Application of Large-Scale Inversion Algorithms to Hydraulic Tomography in an Alluvial Aquifer

P. Fischer, Abderrahim Jardani, A. Soueid Ahmed, Mohamad Abbas, X. Wang, H. Jourde, N. Lecoq

Research Paper/

Application of Large-scale Inversion Algorithms to Hydraulic Tomography in an Alluvial Aquifer

P. Fischer[1], A. Jardani [1], A. Soueid Ahmed[1], M. Abbas[1], X. Wang[2], H. Jourde[3] and N. Lecoq[1]

(1) Normandie Univ, UNIROUEN, UNICAEN, CNRS, M2C, 76000 Rouen, France

(2) Imperial College London, Department of Earth Science and Engineering, London, United Kingdom

(3) Université de Montpellier, CNRS, Laboratoire Hydrosciences, 34000 Montpellier, France

Conflict of interest: None

Corresponding author: Pierre Fischer

E-mail: pierre.fischer1@univ-rouen.fr

Key words: Inverse problems, Large-scale modeling, Hydraulic tomography, Geostatistical inversion

Intended for publication in Groundwater

Abstract

Large-scale inversion methods have been recently developed and permit now to considerably reduce the computation time and memory needed for inversions of models with a large amount of parameters and data. In this work, we have applied a deterministic geostatistical inversion algorithm to a hydraulic tomography investigation conducted in an experimental field site situated within an alluvial aquifer in Southern France. This application aims to achieve a 2-D large-scale modeling of the spatial transmissivity distribution of the site.

The inversion algorithm uses a quasi-Newton iterative process based on a Bayesian approach. We compared the results obtained by using three different methodologies for sensitivity analysis: an adjoint state method, a finite difference method, and a Principal Component Geostatistical Approach (PCGA). The PCGA is a large-scale adapted method which was developed for inversions with a large number of parameters by using an approximation of the covariance matrix, and by avoiding the calculation of the full Jacobian sensitivity matrix.

We reconstructed high resolution transmissivity fields (composed of up to 25,600 cells) which generated good correlations between the measured and computed hydraulic heads. In particular, we show that, by combining the PCGA inversion method and the hydraulic tomography method, we are able to substantially reduce the computation time of the inversions, while still producing high-quality inversion results as those obtained from the other sensitivity analysis methodologies.

Introduction

In hydrogeology, the assessment of hydraulic properties of subsurface aquifers, such as transmissivity, storage coefficient and solute transport parameters, is a key issue to an adequate management and protection of groundwater resources. Generally, aquifer characterization is based on the interpretation of hydraulic observations data collected during pumping, infiltration, or tracer tests (Carrera and Neuman 1986b; Rao et al. 2003; Lee and Kitanidis 2014; Pool et al. 2015). Therefore, hydraulic tomography is considered as one of the most effective approaches for characterizing the spatial distribution of hydraulic transmissivity of an aquifer (Cardiff et al. 2009; Berg and Illman 2013; Cardiff et al. 2013; Soueid Ahmed et al. 2015; Zha et al. 2015; Wang et al. 2016). This method relies on a set of hydraulic head responses recorded during cross-hole pumping experiments. Then, the interpretation can be achieved through the use of an inverse algorithm to image the spatially varying hydraulic properties in the subsurface. The inverse problem for estimating hydraulic parameters involves a formulation of a forward problem, which sets up the link between the hydraulic observations and the unknown hydraulic parameters (Tarantola and Valette 1982). For a hydraulic tomography inversion, the forward modeling is based on a numerical method (e.g. the finite element, finite difference and finite volume methods) used to solve the partial differential equation of the groundwater flow. The forward problem operator can be formulated as:

$$\mathbf{d} = f(\mathbf{s}) + \boldsymbol{\eta}, \tag{1}$$

where $\mathbf{d}$ represents the hydraulic responses of the model, $\mathbf{s}$ is the logarithm of the m unknown hydraulic transmissivities, to be estimated from a set of n observed data $\mathbf{d}_{obs}$ and a nonlinear forward modeling application $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$. $\boldsymbol{\eta}$ is an additive noise of the numerical modeling.

In a probabilistic framework, the inverse problem maximizes a posterior probability density function $\pi_{\text{post}}(\mathbf{s}|\mathbf{d})$. Generally, the problem is ill-posed and the solution is non-unique, therefore additional *a priori* information is required in order to find one physically meaningful solution (Carrera and Neuman 1986a). Furthermore, in our case, the problem is under-determined, it involves a small number of measurements but a large number of unknown parameters ($n \ll m$). The inverse problem corresponds to recovering the 'best fitting' model parameters $\hat{\mathbf{s}}$ which allow the model responses to match the observed data, and at the same time respect the constraints imposed by the *a priori* information on the model. Using the Bayes' formula, the posterior probability density function can be expressed as (Elsheikh et al. 2014):

$$\pi_{\text{post}}(\mathbf{s}|\mathbf{d}) \propto exp\left[-\frac{1}{2}(f(\mathbf{s})-\mathbf{d})^{\mathrm{T}}\mathbf{R}^{-1}(f(\mathbf{s})-\mathbf{d}) - \frac{1}{2}(\mathbf{s}-\mathbf{s}_{\text{prior}})^{\mathrm{T}}\mathbf{Q}^{-1}(\mathbf{s}-\mathbf{s}_{\text{prior}}) \right], \qquad (2)$$

where $\mathbf{s}_{\text{prior}}$ denotes the *a priori* model; $\mathbf{Q}$ is a $m \times m$ covariance matrix of the model parameters, which can include geostatistical information about the distribution trend or pattern of the unknowns parameters (in that case $\mathbf{Q}$ is defined in a matrix where elements of the matrix represent the variogram function associated with the distance between cells of the model) and $\mathbf{R}$ denotes a $n \times n$ diagonal covariance matrix which accounts for the errors occurred in the data measurements.

The aim of the inversion problem is to find a set of parameter which maximizes the density of probability, $\pi_{\text{post}}(\mathbf{s}|\mathbf{d})$. This corresponds to a model of high probability with respect to the measurements and the imposed *a priori* model (Tarantola and Valette 1982). To solve the inverse problem, two main groups of iterative methods are often employed: (1) the deterministic methods which assume that the algorithm converges to a local minimum by performing a linearization of an objective function and (2) the stochastic methods which converge to a global

minimum by selecting different randomly generated parameters fields as probable solutions to the model (Pool et al. 2015). Several deterministic and stochastic inversion algorithms have been widely applied in hydrogeology, but they are time- and memory-expensive, particularly for the cases which involve a fine discretization of the parameter grid and/or a large amount of observational data (Kitanidis and Lee 2014). Thereby, one of the main trends in the development of inversion theory during the last few years was to develop algorithms which are able to solve inversion models with a large number of unknown parameters and data.

Recently, several time- and memory-saving methods have been developed to reduce the memory cost and the computation time of the usually large matrices involved in the inversion algorithms. One way for reducing the computational and memory demands is to use some approximation methods on matrix operations, such as the Fast Multipole Method (FMM) (Greengard and Rokhlin 1987), which is based on Legendre polynomial expansions and spherical harmonics. The FMM was associated with the Hierarchical matrices approach (Hackbusch and Börm 2002) to compute matrix-vector products for a large-scale application in seismic imaging (Ambikasaran et al. 2013). Another way is through using the MINRES Krylov subspace method (Paige and Saunders 1975) which can be combined with the Fast Fourier Transform (FFT) (Nowak et al. 2003) to iteratively solve inversions of large matrix systems. This method has recently been applied to 3-D large-scale transient hydraulic tomography problems (Liu et al. 2014). The two approaches mentioned above avoid the calculation of the full Jacobian matrix of the forward model at each iteration. A new method has recently been developed by Lieberman et al. (2013), in an application of an inversion algorithm for a large-scale 3-D transient contaminant transport. The authors used a Proper Orthogonal Decomposition method (POD) to compute a projection basis with the eigenvectors associated to the highest eigenvalues of the Hessian matrix of the forward problem. The inverse problem was then solved in a reduced projected subspace.

In this paper, we have attempted to achieve an efficient site hydraulic characterization (i.e. to obtain high-resolution transmissivity fields at a low computational cost) by applying a recently developed geostatistical inversion method, the Principal Component Geostatistical Approach (PCGA) (Lee and Kitanidis 2014). This method can considerably reduce the computation time and the memory cost of inversions by using an approximation of the covariance matrix $\mathbf{Q}$ based on a Singular Value Decomposition method (SVD), and by avoiding the computation of the Jacobian matrix through the use of a matrix-free product based on a finite difference method.

This paper presents an application of the PCGA method, combined with a hydraulic tomography investigation, for a large-scale inverse modeling of the hydraulic transmissivity field of an alluvial aquifer. First, we present the methodology of the Geostatistical Approach (GA) algorithm, and the modifications for large-scale application (PCGA). Then, we describe the hydrogeological background of the experimental field site, from which the hydraulic measurements were taken, and present the numerical model setup. Finally, we show our inversion results. In particular, we compare these results to those obtained by applying the classical GA method with two different methods in Jacobian matrix computation (i.e., an adjoint state and a finite difference method), which do not use a covariance matrix approximation. We have evaluated the computation times, and the sensitivities and accuracies of the inversion results for the three different methodologies. Using a hydraulic tomography field application on a porous aquifer, we show the advantages of the PCGA inversion method for efficient large-scale inverse modeling in hydrogeology.

Principal Component Geostatistical Approach

In the Geostatistical Approach (GA) (Kitanidis and Vomvoris 1983; Hoeksema and Kitanidis 1984; Kitanidis 1995), the prior probability density function of the m model parameters $\mathbf{s}$ is set as a multivariate Gaussian with a mean $E(\mathbf{s})=\mathbf{X}\boldsymbol{\beta}$ where $\mathbf{X}$ is an $m \times p$ known matrix and

$\boldsymbol{\beta}$ a $p \times 1$ vector to be determined during the inversion process (generally $p = 1$), and a covariance, $E\left[(\mathbf{s}-\mathbf{X}\boldsymbol{\beta})(\mathbf{s}-\mathbf{X}\boldsymbol{\beta})^{\mathrm{T}}\right] = \mathbf{Q}(\theta)$.

The posterior probability density function $\Psi = -ln\left[\pi(\mathbf{s}|\mathbf{d})\right]$ (also called in inversion problems the objective function) then becomes:

$$\Psi = \frac{1}{2}(f(\mathbf{s})-\mathbf{d})^{\mathrm{T}}\mathbf{R}^{-1}(f(\mathbf{s})-\mathbf{d}) + \frac{1}{2}(\mathbf{s}-\mathbf{X}\boldsymbol{\beta})^{\mathrm{T}}\mathbf{Q}^{-1}(\mathbf{s}-\mathbf{X}\boldsymbol{\beta}) \ . \tag{3}$$

The best approximation $\hat{\mathbf{s}}$ for the model parameters, taking into account the *a priori* information and the observed data, can be found as being the model maximizing the density of probability in (Eq. 2), which is also equivalent to minimize the argument of its exponential. Thus, $\hat{\mathbf{s}}$ is found by minimizing the objective function Ψ (Eq. 3). This minimization can be achieved by using a Newton linearization iterative approach on $\mathbf{s}$. The iterative process initializes at a reasonable $\mathbf{s_0}$. Then, at iteration step $j+1$, the new value $\mathbf{s_{j+1}}$ is found in the vicinity of the previous model $\mathbf{s_j}$ using a first order Taylor approximation:

$$f(\mathbf{s_{j+1}}) = f(\mathbf{s_j}) + \mathbf{F_j}(\mathbf{s_{j+1}} - \mathbf{s_j}) \ . \tag{4}$$

Here $\mathbf{F_j}$ is the $n \times m$ Jacobian matrix of the forward problem f for $\mathbf{s_j}$: $\mathbf{F_j} = \left.\frac{\partial f}{\partial \mathbf{s}}\right|_{\mathbf{s}=\mathbf{s_j}}$.

After some matrix manipulations, the updated solution of the parameters in the iterative process, found by minimizing the objective function, can be written as (Kitanidis 1995):

$$\mathbf{s_{j+1}} = \mathbf{X}\boldsymbol{\beta_j} + \mathbf{Q}\mathbf{F_j}^{\mathrm{T}}\boldsymbol{\xi_j} \ , \tag{5}$$

where the $p \times 1$ matrix $\boldsymbol{\beta}_j$ and the $n \times 1$ matrix $\boldsymbol{\xi}_j$ are solutions of the following matrix system (Kitanidis 1995):

$$\begin{bmatrix} \mathbf{F}_j\mathbf{Q}\mathbf{F}_j^{\mathrm{T}} + \mathbf{R} & \mathbf{F}_j\mathbf{X} \\ (\mathbf{F}_j\mathbf{X})^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{\xi}_j \\ \boldsymbol{\beta}_j \end{bmatrix} = \begin{bmatrix} \mathbf{d} - f(\mathbf{s}_j) + \mathbf{F}_j\mathbf{s}_j \\ \mathbf{0} \end{bmatrix}. \tag{6}$$

Note that here '$\mathbf{0}$' represents a $p \times p$ matrix of zeros on the left-hand side and a $p \times 1$ matrix of zeros on the right-hand side.

At the end of the iterative process, to quantify the model parameter's uncertainty after optimization, we can compute the posterior covariance of $\mathbf{s}$ derived as:

$$\mathbf{Q}_{\mathbf{post}} = \mathbf{Q} - \begin{bmatrix} \mathbf{F}_j\mathbf{Q} \\ \mathbf{X}^T \end{bmatrix}^T \begin{pmatrix} \mathbf{F}_j\mathbf{Q}\mathbf{F}_j^T + \mathbf{R} & \mathbf{F}_j\mathbf{X} \\ (\mathbf{F}_j\mathbf{X})^T & \mathbf{0} \end{pmatrix}^{-1} \begin{bmatrix} \mathbf{F}_j\mathbf{Q} \\ \mathbf{X}^T \end{bmatrix}. \tag{7}$$

The GA method as presented above needs the computation of the Jacobian matrix $\mathbf{F}$ for each iteration in order to solve the system (Eq. 6), which can usually be done by solving the forward problem $m+1$ times using a finite difference method, or $n+1$ times using an adjoint state method. Even if the adjoint state method may considerably decrease the computation time for under-determined problems (see Cardiff and Kitanidis 2008 for a comparison of the finite difference and adjoint state method computation times), it is not efficient for large scale problems with a large number of measurements and parameters. Another problem which appears in the GA method is that when the number of data and/or parameters is high, it requires a significant computational power for the calculation and storage of the covariance matrix $\mathbf{Q}$ (which can be alleviated by FFT, H-matrices or FMM). To overcome these difficulties,

Kitanidis and Lee have developed the Principal Component Geostatistical Approach (PCGA), on the basis of the GA method.

In this new approach, the computational and memory costs associated with the manipulation of matrix $\mathbf{Q}$ are reduced using a low-rank approximation of a chosen truncation order $K \ll m$ through a Singular Value Decomposition (SVD):

$$\mathbf{Q}_K = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathrm{T}} , \tag{8}$$

where $\mathbf{S}$ is a $K \times K$ diagonal matrix containing the singular values of $\mathbf{Q}$ sorted in descending order, $\mathbf{U}$ is a $m \times K$ matrix and $\mathbf{V}$ is an $m \times K$ matrix. As $\mathbf{Q}$ is defined as a symmetric matrix, its SVD simplifies to: $\mathbf{Q}_K = \mathbf{V}\mathbf{S}\mathbf{V}^{\mathrm{T}}$.

This decomposition can also be written as:

$$\mathbf{Q}_K = \sum_{i=1}^{K} \zeta_i \zeta_i^{\mathrm{T}} \quad \text{with } \zeta_i = \sqrt{\lambda_i}\mathbf{V}_i \ . \tag{9}$$

Here λ_i is the i^{th} singular value (also $\mathbf{S}(i,i)$) and $\mathbf{V}_i$ is the i^{th} column vector of $\mathbf{V}$ associated to λ_i. The error arising from this K-rank decomposition equals to the $K+1$ th singular value (λ_{K+1}) of $\mathbf{Q}$.

However, this decomposition is a good approximation only for a matrix $\mathbf{Q}$ in which the most of its information is contained in its few highest eigenvalues and eigenvectors, meaning a relatively smooth pattern. One can also use an eigen-decomposition if $\mathbf{Q}$ is a matrix defined by positive eigenvalues or a randomized decomposition approach which is efficient for high-dimensional matrices with $m \sim 1{,}000{,}000$ (Halko et al. 2011).

In addition, the PCGA method also avoids the full Jacobian matrix calculation at each iteration. When performing a matrix product, such as $\mathbf{F_j u}$ (where $\mathbf{F}_j$ is the $n \times m$ Jacobian matrix and $\mathbf{u}$ is a $m \times 1$ vector), instead of computing it directly, the method finds an approximation to its accurate form using a first order Taylor series :

$$f\left(\mathbf{s}_j+\frac{\delta\|\mathbf{s}_j\|}{\|\mathbf{u}\|}\mathbf{u}\right)=f(\mathbf{s}_j)+\frac{\delta\|\mathbf{s}_j\|}{\|\mathbf{u}\|}\mathbf{F}_j\mathbf{u}+\sigma(\delta^2)\ , \tag{10}$$

$$\mathbf{F}_j\mathbf{u}\approx\frac{\|\mathbf{u}\|}{\delta\|\mathbf{s}_j\|}\left[f\left(\mathbf{s}_j+\frac{\delta\|\mathbf{s}_j\|}{\|\mathbf{u}\|}\mathbf{u}\right)-f(\mathbf{s}_j)\right], \tag{11}$$

where δ is a finite-difference interval and $\|\mathbf{u}\|$ and $\|\mathbf{s}_j\|$ are the norm of the vector $\mathbf{u}$ and $\mathbf{s}_j$. Thus, in the matrix system (Eq. 6) of the GA algorithm the calculation of the full Jacobian $\mathbf{F}$ is avoided by approximating the computation of $\mathbf{F}_j\mathbf{s}_j$, $\mathbf{F}_j\mathbf{X}$, $\mathbf{F}_j\mathbf{Q}\mathbf{F}_j^{\mathrm{T}}$ and $\mathbf{Q}\mathbf{F}_j^{\mathrm{T}}$ with (Kitanidis and Lee 2014):

- $\mathbf{F}_j\mathbf{s}_j=\frac{1}{\delta}\left[f(\mathbf{s}_j+\delta\mathbf{s}_j)-f(\mathbf{s}_j)\right]$, (12)
- $\mathbf{F}_j\mathbf{X}_i=\frac{\|\mathbf{X}_i\|}{\delta\|\mathbf{s}_j\|}\left[f\left(\mathbf{s}_j+\frac{\delta\|\mathbf{s}_j\|}{\|\mathbf{X}_i\|}\mathbf{X}_i\right)-f(\mathbf{s}_j)\right]$ with $\mathbf{X}_i$ the i^{th} column of $\mathbf{X}$, (13)
- $\mathbf{F}_j\mathbf{Q}\mathbf{F}_j^{\mathrm{T}}\approx\mathbf{F}_j\mathbf{Q}_K\mathbf{F}_j^{\mathrm{T}}=\mathbf{F}_j\sum_{i=1}^{K}\zeta_i\zeta_i^{\mathrm{T}}\mathbf{F}_j^{\mathrm{T}}=\sum_{i=1}^{K}(\mathbf{F}_j\zeta_i)(\mathbf{F}_j\zeta_i)^{\mathrm{T}}=\sum_{i=1}^{K}\boldsymbol{\eta}_i\boldsymbol{\eta}_i^{\mathrm{T}}$, (14)
- $\mathbf{Q}\mathbf{F}_j^{\mathrm{T}}\approx\mathbf{Q}_K\mathbf{F}_j^{\mathrm{T}}=\sum_{i=1}^{K}\zeta_i\zeta_i^{\mathrm{T}}\mathbf{F}_j^{\mathrm{T}}=\sum_{i=1}^{K}\zeta_i(\mathbf{F}_j\zeta_i)^{\mathrm{T}}=\sum_{i=1}^{K}\zeta_i\boldsymbol{\eta}_i^{\mathrm{T}}$, (15)

where $\boldsymbol{\eta}_i=\mathbf{F}_j\zeta_i=\frac{\|\zeta_i\|}{\delta\|\mathbf{s}_j\|}\left[f\left(\mathbf{s}_j+\frac{\delta\|\mathbf{s}_j\|}{\|\zeta_i\|}\zeta_i\right)-f(\mathbf{s}_j)\right]$.

If we now consider the number of forward model evaluations needed per iteration for the calculation of the Jacobian matrix, there are $K+p+2$ runs. One run is needed for evaluating $f(\mathbf{s}_j)$, one is needed for assessing $f\left(\mathbf{s}+\delta\mathbf{s}_j\right)$, K runs are needed for calculating $\boldsymbol{\eta}_i$ and p runs are needed for computing $\mathbf{X}_i$. It can be observed that, with this method, the number of

forward model runs per iteration is no longer dependent on m or n. Hence, the number of parameters and observed data can increase without increasing the run time of the algorithm. However, if the number of parameter increases, the low-rank approximation order K might also need to be increased slightly in order to maintain a small truncation error for $\mathbf{Q}_K$.

The algorithm iteratively updates the parameters set in Eq. 5 by solving the matrix system of Eq. 6 with the PCGA approximations until the optimum $\hat{\mathbf{s}} = \mathbf{s}_{\text{post}}$ is achieved, that is, the objective function has iteratively converged to a local minimum.

Application to an experimental site

We have applied the PCGA large-scale method as presented in the previous part to an experimental site, named ‘la Céreirède’. The field site is located in Montpellier in the South of France, on the alluvium of the Lez river, which flows towards the Mediterranean Sea a few kilometers downstream (Figure 1).

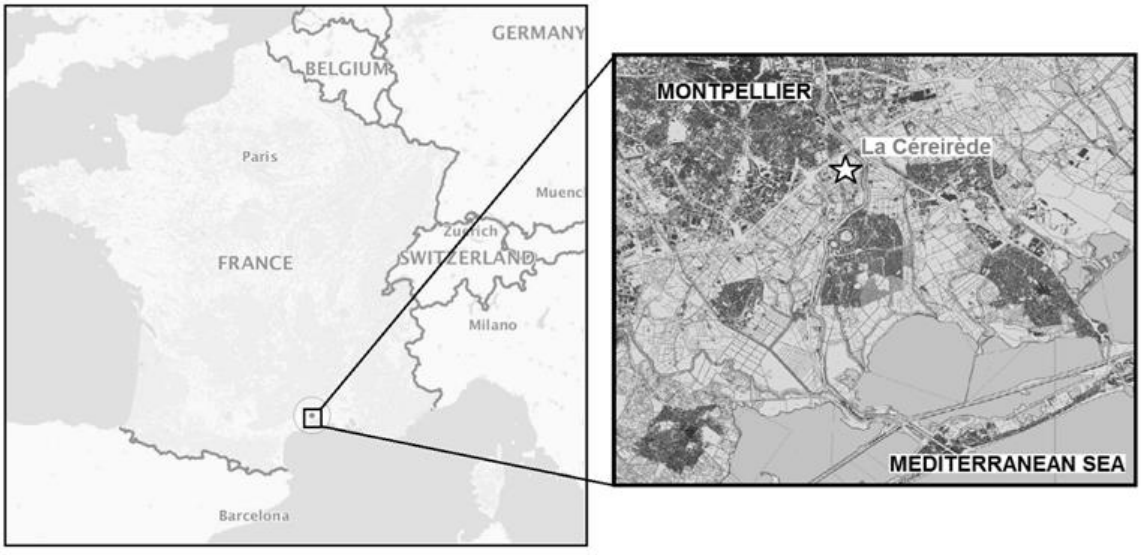

Figure 1 : Location of the studied experimental site 'La Céreirède' (Map and aerial photography from geoportail.fr) occupying an area of 720 m². It is situated in the South of France, near the town of Montpellier and the Mediterranean Sea.

At the field site, the alluvial deposit is composed of a 12 m thick formation of unconsolidated sands and silts lying on a 2 m thick layer of pebbles and gravels. Beneath these Quaternary formations, there exist clayey sands of the Pliocene, marls of the Miocene, and limestone of the Jurassic (Figure 2). Three porous aquifer formations have been characterized in this site:

- a low permeability aquifer in the upper part of the alluvium (unconsolidated sands and silts),
- a semi-confined aquifer in the pebbles and gravel,
- a semi-confined to confined aquifer in the clayey sands of Pliocene.

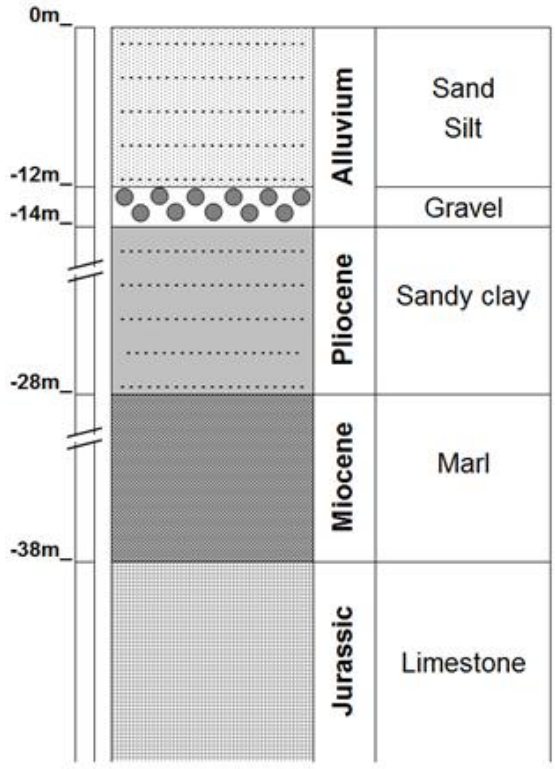

Figure 2 : Schematic geological section of the experimental site 'La Céreirède'. Three aquifers formations have been characterized in the sands and silts alluvium, in the gravels and in the clayey sands.

The field site comprises 12 wells which fully penetrate the three aforementioned aquifers in an area of 36×20 m^2 (Figure 3). The hydraulic data were collected by performing two pumping

tests in PZ 2 (5 L/min), and in PZ 11 (5 L/min), respectively, while measuring hydraulic head variations in the other 10 wells.

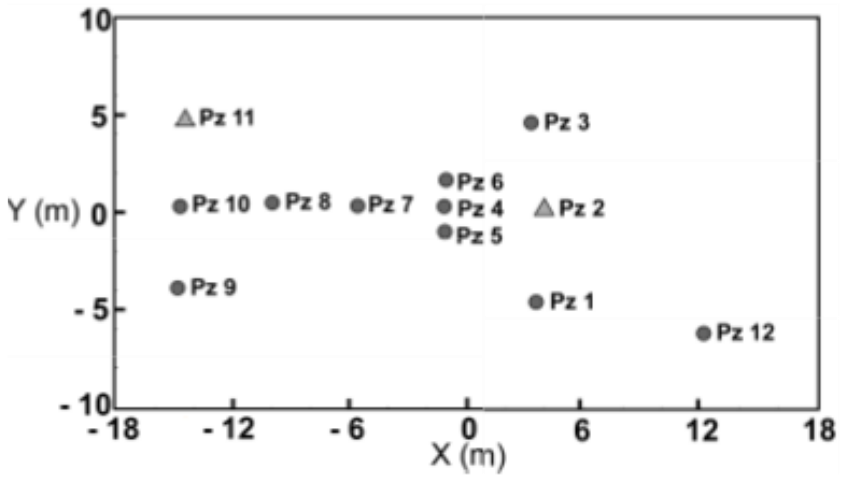

Figure 3 : Well pattern on the experimental site 'La Céréirède' (circles represent the 10 measurement wells and triangles represent the 2 pumping wells). As hydraulic drawdowns in the pumping wells are not measured, the tomography provided 20 observed data.

The pumping were performed at the depth of the pebbles and gravel layer, which is the most productive aquifer, because its transmissivity is considerably higher than the transmissivities in silts and clayed sands. We considered that the contribution of the two others aquifers to the water pumped is negligible compared to the contribution from the pebbles and gravel aquifer. From a classic hydrogeological analysis of the soil of each well, we could also estimate values of the field transmissivity at these points. Using these punctual values of transmissivities and their positions in the field as input in a MATLAB variogram routine ('variogramfit' by W. Schwanghart), we were able to obtain the transmissivity field variogram function, which will be used in the inversion algorithm to create the covariance matrix Q (Table 1). The variogram function is of type exponential with a sill of 0.11 and a range of 8 m.

The PCGA inversion algorithm was implemented in MATLAB and connected to the flow modeling software COMSOL Multiphysics, which solves the forward problem. The inversion

algorithm performs, at the end of each iteration, a linear minimum research ('fminsearch' MATLAB function) to accelerate the convergence of Eq. 3 to a local minima.

The 2-D flow model is discretized as a grid of $\sqrt{m} \times \sqrt{m}$ transmissivity cells in a rectangular zone. This local model is enclosed by a larger buffer zone of 100×100 m² with a constant transmissivity of 10^{-5} m²/s (average value of locally estimated transmissivities at the field site) and a 0 m constant head condition at the boundaries (no drawdown induced by the pumping wells). This buffer zone was set up in order to minimize the impact of the boundary conditions. The flow simulations were performed under steady-state conditions. The inversion of the model is set up using the 2×10 drawdown observed during the pumping tests (10 measurement wells for each of the 2 pumping tests) representing the observed data in the inversion algorithm. The inversion aims to reconstruct the spatially varying *T* distribution in the local region producing the observational data set.

Results

We have applied the PCGA to assess the equivalent transmissivity field of the multi-layered aquifer at the Céreirède field site. The most transmissive part of the aquifer is the pebbles and gravels part, but the alluvium and the clayey sands might also be the cause of some variations in the estimated equivalent transmissivity field. The inputs to the inversion models are given in Table 1.

Table 1 : Values of variables used to perform the PCGA inversion on a model of the site for 25,600 parameters and 20 observed data. Results of this inversion are shown in Figures 4 and 5.

Geometry	X (m) = [-18 , 18] ; Y (m) = [-10 , 10]
Grid (number of parameters)	160 × 160 cells
Uncertainty on data	$\boldsymbol{\sigma}$ = 0.001 m ; $\mathbf{R}=\boldsymbol{\sigma}.Id(n)$

Experimental variogram function	$Var(\text{distance}) = 0.11 \times \left(1 - exp\left(-\frac{\text{distance}}{8}\right)\right)$
Finite-difference step	$\delta = 10^{-5}$

The inversions were performed on a uniform fine-scale discretization grid (160×160). A constant initial transmissivity field was considered in the inverse models. The Kth order of truncation for the covariance matrix **Q** was selected such that the first truncated singular value of the matrix (the low rank approximation error) fall below 1. This corresponds to an order of $K = 128$. We will show that this choice of truncation order is acceptable and allows the significant information about the prior model structure to be preserved. The low-rank covariance matrix was calculated and then imported to the inversion algorithm. The parallelized computation of the low-rank decomposition takes only a few minutes (it even takes less than 1 minute with a 32 cores parallelization). The inversion then converged in 2h 45min on an Intel Xeon QuadCore 2.8GHz with 12Go RAM.

The results from the inversion are presented in Figures 4 and 5. Firstly, from the distribution of model parameters (given as negative log transmissivity in Figure 4), it can be seen that the value of the inverted transmissivity takes the mean value 10^{-5} m^2/s, which is the mean of the transmissivity measurements on the site. A clear contrast in T is observed between the two regions on the east and west sides of boreholes PZ1, PZ2 and PZ3. Overall, the eastern part which is closer to the Lez river, is slightly less transmissive than the western part (T are approximately 2×10^{-6} m²/s and 1×10^{-5} m²/s, respectively). An area with a highest T (3×10^{-5} m^2/s) is also highlighted within the western part, around PZ 7. But it has to be noticed that the boreholes and pumping wells are not homogeneously distributed over the site, thus some parts of the site (especially on the eastern side) might give more uncertain results. Therefore, it is interesting to estimate the values of the transmissivity field uncertainties.

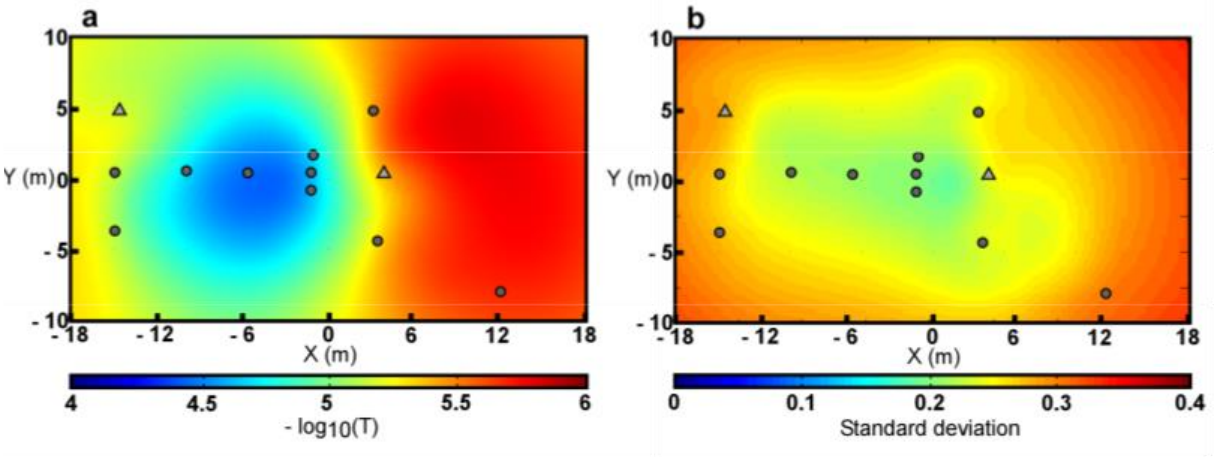

Figure 4 : Maps of the log-transmissivity (a, on the left) and parameter's a posteriori standard deviation (b, on the right) for a PCGA inversion method with 25,600 parameters, 20 observed data and a covariance matrix decomposition of order K=128 applied to the experimental site. The transmissivities vary around a mean of 10^{-5} m²/s which is consistent with transmissivity values estimated from pumping test analysis. The aquifer is less transmissive in the eastern part and more in the western part especially in a zone around PZ 7 (see Figure 3). But we got a better precision in zones with more information: at the center and the western part of the map, while in the eastern part where we didn't have piezometers, the results show a larger standard deviation.

Figure 4b presents a map of the uncertainty for each parameter value (given by the diagonal entries of the posterior covariance matrix $\mathbf{Q}_{post}$). The standard deviation for the log-transmissivity varies between 0.2 for the parameter cells near the investigation wells, and 0.33 in both the area with very few information and close to the model's boundaries. In particular, the uncertainty in the eastern part, where the number of wells is small, is much higher.

Good correlations between the calculated and measured hydraulic heads were obtained (Figure 5). The root mean square error calculated at the end of the inversion was computed as 0.194 m. The hydraulic heads with the most significant difference between the inverted and the measured values are observed on PZ1 and PZ9.

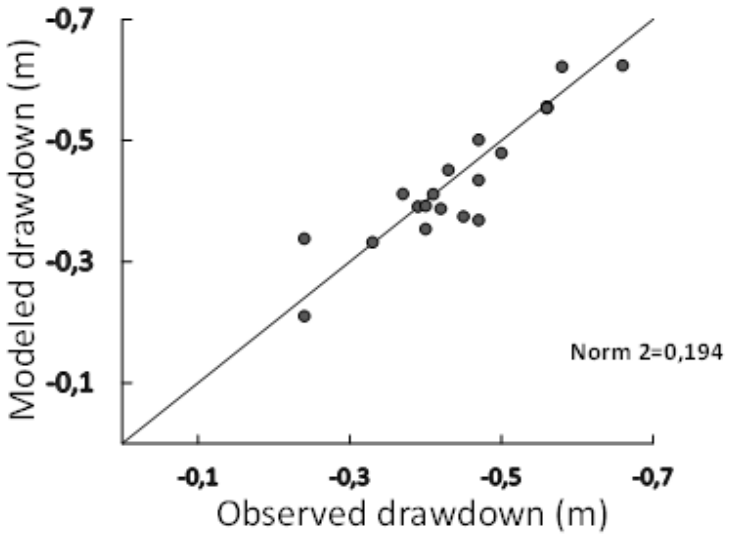

Figure 5 : Graph showing the differences between the 20 observed drawdowns and modeled drawdowns after convergence of a PCGA inversion method with 25,600 parameters and a covariance matrix decomposition of order *K*=128 applied to the experimental site. The drawdowns are globally well reproduced.

It is clear that the PCGA method is reliable for the modeling of the porous aquifer considered in this work because it produces a set of good inversion results with a high resolution (each cell represents a rectangle of 22.5 cm × 12.5 cm on the site) from a few measurements (20 observed data for 720 m²) in less than 3 hours. However, the main problem of this method is that the SVD of the covariance matrix needs a considerably large amount of time and memory. The computational demands increase squarely with the number of model parameters, m. In this work, the decomposition was only performed once before the inversion, and the resulting low-rank covariance matrix was used at all iteration steps throughout the inversion (i.e., the variogram function remaining the same). Otherwise, if a variable variogram model is desired at different iteration steps, the computation of the covariance matrix decomposition can also be accelerated by specific linear algebra methods (FFT, FMM and H-matrices) and a parallelization on several cores to achieve a reasonable computational time (Lee and Kitanidis 2014).

Comparison of results between PCGA simulations with different decomposition order, and between PCGA and GA simulations

In addition to the PCGA experimental application presented above, several other inversions were also conducted for the same field site but with a smaller number of parameters and using different methods for the computation of sensitivity matrix (input for each inversion is shown in Table 1). These numerical experiments allows us to compare (1) the results from PCGA with varied low-rank truncation K; and (2) the results from PCGA to those obtained from the GA method with a 1st order finite-difference Jacobian matrix computation, and an adjoint-state Jabobian matrix computation (the integral was solved using the Gauss-Legendre quadratic method as described in Soueid Ahmed et al. 2014) The inversion results are compared with respect to the total computation time and relative accuracy of the results.

Comparison of inversion results for using different decomposition order

First, we assess the effect of the chosen truncation order for the covariance matrix. Three PCGA inversions, with 10,000 parameters and using different K-order truncations of the covariance matrix corresponding to singular values (truncation errors) of $\lambda_{K+1} \simeq 1$ (K=69), $\lambda_{K+1} \simeq 0.1$ (K=313) and $\lambda_{K+1} \simeq 0.01$ (K=1,532), were performed. Figure 6 shows a relationship between the singular value and the truncation order, on which the position of the three K orders that were adopted in our inversions are indicated. Figure 7 shows the results obtained from these inversions.

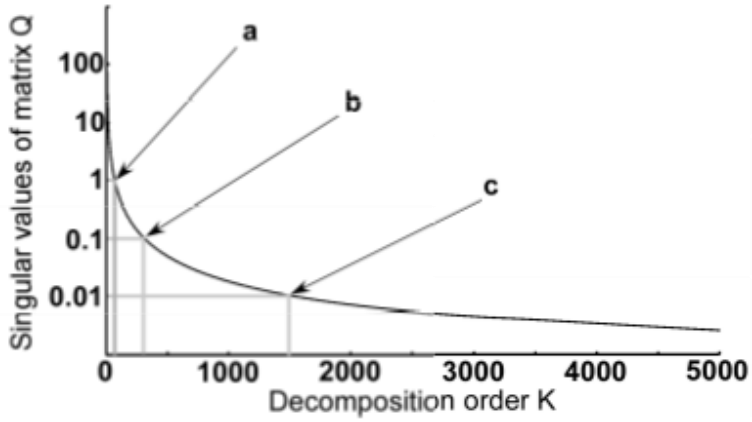

Figure 6 : Covariance matrix singular values decrement curve for 10,000 parameters. Three decomposition order (a to c) corresponding to truncation error of 1, 0.1 and 0.01 have been chosen for the results comparison of the PCGA inversion method (see Figure 7).

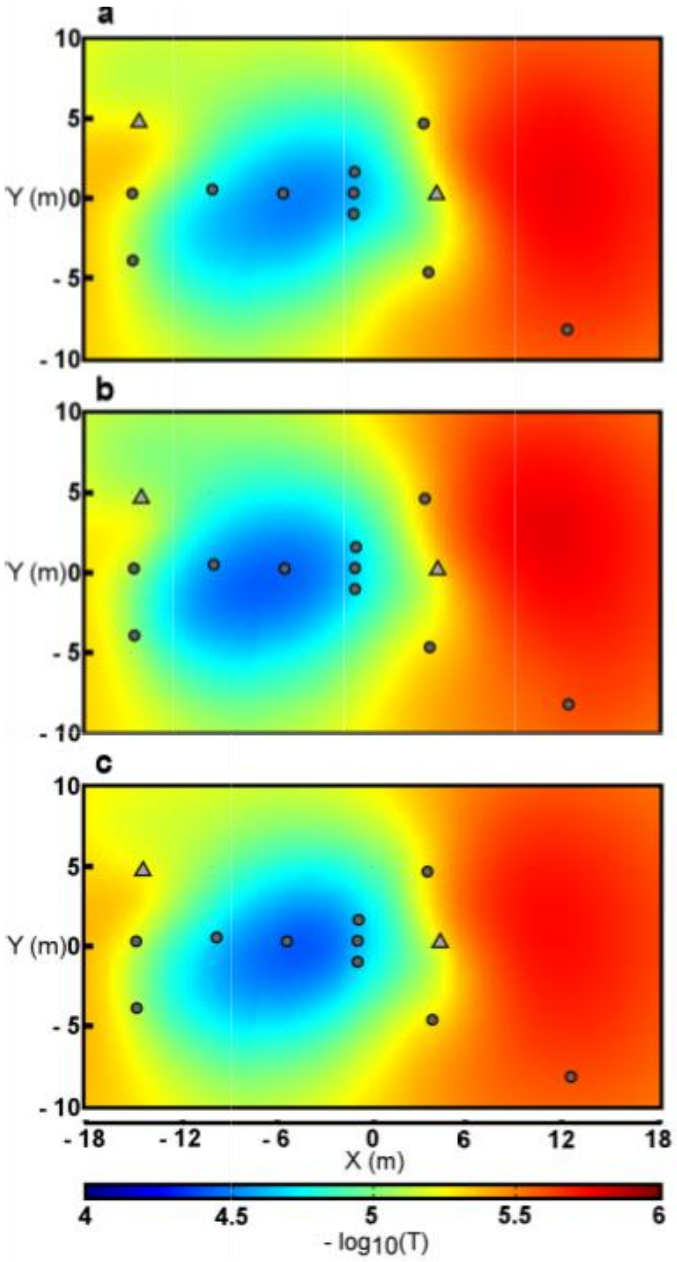

Figure 7 : Maps of the log-transmissivity for a PCGA inversion method with 10,000 parameters, 20 observed data and three different covariance matrix decomposition applied to the experimental site. The map (a) was obtained for *K*=69, the map (b) for *K*=313 and the map (c) for *K*=1,532 (see Figure 6). The results obtained for these three decomposition are relatively the same (same transmissivity values, same zones) so, for this site, there is no significant loss of information when using a truncation order corresponding to an error of 1 (map (a)) for the covariance matrix which allows us to reduce the computation time of the inversion without decreasing the accuracy of the results.

Note that very similar *T* fields (e.g. similar trend and location of high *T* zones) were obtained from the inversions using different truncation order. Therefore, the influence of reducing the K -order, as long as the truncation error is below 1, on the inversion results is mild. The most information regarding the spatial structure of the prior model is preserved in its first few singular values, so it is acceptable to consider a truncation order for a truncation error $\lambda_{K+1} \simeq$ 1 for the covariance matrix.

PCGA and GA results comparison

In this section, we compare the inversion results obtained for the PCGA method using an approximated Q matrix, with those of the GA method using two different methods for Jacobian matrix computation (i.e., the finite-difference and adjoint state methods), where the entirety of the Q matrix were used. All the inverse simulations were performed on a 100 × 100 grid. Figure 8 shows the inverted transmissivity distributions and the corresponding distributions of standard deviation of each model parameter.

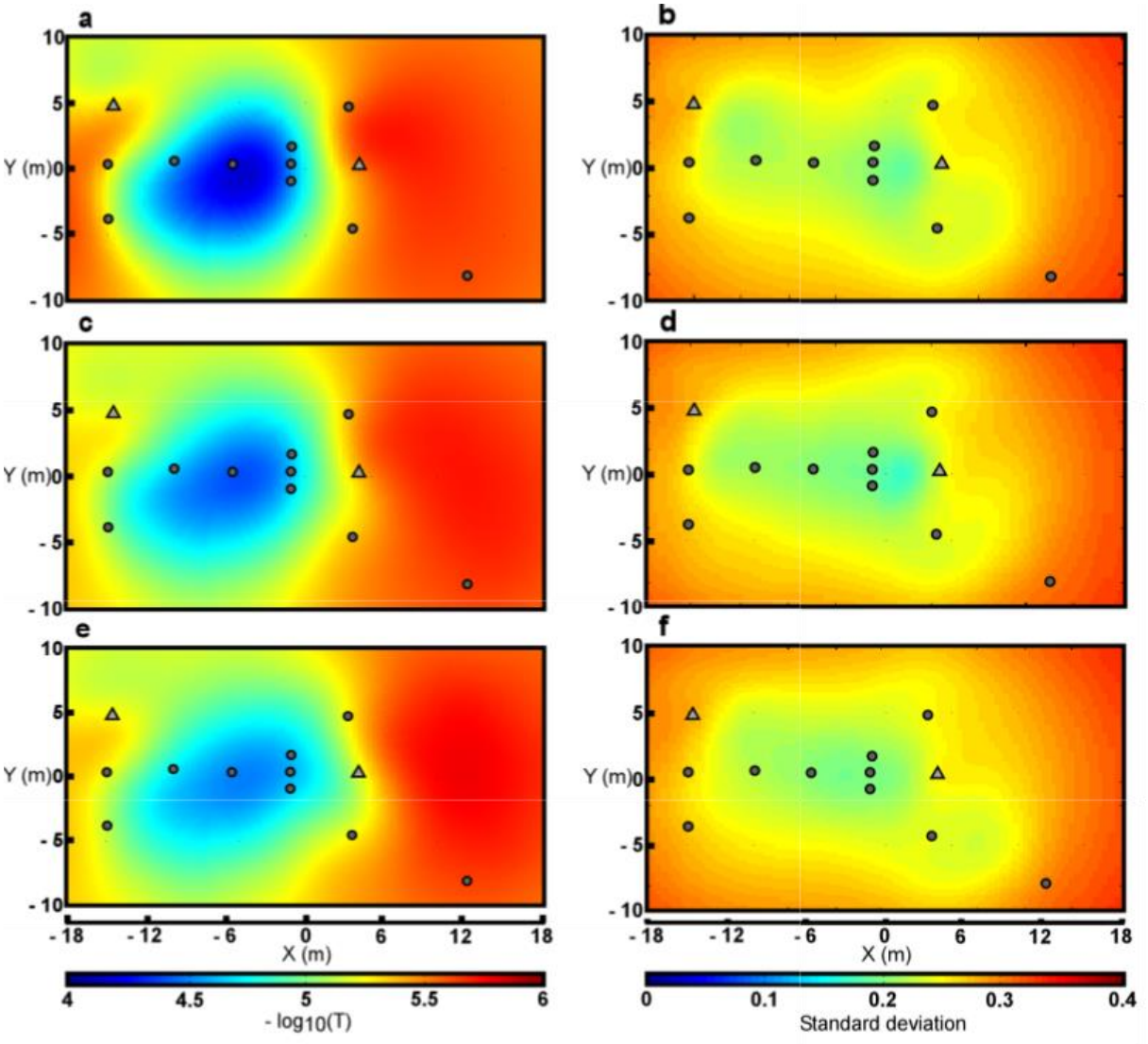

Figure 8 : Maps of the log-transmissivity (a, c, e) and parameter's a posteriori standard deviation (b, d, f) for three different inversion methods with 10,000 parameters and 20 observed data applied to the experimental site. The maps (a) and (b) were obtained with the GA adjoint-state method, the maps (c) and (d) with the GA finite-difference method and the maps (e) and (f) with the PCGA method with a covariance matrix decomposition of order K=69. The results between the three methods are relatively the same for this site, except for the map a which presents a slightly higher contrast of the transmissivity distribution leading to a better data matching (see Figure 9), though the PCGA inversion method is much more efficient for the calculation time (see Table 2).

It can be seen that in the three transmissivity fields, the calculated T value for each cell varies around the mean 10^{-5} m^2/s. In general, the three approaches produced similar spatial distribution of the transmissivity, however the range of the inverted transmissivities from the GA-adjoint state method (i.e., 2×10^{-6} to 1×10^{-4} m^2/s) is larger than that of the PCGA and that of the GA-

finite difference method (i.e., 2×10^{-6} to 3×10^{-5} m^2/s). Since the PCGA method is based on a finite difference matrix vector product approximation, it is expected to obtain a similar parameter range from this method and the GA-finite difference method. The difference between the results from these two methods and the GA-adjoint state method may come from the definition of the finite-difference step δ. In addition, the resolved transmissivity field from PCGA is smoother compared to that from the method where an adjoint state method is used to compute the sensitivity matrix. This is caused by the low-rank truncation in PCGA and also the finite difference approach, which tends to reduce the heterogeneity of the inverted *T* field.

The spatial distributions of the standard deviation of the inverted parameters is presented in Figure 8. For each method, the uncertainty of the reconstruction is mainly dependent on the number and position of the wells. The correlation between inverted and measured hydraulic head data for the three inversion models are presented in the cross-plots of Figure 9.

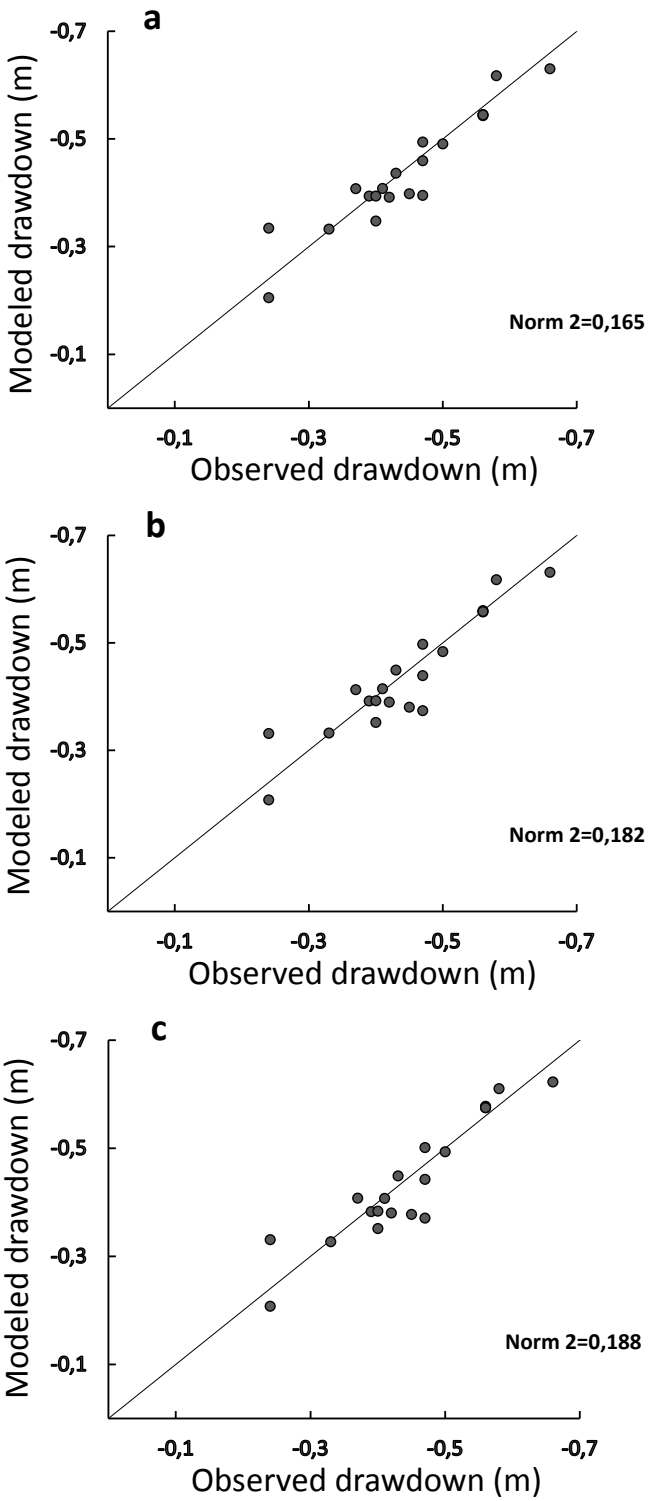

Figure 9 : Graphs showing the differences between the 20 observed drawdowns and modeled drawdowns after convergence of three different inversion methods with 10,000 parameters applied to the experimental site. The graph (a) was obtained with the GA adjoint-state method, the graph (b) with the GA finite-difference method and the graph (c) with the PCGA method with a covariance matrix decomposition of order K=69. Regarding the mathematical norm 2 the GA adjoint-state method has a slightly better convergence on the data than the other methods but the PCGA inversion method is much more efficient for the calculation time (see Table 2).

It can be seen that, with the same inversion inputs, the GA-adjoint-state method generated a slightly better correlation compared to the other two methods, which is reflected by the smaller RMSE values (0.165 m compared to 0.182 m and 0.188 m). The performance of the three methods in terms of simulation time is compared in Table 2.

Table 2 : Comparison of the efficiency between three algorithm of geostatistical inversion methods (GA adjoint-state, GA finite-difference and PCGA) on a same under-determined modeling. Results of these inversions are shown in Figures 8 and 9. The convergence on data was slightly better for an adjoint-state method but the calculation time was considerably reduced by using a PCGA method.

	GA adjoint-state method	**GA finite-difference method**	**PCGA method (truncation order K =69)**
Number of parameters	100×100	100×100	100×100
Computation time	10h 43min	72h 10min	1h 33min
Value of objective function to be minimized after convergence	23.0602	26.6526	27.4641
ℓ_2-norm between observed and modeled data	0.165 m	0.182 m	0.188 m

Table 3 : Convergence times for different methods using different grid sizes. An Intel Xeon QuadCore 2.8GHz with 12Go RAM has been used to perform the computations. The PCGA method (with a truncation error of approximately 1) is always the fastest because it involves less forward problems than the GA finite-difference method and that the Gauss-Legendre resolution of the integral in the GA adjoint-state method requires a calculation of a number of nodes proportional to the number of cells in the grid in each forward problem.

Grid resolution	**GA adjoint-state method**	**GA finite-difference method**	**PCGA method (truncation error $\lambda_{K+1} \simeq 1$)**
10×10	5min	16min	1min
30×30	9min	1h 40min	3min
50×50	45min	8h 21min	6min
80×80	3h 19min	41h 44min	29min
100×100	10h 43min	72h 10min	1h 33min

A significant reduction in computational time is observed for the GA-adjoint-state method compared to the GA-finite difference method (Table 3). This reduction is mainly related to the calculation of the Jacobian sensitivity matrix (Cardiff and Kitanidis 2008). As the grid discretization increases, the significance of reduction in computational time of the GA-adjoint-state method compared to the GA-finite difference based method become more apparent. However, an even more significant time reduction was observed in using the PCGA method. Note that in Table 3 the computation time for PCGA includes the time from both the covariance matrix decomposition and the inversion calculation. The computational time of PCGA is observed to be 10 time less than that of the GA-adjoint state method and 70 time less than that of the GA-finite difference method. Altogether, the advantage of PCGA in obtaining a fast solution without compromising the inversion quality makes it a promising candidate in solving large-scale inversion problems.

Conclusion

The GA finite-difference method is useful and straightforward for inversions with a few parameters and a large number of observational data. In contrast, the GA adjoint-state method is advantageous in dealing with inversion models with a few observational data but a relatively large parameters set. On the contrary, the PCGA is an efficient method for both cases. It is also helpful for extremely under-determined problems where a large number of unknown parameters is present. In fact, the time and memory required by this method to perform the iterative process of the inversion is less sensitive to the number of parameters or measurements, but more dependent on the approximation order of the covariance matrix chosen by the modeler. A higher order approximation will lead to higher computational costs but the error introduced in the inversions will be much smaller.

The comparison of different methodologies has shown that PCGA approach appears to be the most efficient strategy for carrying out large-scale inversions in porous aquifers. We noticed that, in the studied case, the errors introduced by the approximation in the PCGA methods were not significant. Thus, we obtained the same accuracy in results from the PCGA inverse modeling compared to the GA inverse modeling. Thereby, with this strategy, only the principal components of the covariance matrix are kept in the inversion process, and the computational and memory costs necessary for the inversion algorithm are optimized. Additionally, the PCGA method significantly reduces the computational time. With the PCGA method we divided the computation time by seven compared to the GA adjoint state method, and by 50 compared to the GA finite-difference method.

In summary, by applying the PCGA for a hydraulic tomography in a porous aquifer, we found an especially adapted strategy, which produces accurate inversion results with a good resolution in a reduced time, and which manages optimally the computer memory involved in the inversion algorithm. Nevertheless, the PCGA method is efficient specifically for models with a smooth distribution of the targeted parameters (which could typically be used for a good average representation of porous aquifers) so that the covariance matrix can be approximated by much smaller matrices.

Acknowledgments

We would like to thanks Pr. Michael Cardiff and two others anonymous reviewers for their relevant comments on this article. We also thanks the region Normandy for financially supporting the PhD of Pierre Fischer.

References

Ambikasaran, S., J.Y. Li, P.K. Kitanidis and E. Darve. 2013. Large-scale stochastic linear inversion using hierarchical matrices. *Computers & Geosciences* 17: 913-927.

Berg, S.J. and W.A. Illman. 2013. Field study of subsurface heterogeneity with steady-state hydraulic tomography. *Groundwater* 51, no.1: 29-40.

Cardiff, M. and P.K. Kitanidis. 2008. Efficient solution of nonlinear, underdetermined inverse problems with a generalized PDE model. *Computers & Geosciences* 34: 1480-1491.

Cardiff, M., W. Barrash, P.K. Kitanidis, B. Malama, A. Revil, S. Straface and E. Rizzo. 2009. A potential-based inversion of unconfined steady-state hydraulic tomography. *Groundwater* 47, no.2: 259-270.

Cardiff, M., W. Barrash and P.K. Kitanidis. 2013. Hydraulic conductivity imaging from 3-D transient hydraulic tomography at several pumping/observation densities. *Water Resources Research* 49: 7311-7326.

Carrera, J. and S.P. Neuman. 1986a. Estimation of aquifer parameters under transient and steady state conditions: 1.Maximum likelihood method incorporating prior information. *Water Resources Research* 22, no.2: 199-210.

Carrera, J. and S.P. Neuman. 1986b. Estimation of aquifer parameters under transient and steady state conditions: 3.Application to synthetic and field data. *Water Resources Research* 22, no.2: 228-242.

Elsheikh, A.H., M.F. Wheeler and I. Hoteit. 2014. Hybrid nested sampling algorithm for Bayesian model selection applied to inverse subsurface flow problems. *Journal of Computational Physics* 258: 319-337.

Greengard, L. and V. Rokhlin. 1987. A fast algorithm for particle simulations. *Journal of Computational Physics* 73, no.2: 325-348.

Hackbusch, W. and S. Börm. 2002. Data-sparse approximation by adaptive $\mathcal{H}^2$-matrices. *Computing* 69, no.1: 1-35.

Halko, N., P.G. Martinsson and J.A. Tropp. 2011. Finding structure with randomness: Probabilistic algorithms for constructing approximate matrix decompositions. *SIAM Review* 53, no.2: 217-288.

Hoeksema, R.J. and P.K. Kitanidis. 1984. An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling. *Water Resources Research* 20, no.7: 1003-1020.

Kitanidis, P.K. and E.G. Vomvoris. 1983. A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations. *Water Resources Research* 19, no.3: 677-690.

Kitanidis, P.K. 1995. Quasi-linear geostatistical theory for inversing. *Water Resources Research* 31, no.10: 2411-2419.

Kitanidis, P.K. and J. Lee. 2014. Principal component geostatistical approach for large-dimensional inverse problem. *Water Resources Research* 50, no.7: 5428-5443.

Lee, J. and P.K. Kitanidis. 2014. Large-scale hydraulic tomography and joint inversion of head and tracer data using the Principal Component Geostatistical Approach (PCGA). *Water Resources Research* 50, no.7: 5410-5427.

Lieberman, C., K. Fidkowski and K. Willcox, B. van Bloemen Waanders. 2013. Hessian-based model reduction: large-scale inversion and prediction. *International Journal for Numerical Methods in Fluids* 71: 135-150.

Liu, X., Q. Zhou, P.K. Kitanidis and J.T. Birkholzer. 2014. Fast iterative implementation of large-scale nonlinear geostatistical inverse modeling. *Water Resources Research* 50: 198-207.

Nowak, W., S. Tenkleve and O.A. Cirpka. 2003. Efficient computation of linearized cross-covariance and auto-covariance matrices of interdependent quantities. *Mathematical Geology* 35, no.1: 53-66.

Paige, C.C. and M.A. Saunders. 1975. Solution of sparse indefinite systems of linear equations. *SIAM Journal on Numerical Analysis* 12, no.4: 617-629.

Pool, M., J. Carrera, A. Alcolea and E.M. Bocanegra. 2015. A comparison of deterministic and stochastic approaches for regional scale inverse modeling on the Mar del Plata aquifer. *Journal of Hydrology* 531: 214-229.

Rao, S.V.N., B.S. Thandaveswara, S. Murthy Bhallamudi and V. Srinivasulu. 2003. Optimal groundwater management in deltaic regions using simulated annealing and neural networks. *Water Resources Management* 17: 409-428.

Soueid Ahmed, A., A. Jardani, A. Revil and J.P. Dupont. 2014. Hydraulic conductivity field characterization from the joint inversion of hydraulic heads and self-potential data. *Water Resources Research* 50: 1-21.

Soueid Ahmed, A., J. Zhou, A. Jardani, A. Revil and J.P. Dupont. 2015. Image-guided inversion in steady-state hydraulic tomography. *Advances in Water Resources* 82: 83-97.

Tarantola, A. and B. Valette. 1982. Generalized nonlinear inverse problems solved using the least squares criterion. *Reviews of Geophysics and Space Physics* 20, no.2: 219-232.

Wang, X., A. Jardani, H. Jourde, L. Lonergan, J. Cosgrove, O. Gosselin and G., Massonnat. 2016. Characterisation of the transmissivity field of a fractured and karstic aquifer, Southern France. *Advances in Water Resources* 87: 106-121.

Zha, Y., T. J. Yeh, W.A. Illman, T. Tanaka, P. Bruines, H. Onoe and H. Saegusa. 2015. What does hydraulic tomography tell us about fractured geological media? A field study and synthetic experiments. *Journal of Hydrology* 531, no.1:17-30.