Mathematical Geology, Vol. 26, No. 2, 1994

Universal Cokriging Under Intrinsic Coregionalization[1]

Jeffrey D. Helterbrand[2,3] and Noel Cressie[2]

Under the intrinsic coregionalization model, if both primary and secondary measurements are available at all sample locations, the conventional geostatistical wisdom is that cokriging provides exactly the same solution as univariate kriging on the primary process alone. However, recent examples have been given where nonzero secondary cokriging weights have occurred under this spatial dependence structure. This note identifies the conditions under which secondary information is useful under the assumption of intrinsic coregionalization. An illustration is given using a dataset of plutonium and americium concentrations collected from a region of the Nevada Test Site.

KEY WORDS: multivariate spatial dependence models, partial covariance, restricted linear models.

INTRODUCTION

Cokriging produces predictors that use not only information from direct measurements of the spatial component process being considered, but also the information from measurements of a secondary component process. Let $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_n$ be sample locations where measurements on both component processes are available. For the time being, we shall assume no missing data at the sample locations. Let $Z_i(\mathbf{s}_k)$ be the measurement for the ith component at location $\mathbf{s}_k$ ($k = 1, \ldots, n$; $i = 1, 2$), and $\mathbf{Z}(\mathbf{s}_k)$ denote the bivariate vector at location $\mathbf{s}_k$. We shall refer to the first component process $Z_1(\cdot)$ as the primary process and $Z_2(\cdot)$ as the secondary process. Define the data vector $\mathbf{Z} \equiv (\mathbf{Z}(\mathbf{s}_1)^t, \mathbf{Z}(\mathbf{s}_2)^t, \ldots, \mathbf{Z}(\mathbf{s}_n)^t)^t$. Suppose we are interested in predicting the primary process $Z_1(\mathbf{s}_0)$ at unsampled location $\mathbf{s}_0$.

The cokriging predictor for $Z_1(\mathbf{s}_0)$ given the sample information is of the form,

$$\hat{Z}_1(\mathbf{s}_0) = \sum_{i=1}^{2} \sum_{k=1}^{n} \lambda_{ik} Z_i(\mathbf{s}_k) \tag{1}$$

[1]Received 11 November 1992; accepted 6 January 1993.

[2]Department of Statistics, Iowa State University, Ames, Iowa 50011.

[3]Present address: Jeffrey D. Helterbrand, Senior Scientist, Lilly Research Laboratories, Indianapolis, Indiana 46285.

0882-8121/94/0200-0205$07.00/1

where the cokriging weights $\{\lambda_{ik}: k = 1, \ldots, n; i = 1, 2\}$ satisfy the uniform unbiasedness constraint, $E(\Sigma_{i=1}^{2} \Sigma_{k=1}^{n} \lambda_{ik} Z_i(\mathbf{s}_k)) = E(Z_1(\mathbf{s}_0))$. These cokriging weights are selected to minimize

$$E\left(Z_1(\mathbf{s}_0) - \sum_{i=1}^{2} \sum_{k=1}^{n} \lambda_{ik} Z_i(\mathbf{s}_k)\right)^2 \tag{2}$$

subject to the unbiasedness constraint.

To obtain the optimal cokriging weights requires knowledge of models of the spatial correlation for both the primary and the secondary components as well as for the cross-correlation between the two processes. The bivariate spatial dependence model is said to be *intrinsically coregionalized* if the assumed spatial dependence structure is

$$\mathbf{C}(\mathbf{s}_k, \mathbf{s}_m) \equiv \begin{bmatrix} c_{11}(\mathbf{s}_k, \mathbf{s}_m) & c_{12}(\mathbf{s}_k, \mathbf{s}_m) \\ c_{21}(\mathbf{s}_k, \mathbf{s}_m) & c_{22}(\mathbf{s}_k, \mathbf{s}_m) \end{bmatrix} = \begin{bmatrix} 1 & \gamma\rho \\ \gamma\rho & \gamma^2 \end{bmatrix} c(\mathbf{s}_k, \mathbf{s}_m) \tag{3}$$

where $c_{ij}(\mathbf{s}_k, \mathbf{s}_m) = \text{cov}\,[Z_i(\mathbf{s}_k), Z_j(\mathbf{s}_m)]$, $c(\mathbf{s}_k, \mathbf{s}_m)$ is a valid covariance function, $\gamma > 0$, and $|\rho| \leq 1$.

Under the intrinsic coregionalization model, if both primary and secondary measurements are available at all sample data locations, the conventional geostatistical wisdom is that cokriging provides exactly the same solution as univariate kriging on the primary component alone. Specifically, all secondary cokriging weights are zero. In the Summer 1992 issue of *Geostatistics: An Interdisciplinary Geostatistics Newsletter*, an example was given where, under the intrinsic coregionalization assumption with all sample data available, the secondary cokriging weights were nonzero. The question was posed regarding when secondary information is useful under the assumption of intrinsic coregionalization.

In this note we clarify when the secondary cokriging weights may be nonzero. In particular, we investigate properties of the cokriging model under intrinsic coregionalization and determine conditions where cokriging provides the same solution as univariate kriging. We also investigate what happens to the secondary cokriging weights when an additional primary or secondary observation is available at a location $\mathbf{s}_{n+1}$. The results of this paper extend immediately to the m-variate cokriging case (cf. Myers, 1982) if an m-variate intrinsic coregionalization model is assumed. Finally, cokriging under intrinsic coregionalization is illustrated using a dataset of plutonium and americium concentrations collected from a nuclear testing area on the Nevada Test Site. But first, we argue that the intrinsic coregionalization model is very restrictive and often inappropriate. This is done in order to dispel any thoughts one might have that multivariate spatial statistics is an unnecessary complication in which the incorporation of covariate information is not useful.

The Intrinsic Coregionalization Model is Restrictive

The intrinsic coregionalization assumption for a bivariate spatial dependence model is very restrictive and is often not appropriate. Should microscale variation and/or measurement error variation be present in a bivariate spatial dependence model, the model is no longer intrinsically coregionalized. Even when a small nugget effect due to microscale variation is present, the secondary cokriging weights may be nonzero. For a simple example in the plane, suppose that there are two sample locations, $\mathbf{s}_1$, and $\mathbf{s}_2$, one on either side of the point $\mathbf{s}_0$ where the primary value $Z_1(\mathbf{s}_0)$ is to be predicted. The closer location, $\mathbf{s}_1$, is 50 meters due west to $\mathbf{s}_0$, and $\mathbf{s}_2$ is 150 meters due east. We wish to predict $Z_1(\mathbf{s}_0)$ using information $\mathbf{Z}(\mathbf{s}_1)$ and $\mathbf{Z}(\mathbf{s}_2)$. Suppose the bivariate autocovariance function is an intrinsically coregionalized exponential model given by (3) with $\gamma = 1$, $\rho = 0.5$, and spatial-dependence parameter $-3/200$. That is, the bivariate spatial dependence model can be expressed as,

$$\mathbf{C}(\mathbf{s}_k, \mathbf{s}_m) = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \exp\left(-\frac{3}{200}\|\mathbf{s}_k - \mathbf{s}_m\|\right) \tag{4}$$

so that the univariate covariogram is isotropic. The resulting simple cokriging weights (simple cokriging will be discussed momentarily) are $\lambda_{11} = .468$, $\lambda_{12} = .082$, $\lambda_{21} = 0$, and $\lambda_{22} = 0$. We see that the secondary cokriging weights are zero. Now incorporate microscale variation into the bivariate spatial dependence structure. Specifically, consider

$$\mathbf{C}(\mathbf{s}_k, \mathbf{s}_m) = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix} \exp\left(-\frac{3}{200}\|\mathbf{s}_k - \mathbf{s}_m\|\right) + \begin{bmatrix} 0.5 & 0 \\ 0 & 0.25 \end{bmatrix} \exp\left(-\frac{1}{2}\|\mathbf{s}_k - \mathbf{s}_m\|\right) \tag{5}$$

Under this modified spatial dependence model we obtain the simple cokriging weights, $\lambda_{11} = .289$, $\lambda_{12} = .056$, $\lambda_{21} = .072$, and $\lambda_{22} = .011$. Here, the secondary information *is* used in the simple cokriging predictor.

A second remark is that the assumption of intrinsic coregionalization carries with it the assumption of a symmetric spatial cross-covariance function. That is, $c_{ij}(\mathbf{s}_k, \mathbf{s}_m) = c_{ji}(\mathbf{s}_k, \mathbf{s}_m)$. This symmetry need not be true in general (e.g., Ver Hoef and Cressie, 1993).

Third, it is often necessary to estimate unknown parameters present in the valid univariate covariance function $c(\cdot, \cdot)$ given in Eq. (3). In this case, secondary information should be used in conjunction with primary information to obtain the most efficient estimators. Thus, the secondary information should not be ignored.

For these reasons, one should not view multivariate spatial prediction as

an unneccessary complication to a univariate problem. As in standard multivariate analysis, multivariate spatial prediction allows covariate knowledge to yield more precise predictions.

SIMPLE COKRIGING

We can determine the conditions under which the intrinsically coregionalized cokriging model produces zero secondary weights by exploring its properties in the simple cokriging environment. Consider the following decomposition for the component processes. Let

$$Z_i(\mathbf{s}) = \mu_i(\mathbf{s}) + \delta_i(\mathbf{s}); \mathbf{s} \in \mathbf{D}, i = 1, 2 \tag{6}$$

where $\mathbf{D}$ is the spatial domain of interest, $E(Z_i(\mathbf{s})) = \mu_i(\mathbf{s})$, $\mu_i(\mathbf{s})$ is referred to as the large-scale variation term, and $\delta_i(\mathbf{s})$ is a zero-mean process including smooth small-scale variation, microscale variation, and measurement error (see Cressie, 1991, p. 112 for further details). The process δ_i is referred to as the small-scale variation for the ith component.

The large-scale structure for a component process is usually unknown. Often we parameterize $\mu_i(\cdot)$ in a linear manner. That is, we assume

$$\mu_i(\mathbf{s}) = \mathbf{X}_i(\mathbf{s})^T \boldsymbol{\beta}_i \tag{7}$$

where $\mathbf{X}_i(\mathbf{s})$ is most generally a $(p_i \times 1)$ vector of fixed "explanatory" variables for the ith process and $\boldsymbol{\beta}_i$ is a $(p_i \times 1)$ vector of corresponding unknown parameters. The parameters $\boldsymbol{\beta}_i$ are referred to as the large-scale parameters for the ith process; $i = 1, 2$.

In simple cokriging, it is assumed that the $\mu_i(\cdot)$ processes are known in full. For example, in Eq. (7), all $p = p_1 + p_2$ large-scale parameters, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^T, \boldsymbol{\beta}_2^T)^T$, are fixed and known, say equal to $\boldsymbol{\beta}^0 = (\boldsymbol{\beta}_1^{0T}, \boldsymbol{\beta}_2^{0T})^T$. The large-scale parameter space for simple cokriging is thus

$$\Omega = \{\boldsymbol{\beta}^0\} \tag{8}$$

a single point in p-dimensional space. Since $\mu_i(\cdot)$ is known for $i = 1, 2$, simple cokriging focuses on the prediction of $\delta_1(\mathbf{s}_0)$ based on

$$\delta_i(\mathbf{s}_k) = Z_i(\mathbf{s}_k) - \mu_i(\mathbf{s}_k); k = 1, 2, \ldots, n, i = 1, 2 \tag{9}$$

We now establish that the secondary simple cokriging weights will always be zero under the intrinsic coregionalization assumption. We arrive at this result from three perspectives.

Simple Cokriging Equations

Consider the simple cokriging equations for predicting $\delta_1(\mathbf{s}_0)$ based on $\boldsymbol{\delta} = (\boldsymbol{\delta}(\mathbf{s}_1)^T, \boldsymbol{\delta}(\mathbf{s}_2)^T, \ldots, \boldsymbol{\delta}(\mathbf{s}_n)^T)^T$, which are

$$\sum_{m=1}^{n} \{\lambda_{1m} c_{11}(\mathbf{s}_k, \mathbf{s}_m) + \lambda_{2m} c_{12}(\mathbf{s}_k, \mathbf{s}_m)\} = c_{11}(\mathbf{s}_k, \mathbf{s}_0), \qquad k = 1, 2, \ldots, n$$

$$\sum_{m=1}^{n} \{\lambda_{1m} c_{12}(\mathbf{s}_k, \mathbf{s}_m) + \lambda_{2m} c_{22}(\mathbf{s}_k, \mathbf{s}_m)\} = c_{12}(\mathbf{s}_k, \mathbf{s}_0), \qquad k = 1, 2, \ldots, n \tag{10}$$

By the intrinsic coregionalization assumption, we know $c_{11}(\mathbf{s}_k, \mathbf{s}_m) = c(\mathbf{s}_k, \mathbf{s}_m)$, $c_{12}(\mathbf{s}_k, \mathbf{s}_m) = c_{21}(\mathbf{s}_k, \mathbf{s}_m) = \gamma\rho c(\mathbf{s}_k, \mathbf{s}_m)$, and $c_{22}(\mathbf{s}_k, \mathbf{s}_m) = \gamma^2 c(\mathbf{s}_k, \mathbf{s}_m)$, which reduces the simple cokriging equations to

$$\sum_{m=1}^{n} \{\lambda_{1m} + \gamma\rho\lambda_{2m}\} c(\mathbf{s}_k, \mathbf{s}_m) = c(\mathbf{s}_k, \mathbf{s}_0), \qquad k = 1, 2, \ldots, n$$

$$\sum_{m=1}^{n} \{\gamma\rho\lambda_{1m} + \gamma^2\lambda_{2m}\} c(\mathbf{s}_k, \mathbf{s}_m) = \gamma\rho c(\mathbf{s}_k, \mathbf{s}_0), \qquad k = 1, 2, \ldots, n \tag{11}$$

The simple cokriging equations can thus be written

$$\sum_{m=1}^{n} \{\lambda_{1m} + \gamma\rho\lambda_{2m}\} c(\mathbf{s}_k, \mathbf{s}_m) = c(\mathbf{s}_k, \mathbf{s}_0), \qquad k = 1, 2, \ldots, n$$

$$\sum_{m=1}^{n} \left\{\lambda_{1m} + \frac{\gamma}{\rho}\lambda_{2m}\right\} c(\mathbf{s}_k, \mathbf{s}_m) = c(\mathbf{s}_k, \mathbf{s}_0), \qquad k = 1, 2, \ldots, n \tag{12}$$

If we define $\lambda_m^* \equiv \lambda_{1m} + \gamma\rho\lambda_{2m}$ and $\lambda_m^{**} \equiv \lambda_{1m} + (\gamma/\rho)\lambda_{2m}$ $(m = 1, 2, \ldots, n)$, we observe that each set of equations represents exactly the equations for standard univariate simple kriging. Furthermore, since in general $|\rho| \neq 1$ and we must have $\lambda_m^* = \lambda_m^{**}$; $m = 1, 2, \ldots, n$, we see that the secondary weights $\{\lambda_{2m}: m = 1, 2, \ldots, n\}$ must be zero. (If $|\rho| = 1$, Eq. (12) does not have a unique solution).

Least Squares Solution for $\boldsymbol{\lambda}$

Recall that $\boldsymbol{\delta} = (\boldsymbol{\delta}(\mathbf{s}_1)^T, \boldsymbol{\delta}(\mathbf{s}_2)^T, \ldots, \boldsymbol{\delta}(\mathbf{s}_n)^T)^T$. Denote the corresponding variance-covariance matrix for $\boldsymbol{\delta}$ as $\mathbf{T} \equiv \text{cov}(\boldsymbol{\delta})$. Under the intrinsic coregionalization assumption, $\mathbf{T}$ takes on a special form, specifically,

$$\mathbf{T} = \begin{bmatrix} c(\mathbf{s}_1, \mathbf{s}_1) & c(\mathbf{s}_1, \mathbf{s}_2) & \cdots & \\ c(\mathbf{s}_1, \mathbf{s}_2) & c(\mathbf{s}_2, \mathbf{s}_2) & & \\ \vdots & \vdots & \ddots & \\ c(\mathbf{s}_1, \mathbf{s}_n) & \cdots & & c(\mathbf{s}_n, \mathbf{s}_n) \end{bmatrix} \otimes \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix}$$

$$\equiv \mathbf{C} \otimes \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix} \tag{13}$$

where $\otimes$ denotes the Kronecker product [recall $\mathbf{A}_{m \times n} \otimes \mathbf{B}_{p \times q} = ((a_{ij}\mathbf{B}))_{mp \times nq}$]. Next, let $\mathbf{c} \equiv \text{cov}(\boldsymbol{\delta}, \delta_1(s_0))$, which is a $2n \times 1$ vector. Note that, under the intrinsic coregionalization assumption, $\mathbf{c}$ can be expressed as

$$\mathbf{c} = ([c(\mathbf{s}_1, \mathbf{s}_0), c(\mathbf{s}_2, \mathbf{s}_0), \ldots, c(\mathbf{s}_n, \mathbf{s}_0)] \otimes [1, \rho\gamma])^T \equiv \mathbf{c}_0 \otimes [1, \rho\gamma]^T \tag{14}$$

Minimizing Eq. (2) with respect to $\{\lambda_{ik}: k = 1, 2, \ldots, n; i = 1, 2\}$ results in the following least squares solution for $\boldsymbol{\lambda} \equiv (\lambda_{11}, \lambda_{12}, \lambda_{21}, \lambda_{22}, \ldots, \lambda_{n1}, \lambda_{n2})^T$:

$$\hat{\boldsymbol{\lambda}} = \mathbf{T}^{-1}\mathbf{c} = \left(\mathbf{C} \otimes \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix}\right)^{-1} (\mathbf{c}_0 \otimes [1, \rho\gamma]^T)$$

$$= \left(\mathbf{C}^{-1} \otimes \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix}^{-1}\right) (\mathbf{c}_0 \otimes [1, \rho\gamma]^T)$$

$$= \mathbf{C}^{-1}\mathbf{c}_0 \otimes \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix}^{-1} [1, \rho\gamma]^T \tag{15}$$

Thus, we have

$$\hat{\boldsymbol{\lambda}} = \mathbf{C}^{-1}\mathbf{c}_0 \otimes [1, 0]^T \tag{16}$$

Again we see that the secondary simple cokriging weights are zero under the assumption of intrinsic coregionalization.

Partial Covariance

To understand the mechanism generating the null secondary weights under intrinsic coregionalization, it is fruitful to consider the partial covariance of $\delta_1(\mathbf{s}_0)$ and $\delta_2(\mathbf{s}_k)$ given all other measurements. For example, consider the partial covariance between $\delta_1(\mathbf{s}_0)$ and $\delta_2(\mathbf{s}_1)$ given all other observations. We first calculate the partial covariance without conditioning on $\delta_1(\mathbf{s}_1)$, and then we calculate the partial covariance including conditioning on $\delta_1(\mathbf{s}_1)$.

Consider the vector $\boldsymbol{\delta}^0 \equiv [\delta_1(\mathbf{s}_0), \boldsymbol{\delta}(\mathbf{s}_1)^T, \boldsymbol{\delta}(\mathbf{s}_2)^T, \ldots, \boldsymbol{\delta}(\mathbf{s}_n)^T]^T$. Decompose $\boldsymbol{\delta}^0$ into three parts, namely $\boldsymbol{\delta}^0 = (\delta_1(\mathbf{s}_0), \boldsymbol{\delta}(\mathbf{s}_1)^T, \boldsymbol{\delta}_{\text{III}}^T)^T$, where $\boldsymbol{\delta}_{\text{III}} \equiv (\boldsymbol{\delta}(\mathbf{s}_2)^T, \ldots, \boldsymbol{\delta}(\mathbf{s}_n)^T)^T$. The corresponding partitioned covariance matrix is

$$\boldsymbol{\Sigma} \equiv \begin{bmatrix} \text{var } (\delta_1(\mathbf{s}_0)) & \text{cov } (\boldsymbol{\delta}(\mathbf{s}_1), \delta_1(\mathbf{s}_0))^T & \text{cov } (\boldsymbol{\delta}_{\text{III}}, \delta_1(\mathbf{s}_0))^T \\ \text{cov } (\boldsymbol{\delta}(\mathbf{s}_1), \delta_1(\mathbf{s}_0)) & \text{var } (\boldsymbol{\delta}(\mathbf{s}_1)) & \text{cov } (\boldsymbol{\delta}_{\text{III}}, \boldsymbol{\delta}(\mathbf{s}_1))^T \\ \text{cov } (\boldsymbol{\delta}_{\text{III}}, \delta_1(\mathbf{s}_0)) & \text{cov } (\boldsymbol{\delta}_{\text{III}}, \boldsymbol{\delta}(\mathbf{s}_1)) & \text{var } (\boldsymbol{\delta}_{\text{III}}) \end{bmatrix} \tag{17}$$

Now, under the assumption of intrinsic coregionalization we have

$$\boldsymbol{\Sigma} = \begin{bmatrix} c(\mathbf{s}_0, \mathbf{s}_0) & & \\ c(\mathbf{s}_0, \mathbf{s}_1) \otimes [1, \rho\gamma]^T & c(\mathbf{s}_1, \mathbf{s}_1) \otimes \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix} & \\ \mathbf{C}(\mathbf{s}_0, \mathbf{s}_{\text{III}}) \otimes [1, \rho\gamma]^T & \mathbf{C}(\mathbf{s}_1, \mathbf{s}_{\text{III}}) \otimes \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix} & \mathbf{C}(\mathbf{s}_{\text{III}}, \mathbf{s}_{\text{III}}) \otimes \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix} \end{bmatrix} \tag{18}$$

where

$$\mathbf{C}(\mathbf{s}_0, \mathbf{s}_{\text{III}}) \equiv \text{cov } (\delta_1(\mathbf{s}_0), [\delta_1(\mathbf{s}_2), \ldots, \delta_1(\mathbf{s}_n)]^T)^T$$

$$\mathbf{C}(\mathbf{s}_1, \mathbf{s}_{\text{III}}) \equiv \text{cov } (\delta_1(\mathbf{s}_1), [\delta_1(\mathbf{s}_2), \ldots, \delta_1(\mathbf{s}_n)]^T)^T$$

$$\mathbf{C}(\mathbf{s}_{\text{III}}, \mathbf{s}_{\text{III}}) \equiv \text{cov } ([\delta_1(\mathbf{s}_2), \ldots, \delta_1(\mathbf{s}_n)]^T, [\delta_1(\mathbf{s}_2), \ldots, \delta_1(\mathbf{s}_n)]^T)$$

The other entries of the matrix are given by symmetry.

First calculate the partial covariance matrix for $\delta_1(\mathbf{s}_0)$, $\delta_1(\mathbf{s}_1)$, and $\delta_2(\mathbf{s}_1)$, given all other observations, which we will denote $\boldsymbol{\Sigma}_{aa \cdot b}$. Consider the partition

$$\boldsymbol{\Sigma} \equiv \begin{bmatrix} \boldsymbol{\Sigma}_{aa} & \boldsymbol{\Sigma}_{ab} \\ \boldsymbol{\Sigma}_{ba} & \boldsymbol{\Sigma}_{bb} \end{bmatrix} \tag{19}$$

where $\boldsymbol{\Sigma}_{aa}$ is a (3×3) matrix of covariance terms corresponding to $\delta_1(\mathbf{s}_0)$, $\delta_1(\mathbf{s}_1)$, and $\delta_2(\mathbf{s}_1)$. Then, we have

$$\begin{aligned} \boldsymbol{\Sigma}_{aa \cdot b} &= \boldsymbol{\Sigma}_{aa} - \boldsymbol{\Sigma}_{ab}\boldsymbol{\Sigma}_{bb}^{-1}\boldsymbol{\Sigma}_{ba} \\ &= \begin{bmatrix} c(\mathbf{s}_0, \mathbf{s}_0) & c(\mathbf{s}_0, \mathbf{s}_1) \otimes [1, \rho\gamma] \\ c(\mathbf{s}_0, \mathbf{s}_1) \otimes [1, \rho\gamma]^T & c(\mathbf{s}_1, \mathbf{s}_1) \otimes \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix} \end{bmatrix} \\ &\quad - \begin{bmatrix} \mathbf{C}(\mathbf{s}_0, \mathbf{s}_{\text{III}})^T \otimes [1, \rho\gamma] \\ \mathbf{C}(\mathbf{s}_1, \mathbf{s}_{\text{III}})^T \otimes \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix} \end{bmatrix} \left[\mathbf{C}(\mathbf{s}_{\text{III}}, \mathbf{s}_{\text{III}}) \otimes \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix} \right]^{-1} \\ &\quad \begin{bmatrix} \mathbf{C}(\mathbf{s}_0, \mathbf{s}_{\text{III}})^T \otimes [1, \rho\gamma] \\ \mathbf{C}(\mathbf{s}_1, \mathbf{s}_{\text{III}})^T \otimes \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix} \end{bmatrix}^T \end{aligned} \tag{20}$$

If we further decompose

$$\Sigma_{aa \cdot b} \equiv \begin{bmatrix} \omega_{11} & \omega_{12} & \omega_{13} \\ \omega_{21} & \omega_{22} & \omega_{23} \\ \omega_{31} & \omega_{32} & \omega_{33} \end{bmatrix} \tag{21}$$

then

$$\begin{aligned} \omega_{11} &= c(\mathbf{s}_0, \mathbf{s}_0) - (\mathbf{C}(\mathbf{s}_0, \mathbf{s}_{\mathrm{III}})^T \otimes [1, \rho\gamma]) \left(\mathbf{C}(\mathbf{s}_{\mathrm{III}}, \mathbf{s}_{\mathrm{III}}) \otimes \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix} \right)^{-1} \\ &\quad \cdot (\mathbf{C}(\mathbf{s}_0, \mathbf{s}_{\mathrm{III}})^T \otimes [1, \rho\gamma])^T \\ &= c(\mathbf{s}_0, \mathbf{s}_0) - \left\{ \mathbf{C}(\mathbf{s}_0, \mathbf{s}_{\mathrm{III}})^T \mathbf{C}(\mathbf{s}_{\mathrm{III}}, \mathbf{s}_{\mathrm{III}})^{-1} \mathbf{C}(\mathbf{s}_0, \mathbf{s}_{\mathrm{III}}) \right. \\ &\quad \left. \otimes [1, \rho\gamma] \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix}^{-1} [1, \rho\gamma]^T \right\} \\ &= c(\mathbf{s}_0, \mathbf{s}_0) - \{\mathbf{C}(\mathbf{s}_0, \mathbf{s}_{\mathrm{III}})^T \mathbf{C}(\mathbf{s}_{\mathrm{III}}, \mathbf{s}_{\mathrm{III}})^{-1} \mathbf{C}(\mathbf{s}_0, \mathbf{s}_{\mathrm{III}})\} \end{aligned} \tag{22}$$

Similarly,

$$[\omega_{21}, \omega_{31}]^T = (c(\mathbf{s}_0, \mathbf{s}_1) - \mathbf{C}(\mathbf{s}_1, \mathbf{s}_{\mathrm{III}})^T \mathbf{C}(\mathbf{s}_{\mathrm{III}}, \mathbf{s}_{\mathrm{III}})^{-1} \mathbf{C}(\mathbf{s}_0, \mathbf{s}_{\mathrm{III}})) \otimes [1, \rho\gamma]^T \tag{23}$$

and

$$\begin{bmatrix} \omega_{22} & \omega_{23} \\ \omega_{32} & \omega_{33} \end{bmatrix} = (c(\mathbf{s}_1, \mathbf{s}_1) - \mathbf{C}(\mathbf{s}_1, \mathbf{s}_{\mathrm{III}})^T \mathbf{C}(\mathbf{s}_{\mathrm{III}}, \mathbf{s}_{\mathrm{III}})^{-1} \mathbf{C}(\mathbf{s}_1, \mathbf{s}_{\mathrm{III}})) \otimes \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix} \tag{24}$$

We can therefore piece together $\Sigma_{aa \cdot b}$. We note that

$$\omega_{31} = (c(\mathbf{s}_0, \mathbf{s}_1) - \mathbf{C}(\mathbf{s}_1, \mathbf{s}_{\mathrm{III}})^T \mathbf{C}(\mathbf{s}_{\mathrm{III}}, \mathbf{s}_{\mathrm{III}})^{-1} \mathbf{C}(\mathbf{s}_0, \mathbf{s}_{\mathrm{III}}))\rho\gamma \tag{25}$$

which need not be zero in general. That is, the partial covariance between $\delta_1(\mathbf{s}_0)$ and $\delta_2(\mathbf{s}_1)$, conditioned on all other measurements except $\delta_1(\mathbf{s}_1)$, may *not* be zero.

Now consider what happens when we also condition on $\delta_1(\mathbf{s}_1)$. Let $R = \mathrm{cov}\,(\delta_1(\mathbf{s}_0), \delta_2(\mathbf{s}_1) | \delta_1(\mathbf{s}_1), \boldsymbol{\delta}(\mathbf{s}_2)^T, \boldsymbol{\delta}(\mathbf{s}_3)^T, \ldots, \boldsymbol{\delta}(\mathbf{s}_n)^T)$, where $\mathrm{cov}\,(d, e|f)$ denotes the partial covariance of d and e given f. Then,

$$R \equiv \omega_{31} - \frac{\omega_{21}\omega_{32}}{\omega_{22}} \tag{26}$$

From our results in the preceding text,

$$\omega_{31} = (c(\mathbf{s}_0, \mathbf{s}_1) - \mathbf{C}(\mathbf{s}_1, \mathbf{s}_{III})^T\mathbf{C}(\mathbf{s}_{III}, \mathbf{s}_{III})^{-1}\mathbf{C}(\mathbf{s}_0, \mathbf{s}_{III}))\rho\gamma$$

$$\omega_{21} = (c(\mathbf{s}_0, \mathbf{s}_1) - \mathbf{C}(\mathbf{s}_1, \mathbf{s}_{III})^T\mathbf{C}(\mathbf{s}_{III}, \mathbf{s}_{III})^{-1}\mathbf{C}(\mathbf{s}_0, \mathbf{s}_{III}))$$

$$\omega_{32} = (c(\mathbf{s}_1, \mathbf{s}_1) - \mathbf{C}(\mathbf{s}_1, \mathbf{s}_{III})^T\mathbf{C}(\mathbf{s}_{III}, \mathbf{s}_{III})^{-1}\mathbf{C}(\mathbf{s}_1, \mathbf{s}_{III}))\rho\gamma$$

$$\omega_{22} = (c(\mathbf{s}_1, \mathbf{s}_1) - \mathbf{C}(\mathbf{s}_1, \mathbf{s}_{III})^T\mathbf{C}(\mathbf{s}_{III}, \mathbf{s}_{III})^{-1}\mathbf{C}(\mathbf{s}_1, \mathbf{s}_{III}))$$

and hence $R = 0$. Thus, by also conditioning on $\delta_1(\mathbf{s}_1)$, the partial covariance becomes zero. This appears analogous to the *screen effect* encountered in univariate spatial prediction (see, e.g., Journel and Huijbregts, 1978, p. 346, or Cressie, 1991, p. 133). The screen effect says that the influence of a datum is reduced when it is hidden by another datum. Under our current assumptions, we see that the secondary datum at a sample location is completely hidden by the corresponding primary datum.

UNIVERSAL COKRIGING

Recall the decomposition

$$Z_i(\mathbf{s}) = \mu_i(\mathbf{s}) + \delta_i(\mathbf{s}); \mathbf{s} \in \mathbf{D}, i = 1, 2 \tag{27}$$

where

$$\mu_i(\mathbf{s}) = \mathbf{X}_i(\mathbf{s})^T\boldsymbol{\beta}_i \tag{28}$$

$\boldsymbol{\beta}_i$ is a $(p_i \times 1)$ vector of unknown large-scale variation parameters, and $\delta_i(\mathbf{s})$ represents small-scale variation; $i = 1, 2$. We have shown in a variety of ways that when $\boldsymbol{\beta}^0$ is fixed and known, simple cokriging on $\{\delta_i(\mathbf{s}_k)\colon k = 1, 2, \ldots, n;\ i = 1, 2\}$ yields a predictor for $\delta_1(\mathbf{s}_0)$ that gives null weightings to the secondary information. Then, the simple cokriging predictor for $Z_1(\mathbf{s}_0)$ is

$$\hat{Z}_1(\mathbf{s}_0) = \mu_1(\mathbf{s}_0) + \hat{\delta}_1(\mathbf{s}_0) \tag{29}$$

where $\mu_1(\mathbf{s}_0) = \mathbf{X}_1(\mathbf{s}_0)^T\boldsymbol{\beta}_1^0$. The simple cokriging predictor is always unbiased.

Now consider the more general universal cokriging case (Myers, 1982; Clark et al., 1987) when $\boldsymbol{\beta}$ is unknown. The large-scale parameter space can be written as $\Omega = \{\boldsymbol{\beta} \equiv (\boldsymbol{\beta}_1^T, \boldsymbol{\beta}_2^T)^T\colon \boldsymbol{\beta} \in \boldsymbol{R}^p\}$, where $p = p_1 + p_2$. Recall that we are currently assuming that $\boldsymbol{\beta}_1$ are large-scale parameters exclusive to the large-scale structure in $Z_1(\cdot)$, and $\boldsymbol{\beta}_2$ are exclusive to the large-scale structure in $Z_2(\cdot)$. That is, there are no existing relationships or restrictions *between* $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$.

To examine the generalized least-squares estimator of $\boldsymbol{\beta}$ concisely, we redefine $\mathbf{Z} = [\mathbf{Z}_1^T, \mathbf{Z}_2^T]^T$, where $\mathbf{Z}_i \equiv (Z_i(\mathbf{s}_1), \ldots, Z_i(\mathbf{s}_n))^T$; $i = 1, 2$. This is

simply a rearrangement of the elements of the original $\mathbf{Z}$ vector. We note that the optimal universal cokriging predictor for the vector $\mathbf{Z}(\mathbf{s}_0)$ at location $\mathbf{s}_0$ can be written,

$$\hat{\mathbf{Z}}(\mathbf{s}_0) = \mathbf{C}_0^T\mathbf{T}^{-1}(\mathbf{Z} - \mathbf{X}\hat{\boldsymbol{\beta}}_{gls}) + \mathbf{X}(\mathbf{s}_0)\hat{\boldsymbol{\beta}}_{gls} \tag{30}$$

where $\mathbf{C}_0$ is the $(2n \times 2)$ matrix $\text{cov}(\mathbf{Z}, \mathbf{Z}(\mathbf{s}_0))$, $\mathbf{T}$ is the $(2n \times 2n)$ matrix var $(\mathbf{Z})$, $\mathbf{X}$ is a matrix of "explanatory" variables, $\mathbf{X}(\mathbf{s}_0)$ is the $(2 \times p)$ matrix of explanatory variables at location $\mathbf{s}_0$, and $\hat{\boldsymbol{\beta}}_{gls} = (\mathbf{X}^T\mathbf{T}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{T}^{-1}\mathbf{Z}$ is the generalized least-squares estimator of the large-scale parameter vector. (For details, see VerHoef and Cressie, 1993). Under the model given in (3), we have

$$\mathbf{C}_0 = \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix} \otimes \mathbf{c}_0 \tag{31}$$

and we can rewrite

$$\mathbf{T} = \begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix} \otimes \mathbf{C} \tag{32}$$

We note that $\mathbf{C}_0^T\mathbf{T}^{-1}(\mathbf{Z} - \mathbf{X}\hat{\boldsymbol{\beta}}_{gls})$ is equivalent to simple cokriging on $\hat{\boldsymbol{\delta}} \equiv \mathbf{Z} - \mathbf{X}\hat{\boldsymbol{\beta}}_{gls}$, where the unknown large-scale parameters are replaced by their generalized least-squares estimators. We have shown that the secondary simple cokriging weights must be zero under the intrinsic coregionalization assumption with all measurements available at the n sample locations. Hence, it is only possible for nonzero secondary cokriging weights to occur when the estimation of large-scale parameters is necessary, such as in ordinary and universal cokriging.

To determine when secondary information is used to estimate $\boldsymbol{\beta}_1$ we consider two situations. First assume all explanatory variables are common to both component processes, i.e., $\mathbf{X}_1(\mathbf{s}) = \mathbf{X}_2(\mathbf{s}) \equiv \mathbf{X}^*(\mathbf{s})$; $\mathbf{s} \in \mathbf{D}$ and $p_1 = p_2 = p^*$. This model is often encountered in geostatistical problems. In this situation, we note that the matrix of explanatory variables can be written as $\mathbf{X} = (\mathbf{I}_2 \otimes \mathbf{X}^*)$, where $\mathbf{I}_k$ denotes the $(k \times k)$ identity matrix and $\mathbf{X}^*$ is the $(n \times p^*)$ matrix with i-th row $\mathbf{X}^*(\mathbf{s}_i)$. The generalized least squares estimator for $\boldsymbol{\beta}$ can then be written

$$\hat{\boldsymbol{\beta}}_{gls} = \left[(\mathbf{I}_2 \otimes \mathbf{X}^*)^T \left(\begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix} \otimes \mathbf{C} \right)^{-1} (\mathbf{I}_2 \otimes \mathbf{X}^*) \right]^{-1}$$
$$\cdot (\mathbf{I}_2 \otimes \mathbf{X}^*)^T \left(\begin{bmatrix} 1 & \rho\gamma \\ \rho\gamma & \gamma^2 \end{bmatrix} \otimes \mathbf{C} \right)^{-1} \mathbf{Z} \tag{33}$$

which reduces to

$$\hat{\boldsymbol{\beta}}_{gls} = [\mathbf{I}_2 \otimes (\mathbf{X}^{*T}\mathbf{C}^{-1}\mathbf{X}^*)^{-1}\mathbf{X}^{*T}\mathbf{C}^{-1}]\mathbf{Z} \tag{34}$$

Thus we see only the primary observations are used to estimate $\boldsymbol{\beta}_1$ and only the secondary observations are used to estimate $\boldsymbol{\beta}_2$. However, should a *restriction* be placed on the parameter space that involves parameters from $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ jointly, all observations, both primary and secondary, are used to obtain the optimal linear large-scale parameter estimator of $\boldsymbol{\beta}_1$. From (30), the universal cokriging predictor for the primary process is

$$\hat{Z}_1(\mathbf{s}_0) = \mathbf{A}^T\mathbf{Z} + \left(\mathbf{X}_1(\mathbf{s}_0)^T - \mathbf{A}^T\left(\begin{bmatrix}1\\0\end{bmatrix} \otimes \mathbf{X}^*\right)\right)\hat{\boldsymbol{\beta}}_{1,gls} \tag{35}$$

where $\mathbf{A}^T$ is the first row of $\mathbf{C}_0^T\mathbf{T}^{-1}$, and $\mathbf{X}_1(\mathbf{s}_0)^T$ is the first row of $\mathbf{X}(\mathbf{s}_0)$. The first term involves simple cokriging weights which we have shown are zero for the secondary component. The second term involves $\hat{\boldsymbol{\beta}}_{1,gls}$ and it is *here* that *nonzero* weights on $\mathbf{Z}_2$ may enter.

For example, if $\mu_1(\mathbf{s}) = \mu_1$ and $\mu_2(\mathbf{s}) = \mu_2$, and we have the restriction $\mu_1 = \mu_2$, we note that the full-rank reparameterization will lead to a model with one large-scale parameter, $\mu \equiv \mu_1 = \mu_2$. This restriction might arise if the two processes represent two kinds of unbiased measurements of the same variable. The restricted generalized-least-squares estimator $\hat{\mu}$ will be a linear weighting of all $2n$ observations. Thus, in this situation, the secondary information will be used to estimate the mean of the primary process, μ_1, and so will receive nonzero universal cokriging weights in predicting $Z_1(\mathbf{s}_0)$.

Now assume that some of the explanatory variables are unique to one of the component processes and that no restriction is placed on the large-scale parameter space that involves parameters from $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ jointly. In this case $\mathbf{X}_1(\mathbf{s}) \neq \mathbf{X}_2(\mathbf{s})$; $\mathbf{s} \in \mathbf{D}$, but it is still true that the generalized least-squares estimator will reduce to the form

$$\hat{\boldsymbol{\beta}}_{gls} = \begin{bmatrix}\mathbf{D}_1 & 0\\ 0 & \mathbf{D}_2\end{bmatrix}\mathbf{Z} \tag{36}$$

where $\mathbf{D}_1$ and $\mathbf{D}_2$ are $n \times n$ matrices. Hence, secondary information will not be used to estimate $\boldsymbol{\beta}_1$ and therefore the secondary universal cokriging weights will be zero.

We can therefore conclude that when restrictions are placed on the large-scale parameter space that jointly involve parameters in $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, the secondary information will be used to estimate the mean of the primary process and will thus receive non-zero universal cokriging weights in predicting $Z_1(\mathbf{s}_0)$. Otherwise, the secondary cokriging weights will be zero.

An Additional Primary Observation

It is of interest to observe the behavior of the secondary cokriging weights when an additional primary observation, say $Z_1(\mathbf{s}_{n+1})$, is available without $Z_2(\mathbf{s}_{n+1})$. Since we have already resolved that nonzero secondary cokriging

weights can occur when particular restrictions are placed on the large-scale parameter space, we may consider the simple cokriging problem and determine if nonzero secondary cokriging weights can occur when no restrictions are placed on the large-scale parameter space.

Again, consider the simple cokriging equations

$$\sum_{m=1}^{n} \{\lambda_{1m} c_{11}(\mathbf{s}_k, \mathbf{s}_m) + \lambda_{2m} c_{12}(\mathbf{s}_k, \mathbf{s}_m)\} + \lambda_{1,n+1} c_{11}(\mathbf{s}_k, \mathbf{s}_{n+1})$$

$$= c_{11}(\mathbf{s}_k, \mathbf{s}_0), \quad k = 1, 2, \ldots, n+1$$

$$\sum_{m=1}^{n} \{\lambda_{1m} c_{12}(\mathbf{s}_k, \mathbf{s}_m) + \lambda_{2m} c_{22}(\mathbf{s}_k, \mathbf{s}_m)\} + \lambda_{1,n+1} c_{21}(\mathbf{s}_k, \mathbf{s}_{n+1})$$

$$= c_{21}(\mathbf{s}_k, \mathbf{s}_0), \quad k = 1, 2, \ldots, n \tag{37}$$

By the intrinsic coregionalization assumption, we can express these equations as

$$\sum_{m=1}^{n} \lambda_m^* c(\mathbf{s}_k, \mathbf{s}_m) + \lambda_{1,n+1} c(\mathbf{s}_k, \mathbf{s}_{n+1}) = c(\mathbf{s}_k, \mathbf{s}_0), \quad k = 1, 2, \ldots, n+1$$

$$\sum_{m=1}^{n} \lambda_m^{**} c(\mathbf{s}_k, \mathbf{s}_m) + \lambda_{1,n+1} c(\mathbf{s}_k, \mathbf{s}_{n+1}) = c(\mathbf{s}_k, \mathbf{s}_0), \quad k = 1, 2, \ldots, n \tag{38}$$

Again, we see that two sets of equations are identical for $k = 1, 2, \ldots, n$, which implies $\lambda_m^* = \lambda_m^{**}$, $m = 1, 2, \ldots, n$, and hence the secondary cokriging weights are still zero despite the addition of an additional primary observation. It can also be shown that the partial correlation between $\delta_1(\mathbf{s}_0)$ and $\delta_2(\mathbf{s}_1)$ remains zero when we condition additionally on $\delta_1(\mathbf{s}_{n+1})$.

An Additional Secondary Observation

We may also want to consider what happens to the secondary cokriging weights when an additional secondary observation $Z_2(\mathbf{s}_{n+1})$ is available without $Z_1(\mathbf{s}_{n+1})$.

Again, consider the simple cokriging equations

$$\sum_{m=1}^{n} \{\lambda_{1m} c_{11}(\mathbf{s}_k, \mathbf{s}_m) + \lambda_{2m} c_{12}(\mathbf{s}_k, \mathbf{s}_m)\} + \lambda_{2,n+1} c_{12}(\mathbf{s}_k, \mathbf{s}_{n+1})$$

$$= c_{11}(\mathbf{s}_k, \mathbf{s}_0), \quad k = 1, 2, \ldots, n$$

$$\sum_{m=1}^{n} \{\lambda_{1m} c_{12}(\mathbf{s}_k, \mathbf{s}_m) + \lambda_{2m} c_{22}(\mathbf{s}_k, \mathbf{s}_m)\} + \lambda_{2,n+1} c_{22}(\mathbf{s}_k, \mathbf{s}_{n+1})$$

$$= c_{21}(\mathbf{s}_k, \mathbf{s}_0), \quad k = 1, 2, \ldots, n+1 \tag{39}$$

By the intrinsic coregionalization assumption, we have

$$\sum_{m=1}^{n} \lambda_m^* c(\mathbf{s}_k, \mathbf{s}_m) + \gamma\rho\lambda_{2,n+1} c(\mathbf{s}_k, \mathbf{s}_{n+1}) = c(\mathbf{s}_k, \mathbf{s}_0),\ k = 1, 2, \ldots, n$$

$$\sum_{m=1}^{n} \lambda_m^{**} c(\mathbf{s}_k, \mathbf{s}_m) + \frac{\gamma}{\rho} \lambda_{2,n+1} c(\mathbf{s}_k, \mathbf{s}_{n+1}) = c(\mathbf{s}_k, \mathbf{s}_0),\ k = 1, 2, \ldots, n+1 \tag{40}$$

We see that, with the inclusion of $Z_2(\mathbf{s}_{n+1})$, it is no longer necessary for $\lambda_m^* = \lambda_m^{**}$, $m = 1, 2, \ldots, n$. Thus, it is no longer necessary for the secondary simple cokriging weights to be zero. It can also be shown that the partial correlation between $\delta_1(\mathbf{s}_0)$ and $\delta_2(\mathbf{s}_1)$ need not be zero when we condition additionally on $\delta_2(\mathbf{s}_{n+1})$.

AN EXAMPLE—THE NEVADA TEST SITE DATA

In 1957, a device containing plutonium was blown apart by chemical explosives at the Area 13 "safety-shot" location on the Nevada Test Site (NTS). This experiment was performed partly to test for "safety" against fission reactions in an accident situation involving an atomic weapon. A consequence of the test was the contamination of the immediate surrounding desert soil and vegetation with plutonium (Pu) and americium (Am). In 1971, the Nevada Applied Ecology Group (NAEG) began conducting environmental transuranic studies in this area by taking field instrument surveys and collecting soil, vegetation, and animal tissue samples. One goal of these studies was to predict the total amount and spatial distribution of ^{239,240}Pu and ^{241}Am in surface soil.

The Pu concentrations (in μCi/m^2) were determined by wet chemistry on surface (top 5 cm) soil samples taken at random locations. The Am concentrations in surface soil was obtained from Field Instrument for the Detection of Low Energy Radiation (FIDLER) readings (in 10^3 counts per minute (cpm)) at one foot above the surface. In this paper, we use a subset of the data accumulated in Area 13, considering 104 sample locations where measurements are available on both components (Gilbert, 1978). A map of these sample locations and ground zero is displayed in Fig. 1. Previous analyses on the Area 13 data have concluded that there is a good overall correlation, on the log scale, between wet chemistry Pu analyses and Am FIDLER measurements (Church et al., 1975; Gilbert and Simpson, 1985). To consider the effects of cokriging with an intrinsic coregionalization model, a restricted large-scale structure between components, and measurements on both components at all sample locations, we consider the prediction of log Pu at unsampled locations based on the measurements at the locations displayed in Fig. 1. A prediction is made for each of 120 unsampled grid sites in the interior of the rectangular region outlined in Fig. 1.

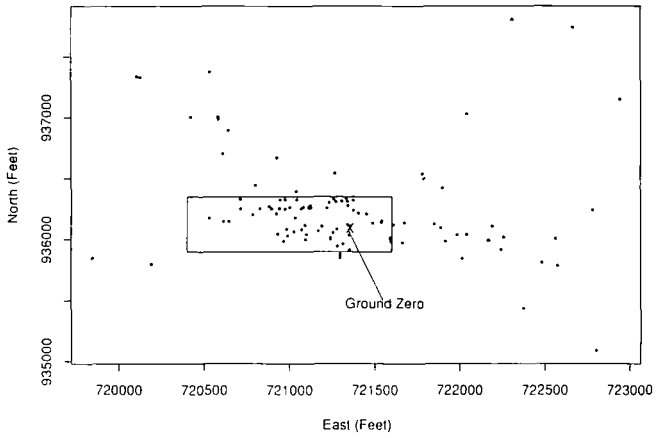

Fig. 1. The sample locations in the inner-fence region of Area 13 where measurements on both Pu and Am are available. The axes are in Nevada feet coordinates and the cross (X) denotes ground zero.

Cokriging is very useful in the under-sampled problem where there are relatively few primary measurements as compared to secondary measurements (see, e.g., Stein and Corsten, 1991; and Zhang et al., 1992). Indeed, in the NTS study, many more FIDLER readings were taken because the cost of a FIDLER reading was approximately 50 times less than that of a Pu analysis on a soil sample (Gilbert, 1978). However, in this intrinsic coregionalization study, we are only interested in the effects of cokriging when measurements are available for both components at all sample locations.

Figure 2 displays a bivariate ray-glyph map of the raw log Pu and log Am measurements (Carr et al., 1992). The rays pointing to the right represent log Pu concentration trends and rays pointing to the left represent log Am concentration trends. The bivariate rap map is an effective technique for showing bivariate associations. The two rays at each location in Fig. 2 generally point down or up together, clearly indicating that the components are positively correlated.

To examine the effects of cokriging under intrinsic coregionalization and a restricted large-scale structure, the following spatial model will be considered for $Z_1 \equiv$ log Pu, and $Z_2 \equiv$ log Am, at spatial locations $\mathbf{s} \in \mathbf{D}$. Assume

$$\begin{aligned} Z_1(\mathbf{s}) &= \mu_1 + \beta \exp(-\|\mathbf{s} - \mathbf{s}_{gz}\|/\theta) + \delta_1(\mathbf{s}) \\ \kappa Z_2(\mathbf{s}) &= \mu_2 + \beta \exp(-\|\mathbf{s} - \mathbf{s}_{gz}\|/\theta) + \delta_2(\mathbf{s}) \end{aligned} \tag{41}$$

where $\mathbf{s}_{gz}$ denotes ground zero, κ and θ are parameters that are assumed known, μ_1, μ_2, and β are unknown large-scale parameters, and $\delta_1(\mathbf{s})$ and $\delta_2(\mathbf{s})$ are as-

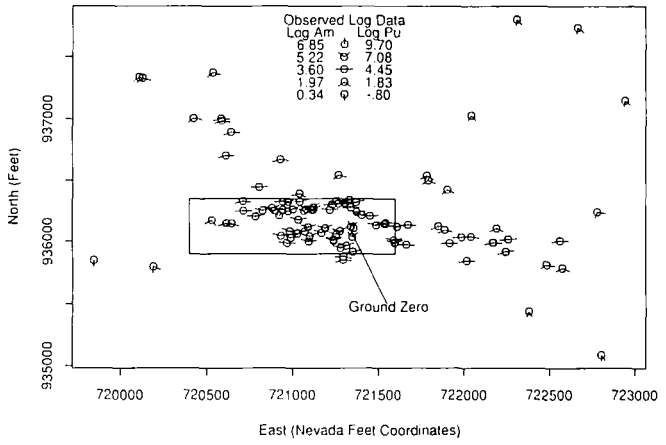

Fig. 2. A bivariate ray-glyph map of the log Pu and Am measurements. Pu is measured in $\mu Ci/m^2$, and Am is measured in 10^3 counts per minute (cpm).

sumed to be second-order stationary processes that are spatially cross-correlated; i.e., the vector $\boldsymbol{\delta}(\mathbf{s}) \equiv [\delta_1(\mathbf{s}), \delta_2(\mathbf{s})]^T$ is a bivariate second-order stationary process.

The large-scale structure for both components in Eq. (41) was selected to model the observed exponential decline away from ground zero. The parameter β is assumed to be shared by the two large-scale models. The parameter κ serves to transform the log Am data to an equivalent scale with log Pu. Though we assume this parameter is known, we actually estimated κ externally by regressing log Pu on log Am, based on all 104 observations. Similarly, the exponential scale parameter θ was estimated externally (for fixed κ) using non-linear least squares. The sample variance of these estimates were relatively small and thus we consider $\kappa = 1.25$ and $\theta = 1880$ fixed for the following analysis.

The ordinary-least-squares residuals from the large-scale models above were used to assist in modeling an isotropic small-scale spatial dependence structure. The experimental (cross-) semivariograms were estimated using the robust fourth-root estimator proposed by Cressie and Hawkins (1980). (The derivation of the fourth-root cross-variogram estimator is exactly as that for the variogram estimator under proper standardization.)

In this example we chose to fit an isotropic exponential intrinsic coregionalization model to the experimental (cross-) semivariograms. A bivariate generalization of the univariate non-linear weighted-least-squares criterion of Cressie (1985) was used to obtain estimates for the parameters of the exponential variogram model. Since variograms and cross-variograms are interrelated, parameter estimation requires simultaneous consideration of the experimental var-

iograms and cross-variograms (Helterbrand and Cressie, 1994). The multivariate spatial dependence structure (in terms of covariograms) was estimated as

$$\mathbf{C}(\mathbf{s}_k, \mathbf{s}_m) = \begin{vmatrix} 2.079 & 1.614 \\ 1.614 & 2.290 \end{vmatrix} \exp\left(-\frac{1}{280}\|\mathbf{s}_k - \mathbf{s}_m\|\right) \tag{42}$$

The fit of this valid model to the experimental (cross-) semivariograms is displayed (standardized) in a matrix display in Fig. 3. The semivariograms for Pu and Am appear on the diagonal, with the cross-semivariogram displayed off the diagonal. Notice that the intrinsic coregionalization model does not have a nugget effect and fails to fit the experimental variogram for the primary component (Pu) at the shorter lags. When a nugget effect parameter is included for the Pu variogram model, the estimated bivariate spatial dependence model is

$$\mathbf{C}(\mathbf{s}_k, \mathbf{s}_m) = \begin{cases} \begin{vmatrix} 0.834 & 1.330 \\ 1.330 & 2.290 \end{vmatrix} \exp\left(-\frac{1}{340}\|\mathbf{s}_k - \mathbf{s}_m\|\right) & \text{if } \mathbf{s}_k \neq \mathbf{s}_m \\ \begin{vmatrix} 1.928 & 1.330 \\ 1.330 & 2.290 \end{vmatrix} & \text{if } \mathbf{s}_k = \mathbf{s}_m \end{cases} \tag{43}$$

This model is no longer of an intrinsic coregionalization form, but provides a satisfactory fit to the experimental variograms (Fig. 4).

To compare the gain in precision of cokriging relative to kriging, the pre-

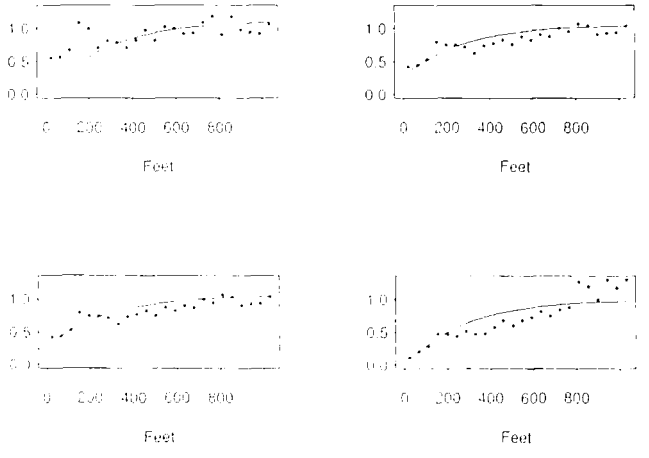

Fig. 3. The fit of the estimated exponential intrinsic coregionalization model to the experimental (cross-) semivariograms.

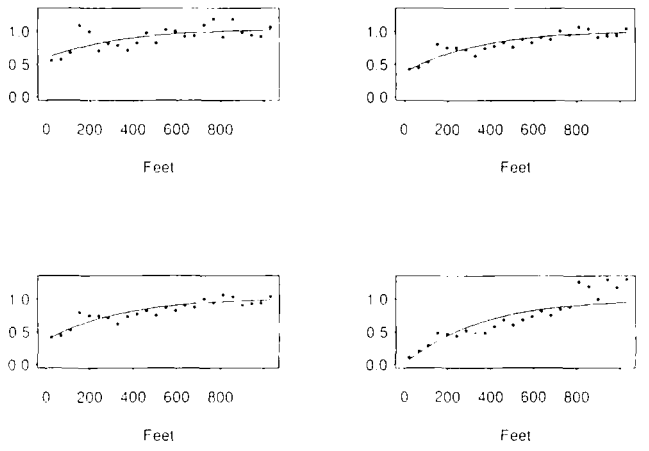

Fig. 4. The fit of the exponential semivariogram model, with a nugget effect for the primary component, to the experimental (cross-) semivariograms.

diction error variance was calculated for the 120 prediction locations based on the universal kriging and cokriging predictors, $\hat{Z}_{1,UK}$ and $\hat{Z}_{1,UCK}$, respectively. The relative efficiency is calculated as

$$\frac{\text{var}\,(Z_1(\mathbf{s}_0) - \hat{Z}_{1,UK}(\mathbf{s}_0))}{\text{var}\,(Z_1(\mathbf{s}_0) - \hat{Z}_{1,UCK}(\mathbf{s}_0))} \tag{44}$$

A kriging neighborhood with a search window of radius 560 feet was used for prediction. If more than 12 sample locations fell in the search window, only the measurements from the 12 nearest locations were used. If fewer than ten locations fell in the search window, the search window was expanded so that each prediction was based on measurements from its nearest ten neighbors (e.g., Harper et al., 1988). Kriging neighborhoods are primarily used to reduce the computational burden demanded by kriging, although Journel and Rossi (1989) also point out that kriging neighborhoods can be used to protect the user from local large-scale model misspecification.

Figures 5 through 7 display the kriging, cokriging, and cokriging variance maps, respectively, based on the fitted exponential intrinsic coregionalization model given in Eq. (42). Though non-zero secondary weights occur and the kriging and cokriging predictions differ, the reduction in the prediction error variance is minimal. The mean relative efficiency for the 120 prediction locations is 1.004, with a maximum of 1.031 for one prediction location. This minimal reduction is not surprising when one notes that the prediction error variance

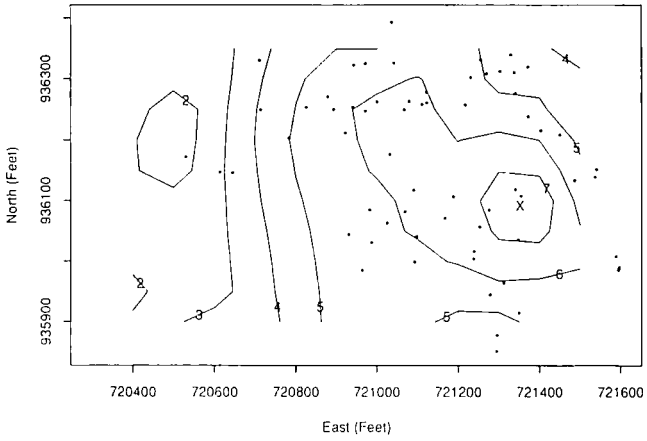

Fig. 5. The kriging map based on the exponential intrinsic coregionalization model. Contour units are in log $\mu Ci/m^2$. The cross (X) denotes ground zero.

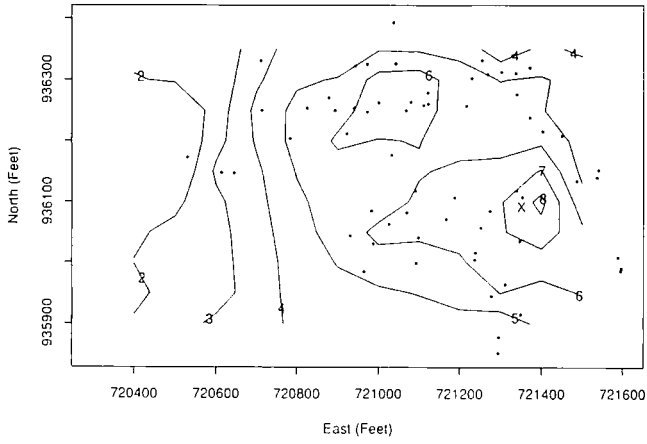

Fig. 6. The cokriging map based on the exponential intrinsic coregionalization model. Contour units are in log $\mu Ci/m^2$. The cross (X) denotes ground zero.

matrix corresponding to Eq. (30) can be decomposed into the sum of two terms, var $(\mathbf{Z}(\mathbf{s}_0) - \mathbf{C}_0^T\mathbf{T}^{-1}\mathbf{Z})$, and var $[(\mathbf{X}(\mathbf{s}_0) - \mathbf{C}_0^T\mathbf{T}^{-1}\mathbf{X})\hat{\boldsymbol{\beta}}_{gls}]$. (For this example, $\boldsymbol{\beta} \equiv (\mu_1, \mu_2, \beta)$.) Under an intrinsic coregionalization model with measurements on both components at all sample locations, $\mathbf{C}_0^T\mathbf{T}^{-1} = \mathbf{I}_2 \otimes \mathbf{c}_0^T\mathbf{C}^{-1}$, and the first variance term is the same for both the kriging and cokriging predictor of

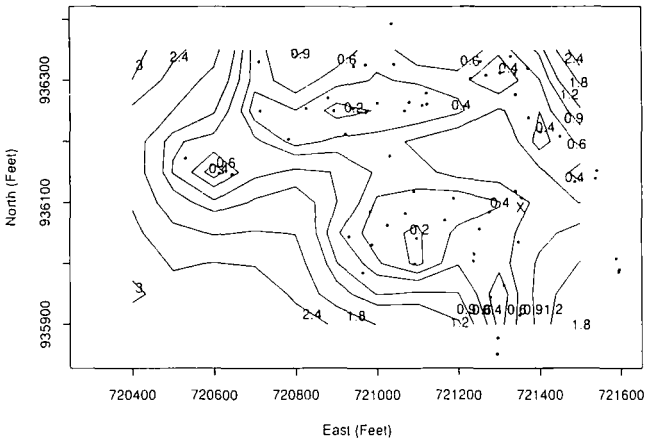

Fig. 7. The cokriging variance map based on the exponential intrinsic coregionalization model. Contour units are in $(\log \mu Ci/m^2)^2$. The cross (X) denotes ground zero.

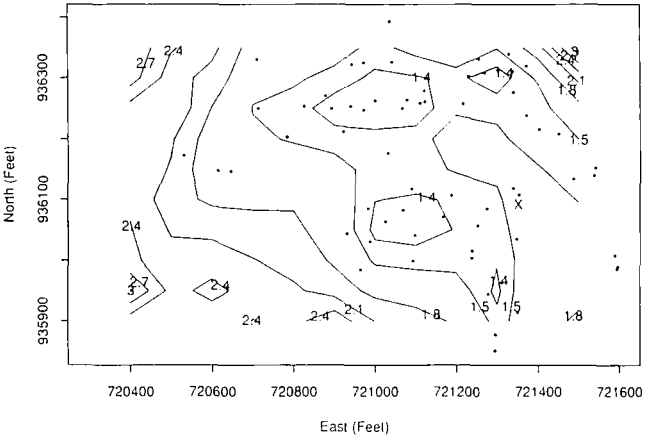

Fig. 8. The kriging variance map based on the exponential variogram model with a nugget effect for Pu. Contour units are in $(\log \mu Ci/m^2)^2$. The cross (X) denotes ground zero.

the primary [secondary] component. Thus, the reduction in the prediction error variance due to cokriging is from the second variance term, so that any additional efficiency attained by cokriging is due to the additional precision in estimating β.

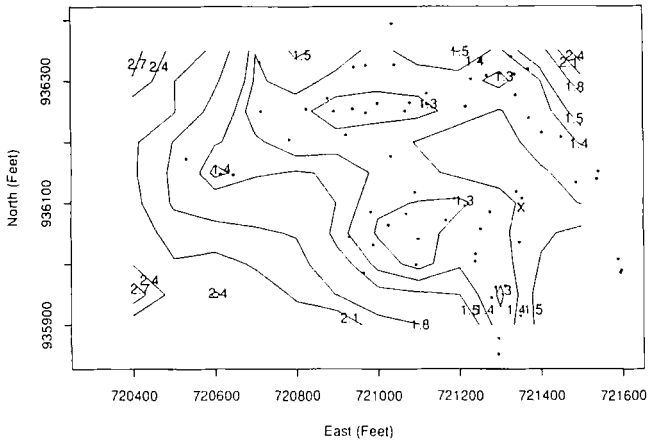

Fig. 9. The cokriging variance map based on the exponential variogram model with a nugget effect for Pu. Contour units are in $(\log \mu Ci/m^3)^2$. The cross (X) denotes ground zero.

Recall that the intrinsic coregionalization model fails to fit the experimental variogram for the primary component (Pu) at the shorter lags. Figures 8 and 9 display, respectively, the kriging variance and cokriging variance maps based on the fitted exponential variogram model with a nugget effect for the Pu variogram. For this model, the first term of the prediction error variance is not the same for kriging and cokriging. The mean relative efficiency for the 120 prediction locations under this more appropriate spatial dependence model is 1.072, with a minimum of 1.013 and maximum of 1.214. Thus, for this example, a reduction of up to 17.6% in the prediction error variance is obtained using cokriging even when measurements are available on both components at all sample locations.

CONCLUSION

We have shown that, under the assumption of intrinsic coregionalization with observations available on both components at each sample location, the large-scale parameter space for uniform unbiasedness determines the allowable values for the secondary cokriging weights. If we require uniform unbiasedness on a parameter space where the large-scale parameters for the primary mean and the large-scale parameters for the secondary mean are restricted to depend on each other in some manner, the secondary information will be used in estimating primary large-scale parameters and thus nonzero secondary cokriging weights will occur. However, the reduction in the prediction error variance may

be minimal. The results here also extend to the m-variate (where $m > 2$) cokriging problem when we assume m-variate intrinsic coregionalization.

We began by arguing that the intrinsic coregionalization assumption is restrictive and often inappropriate. Realistic multivariate spatial statistical models are typically more complex than intrinsic coregionalization, resulting in the *multivariate* universal cokriging predictors (Eq. 30). In general, secondary information is crucial in determining optimal predictors.

ACKNOWLEDGMENTS

This research was supported by the National Science Foundation (DMS-9001862 and DMS-9204521), the National Security Agency (MDA904-92-H-3021), and an Iowa State University Research Grant (Carver Grant). The authors are grateful to Richard O. Gilbert of Battelle, Pacific Northwest Laboratories and the Nevada Field Office of the U.S. Department of Energy for use of the Nevada Test Site data.

REFERENCES

Carr, D., Olsen, A., and White, D., 1992, Hexagon mosaic maps for display of univariate and bivariate geographical data: Cartogr. Geogr. Inf. Syst., v. 19, p. 228–236.

Church, B., Medling, E., and Brady, D., 1975, A different look at area 13 FIDLER survey data, *in* M. White and P. Dunaway (Eds.), *The Radioecology of Plutonium and Other Transuranics in Desert Environments:* USERDA Report, NVO-153, p. 231–235.

Clark, I., Basinger, K., and Harper, W., 1987, MUCK—A novel approach to cokriging, *in* B. Buxton (Ed.), *Proceedings of the Conference on Geostatistical, Sensitivity, and Uncertainty Methods for Ground-Water Flow and Radionuclide Transport Modeling:* Battelle Press, Columbus, p. 473–493.

Cressie, N., 1991, *Statistics for Spatial Data:* Wiley, New York.

Cressie, N., 1985, Fitting variogram models by weighted least squares: Math. Geol., v. 17, p. 563–586.

Cressie, N., and Hawkins, D., 1980, Robust estimation of the variogram, I: Math. Geol., v. 12, p. 115–125.

Gilbert, R., 1978, On the estimation of spatial pattern for environment contaminants, *in* M. White and P. Dunaway (Eds.), *Selected Environmental Plutonium Research Reports of the NAEG:* Nevada Applied Ecology Group, U.S. Department of Energy, Las Vegas, Nevada, p. 319–360.

Gilbert, R., and Simpson, J., 1985, Kriging for estimating spatial pattern of contaminants: Potential and problems: Env. Monitor. Assess., v. 5, p. 113–135.

Harper, W., Basinger, K., and Furr, J., 1988, Geostatistical analysis of potentiometric data in the Pennsylvanian aquifer of the Palo Duro Basin, Texas: Tech. Rep BMI/ONWI-680, Office of Nuclear Waste Isolation, Battelle Memorial Institute, Hereford, Texas.

Helterbrand, J., and Cressie, N., 1994, Models and inference for multivariate spatial processes. In preparation.

Journel, A., and Huijbregts, C., 1978, *Mining Geostatistics:* Academic Press, London.

Journel, A., and Rossi, M., 1989, When do we need a trend model in kriging? Math. Geol., v. 21, p. 715–739.

Myers, D., 1982, Matrix formulation of co-kriging: *Math. Geol.*, v. 14, p. 249–257.

Stein, A., and Corsten, L., 1991, Universal kriging and cokriging as a regression procedure: Biometrics, v. 47, p. 575–587.

Ver Hoef, J., and Cressie, N., 1993, Multivariable spatial prediction: Math. Geol., v. 25, p. 219–240.

Zhang, R., Yates, S., and Shouse, P., 1992, Prediction of soil salinity using cokriging with non-symmetric pseudo-cross-variograms, in *First Conference of the Working Group on Pedometrics of the International Society of Soil Science*: Pedometrics-92: Developments in Spatial Statistics for Soil Science, International Agricultural Centre, Wageningen, The Netherlands, p. 145–166.