Uncertainty modeling and spatial prediction by multi-Gaussian kriging: Accounting for an unknown mean value ☆

Xavier Emery *

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

Keywords:
Geostatistics
Multi-variate normal distribution
Gaussian random fields
Conditional expectation
Conditional distributions

A B S T R A C T

In the analysis of spatial data, one is often interested in modeling conditional probability distributions, in order to assess the uncertainty in the values of the attribute under study and to predict functions of this attribute.

This work examines three geostatistical models in which the attribute is assumed to be, up to a monotonic transformation, a realization of a Gaussian random field. In the first model, the mean of the Gaussian field is a known parameter and the conditional distributions at any set of locations are Gaussian, with expected values equal to simple kriging predictions and covariance matrix equal to that of the prediction errors. In the second model, the mean value is replaced by a random variable adding to the Gaussian field and whose prior variance is infinitely large, indicating a total lack of prior knowledge on the true mean. It is shown that the conditional distributions are still Gaussian, with expected values equal to ordinary kriging predictions and covariance matrix equal to that of the corresponding prediction errors. The third model considers a random drift that adds to the Gaussian field; the conditional distributions are then obtained by substituting universal kriging for simple or ordinary kriging.

A computer program is provided to calculate recovery functions (tonnages, metal contents and mean values above given thresholds) and uncertainty measures (probability intervals and conditional variances) defined at point or block supports. The concepts are illustrated with a case study consisting of evaluating the recoverable resources in a porphyry copper deposit.

1. Introduction

Many disciplines in the geosciences are concerned with the prediction of spatial attributes in relation to threshold values. For instance, environmentalists want to determine whether or not pollutant concentrations in the air, water or soil fall short of given regulatory levels. Pest managers are interested in mapping the probabilities that pest densities or pathogen genotypes exceed critical thresholds, in order to define plant disease control measures. Hydrologists intend to characterize the depth to water table in relation to a threshold depth so as to assess the risk of runoff generation during rainfall events. Soil scientists and land planners are concerned with excesses or deficiencies in soil properties. Mining engineers aim to delineate the volume where mineral grades are greater than an economic cutoff that makes mining profitable.

Because of its smoothing property, linear kriging is ill-suited to predict whether or not the values of an attribute exceed a threshold. Instead, one can use nonlinear kriging (indicator, disjunctive or multi-Gaussian kriging) to characterize the unknown values by conditional probability distributions. Indicator and disjunctive kriging

☆ Code available from server at http://www.iamg.org/CGEditor/index.htm.

* Tel.: +56 2 978 4498; fax: +56 2 978 4985.
E-mail address: xemery@ing.uchile.cl

make hypotheses on the bivariate distributions of the random field representing the attribute of interest. They may produce order relation violations between different thresholds and require corrections to obtain consistent results (Deutsch and Journel, 1998, p. 81; Rivoirard, 1994, p. 60). In contrast, multi-Gaussian kriging considers the full set of multivariate distributions and does not produce order relation problems. It has been widely used in geostatistics applied to recoverable resources assessment (Verly, 1983, 1984; Maréchal, 1984; Schofield, 1988). However, it suffers from several limitations that restrict its scope of application; in particular, it assumes that the mean value of normal scores data is perfectly known and it loses efficiency if this mean value is misspecified (Guibal and Remacre, 1984; Rivoirard, 1994, p. 67, 96).

This article focuses on the multi-Gaussian approach to the modeling of conditional probability distributions and prediction of recoverable resources above given thresholds. Its objective is to weaken the assumption of a known mean and to provide a theoretically sound framework to account for an uncertain mean. The concepts are complemented by a set of computer programs and illustrated with a case study in mineral resources evaluation.

2. Multi-Gaussian kriging with a known mean

2.1. Model assumptions

The attribute of interest is regarded as a realization of a random field over a spatial domain $\mathbf{D}$, say $Z=\{Z(\mathbf{x}),\ \mathbf{x}\in\mathbf{D}\}$, that can be transformed into another random field $Y=\{Y(\mathbf{x}),\ \mathbf{x}\in\mathbf{D}\}$ with Gaussian univariate distribution, by means of a non-decreasing transformation function ϕ:

$$\forall \mathbf{x}\in\mathbf{D},\quad Z(\mathbf{x})=\phi(Y(\mathbf{x})). \tag{1}$$

The following assumptions are made:

(1) Y is a Gaussian random field, i.e. its finite-dimensional distributions are multivariate Gaussian.

(2) Y is second-order stationary and ergodic, with mean 0, variance 1 and semi-variogram γ_Y. The covariance between two variables $Y(\mathbf{x})$ and $Y(\mathbf{x}')$ is equal to the prior variance minus the semi-variogram:

$$\forall \mathbf{x},\mathbf{x}'\in\mathbf{D},\quad C_Y(\mathbf{x},\mathbf{x}')=1-\gamma_Y(\mathbf{x},\mathbf{x}'). \tag{2}$$

(3) The values of Y are known at a set of sampling locations $\{\mathbf{x}_\alpha,\ \alpha=1,\ldots,n\}$:

$$\forall \alpha\in\{1,\ldots n\},\quad Y(\mathbf{x}_\alpha)=y_\alpha. \tag{3}$$

Because of the stationarity assumption, the covariance function and semi-variogram of Y only depend on the lag separation vector $(\mathbf{x}'-\mathbf{x})$, which facilitates their inference.

2.2. Conditional distribution of a vector of unknowns

Let $\boldsymbol{Y}=(Y(\mathbf{u}_1),\ldots,Y(\mathbf{u}_k))$ be a vector associated with a set of locations $\mathbf{u}_1,\ldots,\mathbf{u}_k$. Because of the multivariate Gaussian assumption, the posterior distribution of $\boldsymbol{Y}$ (i.e. its distribution conditioned to the data) is Gaussian, with its moments determined by the linear regression (simple kriging) of $\boldsymbol{Y}$ upon the data (Anderson, 2003, p. 34). Specifically, the conditional mean is the simple kriging vector $\boldsymbol{y}^{SK}$, whose ith component ($i=1,\ldots,k$) is

$$y(\mathbf{u}_i)^{SK}=\sum_{\alpha=1}^{n}\lambda_{\alpha,i}^{SK}y_\alpha, \tag{4}$$

with

$$\sum_{\beta=1}^{n}\lambda_{\beta,i}^{SK}C_Y(\mathbf{x}_\alpha,\mathbf{x}_\beta)=C_Y(\mathbf{x}_\alpha,\mathbf{u}_i)\quad\forall\,\alpha=1,\ldots,n. \tag{5}$$

The conditional variance–covariance matrix is that of the simple kriging errors, which will be denoted by $\boldsymbol{\Sigma}^{SK}$. The generic term of this matrix (covariance between the kriging errors at locations $\mathbf{u}_i$ and $\mathbf{u}_j$) is (Harter, 1994, p. 55; Anderson, 2003, p. 34):

$$\Sigma_{i,j}^{SK}=C_Y(\mathbf{u}_i,\mathbf{u}_j)-\sum_{\beta=1}^{n}\lambda_{\beta,j}^{SK}C_Y(\mathbf{u}_i,\mathbf{x}_\beta)\quad\forall i,j=1,\ldots,k. \tag{6}$$

Eq. (6) can be derived by expressing the kriging predictors in terms of weighted sums of the data and by using the kriging equations (Eqs. (4) and (5)). Because $\boldsymbol{\Sigma}^{SK}$ is a variance–covariance matrix, it is symmetric positive semi-definite and can be written as the product of a matrix by its transpose:

$$\boldsymbol{\Sigma}^{SK}=\mathbf{A}^{SK}(\mathbf{A}^{SK})^t. \tag{7}$$

For instance, $\mathbf{A}^{SK}$ can be the square root of $\boldsymbol{\Sigma}^{SK}$ or a triangular matrix obtained by Choleski factorization. The conditional Gaussian vector $\boldsymbol{Y}$ is therefore decomposed in the following fashion (Davis, 1987):

$$\boldsymbol{Y}=\boldsymbol{y}^{SK}+\mathbf{A}^{SK}\boldsymbol{T}, \tag{8}$$

where $\boldsymbol{T}$ is a standard Gaussian random vector independent of the data and with mutually independent components. In the next subsections, the conditional distribution of $\boldsymbol{Y}$ is used to predict a function of the attribute under study and to assess the uncertainty in the values of this attribute.

2.3. First application: prediction of a transfer function

Let us consider a function $\varphi(\boldsymbol{Y})$ of the unknowns $Y(\mathbf{u}_1),\ldots,Y(\mathbf{u}_k)$ (*transfer* function or, in the mining terminology, *recovery* function). This function can be predicted by its conditional expectation, that is, by the expected value of its conditional distribution:

$$[\varphi(\boldsymbol{Y})]^*=E\{\varphi(\boldsymbol{y}^{SK}+\mathbf{A}^{SK}\boldsymbol{T})\}. \tag{9}$$

In practice, the expectation can be calculated by Monte Carlo integration (Verly, 1984), by putting

$$[\varphi(\boldsymbol{Y})]^*\approx\frac{1}{N}\sum_{i=1}^{N}\varphi(\boldsymbol{y}^{SK}+\mathbf{A}^{SK}\boldsymbol{t}_i), \tag{10}$$

where N is a large positive integer and $\boldsymbol{t}_1,\ldots,\boldsymbol{t}_N$ are realizations of $\boldsymbol{T}$ obtained by simple random sampling or, to speed the rate of convergence of Monte Carlo

integration, by stratified sampling, e.g. Latin hypercube sampling (McKay et al., 1979).

2.4. Example: recoverable resources in an ore deposit

Suppose that the original attribute is the grade of an element of interest in an ore deposit. Let v be a block representing a selective mining unit and $\{\mathbf{u}_1,\ldots,\mathbf{u}_q\}$ a set of locations that discretize this block. Provided that this discretization is fine enough, the true block grade can reasonably be approximated as follows (Eq. (1)):

$$Z(v) \approx \frac{1}{q}\sum_{i=1}^{q} Z(\mathbf{u}_i) = \frac{1}{q}\sum_{i=1}^{q} \phi(Y(\mathbf{u}_i)). \tag{11}$$

The selection of the block as ore or waste material is based on a predicted grade, which is usually derived from a weighted average of production data (e.g. blast hole data) located at $\mathbf{u}_{q+1},\ldots,\mathbf{u}_k$:

$$Z^*(v) = \sum_{i=q+1}^{k} \omega_i Z(\mathbf{u}_i) = \sum_{i=q+1}^{k} \omega_i \phi(Y(\mathbf{u}_i)). \tag{12}$$

From the information on drill hole data (exploration stage), it is of interest to anticipate the resources that will be recovered at the mining stage. One assumes that the pattern of blast hole sampling $\{\mathbf{u}_{q+1},\ldots,\mathbf{u}_k\}$ and the weighting $\{\omega_{q+1},\ldots,\omega_k\}$ are known, but not the blast hole grades. Let us define the vector of unknowns as $\boldsymbol{Y} = (Y(\mathbf{u}_1),\ldots,Y(\mathbf{u}_q),\ Y(\mathbf{u}_{q+1}),\ldots,Y(\mathbf{u}_k))$. The actual recoverable resources above a given cutoff z (incorporating the information effect produced by ore-waste misclassifications at the production stage) can be assessed by the following recovery functions:

- Tonnage above cutoff z:

$$\varphi(\boldsymbol{Y}) = I\left\{\sum_{i=q+1}^{k} \omega_i \phi(Y(\mathbf{u}_i)); z\right\}, \tag{13}$$

  where $I(Z;z)$ is an indicator function, equal to 1 if Z is greater than z, 0 otherwise.
- Metal content above cutoff z:

$$\varphi(\boldsymbol{Y}) = \left\{\frac{1}{q}\sum_{i=1}^{q} \phi(Y(\mathbf{u}_i))\right\} \times I\left\{\sum_{i=q+1}^{k} \omega_i \phi(Y(\mathbf{u}_i)); z\right\}. \tag{14}$$

2.5. Second application: assessment of local uncertainty

Apart from predicting the expected recovery functions, it is also of interest to quantify the uncertainty in the actual block grade $Z(v)$, which is helpful for classifying the mineral resources into measured, indicated and inferred resources. Several measures of uncertainty can be derived from the conditional distribution of $Z(v)$ (in practice, from the empirical distribution of a large set of block grades obtained by Monte Carlo simulation), e.g.:

- conditional variance;
- bounds of the interval in which $Z(v)$ has a given probability to lie.

3. Multi-Gaussian kriging with a constant but uncertain mean

3.1. Model with an unknown mean

The assumption of a known mean is quite restrictive, as it does not allow any uncertainty in the value of this parameter. In particular, the prediction of a transfer function is biased when the mean is misspecified. To weaken the model requisites, one option is to use the following model:

$$\forall \mathbf{x} \in \mathbf{D}, \quad Y(\mathbf{x}) = m + U(\mathbf{x}), \tag{15}$$

in which $U = \{U(\mathbf{x}),\ \mathbf{x} \in \mathbf{D}\}$ is a stationary and ergodic standard Gaussian random field, while m is an unknown scalar parameter. Provided mild assumptions, one can derive an unbiased predictor of any transfer function $\varphi(\boldsymbol{Y})$ based on the ordinary kriging of $\boldsymbol{Y}$ (Emery, 2006a, d). Such a predictor relies on the definition of "pseudo" conditional distributions, instead of the true conditional distributions that would be derived if m were known. However, although they allow unbiased predictions of transfer functions, pseudo conditional distributions cannot be used to derive measures of local uncertainty, as they do not have the same dispersion as the true distributions (Emery, 2006a, b).

3.2. Model with a random mean

To avoid the previous difficulty, a different approach is proposed. It consists of replacing the unknown mean by a random variable constant over space:

$$\forall \mathbf{x} \in \mathbf{D}, \quad Y(\mathbf{x}) = M + U(\mathbf{x}), \tag{16}$$

where M is a Gaussian random variable with mean 0 and variance σ^2, independent of the standard Gaussian field U. The two random fields Y and U have the same semivariogram γ_U, but not the same variance nor the same covariance function. The covariance between $Y(\mathbf{x})$ and $Y(\mathbf{x}')$ is given by

$$\forall \mathbf{x}, \mathbf{x}' \in \mathbf{D}, \quad C_Y(\mathbf{x}, \mathbf{x}') = \sigma^2 + 1 - \gamma_U(\mathbf{x}, \mathbf{x}'). \tag{17}$$

This covariance does not tend to zero as the norm of $\mathbf{x}'-\mathbf{x}$ tends to infinity, indicating that the Gaussian random field Y is not ergodic.

From a practical point of view, the models with unknown mean (Eq. (15)) and with random mean (Eq. (16)) are undistinguishable on the basis of a single realization, hence either can be equally used. They only differ on how the uncertainty in the mean of the Gaussian field Y is accounted for: unknown scalar parameter (m) or random variable (M). Although the true conditional distributions are not accessible in the former case, as the prior distribution of Y is not fully known, we will see that they can be easily determined in the latter case.

3.3. Conditional distribution of a vector of unknowns

Except for the covariance function, the model in Eq. (16) is the same as that analyzed in Section 2. By

using Eq. (17) and putting

$$\mu_i = (1+\sigma^2)\left(\sum_{\beta=1}^{n}\lambda_{\beta,i}^{SK} - 1\right) \quad \forall i = 1,\ldots,k, \tag{18}$$

the simple kriging system at location $\mathbf{u}_i$ (Eq. (5)) can be re-written in the following fashion:

$$\begin{aligned}
&\sum_{\beta=1}^{n}\lambda_{\beta,i}^{SK}\gamma_U(\mathbf{x}_\alpha,\mathbf{x}_\beta) - \mu_i = \gamma_U(\mathbf{x}_\alpha,\mathbf{u}_i) \quad \forall \alpha = 1,\ldots,n,\\
&\sum_{\beta=1}^{n}\lambda_{\beta,i}^{SK} = 1 + \frac{\mu_i}{1+\sigma^2}.
\end{aligned} \tag{19}$$

The variance of the kriging error at this location (Eq. (6) with $j = i$, Eq. (17)) becomes:

$$\Sigma_{i,i}^{SK} = \sum_{\beta=1}^{n}\lambda_{\beta,i}^{SK}\gamma_U(\mathbf{u}_i,\mathbf{x}_\beta) - (1+\sigma^2)\left(\sum_{\beta=1}^{n}\lambda_{\beta,i}^{SK} - 1\right). \tag{20}$$

The case of a known mean value is met when σ^2 is equal to zero, which implies that the outcome of M is almost surely zero. Here, one is interested in the opposite situation, when there is a complete prior ignorance on the mean value. This is done by letting the variance of M become infinite. For the kriging error variance to be finite, it is seen from Eq. (20) that a necessary condition is that the kriging weights add to one. This condition entails that *simple kriging* tends to *ordinary kriging* when σ^2 tends to infinity:

$$\begin{aligned}
&\forall i = 1,\ldots,k, \quad y(\mathbf{u}_i)^{SK} \underset{\sigma^2\to+\infty}{\longrightarrow} y(\mathbf{u}_i)^{OK} = \sum_{\alpha=1}^{n}\lambda_{\alpha,i}^{OK} y_\alpha,\\
&\mathbf{\Sigma}^{SK} \underset{\sigma^2\to+\infty}{\longrightarrow} \Sigma^{OK} = \mathbf{A}^{OK}(\mathbf{A}^{OK})^t,
\end{aligned} \tag{21}$$

with, for $i, j = 1,\ldots,k$:

$$\begin{aligned}
&\sum_{\beta=1}^{n}\lambda_{\beta,i}^{OK}\gamma_U(\mathbf{x}_\alpha,\mathbf{x}_\beta) - \mu_i = \gamma_U(\mathbf{x}_\alpha,\mathbf{u}_i) \quad \forall \alpha = 1,\ldots,n,\\
&\sum_{\beta=1}^{n}\lambda_{\beta,i}^{OK} = 1,\\
&\Sigma_{i,j}^{OK} = -\gamma_U(\mathbf{u}_i,\mathbf{u}_j) + \sum_{\beta=1}^{n}\lambda_{\beta,j}^{OK}\gamma_U(\mathbf{u}_i,\mathbf{x}_\beta) - \mu_j.
\end{aligned} \tag{22}$$

The last identity (covariance between the ordinary kriging errors at locations $\mathbf{u}_i$ and $\mathbf{u}_j$) is obtained by substituting the expression of the prior covariance function (Eq. (17)) into Eq. (6) and by using Eq. (18). One observes that σ^2 is absent from system (22), hence the entries of the covariance matrix $\mathbf{\Sigma}^{OK}$ no longer depend on σ^2 and therefore remain finite when σ^2 tends to infinity. In this limit case, the conditional distribution of $\boldsymbol{Y}$ is Gaussian and its first- and second-order moments are the vector of ordinary kriging predictions ($\boldsymbol{y}^{OK}$) and the variance–covariance matrix of the prediction errors $\mathbf{\Sigma}^{OK}$. Conditionally to the data (Eq. (3)), $\boldsymbol{Y}$ can be decomposed as follows:

$$\boldsymbol{Y} = \boldsymbol{y}^{OK} + \mathbf{A}^{OK}\boldsymbol{T}, \tag{23}$$

with $\boldsymbol{T}$ a standard Gaussian vector independent of the data and with mutually independent components. As in Section 2, Monte Carlo simulation allows calculating the conditional expectation of any transfer function and measures of local uncertainty (probability intervals, conditional variance, etc.).

3.4. Notes

(1) The above results are remarkably simple and easy to implement, since they amount to substituting ordinary kriging for simple kriging in the approach presented in Section 2. However, to avoid misinterpretations, it should be noted that the condition on the sum of kriging weights is *not* an unbiasedness constraint; rather, it is a requirement for the variance of the kriging error to remain finite.

(2) The prior variance of the Gaussian random field Y becomes infinitely large when the variance of the random mean M tends to infinity (Eq. (17)). However, as M is constant in space, the *experimental* variance of the normal data (Eq. (3)) should remain close to the prior variance of U, i.e. 1. Put another way, the random variable M does not introduce spatial variability, but a point-wise variability over the realizations of Y that reflects the uncertainty in the mean. From a practical point of view, one can work with normal scores data with an experimental mean close to 0 and an experimental variance close to 1, as in the traditional multi-Gaussian kriging approach. The only difference is the assumption that the spatial mean of Y is uncertain (therefore, possibly different from the experimental zero mean) and can be represented by a random variable.

(3) The use of ordinary kriging is advantageous when the data are scarce in the kriging neighborhood and the mean value can be considered constant only at the scale of this neighborhood (local stationarity).

4. Multi-Gaussian kriging in the presence of a drift

The results of the previous section can be extended to a more general model, in which the Gaussian random field Y is split into a drift component (M) that varies smoothly in space and a residual component (U) with a standard Gaussian distribution:

$$\forall \mathbf{x} \in \mathbf{D}, \quad Y(\mathbf{x}) = M(\mathbf{x}) + U(\mathbf{x}). \tag{24}$$

The drift is represented by a weighted average of the form:

$$\forall \mathbf{x} \in \mathbf{D}, \quad M(\mathbf{x}) = \sum_{\ell=1}^{L} A_\ell f_\ell(\mathbf{x}), \tag{25}$$

where $\{f_\ell,\ \ell = 1,\ldots,L\}$ is a set of known basis functions, usually chosen as monomials of the coordinates (universal kriging model), cosine functions (trigonometric kriging model) or exhaustively known secondary variables (external drift model) (Matheron, 1971, p. 142; Séguret and Huchon, 1990; Hudson and Wackernagel, 1994). So far, most of these models have been developed with deterministic and unknown coefficients $\{A_\ell,\ \ell = 1,\ldots,L\}$.

Here, we assume that these coefficients are independent Gaussian random variables and are independent of

the residual component U. Let ${\sigma_\ell}^2$ denote the prior variance of A_ℓ and suppose that U is a stationary ergodic Gaussian random field with covariance function C_U. In this context, the covariance between $Y(\mathbf{x})$ and $Y(\mathbf{x}')$ is given by

$$C_Y(\mathbf{x},\mathbf{x}') = C_U(\mathbf{x},\mathbf{x}') + \sum_{\ell=1}^{L} \sigma_\ell^2 f_\ell(\mathbf{x}) f_\ell(\mathbf{x}'). \tag{26}$$

The simple kriging variance of the ith component of $\boldsymbol{Y}$ becomes (Eq. (6) with $j = i$, Eq. (26)):

$$\begin{aligned}\Sigma_{i,i}^{SK} &= C_U(\mathbf{u}_i,\mathbf{u}_i) - \sum_{\beta=1}^{n} \lambda_{\beta,i}^{SK} C_U(\mathbf{u}_i,\mathbf{x}_\beta) \\ &\quad - \sum_{\ell=1}^{L} \sigma_\ell^2 f_\ell(\mathbf{u}_i) \left\{ \sum_{\beta=1}^{n} \lambda_{\beta,i}^{SK} f_\ell(\mathbf{x}_\beta) - f_\ell(\mathbf{u}_i) \right\}. \end{aligned} \tag{27}$$

To account for a complete prior ignorance on the drift, one assumes that the variances of the drift coefficients are infinitely large: ${\sigma_\ell}^2 \to \infty$ for $\ell = 1,\ldots,L$. From Eq. (27), one sees that a necessary condition for the kriging variances to remain finite is that the kriging weights satisfy the following constraints:

$$\sum_{\beta=1}^{n} \lambda_{\beta,i}^{SK} f_\ell(\mathbf{x}_\beta) = f_\ell(\mathbf{u}_i), \quad \forall \ell = 1,\ldots,L \ \forall i = 1,\ldots,k. \tag{28}$$

One recognizes the unbiasedness constraints introduced in universal kriging. Accordingly, if the variances of the drift coefficients tend to infinity, simple kriging (SK) tends to universal kriging (UK). For $i,j = 1,\ldots,k$, the kriging weights and the error covariances are given by (Eqs. (5), (6), (26) and (28))

$$\begin{aligned} &\sum_{\beta=1}^{n} \lambda_{\beta,i}^{UK} C_U(\mathbf{x}_\alpha,\mathbf{x}_\beta) + \sum_{\ell=1}^{L} \mu_{\ell,i} f_\ell(\mathbf{x}_\alpha) = C_U(\mathbf{x}_\alpha,\mathbf{u}_i) \quad \forall \alpha = 1,\ldots,n, \\ &\sum_{\beta=1}^{n} \lambda_{\beta,i}^{UK} f_\ell(\mathbf{x}_\beta) = f_\ell(\mathbf{u}_i) \quad \forall \ell = 1,\ldots,L, \\ &\Sigma_{i,j}^{UK} = C_U(\mathbf{u}_i,\mathbf{u}_j) - \sum_{\beta=1}^{n} \lambda_{\beta,j}^{UK} C_U(\mathbf{u}_i,\mathbf{x}_\beta) - \sum_{\ell=1}^{L} \mu_{\ell,j} f_\ell(\mathbf{u}_i). \end{aligned} \tag{29}$$

This system is established in the same fashion as Eq. (22), which corresponds to the case of a single basis function $f_1 \equiv 1$. Also, the convergence of simple kriging to universal kriging has been demonstrated by Omre and Halvorsen (1989).

In this model, the conditional distribution of $\boldsymbol{Y}$ is Gaussian, with mean vector equal to the universal kriging of $\boldsymbol{Y}$ and variance–covariance matrix equal to that of the universal kriging errors. The results found in the previous sections can therefore be extended by substituting universal kriging for simple or ordinary kriging. As for the ordinary kriging approach, it should be noted that Eq. (28) does not correspond to unbiasedness constraints *stricto sensu*, but to constraints ensuring that the prediction errors have finite variances.

A problem specific to the universal kriging approach is the inference of the residual covariance C_U, as the available data (Eq. (3)) concern the non-stationary Gaussian field Y, not the residual field U. Several options are possible to get out of this problem (Chilès and Delfiner, 1999, p. 115, 281):

- identify a direction along which the drift is approximately constant and calculate the sample semi-variogram along this direction only;
- restrict variogram analysis to short distances, for which the bias between the sample and theoretical semi-variograms is small;
- use the formalism of intrinsic random fields and generalized covariances and work with *generalized increments* that filter the drift.

Another problem in the implementation of multi-Gaussian kriging in the presence of a drift is the difficulty to validate the univariate and multivariate Gaussian assumptions. In practice, checks should be done by analyzing the distributions of generalized increments, which are not "contaminated" by the drift.

To avoid these difficulties, ordinary kriging is often preferred to universal kriging, even if the data reveal clear spatial trends (Journel and Rossi, 1989). The difference between both types of kriging is usually small when the data are abundant and becomes relevant only when data are scarce or in extrapolation situations.

5. Program description

The main program associated with this article (**MK3D.M**) is a Matlab function that uses multi-Gaussian kriging to predict transfer functions and to calculate measures of local uncertainty. It works with data located in a three-dimensional space and requires the following input arguments.

coord coordinates of scattered locations (block centers) targeted for predictions (void for gridded locations)

x0,y0,z0 if gridded locations: minimum grid coordinates along x, y and z directions

nx,ny,nz number of grid nodes along x, y and z directions

dx,dy,dz grid meshes (block size) along x, y and z directions

nd block discretization along x, y and z directions (1×3 vector)

datacoord data coordinates ($n \times 3$ matrix)

ydata Gaussian data values ($n \times 1$ vector)

limits trimming limits (inf and sup) for Gaussian data (1×2 vector)

tableZY conversion table between original (Z) and Gaussian (Y) values (void if no transformation is required). The first column contains the original values, the second column their standard normal score transforms

zmin,zmax minimum and maximum values for the original (Z) variable

tail additional parameters for tail modeling (Emery, 2006c, p. 972)

model semi-variogram model for Gaussian residuals ($nst \times 7$ matrix), see the Matlab file for details

cc sills of nested structures ($nst \times 1$ vector)
b additional parameters for some semi-variogram models ($nst \times 1$ vector)
nugget nugget effect variance
radius maximum search radii along rotated y, x and z directions
angles angles for anisotropic search (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 0)
ktype kriging type: $0 = SK$; $1 = OK$; k+1 $= UK$ with drift of degree k
cutoff cutoffs for recovery calculations ($1 \times ncutoff$ vector)
prob probabilities for probability intervals ($1 \times nprob$ vector)
nrealiz number of realizations for Monte Carlo simulation with Latin hypercube sampling
inf_effect blast hole coordinates (relative to the block center) along x, y and z and weights for blast hole data (Eq. (12)) (void if no information effect has to be considered).

The outputs of MK3D.M consist of three recovery functions that are of interest in mineral resources evaluation: metal content (Eq. (14)), tonnage (Eq. (13)) and mean grade (metal divided by tonnage) associated with the *ncutoff* input cutoffs. Additional outputs are the interval bounds for the *nprob* input probabilities and the local variances of block grades.

Program MK3D.M can also be run with a parameter file: no input argument is required in the Matlab workspace and the user is prompted for the parameter file name. If no name is entered, a default file (MK3D.PAR) is assumed; if this file is not found, a blank file is created. When using the parameter file mode, the coordinates of the locations targeted for prediction (if these locations are not gridded), the data coordinates and data values and the conversion table between original and Gaussian values must be stored in ASCII files without header. The outputs of MK3D.M are written in separate ASCII files with a common base name.

Program MK3D.M uses 10 subroutines:

BACKTR.M back-transformation from Gaussian to original values
COVA.M calculate covariance values
COVMATRIX.M calculate covariance matrices
COVUNIQUE.M calculate covariance matrix and its inverse for kriging in a unique neighborhood (case when the search radii are set to infinity)
CREATE_PARAMFILE.M create default parameter file MK3D.PAR
KRIGE.M calculate kriging weights
PICKSUPR.M build template of super-blocks for search strategy
SEARCH.M search data located in a moving neighborhood
SETROT.M set up matrix for rotation and reduction of coordinates
SUPERBLK.M set up super-block strategy.

6. First case study: mining dataset

In this section, we apply multi-Gaussian kriging to predict the recoverable resources in a porphyry copper deposit located in the Central Andes, Chile. The available information consists of 2811 composite data (10 m length) from a set of exploration drill holes, in which the total copper grade has been assayed. To preserve the confidentiality of the data, the grades have been multiplied by an undisclosed factor. A location map and a histogram of the assayed grades are shown in Fig. 1A and B.

The prediction of mineral resources proceeds according to the following steps.

6.1. Normal scores transformation

Declustering weights are assigned to the data by using the cell method (Deutsch and Journel, 1998, p. 213). Let $z_1,\ldots,z_n$ denote the original data values, ranked in increasing order and $p_1,\ldots,p_n$ their respective declustering weights. The transformed values (Eq. (3)) are

$$\forall \alpha \in \{1,\ldots,n\}, \quad y_\alpha = G^{-1}\left(\sum_{\beta=1}^{\alpha-1} p_\beta + \frac{1}{2}p_\alpha\right), \tag{30}$$

where G^{-1} is the inverse standard Gaussian cumulative distribution function. This step provides normal scores data with experimental mean and variance close to zero and one, respectively.

6.2. Validation of the bivariate Gaussian assumption

Before using multi-Gaussian kriging, it is worthwhile to check whether or not the multivariate Gaussian assumption is compatible with the normal scores data. In practice, the check is limited to bivariate distributions, as the inference of higher-order distributions is beyond reach. A test that does not depend on the (possibly unknown) mean value of the normal scores data consists of examining the semi-variogram of order $\omega > 0$:

$$\forall \mathbf{x},\mathbf{x}' \in \mathbf{D}, \quad \gamma_Y^{(\omega)}(\mathbf{x},\mathbf{x}') = \tfrac{1}{2}E\{|Y(\mathbf{x}) - Y(\mathbf{x}')|^\omega\}. \tag{31}$$

For $\omega = 2$, one finds the traditional semi-variogram, whereas the cases $\omega = 1$ and 0.5 correspond to the semi-madogram and semi-rodogram. If the random field Y has bivariate Gaussian distributions, then one has (Emery, 2005, p. 168):

$$\gamma_Y^{(\omega)}(\mathbf{x},\mathbf{x}') = \frac{2^{\omega-1}}{\sqrt{\pi}}\Gamma\left(\frac{\omega+1}{2}\right)[\gamma_Y^{(2)}(\mathbf{x},\mathbf{x}')]^{\omega/2}, \tag{32}$$

where $\Gamma(\cdot)$ is the gamma function. Eq. (32) indicates that, in log–log coordinates, the points plotting the semi-variogram of order ω as a function of the traditional semi-variogram are aligned with slope $\omega/2$. Once applied to the transformed data, this test is quite satisfactory (Fig. 1C and D), hence the multivariate Gaussian assumption is deemed acceptable for the data under study.

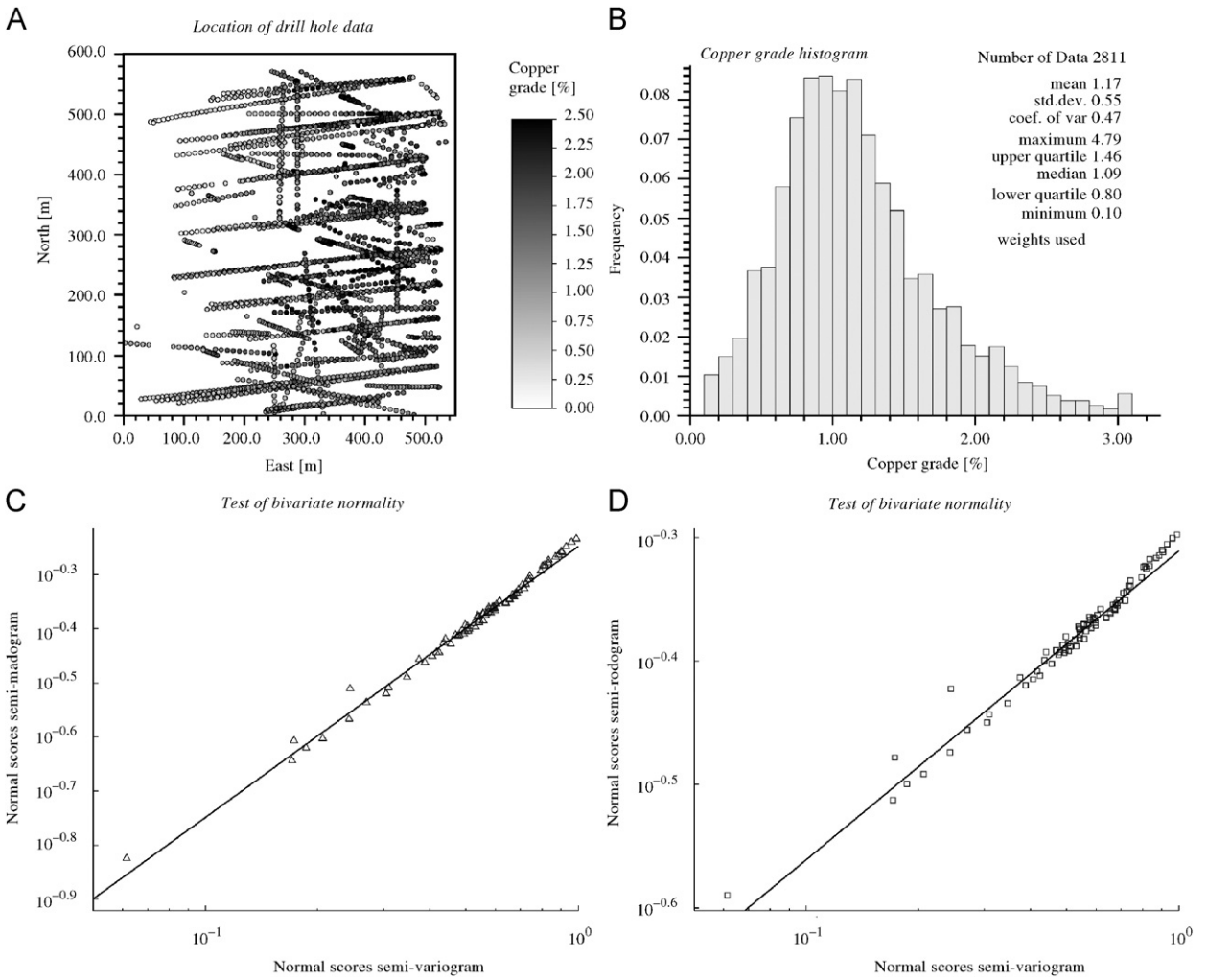

Fig. 1. Exploratory analysis of copper grade data. (A) Map of data locations (projection onto horizontal plane); (B) empirical grade histogram; (C) and (D) sample semi-madogram and semi-rodogram as a function of sample semi-variogram; theoretical lines associated with bivariate Gaussian model are superimposed.

6.3. Variogram analysis of normal scores data

The semi-variogram of the normal scores data is calculated along the main anisotropy directions (north, east and vertical directions). The fitted model is the sum of a nugget effect, a spherical and an exponential model (Fig. 2).

6.4. Prediction of recoverable resources at a block support

Multi-Gaussian kriging is now used to assess the recoverable resources at a $20 \times 20 \times 20\,\mathrm{m}^3$ block support. For Monte Carlo integration, each block is discretized into $6 \times 6 \times 2$ points and 250 realizations are used. Calculations are made by using simple and ordinary kriging of the Gaussian data. The search radii are set to 250, 150 and 200 m along the north, east and vertical directions, respectively, while the optimum number of data per octant is set to 4.

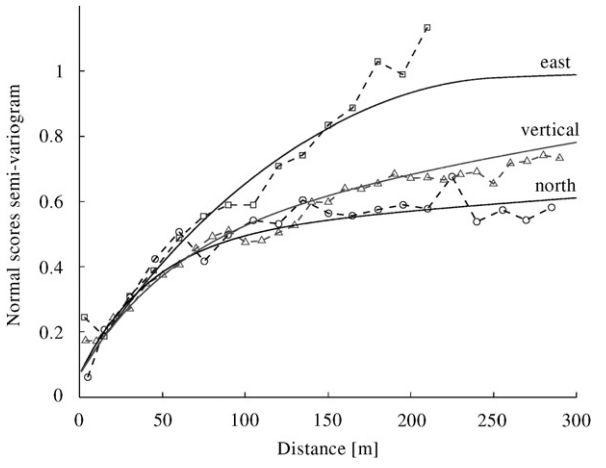

Fig. 2. Sample (dashed lines) and modeled (solid lines) semi-variogram of normal scores data along main anisotropy directions.

For each block, the predicted resources consist of the metal content and tonnage above a cutoff of 0.8%Cu, taking into account the information effect that will occur at the time of ore-waste selection, by assuming that the block grade will be predicted from 16 regularly spaced blast hole data (Figs. 3 and 4).

One observes that predictions obtained by using simple and ordinary kriging are similar, except in the western area of the deposit where drill hole data are scarce (Fig. 1A). In this area, simple multi-Gaussian kriging yields predictions that tend to the average metal content and tonnage above 0.8%, giving the impression that mineral resources can be recovered (Fig. 5A). Such predictions are not warranted by the data values, which decrease when going towards the west (Fig. 5B), but by the model that postulates a constant and known mean grade over space. In contrast, ordinary multi-Gaussian kriging only uses the data located in the moving neighborhood and yields more plausible predictions in the light of the grade distribution observed in Fig. 5B (mineral to the east and waste to the west).

6.5. Cross-validation

To validate the predictions obtained by ordinary multi-Gaussian kriging, a cross-validation is performed. At each data location, the point-support metal contents above given cutoffs (from 0.0% to 2.0%) are assessed by using the data situated at least 50 m away from the target location. This restriction is imposed because the accounting for very close data (in particular, adjacent data in the same drill hole as the data under consideration) often leads to accurate predictions, irrespective of the model, which hinders the validation (Isaaks and Srivastava, 1989, p. 358).

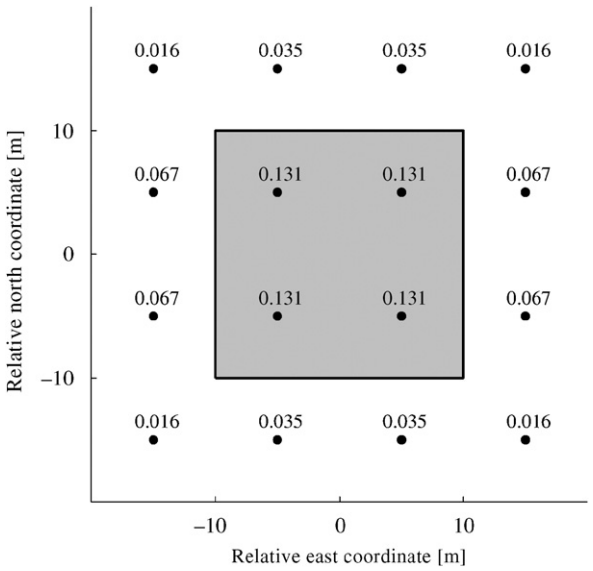

Fig. 3. Pattern of blast hole locations relative to block targeted for prediction (gray). Ordinary kriging weights are indicated above each blast hole location.

The average true and predicted metal contents are given in Table 1, together with the slope of the regression of the former upon the latter. It is seen that the mean error is always close to zero and the slope of the regression close to 1, which suggests that predictions do not suffer from global and conditional bias. A scatter diagram between the true and predicted grades (metal contents at cutoff 0%) is given in Fig. 6A.

One can also validate the uncertainty model derived from the conditional distributions, by constructing probability intervals with, say, probabilities from 0.1 to 0.9, and plotting these probabilities versus the fractions of data that actually belong to the intervals so defined (Fig. 6B). The plot indicates a good match, therefore, a correct modeling of the local uncertainty.

It is worth mentioning that the quality of the validation results depends on the number of conditioning data considered in the kriging neighborhood. For instance, with the current implementation (four data per octant), the slope of the regression of the true grades upon the predicted grades is equal to 0.99 (Fig. 6A). It decreases to 0.94 when considering one datum per octant, and to 0.75 when considering only four data in the neighborhood (without octant division). Following Rivoirard (1987), this regression slope is a helpful tool to choose the size of the ordinary kriging neighborhood.

7. Second case study: sequential simulation

The conditional distributions determined by ordinary kriging can also be used in the context of conditional simulation (Emery, 2007). The simplest approach is the sequential Gaussian algorithm, in which each value is simulated in turn, conditionally to the original data and previously simulated values. In practice, this algorithm requires the definition of a moving neighborhood in which to search the conditioning information.

In the following, we will perform sequential simulation at the nodes of a regular grid with 100,000 nodes, by considering a pure nugget semi-variogram with a unit sill and a single original datum located at the origin and with a zero value. Ten realizations are generated, in each of which the experimental mean and variance are calculated.

A sensitivity analysis is made by changing the number of conditioning data (original data and already simulated values) considered in the kriging neighborhood (Table 2). In the last case, which corresponds to a unique neighborhood implementation, the experimental variance of the simulated values obtained by using ordinary kriging is very close to 1, i.e. it almost matches the expected variance although the mean differs from 0. This situation does not hold if the kriging neighborhood is limited to 1000 conditioning data or less. In contrast, the traditional approach based on simple kriging always leads to a mean close to 0 and a variance close to 1, irrespective of the kriging neighborhood (this is explained because the weights assigned to the conditioning data are zero, since the semi-variogram is a nugget effect).

This exercise proves that, although the use of ordinary kriging for determining conditional distributions and for

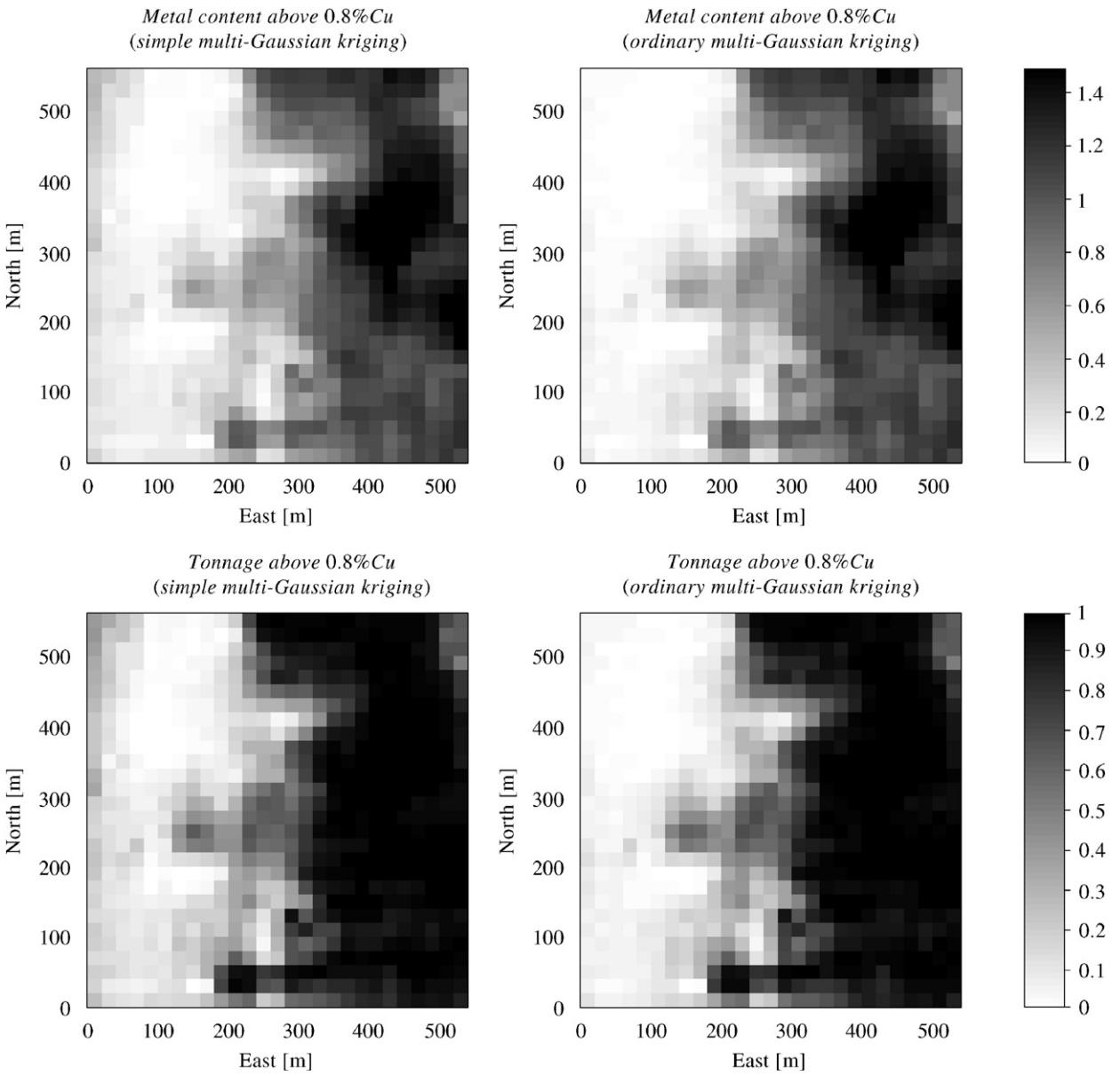

Fig. 4. Multi-Gaussian kriging predictions of metal contents (in %Cu) and tonnages (non-dimensional) above cutoff 0.8% Cu. Representation of bench with elevation 310 m.

sequential conditional simulation is legitimate, care must be taken in the design of the kriging neighborhood. In particular, too restrictive neighborhoods are likely to lead to an overestimation of the true spatial variability. The reason is twofold:

- The moving neighborhood ignores part of the conditioning data and loses information when constructing the conditional distributions. In the case of the pure nugget semi-variogram, this problem is important because closest data do not screen off the influence of farthest data (the ordinary kriging weights of all the data are the same).
- Each simulated value is re-used for conditioning the next ones. Hence, any error made by restricting the neighborhood propagates when the simulation proceeds. Note that this problem is not raised in the multi-Gaussian kriging approach presented in Sections 3 and 6, insofar as the conditioning data required for ordinary kriging only consist of the original data (Eq. (23)).

8. Conclusions

This work addressed the problem of assessing the uncertainty in the values of a spatial attribute and

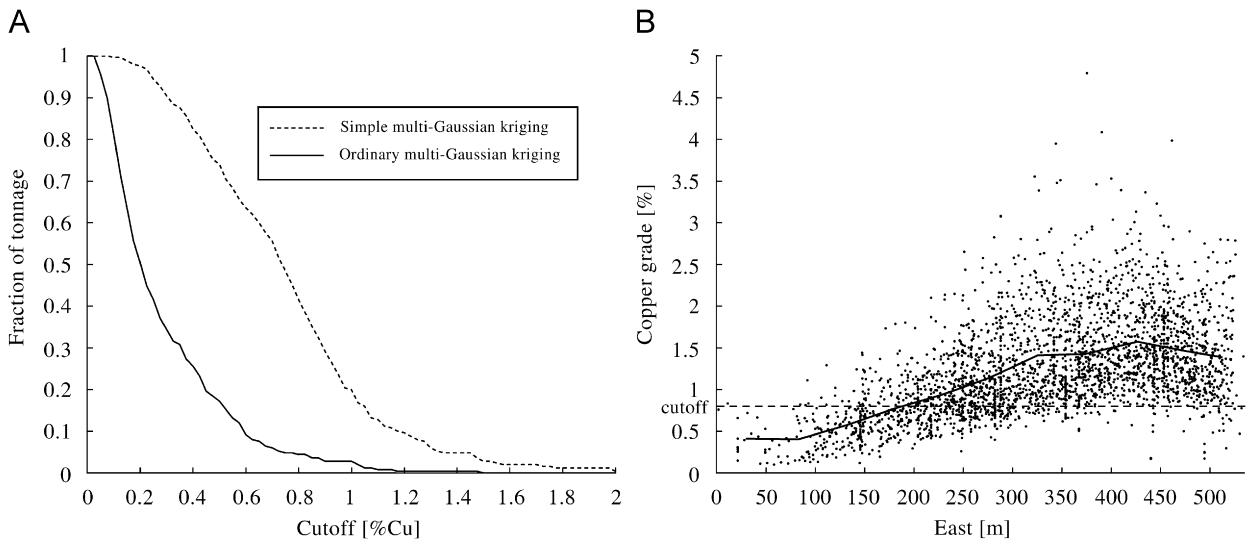

Fig. 5. (A) Grade-tonnage curves for block with coordinates (10,560,310). (B) Scatter plot between copper grades and east coordinate (broken line indicates empirical conditional mean).

Table 1
Cross-validation statistics for ordinary multi-Gaussian kriging

Cutoff grade (% Cu)	Mean of true metal content (Q)	Mean of predicted metal content (Q^*)	Slope of the regression of Q upon Q^*
0.0	1.24	1.24	0.99
0.2	1.24	1.24	0.99
0.4	1.23	1.23	0.99
0.6	1.20	1.21	1.00
0.8	1.12	1.13	1.00
1.0	0.97	0.99	0.96
1.2	0.79	0.81	0.95
1.4	0.63	0.62	0.97
1.6	0.48	0.47	0.98
1.8	0.37	0.34	0.98
2.0	0.25	0.23	0.96

predicting functions of this attribute, at the support of the available data or at a larger support. It focused on the multi-Gaussian model, in which the attribute is represented by the transform of a Gaussian random field.

In the traditional approach, the mean value of the Gaussian random field is supposedly known and its conditional distributions at any set of locations are jointly Gaussian, with means equal to simple kriging predictions and variance–covariance matrix equal to that of the prediction errors. An extension of this model consists of replacing the known mean by an independent random variable constant over space that adds to the Gaussian field. When the prior variance of this random variable becomes infinite, one obtains a model in which there is a total prior ignorance on the true mean value. The conditional distributions are then found by substituting ordinary kriging for simple kriging. Similarly, one can design a more general model that includes a random drift and use universal kriging instead of simple or ordinary kriging.

In all these models, the knowledge of the joint conditional distributions allows calculating the conditional expectation of a transfer function defined at a block support and deriving measures of the uncertainty in the actual block values, by discretizing the block into several points and by resorting to Monte Carlo simulation.

The proposed approach can easily be extended to the stochastic simulation framework, for instance by using the sequential Gaussian algorithm and substituting ordinary or universal kriging for simple kriging when deriving the successive conditional distributions. This constitutes a significant step towards the conditional simulation of locally stationary or non-stationary random fields.

Concerning the implementation of the proposed methods, close attention has to be paid to the design of the kriging neighborhood. In the light of the exercise made in the previous section, this design is critical for the sequential simulation algorithm, since too restrictive neighborhoods lead to an inaccurate reproduction of the spatial variability, especially if the semi-variogram model presents a nugget effect. Additional research is required in this respect to improve the neighborhood definition when using ordinary kriging.

Acknowledgements

This research has been funded by the Chilean Fund for Science and Technology Development (FONDECYT) and is a part of Project 1061103. The author is grateful to Georges Verly and to an anonymous reviewer for their constructive comments and acknowledges Codelco-Chile for providing the dataset used in this work.

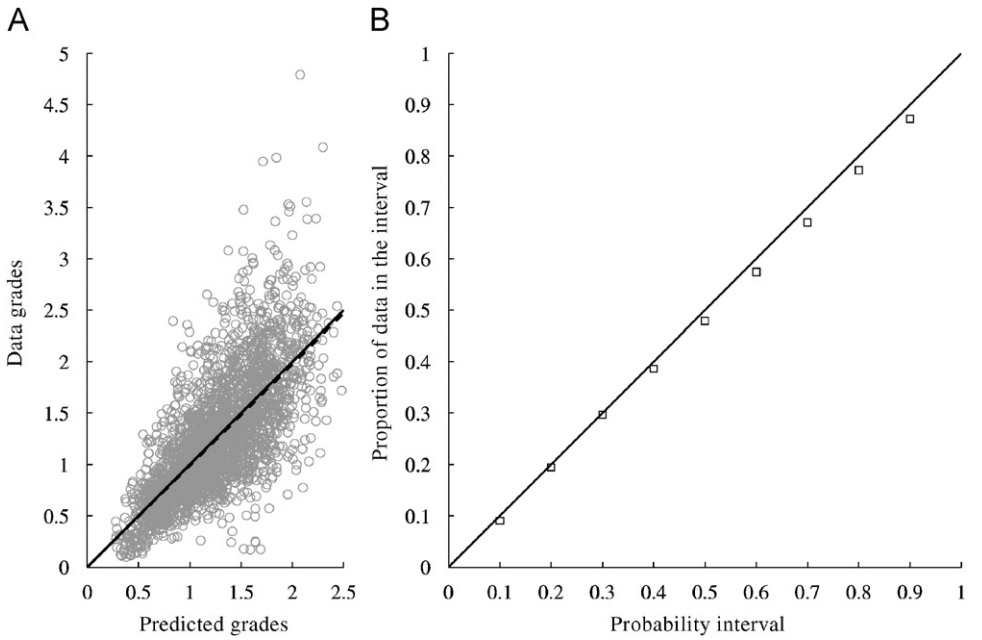

Fig. 6. Cross-validation results. (A) True versus predicted grades (solid line indicates diagonal; dashed line indicates linear regression); (B) Probabilities versus fractions of data that belong to probability intervals.

Table 2
Average means and variances of 10 realizations drawn over a regular 1D grid with 100,000 nodes, by using different kriging types (simple and ordinary) and neighborhoods

Number of conditioning data in kriging neighborhood	Known mean (simple kriging)		Uncertain mean (ordinary kriging)	
	Mean of realization	Variance of realization	Mean of realization	Variance of realization
10	−0.002	0.998	−26.600	445.800
100	−0.002	0.998	−2.971	5.830
1000	−0.002	0.998	−0.530	1.045
10,000	−0.002	0.998	−0.276	0.998
100,000	−0.002	0.998	−0.260	0.998

The program file and implementation parameters are given in SGSIM.M and INSTRUC.M.

Appendix A. Supplementary materials

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.cageo.2007.12.011.

References

Anderson, T.W., 2003. An Introduction to Multivariate Statistical Analysis, third ed. Wiley, Hoboken, NJ, 721pp.

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

Davis, M.W., 1987. Generating large stochastic simulations—the matrix polynomial approximation method. Mathematical Geology 19 (2), 99–107.

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

Emery, X., 2005. Variograms of order ω: a tool to validate a bivariate distribution model. Mathematical Geology 37 (2), 163–181.

Emery, X., 2006a. Multigaussian kriging for point-support estimation: incorporating constraints on the sum of the kriging weights. Stochastic Environmental Research and Risk Assessment 20 (1–2), 53–65.

Emery, X., 2006b. Ordinary multigaussian kriging for mapping conditional probabilities of soil properties. Geoderma 132 (1–2), 75–88.

Emery, X., 2006c. A disjunctive kriging program for assessing point-support conditional distributions. Computers & Geosciences 32 (7), 965–983.

Emery, X., 2006d. Two ordinary kriging approaches to predicting block grade distributions. Mathematical Geology 38 (7), 801–819.

Emery, X., 2007. Conditioning simulations of Gaussian random fields by ordinary kriging. Mathematical Geology 39 (6), 607–623.

Guibal, D., Remacre, A.Z., 1984. Local estimation of the recoverable reserves: comparing various methods with the reality on a porphyry copper deposit. In: Verly, G., David, M., Journel, A.G., Maréchal, A. (Eds.), Geostatistics for Natural Resources Characterization. Reidel, Dordrecht, pp. 435–448.

Harter, T., 1994. Unconditional and conditional simulation of flow and transport in heterogeneous, variably saturated porous media.

Unpublished Ph.D. Dissertation, The University of Arizona, Tucson, AZ, 418pp.
Hudson, G., Wackernagel, H., 1994. Mapping temperature using kriging with external drift: theory and an example from Scotland. International Journal of Climatology 14 (1), 77–91.
Isaaks, E.H., Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. Oxford University Press, New York, NY, 561pp.
Journel, A.G., Rossi, M.E., 1989. When do we need a trend model in kriging? Mathematical Geology 21 (7), 715–739.
Maréchal, A., 1984. Recovery estimation: a review of models and methods. In: Verly, G., David, M., Journel, A.G., Maréchal, A. (Eds.), Geostatistics for Natural Resources Characterization. Reidel, Dordrecht, pp. 385–420.
Matheron, G., 1971. The Theory of Regionalized Variables and its Applications. Ecole Nationale Supérieure des Mines de Paris, Fontainebleau, France, 212pp.
McKay, M.D., Beckman, R.J., Conover, W.J., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21 (2), 239–245.
Omre, H., Halvorsen, K.B., 1989. The Bayesian bridge between simple and universal kriging. Mathematical Geology 21 (7), 767–786.
Rivoirard, J., 1987. Two key parameters when choosing the kriging neighborhood. Mathematical Geology 19 (8), 851–856.
Rivoirard, J., 1994. Introduction to Disjunctive Kriging and Nonlinear Geostatistics. Oxford University Press, Oxford, UK, 181pp.
Schofield, N., 1988. Ore reserve estimation at the Enterprise gold mine, Pine Creek, Northern Territory, Australia. Part 2: the multigaussian kriging model. CIM Bulletin 81 (909), 62–66.
Séguret, S., Huchon, P., 1990. Trigonometric kriging: a new method for removing the diurnal variation from geomagnetic data. Journal of Geophysical Research 95 (B13), 21383–21397.
Verly, G., 1983. The multigaussian approach and its applications to the estimation of local reserves. Mathematical Geology 15 (2), 259–286.
Verly, G., 1984. The block distribution given a point multivariate normal distribution. In: Verly, G., David, M., Journel, A.G., Maréchal, A. (Eds.), Geostatistics for Natural Resources Characterization. Reidel, Dordrecht, pp. 495–515.