Computers & Geosciences 38 (2012) 136–144

Contents lists available at ScienceDirect

Computers & Geosciences

journal homepage: www.elsevier.com/locate/cageo

Cokriging random fields with means related by known linear combinations ☆

Xavier Emery [a,b,*]

[a] Department of Mining Engineering, University of Chile, Avenida Tupper 2069, Santiago, Chile
[b] ALGES Laboratory, Advanced Mining Technology Center, University of Chile, Santiago, Chile

ARTICLE INFO

Article history:
Received 28 January 2011
Received in revised form
7 June 2011
Accepted 8 June 2011
Available online 21 June 2011

Keywords:
Spatial prediction
Vector random field
Multivariate geostatistics
Linear coregionalization model
Standardized ordinary cokriging

ABSTRACT

Traditional approaches to predict a second-order stationary vector random field include simple and ordinary cokriging, depending on whether or not the mean values of the vector components are assumed to be known. This paper explores a variant of cokriging, in which the mean values of the vector components are related by linear combinations with known coefficients. Equations for the cokriging predictor and for the variance–covariance matrix of prediction errors are presented. A set of computer programs is provided and illustrated with applications to mineral resources evaluation, in which the proposed cokriging variant compares favorably with traditional approaches.

1. Introduction

Cokriging is a widely used technique for the spatial prediction of coregionalized variables. Application fields in the geosciences include mineral resources evaluation (Journel and Huijbregts, 1978; Pan et al., 1993), petroleum reservoir modeling (Xu et al., 1992; Hohn, 1999), geochemistry (Wackernagel, 1988), groundwater hydrology (Ahmed and De Marsily, 1987; Kitanidis, 1997), and soil sciences (Yates and Warrick, 1987; Stein et al., 1988).

Considering the coregionalized variables as one realization of a second-order stationary vector random field, the main variants are simple and ordinary cokriging, depending on whether or not the mean of the vector random field is assumed to be known (Myers, 1982; Ver Hoef and Cressie, 1993; Wackernagel, 2003). Nevertheless, a seeming issue of ordinary cokriging is the underweighting of the covariates (the weights of the covariates sum to zero and are often of small magnitude), as though these covariates were of little significance for prediction (Goovaerts, 1998; Hohn, 1999). This underweighting may be explained because the mean of the vector random field (a first-order moment) is assumed to be constant and unknown, while only the spatial correlation structure (direct and cross covariance functions, i.e., second-order moments) is known. In practice, this assumption may be too severe and cause a loss of information, making covariates less relevant for the prediction than they should be.

More weight can be assigned to the covariates by modifying the constraint on the sum of ordinary cokriging weights. One such variant is the so-called "rescaled" or "standardized" ordinary cokriging (Isaaks and Srivastava, 1989; Goovaerts, 1998), in which the weights of all the variables sum to one. However, this new constraint yields unbiased predictions only if all the components of the vector random field have the same constant mean, which is a quite specific situation (e.g., when these components represent the same quantity measured with different devices or on different supports) (Goovaerts, 1997; Pebesma, 2004; Papritz, 2008).

There are many public packages available for cokriging second-order stationary vector random fields, which differ in the programming languages, allowed space dimensions, number of vector components, and number of nested structures used for spatial correlation modeling (Carr et al., 1985; Marcotte, 1991; Deutsch and Journel, 1992; Pan et al., 1992; Pebesma and Wesseling, 1998; Pebesma, 2004). However, these packages are restricted to simple, ordinary, and standardized ordinary cokriging.

In this context, the aim of this paper is threefold: (i) to present cokriging in a more general framework, in which the mean values of the vector components are related by linear combinations with known coefficients (even though these mean values are unknown); (ii) to provide practitioners with computer programs; (iii) to present case studies in mineral resources evaluation, in order to demonstrate the applicability and versatility of the proposed programs.

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

* Corresponding author at: Department of Mining Engineering, University of Chile, Avenida Tupper 2069, Santiago, Chile.
Tel.: +1 56 2 978 4498; fax: +1 56 2 978 4985.
E-mail address: xemery@ing.uchile.cl

0098-3004/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.cageo.2011.06.001

2. Cokriging vector components with linearly dependent means

2.1. Assumptions

Consider a second-order stationary vector random field with N real-valued components defined in $\mathbb{R}^d$, $\mathbf{Z}=(Z_1,\ldots,Z_N)^T$, and denote by $\mathbf{m}=(m_1,\ldots,m_N)^T$ its mean vector (constant in space) and by $\mathbf{C}(\mathbf{h})=(C_{ij}(\mathbf{h})$: $i=1,\ldots,N$; $j=1,\ldots,N)$ its covariance matrix for a given lag separation vector $\mathbf{h}\in\mathbb{R}^d$.

Furthermore, let us assume that the mean values satisfy linear relationships of the form:

$$\forall k\in\{1,\ldots,K\},\ \sum_{i=1}^{N} a_{k,i}m_i=b_k \tag{1}$$

for some known real-valued vector $\mathbf{b}=(b_1,\ldots,b_K)^T$ and known matrix $\mathbf{A}=(a_{k,i}$: $k=1,\ldots,K$; $i=1,\ldots,N)$ of rank $K\leq N$. Since $\mathbf{A}$ is a full-rank matrix, its rows are linearly independent, meaning that there is no redundancy or inconsistency between the relationships in Eq. (1).

Up to an eventual reordering of the components of $\mathbf{Z}$, Eq. (1) is equivalent to

$$\mathbf{A}\,\mathbf{m}=\mathbf{b} \tag{2}$$

with $\mathbf{A}=(\mathbf{A}_1\ \mathbf{A}_2)$, where $\mathbf{A}_1$ is a $K\times(N-K)$ matrix and $\mathbf{A}_2$ is a $K\times K$ matrix of rank K. Without loss of generality, one can assume that $\mathbf{b}=\mathbf{0}$ (it suffices to work with the shifted vector random field $\mathbf{Z}-\boldsymbol{\delta}$, where $\boldsymbol{\delta}$ is a vector such that $\mathbf{A}\boldsymbol{\delta}=\mathbf{b}$, instead of $\mathbf{Z}$).

Accordingly, the last K components of $\mathbf{m}$ (subvector $\mathbf{m}_2$) can be expressed as a function of the first $(N-K)$ components (subvector $\mathbf{m}_1$):

$$\mathbf{m}_2=\boldsymbol{\Omega}\,\mathbf{m}_1 \tag{3}$$

with $\boldsymbol{\Omega}=-\mathbf{A}_2^{-1}\mathbf{A}_1$. Equivalently,

$$\mathbf{m}=\mathbf{F}\,\mathbf{m}_1 \text{ with } \mathbf{F}=\begin{pmatrix}\mathbf{I}_{N-K}\\ \boldsymbol{\Omega}\end{pmatrix} \tag{4}$$

where $\mathbf{I}_{N-K}$ is the identity matrix of size $(N-K)\times(N-K)$.

2.2. Cokriging equations

In the following, it is of interest to predict $\mathbf{Z}$ at a given location $\mathbf{x}_0\in\mathbb{R}^d$, given the values of one or more components of $\mathbf{Z}$ known at surrounding locations $\{\mathbf{x}_\alpha$: $\alpha=1,\ldots,n\}$.

2.2.1. Matrix formulation of cokriging equations

In the first instance, let us suppose that all the components of the vector random field are known at all the data locations (*isotopic sampling*). In such a case, the system of cokriging equations can be written in a simple matrix form. The predictor of $\mathbf{Z}(\mathbf{x}_0)$ is

$$\mathbf{Z}^*(\mathbf{x}_0)=\sum_{\alpha=1}^{n}\boldsymbol{\Lambda}_\alpha^{\mathrm{T}}\mathbf{Z}(\mathbf{x}_\alpha) \tag{5}$$

where $\{\boldsymbol{\Lambda}_\alpha,\ \alpha=1,\ldots,n\}$ are $N\times N$ matrices of weights to be determined. Since these weights actually depend on location $\mathbf{x}_0$, the complete notations should be $\boldsymbol{\Lambda}_\alpha(\mathbf{x}_0)$; however, the shortened notation $\boldsymbol{\Lambda}_\alpha$ will be used hereafter for the sake of simplicity.

The expectation of the prediction error is

$$E\{\mathbf{Z}^*(\mathbf{x}_0)-\mathbf{Z}(\mathbf{x}_0)\}=\left(\sum_{\alpha=1}^{n}\boldsymbol{\Lambda}_\alpha^{\mathrm{T}}-\mathbf{I}_N\right)\mathbf{m}=\left(\sum_{\alpha=1}^{n}\boldsymbol{\Lambda}_\alpha^{\mathrm{T}}-\mathbf{I}_N\right)\mathbf{F}\mathbf{m}_1 \tag{6}$$

where $\mathbf{I}_N$ is the identity matrix of size $N\times N$. Since $\mathbf{m}_1$ is a vector of free components that can take any real value, the only way to ensure that the predictor is unbiased (i.e., that the expected error is zero) is to impose the following constraint:

$$\left(\sum_{\alpha=1}^{n}\boldsymbol{\Lambda}_\alpha^{\mathrm{T}}-\mathbf{I}_N\right)\mathbf{F}=\mathbf{0} \tag{7}$$

The variance–covariance matrix of the prediction error is

$$\begin{aligned}\boldsymbol{\Sigma}&=E\{[\mathbf{Z}^*(\mathbf{x}_0)-\mathbf{Z}(\mathbf{x}_0)][\mathbf{Z}^*(\mathbf{x}_0)-\mathbf{Z}(\mathbf{x}_0)]^T\}\\&=E\left\{\left[\sum_{\alpha=1}^{n}\boldsymbol{\Lambda}_\alpha^{\mathrm{T}}\mathbf{Z}(\mathbf{x}_\alpha)-\mathbf{Z}(\mathbf{x}_0)\right]\left[\sum_{\alpha=1}^{n}\boldsymbol{\Lambda}_\alpha^{\mathrm{T}}\mathbf{Z}(\mathbf{x}_\alpha)-\mathbf{Z}(\mathbf{x}_0)\right]^T\right\}\\&=\sum_{\alpha=1}^{n}\sum_{\beta=1}^{n}\boldsymbol{\Lambda}_\alpha^{\mathrm{T}}\mathbf{C}(\mathbf{x}_\alpha-\mathbf{x}_\beta)\boldsymbol{\Lambda}_\beta-\sum_{\alpha=1}^{n}\mathbf{C}(\mathbf{x}_0-\mathbf{x}_\alpha)\boldsymbol{\Lambda}_\alpha\\&\quad-\sum_{\alpha=1}^{n}\boldsymbol{\Lambda}_\alpha^{\mathrm{T}}\mathbf{C}(\mathbf{x}_\alpha-\mathbf{x}_0)+\mathbf{C}(\mathbf{x}_0-\mathbf{x}_0)\end{aligned} \tag{8}$$

where for $\alpha,\ \beta\in\{0,\ldots,n\}$, $\mathbf{C}(\mathbf{x}_\alpha-\mathbf{x}_\beta)$ is the $N\times N$ matrix whose generic term is $C_{ij}(\mathbf{x}_\alpha-\mathbf{x}_\beta)$.

It remains to minimize the variances of the prediction errors of the components of $\mathbf{Z}$ (diagonal terms of $\boldsymbol{\Sigma}$), under the unbiasedness constraint (Eq. (7)). All calculations done, one obtains the following system of equations:

$$\begin{pmatrix}\mathbf{C}(\mathbf{x}_1-\mathbf{x}_1) & \cdots & \mathbf{C}(\mathbf{x}_1-\mathbf{x}_n) & \mathbf{F}\\ \vdots & \ddots & \vdots & \vdots\\ \mathbf{C}(\mathbf{x}_n-\mathbf{x}_1) & \cdots & \mathbf{C}(\mathbf{x}_n-\mathbf{x}_n) & \mathbf{F}\\ \mathbf{F}^T & \cdots & \mathbf{F}^T & \mathbf{0}\end{pmatrix}\begin{pmatrix}\boldsymbol{\Lambda}_1\\ \vdots\\ \boldsymbol{\Lambda}_n\\ \mathbf{M}\end{pmatrix}=\begin{pmatrix}\mathbf{C}\ (\mathbf{x}_1-\mathbf{x}_0)\\ \vdots\\ \mathbf{C}(\mathbf{x}_n-\mathbf{x}_0)\\ \mathbf{F}^T\end{pmatrix} \tag{9}$$

where $\mathbf{M}$ is a $(N–K)\times N$ matrix of Lagrange multipliers and $\mathbf{0}$ a $(N–K)\times(N–K)$ matrix of zeros.

The variance–covariance matrix of the prediction errors (Eq. (8)) simplifies into

$$\boldsymbol{\Sigma}=\mathbf{C}(\mathbf{x}_0-\mathbf{x}_0)-\sum_{\alpha=1}^{n}\boldsymbol{\Lambda}_\alpha^T[\mathbf{C}(\mathbf{x}_\alpha-\mathbf{x}_0)+\mathbf{F}\mathbf{M}] \tag{10}$$

Note that, in the isotopic sampling case, the cokriging matrix in Eq. (9) is singular if the components of $\mathbf{Z}$ are linearly dependent (Appendix A), for instance if these components sum to one, as found when dealing with regionalized compositions or with class indicators (Tolosana-Delgado et al., 2008). In this case, one or more component(s) must be omitted during cokriging, in order to remove the linear dependence.

In case of a *heterotopic sampling* (i.e., when one or more components of $\mathbf{Z}$ are unknown at some data locations), the previous equations (Eqs. (5)–(10)) must be modified by removing the rows and/or columns of $\mathbf{C}(\mathbf{x}_\alpha-\mathbf{x}_\beta)$, $\boldsymbol{\Lambda}_\alpha$, and $\mathbf{F}$ corresponding to missing data values.

2.2.2. Special cases

Let us note a few specific cases of the former cokriging system:

- $K=N$: simple cokriging (Myers, 1982; Wackernagel, 2003). In this case, $\boldsymbol{\Omega}$, $\mathbf{F}$, and $\mathbf{M}$ are empty matrices.
- $K=0$: ordinary cokriging (Myers, 1982; Wackernagel, 2003). In this case, $\boldsymbol{\Omega}$ is an empty matrix and $\mathbf{F}$ is the $N\times N$ identity matrix.
- $K=N-1$ with $\boldsymbol{\Omega}$ a $(N–1)\times 1$ matrix of ones: rescaled or standardized ordinary cokriging, where all the components of $\mathbf{Z}$ have the same unknown mean value (Isaaks and Srivastava, 1989; Goovaerts, 1998). In this case, $\mathbf{F}$ is a vector of ones.
- $K\in\{1,\ldots,N–1\}$ with $\boldsymbol{\Omega}$ a matrix of zeros: mixed cokriging, in which the K last components of $\mathbf{Z}$ have known means, while the first $N–K$ components have unknown means.

2.2.3. Generalizations

The previous cokriging equations can be generalized in several ways:

- Vector random field $\mathbf{Z}$ with *nonstationary covariance functions*. This situation can be handled using, in Eqs. (9) and (10), covariance matrices $\mathbf{C}(\mathbf{x}_\alpha,\mathbf{x}_\beta)$ that separately depend on the two spatial locations $\{\mathbf{x}_\alpha, \mathbf{x}_\beta\}$ $(\alpha, \beta=0,\ldots,n)$. Several approaches for inferring and modeling nonstationary covariance functions are described by Higdon et al. (1999), Fuentes (2001), and Sampson et al. (2001), among others.
- *Block cokriging*, in which the objective is to predict the spatial average of $\mathbf{Z}$ over a block with a support larger than that of the data. In Eqs. (9) and (10), it suffices to replace $\mathbf{C}(\mathbf{x}_\alpha-\mathbf{x}_0)$ and $\mathbf{C}(\mathbf{x}_0-\mathbf{x}_0)$ by average matrices calculated by letting $\mathbf{x}_0$ discretize the block (Myers, 1984).
- *Cokriging with a moving neighborhood*: The previous equations (Eqs. (5)–(10)) are valid when working in a moving neighborhood, provided that the same set of data is used to predict all the components of $\mathbf{Z}$ at the target location or target block. Otherwise, one can predict only one component at a time, so that several cokriging runs are necessary to predict all the components.
- *Cokriging with variograms*: Subject to a symmetry condition, the cokriging equations (Eqs. (9) and (10)) can be rewritten using direct and cross variograms instead of direct and cross covariance functions (Myers, 1982; Wackernagel, 2003). Explicitly, by assuming that the covariance matrix $\mathbf{C}(\mathbf{x}_\alpha-\mathbf{x}_\beta)$ is symmetric for any $\alpha, \beta\in\{0,\ldots,n\}$, one can replace $\mathbf{C}(\mathbf{x}_\alpha-\mathbf{x}_\beta)$ by $\mathbf{C}(\mathbf{0})-\boldsymbol{\Gamma}(\mathbf{x}_\alpha-\mathbf{x}_\beta)$, where $\boldsymbol{\Gamma}(\mathbf{x}_\alpha-\mathbf{x}_\beta)$ is the $N\times N$ matrix of direct and cross variograms between $\mathbf{Z}(\mathbf{x}_\alpha)$ and $\mathbf{Z}(\mathbf{x}_\beta)$, i.e.,

  $$\boldsymbol{\Gamma}(\mathbf{x}_\alpha-\mathbf{x}_\beta)=\tfrac{1}{2}E\left\{\left[\mathbf{Z}(\mathbf{x}_\alpha)-\mathbf{Z}(\mathbf{x}_\beta)\right]\left[\mathbf{Z}(\mathbf{x}_\alpha)-\mathbf{Z}(\mathbf{x}_\beta)\right]^T\right\} \tag{11}$$

  Note that the symmetry condition holds when the direct and cross covariance functions are fitted with a linear coregionalization model (Journel and Huijbregts, 1978; Chilès and Delfiner, 1999). One may also rewrite the cokriging equations using matrices of direct and pseudo-cross variograms $\mathbf{G}(\mathbf{x}_\alpha-\mathbf{x}_\beta)$, which differ from $\boldsymbol{\Gamma}(\mathbf{x}_\alpha-\mathbf{x}_\beta)$ by a symmetric matrix with zeros on the diagonal (Myers, 1991).

2.3. Inference of the constraints between mean values

In practice, the linear constraints between the mean values of the vector components (Eq. (1)) can be determined in several fashions:

- They can be imposed by the nature of the vector random field under study. For example, if this random field represents a regionalized composition, the mean values of its components should sum to one (Tolosana-Delgado et al., 2008). Also, random field components that represent the same magnitude measured with different devices or on different supports should have the same mean value, unless there is a systematic bias in the measurements.
- They can be inferred by geological knowledge (e.g., stoichiometry of specific minerals or substitution processes in which one element replaces another).
- They can be determined experimentally by fitting a linear or a multilinear regression model between the local mean values of the random field components, assuming that the random field is *locally* second-order stationary (Journel and Huijbregts, 1978). For want of known values, these local means can be estimated by ordinary kriging for each random field component separately.

2.4. Program description

The previous cokriging approach has been implemented in a MATLAB program called COKRIGE. The program allows the prediction of second-order stationary vector random fields defined in Euclidean spaces of one, two, or three dimensions, whose correlation structure corresponds to a linear coregionalization model, i.e., such that the direct and the cross covariances of the vector components are linear combinations of the same set of basic nested structures (Journel and Huijbregts, 1978; Wackernagel, 2003). Each basic structure can have a geometric anisotropy, described by three rotation angles (subroutine SETROT) and three ranges (Deutsch and Journel, 1992). The list of available basic structures is provided in subroutine COVA; any new structure can be added in this subroutine.

The program offers the following functionalities:

(1) No restriction on the number of vector components.
(2) Handling of heterotopic data sets. Missing data values can be codified as not-a-number (*NaN*) or as values outside a trimming limit interval.
(3) No restriction on the number of nested structures used in the linear coregionalization model.
(4) Predictions at scattered or at gridded locations. In the latter case, the grid nodes are ordered point by point to the east, then row by row to the north, and level by level upward (Deutsch and Journel, 1992).
(5) Either point or block cokriging can be used. In the latter case, the block dimensions are equal to the grid mesh.
(6) Either a moving or a unique neighborhood can be used. In the former case, the neighborhood is the region enclosed by an ellipsoid centered on the target location and defined by three rotation angles, three semiaxis lengths, and a maximum number of data points to select (subroutines PICKSUPR, SUPERBLK, and SEARCH) (Deutsch and Journel, 1992). All the $\mathbf{Z}$-components known at the selected data points are then used for cokriging. A unique neighborhood is used if the first semiaxis length is set to infinity.
(7) An option to filter out (remove) the mean, the nugget effect, and/or other basic nested structures in the predictions (Wackernagel, 2003). For instance one can filter out the nugget effect in order to avoid discontinuities in the prediction maps when a target location coincidentally matches a data point. Filtering out all the nested structures of the linear coregionalization model (including the nugget effect) allows predicting the mean value of the vector random field.
(8) An option to apply cross-validation. In the unique neighborhood case, the program code is based on the results presented by Dubrule (1983) and Emery (2009).

The user can input the parameters in the MATLAB workspace or be prompted for an ASCII parameter file. A default parameter file (COKRIGE.PAR, see Table 1) is created if the specified input file does not exist. The output of program COKRIGE is an external ASCII file with the prediction of the vector random field at the target locations (first N columns) and the associated prediction variances (last N columns).

3. Case studies

3.1. First case study: total and soluble copper grade data in an exotic copper deposit

The first case study deals with the prediction of mineral resources in the oxide zone of an exotic copper deposit mined

Table 1
Default parameter file for program COKRIGE.

Parameters for COKRIGE	
*********************	
Start of parameters:	
0	% type of cokriging: 0=gridded locations; 1=scattered locations
locations.prn	% if =1: file with coordinates of locations for cokriging
1 2 3	% columns for location coordinates
5.0 5.0 655.0	% if =0: *x*0, *y*0, *z*0
52 42 20	% *nx*, *ny*, *nz*
10.0 10.0 10.0	% *dx*, *dy*, *dz*
4 4 3	% block discretization (1 1 1 for point-support cokriging)
grades.dat	% file with conditioning data
1 2 3	% columns for coordinates
4 5	% columns for data values
-1 1e21	% trimming limits for data values
1	% number of constraints on the mean values of the form sum{a_ki*m_i}=b_k
0.9 –1 0	% 1st constraint: coefficients a_ki and b_k for *k*=1
2	% number of nested structures
1 35 35 35 0 0 0 1	% 1st structure: it a1 a2 a3 ang1 ang2 ang3 b
0.16 0.20 0.20 0.29	% variance–covariance matrix
2 35 35 15 0 0 0 1	% 2nd structure: it a1 a2 a3 ang1 ang2 ang3 b
2.74 2.51 2.51 2.34	% variance–covariance matrix
2.15 1.89 1.89 1.84	% nugget effect variance–covariance matrix
0 0 0	% filter out mean, nugget and/or other nested structures? 1=yes, 0=no
0	% cross-validation? 1=yes, 0=no
100 100 50	% maximum search radii in the rotated system
0 0 0	% angles for search ellipsoid
1	% divide into octants? 1=yes, 0=no
8	% optimal number of data per octant (if octant=1) or in total (if 0)
cokrige.out	% name of output file
3	% number of decimals for values in the output file
1	% create a header in the output file? 1=yes, 0=no
Available model types:	
1: spherical	
2: exponential	
3: gamma (parameter $b>0$)	
4: stable (parameter $b<2$)	
5: cubic	
6: Gaussian	
7: cardinal sine	
8: J-Bessel (parameter $b>0.5$)	
9: K-Bessel (parameter $b>0$)	
10: generalized Cauchy (parameter $b>0$)	
11: exponential sine	

by open pit. Since the ore material is processed by heap leaching, the total copper grade (*tCu*) is not necessarily representative of the recoverable resources, as there are variations in the metal recovery depending on the mineralogy of the ore material (mainly, chrysocolla, atacamite, and malaquite). In this context, it is of interest to predict the acid-soluble copper (*sCu*), which corresponds to the fraction of total copper that is likely to be recoverable by heap leaching. The predictions of soluble copper grade (primary variable) and total copper grade (secondary variable) provide the data for economic analyses, mine design, and planning.

The database consists of 19,474 exploration drill hole samples, composited at a length of 1 m. A characteristic of the database is that soluble copper grade has been assayed only in part of the samples (preferentially, those for which the total copper grade is greater than a given cutoff grade) (Fig. 1). Because of this preferential sampling, the experimental mean soluble copper grade turns out to be much greater than the experimental mean total copper grade (Table 2) and is an inaccurate estimator of the true mean soluble copper grade.

The total and soluble copper grades are regarded as the components of a second-order stationary vector random field. Their direct and cross variograms have been calculated using the locations with data for both components, and then fitted by a linear coregionalization model with three basic nested structures (namely a nugget effect, an isotropic spherical with range 35 m, and an exponential model with practical ranges 35 m along the horizontal plane and 15 m along the vertical direction):

$$\begin{pmatrix} \gamma_{tCu} & \gamma_{tCu/sCu} \\ \gamma_{tCu/sCu} & \gamma_{sCu} \end{pmatrix} = \begin{pmatrix} 2.15 & 1.89 \\ 1.89 & 1.84 \end{pmatrix} \text{nugget} + \begin{pmatrix} 0.16 & 0.20 \\ 0.20 & 0.29 \end{pmatrix} \text{sph}_{35,35} + \begin{pmatrix} 2.74 & 2.51 \\ 2.51 & 2.34 \end{pmatrix} \exp_{35,15}. \tag{12}$$

Based on the scatter plot displayed in Fig. 2, which shows a proportionality relationship between the total and the soluble copper grades with a correlation coefficient of 0.96, one assumes that the mean grades are related in the following fashion:

$$m_{sCu} = \omega\, m_{tCu} \tag{13}$$

with $0<\omega<1$ (the intercept is assumed to be zero, insofar as the inequality $sCu \le tCu$ implies $sCu=0$ when $tCu=0$). The proportionality coefficient ω (solubility ratio) can be estimated by the quotient of the mean grades calculated with the samples where both grade variables are known, which gives $\omega \approx 0.9$.

Having modeled the relationship between mean values and the spatial correlation structure of the vector random field, three versions of cokriging are used to jointly predict the total and

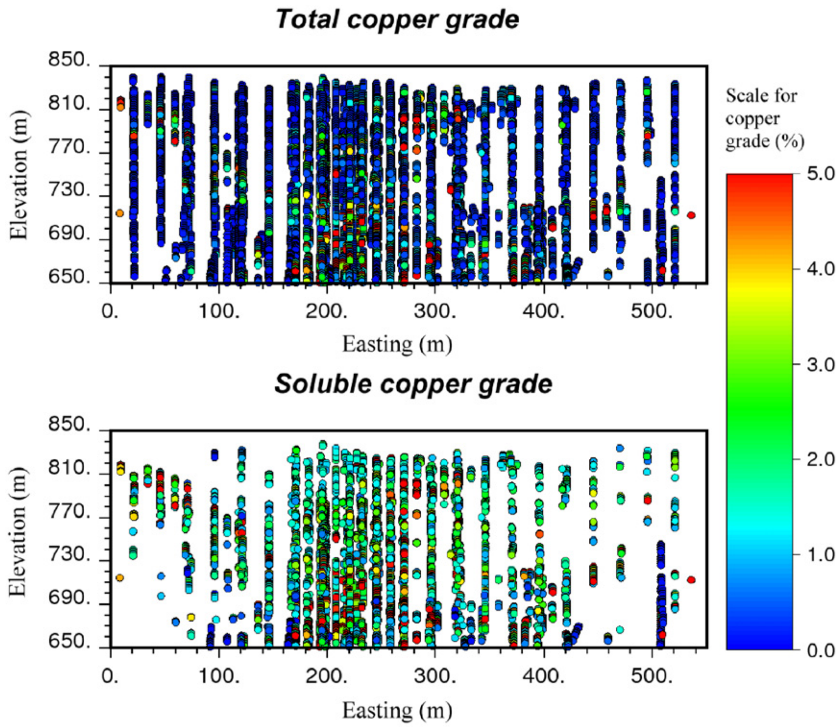

Fig. 1. Location maps (cross section) of total and soluble copper grade data (exotic copper deposit).

Table 2
Declustered statistics of copper grade data, expressed in % (exotic copper deposit).

Variable	Number of data	Minimum	Maximum	Mean	Standard deviation
Total copper grade	19,474	0.05	15.73	1.023	1.742
Soluble copper grade	5773	0.05	14.63	2.681	2.096

For confidentiality reasons, original grade values have been adjusted by a constant scale factor. Declustering has been performed with the cell method.

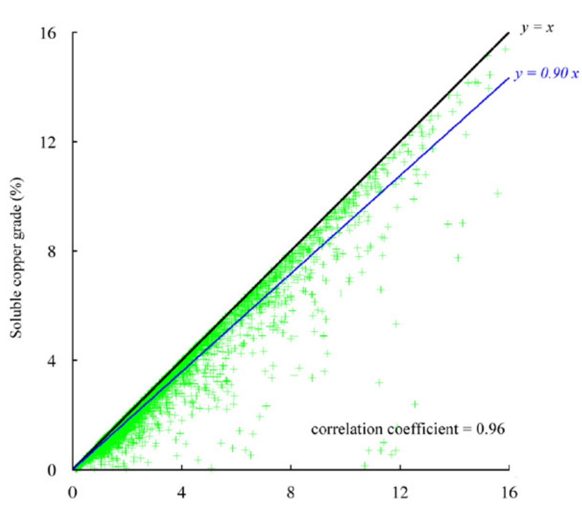

Fig. 2. Scatter plot between total and soluble copper grades (exotic copper deposit).

soluble copper grades: simple cokriging, using the mean value indicated in Table 2 for total copper grade (1.023%) and 0.9 times this mean value for soluble copper grade (i.e., 0.921%); ordinary cokriging; and cokriging with linearly dependent means, based on Eq. (13). In each case, block cokriging is used to predict the average grade on a regular grid containing $52 \times 42 \times 20$ blocks of size $10 \times 10 \times 10$ m, with a block discretization of $4 \times 4 \times 3$. A moving neighborhood with up to 64 data (8 per octant) is used to select the data for cokriging (Table 1).

Table 3
Statistics for predicted total copper grades (exotic copper deposit).

Cokriging type	Minimum	Maximum	Mean	Standard deviation
Simple	0.081	7.332	1.026	0.628
Ordinary	0.058	7.580	0.893	0.824
With linearly dependent means	0.062	7.580	0.900	0.827

The statistics of predicted block grades are indicated in Tables 3–6. The following comments can be made.

(1) Unlike the other two approaches, ordinary cokriging only predicts the grades of 35,655 out of the 43,680 grid blocks. This can be explained by the undersampling of soluble copper grade (primary variable): for the grid blocks with no neighboring primary data, it is not possible to solve the ordinary cokriging equations because of the unbiasedness constraint on the primary cokriging weights. In Tables 3–6, the statistics are calculated on the 35,655 blocks that have been predicted with all three cokriging approaches.
(2) The average of the simple cokriging predictions are close to the assumed mean values of the vector random field components (1.023% for total copper grade; 0.921% for soluble copper grade). However, these assumed mean values have been defined on the basis of the experimental mean total copper grade and do not necessarily match the true mean grades over the volume enclosed in the target grid, presumably because the sampling mesh might be denser in

high-grade areas. In other words, the assumption that the mean values of the vector components are perfectly known is questionable. This is corroborated by observing that the average predicted grades are significantly greater with simple cokriging than with the other two approaches (Table 3).

(3) The predictions of total copper grade are not significantly different between ordinary cokriging and cokriging with linearly dependent means (Tables 3 and 5). This is explained because the total copper grade is known at all the data points and tends to screen out the influence of the collocated soluble copper grade data (Goovaerts, 1997).

(4) Greater differences are observed for the predictions of soluble copper grade (Tables 4 and 6). Cokriging with linearly dependent means better reproduces the proportionality relationship between total and soluble copper grades (Fig. 3c), with a linear regression close to the expected line $y=0.90x$. In contrast, ordinary cokriging leads to a more dispersed scatter plot, a larger number of blocks with inconsistent predictions (soluble copper grade greater than total copper grade or smaller than zero), and a mean soluble copper grade smaller than 0.9 times the mean total copper grade (Fig. 3b; Tables 3 and 4). These poorer results of ordinary cokriging can be explained by the sampling design of soluble copper grade and by the constraint on the sum of ordinary cokriging weights.
Indeed, to supply for the lack of information of the primary variable (soluble copper grade) in undersampled areas, ordinary cokriging uses data of the secondary variable (total copper grade). On the one hand, because the secondary weights sum to zero and because of the strong correlation between the variables, the neighboring secondary data tend to be positively weighted, while the secondary data located farther away tend to be negatively weighted. On the other hand, since the sampling of soluble copper grade is preferential, undersampled areas are likely to be low-grade areas. As a consequence, positive secondary weights tend to be assigned to low total copper grade data, while negative secondary weights tend to be assigned to data with, on average, a higher total copper grade. This dependence between data weighting and data values explains the larger number of blocks with negative values and the smaller average value for the predictions of soluble copper grade.
In contrast, simple cokriging and cokriging with linearly dependent means turn out to be less sensitive to the preferential sampling pattern, insofar as the weights assigned to the secondary data are not forced to sum to zero.
Soluble copper grade predictions can be corrected by setting inconsistent predictions to zero or to the total copper grade, in order to reproduce the inequalities $0 \leq sCu \leq tCu$. This correction slightly modifies the regression line between the total and the soluble copper grades predicted by ordinary cokriging (the intercept and slope become -0.07 and 0.91, respectively), but this regression line is still poorer than the one obtained by cokriging with linearly dependent means.

Table 4
Statistics for predicted soluble copper grades (exotic copper deposit).

Cokriging type	Minimum	Maximum	Mean	Standard deviation
Simple	−0.737	6.959	0.909	0.591
Ordinary	−2.794	7.188	0.723	0.801
With linearly dependent means	−0.730	7.194	0.796	0.754

Table 5
Correlation coefficients between predicted total copper grades (exotic copper deposit).

Cokriging type	Simple	Ordinary	With linearly dependent means
Simple	1	0.902	0.898
Ordinary	0.902	1	0.999
With linearly dependent means	0.898	0.999	1

Table 6
Correlation coefficients between predicted soluble copper grades (exotic copper deposit).

Cokriging type	Simple	Ordinary	With linearly dependent means
Simple	1	0.887	0.901
Ordinary	0.887	1	0.956
With linearly dependent means	0.901	0.956	1

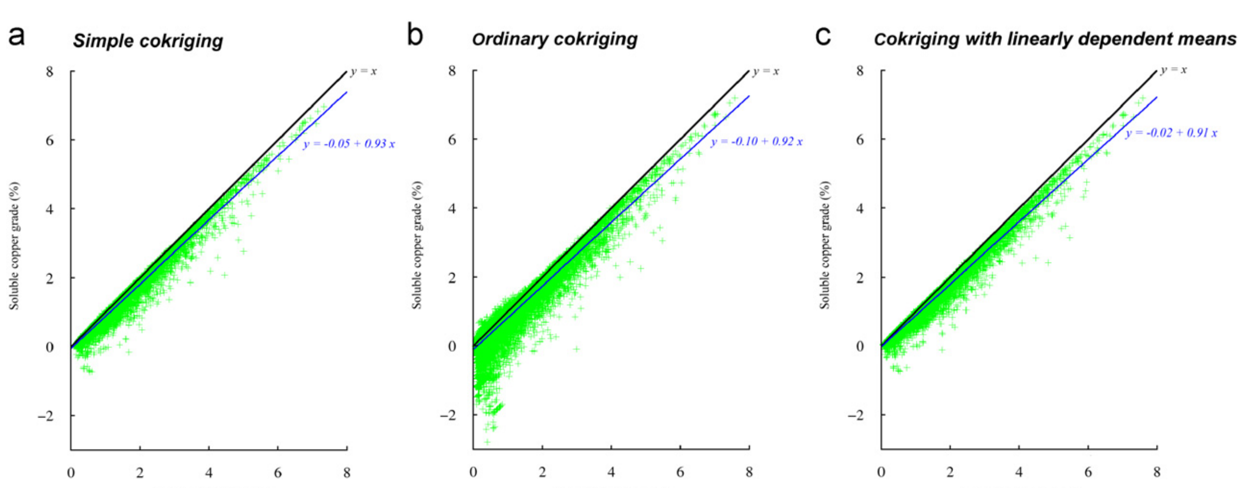

Fig. 3. Scatter plot between total and soluble copper grades predicted at block support, with simple cokriging, ordinary cokriging, and cokriging with linearly dependent means (exotic copper deposit).

3.2. Second case study: drill hole and blast hole copper grade data in a porphyry deposit

The second case study corresponds to a porphyry copper deposit, in which the total copper grade is known on two different sample types: drill hole (DH) composites and blast holes (BH). Due to the differences in volumetric supports and in sampling protocols between drill holes and blast holes, the grades from these two sources of information are associated with two random field components with equal mean value, but possibly different variances and spatial correlation structures (Table 7).

The assumption of equal mean values can be justified by a lagged scatter diagram, by pairing drill hole data with blast hole data at a short distance. For instance, 1455 pairs are found with a separation distance less than 5 m: the average difference in the drill hole and blast hole copper grades is 0.03%, which is not deemed significant.

Table 7
Basic statistics of copper grade data, expressed in % (porphyry copper deposit).

Variable	Number of data	Minimum	Maximum	Mean	Standard deviation
Drill hole copper grade	5228	0.01	3.78	0.673	0.379
Blast hole copper grade	27,827	0.01	4.12	0.763	0.436
Overall	33,055	0.01	4.12	0.748	0.429

For confidentiality reasons, original grade values have been adjusted by a constant scale factor.

Since the sampling scheme is totally heterotopic, the cross correlation structure of the two random field components is inferred by calculating their pseudo-cross variogram (Myers, 1991; Papritz et al., 1993; Cressie and Wikle, 1998). The linear coregionalization model for the direct and pseudo-cross variograms is composed of a nugget effect and nested spherical and exponential structures (Fig. 4):

$$\begin{pmatrix} \gamma_{DH} & \gamma_{DH-BH} \\ \gamma_{DH-BH} & \gamma_{BH} \end{pmatrix} = \begin{pmatrix} 0.02 & 0.045 \\ 0.045 & 0.06 \end{pmatrix} \text{nugget} + \begin{pmatrix} 0.05 & 0.05 \\ 0.05 & 0.05 \end{pmatrix} \text{sph}_{30,30}$$
$$+ \begin{pmatrix} 0.065 & 0.05 \\ 0.05 & 0.04 \end{pmatrix} \text{sph}_{150,150} + \begin{pmatrix} 0.02 & 0.03 \\ 0.03 & 0.05 \end{pmatrix} \exp_{150,\infty}, \tag{14}$$

where the figures in subscript indicate the range or practical range (in meters) along the horizontal and vertical directions

The following comments are pertinent to the linear coregionalization model above:

- The direct variograms of the two random field components are quite similar, except that the blast hole grade variogram has a greater nugget effect and greater sill than the drill hole grade variogram. This can be explained by larger sampling errors (mainly, fundamental, delimitation, extraction, and preparation errors) for blast holes than for drill holes (Pitard, 2008). Accordingly, the drill hole copper grade will be considered as the primary variable (targeted for prediction), while the blast hole copper grade is viewed as a covariate.

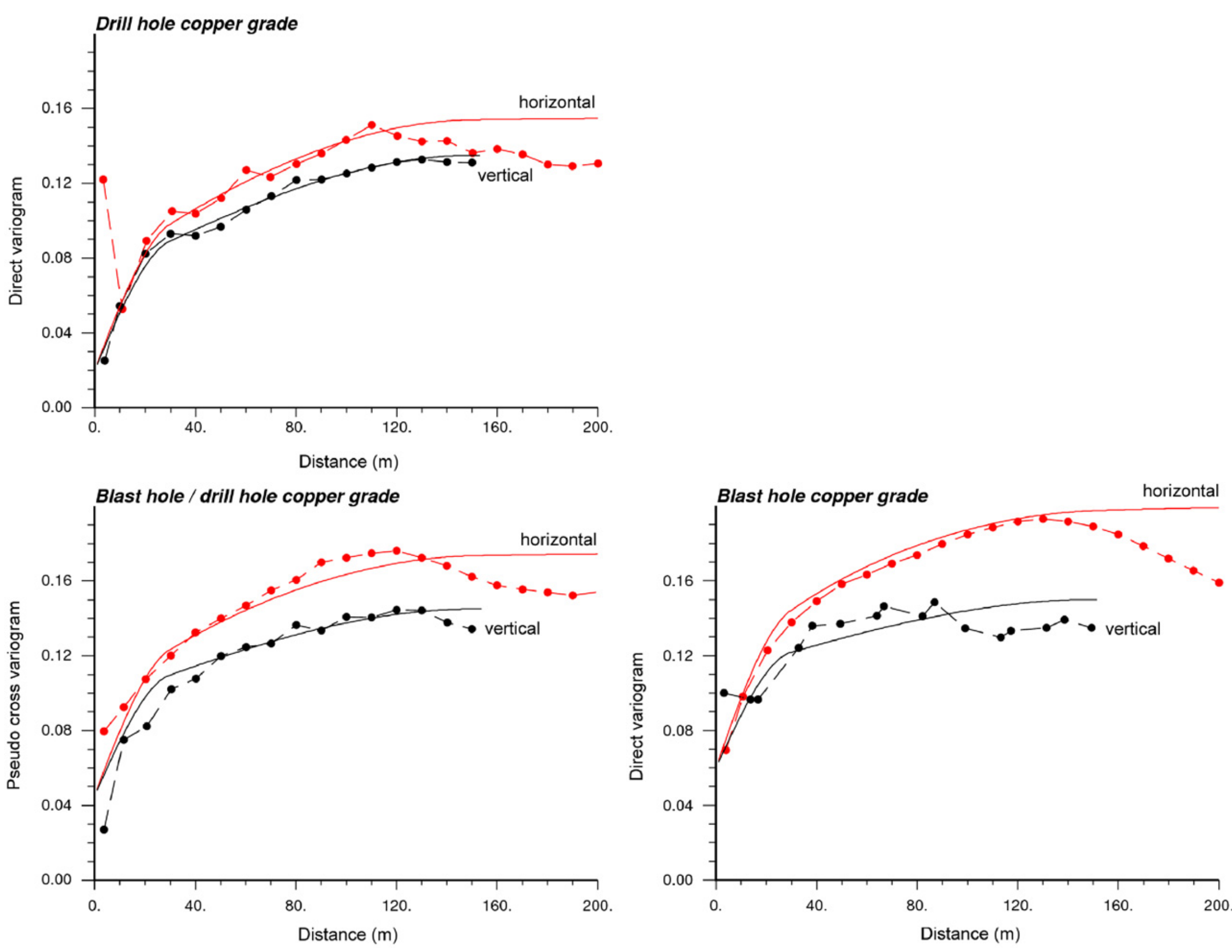

Fig. 4. Sample (dots and dashed lines) and modeled (solid lines) direct and pseudo-cross variograms of drill hole and blast hole copper grades (porphyry copper deposit).

Table 8
Statistics of cross-validation errors for drill hole copper grade (porphyry copper deposit).

Cokriging type	Mean error	Mean absolute error	Root mean square error
Simple	−0.008	0.162	0.233
Ordinary	−0.002	0.182	0.269
Standardized ordinary	−0.006	0.161	0.233

Table 9
Statistics of cross-validation standardized errors for drill hole copper grade (porphyry copper deposit).

Cokriging type	Mean error	Mean absolute error	Root mean square error
Simple	−0.034	0.677	0.985
Ordinary	−0.004	0.734	1.085
Standardized ordinary	−0.027	0.674	0.984

- For the nugget effect structure, the sill matrix of the linear coregionalization model expressed in terms of direct and pseudo-cross variograms (Eq. (14)) may differ from the sill matrix of the model expressed in terms of direct and cross covariance functions, which is the one required for cokriging (Eq. (9)). Specifically, the sill of the cross structure must be revised. If one assumes that the nugget effect of the drill hole grades corresponds to a short-range structure that will also be present in blast hole grades, and that the extra nugget effect of blast hole grade is due to sampling errors, the nugget effect of the cross structure should be only the short-range component, i.e., its sill should be 0.02.

The results from three forms of cokriging are compared by the use of leave-one-out cross-validation: simple cokriging, with the overall mean value (0.748%) for the two random field components; ordinary cokriging; and cokriging with linearly dependent means, as per Eqs. (5)–(10), which in this case study reduces to standardized ordinary cokriging. The basic statistics on prediction errors and standardized errors are presented in Tables 8 and 9, showing that ordinary cokriging leads to poorer predictions, while there is no significant difference between the other two approaches. As for the first case study, these results indicate that not incorporating the linear dependence of the means of the respective components may affect the accuracy of cokriging predictions.

4. Conclusions

The prediction of a second-order stationary vector random field by cokriging relies on the modeling of its mean vector (constant in space) and of its direct and cross covariance functions. Quite often, the latter are fitted via a linear coregionalization model, while the mean values of the vector random field components are assumed either perfectly known (simple cokriging) or unknown and unrelated to each other (ordinary cokriging).

In some cases, however, there may exist linear combinations with known coefficients between the mean values of the vector components, even if these means are unknown. The cokriging equations can be modified to take these linear combinations into account. The applicability and versatility of this approach have been illustrated with case studies in ore body evaluation, in which the prediction by cokriging with linearly dependent means is shown to outperform that of ordinary cokriging.

Possible extensions of the presented work include the cases of components with linearly dependent drifts (universal cokriging), factorial kriging analysis, and multivariate Gaussian cosimulation.

Acknowledgments

This research was funded by the Chilean Commission for Scientific and Technological Research (CONICYT), through FONDECYT Project 1090013. The author acknowledges the support of the Advanced Laboratory for Geostatistical Supercomputing (ALGES) at University of Chile, as well as two anonymous reviewers for their comments on an earlier version of this work.

Appendix A. Second-order stationary vector random field with linearly dependent components

Assume that there exists a linear relationship between the components of $\mathbf{Z}$:

$$\mathbf{aZ} = 0 \tag{A1}$$

for some nonzero vector $\mathbf{a}$ of size $1 \times N$.

On the one hand, the linear relationship also holds for the vector of mean values:

$$\mathbf{am} = 0 \tag{A2}$$

This implies that $\mathbf{a}$ is a row of matrix $\mathbf{A}$ defined in Eq. (2). By writing $\mathbf{a}=(\mathbf{a}_1\ \mathbf{a}_2)$, where $\mathbf{a}_1$ and $\mathbf{a}_2$ are subvectors of respective sizes $1 \times (N-K)$ and $1 \times K$, and by accounting for the definition of $\mathbf{F}$ (Eqs. (3) and (4)), it follows that

$$\mathbf{aF} = \mathbf{a}_1 - \mathbf{a}_2\mathbf{A}_2^{-1}\mathbf{A}_1 = \mathbf{0} \tag{A3}$$

On the other hand, for any pair of locations $\{\mathbf{x}_\alpha, \mathbf{x}_\beta\}$, one has

$$\mathbf{aC}(\mathbf{x}_\alpha - \mathbf{x}_\beta) = \mathbf{a}E\{\mathbf{Z}(\mathbf{x}_\alpha)\mathbf{Z}(\mathbf{x}_\beta)^T\} = E\{\mathbf{aZ}(\mathbf{x}_\alpha)\mathbf{Z}(\mathbf{x}_\beta)^T\} = \mathbf{0} \tag{A4}$$

Consider the following submatrices extracted from the cokriging matrix (Eq. (9)):

$$\forall \alpha \in \{1,\ldots,n\},\ \mathbf{G}_\alpha = (\mathbf{C}\,(\mathbf{x}_\alpha - \mathbf{x}_1) \cdots \mathbf{C}\,(\mathbf{x}_\alpha - \mathbf{x}_n)\mathbf{F}) \tag{A5}$$

From Eqs. (A3) and (A4), it follows that

$$\forall \alpha \in \{1,\ldots,n\}, \mathbf{aG}_\alpha = \mathbf{0} \tag{A6}$$

Accordingly, for any $\alpha = 1,\ldots,n$, $\mathbf{G}_\alpha$ is not a full rank matrix and the cokriging matrix (Eq. (9)) is singular.

Appendix B. Supplementary Material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.cageo.2011.06.001.

References

Ahmed, S., De Marsily, G., 1987. Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity. Water Resources Research 23 (9), 1717–1737.

Carr, J.R., Myers, D.E., Glass, C.E., 1985. Co-kriging—a computer program. Computers & Geosciences 11 (2), 111–127.

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

Cressie, N., Wikle, C.K., 1998. The variance-based cross-variogram: you can add apples and oranges. Mathematical Geology 30 (7), 789–799.

Deutsch, C.V., Journel, A.G., 1992. GSLIB: Geostatistical Software Library and User's Guide. Oxford University Press, New York, NY, 340pp.

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

Emery, X., 2009. The kriging update equations and their application to the selection of neighboring data. Computational Geosciences 13 (3), 269–280.

Fuentes, M., 2001. A high frequency kriging approach for non-stationary environmental processes. Environmetrics 12 (5), 469–483.

Goovaerts, P., 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press, New York, NY, 480pp.

Goovaerts, P., 1998. Ordinary cokriging revisited. Mathematical Geology 30 (1), 21–42.

Higdon, D., Swall, J., Kern, J., 1999. Non-stationary spatial modeling. In: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (Eds.), Bayesian Statistics 6. Oxford University Press, Oxford, UK, pp. 761–768.

Hohn, M.E., 1999. Geostatistics and Petroleum Geology, 2nd edn. Kluwer Academic, Dordrecht, The Netherlands, 235pp.

Isaaks, E.H., Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. Oxford University Press, New York, NY, 560pp.

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

Kitanidis, P.K., 1997. Introduction to Geostatistics: Applications to Hydrology. Cambridge University Press, London, UK, 249pp.

Marcotte, D., 1991. Cokriging with MATLAB. Computers & Geosciences 17 (9), 1265–1280.

Myers, D.E., 1982. Matrix formulation of co-kriging. Mathematical Geology 14 (3), 249–257.

Myers, D.E., 1984. Co-kriging—new developments. In: Verly, G., David, M., Journel, A.G., Maréchal, A. (Eds.), Geostatistics for Natural Resources Characterization. Reidel, Dordrecht, The Netherlands, pp. 295–305.

Myers, D.E., 1991. Pseudo-cross variograms, positive definiteness, and cokriging. Mathematical Geology 23 (6), 805–816.

Pan, G., Gaard, D., Moss, K., Heiner, T., 1993. A comparison between cokriging and ordinary kriging: case study with a polymetallic deposit. Mathematical Geology 25 (3), 377–398.

Pan, G., Moss, K., Heiner, T., Carr, J.R., 1992. A FORTRAN program for three-dimensional cokriging with case demonstration. Computers & Geosciences 18 (5), 557–578.

Papritz, A., 2008. Standardized vs. customary ordinary cokriging: some comments on the article ‘‘The geostatistical analysis of experiments at the landscape-scale’’ by T.F.A. Bishop and R.M. Lark. Geoderma 146 (1–2), 291–396.

Papritz, A., Künsch, H.R., Webster, R., 1993. On the pseudo-cross variogram. Mathematical Geology 25 (8), 1015–1026.

Pebesma, E.J., 2004. Multivariate geostatistics in S: the Gstat package. Computers & Geosciences 30 (7), 683–691.

Pebesma, E.J., Wesseling, C.G., 1998. Gstat, a program for geostatistical modelling, prediction and simulation. Computers & Geosciences 24 (1), 17–31.

Pitard, F.F., 2008. Blasthole sampling for grade control—the many problems and solutions. In: Sampling 2008 Proceedings, The Australasian Institute of Mining and Metallurgy, Carlton, Victoria, pp. 15–22.

Sampson, P.D., Damian, D., Guttorp, P., 2001. Advances in modeling and inference for environmental processes with nonstationary spatial covariance. In: Monestiez, P., Allard, D., Froidevaux, R. (Eds.), GeoENV 2000: Geostatistics for Environmental Applications. Kluwer, Dordrecht, The Netherlands, pp. 17–32.

Stein, A., Van Dooremolen, W., Bouma, J., Bregt, A.K., 1988. Cokriging point data on moisture deficit. Soil Science Society of America Journal 52 (5), 1418–1423.

Tolosana-Delgado, R., Pawlowsky-Glahn, V., Egozcue, J.J., 2008. Indicator kriging without order relation violations. Mathematical Geosciences 40 (3), 327–347.

Ver Hoef, J.M., Cressie, N., 1993. Multivariable spatial prediction. Mathematical Geology 25 (2), 219–240.

Wackernagel, H., 1988. Geostatistical techniques for interpreting multivariate spatial information. In: Chung, C.F., Fabbri, A.G., Sinding-Larsen, R. (Eds.), Quantitative Analysis of Mineral and Energy Resources. Reidel, Dordrecht, The Netherlands, pp. 393–409.

Wackernagel, H., 2003. Multivariate Geostatistics—An Introduction with Applications. Springer, Berlin, Germany, 387pp.

Xu, W. Tran, T.T., Srivastava, R.M., Journel, A.G., 1992. Integrating seismic data in reservoir modeling: the collocated cokriging alternative. In: Proceedings of the 67th Annual Technical Conference and Exhibition, SPE paper 24742, Society of Petroleum Engineers, Washington, DC, pp. 833–842.

Yates, S.R., Warrick, A.W., 1987. Estimating soil water content using cokriging. Soil Science Society of America Journal 51, 23–30.