

# On Some Consistency Conditions for Geostatistical Change-of-Support Models<sup>1</sup>

Xavier Emery<sup>2</sup>

---

*This work deals with a family of geostatistical models used in application fields concerned with a change of support, such as mineral resources evaluation and polluted soil management. Three models are examined: the discrete Gaussian, Hermitian and Laguerrian models, which rely on a transformation of the variable of interest (mineral grade or pollutant concentration) defined on point and block supports into variables with Gaussian or gamma univariate distributions and isofactorial bivariate distributions. The focus is given to the relationships between the transformed variables at both supports, and to the conditions that these relationships imply on the model parameters. Additionally, guidelines are given for improving the variogram analysis of the transformed variables and for validating the change-of-support model.*

---

**KEY WORDS:** discrete Gaussian model, bivariate distributions, isofactorial bivariate models, multi-gaussian kriging, disjunctive kriging, uniform conditioning.

## INTRODUCTION

Change of support is a critical issue in geostatistical applications when the volumetric support of the available information is smaller than the support for which statements are required. Examples of such applications include mineral resources evaluation, reservoir characterization, groundwater hydrology, agricultural land management, remediation of polluted soils and toxic exposure studies (Matheron, 1976b; Chilès and Delfiner, 1999; Gelfand, Zhu, and Carlin, 2001; Webster and Oliver, 2001).

Two main geostatistical approaches can be used to deal with change-of-support problems. The first one is conditional simulation, which consists in drawing multiple realizations of the random field that represents the attribute of interest (e.g. a mineral grade or a pollutant concentration), then in regularizing each realization to the block support (Demougeot-Renard and de Fouquet, 2004; Journel

---

<sup>1</sup>Received 22 November 2005; accepted 14 August 2006; Published online: 12 April 2007.

<sup>2</sup>Department of Mining Engineering, University of Chile, Avenida Tupper 2069, Santiago, Chile; e-mail: xemery@cec.uchile.cl

and Kyriakidis, 2004). However, this approach is demanding because it requires defining the full set of finite-dimensional distributions of the random field and conditioning the realizations to the available data, which is usually not an easy task.

The second approach is to use nonlinear geostatistical methods such as disjunctive kriging or uniform conditioning (Matheron, 1976a; Maréchal, 1984; Remacre, 1989; Webster, 1991), which provide an analytical solution to the change-of-support problem. In general, these methods require defining a model of bivariate distributions, hence they are applicable to a wider range of situations than simulation methods. The goal of this paper is to point out the conditions needed to ensure the internal consistency of the change-of-support model, in particular concerning the variogram analysis step and the relationships between the random fields at different supports.

### GENERAL ASSUMPTIONS FOR GEOSTATISTICAL CHANGE-OF-SUPPORT MODELS

The domain under study  $D$  is divided into non-overlapping blocks, which are identical up to a translation, and the data locations are uniformly randomized within the blocks (Matheron, 1976b, p. 243; Chilès and Delfiner, 1999, p. 439). In the sequel, the bold letters  $\mathbf{x}$  and  $\mathbf{x}'$  will refer to fixed data locations inside blocks  $v$  and  $v'$  respectively, whereas the underlined letters  $\underline{\mathbf{x}}$  and  $\underline{\mathbf{x}'}$  to random locations uniformly distributed in  $v$  and  $v'$ . If two or more data belong to a same block, their random locations are supposed to be independent in this block.

The attribute under study will be represented by three different random fields:

- a point-support variable with fixed locations, denoted by  $\{Z(\mathbf{x}), \mathbf{x} \in D\}$
- the point-support variable  $\{Z(\underline{\mathbf{x}}), \underline{\mathbf{x}} \in D\}$  obtained by randomizing the locations within the blocks
- the regularized variable  $\{Z(v), v \in D\}$ , defined by averaging the point-support values within block  $v$

$$Z(v) = \frac{1}{|v|} \int_v Z(\mathbf{x}) d\mathbf{x} \quad (1)$$

where  $|v|$  represents the volume of  $v$ .

Geostatistical change-of-support models do not work directly on the previous  $Z$ -variables, but on transformed random fields  $\{Y(\underline{\mathbf{x}}), \underline{\mathbf{x}} \in D\}$  and  $\{Y_v, v \in D\}$  such

that

$$\begin{cases} Z(\underline{\mathbf{x}}) = \phi(Y(\underline{\mathbf{x}})) \\ Z(v) = \phi_v(Y_v) \end{cases} \quad (2)$$

where  $\phi$  and  $\phi_v$  are non-decreasing mappings referred to as the point-support and block-support transformation functions respectively. Henceforth, we shall also distinguish the following random fields:

- the point-support transformed variable with fixed locations  $\{Y(\underline{\mathbf{x}}), \underline{\mathbf{x}} \in D\}$
- the corresponding regularized variable  $\{Y(v), v \in D\}$  over block  $v$

$$Y(v) = \frac{1}{|v|} \int_v Y(\underline{\mathbf{x}}) d\underline{\mathbf{x}} \quad (3)$$

Let us introduce the following simple and cross-covariance functions

$$\begin{cases} \rho(\underline{\mathbf{x}}, \underline{\mathbf{x}}') = \text{cov}\{Y(\underline{\mathbf{x}}), Y(\underline{\mathbf{x}}')\} \\ \rho(\underline{\mathbf{x}}, \underline{\mathbf{x}}') = \text{cov}\{Y(\underline{\mathbf{x}}), Y(\underline{\mathbf{x}}')\} \\ \rho(\underline{\mathbf{x}}, v') = \text{cov}\{Y(\underline{\mathbf{x}}), Y_{v'}\} \\ \rho(v, v') = \text{cov}\{Y_v, Y_{v'}\} \end{cases} \quad (4)$$

Note that, except when  $\underline{\mathbf{x}}$  and  $\underline{\mathbf{x}}'$  are the same,  $\rho(\underline{\mathbf{x}}, \underline{\mathbf{x}}')$  is obtained by regularizing  $\rho(\underline{\mathbf{x}}, \underline{\mathbf{x}}')$  to the block support

$$\forall \underline{\mathbf{x}} \in v, \forall \underline{\mathbf{x}}' \in v', \rho(\underline{\mathbf{x}}, \underline{\mathbf{x}}') = \frac{1}{|v|^2} \int_v \int_{v'} \rho(\underline{\mathbf{u}}, \underline{\mathbf{v}}) d\underline{\mathbf{u}} d\underline{\mathbf{v}} \quad \text{if } \underline{\mathbf{x}} \neq \underline{\mathbf{x}}' \quad (5)$$

Such a “randomization” of the covariance function has been introduced in early applications of geostatistics in the scope of the random kriging theory, as a way to deal with data affected by positioning errors or to simplify the kriging equations in the presence of uneven sampling designs (Chilès, 1976). In particular, if  $v$  coincides with  $v'$ , Equation (5) gives

$$\forall \underline{\mathbf{x}}, \underline{\mathbf{x}}' \in v, \rho(\underline{\mathbf{x}}, \underline{\mathbf{x}}') = \text{var}\{Y(v)\} \quad \text{if } \underline{\mathbf{x}} \neq \underline{\mathbf{x}}' \quad (6)$$

This relationship will be used later to find conditions ensuring the consistency of the change-of-support models. The next section examines the example of the discrete Gaussian model. This model is suited to the estimation of recoverable resources by multigaussian kriging, bigaussian disjunctive kriging or uniform

conditioning (Guibal and Remacre, 1984; Emery, 2005b; Emery and Soto Torres, 2005) and has received the most attention in practical applications. Two other models for continuous attributes (the discrete Hermitian and discrete Laguerrian models) will be analyzed in the subsequent section.

### DISCRETE GAUSSIAN MODEL

The discrete Gaussian model relies on the assumption that the two random fields  $\{Y_v, v \in D\}$  and  $\{Y(\underline{x}), \underline{x} \in D\}$  have standard Gaussian univariate distributions (with zero means and unit variances) and jointly stationary bivariate Gaussian distributions. Let us define the change-of-support coefficient as the correlation coefficient between  $Y(\underline{x})$  and  $Y_v$  when  $\underline{x}$  belongs to  $v$

$$r = \text{corr}\{Y(\underline{x}), Y_v\} = \rho(\underline{x}, v) \quad \text{with } \underline{x} \in v \quad (7)$$

By using Cartier's relationship (Matheron, 1984a, p. 425), one can express the block-support transformation function  $\phi_v$  as a function of the change-of-support coefficient  $r$  and the point-support transformation function  $\phi$  (Rivoirard, 1994, p. 83). The following subsections examine the implications of the bivariate Gaussian hypothesis, focusing on variogram analysis and on the approximations implied by the change-of-support model.

#### Relation Between the Covariance Functions of the Gaussian Fields

Due to the randomization of the data locations within the blocks, the following identities are found (Rivoirard, 1994, p. 89)

$$\begin{aligned} \forall \underline{x} \in v, \forall \underline{x}' \in v', \rho(\underline{x}, \underline{x}') &= \begin{cases} r\rho(\underline{x}, v') & \text{if } \underline{x} \neq \underline{x}' \\ 1 & \text{otherwise} \end{cases} \\ \forall \underline{x} \in v, \rho(\underline{x}, v') &= r\rho(v, v') \end{aligned} \quad (8)$$

These relations make the discrete Gaussian model very simple and straightforward to use, as it only depends on the change-of-support coefficient and a single covariance model. Note that the randomization of the data locations within the blocks produces a loss of information, as the exact positions of the data are forgotten. However, without such an assumption, the model leads to mathematical inconsistencies (Matheron, 1976b, p. 241; Chilès and Delfiner, 1999, p. 438). As a particular case of Equation (8), when  $\underline{x}$  and  $\underline{x}'$  belong to the same block  $v$ , one

has Equation (7)

$$\forall \underline{\mathbf{x}}, \underline{\mathbf{x}}' \in v, \rho(\underline{\mathbf{x}}, \underline{\mathbf{x}}') = \begin{cases} r^2 & \text{if } \underline{\mathbf{x}} \neq \underline{\mathbf{x}}' \\ 1 & \text{otherwise} \end{cases} \quad (9)$$

### Relation Between the Block-Gaussian and Regularized Gaussian Variables

Let us now consider the regularized Gaussian variable  $Y(v)$  [Eq. (3)]. As mentioned by Chilès and Delfiner (1999, p. 438), this variable differs from the block-Gaussian variable  $Y_v$ . However both variables are closely related, as it will be shown hereafter. On the one hand, their covariance is equal to the change-of-support coefficient [Eq. (7)]

$$\begin{aligned} \text{cov}\{Y(v), Y_v\} &= \frac{1}{|v|} \int_v \text{cov}\{Y(\mathbf{x}), Y_v\} d\mathbf{x} \\ &= \text{cov}\{Y(\underline{\mathbf{x}}), Y_v\} \quad \text{with } \underline{\mathbf{x}} \in v \\ &= r \end{aligned} \quad (10)$$

On the other hand, by comparing Eq. (9) with Eq. (6), it comes

$$\text{var}\{Y(v)\} = r^2 \quad (11)$$

Because of Eqs. (10) and (11), the correlation coefficient between the regularized Gaussian and block-Gaussian variables is equal to one

$$\text{corr}\{Y(v), Y_v\} = 1 \quad (12)$$

This equality implies that the block-Gaussian variable is the same as the regularized Gaussian variable, up to a multiplicative factor equal to the change-of-support coefficient [Eq. (11)]

$$Y(v) = rY_v \quad (13)$$

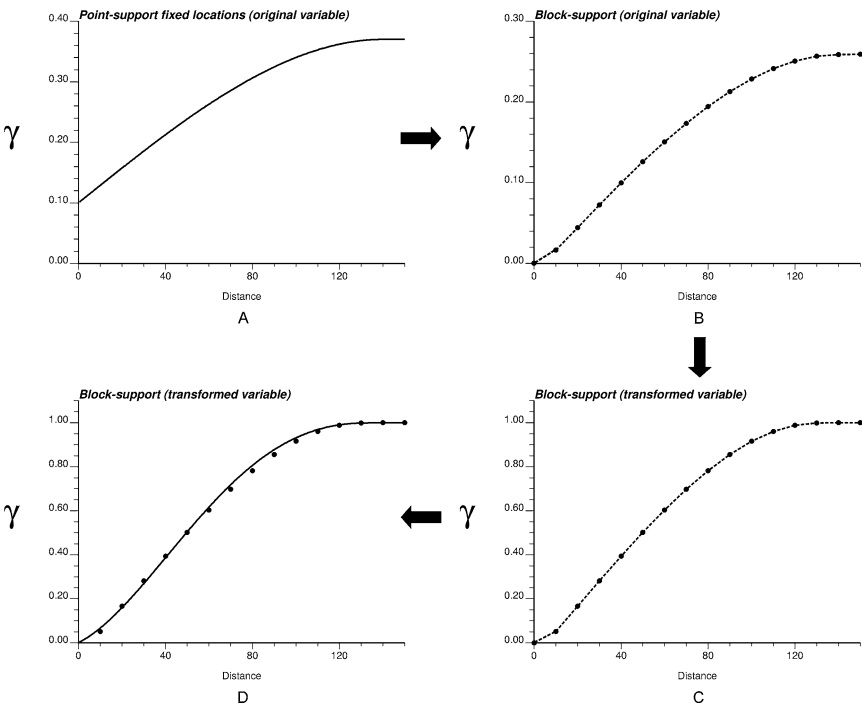
### Comments and Consequences

Equation (13) gives a new insight into the discrete Gaussian model by bringing an interpretation of the block-Gaussian variable  $Y_v$ , which proves to be proportional to the regularized variable  $Y(v)$ . To account for the change of support,

the model implicitly performs a linear averaging of the point-support Gaussian variable. This situation is approximate, unless the transformation function  $\phi$  is linear, since the linear averaging concerns the original  $Z$ -variable [Eq. (1)], not its Gaussian transform. To correct the bias that this approximation could produce on the distribution of  $Z_v$ , the model does not use the point-support transformation function  $\phi$  for back-transforming  $Y_v$ , but a different block-support transformation function  $\phi_v$  [Eq. (2)].

The previous statements are of interest for improving the modeling of the simple and cross-covariances of the Gaussian fields  $\{Y(\mathbf{x}), \mathbf{x} \in D\}$  and  $\{Y_v, v \in D\}$ . In this respect, the traditional approach consists of the following steps (Chilès and Delfiner, 1999, p. 441):

- 1) Model the covariance function (or, equivalently, the variogram) of the original point-support fixed-location variable  $\{Z(\mathbf{x}), \mathbf{x} \in D\}$  (Fig. 1A).



**Figure 1.** Traditional approach for variogram analysis in the discrete Gaussian model: A, variogram of the point-support original variable; B, discretized variogram of the regularized original variable; C, discretized and D, modeled variogram for the block-support Gaussian variable.

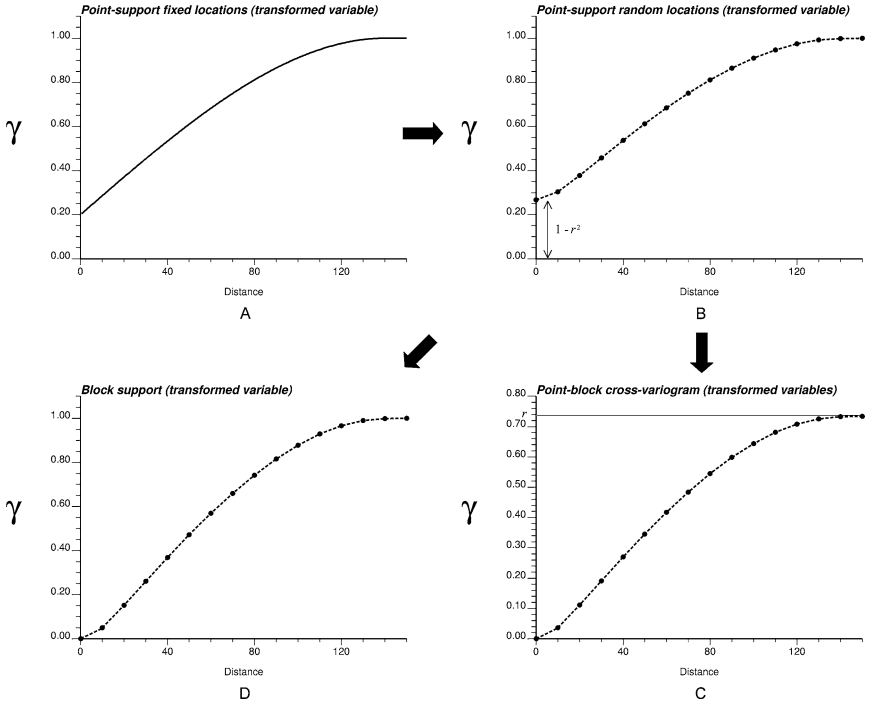
- 2) Regularize this model to the block support and obtain the covariance function (variogram) of  $\{Z(v), v \in D\}$ . In practice, calculations are made for a set of discretized lag separation vectors (Fig. 1B).
- 3) Derive the covariance function (variogram) of  $\{Y_v, v \in D\}$  for these lag vectors, by using the Hermitian expansion of the block-support transformation function  $\phi_v$  (Rivoirard, 1994, p. 89) (Fig. 1C).
- 4) Fit a model to this discretized covariance (variogram) (Fig. 1D).

This approach suffers from three weaknesses:

- It is based on the covariance model of the original point-support variable: this covariance may be poorly known, especially when dealing with long-tailed distributions (Armstrong, 1984). In addition, applications of the discrete Gaussian model (e.g. disjunctive kriging or uniform conditioning) use the  $Y$ -data, not the original  $Z$ -data: the covariance of the former is obtained in an indirect way from that of the latter and may not be suitable for the actual Gaussian data.
- The discretized covariance obtained in step 93) does not always correspond to a positive semi-definite function (Chilès and Delfiner, 1999, p. 442), which indicates that the covariance model chosen in step 1) is likely to be incompatible with the transformation function  $\phi$ , i.e. with the  $Z$ -data univariate distribution, or with the bivariate Gaussian assumption. The conditions to ensure the consistency between a covariance function and a given univariate and bivariate distribution model are still unknown to a great extent; the reader is referred to Matheron (1987, 1989) and Armstrong (1992) for further statements on this issue.
- The set of covariance models allowable for  $\{Y_v, v \in D\}$  is not known. Since this variable corresponds to a block support, some models (e.g. a pure nugget effect) are clearly not acceptable.

An alternative approach for variogram analysis consists in fitting the covariance of the point-support Gaussian field  $\{Y(\mathbf{x}), \mathbf{x} \in D\}$  with fixed locations (Fig. 2A). All the required covariance functions can then be derived from Eqs. (5) and (8) in a consistent manner (Fig. 2B, C and D), avoiding the aforementioned drawbacks. Also, working directly with the Gaussian data facilitates the analysis, as in general the covariance of these data is better known than that of the original  $Z$ -data.

Note that the change-of-support coefficient  $r$  can be obtained directly from the nugget effect component of the covariance function of  $\{Y(\mathbf{x}), \mathbf{x} \in D\}$  (Fig. 2B), or by tuning the Hermitian coefficients of the block-support transformation function in order to match the prior variance of  $Z(v)$  (Rivoirard, 1994, p. 83). These two approaches constitute a way to test the suitability of the discrete Gaussian model, as they should lead to the same value for  $r$ .



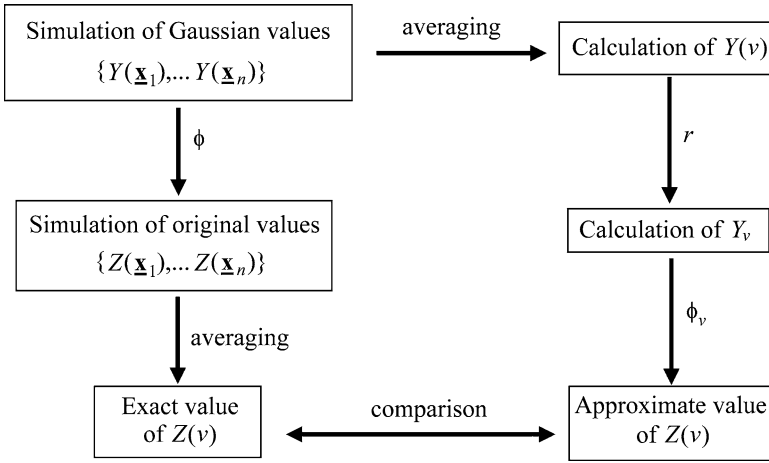
**Figure 2.** Proposed approach for variogram analysis. The point-block cross-variogram and the block-Gaussian variogram are proportional to the continuous component of the point-support randomized variogram.

### Accuracy of the Discrete Gaussian Model

In this section, a numerical exercise is proposed to test the assumptions on which the discrete Gaussian change-of-support model is based. Henceforth, the point-support model is extended to the multivariate distributions, i.e. one assumes that the point-support random-location field  $\{Y(\mathbf{x}), \mathbf{x} \in D\}$  has multivariate Gaussian distributions (this assumption is considered when resorting to the multi-gaussian kriging approach). In addition, the following hypotheses are made:

- 1) The change-of-support coefficient  $r$  is known.
- 2) The original point-support variable  $\{Z(\mathbf{x}), \mathbf{x} \in D\}$  has a lognormal univariate distribution with logarithmic variance  $\sigma^2$ , i.e. the transformation function  $\phi$  is an exponential function

$$\forall y \in \mathbb{R}, \phi(y) = \exp(\sigma y) \quad (14)$$



**Figure 3.** Workflow of the proposed test.

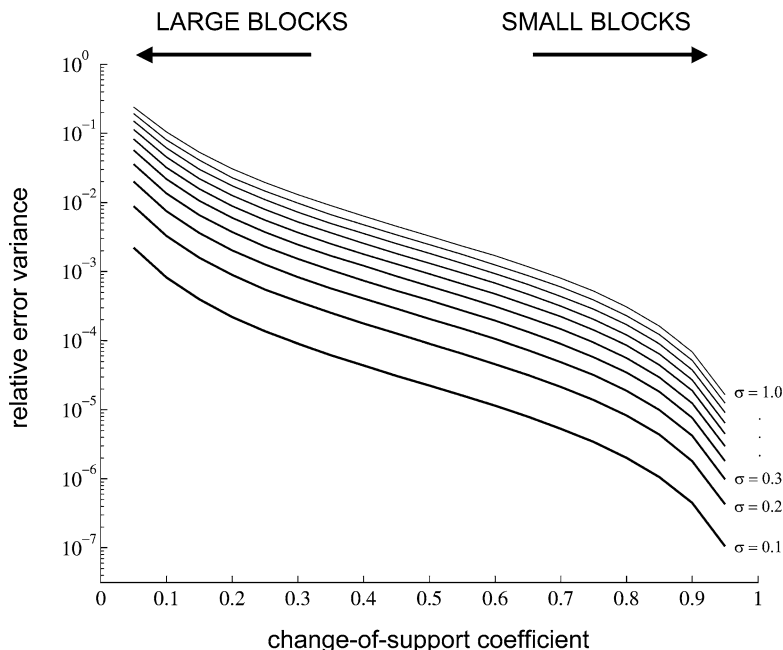
This hypothesis does not play any particular role, except that it makes calculations easier as the block-support transformation function is also an exponential function (permanence of lognormality) (Rivoirard, 1994, p. 82)

$$\forall y \in \mathbb{R}, \phi_v(y) = \exp \left\{ r\sigma y + (1 - r^2) \frac{\sigma^2}{2} \right\} \quad (15)$$

The proposed exercise consists in simulating a set of point-support Gaussian data  $\{Y(\mathbf{x}_i), i = 1, \dots, n\}$  randomly located inside block  $v$ . This can be done by using the LU decomposition method of the covariance matrix (Davis, 1987); according to Eq. (9), the entries of this matrix are equal to  $r^2$ , except the diagonal entries that are equal to one. The block-support variable  $Y(v)$  is calculated by averaging the point-support Gaussian data (this amounts to a numerical integration of the block average by Monte Carlo sampling), from which the block-Gaussian variable  $Y_v$  and the associated regularized variable  $Z(v)$  are derived [Eqs. (2), (13) and (15)].

Concurrently, the simulated Gaussian data  $\{Y(\mathbf{x}_i), i = 1, \dots, n\}$  can be back-transformed to a set of data expressed in the original scale  $\{Z(\mathbf{x}_i), i = 1, \dots, n\}$  [Eq. (2)]. Their average gives the “exact” variable  $Z(v)$  (up to the block discretization), which is compared to the variable obtained from the discrete Gaussian model assumptions (Fig. 3). A simple comparison criterion is the variance of the “error” (difference between the true and approximate variables) divided by the variance of the true variable.

For practical calculations,  $n$  is set to 500, so that the effect of discretizing the block into a finite number of points is expected to be negligible, and the variances



**Figure 4.** Variance of the difference between the exact and approximate  $Z(v)$ -variables divided by the variance of the exact variable (lognormal model with logarithmic variance  $\sigma^2$ ). From top to bottom:  $\sigma = 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2$  and  $0.1$ .

are calculated from a set of 100,000 realizations. The results, reported in Fig. 4 for ten values of the logarithmic standard deviation between 0.1 and 1.0, indicate that the relative error increases when

- i) the block support becomes voluminous (the change-of-support coefficient  $r$  decreases)
- ii) the logarithmic standard deviation increases, i.e. the coefficient of variation of the point-support original distribution increases.

In Fig. 4, this coefficient of variation varies from 0.10 ( $\sigma = 0.1$ ) to 1.31 ( $\sigma = 1$ ). In this range, the discrete Gaussian model approximation proves to be excellent when the change-of-support coefficient  $r$  is greater than 0.7 (the relative error variance is less than 0.001) and remains very good when  $r$  is greater than 0.35 (relative error variance less than 0.01). The error increases if the transformation function becomes highly nonlinear (i.e. if  $\sigma$  increases above 1.0), in which case the application of the discrete Gaussian model has to be restrained to small block supports.

## OTHER ISOFACTORIAL CHANGE-OF-SUPPORT MODELS

Let us now assume that the distributions of the pairs of transformed variables  $\{Y(\underline{\mathbf{x}}), Y(\underline{\mathbf{x}}')\}$  and  $\{Y(\underline{\mathbf{x}}), Y_{v'}\}$  have isofactorial representations (not necessarily bivariate Gaussian) for any  $\underline{\mathbf{x}}, \underline{\mathbf{x}}'$  and  $v'$  in  $D$ . The change-of-support model is characterized by two sets of functions:  $\{\chi_p^{(1)}, p \in \mathbb{N}\}$  (point-support factors) and  $\{\chi_p^{(2)}, p \in \mathbb{N}\}$  (block-support factors), such that (Matheron, 1984b)

- the factors of order 0,  $\chi_0^{(1)}$  and  $\chi_0^{(2)}$ , are constant and equal to 1;
  - for any positive integer  $p$ ,  $\chi_p^{(1)}(Y(\underline{\mathbf{x}}))$  and  $\chi_p^{(2)}(Y_{v'})$  have zero means and unit variances;
  - for any positive integers  $p$  and  $q$ , the covariances between  $\chi_p^{(1)}(Y(\underline{\mathbf{x}}))$  and  $\chi_q^{(1)}(Y(\underline{\mathbf{x}}'))$  and between  $\chi_p^{(1)}(Y(\underline{\mathbf{x}}))$  and  $\chi_q^{(2)}(Y_{v'})$  are zero, unless  $p = q$ .
- Let us put

$$\forall p \in \mathbb{N}^*, \begin{cases} T_p(\underline{\mathbf{x}}, \underline{\mathbf{x}}') = \text{cov}\{\chi_p^{(1)}(Y(\underline{\mathbf{x}})), \chi_p^{(1)}(Y(\underline{\mathbf{x}}'))\} \\ T_p(\underline{\mathbf{x}}, v') = \text{cov}\{\chi_p^{(1)}(Y(\underline{\mathbf{x}})), \chi_p^{(2)}(Y_{v'})\} \end{cases} \quad (16)$$

In the following, one assumes that the first-order factors are polynomials of degree 1, so that  $T_1(\underline{\mathbf{x}}, \underline{\mathbf{x}}')$  and  $T_1(\underline{\mathbf{x}}, v')$  are the correlation coefficients between  $Y(\underline{\mathbf{x}})$  and  $Y(\underline{\mathbf{x}}')$  and between  $Y(\underline{\mathbf{x}})$  and  $Y_{v'}$  respectively. By analogy with the discrete Gaussian model, let us define the *change-of-support coefficient* as the correlation coefficient between  $Y(\underline{\mathbf{x}})$  and  $Y_v$  when  $\underline{\mathbf{x}}$  belongs to  $v$ :

$$r = \text{corr}(Y(\underline{\mathbf{x}}), Y_v) = T_1(\underline{\mathbf{x}}, v) \quad \text{with } \underline{\mathbf{x}} \in v \quad (17)$$

### Relationships Between the Transformation Functions and Factor Covariances at Different Supports

The point-support transformation function can be expanded into the point-support factors

$$\forall y \in \mathbb{R}, \phi(y) = \phi_0 + \sum_{p=1}^{+\infty} \phi_p \chi_p^{(1)}(y) \quad (18)$$

By applying Cartier's relation, it can be established (Matheron, 1984b, p. 451) that the block-support transformation function has the following expansion

$$\forall y \in \mathbb{R}, \phi_v(y) = \phi_0 + \sum_{p=1}^{+\infty} \phi_p T_p(\underline{\mathbf{x}}, v) \chi_p^{(2)}(y) \quad \text{with } \underline{\mathbf{x}} \in v \quad (19)$$

From this identity, the block-support variance is calculated by adding the squared coefficients of order  $p \geq 1$

$$\text{var}\{Z(v)\} = \sum_{p=1}^{+\infty} \phi_p^2 T_p^2(\underline{\mathbf{x}}, v) \quad \text{with } \underline{\mathbf{x}} \in v \quad (20)$$

Also, the point-support factor covariances can be derived from the point-block factor cross-covariances (Matheron, 1984b, p. 451; Emery and Ortiz, 2005):

$$\forall \underline{\mathbf{x}} \in v, \forall \underline{\mathbf{x}}' \in v', \forall p \in \mathbb{N}^*, T_p(\underline{\mathbf{x}}, \underline{\mathbf{x}}') = \begin{cases} T_p(\underline{\mathbf{x}}, v) T_p(\underline{\mathbf{x}}, v') & \text{if } \underline{\mathbf{x}} \neq \underline{\mathbf{x}}' \\ 1 & \text{otherwise} \end{cases} \quad (21)$$

As a particular case, if  $\underline{\mathbf{x}}$  and  $\underline{\mathbf{x}}'$  belong to the same block ( $\underline{\mathbf{x}} \neq \underline{\mathbf{x}}'$ ), one has [Eq. (17)]

$$T_1(\underline{\mathbf{x}}, \underline{\mathbf{x}}') = \text{corr}\{Y(\underline{\mathbf{x}}), Y(\underline{\mathbf{x}}')\} = r^2 \quad (22)$$

### Interpretation of the Block-Transformed Variable

Let  $m$  and  $m_v$  denote the expected values of  $Y(\underline{\mathbf{x}})$  and  $Y_v$  respectively, and  $\sigma^2$  and  $\sigma_v^2$  their prior variances. According to Eqs. (6) and (22), one has

$$\text{var}\{Y(v)\} = r^2 \sigma^2 \quad (23)$$

so the correlation coefficient between  $Y(v)$  and  $Y_v$  is

$$\begin{aligned} \text{corr}\{Y(v), Y_v\} &= \frac{1}{r\sigma\sigma_v} \frac{1}{|v|} \int_v \text{cov}\{Y(\underline{\mathbf{x}}), Y_v\} d\underline{\mathbf{x}} \\ &= \frac{1}{r\sigma\sigma_v} \text{cov}\{Y(\underline{\mathbf{x}}), Y_v\} \quad \text{with } \underline{\mathbf{x}} \in v \end{aligned} \quad (24)$$

Because of Eq. (17), this correlation coefficient is equal to 1, hence the random variables  $Y(v)$  and  $Y_v$  are the same up to an affine correction

$$\frac{Y(v) - m}{r\sigma} = \frac{Y_v - m_v}{\sigma_v} \quad (25)$$

As in the discrete Gaussian model, this identity entails substantial simplifications for variogram analysis, since it allows deriving the simple and cross-covariances (or variograms) of the first-order factors from the covariance function

(variogram) of the point-support transformed variable with fixed locations (Fig. 2). In the following subsections, two examples are presented and a methodology is proposed to fit the model parameters.

### First Example: Discrete Hermitian Model

In this model, the two random fields  $\{Y(\underline{\mathbf{x}}), \underline{\mathbf{x}} \in D\}$  and  $\{Y_v, v \in D\}$  have standard Gaussian univariate distributions, while their bivariate distributions have isofactorial representations with Hermite polynomials as the factors, such that

$$\forall p \in \mathbb{N}^*, \begin{cases} T_p(\underline{\mathbf{x}}, \underline{\mathbf{x}}') = E\{R(\underline{\mathbf{x}}, \underline{\mathbf{x}}')^p\} \\ T_p(\underline{\mathbf{x}}, v') = E\{R(\underline{\mathbf{x}}, v')^p\} \end{cases} \quad (26)$$

where  $R(\underline{\mathbf{x}}, \underline{\mathbf{x}}')$  and  $R(\underline{\mathbf{x}}, v')$  are random variables in  $[-1, 1]$ . Once  $R(\underline{\mathbf{x}}, v')$  has been specified,  $R(\underline{\mathbf{x}}, \underline{\mathbf{x}}')$  is automatically defined by Eq. (21)

$$\forall \underline{\mathbf{x}} \in v, \forall \underline{\mathbf{x}}' \in v', \forall p \in \mathbb{N}^*, T_p(\underline{\mathbf{x}}, \underline{\mathbf{x}}') = E\{R(\underline{\mathbf{x}}, v)^p\}E\{R(\underline{\mathbf{x}}, v')^p\} \text{ if } \underline{\mathbf{x}} \neq \underline{\mathbf{x}}' \quad (27)$$

These terms are identified with the successive moments of a random variable  $R(\underline{\mathbf{x}}, \underline{\mathbf{x}}')$  that is the product of two independent random variables distributed like  $R(\underline{\mathbf{x}}, v)$  and  $R(\underline{\mathbf{x}}, v')$  respectively. For example, Emery and Ortiz (2005) analyze the case when  $R(\underline{\mathbf{x}}, v')$  has a beta distribution with parameters  $\{\beta\rho(\underline{\mathbf{x}}, v'), \beta(1 - \rho(\underline{\mathbf{x}}, v'))\}$  ( $\beta > 0$ ). However, this model is not realistic when  $\beta$  tends to zero since, in this case, all the coefficients  $\{T_p(\underline{\mathbf{x}}, v), p \in \mathbb{N}^*\}$  for  $\underline{\mathbf{x}} \in v$  become equal. Accordingly, the transformation functions at point and block supports [Eqs. (18) and (19)] are linearly related. This means that, at the global scale, the change-of-support model is an affine correction, which lacks realism as it does not account for a symmetrization of the univariate distribution when the support increases. The next subsection presents a more general model that does not suffer from this limitation.

### Second Example: Discrete Laguerrian Model

#### *Presentation of the Model*

Here, the two random fields  $\{Y(\underline{\mathbf{x}}), \underline{\mathbf{x}} \in D\}$  and  $\{Y_v, v \in D\}$  have standard gamma univariate distributions, with respective shape parameters (variances)  $\alpha$  and  $\alpha_v$  (with  $\alpha_v \geq \alpha$ ). Their bivariate distributions have isofactorial representations with Laguerre polynomials as the factors and are characterized by the following

sets of factor covariances (Hu, 1988, p. 69)

$$\forall p \in \mathbb{N}^*, \begin{cases} T_p(\underline{\mathbf{x}}, \underline{\mathbf{x}}') = E\{R(\underline{\mathbf{x}}, \underline{\mathbf{x}}')^p\} \\ T_p(\underline{\mathbf{x}}, v') = \sqrt{\frac{\Gamma(\alpha + p)\Gamma(\alpha_v)}{\Gamma(\alpha_v + p)\Gamma(\alpha)}} E\{R(\underline{\mathbf{x}}, v')^p\} \end{cases} \quad (28)$$

where  $R(\underline{\mathbf{x}}, \underline{\mathbf{x}}')$  and  $R(\underline{\mathbf{x}}, v')$  are random variables lying in  $[0, 1]$ . Again, it is convenient to construct the model by reference to the random variable  $R(\underline{\mathbf{x}}, v')$  that defines the joint distribution of  $Y(\underline{\mathbf{x}})$  and  $Y_{v'}$ . The point-support bivariate distributions are then characterized by the following set of factor covariances [Eq. (21)]

$$\begin{aligned} \forall \underline{\mathbf{x}} \in v, \forall \underline{\mathbf{x}}' \in v', \forall p \in \mathbb{N}^*, \\ T_p(\underline{\mathbf{x}}, \underline{\mathbf{x}}') = \frac{\Gamma(\alpha + p)\Gamma(\alpha_v)}{\Gamma(\alpha_v + p)\Gamma(\alpha)} E\{R(\underline{\mathbf{x}}, v)^p\} E\{R(\underline{\mathbf{x}}, v')^p\} \quad \text{if } \underline{\mathbf{x}} \neq \underline{\mathbf{x}}' \end{aligned} \quad (29)$$

These terms are identified with the successive moments of a random variable  $R(\underline{\mathbf{x}}, \underline{\mathbf{x}}')$  which is the product of three independent random variables, the first one having a beta distribution with parameters  $\{\alpha, \alpha_v - \alpha\}$  and the two last ones having the same distributions as  $R(\underline{\mathbf{x}}, v)$  and  $R(\underline{\mathbf{x}}, v')$ . This ensures that the point-support bivariate distributions are consistent from a mathematical viewpoint [Eq. (28)]. It is noteworthy that the point-support variable  $\{Y(\underline{\mathbf{x}}), \underline{\mathbf{x}} \in D\}$  cannot have diffusion-type bivariate gamma distributions (i.e. such that  $R(\underline{\mathbf{x}}, \underline{\mathbf{x}}')$  has a zero variance for any  $\underline{\mathbf{x}}$  and  $\underline{\mathbf{x}}'$ ) unless  $\alpha_v = \alpha$  and the variance of  $R(\underline{\mathbf{x}}, v')$  is equal to zero for any  $\underline{\mathbf{x}}$  and  $v'$ .

To make the model workable, henceforth the random variable  $R(\underline{\mathbf{x}}, v')$  is supposed to have a beta distribution with parameters  $\{\beta\rho(\underline{\mathbf{x}}, v')/\alpha, \beta(1 - \rho(\underline{\mathbf{x}}, v')/\alpha)\}$  ( $\beta > 0$ )

$$\forall p \in \mathbb{N}^*, E\{R(\underline{\mathbf{x}}, v')^p\} = \frac{\Gamma(\beta)\Gamma\left(\frac{\beta}{\alpha}\rho(\underline{\mathbf{x}}, v') + p\right)}{\Gamma\left(\frac{\beta}{\alpha}\rho(\underline{\mathbf{x}}, v')\right)\Gamma(\beta + p)} \quad (30)$$

Following Equation (28), one has

$$\forall \underline{\mathbf{x}}, v' \in D, \text{corr}\{Y(\underline{\mathbf{x}}), Y_{v'}\} = T_1(\underline{\mathbf{x}}, v') = \frac{\rho(\underline{\mathbf{x}}, v')}{\sqrt{\alpha\alpha_v}} \quad (31)$$

It ensues that  $\rho(\underline{\mathbf{x}}, v')$  is the covariance between  $Y(\underline{\mathbf{x}})$  and  $Y_{v'}$ , in accordance with the definition given in Eq. (4). As a particular case of Eq. (31), for  $\underline{\mathbf{x}} \in v$ , one

finds

$$r = T_1(\underline{\mathbf{x}}, v) = \frac{\rho(\underline{\mathbf{x}}, v)}{\sqrt{\alpha\alpha_v}} \quad \text{with } \underline{\mathbf{x}} \in v \quad (32)$$

### *Parameters Inference*

The following approach is proposed to determine the model parameters:

- 1) Choose the shape parameters  $\alpha$  and  $\alpha_v$ . To facilitate this choice, several criteria have been proposed by Demange and others (1987, p. 200), Hu (1988, p. 16 and 71) and Chilès and Liao (1993, p. 1056).
- 2) Transform the original data to gamma score values with shape parameter  $\alpha$ . Obtain the point-support transformation function  $\phi$  and its expansion into Laguerre polynomials [Eq. (18)].
- 3) Model the covariance function of the point-support fixed-location transformed field  $\{Y(\mathbf{x}), \mathbf{x} \in D\}$  and determine the covariance function of the field with random locations  $\{Y(\underline{\mathbf{x}}), \underline{\mathbf{x}} \in D\}$  [Eq. (5)]. The standardization of the latter gives the covariance function  $T_1(\underline{\mathbf{x}}, \underline{\mathbf{x}}')$  of the first-degree point-support factor.
- 4) Derive the change-of-support coefficient  $r$  [Eq. (22)] and function  $\rho(\underline{\mathbf{x}}, v')$  [Eqs. (17), (21) and (31)]

$$\forall \underline{\mathbf{x}}' \in v', T_1(\underline{\mathbf{x}}, \underline{\mathbf{x}}') = r \frac{\rho(\underline{\mathbf{x}}, v')}{\sqrt{\alpha\alpha_v}} \quad (33)$$

- 5) Choose the value of  $\beta$  so as to honor the block-support variance [Eqs. (20), (28), (30) and (32)]

$$\text{var}\{Z(v)\} = \frac{\Gamma(\alpha_v)\Gamma^2(\beta)}{\Gamma(\alpha)\Gamma^2\left(\beta r \sqrt{\frac{\alpha_v}{\alpha}}\right)} \sum_{p=1}^{+\infty} \phi_p^2 \frac{\Gamma(\alpha + p)\Gamma^2\left(\beta r \sqrt{\frac{\alpha_v}{\alpha}} + p\right)}{\Gamma(\alpha_v + p)\Gamma^2(\beta + p)} \quad (34)$$

The left-hand side of this equality can be calculated after modeling the variogram of the original point-support  $Z$ -data. This variogram model may be poorly fitted if the data have a highly skewed univariate distribution, in which case the variance of  $Z(v)$  and, subsequently, the calculated value of  $\beta$  are known with low precision. An alternative is to perform a sensitivity analysis, by defining several plausible values for  $\text{var}\{Z(v)\}$  and by examining their impact on the model parameters and on the final results (calculation of recoverable resources, delineation of polluted areas, etc.).

Note that the block-support shape parameter  $\alpha_v$  cannot be any positive real number. On the one hand, the construction of the model [Eq. (28)] requires that  $\alpha_v$  is greater than or equal to the point-support shape parameter  $\alpha$ . This condition is compatible with the central limit theorem, since the univariate distribution of the regularized variable  $Y(v)$  (hence, according to Eq. (25), of the block-transformed variable  $Y_v$ ) is less skewed than that of the point-support variable  $Y(\underline{\mathbf{x}})$ . On the other hand, the parameters of the beta distribution of  $R(\underline{\mathbf{x}}, v')$  must be nonnegative, hence

$$\forall \underline{\mathbf{x}}, v' \in D, \rho(\underline{\mathbf{x}}, v') \leq \alpha \quad (35)$$

By using Equation (32), it comes

$$\alpha_v \leq \frac{\alpha}{r^2} \quad (36)$$

The value of  $\alpha_v$  must also ensure the existence of a positive real number  $\beta$  solution to Equation (34).

### *Validation of the Model*

A simple way to validate the previous Laguerrian model is to examine the madogram (first-order variogram) of the point-support transformed random field (Chilès and Liao, 1993, p. 1057)

$$\gamma_1(\underline{\mathbf{x}}, \underline{\mathbf{x}}') = \frac{1}{2} E\{|Y(\underline{\mathbf{x}}) - Y(\underline{\mathbf{x}}')|\} \quad (37)$$

Under the assumption that  $\{Y(\underline{\mathbf{x}}), Y(\underline{\mathbf{x}}')\}$  has an isofactorial Laguerrian bivariate distribution, one has (Emery, 2005a, p. 174):

$$\gamma_1(\underline{\mathbf{x}}, \underline{\mathbf{x}}') = \frac{1}{\Gamma(\alpha)\sqrt{\pi}} \Gamma\left(\alpha + \frac{1}{2}\right) E\{\sqrt{1 - R(\underline{\mathbf{x}}, \underline{\mathbf{x}}')}\} \quad (38)$$

By expanding the square root into a binomial series (Abramowitz, 1972, p. 15), it comes [Eqs. (29) and (30)]

$$\begin{aligned} \gamma_1(\underline{\mathbf{x}}, \underline{\mathbf{x}}') &= \frac{1}{\Gamma(\alpha)\sqrt{\pi}} \Gamma\left(\alpha + \frac{1}{2}\right) \left\{ 1 + \sum_{p \geq 1} \frac{\Gamma(p - 1/2)}{p! \Gamma(-1/2)} T_p(\underline{\mathbf{x}}, \underline{\mathbf{x}}') \right\} \\ &= \frac{1}{\Gamma(\alpha)\sqrt{\pi}} \Gamma\left(\alpha + \frac{1}{2}\right) {}_4F_3\left(-\frac{1}{2}, \alpha, \beta r \sqrt{\frac{\alpha_v}{\alpha}}, \frac{\beta}{\alpha} \rho(\underline{\mathbf{x}}, v'); \alpha_v, \beta, \beta; 1\right) \end{aligned} \quad (39)$$

where  ${}_4F_3$  is a generalized hypergeometric function (Bailey, 1964) and  $v'$  is the block containing  $\mathbf{x}'$ . Because the available gamma score data refer to the random field with fixed locations  $\{Y(\mathbf{x}), \mathbf{x} \in D\}$ , this formula is valid only when the distance between  $\mathbf{x}$  and  $\mathbf{x}'$  is much larger than the block support diameter, for which the randomization of the data locations within the blocks has almost no effect.

## CONCLUSIONS

This paper focused on isofactorial change-of-support models, which are used in geostatistical applications for estimating the local distributions of block-support values. The implications of the model hypotheses have been examined and new conditions have been established to ensure the internal consistency of the change-of-support model. In particular, a simplified approach has been proposed for variogram analysis, which directly works with the transformed data that are of interest for nonlinear kriging methods (multigaussian kriging, disjunctive kriging, uniform conditioning). This constitutes an important difference with the traditional approach, for which the variogram model is obtained in an indirect way and relies on strong assumptions about the univariate and bivariate distributions at point and block supports.

In general, the discrete Gaussian model implies mild approximations, unless the original data have a highly skewed histogram and the change-of-support coefficient is small, indicating an important support effect. However, it may lack flexibility, since the change-of-support coefficient must satisfy two different relationships simultaneously. An alternative is the discrete Laguerrian model, based on the gamma distribution, which uses more parameters and remains quite simple to implement. The presented work can be extended to other isofactorial change-of-support models with polynomial factors, such as the discrete models or the mixed discrete-continuous models based on the negative binomial distribution (Matheron, 1984b; Raspa and Bruno, 1989).

## ACKNOWLEDGMENT

The author acknowledges the sponsoring by Codelco-Chile for supporting this research and the anonymous reviewers for their comments on an earlier version of the manuscript.

## REFERENCES

- Abramowitz, M., 1972, Elementary analytical methods, in Abramowitz, M., and Stegun, I., eds., Handbook of mathematical functions: Dover Publications Inc., New York, p. 9–63.
- Armstrong, M., 1984, Common problems seen in variograms: Math. Geol. v. 16, no. 3, p. 305–313.

- Armstrong, M., 1992, Positive definiteness is not enough: *Math. Geol.* v. 24, no. 1, p. 135–143.
- Bailey, W. N., 1964, Generalized hypergeometric series: Stechert-Hafner Service Agency, New York, 108 p.
- Chilès, J. P., 1976, How to adapt kriging to non-classical problems: Three case studies, *in* Guarascio, M., David, M., and Huijbregts, C. J., eds., *Advanced geostatistics in the mining industry*: Reidel, Dordrecht, p. 69–89.
- Chilès, J. P., and Delfiner, P., 1999, *Geostatistics: Modeling spatial uncertainty*: Wiley, New York, 695 p.
- Chilès, J. P., and Liao, H. T., 1993, Estimating the recoverable reserves of gold deposits: Comparison between disjunctive kriging and indicator kriging, *in* Soares, A., ed., *Geostatistics Tróia'92*: Kluwer Academic, Dordrecht, Vol. 2, p. 1053–1064.
- Davis, M. W., 1987, Production of conditional simulations via the LU triangular decomposition of the covariance matrix: *Math. Geol.* v. 19, no. 2, p. 91–98.
- Demange, C., Lajaunie, C., Lantuéjoul, C., and Rivoirard, J., 1987, Global recoverable reserves: Testing various changes of support models on uranium data, *in* Matheron, G., and Armstrong, M., eds., *Geostatistical Case Studies*: Reidel, Dordrecht, p. 187–208.
- Demougeot-Renard, H., and de Fouquet, C., 2004, Geostatistical approach for assessing soil volumes requiring remediation: Validation using lead-polluted soils underlying a former smelting works: *Environmental Science & Technology*, v. 38, no. 19, p. 5120–5126.
- Emery, X., 2005a, Variograms of order  $\omega$ : A tool to validate a bivariate distribution model: *Math. Geol.* v. 37, no. 2, p. 163–181.
- Emery, X., 2005b, Simple and ordinary multigaussian kriging for estimating recoverable reserves: *Math. Geol.* v. 37, no. 3, p. 295–319.
- Emery, X., and Ortiz, J. M., 2005, Internal consistency and inference of change-of-support isofactorial models, *in* Leuangthong, O., and Deutsch, C.V., eds., *Geostatistics Banff 2004*: Springer, Dordrecht, Vol. 2, p. 1057–1066.
- Emery, X., and Soto Torres, J. F., 2005, Models for support and information effects: A comparative study: *Math. Geol.* v. 37, no. 1, p. 49–68.
- Gelfand, A. E., Zhu, L., and Carlin, B. P., 2001, On the change of support problem for spatio-temporal data: *Biostatistics*, v. 2, no. 1, p. 31–45.
- Guibal, D., and 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., and Maréchal, A., eds., *Geostatistics for natural resources characterization*: Reidel, Dordrecht, Vol. 1, p. 435–448.
- Hu, L. Y., 1988, *Mise en oeuvre du modèle gamma pour l'estimation des distributions spatiales*: Unpubl. Doctoral Thesis, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau, 142 p.
- Journel, A. G., and Kyriakidis, P. C., 2004, *Evaluation of mineral reserves: A simulation approach*: Oxford University Press, Oxford, 216 p.
- Maréchal, A., 1984, Recovery estimation: A review of models and methods, *in* Verly, G., David, M., Journel, A. G., and Maréchal, A., eds., *Geostatistics for natural resources characterization*: Reidel, Dordrecht, Vol. 1, p. 385–420.
- Matheron, G., 1976a, A simple substitute for conditional expectation: The disjunctive kriging, *in* Guarascio, M., David, M., and Huijbregts, C. J., eds., *Advanced geostatistics in the mining industry*: Reidel, Dordrecht, p. 221–236.
- Matheron, G., 1976b, Forecasting block grade distributions: the transfer functions, *in* Guarascio, M., David, M., and Huijbregts, C. J., eds., *Advanced Geostatistics in the Mining Industry*: Reidel, Dordrecht, p. 237–251.
- Matheron, G., 1984a, The selectivity of the distributions and the second principle of geostatistics, *in* Verly, G., David, M., Journel, A. G., and Maréchal, A., eds., *Geostatistics for natural resources characterization*: Reidel, Dordrecht, Vol. 1, p. 421–433.

- Matheron, G., 1984b, Isofactorial models and change of support, *in* Verly, G., David, M., Journel, A. G., and Maréchal A., eds., *Geostatistics for natural Resources characterization*: Reidel, Dordrecht, Vol. 1, p. 449–467.
- Matheron, G., 1987, Suffit-il, pour une covariance, d'être de type positif?: *Sciences de la Terre, Série informatique géologique*, v. 26, p. 51–66.
- Matheron, G., 1989, The internal consistency of models in geostatistics, *in* Armstrong, M., ed., *Geostatistics*: Kluwer Academic, Dordrecht, Vol. 1, p. 21–38.
- Raspa, G., and Bruno, R., 1989, An application of disjunctive kriging: Using the negative binomial model with different change of support models, *in* Armstrong, M., ed., *Geostatistics*: Kluwer Academic, Dordrecht, Vol. 2, p. 935–945.
- Remacre, A. Z., 1989, Uniform conditioning versus indicator kriging: A comparative study with the actual data, *in* Armstrong, M., ed., *Geostatistics*: Kluwer Academic, Dordrecht, Vol. 2, p. 947–960.
- Rivoirard, J., 1994, *Introduction to disjunctive kriging and non-linear geostatistics*: Clarendon Press, Oxford, 181 p.
- Webster, R., 1991, Local disjunctive kriging of soil properties with change of support: *Jour. of Soil Sci.* v. 42, no. 2, p. 301–318.
- Webster, R., and Oliver, M. A., 2001, *Geostatistics for environmental scientists*: Wiley, New York, 271 p.