 0 or to 1. In the third case (window of size $1500 \times 1500 \times 16$ m^3), the zonation of the rock types, especially the porphyry, is lost and the results are similar to those obtained in the scope of the stationary model. The second case (window of size $400 \times 400 \times 16$ m^3) provides a more balanced trade-off that yields geologically plausible probability maps.

However, even if it solves the first abovementioned limitation (reproduction of trends), the use of spatially varying proportions makes the second limitation more critical, insofar as misspecifications in the domain proportions may be more severe locally than globally.

3.4 Plurigaussian model based on truncated IRF-*k*

In this subsection, we explore the possibility of simulating the rock types without using the interpreted lithological model, by truncating intrinsic random fields of order *k* instead of stationary random fields. This idea makes sense not only to reproduce spatial trends or zonal features in the distribution of rock type domains, but also to avoid misspecifying the truncation thresholds. Therefore, the two abovementioned limitations can be addressed in the IRF-*k* model.

3.4.1 Truncation rule and truncation thresholds

The truncation rule remains the same as that presented in the stationary model (Sect. 3.2) and the truncation

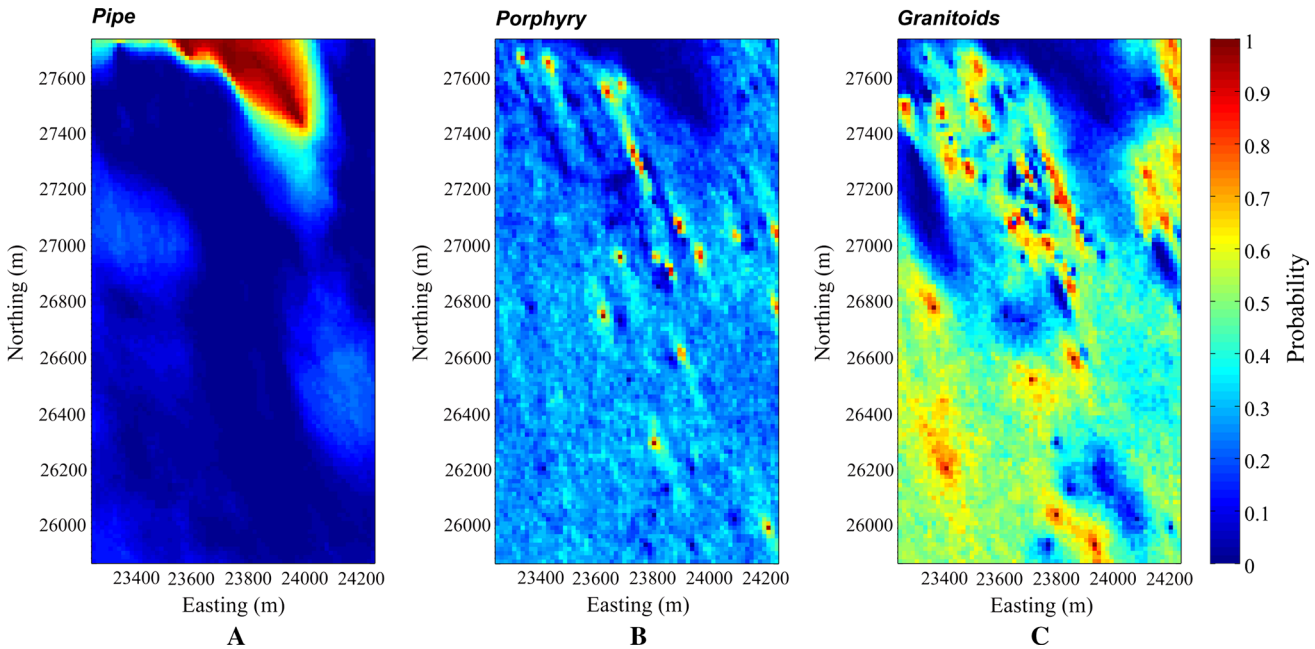

Fig. 4 Plan views showing the probabilities of occurrence of three rock type domains at elevation 3348 m, calculated by using 100 realizations (stationary plurigaussian model)

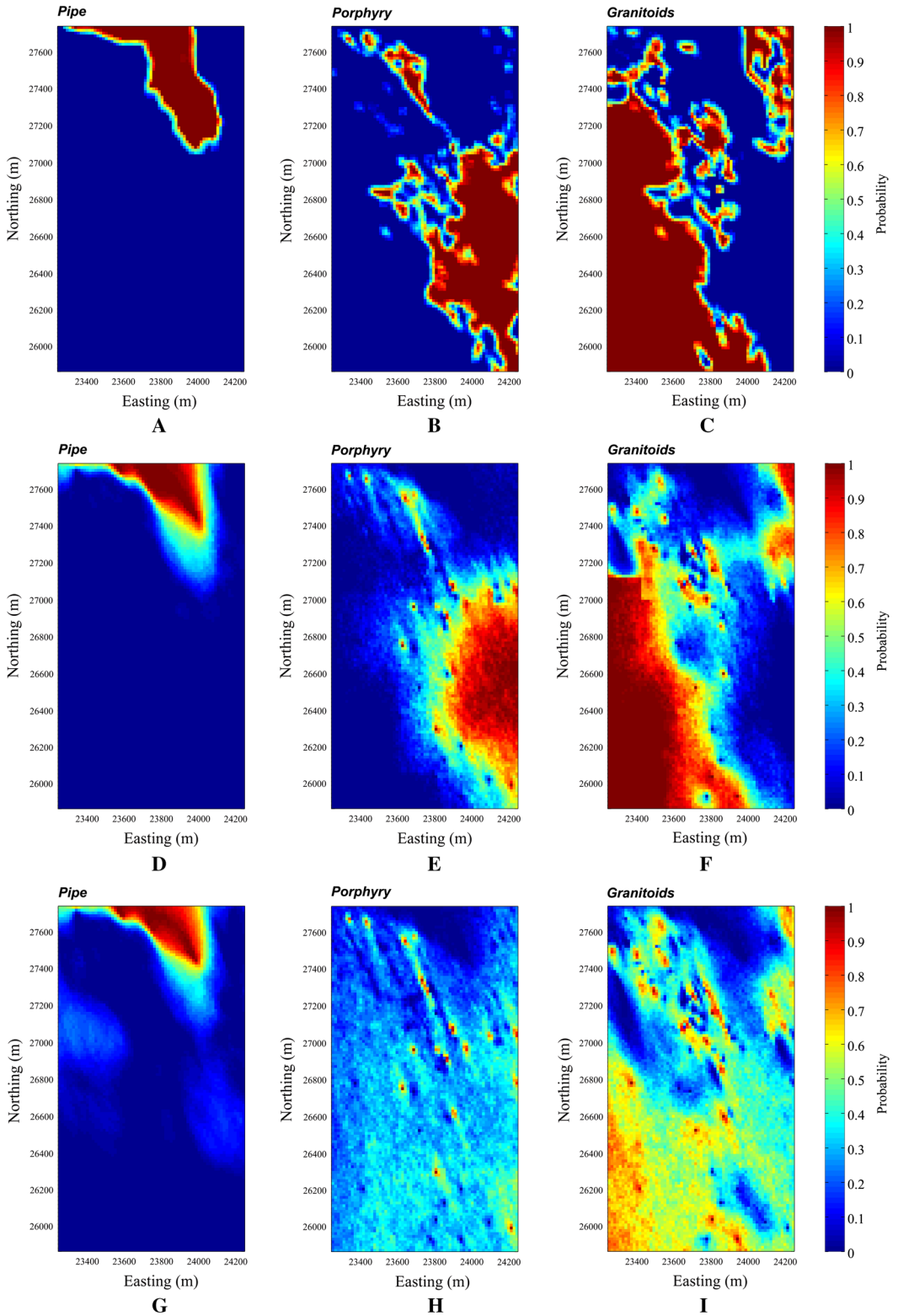

Fig. 5 Plan views showing the probabilities of occurrence of three rock type domains at elevation 3348 m, calculated by using 100 realizations [plurigaussian model with local rock type proportions obtained with a moving window of size $50 \times 50 \times 16$ m^3 (**a**–**c**), $400 \times 400 \times 16$ m^3 (**d**–**f**) and $1500 \times 1500 \times 16$ m^3 (**g**–**i**)]

thresholds can be set to zero, as previously explained in Sect. 2.2.2.

3.4.2 Generalized covariance analysis

To complete the determination of the model parameters, it remains to infer the spatial correlation structure (i.e., the generalized covariance functions) of the intrinsic random fields of order k to be truncated. To this end, the procedure presented in Sect. 2.2.2 is successively applied to each IRF-k:

(1) An order k is chosen, which represents the degree of the polynomial drift.
(2) The number n of locations used to define the internal representation of the IRF-k is set to the minimal allowable value $\left(\text{i.e., } n = L = \frac{(k+3)!}{k!\,3!}\right)$, so as to minimize the size of vector $\mathbf{Z}_0$ and, accordingly, the total number of parameters to be determined by least-squares fitting. Regarding their coordinates, they are drawn randomly within the limits of the region of interest and will remain fixed during the fitting procedure.
(3) The same basic nested structures as those found in the stationary model (Table 5) are considered, together with power generalized covariance models whose exponents and scale factors are chosen after trial and error, so that only the sills of the stationary structures and the multiplicative coefficients of the power structures remain to be found with the proposed semi-automated fitting algorithm.

A criterion for choosing the order k at step (1) and the exponents and scale factors of the power structures at step (3) is to obtain a final fit that minimizes the total sum of squared errors between the experimental and modeled indicator covariance matrices (Eq. 18). As a complement, one could choose the order, exponents and scale factors that yield a good prediction of the rock types through cross-validation at the data locations, or that provide probability maps reproducing the expected zonation of the rock types within the region of interest (pipe located to the north, porphyry mainly to the southeast, and granitoids to the west and southwest). Rather than a detailed geological interpretation as in Fig. 3b, a sketch of such a zonation, based on a global understanding of the deposit, is sufficient for such a purpose.

The random fields no 1, 3, 4 and 5 have been chosen with a power covariance of exponent 1.5 and are therefore IRF-0, while the random fields no 2 and 6 (with a power covariance of exponent 3) are IRF-1. The number of parameters to be fitted with the proposed semi-automated algorithm is equal to the size n of vector $\mathbf{Z}_0$ (1 for an IRF-0, 4 for an IRF-1), plus the number of basic nested structures used in the generalized covariance model, as each structure is associated with a single unknown parameter corresponding to its sill or multiplicative coefficient. This entails a nonlinear least square optimization problem in $\mathbb{R}^d$, with d varying between 3 (for intrinsic random fields no 1 and 3) and 7 (for intrinsic random field no 2). Tables 7 and 8 present the final fit of the generalized covariances and the cross-validation results, respectively.

3.4.3 Conditional simulation and post-processing of the realizations

Having completed the specification of the model parameters, one can simulate the rock type domains over the region of interest conditionally to the drill hole data. This is similar to the stationary case, except for the use of intrinsic kriging of order k instead of simple kriging in the Gibbs sampler and post-conditioning process. One hundred realizations are constructed, with the same implementation parameters as in the stationary case (Table 6).

Figure 6 shows the interpreted lithological model together with a plan view of the first three realizations. In this case, it is noticeable that the realizations bear resemblance to the interpreted model, although the latter is not used as an input in the simulation process. In particular, the layouts of porphyry (POR) and granitoids (GD) are much more realistic than with the stationary model. As before, the realizations can be used to assess the uncertainty locally at each target grid node, through probability maps (Fig. 7). Despite the fluctuations on each individual realization, these maps, which are calculated by processing all the realizations, are strikingly similar to the interpreted lithological model for pipe PIP, porphyry (POR) and granitoids (GD) (Fig. 3b).

From the probability maps, one can construct a lithological model by selecting, for each grid node, the most probable rock type domain. This model can then be compared to the interpreted lithological model in order to identify the grid nodes for which the interpreted rock type domain matches the most probable one, and the grid nodes for which there is a mismatch and the interpreted model may be revised or additional drill holes may be taken to reduce uncertainty (Fig. 8). The latter nodes, mostly located near the boundaries of pipe, porphyry and granitoids, correspond to nodes with a moderate probability (0.3–0.6) for the most probable rock type domain (in other words, none of the rock type domains has a high probability of occurrence). As indicated in Table 9, one obtains 5921 matches over a total of 8500 grid nodes, i.e., the interpreted rock type domain coincides with the most probable domain

Table 7 Fitted generalized covariance models for the non-stationary case

Intrinsic Gaussian random field	Basic nested structure	Sill or multiplicative coefficient	Scale factor along N20°W (m)	Scale factor along N70°E (m)	Scale factor along vertical (m)	Exponent
1	Cubic	0.095	1000	450	900	
	Power	0.021	2000	1000	2000	1.5
2	Cubic	0.046	70	40	80	
	Cubic	0.059	500	60	infinite	
	Power	8.56×10^{-6}	1500	1000	1750	3.0
3	Cubic	0.115	1200	280	600	
	Power	0.393	4500	1000	2000	1.5
4	Cubic	0.182	20	30	20	
	Cubic	0.183	300	250	400	
	Cubic	1.42×10^{-6}	800	250	infinite	
	Power	0.317	1800	1000	3000	1.5
5	Cubic	0.011	10	60	60	
	Cubic	0.004	60	60	60	
	Cubic	0.053	300	200	220	
	Cubic	0.145	infinite	500	2500	
	Power	455.6	18	10	13	1.5
6	Cubic	13.8	700	280	500	
	Power	19.9	2.5	1	2	3.0

Table 8 Cross-validation results for the plurigaussian model based on the truncation of IRF-k

True rock type domain	Most probable rock type domain							
	PIP	POR	MAB	MOB	TOB	GD	AND	Total
PIP	5	1	0	0	0	0	0	6
POR	0	20	0	1	0	0	0	21
MAB	0	0	11	0	1	2	0	14
MOB	0	5	0	5	2	2	0	14
TOB	0	2	0	0	15	7	1	25
GD	0	3	1	1	6	30	0	41
AND	0	0	0	0	0	0	4	4
Total	5	31	12	7	24	41	5	125

For 72% of the drill hole samples (90 out of 125), the most probable rock type (i.e., the rock type that appears most frequently among 100 realizations drawn at the sample location, conditionally to the data observed at the remaining samples) matches the actually observed rock type

for 69.7% of the grid nodes. This result is highly satisfactory, given the very small number of conditioning data. For comparison, in the stationary model presented in Sect. 3.2, the percentage of matches is only 54.0% (Table 10), which demonstrates the significant improvement of the proposed non-stationary model based on IRF-k over the stationary one. Concerning the model with spatially varying proportions (Sect. 3.3), the percentage of matches ranges from 92.0 to 58.7%, depending on the size of the moving window used for calculating the local proportions (Tables 11, 12, 13). However, these percentages are misleading because the local proportions used as an input for constructing the realizations are calculated from the same interpreted lithological model against which the realizations are compared; hence a high percentage of matches is expectable in these cases.

4 Discussion

The plurigaussian model can be extended to the non-stationary framework by truncating intrinsic random fields of order k with Gaussian generalized increments instead of stationary Gaussian random fields. Conditional simulation is performed similarly to the stationary case, except for the type of kriging used in the Gibbs sampler and in the post-

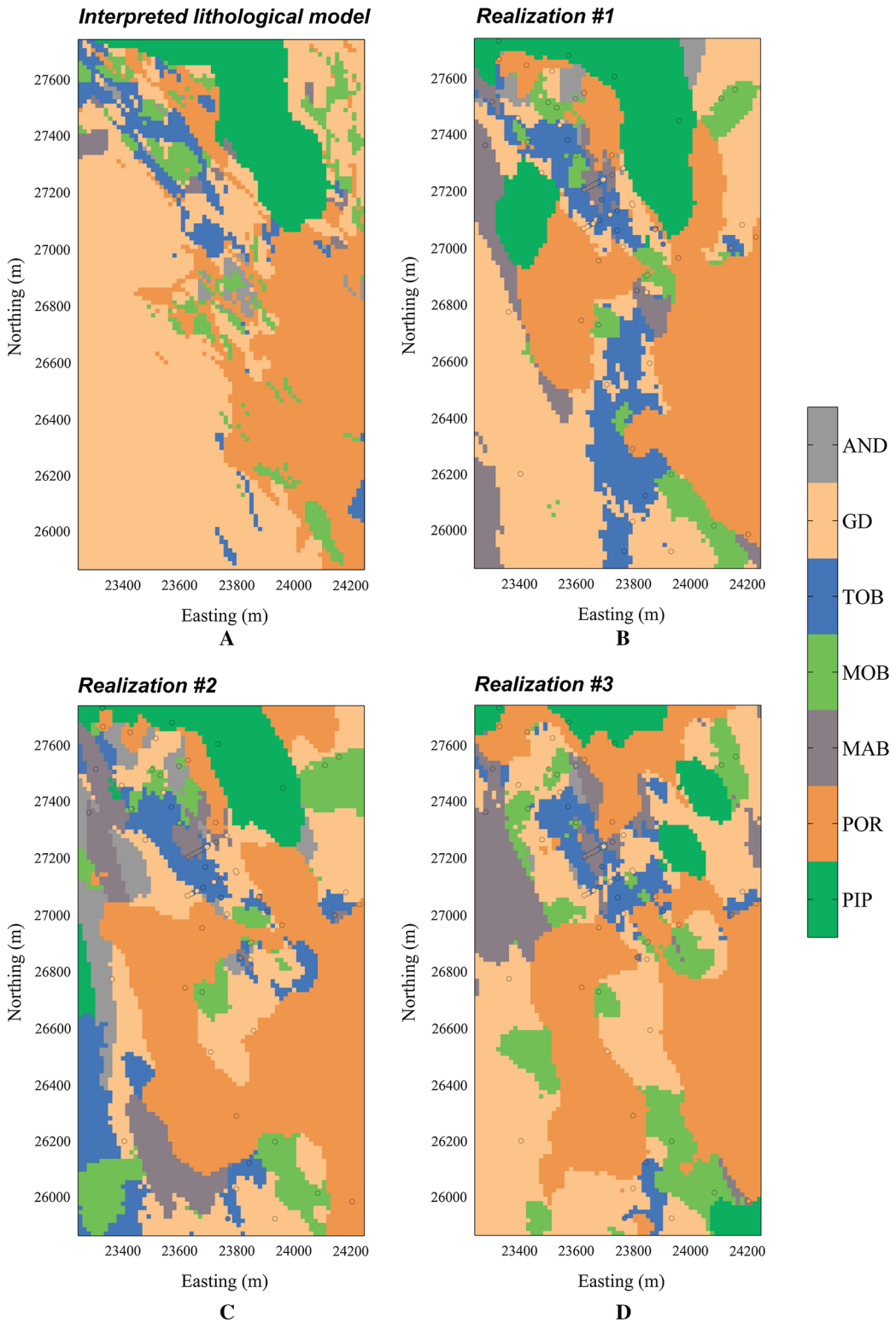

Fig. 6 Plan views of the interpreted lithological model **a** and three realizations of the rock type domains **b–d** for elevation 3348 m (plurigaussian model based on the truncation of IRF-*k*)

conditioning stage, where intrinsic kriging of order *k* replaces simple kriging, which only requires the knowledge of the generalized covariances of the intrinsic random fields.

The benefits of the proposed non-stationary plurigaussian model are twofold. First and foremost, it allows reproducing spatial trends and zonal patterns in the

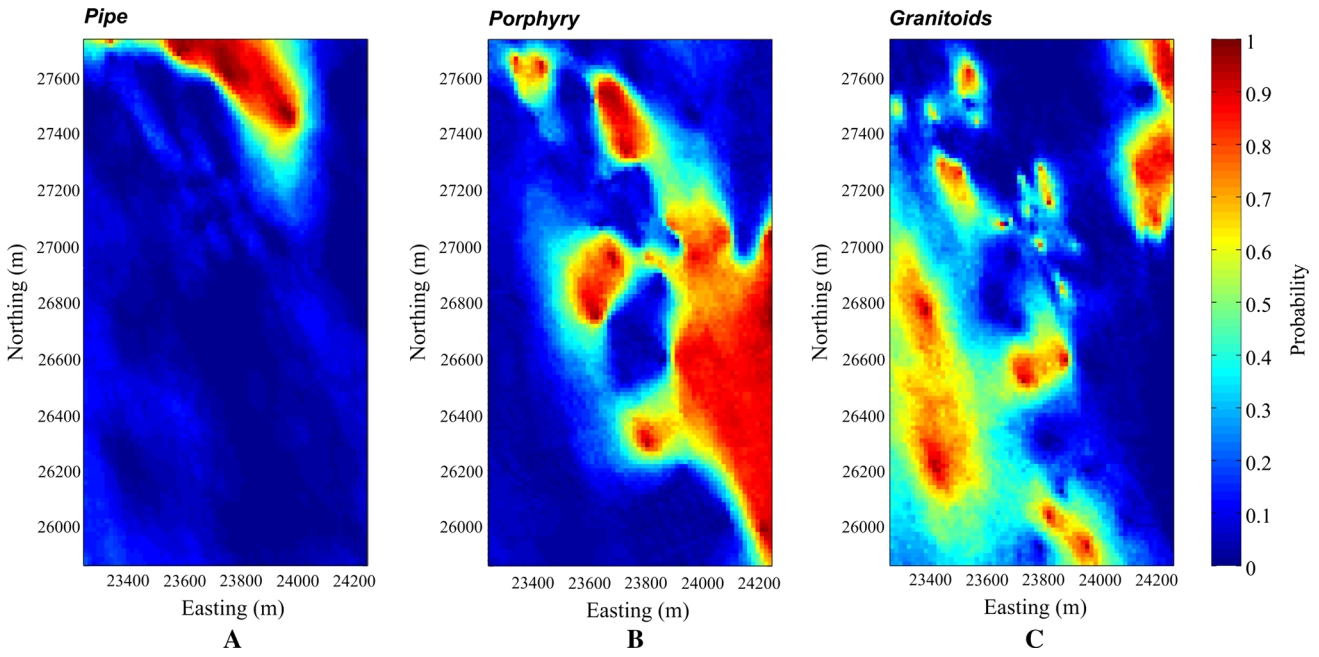

Fig. 7 Plan views showing the probabilities of occurrence of three rock type domains at elevation 3348 m, calculated by using 100 realizations (plurigaussian model based on the truncation of IRF-*k*)

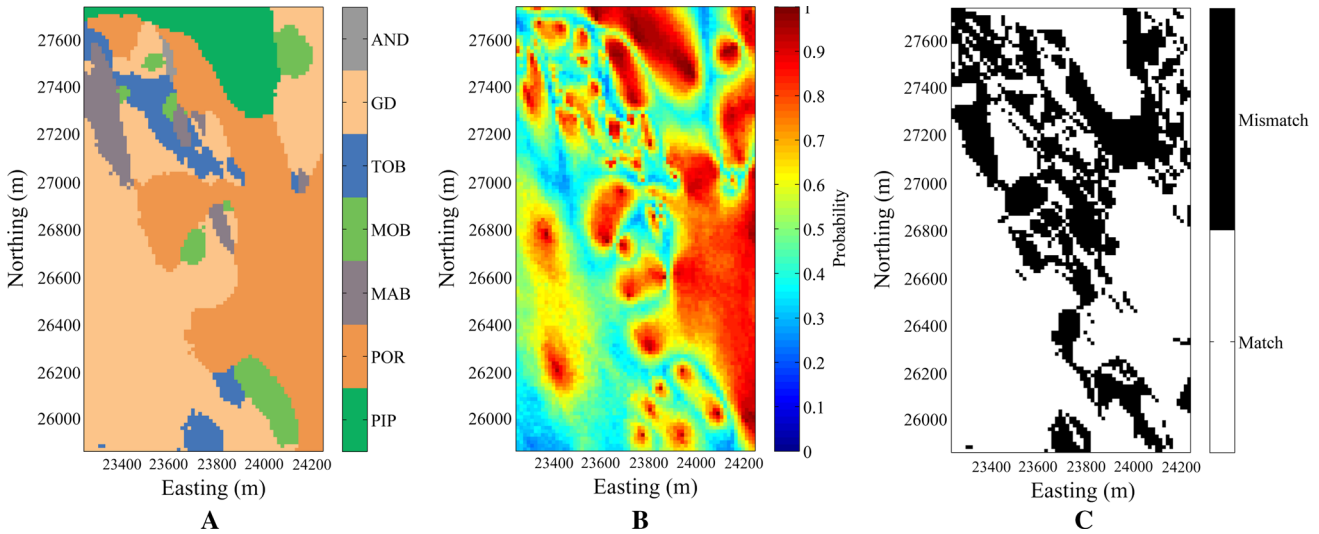

Fig. 8 Plan views showing **a** the most probable rock type domain at elevation 3348 m, **b** the probability of occurrence of the most probable domain, and **c** the matches and mismatches with respect to the interpreted lithological model (plurigaussian model based on the truncation of IRF-*k*)

distribution of the simulated geological domains. This is a common feature when modeling lithological, mineralogical and/or alteration domains due to the parameters that affect the ore body formation, such as genesis, paragenesis and regional structures. The current implementation of the plurigaussian model, based on the truncation of stationary Gaussian random fields, is not able to reproduce these trends and zonal patterns, unless spatially varying truncation thresholds are defined, which amounts to specifying spatially varying proportions of the geological domains under study. However, such an approach suffers from two important limitations when inferring the model parameters: (1) it assumes a perfect knowledge of the local domain proportions, which is unrealistic in practice, and (2) the theoretical background for variogram analysis is not clearly laid out. Here lies the second benefit of our proposal, insofar as the simulation process does not rely on local domain proportions, thus it is not affected by a possible misspecification of these proportions when the available data are scarce or the geology is not well known. Also, a procedure

Table 9 Numbers of grid nodes per rock type domain, according to the interpreted lithological model and to the most probable rock type domain (plurigaussian model based on the truncation of IRF-*k*)

Interpreted lithological model	Most probable rock type domain							
	PIP	POR	MAB	MOB	TOB	GD	AND	Total
PIP	606	234	0	0	0	70	0	910
POR	3	1871	24	131	26	194	15	2264
MAB	3	8	68	5	2	10	0	96
MOB	9	152	32	175	57	117	5	547
TOB	12	160	15	1	197	81	5	471
GD	103	392	247	145	225	3003	0	4115
AND	0	54	20	3	5	14	1	97
Total	736	2871	406	460	512	3489	26	8500

Table 10 Numbers of grid nodes per rock type domain, according to the interpreted lithological model and to the most probable rock type domain (stationary plurigaussian model)

Interpreted lithological model	Most probable rock type domain							
	PIP	POR	MAB	MOB	TOB	GD	AND	Total
PIP	626	58	0	0	0	226	0	910
POR	2	347	53	131	2	1714	15	2264
MAB	8	1	60	0	0	27	0	96
MOB	13	43	37	140	9	299	6	547
TOB	19	32	13	0	101	304	2	471
GD	110	209	290	122	70	3314	0	4115
AND	0	14	23	2	4	51	3	97
Total	778	704	476	395	186	5935	26	8500

Table 11 Numbers of grid nodes per rock type domain, according to the interpreted lithological model and to the model based on the most probable rock type domain (plurigaussian model with rock type proportions calculated by using a moving window of size $50 \times 50 \times 16$ m^3)

Interpreted lithological model	Most probable rock type domain							
	PIP	POR	MAB	MOB	TOB	GD	AND	Total
PIP	895	5	1	0	1	8	0	910
POR	6	2105	2	32	17	91	11	2264
MAB	3	1	88	1	0	3	0	96
MOB	3	67	4	383	22	63	5	547
TOB	1	22	4	5	394	45	0	471
GD	28	41	22	44	69	3909	2	4115
AND	0	15	0	7	8	19	48	97
Total	936	2256	121	472	511	4138	66	8500

Table 12 Numbers of grid nodes per rock type domain, according to the interpreted lithological model and to the most probable rock type domain (plurigaussian model with rock type proportions calculated by using a moving window of size $400 \times 400 \times 16$ m^3)

Interpreted lithological model	Most probable rock type domain							
	PIP	POR	MAB	MOB	TOB	GD	AND	Total
PIP	707	97	0	0	0	106	0	910
POR	8	1925	13	49	28	232	9	2264
MAB	6	5	66	2	0	17	0	96
MOB	13	155	28	134	72	140	5	547
TOB	14	80	14	1	174	188	0	471
GD	118	191	165	132	82	3427	0	4115
AND	0	47	0	8	8	32	2	97
Total	866	2500	286	326	364	4142	16	8500

Table 13 Numbers of grid nodes per rock type domain, according to the interpreted lithological model and to the most probable rock type domain (plurigaussian model with rock type proportions calculated by using a moving window of size $1500 \times 1500 \times 16$ m^3)

Interpreted lithological model	Most probable rock type domain							
	PIP	POR	MAB	MOB	TOB	GD	AND	Total
PIP	678	67	0	0	0	165	0	910
POR	6	767	38	86	2	1348	17	2264
MAB	8	1	65	0	0	22	0	96
MOB	14	80	38	132	15	263	5	547
TOB	16	50	18	0	110	276	1	471
GD	126	287	300	116	54	3232	0	4115
AND	0	27	20	1	4	42	3	97
Total	848	1279	479	335	185	5348	26	8500

has been designed to infer the spatial correlation structure on a sound basis. In the case of spatial prediction, a similar problem occurs with simple kriging when a locally varying mean value has to be defined: this kriging type is often replaced by ordinary kriging, with an unknown mean value at the scale of the kriging neighborhood, or with universal or intrinsic kriging of order k, which assumes a spatially varying and, at the same time, unknown mean value. The proposed plurigaussian model goes in the same line, in the context of simulation rather than prediction.

The case study proves the applicability of this model to simulate multiple rock type domains, whereas the conventional stationary model produces realizations of the domains that do not bear resemblance to the lithological model interpreted by geologists.

In practice, the main difficulty of the proposed non-stationary model stems from the inference of the generalized covariance functions of the underlying intrinsic random fields. Indeed, one does not have access to an experimental variogram on which to fit a variogram model, as in the stationary case. Instead, a semi-automated fitting procedure has been designed, based on the calculation of the indicator covariances between all the pairs of data locations, followed by a nonlinear least-squares optimization to determine the parameters of the generalized covariance functions to be modeled. By the way, a similar approach could be used for the variogram analysis of the plurigaussian model with spatially varying proportions, so as to avoid calculating experimental variograms that are biased estimates of the true underlying variograms. Because of the possibly high dimensionality of this optimization problem and because the results can largely depend on prior decisions of the practitioner (choice of the intrinsic random field orders, of the basic nested structures composing their generalized covariances, of their exponents and scale factors, and of the number and coordinates of the locations used to construct the internal representations), the proposed fitting algorithm actually aims at finding one reasonably good model for the spatial correlation structure of the intrinsic random fields under consideration, keeping in mind that many other solutions could be equally suitable. Some additional tips can help for the fitting, such as considering the stationary fit as an initial guess and adding non-stationary covariance structures (e.g., power generalized covariances with exponents and scale factors chosen by trial and error), or checking that the probability maps derived from a set of realizations reproduce the expected zonation of the domain distribution within the region of interest. The sensitivity of the simulated random fields to the fitted generalized covariance models is a subject of further research.

Another difficulty that has been met relates to the Gibbs sampler, whose convergence is not guaranteed if one works with a moving neighborhood implementation. Solutions exist in the case of stationary models (Lantuéjoul and Desassis 2012; Arroyo et al. 2012; Emery et al. 2014) and should be extended to the conditional simulation of intrinsic random fields of order k.

5 Conclusions

The spatial modeling of geological domains is critical to characterize geological heterogeneities in ore bodies and to further quantify the chemical, physical, mineralogical and geo-metallurgical properties of the material to be mined. As the boundaries of the geological domains are generally uncertain, stochastic spatial models are increasingly used in place of, or in complement to, deterministic models. To date, a variety of stochastic simulation approaches have been developed to model the spatial layout of geological domains. Among them, plurigaussian simulation offers a flexible framework, as it aims at reproducing (1) the topological constraints (contact relationships) on the domains, (2) the domain proportions, (3) the spatial correlation structure of these domains and (4) the available conditioning data. In this respect, one restriction relates to the reproduction of complex contact relationships, such as a directional sequencing of the domains, which cannot be

represented through a truncation rule and requires generalizing the plurigaussian model (Langlais et al. 2008). A second practical restriction relates to the definition of the domain proportions, for which the plurigaussian model is caught in a dilemma: if these proportions are assumed constant in space (stationary model), then the simulation is unable to reproduce spatial trends and zonal patterns; but if the proportions are assumed to vary in space, their inference is likely to be locally inaccurate and the model becomes sensitive to a misspecification of these proportions.

In this context, a non-stationary plurigaussian model based on the truncation of intrinsic random fields of order k instead of stationary random fields has been proposed, together with tools and algorithms for inferring the model parameters and for constructing realizations conditioned to existing data. The application case study demonstrates that, even with few conditioning data (only 125 over a geographical area of 1.87 km^2), it is possible to simulate rock type domains in agreement with the interpretation made by geologists, in a much more successful way than does the conventional stationary plurigaussian model. The proposed model therefore allows a more realistic assessment of the geological uncertainty, with its subsequent gains in the evaluation of mineral resources and ore reserves. This model combines, on the one hand, a sound theoretical framework (in particular, concerning the design of algorithms for fitting the underlying generalized covariance functions or for simulating intrinsic random fields of order k) and, on the other hand, the consideration of qualitative geological knowledge, such as the chronology, contact relationships or spatial trends of the domains to be simulated, for guiding the modeling process and validating it.

Acknowledgements This research was funded by the Chilean Commission for Scientific and Technological Research, through projects CONICYT/FONDECYT/REGULAR/No 1130085 and CONICYT PIA Anillo ACT1407. The authors also acknowledge the support from Mr. Claudio Martínez from Codelco-Chile (Andina Division), who provided the data set used in this work.

References

Armstrong M, Galli A, Beucher H, Le Loc'h G, Renard D, Renard B, Eschard R, Geffroy F (2011) Plurigaussian simulations in geosciences. Springer, Berlin

Arroyo D, Emery X (2015) Simulation of intrinsic random fields of order k with Gaussian generalized increments by Gibbs sampling. Math Geosci 47(8):955–974

Arroyo D, Emery X (2016) Spectral simulation of vector random fields with stationary Gaussian increments in d-dimensional Euclidean spaces. Stoch Environ Res Risk Assess. doi:10.1007/s00477-016-1225-7

Arroyo D, Emery X, Peláez M (2012) An enhanced Gibbs sampler algorithm for non-conditional simulation of Gaussian random vectors. Comput Geosci 46:138–148

Bernstein DS (2009) Matrix mathematics. Princeton University Press, Princeton

Beucher H, Galli A, Le Loc'h G, Ravenne C (1993) Including a regional trend in reservoir modeling using the truncated Gaussian method. In: Soares A (ed) Geostatistics Tróia'92. Kluwer Academic, Dordrecht, pp 555–566

Biver P, Haas A, Bacquet C (2002) Uncertainties in facies proportion estimation II: application to geostatistical simulation of facies and assessment of volumetric uncertainties. Math Geol 34(6):703–714

Chilès JP, Delfiner P (2012) Geostatistics: modeling spatial uncertainty. Wiley, New York

Christakos G (1992) Random field models in earth sciences. Academic Press, San Diego

Delfiner P (1976) Linear estimation of nonstationary spatial phenomena. In: Guarascio M, David M, Huijbregts CJ (eds) Advanced geostatistics in the mining industry. Reidel, Dordrecht, pp 49–68

Emery X (2007) Simulation of geological domains using the plurigaussian model: new developments and computer programs. Comput Geosci 33(9):1189–1201

Emery X (2008) Uncertainty modeling and spatial prediction by multi-Gaussian kriging: accounting for an unknown mean value. Comput Geosci 34(11):1431–1442

Emery X (2010) Iterative algorithms for fitting a linear model of coregionalization. Comput Geosci 36(9):1150–1160

Emery X, Cornejo J (2010) Truncated Gaussian simulation of discrete-valued, ordinal coregionalized variables. Comput Geosci 36(10):1325–1338

Emery X, Lantuéjoul C (2006) TBSIM: a computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. Comput Geosci 32(10):1615–1628

Emery X, Lantuéjoul C (2008) A spectral approach to simulating intrinsic random fields with power and spline generalized covariances. Comput Geosci 12(1):121–132

Emery X, Arroyo D, Peláez M (2014) Simulating large Gaussian vectors subject to inequality constraints by Gibbs sampling. Math Geosci 46:265–283

Galli A, Beucher H, Le Loc'h G, Doligez B (1994) The pros and cons of the truncated Gaussian method. In: Armstrong M, Dowd PA (eds) Geostatistical simulations. Kluwer, Dordrecht, pp 217–233

Geman S, Geman D (1984) Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Trans Pattern Anal Mach Intell 6:721–741

Guilbert JM, Park CF (1986) The geology of ore deposits. Freeman, New York

Langlais V, Beucher H, Renard D (2008) In the shade of truncated Gaussian simulation. In: Ortiz JM, Emery X (eds) Proceedings of the eighth international geostatistics congress. Gecamin Ltda, Santiago, pp 799–808

Lantuéjoul C (2002) Geostatistical simulation, Models and Algorithms. Springer, New York

Lantuéjoul C, Desassis N (2012) Simulation of a Gaussian random vector: a propagative version of the Gibbs sampler. In: Ninth international geostatistics congress, Oslo

Le Loc'h G, Galli A (1997) Truncated plurigaussian method: theoretical and practical points of view. In: Baafi EY, Schofield NA (eds) Geostatistics Wollongong'96. Kluwer Academic, Dordrecht, pp 211–222

Lowell JD, Guilbert JM (1970) Lateral and vertical alteration-mineralization zoning in porphyry ore deposits. Econ Geol 65:373–408

Madani N, Emery X (2015) Simulation of geo-domains accounting for chronology and contact relationships: application to the Río Blanco copper deposit. Stoch Environ Res Risk Assess 29:2173–2191

Matheron G (1971) The theory of regionalized variables and its applications. Ecole Nationale Supérieure des Mines de Paris, Fontainebleau, p 212

Matheron G (1973) The intrinsic random functions and their applications. Adv Appl Probab 5(3):439–468

Ravenne C, Galli A, Doligez B, Beucher H, Eschard R (2002) Quantification of facies relationships via proportion curves. In: Armstrong M, Bettini C, Champigny N, Galli A, Remacre A (eds) Geostatistics Rio 2000. Kluwer Academic, Dordrecht, pp 19–40

Serrano L, Vargas R, Stambuk V, Aguilar C, Galeb M, Holmgren C, Contreras A, Godoy S, Vela I, Skewes MA, Stern CR (1996) The late Miocene to early Pliocene Río Blanco-Los Bronces copper deposit, Central Chilean Andes. In: Camus F, Sillitoe RH, Petersen R (eds) Andean copper deposits: new discoveries, mineralization, styles and metallogeny. Special publication No. 5. Society of Economic Geologists, Littleton, p 119

Skewes MA, Stern CR (1995) Genesis of the late Miocene to Pliocene copper deposits of central Chile in the context of Andean magmatic and tectonic evolution. Int Geol Rev 37(10):893–909

Stambuk V, Aguilar C, Blondel J, Galeb M, Serrano L, Vargas R (1988) Geología del yacimiento Río Blanco. Technical Report. Codelco-Chile, División Andina

Stein ML (2002) Fast and exact simulation of fractional Brownian surfaces. J Comput Graph Stat 11(3):587–599

Xu C, Dowd PA, Mardia KV, Fowell RJ (2006) A flexible true plurigaussian code for spatial facies simulations. Comput Geosci 32(10):1629–1645