AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2016WR019518

Key Points:
- High input dimension is reduced through KL expansion and SIR
- A PCE surrogate is constructed over the low-dimensional subspace
- A two-stage MC simulation is implemented for failure probability analysis

Correspondence to:
L. Zeng,
lingzao@zju.edu.cn

Citation:
Zhang, J., W. Li, G. Lin, L. Zeng, and L. Wu (2017), Efficient evaluation of small failure probability in high-dimensional groundwater contaminant transport modeling via a two-stage Monte Carlo method, *Water Resour. Res.*, *53*, 1948–1962, doi:10.1002/2016WR019518.

Received 16 JUL 2016
Accepted 11 FEB 2017
Accepted article online 15 FEB 2017
Published online 7 MAR 2017

Efficient evaluation of small failure probability in high-dimensional groundwater contaminant transport modeling via a two-stage Monte Carlo method

Jiangjiang Zhang[1], Weixuan Li[2], Guang Lin[3], Lingzao Zeng[1], and Laosheng Wu[4]

[1]Zhejiang Provincial Key Laboratory of Agricultural Resources and Environment, Institute of Soil and Water Resources and Environmental Science, College of Environmental and Resource Sciences, Zhejiang University, Hangzhou, China, [2]Pacific Northwest National Laboratory, Richland, Washington, USA, [3]Department of Mathematics and School of Mechanical Engineering, Purdue University, West Lafayette, Indiana, USA, [4]Department of Environmental Sciences, University of California, Riverside, California, USA

Abstract In decision-making for groundwater management and contamination remediation, it is important to accurately evaluate the probability of the occurrence of a failure event. For small failure probability analysis, a large number of model evaluations are needed in the Monte Carlo (MC) simulation, which is impractical for CPU-demanding models. One approach to alleviate the computational cost caused by the model evaluations is to construct a computationally inexpensive surrogate model instead. However, using a surrogate approximation can cause an extra error in the failure probability analysis. Moreover, constructing accurate surrogates is challenging for high-dimensional models, i.e., models containing many uncertain input parameters. To address these issues, we propose an efficient two-stage MC approach for small failure probability analysis in high-dimensional groundwater contaminant transport modeling. In the first stage, a low-dimensional representation of the original high-dimensional model is sought with Karhunen-Loève expansion and sliced inverse regression jointly, which allows for the easy construction of a surrogate with polynomial chaos expansion. Then a surrogate-based MC simulation is implemented. In the second stage, the small number of samples that are close to the failure boundary are re-evaluated with the original model, which corrects the bias introduced by the surrogate approximation. The proposed approach is tested with a numerical case study and is shown to be 100 times faster than the traditional MC approach in achieving the same level of estimation accuracy.

1. Introduction

Accurate predictions of groundwater flow and solute transport are hampered by many sort of uncertainties originated from the heterogeneity of the subsurface environment, insufficiency of site characterization, and limitation of conceptual-mathematical models [*Anderson et al.*, 2015; *Bear and Cheng*, 2010; *Bolster et al.*, 2009]. Although these uncertainties can be reduced to some extent from available measurements [*Schöniger et al.*, 2012; *Wagner*, 1992; *Yeh et al.*, 2007], they are still ubiquitous due to the limitation of budget, time, and techniques.

Under model structural and parametric uncertainties, the probability of the occurrence of a failure event, also called the failure probability, is a prerequisite for risk assessment, reliability analysis, and decision-making in many areas, including contaminant transport, civil engineering, oil and gas transmission [*Bolster and Tartakovsky*, 2008; *Der Kiureghian and Liu*, 1986; *Hamed and Bedient*, 1997; *Tartakovsky*, 2013; *Yuhua and Datao*, 2005]. In groundwater contaminant transport modeling, it is important to quantify the probability that the contaminant concentration or the total discharge exceeds an allowed level [*Tartakovsky*, 2013], which has been used to determine remediation goals [*Batchelor et al.*, 1998], to guide reliable groundwater remediation strategies [*Barros et al.*, 2013; *Bolster et al.*, 2009; *Fernàndez-Garcia et al.*, 2012; *Troldborg et al.*, 2012; *Wang and McTernan*, 2002], to support policy-makings [*Massmann and Freeze*, 1987a, 1987b] and to construct well-head protection areas [*Enzenhoefer et al.*, 2015; *Rodak and Silliman*, 2012]. For a detailed review of the assessment and management of risks in subsurface hydrology, one can refer to *Tartakovsky* [2013].

The most straightforward approach to obtain an estimate of the failure probability is the Monte Carlo (MC) simulation, which simply evaluates the model at a large number of sample points and counts the number of

failure samples. A rule of thumb for reliable failure probability estimation is to obtain at least 10 samples in the failure region [*Li and Xiu*, 2010]. To achieve a given relative accuracy, the number of samples needed by the MC simulation is proportional to the inverse of the failure probability. For a system with a small failure probability (e.g., $\leq 10^{-3}$), the total number of model evaluations needed in the MC simulation would be very large. In complex large-scale problems, the involved groundwater flow and solute transport models are CPU-demanding [*Asher et al.*, 2015], which makes MC simulation for small failure probability analysis very computationally intensive. To efficiently explore the failure region, some advanced sampling techniques can be adopted, e.g., importance sampling [*Au and Beck*, 1999; *Kanj et al.*, 2006] and subset simulation [*Au and Beck*, 2001; *Elsheikh et al.*, 2014]. If we adopt importance sampling or subset simulation in failure probability analysis, we have to design either a proposal distribution or intermediate events to make it easier to sample from the failure region. In this way, much fewer samples are required. Due to this appealing property, importance sampling and similar methods have been widely used in failure probability analysis [*Bucklew*, 2004].

To alleviate the computational burden encountered in the MC simulation for small failure probability analysis, another promising approach is to replace the original model with an inexpensive surrogate (also called a response surface, an emulator or a metamodel in different literatures). Generally, the techniques of building surrogates can be grouped into three categories, i.e., data-driven, projection, and hierarchical-based methods [*Smith*, 2013]. Among the three methods, the data-driven method, which includes interpolation and regression approaches, is nonintrusive, i.e., the original model can be treated as a black box. Thus, it has gained popularity in groundwater modeling. Examples of data-driven surrogates used in groundwater modeling include polynomial chaos expansion (PCE) [*Ciriello et al.*, 2013; *Laloy et al.*, 2013; *Liao and Zhang*, 2013, 2014, 2016], adaptive sparse grid interpolation [*L. Zeng et al.*, 2012; *X. Zeng et al.*, 2016; *G. Zhang et al.*, 2013; *J. Zhang et al.*, 2015], artificial neural networks [*Chen et al.*, 2014; *Yan and Minsker*, 2006], kriging/Gaussian process emulators [*Marrel et al.*, 2008; *Zhang et al.*, 2016; *Zhao et al.*, 2016], and the autoregressive moving average model [*Felisa et al.*, 2015]. Some reviews about surrogate construction methods in water resources can be found in *Asher et al.* [2015] and *Razavi et al.* [2012]. As demonstrated by many researchers, using surrogate models in failure probability analysis is much more efficient than using the original model-based analysis [*Li and Xiu*, 2010; *Paffrath and Wever*, 2007; *Pulch*, 2010]. When integrated with sampling methods such as importance sampling or subset simulation, the surrogate-based approaches can be even more efficient and applicable for rare failure probability analysis [*Balesdent et al.*, 2013; *Dubourg and Sudret*, 2014; *Dubourg et al.*, 2011; *Dubourg et al.*, 2013; *Li et al.*, 2011]. However, for complex systems, when the dimension of parameters is high (e.g., 200 for the case study tested in the present work), it is almost impossible to design a proper proposal distribution (or intermediate events) for the event with a small failure probability. To guarantee the estimation accuracy in high-dimensional problems, importance sampling or similar methods are not adopted in this paper.

As the failure region is usually far away from the mean values of input parameters, the surrogate constructed over the prior distribution of input parameters usually cannot provide an accurate estimate of the small failure probability when it is directly used in the MC simulation. To circumvent this problem, two different strategies can be adopted. In the first strategy, the surrogate model is adaptively refined around the failure boundary (also called the limit state in different literatures) through a design of computer experiments. For example, *Wang et al.* [2016] implemented the surrogate construction as a Bayesian inference to identify the failure boundary through an information-theoretic optimal experimental design. In the second strategy, one first uses a surrogate constructed over the prior distribution to select a certain number of samples that are close to the failure boundary and then re-evaluates these samples with the original model. By doing so, the inaccuracy of the direct surrogate-based MC simulation can be alleviated. Moreover, this approach is much more efficient than the fully original model-based approach [*Li and Xiu*, 2010, 2014; *Li et al.*, 2011]. In this paper, the second strategy is adopted.

However, the surrogate construction suffers from the so-called "curse of dimensionality," i.e., its computational cost increases dramatically with the input dimension and system nonlinearity. For example, if a surrogate is constructed by interpolation based on the full tensor product scheme, the computational cost will grow exponentially with the input dimension [*Klimke and Wohlmuth*, 2005]. Although using sparse grids can greatly reduce the number of original model evaluations in the surrogate construction [*Klimke and Wohlmuth*, 2005], for problems with high-dimensional inputs, the computational cost is still unacceptable. In large-scale groundwater modeling, the number of unknown parameters is usually huge. In this case, the surrogate-based MC simulation could be even less efficient than the original model-based MC simulation.

To address this issue, a dimension reduction technique should be implemented on the high-dimensional input variables to allow for the easy construction of a surrogate.

Dimension reduction seeks to discover a low-dimensional representation of high-dimensional data, where the low-dimensional subspace can still keep most of the original variability [*Bures*, 2010; *Roweis and Saul*, 2000]. This process can be linear, as in the principle component analysis (PCA) [*Jolliffe*, 1986], or nonlinear, as in the kernel PCA [*Schölkopf et al.*, 1997] and locally linear embedding method [*Roweis and Saul*, 2000]. Due to the ease of implementation, the linear dimension reduction by the PCA has been widely used. In the area of groundwater modeling, the Karhunen-Loève (KL) expansion (another name of the PCA in signal processing) has been widely used to represent the heterogeneous log hydraulic conductivity field with a small number of uncorrelated random variables [*Dai et al.*, 2016; *Laloy et al.*, 2013; *Li and Zhang*, 2007; *Li et al.*, 2009; *Shi et al.*, 2009; *Vermeulen et al.*, 2004; *Yang et al.*, 2004; *Zhang and Lu*, 2004; *Zhang et al.*, 2015]. With the KL expansion, the dimension of parameters for the conductivity field can be reduced from model grid number to the number of truncated KL terms.

For large-scale problems, however, the dimension of input parameters after the KL expansion is usually still too high for surrogate construction. However, it is worth noting that the KL expansion is solely based on the statistical distribution of the input parameters such as the hydraulic conductivity field. On the other hand, the dependent relationship between the quantity of interest (QoI), such as the contaminant concentration, and the input parameters is not considered, thus leaving some potential opportunity for further dimension reduction untouched. For example, the variation of a QoI in many situations is caused only by the model parameters varying in a low-dimensional subspace, known as the central subspace [*Bura and Cook*, 2001], which allows for a low-dimensional representation of the model. A simple yet efficient technique to detect the central subspace is the sliced inverse regression (SIR) [*Li*, 1991]. In a recent work, the SIR is used to enhance the computational efficiency of uncertainty quantification algorithms for high-dimensional problems [*Li et al.*, 2016].

Inspired by the previous works, here, we propose an efficient two-stage MC approach for small failure probability analysis in high-dimensional groundwater contaminant transport modeling. The detailed formulation of the methods is given in section 2. Illustrated with a synthetic example in section 3, the efficiency and accuracy of the proposed methods are well demonstrated. Finally, some conclusions are provided in section 4.

2. Methods

Consider a process where the QoI can be obtained numerically with the following model:

$$q = f(\mathbf{m}), \tag{1}$$

where $\mathbf{m}$ is a $N_m \times 1$ vector of uncertain model parameters described by some distribution function with finite variance. In this study, we consider the case that a failure event occurs when the QoI value exceeds a predefined threshold T_P, and then the failure probability is defined as:

$$P_f = P(f(\mathbf{m}) > T_P). \tag{2}$$

2.1. Monte Carlo Simulation

To calculate the failure probability, the most straightforward method is direct MC simulation. By generating N random samples of the uncertain parameters, $\mathbf{m}_1, \ldots, \mathbf{m}_N$, from the distribution of $\mathbf{m}$, a direct MC estimation of the failure probability is given as:

$$P_f^{MC} = \frac{1}{N}\sum_{i=1}^{N} \delta_{\{f(\mathbf{m})>T_P\}}(\mathbf{m}_i), \tag{3}$$

where

$$\delta_\Omega(\mathbf{m}) = \begin{cases} 1, & \text{if} \quad \mathbf{m} \in \Omega, \\ 0, & \text{if} \quad \mathbf{m} \notin \Omega, \end{cases} \tag{4}$$

is the indicator function.

To measure the estimation accuracy of P_f^{MC}, we can use the mean square error (MSE) of the direct MC estimator approximated with the sample variance:

$$\mathrm{MSE}=\frac{\mathrm{Var}(\delta_\Omega(\mathbf{m}))}{N}\approx\frac{\sum_{i=1}^{N}\left(\delta_\Omega(\mathbf{m}_i)-P_f^{MC}\right)^2}{N^2}. \tag{5}$$

As the failure probability of interest is usually small, it is more reasonable to use the relative error to measure the estimation accuracy, which is defined as:

$$Err=\frac{\sqrt{MSE}}{P_f^{MC}}. \tag{6}$$

For small failure probability, i.e., $P_f \ll 1$, to obtain a desired estimation accuracy, a very large number of MC samples are needed, which could pose a very high computational cost if a CPU-demanding groundwater model is used.

2.2. Polynomial Chaos Expansion

To efficiently estimate the failure probability, a CPU-efficient surrogate can be used to replace the original CPU-demanding model in the MC simulation. In this study, the surrogate $\tilde{f}(\mathbf{m})$ is constructed with the PCE [*Ghanem and Spanos*, 1991; *Xiu and Karniadakis*, 2002]:

$$f(\mathbf{m})\approx\tilde{f}(\mathbf{m})=\sum_{i=0}^{M-1}C_i\psi_i(\mathbf{m}), \tag{7}$$

where $C_i(i=0,\ldots,M-1)$ are deterministic coefficients to be determined, $\psi_i(\mathbf{m})$ are orthogonal polynomials over the distribution of $\mathbf{m}$, and M is the total number of terms. For the Dth degree PCE, we have $M=(N_\mathrm{m}+D)!/(N_\mathrm{m}!D!)$. Note that M grows rapidly as the dimension N_m increases.

To approximate the coefficients, one can use the stochastic collocation method:

$$C_i\approx\frac{\sum_{k=1}^{K}f(\mathbf{m}_k)\psi_i(\mathbf{m}_k)w_k}{\langle\psi_i^2(\mathbf{m})\rangle}, \tag{8}$$

where $\mathbf{m}_k(k=1,\ldots,K)$ are K collocation points in the parameter space and w_k are the corresponding weights. For more details about the implementation of the PCE, one can refer to *Xiu* [2010].

2.3. Dimension Reduction

Since the number of terms in a PCE grows fast with N_m (the dimension of input parameters), the computational cost of constructing a surrogate with the PCE would be unaffordable when N_m is large. To resolve this problem, we combine the KL expansion and the SIR to substantially reduce the input dimension.

2.3.1. Karhunen-Loève Expansion

In this study, a two-dimensional (2-D) subsurface contaminant transport problem is considered. Due to the nature of the subsurface environment, the conductivity field, which is one key parameter in water flow and solute transport simulations, is usually unknown and modeled as a random field [*Zhang*, 2002]. Since it is necessary to employ a fine discretization to resolve the high spatial variability of the conductivity field, this makes the parameters very high-dimensional. Considering the spatial correlation, the random field of log-transformed conductivity $Y(\mathbf{x})$ can be represented with the truncated KL expansion [*Marzouk and Najm*, 2009; *Zhang and Lu*, 2004]:

$$Y(\mathbf{x})\approx\bar{Y}(\mathbf{x})+\sum_{i=1}^{N_{KL}}\sqrt{\tau_i}s_i(\mathbf{x})m_i, \tag{9}$$

where $\bar{Y}(\mathbf{x})$ is the mean component, τ_i and $s_i(\mathbf{x})$ are the eigenvalues and eigenfunctions of the covariance function, respectively. In this work, the log-transformed conductivity field is assumed to be multi-Gaussian, then by definition $\mathbf{m}=[m_1,\ldots,m_{N_{KL}}]$ fit independent standard Gaussian distribution, i.e., $m_i\sim N(0,1)$ $(i=1,\ldots,N_{KL})$. Thus the dimension of the conductivity parameters is reduced from the model grid number to the preserved number of KL terms N_{KL}. For a detailed formulation and discussion of the KL expansion, one can refer to *Marzouk and Najm* [2009] and *Zhang and Lu* [2004].

2.3.2. Sliced Inverse Regression

After the dimension reduction via the KL expansion, the input dimension may be still too high for surrogate construction. Furthermore, the dimension reduction through the KL expansion does not make use of the dependence between the QoI and uncertain parameters. By considering this dependence, the input dimension can be further reduced.

In many problems, the change in the QoI is mainly caused by the variation of the uncertain parameters within a low-dimensional subspace, known as the central subspace. This property can be expressed as:

$$f(\mathbf{m}) \approx f(\mathbf{B}\mathbf{B}^T\mathbf{m}), \tag{10}$$

where $\mathbf{B}=(\boldsymbol{\beta}_1,\ldots,\boldsymbol{\beta}_K)$ is a $N_m \times K$ transformation matrix whose columns consist of a set of orthonormal basis vectors of the central subspace and thus $\mathbf{B}\mathbf{B}^T\mathbf{m}$ represents the projection of $\mathbf{m}$ onto the subspace. Once the central subspace is identified, the original model can be approximated with a low-dimensional model $q \approx g(\eta)=f(\mathbf{B}\eta)$, where $\eta=\mathbf{B}^T\mathbf{m}$ is the low-dimensional parameter vector.

The SIR is an approach that seeks the central subspace by performing a weighted principle component analysis on the covariance matrix for the estimated inverse regression curve, $E(\mathbf{m}|q)$ [*Li*, 1991]. With data $(\mathbf{m}_i, q_i)$ $(i=1,\ldots,n)$, the SIR is implemented as follows:

Step 1: Standardize the input parameters $\mathbf{m}$ to have zero means and identity covariance by the following transformation:

$$\mathbf{m}_i'=\mathbf{C}^{-1/2}(\mathbf{m}_i-\bar{\mathbf{m}}) \quad (i=1,\ldots,n), \tag{11}$$

where $\mathbf{C}$ and $\bar{\mathbf{m}}$ are the sample covariance matrix and sample mean of $\mathbf{m}$, respectively. In this work, we do not need to implement this step as $\mathbf{m}$ are standard Gaussian random variables for the KL expansion.

Step 2: Divide the range of q into H nonoverlapping slices, $J_1,\ldots,J_H$, and in each slice the number of points, $n_h\,(h=1,\ldots,H)$, is roughly equal to each other.
Step 3: To produce a crude estimate of the inverse regression curve $E(\mathbf{m}'|q)$, calculate the sample mean of $\mathbf{m}'$ within each slice:

$$\bar{\mathbf{m}}_h'=\frac{1}{n_h}\sum_{q_j\in J_h}\mathbf{m}_j' \,(h=1,\ldots,H). \tag{12}$$

Step 4: Conduct a weighted principle component analysis for the data $\bar{\mathbf{m}}_h'(h=1,\ldots,H)$ by finding the eigenvalues and eigenvectors of $\hat{\mathbf{V}}=\sum_{h=1}^{H}\frac{n_h}{n}\bar{\mathbf{m}}_h'\bar{\mathbf{m}}_h'^T$.
Step 5: Let the K eigenvectors with the largest eigenvalues be $\hat{\boldsymbol{\alpha}}_k(k=1,\ldots,K)$, then each element in the transformation matrix $\mathbf{B}$ is $\boldsymbol{\beta}_k=\hat{\boldsymbol{\alpha}}_k\mathbf{C}^{-1/2}(k=1,\ldots,K)$.

It should be noted here that we could also implement the SIR on the full random field of conductivity (the dimension of which is model grid number). As the number of data points n needed by the SIR should be at least larger than the dimension of original parameters [*Li and Yin*, 2008], this approach is usually not computationally efficient. Due to spatial correlation, the intrinsic dimensionality of the random field is much smaller than the model grid number. For the sake of computational efficiency, we first use the KL expansion to reduce the dimension from the grid number to the number of truncated KL terms, and then implement the SIR to further reduce the dimension.

An R package for the SIR is available in *Weisberg* [2002], and a complete derivation of the SIR can be found in *Li* [1991]. The convergence of SIR-based uncertainty quantification methods for high-dimensional problems has been analyzed in *Li et al.* [2016].

2.4. Two-Stage Monte Carlo Simulation

Then we can build a PCE surrogate over the low-dimensional subspace, $q \approx \tilde{g}(\eta)=\tilde{g}(\mathbf{B}^T\mathbf{m})$, which can make the MC simulation for small failure probability analysis very efficient. However, as the failure boundary is usually far away from the mean values of model input parameters, the surrogate may not be accurate around the failure boundary. Thus, the direct surrogate-based MC simulation may introduce a significant error in small failure probability analysis. To correct this error, the samples that are around the failure

boundary (identified by the surrogate evaluations) are re-evaluated with the original model. With this two-stage MC simulation, the failure probability can be numerically computed as:

$$P_f^* = \frac{1}{N}\sum_{i=1}^{N} \delta_{\{\tilde{g}(\mathbf{B}^T\mathbf{m}_i) > T_P + \gamma\}}(\mathbf{m}_i) + \frac{1}{N}\sum_{i=1}^{N} \delta_{\{\{|\tilde{g}(\mathbf{B}^T\mathbf{m}_i) - T_P| \le \gamma\} \cap \{f(\mathbf{m}_i) > T_P\}\}}(\mathbf{m}_i), \tag{13}$$

where γ is a threshold that represents the surrogate error. The first sum in equation (13) counts all the samples that lead to a failure event according to the surrogate model, whereas the second sum counts samples where a failure event cannot be confidently identified by the surrogate but need to be confirmed by evaluating the original model.

Intuitively, a larger γ will lead to more original model evaluations but a smaller numerical error, and a smaller γ will lead to fewer original model evaluations but a larger numerical error. If γ=0, the two-stage MC simulation actually reduces to the direct surrogate-based MC simulation. It has been rigorously proven by *Li and Xiu* [2010] that, if γ is well chosen, the result of the two-stage MC simulation will converge to the result of original model-based MC simulation. In this study, γ is chosen as the largest absolute error between the original model-based QoIs and the surrogate-based QoIs around the failure boundary, which is updated in the implementation of the two-stage MC simulation in the following way:

Step 1: For the samples used in the dimension reduction with the SIR, calculate the corresponding surrogate-based QoIs. Then the largest absolute error of the surrogate at these samples is chosen as the initial value of γ.
Step 2: Implement the surrogate-based MC simulation with a large number of samples.
Step 3: With the current value of γ, select the samples that fit $|\tilde{g}(\mathbf{B}^T\mathbf{m}) - T_P| \le \gamma$. Then calculate the QoIs of these samples with the original model.
Step 4: Let $r_{\max}$ be the largest absolute error of the surrogate at these selected samples. If $r_{\max} \le \gamma$, exist; if not, let $\gamma = r_{\max}$.
Step 5: Repeat Steps 3 and 4 until $r_{\max} \le \gamma$. Calculate the failure probability according to equation (13).

The flowchart for the two-stage MC method is shown in Figure 1. Here the original parameters are the truncated KL terms, from which we can obtain the log-transformed conductivity field according to equation (9).

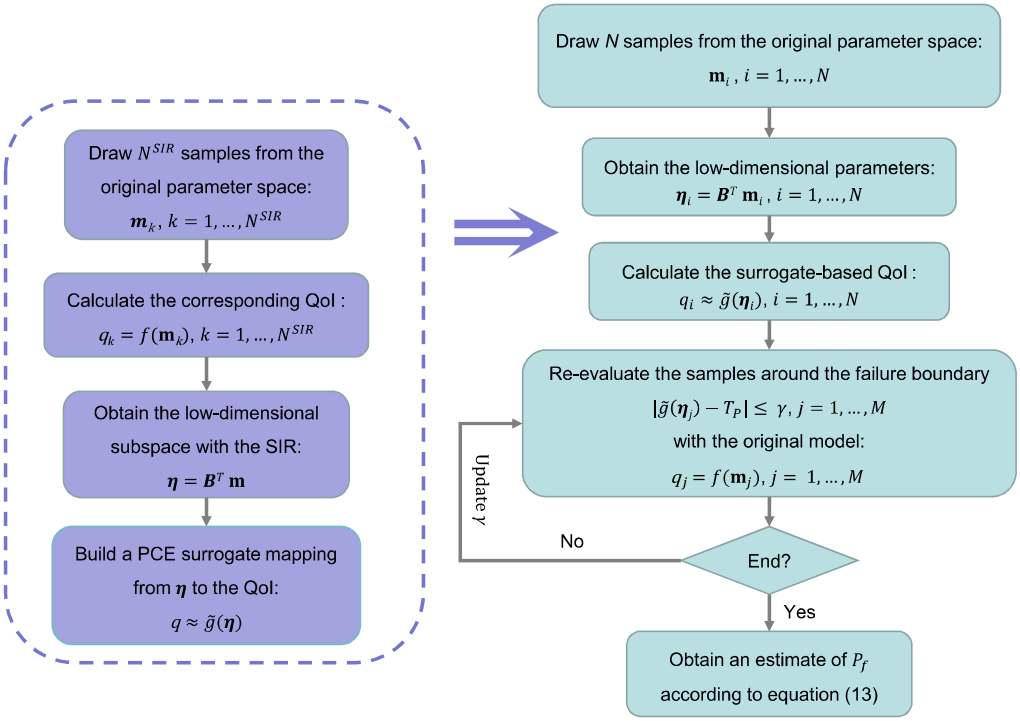

Figure 1. Flowchart for the two-stage MC method.

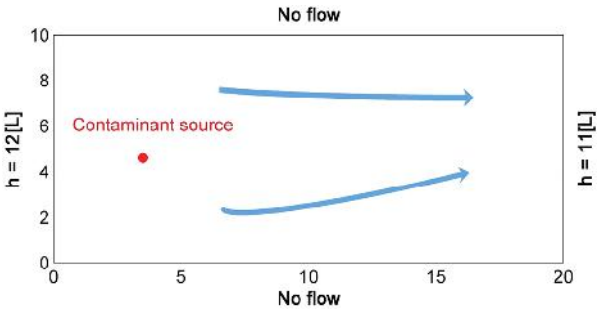

Figure 2. Flow domain for the case study.

3. Case Study

3.1. Case Description

In this study, the proposed method is tested on a 2-D groundwater flow and solute transport model.

In this case, the flow domain is $20[L]$ in x direction and $10[L]$ in y direction (Figure 2). The left and right boundaries have constant pressure heads of $12[L]$ and $11[L]$, respectively, whereas the upper and lower boundaries are no-flow boundaries. The log-transformed conductivity $Y=\log(K)$ at two arbitrary locations (x,y) and (x',y') in the domain are assumed to have a separable exponential correlation:

$$C_Y(x,y;x',y')=\sigma_Y^2\exp\left(-\frac{|x-x'|}{\lambda_x}-\frac{|y-y'|}{\lambda_y}\right), \tag{14}$$

where $\sigma_Y^2=1$ is the variance, $\lambda_x=10[L]$ and $\lambda_y=5[L]$ are the correlation length along x and y directions, respectively. Due to spatial correlation of hydraulic conductivity, we can parameterize the heterogeneous conductivity field using the truncated KL expansion as shown in equation (9). In this case, the mean component is set as $\bar{Y}(\mathbf{x})=2$, and 200 KL terms are kept. The 200 KL terms preserve about 98.8% of the total field variance, i.e., $\sum_{i=1}^{200}\tau_i/\sum_{i=1}^{\infty}\tau_i\approx 0.988$. The sufficient number of truncated KL terms that should be preserved increases with a decreasing ratio of the correlation length to the respective domain length.

In this study, MODFLOW [*Harbaugh et al.*, 2000] is used to numerically solve the governing equation for steady state saturated groundwater flow:

$$\frac{\partial}{\partial x_i}\left(K_i\frac{\partial h}{\partial x_i}\right)=0, \tag{15}$$

and the flow velocity $v_i[LT^{-1}]$ can be obtained by solving:

$$v_i=-\frac{K_i}{\theta}\frac{\partial h}{\partial x_i}, \tag{16}$$

where $x_i[L]$ is the distance along the respective Cartesian coordinate axis, $K_i[LT^{-1}]$ is the principal component of the conductivity tensor, with the assumption that it is aligned with the respective coordinate direction, $h[L]$ is the hydraulic head, and θ is the porosity of the subsurface medium.

Under steady state water flow conditions, some amount of contaminant is released from a known source (denoted by the red dot in Figure 2) during the time interval between 0 [T] and 6 [T]. In this case, the contaminant source is specified as a piecewise-in-time constant concentration cell, whose values are specified in Figure 3. Then MT3DMS [*Zheng and Wang*, 1999] is adopted to numerically solve the advection dispersion equation for the solute transport:

$$\frac{\partial(\theta C)}{\partial t}=\frac{\partial}{\partial x_i}\left(\theta D_{ij}\frac{\partial C}{\partial x_j}\right)-\frac{\partial}{\partial x_i}(\theta v_i C)+q_s C_s, \tag{17}$$

where $C[ML^{-3}]$ is the dissolved concentration of contaminant, $t[T]$ is the time, $D_{ij}[L^2T^{-1}]$ is the hydrodynamic dispersion tensor, $q_s[T^{-1}]$ is the volumetric flow rate per unit volume of aquifer, and $C_s[ML^{-3}]$ is the concentration of the contaminant source. Here the hydrodynamic dispersion tensor is defined as:

$$\begin{cases} D_{xx}=\left(\alpha_L v_x^2+\alpha_T v_y^2\right)/|v|, \\ D_{yy}=\left(\alpha_L v_y^2+\alpha_T v_x^2\right)/|v|, \\ D_{xy}=D_{yx}=(\alpha_L-\alpha_T)v_x v_y/|v|, \end{cases} \tag{18}$$

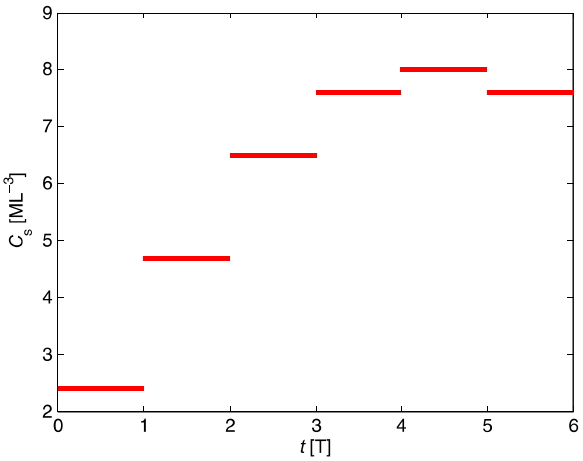

Figure 3. The concentration values of the contaminant source at different times.

where v_x and v_y are the components of pore water velocity $\boldsymbol{v}$, $|\boldsymbol{v}|$ is its magnitude, α_L and α_T are the longitudinal and transverse dispersivities, respectively.

In this case, it is assumed that the uncertainty only stems from the conductivity field. Other parameters, such as porosity, longitudinal, and transverse dispersivities are assumed to be known as $\theta=0.25$, $\alpha_L=0.3[L^2T^{-1}]$, and $\alpha_T=0.03[L^2T^{-1}]$, respectively. The QoI for this case is the integral of the mass flux of the contaminant over the injection period and across the entire right boundary.

Here we have to admit that as we only consider the uncertainty originated from the heterogeneous conductivity field, this case study is a simplification of reality. As the total discharge of a certain contaminant to the downstream is an important index in pollution control, we choose it as the QoI. We think this case study can still describe some real scenarios, i.e., what is the total discharge of a known contaminant source to the downstream given an unknown and heterogeneous conductivity field? Meanwhile, the random dimensionality is as high as 200 (truncated number of the KL expansion), which is extremely challenging for existing surrogate-based modeling methods.

3.2. Surrogate Construction

Note that building a surrogate model, say a PCE, for the 200-dimensional model (after the KL expansion) is still too expensive. To further reduce the input dimension through considering the dependence between the 200 KL terms $\mathbf{m}$ and the QoI, the SIR method is implemented. In this case study, 1000 sample points are used in the SIR, which requires 1000 original model evaluations (i.e., $N_O^{SIR}=1,000$). As described in section 2.3.2, a one-dimensional subspace is found as $\eta=\mathbf{B}^T\mathbf{m}$, which enables the easy construction of a PCE surrogate, $q\approx\tilde{g}(\eta)=\tilde{g}(\boldsymbol{B}^T\boldsymbol{m})$.

The low-dimensional parameter η obtained by the SIR is a linear combination of the original parameters $\mathbf{m}$, i.e., 200 terms for the KL expansion. The relationship between η and the QoI is usually monotonic (Figure 5a). The scalar parameter η itself has no physical meaning, but we can project it back to the original parameter space, i.e., $\mathbf{m}_\eta=\mathbf{B}\eta$, which, however, will inevitably lose a lot of information. To illustrate how this single variable η works, we draw two random samples of the log conductivity field (generated from the 200 KL terms $\mathbf{m}$) and compare them with those obtained from the corresponding low-dimensional subspace found by the SIR in Figure 4. It is obvious that the η-based fields have much simpler patterns compared to the $\mathbf{m}$-based fields, which indicates an information loss during the dimension reduction with the SIR. As the contaminant source has a constant concentration (although its value varies at different time segments), the QoI is mainly determined by the total volumetric water flux through the source, Q_s. As shown in Figures 4c–4d, the log conductivity field obtained from the low-dimensional parameter η can locate the source location, which can produce a similar Q_s as the log conductivity field obtained from the high-dimensional parameters $\mathbf{m}$. As discussed in *de Barros and Nowak* [2010], local hydraulic conditions near the contaminant source have great influence on far-field predictions. In this study, the dimension reduction technique used actually preserves the local hydraulic conditions near the contaminant source, which is most relevant with respect to the QoI, and discard other information that is not relevant. By plotting the values of η against the corresponding QoIs for these 1000 samples (Figure 5a), we can further conclude that the variation in the QoI is largely explained by the single variable η. It is understandable that η is in close relationship with the QoI and the result of the SIR will change if a different QoI is selected.

With only one input variable η and one output variable QoI, the surrogate is easily and accurately constructed with 6th degree PCE on only seven collocation points, which requires seven original model evaluations (i.e., $N_O^{PCE}=7$). The QoI values obtained from this surrogate at the 1000 samples used in the SIR are also

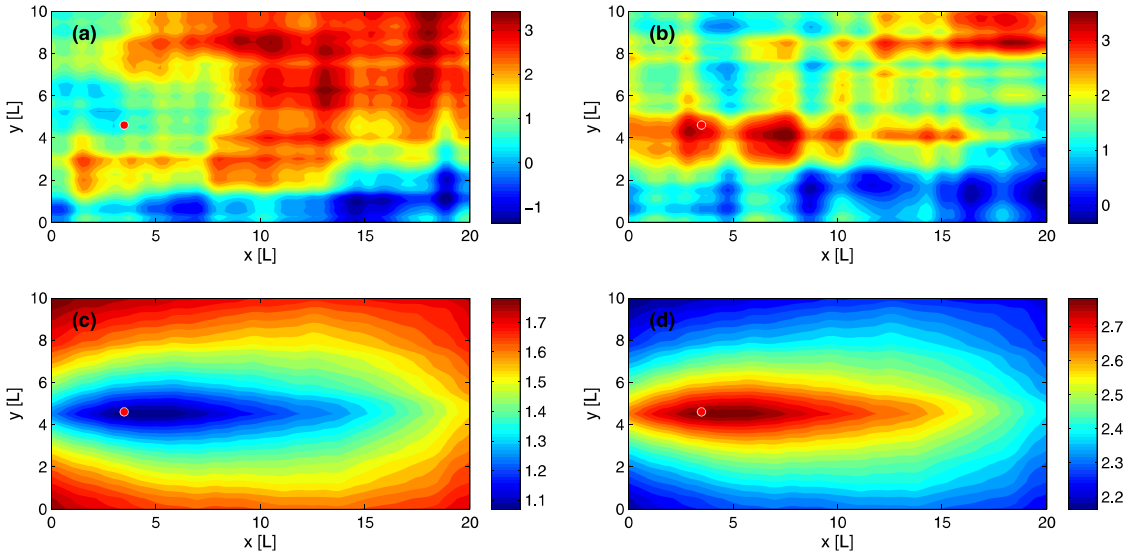

Figure 4. Two random realizations of the log conductivity field (a and b) before and (c and d) after the dimension reduction by the SIR. The contaminant source is represented with red dot in each subplot.

shown in Figure 5a (red line), and the residuals between the original model and surrogate-based QoIs are shown in Figure 5b. To further illustrate the performance of the SIR-based surrogate, we compare the QoI values obtained from the original model and the surrogate at 500 random parameter realizations. Figure 6a summarizes the results of this comparison, which shows a strong correlation between the original model and surrogate-based QoIs. If the dimension reduction with the SIR is not implemented, and only one KL term is kept, the resulting QoI predictions will be very inaccurate, even if the original model evaluations are used (Figure 6b). From Figures 5 and 6a, it is also clear that the surrogate is less accurate where the QoI is larger. Since this is the crucial region for failure probability analysis, the surrogate error cannot be neglected.

3.3. Failure Probability Analysis

To obtain an accurate estimate of a small failure probability (i.e., with a large value of threshold T_P), the two-stage MC simulation as shown in Figure 1 is implemented instead of the direct surrogate-based MC

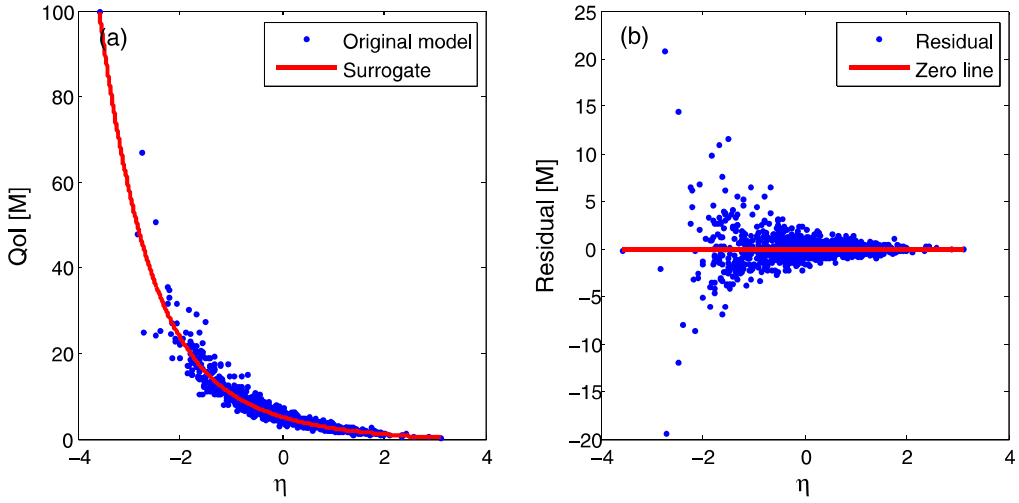

Figure 5. (a) The relationship between the single variable η and the QoI obtained from the original model (blue dots) and the surrogate (red line). (b) The residuals between the original model-based QoIs and surrogate-based QoIs. The 1000 samples are randomly drawn from the prior distribution and used in the dimension reduction by the SIR.

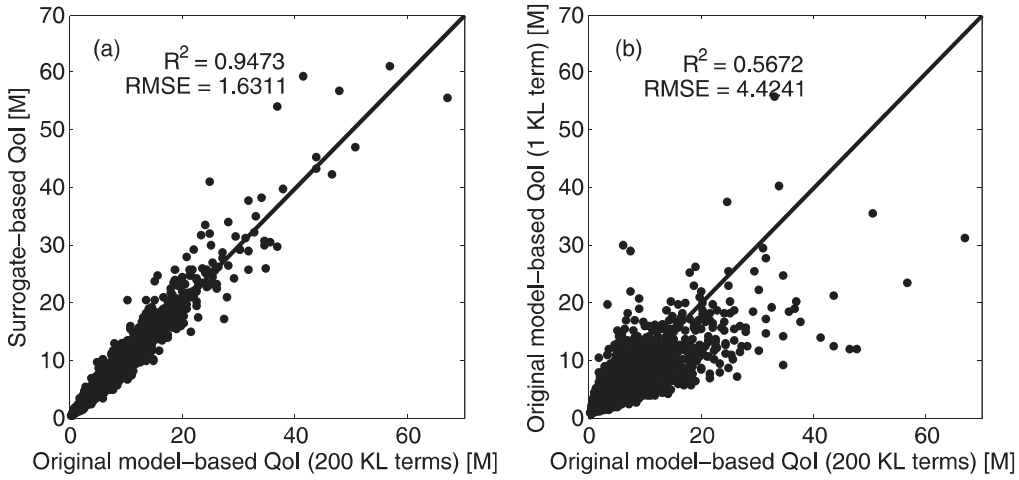

Figure 6. Pairwise comparison of the QoIs (a) between the original model (200 KL terms) and the surrogate, (b) between the original model (200 KL terms) and the original model (1 KL term). The 500 samples are randomly drawn from the prior distribution.

simulation. To provide benchmark results, the original model-based MC simulations are implemented with 1,000,000 samples. To rule out the uncertainty caused by the sampling process, all the two-stage MC simulations implemented in this study use the same batch of 1,000,000 parameter samples as in the original model-based MC simulations. By doing so, the two-stage MC simulations should obtain exactly the same failure probability estimations as the original model-based approach if γ is properly chosen.

To make a reliable estimate of failure probability, we need to obtain enough samples in the failure region. To verify that 1,000,000 samples for direct MC simulation are enough, we plot the decay of the relative error defined in equation (6) with respect to the total number of MC samples (here $T_P=60[M]$). As shown in Figure 7, a reasonable decay is observed. When the total number of MC samples is 1,000,000, the relative error is about 4.24%, which indicates a reliable failure probability estimate.

To implement the two-stage MC simulation with $T_P=60[M]$, the 1,000,000 samples are first evaluated with the surrogate at a negligible computational cost. Then according to the largest absolute error of the surrogate at the 1000 samples used in the SIR (Figure 5b), we choose the initial value of γ as $20.8[M]$, and select the samples within the range of $[T_P-\gamma, T_P+\gamma]$. Then these samples (3425 in total) are re-evaluated with the original model. According to the largest absolute error of the surrogate at these 3425 samples, the value of γ is updated, and some extra samples are re-evaluated with the original model. This process is repeated until the stop criterion described in section 2.4 is satisfied. At last, $\gamma=27.1[M]$ is obtained, which is larger than or equal to all the absolute errors of the surrogate at the 6926 re-evaluated samples in the second stage, as shown in Figure 8. The failure probability for $T_P=60[M]$ is estimated as $P_f^*=5.56\times10^{-4}$ with the two-stage MC simulation, which is equal to the benchmark result of 5.56×10^{-4}. Here the two-stage method requires the model evaluations at only a small fraction of the points sampled, but achieves the same estimation accuracy as if model evaluations were conducted at all points. Thus, the relative error in the failure probability estimation of our approach is the same as the MC estimation based on all sample points, which is well understood and easily computable.

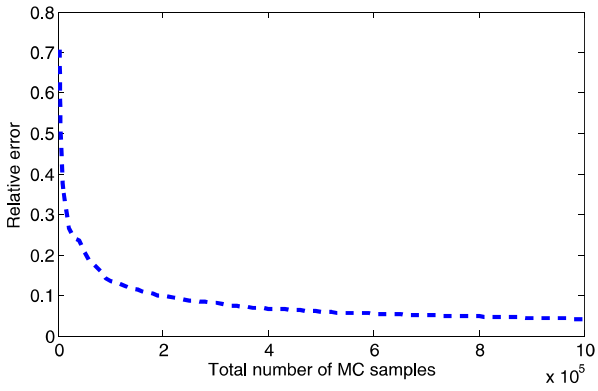

Figure 7. The relationship between the total number of MC samples and relative error of failure probability estimate. Here $T_P=60[M]$.

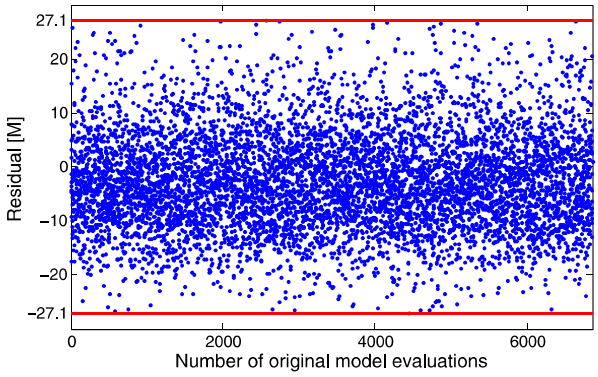

Figure 8. The residuals of QoIs between the original model and the surrogate at the samples around the failure boundary with $T_P=60[M]$. The red lines represent the final value of $\gamma=27.1[M]$.

In this case study, other values of threshold T_P for failure probability analysis are also tested. The values of γ, the total numbers of original model evaluations ($N_O^{Total}=N_O^{SIR}+N_O^{PCE}+N_O^{SS}$, where N_O^{SS} is the number of samples that are re-evaluated in the second stage with the original model) in the two-stage MC simulations, the failure probability estimations from the two-stage MC simulations (P_f^*) and the benchmark results (P_f^{MC}) are listed in Table 1. It shows that the two-stage MC simulations can obtain exactly the same failure probability estimations as the benchmark results with only about one percent of computational cost needed by the latter one. Here N = 1,000,000 can guarantee small relative errors for all the different values of threshold T_P listed in Table 1. If we want to obtain a more accurate failure probability estimation, more samples (i.e., a larger N) in the MC simulation should be evaluated.

In the previous simulations, 1000 samples are used in the dimension reduction by the SIR. If more samples are used in the SIR, the low-dimensional surrogate should be more accurate. With different numbers of samples for the dimension reduction by the SIR (i.e., with different N_O^{SIR}), we test the accuracy of the resulting 6th degree PCE surrogates ($N_O^{PCE}=7$) by calculating the RMSE values between the original model and surrogate-based QoIs at 500 random parameter realizations. As shown in Table 2, a larger N_O^{SIR} generally results in a smaller RMSE value. With these surrogates, we then implement the two-stage MC simulations for failure probability analysis with $T_P=60[M]$, respectively. The values of γ, the numbers of original model evaluations in the second stage simulations (N_O^{SS}) and the total numbers of original model evaluations in the two-stage MC simulations ($N_O^{Total}=N_O^{SIR}+N_O^{PCE}+N_O^{SS}$) are also listed in Table 2. Here all the failure probability estimations are equal to the benchmark result of 5.56×10^{-4}. Considering the total number of original model evaluations, N_O^{Total}, it is not wise to construct a very accurate surrogate. We can also conclude that the two-stage MC simulation for small failure probability analysis does not require a very accurate surrogate.

In this study, both the benchmark simulations and the two-stage simulations adopt the direct MC method. When the failure probability is small, the benchmark results require a very huge number of original model evaluations, which is unacceptable in practice and only used to show the accuracy of the two-stage method. While the computational cost needed by the two-stage MC method is not high, as shown above. However, when the failure probability is relatively large, there will be a considerably large number of samples land around the failure boundary, in which case the two-stage MC method may not be very efficient (but it is still more efficient than the benchmark simulations). In this situation, the direct MC method is usually acceptable, as it does not require a large number of samples. On the other hand, when the event is very rare (e.g., $P_f < 10^{-7}$), there will also be a huge number of samples that need to be re-evaluated with the original model in the second stage, as the error of surrogate evaluation around the failure boundary is rather large. Although to our best knowledge, a failure probability smaller than 10^{-7} is not commonly concerned in groundwater contaminant transport modeling, we have to admit that in this situation the two-stage method is not very efficient for high-dimensional problems. If the input dimension is low, one can integrate importance sampling into the two-stage MC method to improve its efficiency. A similar practice has been implemented in *Li et al.* [2011], where the authors employed cross-entropy to determine the proposal distribution in importance sampling, and constructed a surrogate-based method for rare failure probability computation. Because both

Table 1. Given Different Threshold Settings ($T_P[M]$), the Obtained Values of $\gamma[M]$, the Total Numbers of Original Model Evaluations in the Two-Stage MC Simulations (N_O^{Total}), the Failure Probability Estimations From the Two-Stage MC Simulations (P_f^*), and the Original Model-Based Simulations (P_f^{MC})

$T_P[M]$	60	70	80	90
$\gamma[M]$	27.1	37.9	47.9	47.9
N_O^{Total}	7933	8694	8703	3624
$P_f^*\times10^4$	5.56	2.74	1.24	0.670
$P_f^{MC}\times10^4$	5.56	2.74	1.24	0.670

Table 2. Given the Threshold as $T_P=60\,[M]$, With Different Numbers of Samples for the Dimension Reduction by the SIR (N_O^{SIR}), the RMSE Values of the Resulting Surrogates, the Values of $\gamma\,[M]$, the Numbers of Original Model Evaluations in the Second-Stage MC Simulations (N_O^{SS}), and the Total Numbers of Original Model Evaluations in the Two-Stage MC Simulations (N_O^{Total})[a]

N_O^{SIR}	500	1000	2000	5000
$RMSE\,[M]$	2.314	1.631	1.524	1.383
$\gamma\,[M]$	32.8	27.1	27.2	27.0
N_O^{SS}	13,573	6926	8105	8197
N_O^{Total}	14,080	7933	10,112	13,204

[a]Here all the failure probability estimations are equal to the benchmark result of 5.56×10^{-4}.

surrogate construction and the implementation of importance sampling are extremely difficult in high-dimensional problems, here, we do not extend the two-stage MC method to importance sampling. In fact, the first stage of the two-stage method is similar to the procedure of constructing a proposal distribution in importance sampling. In this stage, we aim to generate sample points that are close to the boundary of the failure region according to surrogate simulations. Then we only re-evaluate these samples with the original model in the second stage to provide an accurate estimate of the failure probability. In this way, we can save a lot of computational cost.

In this study, the two-stage MC simulation is used for efficient and accurate small failure probability analysis. There is another approach that is widely used, i.e., refining the surrogate around the failure boundary through a design of computer experiments, which has also been tested in this study. However, implementing the dimension reduction by the SIR around the failure boundary is not effective. For example, if we set the threshold as $T_P=50\,[M]$ and implement the SIR with 600 samples whose QoI values are within the range of $[40, 60]\,[M]$ (actually these samples are selected from a large number of samples whose QoI values are evaluated with the original model), the resulting single variable η is not highly predictable (Figure 9), which cannot guarantee an accurate surrogate refinement around the failure boundary. If more variables (e.g., 10) are left after the dimension reduction by the SIR, we still cannot build an accurate enough surrogate in the local region near the failure boundary (results not shown here), not to mention the considerable number of original model evaluations for the 10-dimensional surrogate construction. We think it is because in the region with higher QoI values, the relationship between the input parameters and the QoI is more nonlinear, which makes the dimension reduction by the SIR not so effective. Here we should point out that although the SIR is a powerful dimension reduction method, it does not always perform well in strongly nonlinear problems as it is based on the first moment. The SIR performs well only when there exist prominent linear subspaces. For strongly nonlinear problems, one can adopt kernel tricks in the SIR [*Yeh et al.*, 2009]. However, how to properly choose the kernel function for different problems is still an active research topic.

In this study, the two-stage MC approach is used in the small failure probability analysis, a similar approach can also be adopted in MCMC simulations for efficient parameter estimation, i.e., one first uses the surrogate model to filter the unacceptable proposed samples to avoid unnecessary computational cost and then re-evaluates the accepted samples with the original model according to the Metropolis-Hastings algorithm [*Efendiev et al.*, 2005; *Laloy et al.*, 2013; *Ma et al.*, 2008]. Moreover, it is commonly recognized that MCMC simulations struggle in high-dimensional parameter estimation problems. Informed by the QoI (i.e., the log-transformed likelihood function), we can run MCMC on only a few active variables identified by the SIR, which may greatly accelerate MCMC simulations for high-dimensional Bayesian inverse problems. A detailed study of the SIR-based two-stage MCMC will be addressed in our future work.

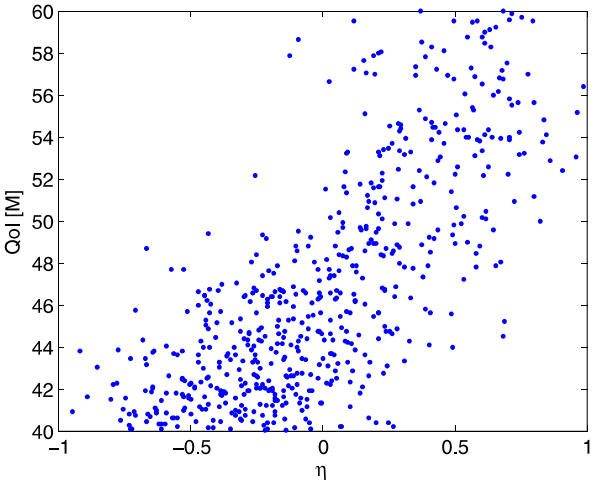

Figure 9. The result of dimension reduction by the SIR within the region where QoI$\in [40, 60]\,[M]$.

4. Conclusions

In groundwater contaminant transport modeling, direct MC simulation for small failure probability analysis usually needs a large number of model evaluations. Under this condition, it is desirable to replace the original model with a surrogate in the MC simulation. However, constructing an accurate surrogate for a model with high-dimensional inputs is challenging. Moreover, approximation errors of the surrogate are inevitable. In this paper, for accurate and efficient small failure probability analysis, a two-stage MC simulation is proposed. In the first stage, we combine the KL expansion and the SIR to substantially reduce the input dimension and construct a PCE-based surrogate over the low-dimensional subspace of the input parameters. Then the surrogate-based MC simulation is efficiently implemented with a large number of samples. In the second stage, the samples around the failure boundary are re-evaluated with the original model. In this way, the bias introduced by the direct surrogate-based MC simulation can be corrected. Illustrated with a high-dimensional groundwater solute transport problem, without sacrificing the estimation accuracy, the proposed method achieves about 100 times of speed-up compared with the traditional MC approach.

Acknowledgments

Computer codes and data used are available upon request to the corresponding author. The authors would like to thank the three anonymous reviewers and the associate editor for their constructive comments on an earlier version, which led to a much improved paper. This work is supported by the National Natural Science Foundation of China (grants 41371237 and 41571215). Jiangjiang Zhang and Weixuan Li would like to acknowledge the support by the Laboratory Directed Research and Development program at Pacific Northwest National Laboratory through the Control of Complex Systems Initiative. Guang Lin would like to acknowledge the support of NSF grant DMS-1555072 and the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Applied Mathematics program as part of the Multifaceted Mathematics for Complex Energy Systems (M^2ACS) project and part of the Collaboratory on Mathematics for Mesoscopic Modeling of Materials project.

References

Anderson, M. P., W. W. Woessner, and R. J. Hunt (2015), *Applied Groundwater Modeling: Simulation of Flow and Advective Transport*, Academic, San Diego, Calif.

Asher, M., B. Croke, A. Jakeman, and L. Peeters (2015), A review of surrogate models and their application to groundwater modeling, *Water Resour. Res.*, *51*, 5957–5973, doi:10.1002/2015WR016967.

Au, S., and J. L. Beck (1999), A new adaptive importance sampling scheme for reliability calculations, *Struct. Saf.*, *21*(2), 135–158, doi:10.1016/S0167-4730(99)00014-4.

Au, S.-K., and J. L. Beck (2001), Estimation of small failure probabilities in high dimensions by subset simulation, *Probab. Eng. Mech.*, *16*(4), 263–277, doi:10.1016/S0266-8920(01)00019-4.

Balesdent, M., J. Morio, and J. Marzat (2013), Kriging-based adaptive importance sampling algorithms for rare event estimation, *Struct. Saf.*, *44*, 1–10, doi:10.1016/j.strusafe.2013.04.001.

Barros, F., D. Fernàndez-Garcia, D. Bolster, and X. Sanchez-Vila (2013), A risk-based probabilistic framework to estimate the endpoint of remediation: Concentration rebound by rate-limited mass transfer, *Water Resour. Res.*, *49*, 1929–1942, doi:10.1002/wrcr.20171.

Batchelor, B., J. Valdés, and V. Araganth (1998), Stochastic risk assessment of sites contaminated by hazardous wastes, *J. Environ. Eng.*, *124*(4), 380–388, doi:10.1061/(ASCE)0733-9372(1998)124:4(380).

Bear, J., and A. H.-D. Cheng (2010), *Modeling Groundwater Flow and Contaminant Transport*, Springer, Dordrecht, Netherlands.

Bolster, D., and D. Tartakovsky (2008), Probabilistic risk analysis of building contamination, *Indoor Air*, *18*(5), 351–364, doi:10.1111/j.1600-0668.2008.00536.x.

Bolster, D., M. Barahona, M. Dentz, D. Fernandez-Garcia, X. Sanchez-Vila, P. Trinchero, C. Valhondo, and D. Tartakovsky (2009), Probabilistic risk analysis of groundwater remediation strategies, *Water Resour. Res.*, *45*, W06413, doi:10.1029/2008WR007551.

Bucklew, J. (2004), *Introduction to Rare Event Simulation*, Springer, New York.

Bura, E., and R. D. Cook (2001), Extending sliced inverse regression: The weighted chi-squared test, *J. Am. Stat. Assoc.*, *96*(455), 996–1003, doi:10.1198/016214501753208979.

Bures, C. J. (2010), Dimension reduction: A guided tour, *Found. Trends Mach. Learn.*, *2*(4), 275–365, doi:10.1561/2200000002.

Chen, C.-W., C.-C. Wei, H.-J. Liu, and N.-S. Hsu (2014), Application of neural networks and optimization model in conjunctive use of surface water and groundwater, *Water Resour. Manage.*, *28*(10), 2813–2832, doi:10.1007/s11269-014-0639-6.

Ciriello, V., V. Di Federico, M. Riva, F. Cadini, J. De 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, *Stochastic Environ. Res. Risk Assess.*, *27*(4), 945–954, doi:10.1007/s00477-012-0616-7.

Dai, C., L. Xue, D. Zhang, and A. Guadagnini (2016), Data-worth analysis through probabilistic collocation-based Ensemble Kalman Filter, *J. Hydrol.*, *540*, 488–503, doi:10.1016/j.jhydrol.2016.06.037.

de Barros, F. P., and W. Nowak (2010), On the link between contaminant source release conditions and plume prediction uncertainty, *J. Contam. Hydrol.*, *116*(1), 24–34, doi:10.1016/j.jconhyd.2010.05.004.

Der Kiureghian, A., and P.-L. Liu (1986), Structural reliability under incomplete probability information, *J. Eng. Mech.*, *112*(1), 85–104, doi:10.1061/(ASCE)0733-9399(1986)112:1(85).

Dubourg, V., and B. Sudret (2014), Meta-model-based importance sampling for reliability sensitivity analysis, *Struct. Saf.*, *49*, 27–36, doi:10.1016/j.strusafe.2013.08.010.

Dubourg, V., B. Sudret, and J.-M. Bourinet (2011), Reliability-based design optimization using kriging surrogates and subset simulation, *Struct. Multidiscip. Optim.*, *44*(5), 673–690, doi:10.1007/s00158-011-0653-8.

Dubourg, V., B. Sudret, and F. Deheeger (2013), Metamodel-based importance sampling for structural reliability analysis, *Probab. Eng. Mech.*, *33*, 47–57, doi:10.1016/j.probengmech.2013.02.002.

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.

Elsheikh, A., S. Oladyshkin, W. Nowak, and M. Christie (2014), Estimating the probability of CO_2 leakage using rare event simulation, paper presented at ECMOR XIV-14th European Conference on the Mathematics of Oil Recovery, EAGE (European Association of Geoscientists & Engineers), Netherlands, doi:10.3997/2214-4609.20141876.

Enzenhoefer, R., P. J. Binning, and W. Nowak (2015), Stakeholder-objective risk model (storm): Determining the aggregated risk of multiple contaminant hazards in groundwater well catchments, *Adv. Water Resour.*, *83*, 160–175, doi:10.1016/j.advwatres.2015.05.015.

Felisa, G., V. Ciriello, M. Antonellini, V. Di Federico, and D. Tartakovsky (2015), Data-driven models of groundwater salinization in coastal plains, *J. Hydrol.*, *531*, 187–197, doi:10.1016/j.jhydrol.2015.07.045.

Fernàndez-Garcia, D., D. Bolster, X. Sanchez-Vila, and D. M. Tartakovsky (2012), A Bayesian approach to integrate temporal data into probabilistic risk analysis of monitored NAPL remediation, *Adv. Water Resour.*, *36*, 108–120, doi:10.1016/j.advwatres.2011.07.001.

Ghanem, R., and P. Spanos (1991), *Stochastic Finite Elements: A Spectral Approach*, Springer, New York.

Hamed, M. M., and P. B. Bedient (1997), On the performance of computational methods for the assessment of risk from ground-water contamination, *Ground Water*, *35*(4), 638–646, doi:10.1111/j.1745-6584.1997.tb00129.x.

Harbaugh, A. W., E. R. Banta, M. C. Hill, and M. G. McDonald (2000), *MODFLOW-2000, the US Geological Survey Modular Ground-Water Model: User Guide to Modularization Concepts and the Ground-Water Flow Process*, U.S. Geol. Surv., Reston, Va.

Jolliffe, I. (1986), *Principal Component Analysis*, Springer, New York.

Kanj, R., R. Joshi, and S. Nassif (2006), Mixture importance sampling and its application to the analysis of SRAM designs in the presence of rare failure events, paper presented at Proceedings Design Automation Conference, ACM, San Francisco, Calif., doi:10.1145/1146909.1146930.

Klimke, A., and B. Wohlmuth (2005), Algorithm 847: Spinterp: Piecewise multilinear hierarchical sparse grid interpolation in MATLAB, *ACM Trans. Math. Software*, *31*(4), 561–579, doi:10.1145/1114268.1114275.

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.

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, J., and D. Xiu (2010), Evaluation of failure probability via surrogate models, *J. Comput. Phys.*, *229*(23), 8966–8980, doi:10.1016/j.jcp.2010.08.022.

Li, J., and D. Xiu (2014), Surrogate based method for evaluation of failure probability under multiple constraints, *SIAM J. Sci. Comput.*, *36*(2), A828–A845, doi:10.1137/120877192.

Li, J., J. Li, and D. Xiu (2011), An efficient surrogate-based method for computing rare failure probability, *J. Comput. Phys.*, *230*(24), 8683–8697, doi:10.1016/j.jcp.2011.08.008.

Li, K.-C. (1991), Sliced inverse regression for dimension reduction, *J. Am. Stat. Assoc.*, *86*(414), 316–327, doi:10.1080/01621459.1991.10475035.

Li, L., and X. Yin (2008), Sliced inverse regression with regularizations, *Biometrics*, *64*(1), 124–131, doi:10.1111/j.1541-0420.2007.00836.x.

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.

Li, W., G. Lin, and B. Li (2016), Inverse regression-based uncertainty quantification algorithms for high-dimensional models: Theory and practice, *J. Comput. Phys.*, *321*, 259–278, doi:10.1016/j.jcp.2016.05.040.

Liao, Q., and D. Zhang (2013), Probabilistic collocation method for strongly nonlinear problems: 1. Transform by location, *Water Resour. Res.*, *49*, 7911–7928, doi:10.1002/2013WR014055.

Liao, Q., and D. Zhang (2014), Probabilistic collocation method for strongly nonlinear problems: 2. Transform by displacement, *Water Resour. Res.*, *50*, 8736–8759, doi:10.1002/2014WR016238.

Liao, Q., and D. Zhang (2016), Probabilistic collocation method for strongly nonlinear problems: 3. Transform by time, *Water Resour. Res.*, *52*, 2366–2375, doi:10.1002/2015WR017724.

Ma, X., M. Al-Harbi, A. Datta-Gupta, and Y. Efendiev (2008), An efficient two-stage sampling method for uncertainty quantification in history matching geological models, *SPE J.*, *13*(1), 77–87, doi:10.2118/102476-PA.

Marrel, A., B. Iooss, F. Van Dorpe, and E. Volkova (2008), An efficient methodology for modeling complex computer codes with Gaussian processes, *Comput. Stat. Data Anal.*, *52*(10), 4731–4744, doi:10.1016/j.csda.2008.03.026.

Marzouk, Y. M., and H. N. Najm (2009), Dimensionality reduction and polynomial chaos acceleration of Bayesian inference in inverse problems, *J. Comput. Phys.*, *228*(6), 1862–1902, doi:10.1016/j.jcp.2008.11.024.

Massmann, J., and R. A. Freeze (1987a), Groundwater contamination from waste management sites: The interaction between risk-based engineering design and regulatory policy: 1. Methodology, *Water Resour. Res.*, *23*(2), 351–367, doi:10.1029/WR023i002p00351.

Massmann, J., and R. A. Freeze (1987b), Groundwater contamination from waste management sites: The interaction between risk-based engineering design and regulatory policy: 2. Results, *Water Resour. Res.*, *23*(2), 368–380, doi:10.1029/WR023i002p00368.

Paffrath, M., and U. Wever (2007), Adapted polynomial chaos expansion for failure detection, *J. Comput. Phys.*, *226*(1), 263–281, doi:10.1016/j.jcp.2007.04.011.

Pulch, R. (2010), Polynomial chaos for the computation of failure probabilities in periodic problems, in *Scientific Computing in Electrical Engineering SCEE 2008*, edited by J. Roos and L. Costa, pp. 191–198, Springer-Verlag Berlin Heidelberg.

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.

Rodak, C., and S. Silliman (2012), Probabilistic risk analysis and fault trees: Initial discussion of application to identification of risk at a wellhead, *Adv. Water Resour.*, *36*, 133–145, doi:10.1016/j.advwatres.2011.02.005.

Roweis, S. T., and L. K. Saul (2000), Nonlinear dimensionality reduction by locally linear embedding, *Science*, *290*(5500), 2323–2326, doi:10.1126/science.290.5500.2323.

Schölkopf, B., A. Smola, and K.-R. Müller (1997), Kernel principal component analysis, in *Artificial Neural Networks—ICANN'97*, pp. 583–588, Springer-Verlag Berlin Heidelberg.

Schöniger, A., W. Nowak, and H. J. Hendricks Franssen (2012), Parameter estimation by ensemble Kalman filters with transformed data: Approach and application to hydraulic tomography, *Water Resour. Res.*, *48*, W04502, doi:10.1029/2011WR010462.

Shi, L., J. Yang, D. Zhang, and H. Li (2009), Probabilistic collocation method for unconfined flow in heterogeneous media, *J. Hydrol.*, *365*(1), 4–10, doi:10.1016/j.jhydrol.2008.11.012.

Smith, R. C. (2013), *Uncertainty Quantification: Theory, Implementation, and Applications*, SIAM, Philadelphia, Pa.

Tartakovsky, D. M. (2013), Assessment and management of risk in subsurface hydrology: A review and perspective, *Adv. Water Resour.*, *51*, 247–260, doi:10.1016/j.advwatres.2012.04.007.

Troldborg, M., W. Nowak, I. V. Lange, M. C. Santos, P. J. Binning, and P. L. Bjerg (2012), Application of Bayesian geostatistics for evaluation of mass discharge uncertainty at contaminated sites, *Water Resour. Res.*, *48*, W09535, doi:10.1029/2011WR011785.

Vermeulen, P., A. Heemink, and C. Te Stroet (2004), Reduced models for linear groundwater flow models using empirical orthogonal functions, *Adv. Water Resour.*, *27*(1), 57–69, doi:10.1016/j.advwatres.2003.09.008.

Wagner, B. J. (1992), Simultaneous parameter estimation and contaminant source characterization for coupled groundwater flow and contaminant transport modelling, *J. Hydrol.*, *135*(1), 275–303, doi:10.1016/0022-1694(92)90092-A.

Wang, H. Q., G. Lin, and J. L. Li (2016), Gaussian process surrogates for failure detection: A Bayesian experimental design approach, *J. Comput. Phys.*, *313*, 247–259, doi:10.1016/j.jcp.2016.02.053.

Wang, T. A., and W. F. McTernan (2002), The development and application of a multilevel decision analysis model for the remediation of contaminated groundwater under uncertainty, *J. Environ. Manage.*, *64*(3), 221–235, doi:10.1006/jema.2001.0470.

Weisberg, S. (2002), Dimension reduction regression in R, *J. Stat. Software*, *7*(1), 1–22.

Xiu, D. (2010), *Numerical Methods for Stochastic Computations: A Spectral Method Approach*, Princeton Univ. Press, Princeton, N. J.

Xiu, D., and G. E. Karniadakis (2002), The Wiener-Askey polynomial chaos for stochastic differential equations, *SIAM J. Sci. Comput.*, *24*(2), 614–644, doi:10.1137/S1064827501387826.

Yan, S., and B. Minsker (2006), Optimal groundwater remediation design using an adaptive neural network genetic algorithm, *Water Resour. Res.*, *42*, W05407, doi:10.1029/2005WR004303.

Yang, J., D. Zhang, and Z. Lu (2004), Stochastic analysis of saturated–unsaturated flow in heterogeneous media by combining Karhunen-Loeve expansion and perturbation method, *J. Hydrol.*, *294*(1), 18–38, doi:10.1016/j.jhydrol.2003.10.023.

Yeh, H.-D., T.-H. Chang, and Y.-C. Lin (2007), Groundwater contaminant source identification by a hybrid heuristic approach, *Water Resour. Res.*, *43*, W09420, doi:10.1029/2005WR004731.

Yeh, Y.-R., S.-Y. Huang, and Y.-J. Lee (2009), Nonlinear dimension reduction with kernel sliced inverse regression, *IEEE Trans. Knowl. Data Eng.*, *21*(11), 1590–1603, doi:10.1109/TKDE.2008.232.

Yuhua, D., and Y. Datao (2005), Estimation of failure probability of oil and gas transmission pipelines by fuzzy fault tree analysis, *J. Loss Prev. Process Ind.*, *18*(2), 83–88, doi:10.1016/j.jlp.2004.12.003.

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.

Zeng, X., M. Ye, J. Burkardt, J. Wu, D. Wang, and X. Zhu (2016), Evaluating two sparse grid surrogates and two adaptation criteria for groundwater Bayesian uncertainty quantification, *J. Hydrol.*, *535*, 120–134, doi:10.1016/j.jhydrol.2016.01.058.

Zhang, D. (2002), *Stochastic Methods for Flow in Porous Media: Coping with Uncertainties*, Academic, San Diego, Calif.

Zhang, D., and Z. Lu (2004), An efficient, high-order perturbation approach for flow in random porous media via Karhunen-Loeve and polynomial expansions, *J. Comput. Phys.*, *194*(2), 773–794, doi:10.1016/j.jcp.2003.09.015.

Zhang, G., D. Lu, M. Ye, M. Gunzburger, and C. Webster (2013), An adaptive sparse-grid high-order stochastic collocation method for Bayesian inference in groundwater reactive transport modeling, *Water Resour. Res.*, *49*, 6871–6892, doi:10.1002/wrcr.20467.

Zhang, J., L. Zeng, C. Chen, D. Chen, and L. Wu (2015), Efficient Bayesian experimental design for contaminant source identification, *Water Resour. Res.*, *51*, 576–598, doi:10.1002/2014WR015740.

Zhang, J., W. Li, L. Zeng, and L. Wu (2016), An adaptive Gaussian process-based method for efficient Bayesian experimental design in groundwater contaminant source identification problems, *Water Resour. Res.*, *52*, 5971–5984, doi:10.1002/2016WR018598.

Zhao, Y., W. Lu, and C. Xiao (2016), A kriging surrogate model coupled in simulation-optimization approach for identifying release history of groundwater sources, *J. Contam. Hydrol.*, *185–186*, 51–60, doi:10.1016/j.jconhyd.2016.01.004.

Zheng, C., and P. P. Wang (1999), MT3DMS: A modular three-dimensional multispecies transport model for simulation of advection, dispersion, and chemical reactions of contaminants in groundwater systems, Documentation and user's guide Rep., DTIC Document, U.S. Army Engineer Research and Development Center and Vicksburg, Miss. [Available at http://www.geology.wisc.edu/courses/g727/mt3dmanual.pdf.]