AGU PUBLICATIONS

Water Resources Research

RESEARCH ARTICLE
10.1002/2016WR019756

A geostatistics-informed hierarchical sensitivity analysis method for complex groundwater flow and transport modeling

Heng Dai[1], Xingyuan Chen[1], Ming Ye[2], Xuehang Song[1], and John M. Zachara[1]

[1]Pacific Northwest National Laboratory, Richland, Washington, USA, [2]Department of Scientific Computing, Florida State University, Tallahassee, Florida, USA

Key Points:
- A new sensitivity analysis method was developed based on a hierarchical uncertainty quantification framework
- Geostatistical tools were used to cope with spatial correlations among model parameters
- The new sensitivity analysis method was tested on a real flow and transport problem in a highly heterogeneous aquifer

Correspondence to:
X. Chen,
Xingyuan.Chen@pnnl.gov

Citation:
Dai, H., X. Chen, M. Ye, X. Song, and J. M. Zachara (2017), A geostatistics-informed hierarchical sensitivity analysis method for complex groundwater flow and transport modeling, *Water Resour. Res., 53*, doi:10.1002/2016WR019756.

Received 6 SEP 2016
Accepted 23 APR 2017
Accepted article online 29 APR 2017

Abstract Sensitivity analysis is an important tool for development and improvement of mathematical models, especially for complex systems with a high dimension of spatially correlated parameters. Variance-based global sensitivity analysis has gained popularity because it can quantify the relative contribution of uncertainty from different sources. However, its computational cost increases dramatically with the complexity of the considered model and the dimension of model parameters. In this study, we developed a new sensitivity analysis method that integrates the concept of variance-based method with a hierarchical uncertainty quantification framework. Different uncertain inputs are grouped and organized into a multilayer framework based on their characteristics and dependency relationships to reduce the dimensionality of the sensitivity analysis. A set of new sensitivity indices are defined for the grouped inputs using the variance decomposition method. Using this methodology, we identified the most important uncertainty source for a dynamic groundwater flow and solute transport model at the Department of Energy (DOE) Hanford site. The results indicate that boundary conditions and permeability field contribute the most uncertainty to the simulated head field and tracer plume, respectively. The relative contribution from each source varied spatially and temporally. By using a geostatistical approach to reduce the number of realizations needed for the sensitivity analysis, the computational cost of implementing the developed method was reduced to a practically manageable level. The developed sensitivity analysis method is generally applicable to a wide range of hydrologic and environmental problems that deal with high-dimensional spatially distributed input variables.

1. Introduction

Numerical modeling has been a vital tool to simulate groundwater flow and solute transport for managing contaminated aquifers and assessing associated environmental risks [*Davis et al.*, 2004; *Liu and Gupta*, 2007; *Volkova et al.*, 2008; *Matott et al.*, 2009; *Tartakovsky*, 2013; *Dai et al.*, 2014]. Uncertainty in numerical modeling arises from various sources, including lack of data to characterize the system of interest, simplification in conceptual models, and unknown nature of system driving forces [*Neuman*, 2003; *Refsgaard et al.*, 2007; *Rubin et al.*, 2010; *Tartakovsky*, 2013]. Sensitivity analysis is needed to identify the most important source of uncertainty so that one can use limited resources to achieve the maximum reduction of uncertainty in model predictions [*Saltelli et al.*, 2010; *Wainwright et al.*, 2014; *Razavi and Gupta*, 2015, 2016a, 2016b; *Song et al.*, 2015].

Sensitivity analysis methods can be broadly divided into two groups—local and global. Local sensitivity analysis evaluates parameter sensitivity within a narrow range of parameter values [*Piggott et al.*, 1996; *Tebes-Stevens et al.*, 2001; *Dai and Samper*, 2004; *Greskowiak et al.*, 2010; *Gedeon et al.*, 2012; *Abdelaziz and Merkel*, 2015], whereas global sensitivity analysis methods provide an overall sensitivity measure of parameters over their entire ranges [*Saltelli and Sobol'*, 1995; *Saltelli*, 2000, *Saltelli et al.*, 2010; *van Griensven et al.*, 2006; *Mishra et al.*, 2009; *Nossent et al.*, 2011; *Pan et al.*, 2011; *Herman, et al.*, 2013; *Shi et al.*, 2014], which has made global sensitivity analysis more appealing in recent modeling practices [*Baroni and Tarantola*, 2014; *Song et al.*, 2015]. Two global sensitivity analysis methods that are most commonly used in the hydrological sciences are the variance-based method [*Sobol'*, 1993; *Saltelli et al.*, 1999; *van Werkhoven et al.*, 2008; *Wagener et al.*, 2009; *Yang*, 2011; *Fajraoui, et al.*, 2012; *Massmann and Holzmann*, 2012; *Ciriello et al.*, 2013; *Zhang et al.*, 2013] and the screening method such as the Morris method [*Morris*, 1991; *van Griensven et al.*, 2006; *Campolongo et al.*, 2007; *Matott*, 2012; *Tong*, 2011; *Song et al.*, 2012]. The screening method can only provide a qualitative importance ranking of parameters or uncertainty sources based on their summary statistics

(e.g., mean and standard deviation) of local sensitivity (e.g., elementary effect) over parameter space. The variance-based method, on the other hand, can provide a quantitative importance ranking of parameters or uncertainty sources using global sensitivity indices.

This study aims to develop a sensitivity analysis method that is suitable for identifying the most significant source of uncertainty from multiple factors that contribute to the state variables simulated by complex multiprocess models such as groundwater flow and solute transport models. While variance-based global sensitivity analysis is considered to be the most suitable approach for this purpose, there are three major challenges associated with its implementation for complex models. First, the existing methods do not explicitly address dependence or hierarchy among the various uncertainty sources. For example, there may be several alternative conceptual models for one system, and each model may have uncertain boundary conditions and parameters. Therefore, it is necessary to first understand the relations between the multiple uncertain factors, and then to accommodate their relations into a suitable framework. Second, variance-based global sensitivity analysis is computationally expensive due to a large number of model simulations required to evaluate the mean and variance of state variables. The computational challenge may limit practical implementation when the runtime of the forward model is relatively long. Third, each uncertainty source may be associated with high-dimensional spatially distributed variables, such as permeability fields of an aquifer or recharge rates to groundwater basins. Methods to explicitly and appropriately deal with spatial correlations among spatially distributed parameters or forcing are limited for sensitivity analysis [*Lilburne and Tarantola*, 2009].

To address the challenges discussed above, this study presents a new sensitivity analysis framework that integrates the concepts of variance-based sensitivity analysis and hierarchical uncertainty quantification with geostatistical tools. The hierarchical uncertainty quantification framework [*Draper et al*., 1999; *Meyer et al*., 2007, 2014; *Rojas et al*., 2010] allows us to decompose the total predictive uncertainty into contributions from different sources that follow their hierarchical structure, similar to that of parametric uncertainty, model uncertainty, and scenario uncertainty [*Dai and Ye*, 2015]. New sensitivity indices are defined in our approach to measure the relative importance of each uncertain source. Geostatistical techniques such as kriging and conditional simulation methods [*Rubin*, 2003] are used within the framework not only to group spatially distributed parameters and variables to reduce the dimensionality of the problem but also to explicitly account for spatial correlations in their realizations.

The new sensitivity analysis framework was implemented to evaluate the relative importance of various model inputs that contribute to the overall uncertainty in groundwater flow and transport at the U.S. Department of Energy's Hanford 300 Area Integrated Field Research Challenge (IFRC) site, where uranium migration within the groundwater-surface water interaction zone has been extensively studied to protect public health and the environment [e.g., *Zachara et al*. 2016; *Chen et al*., 2012, 2013]. The numerical model of the subsurface system developed by *Chen et al*. [2012 and 2013] contains uncertain components in flow boundary conditions, geological structures of high-permeability and low-permeability, and a highly heterogeneous permeability field in the flow domain. All three of these uncertain inputs contain high-dimensional spatially distributed parameters, and the high-computational cost of a single model run limits the number of simulations that can be performed for sensitivity analysis. These conditions of this numerical model make it ideal for testing the new framework of sensitivity analysis proposed in this study.

In the rest of the paper, we present in section 2 the developed hierarchical framework considering different uncertainty sources and the new global sensitivity indices defined based on this framework, followed by the algorithm of estimating the indices using Monte Carlo (MC) method. A real-world test case of groundwater transport modeling considering multiple uncertainty sources for the developed sensitivity analysis is described in detail in section 3. In section 4, we first present the model input samples and results of MC simulations for the test case, and then the sensitivity indices for hydraulic head and tracer concentration. Section 5 summarizes the main findings of this research.

2. Methodology

This section first discusses two configurations of the relation between various uncertainty sources, and then defines new sensitivity indices for the uncertainty sources by applying variance decomposition based on the configurations. Numerical evaluation of the sensitivity indices is given at the end of this section.

2.1. Hierarchical Framework of Uncertainty Quantification

The hierarchical framework of uncertainty quantification used in this study was revised from that developed by *Meyer et al*. [2007, 2014]. As shown in Figure 1a, the framework of *Meyer et al*. [2007, 2014] contains three layers of uncertain factors: scenario, model, and parameter. Scenario uncertainty is aleatory, and originates from the natural driving force of the model system, such as future precipitation scenarios predicted by climate models. The scenarios may influence the model formulation and selection as each model may only be plausible for a limited range of conditions. Model uncertainty is caused by different conceptual or mathematical formulations of the targeted process [*Beven*, 2002, 2006; *Bredehoeft*, 2003, 2005; *Neuman*, 2003; *Ye et al*., 2004, 2005, 2008a, 2008b, 2010a, 2010b], given that characterization data and process understanding is typically limited for any given complex system. Each model may have its distinct set of inputs and parameters that are also subject to uncertainty. This hierarchical framework clearly illustrates the concept of parameter uncertainty depending on both model and scenario uncertainty. Accordingly, the total uncertainty in prediction will first be decomposed into contributions from scenarios, followed by those from models and parameters in sequence.

The hierarchical framework shown in Figure 1a is flexible, and we revised it in Figure 1b for studies that do not have the layer of scenario uncertainty while the layer of model uncertainty is decomposed into various components (corresponding to various processes of a multiphysics model). Different models may have a common component, but the component may have different structures and/or values in different models. Taking the groundwater modeling at the Hanford Site [*Chen et al*., 2013] as an example, each flow and solute transport model requires the topography of the contact between two contrasting geological layers within the aquifer, whereas this topography of contact may vary from one model to another. The explicit expression of model components adds the flexibility to address model structure uncertainty. The model uncertainty is referred to as model component uncertainty for this new configuration. The new hierarchical framework of uncertainty quantification in Figure 1b is used in our case study associated with the groundwater modeling at the Hanford Site.

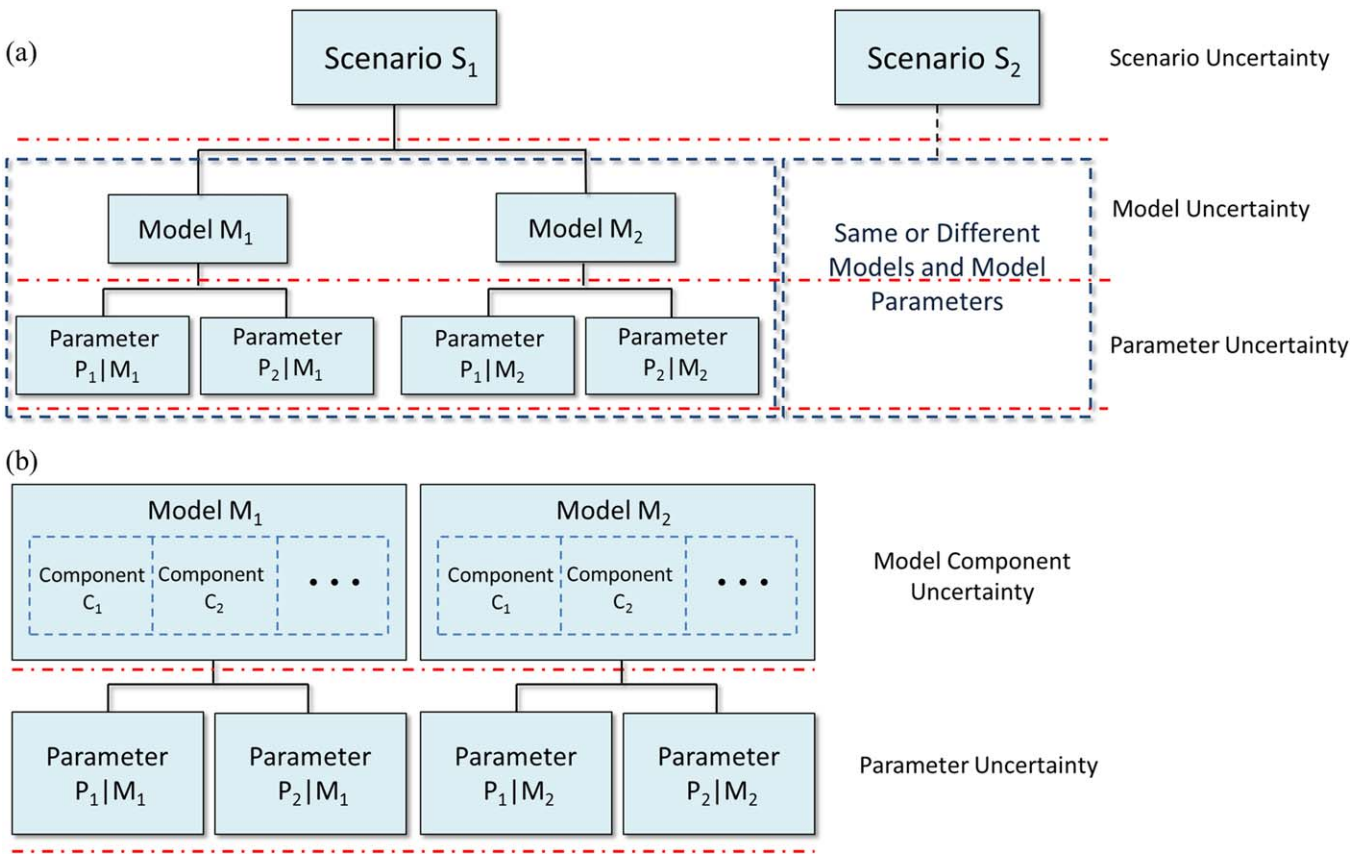

Figure 1. (a) The three-layer hierarchical uncertainty quantification framework, (b) a two-layer uncertainty quantification framework with multiple components included in the model uncertainty.

2.2. Sensitivity Indices for the Hierarchical Sensitivity Analysis Framework

For a model output (Δ) that depends on a set of uncertain model inputs ($\boldsymbol{\theta}=\{\theta_1,\ldots,\theta_k\}$) in the form of $\Delta=f(\boldsymbol{\theta})=f(\theta_1,\ldots,\theta_k)$, its total variance can be decomposed as:

$$V(\Delta)=V_{\theta_i}(E_{\boldsymbol{\theta}_{\sim i}}(\Delta|\theta_i))+E_{\theta_i}(V_{\boldsymbol{\theta}_{\sim i}}(\Delta|\theta_i)), \tag{1}$$

following the principle of variance decomposition. The first term on the right-hand side of equation (1) is the partial variance or the main effect caused by θ_i, with the inner expectation being the mean of the output calculated on all the changing values of $\boldsymbol{\theta}_{\sim i}$, i.e., all model inputs except the fixed subset θ_i. The second term on the right-hand side of equation (1) represents the partial variance contributed by model inputs other than θ_i as well as interactions among all the inputs. The decomposition can be repeated for all model inputs, and consequently the first-order sensitivity index can be defined as $S_i=\frac{V_{\theta_i}(E_{\boldsymbol{\theta}_{\sim i}}(\Delta|\theta_i))}{V(\Delta)}$, representing the percentage of output uncertainty that is contributed by the input subset θ_i. It is thus a quantitative metric for the importance or influence of an individual model input uncertainty source on the model output [*Saltelli et al*., 1999, 1998; *Sobol'*, 1993]. Note that the sum of the sensitivity indices for all model inputs is smaller than or equal to one because it does not include the interaction effects among model inputs.

One aim of this study was to apply the principle of variance decomposition recursively to quantify the contribution of uncertainty from each source in the hierarchical uncertainty framework. For the uncertainty structure shown in Figure 1a, we began by decomposing the total variance at the top layer of scenario uncertainty, i.e.,

$$\begin{aligned} V(\Delta)&=E_{\mathbf{S}}V_{\sim\mathbf{S}|\mathbf{S}}(\Delta|\mathbf{S})+V_{\mathbf{S}}E_{\sim\mathbf{S}|\mathbf{S}}(\Delta|\mathbf{S}) \\ &=E_{\mathbf{S}}V_{\mathbf{M},\boldsymbol{\theta}|\mathbf{S}}(\Delta|\mathbf{S})+V_{\mathbf{S}}E_{\mathbf{M},\boldsymbol{\theta}|\mathbf{S}}(\Delta|\mathbf{S}) \end{aligned}, \tag{2}$$

where $\mathbf{S}$ is the set of multiple scenarios, $\mathbf{M}$ is the set of multiple models, and $\boldsymbol{\theta}=\boldsymbol{\theta}^{(1)}\cup\boldsymbol{\theta}^{(2)}\cdots\cup\boldsymbol{\theta}^{(\mathrm{K})}$ is the parameter set for all the models with $\boldsymbol{\theta}^{(k)}$ being the parameters for a given model M_k, $\sim\mathbf{S}$ represents uncertain inputs excluding $\mathbf{S}$, which are $\mathbf{M},\boldsymbol{\theta}$. The subscript $\mathbf{M},\boldsymbol{\theta}|\mathbf{S}$ refers to the change of model and parameter combinations under certain fixed scenario and implies the hierarchical structure where models and parameters depend on scenarios. The first and the second terms on the right-hand side of equation (2) represent within-scenario and between-scenario variance, respectively [*Draper et al*., 1999]. Between-scenario variance represents the uncertainty induced by multiple alternative scenarios, while the within-scenario variance is due to lower-level uncertainties in models and parameters.

The partial variance within each scenario, $V_{\mathbf{M},\boldsymbol{\theta}|\mathbf{S}}(\Delta|\mathbf{S})$, can be further decomposed as:

$$\begin{aligned} V_{\mathbf{M},\boldsymbol{\theta}|\mathbf{S}}(\Delta|\mathbf{S})=\ & E_{\mathbf{M}|\mathbf{S}}V_{\sim\mathbf{M}|\mathbf{M},\mathbf{S}}(\Delta|\mathbf{M},\mathbf{S})+V_{\mathbf{M}|\mathbf{S}}E_{\sim\mathbf{M}|\mathbf{M},\mathbf{S}}(\Delta|\mathbf{M},\mathbf{S}) \\ &=E_{\mathbf{M}|\mathbf{S}}V_{\boldsymbol{\theta}|\mathbf{M},\mathbf{S}}(\Delta|\mathbf{M},\mathbf{S})+V_{\mathbf{M}|\mathbf{S}}E_{\boldsymbol{\theta}|\mathbf{M},\mathbf{S}}(\Delta|\mathbf{M},\mathbf{S}) \end{aligned}, \tag{3}$$

where the $\sim\mathbf{M}$ represents the uncertain inputs excluding $\mathbf{M}$ in the set of $\mathbf{M},\boldsymbol{\theta}$, which is $\boldsymbol{\theta}$, subscripts $\mathbf{M}|\mathbf{S}$ and $\boldsymbol{\theta}|\mathbf{M},\mathbf{S}$ represent the change of model combinations under certain fixed scenario and change of parameter combinations under certain fixed model and scenario, and they suggest the hierarchical relationships in which models are conditioned on scenarios and parameters are conditioned on models and scenarios. $\mathbf{M}$ contains all models under all scenarios, although each scenario might only associate with a subset of the models. The first and second terms on the right-hand side of the equation are within-model variance and between-model variance, respectively [*Draper*, 1995; *Hoeting et al*., 1999]. The between-model variance is caused by multiple plausible models under a single scenario, and the within-model variance results from parametric uncertainty.

The total variance, after substituting equation (3) into equation (2) and applying the law of total expectation to the first term on the right-hand side of equation (2), is decomposed as:

$$\begin{aligned} V(\Delta)&=E_{\mathbf{s}}\left(E_{\mathbf{M}|\mathbf{S}}V_{\boldsymbol{\theta}|\mathbf{M},\mathbf{S}}(\Delta|\mathbf{M},\mathbf{S})+V_{\mathbf{M}|\mathbf{S}}E_{\boldsymbol{\theta}|\mathbf{M},\mathbf{S}}(\Delta|\mathbf{M},\mathbf{S})\right)+V_{\mathbf{s}}E_{\mathbf{M}|\mathbf{S}}E_{\boldsymbol{\theta}|\mathbf{M},\mathbf{S}}(\Delta|\mathbf{M},\mathbf{S}) \\ &=E_{\mathbf{s}}E_{\mathbf{M}|\mathbf{S}}V_{\boldsymbol{\theta}|\mathbf{M},\mathbf{S}}(\Delta|\mathbf{M},\mathbf{S})+E_{\mathbf{s}}V_{\mathbf{M}|\mathbf{S}}E_{\boldsymbol{\theta}|\mathbf{M},\mathbf{S}}(\Delta|\mathbf{M},\mathbf{S})+V_{\mathbf{s}}E_{\mathbf{M}|\mathbf{S}}E_{\boldsymbol{\theta}|\mathbf{M},\mathbf{S}}(\Delta|\mathbf{M},\mathbf{S}) \\ &=V(\boldsymbol{\theta})+V(\mathbf{M})+V(\mathbf{S}) \end{aligned}, \tag{4}$$

where the total variance is eventually decomposed into variances resulting from parametric uncertainty, model uncertainty, and scenario uncertainty, respectively. A Monte Carlo-based approach to evaluate the

variances contributed by each source is discussed in the next Section. A set of sensitivity indices can then be defined to measure the relative contribution of uncertainty from each source, i.e.,

$$\begin{aligned} S_{\mathbf{S}} &= \frac{V_{\mathbf{S}} E_{\mathbf{M}|\mathbf{S}} E_{\boldsymbol{\theta}|\mathbf{M},\mathbf{S}}(\Delta|\mathbf{M},\mathbf{S})}{V(\Delta)} = \frac{V(\mathbf{S})}{V(\Delta)} \\ S_{\mathbf{M}} &= \frac{E_{\mathbf{S}} V_{\mathbf{M}|\mathbf{S}} E_{\boldsymbol{\theta}|\mathbf{M},\mathbf{S}}(\Delta|\mathbf{M},\mathbf{S})}{V(\Delta)} = \frac{V(\mathbf{M})}{V(\Delta)} \\ S_{\boldsymbol{\theta}} &= \frac{E_{\mathbf{S}} E_{\mathbf{M}|\mathbf{S}} V_{\boldsymbol{\theta}|\mathbf{M},\mathbf{S}}(\Delta|\mathbf{M},\mathbf{S})}{V(\Delta)} = \frac{V(\boldsymbol{\theta})}{V(\Delta)} \end{aligned} \tag{5}$$

The variance decomposition for the uncertainty structure shown in Figure 1b is similar to equation (1) but applying it over the first layer of uncertain inputs of model component $\mathbf{M}_{C_i}$ as:

$$V(\Delta) = E_{\mathbf{M}_{C_i}} V_{\sim\mathbf{M}_{C_i}}(\Delta|\mathbf{M}_{C_i}) + V_{\mathbf{M}_{C_i}} E_{\sim\mathbf{M}_{C_i}}(\Delta|\mathbf{M}_{C_i}), \tag{6}$$

where $\sim\mathbf{M}_{C_i}$ denotes the uncertain inputs except $\mathbf{M}_{C_i}$. Equation (6) assumes that the choice of one model component does not depend on the choice of other components, which is different from the hierarchical dependence between scenario and model in the uncertainty structure of Figure 1a. Further decomposing uncertainty of other sources of $V_{\sim\mathbf{M}_{C_i}}(\Delta|\mathbf{M}_{C_i}, \sim\mathbf{M}_{C_i})$ in equation (6) lead to

$$\begin{aligned} V(\Delta) &= E_{\mathbf{M}_{C_i}}\left(E_{\mathbf{M}_{\sim C_i}} V_{\boldsymbol{\theta}|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i}}(\Delta|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i}) + V_{\mathbf{M}_{\sim C_i}} E_{\boldsymbol{\theta}|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i}}(\Delta|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i})\right) \\ &\quad + V_{\mathbf{M}_{C_i}} E_{\mathbf{M}_{\sim C_i}} E_{\boldsymbol{\theta}|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i}}(\Delta|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i}) \\ &= E_{\mathbf{M}_{C_i}} E_{\mathbf{M}_{\sim C_i}} V_{\boldsymbol{\theta}|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i}}(\Delta|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i}) + E_{\mathbf{M}_{C_i}} V_{\mathbf{M}_{\sim C_i}} E_{\boldsymbol{\theta}|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i}}(\Delta|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i}), \\ &\quad + V_{\mathbf{M}_{C_i}} E_{\mathbf{M}_{\sim C_i}} E_{\boldsymbol{\theta}|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i}}(\Delta|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i}) \\ &= V(\boldsymbol{\theta}) + V(\mathbf{M}_{\sim C_i}) + V(\mathbf{M}_{C_i}) \end{aligned} \tag{7}$$

where the total variance is decomposed into three sources: a given model component $\mathbf{M}_{C_i}$, the other model components $\mathbf{M}_{\sim C_i}$, and the parameters. The only difference between equations (7) and (4) is that the model components $\mathbf{M}_{C_i}$ and $\mathbf{M}_{\sim C_i}$ are independent on each other thus their order is reversible in variance decomposition. Sensitivity indices for this new structure can be defined for model components $\mathbf{M}_{C_i}$s and model parameters as:

$$\begin{aligned} S_{\mathbf{M}_{C_i}} &= \frac{V_{\mathbf{M}_{C_i}} E_{\mathbf{M}_{\sim C_i}} E_{\boldsymbol{\theta}|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i}}(\Delta|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i})}{V(\Delta)} = \frac{V(\mathbf{M}_{C_i})}{V(\Delta)} \\ S_{\boldsymbol{\theta}} &= \frac{E_{\mathbf{M}_{C_i}} E_{\mathbf{M}_{\sim C_i}} V_{\boldsymbol{\theta}|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i}}(\Delta|\mathbf{M}_{C_i},\mathbf{M}_{\sim C_i})}{V(\Delta)} = \frac{V(\boldsymbol{\theta})}{V(\Delta)} \end{aligned} \tag{8}$$

The sensitivity indices defined in equations (8) and (5) are similar, with the partial variance contributed by the parameters averaged over model components in equation (8) rather than over scenario and model as in equation (5).

2.3. Calculation of Sensitivity Indices Using Monte Carlo Approach

The Monte Carlo method was used to evaluate sensitivity indices defined in equations (5) and (8). Since the uncertain sources considered in this research are grouped parameters with strong spatial correlation, to take these correlations into account, geostatistical tools include kriging and conditional simulations were used to generate MC samples based on data at the monitoring wells. To evaluate the sensitivity indices defined in equation (5), the uncertainty quantification starts with the quantification of parametric uncertainty, and then moves to model uncertainty, and ultimately the scenario uncertainty through model and scenario averaging techniques. Specifically speaking, we first calculated the mean and variance of model predictions with respect to uncertain parameters given a model and a scenario, then taking the average or variance of parameter-induced mean or variance over different models, and finally over scenarios.

Assuming that we have L alternative scenarios, K plausible models under each scenario, and n MC generated parameter sets given each scenario and model, the parametric uncertainty can be estimated via

$$V(\boldsymbol{\theta})=\sum_l\sum_k\left(\frac{1}{n}\sum_{j=1}^{n}\Delta^2(\theta_j|M_k,S_l)-\left(\frac{1}{n}\sum_{j=1}^{n}\Delta(\theta_j|M_k,S_l)\right)^2\right)P(M_k|S_l)P(S_l), \tag{9}$$

where $P(M_k|S_l)$ is the weight of model M_k under scenario S_l satisfying $\sum_k P(M_k|S_l)=1$, and $P(S_l)$ is the weight of scenarios satisfying $\sum_l P(S_l)=1$. Similarly, the model and scenario uncertainty can be calculated as:

$$\begin{aligned}V(\mathbf{M})&=\sum_l P(S_l)\begin{pmatrix}E_{\mathbf{M}|S_l}\left(E_{\boldsymbol{\theta}|\mathbf{M},S_l}(\Delta|\mathbf{M},S_l)\right)^2\\-\left(E_{\mathbf{M}|S_l}E_{\boldsymbol{\theta}|\mathbf{M},S_l}(\Delta|\mathbf{M},S_l)\right)^2\end{pmatrix}\\&=\sum_l P(S_l)\begin{pmatrix}\sum_k\left(\frac{1}{n}\sum_{j=1}^{n}\Delta(\boldsymbol{\theta}_j|M_k,S_l)\right)^2P(M_k|S_l)-\\\left(\sum_k\left(\frac{1}{n}\sum_{j=1}^{n}\Delta(\boldsymbol{\theta}_j|M_k,S_l)P(M_k|S_l)\right)\right)^2\end{pmatrix},\end{aligned} \tag{10}$$

and

$$\begin{aligned}V(\mathbf{S})&=E_{\mathbf{S}}\left(E_{\mathbf{M}|\mathbf{S}}E_{\boldsymbol{\theta}|\mathbf{M},\mathbf{S}}(\Delta|\mathbf{M},\mathbf{S})\right)^2-\left(E_{\mathbf{S}}E_{\mathbf{M}|\mathbf{S}}E_{\boldsymbol{\theta}|\mathbf{M},\mathbf{S}}(\Delta|\mathbf{M},\mathbf{S})\right)^2\\&=\sum_l P(S_l)\left(\sum_k P(M_k|S_l)\left(\frac{1}{n}\sum_{j=1}^{n}\Delta_k(\theta_j|M_k,S_l)\right)\right)^2\\&\quad-\left(\sum_l\sum_k P(S_l)P(M_k|S_l)\left(\frac{1}{n}\sum_{j=1}^{n}\Delta_k(\theta_j|M_k,S_l)\right)\right)^2\end{aligned}. \tag{11}$$

Based on equations (9)–(11), evaluating the sensitivity indices defined in equation (5) is straightforward.

The sensitivity indices defined in equation (8) can be evaluated in a similar manner, except that the summation over scenarios is not needed. In addition, the calculation of $V(\mathbf{M}_{C_i})$ needs to consider two situations. If the component is a discrete random variable, the variance can be evaluated in a way similar to equation (11):

$$\begin{aligned}V(\mathbf{M}_{C_i})&=\sum_l P\left((M_{C_i})_l\right)\left(\sum_k P\left((M_{\sim C_i})_k\right)\left(\frac{1}{n}\sum_{j=1}^{n}\Delta(\boldsymbol{\theta}_j|M_{C_i},M_{\sim C_i})\right)\right)^2\\&\quad-\left(\sum_l\sum_k P\left((M_{C_i})_l\right)P\left((M_{\sim C_i})_k\right)\left(\frac{1}{n}\sum_{j=1}^{n}\Delta(\boldsymbol{\theta}_j|M_{C_i},M_{\sim C_i})\right)\right)^2\end{aligned}. \tag{12}$$

Similar to the scenario and model weights, the $P\left((M_{C_i})_l\right)$ and $P\left((M_{\sim C_i})_k\right)$ are the weights of model components satisfying $\sum_l P\left((M_{C_i})_l\right)=1$ and $\sum_k P\left((M_{\sim C_i})_k\right)=1$. If the component is a continuous random variable (e.g., river stage), the variance is evaluated in a manner similar to equation (9) without the summations over models and scenarios follow the same procedure:

$$\begin{aligned}V(\mathbf{M}_{C_i})&=\frac{1}{n_3}\sum_{s=1}^{n_3}\left(\frac{1}{n_2}\sum_{m=1}^{n_2}\left(\frac{1}{n_1}\sum_{j=1}^{n_1}\Delta(\boldsymbol{\theta}_j|(M_{C_i})_m,(M_{\sim C_i})_s)\right)\right)^2\\&\quad-\left(\frac{1}{n_2n_3}\sum_{s=1}^{n_3}\sum_{m=1}^{n_2}\left(\frac{1}{n_1}\sum_{j=1}^{n_1}\Delta(\boldsymbol{\theta}_j|(M_{C_i})_m,(M_{\sim C_i})_s)\right)\right)^2\end{aligned}, \tag{13}$$

where n_1, n_2, n_3 are the numbers of Monte Carlo samples for parameter set $\boldsymbol{\theta}$, model components M_{C_i} and $\sim M_{C_i}$, respectively.

The selection of weights for alternative models or scenarios could be based on the prior confidence on them, or using objective criteria such as their posterior probabilities based on the Bayesian theorem (i.e., $P(M|\mathbf{D})=\frac{p(\mathbf{D}|M)P(M)}{p(\mathbf{D})}$, where $\boldsymbol{D}$ is the data) [*Schoniger et al.*, 2014; *Liu et al.*, 2016]. We used prior equal weights for model components in our test case. We also verified that the selection of unequal weights only slightly changed our results of sensitivity indices while not impacting any of our main findings.

3. Application to Dynamic Groundwater Flow and Transport Modeling

The hierarchical sensitivity analysis framework was implemented to evaluate the relative importance of various model inputs that contribute to the overall uncertainty in groundwater flow and transport modeling at the U.S. Department of Energy's Hanford 300 Area Integrated Field Research Challenge (IFRC) site [*Zachara et al*., 2013]. A persistent uranium groundwater plume exists at this site that discharges to the nearby Columbia River through a hydrologically dynamic groundwater-surface water interaction zone [*Zachara et al*., 2016]. The results of comprehensive field characterization measurements, and numerous field experiments and modeling studies performed at the site [*Hammond and Lichtner,* 2010; *Zachara et al*., 2016] provide a strong technical basis for the test case used herein. This test case is a complex natural system that demonstrates various challenges associated with building a predictive model in the face of uncertainty.

3.1. Site and Experimental Descriptions

The study site is located in southeastern Washington State, about 250 m west of the Columbia River. A well field of 43 wells, mostly concentrated in a triangle of 1600 m^2 (see Figure 2), was drilled to characterize subsurface hydrogeological properties, perform tracer tests, and monitor plume migration [*Chen et al*., 2012, 2013; *Zachara et al*., 2013, 2016]. The aquifer consists of two distinct and generally horizontally bedded geological formations with substantial contrast in their permeability: the highly permeable Hanford formation that is dominated by coarse river cobble, and the less permeable Ringold Formation that contains silty sand [*Chen et al*., 2012]. The Hanford formation is thus the main conduit for groundwater and surface water exchange. The permeability in the Hanford formation is large and very heterogeneous, varying over several orders of magnitude. The complex permeability field combined with dramatic water table fluctuations resulting from Columbia River stage changes creates complex transport flow patterns in the aquifer system. More detailed descriptions of the site can be found in *Bjornstad et al*. [2009] and *Zachara et al*. [2013].

A tracer test was conducted in October 2011 to study the flow and transport process as well as uranium surface complexation. This sensitivity study focuses only on flow and nonreactive (tracer) transport. The injected water was spiked with chloride (Cl^-) to a concentration of 100 mg/L which was then injected into well 2–34 (Figure 2) at a rate of 567 L/min over a period of 5 h. The tracer concentration was monitored at the observation wells for more than 2,000 hours after the injection. However, only the first 192 h of data were considered in this analysis due to subsequent injections of other tracers that could have influenced flow and transport process beyond the 192 h window.

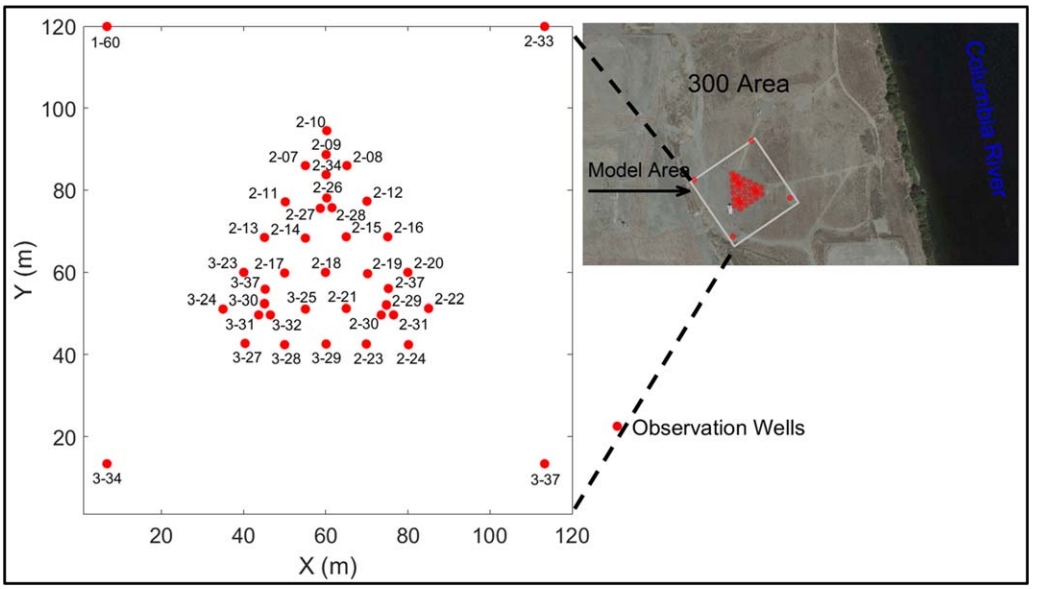

Figure 2. The model domain for the tracer test simulation with observation wells. The right upper subplot presents the location of model domain relative to the Columbia River and Hanford 300 Area.

3.2. Numerical Model and Its Uncertainty Sources

The modeling domain was 120 m × 120 m in area and 15 m in depth, discretized into uniform grids of 1 m × 1 m in the horizontal and 0.5 m in the vertical (Figure 2). The tracer experiment was simulated using PFLOTRAN, a massively parallel three-dimensional flow and reactive transport code [*Hammond et al*., 2014]. We applied transient hydrostatic head boundary conditions at the four lateral boundaries, which were kriged from a set of wells within and beyond the model domain using the approach described by *Chen et al*., [2012, 2013]. A no-flow boundary condition was specified at the bottom of the domain since it is underlain by the Ringold Formation with low permeability. A constant-rate recharge of 55 mm/yr was applied at the top boundary. The tracer boundary conditions were set to be free outflow at the four lateral boundaries, while recharge water at the upper boundary was assumed to be tracer free. A zero-flux of tracer was applied at the lower boundary. The initial flow condition was a hydrostatic pressure field interpolated from the same set of wells that were used to generate the transient lateral flow boundary conditions, and the entire domain was set to be tracer free at the beginning of the simulation. We used hourly time step for the flow and tracer transport simulations. The time series of simulated hydraulic heads and tracer concentrations at the observation well locations (shown in Figure 2) and snapshots of their fields were output for the analyses of spatial and temporal dynamics of sensitivity indices.

We identified three major sources of uncertainty for modeling the tracer transport experiment at the site: (1) the transient hydrostatic head boundary conditions that were interpolated from nearby wells, (2) the topography of the contact surface between the Hanford formation and the underlying Ringold formation (referred to as R-H contact elevations hereinafter), and (3) the permeability field within the Hanford formation. Note that the Ringold Formation is assumed to have a constant permeability field because of its much lower permeability as compared to the Hanford formation. The aim of the sensitivity analysis was to quantify the relative importance of the three sources of uncertainty to the simulated hydraulic heads and solute concentrations in order to prioritize further characterization efforts for predictive model improvement. Because all the three factors were spatially distributed, our study considered each factor as a group with their internal variability and spatial correlation accounted for by geostatistical approaches. The contributions of these factors to the output uncertainty also vary in space, and the most sensitive region of a given factor can be identified after the initial analysis.

All three of the uncertainty sources were placed to the appropriate hierarchy level in the uncertainty framework according to their inherent roles in the numerical model. The boundary condition and R-H contact elevation are considered as two model uncertainty components without hierarchical relationship. The boundary condition is the driving force of the groundwater system, while the R-H contact elevation defines the large-scale geological structure of aquifer. Thus, there is no dependency between these two factors. The permeability field within the Hanford formation was placed at the parameter level, because its influence on model outputs depends on the selected boundary condition and the topography of the R-H contact. The uncertainty structure in Figure 1b fits this case study. For models with two components, $\mathbf{M}_{C_1}$ and $\mathbf{M}_{C_2}$, equation (7) becomes

$$\begin{aligned} V(\Delta) = & E_{\mathbf{M}_{C_2}} E_{\mathbf{M}_{C_1}} V_{\boldsymbol{\theta}|\mathbf{M}_{C_1},\mathbf{M}_{C_2}}(\Delta|\mathbf{M}_{C_1},\mathbf{M}_{C_2}) + E_{\mathbf{M}_{C_1}} V_{\mathbf{M}_{C_2}} E_{\boldsymbol{\theta}|\mathbf{M}_{C_1},\mathbf{M}_{C_2}}(\Delta|\mathbf{M}_{C_1},\mathbf{M}_{C_2}) \\ & + V_{\mathbf{M}_{C_1}} E_{\mathbf{M}_{C_2}} E_{\boldsymbol{\theta}|\mathbf{M}_{C_1},\mathbf{M}_{C_2}}(\Delta|\mathbf{M}_{C_1},\mathbf{M}_{C_2}) \\ = & V(\boldsymbol{\theta}) + V(\mathbf{M}_{C_2}) + V(\mathbf{M}_{C_1}) \end{aligned} \quad . \tag{14}$$

The sensitivity indices $\mathbf{S}_{\mathbf{M}_{C_1}}$, $\mathbf{S}_{\mathbf{M}_{C_2}}$ were defined following the equation (8).

3.3. Sample Generation for Sensitivity Analysis

For all three uncertain factors, we used point measurements from our site to build a spatial correlation model for interpolations to the entire model domain. For example, geological logs from the monitoring wells define the elevations of the Ringold-Hanford contact (R-H contact), monitored water elevations can be used to interpolate hourly head boundary conditions, and permeability estimated from pumping tests can be used to map the hydraulic conductivity field in the Hanford formation. Geostatistical approaches such as kriging and conditional simulation [*Rubin*, 2003] have enabled the assimilation of such point measurements to estimate their spatial distribution and to generate geostatistical samples that represent the remaining uncertainty of the factor. Geostatistical realizations also enable us to analyze the spatially correlated

uncertain factors which represent the same system characteristic as a group instead of treating each location as an independent individual.

We generated 100 realizations of the 3-D permeability field from realizations of spatial correlation parameters and conditioning points determined in previous studies [*Murakami et al*., 2010; *Chen et al*., 2012] using the kriging method, following the method described in detail by *Chen et al*. [2012]. For the boundary condition at each time step and R-H contact elevation, we first generated 100 realizations using conditional simulation of a Gaussian random field, with conditioning data obtained from hourly well monitoring data of hydraulic head or from drilling logs of 300 Area wells queried from the Hanford Environmental Information System (HEIS), respectively. We then selected a subset of nine and eight representative realizations, respectively, under the guiding principle of preserving the level of uncertainty reflected in 100 realizations. In doing so, we conducted bootstrapping with replacements for 1,000,000 times, and then selected the combination that produced uncertainty closest to that of the 100 samples. These two subsets of realizations were combined with the kriged fields for the both factors and an alternative expert interpretation for R-H contact elevation to form 10 representative samples for both model components to keep the computational cost manageable. The kriged field represents an alternative approach that is often used when multiple realizations are not considered and the resulting estimate is equivalent to the mean field of a conditional simulation. It was thus included as a plausible sample. The sample sizes for the three uncertain factors reflected their differences in data dimensionality, with permeability field containing a significant portion of the 30 $\times$ 120 $\times$ 120 grid cells in Hanford formation, R-H contact elevation containing 120 $\times$ 120 grid cells, and boundary conditions containing 4 $\times$ 120 grid cells for each time step.

4. Results and Discussions

4.1. Selection of Representative Samples for Boundary Conditions and R-H Contact

An example of 10 samples of boundary conditions at 100 h after the start of simulation is shown in Figure 3. The intrinsic variability represented by the nine realizations generated from conditional simulation was confirmed to be comparable to that of the 100 realizations (results not shown). We chose eight realizations of R-H contact elevations out of the 100 realizations generated by conditional simulation for the sensitivity analysis (Figure 4). In addition to the kriged sample, an R-H contact elevation map generated by an experienced hydrogeologist using the EarthVision software was also included, which is referred to as the expert interpretation hereafter. The conditionally simulated samples and their mean field displayed significantly

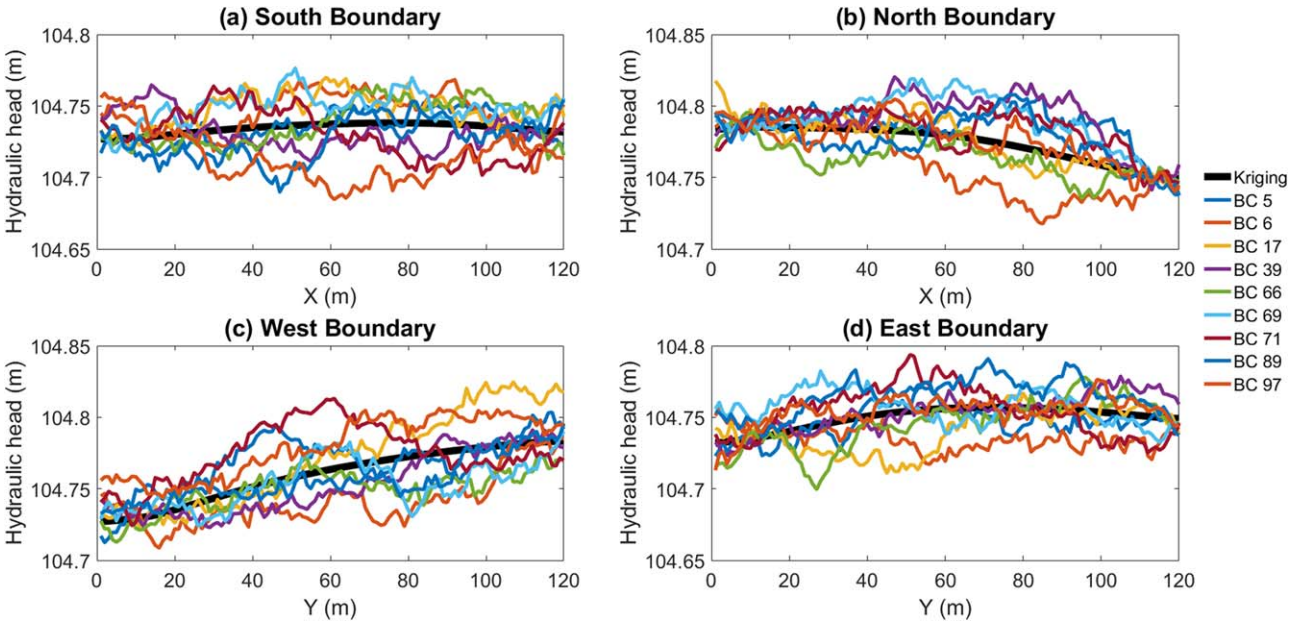

Figure 3. The 10 selected boundary condition samples for the model domain (a) south boundary, (b) north boundary, (c) west boundary, and (d) east boundary at the simulation time of 100 h.

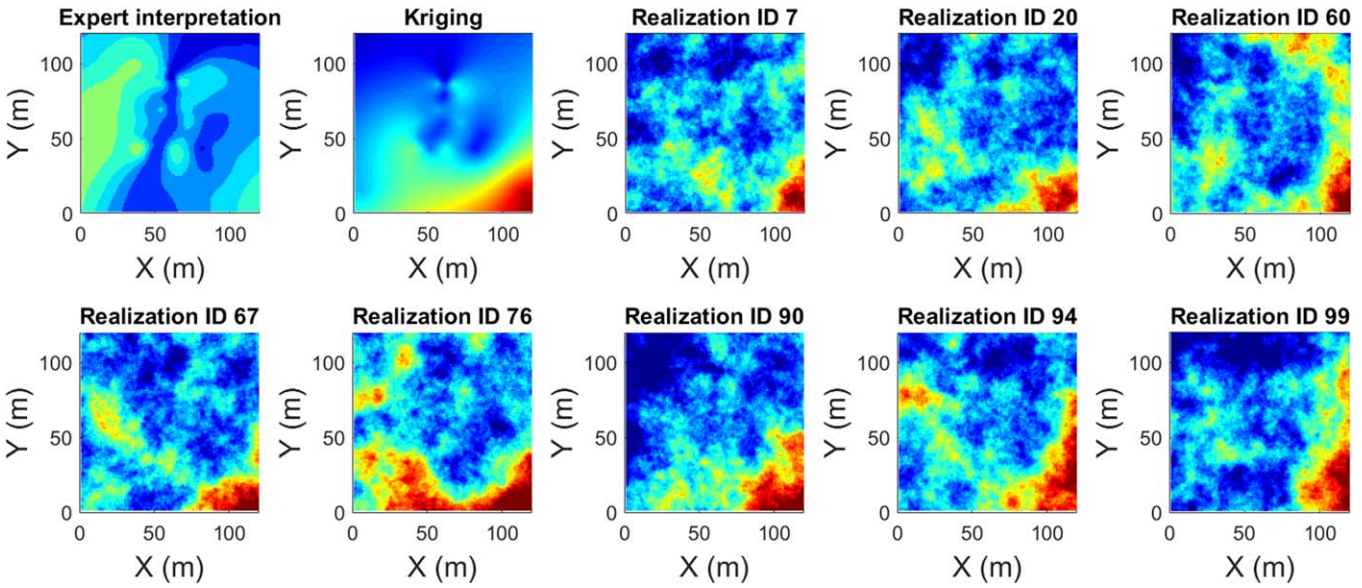

Figure 4. The 10 selected samples of the R-H contact elevations for the model domain.

higher R-H contact than the expert interpretation at the southeast corner (lower right), which was caused by two possible interpretations of the R-H contact elevation based on the drilling log of well 3–37. The expert chose the lower elevation, while we chose the higher elevation to maximize the uncertainty in this geological setting.

4.2. PFLOTRAN Simulations of Flow and Transport

PFLOTRAN simulations were run on 10,000 combinations of the three input factors, i.e., 100 permeability field realizations with 10 samples of boundary conditions and 10 samples of R-H contact elevations. The wall clock time for each realization was approximately 15 min using 128 cores on the Hopper supercomputer at the National Energy Research Scientific Computing Center (NERSC).

The simulated hydraulic head and tracer breakthrough resulting from these 10,000 realizations were first summarized and compared to observations to assess the overall uncertainty (Figures 5 and 6 for select wells only). Fewer wells were monitored for hydraulic head because of the flat hydraulic gradient within the model domain. The hydraulic head was accurately reproduced during the course of simulation, as represented by wells 3–29 and 2–10, which are located in the south and north ends of the well field. The absolute errors in hydraulic head (Figures 5b and 5d) were less than 5 cm, which implied insensitivity of hydraulic head to variability in model parameters.

In contrast to the hydraulic head, simulated tracer concentrations (Figure 6) exhibited substantial variability among the 10,000 realizations. Note that both the observed and simulated concentrations were normalized by the injection concentration of the tracer to facilitate comparisons between wells. While the 10,000 realizations of simulated tracer breakthrough encapsulated the observed tracer behavior for most wells, significant uncertainty was resulted in the model simulations, which renders it necessary to quantify the contribution of uncertainty from individual sources for uncertainty reduction.

4.3. Sensitivity Indices for Hydraulic Head

The sensitivity indices defined in equation (14) were first calculated for the hydraulic head predictions in all observation wells for each time step. Two observation wells with different distances from the injection well were chosen as examples, 2–18 and 2–26, where their sensitivity indices were plotted at four different time steps (Figure 7). Despite the differences in specific contributions from each uncertainty source in time and space, the boundary condition was the dominant contributor of uncertainty to the hydraulic head, accounting for 50–84% of its total uncertainty in the head predictions at these two locations. The R-H contact elevations contributed the least uncertainty in general, whereas the permeability field is the second most important uncertainty source. The permeability field and R-H contact elevation contributed more

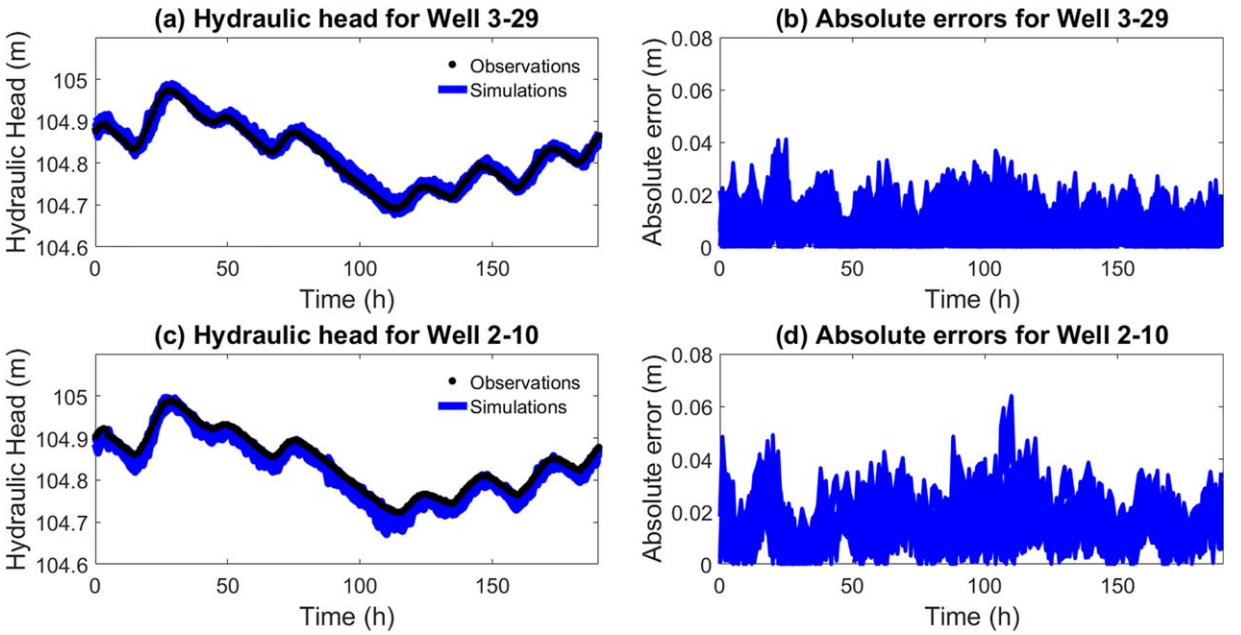

Figure 5. PFLOTRAN hydraulic head simulation results of 10,000 realizations for (a) well 3–29 and (c) well 2–10 compared to field observations, and the absolute errors of PFLOTRAN hydraulic head simulation results for (b) well 3–29 and (d) well 2–10.

uncertainty to well 2–18 than to well 2–26 because the former is farther downstream from the injection point, thus experienced a greater accumulated impact of variability in flow paths. The increasing variability in flow paths over time also explains the increasing contributions from the permeability field over time.

The sensitivity indices were calculated for all the grid cells in the model domain at a few selected time steps, from which we generated sensitivity maps at those time steps (Figure 8) to demonstrate the spatial

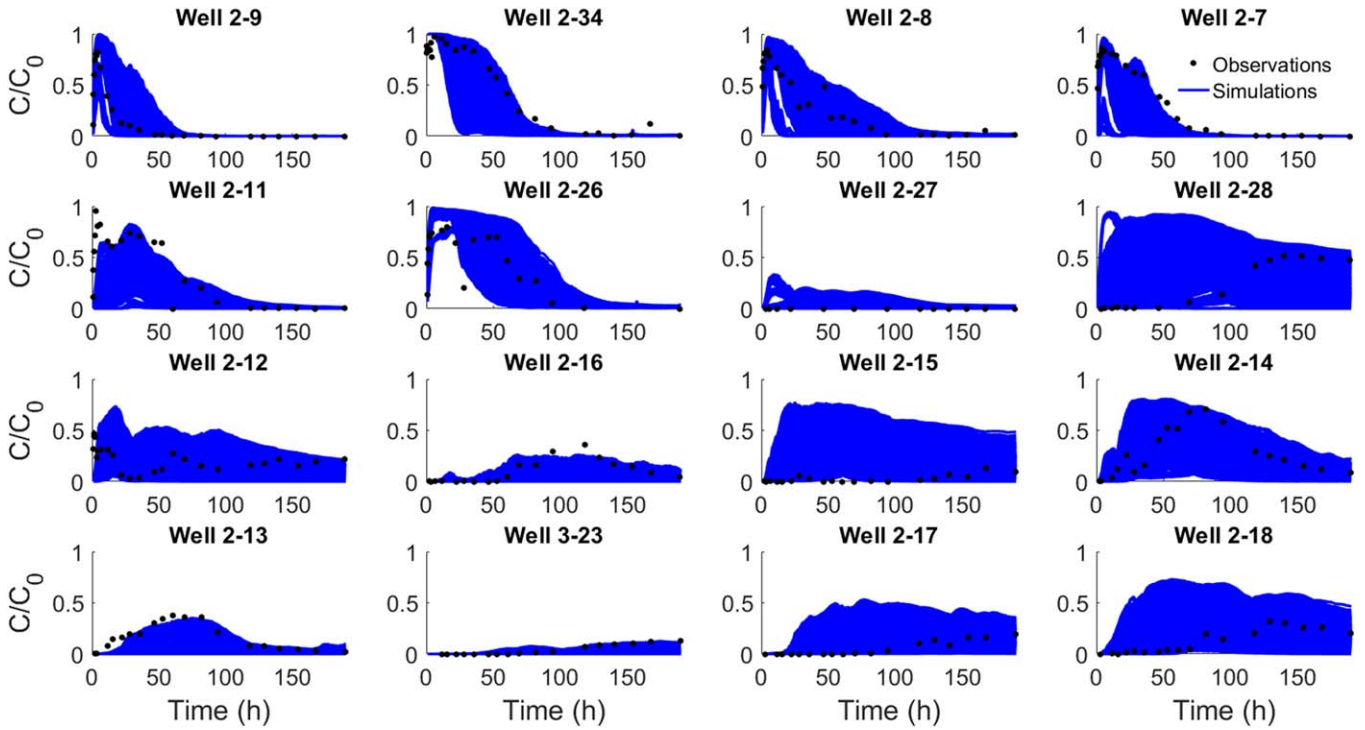

Figure 6. PFLOTRAN tracer concentration simulation results of 10,000 realizations compared to field observations for 16 wells.

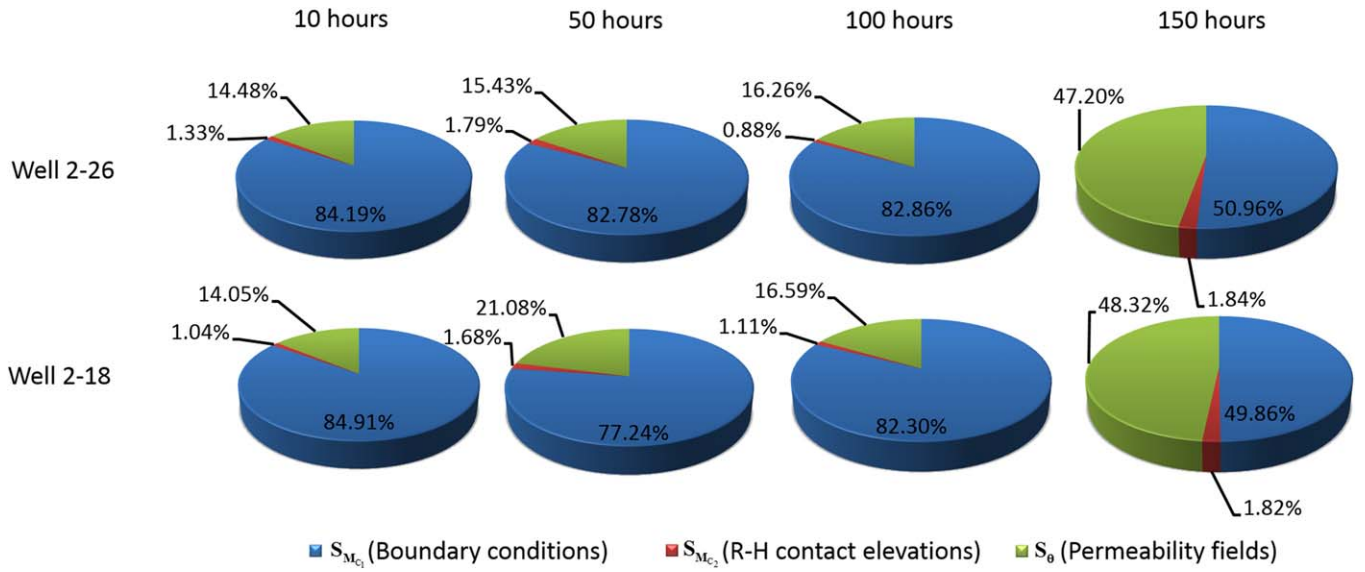

Figure 7. Estimated sensitivity indices for hydraulic head predictions at wells 2–26 and 2–18 at four different elapsed times.

variability of sensitivity contribution from each source. The maps demonstrated that the boundary condition was the most important contributor of uncertainty to the hydraulic head predictions for the entire domain at all times. It was also observed that the contribution of uncertainty from the boundary condition

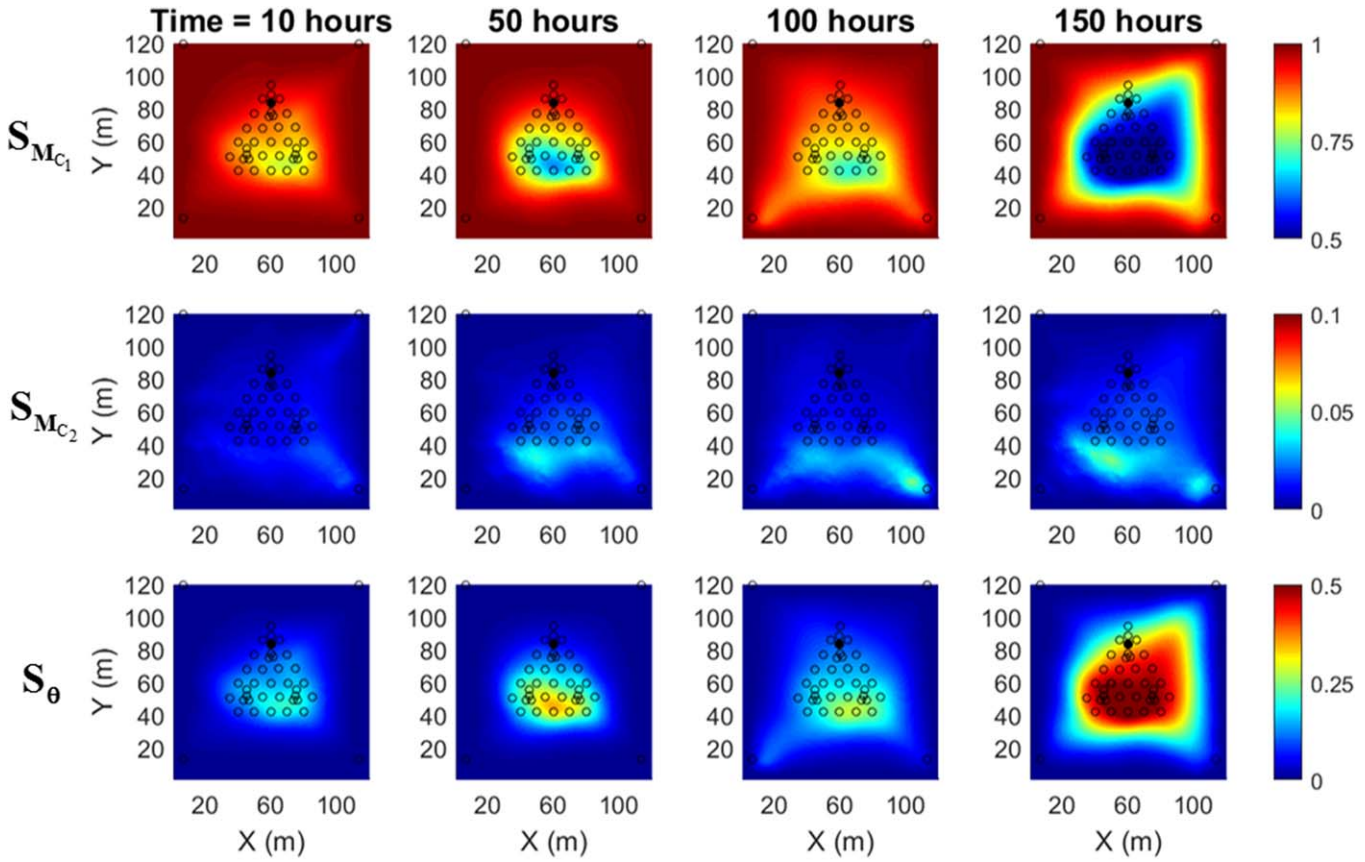

Figure 8. The maps of sensitivity indices for hydraulic head predictions at different elapsed times. Note that different color scales were used for each uncertainty source to show spatial variability. The hollow circles represent wells and the solid circle is the location of tracer injection.

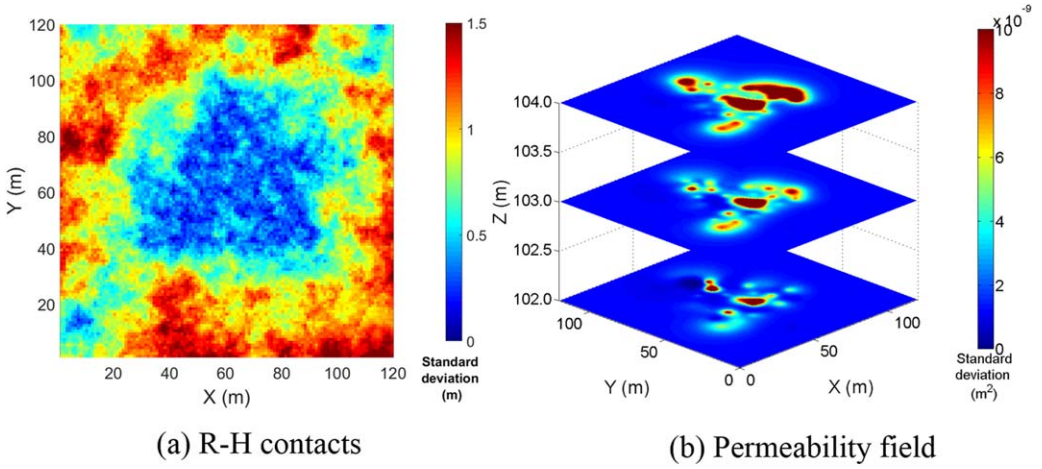

Figure 9. The standard deviations of (a) 10 selected samples of R-H contact elevations, and (b) 100 permeability field realizations at elevations of 102, 103, 104 m.

decreased towards the central part of the model domain; whereas the permeability field contributed more to uncertainty in this central area compared to the boundary area. This pattern was especially prominent at the 150 h, when the permeability field contributed to more than 40% of the total uncertainty in the central area of the domain. The R-H contact elevation contributed the least to the total uncertainty in the head predictions with peak contributions occurring at the south end of domain at later times. These spatial patterns were combined results of the spatial distribution of uncertainties from each source (Figure 9) and the propagation of pressure wave from the boundary to the central domain over time.

4.4. Sensitivity Indices for Tracer Concentration

The sensitivity indices for tracer concentrations at the observation wells were computed following the same analyses procedure for the hydraulic head. The results for tracer concentrations, as shown, for example, wells 2–18 and 2–26 (Figure 10), differed significantly from the results of hydraulic head predictions at these same wells. The permeability field contributed more than 94% and 99% of total uncertainty to simulated tracer concentrations at wells 2–26 and 2–18, respectively. The boundary condition was the second most important source of uncertainty, in general, with increasing contribution to wells closer to the model boundaries. The R-H contact elevation contributed minimal uncertainty to tracer concentration predictions

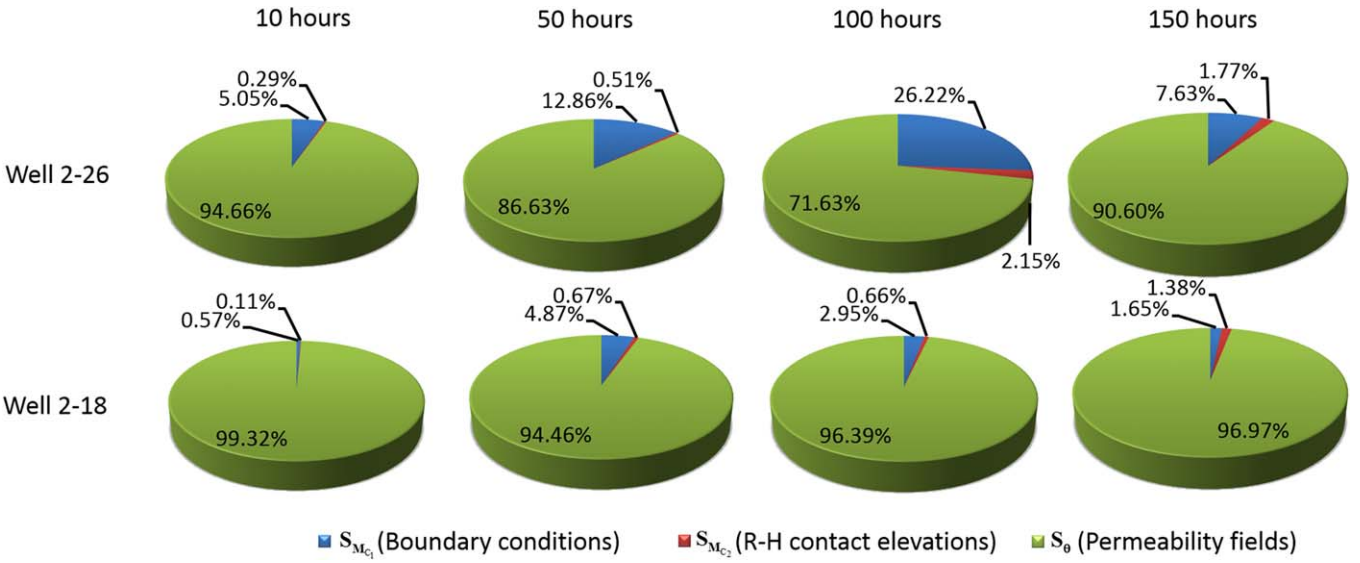

Figure 10. Sensitivity indices for tracer concentration predictions at well 2–26 and 2–18 at different elapsed times.

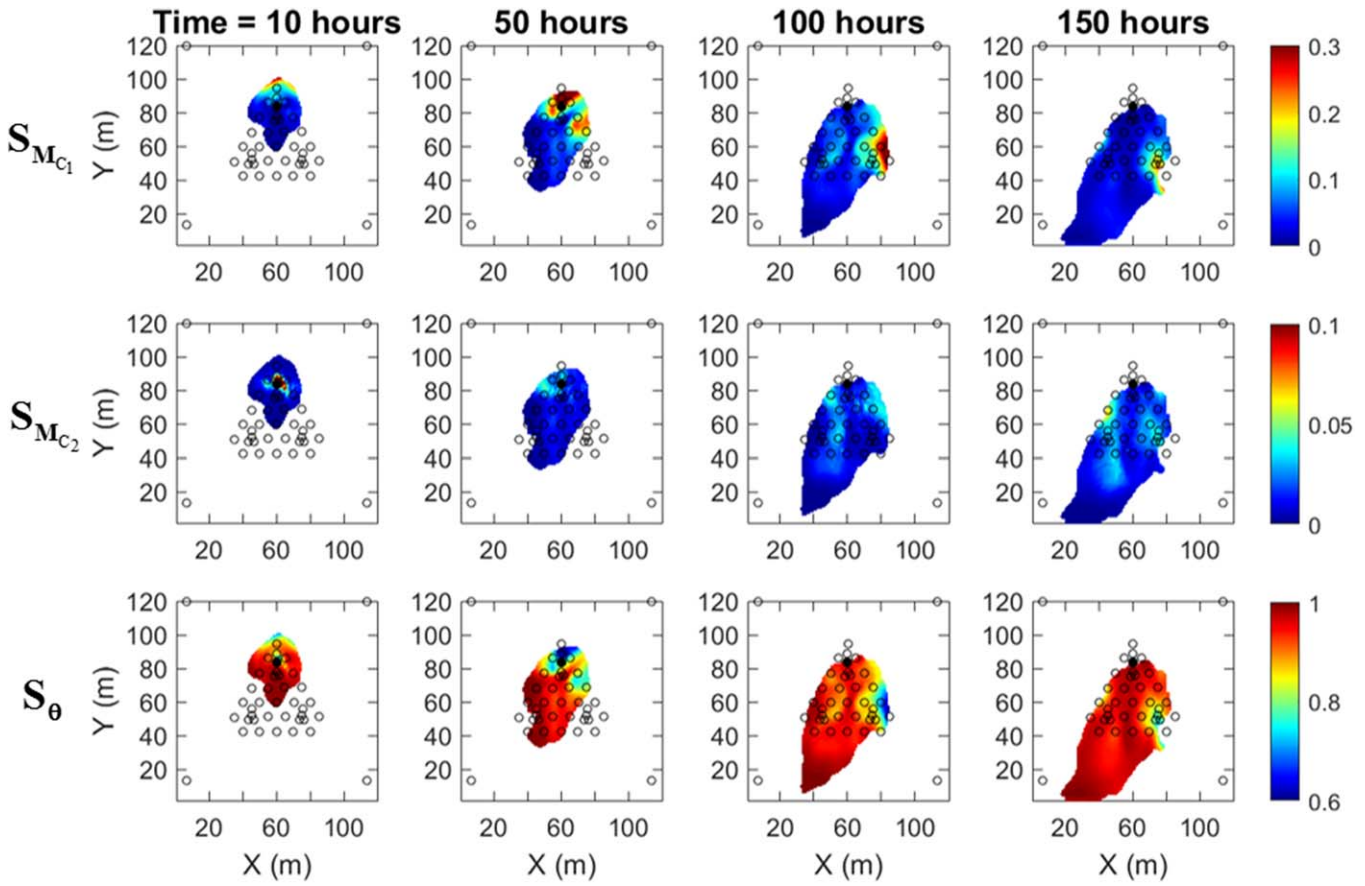

Figure 11. Sensitivity maps for predicted tracer concentrations only considering locations where significant tracer breakthrough (>5% injection concentration) was observed at different elapsed times. Note that different color scales were used for each uncertainty source to show spatial variability. The hollow circles represent observation wells and the solid circle the location of tracer injection.

in these two wells. The large difference in sensitivity indices between predicted hydraulic heads and tracer concentrations illustrates the importance of dispersion on the transport processes, especially the macrodispersion caused by heterogeneous permeability field.

The whole-domain sensitivity maps for tracer concentration displayed a clear effect of southwestern migration direction (Figure 11). The permeability field is the most important uncertainty source at all the times. Boundary conditions contributed a significant amount of uncertainty to simulated tracer concentrations in northern and eastern part of the plume, decreasing over time as the tracer plume moved toward the central part of the domain. This trend was consistent with that of the simulated hydraulic head. In contrast, the permeability field contributed increasing amount of uncertainty to the simulated tracer plume over time, which again confirmed the increasing influence of macrodispersion on transport. The limited contribution of uncertainty from the R-H contact might be partially explained by the shallow depths at which our observations wells were screened (103–105 m in elevation) as compared to the R-H contact elevations (lower than 102 m) over the model domain.

4.5. Impacts of Sample Size

An important question that remains is how many realizations are needed to perform reliable sensitivity analysis. We ran 10,000 simulations to perform the analyses in this study, as it is computationally affordable for our problem size using a user-facility supercomputer. However, this number may be too large for other modeling problems that have more processes over bigger domains. To provide a preliminary answer to this question, we reduced the number of realizations to four for boundary conditions and R-H contact elevations applying the same sample selection strategy described in section 3.3. These choices led to 1600 model simulations. The results from this reduced set of samples (sensitivity maps that are not shown), revealed only slight differences from the full set of 10,000 realizations. Significantly, the importance ranking of all the

uncertain factors remained the same. This efficiency test demonstrated that an affordable number of simulations may achieve desirable accuracy given a deliberate sample selection strategy.

5. Conclusions

We developed a new variance-based sensitivity analysis method based on the hierarchical uncertainty quantification framework to identify important uncertainty sources for real-world complex groundwater transport modeling. Our work extended the hierarchical uncertainty quantification framework to variance-based sensitivity analysis, recognizing that sensitivity analysis is a natural step in uncertainty quantification and that variance decomposition is the mathematical foundation for both techniques. This new sensitivity analysis approach first divides uncertain input factors into a few major groups or sources based on their characteristics, such as parameters, models, and scenarios, and then builds a hierarchy within the uncertainty quantification framework for analyses that will follow. Geostatistical tools were used to characterize spatial correlations within a group/source and reduce data dimensionality. The new sensitivity indices were computed for each source using realizations of model simulations with input parameters generated from a Monte Carlo-based sampling scheme combined with geostatistics. The method is particularly useful for identifying the most important uncertainty source of a mechanistic model before refining parameters at specific locations.

We demonstrated and evaluated the new sensitivity analysis method in a real-world case of a dynamic groundwater flow and transport model, which was applied to a conservative tracer test in the groundwater-surface water interaction zone along the Columbia River shoreline in Washington State. Among the three main uncertainty sources of the numerical model: dynamic flow boundary conditions, aquifer-aquitard contact elevations, and permeability field, our analyses identified boundary conditions and the permeability field as the most important uncertainty source for the simulated hydraulic head and tracer concentration. These sensitivity results indicate that macrodispersion caused by heterogeneity was the most important mechanism for uncertainty in tracer plume movement in our system.

The sensitivity indices were calculated on different time steps for all the grid cells of the model domain to illustrate their spatial and temporal variations. These variations were significant for both hydraulic head and tracer concentration although the importance ranking of the uncertainty sources remained the same. Spatial maps of sensitivity changed over time, depending on where and when the most active flow and transport processes occurred. Our analyses imply that a simple analysis at a certain location or time point cannot capture the overall sensitivity information of such dynamic and heterogeneous problems. Other statistical tools may be needed when the dominant sources of uncertainty vary over space and time.

Although we demonstrated our sensitivity analysis framework using an example with a hierarchical structure of two layers of uncertainty in conceptual model (composed of two components—boundary condition and aquifer-aquitar topography) and parameters, the proposed framework is flexible and can be applied to problems with a different number of layers and structures, owing to the recursive nature of variance decomposition. This flexibility facilitates application to models where a definitive hierarchy may not exist. Furthermore, the efficiency of our sensitivity analysis method can be achieved by selecting a small number of representative samples at each layer to maintain the largest variability of a larger pool of samples. This selection process was demonstrated to be effective for our study, yielding similar sensitivity results from 1600 samples as compared to 10,000 samples. Managing the number of model simulations for sensitivity analysis is an essential step for most of the real-world models as they are computationally expensive, and our study has provided a useful strategy for reducing the total sample size. The sensitivity results may also depend on the strategy of assigning weights to each representative sample. Although we have tested the sensitivity analysis with a different set of weights for samples and the results showed negligible difference between two weighting schema, future study is needed to provide further guidance. The insights gained from our case study are transferrable to other environmental problems that may involve other processes or occur at larger scales.

Acknowledgments
This research was supported by the U.S. Department of Energy (DOE), Office of Biological and Environmental Research (BER), as part of BER's Subsurface Biogeochemical Research Program (SBR). The third author was supported in part by DOE Early Career award DE-SC0008272 and NSF-EAR grant 1552329. This research used resources of the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH11231. All the data used to produce the results in this study can be requested by contacting the corresponding author.

References

Abdelaziz, R., and B. Merkel (2015), Sensitivity analysis of transport modeling in a fractured gneiss aquifer, *J. Afr. Earth Sci.*, *103*, 121–127.
Baroni, G., and S. Tarantola (2014), A general probabilistic framework for uncertainty and global sensitivity analysis of deterministic models: A hydrological case study, *Environ. Modell. Software*, *51*, 26–34.
Beven, K. (2002), Towards a coherent philosophy for modeling the environment, *Proc. R. Soc. London, Ser. A*, *458*(2026), 2465–2484.
Beven, K. (2006), A manifesto for the equifinality thesis, *J. Hydrol.*, *320*, 18–36.

Bjornstad B., J. Horner, V. Vermeul, D. Lanigan, and P. Thorne (2009), Borehole completion and conceptual hydrogeologic model for the IFRC well field, 300 area, Hanford site, *PNNL-18340*, Pac. Northwest Natl, Lab., Richland, Wash.
Bredehoeft, J. D. (2003), From models to performance assessment: The conceptualization problem, *Ground Water*, *41*(5), 571–577.
Bredehoeft, J. D. (2005), The conceptualization model problem-surprise, *Hydrogeol. J.*, *13*, 37–46.
Campolongo, F., J. Cariboni, and A. Saltelli (2007), An effective screening design for sensitivity analysis of large models, *Environ. Modell. Software*, *22*, 1509–1518.
Chen, X., H. Murakami, M. S. Hahn, G. E. Hammond, M. L. Rockhold, J. M. Zachara, and Y. Rubin (2012), Three dimensional Bayesian geostatistical aquifer characterization at the Hanford 300 Area using tracer test data, *Water Resour. Res.*, *48*, W06501, doi:10.1029/2011WR010675.
Chen, X., G. E. Hammond, C. J. Murray, M. L. Rockhold, V. R. Vermeul, and J. M. Zachara (2013), Application of ensemble-based data assimilation techniques for aquifer characterization using tracer data at Hanford 300 area, *Water Resour. Res.*, *49*, 7064–7076, doi:10.1002/2012WR013285.
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, *Stochastic Environ. Res. Risk. Assess.*, *27*, 945–954.
Dai, Z., and J. Samper (2004), Inverse problem of multicomponent reactive chemical transport in porous media: Formulation and applications, *Water Resour. Res.*, *40*, W07407, doi:10.1029/2004WR003248.
Dai, H., and M. Ye (2015), Variance-based global sensitivity analysis for multiple scenarios and models with implementation using sparse grid collocation, *J. Hydrol.*, *528*, 286–300, doi:10.1016/j.jhydrol.2015.06034.
Dai, Z., P. H. Stauffer, J. W. Carey, R. S. Middleton, Z. Lu, J. F. Jacobs, K. Hnottavange-Telleen, and L. Spangle (2014), Pre-site characterization risk analysis for commercial-scale carbon sequestration, *Environ. Sci. Technol.*, *48*, 3908−3915.
Davis, J. A., S. B. Yabusaki, C. I. Steefel, J. M. Zachara, G. P. Curtis, G. D. Redden, L. J. Criscenti, and B. D. Honeyman (2004), Assessing conceptual models for subsurface reactive transport of inorganic contaminants, *Eos Trans. AGU*, *85*(4), 449–455.
Draper, D. (1995), Assessment and propagation of model uncertainty, *J. R. Stat. Soc., Ser. B*, *57*(1), 45–97.
Draper, D., A. Pereira, P. Prado, A. Saltelli, R. Cheal, S. Eguilior, B. Mendes, and S. Tarantola (1999), Scenario and parametric uncertainty in GESAMAC: A methodological study in nuclear waste disposal risk assessment, *Comput. Phys. Commun.*, *117*, 142–155.
Fajraoui, N., T. A. Mara, A. Younes, and R. Bouhliha (2012), Reactive transport parameter estimation and global sensitivity analysis using sparse polynomial chaos expansion, *Water Air Soil Pollut.*, *223*, 4183–4197.
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.
Greskowiak, J., H. Prommer, C. Liu, V. E. A. Post, R. Ma, C. Zheng, and J. M. Zachara (2010), Comparison of parameter sensitivities between a laboratory and field-scale model of uranium transport in a dual domain, distributed rate reactive system, *Water Resour. Res.*, *46*, W09509, doi:10.1029/2009WR008781.
Hammond, G. E., and P. C. Lichtner (2010), Field-scale model for the natural attenuation of uranium at the Hanford 300 Area using high performance computing, *Water Resour. Res.*, *46*, W09527, doi:10.1029/2009WR008819.
Hammond, G. E., P. C. Lichtner, and R. T. Mills (2014), Evaluating the performance of parallel subsurface simulators: An illustrative example with PFLOTRAN, *Water Resour. Res.*, *50*, 208–228, doi:10.1002/2012WR013483.
Herman, J. D., P. M. Reed, and T. Wagener (2013), Time-varying sensitivity analysis clarifies the effects of watershed model formulation on model behavior, *Water Resour. Res.*, *49*, 1400–1414, doi:10.1002/wrcr.20124.
Hoeting, J. A., D. Madigan, A. E. Raftery, and C. T. Volinsky (1999), Bayesian model averaging: A tutorial, *Stat. Sci.*, *14*(4), 382–417.
Lilburne, L., and S. Tarantola (2009), Sensitivity analysis of spatial models, *Int. J. Geogr. Inf. Sci.*, *23*(2), 151–168, doi:10.1080/13658810802094995.
Liu, Y., and H. V. Gupta (2007), Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework, *Water Resour. Res.*, *43*, W07401, doi:10.1029/2006WR005756.
Liu, P., A. S. Elshall, M. Ye, P. Beerli, X. Zeng, D. Lu, and Y. Tao (2016), Evaluating marginal likelihood with thermodynamic integration method and comparison with several other numerical methods, *Water Resour. Res.*, *52*, 734–758, doi:10.1002/2014WR016718.
Massmann, C., and H. Holzmann (2012), Analysis of the behavior of a rainfall-runoff model using three global sensitivity analysis methods evaluated at different temporal scales, *J. Hydrol.*, *475*, 97–110.
Matott, L. S., J. E. Babendreier, and S. T. Purucker (2009), Evaluating uncertainty in integrated environmental models: A review of concepts and tools, *Water Resour. Res.*, *45*, W06421, doi:10.1029/2008WR007301.
Matott, L. S. (2012), Screening-level sensitivity analysis for the design of pump-and-treat systems, *Ground Water Monit. Rem.*, *32*(2), 66–80.
Meyer, P. D., M. Ye, M. L. Rockhold, S. P. Neuman, and K. J. Cantrell (2007), Combined estimation of hydrogeologic conceptual model, parameter, and scenario uncertainty with application to uranium transport at the Hanford site 300 area, *NUREG/CR-6940 (PNNL-16396)*, U.S. Nucl. Regul. Comm., Washington, D. C.
Meyer, P. D., M. Ye, T. Nicholson, S. P. Neuman, and M. Rockhold (2014), Incorporating scenario uncertainty within a hydrogeologic uncertainty assessment methodology, in *Proceedings of the International Workshop on Model Uncertainty: Conceptual and Practical Issues in the Context of Risk-Informed Decision Making*, edited by A. Mosleh and J. Wood, Int. Workshop Ser. on Adv. Top. in Reliab. and Risk Anal., Cent. for Risk and Reliab., Univ. of Maryland, Md., ISSN: 1084-5658.
Mishra, S., N. Deeds, and G. Ruskauff, (2009), Global sensitivity analysis techniques for probabilistic ground water modeling, *Ground Water*, *47*, 727–744.
Morris, M. D. (1991), Factorial sampling plans for preliminary computational experiments, *Technometrics*, *33*(2), 161–174.
Murakami, H., X. Chen, M. S. Hahn, Y. Liu, M. L. Rockhold, V. R. Vermeul, J. M. Zachara, and Y. Rubin (2010), Bayesian approach for three-dimensional aquifer characterization at the Hanford 300 Area, *Hydrol. Earth Syst. Sci.*, *14*(10), 1989–2001.
Neuman, S. P. (2003), Maximum likelihood Bayesian averaging of alternative conceptual-mathematical models, *Stochastic Environ. Res. Risk A*, *17*(5), 291–305, doi:10.1007/s00477-003-0151-7.
Nossent, J., P. Elsen, and W. Bauwens (2011), Sobol' sensitivity analysis of a complex environmental model, *Environ. Modell. Software*, *26*, 1515–1525.
Pan, F., J. Zhu, M. Ye, Y. A. Pachepsky, and Y.-S. Wu (2011), Sensitivity analysis of unsaturated flow and contaminant transport with correlated parameters, *J. Hydrol.*, *397*, 238–249, doi:10.1016/j.jhydrol.2010.11.045.
Piggott, J. H., and J. D. Cawlfield (1996), Probabilistic sensitivity analysis for one-dimensional contaminant transport in the vadose zone, *J. Contam. Hydrol.*, *24*, 97–115.
Razavi, S., and H. V. Gupta (2015), What do we mean by sensitivity analysis? The need for comprehensive characterization of "global" sensitivity in Earth and Environmental systems models, *Water Resour. Res.*, *51*, 3070–3092, doi:10.1002/2014WR016527.
Razavi, S., and H. V. Gupta (2016a), A new framework for comprehensive, robust, and efficient global sensitivity analysis: 1. Theory, *Water Resour. Res.*, *52*, 423–439, doi:10.1002/2015WR017558.

Razavi, S., and H. V. Gupta (2016b), A new framework for comprehensive, robust, and efficient global sensitivity analysis: 2. Application, *Water Resour. Res.*, *52*, 440–455, doi:10.1002/2015WR017559.
Refsgaard, J. C., S. Christensen, T. O. Sonnenborg, D. Seifert, A. L. Hojberg, and L. Troldborg (2012), Review of strategies for handling geological uncertainty in groundwater flow and transport modeling, *Adv. Water Resour.*, *36*, 36–50.
Rojas, R., S. Kahunde, L. Peeters, O. Batelaan, L. Feyen, and A. Dassargues (2010), Application of a multimodel approach to account for conceptual model and scenario uncertainties in groundwater modelling, *J. Hydrol.*, *394*(3–4), 416–435.
Rubin, Y. (2003), *Applied Stochastic Hydrogeology*, Oxford Univ. Press, Cambridge, U. K.
Rubin, Y., X. Chen, H. Murakami, and M. Hahn (2010), A Bayesian approach for inverse modeling, data assimilation, and conditional simulation of spatial random fields, *Water Resour. Res.*, *46*, W10523, doi:10.1029/2009WR008799.
Saltelli, A. (2000), What is sensitivity analysis? in *Sensitivity Analysis*, A. Saltelli, K. Chan, and M. Scott, pp. 3–14, Wiley, Chichester.
Saltelli, A., and I. M. Sobol' (1995), About the use of rank transformation in sensitivity analysis of model output, *Reliab. Eng. Syst. Saf.*, *50*(3), 225–239.
Saltelli, A., S. Tarantola, and K. Chan (1998), Presenting results from model based studies to decision makers: Can sensitivity analysis be a defogging agent?, *Risk Anal.*, *18*, 799–803.
Saltelli, A., S. Tarantola, and K. P.-S. Chan (1999), A quantitative model independent method for global sensitivity analysis of model output, *Technometrics*, *41*, 39–56.
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*(2), 259–270.
Schoniger, A., T. Wholing, L. Samaniego, and W. Nowak (2014), Model selection on solid ground: Rigorous comparison of nine ways to evaluate Bayesian model evidence, *Water Resour. Res.*, *50*, 9484–9513, doi:10.1002/2014WR016062.
Shi, X., M. Ye, G. P. Curtis, G. L. Miller, P. D. Meyer, M. Kohler, S. Yabbusaki, and J. Wu (2014), Assessment of parametric uncertainty for groundwater reactive transport modeling, *Water Resour. Res.*, *50*, 4416–4439, doi:10.1002/2013WR013755.
Sobol', I. M. (1993), Sensitivity analysis for nonlinear mathematical models, *Math. Model. Comput. Exp.*, *1*(4), 407–414.
Song, X., C. Zhan, J. Xia, and F. Kong (2012), An efficient global sensitivity analysis approach for distributed hydrological model, *J. Geogr. Sci.*, *22*(2), 209–222.
Song, X., J. Zhang, C. Zhan, Y. Xuan, M. Ye, and C. Xu (2015), Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications, *J. Hydrol.*, *523*, 739–757.
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.advwater.2012.04.007.
Tebes-Stevens, C. L., F. Espinoza, and A. J. Valocchi (2001), Evaluating the sensitivity of a subsurface multicomponent reactive transport model with respect to transport and reaction parameters, *J. Contam. Hydrol.*, *52*, 3–27.
Tong, C. (2011). *PSUADE User's Manual (Version 1.2.0)*, *LLNL-SM-407882*, Lawrence Livermore Natl. Lab., Livermore, Calif.
van Griensven, A., T. Meixner, S. Grunwald, T. Bishop, M. Diluzio, and R. Srinivasan (2006), A global sensitivity analysis tool for the parameters of multi-variable catchment models, *J. Hydrol.*, *324*, 10–23.
van Werkhoven, K., T. Wagener, P. Reed, and Y. Tang (2008), Characterization of watershed model behavior across a hydroclimatic gradient, *Water Resour. Res.*, *44*, W01429, doi:10.1029/2007WR006271.
Volkova, E., B. Iooss, and F. Van Dorpe (2008), Global sensitivity analysis for a numerical model of radionuclide migration from the RRC "Kurchatov Institute" radwaste disposal site, *Stochastic Environ. Res. Risk. Assess.*, *22*, 17–31.
Wagener, T., K. van Werkhoven, P. Reed, and Y. Tang (2009), Multiobjective sensitivity analysis to understand the information content in streamflow observations for distributed watershed modeling, *Water Resour. Res.*, *45*, W02501, doi:10.1029/2008WR007347.
Wainwright, H. M., S. Finsterle, Y. Jung, Q. Zhou, and J. T. Birkholzer (2014), Making sense of global sensitivity analyses, *Comput. Geosci.*, *65*, 84–94.
Yang, J. (2011), Convergence and uncertainty analyses in Monte-Carlo based sensitivity analysis, *Environ. Modell. Software*, *26*(4), 444–457.
Ye, M., S. P. Neuman, and P. D. Meyer (2004), Maximum likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff, *Water Resour. Res.*, *40*, W05113, doi:10.1029/2003WR002557.
Ye, M., S. P. Neuman, P. D. Meyer, and K. F. Pohlmann (2005), Sensitivity analysis and assessment of prior model probabilities in MLBMA with application to unsaturated fractured tuff, *Water Resour. Res.*, *41*, W12429, doi:10.1029/2005WR004260.
Ye, M., P. D. Meyer, and S. P. Neuman (2008a), On model selection criteria in multimodel analysis, *Water Resour. Res.*, *44*, W03428, doi:10.1029/2008WR006803.
Ye, M., K. F. Pohlmann, and J. B. Chapman (2008b), Expert elicitation of recharge model probabilities for the Death Valley regional flow system, *J. Hydrol.*, *354*, 102–115.
Ye, M., K. F. Pohlmann, J. B. Chapman, G. M. Pohll, and D. M. Reeves (2010a), A model-averaging method for assessing groundwater conceptual model uncertainty, *Ground Water*, *48*(5), 716–728, doi:10.1111/j.1745-6584.2009.00633.x.
Ye, M., D. Lu, S. P. Neuman, and P. D. Meyer (2010b), Comment on "Inverse groundwater modeling for hydraulic conductivity estimation using Bayesian model averaging and variance window" by Frank T.-C. Tsai and Xiaobao Li, *Water Resour. Res.*, *46*, W02801, doi:10.1029/2009WR008501.
Zachara J., P. Long, J. Bargar, J. Davis, P. Fox, J. Fredrickson, M. Freshley, A. Konopka, C. Liu, J. McKinley, M. Rockhold, K. Williams, and S. Yabusaki (2013), Persistence of uranium groundwater plumes: Contrasting mechanisms at two DOE sites in the groundwater-river interaction zone, *J. Contam. Hydrol.*, *147*, 45–72.
Zachara, J. M., X. Chen, C. Murray, and G. Hammond (2016), River stage influences on uranium transport in a hydrologically dynamic groundwater-surface water transition zone, *Water Resour. Res.*, *52*, 1568–1590, doi:10.1002/2015WR018009.
Zhang, C., J. Chu, and G. Fu (2013), Sobol's sensitivity analysis for a distributed hydrological model of Yichun river basin, *China J. Hydrol.*, *480*, 58–68.