Journal of Hydrology 551 (2017) 29–46

Contents lists available at ScienceDirect

Journal of Hydrology

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

Research papers

A hybrid inverse method for hydraulic tomography in fractured and karstic media

Xiaoguang Wang [a,*], Abderrahim Jardani [b], Hervé Jourde [a]

[a] Laboratoire Hydrosciences Montpellier (HSM), Université de Montpellier, CNRS, IRD, Montpellier, France
[b] CNRS UMR 6143, M2C, Morphodynamique Continentale et Côtière, Université de Rouen, Mont-Saint-Aignan, France

ARTICLE INFO

Article history:
Received 24 January 2017
Received in revised form 24 May 2017
Accepted 25 May 2017
Available online 27 May 2017
This manuscript was handled by Corrado Corradini, Editor-in-Chief, with the assistance of Stephen Worthington, Associate Editor

Keywords:
Fractures
Karstified aquifers
Transition probability
Markov chain Monte Carlo
Stochastic Newton
Hydraulic tomography

ABSTRACT

We apply a stochastic Newton (SN) approach to solve a high-dimensional hydraulic inverse problem in highly heterogeneous geological media. By recognizing the connection between the cost function of deterministic optimizations and the posterior probability density of stochastic inversions, the Markov chain Monte Carlo (MCMC) sampler of SN is constructed by two parts: a deterministic part, which corresponds to a Newton step of deterministic optimization, and a stochastic part, which is a Gaussian distribution with the inverse of the local Hessian as the covariance matrix. The hybrid inverse method exploits the efficient tools for fast solution of deterministic inversions to improve the efficiency of the MCMC sampler. To address the ill-posedness of the inverse problem, *a priori* models, generated by a transition-probability geostatistical method, and conditioned to inter-well connection data, are used as regularization constraints.

The effectiveness of the stochastic Newton method is first demonstrated by a synthetic test. The transmissivity field of the synthetic model is highly heterogeneous, and includes sharp variations. The inverse approach was then applied to a field hydraulic tomography investigation in a fractured and karstified aquifer to reconstruct its transmissivity field from a collection of real hydraulic head measurements. From the inversions, a series of transmissivity fields that produce good correlations between the inverted and the measured hydraulic heads were obtained. The inverse approach produced slightly different *a posteriori* transmissivity patterns for different *a priori* structure models of transmissivity; however, the trend and location of the high-transmissivity channels are consistent among various realizations. In addition, the uncertainty associated with each realization of the inverted transmissivity fields was quantified.

1. Introduction

The hydrodynamic prediction of fluid flow and solute transport in subsurface aquifers requires the spatial distributions of hydraulic properties, *e.g.* transmissivity (T) and storativity (S), to be accurately defined. Recently, hydraulic tomography (HT) has been proven to be an efficient technique for characterizing site-specific, spatial distributions of transmissivity (*e.g.* Gottlieb and Dietrich, 1995; Butler et al., 1999; Yin and Illman, 2009; Castagna et al., 2011; Klepikova et al., 2013; Illman, 2014). This method jointly analyses several sets of cross-hole pumping test data, and therefore integrates more hydraulic information contained in the recorded data (Illman, 2014). The efficiency of HT has been demonstrated by numerical (Hao et al., 2008), laboratory (Illman et al., 2007; Liu et al., 2007; Sharmeen et al., 2012) and field experiments (Straface et al., 2007; Illman et al., 2009; Wang et al., 2016).

Numerous inverse methods have been developed for conditioning hydrogeology models to observed hydraulic data and assessing the associated uncertainties (see *e.g.* Carrera et al., 2005 and Illman, 2014). Among others, the most popular ones are the geostatistical inverse methods and the stochastic sampling methods. The geostatistical approaches (*e.g.* Kitanidis and Vomvoris, 1983; Kitanidis, 1995; Yeh et al., 1995, 1996; Kitanidis and Lee, 2014), which formulate the inverse problem as a regularized optimization, are efficient for high-dimensional problems. They have been widely used in many hydrogeological studies (*e.g.* Zhang and Yeh, 1997; Hanna and Yeh, 1998; Zhu and Yeh, 2005; Soueid Ahmed et al., 2014), and has recently been successfully applied to fractured media (Hao et al., 2008; Illman et al., 2009; Castagna et al., 2011; Sharmeen et al., 2012; Zha et al., 2015, 2016). Even though these inverse practices turn out to be successful, the solutions given by the geostatis-

* Corresponding author.
E-mail addresses: xiaoguang.wang@umontpellier.fr (X. Wang), abderrahim.jardani@univ-rouen.fr (A. Jardani), herve.jourde@umontpellier.fr (H. Jourde).

http://dx.doi.org/10.1016/j.jhydrol.2017.05.051
0022-1694/

tical approaches, in many cases, correspond to local minima, and often suffer from inadequate exploration of the parameter space (Bates and Campbell, 2001; Keating et al., 2010).

In recent years, stochastic approaches based on sample-based Bayesian inference have become popular in aquifer characterization (Fu and Gomez-Hernandez, 2009; Vrugt et al., 2008; Mondal et al., 2010; Cui et al., 2011). These methods, using Markov chain Monte Carlo (MCMC) simulations to generate samples from the posterior target distribution of the model parameters, allowing for inference of both the unknown parameters and the uncertainties associated with the analysis. Comparing to geostatistical approach, MCMC simulations allow to avoid local minima and generate global statistical solutions that do not depend on *a priori* models.

In a traditional MCMC simulation that relies on a single-site updating scheme, a large number of forward model runs (equal to the number of parameters) are required to perform a single parameter sweep (Metropolis et al., 1953). The required number of simulations becomes prohibitive for high-dimensional inverse models that include a large number of model parameter. A popular approach to make the sampling adaptive for high-dimensional models is to construct surrogate models or emulators based on polynomial approximations (*e.g.* polynomial chaos or Karhunen–Loève expansions) that reduce the dimension of the problem (*e.g.* Marzouk and Xiu, 2009; Laloy et al., 2013). The approach is often improved with a two-stage scheme to achieve better sampling efficiency (Liu, 2001; Dostert et al., 2006; Cui et al., 2011). However, because of the requirement of smoothness in the model response fields, the emulator-based methods cannot be applied to highly heterogeneous media, such as fractured and karstic aquifers where sharp changes in the spatial distribution of hydraulic head/flux are commonly observed. Another approach for improving the efficiency of MCMC sampler is to adopt a multivariate proposal density (*e.g.* Haario et al., 2006; Fu and Gomez-Hernandez, 2009). For instance, the delayed rejection adaptive metropolis (DRAM) MCMC resorts to constructing an approximation to the posterior covariance matrix to guide the sampling process. Recently, many authors have proposed hybrid MCMC algorithms, in which local derivative information (gradient and Hessian for the forward model) is used in the construction of the proposal density (Girolami and Calderhead, 2011; Flath et al., 2011; Martin et al., 2012). Because the proposal distributions are tuned using derivative information contained in the past chain, the proposal distribution evolves to the target density much more efficiently. Although using Hessian information in conditioning MCMC samplers has been previously considered, most of the applications are restricted to models with a relatively small number of parameters, or to high-dimensional linear problems where the derivative information (gradient and Hessian) is easy to obtain (Flath et al., 2011). The applicability of the method to complex, nonlinear, hydrogeological models remains to be demonstrated. Also, there are only a few MCMC applications in the context of hydraulic tomography (Oliver et al., 1997; Fu and Gomez-Hernandez, 2009; Jardani et al., 2012). Therefore, in the present study, we have attempted to apply a Hessian-based MCMC method, originally proposed in Martin et al. (2012), to infer the posterior distribution of the unknown and uncertain parameters in a high-dimensional groundwater model.

No matter which type of inverse method is applied, inverse problems have an essential difficulty: non-uniqueness (Tarantola, 2005). Consequently, there are many models can fit the data equally well. To overcome this difficulty, the inversion models have to be constrained by *a priori* information. Conventionally, variogram models are adopted to pose assumptions about the spatial pattern of the model parameter field, in order to constrain the solution. However, the classical variogram models based on kriging or cokriging processes generate smooth spatial patterns. Therefore, they are not suitable for modelling fractured and karstified media. To capture the prominent heterogeneous nature and sharp property variation of fractured aquifers, indicator geostatistical methods based on categorical classification are developed (Goovaerts, 1996; Day-Lewis et al., 2000). However, the traditional indicator approaches require a large amount of transmissivity measurements. In the case where only sparse transmissivity data are available, the emergent categorical realizations are often geologically unrealistic (Illman, 2014). An alternative and promising method to model highly heterogeneous media is the transitional-probability (TP) based categorical approach (Ritzi et al., 1995; Carle and Fogg, 1996). The approach employs measured transition probabilities between different rock types as the weights of their spatial cross-correlation. The method has been extensively used in soil surveys and successfully applied to alluvial deposits (Carle and Fogg, 1997; Weissmann et al., 1999; Weissmann and Fogg, 1999; Lee et al., 2007), glacial deposits (He et al., 2014), and recently to fractured media (Park et al., 2004; Blessent et al., 2011). This method has also been used in the present work to generate *a priori* transmissivity structure fields, where the salient heterogeneity and anisotropy of fracture and karst networks are captured.

Our analysis differs from previous studies (*e.g.* Park et al., 2004 and Blessent et al., 2011) in how the spatial patterns are generated from TP geostatistical approach, and are subsequently calibrated to match hydraulic data. In Park et al. (2004) and Blessent et al. (2011) a classical two-stage approach (*e.g.* McKenna and Poeter, 1995; Day-Lewis et al., 2000) is used. In the first stage, the TP spatial patterns generated by fitting geostatistical models (transiograms) and hard data are treated as stochastic continuum models (i.e. rock type bodies being transformed into stochastic transmissivity zones for a flow model). In the second stage, for each realization, a set of mean hydraulic properties of transmissivity zones is then calibrated to match hydraulic data. In other words, variation of hydraulic properties within each rock type is not allowed. In such cases, history matching of hydraulic data can only be achieved to a limited level where the general trend displayed by the data is preserved, while disregarding a large amount of valuable information that may be crucial to infer fine-scale heterogeneity.

In our approach, we still follow the two-stage strategy. But, in the first stage, instead of fitting available sparse transmissivity estimates to transiogram models, we specify the explicitly the embedded transition probability matrix, based on integrated hydrogeological interpretation, to derive Markovian categorical rock-type realizations (Section 3.1). The stochastic realizations are conditioned to transmissivity estimates and inter-borehole connectivity both are determined from field cross-hole pumping tests. In the second stage, the generated TP realizations is used as *a priori* structural model of transmissivity, and we then apply a hybrid MCMC method (Section 3.2) to statistically infer the spatially varying transmissivity (mean and variance) of grid cells in selected parameter zones. In this sense, our inversion approach is based on an equivalent porous medium representation.

2. Field site and numerical model

2.1. Site description and main hydraulic dataset

The location of this study is the Terrieu experimental site, which is located in the Montpellier region, Southern France (Fig. 1). The field site has a surface area of 1500 m^2 (30 m $\times$ 50 m), and contains 22 boreholes that were drilled through the Cretaceous marly limestones of the Lez aquifer (total