Simulation of geological domains using the plurigaussian model: New developments and computer programs [☆]

Xavier Emery*

Department of Mining Engineering, University of Chile, Avenida Tupper 2069, Santiago 837 0451, Chile

Abstract

The plurigaussian model is currently used for simulating geological domains (facies) in petroleum reservoirs and mineral deposits, with the aim of assessing the uncertainty in the domain boundaries and of improving the geological controls in the characterization of quantitative attributes.

This paper discusses the main aspects of the model and provides a set of computer programs to perform its inference and conditional simulation. Two types of conditioning information are allowed: hard data for which one has an exact knowledge of the actual domain at sample locations, and soft data consisting of inequality constraints on the local domain proportions (probabilities of occurrence) at control points chosen by the mining or reservoir geologist.

An application to a Chilean porphyry copper deposit is finally presented, in which three Gaussian fields are used to simulate the spatial distribution of five mineralogical domains: gravels, leached capping, oxides, primary and secondary sulfides. The model is constructed so as to honor the topological contacts between mineralogical domains, their spatial continuity, the information logged at exploration drill holes, as well as the vertical proportion curves that indicate the mineralization profile with depth.

Keywords: Geostatistics; Stochastic simulation; Gaussian random fields; Random sets; Lithofacies; Categorical variables

1. Introduction

The geostatistical simulation of random sets and categorical variables is used for modeling the spatial distribution of geological domains, e.g. lithofacies in an oil reservoir or a sedimentary formation, hydrofacies in an aquifer, mineralogical domains in an ore body or soil types in a landscape. It provides a probabilistic description of the geological domains and contributes to enhanced geological control for the quantitative petrophysical variables of interest (porosity, permeability, concentration, mineral grade, etc.), when these variables are homogeneous within each geological domain but the layout of the domain boundaries is uncertain (Dowd, 1994; Dubrule, 1993; Emery and González, 2007).

Two main families of geostatistical methods can be used for categorical simulation. The first one corresponds to Boolean and object-oriented models (Lantuéjoul, 2002; Stoyan et al., 1996), the second one to pixel-based methods. Among the latter,

sequential indicator simulation provides a flexible framework, but realizations often lack realism from a geological viewpoint as the method cannot incorporate topological constraints on the simulated domains. Truncated Gaussian and plurigaussian models are an alternative to the indicator approach and have been applied for more than a decade to the characterization of oil reservoirs (Armstrong et al., 2003; Beucher et al., 1993) and ore deposits (Betzhold and Roth, 2000; Skvortsova et al., 2001, 2002).

This work aims at presenting new developments on the plurigaussian model and on the conditioning process, and at providing a set of computer programs that include the following features:

- no restriction on the number of simulated facies;
- no restriction on the number of truncated Gaussian random fields;
- truncation rule associated with any partition of the space spanned by the Gaussian random fields into cuboids;
- availability of the most common variogram models, with any number of nested structures and possible geometric or zonal anisotropies;
- possibility to handle soft conditioning data quantifying known geological features;
- ability to deal with large simulation domains; and
- improvement in the construction of conditioning Gaussian data with respect to existing softwares.

For the sake of simplicity, the plurigaussian model considered in this work and the related computer programs are limited to a set of independent Gaussian random fields. The use of dependent random fields can further enrich the method (Dowd et al., 2003; Le Loc'h and Galli, 1997), but increases the required efforts for parameter inference.

2. The plurigaussian model

2.1. Principles

Consider a set of independent standard Gaussian random fields $\{Y_1,\ldots,Y_m\}$ in R^3 with correlograms $\rho_1,\ldots,\rho_m$, respectively, and define a vectorial random field Y such that

$$\forall \mathbf{x} \in \mathsf{R}^3, \quad Y(\mathbf{x}) = (Y_1(\mathbf{x}),\ldots,Y_m(\mathbf{x})). \quad (1)$$

Let $(D_1,\ldots,D_n)$ be a partition of R^m into n disjoint subdomains. A categorical random field with n

categories (*facies*) is obtained by putting

$$\forall \mathbf{x} \in \mathsf{R}^3, \quad I(\mathbf{x}) = i \text{ if and only if } Y(\mathbf{x}) \in D_i. \quad (2)$$

2.2. Truncation rule

The choice of the partition $(D_1,\ldots,D_n)$ has implications on the spatial relationships between the facies, as it defines the permissible and forbidden contacts between pairs of facies and also allows reproducing the chronological ordering of the facies (Armstrong et al., 2003, p. 50; Lantuéjoul, 2002, p. 212; Le Loc'h et al., 1994).

Henceforth, we will assume that the subdomains $(D_1,\ldots,D_n)$ are cuboids of R^m or unions of cuboids. This situation is still quite general, as we do not make any restriction on the number of Gaussian fields (m), the number of facies (n) or the configuration of the cuboids in R^m. Fig. 1 gives two examples of plurigaussian realizations obtained by using two Gaussian fields with cubic correlograms; the partition (truncation rule) is symbolized by a flag below each realization.

2.3. Variogram inference

For any separation vector $\mathbf{h} \in \mathsf{R}^3$, the indicator (simple or cross) variogram between facies no. i and facies no. j is derived from the corresponding non-centered covariance

$$\gamma_{ij}(\mathbf{h}) = C_{ij}(\mathbf{0}) - \frac{C_{ij}(\mathbf{h}) + C_{ji}(\mathbf{h})}{2}, \quad (3)$$

with

$$C_{ij}(\mathbf{h}) = \mathrm{Prob}(Y(\mathbf{x}) \in D_i, Y(\mathbf{x}+\mathbf{h}) \in D_j). \quad (4)$$

Because the components of Y are mutually independent, the previous probability is a product or a sum of products of the following form:

$$\prod_{k=1}^m \mathrm{Prob}(Y_k(\mathbf{x}) \in [a_i,b_i], Y_k(\mathbf{x}+\mathbf{h}) \in [a_j,b_j])$$
$$= \prod_{k=1}^m \int_{a_i}^{b_i} \int_{a_j}^{b_j} g_{\rho_k(\mathbf{h})}(u,v)\,\mathrm{d}u\,\mathrm{d}v, \quad (5)$$

where $g_r(.,.)$ is the probability density function of a standard bivariate Gaussian pair with correlation coefficient r. The integral is usually calculated by numerical integration (e.g. Dowd et al., 2003). Here, following Matheron et al. (1988), we prefer to use expansions into the normalized Hermite polynomials
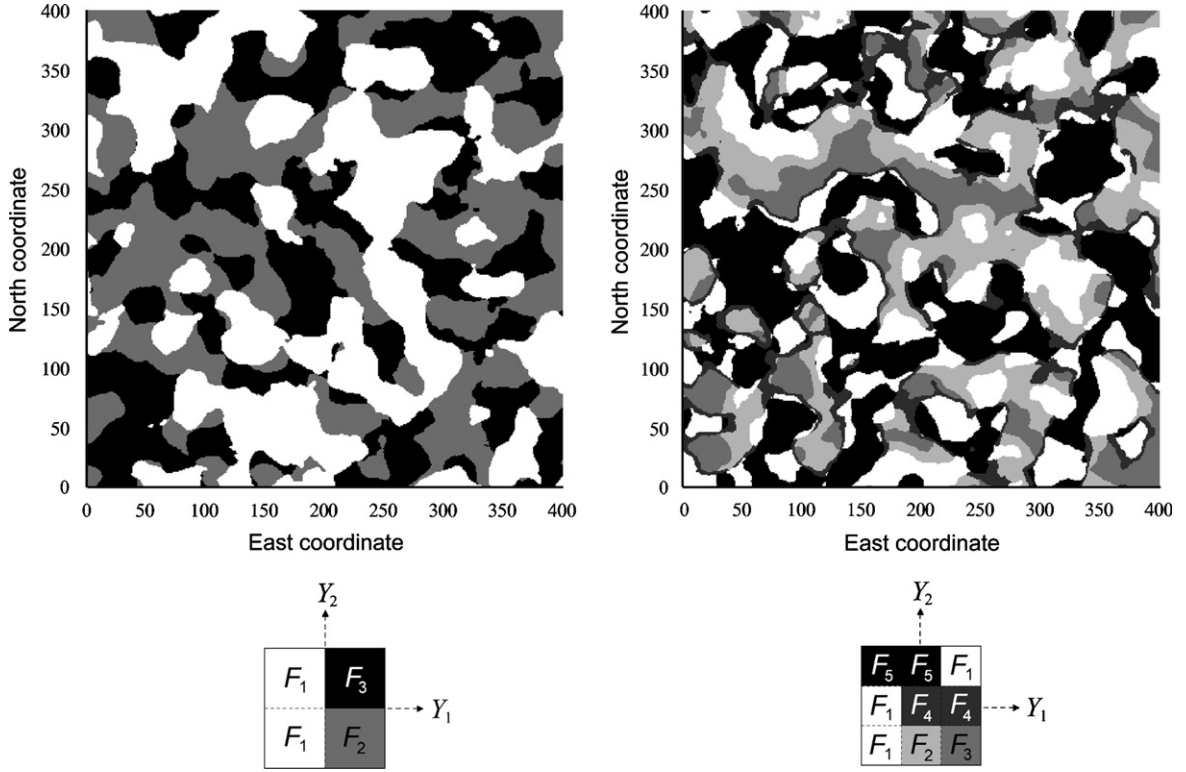
Fig. 1. Examples of plurigaussian realizations obtained by truncating two independent Gaussian fields with cubic correlograms. The truncation rule is represented by a flag below the realization. Each facies (F_1–F_3 or F_1–F_5) is associated with a specific gray tone.

$\{H_p, p \in \mathsf{N}\}$, which are orthonormal functions for the Gaussian distribution. Let $g(.)$ and $G(.)$ be the standard Gaussian probability density function and cumulative density function, respectively. Then one has

$$\int_{a_i}^{b_i} \int_{a_j}^{b_j} g_r(u, v)\, \mathrm{d}u\, \mathrm{d}v = \sum_{p=0}^{+\infty} r^p \varphi_p(a_i, b_i)\varphi_p(a_j, b_j), \qquad (6)$$

where $\{\varphi_p, p \in \mathsf{N}\}$ are the coefficients of the Hermitian expansion of the indicator of the interval $[a,b]$ (Chilès and Delfiner, 1999, p. 399):

$$\begin{cases} \varphi_0(a, b) = G(b) - G(a), \\ \forall p \in \mathsf{N}^*, \ \ \varphi_p(a, b) = \frac{H_{p-1}(b)g(b) - H_{p-1}(a)g(a)}{\sqrt{p}}. \end{cases} \qquad (7)$$

For practical calculations, the expansion in Eq. (6) has to be truncated at a high order, e.g. 500 or 1000. Relations (3) to (5) establish a link between the Gaussian field correlograms and the indicator variograms, which are accessible experimentally from the facies data. The former can

therefore be determined according to the fitting of the latter (Le Loc'h and Galli, 1997).

3. Conditioning the realizations to hard and soft data

The conditioning process of plurigaussian simulations is performed in three steps. First, the available conditioning information (e.g. facies data at sample locations) is converted into a set of Gaussian data that reproduce the spatial continuity of the vectorial random field Y. Second, the Gaussian data are used to construct conditional realizations of Y. Finally, the categorical realizations are obtained by applying the truncation rule (Eq. (2)). Since many algorithms are available to simulate Gaussian random fields (e.g. Chilès and Delfiner, 1999, pp. 462–504) and the last step is straightforward, only the first step needs to be examined.

3.1. Conditioning to hard data

Hard data consist of an exact knowledge of the prevailing facies at given locations of the deposit or reservoir, codified as categorical data indicating the

facies number at such locations. To transform these data into Gaussian ones, Dowd et al. (2003) proposed an acceptance–rejection method embedded in a sequential algorithm, in which the Gaussian values are generated at each data location conditionally to the categorical information at this location and to the previously simulated Gaussian values. However, such a method fails at accounting for all the original categorical information and is only approximate. In particular, the Gaussian values at the first data location are simulated so as to be consistent with the categorical information at this location, but the information brought by the neighboring data is ignored.

To account for the full amount of conditioning information, Freulon and de Fouquet (1993) and Freulon (1994) designed an iterative algorithm to simulate a truncated Gaussian vector, based on the Gibbs sampler (Geman and Geman, 1984). The extension to the plurigaussian case is a generalization of this algorithm and proceeds as follows.

3.1.1. Initialization

Let $\{\mathbf{x}_\alpha, \alpha \in A\}$ be the set of hard data locations and $\{i_\alpha, \alpha \in A\}$ the facies numbers at these locations. For $\alpha \in A$, simulate an independent random vector $\mathbf{Y}(\mathbf{x}_\alpha)$ that belongs to D_{i_α} (Eq. (2)), e.g. by using an acceptance–rejection method.

3.1.2. Iterative procedure

(a) Select an index δ at random, uniformly in A. Define $A' = A - \{\delta\}$.
(b) Calculate the distribution of $\mathbf{Y}(\mathbf{x}_\delta)$ conditionally to the $\mathbf{Y}$-values at the other locations $\{\mathbf{Y}(\mathbf{x}_\alpha), \alpha \in A'\}$. Since the components of $\mathbf{Y}$ are independent Gaussian random fields, the conditional distribution of $Y_k(\mathbf{x}_\delta)$ (with $k = 1,\dots,m$) is Gaussian, with mean equal to its simple kriging estimate from $\{Y_k(\mathbf{x}_\alpha), \alpha \in A'\}$, and variance equal to its simple kriging variance. If a unique neighborhood is used, the kriging weights and variances for all the data locations can be determined simultaneously from the inverse of the Gaussian data covariance matrix (Dubrule, 1983).
(c) Simulate a new random vector $\mathbf{Y}'(\mathbf{x}_\delta)$ with the conditional distribution of $\mathbf{Y}(\mathbf{x}_\delta)$.
(d) If $\mathbf{Y}'(\mathbf{x}_\delta)$ belongs to D_{i_δ}, substitute $\mathbf{Y}'(\mathbf{x}_\delta)$ for $\mathbf{Y}(\mathbf{x}_\delta)$.
(e) Go back to step (a) and loop many times.

The rate of convergence of this iterative algorithm depends on the correlation between the Gaussian data and is still unknown to a great extent. Some theoretical and experimental results on this subject can be found in Armstrong et al. (2003, pp. 82–105), Chan (1993), Galli and Gao (2001) and Roberts and Sahu (1997). In practice, to ensure convergence, it is recommended that each datum is updated several tens or hundreds of times, unless the Gaussian field correlograms have an important nugget effect (in case of a pure nugget effect, the Gibbs sampler would converge after a single iteration).

3.2. Conditioning to soft data

A second kind of conditioning information considered in this work corresponds to qualitative or "soft" data that express the background knowledge of the geologist about the deposit/reservoir under study. This information is useful to preclude that given facies are simulated in regions of space where hard data are sparse, but where the geologist is almost sure that such facies are absent. To incorporate this geological knowledge into the simulation process, the idea is to split each Gaussian field $\{Y_1,\dots, Y_m\}$ into two independent Gaussian components with the same correlogram and half the variance, and to translate the soft information into a constraint on the first component. More formally, let us write

$$\forall k \in \{1,\dots,m\}, Y_k = \frac{U_k + V_k}{\sqrt{2}}, \tag{8}$$

where U_k and V_k are two independent standard Gaussian random fields with correlogram ρ_k. In terms of vectorial random field:

$$\mathbf{Y} = \frac{\mathbf{U} + \mathbf{V}}{\sqrt{2}}. \tag{9}$$

Soft information consists of inequality constraints on the local prior (pre-posterior) probabilities of occurrence of given facies, for instance

- "There is more than 90% chance of finding facies no. 1 at location $\mathbf{x}$".
- "There is less than 10% chance of finding facies no. 1 or facies no. 2 at location $\mathbf{x}$".
- "Facies no. 1 has between 40% and 60% chance of occurrence at location $\mathbf{x}$". This constraint may correspond to a tolerance interval around an estimate of the facies proportion at the elevation under consideration, given by the vertical proportion curves (Armstrong et al., 2003, p. 27).

This information can be codified by one or several inequalities of the form

$$\sum_{i=1}^{n} \omega_i(\mathbf{x})\text{Prob}(\boldsymbol{Y}(\mathbf{x}) \in D_i) < \eta(\mathbf{x}), \qquad (10)$$

with numerical coefficients $\{\omega_1(\mathbf{x}),\ldots,\omega_n(\mathbf{x}),\eta(\mathbf{x})\}$ specified by the geologist according to its understanding or its intuition on the deposit/reservoir, or according to statistical tools such as vertical proportion curves or 3D proportion matrices (Ravenne et al., 2002).

Here the Gibbs sampler consists of simulating the random field $\boldsymbol{U}$ at the hard and soft data locations, and the random field $\boldsymbol{V}$ at the hard data locations only, such that:

- at each hard data location, the pair $(\boldsymbol{U},\boldsymbol{V})$ is consistent with the hard information;
- at each soft data location, vector $\boldsymbol{U}$ is consistent with the soft information.

A description of the simulation algorithm is given in Appendix A.

4. Program description

4.1. Main program: PLURISIM.M

The proposed programs are ASCII files written in MATLAB language (version 5.0 or later). The main routine (**PLURISIM.M**) allows drawing plurigaussian realizations over a regular grid or over scattered locations in R^3, conditional to hard and/or soft data. The realizations are obtained by truncating any number of Gaussian fields, the simulation of which uses the turning bands method and is based on a program developed by Emery and Lantuéjoul (2006).

The input parameters of PLURISIM.M are the following ones:

- *simucoord*: coordinates of the locations to simulate in R^3 ($m \times 3$ matrix; void if a grid simulation is sought after);
- *x0*, *y0* and *z0*: minimum grid coordinates along x, y and z directions;
- *nx*, *ny* and *nz*: numbers of grid nodes along x, y and z directions;
- *dx*, *dy* and *dz*: grid meshes along x, y and z directions;
- *nfield*: number of Gaussian random fields to truncate

- *nthres*: number of thresholds for each Gaussian field ($1 \times nfield$ vector);
- *thresholds*: threshold values (first Gaussian field first, then second field, and so on);
- *flag*: matrix codifying the truncation rule and containing the facies numbers. The flag representing the truncation rule is divided into prod($nthres + 1$) cuboids and the matrix entries (facies numbers in the cuboids) are ordered by cycling along Y_1, then along Y_2, and so on. For instance, the flags in Fig. 1 would be entered as [1 2 1 3] and [1 2 3 1 4 4 5 5 1], respectively;
- *nst*: number of nested structures for each Gaussian field ($1 \times nfield$ vector);
- *model*: covariance model for the Gaussian random fields (sum(nst) $\times$ 7 matrix). Each row corresponds to a nested structure and is codified as [type, scale factors, angles]. There are three scale factors (along the rotated y, x and z directions, respectively) and three angles to define the coordinate rotation (azimuth, dip and plunge) (Deutsch and Journel, 1998, p. 27). The available types are described in the program file;
- *cc*: sills of the nested structures (sum(nst) $\times$ 1 vector);
- *b*: third parameters, used for some of the covariance types (sum(nst) $\times$ 1 vector);
- *nlines*: number of lines to use for simulating each nested structure with the turning bands method (sum(nst) $\times$ 1 vector);
- *nugget*: nugget effect variances ($1 \times nfield$ vector);
- *harddatacoord*: hard data coordinates (void for non-conditional simulations);
- *harddata*: categorical data (facies numbers) (column vector);
- *softdatacoord*: soft data coordinates (void if soft data are not considered; $nsoftdata \times 3$ matrix otherwise);
- *softdata*: coefficients $\{\omega_1,\ldots,\omega_n,\eta\}$ of the inequality constraints (Eq. (10)) ($nsoftdata \times (n+1) \times nconstraints$ matrix);
- *nrealiz*: number of realizations to draw;
- *seed*: seed number for generating random values;
- *radius*: maximum search radii along y, x and z (rotated system) for the conditioning data; if the radius along y is set to infinity, a unique neighborhood is assumed for the Gibbs sampler and for the conditioning kriging step;
- *angles*: angles to define the rotated system (azimuth, dip and plunge), according to GSLIB conventions (Deutsch and Journel, 1998, p. 27);

- *octant*: divide the neighborhood into octants? 1 = yes, 0 = no;
- *ndata*: number of conditioning data per octant (if octant = 1) or in total (if octant = 0);
- *name*: name of output file;
- *header*: create a GSLIB header in the output file? 1 = yes, 0 = no;
- *ntok*: maximum number of locations/nodes to simulate simultaneously. This number defines how many locations are projected onto the lines at each step of the turning bands simulation; and
- *niteration*: number of iterations to consider for the Gibbs sampler.

Alternatively, program PLURISIM.M can be used with a parameter file: no input argument is then required in the Matlab workspace and the user is prompted for the name of the parameter file. If no name is entered, a default file (*plurisim.par*) is assumed. In case the file does not exist, a blank file *plurisim.par* is created. When resorting to the parameter file mode, the coordinates of the locations to simulate (if these locations are not gridded) the data coordinates and data values must be stored in external ASCII files without headers, so as to be loadable in the Matlab workspace.

PLURISIM.M creates an external ASCII file with the simulated categorical values (one realization per column). For grid simulations, the ordering of the grid nodes is the same as in GSLIB (Deutsch and Journel, 1998): point by point to the east, then row by row to the north, and finally level by level upward.

4.2. Auxiliary programs for parameter inference

Three other programs are provided to help for parameter inference:

- CALCULATE_THRESHOLDS.M: calculate the threshold values associated with a given flag (truncation rule) and a set of facies proportions. This program examines successively each Gaussian random field and attempts to determine the thresholds associated with this field by grouping the facies, following the procedure described by Armstrong et al. (2003, p. 54). Once a threshold is determined, the initial flag is split into two sub-flags and the program is used recursively. If the calculation fails (non-unique solution or too complex flag), the output thresholds are set to minus infinity. In such a case, the thresholds can

be determined via trial-and-error or iterative optimization methods (not implemented in this work), in order to reproduce the facies proportions.
- PROPCURVES.M: calculate smoothed proportion curves along a given direction, from a set of categorical data. The data coordinates are discretized and assigned to given levels, then the empirical domain proportions are calculated at each level. The user can introduce a smoothing parameter corresponding to half the size of a moving window (expressed in number of levels) over which the raw proportion curves are averaged.
- VMODEL_IND.M: calculate the indicator simple and cross variograms associated with a truncation rule, facies proportions and Gaussian variograms. The calculation is based on expansions of the indicator function into Hermite polynomials (Eqs. (3)–(7)). A graphical comparison between the resulting indicator variograms and the empirical ones (derived from the available hard data) allows determining the variogram models of the Gaussian random fields.

5. A case study in mining engineering

5.1. Presentation of data

A case study is now proposed, dealing with the simulation of mineralogical domains in a porphyry copper deposit (*Radomiro Tomic*) located in northern Chile, at 40 km from Calama city and at an average elevation of approximately 3000 m above mean sea level. The ore body is approximately 5 km long, 1 km large and 0.4 km deep.

The mineralization profile can be divided into five main domains (Cuadra and Rojas, 2001): alluvial gravels, leached capping, oxidized copper mineralization, primary sulfide mineralization and secondary sulfide enrichment. The deposit has been recognized by diamond drill holes, along which the mineralogy has been logged. A map showing the distribution of the data over a cross-section is presented in Fig. 2. One observes that the direction of greatest geological continuity is almost horizontal (it is dipping about $0.7°$ downwards with respect to the north direction).

5.2. Truncation rule and threshold determination

The truncation rule must be defined so as to honor the topological contacts between mineralogical
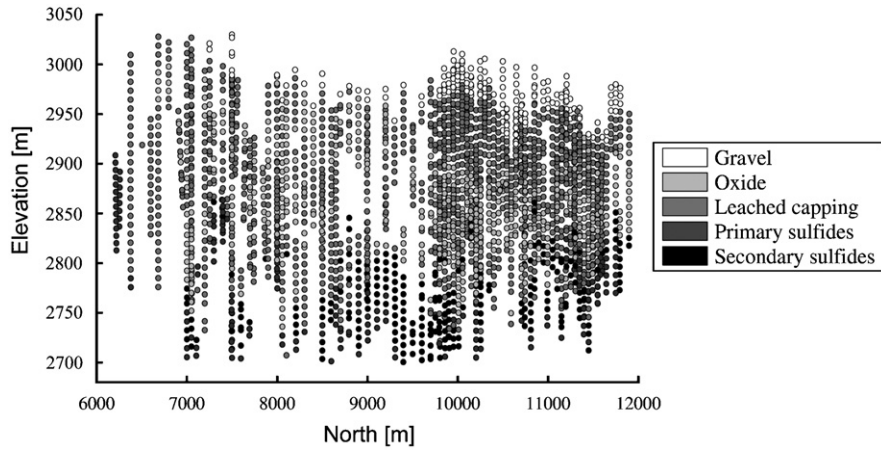
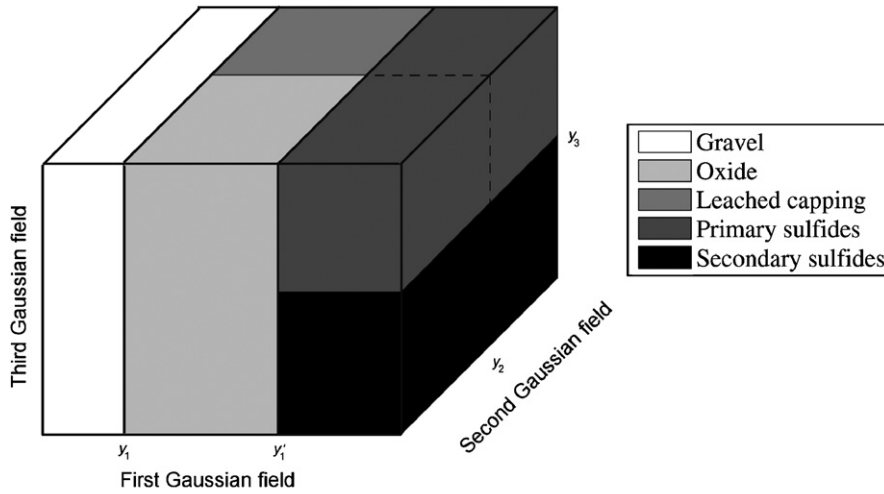Fig. 2. Distribution of available data over cross-section with east coordinate 4300 m.

Fig. 3. Three-dimensional flag indicating truncation rule used for converting Gaussian values into categorical values.

domains. In the present case, gravels are located near the surface and are in contact with the oxide and leach domains but never with sulfides, located beneath the phreatic level at an elevation of 2850 m. In contrast, the other four domains (leached capping, oxides, primary and secondary sulfides) are in contact with each other. This situation motivates the use of three Gaussian fields $\{Y_1, Y_2, Y_3\}$ and the definition of the following truncation rule (Fig. 3):

● location $\mathbf{x}$ belongs to gravel

$$\Leftrightarrow Y_1(\mathbf{x}) < y_1,$$

● location $\mathbf{x}$ belongs to oxide

$$\Leftrightarrow \begin{cases} y_1 \leqslant Y_1(\mathbf{x}) < y_1', \\ \quad Y_2(\mathbf{x}) < y_2, \end{cases}$$

● location $\mathbf{x}$ belongs to leached capping

$$\Leftrightarrow \begin{cases} y_1 \leqslant Y_1(\mathbf{x}) < y_1', \\ y_2 \leqslant Y_2(\mathbf{x}), \end{cases}$$

● location $\mathbf{x}$ belongs to secondary sulfide

$$\Leftrightarrow \begin{cases} y_1' \leqslant Y_1(\mathbf{x}), \\ Y_3(\mathbf{x}) < y_3, \end{cases}$$

- location **x** belongs to primary sulfide

$$\Leftrightarrow \begin{cases} y_1' \leqslant Y_1(\mathbf{x}), \\ y_3 \leqslant Y_3(\mathbf{x}). \end{cases}$$

The thresholds $\{y_1, y_1', y_2, y_3\}$ are linked to the global proportions of the mineralogical domains. Instead of estimating these proportions directly from the available categorical data, we prefer to define them from the variogram sills of the domain indicators, in order to improve the consistency of the subsequent variogram models. Recall that, if p is the proportion of a spatial domain, then the sill of the corresponding indicator variogram is $p(1-p)$. The analysis of the sample indicator variograms leads to the following threshold values:

$y_1 = -1.20, \; y_1' = 1.15, \; y_2 = 0.17 \; \text{and} \; y_3 = 0.22.$

5.3. Variogram analysis

We now turn to the modeling of the Gaussian field variograms. As mentioned in Section 2.3, these are determined through their impact on the variograms of the domain indicators. For practical calculations, only the sub-horizontal ($-0.7°$ dip) and sub-vertical ($89.3°$ dip) directions, corresponding to the main anisotropy directions, are considered. A trial-and-error procedure based on program VMODEL_IND.M is adopted. The following Gaussian variogram models provide a satisfactory fitting of the indicator simple variograms (Fig. 4) (for the sake of simplicity, the indicator cross-variograms are not examined):

- Variogram model for the first Gaussian field:

$\gamma_1 = 0.04$ cubic $(100\,\text{m}; 100\,\text{m})$
  $+ 0.06$ cubic $(250\,\text{m}; 200\,\text{m})$
  $+ 0.90$ cubic $(2500\,\text{m}; 200\,\text{m}).$

- Variogram model for the second Gaussian field:

$\gamma_2 = 0.40$ spherical $(100\,\text{m}; 100\,\text{m})$
  $+ 0.30$ spherical $(2000\,\text{m}; 120\,\text{m})$
  $+ 0.15$ spherical $(100\,\text{m}; \infty)$
  $+ 0.15$ spherical $(2000\,\text{m}; \infty).$

- Variogram model for the third Gaussian field:

$\gamma_3 = 0.40$ cubic $(200\,\text{m}; 200\,\text{m})$
  $+ 0.60$ cubic $(25{,}000\,\text{m}; 10{,}000\,\text{m}).$

In the previous equations, the distances in brackets represent the ranges along the sub-horizontal and sub-vertical directions. The shape of the variograms near the origin is linked to the regularity of the boundaries between mineralogical domains (Lantuéjoul, 2002, p. 26). In particular, the cubic variogram model is smooth at the origin and is associated with a regular boundary. In contrast, the spherical model (used for the second Gaussian field to separate oxides from leached capping) increases linearly and entails a more erratic boundary between these two domains.

5.4. Definition of soft conditioning constraints

To be realistic from a geological viewpoint, the plurigaussian realizations must honor the mineralogical profile with depth: gravels should be found near the surface, whereas sulfides should appear below the phreatic level. Now, these conditions are not necessarily fulfilled if one only uses the drill hole data as conditioning information. The reason is that the realizations are almost unconstrained in under-sampled areas, for which the drill hole data are scarce.

To account for the geological information, ancillary data are defined over a regular grid covering the area to simulate. These data consist of interval constraints on the local domain proportions; the interval midpoints correspond to the proportion estimates given by the vertical proportion curves (Fig. 5), whereas the interval widths are set to 20%. For instance, if the curves indicate an oxide proportion of 30% at a given elevation, then the actual oxide proportion is constrained between 20% and 40%.

5.5. Remark on the use of a stationary prior model

The vertical proportion curves (Fig. 5) indicate the presence of trends along the vertical direction. These trends can be interpreted as (i) an effect of non-stationarity (presence of prior *drifts*), (ii) a result of randomness or chance fluctuations (Chilès and Delfiner, 1999, p. 233) or (iii) an effect of conditioning data.

The first interpretation (drift models) is associated with the definition of regionalized facies proportions and spatially varying truncation thresholds, which raises difficulties for variogram analysis (Armstrong et al., 2003; Beucher et al., 1993). In general, the second interpretation (chance fluctuations) does not allow a systematic reproduction of the trends, which
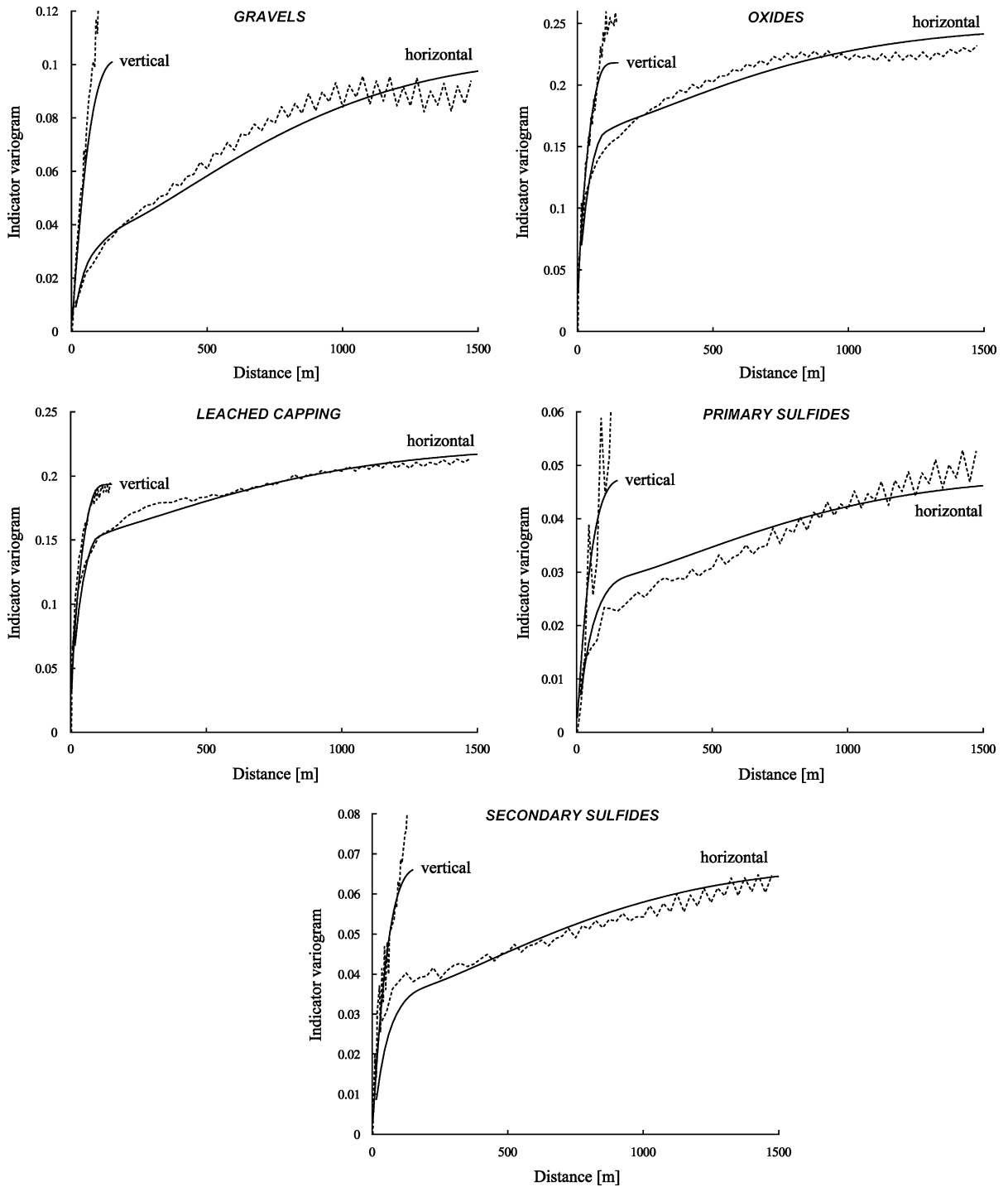
Fig. 4. Sample (dashed lines) and modeled (solid lines) indicator variograms along sub-horizontal and sub-vertical directions.

appear only in some realizations. The third interpretation assumes that trends are implicitly defined by the presence of conditioning data: the simulated random fields are forced to honor the conditioning constraints, which alter their stationary prior distributions into non-stationary posterior (conditional) distributions and produce patterns such as local drifts (Journel and Huijbregts, 1978, p. 501).
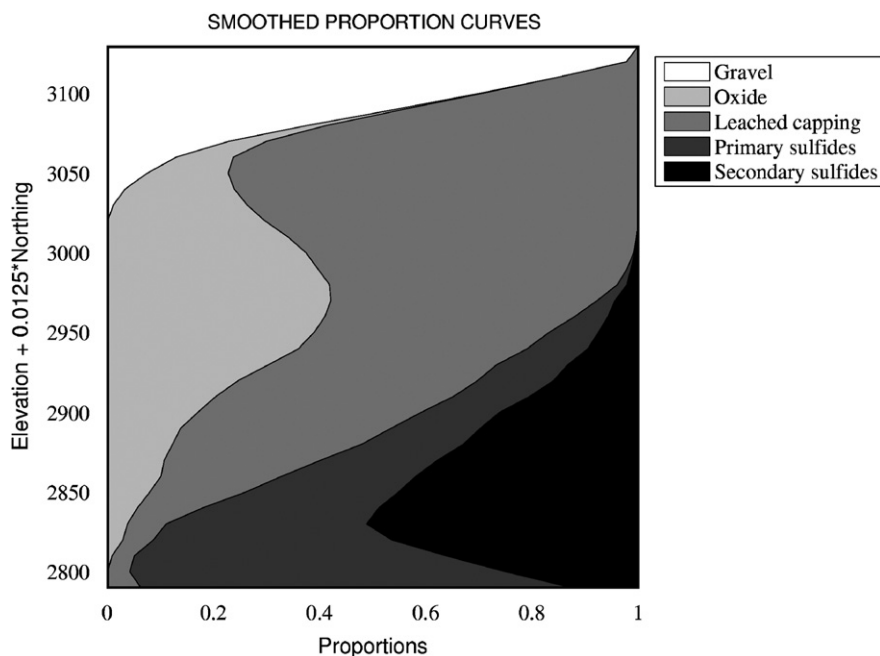
SMOOTHED PROPORTION CURVES

Fig. 5. Smoothed proportion curves, giving estimates of domain proportions along the direction of maximum geological variability (sub-vertical).

Prior drift models actually have little impact when conditioning data abound and are relevant only in extrapolation situations or when conditioning data are scarce (Journel and Rossi, 1989).

In the present case study, soft data locations have been placed in order to achieve a regular coverage of the area of interest, and the soft constraints at these locations have been defined by reference to the vertical proportion curves. Such constraints are therefore expected to impose the reproduction of the trends along the vertical direction, without the need for a prior drift model. Working with a stationary prior model greatly facilitates the inference of model parameters, as the thresholds defining the truncation rule are constant in space and the Gaussian variograms only depend on the separation vector.

It is worthy to note that the use of inequality constraints (Eq. (10)) gives flexibility to the proposed methodology, as it allows accounting for an uncertainty in the local facies proportions (defined as the probabilities of occurrence of the facies at a given location). In contrast, with non-stationary models based on vertical or on regionalized proportion curves, the local facies proportions are considered as "hard" values with no uncertainty, although they have been estimated from a limited set of neighboring data and, therefore, are approximate. Accordingly, the proposed approach is likely to simulate a larger range of situations, with more varied facies proportions between realizations, and to better reflect the actual uncertainty in the geology of the deposit.

5.6. Conditional simulation

Once the model parameters (truncation rule, Gaussian thresholds and variograms) and conditioning data are ready, program PLURISIM.M can be used to create conditional realizations of the mineralogical domains in the deposit. Two of them are displayed in Fig. 6 as an illustration. One can check that both realizations reproduce the drill hole data (Fig. 2), as well as the vertical proportion curves (Fig. 5), through the incorporation of the soft conditioning data (interval constraints).

All in all, the plurigaussian model is relatively simple and straightforward to put in practice, and is able to account for diverse sources of information: contacts between the domains (truncation rule), mineralogical profile with depth (vertical proportion curves, converted into soft data constraints), geological continuity (indicator variograms) and drill hole information (conditioning hard data). The application of this model requires a close collaboration between the geostatistician and the
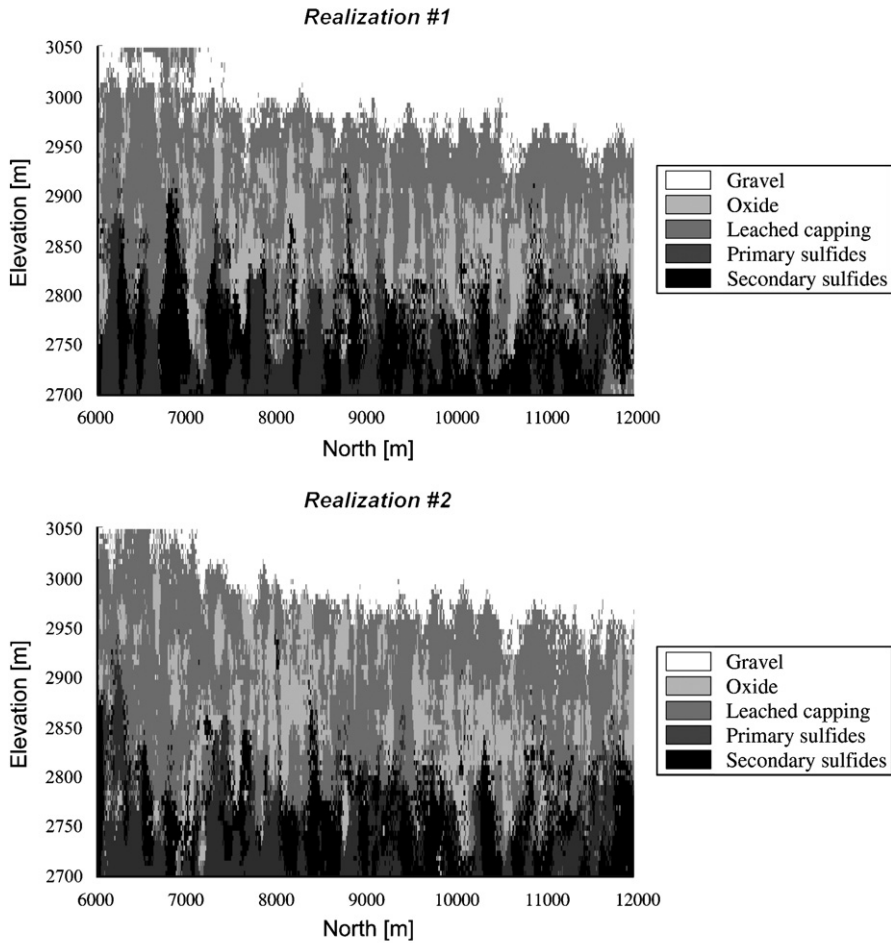
Realization #1

Realization #2

Fig. 6. Two conditional realizations of mineralization domains. Representation of cross-section with east coordinate 4300 m.

geologist, in particular when defining the truncation rule and the soft conditioning data.

6. Conclusions

The plurigaussian model is a promising approach for simulating categorical variables representing spatial domains (facies) and for assessing the uncertainty in their boundaries. It is simple to implement, mathematically consistent and data-charged, as the realizations honor the topological constraints (contacts and ordering) between facies, their spatial continuity (indicator variograms) as well as prior geological knowledge, e.g. information of the vertical proportion curves for characterizing a mineralization profile with depth. All these features have been included in a set of computer programs,

which also allow a great flexibility in the choice of the number of Gaussian fields to truncate, the number of facies and the truncation rule. The presentation of the concepts has been accompanied by a case study, so as to familiarize practitioners with the steps involved in the modeling and simulation, and to assess the performance of the proposed programs.

Acknowledgments

Appendix A

This appendix gives the steps for plurigaussian simulation in the presence of hard and soft conditioning data. The simulation procedure is based on the decomposition of the vectorial random field Y into two independent components U and V (Eq. (9)).

A.1. Initialization

(a) Let $\{\mathbf{x}_\alpha, \alpha \in A\}$ be the set of hard data locations and $\{i_\alpha, \alpha \in A\}$ the facies numbers at these locations. For any α in A, simulate an independent random vector $Y(\mathbf{x}_\alpha)$ that belongs to D_{i_α} (Eq. (2)) and set

$$U(\mathbf{x}_\alpha) = V(\mathbf{x}_\alpha) = \frac{Y(\mathbf{x}_\alpha)}{\sqrt{2}}. \qquad (A.1)$$

(b) Let $\{\mathbf{x}_\beta, \beta \in B\}$ be the set of soft data locations and W a standard Gaussian random vector with m independent components. For any β in B, draw an independent random vector $U(\mathbf{x}_\beta)$ such that (Eq. (10))

$$\sum_{i=1}^n \omega_i(\mathbf{x}_\beta) \mathrm{Prob}\left(\frac{U(\mathbf{x}_\beta) + W}{\sqrt{2}} \in D_i \right) < \eta(\mathbf{x}_\beta). \qquad (A.2)$$

A.2. Simulation at the data locations (Gibbs sampler)

(c) Select an index δ at random, uniformly in $A \cup B$.
(d) If $\delta \in A$ (hard datum, associated with facies number i_δ)
  ○ Define $A' = A - \{\delta\}$.
  ○ Calculate the distributions of $U(\mathbf{x}_\delta)$ and $V(\mathbf{x}_\delta)$ conditionally to $\{U(\mathbf{x}_\varepsilon), \varepsilon \in A' \cup B\}$ and $\{V(\mathbf{x}_\alpha), \alpha \in A'\}$. For any k in $\{1,\ldots,m\}$, the conditional distributions of $U_k(\mathbf{x}_\delta)$ and $V_k(\mathbf{x}_\delta)$ are Gaussian, with means equal to simple kriging estimates and variances equal to simple kriging variances.
  ○ Simulate two random vectors $U'(\mathbf{x}_\delta)$ and $V'(\mathbf{x}_\delta)$, with the conditional distributions of $U(\mathbf{x}_\delta)$ and $V(\mathbf{x}_\delta)$, respectively.
  ○ Define $Y'(x_\delta) = \dfrac{U'(\mathbf{x}_\delta) + V'(\mathbf{x}_\delta)}{\sqrt{2}}$.
  ○ If $Y'(\mathbf{x}_\delta)$ belongs to D_{i_δ}, substitute $U'(\mathbf{x}_\delta)$ for $U(\mathbf{x}_\delta)$ and $V'(\mathbf{x}_\delta)$ for $V(\mathbf{x}_\delta)$.
(e) If $\delta \in B$ (soft datum)
  ○ Define $B' = B - \{\delta\}$.

  ○ Calculate the distribution of $U(\mathbf{x}_\delta)$ conditionally to $\{U(\mathbf{x}_\varepsilon), \varepsilon \in A \cup B'\}$. Again, each component $U_k(\mathbf{x}_\delta)$ has a Gaussian distribution, with mean and variance equal to its simple kriging estimate and simple kriging variance.
  ○ Simulate a new random vector $U'(\mathbf{x}_\delta)$ with the conditional distribution of $U(\mathbf{x}_\delta)$.
  ○ Let W be an independent standard Gaussian random vector with m independent components. For any i in $\{1,\ldots,n\}$, calculate the probability

$$P_i = \mathrm{Prob}\left(\frac{U'(\mathbf{x}_\delta) + W}{\sqrt{2}} \in D_i \right). \qquad (A.3)$$

  ○ If $\sum_{i=1}^n \omega_i(\mathbf{x}_\delta) P_i < \eta(\mathbf{x}_\delta)$ (Eq. (10)), substitute $U'(\mathbf{x}_\delta)$ for $U(\mathbf{x}_\delta)$.
(f) Go back to step (c) and loop many times.

At the end of this iterative procedure, one obtains vectors that honor the conditioning information: the hard information at $\mathbf{x}_\alpha$ ($\alpha \in A$) is accounted for by the pair $(U(\mathbf{x}_\alpha), V(\mathbf{x}_\alpha))$, while the soft information at $\mathbf{x}_\beta$ ($\beta \in B$) is accounted for by vector $U(\mathbf{x}_\beta)$.

A.3. Simulation at other locations in R^3

The remaining steps for conditional simulation of facies are

(g) Simulate the random field V at the soft data locations $\{\mathbf{x}_\beta, \beta \in B\}$, conditionally to $\{V(\mathbf{x}_\alpha), \alpha \in A\}$.
(h) Calculate Y at the hard and soft data locations (Eq. (9)).
(i) Simulate Y at other locations in R^3, conditionally to its values at the data locations.
(j) Derive the facies simulation by applying the truncation rule (Eq. (2)).

Appendix B. Supporting Information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.cageo.2007.01.006.

References

Armstrong, M., Galli, A., Le Loc'h, G., Geffroy, F., Eschard, R., 2003. Plurigaussian Simulations in Geosciences. Springer, Berlin, 160pp.

Betzhold, J., Roth, C., 2000. Characterizing the mineralogical variability of a Chilean copper deposit using plurigaussian simulations. Journal of the South African Institute of Mining and Metallurgy 100 (2), 111–120.

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

Chan, K.S., 1993. Asymptotic behavior of the Gibbs sampler. Journal of the American Statistical Association 88 (421), 320–326.

Chilès, J.P., Delfiner, P., 1999. Geostatistics: Modeling Spatial Uncertainty. Wiley, New York, 695pp.

Cuadra, P., Rojas, G., 2001. Oxide mineralization at the Radomiro Tomic copper deposit, Northern Chile. Economic Geology 96 (2), 387–400.

Deutsch, C.V., Journel, A.G., 1998. GSLIB: Geostatistical Software Library and User's Guide, second ed. Oxford University Press, New York, 369pp.

Dowd, P.A., 1994. Geological controls in the geostatistical simulation of hydrocarbon reservoirs. Arabian Journal for Science and Engineering 19 (2B), 237–247.

Dowd, P.A., Pardo-Igúzquiza, E., Xu, C., 2003. Plurigau: a computer program for simulating spatial facies using the truncated plurigaussian method. Computers & Geosciences 29 (2), 123–141.

Dubrule, O., 1983. Cross-validation of kriging in a unique neighbourhood. Mathematical Geology 15 (6), 687–699.

Dubrule, O., 1993. Introducing more geology in stochastic reservoir modelling. In: Soares, A. (Ed.), Geostatistics Tróia'92. Kluwer Academic, Dordrecht, pp. 351–369.

Emery, X., González, K.E., 2007. Incorporating the uncertainty in geological boundaries into mineral resources evaluation. Journal of the Geological Society of India 69 (1), 29–38.

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

Freulon, X., 1994. Conditional simulation of a Gaussian random vector with nonlinear and/or noisy observations. In: Armstrong, M., Dowd, P.A. (Eds.), Geostatistical Simulations. Kluwer Academic, Dordrecht, pp. 57–71.

Freulon, X., de Fouquet, C., 1993. Conditioning a Gaussian model with inequalities. In: Soares, A. (Ed.), Geostatistics Tróia'92. Kluwer Academic, Dordrecht, pp. 201–212.

Galli, A., Gao, H., 2001. Rate of convergence of the Gibbs sampler in the Gaussian case. Mathematical Geology 33 (6), 653–677.

Geman, S., Geman, D., 1984. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6 (6), 721–741.

Journel, A.G., Huijbregts, C.J., 1978. Mining Geostatistics. Academic Press, London, 600pp.

Journel, A.G., Rossi, M.E., 1989. When do we need a trend model in kriging? Mathematical Geology 21 (7), 715–739.

Lantuéjoul, C., 2002. Geostatistical Simulation, Models and Algorithms. Springer, Berlin, 256pp.

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

Le Loc'h, G., Beucher, H., Galli, A., Doligez, B., 1994. Improvement in the truncated Gaussian method: combining several Gaussian functions. In: ECMOR IV, Fourth European Conference on the Mathematics of Oil Recovery. Røros, Norway, 13pp. (unpublished conference proceedings).

Matheron, G., Beucher, H., de Fouquet, C., Galli, A., Ravenne, C., 1988. Simulation conditionnelle à trois faciès d'une falaise de la formation du Brent. Sciences de la Terre, Série Informatique Géologique 28, 213–249.

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.

Roberts, G.O., Sahu, S.K., 1997. Updating scheme, correlation structure, blocking and parametrization for the Gibbs sampler. Journal of the Royal Statistical Society B 59 (2), 291–317.

Skvortsova, T., Armstrong, M., Beucher, H., Forkes, J., Thwaites, A., Turner, R., 2001. Applying plurigaussian simulations to a granite-hosted orebody. In: Kleingeld, W.J., Krige, D.G. (Eds.), Geostats 2000 Cape Town. Geostatistical Association of Southern Africa, Johannesburg, pp. 904–911.

Skvortsova, T., Beucher, H., Armstrong, M., Forkes, J., Thwaites, A., Turner, R., 2002. Simulating the geometry of a granite-hosted uranium orebody. In: Armstrong, M., Bettini, C., Champigny, N., Galli, A., Remacre, A. (Eds.), Geostatistics Rio 2000. Kluwer Academic, Dordrecht, pp. 85–99.

Stoyan, D., Kendall, W.S., Mecke, J., 1996. Stochastic Geometry and Its Applications, second ed. Wiley, New York, 456pp.