Journal of Hydrology 384 (2010) 14–25

Contents lists available at ScienceDirect

Journal of Hydrology

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

Universal cokriging of hydraulic heads accounting for boundary conditions

Kristopher L. Kuhlman [a,*], Eulogio Pardo Igúzquiza [b]

[a] Repository Performance Department, Sandia National Laboratories, 4100 National Parks Highway, Carlsbad, NM 88220, USA
[b] Instituto Geológico y Minero de España, Calle Rios Rosas 23, 28003 Madrid, Spain

ARTICLE INFO

Article history:
Received 27 March 2009
Received in revised form 17 November 2009
Accepted 6 January 2010

This manuscript was handled by P. Baveye, Editor-in-Chief, with the assistance of Chunmiao Zheng, Associate Editor

Keywords:
Kriging
Cokriging
Boundary conditions
Matérn covariance
Cauchy covariance
Gaussian covariance

SUMMARY

When contouring scalar potentials from point observations the process can often benefit from including the known effects of boundary curves with specified potential or gradient. Here we consider the hydraulic head in an aquifer and both no-flow and constant-head boundary conditions. We present a new approach to enforcing that equipotential contours be normal to no-flow boundaries. A constant-head boundary, with unknown head, can be included through the same process by rotating the boundary vector by 90°. Collocated observations of heads and boundaries can specify a constant-head boundary of known value. We estimate head given both head and boundary condition observations, cokriging with both types of information. Our new approach uses gradient vectors in contrast with previous approximate finite-difference methods that include boundary conditions in kriging. Either the approach given here or the finite-difference method must be implemented with smooth covariance models, e.g., Gaussian, generalized Cauchy, and Matérn.

Introduction

Kriging can be used to estimate hydraulic head between observation wells (e.g., on a grid) for the construction of 2D equipotential contour maps (Kitanidis, 1997). Although kriging is the best linear unbiased estimator, it does not include geologic or physical knowledge of the system (beyond structure embodied in the variogram), that a practitioner would likely use when contouring the same potentials by hand. We illustrate a method to include boundary condition information in the kriging of 2D potentials. These boundary conditions may include no-flow conditions along faults or hydrologic contacts, or constant head conditions.

Cokriging is the multi-variable extension to kriging, and was developed in mining to address the common problem of estimating an under-sampled variable (e.g., Chilés and Delfiner, 1999, Chapter 5). Often an allied variable is estimated more frequently than the variable of interest, and the correlation between the two is used to improve the quality of the final estimate (e.g., Chilés and Delfiner, 1999; Goovaerts, 1997; Isaaks and Srivastava, 1989; Kitanidis, 1997). Frequently the information contained in a second variable can be used to enhance estimates of the primary variable. The estimation of cross-covariance functions, which describe the spatial correlation between variables at different locations, is a hindrance to more widespread use of cokriging (Isaaks and Srivastava, 1989). Unless the dataset is exhaustive, cross-covariance models estimated solely from data are questionable.

Our approach derives the required cross-covariance functions from the mathematical relationship followed by a potential and its gradient. The benefits are twofold; first, only the direct covariance (or equivalent variogram) function need be estimated (as is done for single-variable kriging) and secondly, the cokriging now honors a portion of the underlying physical process (i.e., the spatial relationship between the potential and its gradient), which single-variable kriging cannot.

Pardo Igúzquiza and Chica Olmo (2004) and Brochu and Marcotte (2003) discuss covariance models which can be used to represent a function known to be second-order continuous, i.e., a potential governed by a second-order differential equation. The covariance model must be continuous at the origin (zero lag); most common covariance models do not satisfy this requirement (e.g., exponential and spherical), we will discuss three that do.

Chilés and Delfiner (1999, p. 319) introduced a finite-difference approximation to no-flow boundary condition information when kriging hydraulic heads, referring to an unpublished presentation by Delhomme from 1979. More recently Brochu and Marcotte (2003) give a finite-difference example in terms of a dual kriging formulation. We develop the cokriging equations using the true

* Corresponding author. Tel.: +1 575 234 0084; fax: +1 575 234 0061.
E-mail address: klkuhlm@sandia.gov (K.L. Kuhlman).

0022-1694/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhydrol.2010.01.002

gradients of head and the cross-covariance models required from the covariance function used for heads. We compare and contrast this with the finite-difference approach.

True derivative approach

Cokriging is used to include boundary condition information when kriging potentials. Pardo Igúzquiza and Chica Olmo (2004, 2007) extended the related procedure of estimating the gradients using head data alone (Philip and Kitanidis, 1989), deriving the covariance and cross-covariance functions analytically from the Gaussian covariance which models the variability of head. The cokriging linear estimator, Z^*, is

$$Z^*(\mathbf{x}_0) = \sum_{\alpha=1}^{N} \lambda_\alpha Z(\mathbf{x}_\alpha) + \sum_{\beta=1}^{M} \delta_\beta[\hat{\nu}(\mathbf{x}_\beta) \cdot \nabla Z(\mathbf{x}_\beta)], \quad (1)$$

where $Z(\mathbf{x}_\alpha)$ is the potential observed at $\mathbf{x}_\alpha$ ($\mathbf{x}$ a 2D Cartesian coordinate), $\mathbf{x}_0$ is the location where the potential is to be estimated, λ_α and δ_β are cokriging weights, α and β are dummy variables, N and M are the number of head and boundary observations respectively, and $\hat{\nu}(\mathbf{x}_\beta)$ is a unit vector normal to the no-flow boundary (or tangent to a constant-head boundary), see Fig. 1. Using the shorthand notation $Z(\mathbf{x}_\alpha) = Z_\alpha$, we denote the directional derivative of Z_α, in the direction $\hat{\nu}$, as $(\hat{\nu} \cdot \nabla Z)_\alpha = Z_\alpha^{\hat{\nu}}$.

Kriging with a trend

The universal kriging drift is assumed to be polynomial in form, specified as

$$E[Z_0] = m(\mathbf{x}_0) = \sum_{\ell=0}^{L} \gamma_\ell f_0^\ell, \quad (2)$$

where E is the expectation operator, $m(\mathbf{x}_0)$ is the mean (a smooth function), γ_ℓ are free coefficients, L is the order of the polynomial approximation, and the monomial basis functions are $f^0 = 1,\ f^1 = x,\ f^2 = y,\ f^3 = x^2,\ f^4 = y^2,\ f^5 = xy$, etc. For $L = 0$, the universal kriging system simplifies to that of ordinary kriging (constant unknown mean). $L = 2$ corresponds to a linear trend, while $L = 5$ corresponds to a quadratic trend. Brochu and Marcotte (2003) indicate how other types of drift basis functions (e.g., the Thiem steady-state well solution) can be used in the contouring of hydraulic heads in special circumstances where additional information is known about the flow system.

Unbiasedness condition

The kriging weights λ_α and δ_α are sought to minimize the variance, while producing an unbiased solution. The unbiasedness condition is

$$E\left[Z_0^* - Z_0\right] = 0 \quad (3)$$

where Z_0 is the unknown true value at the desired estimation location. This can be expanded using (1) and (2) as

$$\begin{aligned} E[Z_0] &= \sum_\alpha \lambda_\alpha E[Z_\alpha] + \sum_\beta \delta_\beta E\left[Z_\beta^{\hat{\nu}}\right] \\ \sum_\ell \gamma_\ell f_0^\ell &= \sum_\ell \gamma_\ell \sum_\alpha \lambda_\alpha f_\alpha^\ell + \sum_\ell \gamma_\ell \sum_\beta \delta_\beta (\hat{\nu} \cdot \nabla f^\ell)_\beta \\ f_0^\ell &= \sum_\alpha \lambda_\alpha f_\alpha^\ell + \sum_\beta \delta_\beta (f^\ell)_\beta^{\hat{\nu}} \quad \ell = 0, \ldots, L \end{aligned} \quad (4)$$

where the $\hat{\nu}$ component of the gradient of the ℓth monomial drift term, $(f^\ell)_\beta^{\hat{\nu}}$, at location $\mathbf{x}_\beta$, can be computed explicitly. These gradients are $\nabla f_\beta^\ell = 0,\ \hat{\imath},\ \hat{\jmath},\ 2x\hat{\imath},\ 2y\hat{\jmath},\ y\hat{\imath} + x\hat{\jmath}$, for $\ell = 0, \ldots, 5$, where $\hat{\imath}$ and $\hat{\jmath}$ are the Cartesian unit vectors. These gradient vectors are projected onto $\hat{\nu}$. To ensure the expected value of the prediction is equal to the mean, $m(\mathbf{x})$, we enforce (4) while minimizing the estimation variance.

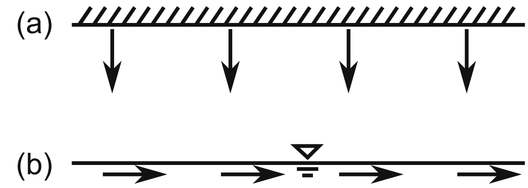

Fig. 1. No-flow (a) and constant head (b) boundary conditions, represented with boundary vectors (tail at point of application). Arrows offset from boundary for clarity.

Estimation variance

The variance of the estimation error R due to the linear estimator is

$$\mathrm{Var}[R] = E\{\left[Z_0^* - Z_0\right]^2\} \quad (5)$$

Following a procedure akin to that used to derive the standard kriging and cokriging equations (e.g., Goovaerts, 1997; Isaaks and Srivastava, 1989; Kitanidis, 1997), we substitute (1) and expand (5) as

$$\begin{aligned} \mathrm{Var}[R] = {} & \lambda_\alpha \lambda_\beta E[Z_\alpha Z_\beta] + \delta_\alpha \delta_\beta E\left[Z_\alpha^{\hat{\nu}} Z_\beta^{\hat{\nu}}\right] + E[Z_0 Z_0] + 2\lambda_\alpha \delta_\beta E\left[Z_\alpha Z_\beta^{\hat{\nu}}\right] \\ & - 2\lambda_\alpha E[Z_\alpha Z_0] - 2\delta_\alpha E\left[Z_0 Z_\alpha^{\hat{\nu}}\right] \end{aligned} \quad (6)$$

For brevity, Einstein summation convention is used; pairs of dummy subscripts within a product imply summation. The expected value of the product of random variables is their covariance, $C(h)$; here it is assumed the covariance is isotropic and only a function of the distance or lag h between the two points. For example, $E[Z_\alpha Z_\beta] = C(h_{\alpha\beta}) = \sigma_{\alpha\beta}$, where $\sigma_{\alpha\beta}$ is element (α, β) from the covariance matrix.

Parzen (1962, Section 3.3) illustrates how the expected value of the derivative of a stochastic process is the derivative of the expected value; the order of the E and ∇ operators can be switched, by assuming $m(\mathbf{x})$ is differentiable and the second derivative of the covariance exists. Doing so yields

$$\begin{aligned} \mathrm{Var}[R] = {} & \lambda_\alpha \lambda_\beta \sigma_{\alpha\beta} + \delta_\alpha \delta_\beta \sigma_{\alpha\beta}^{\hat{\nu}\hat{u}} + \sigma_{00}^2 + 2\lambda_\alpha \delta_\beta \sigma_{\alpha\beta}^{\hat{\nu}} - 2\lambda_\alpha \sigma_{\alpha 0} \\ & - 2\delta_\alpha \sigma_{0\alpha}^{\hat{u}} \end{aligned} \quad (7)$$

where $\sigma_{\alpha\beta}^{\hat{\nu}}$ and $\sigma_{\alpha\beta}^{\hat{\nu}\hat{u}}$ are the first and second directional derivatives of the covariance, respectively.

An objective function, Q, which incorporates both the minimization of the variance (7) and the unbiasedness condition (4), can be defined as (Chilés and Delfiner, 1999, p. 167)

$$Q = \mathrm{Var}[R] + 2\mu_\ell\left[\lambda_\alpha f_\alpha^\ell + \delta_\beta (f^\ell)_\beta^{\hat{\nu}} - f_0^\ell\right] \quad (8)$$

where $\mu_\ell : \ell = 0, \ldots, L$ are Lagrange multipliers.

Minimize variance of residual

To minimize Q, we take derivatives of (8) with respect to each weight and Lagrange multiplier (e.g., Chilés and Delfiner, 1999, Section 3.3.1)

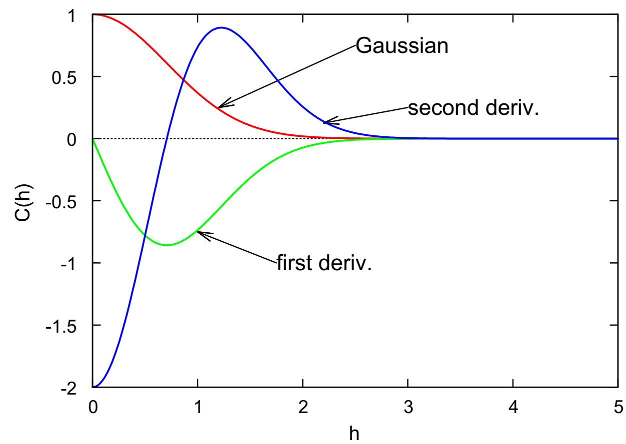

Fig. 2. Gaussian covariance model and its first two derivatives, $\sigma^2 = 1$, $a = 1$.

$$\frac{\partial Q}{\partial \lambda_\alpha} = 2\left[\lambda_\beta \sigma_{\beta\alpha} + \delta_\beta \sigma_{\beta\alpha}^{\hat{\nu}} - \sigma_{\alpha 0} + \mu_\ell f_\alpha^\ell\right] \quad \alpha = 1,\ldots,N \tag{9}$$

$$\frac{\partial Q}{\partial \delta_\beta} = 2\left[\delta_\alpha \sigma_{\alpha\beta}^{\hat{\nu}\hat{u}} - \sigma_{\beta 0}^{\hat{u}} + \lambda_\alpha \sigma_{\alpha\beta}^{\hat{u}} + \mu_\ell (f^\ell)_\beta^{\hat{u}}\right] \quad \beta = 1,\ldots,M \tag{10}$$

$$\frac{\partial Q}{\partial \mu_\ell} = 2\left[\lambda_\alpha f_\alpha^\ell + \delta_\beta (f^\ell)_\beta^{\hat{u}} - f_0^\ell\right] \quad \ell = 0,\ldots,L \tag{11}$$

The system of equations for λ_α, δ_β, and μ_ℓ is obtained by setting (9)–(11) to zero. In matrix notation (e.g., Myers, 1982) the system of universal cokriging equations are

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} & \mathbf{F} \\ \mathbf{B}^T & \mathbf{D} & \mathbf{G} \\ \mathbf{F}^T & \mathbf{G}^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\delta} \\ \boldsymbol{\mu} \end{bmatrix} = \begin{bmatrix} \boldsymbol{\sigma}_0 \\ \boldsymbol{\Delta}_0 \\ \boldsymbol{f}_0 \end{bmatrix} \tag{12}$$

where ^{T} is matrix transpose, and $_0$ indicates terms evaluated at the estimation location, $\mathbf{x}_0$. The submatrices in (12) are

$$\begin{aligned} \mathbf{A} &= \sigma_{\alpha\beta} \quad \alpha = 1,\ldots,N; \quad \beta = 1,\ldots,N \\ \mathbf{B} &= \sigma_{\alpha\beta}^{\hat{\nu}} \quad \alpha = 1,\ldots,M; \quad \beta = 1,\ldots,N \\ \mathbf{D} &= \sigma_{\alpha\beta}^{\hat{\nu}\hat{u}} \quad \alpha = 1,\ldots,M; \quad \beta = 1,\ldots,M \\ \mathbf{F} &= f_\alpha^\ell \quad \alpha = 1,\ldots,N; \quad \ell = 0,\ldots,L \\ \mathbf{G} &= (f^\ell)_\alpha^{\hat{\nu}} \quad \alpha = 1,\ldots,M; \quad \ell = 0,\ldots,L \end{aligned} \tag{13}$$

and the vectors appearing on the right-hand side of (12) are

$$\begin{aligned} \boldsymbol{\sigma}_0 &= \sigma_{\alpha 0} \quad \alpha = 1,\ldots,N \\ \boldsymbol{\Delta}_0 &= \sigma_{\alpha 0}^{\hat{\nu}} \quad \alpha = 1,\ldots,M \\ \boldsymbol{f}_0 &= f_0^\ell \quad \ell = 0,\ldots,L \end{aligned} \tag{14}$$

Next, we explore three valid covariance models that can be used in this cokriging approach.

Covariance models

The kriging equations are presented here in terms of covariance; they can be related to the equivalent variogram-form equations using the relationship

$$\gamma(h) = \sigma^2 - C(h) \tag{15}$$

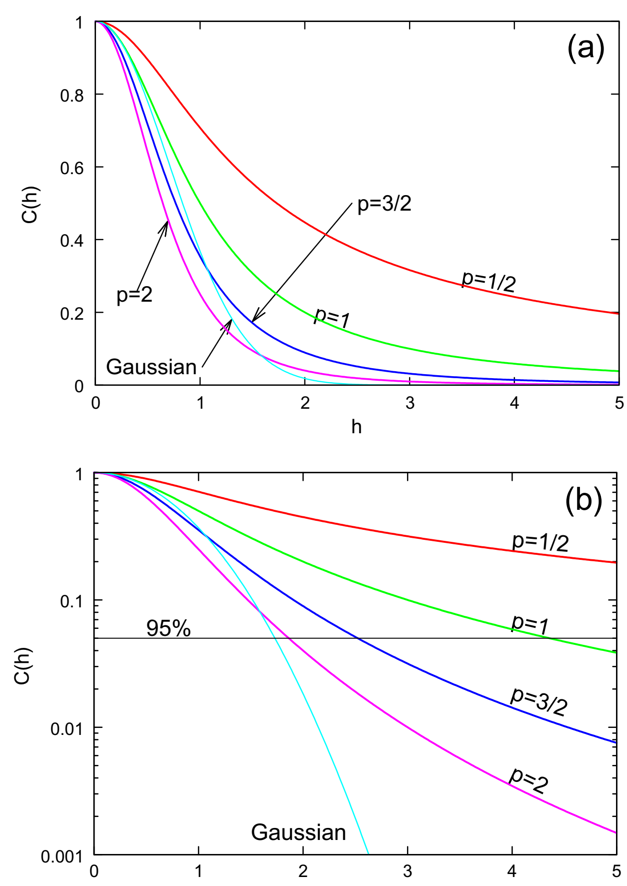

Fig. 3. Generalized Cauchy model with different exponents plotted on linear (a) and semilog (b) scale; $\sigma^2 = 1$, $a = 1$. Ninty-five percent line indicates effective range.

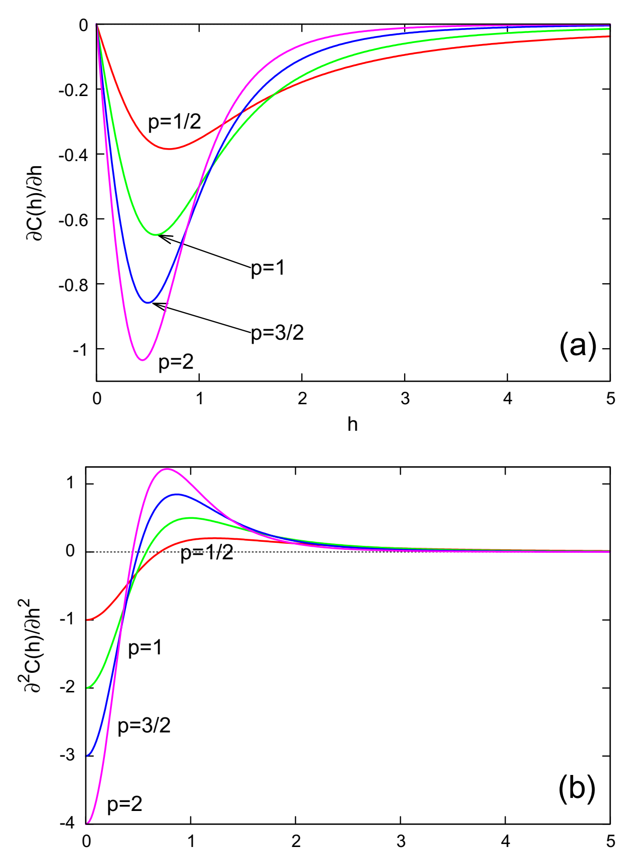

Fig. 4. First (a) and second (b) derivatives of the generalized Cauchy model with different exponents; $\sigma^2 = 1$, $a = 1$.

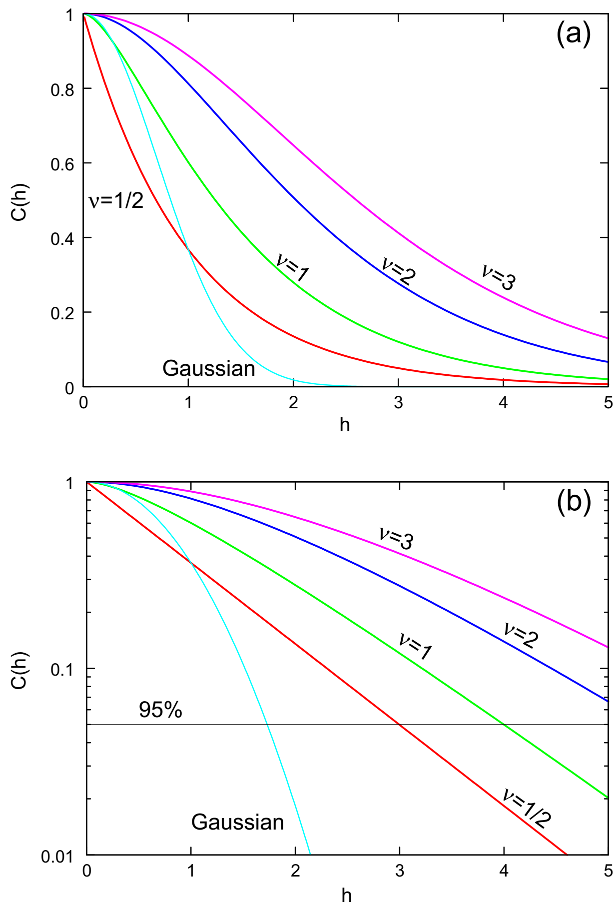

Fig. 5. Matérn model plotted with different smoothness parameters on linear (a) and semilog (b) scale; $\sigma^2 = 1$, $a = 1$. Ninty-five percent line indicates effective range.

when it applies. Here, $C(0) = \sigma^2$ is the variance when it exists (e.g., Isaaks and Srivastava, 1989, p. 171).

The limitations and requirements on a covariance model used to represent a differentiable function are discussed in detail in the literature (e.g., Philip and Kitanidis, 1989; Brochu and Marcotte, 2003; Pardo Igúzquiza and Chica Olmo, 2004, 2007); the covariance model must be twice differentiable at the origin. The commonly-used exponential, power and spherical variogram models do not satisfy this differentiability requirement.

As stated in the derivation of the estimation variance, the expected value of the gradient of a stochastic process (e.g., hydraulic heads) is equivalent to the gradient of the expected value of the same process. This allows gradient-based boundary condition information to be included through the derivative of the covariance function used for the potentials themselves. This obviates the need to estimate additional cross-covariance functions needed for cokriging, as they are found through functional relationships with the covariance modeled after the observed potentials.

Gaussian

The Gaussian covariance model is well known (e.g., Isaaks and Srivastava, 1989; Kitanidis, 1997); its directional derivatives are given in Pardo Igúzquiza and Chica Olmo (2004, Eqs. (33)–(35)), and are plotted in Fig. 2. For contouring hydraulic heads, this model is often too smooth, being infinitely differentiable at

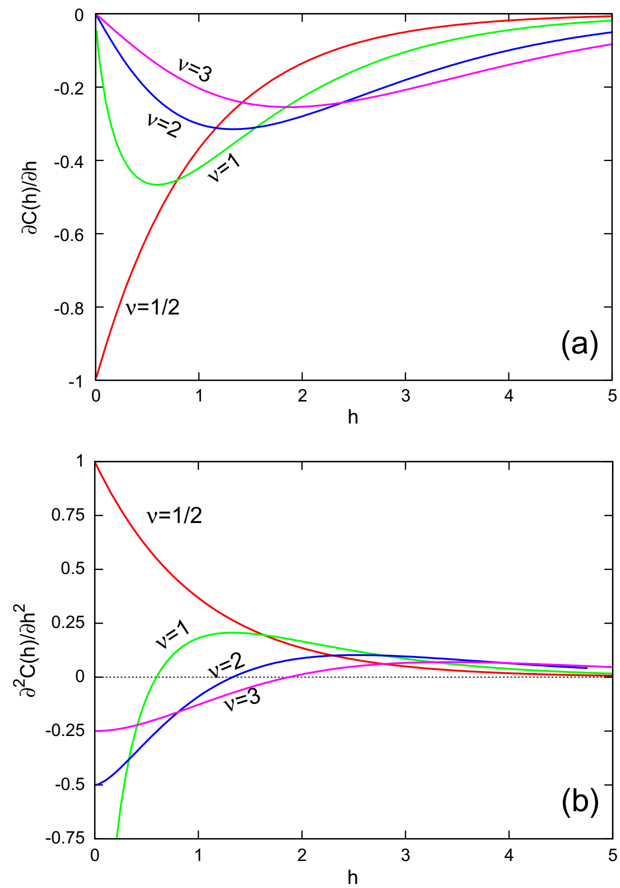

Fig. 6. First (a) and second (b) derivatives of the Matérn model with different smoothness parameters; $\sigma^2 = 1$, $a = 1$.

$h = 0$ (e.g., Schabenberger and Gotway, 2005, p. 144). The over-smoothness of the otherwise useful Gaussian model is usually treated operationally by adding a small artificial nugget (Goovaerts, 1997, p. 102).

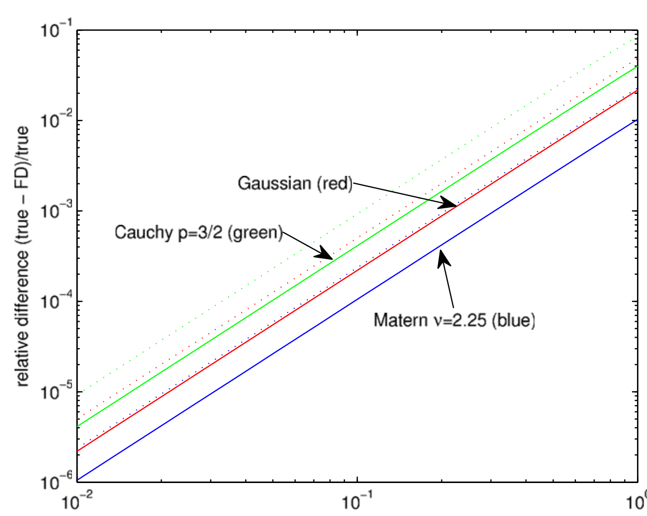

Fig. 7. Relative difference between true derivative and approximate finite difference covariance functions over a range of ρ. First derivatives are solid lines, second derivatives are dotted; $a = 6$.

Generalized Cauchy

The generalized Cauchy covariance model is used for analyzing gravity or magnetic geophysical data (e.g., Brochu and Marcotte, 2003; Chilés and Delfiner, 1999, p. 85); both fields are governed by second-order differential equations. The generalized Cauchy model is

$$\sigma(h) = \sigma^2\left(1+\frac{h^2}{a^2}\right)^{-p} \tag{16}$$

where a is the range, p is the exponent, $h^2 = \mathbf{h}\cdot\mathbf{h}$ is the squared lag, and $\mathbf{h}$ is the separation vector between two locations.

Chilés and Delfiner (1999) physically justified the choices of $p = 1/2$ for gravity fields and $p = 3/2$ for magnetic fields using statistical models derived from random sources buried at an average depth of $a/2$. The case $p = 1$ corresponds to the standard Cauchy model (see Fig. 3). The generalized Cauchy model is infinitely differentiable at the origin, but has an additional degree of freedom compared to the Gaussian model.

The first directional derivative of (16) in the direction $\hat{v}$ is

$$\sigma^{\hat{v}}(h) = -\frac{2p\sigma^2 h^{\hat{v}}}{a^2}\left(1+\frac{h^2}{a^2}\right)^{-p-1} \tag{17}$$

where $h^{\hat{v}} = \mathbf{h}\cdot\hat{v}$ is the scalar projection of the separation vector onto the boundary unit vector. The second directional derivative in the direction $\hat{u}$ is

$$\sigma^{\hat{u}\hat{v}}(h) = \frac{2p\sigma^2}{a^2}\left[\frac{2(p+1)h^{\hat{u}}h^{\hat{v}}}{a^2}\left(1+\frac{h^2}{a^2}\right)^{-p-2} - \left(1+\frac{h^2}{a^2}\right)^{-p-1}\right] \tag{18}$$

See Fig. 4 for plots of (17) and (18) for various exponents.

Matérn

The Matérn or K-Bessel covariance model is a viable model for geostatistical analysis of physically-based smooth fields (e.g.,

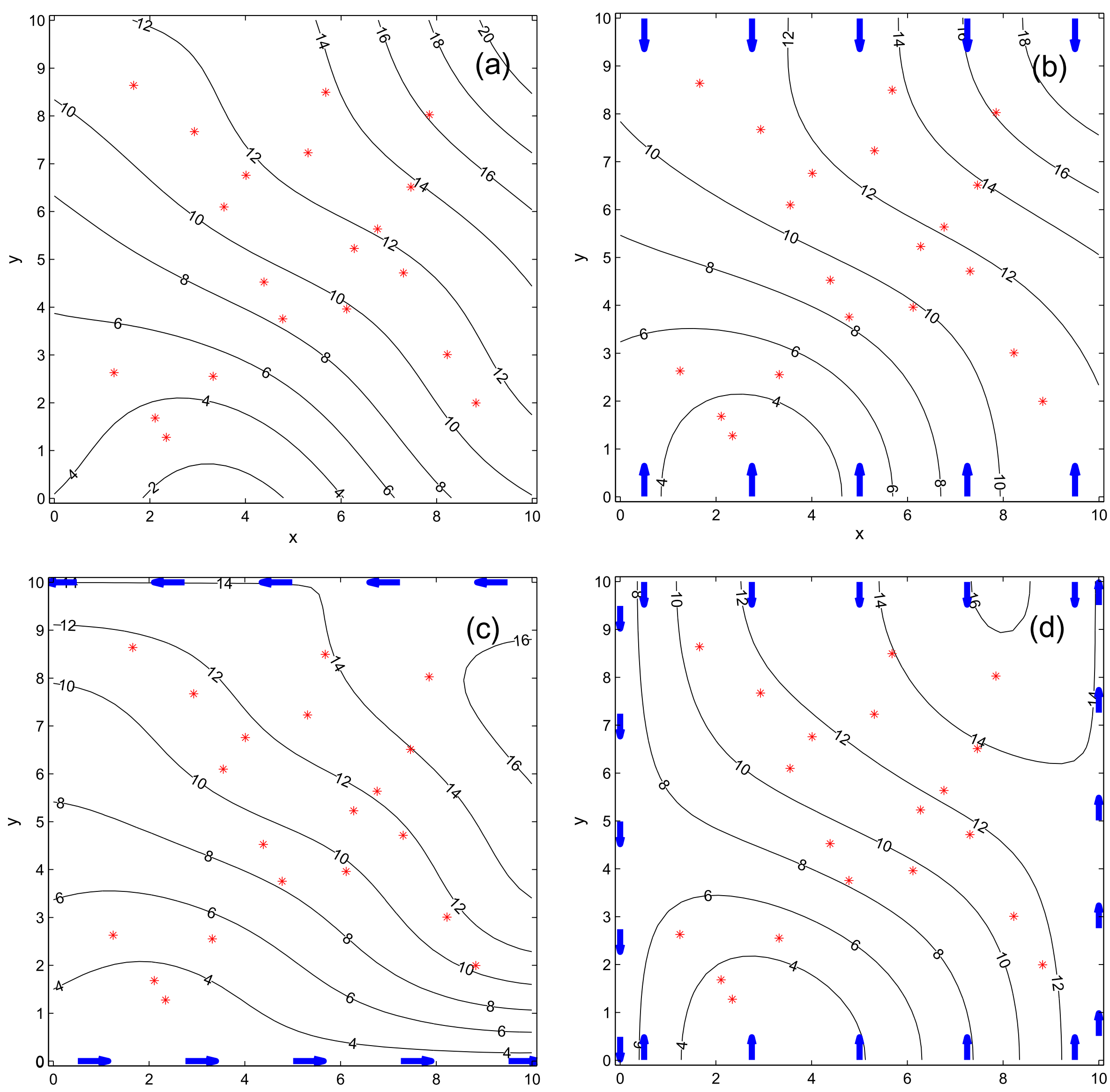

Fig. 8. Kriged head for synthetic example with head observations (red stars) and boundary conditions (blue arrows). (a) No boundaries, (b) no-flow along top and bottom, (c) constant head along top and bottom, (d) no-flow top and bottom and constant head along left and right. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Chilés and Delfiner, 1999, p. 86; Pardo Igúzquiza and Chica Olmo, 2008; Pardo Igúzquiza et al., 2009; Bras and Rodríguez Iturbe, 1993, p. 287). A benefit of using the Matérn model is the adjustable smoothness parameter; for second-order smooth fields $\nu \geqslant 2$. For example, ν could be increased or decreased to reflect the knowledge of relative aquifer homogeneity. The Matérn model is plotted in Fig. 5 and is defined as

$$\sigma(h) = \frac{\sigma^2 2^{1-\nu}}{\Gamma(\nu)}\left(\frac{h}{a}\right)^{\nu} \mathrm{K}_{\nu}\left(\frac{h}{a}\right) \tag{19}$$

where ν is the shape or smoothness parameter and $\Gamma(x)$ and $K_\nu(x)$ are the gamma and ν-order second-kind modified Bessel functions (see Abramowitz and Stegun (1964) and Andrews (1998) for properties and characteristic plots). For $\nu = \frac{1}{2}$ (19), reduces to the familiar exponential model, a straight line in Fig. 5b. As $\nu \to \infty$ the Matérn model converges towards the Gaussian model (it becomes infinitely differentiable at $h = 0$) (Schabenberger and Gotway, 2005, p. 143).

The first directional derivative of (19) in the direction $\hat{\nu}$ is

$$\sigma^{\hat{\nu}}(h) = -\hat{h}^{\hat{\nu}} \frac{\sigma^2 2^{1-\nu}}{a\Gamma(\nu)}\left(\frac{h}{a}\right)^{\nu} \mathrm{K}_{\nu-1}\left(\frac{h}{a}\right) \tag{20}$$

where $\hat{h}^{\hat{\nu}} = \hat{\nu}\cdot\mathbf{h}/h$ is the scalar product of two unit vectors. The derivative of $x^\nu \mathrm{K}_\nu(x)$ is evaluated in terms of a recurrence relationship involving different orders of the same function (McLachlan, 1955, p. 204, Eq. (216)).

The second directional derivative of the Matérn covariance function in the direction $\hat{u}$ is

$$\sigma^{\hat{\nu}\hat{u}}(h) = \hat{h}^{\hat{\nu}}\hat{h}^{\hat{u}} \frac{\sigma^2 2^{1-\nu}}{a^2\Gamma(\nu)}\left(\frac{h}{a}\right)^{\nu}\left\{\frac{a}{h}\mathrm{K}_{\nu-1}\left(\frac{h}{a}\right) - \mathrm{K}_{\nu-2}\left(\frac{h}{a}\right)\right\} \tag{21}$$

See Fig. 6 for plots of (20) and (21) for various values of smoothness parameter.

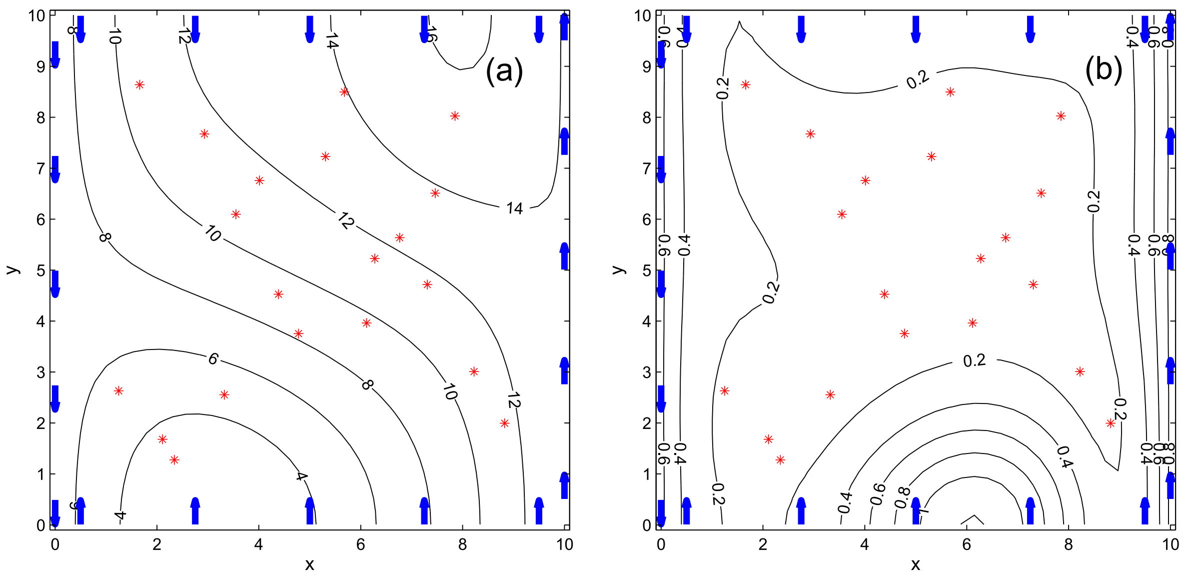

Fig. 9. Kriged head (a) and estimation error (b) for case without collocated head and boundary observations (same conditions as Fig. 8d).

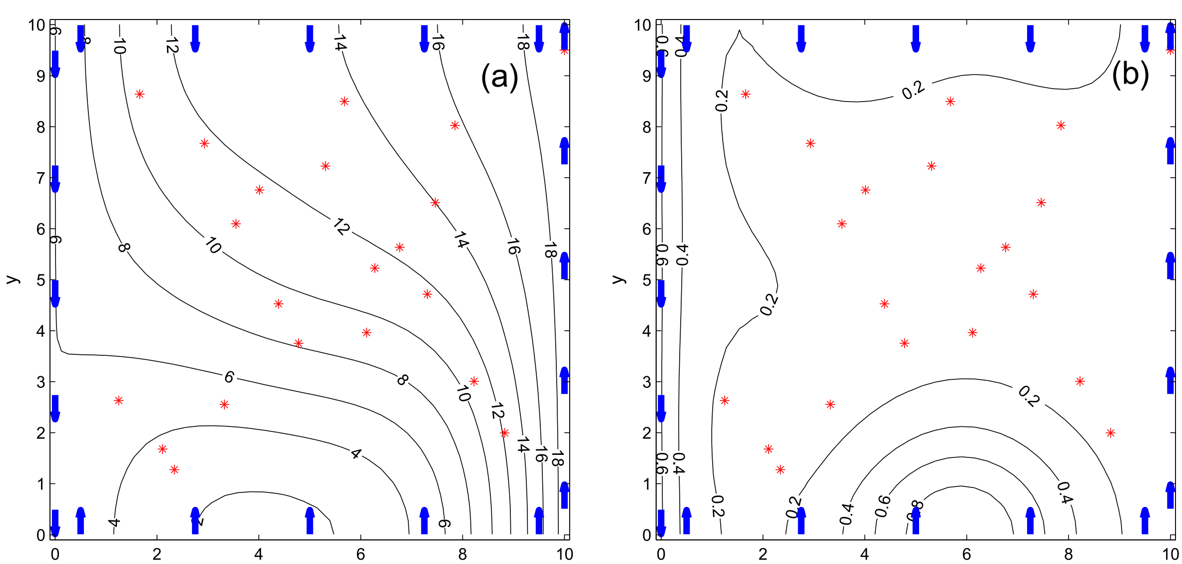

Fig. 10. Kriged head (a) and estimation error (b) for case with collocated head and flux observation at (10,9.5) – note red star and tail of blue arrow at same location. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Comparison of covariance models

Of the three sufficiently smooth covariance models discussed here, the Gaussian model is the most widely known and implemented, but both the generalized Cauchy and Matérn models have an additional degree of freedom in their formulation. The Matérn smoothness parameter potentially has more physical relevance than the generalized Cauchy exponent, although the Cauchy exponent has been related to the type of governing equation (Chilés and Delfiner, 1999).

Of these three models, the Matérn model (for $\nu \geqslant 2$) is least smooth, being at least twice differentiable, smoothness increasing with ν. Both the Gaussian and generalized Cauchy are infinitely smooth, which is at variance with what we observe in nature for these fields (Schabenberger and Gotway, 2005, p. 144). The Matérn model has a larger effective range than either the Gaussian or generalized Cauchy models (see Figs. 3b and 5b).

Finite difference approach

A simplified form of cokriging, specific to the current problem, is outlined. Rather than rigorously treating the gradient information as a secondary variable (as done in cokriging), the gradient is approximated by using a centered finite difference in place of the true derivative. Two points are placed straddling the boundary, in approximation of the boundary vectors. The approach has been discussed in Chilés and Delfiner (1999) and an example is shown in Brochu and Marcotte (2003) in terms of the dual kriging approach.

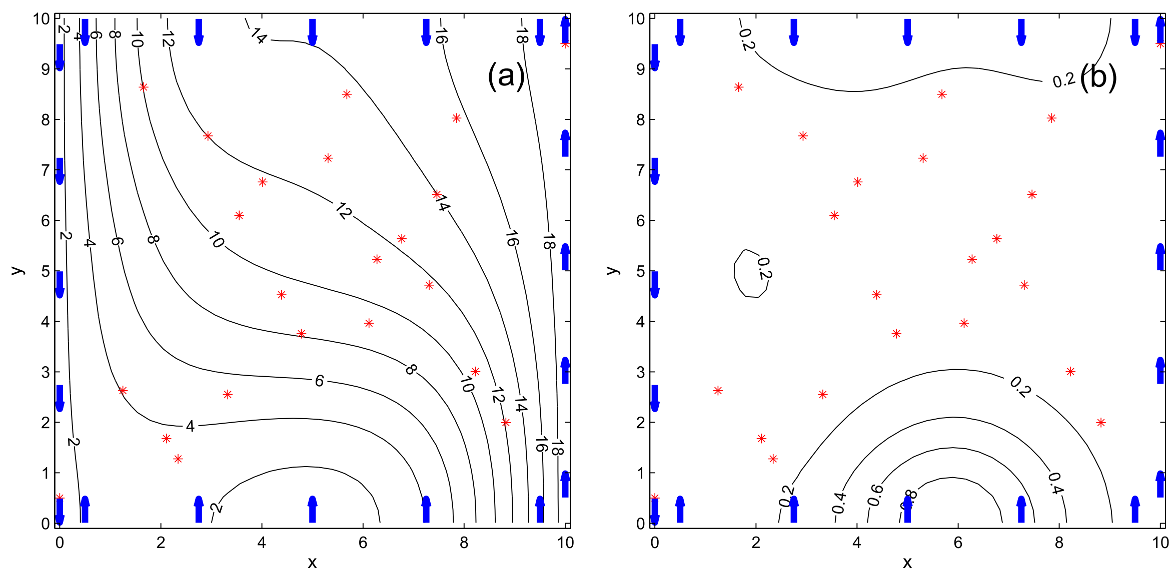

Fig. 11. Kriged head (a) and estimation error (b) for case with collocated head and flux observations at (10,9.5) and (0,0.5) – note red stars and tail of blue arrows at same locations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

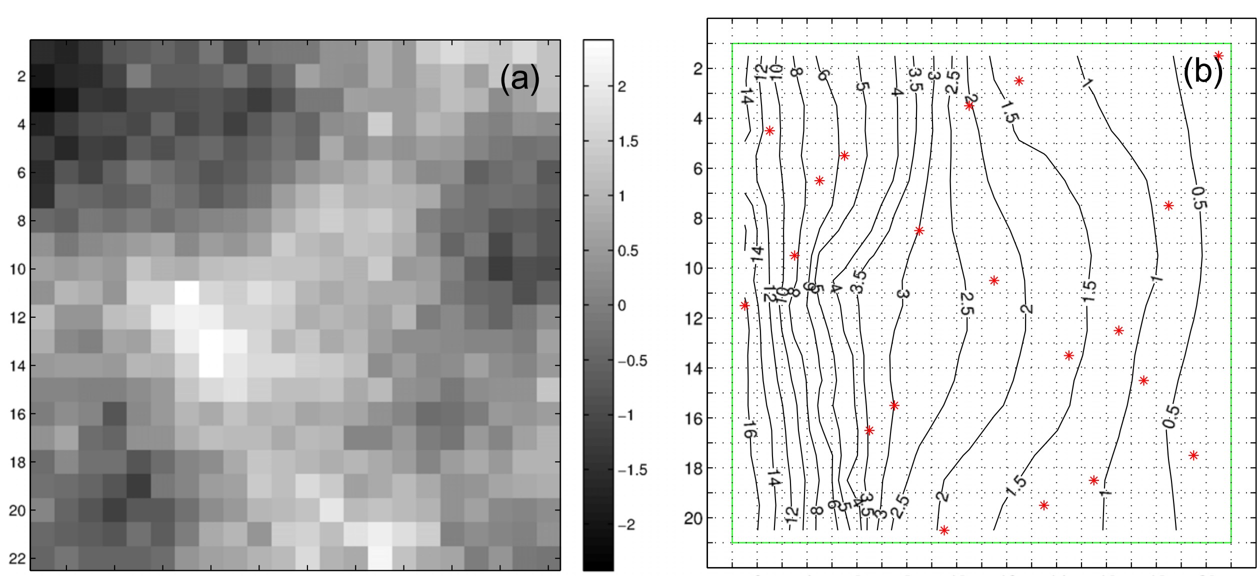

Fig. 12. Log-transmissivity (a) and modeled steady-state heads (b) for 22 × 22 flow model domain. Top and bottom rows are no-flow; left and right columns are constant head, set to 20 and 0 respectively. Sampling locations indicated with stars in (b).

Following Chilés and Delfiner (1999, Section 5.5.4), the linear estimator for the head, Z^{**}, is

$$Z_0^{**} = \lambda_\alpha Z_\alpha + \frac{\delta_\beta}{2\rho}[Z(\mathbf{x}+\rho\hat{v})_\beta - Z(\mathbf{x}-\rho\hat{v})_\beta] \tag{22}$$

where ρ is a small displacement used in the centered finite difference approximation to the true derivative. Chilés and Delfiner (1999) give (22) without the $1/2\rho$ term, as this constant does not affect the ordinary kriging estimation process. Retaining this term leads to numerically comparable kriging matrices for both approaches given here, allowing derivatives to be checked against their finite-difference equivalents.

The variance R due to (22), analogous to (6), is

$$\begin{aligned}\mathrm{Var}[R] = {} & \lambda_\alpha\lambda_\beta E[Z_\alpha Z_\beta] + \delta_\alpha\delta_\beta E[(Z_{\alpha^+} - Z_{\alpha^-})(Z_{\beta^+} - Z_{\beta^-})] \\ & + E[Z_0 Z_0] + \lambda_\alpha\delta_\beta E[Z_\alpha(Z_{\beta^+} - Z_{\beta^-})] - 2\lambda_\alpha E[Z_\alpha Z_0] \\ & - 2\delta_\alpha E[Z_0(Z_{\alpha^+} - Z_{\alpha^-})]\end{aligned} \tag{23}$$

where the shorthand $Z_{\beta^\pm} = Z(\mathbf{x}_\beta \pm \rho\hat{v})$ is used. Expanding the products in (23), and expressing the results in terms of covariance matrices, leads to

$$\begin{aligned}\mathrm{Var}[R] = {} & \lambda_\alpha\lambda_\beta\sigma_{\alpha\beta} + \delta_\alpha\delta_\beta[\sigma_{\alpha^+\beta^+} - \sigma_{\alpha^-\beta^+} - \sigma_{\alpha^+\beta^-} + \sigma_{\alpha^-\beta^-}] + \sigma^2 \\ & + 2\lambda_\alpha\delta_\beta[\sigma_{\alpha\beta^+} - \sigma_{\alpha\beta^-}] - 2\lambda_\alpha\sigma_{\alpha 0} - 2\delta_\alpha[\sigma_{\alpha^+0} - \sigma_{\alpha^-0}]\end{aligned} \tag{24}$$

In this finite-difference approach, only the difference of covariance functions appear; there are no covariance derivatives representing cross-covariance between head and its gradient.

Fig. 7 is a numerical comparison between the true gradient and finite-difference approximation forms, keeping the $1/2\rho$ term in (22). The uniform slope for all the curves is related to the accuracy of the finite-difference approximation, here $\mathcal{O}(h)$; the error for the first-order derivatives (solid lines) is less than that for the corresponding second-order ones (dotted lines). Clearly, the error associated with the finite-difference approximation is insignificant

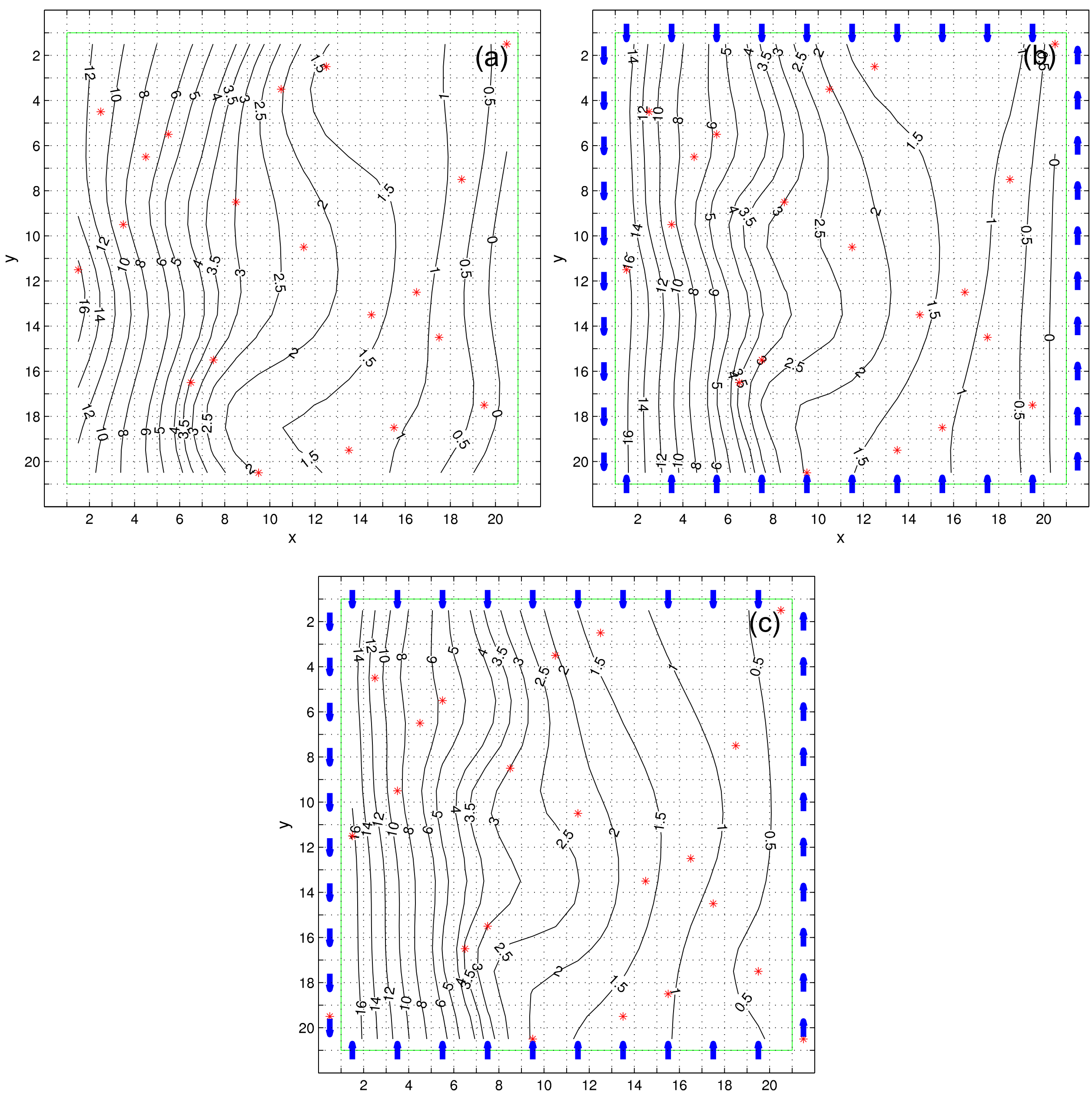

Fig. 13. Kriged head contours without boundary conditions (a), with boundary conditions and no collocated head observations, and (c) with both boundary conditions and collocated head data.

compared to the uncertainty in the entire estimation process (for reasonably small ρ).

An analogous unbiasedness condition to (4) is derived for (24), namely

$$f_0^\ell = \lambda_\alpha f_\alpha^\ell + \delta_\beta\left(f_{\beta^+}^\ell - f_{\beta^-}^\ell\right) \quad \ell = 0,\ldots,L \tag{25}$$

which does not involve derivatives of the drift functions, only function evaluations. This could be beneficial if using drift functions more complex than monomials (e.g., Brochu and Marcotte, 2003).

Examples

Non-collocated data

We give a simple synthetic example that illustrates the effects boundary conditions can have on sparse data; it is comprised of a 10×10 domain with 20 head observations picked with a Latin hypercube sampling strategy in the region $1 \leqslant x \leqslant 9$, $1 \leqslant y \leqslant 9$ (see Fig. 8). Heads at sample locations are $h = x + y + N(0,1)$, where $N(0,1)$ is an unbiased unit variance random normal deviate. A linear trend $(L = 2)$, and a Gaussian covariance model with $a = 6.0$, $\sigma^2 = 13.0$, and nugget = 0.1 was used to krige the data onto a regular 1×1 grid for contouring. A small nugget value can represent measurement error, and is often operationally included even when there is no measurement error to counteract the overly smooth nature of the Gaussian model (Goovaerts, 1997, p. 102).

Adding five no-flow (8b) or constant head (8c) observation points to represent hypothetical upper and lower boundaries clearly changes the nature of the prediction at the edges of the potential data. Imposing a no-flow condition on the top and bottom, along with an unknown constant head along the left and right (8d) gives yet another possible interpretation of the same potential data with non-collated head and boundary data.

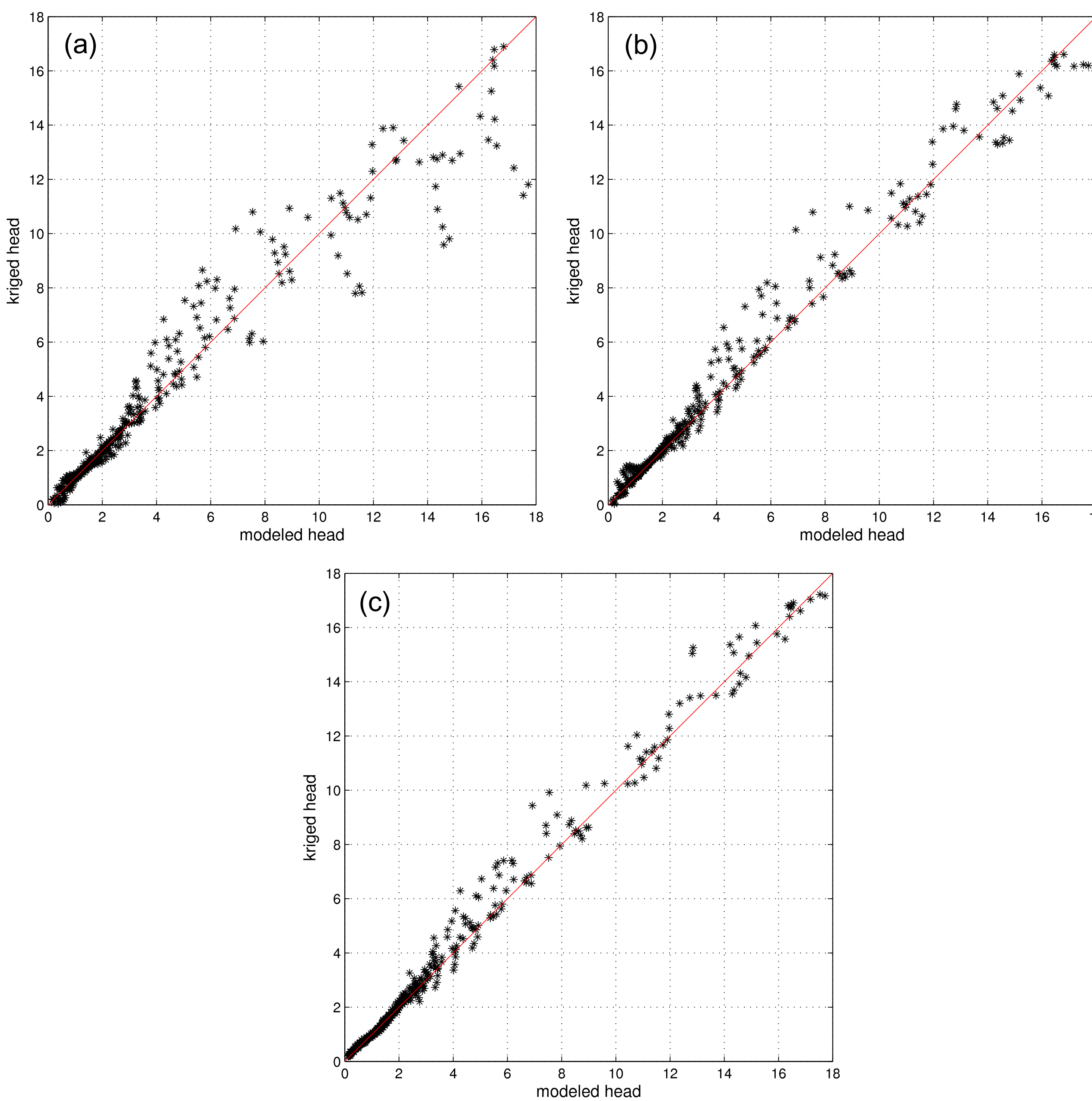

Fig. 14. Scatter plot of observed (model-generated) and kriged heads for all 400 model element centers, for the same cases given in Fig. 13.

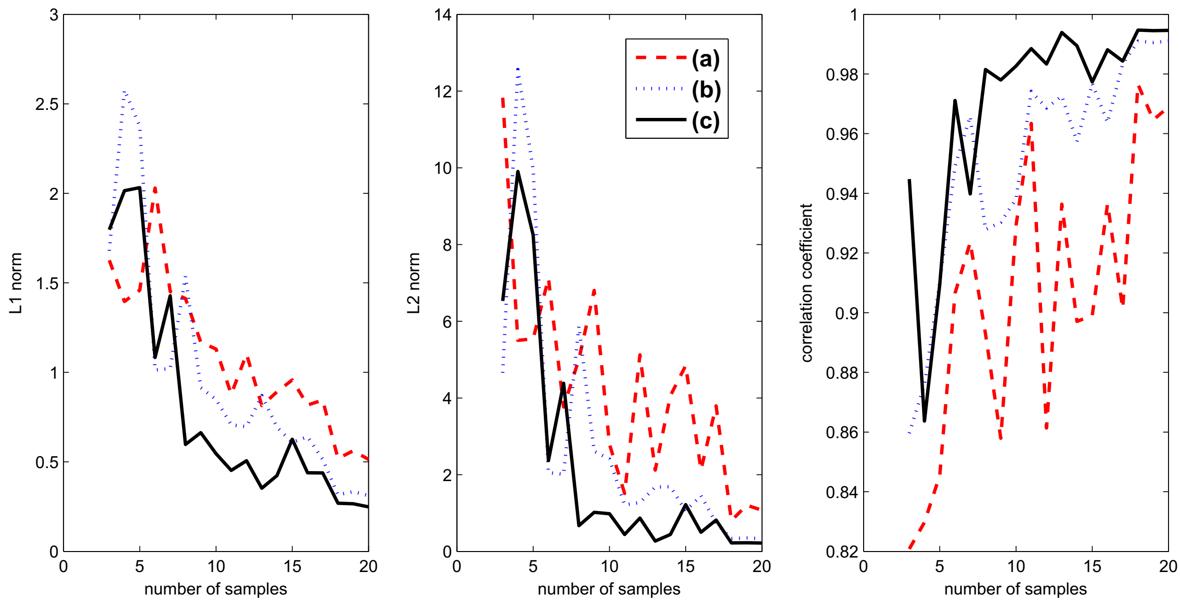

Fig. 15. L1 and L2 error norms and correlation coefficient between kriged and model-generated heads at all 400 element centers as a function of sample size, for the same three cases given in Figs. 13 and 14.

Collocated data

Here we consider the effects of collocated data i.e., both head and boundary condition observations at the same location, using the simple synthetic problem introduced above. Fig. 9 shows the effect this has on both the value (a) predicted near the boundary and the associated cokriging estimation error (b), given by *R*. The observed heads are the same used in Fig. 8d.

Including one additional head observation at $x = 10$, $y = 9.5$ along the right specified-head boundary (upper right of Fig. 10a and b) both decreases the estimation error along the entire boundary (compare Figs. 9b and 10b) and increases the head along the right constant-head boundary from approximately 13 to about 19 (comparing Figs. 9a and 10a). Additionally specifying the head along the lower left boundary has a similar effect, and the constant-head boundary is lowered from about 7 (Fig. 9a) or 5 (Fig. 10a) to less than 2 (Fig. 11a). In this simple example it is clear that kriging tends to produce results at the boundaries averaged from the whole domain; often physical boundary conditions are extreme values, driving system behavior. Collocated boundary data are required to force kriging-predicted constant-head boundary values that are greater or less than head observations interior to the domain.

Flow model comparison

In this example, a 22 × 22 confined flow model with log transmissivity plotted in Fig. 12a is used to generate steady-state heads (Fig. 12b); modeled heads are compared with the results of kriging using the no-flow (top and bottom) and constant head (head = 0 left and head = 20 right) boundary conditions, with and without collocated data.

Twenty heads are sampled from the 20 × 20 active portion of the model domain using a Latin hypercube approach, Fig. 13 shows contours of kriged heads illustrating the benefits the boundary condition data have on the kriging at the edges of the domain (especially along the top and bottom). Additionally including head observations on the constant-head boundaries improves the kriged predictions there. The MLMATERN maximum likelihood variogram estimator (Pardo Igúzquiza et al., 2009) was used to estimate Matérn variogram parameters from the 20 sampled heads; $\sigma^2 = 13.934$, $a = 2.9141$, nugget = 0, and $\nu = 2.5450$.

Scatter plots in Fig. 14 shows that all three cases illustrated in Fig. 13 do a good job estimating heads in the flat gradient on the right of the domain. On the left side of the domain, where the gradient is steeper, both non-collocated methods predict results scattered about the modeled values, with more bias in the case without boundary conditions; collocated head and boundary observations improve fits near both boundaries. Including the left constant-head boundary condition improves the estimation in areas where there are no observed heads, including the lower left corner. The constant head condition gives the contours the correct shape, but there is still a bias at the boundary; there are not enough observations in the area where steep gradients exist.

To explore effects due to sample size, the approach given in Figs. 13 and 14 is repeated for random subsamples of 3–20 observations from the original 20 samples shown in 12b (same variogram parameters used for all sample sizes). The absolute error

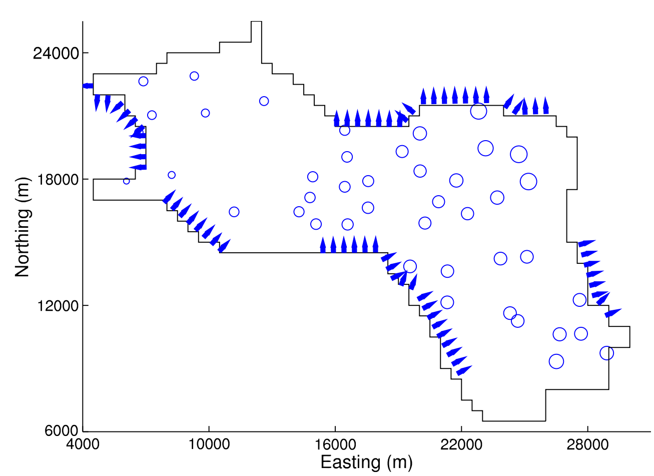

Fig. 16. Observed heads (circle size indicates relative value of head) for the Vega de Granada aquifer; no-flow boundaries of aquifer are indicated with boundary normal arrows.

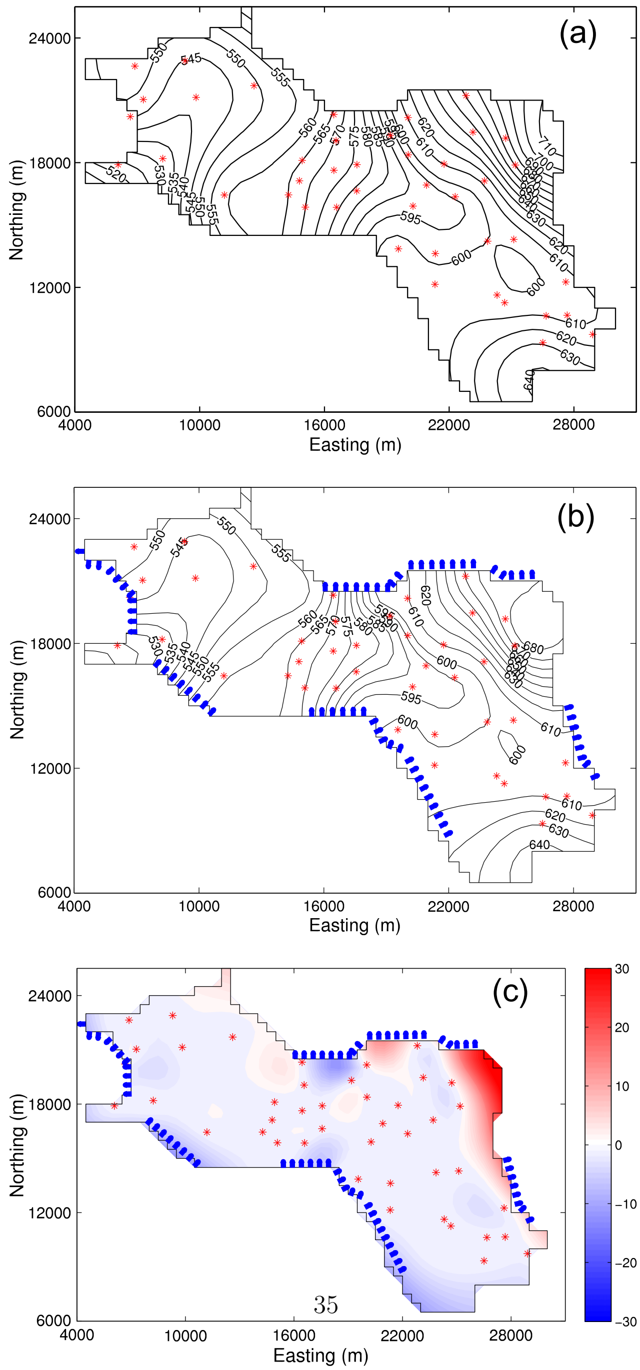

Fig. 17. Kriged head in meters without (a) and with (b) boundary condition information for the Vega de Granada aquifer; boundary vectors indicated with blue arrows; difference (without BC – with BC) given in (c) with 1 m steps in shading. (For interpretation of the references to colours in this figure legend, the reader is referred to the web version of this paper.)

norm ($L1$), mean-square error norm ($L2$) (e.g., Yeh and Liu, 2000, Eq. (14)), and correlation coefficient are used to quantify goodness of fit between the full populations of 400 kriged and modeled head values for each observation sample size used in the kriging (Fig. 15). Lower norm values and higher correlation coefficient values indicate greater similarity between the model-generated and kriged fields. Clearly, including both boundary conditions and collocated observations (case c) performs best for most sample sizes. Including boundary conditions without collocated observations (case b) performs slightly worse for most sample sizes. Kriging without boundary conditions (case a) almost always performs the worst across all sample sizes.

Case study

The study area is the Vega de Granada aquifer in southern Spain; an unconfined aquifer comprised mainly of fluvial deposits (see Castillo (1986), Luque-Espinar (2001), Pardo Igúzquiza and Chica Olmo (2004), and Pardo Igúzquiza et al. (2009) for additional information and the results of previous studies). Observed heads are illustrated in Fig. 16 in a pictogram; higher heads are larger symbols. Portions of the domain boundary known from previous geologic and hydrologic studies to correspond to no-flow conditions; they are indicated with blue arrows perpendicular to the boundary. Pardo Igúzquiza et al. (2009) give the Matérn model parameters for this dataset, fit using MLMATERN. We used a linear trend ($L = 2$) and the isotropic Matérn covariance model with the parameters $a = 2048.2$ m, $\sigma^2 = 641.18$ m^2, nugget $= 0.70357$ m^2, and $\nu = 2.1613$.

The addition of no-flow boundary condition information (Fig. 17b) improves the hydrologic quality of the kriged contour map compared to kriging with heads alone (Fig. 17a). The difference between the two methods is plotted in Fig. 17c; the difference (without boundary conditions minus with boundary conditions) ranges from 36.3 to -12.3 m across the domain shown, although nearly all large differences are confined to the perimeter of the domain. Specific effects of including the no-flow boundaries can be seen along the southern boundary, especially near the downstream outlet of the basin. In the interior of the domain, the differences between the contours for the two methods are slight.

Discussion and conclusions

We present an approach for including both no-flow and constant-head boundary conditions in the estimation of a potential field in this paper. Our approach produces nearly identical results to the finite difference approach given elsewhere. The approximate approach is easier to derive and implement, while the true derivative approach executes marginally quicker. Fewer covariance function evaluations are needed in our approach; each $\sigma_{\alpha\beta}^{\hat{\nu}\hat{u}}$ evaluation requires four covariance function calls in the finite-difference approach, and $\sigma_{\alpha\beta}^{\hat{\nu}}$ requires two covariance functions calls (see (7) and (24)). Because twice-differentiable covariance functions are already very smooth, the finite-difference approximation to the derivative is quite accurate (see Fig. 7).

Any covariance function that is at least twice differentiable at $h = 0$ can be used in this approach, but when the Gaussian model is used a nugget must often be included (Goovaerts, 1997). While the Matérn model is more complicated to implement (requiring modified Bessel function evaluations) it is the most flexible model described (Schabenberger and Gotway, 2005, p. 143). The covariance model chosen depends on the application and desired outcome. Is the goal smoothness (infinitely differentiable models like Gaussian are good for this), exact replication of observed data (zero nugget), no extraneous high and low values? (Models that are not overly-smooth like Mátern are good for this.) Our approach is general, and while illustrated in terms of heads and gradient boundary conditions, it would work for any field where a potential and a gradient-based constitutive law (here Darcy's law) govern the process (e.g., Moore, 1964). The examples illustrate the benefits of including both no-flow and constant-head boundary condition information in the kriging process. As illustrated in the examples, constant-head boundary conditions often require collocated head and boundary observations to force the kriging process to estimate heads commensurate with our hydrologic knowledge of the system.

Acknowledgements

The authors would like to thank three anonymous reviewers for their comments, which improved the examples and notation. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

References

Abramowitz, M., Stegun, I.A. (Eds.), 1964. Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables. No. 55 in Applied Mathematics Series. National Bureau of Standards.
Andrews, L.C., 1998. Special Functions of Mathematics for Engineers, second ed. SPIE Press.
Bras, R.L., Rodríguez Iturbe, I., 1993. Random Functions and Hydrology. Dover.
Brochu, Y., Marcotte, D., 2003. A simple approach to account for radial flow and boundary conditions when kriging hydraulic head fields for confined aquifers. Mathematical Geology 35 (2), 111–139.
Castillo, A., 1986. Estudio hidroquímico del acuífero de la Vega de Granada. Ph.D. Thesis, University of Granada, Spain.
Chilés, J.-P., Delfiner, P., 1999. Geostatistics: Modeling Spatial Uncertainty. Wiley Interscience.
Goovaerts, P., 1997. Geostatistics for Natural Resource Evaluation. Oxford University Press.
Isaaks, E.H., Srivastava, R.M., 1989. An Introduction to Applied Geostatistics. Oxford University Press.
Kitanidis, P.K., 1997. Introduction to Geostatistics: Applications in Hydrogeology. Cambridge University Press.
Luque-Espinar, J., 2001. Analisis geoestadistico espacio-temporal de la variabilidad piezometrica. aplicacion a la Vega de Granada. Ph.D. Thesis, University of Granada, Spain.
McLachlan, N.W., 1955. Bessel Functions for Engineers, second ed. Oxford University Press.
Moore, R.K., 1964. Wave and Diffusion Analogies. McGraw-Hill.
Myers, D.E., 1982. Matrix formulation of co-kriging. Mathematical Geology 14 (3), 249–257.
Pardo Igúzquiza, E., Chica Olmo, M., 2004. Estimation of gradients from sparse data by universal kriging. Water Resources Research 40, W12418.
Pardo Igúzquiza, E., Chica Olmo, M., 2007. KRIGRADI: a cokriging program fro estimating the gradient of spatial variables from sparse data. Computers and Geosciences 33, 497–512.
Pardo Igúzquiza, E., Chica Olmo, M., 2008. Geostatistics with the Matérn semivariogram model: a library of computer programs for inference, kriging and simulation. Computers and Geosciences 34, 1073–1079.
Pardo Igúzquiza, E., Mardia, K.V., Chica Olmo, M., 2009. MLMATERN: a computer program for maximum likelihood inference with the spatial Matérn covariance model. Computers and Geoscience 35, 1139–1150.
Parzen, E., 1962. Stochastic Processes. Holden-Day.
Philip, R.D., Kitanidis, P.K., 1989. Geostatistical estimation of hydraulic head gradients. Ground Water 27 (6), 855–865.
Schabenberger, O., Gotway, C.A., 2005. Statistical Methods for Spatial Data Analysis. Texts in Statistical Science. Chapman & Hall.
Yeh, T.-C.J., Liu, S., 2000. Hydraulic tomography: development of a new aquifer test method. Water Resources Research 36 (8), 2095–2105.