



# The Many Forms of Co-kriging: A Diversity of Multivariate Spatial Estimators

Peter A. Dowd<sup>1</sup> · Eulogio Pardo-Igúzquiza<sup>2</sup>

Received: 21 June 2023 / Accepted: 4 September 2023 / Published online: 12 October 2023  
© The Author(s) 2023

## Abstract

In this expository review paper, we show that co-kriging, a widely used geostatistical multivariate optimal linear estimator, has a diverse range of extensions that we have collected and illustrated to show the potential of this spatial interpolator. In the context of spatial stochastic processes, this paper covers scenarios including increasing the spatial resolution of a spatial variable (downscaling), solving inverse problems, estimating directional derivatives, and spatial interpolation taking boundary conditions into account. All these spatial interpolators are optimal linear estimators in the sense of being unbiased and minimising the variance of the estimation error.

**Keywords** Co-kriging · Upscaling · Downscaling · Spatial interpolators · Inverse problems · Estimating directional derivatives · Spatial boundary conditions

## 1 Introduction

Geostatistical spatial estimators are standard techniques used in spatial analyses in the geosciences and geographical information systems. They are applied in remote sensing (Van der Meer 2012), meteorology (Teegavarapu et al. 2015), machine learning (De Iaco et al. 2022), medical geology (Zhang et al. 2021) and epidemiology (Graham et al. 2004), among others. In this context, co-kriging (Matheron 1963) is a multivariate geostatistical interpolation method that is used to generate maps of a primary variable by using experimental data of that variable and experimental data from auxiliary variables correlated with the former. Without loss of generality, and for simplicity, we

---

✉ Peter A. Dowd  
peter.dowd@adelaide.edu.au  
Eulogio Pardo-Igúzquiza  
e.pardo.iguzquiza@csic.es

<sup>1</sup> Faculty of Sciences, Engineering and Technology, The University of Adelaide, Adelaide, Australia

<sup>2</sup> Instituto de Geociencias (IGEO, CSIC), Calle Severo Ochoa, 7, 28040 Madrid, Spain

consider the case of only one auxiliary variable, which is also the most frequent case encountered in practice.

In geostatistics, an observed value at a spatial location  $\mathbf{u}_i$  is modelled as a realisation of a random variable  $Z(\mathbf{u}_i)$ . For example,  $Z$  may represent a spatial variable, such as temperature, and  $\mathbf{u}_i = \{x_i, y_i\}$  is a spatial location on the plane; thus  $Z(\mathbf{u}_i)$  models the temperature at that location. The set of all random variables  $Z(\mathbf{u}_i)$  in a region  $\chi$  of the space  $\mathbf{u}_i \in \chi$  is a random function, or a random field,  $Z(\mathbf{u})$ . For  $\chi \subset \mathfrak{R}^d$  and  $d = 2$ , the problem is two-dimensional, which is the case in the work presented here.

It is assumed that the random function  $Z(\mathbf{u})$  is second-order stationary, that is, with constant spatial mean and covariance function  $C_Z(\mathbf{u}, \mathbf{u} + \mathbf{h})$  that depends only on the vector  $\mathbf{h}$

$$E\{Z(\mathbf{u})\} = m_Z, \quad (1)$$

$$C_Z(\mathbf{h}) = E\{Z(\mathbf{u})Z(\mathbf{u} + \mathbf{h})\} - m_Z^2, \quad (2)$$

$$C_Z(0) = \sigma_Z^2, \quad (3)$$

where  $m_Z$ ,  $\sigma_Z^2$  and  $C_Z(\mathbf{h})$  are the mean, variance and covariance, respectively, of the random function  $Z(\mathbf{u})$  and  $E\{\cdot\}$  is the mathematical expectation operator. Although the variogram is most often used in geostatistics, only covariances are used in the work presented here.

Co-kriging is a linear estimator that can be written as (e.g., Journel and Huijbregts 1978; Ver Hoef and Cressie 1993)

$$Z^*(\mathbf{u}_0) = \sum_{i=1}^n \lambda_i^0 Z(\mathbf{u}_i) + \sum_{j=1}^m \beta_j^0 Y(\mathbf{u}_j), \quad (4)$$

where  $Z^*(\mathbf{u}_0)$  is the estimated value of the primary variable at the spatial location  $\mathbf{u}_0 = \{x_0, y_0\}$ , that is,  $Z(\mathbf{u}_0)$ , where  $\mathbf{u}_0$  is a point on the plane with coordinates  $x_0$  as easting and  $y_0$  as northing.  $\{Z(\mathbf{u}_i); i = 1, \dots, n\}$  is the set of  $n$  experimental data of the primary variable in Eq. (4), and  $\{Y(\mathbf{u}_j); j = 1, \dots, m\}$  is the set of  $m$  experimental data of the auxiliary variable used in Eq. (4).  $\lambda_i^0$  is the weight applied to the primary variable  $Z(\mathbf{u}_i)$  in the estimation of  $Z(\mathbf{u}_0)$ , and  $\beta_j^0$  is the weight applied to the auxiliary variable  $Z(\mathbf{u}_j)$  in the estimation of  $Z(\mathbf{u}_0)$ .

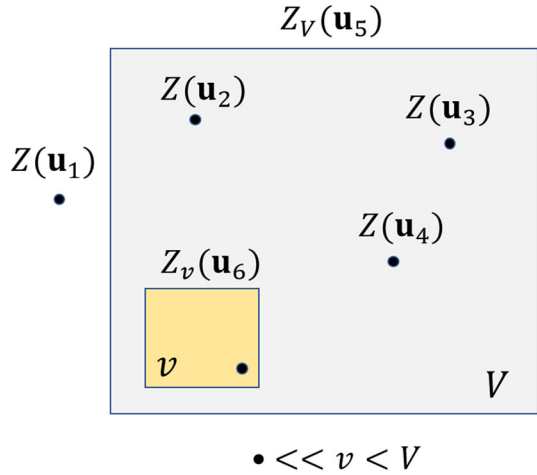
The set of optimal weights  $\{\lambda_i^0; i = 1, \dots, n\}$  and  $\{\beta_j^0; i = 1, \dots, n\}$  are obtained by minimising the variance of the estimation error

$$\text{Var}\{Z^*(\mathbf{u}_0) - Z(\mathbf{u}_0)\}, \quad (5)$$

subject to the unbiasedness condition

$$E\{Z^*(\mathbf{u}_0) - Z(\mathbf{u}_0)\} = 0. \quad (6)$$

**Fig. 1** Spatial support of the random variable. In two-dimensional problems, the spatial support is the area to which the experimental information is assigned, that is, the pixel size or the area covered by the values

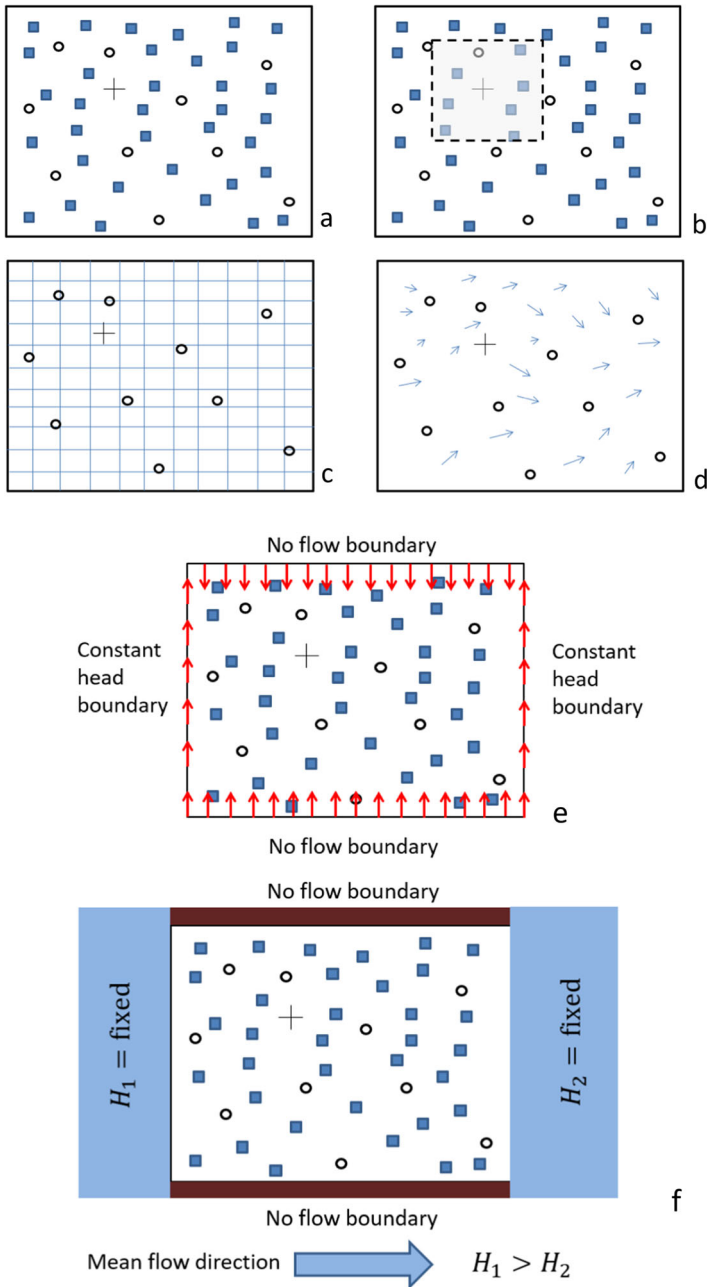


The position of  $\mathbf{u}_0$  can vary in order to define a grid (raster image), a polygon, lineation, and so on.

An important aspect to be considered in co-kriging is the spatial support of a random variable as shown in Fig. 1, which shows six random variables with three different types of support: point support,  $\bullet()$ , and two different sizes of areal support ( $v$  and  $V$ ) with  $\bullet \ll v < V$ . For example,  $Z_V(\mathbf{u}_5)$  represents the mean value of the random variable over the surface  $V$ .

In two dimensions, the spatial support of a random variable is the area over which the random variable is measured. In remote sensing, for example, the spatial support may be the pixel size or the spatial resolution of a satellite image. A rain gauge, for example, measures rainfall on a spatial support that is very small with respect to the area over which the rainfall is to be estimated, and is sometimes referred to as a point support. If, for example, rainfall is measured by the variation in the height of water in a swimming pool, the support would be the area of the swimming pool. If the mean rainfall over a river basin is known, the spatial support is the area of the river basin. When using data for the same variable measured on different spatial supports, the support effect must be taken into account.

Figure 2 provides two-dimensional illustrative representations of the various forms of co-kriging. Figure 2a is the classical application of (ordinary) co-kriging to estimate the primary variable at an unsampled spatial location by using the available primary variable data (shown as open circles) and the available auxiliary data (shown as solid squares). The cross represents a location at which the primary variable has not been measured and is to be estimated. The black cross can be located arbitrarily, and thus the primary variable can be estimated at a grid of locations to generate contour maps or colour-coded maps. Figure 2b provides an example of block ordinary co-kriging or upscaling co-kriging where the mean value of the grey square, or any other arbitrary polygon, is to be estimated. The areal average represents the mean value of the primary variable inside the polygon. The black cross at the centre of the square is a reference point and not necessarily a location to be estimated. The spatial support of



**Fig. 2** **a** The general ordinary co-kriging case. **b** Block ordinary co-kriging or upscaling co-kriging. **c** The case in which the auxiliary variable has been measured at every location (blue square) of a grid or raster image. **d** The objective here is to estimate a directional derivative of the primary variable. **e** Co-kriging when the primary variable is to be estimated using the boundary conditions of the problem expressed as directional derivatives. **f** An example of inverse problem co-kriging. **g** Downscaling co-kriging

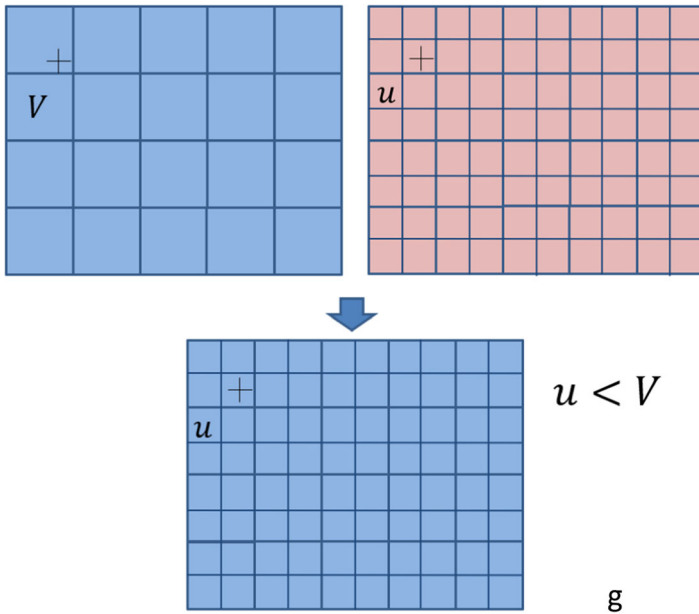
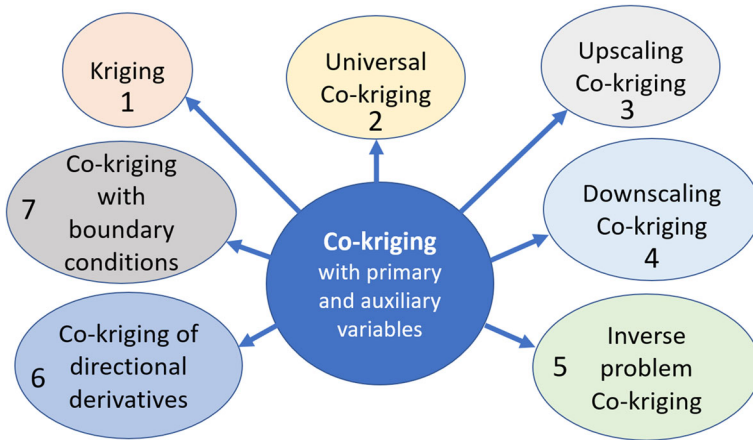


Fig. 2 continued

the estimation is larger than the support of the experimental data and is an example of upscaling the primary variable. In Fig. 2c, the auxiliary variable has been measured at every location (blue square) of a grid or raster image, such as a satellite image or a digital elevation model. Often the primary variable is to be estimated on the same grid. An example of this is estimating rainfall from rainfall data measured in rain gauges together with altitude data as a secondary variable obtained from a digital elevation model (a raster image). In Fig. 2d, the objective is to estimate a directional derivative of the primary variable (open circles) by using the secondary data provided by the directional derivatives of the variable (arrows). In this way, the gradient of the primary variable can be estimated, and the auxiliary data could also be the known gradient values. In the co-kriging problem in Fig. 2e, the primary variable (a scalar) is to be estimated using the boundary conditions of the problem (no flow and constant value) expressed as directional derivatives. The derivative perpendicular to the boundary for the no-flow case and the derivative parallel to the boundary for the constant value case are null. Figure 2f is an example of inverse problem co-kriging. Although a groundwater flow model has been used here, the procedure is completely generalisable to other partial differential equation problems. In inverse problem co-kriging, the parameters of the model (primary variable) are estimated from the experimental values of the parameters and the model output values (auxiliary variables). In the groundwater flow problem, these are log-transmissivity and water head for the primary and auxiliary variables, respectively. In Fig. 2g, downscaling co-kriging can be used in a setting such as that shown in Fig. 2a, but it has a clearer application in increasing the spatial resolution of satellite images where the image is to be estimated at the same resolution



**Fig. 3** The many forms of co-kriging

as that of the secondary information. The spatial resolution to be estimated could even be a point support.

Extensions of co-kriging considered in the work presented here are as follows:

- Co-kriging to estimate the directional derivative of the primary variable at specified locations by using the available directional derivative data and the gradient of the primary variable (Fig. 2d).
- Boundary co-kriging, in which the primary variable is estimated using primary variable data together with secondary variable boundary conditions in the form of directional derivatives (Fig. 2e).
- Inverse problem co-kriging, in which the primary and the auxiliary variables are physically linked by a differential equation with given boundary conditions. Inverse problem co-kriging is used to solve an inverse problem (Fig. 2f).
- Downscaling co-kriging or estimating the primary variable on a support smaller than the support of the experimental data of the primary variable (Fig. 2g).

The different aspects of co-kriging are summarised in Fig. 3 and Table 1. In this paper, without any loss of generality, we consider the interpolation of a primary variable using primary variable data and auxiliary variable data, that is, ordinary co-kriging with two variables. If there are no available auxiliary variable data, co-kriging reduces to kriging (1). Ordinary co-kriging assumes that the spatial means of the variables are constant. If there is a spatially variable mean (a trend) in the primary or secondary variables, the universal co-kriging estimator is used (2). If the support (or pixel size) of the estimation is larger than that of the experimental support, then up-scaling co-kriging is applied (3). When the support to be estimated is smaller than the experimental support of the primary variable, then downscaling co-kriging is the interpolator (4). If the primary and auxiliary variables are related by a physical model described by a partial differential equation, then the interpolator is inverse problem co-kriging (5). Co-kriging can be adapted to estimate the directional derivatives of a scalar variable by using the directional derivatives as auxiliary data (6). The boundary conditions of no

**Table 1** The many forms of co-kriging shown as a diversity of interpolation problems

Interpolator	Main particularities
Ordinary co-kriging	The primary variable is to be estimated on the same spatial support as the experimental data
Ordinary Kriging	There is no auxiliary variable
Universal co-kriging	There is a trend, or spatially variable mean, in the primary and/or auxiliary variable
Upscaling co-kriging	The spatial support of the estimate is larger than the spatial support of the primary variable
Downscaling co-kriging	The spatial support of the estimate is smaller than the spatial support of the primary variable
Inverse problem co-kriging	The primary and auxiliary variables are physically linked by a differential equation
Co-kriging of directional derivatives	The variable to be estimated is the directional derivative of the primary variable
Co-kriging with boundary conditions	The auxiliary variable is given in the form of directional derivatives from the boundary conditions of no-flow and fixed flow

flow and fixed value can be incorporated into co-kriging with boundary conditions. The following sections provide a brief review of ordinary co-kriging before dealing with upscaling co-kriging, downscaling co-kriging, inverse problem co-kriging, co-kriging of directional derivatives and co-kriging with boundary conditions.

## 2 Ordinary Co-kriging

In the simplest case, a spatial variable of interest  $Z(\mathbf{u})$ , or primary variable, is to be estimated at a location  $\mathbf{u}_0$  at which the variable was not sampled. The variable is to be estimated by using the experimental values of the primary variable  $\{Z(\mathbf{u}_i); i = 1, \dots, n\}$  and the experimental values of an auxiliary variable  $\{Y(\mathbf{u}_j); j = 1, \dots, m\}$ . The ordinary co-kriging estimator is given in Eq. (4), and the optimal weights are obtained by minimising the variance of the estimation error given in Eq. (5) subject to the unbiasedness condition in Eq. (6). The unbiasedness condition of the co-kriging in Eq. (6) implies that the following conditions must be satisfied

$$\sum_{i=1}^n \lambda_i^0 = 1, \tag{7}$$

$$\sum_{j=1}^m \beta_j^0 = 0. \tag{8}$$

Isaaks and Srivastava (1989) showed that a single unbiased condition in co-kriging

$$\sum_{i=1}^n \lambda_i^0 + \sum_{j=1}^m \beta_j^0 = 1, \quad (9)$$

may be preferable to using the two individual unbiased conditions in (7) and (8). In particular, the condition in Eq. (8) that the sum of the weights assigned to the secondary variable must be zero implies that some of the weights must be negative, which can produce problematic estimates. This problem is avoided by using the unique unbiasedness condition given in Eq. (9).

The ordinary co-kriging system is given in “Appendix A”. The constant mean condition in Eq. (1) can be relaxed to consider a spatially variable mean. In the latter case, the interpolator is given by universal co-kriging, which is reviewed in “Appendix B”.

The variance of the estimation error, or estimation variance, can be written as (Myers 1982, 1983, 1991; Isaaks and Srivastava 1989; Wackernagel 2003)

$$\begin{aligned} \text{Var}\{Z^*(\mathbf{u}_0) - Z(\mathbf{u}_0)\} &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i^0 \lambda_j^0 C_Z(\mathbf{h}_{ij}) + \sum_{i=1}^m \sum_{j=1}^m \beta_i^0 \beta_j^0 C_Y(\mathbf{h}_{ij}) \\ &+ \sum_{i=1}^n \sum_{j=1}^m \lambda_i^0 \beta_j^0 C_{ZY}(\mathbf{h}_{ij}) + \sum_{j=1}^m \sum_{i=1}^n \beta_j^0 \lambda_i^0 C_{YZ}(\mathbf{h}_{ji}) \\ &- 2 \sum_{i=1}^n \lambda_i^0 C_Z(\mathbf{h}_{i0}) - 2 \sum_{j=1}^m \beta_j^0 C_Y(\mathbf{h}_{j0}) + C_Z(\mathbf{h}_{00}), \end{aligned} \quad (10)$$

where  $C_Z(\mathbf{h}_{ij})$  is the covariance between the random variables  $Z(\mathbf{u}_i)$  and  $Z(\mathbf{u}_j)$  for which the spatial distance is equal to the vector  $\mathbf{h}_{ij} = \mathbf{u}_j - \mathbf{u}_i$ . In a similar way,  $C_{ZY}(\mathbf{h}_{ij})$  is defined as the cross-covariance between the random variables  $Z(\mathbf{u}_i)$  and  $Y(\mathbf{u}_j)$

$$C_{ZY}(\mathbf{h}_{ij}) = E\{Z(\mathbf{u}_i)Y(\mathbf{u}_j)\} - m_Z m_Y. \quad (11)$$

Similarly,  $C_{YZ}(\mathbf{h}_{ji})$  is the cross-covariance between the random variables  $Y(\mathbf{u}_j)$  and  $Z(\mathbf{u}_i)$ .

If the number of auxiliary data,  $m$ , in Eq. (4) is zero, then ordinary co-kriging reduces to ordinary kriging. The main additional requirement is the inclusion of the auxiliary information in the estimator in Eq. (4). This is a type of statistical inference because, in addition to the inference of the covariance of the primary variable, it is necessary to infer the covariance of the auxiliary variable and the cross-covariance between the two variables, although in special cases some simplifications could be assumed (Journal 1999; Babak and Deutsch 2009). The co-kriging system is given in “Appendix A” and is known as the ordinary co-kriging system (Journal and Huijbregts 1978; Isaaks and Srivastava 1989; Deutsch and Journal 1992; Goovaerts 1997; Chilès and Delfiner 1999;



Remy et al. 2009) and is the most often used in practice, as in mining (e.g., Journel and Huijbregts 1978), hydrogeology (e.g., Hoeksema et al. 1989), remote sensing (e.g., Atkinson et al. 1994), soil science (e.g., Lesch et al. 1995), geophysics (e.g., Doyen 1988), meteorology and climatology (e.g., Pardo-Igúzquiza 1998), among many other earth science disciplines.

### 3 Upscaling Co-kriging

A random variable  $Z(\mathbf{u}_0)$  with point support is a random variable at the spatial location  $\mathbf{u}_0 = \{x_0, y_0\}$ , that is, a point in the plane. However,  $Z(\mathbf{u}_0)$  could represent the mean value of an area (areal support), for example, a square centred at the spatial location  $\mathbf{u}_0 = \{x_0, y_0\}$ . In this case, the value of the random variable at any location is an areal average and is denoted  $Z_V(\mathbf{u}_0)$ , referring to the mean value of the random function  $Z(\mathbf{u})$  over the polygon  $V$

$$Z_V(\mathbf{u}_0) = \int_V Z(x)dx. \tag{12}$$

In the notation  $Z_V(\mathbf{u}_0)$ ,  $\mathbf{u}_0$  is an arbitrary point denoting the polygon  $V$ , for example, the centroid.

The ordinary co-kriging estimator  $Z_V^*(\mathbf{u}_0)$  is similar to the estimator in Eq. (4), and the weights are obtained by solving the block co-kriging system, similar to that given in “Appendix A”, but with a new vector  $\mathbf{B}$

$$\mathbf{B} = \begin{bmatrix} C_Z(\mathbf{h}_{1V}) \\ \vdots \\ C_Z(\mathbf{h}_{nV}) \\ C_{ZY}(\mathbf{h}_{V1}) \\ \vdots \\ C_{ZY}(\mathbf{h}_{Vm}) \\ 1 \\ 0 \end{bmatrix}, \tag{13}$$

where  $C_Z(\mathbf{h}_{iV})$  is the mean covariance between the  $i$ th experimental location and the polygon  $V(\mathbf{u}_0)$

$$C_Z(\mathbf{h}_{iV}) \approx \frac{1}{M} \sum_{j=1}^M C_Z(\mathbf{h}_{ij}), \tag{14}$$

in which the polygon  $V$  has been approximated by  $M$  points, for example, points on a regular grid inside the polygon. In the same manner, the cross-covariance  $C_{ZY}(\mathbf{h}_{Vj})$

can be defined between the primary variable with support  $V$  and the secondary variable with point support.

Note that the support of the estimate is equal to, or greater than, the support of the experimental data. Equation (12) can be made more general by introducing a function that gives a different weight to every point that defines the support  $V$ , for example, in remote sensing, in which a satellite sensor has a point spread function that, in general, gives more weight to the centre of a pixel than to the borders.

### 4 Downscaling Co-kriging

In downscaling co-kriging, the estimation support of the primary variable is smaller than its experimental support. A typical application is in remote sensing where the spatial resolution of a satellite image, for a particular spectral band, is to be increased. The ordinary co-kriging estimator in Eq. (4) can be rewritten to take explicit account of the support of each random function, as

$$Z_v^*(\mathbf{u}_0) = \sum_{i=1}^n \lambda_i^0 Z_V(\mathbf{u}_i) + \sum_{j=1}^m \beta_j^0 Y_v(\mathbf{u}_j), \tag{15}$$

where  $v$  is the high spatial resolution support (or pixel size), and  $V$  is the low spatial resolution pixel size, expressed as

$$\bullet \ll v < V, \tag{16}$$

where  $\bullet$  represents the point support.

The resolution of the downscaling co-kriging system provides the weights of the estimator given in Eq. (15) and is similar to the co-kriging system given in “Appendix A”, but takes into account the different supports as, for example, the following matrix **B**

$$\mathbf{B} = \begin{bmatrix} C_Z^{Vv}(\mathbf{h}_{10}) \\ \vdots \\ C_Z^{Vv}(\mathbf{h}_{n0}) \\ C_{ZY}^{vv}(\mathbf{h}_{01}) \\ \vdots \\ C_{ZY}^{vv}(\mathbf{h}_{0m}) \\ 1 \\ 0 \end{bmatrix}. \tag{17}$$

$C_Z^{Vv}(\mathbf{h}_{10})$  is the covariance between the random variables  $Z_V(\mathbf{u}_1)$  and  $Z_v(\mathbf{u}_0)$ , and  $C_{ZY}^{vv}(\mathbf{h}_{01})$  is the covariance between  $Z_v(\mathbf{u}_0)$  and  $Y_v(\mathbf{u}_1)$ .

There are no experimental data for the random variable  $Z_v(\mathbf{u})$  because it has not been observed at that high spatial resolution. Thus, the matrix in Eq. (17) cannot be

experimentally estimated. The solution proposed in Pardo-Igúzquiza et al. (2006) and Atkinson et al. (2008) is a numerical one that implies convolutions and deconvolutions. The method consists of proposing covariance models with point support  $C_Z^{\bullet\bullet}(\mathbf{h})$  that are introduced in the equation (Matheron 1963)

$$\tilde{C}_Z^{VV}(\mathbf{h}) = C_Z^{\bullet\bullet}(\mathbf{h}) * \rho_V(\mathbf{h}), \tag{18}$$

which produces the induced covariance model  $\tilde{C}_Z^{VV}(\mathbf{h})$  that can be compared with the experimental covariance  $\hat{C}_Z^{VV}(\mathbf{h})$ . In Eq. (18),  $\rho_V(\mathbf{h})$  is the geometric covariogram in Matheron (1975), and  $*$  is the convolution operator. The iterative process consists in finding the covariance model with point support that minimises the difference between the covariance with support  $V$  induced by Eq. (18) and the experimental covariance with support  $V$ . Once the covariance model with point support has been estimated, the covariance over any other support can be calculated using Eq. (18). In particular, all the covariances required in the downscaling co-kriging system can be calculated. A detailed account of downscaling co-kriging is given in Pardo-Igúzquiza et al. (2006) and Pardo-Igúzquiza and Atkinson (2007), and a computer program for downscaling co-kriging is given in Pardo-Igúzquiza et al. (2010). In remote sensing, downscaling co-kriging can be used for image sharpening and image fusion.

### 5 Co-kriging as a Solution to the Inverse Problem

Another application of co-kriging is related to the geostatistical solution of the inverse problem. For a physical system, the direct problem is to predict the system given the values of the parameters that characterise the system. The inverse problem uses experimental measurements to infer the values of the parameters that characterise the system. For example, solving the inverse problem in groundwater hydrology consists of gathering transmissivity and water head data and taking into account that both variables are related by the groundwater flow equations (Hoeksema and Kitanidis 1984; Dagan 1985; Ahmed and De Marsily 1993). Transmissivity is the primary variable that defines the parameters of the system. Water head, the output of the direct problem, is the auxiliary variable. The co-kriging solution of the inverse problem consists of using co-kriging with a cross-covariance between transmissivity and water head obtained from theoretical considerations by taking into account the groundwater flow equations and the aquifer boundary conditions; in particular, taking into account the steady-state groundwater flow equations (Kitanidis 1997)

$$\frac{\partial Z(\mathbf{u})}{\partial x} \frac{\partial Y(\mathbf{u})}{\partial x} + \frac{\partial Z(\mathbf{u})}{\partial y} \frac{\partial Y(\mathbf{u})}{\partial y} + \frac{\partial^2 Y(\mathbf{u})}{\partial x^2} + \frac{\partial^2 Y(\mathbf{u})}{\partial y^2} = 0, \tag{19}$$

where  $Z(\mathbf{u})$  is the logarithm of transmissivity and  $Y(\mathbf{u})$  is the water head. The cross-covariance between the primary and auxiliary variables in Eq. (19), together with any boundary conditions, can be found analytically or by a Monte Carlo methodology (Kitanidis 1997). Dagan’s (1985) analytical solution for Eq. (19) is given in “Appendix C”.

## 6 Co-kriging for Estimating Directional Derivatives

For some problems in the geosciences and in other disciplines, the interest is in calculating the directional derivative of a given scalar variable. For example, the hydraulic gradient is a vector variable, the two components of which are the directional derivatives of the water head in the directions of the principal axes. Co-kriging can be adapted for this type of estimation. Furthermore, by using linear systems theory, the covariance of the directional derivative and the cross-covariance between the directional derivative and the scalar variable can be theoretically obtained from the covariance of the scalar variable. Other earth science applications can be found in terrain analysis, remote sensing, geophysics and meteorology, among others.

The directional derivative of the random function  $Z(\mathbf{u})$  can be estimated by extending the trend model in equation (B2) to distinguish two components in the residual  $R(\mathbf{u})$ , namely a correlated stochastic component  $\tilde{R}(\mathbf{u})$ , and a non-correlated stochastic component  $N(\mathbf{u})$ . The geostatistical model is then

$$Z(\mathbf{u}) = m(\mathbf{u}) + \tilde{R}(\mathbf{u}) + N(\mathbf{u}). \quad (20)$$

There is no correlation between the three components in Eq. (20). The non-correlated stochastic component  $N(\mathbf{u})$  is a zero-mean random function with no spatial correlation and is discontinuous at every point, is non-differentiable, and represents the variability associated with a second-order stationary nugget covariance  $C_N(\mathbf{h})$

$$C_N(\mathbf{h}) = \begin{cases} C_0 & |\mathbf{h}| = 0 \\ 0 & |\mathbf{h}| > 0 \end{cases}. \quad (21)$$

In applications, there will be a component  $N(\mathbf{u})$  if there is a nugget variance in the experimental covariance.

The correlated stochastic component  $\tilde{R}(\mathbf{u})$  is a zero-mean random function with spatial covariance  $\tilde{C}(\mathbf{h})$ . Thus, the spatial correlation  $C(\mathbf{h})$  of the random function  $Z(\mathbf{u})$  is given by

$$C(\mathbf{h}) = \begin{cases} C_0 + \tilde{C}(0) & |\mathbf{h}| = 0 \\ \tilde{C}(\mathbf{h}) & |\mathbf{h}| > 0 \end{cases}. \quad (22)$$

A random function,  $\tilde{Z}(\mathbf{u})$ , that is continuous and differentiable can be obtained from  $Z(\mathbf{u})$  by filtering the nugget component,

$$\tilde{Z}(\mathbf{u}) = m(\mathbf{u}) + \tilde{R}(\mathbf{u}), \quad (23)$$

the covariance of which is  $\tilde{C}(\mathbf{h})$ . The first requirement is for  $\tilde{Z}(\mathbf{u})$  to be differentiable for which the trend,  $m(\mathbf{u})$ , must be differentiable (Parzen 1972), which is the case for the polynomial drift considered here. A second requirement is that the covariance,  $\tilde{C}(\mathbf{h})$ , must be at least twice differentiable. In particular, the Gaussian covariance and the Matérn covariance (Pardo-Igúzquiza et al. 2009) are appropriate.

We use  $D_{\mathbf{e}}\{\tilde{Z}(\mathbf{u})\}$  to denote the directional derivative of  $\tilde{Z}(\mathbf{u})$  in the direction described by the unitary vector  $\mathbf{e} = \cos(\varphi)\mathbf{i} + \sin(\varphi)\mathbf{j}$ , where  $\varphi$  is the counterclockwise angle between the unitary vector  $\mathbf{e}$  and the  $X$  axis, and  $\mathbf{i}$  and  $\mathbf{j}$  are the unitary vectors of the coordinate axes  $X$  and  $Y$ . The directional derivative is defined as

$$D_{\mathbf{e}}\{\tilde{Z}(\mathbf{u})\} = \lim_{h \rightarrow 0} \left[ \frac{\tilde{Z}(\mathbf{u} + h\mathbf{e}) - \tilde{Z}(\mathbf{u})}{h} \right]. \tag{24}$$

The gradient of a scalar random function  $\tilde{Z}(\mathbf{u})$ , denoted by  $\nabla\tilde{Z}(\mathbf{u})$ , is a vector field defined at any point  $\mathbf{u}$  by its components

$$\nabla\tilde{Z}(\mathbf{u}) = \left( D_{\mathbf{i}}\{\tilde{Z}(\mathbf{u})\}, D_{\mathbf{j}}\{\tilde{Z}(\mathbf{u})\} \right). \tag{25}$$

The relationship between the directional derivative and the gradient is (Bradley and Smith 1989; Meyer et al. 2001)

$$D_{\mathbf{e}}\{\tilde{Z}(\mathbf{u})\} = \nabla\tilde{Z}(\mathbf{u}) \cdot \mathbf{e}, \tag{26}$$

where  $\mathbf{a} \cdot \mathbf{b}$  is the scalar product of the vectors  $\mathbf{a}$  and  $\mathbf{b}$ .

Direct measurements of the gradient can be incorporated in the estimation of the directional derivative by using co-kriging

$$\hat{D}_{\mathbf{e}_0}\{\tilde{Z}(\mathbf{u}_0)\} = \sum_{i=1}^n \lambda_i^0 Z(\mathbf{u}_i) + \sum_{j=1}^m \beta_j^0 D_{\mathbf{e}_j}\{\tilde{Z}(\mathbf{u}_j)\}, \tag{27}$$

where  $\hat{D}_{\mathbf{e}_0}\{\tilde{Z}(\mathbf{u}_0)\}$  is the estimator of the directional derivative  $D_{\mathbf{e}_0}\{\tilde{Z}(\mathbf{u}_0)\}$  at the spatial location  $\mathbf{u}_0$  in direction  $\mathbf{e}_0$  using  $n$  experimental values  $Z(\mathbf{u}_i)$  and  $m$  experimental measurements of the directional derivatives of  $Z(\mathbf{u})$ .  $D_{\mathbf{e}_j}\{\tilde{Z}(\mathbf{u}_j)\}$  provides the auxiliary information in the same way that  $Y(\mathbf{u}_j)$  does in Eq. (4).

The concept of estimating the gradient by ordinary kriging is given in Philip and Kitanidis (1989), and the estimation of the directional derivative by co-kriging is given in Pardo-Igúzquiza and Chica-Olmo (2004, 2007) in which co-kriging has been extended to take into account the trend that is observed in many spatial variables when considered at the regional scale. This is the case, for example, for the piezometric level or the water head in an aquifer, the rainfall in an area of rough orography, and the trend shown by geophysical variables associated with a sedimentary basin or a tectonic structure.

## 7 Co-kriging Taking Boundary Conditions into Account

This extension of co-kriging is related to the previous one, but the variable of interest is not the derivative of the primary variable but rather the primary variable itself. Thus, boundary conditions can be included in the co-kriging estimation by expressing

the boundary conditions in the form of directional derivatives that play the role of the auxiliary variable. For example, no-flow boundary conditions in an aquifer indicate that the directional derivatives of water head perpendicular to the boundary are null. In a similar manner, constant-head boundary conditions indicate that the directional derivatives parallel to the boundary are null (Fig. 1f). Thus, the new co-kriging estimator is given by

$$\tilde{Z}^*(\mathbf{u}_0) = \sum_{i=1}^n \lambda_i^0 Z(\mathbf{u}_i) + \sum_{j=1}^m \beta_j^0 D_{\mathbf{e}_0} \left\{ \tilde{Z}(\mathbf{u}_j) \right\}. \quad (28)$$

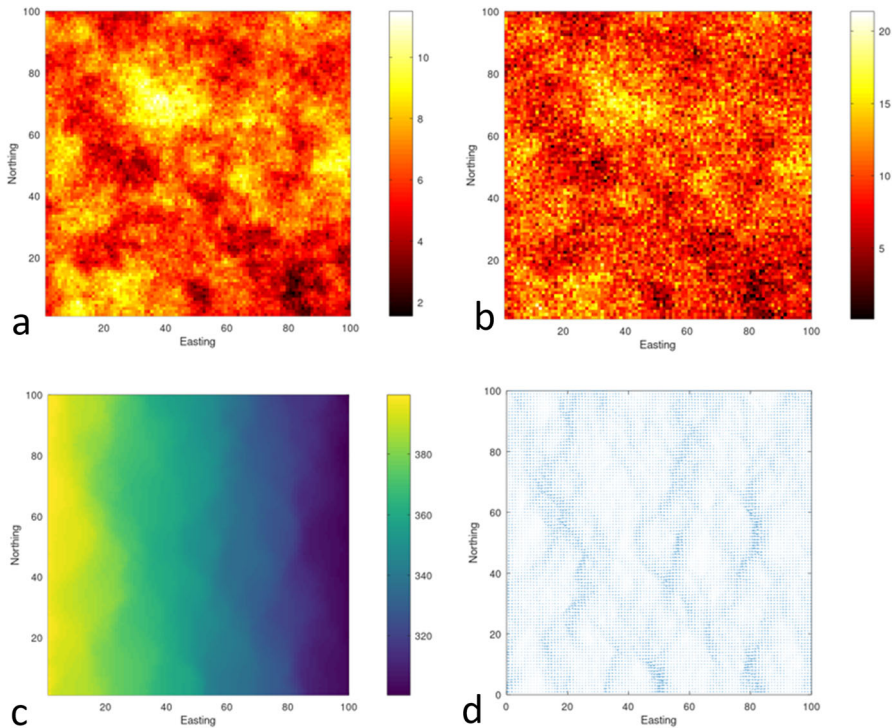
Although the right-hand sides of Eqs. (27) and (28) are the same, the weights will be different because the co-kriging systems are different. For example, the  $\mathbf{B}$  matrix is given by

$$\mathbf{B} = \begin{bmatrix} \tilde{C}\{\mathbf{h}_{10}\} \\ \vdots \\ \tilde{C}\{\mathbf{h}_{m0}\} \\ \tilde{C}_{\mathbf{e}_0}^*\{\mathbf{h}_{10}\} \\ \vdots \\ \tilde{C}_{\mathbf{e}_0}^*\{\mathbf{h}_{m0}\} \\ f_1(\mathbf{u}_0) \\ \vdots \\ f_L(\mathbf{u}_0) \end{bmatrix}. \quad (29)$$

The implementation of the methodology is given in detail in Kuhlman and Pardo-Igúzquiza (2010).

## 8 Illustrative Examples

We use a simulation to illustrate the different applications of co-kriging because, when the target field is known, the true errors can be evaluated and assessed. A  $100 \times 100$  grid of log-transmissivity values was simulated using the spectral simulation method for generating realisations of correlated random fields (Pardo-Igúzquiza and Chica-Olmo 1994a, b). A realisation of a random field of the log-transmissivity of an aquifer is shown in Fig. 4a. The auxiliary variable shown in Fig. 4b was obtained from Fig. 4a by adding noise (an uncorrelated variable) so that both random fields reproduce the same covariance model but with a nugget variance in Fig. 4b. The coefficient of correlation between the realisations in Fig. 4a, b is 0.718. Another auxiliary variable is provided in Fig. 4c which was obtained as the solution of the direct problem of the groundwater flow in an unconfined aquifer given Eq. (19) and the boundary conditions shown in Fig. 2f, with  $H_1 = 400$  m and  $H_2 = 300$  m and using as parameters the

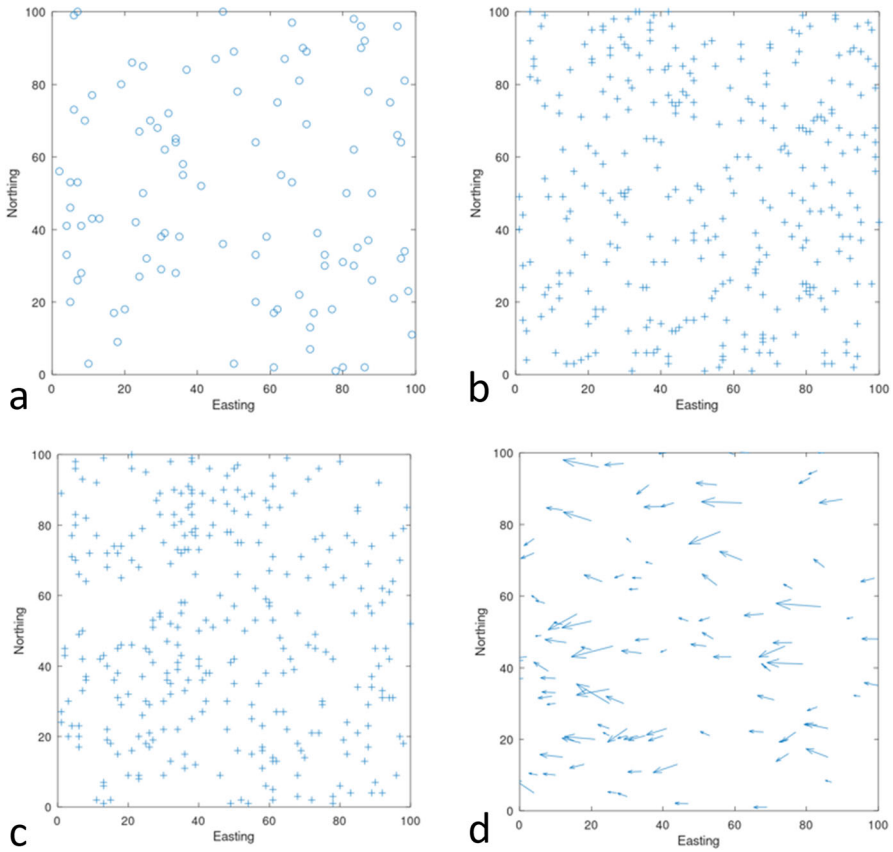


**Fig. 4** **a** Realisation of a random field representing the primary variable. **b** Realisation of the auxiliary variable obtained as a noisy version of **(a)**. **c** Realisation of a random field obtained as the solution of a direct problem in groundwater flow for which **(a)** is the parameter field. **d** Gradient field of **(c)**. This representation consists of small arrows that represent the magnitude and orientation of the gradient vector field. Each of the four realisations comprises  $100 \times 100$  elements

log-transmissivity field given in Fig. 4a. The gradient field of the scalar field of water head data in Fig. 4c is shown in Fig. 4d.

The experimental data were obtained by randomly sampling the gridded data in Fig. 4. In this way 100, 300, 300 and 100 data are sampled for log-transmissivity, porosity, water head and gradient, respectively, and are shown in Fig. 5a to d, respectively. The covariance models of log-transmissivity, porosity, and water head were inferred by maximum likelihood (Pardo-Igúzquiza 1997) assuming a constant mean and exponential covariance for log-transmissivity and porosity. A linear trend and a Gaussian covariance were assumed for the water head data. The cross-covariance between log-transmissivity and porosity was estimated using *sgems* (Remy et al. 2009). The estimated covariance parameters are given in Table 2.

When there are no auxiliary data, ordinary co-kriging reduces to ordinary kriging. The kriging estimates of the log-transmissivity field in Fig. 4a obtained by using only the experimental data of the primary variable (Fig. 5a) are shown in Fig. 6a. Validation statistics were used to compare the target field (Fig. 4a) and the interpolated field (Fig. 6a). The validation statistics used are the mean error (ME), mean squared error



**Fig. 5** Experimental data obtained by sampling from realisations on a  $100 \times 100$  grid. **a** 100 data obtained by random sampling from Fig. 2a. **b** 300 data obtained by random sampling from Fig. 2b. **c** 300 data obtained by random sampling from Fig. 2c. **d** 100 data obtained by random sampling from Fig. 2d

(MSE) and the mean absolute error (MAE),

$$ME = \frac{1}{N} \sum_{i=1}^N (Z^*(\mathbf{u}_i) - Z(\mathbf{u}_i)), \tag{30}$$

$$MSE = \frac{1}{N} \sum_{i=1}^N (Z^*(\mathbf{u}_i) - Z(\mathbf{u}_i))^2, \tag{31}$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |Z^*(\mathbf{u}_i) - Z(\mathbf{u}_i)|, \tag{32}$$

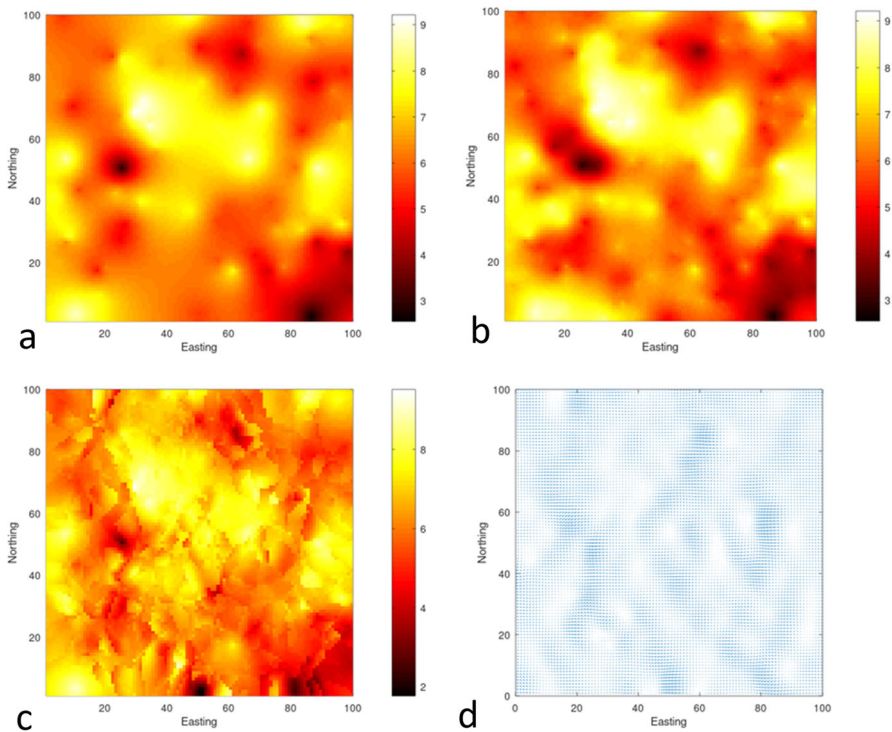
where  $(\mathbf{u}_i)$  is the target or true value at spatial location  $\mathbf{u}_i$ ,  $Z^*(\mathbf{u}_i)$  is an interpolated or estimated value at spatial location  $\mathbf{u}_i$ ,  $\sigma_*^2(\mathbf{u}_i)$  is the estimation variance at spatial



**Table 2** Estimated parameters of covariances and cross-covariances of the experimental data using MLREML (Pardo-Igúzquiza 1997) for the covariances and sgems (Remy et al. 2009) for the cross-covariance

Spatial variable(s)	Type	Nugget variance	Partial sill	Range
Covariance of log-transmissivity	Exponential $k = 0$	0	2.07	9.28
Covariance of porosity	Exponential $k = 0$	4.38	5.35	10.59
Cross-covariance between log-transmissivity and porosity	Exponential $k = 0$	0	3	10.00
Covariance of water head	Gaussian $k = 1$	0.25	1.99	8.41

The order ( $k$ ) of the trend implies a constant spatial mean ( $k = 0$ ) or a linear trend ( $k = 1$ ). The linear trend estimated by maximum likelihood for the water head is  $m(x, y) = 400.65 - 0.999x - 0.009y$



**Fig. 6** Interpolated fields on a  $100 \times 100$  grid. **a** Ordinary kriging map using the data in Fig. 3a. **b** Ordinary co-kriging map using the data in Fig. 3a and b. **c** Inverse problem universal co-kriging map using the data in Fig. 3a and c. **d** Gradient co-kriging using the data in Fig. 3a and d

**Table 3** Validation statistics for the estimated log-transmissivity for the different interpolators: ordinary kriging (O-K), ordinary co-kriging (O-COK) and inverse problem universal co-kriging (IP-U-COK)

Interpolator	ME	MSE	MAE
O-K	−0.061	1.078	0.828
O-COK	−0.072	0.932	0.766
IP-U-COK	−0.0772	1.115	0.839

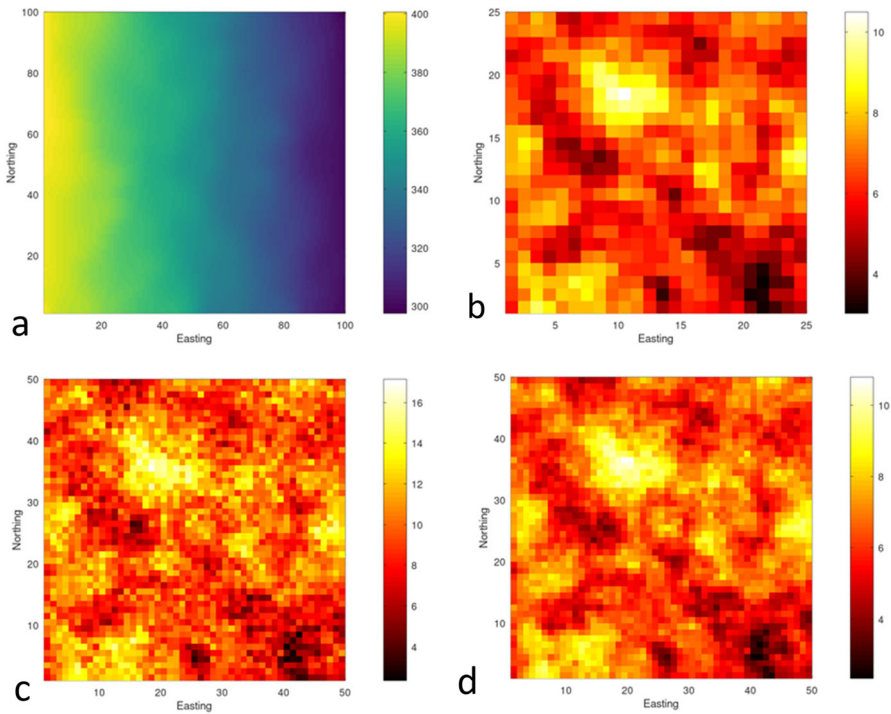
**Table 4** Validation statistics using the direct problem of Eq. (18). Ordinary kriging (O-K), ordinary co-kriging (O-COK) and inverse problem universal co-kriging (IP-U-COK)

Interpolator	ME	MSE	MAE
True field	0	0	0
O-K	0.218	2.591	1.239
O-COK	0.138	1.871	1.041
IP-U-COK	0.163	2.117	1.135

location  $\mathbf{u}_i$ , and  $N$  is the number of spatial locations at which the variable was estimated. For an unbiased model, ME should be zero. MSE and MAE should be as small as possible.

Table 3 shows the validation statistics for the estimation of the log-transmissivity field by ordinary kriging (O-K) using only log-transmissivity data (Fig. 5a); ordinary co-kriging (O-COK) using log-transmissivity data (Fig. 5a) and the auxiliary porosity variable (Fig. 5b); and inverse problem universal co-kriging (IP-U-COK) using the water head data (Fig. 5c) as the secondary variable. The latter case is a different way of incorporating a different type of auxiliary data. A comparison of the MSE values shows that co-kriging is a clear improvement over kriging, but that is not the case for IP-U-COK. The latter case should be compared with the output of the direct problem of estimating water head from a log-transmissivity field, the results of which are given in Table 4, which clearly show the improvement of IP-U-COK over kriging.

When estimating water head (Fig. 4c) as the primary variable, the universal kriging results using the experimental data in Fig. 5c can be compared with those of universal co-kriging taking the boundary conditions in Fig. 2e into account. This gives the estimated field in Fig. 7a. Table 5 lists the validation statistics that show the improvement of co-kriging with boundary conditions. With respect to the estimation of gradients, Fig. 6d shows the estimated gradient field that can be compared with the true gradient field in Fig. 4d to obtain the validation statistics given in Table 6. Finally, the log-transmissivity image in Fig. 4a has been degraded to a resolution image by a factor of 4 and to a high-resolution image by a factor of 2. Similarly, the auxiliary image in Fig. 4b has been degraded to a high-resolution image by a factor of 2. The purpose of downscaling co-kriging is to estimate a high-resolution image of the primary variable by using the low-resolution image of the primary variable and the high-resolution image of the auxiliary variable. The results are shown in Fig. 7, and the validation statistics are given in Table 7. Whilst every study must be considered in its own right, the results shown here demonstrate the potential of the methodology. Each estimation



**Fig. 7** Interpolated fields on a  $100 \times 100$  grid. **a** Boundary co-kriging map using the data in Figs. 1c and 3a. **b** Low-resolution image from Fig. 2a. **c** High-resolution image from Fig. 2a. **d** High-resolution image estimated by downscaling co-kriging

**Table 5** Validation statistics of the estimated water head for the different interpolators

Interpolator	ME	MSE	MAE
U-K	-0.049	1.165	0.894
B-COK	-0.054	1.015	0.717

Universal kriging (U-K) and co-kriging with boundary conditions (B-COK)

**Table 6** Validation statistics for the estimated gradient (directional derivatives in the X direction and the Y direction)

Derivative	ME	MSE	MAE
X direction	0.001	0.162	0.306
Y direction	-0.010	0.095	0.232

**Table 7** Validation statistics for downscaling co-kriging (DS-COK) for the result shown in Fig. 7d

Interpolator	ME	MSE	MAE
DS-COK	0.123	0.959	0.788

is accompanied by a map of the associated estimation variance which can be used as a measure of the uncertainty of the estimates.

## 9 Conclusions

The ordinary co-kriging estimator that has traditionally been used to estimate a primary variable using data from that variable and data from an auxiliary variable can be modified to take into account a variety of aspects of significant practical interest as shown in Fig. 3. These aspects include accounting for different supports of the primary and auxiliary variables and the mathematical link between the primary and auxiliary variables which provide a means of solving problems such as the following:

- Downscaling co-kriging to increase the spatial resolution of the primary variable
- Solving inverse problems
- Estimating directional derivatives and hence the gradient of a scalar variable
- Including boundary conditions in the estimation of a scalar variable.

This is not an exhaustive list of applications, and more are likely to arise in future research. In addition, there are cases in which diverse forms of co-kriging are required because of the type of data. Examples include indicator co-kriging (Pardo-Igúzquiza and Dowd 2005), compositional co-kriging (Pardo-Igúzquiza et al. 2015) and estimating spatial factors or spatial components using factorial co-kriging (Pardo-Igúzquiza and Dowd 2002).

All the methods discussed here include an evaluation of the uncertainty of the estimates given by the estimation variance in Eq. (9). For each map of estimates, there is a corresponding map of the associated estimation variance. Among other applications, the evaluation of uncertainty is important for establishing interval estimates, for optimal sampling and for the propagation of the uncertainty of measurements.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11004-023-10104-7>.

**Acknowledgements** The work reported here was supported by research project PID2019-106435GB-I00 of the Ministerio de Ciencia e Innovación of Spain and by the Australian Research Council Industrial Transformation Training Centre for Integrated Operations for Complex Resources.

**Funding** Open Access funding enabled and organized by CAUL and its Member Institutions. The authors have no relevant financial or non-financial interests to disclose. The authors certify that they have no affiliations with, or involvement in, any organisation or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no financial or proprietary interests in any material discussed in this article.

### Declarations

**Conflict of interest** The authors have no conflicts of interest to declare that are relevant to the content of this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative

Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

### Appendix A. Ordinary co-kriging

The co-kriging system can be written in matrix form as

$$\mathbf{C}\lambda = \mathbf{B}, \tag{A1}$$

with

$$\mathbf{C} = \begin{bmatrix} C_Z(\mathbf{h}_{11}) & \cdots & C_Z(\mathbf{h}_{1n}) & C_{ZY}(\mathbf{h}_{11}) & \cdots & C_{ZY}(\mathbf{h}_{nm}) & 1 & 0 \\ \vdots & & \ddots & \vdots & & \ddots & \vdots & \vdots \\ C_Z(\mathbf{h}_{n1}) & \cdots & C_Z(\mathbf{h}_{nn}) & C_{ZY}(\mathbf{h}_{n1}) & \cdots & C_{ZY}(\mathbf{h}_{nm}) & 1 & 0 \\ C_{YZ}(\mathbf{h}_{11}) & \cdots & C_{YZ}(\mathbf{h}_{1n}) & C_Y(\mathbf{h}_{11}) & \cdots & C_Y(\mathbf{h}_{1m}) & 0 & 1 \\ \vdots & & \ddots & \vdots & & \vdots & \vdots & \vdots \\ C_{YZ}(\mathbf{h}_{m1}) & \cdots & C_{YZ}(\mathbf{h}_{mn}) & C_Y(\mathbf{h}_{m1}) & \cdots & C_Y(\mathbf{h}_{mm}) & 0 & 1 \\ 1 & \cdots & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & \cdots & 1 & 1 & \cdots & 1 & 0 & 0 \end{bmatrix} \tag{A2}$$

$$\lambda = \begin{bmatrix} \lambda_1^0 \\ \vdots \\ \lambda_n^0 \\ \beta_1^0 \\ \vdots \\ \beta_m^0 \\ \mu_1 \\ \mu_2 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} C_Z(\mathbf{h}_{10}) \\ \vdots \\ C_Z(\mathbf{h}_{n0}) \\ C_{ZY}(\mathbf{h}_{01}) \\ \vdots \\ C_{ZY}(\mathbf{h}_{0m}) \\ 1 \\ 0 \end{bmatrix} \tag{A3}$$

where  $\mu_1$  and  $\mu_2$  are two Lagrange multipliers or parameters in the minimization of the estimation variance subject to the constraints given in Eqs. (7) and (8). Note that  $C_{ZY}(h_{ij}) = C_{YZ}(h_{ji})$  always so that the matrix  $\mathbf{C}$  is symmetric.

The solution of the co-kriging system

$$\lambda = \mathbf{C}^{-1}\mathbf{B}, \tag{A4}$$

provides the weights required in the estimator in Eq. (4).

## Appendix B. Universal cokriging: spatially variable mean

The primary variable or the auxiliary variable or both variables often show a spatial trend or a spatially variable mean. In this case, Eq. (1) should be replaced by

$$E\{Z(\mathbf{u})\} = m(\mathbf{u}), \quad (\text{B1})$$

and the geostatistical model is a random function  $Z(\mathbf{u})$  with two components (Wackernagel 2003)

$$Z(\mathbf{u}) = m(\mathbf{u}) + R(\mathbf{u}), \quad (\text{B2})$$

where  $m(\mathbf{u})$  is a deterministic component that represents the trend or spatially variable mean, i.e., a smooth component of low frequency variability. Equation (B1) expresses the trend as the mathematical expectation of the random function, which is usually modelled as a low order polynomial, although there are other alternatives (see for example Brochu and Marcotte 2003)

$$m(\mathbf{u}) = \sum_{i=0}^p b_i f_i(\mathbf{u}), \quad (\text{B3})$$

where  $\{b_i; i = 0, \dots, p\}$ : is a set of coefficients to be estimated.  $\{f_i(\mathbf{u}); i = 0, \dots, p\}$ : is a set of known base functions, usually monomial, of the co-ordinates. For example,  $\{1, x, y, xy, x^2, y^2\}$  are the six base functions for the case of a quadratic polynomial trend on the plane.

$R(\mathbf{u})$  is the residual, a second order stationary random function (Chilès and Definer, 1999) with zero mean, covariance  $C_R(\mathbf{h})$  and variogram  $\gamma_R(\mathbf{h}) = C_R(0) - C_R(\mathbf{h})$ . The residual represents a component of high frequency variability around the trend. Note that the covariance of the residual  $R(\mathbf{u})$  is identical to the covariance of the random function  $Z(\mathbf{u})$

$$\begin{aligned} C_R(\mathbf{h}) &= E\{R(\mathbf{u})R(\mathbf{u} + \mathbf{h})\} \\ &= E\{[Z(\mathbf{u}) - m(\mathbf{u})][Z(\mathbf{u} + \mathbf{h}) - m(\mathbf{u} + \mathbf{h})]\} = C_Z(\mathbf{h}) \end{aligned} \quad (\text{B4})$$

Also note that the case of a constant mean in Eq. (1) is accommodated in the polynomial drift model of equation (B3) by setting,  $p = 0$ ,  $b_0 = m_Z$  and  $f_0 = 1$ .

The universal kriging estimator is equal to the ordinary co-kriging estimator of Eq. (4) but the weights are obtained from the universal co-kriging system. As an example, when there is a linear drift only in the primary variable, the co-kriging system must include the following unbiasedness conditions

$$\sum_{i=1}^n \lambda_i^0 = 1, \quad (\text{B5})$$

$$\sum_{i=1}^n \lambda_i^0 x_i = x_0, \tag{B6}$$

$$\sum_{i=1}^n \lambda_i^0 y_i = y_0, \tag{B7}$$

$$\sum_{j=1}^m \beta_j^0 = 0 \tag{B8}$$

The universal co-kriging system in matrix form is similar to equations (A2) and (A3) and can be found elsewhere (for example Goovaerts 1997). Universal co-kriging can be found in the geostatistical literature (Goovaerts 1997; Olea 1999; Chilès and Delfiner 1999).

### Appendix C. Co-kriging as a solution of an inverse problem

For an unbounded aquifer, in which the water head has a linear trend along one of the main co-ordinate axes and using an exponential covariance for the log-transmissivity, Dagan (1985) found that the cross-covariance between log-transmissivity and water head is equal to

$$C_{ZY}(\mathbf{h}) = \sigma_Y^2 b_1 \ell \frac{h_x}{h^2} [1 - (1 + h) \exp(-h)], \tag{C1}$$

where  $\sigma_Y^2$  : is the variance, or total sill, of the log-transmissivity  $Z(\mathbf{u})$ .  $\uparrow$  : is the range of the exponential covariance of log-transmissivity  $Z(\mathbf{u})$ . The covariance of log-transmissivity is

$$C_Z(\mathbf{h}) = C_0 + C_1 \exp(-h/\ell). \tag{C2}$$

$$\sigma_Y^2 = C_0 + C_1. \tag{C3}$$

where  $C_0$  : is the nugget variance.  $C_1$  : is the partial sill variance.  $\mathbf{h} = (h_x, h_y) = \left( \frac{x_1 - x_2}{\uparrow}, \frac{y_1 - y_2}{\uparrow} \right)$  is the distance vector standardized by the range of log-transmissivity.  $h = \sqrt{h_x^2 + h_y^2}$  is the magnitude of the standardized distance vector and  $m(x, y) = b_0 + b_1 x$  is the linear trend of the water head which, in this case, is an east–west trend.

## Appendix D. Co-kriging for estimating directional derivatives

In the estimation of directional derivatives, the non-bias conditions for the co-kriging estimator are

$$\sum_{i=1}^n \lambda_i f_\ell(\mathbf{u}_i) + \sum_{j=1}^m w_j D_{\mathbf{e}_j} \{f_\ell(\mathbf{u}_j)\} = D_{\mathbf{e}_0} \{f_\ell(\mathbf{u}_0)\}, \ell = 0, \dots, p, \quad (\text{D1})$$

and the estimation variance is

$$\begin{aligned} & \text{Var} \left\{ \hat{D}_{\mathbf{e}_0} \left\{ \tilde{Z}(\mathbf{u}_0) \right\} - D_{\mathbf{e}_0} \left\{ \tilde{Z}(\mathbf{u}_0) \right\} \right\} \\ &= \tilde{C}_{\mathbf{e}_0}^{**}(0) + \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C(\mathbf{h}_{ij}) + \sum_{k=1}^m \sum_{\ell=1}^m w_k w_\ell \tilde{C}_{\mathbf{e}_0}^{**}(\mathbf{h}_{k\ell}) \\ &+ 2 \sum_{s=1}^n \sum_{t=1}^m \lambda_s w_t \tilde{C}_{\mathbf{e}_0}^*(\mathbf{h}_{st}) - 2 \sum_{i=1}^n \lambda_i \tilde{C}_{\mathbf{e}_0}^*(\mathbf{h}_{i0}) - 2 \sum_{j=1}^m w_j \tilde{C}_{\mathbf{e}_0}^{**}(\mathbf{h}_{j0}), \end{aligned} \quad (\text{D2})$$

where  $D_{\mathbf{e}_0} \{f_\uparrow(\mathbf{u}_0)\}$  is the directional derivative of the basis functions  $f_\uparrow(\cdot)$  evaluated in direction  $\mathbf{e}_0$  for position  $\mathbf{u}_0$ .  $C(\mathbf{h})$  is the covariance of  $Z(\mathbf{u})$ .  $\tilde{C}(\mathbf{h})$  is the covariance of  $\tilde{Z}(\mathbf{u}) = m(\mathbf{u}) + R(\mathbf{u})$ , that is, the covariance  $C(\mathbf{h})$  but without the nugget variance. The covariance of the directional derivative of the random function is

$$\tilde{C}_{\mathbf{e}_0}^{**}(\mathbf{h}) = \text{Cov} \left\{ D_{\mathbf{e}_0} \left\{ \tilde{Z}(\mathbf{u}) \right\}, D_{\mathbf{e}_0} \left\{ \tilde{Z}(\mathbf{u} + \mathbf{h}) \right\} \right\} = -D_{\mathbf{e}_0}^2 \left\{ \tilde{C}(\mathbf{h}) \right\}, \quad (\text{D3})$$

and is equal to minus the second directional derivative of the covariance  $\tilde{C}(\mathbf{h})$ . Finally, the cross-covariance between the random function  $\tilde{Z}(\mathbf{u})$  and its directional derivative is given by

$$\tilde{C}_{\mathbf{e}_0}^*(\mathbf{h}) = \text{Cov} \left\{ \tilde{Z}(\mathbf{u}), D_{\mathbf{e}_0} \left\{ \tilde{Z}(\mathbf{u} + \mathbf{h}) \right\} \right\} = -D_{\mathbf{e}_0} \left\{ \tilde{C}(\mathbf{h}) \right\} \quad (\text{D4})$$

and is equal to minus the directional derivative of the covariance  $\tilde{C}(\mathbf{h})$ .

The last two results are from linear systems theory (Parzen 1972; Papoulis 1984) and it can be concluded that the only covariance that must be inferred experimentally is the covariance  $\tilde{C}(\mathbf{h})$  of the scalar variable.

The optimal weights in Eq. (26) are obtained by minimizing the estimation variance of equation (D2) subject to the unbiased conditions in equation (D1). Thus, the optimal weights are obtained by solving the co-kriging system  $\mathbf{C} = \lambda \mathbf{B}$  which in matrix form are



$$\mathbf{C} = \begin{bmatrix} C(\mathbf{h}_{11}) & \cdots & C(\mathbf{h}_{1n}) & \tilde{C}_{\mathbf{e}_0}^*(\mathbf{h}_{11}) & \cdots & \tilde{C}_{\mathbf{e}_0}^*(\mathbf{h}_{1m}) & f_1(\mathbf{u}_1) & \cdots & f_L(\mathbf{u}_1) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ C(\mathbf{h}_{n1}) & \cdots & C(\mathbf{h}_{nn}) & \tilde{C}_{\mathbf{e}_0}^*(\mathbf{h}_{n1}) & \cdots & \tilde{C}_{\mathbf{e}_0}^*(\mathbf{h}_{nm}) & f_1(\mathbf{u}_n) & \cdots & f_L(\mathbf{u}_n) \\ \tilde{C}_{\mathbf{e}_0}^*(\mathbf{h}_{11}) & \cdots & \tilde{C}_{\mathbf{e}_0}^*(\mathbf{h}_{1n}) & \tilde{C}_{\mathbf{e}_0}^{**}(\mathbf{h}_{11}) & \cdots & \tilde{C}_{\mathbf{e}_0}^{**}(\mathbf{h}_{1m}) & D_{\mathbf{e}_0}\{f_1(\mathbf{u}_1)\} & \cdots & D_{\mathbf{e}_0}\{f_L(\mathbf{u}_1)\} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \tilde{C}_{\mathbf{e}_0}^*(\mathbf{h}_{m1}) & \cdots & \tilde{C}_{\mathbf{e}_0}^*(\mathbf{h}_{mn}) & \tilde{C}_{\mathbf{e}_0}^{**}(\mathbf{h}_{m1}) & \cdots & \tilde{C}_{\mathbf{e}_0}^{**}(\mathbf{h}_{mm}) & D_{\mathbf{e}_0}\{f_1(\mathbf{u}_m)\} & \cdots & D_{\mathbf{e}_0}\{f_L(\mathbf{u}_m)\} \\ f_1(\mathbf{u}_1) & \cdots & f_1(\mathbf{u}_n) & D_{\mathbf{e}_0}\{f_1(\mathbf{u}_1)\} & \cdots & D_{\mathbf{e}_0}\{f_1(\mathbf{u}_m)\} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ f_L(\mathbf{u}_1) & \cdots & f_L(\mathbf{u}_n) & D_{\mathbf{e}_0}\{f_L(\mathbf{u}_1)\} & \cdots & D_{\mathbf{e}_0}\{f_L(\mathbf{u}_m)\} & 0 & \cdots & 0 \end{bmatrix} \tag{D5}$$

$$\boldsymbol{\lambda} = \begin{bmatrix} \lambda_1^0 \\ \vdots \\ \lambda_n^0 \\ \beta_1^0 \\ \vdots \\ \beta_m^0 \\ \mu_1 \\ \vdots \\ \mu_L \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \tilde{C}_{\mathbf{e}_0}^*\{\mathbf{h}_{10}\} \\ \vdots \\ \tilde{C}_{\mathbf{e}_0}^*\{\mathbf{h}_{n0}\} \\ \tilde{C}_{\mathbf{e}_0}^{**}\{\mathbf{h}_{10}\} \\ \vdots \\ \tilde{C}_{\mathbf{e}_0}^{**}\{\mathbf{h}_{m0}\} \\ D_{\mathbf{e}_0}\{f_1(\mathbf{u}_0)\} \\ \vdots \\ D_{\mathbf{e}_0}\{f_L(\mathbf{u}_0)\} \end{bmatrix} \tag{D6}$$

**References**

Ahmed S, De Marsily G (1993) Cokriged estimation of aquifer transmissivity as an indirect solution of the inverse problem: a practical approach. *Water Resour Res* 29(2):521–530

Atkinson PM, Webster R, Curran PJ (1994) Co-kriging with airborne MSS imagery. *Remote Sens Environ* 50:335–345

Atkinson P, Pardo-Igúzquiza E, Chica-Olmo M (2008) Downscaling co-kriging for super-resolution mapping of continua in remotely sensed images. *TGARS Trans Geosci Remote Sens* 46(2):573–580

Babak O, Deutsch CV (2009) Collocated co-kriging based on merged secondary attributes. *Math Geosci* 41:921–926

Bradley GL, Smith KL (1989) *Calculus*. Prentice Hall, p 1239

Brochu Y, Marcotte D (2003) A simple approach to account for radial flow and boundary conditions when kriging hydraulic head fields for conditioned aquifers. *Math Geol* 35(2):111–139

Chilès J-P, Delfiner P (1999) *Geostatistics: modelling spatial uncertainty*. Wiley, New York, p 695

Dagan G (1985) Stochastic modelling of groundwater flow by unconditional and conditional probabilities: the inverse problem. *Water Resour Res* 21(1):65–72

De Iaco S, Hristopulos DT, Lin G (2022) Special issue: geostatistics and machine learning. *Math Geosci* 54:459–465

Deutsch CV, Journel AG (1992) *GSLIB. Geostatistical software library and user’s guide*. Oxford University Press, New York, p 340

Doyen PM (1988) Porosity from seismic data: a geostatistical approach. *Geophysics* 53:1263–1275

Goovaerts P (1997) *Geostatistics for natural resources evaluation*. Oxford University Press, New York, p 483

- Graham AJ, Atkinson PM, Danson FM (2004) Spatial analysis for epidemiology. *Acta Trop* 91(3):219–225
- Hoeksema RJ, Kitanidis PK (1984) An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modelling. *Water Resour Res* 20(7):1003–1020
- Hoeksema RJ, Clapp RB, Thomas AL, Hunley AE, Farrow ND, Dearstone KC (1989) Co-kriging model for estimation of water table elevation. *Water Resour Res* 25(3):429–438
- Isaaks EH, Srivastava RM (1989) *Applied geostatistics*. Oxford University Press, New York, p 561
- Journel AG (1999) Markov models for cross-covariances. *Math Geol* 31(8):931–953
- Journel A, Huijbregts C (1978) *Mining geostatistics*. Academic Press, New York, p 600
- Kitanidis PK (1997) *Introduction to geostatistics. Applications in hydrogeology*. Cambridge University Press, p 320
- Kuhlman K, Pardo-Igúzquiza E (2010) Universal co-kriging of hydraulic heads accounting for boundary conditions. *J Hydrol* 384(1–2):14–25
- Lesch SM, Strauss DJ, Rhoades JD (1995) Spatial prediction of soil salinity using electromagnetic induction techniques: 1. Statistical prediction models: a comparison of multiple linear regression and co-kriging. *Water Resour Res* 31(2):373–386
- Matheron G (1975) *Random sets and integral geometry*. Wiley, New York
- Matheron G (1963) *Traité de Géostatistique Appliquée: le krigeage*. Tome 2. Paris, Editions Technip
- Meyer TH, Eriksson M, Maggio RC (2001) Gradient estimation from irregularly spaced data sets. *Math Geol* 33(6):693–717
- Myers DE (1982) Matrix formulation of co-kriging. *Math Geol* 14(3):249–257
- Myers DE (1983) Estimation of linear combinations and co-kriging. *Math Geol* 15(5):633–637
- Myers DE (1991) Pseudo-cross variograms, positive-definiteness and co-kriging. *Math Geol* 23(6):805–816
- Olea RA (1999) *Geostatistics for engineers and earth scientists*. Kluwer Academic Publishers, Boston, p 303
- Pardo-Igúzquiza E (1997) MLREML: a computer program for the inference of spatial covariance parameters by maximum likelihood and restricted maximum likelihood. *Comput Geosci* 23(2):153–162
- Papoulis A (1984) *Probability, random variables and stochastic processes*, 2nd edn. McGraw-Hill International Editions, Singapore
- Pardo-Igúzquiza E (1998) Comparison of geostatistical methods for estimating the areal average climatological rainfall mean using information of precipitation and topography. *Int J Climatol* 18:1031–1047
- Pardo-Igúzquiza E, Atkinson PM (2007) Modelling the semi-variograms and cross-semi-variograms required in downscaling co-kriging by numerical convolution-deconvolution. *Comput Geosci* 33(10):1273–1284
- Pardo-Igúzquiza E, Chica-Olmo M (1994a) Spectral simulation of multivariable stationary random functions using covariance Fourier transforms. *Math Geol* 26(3):277–299
- Pardo-Igúzquiza E, Chica-Olmo M (1994b) SPECSIM: a program for simulating random fields by an improved spectral approach. *Comput Geosci* 20(4):597–613
- Pardo-Igúzquiza E, Chica-Olmo M (2004) Estimation of gradients from sparse data by universal kriging. *Water Resour Res* 40(W12418):1–17
- Pardo-Igúzquiza E, Chica-Olmo M (2007) KRIGRADI: a co-kriging program for estimating the gradient of spatial variables from sparse data. *Comput Geosci* 33(4):497–512
- Pardo-Igúzquiza E, Dowd PA (2002) FACTOR2D: a computer program for factorial co-kriging. *Comput Geosci* 28:857–875
- Pardo-Igúzquiza E, Dowd PA (2005) Multiple indicator co-kriging with application to optimal sampling for environmental monitoring. *Comput Geosci* 31(1):1–13
- Pardo-Igúzquiza E, Chica-Olmo M, Atkinson P (2006) Downscaling co-kriging for image sharpening. *Remote Sens Environ* 102:86–98
- Pardo-Igúzquiza E, Chica-Olmo M, Mardia KV (2009) MLMATERN: a computer program for maximum likelihood inference with the spatial Matérn covariance model. *Comput Geosci* 35:1139–1150
- Pardo-Igúzquiza E, Atkinson PM, Chica-Olmo M (2010) DSCOKRI: a library of computer programs for downscaling co-kriging in support of remote sensing applications. *Comput Geosci* 36:881–894
- Pardo-Igúzquiza E, Chica-Olmo M, Luque-Espinar JA, Rodríguez-Galiano V (2015) Compositional co-kriging for mapping the probability risk of groundwater contamination by nitrates. *Sci Total Environ* 532:162–175
- Parzen E (1972) *Stochastic processes*. Holden-Day Inc., San Francisco, p 370
- Philip RD, Kitanidis PK (1989) Geostatistical estimation of hydraulic head gradients. *Ground Water* 27(6):855–865

- Remy N, Boucher A, Wu J (2009) *Applied geostatistics with SGeMS: a user's guide*. Cambridge University Press, Cambridge, p 264
- Teegavarapu RSV, Pathak CS, Mecikalski JR, Srikishen J (2015) Optimal solar radiation sensor network design using spatial and geostatistical analyses. *J Spat Sci* 61(1):69–97
- Van der Meer F (2012) Remote-sensing image analysis and geostatistics. *Int J Remote Sens* 33(18):5644–5676
- Ver Hoef JM, Cressie N (1993) Multivariate spatial prediction. *Math Geol* 25(2):219–240
- Wackernagel H (2003) *Multivariate geostatistics. An introduction with applications*, 3rd edn. Springer, Berlin, p 386
- Zhang C, Zuo R, Xiong Y, Shi X, Donnelly C (2021) GIS, geostatistics and machine learning in medical geology. In: Siegel M, Selinus O, Finkelman R (eds) *Practical applications of medical geology*, chapter 7, pp 215–234