Modeling the Coregionalization of Multiscale Data

S. N. Elogne and Oy Leuangthong

An essential task in geostatistics is integrating various data sources to improve numerical models. Modeling the spatial dependence and codependence between attributes is an important issue in multivariate spatial data analysis. Cokriging with large-scale secondary data is not often used because of the challenge of obtaining a valid model of coregionalization. The difficulty in using the linear model of coregionalization lies with the different smoothness properties of the data sets. For instance, standard methods for cosimulation cannot incorporate one field with a spherical auto-covariance, the other with a Gaussian auto-covariance, and their cross-covariance with a Gaussian model. Our main goal in this paper is to derive a valid linear model of coregionalization from data collected at different scales for cokriging prediction and cosimulation. In addition, we discuss some conditions under which the positive semidefiniteness/definiteness properties of the proposed model of coregionalization are satisfied. Finally, the proposed model is applied to a cobalt and nitrate data sets.

Introduction

With the increasing interest in accurate simulation of reservoir properties, subsurface-characterization based on limited well and seismic data takes on greater importance in reservoir modeling. Reducing the uncertainty in reservoir prediction can be achieved by integrating additional information such as core and seismic data. One of the major challenges in spatial statistics is the modeling of the auto- and cross-dependence structures by analyzing and integrating information collected from different sources and at different resolutions. The determination of the cross-spatial dependence between attributes with different measurement scales plays a significant role in data integration. The difference in volumes makes some classical geostatistical methods nontrivial as the correlation coefficient between the attributes cannot be computed in a scale-consistent manner. In general, we have hard data such as well logs and core plugs, and soft data such as larger-scale seismic data (Xu et al. 1992), (Tjolsen et al. 1995), (Doyen et al. 1996), (Doyen et al. 1997), (Frykman and Deutsch 1999), (Yao and Journel 2000).

In the practice of cosimulation it is usually assumed that two correlated random fields have identical structures. The most-used approach (Goovaerts 1997) characterizes the coregionalization matrix in the following way:

$$\begin{pmatrix} c_1(\mathbf{h}) & c_{1,2}(\mathbf{h}) \\ c_{1,2}(\mathbf{h}) & c_2(\mathbf{h}) \end{pmatrix} = \begin{pmatrix} b_{11}^{(1)} & b_{12}^{(1)} \\ b_{12}^{(1)} & b_{22}^{(1)} \end{pmatrix} \varphi_1(\mathbf{h}) + \cdots + \begin{pmatrix} b_{11}^{(L)} & b_{12}^{(L)} \\ b_{12}^{(L)} & b_{22}^{(L)} \end{pmatrix} \varphi_L(\mathbf{h}) \tag{1}$$

where $\varphi_j(\mathbf{h})$, *j=1, …, L* are covariance models and where the matrices of coefficients are positive semidefinite (Journel and Huijbregts 1978). One reason why cokriging methods are not often considered in a multivariate geostatistical framework is because of the challenges associated with modeling the coregionalization (Myers 1982). This process is already difficult with data collected at the same scale and is made even more so when different resolutions are considered (Kupfersberger et al. 1998). This is because every aspect of the framework needs to be incorporated in a more robust approach without ignoring their difference in resolution. To our knowledge, the only technique that considers this kind of modeling is a Markov model approach that is based on screening hypotheses (Journel 1999). While the Markov model is a particular case of the LMC (1), it is given in such a way that the cross-covariance model $c_{1,2}(\mathbf{h})$ is proportional to the secondary field covariance $c_2(\mathbf{h})$ and does not need to be inferred. It follows that the covariance of the primary field also contains the secondary-field covariance since sufficient conditions must be imposed on both a cross-covariance model and its associated auto-covariance model.

It is likely that the spatial structure of the attribute at the coarser scale will be smooth. It is also expected that the auto-covariance structure of the attribute at the finer scale will present more variability at small distance lags, and that the commonly-used linear model of coregionalization (1) will not be adequate for modeling the coregionalization matrix. This is because finding the same set of basic

covariance models will result in drastic approximations and subsequent errors and inaccuracies in the estimation. As Oliver (2003) pointed out: "While it may be unlikely for the auto-covariance of porosity to be much different structurally from the auto-covariance of permeability, it is certainly clear that the auto-covariance of seismic impedance might be much different from the auto-covariance of porosity, simply because of the averaging inherent in the data acquisition and processing." Furthermore, the experimental cross-covariance $\hat{c}_{1,2}(\mathbf{h})$ might display a different effective range from the experimental auto-covariance of the soft data $\hat{c}_2(\mathbf{h})$. This is partly due to the use of collocated data for computing $\hat{c}_{1,2}(\mathbf{h})$ as shown in the proposed example (see Section 4).

The aim of this paper is to demonstrate that conditions may be imposed on the cross-covariance structure separately from the auto-covariance models. The construction of a positive definite coregionalization model is considered in a way that allows independent fitting of the auto/cross-covariance from the data. Our technique is mainly based on the spectral representation (Fourier transform) of positive semidefinite covariance function. Although modeling the spatial co-variation between attributes is not straightforward, it is a relevant approach to direct sequential simulation for multiscale data integration (Soares 2001), (Leuangthong 2005). Figure 1 gives an schematic example of three different scales/shapes that occur in practice.

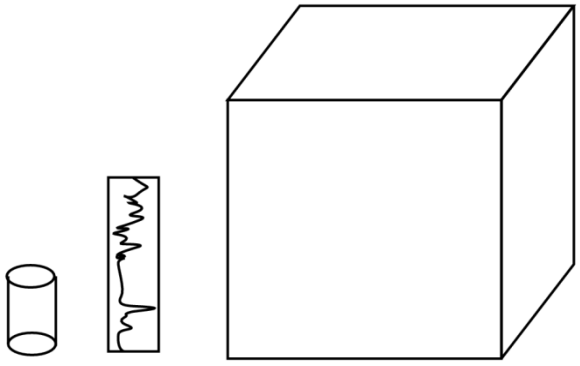

Figure 1: Illustration of volume measures for different length-scales data and shapes that need to be analyzed and incorporated in a multivariate spatial co-variation.

This paper is comprised of five sections. Section 1 defines the purpose of the paper. Section 2 describes related research that has been conducted, while Section 3 illustrates how to preserve the positive semidefiniteness property of the coregionalization matrix. Section 4 is devoted to real data application. Finally, Section 5 presents conclusions and suggests directions for further research.

Background

As previously done by other investigators, we use two scales v and V and consider one primary attribute $Z_1(v_\alpha)$ defined at a finer scale v and one secondary soft data $Z_2(V_\beta)$ defined at a larger scale V (Xu et al. 1992), (Tjolsen et al. 1995), (Doyen et al. 1997), (Yao and Journel 2000). The random processes $Z_1(v_\alpha)$ and $Z_2(V_\beta)$ are further assumed to be second-order stationary and observed at blocks $v_1,\ldots,v_n$ and $V_1,\ldots,V_N$ respectively with $n,N>1$. For modeling the coregionalization covariance structure, it is assumed that the data have been standardized (Cressie and Wikle 1998). This paper focuses exclusively on the problems of uniform volume blocks. Blocks at the supports v and V will be denoted by v_α and V_β, respectively. The center of the volume will be represented by $\langle \mathbf{u}^v \rangle$ and

$\|v\|_x$, $\|v\|_y$ and $\|v\|_z$ will denote its length scales in the x, y and z directions, respectively. In classical geostatistics, one considers that

$$v_\alpha - v_{\alpha'} = \left\langle \mathbf{u}^{v_\alpha} \right\rangle - \left\langle \mathbf{u}^{v_{\alpha'}} \right\rangle \tag{2}$$

in a way that auto-covariance models $\overline{c}_1\left(v_\alpha, v_{\alpha'}\right)$ and $\overline{c}_2\left(V_\beta, V_{\beta'}\right)$ of the processes $Z_1\left(v_\alpha\right)$ and $Z_2\left(V_\beta\right)$ are consistently defined. The associated cross-spatial dependence will be denoted by $\overline{c}_{1,2}\left(v_\alpha, V_\beta\right)$. Consider that $Z_2\left(V_\beta\right)$ is defined through a second order stationary pseudo/point scale random field $Y\left(\mathbf{s}\right)$ with covariance function $c_{\mathrm{y}}\left(\mathbf{h}\right)$ as

$$Z_2\left(V_\beta\right) = N_V^{-1} \sum_{\alpha=1}^{N_V} Y\left(\mathbf{s}_\alpha \in V_\beta\right),$$

where N_V is a discretization number. Let $c_{\mathrm{y},\mathrm{z}_1}$ denote the cross-covariance function between the random fields Y and Z_1. Assume that the coregionalization matrix of the vector (Z_1, Y) can be modeled through a linear model in the form

$$\begin{pmatrix} \overline{c}_1\left(\mathbf{h}\right) & \overline{c}_{\mathrm{y},\mathrm{z}_1}\left(\mathbf{h}\right) \\ \overline{c}_{\mathrm{y},\mathrm{z}_1}\left(\mathbf{h}\right) & c_{\mathrm{y}}\left(\mathbf{h}\right) \end{pmatrix} = \sum_{k=1}^{L} \begin{pmatrix} b_{11}^{(k)} & b_{12}^{(k)} \\ b_{12}^{(k)} & b_{22}^{(k)} \end{pmatrix} \varphi_k\left(\mathbf{h}\right)$$

where the nested basic structures $\varphi_1\left(\mathbf{h}\right), \ldots, \varphi_L\left(\mathbf{h}\right)$ are all differentiable covariance models at the origin. It follows that the covariance $\overline{c}_2\left(V_\beta, V_{\beta'}\right)$ can be written as

$$\begin{aligned} \overline{c}_2\left(V_\beta, V_{\beta'}\right) &= N_V^{-2} \sum_{\alpha=1}^{N_V} \sum_{\alpha'=1}^{N_V} c_{\mathrm{y}}\left(\mathbf{s}_\alpha - \mathbf{s}_{\alpha'}\right) \\ &= N_V^{-2} \sum_{k=1}^{L} b_{22}^{(k)} \left(\sum_{\alpha=1}^{N_V} \sum_{\alpha'=1}^{N_V} \varphi_k\left(\mathbf{s}_\alpha - \mathbf{s}_{\alpha'}\right) \right) \end{aligned} \tag{3}$$

Similarly the cross covariance $\overline{c}_{1,2}\left(v_\alpha, V_\beta\right)$ is obtained through

$$\begin{aligned} \overline{c}_{1,2}\left(v_\alpha, V_\beta\right) &= N_V^{-1} \sum_{\alpha'=1}^{N_V} c_{\mathrm{y},\mathrm{z}_1}\left(\left\langle \mathbf{u}^{v_\alpha} \right\rangle - \mathbf{s}_{\alpha'}\right) \\ &= N_V^{-1} \sum_{k=1}^{L} b_{12}^{(k)} \left(\sum_{\alpha'=1}^{N_V} \varphi_k\left(\left\langle \mathbf{u}^{v_\alpha} \right\rangle - \mathbf{s}_{\alpha'}\right) \right) \end{aligned} \tag{4}$$

It is then likely that the resulting coregionalization matrix

$$\begin{pmatrix} \overline{c}_1\left(v_\alpha, v_{\alpha'}\right) & \overline{c}_{1,2}\left(v_\alpha, V_{\beta'}\right) \\ \overline{c}_{1,2}\left(v_\alpha, V_{\beta'}\right) & \overline{c}_2\left(V_\beta, V_{\beta'}\right) \end{pmatrix}$$

cannot be obtained by a linear model since $\overline{c}_2\left(V_\beta, V_{\beta'}\right)$ and $\overline{c}_{1,2}\left(v_\alpha, V_{\beta'}\right)$ may follow differentiable models due to the averaging given by (3) and (4). This is aptly demonstrated by the auto-covariance of the seismic impedance (Oliver, 2003). With this in mind, there is a crucial need to derive conditions under which a coregionalization model can be positive semidefinite.

The accepted Markov model for coregionalization defines the secondary variable on a much larger support than the primary variable (Journel 1999). The theory behind this assumes that

$$\mathrm{E}\left[Z_1(\mathbf{u}) \mid Z_2(\mathbf{u}) = z_2, Z_2(\mathbf{u}') = z_2' \right] = \mathrm{E}\left[Z_1(\mathbf{u}) \mid Z_2(\mathbf{u}) = z_2 \right] \quad \forall \mathbf{u}, \mathbf{u}' \tag{5}$$

and $\overline{c}_{1,2}(\mathbf{h}) = \rho \overline{c}_2(\mathbf{h})$ and $\overline{c}_1(\mathbf{h}) = \rho^2 \overline{c}_2(\mathbf{h}) + (1-\rho^2) c_r(\mathbf{h})$ for some covariance $c_r(\mathbf{h})$ where ρ denotes the correlation coefficient between the collocated secondary data. In addition to the fact that the Markov assumption (5) does not consider the positioning of the location $\mathbf{u}$ for $Z_1(\mathbf{u}) = z_1(\mathbf{u})$ with respect to the location $\mathbf{u}'$ for $Z_2(\mathbf{u}') = z_2(\mathbf{u}')$, there are at least three drawbacks to the theory. First, the cross-covariance is not modeled from the data. Second, the theory calls for the use of collocated data despite the difference in scale. To further complicate matters, more than four observed locations can fall within one volume of the secondary variable and could lead to many different measurements of the correlation coefficient. In addition, an excessive number of data points are used to compute the Pearson's product moment correlation coefficient (Shmaryan and Journel (1999). Finally, the modeling of the covariance of the primary variable must contain the prior model $\overline{c}_2(\mathbf{h})$, which could both produce inexact calculations and be time-consuming for the practitioner. In this paper we argue that there is a need to obtain the cross-covariance model from the data in a consistent manner that considers the difference in scale. Our inference approach will take advantage of the usually abundant secondary data $z_2(V_1), \cdots, z_2(V_N)$ to model the cross-spatial dependence.

Proposed Methodology

The framework considered in the paper uses the data sets in their original space to demonstrate that the collocated data sets have not been transformed into their normal score counterparts. Cross-covariance functions describe the linear relationship between spatial variables and are crucial for stochastic prediction and simulation. To integrate different data sources, one needs to consider the difference in resolution and define cross-covariance functions between all the attributes of interest without ignoring any relevant information.

Cross covariance modeling

The method of moments estimate of the cross-variogram

$$2\gamma_{1,2}(\mathbf{h}) = \mathrm{E}\left[\left\{ Z_1\left(\left\langle \mathbf{u}^v \right\rangle\right) - Z_1\left(\left\langle \mathbf{u}^v \right\rangle + \mathbf{h}\right) \right\} \left\{ Z_2\left(\left\langle \mathbf{u}^v \right\rangle\right) - Z_2\left(\left\langle \mathbf{u}^v \right\rangle + \mathbf{h}\right) \right\} \right]$$

is given

$$2\hat{\gamma}_{1,2}(\mathbf{h}) = \frac{1}{\left| N_{1,2}(\mathbf{h}) \right|} \sum_{N_{1,2}(\mathbf{h})} \left\{ z_1\left(\left\langle \mathbf{u}^v \right\rangle\right) - z_1\left(\left\langle \mathbf{u}^v \right\rangle + \mathbf{h}\right) \right\} \left\{ z_2\left(\left\langle \mathbf{u}^v \right\rangle\right) - z_2\left(\left\langle \mathbf{u}^v \right\rangle + \mathbf{h}\right) \right\} \tag{6}$$

where $N_{1,2}(\mathbf{h})$ represents the pairs of observations separated by lag $\mathbf{h}$ and $\left| N_{1,2}(\mathbf{h}) \right|$ its cardinality. Because of major scale differences between attributes, geostatistical techniques for obtaining the experimental cross-spatial dependence may be inefficient for small distances presented in the finer scale. There is a difficulty in calculating the cross-variogram $2\hat{\gamma}_{1,2}$ because the centers of blocks are not defined at exactly the same locations. Consider the recorded data of the secondary attribute $z_2(V_1), \cdots, z_2(V_N)$ and let $z_1(v_1), \cdots, z_1(v_n)$ be those of the primary attribute. Since $|v| < |V|$, we define a new set of observed secondary data $z_2'(v_1), \cdots, z_2'(v_M)$ based on the locations of the primary data where

$$z_2'(v_k) = z_2(V_\beta) \quad \text{if } v_k \subset V_\beta . \tag{7}$$

Thus, the sample cross-variogram estimates can be obtained using the standard estimator

$$2\hat{\gamma}_{1,2}(\mathbf{h}) = \frac{1}{\left|N_{1,2}(\mathbf{h})\right|} \sum_{N_{1,2}(\mathbf{h})} \left\{ z_1(v_\alpha) - z_1(v_\beta) \right\} \left\{ z_2'(v_\alpha) - z_2'(v_\beta) \right\} \tag{8}$$

where $N_{1,2}(\mathbf{h})$ represents the pairs of observations separated by lag $\mathbf{h}$ and $\left|N_{1,2}(\mathbf{h})\right|$ its cardinality. The consistency of the approach comes from (i) using the same number of data for calculating the sample variogram estimates of the primary variable and (ii) the similarity between the sample variogram estimates obtained from the data set $z_2'(v_1), \cdots, z_2'(v_M)$ and those from the data set $z_2(V_1), \cdots, z_2(V_N)$. This permits an acceptable cross-variogram estimate which will be analyzed to produce a valid model of coregionalization.

Sufficient conditions to achieve positive semidefiniteness

In this section we describe the sufficient conditions needed to ensure the positive definiteness of the matrix of coregionalization $\mathbf{\Sigma}_{n+m}$. The matrix is given by the following:

$$\mathbf{\Sigma}_{n+m} = \begin{pmatrix} \left[\overline{c}_1(v_i, v_j)\right]_{1\le i,\, j\le n} & \left[\overline{c}_{1,2}(v_j, V_k)\right]_{\substack{1\le j\le n \\ 1\le k\le m}} \\ \left[\overline{c}_{1,2}(V_k, v_j)\right]^T_{\substack{1\le k\le m \\ 1\le j\le n}} & \left[\overline{c}_2(V_k, V_q)\right]_{1\le k,\, q\le m} \end{pmatrix} \tag{9}$$

To ensure the nonnegativity of the variance of any combination $\mathrm{T}_{n,m}$ of the form

$$\mathrm{T}_{n,m} = \sum_{i=1}^{n} \lambda_i Z_1(v_i) + \sum_{j=1}^{m} \lambda_{n+j} Z_2(V_j), \tag{10}$$

for any $n, m \ge 1$, any volumes $v_1, \cdots, v_n$ and $V_1, \cdots, V_m$, and any coefficients $\lambda_1, \cdots, \lambda_{n+m}$, the matrix may then be written as:

$$\mathbf{\Sigma}_{n+m} = \begin{pmatrix} \mathbf{A}_{n,n} & \mathbf{B}_{n,m} \\ \mathbf{B}^{\mathbf{t}}_{m,n} & \mathbf{C}_{m,m} \end{pmatrix} \tag{11}$$

where $\mathbf{A}_{n,n} = \left[\overline{c}_1(v_i, v_j)\right]_{1\le i,\, j\le n}$, $\mathbf{B}_{n,m} = \left[\overline{c}_{1,2}(v_j, V_k)\right]_{\substack{1\le j\le n \\ 1\le k\le m}}$, and $\mathbf{C}_{m,m} = \left[\overline{c}_2(V_k, V_q)\right]_{1\le k,\, q\le m}$.

While the auto-covariance functions $\overline{c}_1$ and $\overline{c}_2$ of the attributes must be positive definite, the cross-covariance function $\overline{c}_{1,2}$ does not need to satisfy this property. The practice of modeling the coregionalization matrix through (1) uses this requirement by default. We can exploit the positive definite property of $\overline{c}_{1,2}$ to derive sufficient conditions on the validity of the matrix $\mathbf{\Sigma}_{n+m}$ for any $n, m \ge 1$.
Where the following conditions are true:

1. $\overline{c}_{1,2}$ is positive semidefinite,
2. $\eta_1 > 0$ and $\eta_2 > 0$ such that $\eta_1\eta_2 > 1$ and $\overline{c}_1(\mathbf{h}) - \eta_1\overline{c}_{1,2}(\mathbf{h})$ and $\overline{c}_2(\mathbf{h}) - \eta_2\overline{c}_{1,2}(\mathbf{h})$ are positive semidefinite,

then the matrix $\mathbf{\Sigma}_{n+m}$ given in (9) is positive semidefinite. Indeed, for any vector $\boldsymbol{\lambda}_{n+m} = (\lambda_1, \cdots, \lambda_{n+m})$, one has

$$\boldsymbol{\lambda}_{n+m}^{\mathbf{t}}\boldsymbol{\Sigma}_{n+m}\boldsymbol{\lambda}_{n+m}=\boldsymbol{\lambda}_{n+m}^{\mathbf{t}}\begin{pmatrix}\left[\overline{c}_{1,2}\left(v_i,v_j\right)\right]_{1\le i,j\le n} & \left[\overline{c}_{1,2}\left(v_j,V_k\right)\right]_{\substack{1\le j\le n\\1\le k\le m}}\\ \left[c\overline{c}_{1,2}\left(V_k,v_j\right)\right]^{T}_{\substack{1\le k\le m\\1\le j\le n}} & \left[\overline{c}_{1,2}\left(V_k,V_q\right)\right]_{1\le k,q\le m}\end{pmatrix}\boldsymbol{\lambda}_{n+m}$$

$$+\boldsymbol{\lambda}_{n+m}^{\mathbf{t}}\begin{pmatrix}\left[\overline{c}_1\left(v_i,v_j\right)-\eta_1\overline{c}_{1,2}\left(v_i,v_j\right)\right]_{1\le i,j\le n} & \left[\mathbf{0}\right]_{\substack{1\le j\le n\\1\le k\le m}}\\ \left[\mathbf{0}\right]^{T}_{\substack{1\le k\le m\\1\le j\le n}} & \left[\overline{c}_2\left(V_k,V_q\right)-\eta_2\overline{c}_{1,2}\left(V_k,V_q\right)\right]_{1\le k,q\le m}\end{pmatrix}\boldsymbol{\lambda}_{n+m}$$

$$\ge 0$$

where the latter comes from the positive semidefinite assumptions. This is because the coregionalization matrix can be written as

$$\begin{pmatrix}\overline{c}_1(\mathbf{h}) & \overline{c}_{1,2}(\mathbf{h})\\ \overline{c}_{1,2}(\mathbf{h}) & \overline{c}_2(\mathbf{h})\end{pmatrix}=\begin{pmatrix}\eta_1\overline{c}_{1,2}(\mathbf{h}) & \overline{c}_{1,2}(\mathbf{h})\\ \overline{c}_{1,2}(\mathbf{h}) & \eta_2\overline{c}_{1,2}(\mathbf{h})\end{pmatrix}+\begin{pmatrix}\overline{c}_1(\mathbf{h})-\eta_1\overline{c}_{1,2}(\mathbf{h}) & 0\\ 0 & \overline{c}_2(\mathbf{h})-\eta_2\overline{c}_{1,2}(\mathbf{h})\end{pmatrix}$$

$$=\begin{pmatrix}\eta_1 & 1\\ 1 & \eta_2\end{pmatrix}\overline{c}_{1,2}(\mathbf{h})+\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\left[\overline{c}_1(\mathbf{h})-\eta_1\overline{c}_{1,2}(\mathbf{h})\right]+\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}\left[\overline{c}_2(\mathbf{h})-\eta_2\overline{c}_{1,2}(\mathbf{h})\right] \quad (12),$$

and with $\begin{pmatrix}\eta_1 & 1\\ 1 & \eta_2\end{pmatrix}$, becomes positive definite. Since the functions $\eta_1\overline{c}_{1,2}(\mathbf{h})$ and $\eta_2\overline{c}_{1,2}(\mathbf{h})$ are positive definite, investigating the positive semidefiniteness of the functions $\overline{c}_1(\mathbf{h})-\eta_1\overline{c}_{1,2}(\mathbf{h})$ and $\overline{c}_2(\mathbf{h})-\eta_2\overline{c}_{1,2}(\mathbf{h})$ is relevant under the notion of the positive semidefiniteness ordering (Horn and Johnson 1985 p. 469). Bochner's theorem (Christakos 1992, p. 64) can be exploited to ensure that $\overline{c}_1(\mathbf{h})-\eta_1\overline{c}_{1,2}(\mathbf{h})$ and $\overline{c}_2(\mathbf{h})-\eta_2\overline{c}_{1,2}(\mathbf{h})$ are positive semidefinite functions as

$$g_k(\boldsymbol{\omega},\eta_k)=(2\pi)^{-d}\int_{\mathbb{R}^d}e^{-i\boldsymbol{\omega}\bullet\mathbf{h}}\left[\overline{c}_k(\mathbf{h})-\eta_k\overline{c}_{1,2}(\mathbf{h})\right]d\mathbf{h}\ge 0\ \ \forall\boldsymbol{\omega}\ \text{for}\ k=1,2. \quad (13)$$

If we restrict ourselves to isotropic covariance functions, Schoenberg (1938) arrived at

$$g_k(\omega,\eta_k)=\frac{\omega^{1-d/2}}{(2\pi)^{d/2}}\int_{[0,\infty)}r^{d/2}J_{(d-2)/2}(r\omega)\left[\overline{c}_k(r)-\eta_k\overline{c}_{1,2}(r)\right]dr\ge 0\ \ \forall\omega\ge 0 \quad (14)$$

where J_ν denotes the Bessel function of first kind of order ν. The spectral density can be a useful tool in verifying valid covariance models (Christakos, 1984). We have the following lemma.

Lemma Assume a positive number η_k exists such that $g_k(\boldsymbol{\omega},\eta_k)\ge 0$ and $\overline{c}_k(\mathbf{h})-\eta_k\overline{c}_{1,2}(\mathbf{h})$ is positive semidefinite. Then where $0\le\eta<\eta_k$, one has $g_k(\boldsymbol{\omega},\eta)\ge 0$ and $g_k(\boldsymbol{\omega},\eta_k)\ge g_k(\boldsymbol{\omega},\eta)$ for any $\boldsymbol{\omega}$.

Proof. By definition, we have

$$
\begin{aligned}
g_k(\boldsymbol{\omega},\eta) &= (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-i\boldsymbol{\omega}\bullet\mathbf{h}} \left[\overline{c}_k(\mathbf{h}) - (\eta_k - \eta_k + \eta)\overline{c}_{1,2}(\mathbf{h}) \right] d\mathbf{h} \\
&= (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-i\boldsymbol{\omega}\bullet\mathbf{h}} \left[\overline{c}_k(\mathbf{h}) - \eta_k \overline{c}_{1,2}(\mathbf{h}) \right] d\mathbf{h} + (2\pi)^{-d} \int_{\mathbb{R}^d} e^{-i\boldsymbol{\omega}\bullet\mathbf{h}} \left[(\eta_k - \eta)\overline{c}_{1,2}(\mathbf{h}) \right] d\mathbf{h}. \\
&\leq g_k(\boldsymbol{\omega},\eta_k)
\end{aligned}
$$

This shows that $g_k(\boldsymbol{\omega},\eta) \geq 0$ since $g_k(\boldsymbol{\omega},\eta_k) \geq 0$. The proof is completed by the fact that $\overline{c}_{1,2}(\mathbf{h})$ is positive semidefinite. □

Algorithm for positive definiteness

According to equation (11), the technique to ensure positive definiteness is

$$g_k(\boldsymbol{\omega},\eta_k) \geq 0 \quad \text{for } k = 1,2. \tag{15}$$

Let $S_k \subset \mathbb{R}_+$ be defined as the set of numbers $\eta_k \geq 0$ such that $g_k(\boldsymbol{\omega},\eta_k) \geq 0$, that is

$$S_k = \left\{ \eta_k \geq 0 \quad \text{such that } g_k(\boldsymbol{\omega},\eta_k) \geq 0 \right\}. \tag{16}$$

Clearly, $S_k \neq \varnothing$ since $0 \in S_k$ for $k = 1,2.$ We formulate the positive definiteness property to find the subsets S_1 and S_2.

To this end, and using the result given in Lemma 1, we propose an iterative procedure that focuses on each parameter $\eta_k > 0$ as follows:

Problem Find $\eta_k^{[\max]} > 0$ such that $g_k\left(\boldsymbol{\omega},\eta_k^{[\max]}\right) \geq 0$ for $k = 1,2$ where

$$\eta_k^{[\max]} = \max_{\eta_k > 0}(S_k). \tag{17}$$

Then, for the positive definiteness property to be ensured,

$$\eta_1^{[\max]} \eta_2^{[\max]} \geq 1. \tag{18}$$

In case of the Markov model of coregionalization, if $g(\boldsymbol{\omega},\eta)$ denotes the spectral density of $\overline{c}_1(\mathbf{h}) - \eta\overline{c}_2(\mathbf{h})$, we ensure the validity of the model if

$$\eta^{[\max]} > \rho^2 \quad \text{and} \quad g\left(\boldsymbol{\omega},\eta^{[\max]}\right) \geq 0. \tag{19}$$

Spatial statistics may be used to demonstrate the necessary and sufficient conditions to make the following coregionalization matrix positive semidefinite:

$$\boldsymbol{\Sigma}_{n+m} = \begin{pmatrix} \mathbf{A}_{n,n} & \mathbf{B}_{n,m} \\ \mathbf{B}_{m,n}^{\mathbf{t}} & \mathbf{C}_{m,m} \end{pmatrix}$$

where $\mathbf{A}_{n,n} = \left[\overline{c}_1(v_i, v_j) \right]_{1 \leq i,j \leq n}$, $\mathbf{B}_{n,m} = \left[\overline{c}_{1,2}(v_j, V_k) \right]_{\substack{1 \leq j \leq n \\ 1 \leq k \leq m}}$ and $\mathbf{C}_{m,m} = \left[\overline{c}_2(V_k, V_q) \right]_{1 \leq k,q \leq m}$. This concurs with Cressie (1993, p. 141), who noted that building valid, flexible models for auto- and cross-covariance and fitting them to available data is a challenge that requires further research.
Assume the following conditions:

1. $c_1(\mathbf{h})$ is positive definite or $\mathbf{A}_{n,n}$ is positive definite.
2. $c_2(\mathbf{h})$ is positive definite or $\mathbf{C}_{m,m}$ is positive definite.

Then a *necessary and sufficient condition* (Horn and Johnson 1985, p. 473) to ensure that the matrix $\mathbf{\Sigma}_{n+m}$ is positive semidefinite may be obtained by ensuring that, for any $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^m$, the following inequality holds

$$\left(\mathbf{x}^T \mathbf{A}_{n,n} \mathbf{x}\right)\left(\mathbf{y}^T \mathbf{C}_{m,m} \mathbf{y}\right) \geq \left(\mathbf{x}^T \mathbf{B}_{n,m} \mathbf{y}\right)^2 . \tag{20}$$

Finding mild conditions under which the inequality (20) is satisfied can be investigated using the spectral densities of the random fields.

Generalizing the Markov model of coregionalization

The Markov model of coregionalization as described in (5) is written as

$$\begin{pmatrix} \overline{c}_1(\mathbf{h}) & \overline{c}_{1,2}(\mathbf{h}) \\ \overline{c}_{1,2}(\mathbf{h}) & \overline{c}_2(\mathbf{h}) \end{pmatrix} = \begin{pmatrix} \rho^2 & \rho \\ \rho & 1 \end{pmatrix} \overline{c}_2(\mathbf{h}) + \begin{pmatrix} 1-\rho^2 & 0 \\ 0 & 0 \end{pmatrix} c_r(\mathbf{h}) \tag{21}$$

under the regression model

$$Z_1(\mathbf{u}) = \rho Z_2(\mathbf{u}) + R(\mathbf{u})$$

where $Z_2(\mathbf{u})$ is independent of the process $R(\mathbf{u})$. This model, then, requires that the auto-covariance model $c_1(\mathbf{h})$ be described by a precise linear of $c_2(\mathbf{h})$ and $c_r(\mathbf{h})$. This is obviously too restrictive and may lead to unrealistic models of coregionalization if the observed range of the sample estimates of $\overline{c}_1(\mathbf{h})$ is close to that of $\overline{c}_2(\mathbf{h})$ and if the correlation coefficient ρ is high. Our formulation of a valid model of coregionalization in the Markov model framework is

$$\begin{pmatrix} \overline{c}_1(\mathbf{h}) & \overline{c}_{1,2}(\mathbf{h}) \\ \overline{c}_{1,2}(\mathbf{h}) & \overline{c}_2(\mathbf{h}) \end{pmatrix} = \begin{pmatrix} \eta & \rho \\ \rho & 1 \end{pmatrix} \overline{c}_2(\mathbf{h}) + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \left[\overline{c}_1(\mathbf{h}) - \eta \overline{c}_2(\mathbf{h})\right] \tag{22}$$

where η is such that

$$1 > \eta \geq \rho^2 \quad \text{and} \quad \overline{c}_1(\mathbf{h}) - \eta \overline{c}_2(\mathbf{h}) \quad \text{is positive definite}. \tag{23}$$

This significantly improves the Markov model in that $\overline{c}_1(\mathbf{h})$ is not required to contain $\overline{c}_2(\mathbf{h})$. The positive definiteness property can be observed by writing the coregionalization matrix as

$$\begin{pmatrix} \overline{c}_1(\mathbf{h}) & \overline{c}_{1,2}(\mathbf{h}) \\ \overline{c}_{1,2}(\mathbf{h}) & \overline{c}_2(\mathbf{h}) \end{pmatrix} = \begin{pmatrix} \eta & 1 \\ 1 & \rho^{-1} \end{pmatrix} \overline{c}_{1,2}(\mathbf{h}) + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \left[\overline{c}_1(\mathbf{h}) - \eta \overline{c}_{1,2}(\mathbf{h})\right].$$

If $1 > \eta \geq \rho$, the result is a positive semidefinite matrix with $\overline{c}_{1,2}$ and $\overline{c}_1(\mathbf{h}) - \eta \overline{c}_{1,2}(\mathbf{h})$ being valid covariance functions.

More generally, the linear model of coregionalization (1) writes

$$\overline{c}_1(\mathbf{h}) = \sum_{k=1}^{L} b_{11}^{(k)} \varphi_k(\mathbf{h}), \quad \overline{c}_2(\mathbf{h}) = \sum_{k=1}^{L} b_{22}^{(k)} \varphi_k(\mathbf{h}) \quad \text{and } \overline{c}_{12}(\mathbf{h}) = \sum_{k=1}^{L} b_{12}^{(k)} \varphi_k(\mathbf{h})$$

with the positive definiteness property satisfies under:

- (A1) $\left|b_{11}^k b_{22}^k\right|^{\frac{1}{2}} \geq \left|b_{12}^k\right|$ for each $k \geq 1$
- (A2) b_{11}^k and b_{22}^k having the same sign for each $k \geq 1$

The requirements (A1) and (A2) are too restrictive. Indeed, consider the following coregionalization model

$$\mathbf{M}(\mathbf{h}) = \begin{pmatrix} \overline{c}_1(\mathbf{h}) & \overline{c}_{1,2}(\mathbf{h}) \\ \overline{c}_{1,2}(\mathbf{h}) & \overline{c}_2(\mathbf{h}) \end{pmatrix}$$

with

$$\begin{aligned} \overline{c}_1(\mathbf{h}) &= 0.75\ \mathrm{Exp}\left(\frac{\|\mathbf{h}\|}{5}\right) + 0.25\ \mathrm{Gaus}\left(\frac{\|\mathbf{h}\|}{10}\right), \\ \overline{c}_2(\mathbf{h}) &= 0.4\ \mathrm{Exp}\left(\frac{\|\mathbf{h}\|}{5}\right) + 0.6\ \mathrm{Gaus}\left(\frac{\|\mathbf{h}\|}{10}\right) \\ \overline{c}_{12}(\mathbf{h}) &= 0.025\ \mathrm{Exp}\left(\frac{\|\mathbf{h}\|}{5}\right) + 0.7\ \mathrm{Gaus}\left(\frac{\|\mathbf{h}\|}{10}\right) \end{aligned} \tag{24}$$

The positive definiteness property of the matrix $\mathbf{M}(\mathbf{h})$ cannot be obtained through the LMC since

$$\begin{vmatrix} 0.25 & 0.7 \\ 0.7 & 0.6 \end{vmatrix} < 0 .$$

Although the Markov model of coregionalization cannot be used since $\overline{c}_{12}$ is not proportional to $\overline{c}_2$, one can easily find real numbers satisfying $\eta_1\eta_2 > 1$ such that $\overline{c}_1(\mathbf{h}) - \eta_1\overline{c}_{1,2}(\mathbf{h})$ and $\overline{c}_2(\mathbf{h}) - \eta_2\overline{c}_{1,2}(\mathbf{h})$ are positive semidefinite functions.

Numerical results

In this section we consider (1) a theoretical example of a model of coregionalization with a relatively high correlation coefficient, and (2) real data and show applications of the technique.

4.1 Theoretical example of valid model of coregionalization

Consider the following example as plotted in Figure 2:

$$\begin{aligned} \overline{c}_1(\mathbf{h}) &= \mathrm{Sph}\left(\frac{\|\mathbf{h}\|}{5}\right), \quad \overline{c}_2(\mathbf{h}) = \mathrm{Gaus}\left(\frac{\|\mathbf{h}\|}{5}\right) \\ \overline{c}_{1,2}(\mathbf{h}) &= 0.75 c_2(\mathbf{h}) \end{aligned} \tag{25}$$

with a high correlation coefficient $\rho = 0.75$ between the attributes. The correlation length $\xi = 5$ is the same for all the auto/cross-covariance models and the exact values of ξ are not important since they only lead to a rescaling of the distance. We also observe that covariance $\overline{c}_1(\mathbf{h})$ has a different degree of smoothness at the origin than $\overline{c}_2(\mathbf{h})$. Furthermore, it goes beyond the Markov model requirement which would express $\overline{c}_1(\mathbf{h})$ as

$$\overline{c}_1(\mathbf{h}) = 0.5625\mathrm{Gaus}\left(\frac{\|\mathbf{h}\|}{5}\right) + 0.4375 c_r(\mathbf{h})$$

for a specified covariance model $c_r(\mathbf{h})$ (exponential, spherical, Gaussian etc.) and could lead to an arbitrarily modeling of $\overline{c}_1(\mathbf{h})$.

Write that

$$\begin{pmatrix} \overline{c}_1(\mathbf{h}) & \overline{c}_{1,2}(\mathbf{h}) \\ \overline{c}_{1,2}(\mathbf{h}) & \overline{c}_2(\mathbf{h}) \end{pmatrix} = \begin{pmatrix} \eta & 0.75 \\ 0.75 & 1 \end{pmatrix} \mathrm{Gaus}\left(\frac{\|\mathbf{h}\|}{5}\right) + \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \left[\mathrm{Sph}\left(\frac{\|\mathbf{h}\|}{5}\right) - \eta \mathrm{Gaus}\left(\frac{\|\mathbf{h}\|}{5}\right) \right].$$

Thus our investigation of the positive definiteness is reduced to find a positive number η such that

$$\eta > \rho^2 = 0.5625 \quad \text{and} \quad \mathrm{Sphe}\left(\frac{\|\mathbf{h}\|}{5}\right) - \eta \mathrm{Gaus}\left(\frac{\|\mathbf{h}\|}{5}\right) \quad \text{is positive semidefinite}.$$

To do so, we start the convergence of our algorithm with an arbitrary value of $\eta^{(0)} = 0.1$ and gradually increase this value while $g\left(\boldsymbol{\omega}, \eta^{(0)}\right) \geq 0$. The algorithm continues until a value of η is obtained where $\boldsymbol{\omega}_0 \in \mathbb{R}^d$ such that $g\left(\boldsymbol{\omega}_0, \eta\right) < 0$.

The plot of the spectral density of the function $\mathrm{Sphe}\left(\frac{\|\mathbf{h}\|}{5}\right) - \eta^{[\max]} \mathrm{Gaus}\left(\frac{\|\mathbf{h}\|}{5}\right)$ is nonnegative for values of $\eta^{[\max]} = 0.6$. This indicates that the coregionalization model (25) is positive semidefinite.

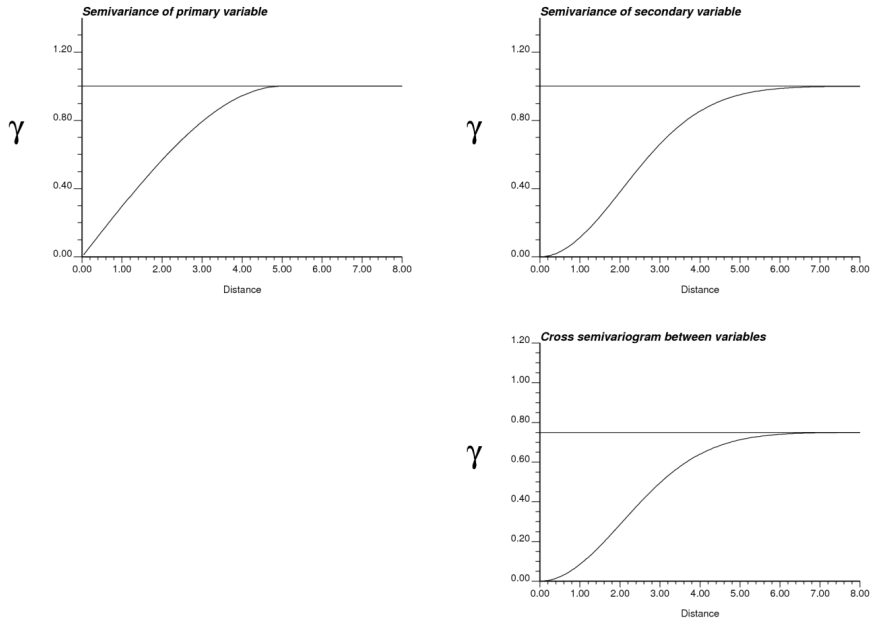

Figure 2: Theoretical example of a generalized Markov model of coregionalization given in (25) with a high correlation coefficient $\rho = 0.75$.

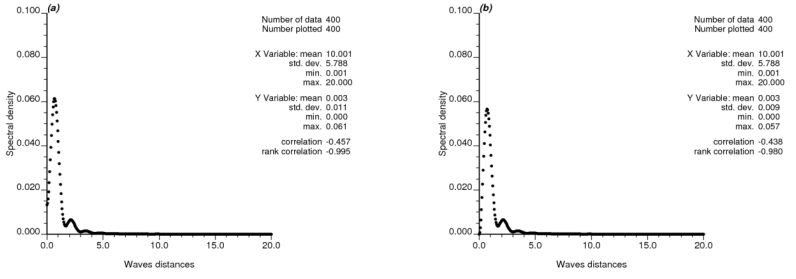

Figure 3: Spectral density of $\mathrm{Sphe}\left(\frac{\|\mathbf{h}\|}{5}\right) - \eta \mathrm{Gaus}\left(\frac{\|\mathbf{h}\|}{5}\right)$; **(a)** $\eta = 0.58$ and **(b)** $\eta = 0.6$.

Soil data application

To illustrate applications of the proposed technique for ensuring the validity of a linear model of coregionalization across scales, we consider the well-known Jura data (Goovaerts, 1997) using the cobalt (Co) and nitrate (Ni) data sets as primary and secondary variables of sample size $n = 100$. Using classical geostatistical techniques, we simulate the nitrate attribute at the point scale using SGSIM (Deutsch and Journel, 1992) and obtain regularization at blocks support of size $0.48 \times 0.5 \mathrm{m}^2$. We use the resulting block scale nitrate datasets as secondary data in simple cokriging prediction for the cobalt data. The location maps of the data sets are shown in Figure 4 below. The omni-directional framework and the hypothesis of isotropy will be considered as the data sets do not indicate a preferential direction of variability.

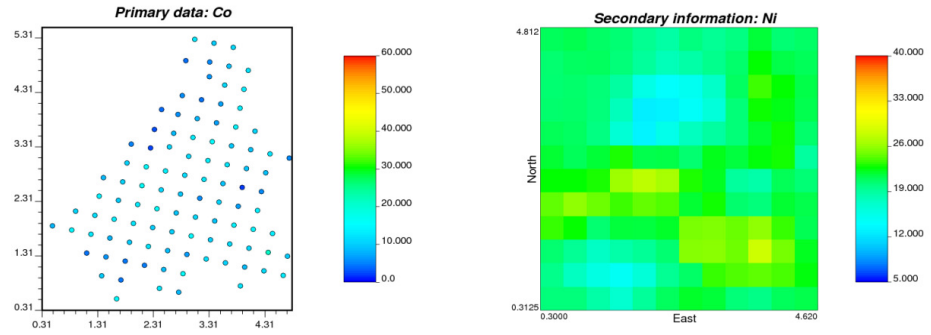

Figure 4: Location maps of the cobalt and nitrate data sets.

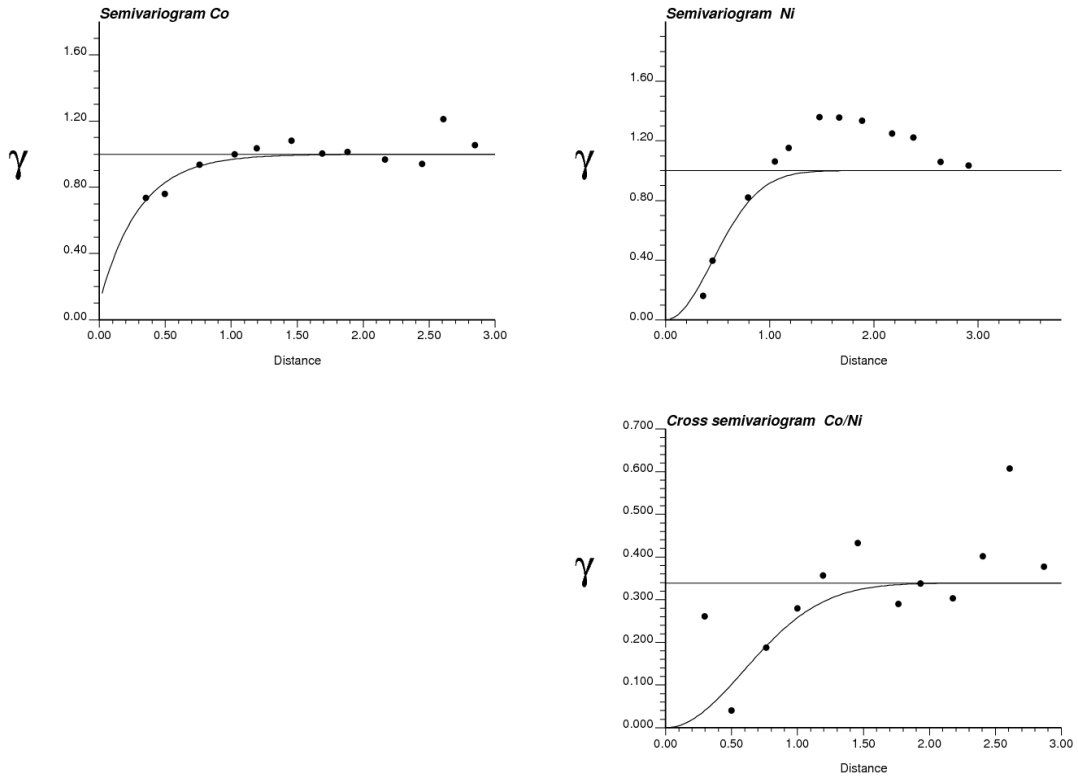

Figure 5: Variography of the data sets and modeling.

Consider the following methodology to infer the coregionalization matrix. The auto- and cross-semivariogram models of the standardized data sets are displayed in Figure 5. The covariance models are given by

$$\overline{c}_1(\mathbf{h}) = 0.1\ \delta(\mathbf{h}) + 0.9\ \mathrm{Exp}\left(\frac{\|\mathbf{h}\|}{0.9}\right),\ \ \overline{c}_2(\mathbf{h}) = 0.00\delta(\mathbf{h}) + \mathrm{Gaus}\left(\frac{\|\mathbf{h}\|}{1.1}\right)$$
$$\overline{c}_{1,2}(\mathbf{h}) = 0.00\delta(\mathbf{h}) + 0.3392\ \mathrm{Gaus}\left(\frac{\|\mathbf{h}\|}{1.45}\right) \tag{26}$$

where Exp and Gaus denote the exponential and Gaussian semivariogram models of range and sill equal to one and $\delta(\mathbf{h})$ represents the Kronecker delta function. The cross-spatial dependence is modeled through a positive definite Gaussian model with a longer range than those of the auto-covariance models as given by (26). Hence the Markov model approach cannot be applied in this case. Consequently, the positive definiteness of the matrix

$$\mathbf{\Sigma}_{n+m} = \begin{pmatrix} \mathbf{A}_{n,n} & \mathbf{B}_{n,m} \\ \mathbf{B}_{m,n}^{\mathbf{t}} & \mathbf{C}_{m,m} \end{pmatrix}$$

as defined in Equation (11) will have less zeros in the matrix $\mathbf{B}_{n,m}$ than in the square matrices $\mathbf{A}_{n,n}$ and $\mathbf{C}_{m,m}$. Ensuring the validity of the model, then, presents a challenge. As described in (12), positive numbers η_1 and η_2 satisfying $\eta_1\eta_2 \geq 1$ must be used such that the functions $\overline{c}_1(\mathbf{h}) - \eta_1\overline{c}_{1,2}(\mathbf{h})$ and $\overline{c}_2(\mathbf{h}) - \eta_2\overline{c}_{1,2}(\mathbf{h})$ satisfy the positive semidefinite requirement. An algorithm with the following starting point $\eta_1^{[0]} = \eta_2^{[0]} = 0.1$ is then employed by gradually increasing the value of η_k. The search stops when the first value of η_k is reached such that $\boldsymbol{\omega}_0 \in \mathbb{R}^d$ where $g_k(\boldsymbol{\omega}_0, \eta_k) < 0$. Ultimately, we obtain

$$\eta_1^{[\max]} = 0.68 \quad \text{and} \quad \eta_2^{[\max]} = 1.695\,. \tag{27}$$

Since $\eta_1^{[\max]}\eta_2^{[\max]} > 1$, we conclude that the coregionalization model (26) is positive definite. Figure 6 gives the plots of the spectral densities $g_k(\boldsymbol{\omega}_0, \eta_k^{[\max]})$ for $k = 1, 2.$

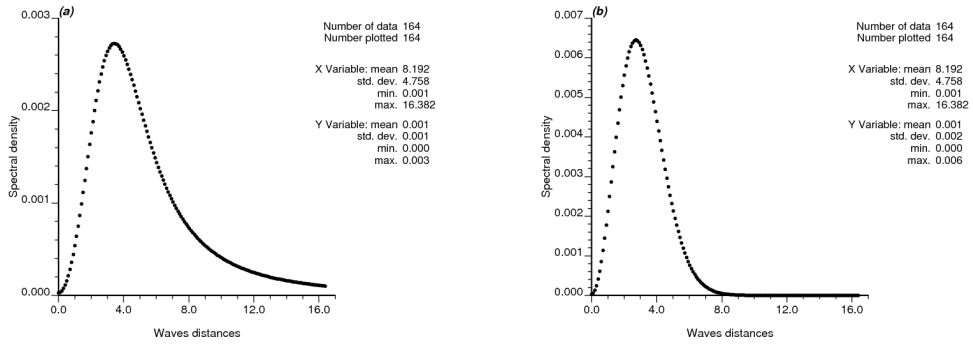

Figure 6: Plot of the spectral density functions; (a) $\overline{c}_1(\mathbf{h}) - \eta_1^{[\max]}\overline{c}_{1,2}(\mathbf{h})$ and (b) $\overline{c}_2(\mathbf{h}) - \eta_2^{[\max]}\overline{c}_{1,2}(\mathbf{h})$.

We perform leave-one-out cross-validation using the proposed method where each observation is removed in turn and the remaining observations are used for prediction. The performance of the predictors is evaluated using the following statistical measures: mean error (ME), mean absolute error (MAE), root mean square error (RMSE), Pearson's linear correlation coefficient (R) and the mean of the mean standardized squared deviation ratio (MSSDR). The criterions are computed as:

$$\text{MAE}=n^{-1}\sum_{i=1}^{n}\left|z(\mathbf{s}_i)-\hat{z}(\mathbf{s}_i)\right|,\quad \text{RMSE}=n^{-1/2}\left[\sum_{i=1}^{n}\left(z(\mathbf{s}_i)-\hat{z}(\mathbf{s}_i)\right)^2\right]^{1/2}$$
$$\text{ME}=n^{-1}\sum_{i=1}^{n}\left(z(\mathbf{s}_i)-\hat{z}(\mathbf{s}_i)\right),\quad \text{MSSDR}=n^{-1}\sum_{i=1}^{n}\sigma^{-2}(\mathbf{s}_i)\left(z(\mathbf{s}_i)-\hat{z}(\mathbf{s}_i)\right)^2 \tag{28}$$

where $z(\mathbf{s}_i)$ is the observed data, $\hat{z}(\mathbf{s}_i)$ is the cokriged/kriged estimate in the cross-validation and $\sigma^2(\mathbf{s}_i)$ is the cokriging/kriging variance. The results are reported in Table 1 and 2 and show improvement of the cross-validation statistics when using the valid model of coregionalization (26) for both data sets. The goodness measure of the fit MSSDR shows encouraging results for the primary data.

Table 1: Statistics of the leave-one-out cross-validation performance through the simple cokriging/kriging prediction on the secondary data.

Criterion	ME	MAE	RMSE	R	MSSDR
SK using $\overline{c}_2(\mathbf{h})$	0.00	0.57	0.75	0.97	2.45
SCK using (26)	0.00	0.55	0.73	0.97	2.47

Table 2: Statistics of the leave-one-out cross-validation performance through the simple cokriging/kriging prediction on the primary data.

Criterion	ME	MAE	RMSE	R	MSSDR
SK using $\overline{c}_1(\mathbf{h})$	0.01	2.53	3.11	0.46	0.98
SCK using (26)	-0.06	2.47	3.03	0.51	0.95

As expected, there is not much improvement for the exhaustive secondary data. However, the estimates of the cobalt data are significantly improved with the leave-one-out cross-validation method producing the lowest MAE and RMSE and a higher correlation coefficient. In addition, Figure 7 demonstrates the prediction and variance maps obtained for the cobalt data. Although there is a lack of comparison with other relevant techniques, the overall pattern is promising.

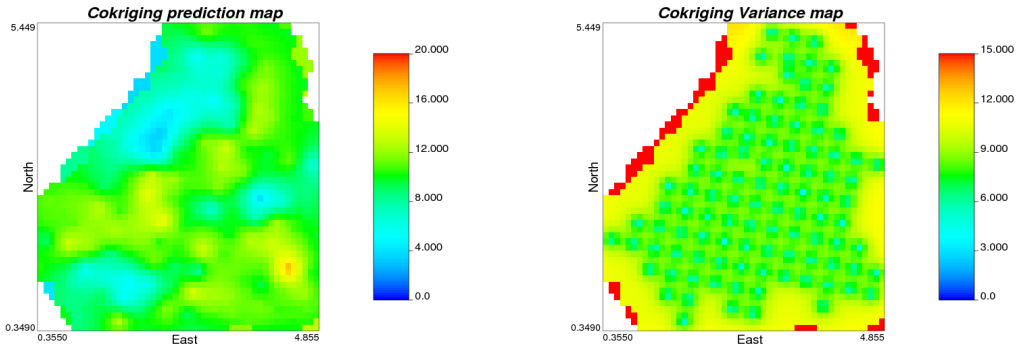

Figure 7: Simple cokriging prediction and the associated variance map of the primary data.

Conclusions

In this work we propose a method for obtaining a valid model of coregionalization for attributes observed at different resolutions. The main goal of this paper is to avoid the somewhat drastic approximation produced by the Markov model of coregionalization upon integration of larger-scale secondary attributes. The methodology described herein generalizes the linear model of coregionalization classically used for modeling the auto- and cross-spatial dependence. The flexibility of this approach includes the ability to accurately model the co-dependence structures in a careful manner to satisfy the positive definiteness requirement. We apply simple cokriging using the proposed model of coregionalization, to the cobalt nitrate data without having to build any set of collocated data and then computed the associated correlation coefficient. Finally, with simple cokriging, we use leave-one-out cross-validation data to assess the quality of the results.

References

Christakos G (1984) On the problem of permissible covariance and variogram models. Wat Resour Res 20: 251-265.

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

Cressie NAC (1993) Statistics for Spatial data: Wiley, New York.

Cressie N, Wikle CK (1998) The variance-based cross variogram: You can add apples and oranges. Math Geol, 30(7): 789-799.

Deutsch CV, Journel AG (1998) GSLIB: Geostatistical Software Library and User's Guide, 2nd edition, Oxford University Press, New York.

Doyen PM, Den Boer LD, Pillet WR (1996) Seismic porosity mapping in the Ekofisk Field Using a New Form of Collocated Kriging. In: SPE paper 36498.

Doyen PM, Psaila DE, Den Boer LD (1997) Reconciling data at seismic and well log scales in 3D Earth modeling. In: SPE paper 38698.

Frykman P, Deutsch CV (1999) Geostatistical scaling laws applied to core and log data. SPE 56822 presented at the SPE Annual Technical Conference and Exhibition, Houston, TX.

Goovaerts P (1997) Geostatistics for natural resources evaluation: Oxford University Press, New York.

Haas A, Dubrule O (1994) Geostatistical inversion: A sequential method of stochastic reservoir modeling constrained by seismic data. First Break 12(11), p. 561-569.

Horn R, Johnson C (1985) Matrix Analysis. Cambridge University Press.

Johannesson G, Cressie N (2004), Inference in a coregionalization model, in Sanchez et al. (eds.) GeoENV IV-Geostatistics for Environmental Applications: Kluwer Academic, Dordrecht, p. 319–330.

Journel AG and Huijbregts J (1978) Mining Geostatiscis, London: Academic Press.

Journel AG (1999) Markov models for cross covariances. Mathematical Geology, 31(8): 955-964.

Kupfersberger H, Deutsch CV and Journel AG (1998) Deriving constraints on small-scale variograms due to variograms of large scale-data. Mathematical Geology, 38(7): 801-819.

Leuangthong O (2005). The promises and pitfalls of direct simulation, In: O. Leuangthong and C. V. Deutsch (eds.), Geostatistiscs Banff 2004: 305-314.

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

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

Myers DE (1983) Estimation of linear combinations and co-kriging: Mathematical Geology, 15(5): 633-637.

Oliver DS (2003) Gaussian Cosimulation: Modelling of the Cross-Covariance. Mathematical Geology, 35(6): 681-698.

Shmaryan LE, Journel AG (1999) Two Markov models and their application. Math Geol, 31(8): 965-988.

Soares A (2001) Direct Sequential Simulation and Cosimulation, Mathematical Geology, 33(8): 911-926.

Tjolsen CB, Johnsen G, Rtseth A, Damsleth E (1995). Seismic data can improve the stochastic facies modeling significantly. In: SPE 30567, p. 373-379.

Yao T, Journel AG (2000) Integrating seismic attributes maps and well logs for porosity modeling in a west Texas carbonate reservoir: addressing the scale and precision problem. Journal of Petroleum Science and Engineering, 28, p. 65-79.