



## UvA-DARE (Digital Academic Repository)

### Efficient posterior exploration of a high-dimensional groundwater model from two-stage MCMC simulation and polynomial chaos expansion

Laloy, E.; Rogiers, B.; Vrugt, J.A.; Mallants, D.; Jacques, D.

**DOI**

[10.1002/wrcr.20226](https://doi.org/10.1002/wrcr.20226)

**Publication date**

2013

**Document Version**

Final published version

**Published in**

Water Resources Research

[Link to publication](#)

**Citation for published version (APA):**

Laloy, E., Rogiers, B., Vrugt, J. A., Mallants, D., & Jacques, D. (2013). Efficient posterior exploration of a high-dimensional groundwater model from two-stage MCMC simulation and polynomial chaos expansion. *Water Resources Research*, 49(5), 2664-2682. <https://doi.org/10.1002/wrcr.20226>

**General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

**Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

*UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)*

# Efficient posterior exploration of a high-dimensional groundwater model from two-stage Markov chain Monte Carlo simulation and polynomial chaos expansion

Eric Laloy,<sup>1</sup> Bart Rogiers,<sup>1,2</sup> Jasper A. Vrugt,<sup>3,4</sup> Dirk Mallants,<sup>5</sup> and Diederik Jacques<sup>1</sup>

Received 18 November 2012; revised 7 March 2013; accepted 28 March 2013; published 28 May 2013.

[1] This study reports on two strategies for accelerating posterior inference of a highly parameterized and CPU-demanding groundwater flow model. Our method builds on previous stochastic collocation approaches, e.g., Marzouk and Xiu (2009) and Marzouk and Najm (2009), and uses generalized polynomial chaos (gPC) theory and dimensionality reduction to emulate the output of a large-scale groundwater flow model. The resulting surrogate model is CPU efficient and serves to explore the posterior distribution at a much lower computational cost using two-stage MCMC simulation. The case study reported in this paper demonstrates a two to five times speed-up in sampling efficiency.

**Citation:** Laloy, E., B. Rogiers, J. A. Vrugt, D. Mallants, and D. Jacques (2013), Efficient posterior exploration of a high-dimensional groundwater model from two-stage Markov Chain Monte Carlo simulation and polynomial chaos expansion, *Water Resour. Res.*, 49, 2664–2682, doi:10.1002/wrcr.20226.

## 1. Introduction

[2] Parameterization of groundwater flow models can be a challenging task, especially for heterogeneous media [e.g., Koltermann and Gorelick, 1996]. Hydraulic conductivity and porosity are difficult to measure in the field and poorly known a priori. This necessitates the use of inverse methods that indirectly estimate these properties from other observed subsurface variables. In this procedure, one iteratively proposes new parameter values until the model fits the observed data up to a prespecified precision. Yet, high-parameter dimensionality and significant model nonlinearities pose considerable challenges to inverse methods. Also, incomplete knowledge of the hydrogeology of the site of interest and presence of measurement errors introduce considerable uncertainty in the model parameters and predictions. This is in part due to the limited information content of the available calibration data, typically groundwater head observations.

[3] Furthermore, another important source of uncertainty originates from sparse measurement data that do not contain sufficient information to reliably constrain flow and transport properties of the hydrogeologic domain under

consideration. This might result in an ill-posed inverse problem with multiple (parameter) solutions that fit the data equally well. Quantification of uncertainty is not only of crucial importance in decision making but also helps diagnose which components of the model are well constrained by the available data, and which processes in the model are not very well resolved [e.g., Carrera and Neuman, 1986a, 1986b; Kitanidis, 1986, 1997; Carrera et al., 2005; Cooley and Christensen, 2006; Hernandez et al., 2006; Tonkin and Doherty, 2009; Keating et al., 2010; Trolborg et al., 2010; Nowak et al., 2010; Zeng and Zhang, 2010, and many others].

[4] In recent years, Bayesian approaches are becoming increasingly popular for aquifer and reservoir characterization, and parameter and model predictive uncertainty analysis [e.g., Oliver et al., 1997; Efendiev et al., 2005; Fu and Gomez-Hernandez, 2009; Mariethoz et al., 2010; Mondal et al., 2010; Cui et al., 2011]. These methods allow for a rigorous statistical analysis of uncertainty [Kennedy and O'Hagan, 2001], and use Markov chain Monte Carlo (MCMC) simulation to generate samples from the posterior target distribution. Despite significant advances in sampling efficiency [e.g., Haario et al., 2001; Vrugt et al., 2003; Christen and Fox, 2005; Haario et al., 2006; ter Braak, 2006; ter Braak and Vrugt, 2008; Vrugt et al., 2008; Cui et al., 2011; Laloy and Vrugt, 2012], MCMC simulation might require a prohibitively large number of model calls, especially in the presence of high-parameter dimensionality. This is particularly true for CPU-intensive hydrogeologic models, and hence MCMC simulation becomes a daunting task [e.g., Tonkin and Doherty, 2005, 2009; Keating et al., 2010].

[5] Marzouk and Xiu [2009] recently proposed to extend the stochastic collocation approach of Xiu [2007] to Bayesian solution of nonlinear inverse problems. This method uses the generalized polynomial chaos (gPC) framework to construct a polynomial approximation of the forward

<sup>1</sup>Belgian Nuclear Research Centre, Institute for Environment, Health and Safety Mol, Belgium.

<sup>2</sup>Department of Earth and Environmental Sciences, KU Leuven, Heverlee, Belgium.

<sup>3</sup>Department of Civil and Environmental Engineering, University of California Irvine, Irvine, USA.

<sup>4</sup>Institute for Biodiversity and Ecosystems Dynamics, University of Amsterdam, Amsterdam, Netherlands.

<sup>5</sup>CSIRO Land and Water, Urrbrae South Australia, Australia.

Corresponding author: E. Laloy, Belgian Nuclear Research Centre, Institute for Environment, Health and Safety, Mol 2400, Belgium. (elaloy@sckcen.be)

solution over the support of the prior parameter distribution. This approximation defines a surrogate posterior distribution that can be derived at a minimal computational cost. Convergence properties of this approach have been studied in detail by *Marzouk and Xiu* [2009]. The rate of convergence of the surrogate posterior to the true posterior depends in large part on the smoothness of the original forward model. For relatively simple low-dimensional models significant speed ups in sampling efficiency have been reported [*Marzouk and Xiu*, 2009]. Yet, applicability of this methodology to complex parameter-rich models, common in groundwater hydrology, has still to be demonstrated. Such models pose specific difficulties, in particular because the (spatially) simulated head and fluxes might not vary smoothly enough with changes in the parameter values, and the parameter space might simply be too large to develop a sufficiently accurate emulator. Another complicating factor is that groundwater models are typically CPU-demanding. This severely limits our ability to exhaustively sample the parameter space and build a reliable surrogate model that provides an accurate characterization of the posterior distribution. The scope of the present paper is thus to extend the technique further in order to derive the posterior distribution of a highly parameterized, and computationally expensive groundwater model.

## 2. Approach

[6] In this work we propose and evaluate two strategies to accelerate Bayesian inference of a highly parameterized and CPU-demanding groundwater flow model. Our first approach proceeds in the two following steps. First, a gPC-based surrogate model is constructed using the least possible amount of deterministic forward simulations, and the corresponding surrogate posterior distribution is subsequently derived with MCMC simulation at a negligible computational cost. Next, samples from this surrogate distribution are then used to initialize the  $\text{DREAM}_{(ZS)}$  algorithm [*ter Braak and Vrugt*, 2008; *Vrugt et al.*, 2009; *Laloy and Vrugt*, 2012] and MCMC simulation proceeds with the original flow simulator. Note that this first strategy is similar in spirit to the recently published approach of *Zeng et al.* [2012] in which draws from an interpolation-based surrogate posterior is employed as initial position for posterior exploration of a lowly parameterized transport model using the DRAM [*Haario et al.*, 2006] sampler.

[7] The second strategy uses the gPC-based surrogate model derived following *Xiu* [2007] within a two-stage MCMC simulation scheme [e.g., *Liu*, 2001; *Christen and Fox*, 2005; *Efendiev et al.*, 2005; *Kaipio and Somersalo*, 2007; *Mondal et al.*, 2010; *Cui et al.*, 2011], using  $\text{DREAM}_{(ZS)}$ . In this two-stage MCMC framework, a reduced-order (either coarse or approximated) model is typically used that approximates the original model but operates at a much coarser spatial resolution, possibly with simplified processes. Then, each time a proposal is generated using MCMC simulation, the reduced-order model is first evaluated and only those proposals that are accepted are being evaluated by the original model. We combine herein the latest advances in two-stage MCMC simulation [*Cui et al.*, 2011] with polynomial chaos expansion (PCE) to create a reduced-order model in the first evaluation step.

This approach has some common features with the study of *Dostert et al.* [2009] who used standard two-stage MCMC simulation with a reduced order model based on stochastic collocation and Lagrange interpolation. A salient feature of the two-stage MCMC methodology used herein is the adaptive construction of a stochastic model for the reduced order to original model error [*Cui et al.*, 2011], which considerably improves efficiency over the classical two-stage MCMC scheme.

[8] We illustrate our methodology using the hydrogeological characterization of a three-dimensional multilayered aquifer. Steady-state groundwater flow is simulated using MODFLOW-2005 [*Harbaugh*, 2005] and measured heads serve as calibration data. As the inverse problem involves more than 26,000 unknown parameters, we take advantage of the Karhunen-Loève [*Loève*, 1977] transform to significantly reduce the dimensionality of the parameter space. Note that this dimensionality reduction technique is commonly employed to represent Gaussian spatial fields such as observed for the log conductivity, and several studies have used this approach in conjunction with MCMC simulation [e.g., *Dostert et al.*, 2006, 2009; *Marzouk and Najm*, 2009].

[9] We would like to stress that polynomial chaos based surrogates have mainly been used in the hydrologic literature for stochastic analysis with known parameter distributions [e.g., *Li and Zhang*, 2007; *Li et al.*, 2009] or global sensitivity analysis [e.g., *Fajraoui et al.*, 2011; *Ciriello et al.*, 2013], but rarely within the context of parameter estimation [some notable exceptions are *Saad and Ghanem*, 2009; *Zeng and Zhang*, 2010; *Fajraoui et al.*, 2012] and MCMC sampling. Indeed, the recent review of *Razavi et al.* [2012] on surrogate modeling in water resources does not address polynomial chaos expansion. Those studies that have used gPC with MCMC simulation [*Dostert et al.*, 2009; *Marzouk and Najm*, 2009; *Marzouk and Xiu*, 2009] were limited to relatively simple case studies, often involving synthetic data, low-dimensional models and simple search methods.

[10] Thus the main novelty of our work lies in the state-of-the-art techniques implemented for speeding up MCMC simulation of hydrologic models with gPC, together with a challenging CPU demanding and highly parameterized real-world application. This paper therefore presents an important step forward as it bridges the gap between earlier theoretical work on relatively simple simulation models [e.g., *Dostert et al.*, 2009; *Marzouk and Najm*, 2009; *Marzouk and Xiu*, 2009] and more complex real-world inference problems.

[11] We foresee several potential advantages of gPC with two-stage MCMC simulation. While a surrogate model based on a coarser mesh and/or model abstraction may involve difficult decisions about the extent to which the physics of the model can be consistently simplified, gPC expansion is relatively straightforward. Second, the gPC-based surrogate model is computationally cheap compared to the original model. This may not necessarily be the case for reduced-order models that nevertheless require numerical solution of partial differential equations in space and/or time. Third, the gPC surrogate can be setup such that it uses the exact same calibration parameters as the original model. The surrogate posterior derived from

MCMC simulation is therefore easily analyzed and used in the original model. Finally, it is worth noting that we employ herein a linear gPC, that is, a gPC of order 1. This approximation suffices for the purpose and high dimensionality of the present study, but arguably is rather simple. Constructing accurate higher-order approximations was found to be computationally intractable.

[12] This paper is organized as follows. Section 3 presents the theoretical underpinning of the inversion strategy used herein. This is followed in section 4 with application to a highly parameterized groundwater model. In this section, we are especially concerned with the CPU time and compare the results of the proposed inversion techniques against a classical MCMC inversion of the original model. Finally, in section 5 we provide a summary with conclusions.

### 3. Methods

[13] This section describes the inversion methodologies followed herein. We first provide a brief description of the Karhunen-Loève (KL) representation of a spatial Gaussian process such as the log conductivity field. We then give an overview of the generalized polynomial chaos, and the stochastic collocation method of *Xiu* [2007]. This section concludes with a description of the traditional and two-stage MCMC strategies used herein to sample the posterior parameter distributions of the surrogate and original flow models.

#### 3.1. Karhunen-Loève Expansion

[14] Let's consider a three dimensional spatial domain  $\mathbf{D} \subset \mathbb{R}^3$ , let  $\mathbf{x}$  be the spatial location,  $\mathbf{Y}(\mathbf{x})$  be a real-valued random field with mean  $\mu(\mathbf{x})$  and a parametric covariance kernel  $C$  that is continuous on  $\mathbf{D} \times \mathbf{D}$ . Then, the KL expansion of  $\mathbf{Y}(\mathbf{x})$  is given by [e.g., *Marzouk and Najm*, 2009]

$$\mathbf{Y}(\mathbf{x}) = \mu(\mathbf{x}) + \sum_{i=1}^{\infty} \sqrt{\lambda_i} c_i \phi_i(\mathbf{x}), \quad (1)$$

where  $\phi_i(\mathbf{x})$  and  $\lambda_i$  are eigenfunctions and eigenvalues of the covariance kernel  $C$

$$\int_{\mathbf{D}} C(\mathbf{x}_1, \mathbf{x}_2) \phi_i(\mathbf{x}_2) d\mathbf{x}_2 = \lambda_i \phi_i(\mathbf{x}_1). \quad (2)$$

[15] The covariance kernel  $C$  is symmetric and positive semidefinite and thus we have

$$C(\mathbf{x}_1, \mathbf{x}_2) = \sum_{i=1}^{\infty} \lambda_i \phi_i(\mathbf{x}_1) \phi_i(\mathbf{x}_2), \quad (3)$$

where the eigenfunctions  $\phi_i(\mathbf{x})$  form a complete orthonormal system in  $L^2(\mathbf{D})$ . The random variables  $c_i$ , hereafter referred as KL modes, are uncorrelated with zero mean and unit variance. If  $\mathbf{Y}(\mathbf{x})$  is multi-Gaussian, then by definition the  $c_i$  are standard normal,  $c_i \sim N(0, 1)$ .

[16] For a concise representation of the considered random field, one can truncate the KL expansion up to a desired accuracy using the  $k$  largest eigenvalues

$$\mathbf{Y}_k(\mathbf{x}) = \mu(\mathbf{x}) + \sum_{i=1}^k \sqrt{\lambda_i} c_i \phi_i(\mathbf{x}), \quad (4)$$

with covariance function of  $\mathbf{Y}_k(\mathbf{x})$  given by

$$C_k(\mathbf{x}_1, \mathbf{x}_2) = \sum_{i=1}^k \lambda_i \phi_i(\mathbf{x}_1) \phi_i(\mathbf{x}_2). \quad (5)$$

[17] The number  $k$  is commonly taken so that 95% of the total variance or “energy” of  $\mathbf{Y}(\mathbf{x})$  is preserved

$$\int_{\mathbf{D}} E(\mathbf{Y}(\mathbf{x}) - \mu(\mathbf{x}))^2 d\mathbf{x} = \int_{\mathbf{D}} C(\mathbf{x}, \mathbf{x}) d\mathbf{x} = \sum_{i=1}^{\infty} \lambda_i, \quad (6)$$

$$\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^{\infty} \lambda_i} = 0.95.$$

[18] A complete derivation of the KL expansion, can be found in, e.g., *Marzouk and Najm* [2009] and interested readers are referred to this publication for more details.

[19] In this work, we assume a Gaussian covariance model of the form

$$C(\mathbf{x}_1, \mathbf{x}_2) = \sigma^2 \exp \left( -3 \left[ \frac{(x_{1,x} - x_{2,x})^2}{L_x^2} + \frac{(x_{1,y} - x_{2,y})^2}{L_y^2} + \frac{(x_{1,z} - x_{2,z})^2}{L_z^2} \right] \right), \quad (7)$$

where  $\sigma^2$  is referred to in the hydrogeologic literature as the sill and taken as the field variance, and  $L_x, L_y$ , and  $L_z$  denote the correlation lengths or ranges along the spatial dimensions  $x, y$ , and  $z$ , respectively.

#### 3.2. Generalized Polynomial Chaos

[20] Generalized polynomial chaos (gPC) is an approximation methodology that emulates the output of a simulation model using orthogonal polynomials. The parameters in these polynomials are transformations of the parameters of the original CPU-intensive model. The gPC has been introduced in the field of engineering by *Ghanem and Spanos* [1991] and subsequently used by various authors to build simple CPU-efficient surrogate models [*Xiu and Karniadakis*, 2002; *Le Maître and Knio*, 2010]. We provide a general overview of the gPC approach herein and refer the reader to the cited literature and references therein for a more comprehensive description of this methodology.

[21] Let  $\mathbf{z} = (z_1, \dots, z_n) \in \mathbb{R}^n, n \geq 1$  be a set of independent continuous random variables characterizing the yet unknown parameter values of a system of partial differential equations. We assume each  $z_i, i = 1, \dots, n$  to have a prior probability density function (pdf)  $p(z_i)$ . The joint prior pdf of  $\mathbf{z}$  therefore becomes

$$p(\mathbf{z}) = \prod_{i=1}^n p(z_i). \tag{8}$$

[22] Let  $\mathbf{d} = (d_1, \dots, d_N) \in \mathbb{R}^N, N \geq 1$  be a set of true measurements of one variable, and  $F$  be a deterministic, error-free “forward model” that expresses the relation between the uncertain parameters  $\mathbf{z}$  and  $\mathbf{d}$

$$F(\mathbf{z}) = \mathbf{d}. \tag{9}$$

[23] In the context of groundwater flow and transport modeling,  $F$  typically involves numerical solution of a large system of partial differential equations (PDEs). Note that model and measurement errors will be considered later on.

[24] In the following, we describe the gPC expansion of equation (9) for  $N = 1$ . If  $N > 1$ , it is straightforward to apply the approximation separately to each component of  $F$ . The  $P$ -degree gPC approximation of  $F(\mathbf{z})$  is given by

$$F_P(\mathbf{z}) = \sum_{j=0}^{U-1} a_j \Psi_j(\mathbf{z}), \tag{10}$$

where the deterministic coefficients  $a_1, \dots, a_U$  are unknown, and  $U$  denotes the total number of  $n$ -dimensional orthogonal polynomials  $\Psi_j(\mathbf{z})$  of degree not exceeding  $P$  in a total-order expansion. The value of  $U$  is simply computed as  $U = \frac{(P+n)!}{P!n!}$ . It is worth noting that a full tensor-product expansion may also be employed. In this case, the polynomial order in each dimension of the parameter space is at most  $P$ , with total number of polynomial basis functions  $(P + 1)^n$ . The latter grows prohibitively fast with parameter dimensionality and is therefore rarely used in practice. The  $\Psi_j(\mathbf{z})$  are products of the monodimensional polynomials for expansion terms  $j = 1, \dots, U$ :  $\Psi_j(\mathbf{z}) = \psi_{j,1}(z_1) \times \psi_{j,2}(z_2), \dots, \times \psi_{j,n}(z_n)$ .

[25] Various polynomial types can be used in equation (10) depending on the available prior information about the variable  $z_i$ . For instance, Jacobi polynomials are typically associated with a beta distribution, Hermite polynomials with a Gaussian distribution, Laguerre polynomials with a gamma distribution, and Legendre polynomials, that represent a special case of the Jacobi polynomials, with a uniform distribution [Xiu and Karniadakis, 2002]. Different polynomial bases can also be combined, which is done in the present study, where a mixture of Legendre and Hermite polynomials is used (see later).

[26] According to classical approximation theory, the convergence of  $F_P(\mathbf{z})$  to  $F(\mathbf{z})$  as  $P \rightarrow \infty$  requires  $F(\mathbf{z})$  to be square integrable with respect to  $p(\mathbf{z})$ . The rate of convergence is controlled by the regularity of  $F(\mathbf{z})$  [e.g., Le Maître and Knio, 2010], and hence the smoother the forward model, the faster the convergence.

### 3.3. Spectral Projection

[27] Linear regression and spectral projection are two common methods to compute the expansion coefficients  $a_j$  in equation (10) without affecting the original model equations. The first method uses a single least squares solution to solve for the complete set of coefficients that best match the respective predictions of the original CPU-intensive

model at predefined grid points. This requires the use of at least  $\frac{(P+n)!}{P!n!}$  parameter combinations, and thus forward model simulation runs, to warrant a well posed problem. The spectral approach, which is used herein, obtains the  $a_j$  expansion coefficients by projecting the model response against each individual basis function. The  $a_j$  are defined as follows

$$a_j = \frac{\langle F(\mathbf{z}), \Psi_j(\mathbf{z}) \rangle}{\langle \Psi_j^2(\mathbf{z}) \rangle} = \frac{1}{\langle \Psi_j^2(\mathbf{z}) \rangle} \int F(\mathbf{z}) \Psi_j p(\mathbf{z}) d\mathbf{z}. \tag{11}$$

[28] The denominator in equation (11) is the inner product of the respective multivariate orthogonal polynomial, which can be computed for  $j = 0, \dots, U - 1$  using

$$\langle \Psi_j^2(\mathbf{z}) \rangle = \prod_{i=1}^n \langle \psi_{i,j}^2 \rangle, \tag{12}$$

where the univariate inner products of each respective polynomial have a simple closed form solution for which analytical expressions are readily available. The computational effort therefore resides in the time it takes to evaluate the numerator in equation (11).

[29] In the spectral projection framework of Xiu [2007], the expansion coefficients,  $\tilde{a}_j$ , are approximated using

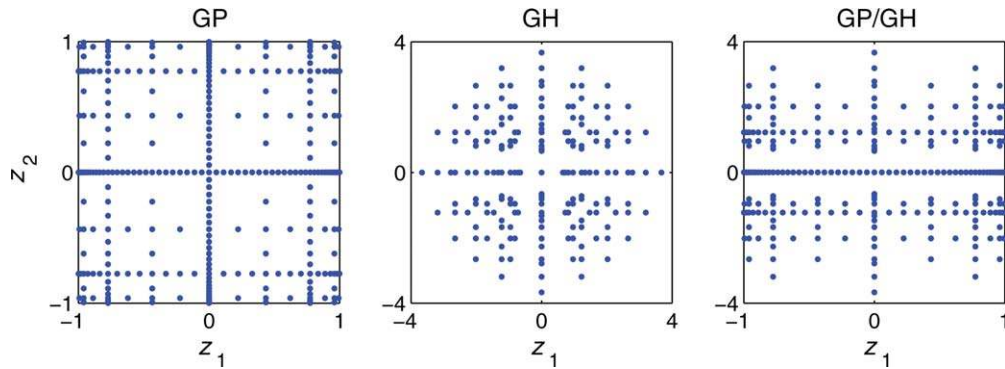
$$\tilde{a}_j = \frac{\sum_{k=1}^Q F(\mathbf{z}_k) \Psi_j(\mathbf{z}_k) w_k}{\langle \Psi_j^2(\mathbf{z}) \rangle}, \tag{13}$$

where the vector  $\mathbf{z}_k$  denotes one of the  $Q$  points in the model parameter space, and the  $w_k$ 's are the associated weights of an integration (cubature) rule on  $\mathbb{R}^n$ . This leads to the following approximation of  $F_P$  in equation (10)

$$\tilde{F}_P(\mathbf{z}) = \sum_{j=0}^{U-1} \tilde{a}_j \Psi_j(\mathbf{z}). \tag{14}$$

[30] The use of an integration rule in equation (13) causes  $F_P$  to differ from  $\tilde{F}_P$ . This deviation is also referred to as the “aliasing” error [Xiu, 2007] and decreases with  $Q$  in equation (13). Hence,  $\tilde{F}_P$  converges to the true forward solver,  $F$ , as  $P \rightarrow \infty$  and  $Q \rightarrow \infty$ . This introduces a tradeoff between the accuracy of the surrogate model  $\tilde{F}_P$  and associated computational burden.

[31] A key issue in the use of the spectral projection framework is the choice of the nodal set  $\{\mathbf{z}_k, w_k\}$  in equation (13). Ideally, the integration rule should be as accurate and efficient as possible using the least number of nodal points and thus deterministic forward simulations. Several approaches have been proposed in the literature to approximate a multidimensional integral by a weighted discrete sum. Tensor product formulas, e.g., Gauss quadratures, are probably the most effective but are rather inefficient and cumbersome in parameter spaces of more than five to six dimensions. Indeed, if  $m$  denotes the prescribed number of points in each parameter dimension, the total number of sampling points obeys the simple quadrature rule  $Q = m^n$



**Figure 1.** Illustration of a mixed sparse grid in two dimensions,  $z_1$  and  $z_2$ . (a) Level five (321 nodes) Gauss-Patterson (GP) sparse grid for uniformly distributed variables, (b) level five (181 nodes) Gauss-Hermite (GH) sparse grid for standardized normally distributed variables, and (c) associated level five (261 nodes) mixed Gauss-Patterson/Gauss-Hermite (GP/GH) sparse grid in which  $z_1$  has a uniform distribution and  $z_2$  a standard normal distribution.

and therefore grows exponentially with parameter dimensionality. In contrast, sparse grid approximation methods [e.g., Smolyak, 1963; Gerstner and Griebel, 2003; Xiu and Hesthaven, 2005; Babuska et al., 2007; Nobile et al., 2008a, 2008b] can drastically reduce the required number of collocation points while preserving a high level of accuracy for moderately large dimensional parameter spaces.

[32] As we consider both uniform and standard normal prior parameter distributions (more details to follow), we employ a mixture of Legendre and Hermite polynomials using a mixed sparse grid whose factors are distinct one-dimensional quadrature rules. More specifically, we combine a Gauss-Patterson (GP) rule on the interval  $[-1, 1]$  for the uniform variables to be employed with the Legendre polynomials, with a Gauss-Hermite (GH) rule defined over  $[-\infty, \infty]$  for the standard normal random parameters of the Hermite polynomials. This is schematically illustrated in Figure 1. The computer program and implementation details can be downloaded from [http://people.sc.fsu.edu/~jburkardt/m\\_src/sparse\\_grid\\_mixed/sparse\\_grid\\_mixed.html](http://people.sc.fsu.edu/~jburkardt/m_src/sparse_grid_mixed/sparse_grid_mixed.html).

### 3.4. Bayesian Inference

[33] Acknowledging that measurement error and model inadequacy are inevitable, a common stochastic representation of the inverse problem is

$$F(\mathbf{z}) = \mathbf{d} + \mathbf{e}, \quad (15)$$

where the noise term  $\mathbf{e}$  lumps measurement and model errors. In this work, it is assumed that  $\mathbf{e}$  follows a zero-mean multivariate Gaussian distribution, i.e.,  $\mathbf{e} \sim N(0, \Sigma_e)$ . Prior information can help refine this assumption. Given reasonably sparse measurements, we further consider a constant error variance,  $\sigma_e^2$ , and thus  $\Sigma_e = \sigma_e^2 \mathbf{I}$ , where  $\mathbf{I}$  signifies the identity matrix. These assumptions will be verified a posteriori.

[34] In the Bayesian paradigm, the unknown model parameters  $\mathbf{z}$  are viewed as random variables with posterior pdf  $p(\mathbf{z}|\mathbf{d})$  given by

$$p(\mathbf{z}|\mathbf{d}) = \frac{p(\mathbf{z})p(\mathbf{d}|\mathbf{z})}{p(\mathbf{d})} \propto p(\mathbf{z})L(\mathbf{z}|\mathbf{d}), \quad (16)$$

where  $L(\mathbf{z}|\mathbf{d}) \equiv p(\mathbf{d}|\mathbf{z})$  signifies the likelihood function of  $\mathbf{z}$ . The normalization factor  $p(\mathbf{d}) = \int p(\mathbf{z})p(\mathbf{d}|\mathbf{z})d\mathbf{z}$  is obtained from numerical integration over the parameter space so that  $p(\mathbf{z}|\mathbf{d})$  scales to one. In the remainder of this paper, we will focus on the unnormalized posterior  $p(\mathbf{z}|\mathbf{d}) \propto p(\mathbf{z})L(\mathbf{z}|\mathbf{d})$ . If we assume  $\mathbf{e}$  to be Gaussian distributed, uncorrelated and with constant variance, the likelihood function can be written as

$$L(\mathbf{d}|\mathbf{z}) = \left( \frac{1}{\sqrt{2\pi\sigma_e^2}} \right)^N \exp\left( -\frac{1}{2}\sigma_e^{-2} \sum_{i=1}^N [d_i - F_i(\mathbf{z})]^2 \right). \quad (17)$$

### 3.5. MCMC Simulation of $p(\mathbf{z}|\mathbf{d})$

[35] An exact analytical solution of  $p(\mathbf{z}|\mathbf{d})$  is not available in many practical cases, and we therefore resort to MCMC simulation to generate samples from the posterior distribution. The basis of this technique is a Markov chain that generates a random walk through the search space and iteratively finds parameter sets with stable frequencies stemming from the parameter posterior pdf. To explore the posterior target, an MCMC algorithm generates trial moves from the current state of the Markov chain,  $\mathbf{z}$ , to a new configuration  $\mathbf{z}'$ . The earliest MCMC approach is probably the well-known random walk Metropolis (RWM) sampler [Metropolis et al., 1953]. Assuming that a random walk has already sampled the points  $\{\mathbf{z}_0, \dots, \mathbf{z}\}$ , this algorithm proceeds in the following three steps. First, a candidate point  $\mathbf{z}'$  is sampled from a proposal distribution  $q(\cdot, \cdot)$  that is symmetric,  $q(\mathbf{z}, \mathbf{z}') = q(\mathbf{z}', \mathbf{z})$  and may depend on the present location,  $\mathbf{z}$ . Second, the candidate point is either accepted or rejected using the Metropolis probability ratio:

$$\alpha(\mathbf{z}, \mathbf{z}') = \min\left[ \frac{p(\mathbf{z}'|\mathbf{d})}{p(\mathbf{z}|\mathbf{d})}, 1 \right]. \quad (18)$$

[36] Finally, if the proposal is accepted, the chain moves to  $\mathbf{z}'$ ; otherwise the chain remains at  $\mathbf{z}$ .

[37] This produces a Markov chain which, under certain regularity conditions, has a unique stationary distribution with pdf  $p(\mathbf{z}|\mathbf{d})$ . *Hastings* [1970] extended equation (18) to include nonsymmetrical proposal distributions, i.e.,  $q(\mathbf{z}, \mathbf{z}') \neq q(\mathbf{z}', \mathbf{z})$ , in which a proposal jump to  $\mathbf{z}'$  and the reverse jump do not have equal probability. The resulting algorithm is called the Metropolis-Hastings (MH) algorithm, and has become the main building block of many existing MCMC sampling schemes.

[38] In practice, the MCMC sampling efficiency strongly depends on the assumed proposal distribution,  $q(\cdot, \cdot)$ , used to generate transitions in the Markov chain. When the proposal distribution is too wide, very few candidate points will be accepted, and the chain will converge rather slowly to the posterior target distribution. On the other hand, if the proposal distribution is too narrow, the chain will remain in close vicinity of its current location, and a prohibitively large number of iterations is needed before the entire posterior distribution has been explored. The choice of the proposal distribution is therefore of crucial importance in determining the efficiency of MCMC simulation.

### 3.6. Accelerating Posterior Exploration of a Complex CPU-Intensive Flow Model

[39] The gPC-based Bayesian algorithm of *Marzouk and Xu* [2009] simply replaces the true but unknown forward solution of equation (15) with the surrogate solution (14) to compute the likelihood of  $\mathbf{z}$ . This leads to an approximated (surrogate) posterior distribution  $\tilde{p}(\mathbf{z}|\mathbf{d})$ . If  $\tilde{F}_P$  is sufficiently close to the true model,  $\tilde{p}(\mathbf{z}|\mathbf{d})$  will converge to  $p(\mathbf{z}|\mathbf{d})$  and posterior inference can thus be achieved at negligible computational cost. Unfortunately, our preliminary tests (not shown herein) demonstrated that mildly and highly parameterized groundwater flow models may be hard to approximate accurately at an affordable computational cost for constructing  $\tilde{F}_P$ . The assumption that  $p(\mathbf{z}|\mathbf{d}) \approx \tilde{p}(\mathbf{z}|\mathbf{d})$  may hence be invalid, especially with increasing dimensionality of the parameter space. This is the motivation for the two different strategies detailed below.

#### 3.6.1. Strategy A

[40] Our first methodology capitalizes on the fact that, though not exactly similar, the surrogate posterior,  $\tilde{p}(\mathbf{z}|\mathbf{d})$ , is in most cases likely to be closer to the desired posterior,  $p(\mathbf{z}|\mathbf{d})$ , than the prior distribution,  $p(\mathbf{z})$ . Our first strategy is therefore to (1) build a surrogate solution  $\tilde{F}_P$  using the least amount of computational time, (2) determine  $\tilde{p}(\mathbf{z}|\mathbf{d})$  using MCMC simulation, and (3) explore the posterior distribution of the original model,  $p(\mathbf{z}|\mathbf{d})$ , using samples from (2) as a starting point. In the remainder of this paper, we refer to this approach as strategy A.

[41] The state-of-the-art DREAM<sub>(ZS)</sub> [*ter Braak and Vrugt*, 2008; *Vrugt et al.*, 2009; *Laloy and Vrugt*, 2012] algorithm is used to generate posterior samples. A detailed description of this sampling scheme including a proof of ergodicity and detailed balance can be found in the cited literature. We construct the initial archive of parameter sets used by DREAM<sub>(ZS)</sub> by drawing  $m_0 = 10n$  samples from  $\tilde{p}(\mathbf{z}|\mathbf{d})$ . Consequently, the search is initiated in the neighborhood of the mode of  $\tilde{p}(\mathbf{z}|\mathbf{d})$ , and the initial proposals

honor the correlation structure of  $\tilde{p}(\mathbf{z}|\mathbf{d})$ . We purposely do not approximate  $p(\mathbf{z})$  with a mixture distribution of  $\tilde{p}(\mathbf{z}|\mathbf{d})$ . Reasons for this are twofold. First, we prefer a data-driven problem that is not overly constrained by the prior distribution,  $p(\mathbf{z})$ . Second, the marginal prior distribution of the KL modes is known beforehand to be standard normal.

#### 3.6.2. Strategy B

[42] In our second sampling scheme, herein referred to as strategy B, we adopt a state-of-the-art two-stage MCMC approach [*Cui et al.*, 2011], in which  $\tilde{F}_P$  is constructed beforehand and serves as reduced-order model in the first stage. The DREAM<sub>(ZS)</sub> scheme is again used to propose candidate points, but without the optional snooker update [see *ter Braak and Vrugt*, 2008]. In addition, the initial archive of parameter samples is constructed similarly as in strategy A. An excellent description of the two-stage or delayed acceptance MCMC technique, including its extension to the reversible jump method of *Green* [1995], can be found in *Cui et al.* [2011]. We therefore present a short summary of the main elements of the methodology.

[43] In the two-stage MCMC approach, one first samples a candidate point  $\mathbf{z}'$  from a proposal distribution  $q(\mathbf{z}, \mathbf{z}')$  and calculates an approximate posterior,  $\tilde{p}^*(\mathbf{z}|\mathbf{d})$ , that accounts for the model reduction error,  $R(\mathbf{z})$  (see later), using the standard Metropolis-Hastings (MH) acceptance probability [*Metropolis et al.*, 1953; *Hastings*, 1970] of equation (18). This filtering of the unacceptable proposals  $\mathbf{z}'$  essentially defines a modified jump probability  $q^*(\mathbf{z}, \mathbf{z}')$ . In the second step, the accepted draws from  $q(\mathbf{z}, \mathbf{z}')$  are evaluated with the original model in conjunction with the MH rule applied to  $q^*(\mathbf{z}, \mathbf{z}')$ . This avoids unnecessary and expensive computation of  $p(\mathbf{z}|\mathbf{d})$  when proposal points are rejected in the first step. The lower the acceptance rate (*AR*) in the first step and the higher the *AR* in the second step, the larger the computational gain. Note however, that this gain in computational effort compared to standard MCMC simulation schemes comes at the expense of a potential increase in autocorrelation between the posterior samples [*Christen and Fox*, 2005; *Cui et al.*, 2011].

[44] If we assume  $R$  to be multivariate normally distributed and independent of  $\mathbf{z}$ , that is,  $R \sim N(\boldsymbol{\mu}_R, \boldsymbol{\Sigma}_R)$ , we derive

$$\begin{aligned} \tilde{p}^*(\mathbf{z}|\mathbf{d}) &= p(\mathbf{z}) \\ &\times \frac{1}{(2\pi)^{N/2} |\boldsymbol{\Sigma}_c|^{1/2}} \exp\left(-\frac{1}{2} \|\boldsymbol{\Sigma}_c^{-1/2} (\mathbf{d} - \tilde{F}_P(\mathbf{z}) - \boldsymbol{\mu}_R)\|^2\right), \end{aligned} \quad (19)$$

where  $\|\cdot\|$  denotes the Euclidean norm,  $\boldsymbol{\Sigma}_c = \boldsymbol{\Sigma}_R + \boldsymbol{\Sigma}_e$ , with  $\boldsymbol{\mu}_R$  and  $\boldsymbol{\Sigma}_R$  derived from the  $Q$  prior parameter samples. If the analysis of the relationship between  $\tilde{F}_P$  and  $F$  supports the assumptions of  $\boldsymbol{\mu}_R = \mathbf{0}$  and a constant variance,  $\sigma_R^2$ , of  $R$ , that is,  $\boldsymbol{\Sigma}_R = \sigma_R^2 \mathbf{I}$ , equation (19) can be reformulated as

$$\begin{aligned} \tilde{p}^*(\mathbf{z}|\mathbf{d}) &= p(\mathbf{z}) \\ &\times \left(\frac{1}{\sqrt{2\pi\sigma_c^2}}\right)^N \exp\left(-\frac{1}{2} \sigma_c^{-2} \sum_{i=1}^N [d_i - \tilde{F}_{P,i}(\mathbf{z})]^2\right), \end{aligned} \quad (20)$$

where  $\sigma_c^2 = \sigma_R^2 + \sigma_e^2$ .

[45] For symmetric proposal distributions such as used in DREAM<sub>(ZS)</sub>,  $q(\mathbf{z}, \mathbf{z}') = q(\mathbf{z}', \mathbf{z})$  and the MH ratio,  $\alpha(\mathbf{z}, \mathbf{z}')$  reduces to

$$\alpha(\mathbf{z}, \mathbf{z}') = \min \left[ 1, \frac{\tilde{p}^*(\mathbf{z}'|\mathbf{d})}{\tilde{p}^*(\mathbf{z}|\mathbf{d})} \right]. \quad (21)$$

[46] Furthermore, for an approximate posterior of the form of equations (19) and (20), the MH acceptance probability in the second step, hereafter referred to as  $\beta$ , can be written as [Cui et al., 2011]

$$\beta(\mathbf{z}, \mathbf{z}') = \min \left[ 1, \frac{p(\mathbf{z}'|\mathbf{d})\tilde{p}^*(\mathbf{z}|\mathbf{d})}{p(\mathbf{z}|\mathbf{d})\tilde{p}^*(\mathbf{z}'|\mathbf{d})} \right]. \quad (22)$$

[47] The acceptance rate of this second step is controlled by the validity of the approximate posterior  $\tilde{p}^*(\mathbf{z}|\mathbf{d})$ . This inspired Cui et al. [2011] to locally adapt  $\tilde{p}^*(\mathbf{z}|\mathbf{d})$  during sampling, to better mimic the true posterior  $p(\mathbf{z}|\mathbf{d})$ . For a proposal  $\mathbf{z}'$ , the reduced-order model becomes

$$\tilde{F}_{P,\mathbf{z}}(\mathbf{z}') = \tilde{F}_P(\mathbf{z}') + [F(\mathbf{z}) - \tilde{F}_P(\mathbf{z})], \quad (23)$$

which leads to the following form of the approximate posterior

$$\tilde{p}_{\mathbf{z}}^*(\mathbf{z}'|\mathbf{d}) = p(\mathbf{z}) \times \left( \frac{1}{\sqrt{2\pi\sigma_{c,\mathbf{z}}^2}} \right)^N \exp \left( -\frac{1}{2}\sigma_{c,\mathbf{z}}^{-2} \sum_{i=1}^N [d_i - \tilde{F}_{P,\mathbf{z},i}(\mathbf{z}')]^2 \right), \quad (24)$$

where  $\sigma_{c,\mathbf{z}}^2 = \sigma_{R,\mathbf{z}}^2 + \sigma_e^2$ , and  $\sigma_{R,\mathbf{z}}^2$  signifies the error between  $\tilde{F}_{P,\mathbf{z}}(\mathbf{z}')$  and  $F(\mathbf{z}')$  which is adaptively adjusted [Cui et al., 2011]. Using this local correction, equation (21) can be written as

$$\alpha(\mathbf{z}, \mathbf{z}') = \min \left[ 1, \frac{\tilde{p}_{\mathbf{z}}^*(\mathbf{z}'|\mathbf{d})}{\tilde{p}_{\mathbf{z}}^*(\mathbf{z}|\mathbf{d})} \right], \quad (25)$$

and equation (22) takes the following form

$$\beta(\mathbf{z}, \mathbf{z}') = \min \left\{ 1, \frac{p(\mathbf{z}'|\mathbf{d}) \times \min \left[ 1, \frac{\tilde{p}_{\mathbf{z}}^*(\mathbf{z}|\mathbf{d})}{\tilde{p}_{\mathbf{z}}^*(\mathbf{z}'|\mathbf{d})} \right]}{p(\mathbf{z}|\mathbf{d}) \times \min \left[ 1, \frac{\tilde{p}_{\mathbf{z}}^*(\mathbf{z}'|\mathbf{d})}{\tilde{p}_{\mathbf{z}}^*(\mathbf{z}|\mathbf{d})} \right]} \right\}, \quad (26)$$

the proof of which is given in appendix A.

[48] If the magnitude of  $\sigma_{R,\mathbf{z}}^2$  is negligibly small compared to the magnitude of  $\sigma_e^2$ , one can further simplify equation (26) by assuming  $\tilde{p}_{\mathbf{z}}^*(\mathbf{z}'|\mathbf{d}) = p(\mathbf{z}'|\mathbf{d})$  and  $\tilde{p}_{\mathbf{z}}^*(\mathbf{z}|\mathbf{d}) = p(\mathbf{z}|\mathbf{d})$ . In this study,  $\sigma_{R,\mathbf{z}}^2 \in [0.05\sigma_e^2, \sigma_e^2]$ . Consequently, we do not use this simplification as it would lead to corrupted values of  $\alpha$  and  $\beta$ .

[49] To account for the loss of “statistical efficiency” or increase in posterior sample autocorrelation that accompanies the computational gain in two-stage MCMC, the overall computational speed up can be assessed as [Cui et al., 2011]

$$\frac{\tau}{\tau^*} \frac{1}{\hat{\alpha} + \frac{t_s^*}{t_s}}, \quad (27)$$

where  $\hat{\alpha}$  is the average acceptance rate in the first step, and  $t_s^*$  and  $t_s$  signify the computational costs of calculating  $\tilde{p}^*(\mathbf{z}|\mathbf{d})$  and  $p(\mathbf{z}|\mathbf{d})$ , respectively. The ratio  $\frac{\tau}{\tau^*}$  is smaller than one and measures the increase in posterior sample autocorrelation that directly follows from  $\beta < 1$ . The variables  $\tau$  and  $\tau^*$  can be estimated for some statistics of interest of the Markov chain using, e.g., the inverse of the relative numerical efficiency of Geweke [1992], or the integrated autocorrelation time [IACT; Wolff, 2004]. In this work, we follow Cui et al. [2011] and benchmark strategy B using the IACT of log-likelihood values of the sampled parameter values. The IACT values reported in this paper represent an average of the different Markov chains created with DREAM<sub>(ZS)</sub>.

[50] Ideally, a two-stage MCMC simulation should have a relatively low value of  $\alpha$  for the highest possible  $\beta$  value, while preserving that the product  $\alpha \times \beta$  is sufficiently large, say not less than 10%, to produce a reasonable mixing of the chains. In this respect, it is worth noting that there is an intrinsic tradeoff between  $\alpha$  and  $\beta$  in equations (19)–(26). The smaller the distance between the current state in the Markov chain,  $\mathbf{z}$ , and the proposal point,  $\mathbf{z}'$ , the closer  $\tilde{F}_P(\mathbf{z}')$  to  $\tilde{F}_P(\mathbf{z})$  and  $F(\mathbf{z}')$  to  $F(\mathbf{z})$ , respectively, and hence the better  $\tilde{F}_{P,\mathbf{z}}(\mathbf{z}')$  approximates  $F(\mathbf{z}')$ . This leads to a relatively large acceptance rate in the second stage. However, candidates  $\mathbf{z}'$  that differ only marginally from their current position  $\mathbf{z}$  will also have a high chance to be accepted in the first step. This will cause the Markov chain(s) to converge slowly, hence decreasing search efficiency. Striking a appropriate balance between  $\alpha$  and  $\beta$  may therefore not be easy.

[51] Besides autocorrelation, it is also important to assess when convergence of the sampled Markov chains to a limiting (posterior) distribution has been achieved. The Gelman and Rubin [1992] and Raftery and Lewis [1992] diagnostics help to monitor convergence of the sampled chains. Note that these statistics essentially measure efficiency, without recourse to estimating the validity of the sampled posterior distribution. To quantify search effectiveness, one can either compute the distance to the actual target distribution if this is known beforehand [e.g., Laloy and Vrugt, 2012], or simply assess whether the posterior mode coincides with the maximum a posteriori parameter values derived separately with an optimization method.

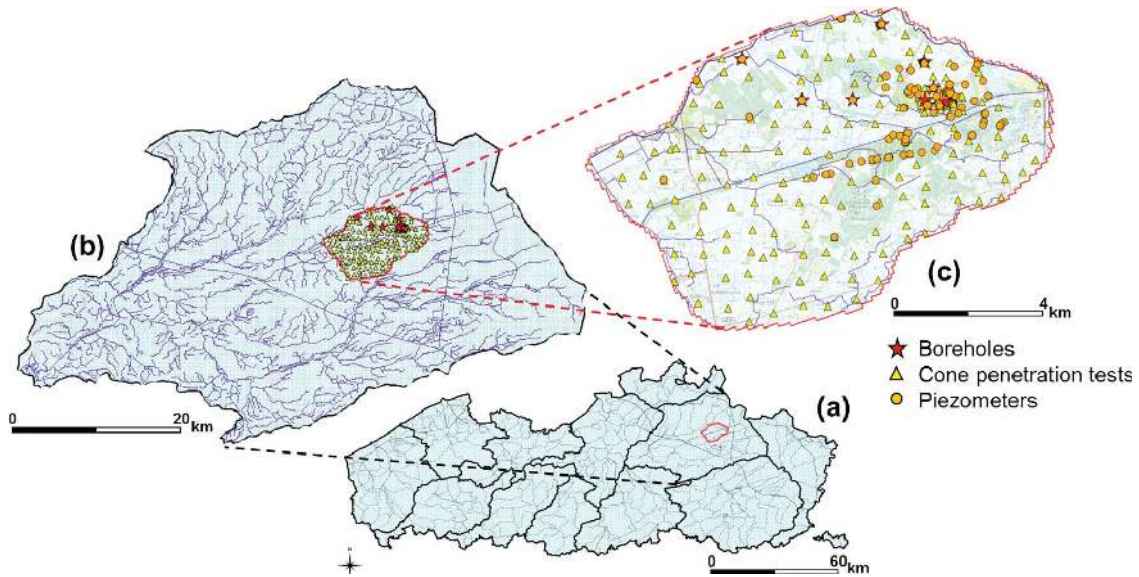
## 4. Case Study

### 4.1. Site and Model Description

[52] We test our two strategies A and B by calibration and predictive uncertainty analysis of a highly parameterized groundwater model involving a subbasin of the Kleine Nete catchment in the region of Mol/Dessel, Belgium (Figures 2a and 2b). A hydrogeological site characterization to depths up to approximately 180 m, of which the first 40–50 m was very detailed, has been carried out in 2008–2009 [Beerten et al., 2010; Rogiers et al., 2010, 2012a, 2012b], using several cored boreholes, and more than 200 cone penetration tests (CPT) (Figures 2b and 2c). This site characterization complements monthly head data from 86 piezometers covering a 25 year period (Figure 2c).

[53] The hydrostratigraphy of the site can be summarized as follows. The upper aquifer with an average thickness



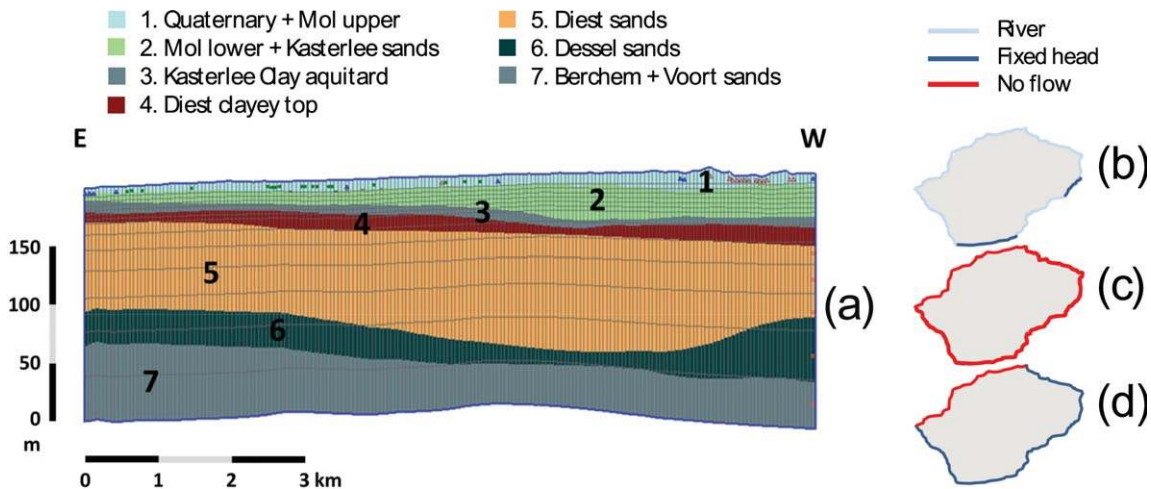


**Figure 2.** Geographical location of the study area and site investigations, including location of the cored boreholes, cone penetration tests, and piezometers. (a) Flanders region (Northern Belgium), (b) Nete basin, and (c) groundwater model extent.

ranging from 25 to 35 m, consists of Quaternary, Mol and Kasterlee Sands (see layers 1 and 2 in Figure 3a). This is followed by an highly heterogeneous 5–10 m thick aquitard known as the Kasterlee Clay (layer 3 in Figure 3a)), and a lower aquifer about 150 m thick, consisting of the Diest, Dessel, Berchem and Voort Sands (layers 5–7 in Figure 3a)), with a less permeable top of a few meters (Diest clayey top, layer 4 in Figure 3a) at the interface with the aquitard. The lower boundary of the lower aquifer corresponds to the Boom Clay aquitard [Yu *et al.*, 2013]. The geometry of the hydrogeological units depicted in Figure 3a was derived from classification and geostatistical

modeling of the CPT data [see Rogiers *et al.*, 2010, for details]. A very thin additional unit, referred to as Mol Clay with unit index 1', occurs locally between units 1 (Quaternary + Mol upper) and 2 (Mol lower + Kasterlee sands) but is too small to be conveniently represented in Figure 3a.

[54] Core samples from the boreholes were used to define the upper and lower bounds of the conductivity ( $K$ ) values of each hydrogeological unit. In the absence of detailed prior information about the parameters at the scale of interest, a uniform prior distribution was used. Since the heterogeneous aquitard (layer 3) controls the flow and transport between the upper and lower aquifers, the hydraulic conductivity of



**Figure 3.** (a) Hydrostratigraphy and model grid of a two-dimensional East-West profile through the study area, (b) upper layer boundary conditions, (c) upper aquifer boundary conditions (not including the upper most numerical layer), and (d) lower aquifer boundary conditions (Diest clayey top and below). A very thin additional unit, referred to as Mol Clay with unit index 1', occurs locally between units 1 (Quaternary + Mol upper) and 2 (Mol lower + Kasterlee sands) but is too small to be conveniently represented in Figure 3a.

this unit is of major importance in our model-data synthesis efforts and will be treated as a spatially variable parameter. A joint analysis of the core samples and CPT measurements showed that the vertical component,  $\mathbf{k}_v$ , of the three-dimensional aquitard conductivity tensor field can be adequately described with an isotropic Gaussian covariance function with sill of about 0.32 (in base 10 logarithmic scale of the conductivity in  $\text{ms}^{-1}$ ), and range,  $L$ , of about 2000 m [Rogiers *et al.*, 2010, 2012a]. The other stratigraphical units were observed to be fairly homogeneous and hence a single isotropic  $K$  value was used to characterize the conductivity of each of these units.

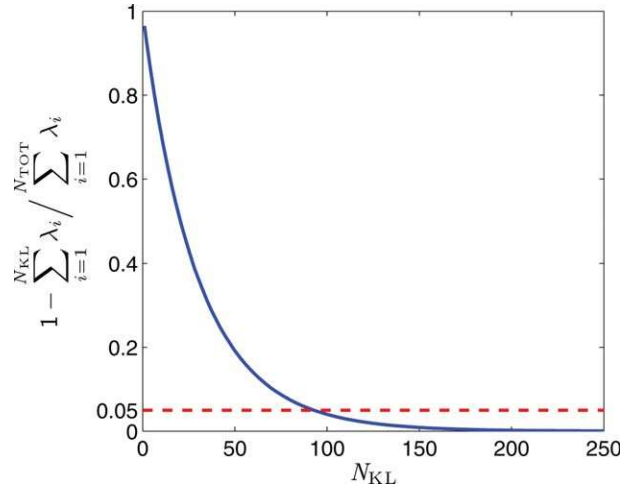
[55] MODFLOW-2005 [Harbaugh, 2005] was used to solve the governing partial differential equation

$$\frac{\partial}{\partial x} \left[ K_{xx} \frac{\partial h}{\partial x} \right] + \frac{\partial}{\partial y} \left[ K_{yy} \frac{\partial h}{\partial y} \right] + \frac{\partial}{\partial z} \left[ K_{zz} \frac{\partial h}{\partial z} \right] + W = 0, \quad (28)$$

in its finite difference form in the discretized model domain. The variables  $K_{xx}$ ,  $K_{yy}$ , and  $K_{zz}$  represent the hydraulic conductivity along the  $x$ ,  $y$ , and  $z$  coordinate axes (assumed to be parallel to the major axes of hydraulic conductivity), respectively,  $h$  is the potentiometric head, and  $W$  represents sources and/or sinks of water. The model domain of 12,500 m by 9000 m by 200 m in the  $x$ – $y$ – $z$  domain was discretized into  $250 \times 180 \times 14 = 630,000$  cuboid voxels with a horizontal side length of 50 m and variable height in the vertical (depth) direction.

[56] The initial heads were obtained by interpolation from the coarser regional Neogene Aquifer Model [Gedeon, 2008], with boundary conditions illustrated in Figures 3b and 3c. No flow conditions were represented by a Neumann boundary condition (in this case a zero flux), and Dirichlet boundary conditions were used to specify the fixed head boundaries. The latter were also interpolated from the calculated heads using the regional Neogene Aquifer Model. The river boundaries are a combination of no flow boundaries with an additional source/sink in the uppermost voxel. At the few locations where the model boundary does not follow river paths, fixed heads are used in the uppermost voxels to bridge these gaps (Figures 3b and 3c). Water flow between the river and underlying aquifer is calculated from Darcy's law using a fixed hydraulic head within the stream, and a spatially variable hydraulic conductivity for each individual cell of the aquifer. Due to the presence of the Kasterlee Clay aquitard, only the main river is assumed to significantly influence the heads of the lower aquifer. Different boundary conditions were therefore specified for the layers above (Figures 3b and 3c) and below (Figure 3d) the aquitard. The bottom layer of the deepest hydrogeologic unit is assumed to be impermeable, which is represented by a zero flux boundary condition (Neumann). As upper model boundary, we used a constant recharge of  $307 \text{ mm y}^{-1}$ , with a complex network of rivers and drains (Figure 3b). For a more detailed model description, we refer the reader to Gedeon *et al.* [2011] and Gedeon and Mallants [2012].

[57] Under steady-state flow conditions, a single model run required about 2–3 min of CPU time on a 64-bit Intel® Core™ i7–2820 CPU 2.30 GHz processor, but sometimes lasted up to 70 min depending on the values of the model parameters.



**Figure 4.** Decay of the KL eigenvalues,  $\lambda_i$ , with the selected covariance model.  $N_{\text{KL}}$  and  $N_{\text{TOT}}$  denote the number of KL modes and total number of field voxels, respectively. The blue solid line depicts the fraction of the unresolved field variance as function of the number of KL modes used in the analysis. A total of 92 KL modes is required to preserve 95% of the field variance (intersection between the blue and red dashed line).

## 4.2. Model Parameterization

[58] The observational data set comprises monthly piezometric heads from  $N = 86$  locations averaged over a 25 year period. The inferred model parameters are the homogeneous and isotropic  $K$  values of the aquifer units 1, 1', 2, and 4–7 shown in Figure 2a, and the  $250 \times 180 = 45,000$  dimensional  $\log_{10}(\mathbf{k}_h)$  and  $\log_{10}(\mathbf{k}_v)$  fields of the aquitard unit 3. The  $\mathbf{k}_h$  component of unit 3 was assumed to be isotropic, that is,  $\mathbf{k}_{(h,x)} = \mathbf{k}_{(h,y)}$ , and linearly related to  $\mathbf{k}_v$  by an anisotropy factor,  $ANI_{3,h} = \mathbf{k}_h/\mathbf{k}_v$ , which is estimated jointly with the other unknown parameters. In practice, we only estimate  $\mathbf{k}_v$  and derive  $\mathbf{k}_h$  for those grid cells that appear within the active modeling domain of the basin considered herein (Figure 3c). This is equivalent to 26,112 hydraulic conductivity values for both spatial fields.

[59] We take advantage of the KL transform described in equations (1)–(7) to substantially decrease the dimensionality of  $\log_{10}(\mathbf{k}_v)$ . For the assumed correlation length and covariance model of equation (7), a total of 92 KL modes were deemed sufficient to preserve 95% of the total field variance (see Figure 4). We numerically evaluated the eigenfunctions,  $\phi_i(\mathbf{x})$ , and eigenvalues,  $\lambda_i$ , of the kernel  $C$  (equations (1)–(6)) based on the observed covariance model (equation (7)). For computational tractability,  $\phi_i(\mathbf{x})$  and  $\lambda_i$  were computed using a voxel horizontal side length of 100 m instead of 50 m, thereby reducing the size of the numerical domain from  $250 \times 180$  to  $125 \times 90$ . Each  $\mathbf{k}_v$  field was then interpolated from the coarser  $125 \times 90$  grid to the actual finer  $250 \times 180$  mesh.

[60] The Gaussian covariance function and correlation length used herein is supported by geostatistical analysis of field data [Rogiers *et al.*, 2010, 2012a]. The Gaussian covariance function has important computational advantages in that only a relatively small number of KL modes are needed to preserve 95% of the field variance. The use

**Table 1.** Description of the Inferred Model Parameters With Their Prior Uncertainty Interval<sup>a</sup>

Parameter	Units	Prior
<i>Units 1, 1', 2, and 4–7</i>		
$K_1$	m d <sup>-1</sup>	1–25
$K_{1'}$	m d <sup>-1</sup>	0.1–1
$K_2$	m d <sup>-1</sup>	1–25
$K_4$	m d <sup>-1</sup>	0.1–1
$K_5$	m d <sup>-1</sup>	1–25
$K_6$	m d <sup>-1</sup>	0.1–2.5
$K_7$	m d <sup>-1</sup>	$1 \times 10^{-3}$ – $2.5 \times 10^{-2}$
<i>Unit 3</i>		
$ANI_{3,h}$		$1 \times 10^{-4}$ – $1 \times 10^{-3}$
$MK_{3,v}$		–9.3 to –6.7
$SK_{3,v}$		0.02–0.75
$KL_1, \dots, KL_{92}$		–3 to 3

<sup>a</sup>The subscripts 1–7 refer to the layering shown in Figure 3a. Subscript 1' signifies the very thin and locally occurring clay layer between horizons 1 and 2 in Figure 3a.  $MK_{3,v}$  and  $SK_{3,v}$  denote the mean and variance of  $\log_{10}(\mathbf{k}_v)$ , respectively, and  $KL_1, \dots, KL_{92}$  refer to the associated KL modes. Other parameters are defined in the main text.

of a less smooth covariance model, such as the exponential covariance function, on the contrary, would allow a much higher level of spatial detail and variability, but at the expense of several hundreds of KL modes to preserve 95% of the variance of the log conductivity field [e.g., Li and Cirpka, 2006].

[61] Note that inferring the covariance function and correlation length along with the unknown KL modes from the available piezometric data would pose significant computational challenges, whereas the final results might not honor the prior variogram analysis of the field data. For instance, the calculation of the KL representation associated with each proposed parameter set might take a prohibitively long time for large parameter spaces. Indeed, derivation of a single set of KL eigenvalues and eigenfunctions requires more than 12 h of CPU time for the Gaussian covariance kernel considered herein. Moreover, each time a different covariance function and correlation length is used, the number of KL modes needs to be rederived to preserve 95% of the field variance. The MCMC inference then becomes trans-dimensional in a sense that the dimensionality of the parameter space varies from one proposal to the next. In principle, this can be solved by reversible jump MCMC simulation [Green, 1995], but this methodology is not straightforward to implement in practice.

[62] The standard deviation of  $\log_{10}(\mathbf{k}_v)$ ,  $\sigma$ , simply acts as a scaling factor of the covariance matrix  $\mathbf{C}$  in equation (7), and therefore does not affect the shape of the decay of the KL eigenvalues displayed in Figure 4. For greater flexibility, we therefore use  $\sigma = 1$  in equations (1)–(6), and subsequently estimate the  $\sigma$  value of the field along with the other parameters. Note that for a fixed value of  $\sigma$ , a slight deviation of the KL mode distribution from a distribution with zero mean and unit variance will also impact the variance of the reconstructed field. Hence, we compute the field variance from a range of posterior realizations. In practice, we multiply  $\sqrt{\lambda_i}$  derived from  $\sigma = 1$  in equation (1) by the estimated value of  $\sigma$ . The  $\log_{10}(\mathbf{k}_v)$  field is hence characterized by 94 parameters including the mean and variance parameter,  $MK_{3,v}$  and  $SK_{3,v}$ , respectively, and the 92 KL modes that honor the observed Gaussian covariance

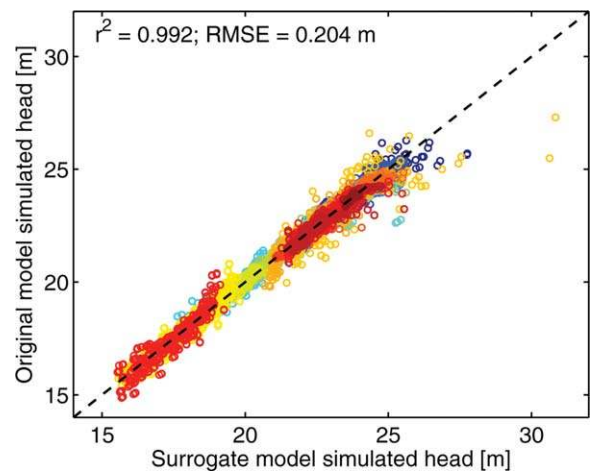
structure with a range of 2000 m. The vector  $\mathbf{z}$  thus contains  $n = 102$  parameters. The prior distribution of the  $N_{\text{KLT}} = 92$  KL modes is given by a multinormal distribution where the individual KL modes are initially assumed to be uncorrelated. The prior distribution of the ten remaining parameters is taken to be “uniform” with ranges listed in Table 1. The prior distribution of  $\mathbf{z}, p(\mathbf{z})$ , thus reduces to

$$p(\mathbf{z}) = \left(\frac{1}{\sqrt{2\pi}}\right)^{N_{\text{KLT}}} \exp\left(-\frac{1}{2} \sum_{i=1}^{N_{\text{KLT}}} KL_i^2\right). \quad (29)$$

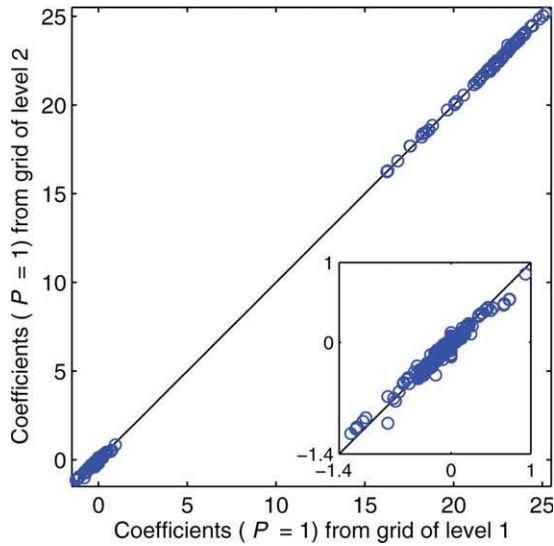
### 4.3. Surrogate Construction and Posterior Exploration Using MCMC Simulation

[64] To construct  $\tilde{F}_P$  in equations (8)–(14), we combine  $p(\mathbf{z})$  with Legendre polynomials for  $K_1, K_{1'}, K_2, K_4, K_5, K_6, K_7, ANI_{3,h}, MK_{3,v}$ , and  $SK_{3,v}$ , respectively, and Hermite polynomials for  $KL_1, \dots, KL_{92}$ . For the nodal set  $\{\mathbf{z}_k, w_k\}$  in equation (13), we use the GP/GH mixed sparse grid depicted in Figure 1 with a sparse grid of level one. For 102 dimensions this leads to a total of 205 sparse grid points ( $Q$  in equation (13)) for which the forward model has to be evaluated.

[65] In all our numerical experiments, we use an expansion order of  $P = 1$ . While this order does not account for correlation between the different parameters, its value is appropriate for the following reasons. In the first place,  $P = 1$  performs well for the surrogate model considered herein. An average RMSE of approximately 0.20 m is found in Figure 5 between the simulated heads of the surrogate and original hydrogeologic model. The use of  $P = 2$  does not improve the results, but rather deteriorates the RMSE because the number of 205 nodes used in equation (13) becomes insufficient to adequately capture the behavior of the original CPU-intensive model. Moreover, a larger number of 21,217 grid points and corresponding model calls are required to increase the resolution of the selected mixture of GH/GH quadrature rules, that is, to reach a



**Figure 5.** Pair wise comparison of the simulated heads of the original ( $F$ ) and surrogate model ( $\tilde{F}_P$ ) for 200 different parameter combinations randomly drawn from the prior distribution. The ensemble of simulated heads at each measurement location is coded with a different color.



**Figure 6.** Comparison of the 103 coefficients of the order-one polynomials derived from the sparse grids of level 1 ( $x$  axis) and 2 ( $y$  axis), respectively, for the 86 model outputs.

sparse grid of level two. This defeats the purpose of our method. Also the computational burden of 21,217 model runs is much higher than the total CPU time required for strategy B to converge to the posterior parameter distribution (about 2500 additional original model simulations, see later). Nevertheless, for illustrative purposes we did compute the associated results of  $P = 2$  using distributed computing of the 21,217 deterministic simulations. The resulting aliasing error of  $\bar{F}_P$  remains too high to accurately reproduce the original model behavior. In other words, such a large number of sparse grid nodes is still insufficient to derive accurate degree-two expansion coefficients in equation (13), that properly capture the pairwise first-order mixed effects between the model parameters. Further increase of the sparse grid resolution to a grid level of three necessitates numerical solution of such a large number of grid points (1,477,865) that computation becomes virtually impossible. Note, however, that the proposed approach is not limited to  $P = 1$ . It can be employed with a gPC approximation up to any order, if the required computational effort for constructing an accurate expansion is affordable.

[66] To verify whether the 205 points of the sparse grid of level 1 sufficiently well resolve the order-one expansion coefficients, we also compute the coefficients of the order-one polynomials on the basis of the 21,217 points of the grid of level 2. A significant difference between the coefficients obtained from both levels of the sparse grid would demonstrate that the order-one gPC model resulting from level 1 is under resolved and that the level of the sparse grid should be increased. Figure 6 summarizes the results of this comparison. Fortunately, the coefficients of both grid levels appear to be relatively similar. Evidence of this is given by the RMSE between both sets of coefficients. This value is about 0.014 with coefficients ranging between  $-1.5$  and  $25.5$  for the 86 model outputs. The average mismatch (in absolute value) between the two sets of

coefficients is 0.0021. The main difference between both sets of coefficients occurs for parameter  $SK_{3,v}$ . The associated coefficients derived from the grid of level 1 are zero, whereas the grid of level 2 results in nonzero coefficients for  $SK_{3,v}$ , demonstrating some sensitivity of the gPC model to this parameter.

[67] The correspondence between the coefficient values calculated from both grid levels is further confirmed by a pair wise analysis of the simulated heads of the original and surrogate models for the 200 different prior parameter realizations used previously in Figure 5. The associated RMSE is about 0.20 m for the coefficients obtained from the sparse grid of level 1 (Figure 5), whereas this value reduces to 0.19 m for the coefficients derived from the sparse grid of level 2 (not shown). These findings inspire confidence that a sparse grid of level 1 provides sufficiently accurate coefficients of the order-one gPC expansion used herein.

[68] The stochastic expansion used herein assumes independence between the uncertain parameters. This assumption is consistent with the selected prior pdf, which is given by the product of independent uniform distributions and independent standard normal distributions (see equation (29)). If available prior information reveals the presence of correlation between individual parameters, the *Rosenblatt* [1952] and Nataf [*Der Kiureghian and Liu, 1986*] transformations can be used for decorrelation [*Eldred, 2011*].

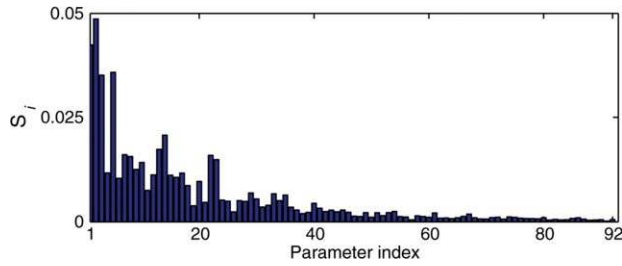
[69] The gPC technique is well suited for global sensitivity analysis (GSA) [e.g., *Saltelli et al., 2006*] because the associated computational cost becomes practically equivalent to the time it takes to estimate the expansion coefficients [e.g., *Sudret, 2008*]. To investigate the sensitivity of  $\bar{F}_P$  to each individual model parameter, we computed their first ( $S$ ) order Sobol' [1993] sensitivity. The scalars  $S_i, i = 1, \dots, n$  numerically quantify the influence of the model parameters on the output of  $\bar{F}_P$ . For additive (linear) models the sum of  $S_i, i = 1, \dots, n$  is 1.

[70] Table 2 lists the average  $S_i$  values for the simulated head values at 86 different locations within the modeling domain. These indices were obtained from two different

**Table 2.** First-Order Sobol Sensitivity Index of the Surrogate Model ( $\bar{F}_P$ ) Parameters<sup>a</sup>

Parameter	Classical GSA	PCE-GSA
$K_1$	0.087	0.052
$K_{1'}$	$3 \times 10^{-4}$	$1 \times 10^{-4}$
$K_2$	0.343	0.241
$K_4$	$0.3 \times 10^{-4}$	$1 \times 10^{-4}$
$K_5$	0.084	0.040
$K_6$	$2 \times 10^{-4}$	$1 \times 10^{-4}$
$K_7$	$8 \times 10^{-7}$	$9 \times 10^{-6}$
$ANI_{3,h}$	0.052	0.024
$MK_{3,v}$	0.246	0.1122
$SK_{3,v}$	0	0
$KL_1, \dots, KL_{92}$	$9 \times 10^{-5} - 0.017$	$2 \times 10^{-4} - 0.049$
$\sum_{i=1}^{92} KL_i$	0.187	0.531

<sup>a</sup>Listed values represent averages over the 86 model outputs. Classical GSA represents the standard Sobol sensitivity analysis of the surrogate model using a range of  $[-3, 3]$  for the 92 KL modes. The *Sudret* [2008] approach, which for the case study considered herein uses a KL mode range of  $[-1.2247, 1.2247]$ , is denoted by PCE-GSA.



**Figure 7.** Sobol's sensitivity of the surrogate model ( $\tilde{F}_P$ ) to the  $i = 1, \dots, 92$  KL modes derived from the classical GSA approach. Each index  $S_i$  represents an average of the sensitivity of the 86 simulated head values.

approaches. First, we applied the *Sudret* [2008] methodology hereafter conveniently referred to as gPC-GSA. This technique approximates the Sobol' sensitivity estimates from analytical postprocessing of the gPC coefficients. Note that with a sparse grid level of one, at most two points are placed in each grid dimension and the investigated range for the parameters of the Hermite polynomials, that is, the KL modes, is  $[-1.2247, 1.2247]$ , respectively. This rather narrow range conditions the values of the sensitivity indices obtained by the gPC-PCA approach. But this interval is much smaller than the expected range of the KL modes that should be normally distributed with values that differ at least between  $-3$  and  $3$ . To avoid underestimating the sensitivity of  $\tilde{F}_P$  to the KL modes, we performed a second standard Sobol' sensitivity analysis of  $\tilde{F}_P$  but this time using Jansen formulations [see, e.g., *Saltelli et al.*, 2010], with each KL mode varying between  $-3$  and  $3$ . This procedure is hereafter referred to as classical GSA and incurs an extra CPU-time cost of less than 20 minutes for a total of  $1000 \times (n + 2) = 104,000$  surrogate model runs.

[72] Among the homogeneous hydrogeological units, that is, all but three (see Figure 3), the most sensitive surrogate model parameters are  $K_1, K_2$  and  $K_5$ , whereas  $K_1, K_4$  and  $K_6$  appear much less influential. For gPC-GSA, the sum of the sensitivity indices associated with unit 3 represents about 49% of the total variance of  $\tilde{F}_P$ . On the contrary, the classical GSA analysis assigns about 67% of the total variance of  $\tilde{F}_P$  to the parameters associated with unit 3. The sum of the KL mode sensitivity estimates equals 19% for the former analysis and 53% for the latter. The sensitivity of the gPC model to each KL mode decreases exponentially with mode number (Figure 7). Whereas the first 10–20 KL modes strongly affect the simulated head values of the surrogate model, the last 20–30 KL modes appear rather noninfluential. This lack of influence does not pose problems with posterior parameter estimation due to the standard normal informative prior of equation (29). Indeed, the less influential the KL mode, the closer its marginal posterior distribution to the normal prior pdf. Lastly, note that the parameter  $SK_{3,v}$  has no influence at all on  $\tilde{F}_P$ . These surrogate sensitivity estimates will be compared later to the output of the two-stage MCMC simulation of the original model (see section 4.5). Note that a degree-one gPC approximation does not consider the effect of parameter interaction, likely underestimating the sensitivity of the original model ( $F$ ).

[73] It is worth noting that an anisotropic sparse grid may be used if the parameters demonstrate a differing sensitivity to the model output [see *Nobile et al.*, 2008b, for details]. The underlying idea is that one can significantly reduce the number of evaluation nodes and hence CPU time by reducing the number of sparse grid points for the least sensitive parameters. Limited experimentation with this technique for the case study considered herein did however not show any noticeable improvements over the isotropic mixed sparse grid. We posit that the number of points in a sparse grid of level one is insufficient to account effectively for anisotropy.

[74] After  $\tilde{F}_P$  is built, the gPC-based posterior,  $\tilde{p}(\mathbf{z}|\mathbf{d}) \propto p(\mathbf{z})\tilde{L}(\mathbf{z}|\mathbf{d})$ , can be derived. Preliminary least squares analysis suggests a value of  $\sigma_e = 0.20$  m in the likelihood functions described in equations (17), (20), and (24). Alternatively, one can sample  $\sigma_e$  along with the other parameters. To draw samples from  $\tilde{p}(\mathbf{z}|\mathbf{d})$ , we employ the adaptive DREAM<sub>(ZS)</sub> algorithm with three different Markov chains. A maximum total of 50,000 surrogate model evaluations is used to explore  $\tilde{p}(\mathbf{z}|\mathbf{d})$ . This takes less than five minutes on our processor. The *Gelman and Rubin* [1992] convergence diagnostic,  $\hat{R}$ , is satisfied after about 39,000  $\tilde{F}_P$  evaluations, with mean, minimum and maximum posterior RMSE values of 0.219, 0.200, and 0.265 m, respectively.

#### 4.4. Performance of Strategy A and B

[75] The first strategy “A” is hereafter referred to as DR-A, where the abbreviation DR stands for DREAM<sub>(ZS)</sub>. Similarly, DR-B refers to strategy B. For comparative purposes we include a separate DREAM<sub>(ZS)</sub> trial for the original hydrogeologic model that does not use any information from the surrogate posterior,  $\tilde{p}(\mathbf{z}|\mathbf{d})$ . A maximum total of 10,000 original model (function) evaluations ( $FE$ ) is used to test the three different posterior estimation methods.

[76] Table 3 lists summary statistics of the different MCMC trials considered herein. These statistics include the mean and average  $AR$  of the first ( $\hat{\alpha}$ ) and second ( $\hat{\beta}$ ) step of the two-stage MCMC simulation ( $AR = \hat{\alpha} \times \hat{\beta}$ ), and the convergence diagnostics of *Raftery and Lewis* [1992] and *Gelman and Rubin* [1992], respectively. The latter statistics, referred to as  $RL_{FE}$  and  $\hat{R}_{FE}$ , summarize the total number of  $FE$  needed to reach convergence of the sampled chains. Note that these values do not include the 205  $FE$  required for construction of the surrogate model.

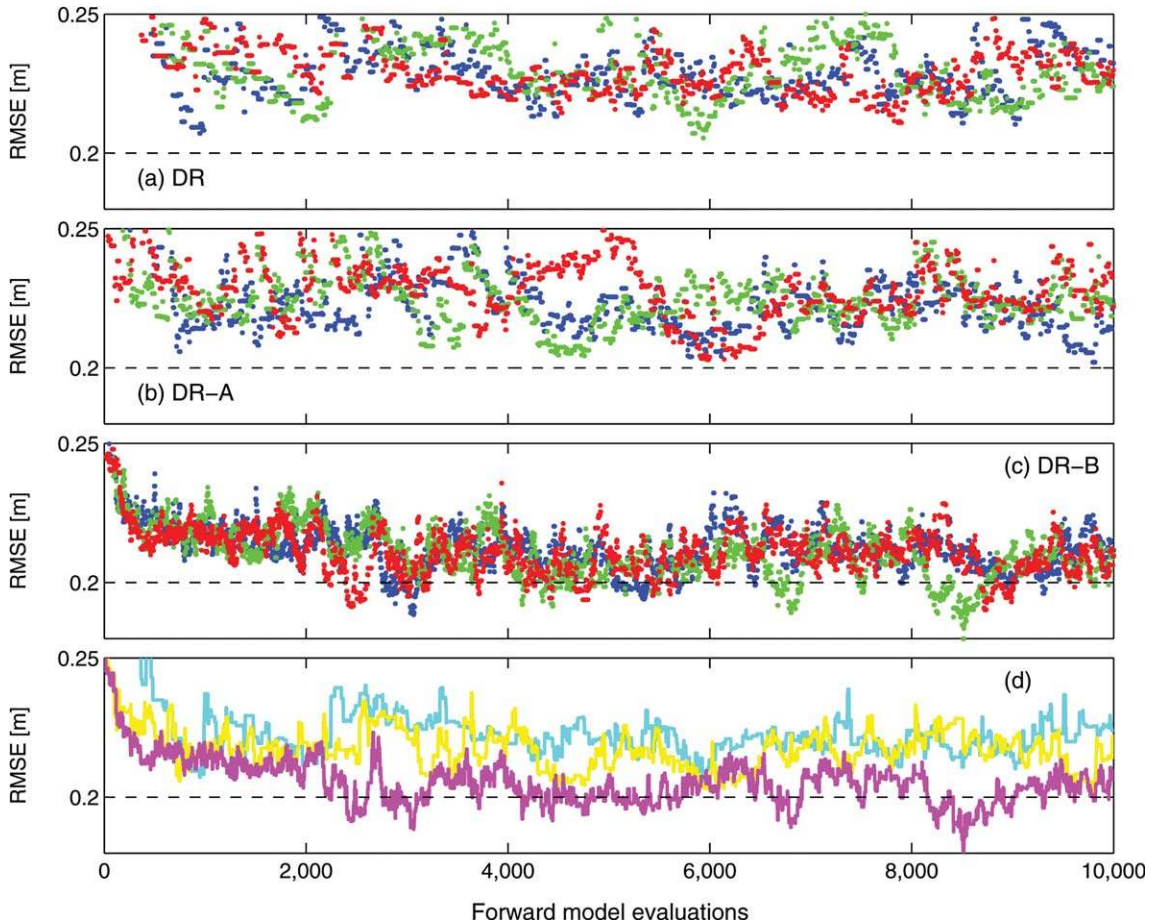
**Table 3.** Comparison of the Different Sampling Acceleration Strategies for a Maximum Total of 10,000 Forward Model Evaluations ( $FE$ )<sup>a</sup>

	$RL_{FE}$	$\hat{R}_{FE}$ ( $FE$ )	$AR$ (%)	$\hat{\alpha}$ (%)	$\hat{\beta}$ (%)	IACT	$STD_{IACT}$	$IACT_{FE}$ ( $FE$ )
DR	5622	$N/C^b$	17	$N/A^c$	$N/A^c$	78	43	78
DR-A	6474	$N/C^b$	20	$N/A^c$	$N/A^c$	76	22	76
DR-B	957	$N/C^b$	15	26	56	160	70	42

<sup>a</sup>The variables  $N, RL, \hat{R}_{FE}, AR, \hat{\alpha}, \hat{\beta}$ , IACT and  $IACT_{FE}$  have been defined in the main text. Listed statistics represent averages from the three interacting Markov chains, and  $STD_{IACT}$  is the standard deviation of the mean IACT. IACT,  $STD_{IACT}$ , and  $IACT_{FE}$  are estimated using the last 2500 samples generated by each method within the allowed computational time.

<sup>b</sup>Not satisfied within the allowed 10,000  $FE$ .

<sup>c</sup>Not applicable.



**Figure 8.** Evolution of sampled RMSE (m) values for each of the MCMC simulation schemes considered herein: (a) DREAM<sub>(ZS)</sub> (DR), (b) strategy A with DREAM<sub>(ZS)</sub> (DR-A), and (c) two-stage MCMC simulation (strategy B) with DREAM<sub>(ZS)</sub> (DR-B). Each chain is coded with a different color. (d) Evolution of the minimum RMSE value of the three different chains derived with DR (cyan), DR-A (yellow), and DR-B (magenta).

Finally, the IACT statistic reports how many steps are required in each chain to produce an independent sample, and the associated number of  $FE$  are indicated with  $IACT_{FE}$ , respectively.

[80] An adequate mixing of the chains is observed for all approaches, with a net  $AR$  ranging between 15% and 20%. While all trials satisfy the more tolerant  $RL$  statistic, none of the approaches formally converge ( $\hat{R}_j < 1.2; j = 1, \dots, n$ ), within the allowed computational budget of 10,000 hydro-geologic model evaluations. About 12,000 additional  $FE$  are needed for DR-B to converge to a limiting distribution, whereas DR and DR-A do not converge even after 20,000  $FE$  (not shown herein). This difference in convergence speed is mainly due to the efficiency of two-stage MCMC simulation with DR-B creating about  $1/\hat{\alpha} \approx 3.8$  samples for each  $FE$ . On the contrary, DR and DR-A draw only one sample per  $FE$ . Whereas it remains difficult to exactly pinpoint when convergence of a MCMC sampler to the posterior distribution has been achieved, the listed statistics enable a qualitative comparison of the efficiency of each different strategy used herein. Trace plots of individual parameters and sampled RMSE values can help to visually

assess when convergence of the MCMC sampler has taken place [e.g., Geyer, 2010].

[81] Figure 8 presents trace plots of the sampled RMSE values for each of the three methods used herein. Strategy A appears to be of little added value compared to traditional MCMC simulation. The RMSE trajectory of DR-A is very similar to that of the (standard) DR run, with only marginal improvements in the sampled RMSE values. On the contrary, DR-B converges rapidly to RMSE values of about 0.20 after about 2500  $FE$ . The posterior density of these RMSE values is more than four orders higher than their counterparts sampled with DR and DR-A (Table 4).

[82] Regarding statistical efficiency, we observe an important 105% increase in IACT between DR-B and DR (Table 3). Hence, the ratio  $\frac{\bar{\tau}}{\bar{\tau}^*}$  in equation (27) is on the order of  $\frac{78}{160} \approx 0.5$ . This mainly results from the fact that  $\hat{\beta} \approx 56\%$  in DR-B. If computational effort is of main importance, however, then DR-B is most efficient requiring the least amount of CPU time to produce an independent sample from  $p(\mathbf{z}|\mathbf{d})$ , with an  $IACT_{FE}$  of 42.

[83] A single evaluation of  $\tilde{p}(\mathbf{z}|\mathbf{d})$ ,  $\tilde{p}^*(\mathbf{z}|\mathbf{d})$  or  $\tilde{p}_z^*(\mathbf{z}'|\mathbf{d})$  takes less than 0.01 s, and hence the speed-up in sampling

**Table 4.** Summary of the Root Mean Square Error (RMSE), and Unnormalized Posterior Density Values,  $p(\mathbf{z}|\mathbf{d})$ , of the Last 50% Sample Outputs Generated with Each Method for a Maximum Total of 10,000 Forward Model Evaluations ( $FE$ )

	RMSE (m)			$p(\mathbf{z} \mathbf{d})$ (base 10 logarithmic scale [-])		
	Mean	Minimum	Maximum	Mean	Minimum	Maximum
DR	0.228	0.206	0.343	-55.98	-108.40	-43.71
DR-A	0.223	0.202	0.248	-54.72	-70.44	-44.32
DR-B	0.208	0.180	0.232	-45.81	-54.67	-38.83

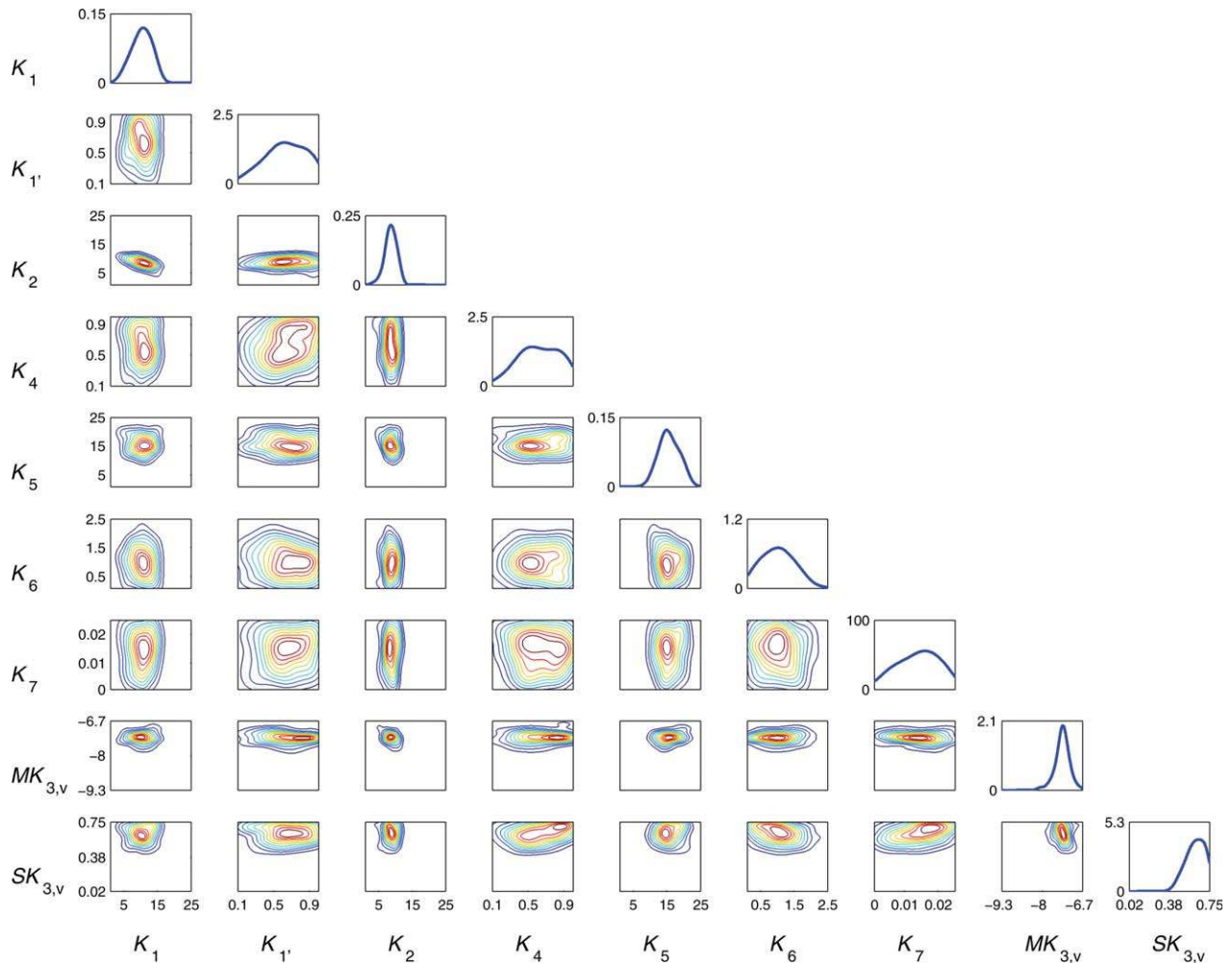
efficiency for strategy DR-B is about 2 according to equation (27). Though not spectacular, this constitutes an important reduction in CPU time. The effective speed-up of DR-B is actually much higher. Indeed, approximately 13,000 (11,000)  $FE$  are required with DR (DR-A) to converge to a limiting distribution of RMSE values that gather around 0.20 (not shown herein). Strategy DR-B however, converges much more rapidly to RMSE values of 0.20. The trace plots illustrate that about 2500  $FE$  are required with DR-B to sample from a limiting distribution. The effective

computational speed up of DR-B is therefore about four to five times. We hence conclude that strategy DR-B is most efficient and effective in sampling from  $p(\mathbf{z}|\mathbf{d})$ .

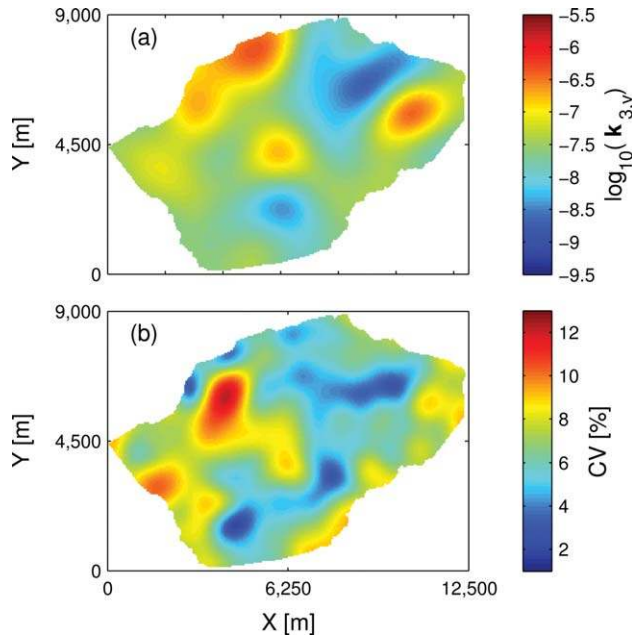
**4.5. Analysis of the Posterior Parameter Uncertainty**

[84] Given the superior performance of DR-B, we now turn our attention to the analysis of the posterior parameter pdf. Figure 9 depicts univariate and bivariate marginal distributions of a set of nine parameters (from  $n = 102$ ) derived by sampling  $p(\mathbf{z}|\mathbf{d})$  using two-stage MCMC simulation with DR-B.

[85] All parameters appear reasonably well identified, with the exception of parameters  $K_{1'}$ ,  $K_4$ , and  $K_7$  that exhibit a much larger posterior uncertainty with probability mass accumulated across the entire prior ranges. This is consistent with the previous sensitivity analysis of  $\tilde{F}_P$  (see section 4.3), for which these three parameters were found to be weakly influential. The relatively narrow marginal posterior distributions of  $K_1, K_2, K_5$ , and  $MK_{3,v}$  reiterate their importance in  $\tilde{F}_P$ . It is interesting to observe that parameters  $K_6$  and  $SK_{3,v}$  are well resolved, though they are either weakly influential or noninfluential with respect to  $\tilde{F}_P$ . This demonstrates that their actual importance is



**Figure 9.** Univariate and bivariate marginal posterior distributions derived with DR-B for a set of nine parameters (out of 102). Kernel density smoothing is used prior to plotting using the last 50% of the sampled parameter values in each chain.

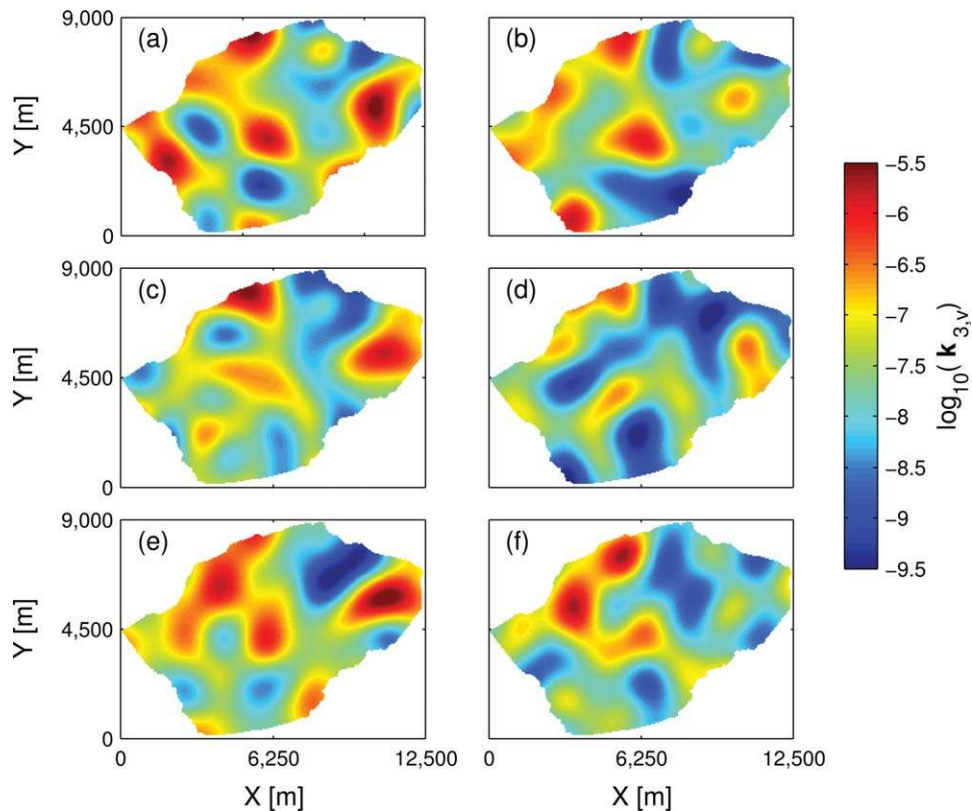


**Figure 10.** Posterior (a) mean and (b) coefficient of variation, CV (%) of the posterior distribution of the vertical component ( $k_v$ ) of the three-dimensional aquitard conductivity tensor field.

largely underestimated by  $\bar{F}_P$ . Also notice that the mode of  $SK_{3,v}$  is in close proximity of its upper bound derived from prior information. Indeed, the field variance is better derived from the posterior field realizations. We follow this approach in the remainder of this paper and summarize the corresponding findings below.

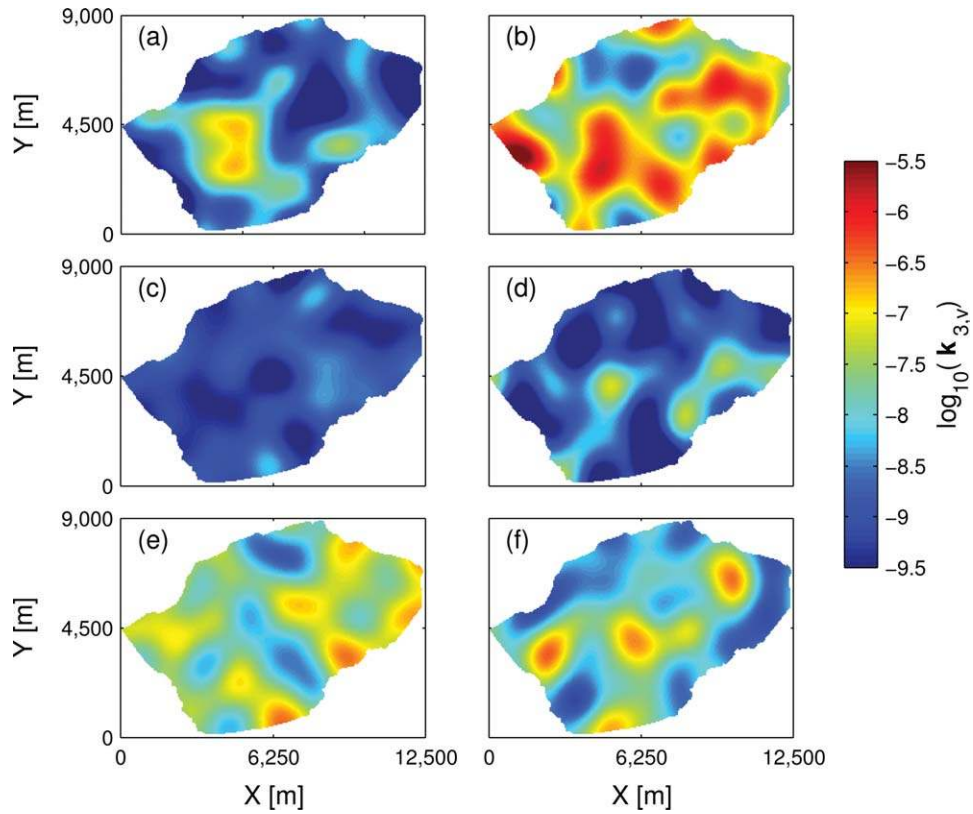
[86] Now that we have samples of  $p(\mathbf{z}|\mathbf{d})$ , we derive estimates of the posterior uncertainty of the hydraulic conductivity field. Figure 10 presents the posterior mean and coefficient variation of the  $k_v$  conductivity field, whereas Figure 11 depicts six possible posterior realizations of  $k_v$  derived by randomly drawing samples from  $p(\mathbf{z}|\mathbf{d})$ . The highest conductivity values are being found at the North-West boundary of the considered spatial domain. Along the South-West-North-East axis much lower hydraulic conductivities are being found. The mean posterior conductivity values depicted in Figure 10 agree well with the results of Rogiers *et al.* [2012a], obtained with a deterministic inversion method and extended data set. The posterior variance of the log conductivity field is approximately Gaussian and centered around 0.40 (not shown). This finding is in good agreement with the value of 0.32 derived from data analysis by Rogiers *et al.* [2010, 2012a].

[87] To illustrate the effect of MCMC simulation, Figure 12 presents six different realizations of the hydraulic conductivity field by sampling from the prior parameter



**Figure 11.** Six posterior realizations of the vertical component ( $k_v$ ) of the three-dimensional aquitard conductivity tensor field, with RMSE values of (a) 0.180 m, (b) 0.198 m, (c) 0.213 m, (d) 0.212 m, (e) 0.194 m, and (f) 0.214 m.



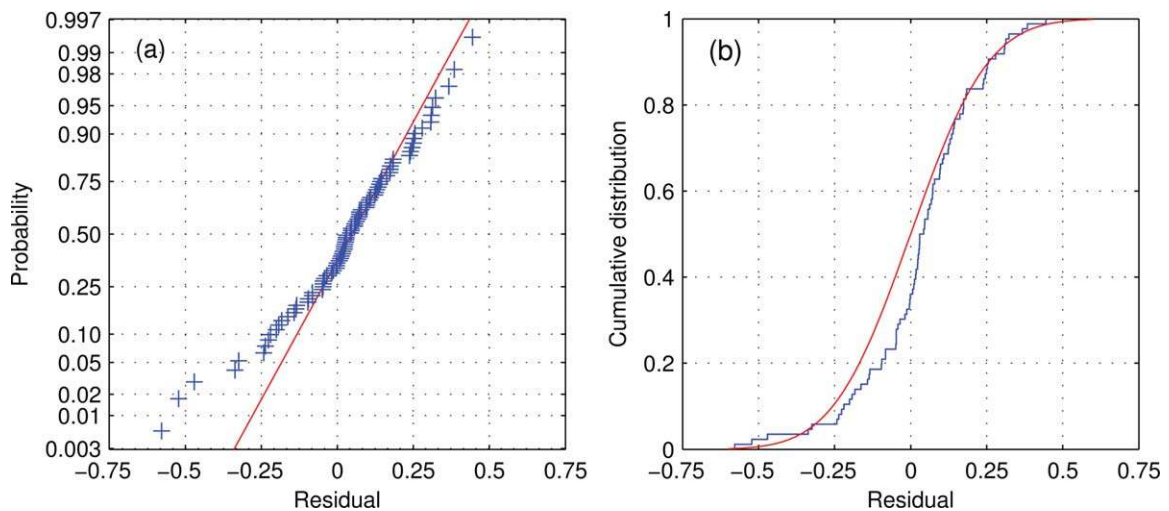


**Figure 12.** Six possible realizations of the vertical component ( $k_v$ ) of the three-dimensional aquitard conductivity tensor field, randomly drawn from the prior parameter distribution. The corresponding RMSE values are (a) 0.562 m, (b) 0.637 m, (c) 0.784 m, (d) 0.592 m, (e) 0.518 m, and (f) 0.824 m.

distribution. These realizations not only differ substantially from each other (large uncertainty) but also deviate significantly from their posterior counterparts with RMSE values that are about three to five times higher than their MCMC-derived maximum a posteriori density values (Figures 11 and 12).

**4.6. A Posteriori Check of the Residual Properties**

[88] We now check if the posterior residual distribution follows our prior assumptions made in the selection of the likelihood function. Figure 13 presents a normal probability plot and cumulative distribution function of the posterior mean residuals. Figure 13a shows that the residuals follow



**Figure 13.** (a) Normal probability and (b) cumulative distribution function plots of the residuals. The blue crosses (Figure 13a) and blue solid line (Figure 13b) signify the empirical estimates, and the red line indicates the assumed Gaussian distribution  $\sim N(0, 0.2)$ .

a Gaussian distribution with  $\sigma_e = 0.20$  in the interval between  $-0.1$  and  $0.5$ . This region contains about 81% of the residuals. Outside this range the posterior mean residual distribution deviates from normality. Yet, this deviation is acceptable, with minor differences between the assumed and MCMC derived cumulative residual distribution function (Figure 13b). We posit that data transformations [e.g., *Box and Cox*, 1964], or a more flexible likelihood function that is especially designed to handle nontraditional residual distributions [e.g., *Schoups and Vrugt*, 2010; *Erdal et al.*, 2012] will further improve the modeling results. This is beyond the scope of the current paper, and should be explored in future studies.

### 5. Summary and Conclusions

[89] This study presents two strategies to speed up Bayesian inference of a highly parameterized and CPU-demanding groundwater flow model. Both methods build on the stochastic collocation approach of *Marzouk and Xiu* [2009] and use the generalized polynomial chaos (gPC) theory to construct a CPU-efficient approximation of the original porous flow simulator. This surrogate model serves to efficiently approximate the posterior parameter distribution of the original model using MCMC simulation with DREAM<sub>(ZS)</sub>. This distribution accurately represents the actual posterior parameter distribution if the polynomial approximation used in the surrogate model accurately summarizes the behavior of the original CPU-demanding model. In practice, however, perfect emulation of a groundwater flow simulator is difficult, and in this paper we therefore introduced two alternative computational strategies to estimate as closely and consistently as possible the posterior parameter distribution. To maximize computational efficiency both approaches were implemented using KL transformation of 26,112 dimensions of the original 26,120 dimensional parameter space, leading to a total of 102 unknown parameters. The first strategy (DR-A) used the posterior samples derived from a linear gPC-based surrogate model to precondition MCMC simulation of the original flow model. The second strategy (DR-B) combined the first strategy with state-of-the-art two-stage MCMC simulation [*Cui et al.*, 2011] in which the error introduced by the use of a reduced-order model is adaptively modeled.

[90] Our results demonstrate the advantages of the proposed two-stage MCMC inversion strategy to estimate the posterior parameter distribution of a CPU-intensive groundwater flow model. A speed up in computational efficiency on the order of two to five times was observed. Future work will investigate the applicability of the proposed framework to MCMC simulation of transient modeling problems involving nonlinear variably saturated flow conditions. In this situation, a linear gPC emulation might not be optimal. This would require a nonlinear gPC emulator, significantly enhancing the computational demands, in particular for parameter-rich CPU-demanding hydrogeologic models. To be practically useful this might necessitate a gPC approximation that considers only the most sensitive parameters (dimensions) of the original CPU intensive models.

### Appendix A

[91] We derive equation (26) of the paper. This derivation follows *Cui et al.* [2011].

[92] If we denote the proposal distribution with  $q(\mathbf{z}, \mathbf{z}')$ , the Metropolis acceptance probability in the first step,  $\alpha(\mathbf{z}, \mathbf{z}')$  becomes:

$$\alpha(\mathbf{z}, \mathbf{z}') = \min \left[ 1, \frac{\tilde{p}_{\mathbf{z}}^*(\mathbf{z}'|\mathbf{d}) q(\mathbf{z}', \mathbf{z})}{\tilde{p}_{\mathbf{z}'}^*(\mathbf{z}|\mathbf{d}) q(\mathbf{z}, \mathbf{z}')} \right]. \quad (\text{A1})$$

[93] For symmetric proposal distributions such as used in DREAM<sub>(ZS)</sub>,  $q(\mathbf{z}, \mathbf{z}') = q(\mathbf{z}', \mathbf{z})$  and equation (A1) reduces to equation (21).

[94] The acceptance probability ratio in the second step,  $\beta(\mathbf{z}, \mathbf{z}')$  is given by

$$\beta(\mathbf{z}, \mathbf{z}') = \min \left[ 1, \frac{p(\mathbf{z}'|\mathbf{d}) q^*(\mathbf{z}', \mathbf{z})}{p(\mathbf{z}|\mathbf{d}) q^*(\mathbf{z}, \mathbf{z}')} \right], \quad (\text{A2})$$

with modified jump probability

$$q^*(\mathbf{z}, \mathbf{z}') = q(\mathbf{z}, \mathbf{z}') \alpha(\mathbf{z}, \mathbf{z}'), \quad (\text{A3})$$

$$q^*(\mathbf{z}', \mathbf{z}) = q(\mathbf{z}', \mathbf{z}) \alpha(\mathbf{z}', \mathbf{z}). \quad (\text{A4})$$

[95] If we insert equations (A3) and (A4) into equation (A2) we derive

$$\beta(\mathbf{z}, \mathbf{z}') = \min \left[ 1, \frac{p(\mathbf{z}'|\mathbf{d}) q(\mathbf{z}', \mathbf{z}) \alpha(\mathbf{z}', \mathbf{z})}{p(\mathbf{z}|\mathbf{d}) q(\mathbf{z}, \mathbf{z}') \alpha(\mathbf{z}, \mathbf{z}')} \right], \quad (\text{A5})$$

which for a symmetric proposal distribution becomes equivalent to equation (26) of our paper

$$\beta(\mathbf{z}, \mathbf{z}') = \min \left\{ 1, \frac{p(\mathbf{z}'|\mathbf{d}) \times \min \left[ 1, \frac{\tilde{p}_{\mathbf{z}'}^*(\mathbf{z}|\mathbf{d})}{\tilde{p}_{\mathbf{z}}^*(\mathbf{z}'|\mathbf{d})} \right]}{p(\mathbf{z}|\mathbf{d}) \times \min \left[ 1, \frac{\tilde{p}_{\mathbf{z}}^*(\mathbf{z}'|\mathbf{d})}{\tilde{p}_{\mathbf{z}'}^*(\mathbf{z}|\mathbf{d})} \right]} \right\}, \quad (\text{A6})$$

[96] This concludes the derivation.

[97] **Acknowledgments.** We would like to thank the Associate Editor and three anonymous referees for their useful comments and suggestions that have greatly improved the current version of this manuscript. We thank Youssef Marzouk for helpful discussions. John Burkardt, Paul G. Constantine and Alexander Ihler are gratefully acknowledged for their online MATLAB toolboxes which are used extensively in the present paper.

### References

Babuska, I., F. Nobile, and R. Tempone (2007), A stochastic collocation method for elliptic partial differential equations with random input data, *SIAM J. Numer. Anal.*, 47, 1005–1034, doi:10.1137/050645142.

Beerten, K., I. Wemaere, M. Gedeon, S. Labat, B. Rogiers, D. Mallants, S. Salah, and B. Leterme (2010), Geological, hydrogeological and hydrological data for the Dessel disposal site, in Project Near Surface Disposal of category A waste at Dessel, STB-SIE(HYD)–Version 1, NIROND-TR 2009-05 E, 261 pp., NIRAS/ONDRAF, Brussels, Belgium.

Box, G. E. P., and D. R. Cox (1964), An analysis of transformations, *J. R. Stat. Soc. Ser. B*, 26, 211–246.

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 Resour. Res.*, 22(2), 199–210, doi:10.1002/WR022i002p00199.

Carrera, J., and S. P. Neuman (1986b), Estimation of aquifer parameters under transient and steady state conditions: 2. Uniqueness, stability, and solution algorithms, *Water Resour. Res.*, 22(2), 211–227, doi:10.1002/WR022i002p00211.

- Carrera, J., A. Alcolea, A. Medina, J. Hidalgo, and J. Luit (2005), Inverse problem in hydrogeology, *Hydrogeol. J.*, 13, 206–222.
- Christen, A., and C. Fox (2005), MCMC using an approximation. *J. Comput. Graphical Stat.*, 14(4), 795–810, doi:10.1198/106186005X76983.
- Ciriello, V., V. D. Federico, M. Riva, F. Cadini, J. D. Sanctis, E. Zio, and A. Guadagnini (2013), Polynomial chaos expansion for global sensitivity analysis applied to a model of radionuclide migration in a randomly heterogeneous aquifer, *Stoch. Environ. Res. Risk Assess.*, 27, 945–954, doi:10.1007/s00477-012-0616-7.
- Cooley, R. L., and S. Christensen (2006), Bias and uncertainty in regression calibrated models of groundwater flow in heterogeneous media, *Adv. Water Resour.*, 29, 639–656, doi:10.1016/j.advwatres.2005.07.012.
- Cui, T., C. Fox, and M. J. O’Sullivan (2011), Bayesian calibration of a large-scale geothermal reservoir model by a new adaptive delayed acceptance Metropolis Hastings algorithm, *Water Resour. Res.*, 47, W10521, doi:10.1029/2010WR010352.
- Der Kiureghian, A., and P. L. Liu (1986), Structural reliability under incomplete probability information, *J. Eng. Mech.*, 112, 85–104.
- Dostert P., Y. Efendiev, T. Hou, and W. Luo (2006), Coarse-gradient Langevin algorithms for dynamic data integration and uncertainty quantification, *J. Comput. Phys.*, 217, 123–142, doi:10.1016/j.jcp.2006.03.012.
- Dostert, P., Y. Efendiev, and B. Mohanty (2009), Efficient uncertainty quantification techniques in inverse problems for Richards’ equation using coarse-scale simulation models, *Adv. Water Resour.*, 32, 329–339, doi:10.1016/j.advwatres.2008.11.009.
- Efendiev, Y., A. Datta-Gupta, V. Ginting, X. Ma, and B. Mallick (2005), An efficient two-stage Markov chain Monte Carlo method for dynamic data integration, *Water Resour. Res.*, 41, W12423, doi:10.1029/2004WR003764.
- Eldred, M. S. (2011), Design under uncertainty employing stochastic expansion methods, *Int. J. Uncertainty Quantification*, 1(2), 119–146, doi:10.1615/IntJUncertaintyQuantification.v1.i2.
- Erdal, D., I. Neuweiler, and J. A. Huisman (2012), The use of error models to estimate effective unsaturated flow parameters, *Water Resour. Res.*, 48, W06530, doi:10.1029/2011WR011062.
- Fajraoui, N., F. Ramasomanana, A. Younes, T. A. Mara, P. Ackerer, and A. Guadagnini (2011), Use of global sensitivity analysis and polynomial chaos expansion for interpretation of nonreactive transport experiments in laboratory-scale porous media, *Water Resour. Res.*, 47, W02521, doi:10.1029/2010WR009639.
- Fajraoui, N., T. A. Mara, A. Younes, and R. Bouhlila (2012), Reactive transport parameter estimation and global sensitivity analysis using sparse polynomial chaos expansion, *Water Air Soil Pollut.*, 223, 4183–4197, doi:10.1007/s11270-012-1183-8.
- Fu, J., and J. Gomez-Hernandez (2009), Uncertainty assessment and data worth in groundwater flow and mass transport modeling using a blocking Markov chain Monte Carlo method, *J. Hydrol.*, 364, 328–341.
- Gedeon, M. (2008), Neogene aquifer model, geological disposal of conditioned high-level and long-lived radioactive waste, Rep. SCK•CEN-ER-48, 110 pp., SCK-CEN. Mol, Belgium.
- Gedeon, M., and D. Mallants (2012), Sensitivity analysis of a combined groundwater flow and solute transport model using local grid refinement: A case study, *Math. Geosci.*, 44, 881–899. doi:10.1007/s11004-012-9416-3.
- Gedeon, M., D. Mallants, K. Vandersteen, E. Laloy, and B. Rogiers (2011), Hydrogeological modelling of the Dessel site, Overview Rep., NIR-OND-TR 2008-15 E V2, 220 pp., NIRAS/ONDRAF, Brussels, Belgium.
- Gelman, A. G., and D. B. Rubin (1992), Inference from iterative simulation using multiple sequences, *Stat. Sci.*, 7, 457–472.
- Gerstner, T., and M. Griebel (2003), Dimension-adaptive tensor-product quadrature, *Computing*, 71, 65–87, doi:10.1007/s00607-003-0015-5.
- Geweke, J. (1992), Evaluating the accuracy of sampling-based approaches to calculating posterior moments, in *Bayesian Statistics 4*, edited by J. M. Bernardo et al., pp. 169–194, Clarendon, Oxford, U. K.
- Geyer, C. J. (2010), Introduction to Markov Chain Monte Carlo, in *Handbook of Markov Chain Monte Carlo: Methods and Applications*, edited by A. Gelman, S. Brooks, G. Jones and X. L. Meng, CRC Press, London.
- Ghanem R. G., and P. D. Spanos (1991), *Stochastic Finite Elements: A Spectral Approach*, Springer, New York, Berlin, Heidelberg, London, Paris, Tokyo, Barcelona.
- Green, P. J. (1995), Reversible jump Markov chain Monte Carlo computation and Bayesian model determination, *Biometrika*, 82, 711–732.
- Haario, H., E. Saksman, and J. Tamminen (2001), An adaptive metropolis algorithm, *Bernoulli*, 7, 223–242.
- Haario H., M. Laine, A. Mira, and E. Saksman (2006), DRAM: Efficient adaptive MCMC, *Stat. Comput.*, 16, 339–354.
- Harbaugh, A. W. (2005), MODFLOW-2005, the U.S. Geological Survey modular ground-water model—The Ground-Water Flow Process, Tech. and Methods 6–A16, U. S. Geol. Surv., Reston, Va.
- Hastings, H. (1970), Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, 57, 97–109.
- Hernandez, A., S. Neuman, A. Guadagnini, and J. Carrera (2006), Inverse stochastic moment analysis of steady state flow in randomly heterogeneous media, *Water Resour. Res.*, 42, W05425, doi:10.1029/2005WR004449.
- Kaipio, J. P., and E. Somersalo (2007), Statistical inverse problems: Discretization, model reduction and inverse crimes, *J. Comput. Appl. Math.*, 198(2), 493504.
- Keating, E. H., J. Doherty, J. A. Vrugt, and Q. Kang (2010), Optimization and uncertainty assessment of strongly nonlinear groundwater models with high parameter dimensionality, *Water Resour. Res.*, 46, W10517, doi:10.1029/2009WR008584.
- Kennedy, M. C., and A. O’Hagan (2001), Bayesian calibration of computer models, *J. R. Statist. Soc. Ser. B*, 63, 425–464.
- Kitanidis, P. K. (1986), Parameter uncertainty in estimation of spatial functions: Bayesian analysis, *Water Resour. Res.*, 22, 499–507, doi:10.1029/WR022i004p00499.
- Kitanidis, P. K. (1997), The minimum structure solution to the inverse problem, *Water Resour. Res.*, 33, 2263–2272, doi:10.1029/97WR01619.
- Koltermann, C., and S. Gorelick (1996), Heterogeneity in sedimentary deposits: A review of structure imitating, process imitating, and descriptive approaches, *Water Resour. Res.*, 32(9), 2617–2658.
- Laloy, E., and J. A. Vrugt (2012), High-dimensional posterior exploration of hydrologic models using multiple-try DREAM<sub>(ZS)</sub> and high performance computing, *Water Resour. Res.*, 48, W01526, doi:10.1029/2011WR010608.
- Le Maître, O. P., and O. M. Knio (2010), *Spectral Methods for Uncertainty Quantification*, Springer, Dordrecht, Heidelberg, London, and New York.
- Li, W. and O. A. Cirpka (2006), Efficient geostatistical inverse methods for structured and unstructured grids, *Water Resour. Res.*, 42, W06402, doi:10.1029/2005WR004668.
- Li, H., and D. Zhang (2007), Probabilistic collocation method for flow in porous media: Comparisons with other stochastic methods, *Water Resour. Res.*, 43, W09409, doi:10.1029/2006WR005673.
- Li, W., Z. Lu, and D. Zhang (2009), Stochastic analysis of unsaturated flow with probabilistic collocation method, *Water Resour. Res.*, 45, W08425, doi:10.1029/2008WR007530.
- Liu, J. S. (2001), *Monte Carlo Strategies in Scientific Computing*, Springer, New York.
- Loève, M. (1977), *Probability Theory*, 4th ed., Springer, New York.
- Mariethoz, G., P. Renard, and J. Caers (2010), Bayesian inverse problem and optimization with iterative spatial resampling, *Water Resour. Res.*, 46, W11530, doi:10.1029/2010WR009274.
- Marzouk, Y. M., and H. N. Najm (2009), Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems, *J. Comput. Phys.*, 118.6, 862–1902, doi:10.1016/j.jcp.2008.11.024.
- Marzouk, Y. M., and D. B. Xiu (2009), A stochastic collocation approach to Bayesian inference in inverse problems, *Commun. Comput. Phys.*, 6, 826–847.
- Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953), Equation of state calculations by fast computing machines, *J. Chem. Phys.*, 21, 1087–1092.
- Mondal, A., Y. Efendiev, B. Mallick, and A. Datta-Gupta (2010), Bayesian uncertainty quantification for flows in heterogeneous porous media using reversible jump Markov chain Monte Carlo methods, *Adv. Water Resour.*, 33(3), 241–256.
- Nobile, F., R. Tempone, and C. G. Webster (2008a), A sparse grid stochastic collocation method for partial differential equations with random input data, *SIAM J. Numer. Anal.*, 46, 2411–2442, doi:10.1137/060663660.
- Nobile, F., R. Tempone, and C. G. Webster (2008b), An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data, *SIAM J. Numer. Anal.*, 46, 2411–2442, doi:10.1137/070680540.
- Nowak, W., F. P. J. de Barros, and Y. Rubin (2010), Bayesian geostatistical design: Task driven optimal site investigation when the geostatistical

- model is uncertain, *Water Resour. Res.*, *46*, W03535, doi:10.1029/2009WR008312.
- Oliver, D., L. Cunha, and A. Reynolds (1997), Markov chain Monte Carlo methods for conditioning a logpermeability field to pressure data, *Math. Geosci.*, *29*, 61–91.
- Raftery, A. E., and S. M. Lewis (1992), One long run with diagnostics: Implementation strategies for Markov chain Monte Carlo, *Stat. Sci.*, *7*, 493–497.
- Razavi, S., B. A. Tolson, and D. H. Burn (2012), Review of surrogate modeling in water resources, *Water Resour. Res.*, *48*, W07401, doi:10.1029/2011WR011527.
- Rogiers, B., M. Schiltz, K. Beerten, M. Gedeon, D. Mallants, O. Batelaan, A. Dassargues, and M. Huysmans (2010), Groundwater model parameter identification using a combination of cone-penetration tests and borehole data, paper presented at International Association for Hydro-Environment Engineering and Research (IAHR) International Groundwater Symposium, Valencia.
- Rogiers, B., D. Mallants, O. Batelaan, M. Gedeon, M. Huysmans, and A. Dassargues (2012a), The usefulness of CPT data for deterministic spatially heterogeneous, large-scale aquitard parameterization, in *Models—Repositories of Knowledge, Proceedings ModelCARE2011*, edited by S. Oswald, IAHS Publ. 355, IAHS Press, Wallingford, Oxfordshire, 2011, pp. 41–47, ISBN:978-1-907161-34-6.
- Rogiers, B., D. Mallants, O. Batelaan, M. Gedeon, M., H. Huysmans, and A. Dassargues (2012b), Estimation of hydraulic conductivity and its uncertainty from grain-size data using GLUE and artificial neural networks, *Math. Geosci.*, *44*, 739–763, doi:10.1007/s11004-012-9409-2.
- Rosenblatt, M. (1952), Remarks on a multivariate transformation, *Ann. Math. Stat.*, *23*, 470–472.
- Saad, G., and R. Ghanem (2009), Characterization of reservoir simulation models using a polynomial chaos-based ensemble Kalman filter, *Water Resour. Res.*, *45*, W04417, doi:10.1029/2008WR007148.
- Saltelli, A., M. Ratto, S. Tarantola, and F. Campolongo (2006), Sensitivity analysis practices: Strategies for model-based inference, *Reliability Eng. Syst. Safety*, *91*, 1109–1125, doi:10.1016/j.res.2005.11.014.
- Saltelli, A., P. Annoni, I. Azzini, F. Campolongo, M. Ratto, and S. Tarantola (2010), Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, *Comput. Phys. Commun.*, *181*, 259–270, doi:10.1016/j.cpc.2009.09.018.
- Schoups, G., and J. A. Vrugt (2010), A formal Likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic and non-Gaussian errors, *Water Resour. Res.*, *46*, W10531, doi:10.1029/2009WR008933.
- Smolyak, S. (1963), Quadrature and interpolation formulas for tensor products of certain classes of functions, *Soviet Math. Doklady*, *4*, 240–243.
- Sobol', I. M. (1993), Sensitivity estimates for nonlinear mathematical models, *Math. Model. Comput. Exp.*, *1*, 407–414.
- Sudret, B. (2008), Global sensitivity analysis using polynomial chaos expansions, *Reliability Eng. Syst. Safety*, *93*, 964–979, doi:10.1016/j.res.2007.04.002.
- ter Braak, C. J. F. (2006), A Markov Chain Monte Carlo version of the genetic algorithm differential evolution: Easy Bayesian computing for real parameter space, *Stat. Comput.*, *16*, 239–249, doi:10.1007/s11222-006-8769-1.
- ter Braak, C. J. F., and J. A. Vrugt (2008), Differential evolution Markov chain with snooker updater and fewer chains, *Stat. Comput.*, *18*, 435–446, doi:10.1007/s11222-008-9104-9.
- Tonkin, M., and J. Doherty (2005), A hybrid regularized inversion methodology for highly parameterized environmental models, *Water Resour. Res.*, *41*, W10412, doi:10.1029/2005WR003995.
- Tonkin, M., and J. Doherty (2009), Calibration constrained Monte Carlo analysis of highly parameterized models using subspace techniques, *Water Resour. Res.*, *45*, W00B10, doi:10.1029/2007WR006678.
- Troldborg, M., W. Nowak, N. Tuxen, P. L. Bjerg, R. Helmig, and P. J. Binning (2010), Uncertainty evaluation of mass discharge estimates from a contaminated site using a fully Bayesian framework, *Water Resour. Res.*, *46*, W12552, doi:10.1029/2010WR009227.
- Vrugt, J. A., H. V. Gupta, W. Bouten, and S. Sorooshian (2003), A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, *Water Resour. Res.*, *39*(8), 1201, doi:10.1029/2002WR001642.
- Vrugt, J. A., C. J. F. ter Braak, M. P. Clark, J. M. Hyman, and B. A. Robinson (2008), Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation, *Water Resour. Res.*, *44*, W00B09, doi:10.1029/2007WR006720.
- Vrugt, J. A., C. J. F. ter Braak, C. G. H. Diks, D. Higdon, B. A. Robinson, and J. M. Hyman (2009), Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling, *Int. J. Nonlinear Sci. Numer. Simul.*, *10*, 273–290.
- Wolff, U. (2004), Monte Carlo errors with less errors, *Comput. Phys. Commun.*, *156*(2), 143–153, doi:10.1016/S0010-4655(03)00467-3.
- Xiu, D., and G. E. Karniadakis (2002), The Wiener-Askey polynomial chaos for stochastic differential equations, *SIAM J. Sci. Comput.*, *24*, 619–644.
- Xiu, D., and J. S. Hesthaven (2005), High-order collocation methods for differential equations with random inputs, *SIAM J. Sci. Comput.*, *27*, 1118–1139.
- Xiu, D. B. (2007), Efficient collocational approach for parametric uncertainty analysis, *Commun. Comput. Phys.*, *2*(2), 293–309.
- Yu, L., B. Rogiers, M. Gedeon, I. Wemaer, J. Marivoet, M. De Craen, and D. Mallants (2013), Hydraulic conductivity of “undisturbed” Boom Clay, *Appl. Clay Sci.*, *75–76*, 1–12.
- Zeng, L., and D. Zhang (2010), A stochastic collocation based Kalman filter for data assimilation, *Comput. Geosci.*, *14*, 721–744, doi:10.1007/s10596-010-9183-5.
- Zeng, L., L. Shi, D. Zhang, and L. Wu (2012), A sparse grid based Bayesian method for contaminant source identification, *Adv. Water Resour.*, *37*, 1–9, doi:10.1016/j.advwatres.2011.09.011.