Stochastic Environmental Research and Risk Assessment 13 (1999) 416–435 © Springer-Verlag 1999

Sequential kriging and cokriging: Two powerful geostatistical approaches

J. A. Vargas-Guzmán, T.-C. Jim Yeh

Abstract. A sequential linear estimator is developed in this study to progressively incorporate new or different spatial data sets into the estimation. It begins with a classical linear estimator (i.e., kriging or cokriging) to estimate means conditioned to a given observed data set. When an additional data set becomes available, the sequential estimator improves the previous estimate by using linearly weighted sums of differences between the new data set and previous estimates at sample locations. Like the classical linear estimator, the weights used in the sequential linear estimator are derived from a system of equations that contains covariances and cross-covariances between sample locations and the location where the estimate is to be made. However, the covariances and cross-covariances are conditioned upon the previous data sets.

The sequential estimator is shown to produce the best, unbiased linear estimate, and to provide the same estimates and variances as classic simple kriging or cokriging with the simultaneous use of the entire data set. However, by using data sets sequentially, this new algorithm alleviates numerical difficulties associated with the classical kriging or cokriging techniques when a large amount of data are used. It also provides a new way to incorporate additional information into a previous estimation.

Key words: Sequential linear estimator, successive linear estimator, conditional covariance, interpolation with large data sets.

1 Introduction

During the last decades, kriging and cokriging techniques have been applied extensively to many studies of subsurface hydrology. For instance, Kitanidis and Vomvoris (1983), and Hoeksema and Kitanidis (1984) applied cokriging technique to one- and two-dimensional saturated, steady flow problems for estimating hydraulic conductivity of geological media. Using cokriging, Yates and Warrick (1987) utilized soil temperature to estimate the spatial distribution of moisture content in the subsurface. Sun and Yeh (1992) extended the method to estimate conductivity using information on hydraulic head under transient saturated flow conditions. Harter and Yeh (1996) used the cokriging technique to investigate effects of conditioning using head and conductivity measurements on solute transport in the vadose zone. The same technique was also employed by Yeh and Zhang (1996) to estimate parameters of unsaturated conductivity based on

J.A. Vargas-Guzmán (✉), T.-C. Jim Yeh
Department of Hydrology and Water resources,
University of Arizona, Tucson, AZ 85721, USA

This research is funded by a DOE/Sandia contract AV-0655#1

moisture content and head measurements. Tong (1996) applied cokriging to estimate the saturated conductivity of geological media using tracer concentration measurements. Harvey and Gorelick (1995) developed a sequential approach for cokriging in which information such as head and solute arrival time was used consecutively to improve the estimates of the heterogeneous conductivity field. Li (1998) used a sequential approach to cokrige the saturated conductivity of geological media using some conductivity data first then the tracer concentration measurements. On the other hand, Yeh et al. (1995 and 1996) proposed iterative cokriging techniques for nonlinear systems in which the requirement of unbiasness and minimum variance were imposed in each iteration. This iterative approach was further extended to unsaturated flow by Zhang and Yeh (1997) to estimate parameters for unsaturated hydraulic conductivity in the vadose zone. Similarly, a quasi-linear geostatistical approach was presented by Kitanidis (1995) in an attempt to incorporate the nonlinear relationship between the parameter and secondary information of the subsurface flow system.

Despite their popularity, kriging and cokriging with large data sets can be a non-trivial problem because of numerical instabilities associated with solving large systems of equations (Davis and Grivet, 1984, Dietrich and Newsam, 1989). Davis (1975) introduced an alternative called "moving or local neighborhood kriging" which is a circular or elliptic moving window that allows to krige a central point using only the data within such a local neighborhood. As mentioned by Davis and Culhane (1984), local neighborhoods produce an effect of spurious behavior in the estimates. They also proposed an alternative that uses the covariance instead of the variogram and rearranges rows and columns in the kriging matrix to yield a symmetrically banded kriging matrix that is easier to invert. Nevertheless, such an approach still deals with large matrices and has the disadvantage that the problem is not entirely solved if the range of spatial correlation is very large.

In this paper, we present a sequential approach that resolves the numerical difficulties associated with interpolation with large data sets, using kriging or cokriging. Such a new approach allows us to incorporate a small group of data at a time or sequentially during the estimation. Therefore, our new approach is expected to be very useful for solving many problems in hydrology that require estimation with large amount of data or that require integration of many different types of information in both time and space. While the key of sequential kriging and cokriging is the successive linear estimation empirically introduced by Yeh et al. (1996), our objectives here are to provide the theory for sequential kriging and cokriging and to prove their validity.

2
Background

2.1
The successive estimator for non linearity

While applying a geostatistical approach (i.e., cokriging) to inverse problems in subsurface hydrology, Gutjahr et al. (1994) and Yeh et al. (1995, 1996) recognized the linear predictor nature of cokriging and the nonlinear relationship between hydrologic properties and response of the subsurface. To allow the geostatistical approach to consider the nonlinearity, Yeh et al. (1996) introduced a successive linear estimator of the form

$$\hat{f}^{(r+1)}(x_0) = \hat{f}^{(r)}(x_0) + \sum_{j=1}^{n} \varpi_j^{r+1} \Delta h(x_j) \qquad (1)$$

where $\hat{f}^{(r+1)}$ is the conductivity estimate at $r+1$ iteration, and $\hat{f}^{(r)}$ is the estimate at the previous iteration, r, and $\Delta h(x_j)$ is the residual between the predicted head (i.e., $h^{(r)}(x_j) = G[\hat{f}^{(r)}(x_j)]$, where G represents the flow equation) and the observed head at location j. That is,

$$\Delta h(x_j) = h(x_j) - h^{(r)}(x_j) = h(x_j) - G(\hat{f}^{(r)}(x_j)) \tag{2}$$

when $r = 1, f^{(r)}$ is the classical cokriging estimate based on some f and h measurements, and weights derived from cokriging equations that involve covariance of h and f and their cross-covariance. Once the iteration $(r > 1)$, the successive

linear estimator solves a system of equations similar to that in the classical cokriging to determine the weights, ϖ_i^{r+1}. However, it uses the covariance and cross-covariance of h and f conditioned to their previous estimates in the system of equations, instead of unconditional covariances. In other words, the successive linear estimation approach propagates conditional mean and covariances during each iteration.

The successive linear estimator propagates conditional moments to incorporate the nonlinear relationship between h and f. The concept of moment propagation can also be applied to space or time domains to reduce the size of kriging system of equations or to incorporate new data sets – sequential kriging or cokriging.

2.2 Estimation by simple kriging

Consider a spatial random field, $Z(x)$, which is second-order stationary and assume values of the random field are known at sample locations $x_1, \ldots, x_n$. The best unbiased linear estimator for the value of $z(x_o)$ at some location x_o, using all the samples simultaneously is given by the well-known classic simple kriging estimator set (e.g., Journel and Huijbregts, 1978). That is,

$$\hat{z}(x_o) = \vec{\lambda}^T \vec{z} \tag{3}$$

where $\vec{\lambda}$ is the vector of kriging weights and the vector $\vec{z}$ represents the data at sample locations $x_1, \ldots, x_n$.

Before developing the sequential kriging theory, we will discuss the classic simultaneous simple kriging expressed in terms of partitioned matrices. This result will be useful for comparing the classical kriging to our sequential approach. Suppose we split the data into two sets: $\vec{z}_p$ and $\vec{z}_s$. That is,

$$\vec{z} = \begin{bmatrix} \vec{z}_1 \\ \vec{z}_2 \\ \vdots \\ \vec{z}_{s-1} \\ \vec{z}_s \end{bmatrix} = \begin{bmatrix} \vec{z}_p \\ \vec{z}_s \end{bmatrix} \tag{4}$$

The simultaneous simple kriging estimator of Eq. (3) can be written as follows:

$$\hat{z}(x_o) = [\vec{\lambda}_{po}]^T \vec{z}_p + [\vec{\lambda}_{so}]^T \vec{z}_s \tag{5}$$

The system of kriging equations for determining the weights in the equation is

$$\mathbf{c}_{zz}\vec{\lambda}_{zo} = \vec{c}_{zo} \tag{6}$$

where

$$\mathbf{c}_{zz} = \begin{bmatrix} \mathbf{c}_{pp} & \mathbf{c}_{ps} \\ \mathbf{c}_{sp} & \mathbf{c}_{ss} \end{bmatrix} \tag{7}$$

and

$$\vec{c}_{zo} = \begin{bmatrix} \vec{c}_{po} \\ \vec{c}_{so} \end{bmatrix} \tag{8}$$

Then, the solution to the kriging system is

$$\begin{bmatrix} \vec{\lambda}_{po} \\ \vec{\lambda}_{so} \end{bmatrix} = \begin{bmatrix} \mathbf{c}_{pp} & \mathbf{c}_{ps} \\ \mathbf{c}_{sp} & \mathbf{c}_{ss} \end{bmatrix}^{-1} \begin{bmatrix} \vec{c}_{po} \\ \vec{c}_{so} \end{bmatrix} \tag{9}$$

From basic matrix theory, the inverse of a square, non-singular partitioned matrix is

$$\begin{bmatrix} \mathbf{c}_{pp} & \mathbf{c}_{ps} \\ \mathbf{c}_{sp} & \mathbf{c}_{ss} \end{bmatrix}^{-1} = \begin{bmatrix} [\mathbf{c}_{pp} - \mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}]^{-1} & -\mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}[\mathbf{c}_{ss} - \mathbf{c}_{sp}\mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}]^{-1} \\ -\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}[\mathbf{c}_{pp} - \mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}]^{-1} & [\mathbf{c}_{ss} - \mathbf{c}_{sp}\mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}]^{-1} \end{bmatrix} \tag{10}$$

Therefore, the classical simple kriging estimator following Eq. (4) expressed in a partition form for two data sets is

$$\begin{aligned} \hat{z}(x_o) = {} & ([\mathbf{c}_{pp} - \mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}]^{-1}\vec{c}_{po} - \mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}[\mathbf{c}_{ss} - \mathbf{c}_{sp}\mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}]^{-1}\vec{c}_{so})^T\vec{z}_p \\ & + ([\mathbf{c}_{ss} - \mathbf{c}_{sp}\mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}]^{-1}\vec{c}_{so} - \mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}[\mathbf{c}_{pp} - \mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}]^{-1}\vec{c}_{po})^T\vec{z}_s \end{aligned} \tag{11}$$

and the estimation variance $\mathrm{var}(\Delta z(x_o))$ is

$$\begin{aligned} \mathrm{var}(\Delta z(x_o)) = {} & [\vec{\lambda}_{po}]^T\mathbf{c}_{pp}\vec{\lambda}_{po} + 2[\vec{\lambda}_{po}]^T\mathbf{c}_{ps}\vec{\lambda}_{so} - 2[\vec{\lambda}_{po}]^T\vec{c}_{po} \\ & + [\vec{\lambda}_{so}]^T\mathbf{c}_{ss}\vec{\lambda}_{so} - 2[\vec{\lambda}_{so}]^T\vec{c}_{so} + c_{oo} \end{aligned} \tag{12}$$

3 Theory

3.1 The sequential kriging estimator

To avoid the simultaneous usage of the large number of samples, our sequential estimator will partition the data into subsets. For instance, the data set, $\vec{z}$, is partitioned into s subsets as Eq. (4). Hereafter, subscripts denote the data sets or spatial locations and superscripts denote the steps in the sequential approach described in the following sections. Instead of using all the s subsets simultaneously, our sequential estimator estimates, $z(x_o)$, at some location x_o in s steps, in vector form, that is

$$
\begin{aligned}
&\hat{z}^{(1)}(x_o) = [\vec{\varpi}^{(1)}]^T(\vec{z}_1)\\
&\hat{z}^{(2)}(x_o) = \hat{z}^{(1)}(x_o) + [\vec{\varpi}^{(2)}]^T\left(\vec{z}_2 - \hat{\vec{z}}_2^{(1)}\right)\\
&\cdots\\
&\hat{z}^{(r)}(x_o) = \hat{z}^{(r-1)}(x_o) + [\vec{\varpi}^{(r)}]^T\left(\vec{z}_r - \hat{\vec{z}}_r^{(r-1)}\right)\\
&\hat{z}^{(r+1)}(x_o) = \hat{z}^{(r)}(x_o) + [\vec{\varpi}^{(r+1)}]^T\left(\vec{z}_{r+1} - \hat{\vec{z}}_{r+1}^{(r)}\right)\\
&\cdots\\
&\hat{z}^{s}(x_o) = \hat{z}^{(s-1)}(x_o) + [\vec{\varpi}^{(s)}]^T\left(\vec{z}_s - \hat{\vec{z}}_s^{(s-1)}\right)
\end{aligned}
\tag{13}
$$

where $\vec{\varpi}^r$ are the kriging weights for the residuals of the new data included at step r. The vector estimate at new data locations from previous step is denoted by $\hat{\vec{z}}_s^{(s-1)}$. Following Eq. (13), the sequential estimator at any step is the previous estimate plus the contribution from the new data set.

This sequential estimator of Eq. (13) can be analyzed at any step in terms of the classic simultaneous simple kriging. For example, consider the partition in Eq. (4), then at the last sequential kriging step in Eq. (13) the estimate from the previous step is equivalent to the simultaneous kriging estimate when only $\vec{z}_p$ is utilized as follows:

$$
\hat{z}^{(s-1)}(x_o) = [\vec{\theta}_{po}]^T\vec{z}_p \tag{14}
$$

where the "old" data vector $\vec{z}_p$ represents the data set up to step $s-1$ and weights $\vec{\theta}_{po}$ are the simultaneous kriging weights associated with the data vector $\vec{z}_p$ and the location x_o. similarly, the term, $\hat{\vec{z}}_s^{(s-1)}$, in Eq. (13) can be expressed as

$$
\hat{\vec{z}}_s^{(s-1)}(x_s) = \left[\bar{\theta}_{ps}\right]^T\vec{z}_p \tag{15}
$$

where $\bar{\theta}_{ps}$ is the weights matrix for estimating the subset $\hat{\vec{z}}_s$ at new data locations using the old data set, $\vec{z}_p$. Application of Eqs. (14) and (15) to the last step of the sequential kriging estimator of Eq. (13) leads to

$$
\hat{z}^{(s)}(x_o) = [\vec{\theta}_{po}]^T\vec{z}_p + [\vec{\varpi}^{(s)}]^T(\vec{z}_s - \left[\bar{\theta}_{ps}\right]^T\vec{z}_p) \tag{16}
$$

or

$$
\hat{z}^{(s)}(x_o) = \left[\left[\vec{\theta}_{po}\right]^T - [\vec{\varpi}^{(s)}]^T[\bar{\theta}_{ps}]^T\right]\vec{z}_p + [\vec{\varpi}^{(s)}]^T\vec{z}_s \tag{17}
$$

Comparing Eq. (17) to Eq. (5) leads to

$$
\vec{\varpi}^{(s)} = \vec{\lambda}_{so} \tag{18}
$$

We will prove in Sect. 3.3 that the weights $\vec{\varpi}^{(s)}$ of the last sequential step s are the same as the corresponding (sub)vector of weights $\vec{\lambda}_s$ of the simultaneous simple kriging. Now, drop out the data $\vec{z}_s$ from the whole data set, so one step before the available data vector $\vec{z}_p$ is

$$\vec{z}_p = \begin{bmatrix} \vec{z}_1 \\ \vec{z}_2 \\ \vdots \\ \vec{z}_{s-1} \end{bmatrix} = \begin{bmatrix} \vec{z}_q \\ \vec{z}_{s-1} \end{bmatrix} \tag{19}$$

We now apply the same analysis as before. The simultaneous kriging estimate is equivalent to the updated estimate at the $s-1$ step in the sequential approach. In a form analogous to Eq. (17), the result is

$$\hat{z}^{(s-1)}(x_o) = \left[[\vec{\vartheta}_{qo}]^T - [\vec{\varpi}^{(s-1)}]^T [\bar{\vartheta}_{q(s-1)}]^T \right] \vec{z}_q + [\vec{\varpi}^{(s-1)}]^T \vec{z}_{s-1} \tag{20}$$

From Eqs. (14) and (19) we have

$$\vec{\theta}_{po} = \begin{bmatrix} \vec{\theta}_{qo} \\ \vec{\theta}_{(s-1)o} \end{bmatrix}$$

and applying Eq. (18) to the case of the estimator in Eq. (20) implies

$$\vec{\theta}_{(s-1)o} = \vec{\varpi}^{(s-1)} \tag{21}$$

The same analysis can be applied for a previous step. This gives

$$\vec{\vartheta}_{(s-2)o} = \vec{\varpi}^{(s-2)} \tag{22}$$

Notice that every time data are dropped out of the system, a new system of kriging equations can be solved. Eventually, the first vector of data $\vec{z}_1$ is reached after continuing backward, dropping the sets of data. Sequential kriging is opposite to this backward step analysis. As soon will be apparent, the whole analysis for a large number of successive steps is not needed. The development of the subsequent theory only needs to consider a partition of the data into two sets: the old or previous data at given locations and the new or additional data set at different locations. Extension to the sequential approach will be made by using this backward step analysis.

3.2. Derivation of the sequential kriging equations

As shown in Eqs. (17) and (20) at any stage of the sequential approach, the estimator can be split into two parts: one associated with the old data $\vec{z}_p$, and the second which corresponds to an additional or new data set $\vec{z}_s$. Utilizing Eqs. (12) and (17), the variance of the estimate can be written as

$$\begin{aligned} \mathrm{var}(\Delta z(x_o)) = {} & [\vec{\theta}_{po}^T - \vec{\varpi}_{so}^T \bar{\theta}_{ps}^T] \mathbf{c}_{pp} [\vec{\theta}_{po}^T - \vec{\varpi}_{so}^T \bar{\theta}_{ps}^T]^T + 2[\vec{\theta}_{po}^T - \vec{\varpi}_{so}^T \bar{\theta}_{ps}^T] \mathbf{c}_{ps} \vec{\varpi}_{so} \\ & - 2[\vec{\theta}_{po}^T - \vec{\varpi}_{so}^T \bar{\theta}_{ps}^T] \vec{c}_{po} + [\vec{\varpi}_{so}]^T \mathbf{c}_{ss} \vec{\varpi}_{so} - 2[\vec{\varpi}_{so}]^T \vec{c}_{so} + c_{oo} \end{aligned} \tag{23}$$

After a few operations, Eq. (23) can be expressed as

$$\begin{aligned}\mathrm{var}(\Delta z(x_o)) = {} & \vec{\theta}_{po}^{T}\mathbf{c}_{pp}\vec{\theta}_{po} - \vec{\varpi}_{so}^{T}\bar{\theta}_{ps}^{T}\mathbf{c}_{pp}\vec{\theta}_{po} - \vec{\theta}_{po}^{T}\mathbf{c}_{pp}\bar{\theta}_{ps}\vec{\varpi}_{so} \\ & + \vec{\varpi}_{so}^{T}\bar{\theta}_{ps}^{T}\mathbf{c}_{pp}\bar{\theta}_{ps}\vec{\varpi}_{so} + 2\vec{\theta}_{po}^{T}\mathbf{c}_{ps}\vec{\varpi}_{so} - 2\vec{\varpi}_{so}^{T}\bar{\theta}_{ps}^{T}\mathbf{c}_{ps}\vec{\varpi}_{so} \\ & - 2\vec{\theta}_{po}^{T}\vec{c}_{po} + 2\vec{\varpi}_{so}^{T}\bar{\theta}_{ps}^{T}\vec{c}_{po} + \vec{\varpi}_{so}^{T}\mathbf{c}_{ss}\vec{\varpi}_{so} - 2\vec{\varpi}_{so}^{T}\vec{c}_{so} + c_{oo} \end{aligned} \tag{24}$$

where $\mathbf{c}_{pp}$, $\mathbf{c}_{ss}$ and $\mathbf{c}_{ps}$ are the covariance matrices between pairs of point random variables of the previous and additional data. $\vec{c}_{po}, \vec{c}_{so}$ and c_{oo} are the covariances with respect to the estimated point location. Differentiating Eq. (24) with respect to the weights $\bar{\theta}_{po}$ and setting the result equal to zero leads to a first kriging:

$$\mathbf{c}_{pp}\vec{\theta}_{po} - \vec{c}_{po} = 0 \tag{25}$$

Notice that the following condition,

$$\frac{\partial^2(\mathrm{var}(\Delta z(x_o)))}{\partial\vec{\theta}_{po}^{2}} \geq 0 \tag{26}$$

assures that the estimation minimizes the previous estimation variance.

Differentiating (24) with respect to the other set of weights, ϖ_{so}, and letting the result equal to zero, we obtain

$$\begin{aligned} & - \bar{\theta}_{ps}^{T}\mathbf{c}_{pp}\vec{\theta}_{po} - \bar{\theta}_{ps}^{T}\mathbf{c}_{pp}\vec{\theta}_{po} + 2\bar{\theta}_{ps}^{T}\mathbf{c}_{pp}\bar{\theta}_{ps}\vec{\varpi}_{so} + 2\mathbf{c}_{ps}^{T}\vec{\theta}_{po} - 4\bar{\theta}_{ps}^{T}\mathbf{c}_{ps}\vec{\varpi}_{so} \\ & + 2\bar{\theta}_{ps}^{T}\vec{c}_{po} + 2\mathbf{c}_{ss}\vec{\varpi}_{so} - 2\vec{c}_{so} = 0 \end{aligned} \tag{27}$$

After rearrangement, Eq. (28) can be rewritten as:

$$[2\bar{\theta}_{ps}^{T}\mathbf{c}_{pp}\bar{\theta}_{ps} - 4\bar{\theta}_{ps}^{T}\mathbf{c}_{ps} + 2\mathbf{c}_{ss}]\vec{\varpi}_{so} = 2\bar{\theta}_{ps}^{T}\mathbf{c}_{pp}\vec{\theta}_{po} - 2\mathbf{c}_{ps}^{T}\vec{\theta}_{po} - 2\bar{\theta}_{ps}^{T}\vec{c}_{po} + 2\vec{c}_{so} \tag{28}$$

From solving Eq. (25) for estimating the locations s using the previous data, recall $\bar{\theta}_{ps} = \mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}$. Then, Eq. (28) can be expressed as:

$$[\mathbf{c}_{ss} - \bar{\theta}_{ps}^{T}\mathbf{c}_{ps}]\vec{\varpi}_{so} = [\vec{c}_{so} - \bar{\theta}_{ps}^{T}\mathbf{c}_{po}] \tag{29}$$

and in more abbreviated form is

$$\bar{\varepsilon}_{ss}\vec{\varpi}_{so} = \vec{\varepsilon}_{so} \tag{30}$$

where $\bar{\varepsilon}_{ss}$ denotes the conditional covariance matrix for residuals of the new data, i.e.,

$$\bar{\varepsilon}_{ss} = [\mathbf{c}_{ss} - \bar{\theta}_{ps}^{T}\mathbf{c}_{ps}] \tag{31}$$

and $\bar{\varepsilon}_{so}$ represents the conditional covariance vector for the residuals of the new data respect to the residuals at the location to be estimated. That is,

$$\vec{\varepsilon}_{so} = [\vec{c}_{so} - \bar{\theta}_{ps}^{T}\mathbf{c}_{po}] \tag{32}$$

From the previous estimation the matrix $\bar{\theta}_{ps} = \mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}$, this leads to

$$\bar{\varepsilon}_{ss} = [\mathbf{c}_{ss} - \mathbf{c}_{ps}^{T}[\mathbf{c}_{pp}^{-1}]^{T}\mathbf{c}_{ps}] \tag{33}$$

and

$$\vec{\varepsilon}_{so} = [\vec{c}_{so} - \mathbf{c}_{ps}^{T}[\mathbf{c}_{pp}^{-1}]^{T}\vec{c}_{po}] \tag{34}$$

Thus, following Eq. (30). the kriging weights from the sequential approach can be expressed as

$$\vec{\varpi}_{so} = [\mathbf{c}_{ss} - \mathbf{c}_{ps}^{T}[\mathbf{c}_{pp}^{-1}]^{T}\mathbf{c}_{ps}]^{-1}[\vec{c}_{so} - \mathbf{c}_{ps}^{T}[\mathbf{c}_{pp}^{-1}]^{T}\vec{c}_{po}] \tag{35}$$

This result will allow us to find a link between the classic simultaneous simple kriging and the sequential kriging.

Moreover, the second order derivative is written as:

$$\frac{\partial^2(\mathrm{var}(\Delta z(x_o)))}{\partial\vec{\varpi}_{so}^2} = 2\bar{\theta}_{ps}^{T}\mathbf{c}_{ps} - 4\bar{\theta}_{ps}^{T}\mathbf{c}_{ps} + 2\mathbf{c}_{ss} \tag{36}$$

From Eqs. (31) and (36), we realize that the condition $\|\mathbf{c}_{ss}\| \geq \|\bar{\varepsilon}_{ss}\|$ must be met to get a minimal estimation variance.

At this point, we have derived a conditional kriging system, given by Eq. (30) and the residuals conditional covariances of Eqs. (31) and (32). The above results may be generalized by considering any two steps r and $r+1$ in the sequential approach.

Moreover, the whole data set at step $r+1$ is $\vec{z} = \begin{bmatrix} \vec{z}_p \\ \vec{z}_{s^{r+1}} \end{bmatrix}$ which following Eqs. (4) and (19) also can be written as

$$\vec{z} = \begin{bmatrix} \vec{z}_q \\ \vec{z}_{s^r} \\ \vec{z}_{s^{r+1}} \end{bmatrix} \tag{37}$$

From Eq. (20) and the derivation of conditional kriging above, the residual conditional covariance for a couple of locations (x_i, x_j) at step r is

$$\varepsilon_{ij}^{r} = [c_{ij} - \vec{\vartheta}_{q_i}^{T}\vec{c}_{qj}] \tag{38}$$

and from Eq. (32) at the next step $r+1$, this is

$$\varepsilon_{ij}^{r+1} = [c_{ij} - \vec{\theta}_{pi}^{T}\vec{c}_{pj}] \tag{39}$$

where p implies the estimation is made with all available data prior to step $r+1$, and q implies the estimation is made with all available data prior to step r. The new information at step $r+1$ is $\vec{z}_{s^{r+1}}$. Due to the partitioning, we can write Eq. (39) as:

$$\varepsilon_{ij}^{r+1} = \left[c_{ij} - \begin{bmatrix} \vec{\theta}_{qi}^{T} \\ \vec{\theta}_{s^r i}^{T} \end{bmatrix}^{T} \begin{bmatrix} \vec{c}_{qj} \\ \vec{c}_{s^r j} \end{bmatrix} \right] \tag{40}$$

Replacing the weights by the sequential weights according to Eq. (20) yields,

$$\varepsilon_{ij}^{r+1} = \left[c_{ij} - \begin{bmatrix} \vec{\vartheta}_{qi}^T - \vec{\varpi}_{s^r i}^T \bar{\vartheta}_{s^r q}^T \\ \vec{\varpi}_{s^r i}^T \end{bmatrix}^T \begin{bmatrix} \vec{c}_{qj} \\ \vec{c}_{s^r j} \end{bmatrix} \right] \tag{41}$$

Developing the product, we get

$$\varepsilon_{ij}^{r+1} = c_{ij} - \vec{\vartheta}_{qi}^T \vec{c}_{qj} - \vec{\varpi}_{s^r i}^T \left[\vec{c}_{s^r j} - \bar{\vartheta}_{s^r q}^T \vec{c}_{qj} \right] \tag{42}$$

which sequentially leads to an updated residual conditional covariance

$$\varepsilon_{ij}^{r+1} = \varepsilon_{ij}^{r} - \vec{\varpi}_{s^r i}^T \vec{\varepsilon}_{s^r j}^{r} \tag{43}$$

This finding allows generalizing conditional residual kriging to a sequential approach in which the variance is minimized by successive steps. The condition to minimize the estimation variance, expressed above, is also generalized as $\|\bar{\varepsilon}_{ss}^{r}\| \geq \|\bar{\varepsilon}_{ss}^{r+1}\|$.

Equation (25) gives the kriging system for the first step and Eq. (30) is used in all successive steps using updated conditional covariances given by Eq. (43). The next equations represent this new explicit form of kriging

$$\begin{aligned} &\mathbf{c}_{pp}^{1} \vec{\varpi}_{so}^{1} = \vec{c}_{po}^{1} \\ &\vdots \\ &\bar{\varepsilon}_{s^r s^r}^{r} \vec{\varpi}_{s^r o}^{r} = \vec{\varepsilon}_{s^r o}^{r} \\ &\bar{\varepsilon}_{s^{r+1} s^{r+1}}^{r+1} \vec{\varpi}_{s^{r+1} o}^{r+1} = \vec{\varepsilon}_{s^{r+1} o}^{r+1} \\ &\vdots \\ &\bar{\varepsilon}_{ss}^{s} \vec{\varpi}_{so}^{s} = \vec{\varepsilon}_{so}^{s} \end{aligned} \tag{44}$$

The assumption made in the form of the estimator in Eqs. (17) and (20) are justified by results in the following section which prove that results obtained by this new approach give the same results as simultaneous simple global kriging.

3.3
Equivalence between classic simple kriging and sequential kriging

In this section, we will demonstrate that the sequential kriging estimator and the simple kriging estimator both produce the same estimate.

The simple kriging estimator from the partitioned matrix in Eq. (11) is

$$\begin{aligned} \hat{z}_{\text{classic}}(x_o) = &\left(\left[\mathbf{c}_{pp} - \mathbf{c}_{ps} \mathbf{c}_{ss}^{-1} \mathbf{c}_{sp} \right]^{-1} \vec{c}_{po} - \mathbf{c}_{pp}^{-1} \mathbf{c}_{ps} \left[\mathbf{c}_{ss} - \mathbf{c}_{sp} \mathbf{c}_{pp}^{-1} \mathbf{c}_{ps} \right]^{-1} \vec{c}_{so} \right)^T \vec{z}_p \\ &+ \left(\left[\mathbf{c}_{ss} - \mathbf{c}_{sp} \mathbf{c}_{pp}^{-1} \mathbf{c}_{ps} \right]^{-1} \vec{c}_{so} - \mathbf{c}_{ss}^{-1} \mathbf{c}_{sp} \left[\mathbf{c}_{pp} - \mathbf{c}_{ps} \mathbf{c}_{ss}^{-1} \mathbf{c}_{sp} \right]^{-1} \vec{c}_{po} \right)^T \vec{z}_s \end{aligned} \tag{45}$$

On the other hand, from Eqs. (17), (25) and (35), the sequential kriging estimator in terms of the partitioned matrix is

$$\hat{z}_{\text{sequential}}(x_o) = \left((\mathbf{c}_{pp}^{-1}\vec{c}_{po})^T - [\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}[\vec{c}_{so} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po}][\mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}]\right)^T \vec{z}_p + \left([\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}[\vec{c}_{so} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po}]\right)^T \vec{z}_s \quad (46)$$

From the second terms in the right hand side of these two equations, a condition for equivalence is

$$[\mathbf{c}_{ss} - \mathbf{c}_{sp}\mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}]^{-1}\vec{c}_{so} - \mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}\left[\mathbf{c}_{pp} - \mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}\right]^{-1}\vec{c}_{po} = [\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}[\vec{c}_{so} - \mathbf{c}_{ps}^T.[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po}] \quad (47)$$

Also consider, $\mathbf{c}_{pp} = \mathbf{c}_{pp}^T$ and $\mathbf{c}_{ps} = \mathbf{c}_{sp}^T$ are square here. Then,

$$-\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}\left[\mathbf{c}_{pp} - \mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}\right]^{-1} = -[\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}[\mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T] \quad (48)$$

Introducing all terms within the inverse matrices and making few simplifications, yields

$$-[\mathbf{c}_{ss}\mathbf{c}_{sp}^{-1}\mathbf{c}_{pp} - \mathbf{c}_{ss}\mathbf{c}_{sp}^{-1}\mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}]^{-1} = -[\mathbf{c}_{ss}[\mathbf{c}_{ps}^{-1}]^T\mathbf{c}_{pp}^T - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}[\mathbf{c}_{ps}^{-1}]^T\mathbf{c}_{pp}^T]^{-1} \quad (49)$$

The first equivalence condition is satisfactorily accomplished because,

$$[\mathbf{c}_{ss}\mathbf{c}_{sp}^{-1}\mathbf{c}_{pp} - \mathbf{c}_{ps}]^{-1} = [\mathbf{c}_{ss}[\mathbf{c}_{ps}^{-1}]^T\mathbf{c}_{pp}^T - \mathbf{c}_{ps}]^{-1} \quad (50)$$

Notice, this identify provides a proof of Eq. (18)

From the first terms of Eqs. (45) and (46), the next condition for equivalence is

$$\left[\mathbf{c}_{pp} - \mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}\right]^{-1}\vec{c}_{po} - \mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}[\mathbf{c}_{ss} - \mathbf{c}_{sp}\mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}]^{-1}\vec{c}_{so} = (\mathbf{c}_{pp}^{-1}\vec{c}_{po}) - [\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}[\vec{c}_{so} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po}][\mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}] \quad (51)$$

calculating some products, the terms containing $\vec{c}_{so}$ cancel out and this leads to,

$$\left[\mathbf{c}_{pp} - \mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}\right]^{-1}\vec{c}_{po} = \left[(\mathbf{c}_{pp}^{-1}) + [\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}\mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}\right]\vec{c}_{po} \quad (52)$$

making simplifications and introducing terms within the inverse matrix yields,

$$\left[\mathbf{c}_{pp} - \mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}\right]^{-1} = (\mathbf{c}_{pp}^{-1}) + [\mathbf{c}_{ss}[\mathbf{c}_{ps}^{-1}]^T\mathbf{c}_{pp}^T\mathbf{c}_{pp}\mathbf{c}_{ps}^{-1} - \mathbf{c}_{pp}]^{-1} \quad (53)$$

To prove this last identity, we multiply both terms by the inverse of the first term. Then, developing products and introducing terms within the inverse matrices give

$$-\mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}[\mathbf{c}_{ss}[\mathbf{c}_{ps}^{-1}]^T\mathbf{c}_{pp}^T\mathbf{c}_{ps}^{-1} - \mathbf{I}] + \mathbf{I} = \mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp} \quad (54)$$

Finally,

$$-\mathbf{I} + \mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp} = -\mathbf{I} + \mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp} \tag{55}$$

Identities, Eqs. (50) and (55), prove that Eqs. (45) and (46) are equal, implying that sequential kriging will yield an identical estimate as simple kriging using all the data set simultaneously. Notice that the order of the data set being used does not affect the final estimate.

3.4 Sequential kriging variance

Now we develop a mathematical proof that the simple kriging variance of the simultaneous classic approach will be equal to the total or final estimation variance of the sequential approach. This also shows that the sequential estimation variance will be successively reduced as more data are utilized in the refinement of the estimation made.

The classic kriging variance in matrix form is

$$K_v(x_o) = \sigma^2 - \vec{\lambda}^T\vec{c}_{zo} = \sigma^2 - v(x_o) \tag{56}$$

where σ^2 is the variance. In classic simultaneous simple kriging, the variance of the estimated values can be determined from Eqs. (8) and (11). This is

$$\begin{aligned} v(x_o) = {} & \left(\left[\mathbf{c}_{pp} - \mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}\right]^{-1}\vec{c}_{po} - \mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}\left[\mathbf{c}_{ss} - \mathbf{c}_{sp}\mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}\right]^{-1}\vec{c}_{so}\right)^T\vec{c}_{po} \\ & + \left(\left[\mathbf{c}_{ss} - \mathbf{c}_{sp}\mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}\right]^{-1}\vec{c}_{so} - \mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}\left[\mathbf{c}_{pp} - \mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}\right]^{-1}\vec{c}_{po}\right)^T\vec{c}_{so} \end{aligned} \tag{57}$$

In sequential kriging, we also have the cumulated variance of the estimated values from previous steps. This is equivalent to the smooth variance of the estimates from previous data:

$$v_{pp}(x_o) = [\vec{\theta}_{po}]^T\vec{c}_{po} \tag{58}$$

or

$$v_{pp}(x_o) = \vec{c}_{po}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po} \tag{59}$$

with sequential kriging, the variance of the estimated residuals at any step is

$$v_{ss}(x_o) = \vec{\varpi}^T\vec{\varepsilon}_{so} \tag{60}$$

using Eqs. (34) and (35), this is

$$v_{ss}(x_o) = \left[\left[\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}\right]^{-1}\left[\vec{c}_{so} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po}\right]\right]^T\left[\vec{c}_{so} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po}\right] \tag{61}$$

The residuals are part of additional data that are independent of the previous data and the kriging variance of estimated residuals is independent of the variance of the previous estimates. Accordingly, the following algebraic sum should be true:

$$v(x_o) = v_{pp}(x_o) + v_{ss}(x_o) \tag{62}$$

or

$$v(x_o) = \vec{c}_{po}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po} + [\vec{c}_{so} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po}]^T \times [[\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}]^T[\vec{c}_{so} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po}] \quad (63)$$

calculating some products leads to

$$\begin{aligned} v(x_o) = {} & \vec{c}_{po}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po} + \vec{c}_{so}^T[[\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}]^T\vec{c}_{so} \\ & - [\mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po}]^T[[\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}]^T\vec{c}_{so} \\ & - \vec{c}_{so}^T[[\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}]^T\mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po} \\ & + [\mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po}]^T[[\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}]^T\mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po} \end{aligned} \quad (64)$$

The second and third terms do not need any operation. Introducing covariance matrices within the inverse in the last term yields

$$\begin{aligned} v(x_o) = {} & \vec{c}_{so}^T[[\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}]^T\vec{c}_{so} - [\mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po}]^T[[\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}]^T\vec{c}_{so} \\ & + \vec{c}_{po}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po} - \vec{c}_{so}^T[[\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}]^T\mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po} \\ & + [\vec{c}_{po}]^T[[[\mathbf{c}_{ps}^{-1}]\mathbf{c}_{pp}\mathbf{c}_{ss}[\mathbf{c}_{ps}^{-1}]^T\mathbf{c}_{pp}^T - [\mathbf{c}_{ps}^{-1}]\mathbf{c}_{pp}\mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}[\mathbf{c}_{ps}^{-1}]^T\mathbf{c}_{pp}^T]^{-1}]^T\vec{c}_{po} \end{aligned} \quad (65)$$

Then, terms can be arrayed and after few operations this is,

$$\begin{aligned} v(x_o) = {} & \vec{c}_{so}^T[[\mathbf{c}_{ss} - \mathbf{c}_{sp}[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}]^T\vec{c}_{so} - [\mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po}]^T[[\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}]^T\vec{c}_{so} \\ & - \vec{c}_{so}^T\mathbf{c}_{sp}[\mathbf{c}_{pp}^{-1}]^T[[\mathbf{c}_{ss} - \mathbf{c}_{sp}[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}]^T\vec{c}_{po} \\ & + \vec{c}_{po}^T[[\mathbf{c}_{pp}^{-1}]^T + [[\mathbf{c}_{ps}^{-1}]\mathbf{c}_{pp}\mathbf{c}_{ss}[\mathbf{c}_{ps}^{-1}]^T\mathbf{c}_{pp}^T - \mathbf{c}_{ps}[\mathbf{c}_{ps}^{-1}]\mathbf{c}_{pp}]^{-1}]^T\vec{c}_{po} \end{aligned} \quad (66)$$

The last term on the right hand side can be modified similarly to the identity shown in Eq. (53), this is

$$\begin{aligned} & \vec{c}_{po}^T[[\mathbf{c}_{pp}^{-1}]^T + [[\mathbf{c}_{ps}^{-1}]\mathbf{c}_{pp}\mathbf{c}_{ss}[\mathbf{c}_{ps}^{-1}]^T\mathbf{c}_{pp}^T - \mathbf{c}_{ps}[\mathbf{c}_{ps}^{-1}]^T\mathbf{c}_{pp}^T]^{-1}]^T\vec{c}_{po} \\ & \quad = ([\mathbf{c}_{pp} - \mathbf{c}_{ps}\mathbf{c}_{ss}^{-1}\mathbf{c}_{sp}]^{-1}\vec{c}_{po})^T\vec{c}_{po} \end{aligned} \quad (67)$$

Then,

$$-[\mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\vec{c}_{po}]^T[[\mathbf{c}_{ss} - \mathbf{c}_{ps}^T[\mathbf{c}_{pp}^{-1}]^T\mathbf{c}_{ps}]^{-1}]^T\vec{c}_{so} = (-\mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}[\mathbf{c}_{ss} - \mathbf{c}_{sp}\mathbf{c}_{pp}^{-1}\mathbf{c}_{ps}]^{-1}\vec{c}_{so})^T\vec{c}_{po} \quad (68)$$

Thus, Eq. (66) becomes the same as Eq. (57). This proves that the simultaneous and the final sequential kriging variances are the same, and also implies that the assumption of independence between residuals and old estimations is correct. Considering the sequential approach, the sequential kriging variance is $K_v(x_o) = \sigma^2 - \sum_{i-1}^{r} v_{s_i s_i}^i(x_o)$ and decreases as more data are available for estima-

tion. In other words, residuals will become smooth and estimates will increase their variance as more data are included in surrounding space of the estimated location.

3.5
Sequential cokriging

Most of the kriging theory is applicable to the multivariate case or cokriging case. Similarities in notation between kriging and cokriging arise particularly when matrix notation is utilized (Myers, 1982). Here, we present a new form of cokriging, which we term sequential cokriging. This type of technique was empirically suggested by Harvey and Gorelick (1995) to estimate hydraulic conductivity fields conditioned on head and solute arrival time consecutively. Similarly, Li (1998) empirically used this technique to estimate conductivity in variably saturated flow field using head and concentration measurements.

A vector random function $\vec{Z}(x)$ is made of a set of spatially cross-correlated random functions, each one of which corresponds to a particular statistical attribute. Then, there is one multivariate random variable $\vec{Z}(x_i)$ at each location. If data are available at a point, the multivariate random variable takes constant values for the known attributes. Since the random variables are spatially correlated, the available data can be interpolated by cokriging to provide estimates at any non-sampling location.

For cokriging, we need multivariate covariance matrices. The set p of old data is at n locations and the set s of new data is at m locations. The numerical multivariate covariance matrices for attributes $\{z, w, \ldots, q\}$ are $\mathbf{C}_{pp}, \mathbf{C}_{ss}$ and $\mathbf{C}_{ps}$, where (capital) $\mathbf{C}$ means multivariate covariance. These are,

$$\mathbf{C}_{pp} = \begin{bmatrix} \mathbf{c}_{z_p z_p} & \mathbf{c}_{z_p w_p} & \cdots & \mathbf{c}_{z_p q_p} \\ \mathbf{c}_{w_p z_p} & \mathbf{c}_{w_p w_p} & \cdots & \mathbf{c}_{w_p q_p} \\ & \vdots & & \\ \mathbf{c}_{q_p z_p} & \mathbf{c}_{q_p w_p} & \cdots & \mathbf{c}_{q_p q_p} \end{bmatrix} \qquad \mathbf{C}_{ss} = \begin{bmatrix} \mathbf{c}_{z_s z_s} & \mathbf{c}_{z_s w_s} & \cdots & \mathbf{c}_{z_s q_s} \\ \mathbf{c}_{w_s z_s} & \mathbf{c}_{w_s w_s} & \cdots & \mathbf{c}_{w_s q_s} \\ & \vdots & & \\ \mathbf{c}_{q_s z_s} & \mathbf{c}_{q_s w_s} & \cdots & \mathbf{c}_{q_s q_s} \end{bmatrix}$$

$$\mathbf{C}_{ps} = \begin{bmatrix} \mathbf{c}_{z_p z_s} & \mathbf{c}_{z_p w_s} & \cdots & \mathbf{c}_{z_p q_s} \\ \mathbf{c}_{w_p z_s} & \mathbf{c}_{w_p w_s} & \cdots & \mathbf{c}_{w_p q_s} \\ & \vdots & & \\ \mathbf{c}_{q_p z_s} & \mathbf{c}_{q_p w_s} & \cdots & \mathbf{c}_{q_p q_s} \end{bmatrix} \tag{69}$$

Notice, all the diagonal terms are (auto)covariance matrices and the off-diagonal terms are matrices of cross-covariances. The covariances between data and any estimated locations are

$$\mathbf{C}_{po} = \begin{bmatrix} \vec{c}_{z_p z_o} & \vec{c}_{z_p w_o} & \cdots & \vec{c}_{z_p q_o} \\ \vec{c}_{w_p z_o} & \vec{c}_{w_p w_o} & \cdots & \vec{c}_{w_p q_o} \\ & \vdots & & \\ \vec{c}_{q_p z_o} & \vec{c}_{q_p w_o} & \cdots & \vec{c}_{q_p q_o} \end{bmatrix} \qquad \mathbf{C}_{so} = \begin{bmatrix} \vec{c}_{z_s z_o} & \vec{c}_{z_s w_o} & \cdots & \vec{c}_{z_s q_o} \\ \vec{c}_{w_s z_o} & \vec{c}_{w_s w_o} & \cdots & \vec{c}_{w_s q_o} \\ & \vdots & & \\ \vec{c}_{q_s z_o} & \vec{c}_{q_s w_o} & \cdots & \vec{c}_{q_s q_o} \end{bmatrix} \tag{70}$$

Also, $\mathbf{C}_{sp} = \left[\mathbf{C}_{ps}\right]^T$, and consider $\mathbf{C}_{oo}$ as the zero lag distance multivariate covariance matrix for the attributes.

The matrix form of the estimation variance in terms of the cokriging weights and covariances is given as follows:

$$[\mathrm{var}(\Delta z(x_o))] = \begin{bmatrix} \bar{\Gamma}_{po} \\ \bar{\Gamma}_{so} \\ -\mathbf{I} \end{bmatrix}^T \begin{bmatrix} \mathbf{C}_{pp} & \mathbf{C}_{ps} & \mathbf{C}_{po} \\ \mathbf{C}_{sp} & \mathbf{C}_{ss} & \mathbf{C}_{so} \\ \mathbf{C}_{op} & \mathbf{C}_{os} & \mathbf{C}_{oo} \end{bmatrix} \begin{bmatrix} \bar{\Gamma}_{po} \\ \bar{\Gamma}_{so} \\ -\mathbf{I} \end{bmatrix} \tag{71}$$

and

$$\begin{aligned}[\mathrm{var}(\Delta z(x_o))] = {} & \left[\bar{\Gamma}_{po}\right]^T \left[\mathbf{C}_{pp}\right] \bar{\Gamma}_{po} + 2\left[\bar{\Gamma}_{po}\right]^T \left[\mathbf{C}_{ps}\right] \bar{\Gamma}_{so} \\ & - 2\left[\bar{\Gamma}_{po}\right]^T \mathbf{C}_{po} + [\bar{\Gamma}_{so}]^T \mathbf{C}_{ss} \bar{\Gamma}_{so} - 2[\bar{\Gamma}_{so}]^T \mathbf{C}_{so} + \mathbf{C}_{oo}\end{aligned} \tag{72}$$

The matrices of weights are composed of column vectors, and each vector estimates a particular attribute. Minimization of Eq. (72) leads to the classical cokriging expressed in partitioned form. Instead of doing that, we extend the sequential estimator of Eq. (17) to the cokriging case. This is

$$\hat{\vec{z}}^{(s)}(x_o) = [[\bar{\theta}_{po}]^T - [\bar{\varpi}_{so}]^T [\bar{\theta}_{ps}]^T] z_p + [\bar{\varpi}_{so}]^T z_s \tag{73}$$

where $\left[\bar{\Gamma}_{po}\right]^T = \left[\bar{\theta}_{po}\right]^T - [\bar{\varpi}_{so}]^T \left[\bar{\theta}_{ps}\right]^T$ and $[\bar{\Gamma}_{so}]^T = [\bar{\varpi}_{so}]^T$. It is important to notice that vectors in the univariate approach are matrices in cokriging. The data matrices $\mathbf{z}_p$ and $\mathbf{z}_s$ have q column vectors, one for each attribute. Matrices of weights for the previous and additional data sets are as follows:

$$\bar{\theta}_{po} = \begin{bmatrix} \vec{\alpha}_{z_p z_o} & \vec{\mu}_{z_p w_o} & \cdots & \vec{\zeta}_{z_p q_o} \\ \vec{\beta}_{w_p z_o} & \vec{\nu}_{w_p w_o} & \cdots & \vec{\eta}_{w_p q_o} \\ \vdots & & & \\ \vec{\gamma}_{q_p z_o} & \vec{\varphi}_{q_p w_o} & \cdots & \vec{\tau}_{q_p q_o} \end{bmatrix} \tag{74}$$

and

$$\bar{\varpi}_{so} = \begin{bmatrix} \vec{\alpha}_{z_s z_o} & \vec{\mu}_{z_s w_o} & \cdots & \vec{\zeta}_{z_s q_o} \\ \vec{\beta}_{w_s z_o} & \vec{\nu}_{w_s w_o} & \cdots & \vec{\eta}_{w_s q_o} \\ \vdots & & & \\ \vec{\gamma}_{q_s z_o} & \vec{\varphi}_{q_s w_o} & \cdots & \vec{\tau}_{q_s q_o} \end{bmatrix} \tag{75}$$

The vectors in the major diagonal are weights from an attribute that are used to estimate the same attribute. On the other hand, the off-diagonal vectors are weights of one attribute that are used to estimate a different attribute. Notice that the elements of $\bar{\theta}_{ps}$ for old versus new data are n by m matrices. Replacing matrices defined in Eqs. (69) and (70) into Eq. (71) gives

$$[\mathrm{var}(\Delta z(x_o))] = \begin{bmatrix} \left[\bar{\theta}_{po}\right] - \left[\bar{\theta}_{ps}\right][\bar{\varpi}_{so}] \\ \bar{\varpi}_{so} \\ -\mathbf{I} \end{bmatrix}^T \begin{bmatrix} \mathbf{C}_{pp} & \mathbf{C}_{ps} & \mathbf{C}_{po} \\ \mathbf{C}_{sp} & \mathbf{C}_{ss} & \mathbf{C}_{so} \\ \mathbf{C}_{op} & \mathbf{C}_{os} & \mathbf{C}_{oo} \end{bmatrix} \begin{bmatrix} \left[\bar{\theta}_{po}\right] - \left[\bar{\theta}_{ps}\right][\bar{\varpi}_{so}] \\ \bar{\varpi}_{so} \\ -\mathbf{I} \end{bmatrix} \tag{76}$$

The minimum variance is obtained by taking derivatives of the matrix of estimation variance with respect to the weight vectors: $\frac{\partial[\mathrm{var}(\Delta z(\mathbf{x}_o))]}{\partial \bar{\theta}_{po}} = 0$. This provides a classical cokriging system of equations in matrix form as follows:

$$\mathbf{C}_{pp}\bar{\theta}_{po} = \mathbf{C}_{po} \tag{77}$$

In the sequential kriging approach, this system of equations is used only in the first step.

The derivative of the variance with respect to the weight vectors of the additional or new information $\frac{\partial[\mathrm{var}(\Delta z(\mathbf{x}_o))]}{\partial \bar{\varpi}_{so}} = 0$, in analogous form to the univariate case, gives the next conditional cokriging system

$$[\mathbf{C}_{ss} - \bar{\theta}_{ps}^T \mathbf{C}_{ps}]\bar{\varpi}_{so} = [\mathbf{C}_{so} - \bar{\theta}_{ps}^T \mathbf{C}_{po}] \tag{78}$$

This is the equation used for cokriging after a previous estimation. This cokriging of previous and new additional data is a two-step approach. Another way of writing conditional cokriging of residuals using the matrix of weights given in Eq. (74) and introducing matrices of residual conditional covariances is

$$\begin{bmatrix} \bar{\varepsilon}_{z_s z_s} & \bar{\varepsilon}_{z_s w_s} & \dots & \bar{\varepsilon}_{z_s q_s} \\ \bar{\varepsilon}_{w_s z_s} & \bar{\varepsilon}_{w_s w_s} & \dots & \bar{\varepsilon}_{w_s q_s} \\ & \vdots & & \\ \bar{\varepsilon}_{q_s z_s} & \bar{\varepsilon}_{q_s w_s} & \dots & \bar{\varepsilon}_{q_s q_s} \end{bmatrix} \begin{bmatrix} \vec{\alpha}_{z_s z_o} & \vec{\mu}_{z_s w_o} & \dots & \vec{\zeta}_{z_s q_o} \\ \vec{\beta}_{w_s z_o} & \vec{\nu}_{w_s w_o} & \dots & \vec{\eta}_{w_s q_o} \\ & \vdots & & \\ \vec{\gamma}_{q_s z_o} & \vec{\varphi}_{q_s w_o} & \dots & \vec{\tau}_{q_s q_o} \end{bmatrix}$$
$$= \begin{bmatrix} \bar{\varepsilon}_{z_s z_o} & \bar{\varepsilon}_{z_s w_o} & \dots & \bar{\varepsilon}_{z_s q_o} \\ \bar{\varepsilon}_{w_s z_o} & \bar{\varepsilon}_{w_s w_o} & \dots & \bar{\varepsilon}_{w_s q_o} \\ & \vdots & & \\ \bar{\varepsilon}_{q_s z_o} & \bar{\varepsilon}_{q_s w_o} & \dots & \bar{\varepsilon}_{q_s q_o} \end{bmatrix} \tag{79}$$

Notice the cokriging matrices are more complicated than in the kriging case. A numerical multivariate residual conditional covariance matrix can be defined from Eq. (78) as follows:

$$\bar{\varepsilon}_{ij} = \begin{bmatrix} \varepsilon_{z_i z_j} & \varepsilon_{z_i w_j} & \dots & \varepsilon_{z_i q_j} \\ \varepsilon_{w_i z_j} & \varepsilon_{w_i w_j} & \dots & \varepsilon_{w_i q_j} \\ & \vdots & & \\ \varepsilon_{q_i z_j} & \varepsilon_{q_i w_j} & \dots & \varepsilon_{q_i q_j} \end{bmatrix}$$
$$= \begin{bmatrix} c_{z_i z_j} & c_{z_i w_j} & \dots & c_{z_i q_j} \\ c_{w_i z_j} & c_{w_i w_j} & \dots & c_{w_i q_j} \\ & \vdots & & \\ c_{q_i z_j} & c_{q_i w_j} & \dots & c_{q_i q_j} \end{bmatrix} - \begin{bmatrix} \vec{\alpha}_{z_p z_i} & \vec{\mu}_{z_p w_i} & \dots & \vec{\zeta}_{z_p q_i} \\ \vec{\beta}_{w_p z_i} & \vec{\nu}_{w_p w_i} & \dots & \vec{\eta}_{w_p q_i} \\ & \vdots & & \\ \vec{\gamma}_{q_p z_i} & \vec{\varphi}_{z_p w_i} & \dots & \vec{\tau}_{q_p q_i} \end{bmatrix}^T$$

$$\times \begin{bmatrix} \vec{c}_{z_p z_j} & \vec{c}_{z_p w_j} & \dots & \vec{c}_{z_p q_j} \\ \vec{c}_{w_p z_j} & \vec{c}_{w_p w_j} & \dots & \vec{c}_{w_p q_j} \\ & \vdots & & \\ \vec{c}_{q_p z_j} & \vec{c}_{q_p w_j} & \dots & \vec{c}_{q_p q_j} \end{bmatrix} \tag{80}$$

The off-diagonal terms are residual conditional cross-covariances and the major diagonal terms are the already used residual conditional (auto)covariances. Equation (80) expresses the residual conditional covariance between two point locations and weights may be expressed in terms of the previous cokriging step. In the case of several covariated locations, the terms in the residual conditional covariance matrix will be matrices. If several locations are covariated to a single location, the terms on the matrix will be vectors.

Generalization of the sequential approach comes from considering any two steps r and $r+1$. In a form analogous to the univariate case, the conditional cross-covariances at step $r+1$ can be computed with the cross-covariance at step r. Extending Eq. (43) to the multivariate case, the updated residual conditional covariances and cross-covariances are

$$\bar{\varepsilon}_{ij}^{(r+1)} = \begin{bmatrix} \varepsilon^r_{z_i z_j} & \varepsilon^r_{z_i w_j} & \dots & \varepsilon^r_{z_i q_j} \\ \varepsilon^r_{w_i z_j} & \varepsilon^r_{w_i w_j} & \dots & \varepsilon^r_{w_i q_j} \\ & \vdots & & \\ \varepsilon^r_{q_i z_j} & \varepsilon^r_{q_i w_j} & \dots & \varepsilon^r_{q_i q_j} \end{bmatrix} - \begin{bmatrix} \vec{\alpha}_{z_s^r z_i} & \vec{\mu}_{z_s^r w_i} & \dots & \vec{\zeta}_{z_s^r q_i} \\ \vec{\beta}_{w_s^r z_i} & \vec{\nu}_{w_s^r w_i} & \dots & \vec{\eta}_{w_s^r q_i} \\ & \vdots & & \\ \vec{\gamma}_{q_s^r z_i} & \vec{\varphi}_{q_s^r w_i} & \dots & \vec{\tau}_{q_s^r q_i} \end{bmatrix}^T$$
$$\times \begin{bmatrix} \vec{\varepsilon}^r_{z_s^r z_j} & \vec{\varepsilon}^r_{z_s^r w_j} & \dots & \vec{\varepsilon}^r_{z_s^r q_j} \\ \vec{\varepsilon}^r_{w_s^r z_j} & \vec{\varepsilon}^r_{w_s^r w_j} & \dots & \vec{\varepsilon}^r_{w_s^r q_j} \\ & \vdots & & \\ \vec{\varepsilon}^r_{q_s^r z_j} & \vec{\varepsilon}^r_{q_s^r w_j} & \dots & \vec{\varepsilon}^r_{q_s^r q_j} \end{bmatrix} \tag{81}$$

Updated residual conditional cross-covariances will follow the same properties as demonstrated before for the unvariate case, and we can subsequently generalize the residual cokriging to a sequential approach algorithm in explicit form as follows:

$$\begin{aligned} & \mathbf{c}^1_{pp} \bar{\varpi}^1_{so} = c^1_{po} \\ & \vdots \\ & \bar{\varepsilon}^r_{s^r s^r} \bar{\varpi}^r_{s^r o} = \bar{\varepsilon}^r_{s^r o} \\ & \bar{\varepsilon}^{r+1}_{s^{r+1} s^{r+1}} \bar{\varpi}^{r+1}_{s^{r+1} o} = \bar{\varepsilon}^{r+1}_{s^{r+1} o} \\ & \vdots \\ & \bar{\varepsilon}^s_{ss} \bar{\varpi}^s_{so} = \bar{\varepsilon}^s_{so} \end{aligned} \tag{82}$$

These sequential cokriging equations are managed in the same way as classic cokriging. The case of available data for all attributes at each sampling location is called "isotopy", and implies that matrices will be full. The case of missing some attributes at some or all sampling locations is called "heterotopy", and implies the rows and columns corresponding to missing data must be erased from the cokriging equations.

Sequential cokriging is highly attractive because it allows one to deal with several attributes and a large number of data points. Sequential cokriging can allow one to initially cokrige using just the primary attribute, and in a second part cokrige the secondary attribute. Notice that in all cases, it is desirable to have the new data or second attribute data at different locations than the old or first attribute data to avoid zero cross-covariances.

3.6
Residual conditional cross-covariances

The introduction of residuals conditional cross-covariance computations by matrix operations in Eq. (80) is a consequence of the derivation of the cokriging equations in which old and additional data are utilized in a successive fashion. Moreover, we show that the residual conditional cross-covariance concept is a generalization of the known conditional (auto) covariance. The sample cross-covariance for two attributes, z and w is

$$c_{z_u w_v} = E[(z_u - \mu_z)(w_v - \mu_w)] \tag{83}$$

If we assume that the random functions are residuals with zero mean. Then residuals cross-covariance is defined as:

$$\varepsilon_{z_u w_v} = E[(\hat{z}_u - z_u)(\hat{w}_v - w_v)] \tag{84}$$

where z_u and w_v are conditional means of the random variables at locations u and v. The hat values are the respective estimates made with cokriging with old data. Since,

$$\hat{w}_v = \vec{\mu}_{iv}^T \vec{z}_i + \hat{\nu}_{jv}^T \vec{w}_j \tag{85}$$

$$\hat{z}_u = \vec{\alpha}_{iu}^T \vec{z}_i + \vec{\beta}_{ju}^T \vec{w}_j \tag{86}$$

Then,

$$\varepsilon_{z_u w_v} = E[((\vec{\alpha}_{iu}^T \vec{z}_i + \vec{\beta}_{ju}^T \vec{w}_j) - z_u)((\vec{\mu}_{iv}^T \vec{z}_i + \hat{\nu}_{jv}^T \vec{w}_j) - w_v)] \tag{87}$$

Making operations yields,

$$\begin{aligned}\varepsilon_{z_u w_v} = E[&\vec{\alpha}_{iu}^T (\vec{z}_i \vec{z}_i^T) \vec{\mu}_{iv} + \vec{\alpha}_{iu}^T (\vec{z}_i \vec{w}_i^T) \vec{\nu}_{jv} + \vec{\beta}_{ju}^T (\vec{w}_j \vec{z}_i^T) \vec{\mu}_{iv} + \vec{\beta}_{ju}^T (\vec{w}_j \vec{w}_j^T) \vec{\nu}_{jv} \\ &- \vec{\alpha}_{iu}^T (w_v \vec{z}_i^T) - \vec{\beta}_{ju}^T (w_v \vec{w}_j^T) - \vec{\mu}_{iv}^T (z_u \vec{z}_i^T) - \vec{\nu}_{jv}^T (z_u \vec{w}_v^T) + z_u w_v^T]\end{aligned} \tag{88}$$

Taking the expected value, this is

$$\begin{aligned}\varepsilon_{z_u w_v} = &\vec{\alpha}_{iu}^T \mathbf{c}_{\vec{z}_i \vec{z}_i} \vec{\mu}_{iv} + \vec{\alpha}_{iu}^T \mathbf{c}_{\vec{z}_i \vec{w}_i} \vec{\nu}_{jv} + \vec{\beta}_{jv}^T \mathbf{c}_{\vec{w}_j \vec{z}_i} \vec{\mu}_{iv} + \vec{\beta}_{jv}^T \mathbf{c}_{\vec{w}_j \vec{w}_j} \vec{\nu}_{jv} \\ &- \vec{\alpha}_{iu}^T \vec{c}_{w_v \vec{z}_i} - \vec{\beta}_{ju}^T \vec{c}_{w_v \vec{w}_j} - \vec{\mu}_{iv}^T \vec{c}_{z_u \vec{z}_i} - \vec{\nu}_{jv}^T \vec{c}_{z_u \vec{w}_j} + c_{z_u w_v}\end{aligned} \tag{89}$$

To get further simplifications we appeal to the following cokriging equations for two attributes which have been multiplied by vectors of weights. Those are

$$\begin{aligned}\vec{\alpha}_{iu}^{T}\mathbf{c}_{\vec{z}_i\vec{z}_i}\vec{\mu}_{iv}+\vec{\alpha}_{iu}^{T}\mathbf{c}_{\vec{z}_i\vec{w}_i}\vec{\nu}_{jv}&=\vec{\alpha}_{iu}^{T}\vec{c}_{\vec{z}_i w_v}\\ \vec{\beta}_{ju}^{T}\mathbf{c}_{\vec{w}_j\vec{z}_i}\vec{\mu}_{iv}+\vec{\beta}_{ju}^{T}\mathbf{c}_{\vec{w}_j\vec{w}_j}\vec{\nu}_{jv}&=\vec{\beta}_{ju}^{T}\vec{c}_{\vec{w}_j w_v}\end{aligned} \tag{90}$$

also

$$\begin{aligned}\vec{\alpha}_{iu}^{T}\mathbf{c}_{\vec{z}_i\vec{z}_i}\vec{\mu}_{iv}+\vec{\beta}_{ju}^{T}\mathbf{c}_{\vec{w}_j\vec{z}_i}\vec{\mu}_{iv}&=\vec{\mu}_{iv}^{T}\vec{c}_{\vec{z}_i z_u}\\ \vec{\alpha}_{iu}^{T}\mathbf{c}_{\vec{z}_i\vec{w}_j}\vec{\nu}_{iv}+\vec{\beta}_{ju}^{T}\mathbf{c}_{\vec{w}_j\vec{w}_j}\vec{\nu}_{jv}&=\vec{\nu}_{ju}^{T}\vec{c}_{\vec{w}_j z_u}\end{aligned} \tag{91}$$

Identities in Eqs. (90) and (91) are substituted into Eq. (89). When the first two identities are utilized, it follows that

$$\varepsilon_{z_u w_v}=c_{z_u w_v}-\vec{\mu}_{iv}^{T}\vec{c}_{z_u\vec{z}_i}-\vec{\nu}_{jv}^{T}\vec{c}_{z_u\vec{w}_j} \tag{92}$$

If the second two identifies are utilized, then

$$\varepsilon_{z_u w_v}=c_{z_u w_v}-\vec{\alpha}_{iu}^{T}\vec{c}_{w_v\vec{z}_i}-\vec{\beta}_{ju}^{T}\vec{c}_{w_v\vec{w}_j} \tag{93}$$

Notice that Eqs. (92) and (93) give the same numerical results. From this, it is evident that conditional cross-covariance matrix is symmetric when the previous covariance matrix is symmetric. We can also look at the matrix computation of this case by using Eq. (80). This leads to,

$$\vec{\varepsilon}_{uv}=\begin{bmatrix}c_{z_u z_u} & c_{z_u w_v}\\ c_{w_v z_u} & c_{w_v w_v}\end{bmatrix}-\begin{bmatrix}\vec{\alpha}_{z_i z_u} & \vec{\mu}_{z_i w_v}\\ \vec{\beta}_{w_j z_u} & \vec{\nu}_{w_j w_v}\end{bmatrix}^{T}\begin{bmatrix}\vec{c}_{z_i z_u} & \vec{c}_{z_i w_v}\\ \vec{c}_{w_j z_u} & \vec{c}_{w_j w_v}\end{bmatrix} \tag{94}$$

The residual conditional cross-covariances are easier to compute using matrix operations. At zero lag distances, conditional covariances and cross-covariances are the kriging or cokriging variances.

4 Discussion

From the practical point of view, we have provided an alternative for solving kriging and cokriging with large data sets. Most of the currently available software that kriges data sets applies the local neighborhood algorithm which may not provide unique results. The sequential approach, provided here, is an algorithm which allows for unique results and requires less effort in terms of inversion of matrices. Consider, for example, 100 data locations and 1000 grid points to krige. Using the local neighborhood approach, one must invert a covariance matrix at each of the 1000 grid locations, since the local neighborhood is a moving window and the available data array changes from one point to another. The second part of the interpolation is a simple matrix multiplication. In the sequential approach proposed here, the slowest case may be that one takes just two data locations and invert a two by two matrix of covariances. The matrix product is needed to estimate the 1000 locations. In a subsequent step, two additional data points are used to compute residual conditional covariances by using a matrix product. Then, a two by two matrix of residual conditional covariances is inverted. At this

step, the result will yield 1000 estimated residuals. In this particular example, the sequential approach requires only 50 inversion of small matrices, while the local neighborhood requires 1000 inversions of larger matrices. However, the sequential approach may need much more matrix products. The major advantage of the sequential approach is that it gives a unique numerical solution. On the other hand, local neighborhood kriging exhibits spurious behavior, and results may depend of the size and shape of the neighborhood.

The sequential kriging and cokriging approaches are the result of a successive minimization of the estimation variance. The equivalence of results of these new approaches and the classical counterparts validates our sequential algorithm which is a desirable solution for interpolation with large data sets and/or when all data are not available at once. The covariance matrices to be managed may be reasonably small. However, the covariances between old and new data at each step are still required for computation of the residual conditional covariance while conditional cross-covariances are computed analogously to univariate residual conditional covariances. Another important finding is that updated residual conditional covariances can be easily computed from previous residual conditional covariances. The use of matrix notation greatly facilitates the required operations, and the smoothing effect of kriging is proved to be reduced sequentially. In the cokriging case, a cokriging estimation covariance matrix measures the smoothing effect. Sequential kriging for weaker stationary random functions may be derived using the same procedure. Moreover, sequential cokriging is expected to become useful for solving the problems in hydrology and interpolations in the space and time domains.

References

Davis MW (1975) The practice of kriging. In Advanced Geostatistics in the Mining Industry, M.G. Guarascio, M. Davis, and C. Huijbregts (eds). D. Reidel Publishing Company

Davis MW, Culbane PG (1984) Contouring very large data sets using kriging. In Geostatistics for Natural Resources Characterization, Part 2, Verly G. et al., (eds) NATO ASI series

Davis MW, Grivet C (1984) Kriging in a global neighborhood. Math. Geology 16(3): 249–265

Dietrich CR, Newsam GN (1989) A stability analysis of the geostatistical approach to aquifer transmissivity identification. Stochasctic Hydrol. Hydraul., 3: 293–316

Gutjahr A, Bullard B, Hatch S, Hughson L (1994) Joint conditional simulations and the spectral approach to flow modeling. Stochastic Hydrol. Hydraul., 8(1): 79–108

Harter Th, Yeh T-C J (1996) Conditional stochastic analysis of solute transport in heterogeneous, variably saturated soils. Water Resour. Res., 32(6): 1597–1610

Harvey CF, Gorelick S (1995) Mapping hydraulic conductivity: Sequential conditioning with measurements of solute arrival time, hydraulic head, and local conductivity. Water Resour. Res., 31: 1615–1626

Hoeksema RJ, Kitanidis PK (1984) An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling. Water Resour. Res., 20(7): 1003–1020

Journal AG, Huijbregts ChJ (1978) Mining Geostatistics. Academic Press Inc., New York, pp 600

Kitanidis PK (1995) Quasi-linear geostatistical theory for inversing. Water Resour. Res. 31(10): 2411–2419

Kitanidis PK, Vomboris EG (1983) A geostatistical approach to the inverse problem in ground water modeling (steady state) an one dimensional simulations. Water Resour. Res. 19(3): 677–690

Li B (1998) Unconditional and conditional analysis of flow and solute transport in variably saturated porous media, Ph.D. Dissertation, University of Arizona, Tucson

Myers DE (1982) Matrix formulation of cokriging, Math. Geology 14(3): 249–257

Sun N-Z, Yeh WW-G (1992) A stochastic inverse solution for transient groundwater flow: Parameter identification and reliability analysis. Water Resour. Res., 28(12): 3269–3280

Tong AYS (1996) Stochastic parameter estimation, reliability analysis, and experimental design in ground water modeling Ph.D. dissertation, University of California at Los Angeles
Yates SR, Warrick AW (1987) Estimating soil water content using cokriging Soil. Sci. Soc. Am. J. 51: 23–30
Yeh T-CJ, Gutjahr A, Jin M (1995) An iterative cokriging like technique for ground water flow modeling. Groundwater 33(1): 33–41
Yeh T-CJ, Jin M, Hanna S (1996) An iterative stochastic inverse method: Conditional effective transmissivity and hydraulic head fields. Water Resour. Res. 32: 85–92
Yeh T-CJ, Zhang J (1996) A geostatistical inverse method for variably saturated flow in the vadose zone. Water Resour. Res. 32(9): 2757–2766
Zhang J, Yeh T-CJ (1997) An iterative geostatistical inverse method for steady flow in the vadose zone. Water Resour. Res. 33(1): 63–71