Efficient computational methods for iterative cokriging

Wolfgang Nowak [1] and Olaf A. Cirpka

Institut für Wasserbau, Universität Stuttgart, Pfaffenwaldring 61, 70569 Stuttgart, Germany. Phone: +49 (711) 685-8220. Fax: +49 (711) 685-7020. email: wolfgang.nowak@iws.uni-stuttgart.de

Abstract

Cokriging is a powerful tool for geostatistical parameter identification. The unknown parameter field, e.g. hydraulic conductivity, is considered a stationary random space function, which then is conditioned on observations of dependent quantities, such as hydraulic head and the arrival time of conservative tracers. Discretizing the parameter field by many element-, cell- or point-related values, the underlying problem is underdetermined since only few measurements are available. The problem of under-determination is overcome by the introduction of a priori knowledge, allowing a rigorous uncertainty analysis in the Bayesian framework. Cokriging is, however, often restricted by its computational costs. We show how to increase the computational efficiency of iterative cokriging by using a combination of both well-known and newly developed mathematical methods.

Introduction

Cokriging is a geostatistically based technique to identify an unknown spatial parameter field given observations of a correlated quantity. Originally developed for mining exploration, cokriging has successfully been used as tool for inverse modeling in other fields such as hydrogeology (Kitanidis 1983, Harvey & Gorelick 1995, Sun 1994). In geostatistical inverse modeling, we consider the unknown parameters $\mathbf{s}$, such as the hydraulic conductivity field of a porous formation, as a random space function which is conditioned on observations $\mathbf{y}$ of dependent quantities, such as the hydraulic head or the travel time of a solute (Cirpka & Kitanidis 2001). The dependency of the observations on the unknown parameters, called the transfer function, is in most hydrological applications a *pde*, such as the groundwater flow equation or the transport equation.

It has been shown, that universal cokriging is identical to a Bayesian analysis for the conditional mean of the unknowns with diffuse prior information about the mean (Kitanidis 1986). The rigorous Bayesian context allows an accurate quantification of the parameter uncertainty while imposing a minimum of structural assumptions onto the unknowns.

Rather than considering a continuous space function, the unknowns are discretized on a computational grid, with n grid cells corresponding to n unknown values. The resolution for this grid is in most applications given by numerical groundwater flow and transport models, which are, e.g., discretized by 10^4 -

10^5 cells, or by at least 10^6 cells for well resolved 3-D computations. The number m of observations, by contrast, is typically at the order of 10^1 to 10^2.

A disadvantage of traditional cokriging techniques lies in the computational costs, increasing with the square or cube of the number of unknown values n, depending on the algorithms used. Conventional implementations of iterative cokriging have been restricted to a certain maximum problem size for reasons of computational costs or storage requirements (Zimmerman et al. 1998). Among other reasons, this has led to the development of alternative geostatistical methods of inversing. These alternative methods avoid the computationally most expensive steps or execute them only partially (RamaRao et al. 1995, Gomez-Hernandez et al. 1997, Zimmerman et al. 1998) at the cost of sacrificing the rigor in determining the parameter uncertainty.

In our present work, we have collected mathematical tools that reduce the computational costs of iterative cokriging to make it competitive with these alternative methods. This paper is structured as follows: in the next section, we give a short derivation of iterative cokriging. After that, we discuss which steps are the computationally most expensive ones and present the methods that are apt to reduce these costs. Finally, we present a test case of iterative cokriging using a combination of all methods available to demonstrate the effectivity of these methods.

Basic Equations

Consider the unknown values of the hydraulic log-conductivity a $n \times 1$ random vector $\mathbf{s}$ distributed multi-Gaussian with mean $\mathbf{X}\boldsymbol{\beta}$ and covariance $\mathbf{Q}$, $\mathbf{s} \sim \mathbf{N}(\mathbf{X}\boldsymbol{\beta}, \mathbf{Q})$. The covariance matrix $\mathbf{Q}$ is sized $n \times n$, $\mathbf{X}$ is a $n \times p$ matrix of known base functions, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown drift coefficients. The transfer function that relates the observed hydraulic heads and arrival times to the conductivity field is denoted by:

$$\mathbf{y} \;=\; \mathbf{f}(\mathbf{s}) + \mathbf{r}\,, \tag{1}$$

in which $\mathbf{y}$ is the $m \times 1$ vector of observations, $\mathbf{H}$ is the sensitivity matrix sized $m \times n$, and $\mathbf{r}$ is a $m \times 1$ vector of measurement errors, so that $\mathbf{y}$ is distributed $\mathbf{y} \sim \mathbf{N}(\mathbf{f}(\mathbf{s}), \mathbf{R})$ for a given value of $\mathbf{s}$. Since, in the classical case of estimating conductivity from heads and arrival times, the transfer function is non-linear, $\mathbf{f}(\mathbf{s})$ is linearized:

$$\mathbf{f}(\mathbf{s}) \;\approx\; \mathbf{f}(\mathbf{s}_k) + \mathbf{H}_k(\mathbf{s} - \mathbf{s}_k)\,, \tag{2}$$

in which $\mathbf{H}_k$ is the $m \times n$ sensitivity matrix evaluated at $\mathbf{s}_k$. Then, a modified vector of observations is introduced

$$\mathbf{y}'_k = \mathbf{y} - \mathbf{f}(\mathbf{s}_k) + \mathbf{H}_k\mathbf{s}_k\,, \tag{3}$$

so that $\mathbf{y}'_k \sim \mathbf{N}(\mathbf{H}_k\mathbf{s}_k, \mathbf{R})$. In the following, the index k will be omitted unless necessary.

Following Bayes theorem, the prior distribution $\mathbf{s} \sim \mathbf{N}(\mathbf{X}\boldsymbol{\beta}, \mathbf{Q})$ and the likelihood of the measurements $\mathbf{y}' \sim \mathbf{N}(\mathbf{H}\mathbf{s}, \mathbf{R})$ are combined to yield the posterior distribution of $\mathbf{s}$ given the measurements,

specified by $\mathbf{s} \sim \mathbf{N}\left(\hat{\mathbf{s}}, \hat{\mathbf{Q}}\right)$. The cokriging estimator is identical to the mean $\hat{\mathbf{s}}$ of the posterior distribution, and can be found by minimizing the negative logarithm of the corresponding Gaussian probability density function:

$$L \;=\; \frac{1}{2}\left(\mathbf{s}-\mathbf{X}\boldsymbol{\beta}\right)^T \mathbf{Q}^{-1}\left(\mathbf{s}-\mathbf{X}\boldsymbol{\beta}\right) + \frac{1}{2}\left(\mathbf{y}'-\mathbf{H}\mathbf{s}\right)^T \mathbf{R}^{-1}\left(\mathbf{y}'-\mathbf{H}\mathbf{s}\right)\,, \tag{4}$$

hereafter referred to as the objective function. The resulting estimator is (Kitanidis 1996):

$$\mathbf{s}_{k+1} \;=\; \mathbf{X}\hat{\boldsymbol{\beta}}_{k+1} + \mathbf{Q}\mathbf{H}_k^T\boldsymbol{\xi}_{k+1}\,, \tag{5}$$

in which $\boldsymbol{\xi}_{k+1}$ and $\hat{\boldsymbol{\beta}}_{k+1}$ are parameters obtained by solving the cokriging system:

$$\left[\begin{array}{cc} \mathbf{H}_k\mathbf{Q}\mathbf{H}_k^T+\mathbf{R} & \mathbf{H}_k\mathbf{X} \\ \mathbf{X}^T\mathbf{H}_k^T & \mathbf{0} \end{array}\right]\left[\begin{array}{c} \boldsymbol{\xi}_k \\ \hat{\boldsymbol{\beta}}_k \end{array}\right] = \left[\begin{array}{c} \mathbf{y}'_k \\ \mathbf{0} \end{array}\right]. \tag{6}$$

Since for non-linear problems $\mathbf{H}_k$ depends on $\mathbf{s}_k$, eq. (5) and eq. (6) are applied repeatedly. The uncertainty of the estimate is specified by the posterior covariance $\hat{\mathbf{Q}}$ (Kitanidis 1996):

$$\hat{\mathbf{Q}} \;\approx\; \mathbf{Q} - \left[\begin{array}{c} \mathbf{H}^T\mathbf{Q} \\ \mathbf{X}^T \end{array}\right]^T \left[\begin{array}{cc} \mathbf{H}\mathbf{Q}\mathbf{H}^T+\mathbf{R} & \mathbf{H}\mathbf{X} \\ \mathbf{X}^T\mathbf{H}^T & \mathbf{0} \end{array}\right]\left[\begin{array}{c} \mathbf{H}^T\mathbf{Q} \\ \mathbf{X}^T \end{array}\right], \tag{7}$$

which is exact for linear problems.

The greatest contributions to the computational costs of cokriging arise from a few basic operations. (1) computing the sensitivity matrix $\mathbf{H}$, (2) evaluating the matrix products $\mathbf{Q}\mathbf{H}^T$ and $\mathbf{H}\mathbf{Q}\mathbf{H}^T$, (3) evaluating the objective function and (4) repeating these steps within an iterative algorithm. In the next section, we discuss methods for each of these operations that drastically reduce their computational costs.

Efficient algorithms for cokriging

Computing the sensitivity matrix. To numerically compute the derivative $\mathbf{H} = \frac{\partial \mathbf{f}(\mathbf{s})}{\partial \mathbf{s}}$ about a current estimate $\mathbf{s}_k$, takes one evaluation of $\mathbf{f}\left(\mathbf{s}_k\right)$ and n evaluations of $\mathbf{f}\left(\mathbf{s}+\delta\mathbf{s}_i\right)$, $i = 1 \ldots n$. Since one evaluation of $\mathbf{f}\left(\mathbf{s}\right)$ has costs up to the order of n^2 if a standard solver is employed, the overall computational costs may sum up to $\mathcal{O}\left(n^3\right)$. Sun & Yeh (1990) have demonstrated that the adjoint state method allows to obtain $\mathbf{H}$ by solving $\mathcal{O}\left(m\right)$ problems that are structurally identical to $\mathbf{f}\left(\mathbf{s}\right)$. Using an algebraic multigrid solver further reduces the costs of evaluating $\mathbf{f}\left(\mathbf{s}\right)$ to $\mathcal{O}\left(n \log n\right)$. Then, the overall costs for computing $\mathbf{H}$ are merely $\mathcal{O}\left(mn \log n\right)$. The full mathematical background and derivation is covered by Sun (1994).

Matrix Products. Using standard methods, evaluating the matrix-matrix product $\mathbf{Q}\mathbf{H}^T$ ($\mathbf{H}$ sized $m \times n$ and $\mathbf{Q}$ sized $n \times n$, $n \gg m$) requires storage $\mathcal{O}\left(n^2\right)$ and computations $\mathcal{O}\left(mn^2\right)$. For large problems, e.g. $n = 10^6$, storage of $\mathbf{Q}$ takes 8.000 GByte. However, if the unknowns are a stationary Gaussian random vector, $\mathbf{Q}$ is a symmetric Toeplitz (ST) matrix or a symmetric block Toeplitz matrix with Toeplitz blocks (STT), depending on the dimensionality of the underlying problem (Zimmerman 1989). ST and STT

matrices have a strict structure that can be exploited to store $\mathbf{Q}$ with only $\mathcal{O}(n)$ entries. Further, ST and STT matrices can be embedded in larger circulant matrices, which then allows to apply FFT-based methods that reduce the computations for $\mathbf{QH}^T$ and $\mathbf{HQH}^T$ to $\mathcal{O}(mn \log_2 n)$ (Nowak et al. 2003, van Loan 1992).

Evaluating the objective function. Evaluating the objective function consists of two parts: evaluating the a priori term

$$L_p = \frac{1}{2} (\mathbf{s} - \mathbf{X}\boldsymbol{\beta})^T \mathbf{Q}^{-1} (\mathbf{s} - \mathbf{X}\boldsymbol{\beta}) , \tag{8}$$

and the likelihood term before linearization

$$L_\ell = \frac{1}{2} (\mathbf{y} - \mathbf{f}(\mathbf{s}))^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{f}(\mathbf{s})) . \tag{9}$$

The latter has computational costs $\mathcal{O}(n \log n)$ for evaluating $\mathbf{f}(\mathbf{s})$. The a priori term in the form of eq. (8) requires explicit storage of $\mathbf{Q}$, again resulting in extreme memory consumption. However, it can be shown that the a priori term can be simplified to (Nowak & Cirpka 2003):

$$L_p = \frac{1}{2} \boldsymbol{\xi}^T \mathbf{HQH}^T \boldsymbol{\xi} , \tag{10}$$

which minimizes storage requirements and computational costs since $\mathbf{HQH}^T$ has already been computed in previous steps. Evaluating the quadratic form with $\boldsymbol{\xi}$ merely takes $\mathcal{O}(m^2)$ floating point operations.

Iteration algorithm. The standard algorithm used for iterative cokriging is a procedure formally similar to the well-known Gauss-Newton algorithm (Kitanidis 1995, McLaughlin & Townley 1996). Some differences arise from the fact that cokriging is the solution of an underdetermined problem. Like the Gauss-Newton algorithm (GN), the standard iterative cokriging algorithm (IC) is a powerful tool for quasi-linear problems, but fails to converge for problems with higher degrees of non-linearity.

Additionally, the IC has other drawbacks based on the underdetermined character of the underlying problem. The cokriging estimate is defined by the value of $\mathbf{s}$ that minimizes the sum of the a priori term (eq. 8) and the likelihood term (eq. 9). The graphical interpretation of the objective function is, that its minimum is both very smooth in the sense the prior distribution of $\mathbf{s}$, and at the same time honors the observations as specified by the likelihood of the measurements. In many cases where the degree of non-linearity is sufficiently high to cause the existence of local minima of the objective function, the IC algorithm converges to solutions that fail to comply with the smoothness condition (Nowak & Cirpka 2003).

The Levenberg-Marquardt algorithm (LM) is a popular derivative of the GN that enjoys a great popularity due to its robustness for non-linear problems in all fields of engineering. However, since IC differs from GN in some aspects, LM has to be modified on a basic level to be adapted for cokriging. In an ongoing study, the authors of this paper have developed a modified LM algorithm for iterative cokriging (ICLM) (Nowak & Cirpka 2003). The resulting algorithm is:

Algorithm 1 *(Levenberg-Marquardt algorithm for iterative cokriging, ICLM): The problem statement is as specified in section two. Error analysis of the linearization yields that the error of linearization is acceptable for* $|\mathbf{s} - \mathbf{s}_k| < \Delta\mathbf{s}_1$*, and negligible for* $|\mathbf{s} - \mathbf{s}_k| < \Delta\mathbf{s}_2$*. Define an initial guess* $\mathbf{s}_0 = \mathbf{X}\boldsymbol{\beta}_0$ *and initialize* λ *with* $\lambda > 0$*.*

1. *compute* $\tilde{\mathbf{H}}_k$ *unless the error of linearization is negligible.*

2. *Find* $\mathbf{s}_{k+1}$ *by solving the following equations:*

$$\mathbf{s}_{k+1} = \mathbf{X}\left(\hat{\boldsymbol{\beta}}_{rep} + \hat{\boldsymbol{\beta}}_{in}\right) + \tilde{\mathbf{Q}}_{\mathbf{sy}}\left(\boldsymbol{\xi}_{rep} + \boldsymbol{\xi}_{in}\right) \tag{11}$$

$$\begin{bmatrix} \tilde{\mathbf{Q}}_{\mathbf{yy},k} + \lambda\mathbf{R} & \tilde{\mathbf{H}}_k\mathbf{X} \\ \mathbf{X}^T\tilde{\mathbf{H}}_k^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi}_{in} \\ \hat{\boldsymbol{\beta}}_{in} \end{bmatrix} = (1+\lambda)^{-\alpha} \begin{bmatrix} \mathbf{y} - \mathbf{f}\left(\mathbf{s}_{k,l}\right) \\ \mathbf{0} \end{bmatrix} \tag{12}$$

$$\begin{bmatrix} \tilde{\mathbf{Q}}_{\mathbf{yy},k} - \tau\mathbf{R} & \tilde{\mathbf{H}}_k\mathbf{X} \\ \mathbf{X}^T\tilde{\mathbf{H}}_k^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi}_{rep} \\ \hat{\boldsymbol{\beta}}_{rep} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{H}}_k\mathbf{s}_k \\ \mathbf{0} \end{bmatrix} \tag{13}$$

$$\tau = 1 - (\lambda + 1)^{-\gamma} \tag{14}$$

   If $|\mathbf{s}_{k+1} - \mathbf{s}_k| >= \Delta\mathbf{s}_1$ *or if the objective function does not improve, increase* λ *and repeat step 2. Otherwise decrease* λ*.*

3. *Increase* k *by one and repeat until convergence.*

The parameters γ and α are chosen depending on the non-linearity of the problem. Larger values of γ and small values of $\alpha > 0$ ensure robustness and fast convergence. Starting at the prior mean, the ICLM algorithm screens the solution space with a preference to stay close to the prior mean by artificially increasing the measurement error whenever convergence is poor. This reduces the risk of finding solutions that do not have the required smoothness. By suppressing oscillations and excessive step sizes that frequently occur in GN, the ICLM takes less iteration steps for quasi-linear problems and allows for larger degrees of non-linearity through its increased robustness.

Case study

To test the computational efficiency of iterative cokriging when applying this combination of methods, we demonstrate for a test case that computing time even on a desktop PC purchased in 2001 is well acceptable. We seek the unknown log-conductivity distribution for a groundwater and transport model with a resolution of 200 times 200 cells, setting up an average-sized problem with $n = 40,000$ unknowns. The model parameters are displayed in table 1, and details about the test computer are listed in table 2. The derivation of the adjoint states and sensitivities for measurements of hydraulic head and arrival, as well as the boundary conditions chosen are described in detail by Cirpka & Kitanidis (2001).

parameter	units	value	parameter	units	value
mean conductivity $\bar{K}$	$\frac{m^2}{s}$	$\exp(-10)$	No. of wells (head h)	-	20
correlation length λ_x	m	10	No. of wells (t_{50})	-	6
correlation length λ_y	m	5	error σ_h^2 for h	m^2	0.002
$\log K$ variance σ^2	-	1	error σ_t^2 for t_{50}	s^2	5%
domain length l_x	m	50	porosity	-	0.3
domain length l_y	m	25	dispersivity a_l	m	0.01
grid spacing d_x	m	0.2	dispersivity α_t	m	0.001
grid spacing d_y	m	0.1	diffusion D_m	$\frac{m^2}{s}$	10^{-9}

Table 1: model parameters used

CPU type and speed, RAM size	1.2 GHz AMD Athlon, 1.5GByte RAM
Environment	MatlqabR13 under Linux (Suse8)
Buildt-in FFT algorithm	FFTW (Frigo & Johnson 1998)
Solver plug-in	AMG(Ruge & Stueben 1986)
Flow implementation type	Standard Galerkin FEM, bilinear rectangular elements
Transport implementation type	SUPG FEM (Brooks & Hughes 1982), bilinear rectangular elements

Table 2: test conditions used

We generated an unconditional realization of log conductivity using the method by Dietrich & Newsam (1997) and evaluated hydraulic heads and arrival times of a conservative tracer. Values at selected locations plus white noise were used as an artificial measurement data set. After that, we 'forgot' the unconditional realization and used iterative cokriging to determine the unknown conductivity field from the artificial data.

The unconditional realization of $\log K$ together with the head and arrival time distribution and measurement locations is displayed in Figure 1, (a) through (c). The hydraulic boundary conditions were fixed-head at the left and the right-hand side and no-flow conditions on top and on bottom. The transport boundary conditions were a Dirichlet pulse of the conservative tracer at $t = 0$ a the left boundary, zero flux on top and on bottom, and the right boundary was closed for dispersive flux.

Figures 1 (d) through (f) show the result of iterative cokriging, and the corresponding head and arrival time distributions. The solution was obtained after 1 hour and 45 minutes CPU time with six iteration steps. No oscillations occurred during the iteration, although the problem is considered to be well out of the quasi-linear range of $\sigma^2 = 0.25$. The minimized value of the objective function is $L_p = 24.7$ for the prior term and $L_\ell = 39.2$ for the likelihood term. Statistical tests based on the χ^2-distribution with $m - 1$ degrees of freedom for each term indicate that this is an acceptable solution on a confidence level of 90%.

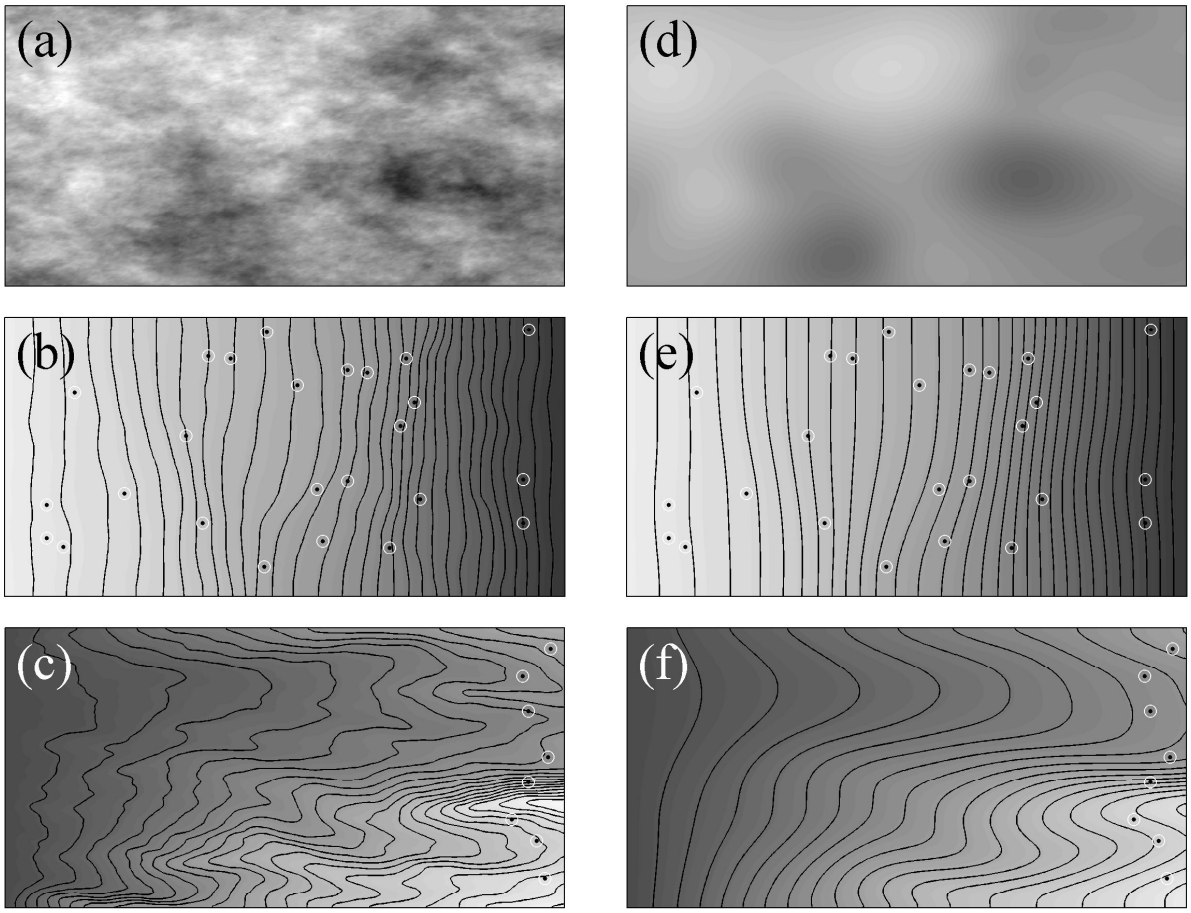

Figure 1: (a): Unconditional realization of $\log K$. (b): Hydraulic heads. (c): arrival times. (d): Cokriged $\log K$ field. (e): Hydraulic heads. (f): Arrival times. Lighter colors correspond to higher values. identical color scales apply for (a)/(d), (b)/(e), and (c)/(f), respectively.

Conclusions

We have presented a collection of methods that drastically decrease the computational costs of iterative cokriging. In a test case, we demonstrated the efficiency of these methods. Iterative cokriging was used for geostatistical inversion with $n = 40.000$ unknown values of hydraulic log-conductivity given 25 measurements of hydraulic head and 6 measurements of arrival time of a conservative tracer. The solution was obtained in 1 hour 45 minutes on a desktop PC purchased in 2001. This test case indicates that, if using highly efficient algorithms whenever possible, the computational costs of iterative cokriging can be reduced to a well acceptable level, making it competitive to alternative methods of geostatistical inversion that were developed to reduce computational costs.

Acknowledgments

This study has been funded by the Emmy Noether program of the Deutsche Forschungsgemeinschaft under the grant Ci 26/3-3.

References

Brooks, A. & Hughes, T. (1982), 'Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations', *Computational Methods in Applied Mechanical Engineering* **32**(1–3), 199–259.

Cirpka, O. & Kitanidis, P. (2001), 'Sensitivity of temporal moments calculated by the adjoint-state method, and joint inversing of head and tracer data', *Advances in Water Resources* **24**(1), 89–103.

Dietrich, C. & Newsam, G. (1997), 'Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix', *SIAM Journal on Scientific Computing* **18**(4), 1088–107.

Frigo, M. & Johnson, S. (1998), FFTW: An adaptive software architecture for the FFT, *in* 'Proc. ICASSP', Vol. 3. IEEE, Seattle, WA, pp. 1381–1384. http://www.fftw.org.

Gomez-Hernandez, J., Sahuquillo, A. & Capilla, J. (1997), 'Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data, 1, theory', *Journal of Hydrology* **203**, 162–174.

Harvey, C. & Gorelick, S. (1995), 'Mapping hydraulic conductivity: Sequential conditioning with measurements of solute arrival time, hydraulic-head, and local conductivity', *Water Resources Research* **31**(7), 1615–26.

Kitanidis, P. (1983), 'A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations', *Water Resources Research* **19**(3), 677–90.

Kitanidis, P. (1986), 'Parameter uncertainty in estimation of spatial functions: Bayesian analysis', *Water Resources Research* **22**(4), 499–507.

Kitanidis, P. (1995), 'Quasi-linear geostatistical theory for inversing', *Water Resources Research* **31**(10), 2411–19.

Kitanidis, P. (1996), 'Analytical expressions of conditional mean, covariance, and sample functions in geostatistics', *Stochastic Hydrology and Hydraulics* **12**, 279–94.

McLaughlin, D. & Townley, L. (1996), 'A reassessment of the groundwater inverse problem', *Water Resources Research* **32**(5), 1131–61.

Nowak, W. & Cirpka, O. (2003), 'A modified Levenberg-Marquardt algorithm for the application to iterative cokriging', *in process* .

Nowak, W., Tenkleve, S. & Cirpka, O. (2003), 'Efficient computation of linearized cross-covariance and auto-covariance matrices of interdependent quantities', *Mathematical Geology* **35**(1), 53–66.

RamaRao, B., LaVenue, A., de Marsily, G. & Marietta, M. (1995), 'Pilot point methodology for automated calibration of an ensemble of conditionally simulated transmissivity fields, 1, theory and computational experiments', *Water Resources Research* **31**(3), 475–493.

Ruge, J. & Stueben, K. (1986), *"Algebraic Multigrid (AMG)" in Multigrid Methods*, Frontiers in Applied Mathematics, vol. 5, SIAM, Philadelphia.

Sun, N.-Z. (1994), *Inverse problems in groundwater modeling*, Theory and applications of transport in porous media, Kluwer Academic Publishers, Dordrecht.

Sun, N.-Z. & Yeh, W.-G. (1990), 'Coupled inverse problems in groundwater modeling. 1. sensitivity analysis and parameter identification', *Water Resources Research* **26**(10), 2507–25.

van Loan, C. (1992), *Computational Frameworks for the Fast Fourier Transform*, SIAM Publications, Philadelphia, PA.

Zimmerman, D. (1989), 'Computationally exploitable structure of covariance matrices and generalized covariance matrices in spatial models', *Journal of Statistical Computation and Simulation* **32**(1/2), 1–15.

Zimmerman, D., de Marsily, G., Gotaway, C. & Marietta, M. (1998), 'A comparison of seven geostatistically based inverse approaches to estimate transmissivities for modeling advective transport by groundwater flow', *Water Resources Research* **34**(6), 1373–1413.