A Computationally Efficient Method for Estimating Multi-model Process Sensitivity Index

Heng Dai[1,2], Fangqiang Zhang[1,2], Ming Ye[3,*], Alberto Guadagnini[4], Qi Liu[5], Bill Hu[6], Songhu Yuan[1,2]

[1]State Key Laboratory of Biogeology and Environmental Geology, China University of Geosciences, Wuhan, Hubei 430074, China.

[2]Hubei Key Laboratory of Yangtze Catchment Environmental Aquatic Science, China University of Geosciences, Wuhan, Hubei 430074, China.

[3]Department of Earth, Ocean, and Atmospheric Science, Florida State University, Tallahassee, FL 32306, USA.

[4]Dipartimento di Ingegneria Civile e Ambientale, Politecnico di Milano, Piazza L. Da Vinci, 32, Milano, I-20133, Italy.

[5]College of Life Science and Technology, Jinan University, Guangzhou 510632, China.

[6]School of Water Conservancy and Environment, University of Jinan, Jinan, Shangdong, 250022, China.

*Corresponding Author, Email: mye@fsu.edu, Phone: (850) 645-4987.

Abstract

Identification of important processes of a hydrologic system is critical for improving process-based hydrologic modeling. To identify important processes while jointly considering parametric and model uncertainty, Dai et al. (2017) developed a multi-model process sensitivity index. Numerical evaluation of the index using a brute force Monte Carlo (MC) simulation is computationally expensive, because it requires a nested structure of parameter sampling and the number of model simulations is on the order of N^2 (N being the number of parameter samples). To reduce computational cost, develops a new method (here denoted as quasi-MC for brevity) that uses triple sets of parameter samples to remove the nested structure of parameter sampling in a theoretically rigorous way. It then illustrates the way the method is implemented using a quasi-MC algorithm. It reduces the number of model simulations from the order of N^2 to 2N. The performance of the quasi-MC method is assessed against the brute force MC approach and the recent binning method developed by Dai et al. (2017) through two synthetic cases of groundwater flow and solute transport modeling. Due to its rigorous theoretical foundation, the quasi-MC method overcomes the limitations imposed by the inherently empirical nature of the binning approach. We find that the quasi-MC method outperforms both the brute force Monte Carlo and the binning method in terms of computational requirements and theoretical aspects, thus strengthening its potential for the assessment of process sensitivity indices subject to various sources of uncertainty.

Key words: Multi-model process sensitivity index; Global sensitivity analysis; Quasi-MC method; Binning method; Process model uncertainty; Parametric uncertainty

1. Introduction

Development of process-based models is a key research focus in water related research areas. In this context, the functioning of a hydrologic system is often depicted through a model that embeds various components, each associated with a mathematical formulation representing a given system process (Montanari & Koutsoyiannis, 2012; Clark et al., 2015; Antonetti et al., 2016, 2017; Zhang 2019). Improving the performance of process-based models requires identification of important processes. One can then enhance their characterization through data collection and/or model calibration (Grayson & Bloschl, 2000; Bloschl, 2001; Beven, 2002; Razavi & Gupta, 2015; Antonetti et al., 2016). Important processes can be identified through a global sensitivity analysis targeting the assessment of the relative importance of model parameters. In this sense, processes associated with important parameters are typically considered to be important (Wainwright et al., 2014; Guse et al., 2016; Dell'Oca et al., 2017; Ceriotti et al., 2018; Melsen & Guse, 2019). Such an approach to sensitivity analysis is only geared towards the assessment of model parameters, while otherwise not considering model uncertainty, which arises when limited data and/or knowledge lead to multiple plausible conceptual-mathematical models (Neuman, 2003; Beven, 2006). New approaches to global sensitivity analysis have been recently developed to address model uncertainty for the identification of important system processes (e.g., Walker et al., 2018; Mai et al., 2020). Dai et al. (2017) develop a multi-model process sensitivity analysis method that relies on the integration of variance-based global sensitivity analysis (Sobol', 1993; Saltelli et al., 1999, 2010) with model averaging methods (Draper, 1995; Neuman, 2003; Ye et al., 2008, 2010). The multi-model process sensitivity analysis enables one to jointly address uncertainty in process models as well as parametric uncertainty within

each process model. The approach yields a so-called *process sensitivity index* for each system process. A process associated with a larger value of the index is considered to be more important than other processes.

The multi-model process sensitivity analysis method has been incorporated into a multi-assumption architecture and testbed (MATT) (Walker et al., 2018), and the process sensitivity index has been used in several studies (Walker et al., 2018, 2020; Xu et al., 2019; Yang et al., 2022; Yang and Ye, 2022). The current way of estimating the process sensitivity index is to use brute force Monte Carlo (MC) simulations, which is well known to be computationally expensive. Computational cost for estimating the index still constitutes a serious barrier to increase the index's potential for applications. This is related to the observation that the total number of model simulations for estimating the index corresponds to $n_K n_{\sim K} N^2$ (here, N is the number of MC simulations for a single process model, and n_K and $n_{\sim K}$ are the numbers of plausible models of system process K and other processes, denoted as $\sim K$, respectively). As we illustrate in Section 2, the dependence on N^2 is caused by a nested structure that multiplies N randomly selected parameter samples for a model of process K by N parameter samples of a model (or a set of models) of the remaining processes $\sim K$. Such a dependence is the key reason of the high computational cost. For example, tens of millions of model simulations are needed for a collection of $N = 1,000$ samples randomly selected across the parameter space.

To reduce the computational cost of the brute force MC method, Dai et al. (2017) develop a binning approach that removes the nested structure of parameter sampling. Doing so enables one to decrease the number of model simulations from the order of N^2 to the order

85 of N. These authors show that the results obtained by using 36,000,000 model simulations
86 of the brute force MC method can be obtained by using 16,000 model simulations framed in
87 the context of their binning approach. A major drawback of the binning method of Dai et al.
88 (2017) is that it lacks a mathematically rigorous foundation due to its reliance on an empirical
89 selection of the number of bins as well as of the number of parameter samples across each
90 bin. Tuning these settings by trial and error is time consuming. Additionally, an accurate
91 evaluation of process sensitivity indices is not guaranteed. Therefore, the development of a
92 theoretically rigorous method to overcome these issues and obtain an efficient and reliable
93 estimate of process sensitivity index is still an open research challenge. This is the main
94 objective of the current study.

95 The key to reducing the computational cost associated with the brute force MC method is
96 to remove the nested structure of parameter sampling for two process models (i.e., processes
97 K and $\sim K$). In the context of parameter sensitivity analysis, Ishigami and Homma (1990)
98 and Saltelli et al. (2010) introduce a method to remove a nested sampling structure for two
99 parameters and illustrate its implementation upon relying on a quasi-MC sampling algorithm.
100 Inspired by these studies for parameter sensitivity analysis, we develop here an original
101 method that, for the first time, uses triple sets of parameter samples to remove the nested
102 structure of parameter sampling for estimating the process sensitivity index in a
103 computationally efficient manner and according to a theoretically rigorous approach. It
104 should be noted that the our derivation of the method of triple sets of parameter samples is
105 novel and different from those of Ishigami and Homma (1990) and Saltelli et al. (2010),
106 because our derivation is set in the context of process (not parameter) sensitivity analysis. As

shown in Section 2, the method of triple sets of parameter samples yields a marked reduction of the number of parameter samples (i.e., the number of model executions). i.e., from $n_K n_{\sim K} N^2$ to $2n_K n_{\sim K} N$. We implement the method using a quasi-MC sampling algorithm to achieve an enhanced rate of convergence. As opposed to brute force MC that uses pseudorandom sequence of parameter samples and has a convergence rate of $O(N^{-0.5})$, quasi-MC uses a low-discrepancy sequence (e.g., Halton sequence, Sobol sequence, and Faure sequence) of parameter samples, and is characterized by a rate of convergence close to $O(N^{-1})$. Hereinafter, our method, which uses the triple sets of parameter samples and is implemented through quasi-MC, is referred to as the quasi-MC method for process sensitivity index (or quasi-MC method for brevity).

The quasi-MC method is appraised by way of two synthetic cases associated with groundwater flow and transport to provide a transparent way of analysis. The first synthetic case corresponds to the groundwater reactive transport setting considered by Dai and Ye (2015) and Yang et al. (2022) to analyze model uncertainty in the representation of recharge, geology, and snowmelt processes. The other setting considers zinc sorption in a heterogeneous porous medium and is designed on the basis of the studies of Duan et al. (2020) and Maina et al. (2018). Here, we consider model uncertainty in the way geology and sorption processes are represented. Each of the process models associated with the two cases is characterized by parametric uncertainty (i.e., model parameters are uncertain and treated as random quantities). The process sensitivity index is estimated using the brute force MC, binning, and quasi-MC methods for both cases. Considering the results of the brute force MC method as a reference, the results of the latter two approaches are then assessed in terms of

their accuracy and convergence.

2. Methodology

We provide for completeness a brief introduction of the process sensitivity index and its estimation using the brute force MC and binning methods in Section 2.1. Section 2.2 is devoted to introducing the derivation of the new method that uses triple sets of parameter samples to remove the nested structure of parameter sampling. Section 2.3 discusses the implementation of the method through quasi-MC and its computational cost in terms of the number of parameter samples, which we note being equal to the number of model executions.

2.1. Definition and Estimation of Process Sensitivity Index

We consider a system of interest that is driven by various processes (e.g., $A, B, \cdots$). Each of these processes is prone to multiple representations, i.e., alternative process models. If process A can be characterized through several alternative models (e.g., $M_{A_1}, M_{A_2}, \cdots$) with associated parameters $\theta_{A_1}, \theta_{A_2}, \cdots$, the model set of process A can be denoted as $\mathbf{M}_A\left(\theta_A\right) = \left\{M_{A_1}\left(\theta_{A_1}\right), M_{A_2}\left(\theta_{A_2}\right), \cdots\right\}$. The system model can then be viewed as an integration, $\cup\left(\mathbf{M}_A\left(\theta_A\right), \mathbf{M}_B\left(\theta_B\right), \cdots\right)$, of process models. Dai et al. (2017) propose to evaluate the importance of process A (the collection of the other processes being hereafter denoted via $\sim A$) through the following first-order process sensitivity index of process A

$$PS_A = \frac{V_{\mathbf{M}_A}\left(E_{\mathbf{M}_{\sim A}}\left[\Delta \mid M_A\right]\right)}{V\left(\Delta\right)}, \tag{1}$$

where $V\left(\Delta\right)$ is the total variance of the system output of interest Δ, M_A is a single process model in model set $\mathbf{M}_A$, and $\mathbf{M}_{\sim A}$ denotes the set of alternative models for all of the other processes except process A. The term $E_{\mathbf{M}_{\sim A}}\left[\Delta \mid M_A\right]$ is the mean of Δ taken across all possible models $\mathbf{M}_{\sim A}$ of process $\sim A$, Δ being conditioned on the single model

151 M_A of process A, and $V_{\mathbf{M}_A}\left(E_{\mathbf{M}_{\sim A}}\left[\Delta \mid M_A\right]\right)$ is the variance of $E_{\mathbf{M}_{\sim A}}\left[\Delta \mid M_A\right]$ taken over

152 all possible models in the model set $\mathbf{M}_A$ (Dai et al., 2017). The first-order process

153 sensitivity index (1) measures the average reduction of the variance of Δ when the

154 representation of process A is fixed according to a given (individual) model. An important

155 process yields a large variance reduction, and the process with the largest value of PS_A

156 value is deemed as the most important. Thus, index PS_A can be used to quantitatively rank

157 the importance of multiple system processes.

158 Equation 1 explicitly considers process model uncertainty. It can be expanded to further

159 consider parametric uncertainty. To this end, let us consider a system with multiple processes

160 and denote one process as process A while denoting other processes as process B (which

161 can be either a single process or a combination of multiple processes). Based on the definition

162 of variance and the law of total expectation, the term $V_{\mathbf{M}_A}\left(E_{\mathbf{M}_{\sim A}}\left[\Delta \mid M_A\right]\right)$ appearing in

163 Equation 1 for process A can be rewritten as

$$
\begin{aligned}
V_{\mathbf{M}_A}\left(E_{\mathbf{M}_{\sim A}}\left[\Delta \mid M_A\right]\right) &= E_{\mathbf{M}_A} E_{\boldsymbol{\theta}_A \mid M_A}\left(E_{\mathbf{M}_{\sim A}} E_{\boldsymbol{\theta}_{\sim A} \mid M_{\sim A}}\left[\Delta \mid \theta_A, M_A, \theta_{\sim A}, M_{\sim A}\right]\right)^2 \\
&\quad -\left(E_{\mathbf{M}_A} E_{\boldsymbol{\theta}_A \mid M_A} E_{\mathbf{M}_{\sim A}} E_{\boldsymbol{\theta}_{\sim A} \mid M_{\sim A}}\left[\Delta \mid \theta_A, M_A, \theta_{\sim A}, M_{\sim A}\right]\right)^2 \\
&= E_{\mathbf{M}_A} E_{\boldsymbol{\theta}_A \mid M_A}\left(E_{\mathbf{M}_B} E_{\boldsymbol{\theta}_B \mid M_B}\left[\Delta \mid \theta_A, M_A, \theta_B, M_B\right]\right)^2 \\
&\quad -\left(E_{\mathbf{M}_A} E_{\boldsymbol{\theta}_A \mid M_A} E_{\mathbf{M}_B} E_{\boldsymbol{\theta}_B \mid M_B}\left[\Delta \mid \theta_A, M_A, \theta_B, M_B\right]\right)^2
\end{aligned}
\tag{2}
$$

165 where $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_B$ are the parameter sets corresponding to model M_A and M_B,

166 respectively, and θ_A and θ_B represent a single realization of $\boldsymbol{\theta}_A$ and $\boldsymbol{\theta}_B$, respectively.

167 Applying the model averaging techniques (e.g., Draper, 1995; Neuman, 2003; Ye et al., 2008)

168 to estimate $E_{\mathbf{M}_A}$ and $E_{\mathbf{M}_B}$, the two terms at the right hand side of Equation 2 can be

169 rewritten as

$$E_{\mathbf{M}_A}E_{\boldsymbol{\theta}_A|M_A}\left(E_{\mathbf{M}_B}E_{\boldsymbol{\theta}_B|M_B}\left[\Delta\,|\,\theta_A,M_A,\theta_B,M_B\right]\right)^2$$

$$=\sum_{M_A}E_{\boldsymbol{\theta}_A|M_A}\left(\sum_{M_B}E_{\boldsymbol{\theta}_B|M_B}\left[\Delta\,|\,\theta_A,M_A,\theta_B,M_B\right]P\left(M_B\right)\right)^2P\left(M_A\right), \tag{3}$$

and

$$\left(E_{\mathbf{M}_A}E_{\boldsymbol{\theta}_A|M_A}E_{\mathbf{M}_B}E_{\boldsymbol{\theta}_B|M_B}\left[\Delta\,|\,\theta_A,M_A,\theta_B,M_B\right]\right)^2$$

$$=\left(\sum_{M_A}E_{\boldsymbol{\theta}_A|M_A}\left(\sum_{M_B}E_{\boldsymbol{\theta}_B|M_B}\left[\Delta\,|\,\theta_A,M_A,\theta_B,M_B\right]P\left(M_B\right)\right)P\left(M_A\right)\right)^2, \tag{4}$$

where $P\left(M_A\right)$ is the probability associated with process model M_A in model set $\mathbf{M}_A$ and satisfies the condition $\sum_{l=1}^{n_A}P\left(M_{A_l}\right)=1$ (n_A being the number of alternative process models considered). The same reasoning also holds for probability $P\left(M_B\right)$ related to process B.

The terms related to $E_{\boldsymbol{\theta}_A|M_A}$ and $E_{\boldsymbol{\theta}_B|M_B}$ are estimated using a brute force MC method in Dai et al. (2017), and its pseudo code being given in Figure 1(a). The method is structured across four nested loops, i.e., loops [1] and [3] for process models and loops [2] and [4] for process model parameters. If there are n_A and n_B process models for processes A and B, respectively, and N parameter samples are used for each process model, then the number of model simulations associated with the approach is $n_A\times N\times n_B\times N=n_An_BN^2$. We remark that the N^2 term due to nested loops [2] and [4] is the computational barrier to be removed.

To remove the nested loops [2] and [4] of parameter sampling, Dai et al. (2017) develop a binning method whose pseudo code implementation is shown in Figure 1(b). In the binning method, MC simulations are performed in loops [1] - [3] for paired samples of $\left\{\boldsymbol{\theta}_A,\boldsymbol{\theta}_B\right\}$ that can be sampled in one loop, i.e., loop [3] in Figure 1(b), without using the nested loops [2] and [4] shown in Figure 1(a). After the MC simulations, the parameter space across which

parameter vector $\boldsymbol{\theta}_A$ is defined is divided into multiple bins. Subsequently, $E_{\boldsymbol{\theta}_B|M_B}\left[\Delta \mid \theta_A, M_A, \theta_B, M_B\right]$ is approximated by $E_{\boldsymbol{\theta}_B|M_B}\left[\Delta \mid \theta_A^{bin}, M_A, \theta_B, M_B\right]$ in loop [6] of Figure 1(b) based on the model simulations for the values of $\boldsymbol{\theta}_A$ comprised within each bin. The expectation, $E_{\boldsymbol{\theta}_A|M_A}$, is subsequently approximated in loop [5] by using $E_{\boldsymbol{\theta}_A^{bin}|M_A}$ through averaging over the values of $E_{\boldsymbol{\theta}_B|M_B}\left[\Delta \mid \theta_A^{bin}, M_A, \theta_B, M_B\right]$ obtained for different bins associated with $\boldsymbol{\theta}_A$. The number of model runs for the binning method is $n_A n_B N^{bin}$, where N^{bin} is the number of parameter samples for $\{\boldsymbol{\theta}_A, \boldsymbol{\theta}_B\}$ in Loop [3] of Figure 1(b). Note that N^{bin} is generally larger than N and smaller than N^2.

While the binning method is computationally more efficient than the MC method, the selection of the number of bins is purely empirical, and there are no theoretically firm guidelines to drive it. The bin width in the binning method can be selected according to two algorithms based on the concept of (*a*) equal width or (*b*) equal depth, respectively. The equal width algorithm subdivides the support of parameter variability into bins of equal width. When the number of parameter realizations is small and/or the bin width is narrow, it may happen that a bin is empty (i.e., no parameter realization is comprised in the bin) or thin (i.e., only very few, e.g., one or two, parameter samples can be found in it). The presence of empty and/or thin bins leads to inaccurate estimates of the expectations required for the evaluation of the sensitivity indices. Otherwise, the equal depth algorithm subdivides the support of parameter variability into bins containing (approximately) the same number of random samples. We remark that the number of parameter realizations and/or the number of bins must be adjusted empirically for both algorithms, a procedure which can be markedly time consuming.

Loop [1] over models of process A ($\mathbf{M}_A$, the set of process model M_A) (a)
Loop [2] over parameter realizations θ_A of model M_A
Loop [3] over models of process B ($\mathbf{M}_B$, the set of process model M_B)
Loop [4] over parameter realizations θ_B of model M_B
Compute $\Delta \mid \theta_A, M_A, \theta_B, M_B$
End loop [4]
Compute $E_{\boldsymbol{\theta}_B \mid M_B}[\Delta \mid \theta_A, M_A, \theta_B, M_B]$
End loop [3]
Compute $(E_{\mathbf{M}_B} E_{\boldsymbol{\theta}_B \mid M_B}[\Delta \mid \theta_A, M_A, \theta_B, M_B])^2$ and $E_{\mathbf{M}_B} E_{\boldsymbol{\theta}_B \mid M_B}[\Delta \mid \theta_A, M_A, \theta_B, M_B]$ using model averaging
End loop [2]
Compute $E_{\boldsymbol{\theta}_A \mid M_A}(E_{\mathbf{M}_B} E_{\boldsymbol{\theta}_B \mid M_B}[\Delta \mid \theta_A, M_A, \theta_B, M_B])^2$ and $E_{\boldsymbol{\theta}_A \mid M_A} E_{\mathbf{M}_B} E_{\boldsymbol{\theta}_B \mid M_B}[\Delta \mid \theta_A, M_A, \theta_B, M_B]$
End loop [1]
Compute $E_{\mathbf{M}_A} E_{\boldsymbol{\theta}_A \mid M_A}(E_{\mathbf{M}_B} E_{\boldsymbol{\theta}_B \mid M_B}[\Delta \mid \theta_A, M_A, \theta_B, M_B])^2$ and $(E_{\mathbf{M}_A} E_{\boldsymbol{\theta}_A \mid M_A} E_{\mathbf{M}_B} E_{\boldsymbol{\theta}_B \mid M_B}[\Delta \mid \theta_A, M_A, \theta_B, M_B])^2$ using model averaging

Loop [1] over models of process A ($\mathbf{M}_A$, the set of process model M_A) (b)
Loop [2] over models of process B ($\mathbf{M}_B$, the set of process model M_B)
Loop [3] over parameter realizations θ_A of model M_A and θ_B of model M_B
Compute $\Delta \mid \theta_A, M_A, \theta_B, M_B$
End loop [3]
End Loop [2]
End Loop [1]
Loop [4] over models of process A ($\mathbf{M}_A$, the set of process model M_A)
Loop [5] over parameter bins θ_A^{bin} of model M_A
Loop [6] over models of process B ($\mathbf{M}_B$, the set of process model M_B)
Compute $E_{\boldsymbol{\theta}_B \mid M_B}[\Delta \mid \theta_A^{bin}, M_A, \theta_B, M_B]$ for all θ_B realizations of each bin
End loop [6]
Compute $(E_{\mathbf{M}_B} E_{\boldsymbol{\theta}_B \mid M_B}[\Delta \mid \theta_A^{bin}, M_A, \theta_B, M_B])^2$ and $E_{\mathbf{M}_B} E_{\boldsymbol{\theta}_B \mid M_B}[\Delta \mid \theta_A^{bin}, M_A, \theta_B, M_B]$ using model averaging
End loop [5]
Compute $E_{\boldsymbol{\theta}_A^{bin} \mid M_A}(E_{\mathbf{M}_B} E_{\boldsymbol{\theta}_B \mid M_B}[\Delta \mid \theta_A^{bin}, M_A, \theta_B, M_B])^2$ and $E_{\boldsymbol{\theta}_A^{bin} \mid M_A} E_{\mathbf{M}_B} E_{\boldsymbol{\theta}_B \mid M_B}[\Delta \mid \theta_A^{bin}, M_A, \theta_B, M_B]$
End loop [4]
Compute $E_{\mathbf{M}_A} E_{\boldsymbol{\theta}_A^{bin} \mid M_A}(E_{\mathbf{M}_B} E_{\boldsymbol{\theta}_B \mid M_B}[\Delta \mid \theta_A^{bin}, M_A, \theta_B, M_B])^2$ and $(E_{\mathbf{M}_A} E_{\boldsymbol{\theta}_A^{bin} \mid M_A} E_{\mathbf{M}_B} E_{\boldsymbol{\theta}_B \mid M_B}[\Delta \mid \theta_A^{bin}, M_A, \theta_B, M_B])^2$ using model averaging

Figure 1. Pseudo codes for evaluating Equations 3 and 4 for sensitivity index of process A using model averaging and (a) the brute force MC method and (b) the binning method.

2.2. Using Triple Sets of Parameter Samples for Estimating Process Sensitivity Index

The nested sampling structure can be removed in a theoretically rigorous way by using the method of triple sets of parameter samples. We note that our derivation, while being in spirit similar to the one given by Saltelli et al. (2010), is set in the context of process uncertainty. Thus, it markedly differs from the work of Saltelli et al. (2010). Starting from Equation 3 and expanding the squared term yields

$$
\begin{aligned}
&E_{\mathbf{M}_A}E_{\boldsymbol{\theta}_A|M_A}\left(E_{\mathbf{M}_B}E_{\boldsymbol{\theta}_B|M_B}\left[\Delta\,|\,\theta_A,M_A,\theta_B,M_B\right]\right)^2\\
&=E_{\mathbf{M}_A}E_{\boldsymbol{\theta}_A|M_A}\left(\begin{array}{l}\displaystyle\sum_{j=1}^{n_B}\left(E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_A,M_A,\theta_{B_j},M_{B_j}\right]\right)^2P\left(M_{B_j}\right)^2\\ \displaystyle+\sum_{j=1}^{n_B}\sum_{\substack{l=1\\l\neq j}}^{n_B}\left(\begin{array}{l}E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_A,M_A,\theta_{B_j},M_{B_j}\right]\\ \times E_{\boldsymbol{\theta}_{B_l}|M_{B_l}}\left[\Delta\,|\,\theta_A,M_A,\theta_{B_l},M_{B_l}\right]\end{array}\right)P\left(M_{B_j}\right)P\left(M_{B_l}\right)\end{array}\right)\\
&=\sum_{i=1}^{n_A}E_{\boldsymbol{\theta}_{A_i}|M_{A_i}}\left(\begin{array}{l}\displaystyle\sum_{j=1}^{n_B}\left(E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\right)^2P\left(M_{B_j}\right)^2\\ \displaystyle+\sum_{j=1}^{n_B}\sum_{\substack{l=1\\l\neq j}}^{n_B}\left(\begin{array}{l}E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\\ \times E_{\boldsymbol{\theta}_{B_l}|M_{B_l}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_l},M_{B_l}\right]\end{array}\right)P\left(M_{B_j}\right)P\left(M_{B_l}\right)\end{array}\right)P\left(M_{A_i}\right)\\
&=\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}E_{\boldsymbol{\theta}_{A_i}|M_{A_i}}\left(E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\right)^2P\left(M_{B_j}\right)^2P\left(M_{A_i}\right)\\
&+\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}\sum_{\substack{l=1\\l\neq j}}^{n_B}E_{\boldsymbol{\theta}_{A_i}|M_{A_i}}\left(\begin{array}{l}E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\\ \times E_{\boldsymbol{\theta}_{B_l}|M_{B_l}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_l},M_{B_l}\right]\end{array}\right)P\left(M_{B_j}\right)P\left(M_{B_l}\right)P\left(M_{A_i}\right)
\end{aligned}
\qquad (5)
$$

where $\boldsymbol{\theta}_{A_i}$, $\boldsymbol{\theta}_{B_j}$, and $\boldsymbol{\theta}_{B_l}$ are the parameter sets of models M_{A_i}, M_{B_j}, and M_{B_l}, respectively, and M_{B_j} and M_{B_l} are two individual models of process B.

The next step is to evaluate the expectation terms appearing at the right hand side of Equation 5, i.e., $E_{\boldsymbol{\theta}_{A_i}|M_{A_i}}\left(E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\right)^2$ and $E_{\boldsymbol{\theta}_{A_i}|M_{A_i}}\left(E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\times E_{\boldsymbol{\theta}_{B_l}|M_{B_l}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_l},M_{B_l}\right]\right)$. We start by writing

$$E_{\boldsymbol{\theta}_{A_i}|M_{A_i}}\left(E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\right)^2 = \int_{\boldsymbol{\theta}_{A_i}}\left(E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\right)^2 p\left(\theta_{A_i}|M_{A_i}\right)d\theta_{A_i} \quad (6)$$

where $p\left(\theta_{A_i}|M_{A_i}\right)$ is the probability density function for parameter θ_{A_i} conditional on process model M_{A_i}. Following Saltelli et al. (2010), the term $\left(E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\right)^2$ in Equation 6 can be expressed as

$$\begin{aligned}&\left(E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\right)^2\\ &= E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\times E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right], \quad (7)\\ &= \iint_{\boldsymbol{\theta}_{B_j},\boldsymbol{\theta}_{B_j}'}\left(\begin{array}{c}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j}',M_{B_j}\right]\\ \times\left(p\left(\theta_{B_j}|M_{B_j}\right)d\theta_{B_j}\right)\left(p\left(\theta_{B_j}'|M_{B_j}\right)d\theta_{B_j}'\right)\end{array}\right)\end{aligned}$$

where $p\left(\theta_{B_j}|M_{B_j}\right)$ and $p\left(\theta_{B_j}'|M_{B_j}\right)$ are the probability density functions for parameter θ_{B_j} (conditional on M_{B_j}) and θ_{B_j}' (conditional on M_{B_l}), respectively. The double integral appearing in Equation 7 accounts for duplicate parameter sets of $\left\{\boldsymbol{\theta}_{B_j},\boldsymbol{\theta}_{B_j}'\right\}$, where these two parameter sets are all for model M_{B_j}. Parameter sets $\boldsymbol{\theta}_{B_j}$ and $\boldsymbol{\theta}_{B_j}'$ are sampled from $p\left(\theta_{B_j}|M_{B_j}\right)$ and $p\left(\theta_{B_j}'|M_{B_j}\right)$ (that is the same as $p\left(\theta_{B_j}|M_{B_j}\right)$), respectively. Substituting Equation 7 into 6 leads to

$$\begin{aligned}&E_{\boldsymbol{\theta}_{A_i}|M_{A_i}}\left(E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\right)^2\\ &= \int_{\boldsymbol{\theta}_{A_i}}\left(\iint_{\boldsymbol{\theta}_{B_j},\boldsymbol{\theta}_{B_j}'}\begin{array}{c}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j}',M_{B_j}\right]\\ \times\left(p\left(\theta_{B_j}|M_{B_j}\right)d\theta_{B_j}\right)\left(p\left(\theta_{B_j}'|M_{B_j}\right)d\theta_{B_j}'\right)\end{array}\right)p\left(\theta_{A_i}|M_{A_i}\right)d\theta_{A_i}, \quad (8)\\ &= \iiint_{\boldsymbol{\theta}_{A_i},\boldsymbol{\theta}_{B_j},\boldsymbol{\theta}_{B_j}'}\begin{array}{c}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j}',M_{B_j}\right]\\ \times\left(p\left(\theta_{B_j}|M_{B_j}\right)d\theta_{B_j}\right)\left(p\left(\theta_{B_j}'|M_{B_j}\right)d\theta_{B_j}'\right)\left(p\left(\theta_{A_i}|M_{A_i}\right)d\theta_{A_i}\right)\end{array}\end{aligned}$$

where the triple integral $\iiint_{\boldsymbol{\theta}_{A_i},\boldsymbol{\theta}_{B_j},\boldsymbol{\theta}_{B_j}'}$ denotes integration across the support of parameter set $\left\{\boldsymbol{\theta}_{A_i},\boldsymbol{\theta}_{B_j},\boldsymbol{\theta}_{B_j}'\right\}$. The term $(\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j})$ indicates that the system model output, Δ, is

conditioned on the system model formed with the i^{th} model of process A, the j^{th} model of process B, and their corresponding parameter realizations θ_{A_i} and θ_{B_j}. The only difference between $\left[\Delta \mid \theta_{A_i}, M_{A_i}, \theta_{B_j}, M_{B_j}\right]$ and $\left[\Delta \mid \theta_{A_i}, M_{A_i}, \theta_{B_j}', M_{B_j}\right]$ is that the latter term is evaluated with θ_{B_j}', which is a random sample for the parameters of the j^{th} model of process B.

We denote the product, $\left[\Delta \mid \theta_{A_i}, M_{A_i}, \theta_{B_j}, M_{B_j}\right]\left[\Delta \mid \theta_{A_i}, M_{A_i}, \theta_{B_j}', M_{B_j}\right]$, of two model outputs as

$$f\left(\theta_{A_i}, \theta_{B_j}, \theta_{B_j}' \mid M_{A_i}, M_{B_j}\right) = \left[\Delta \mid \theta_{A_i}, M_{A_i}, \theta_{B_j}, M_{B_j}\right]\left[\Delta \mid \theta_{A_i}, M_{A_i}, \theta_{B_j}', M_{B_j}\right], \tag{9}$$

Equation 8 then becomes

$$\begin{aligned} & E_{\boldsymbol{\theta}_{A_i} \mid M_{A_i}}\left(E_{\boldsymbol{\theta}_{B_j} \mid M_{B_j}}\left[\Delta \mid \theta_{A_i}, M_{A_i}, \theta_{B_j}, M_{B_j}\right]\right)^2 \\ & = \iiint_{\boldsymbol{\theta}_{A_i}, \boldsymbol{\theta}_{B_j}, \boldsymbol{\theta}_{B_j}'} \left(\begin{array}{l} f\left(\theta_{A_i}, \theta_{B_j}, \theta_{B_j}' \mid M_{A_i}, M_{B_j}\right) \\ \times\left(p\left(\theta_{B_j} \mid M_{B_j}\right) d\theta_{B_j}\right)\left(p\left(\theta_{B_j}' \mid M_{B_j}\right) d\theta_{B_j}'\right)\left(p\left(\theta_{A_i} \mid M_{A_i}\right) d\theta_{A_i}\right) \end{array} \right). \end{aligned} \tag{10}$$

Since the integral in Equation 10 corresponds to the expectation of $f\left(\theta_{A_i}, \theta_{B_j}, \theta_{B_j}' \mid M_{A_i}, M_{B_j}\right)$, Equation 10 can be rewritten as

$$\begin{aligned} & E_{\boldsymbol{\theta}_{A_i} \mid M_{A_i}}\left(E_{\boldsymbol{\theta}_{B_j} \mid M_{B_j}}\left[\Delta \mid \theta_{A_i}, M_{A_i}, \theta_{B_j}, M_{B_j}\right]\right)^2 \\ & = E_{\boldsymbol{\theta}_{A_i}, \boldsymbol{\theta}_{B_j}, \boldsymbol{\theta}_{B_j}'}\left(f\left(\theta_{A_i}, \theta_{B_j}, \theta_{B_j}' \mid M_{A_i}, M_{B_j}\right)\right) \end{aligned}. \tag{11}$$

This is a key equation of our new method, because it transforms the nested expectation $E_{\boldsymbol{\theta}_{A_i} \mid M_{A_i}}\left(E_{\boldsymbol{\theta}_{B_j} \mid M_{B_j}}\left[\Delta \mid \theta_{A_i}, M_{A_i}, \theta_{B_j}, M_{B_j}\right]\right)^2$ into a single expectation $E_{\boldsymbol{\theta}_{A_i}, \boldsymbol{\theta}_{B_j}, \boldsymbol{\theta}_{B_j}'}\left(f\left(\theta_{A_i}, \theta_{B_j}, \theta_{B_j}' \mid M_{A_i}, M_{B_j}\right)\right)$. In other words, the nested structure of parameter sampling is removed in a rigorous way by using the triple sets of parameter samples, $\left\{\boldsymbol{\theta}_{A_i}, \boldsymbol{\theta}_{B_j}, \boldsymbol{\theta}_{B_j}'\right\}$, and the parameter samples can be generated separately (rather than requiring a nested loop), as we discuss in the following.

Similar to the derivation of Equation 11, the second term of Equation 5 is expressed as

$$
\begin{aligned}
&E_{\boldsymbol{\theta}_{A_i}|M_{A_i}}\left(E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\times E_{\boldsymbol{\theta}_{B_l}|M_{B_l}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_l},M_{B_l}\right]\right)\\
&=\int_{\boldsymbol{\theta}_{A_i}}\left(\begin{array}{l}\left(\int_{\boldsymbol{\theta}_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\left(p\left(\theta_{B_j}\,|\,M_{B_j}\right)d\theta_{B_j}\right)\right)\\ \times\left(\int_{\boldsymbol{\theta}_{B_l}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_l},M_{B_l}\right]\left(p\left(\theta_{B_l}\,|\,M_{B_l}\right)d\theta_{B_l}\right)\right)\end{array}\right)p\left(\theta_{A_i}\,|\,M_{A_i}\right)d\theta_{A_i}\\
&=\int_{\boldsymbol{\theta}_{A_i}}\left(\iint_{\boldsymbol{\theta}_{B_j},\boldsymbol{\theta}_{B_l}}\left(\begin{array}{l}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_l},M_{B_l}\right]\\ \times\left(p\left(\theta_{B_j}\,|\,M_{B_j}\right)d\theta_{B_j}\right)\left(p\left(\theta_{B_l}\,|\,M_{B_l}\right)d\theta_{B_l}\right)\end{array}\right)\right)p\left(\theta_{A_i}\,|\,M_{A_i}\right)d\theta_{A_i}\\
&=\iiint_{\boldsymbol{\theta}_{A_i},\boldsymbol{\theta}_{B_j},\boldsymbol{\theta}_{B_l}}\left(\begin{array}{l}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_l},M_{B_l}\right]\left(p\left(\theta_{B_j}\,|\,M_{B_j}\right)d\theta_{B_j}\right)\\ \times\left(p\left(\theta_{B_l}\,|\,M_{B_l}\right)d\theta_{B_l}\right)\left(p\left(\theta_{A_i}\,|\,M_{A_i}\right)d\theta_{A_i}\right)\end{array}\right)\\
&=\iiint_{\boldsymbol{\theta}_{A_i},\boldsymbol{\theta}_{B_j},\boldsymbol{\theta}_{B_l}}\left(\begin{array}{l}g\left(\theta_{A_i},\theta_{B_j},\theta_{B_l}\,|\,M_{A_i},M_{B_j},M_{B_l}\right)\left(p\left(\theta_{B_j}\,|\,M_{B_j}\right)d\theta_{B_j}\right)\\ \times\left(p\left(\theta_{B_l}\,|\,M_{B_l}\right)d\theta_{B_l}\right)\left(p\left(\theta_{A_i}\,|\,M_{A_i}\right)d\theta_{A_i}\right)\end{array}\right)\\
&=E_{\boldsymbol{\theta}_{A_i},\boldsymbol{\theta}_{B_j},\boldsymbol{\theta}_{B_l}}\left(g\left(\theta_{A_i},\theta_{B_j},\theta_{B_l}\,|\,M_{A_i},M_{B_j},M_{B_l}\right)\right) \qquad , \qquad (12)
\end{aligned}
$$

where $g\left(\theta_{A_i},\theta_{B_j},\theta_{B_l}\,|\,M_{A_i},M_{B_j},M_{B_l}\right)=\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_l},M_{B_l}\right]$.

This equation enables one to reduce the nested expectation $E_{\boldsymbol{\theta}_{A_i}|M_{A_i}}\left(E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\times E_{\boldsymbol{\theta}_{B_l}|M_{B_l}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_l},M_{B_l}\right]\right)$ to a single expectation, i.e., $E_{\boldsymbol{\theta}_{A_i},\boldsymbol{\theta}_{B_j},\boldsymbol{\theta}_{B_l}}\left(g\left(\theta_{A_i},\theta_{B_j},\theta_{B_l}\,|\,M_{A_i},M_{B_j},M_{B_l}\right)\right)$. This is another example of using triple sets of parameter samples to remove the nested structure of parameter sampling.

2.3.Quasi-MC Implementation

The expectations in Equations 11 and 12 can be estimated using either brute force MC with pseudorandom samples or quasi-MC with a low-discrepancy sequence (also known as quasi-random sequence, e.g., Halton sequence, Sobol sequence, and Faure sequence) of parameter samples. Quasi-random sequences can cover the parameter space more quickly and uniformly than pseudorandom samples (Caflisch, 1998), thus achieving a faster convergence rate (Hou et al., 2019). In this study, we used the Sobol sequence for the estimation of

expectations, and rewrote Equation 12 as

$$E_{\boldsymbol{\theta}_{A_i}|M_{A_i}}\left(E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\right)^2=\frac{1}{N}\sum_{r=1}^{N}f\left(\theta_{A_i}^r,\theta_{B_j}^r,\theta_{B_j}^r{}'\,|\,M_{A_i},M_{B_j}\right), \tag{13}$$

and

$$\begin{aligned}&E_{\boldsymbol{\theta}_{A_i}|M_{A_i}}\left(E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\times E_{\boldsymbol{\theta}_{B_l}|M_{B_l}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_l},M_{B_l}\right]\right)\\&=\frac{1}{N}\sum_{r=1}^{N}g\left(\theta_{A_i}^r,\theta_{B_j}^r,\theta_{B_l}^r\,|\,M_{A_i},M_{B_j},M_{B_l}\right)\end{aligned}, \tag{14}$$

where N is the number of Sobol sequence parameters. As a result, the first and second terms on the right side of Equation 5 become

$$\begin{aligned}&\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}E_{\boldsymbol{\theta}_{A_i}|M_{A_i}}\left(E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\right)^2P\left(M_{B_j}\right)^2P\left(M_{A_i}\right)\\&=\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}\frac{1}{N}\sum_{r=1}^{N}f\left(\theta_{A_i}^r,\theta_{B_j}^r,\theta_{B_j}^r{}'\,|\,M_{A_i},M_{B_j}\right)P\left(M_{B_j}\right)^2P\left(M_{A_i}\right)\end{aligned}, \tag{15}$$

and

$$\begin{aligned}&\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}\sum_{\substack{l=1\\l\neq j}}^{n_B}E_{\boldsymbol{\theta}_{A_i}|M_{A_i}}\left(\begin{array}{l}E_{\boldsymbol{\theta}_{B_j}|M_{B_j}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_j},M_{B_j}\right]\\\times E_{\boldsymbol{\theta}_{B_l}|M_{B_l}}\left[\Delta\,|\,\theta_{A_i},M_{A_i},\theta_{B_l},M_{B_l}\right]\end{array}\right)P\left(M_{B_j}\right)P\left(M_{B_l}\right)P\left(M_{A_i}\right)\\&=\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}\sum_{\substack{l=1\\l\neq j}}^{n_B}\frac{1}{N}\sum_{r=1}^{N}g\left(\theta_{A_i}^r,\theta_{B_j}^r,\theta_{B_l}^r\,|\,M_{A_i},M_{B_j},M_{B_l}\right)P\left(M_{B_j}\right)P\left(M_{B_l}\right)P\left(M_{A_i}\right)\end{aligned}. \tag{16}$$

Therefore, Equation 3 can be estimated via

$$\begin{aligned}&E_{\mathbf{M}_A}E_{\boldsymbol{\theta}_A|M_A}\left(E_{\mathbf{M}_B}E_{\boldsymbol{\theta}_B|M_B}\left[\Delta\,|\,\theta_A,M_A,\theta_B,M_B\right]\right)^2\\&=\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}\frac{1}{N}\sum_{r=1}^{N}f\left(\theta_{A_i}^r,\theta_{B_j}^r,\theta_{B_j}^r{}'\,|\,M_{A_i},M_{B_j}\right)P\left(M_{B_j}\right)^2P\left(M_{A_i}\right)\\&+\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}\sum_{\substack{l=1\\l\neq j}}^{n_B}\frac{1}{N}\sum_{r=1}^{N}g\left(\theta_{A_i}^r,\theta_{B_j}^r,\theta_{B_l}^r\,|\,M_{A_i},M_{B_j},M_{B_l}\right)P\left(M_{B_j}\right)P\left(M_{B_l}\right)P\left(M_{A_i}\right)\end{aligned}. \tag{17}$$

Estimation of the terms included in Equation 4 is relatively straightforward. It essentially corresponds to the total expectation of Δ over all of the models and model parameters, according to

$$
\begin{aligned}
&\left(E_{\mathbf{M}_A}E_{\boldsymbol{\theta}_A|M_A}E_{\mathbf{M}_B}E_{\boldsymbol{\theta}_B|M_B}\left[\Delta\,|\,\theta_A,M_A,\theta_B,M_B\right]\right)^2\\
&=\left(\sum_{\mathbf{M}_A}E_{\boldsymbol{\theta}_A|M_A}\left(\sum_{\mathbf{M}_B}E_{\boldsymbol{\theta}_B|M_B}\left[\Delta\,|\,\theta_A,M_A,\theta_B,M_B\right]P\left(M_B\right)\right)P\left(M_A\right)\right)^2\\
&=\left(\sum_{\mathbf{M}_A}\sum_{\mathbf{M}_B}E_{\boldsymbol{\theta}_A|M_A}E_{\boldsymbol{\theta}_B|M_B}\left[\Delta\,|\,\theta_A,M_A,\theta_B,M_B\right]P\left(M_B\right)P\left(M_A\right)\right)^2\\
&=\left(\sum_{\mathbf{M}_A}\sum_{\mathbf{M}_B}E_{\boldsymbol{\theta}_A,\boldsymbol{\theta}_B|M_A,M_B}\left[\Delta\,|\,\theta_A,M_A,\theta_B,M_B\right]P\left(M_B\right)P\left(M_A\right)\right)^2 \qquad (18)\\
&=\left(\sum_{\mathbf{M}_A}\sum_{\mathbf{M}_B}E\left[\Delta\,|\,M_A,M_B\right]P\left(M_B\right)P\left(M_A\right)\right)^2\\
&=\left(E\left(\Delta\right)\right)^2\\
&=\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}\frac{1}{N}\sum_{r=1}^{N}\left[\Delta\,|\,\theta_{A_i}^r,M_{A_i},\theta_{B_j}^r,M_{B_j}\right]P\left(M_{B_j}\right)P\left(M_{A_i}\right)
\end{aligned}
$$

Therefore, the process sensitivity index defined through Equation 2 can be estimated via

$$
\begin{aligned}
&V_{\mathbf{M}_A}\left(E_{\mathbf{M}_B}\left[\Delta\,|\,M_A\right]\right)\\
&=E_{\mathbf{M}_A}E_{\boldsymbol{\theta}_A|M_A}\left(E_{\mathbf{M}_B}E_{\boldsymbol{\theta}_B|M_{\sim K}}\left[\Delta\,|\,\theta_A,M_A,\theta_B,M_B\right]\right)^2\\
&-\left(E_{\mathbf{M}_A}E_{\boldsymbol{\theta}_A|M_A}E_{\mathbf{M}_B}E_{\boldsymbol{\theta}_B|M_{\sim K}}\left[\Delta\,|\,\theta_A,M_A,\theta_B,M_B\right]\right)^2\\
&=\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}\frac{1}{N}\sum_{r=1}^{N}f\left(\theta_{A_i}^r,\theta_{B_j}^r,\theta_{B_j}^{r\prime}\,|\,M_{A_i},M_{B_j}\right)P\left(M_{B_j}\right)^2P\left(M_{A_i}\right) \qquad (19)\\
&+\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}\sum_{\substack{l=1\\l\neq j}}^{n_B}\frac{1}{N}\sum_{r=1}^{N}g\left(\theta_{A_i}^r,\theta_{B_j}^r,\theta_{B_l}^r\,|\,M_{A_i},M_{B_j},M_{B_l}\right)P\left(M_{B_j}\right)P\left(M_{B_l}\right)P\left(M_{A_i}\right)\\
&-\left(\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}\frac{1}{N}\sum_{r=1}^{N}\left[\Delta\,|\,\theta_{A_i}^r,M_{A_i},\theta_{B_j}^r,M_{B_j}\right]P\left(M_{B_j}\right)P\left(M_{A_i}\right)\right)^2
\end{aligned}
$$

It should be noted that, since the quasi-MC algorithm is only applied to the expectation terms, model weights (e.g., $P(M_A)$ and $P(M_B)$) do not affect computational performance for the estimation of such terms.

The pseudo code for the evaluation of Equation 19 is given in Figure 2. The figure shows that, while the nested loops [1] and [2] for process models remain, the new quasi-MC method removes the nested structure of parameter sampling, and parameter sampling is performed in

298 only one loop, i.e., loop [3]. Three sets of parameter samples are needed in this loop. These
299 correspond to N realizations of $\theta_{A_i}^r$ for process model M_{A_i}, N realizations of $\theta_{B_j}^r$ for
300 process model M_{B_j}, and N realizations of $\theta_{B_j}^r{}'$ for process model M_{B_j}. The quantity
301 $\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^r, M_{B_j}$ is evaluated by running the system model (denoted as $M_{A_i} \cup M_{B_j}$) for
302 N times based on the N realizations of $\theta_{A_i}^r$ and $\theta_{B_j}^r$. Evaluation of $\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^r{}', M_{B_j}$ is
303 based on running the same system model ($M_{A_i} \cup M_{B_j}$) for N times, considering the N
304 realizations of $\theta_{A_i}^r$ and $\theta_{B_j}^r{}'$. Therefore, the total number of model simulations associated
305 with the quasi-MC method is $2n_A n_B N$, which is substantially smaller than the corresponding
306 number of simulations of the MC method (i.e., $n_A n_B N^2$).
307 The formulation of the quasi-MC method is rigorous and general. It can be applied to
308 complex problems with a large number of processes and process model parameters. It is
309 noted that the computational cost of the quasi-MC method is independent of the number of
310 processes. The pseudo code in Figure 2 shows that Loop [1] is for one process and Loop [2]
311 is for other processes (they are denoted as process *B* in Figure 2). In other words, the quasi-
312 MC method only needs two loops regardless of the number of processes. We note that, while
313 the computational cost of our quasi-MC method for process model parameters can in
314 principle be reduced upon development of more efficient quasi-MC sampling algorithms, cost
315 reduction for model parameters is beyond the scope of this study.
316
317

Loop [1] over models of process A ($\mathbf{M}_A$, the set of process models M_{A_i})

Loop [2] over model of process B ($\mathbf{M}_B$, the set of process models M_{B_j})

Loop [3] over parameter realizations $\theta_{A_i}^r$ of model M_{A_i}, $\theta_{B_j}^r$ of model M_{B_j}, and realizations $\theta_{B_j}^{r\prime}$ of model M_{B_j}

Compute $\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^r, M_{B_j}$ and $\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^{r\prime}, M_{B_j}$

End loop [3]

Compute $\frac{1}{N}\sum_{r=1}^{N}\left[\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^r, M_{B_j}\right]$,

$\frac{1}{N}\sum_{r=1}^{N}\left[\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^r, M_{B_j}\right]\left[\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^{r\prime}, M_{B_j}\right]$, and

$\frac{1}{N}\sum_{r=1}^{N}\left[\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^r, M_{B_j}\right]\left[\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_l}^r, M_{B_l}\right]$

End loop [2]

Compute $\sum_{j=1}^{n_B}\frac{1}{N}\sum_{r=1}^{N}\left[\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^r, M_{B_j}\right]P\left(M_{B_j}\right)$,

$\sum_{j=1}^{n_B}\frac{1}{N}\sum_{r=1}^{N}\left[\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^r, M_{B_j}\right]\left[\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^{r\prime}, M_{B_j}\right]P\left(M_{B_j}\right)^2$, and

$\sum_{j=1}^{n_B}\sum_{\substack{l=1\\ l\neq j}}^{n_B}\frac{1}{N}\sum_{r=1}^{N}\left[\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^r, M_{B_j}\right]\left[\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_l}^r, M_{B_l}\right]P\left(M_{B_j}\right)P\left(M_{B_l}\right)$

End loop [1]

Compute $\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}\frac{1}{N}\sum_{r=1}^{N}\left[\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^r, M_{B_j}\right]P\left(M_{B_j}\right)P\left(M_{A_i}\right)$,

$\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}\frac{1}{N}\sum_{r=1}^{N}\left[\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^r, M_{B_j}\right]\left[\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^{r\prime}, M_{B_j}\right]P\left(M_{B_j}\right)^2P\left(M_{A_i}\right)$

and $\sum_{i=1}^{n_A}\sum_{j=1}^{n_B}\sum_{\substack{l=1\\ l\neq j}}^{n_B}\frac{1}{N}\sum_{r=1}^{N}\left[\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_j}^r, M_{B_j}\right]\left[\Delta \mid \theta_{A_i}^r, M_{A_i}, \theta_{B_l}^r, M_{B_l}\right]P\left(M_{B_j}\right)P\left(M_{B_l}\right)P\left(M_{A_i}\right)$

Figure 2. Pseudo code for estimating Equation 19 for the sensitivity index of process A using the quasi-MC method.

3. Assessment of the Quasi-MC Method

We compare our quasi-MC method against the brute force MC and the binning method of Dai et al. (2017) upon performing process sensitivity analysis associated with two synthetic cases encompassing groundwater flow and solute transport. This yields transparent comparative analyses to assess the quality of the performance of the quasi-MC approach we

develop. The first case is related to a one-dimensional setting where reactive transport in a river-aquifer system takes place. Three key processes, a set of eight alternative system models, and eight uncertain parameters are considered. We then consider a scenario associated with a two-dimensional heterogeneous porous medium across which we simulate migration of zinc in the presence of sorption. In this case, the effects of two processes, a set of six alternative system models, and five uncertain parameters are analyzed. Numerical results associated with the classical MC method whose implementation is based on a markedly large number of forward model simulations are employed as a reference against which the performance of the quasi-MC and binning methods are assessed.

3.1. One-dimensional reactive transport in a river-groundwater system

We consider a one-dimensional scenario that is adapted from the synthetic setting introduced by Dai and Ye (2015) and Yang et al. (2022) and sketched in Figure 3. Steady-state flow takes place across the unconfined aquifer of length L=10，000 m depicted in Figure 3. A constant precipitation $PP = 1,524$ mm/year is set across the entire domain. A constant head $h_1 = 330$ m is set on the left boundary. A constant head h_2 determined by a snowmelt process, whose characterization we discuss below, is set at the river boundary. A continuous contaminant source is placed at the center of the domain ($x = 5,000$ m). The source involves chain reactions of the following five chemical species: perchloroethene (PCE), trichloroethene (TCE), dichloroethane (DCE), vinyl chloride (VC) and ethene (ETH). Additional details of the chain reactions considered are offered by Dai and Ye (2015) and Aziz et al (2000). An analytical solution for the evolution of the considered chemical species across the domain is given by Sun et al. (1999).

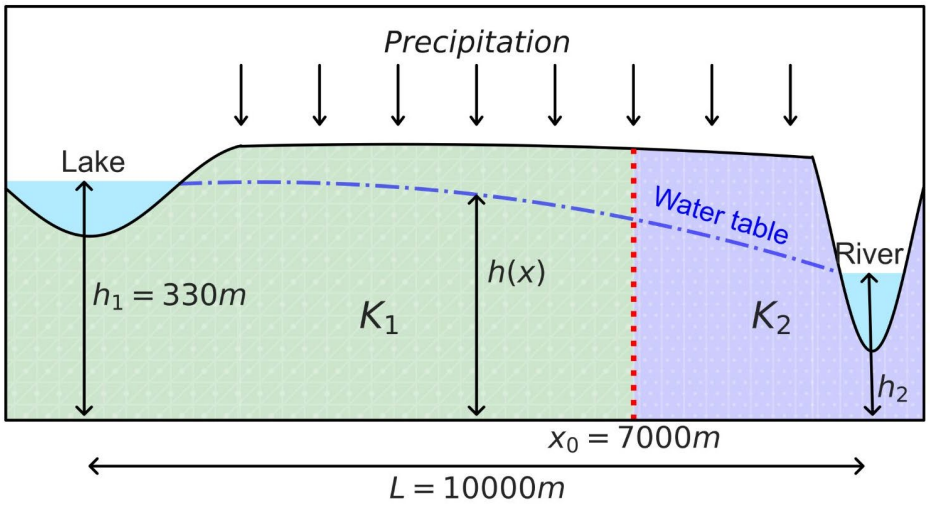


Figure 3. Sketch of the one-dimensional domain considered, including the key geometrical features and boundary conditions. A continuous contaminant source is located at the domain center. The demarcation of the two regions with differing hydraulic conductivity values considered in one of the representations of the geological setting is highlighted (vertical red dash line).

Two alternative mathematical models (hereafter termed R_1 and R_2) are considered to depict the recharge process, i.e.,

$$\begin{aligned} R_1 &: w = a\left(PP - 355.6\right)^{0.50} \\ R_2 &: w = b\left(PP - 399.8\right) \end{aligned} \tag{20}$$

Parameter a of process model R_1 and parameter b of process model R_2 are two arbitrary coefficients for the recharge estimation. Here, these are assumed to be characterized by normal and uniform distributions, respectively. The distributions of these two parameters and other parameters discussed below are listed in Table 1, for completeness. Values of deterministic parameters discussed below are listed in Table 2.

Two alternative representations (hereafter termed G_1 and G_2) of the setting stemming from the action of geological processes are considered, i.e.,

$$\begin{aligned} &G_1 : \text{aquifer is homogeneous} \\ &G_2 : K = \begin{cases} K_1 \text{ for } x < 7{,}000 \\ K_2 \text{ for } x \geq 7{,}000 \end{cases} \end{aligned} \tag{21}$$

366 Hydraulic conductivity K [m/d] in model G_1 is homogeneous across the model domain.
367 Model G_2 includes a zonation, as shown in Figure 3. Hydraulic conductivity is
368 homogeneous (while uncertain) across each of these regions. Conductivities of Zone 1 and
369 Zone 2 are denoted as K_1 and K_2, respectively.
370 The river stage h_2 is considered to be driven by a snowmelt process and is
371 characterized through an empirical rating curve of the kind:
372 $$h_2 = a_1 Q^{a_2} + a_3 \tag{22}$$
373 where a_1, a_2, and a_3 are arbitrary coefficients. The river discharge Q [m^3/s] is estimated
374 through snowmelt runoff as:
375 $$Q = C_1 \times C_{sn} \times M \times SVC \times A \tag{23}$$
376 where parameters C_{sn}, M, A, and SVC are runoff coefficient, snowmelt rate [mm/d],
377 watershed area [m^2], ratio of snow-covered area to watershed area, respectively, and C_1 is
378 a unit conversion factor from [mm/d] to [m^3/s]. We use two process models to evaluate the
379 daily snowmelt rate. These are hereafter termed as M_1 and M_2, and they correspond to the
380 degree-day method and the restricted degree-day radiation balance method, respectively, i.e.,
381 $$\begin{aligned} M_1 &: M = f_1 (T_a - T_m) \\ M_2 &: M = f_2 (T_a - T_m) + rR_n \end{aligned} \tag{24}$$
382 Here, f_1 and f_2 [$mm \cdot °C^{-1} \cdot d^{-1}$] represent the snowmelt factors; T_a [°C] and T_m [°C]
383 represent the average temperature for one day and the temperature threshold for snow
384 melting, respectively. Process model M_2 considers the effects of surface radiation budget,
385 R_n [W/m^{-2}], and relies on the transformation coefficient r [(mm/d)/(W/m^2)] to estimate the
386 snowmelt rate by energy flux.
387

Table 1. Uncertain parameters and their distributions used in the first case.

Process	Model	Parameter	Distribution	Unit
Recharge	R_1	a	$\mathcal{N}(16.88, 1)$	
	R_2	b	$U(0.1, 0.2)$	
Geology	G_1	K	$\mathcal{N}(15, 1)$	m/d
	G_2	K_1	$\mathcal{N}(20, 1)$	m/d
		K_2	$\mathcal{N}(10, 1)$	m/d
Snowmelt	M_1	f_1	$\mathcal{N}(3.5, 0.75)$	mm·℃$^{-1}$·d^{-1}
	M_2	f_2	$\mathcal{N}(2.5, 0.3)$	mm·℃$^{-1}$·d^{-1}
		r	$\mathcal{N}(0.3, 0.05)$	(mm/d)/(W/m^2)

Table 2. Values of deterministic parameters used in the first case.

Parameter	Value	Unit
Coefficient a_1	0.3	
Coefficient a_2	0.6	
Coefficient a_3	289	
Runoff coefficient (C_{sn})	0.8	
Watershed area (A)	2000	km^2
Ratio of snow-covered area to watershed area (SVC)	0.7	
Unit conversion factor (C_1)	0.001/86400	
Average temperature for one day (T_a)	7	℃
Temperature threshold for snow melting (T_m)	0	℃
Surface radiation budget (R_n)	80	W/m^{-2}

Based on the (*i*) two recharge process models, (*ii*) two geology process models, and (*iii*) two snowmelt process models described above, a total of eight alternative system models are here developed (hereafter denoted as $R_1G_1M_1$, $R_1G_1M_2$, $R_1G_2M_1$, $R_1G_2M_2$, $R_2G_1M_1$, $R_2G_1M_2$, $R_2G_2M_1$, and $R_2G_2M_2$). Equal weights are used for these process models, i.e., $P(R_1) = P(R_2) = 0.5$, $P(G_1) = P(G_2) = 0.5$, and $P(M_1) = P(M_2) = 0.5$. Ethene concentrations at seven locations within the range of $5,400\ \text{m} \le x \le 6,000\text{m}$ are the model outputs of interest (Δ) for the evaluation of the process sensitivity index.

The brute force MC method is implemented using a total of NMC = 3,920,000 model executions for the evaluation of the process sensitivity index based on the pseudo code shown

400 in Figure 1(a) for each of the processes analyzed. Considering the recharge process as an
401 example, the number of MC simulations is structured according to 2 recharge models × 700
402 parameter realizations for each recharge model × 2 geology models × 2 snowmelt models ×
403 700 parameter realizations for each combination of geology and snowmelt models. The total
404 number MC simulations for the combination of these three processes is then
405 $NMC = 11,760,000 = 3 \times 3,920,000$. Note that such a large number of MC simulations is
406 unnecessary from a practical standpoint (as discussed below, sufficiently accurate results of
407 process sensitivity index can be obtained through about 2,000,000 MC simulations). We
408 resort to such a high number of MC iterations only to ensure that the MC results can be used
409 as a highly accurate benchmark against which the main features of the quasi-MC and the
410 binning methods can be assessed. In this sense, we take the MC results for the value of the
411 process sensitivity index obtained at NMC = 3,920,000 as the reference to evaluate the
412 absolute relative error (%) for the three computational methods as a function of the number of
413 model simulations. The binning and quasi-MC methods are implemented with 400,000 and
414 200,000 model executions, respectively. For the quasi-MC method, 28,000 model executions
415 (corresponding to 1,750 parameter samples) are adequate for reaching convergence, as
416 discussed in the following.
417 Figure 4 shows the difference between the results obtained through 700 pseudorandom
418 samples of brute force MC and 1,750 quasi-random samples of quasi-MC, using model
419 $R_2G_2M_2$ as an example (only parameters of G_2 and M_2 are shown in Figure 4). Figure 4(a)
420 depicts scatterplots of 700 sample points for pairs of parameters of the same process model
421 (e.g., K_1 and K_2 of model G_2). One can note that the nested sampling structure of brute force

MC requires generating 700 × 700 = 490,000 samples for two parameters of different process models. For example, for each of the 700 samples of K_1, 700 samples of b, f_2, or r values need to be generated, according to loops [2] and [4] in Figure 1(a). Thus, the subplots for K_1-b, K_1-f_2, and K_1-r include a total of 490,000 sample points. Figure 4(b) shows that, after removing the nested sampling structure, application of our quasi-MC method leads to generating 1,750 samples for each parameter. This constitutes a clear advantage of our quasi-MC method, because more samples are generated for parameters (e.g., K_1 and K_2) of the same process model, while less samples are generated for parameters (e.g., K_1-b, K_1-f_2, and K_1-r) of different process models without sacrificing computational accuracy as shown below.

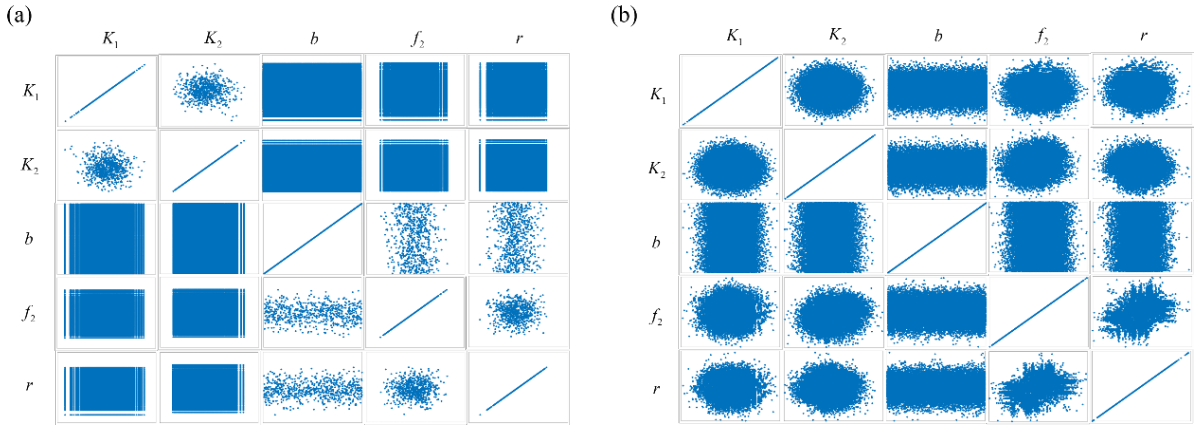
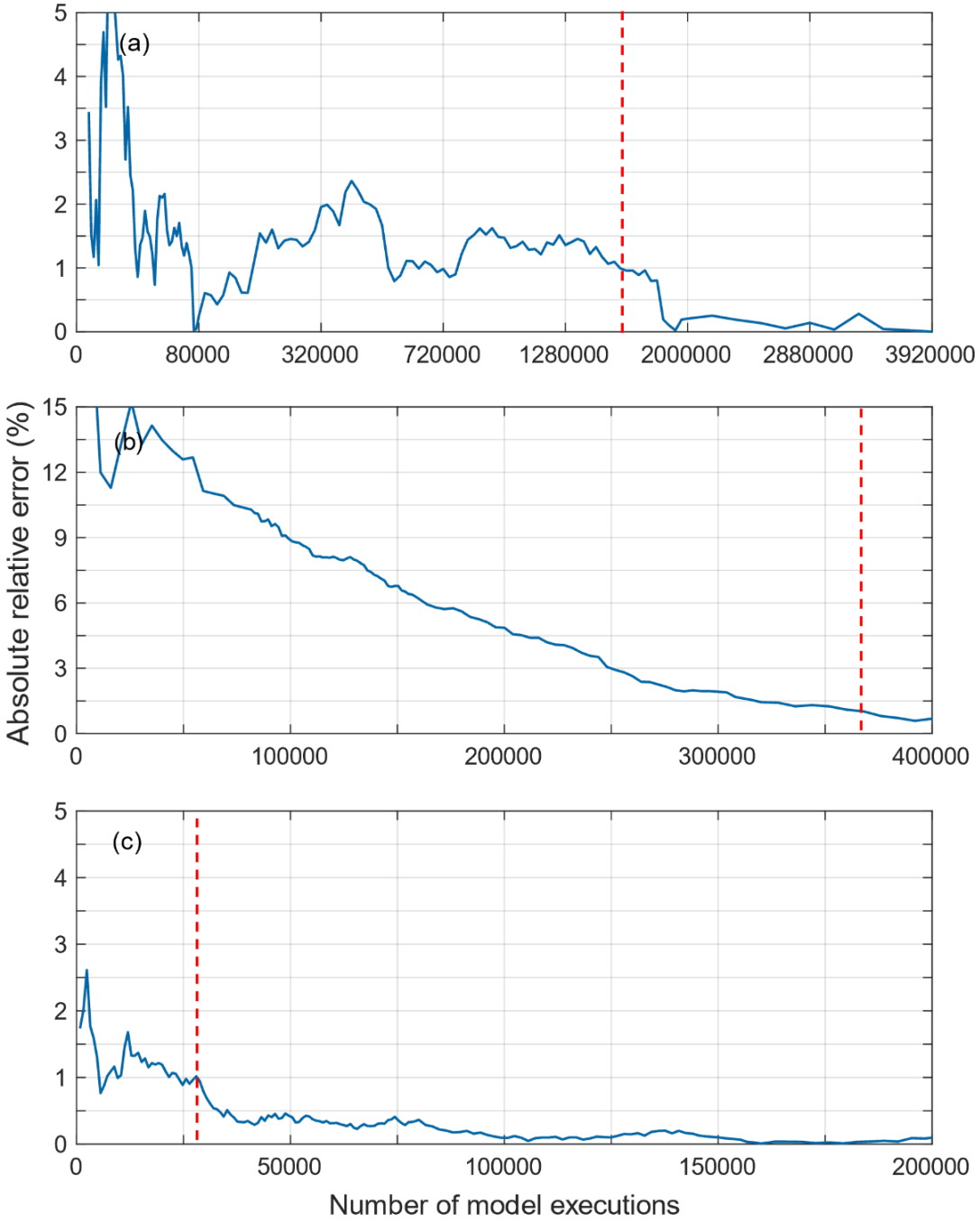


Figure 4. Random parameter samples associated with model $R_2G_2M_2$ for the process sensitivity index estimation of the first test scenario generated through (a) the brute force MC and (b) the quasi-MC methods.

Figure 5 depicts the results of absolute relative errors referred to the geology process sensitivity index obtained at location x = 5,700 m. Figure 5(a) enables one to examine convergence of the brute force MC method. These results suggest that the brute force MC method stabilizes at the reference value after approximately 2,000,000 model simulations (with an absolute relative error less than 1%). Figure 5(b) focuses on the binning method. Note that here the maximum number of model executions is 400,000. The latter stems from

441 considering 2 recharge models × 2 geology models × 2 snowmelt models × a total of 50,000
442 sets of parameter values (each of these sets corresponds to a combination of the eight process
443 models considered). These results suggest that the absolute relative error shows a consistently
444 decreasing trend and becomes less than 1% after about 360,000 model simulations. The
445 results depicted in Figure 5(c) for the quasi-MC method shows that the absolute relative error



446
447 Figure 5. Absolute relative error (%) associated with the process sensitivity index estimated
448 for the geology process using (a) MC method, (b) binning method and (c) new quasi-MC

method for the ethene concentration at location $x = 5,700$ m. The process sensitivity index value obtained through the MC method with 3,920,000 model simulations is used as the reference value for the evaluation of the absolute relative error. The red dashed lines mark the numbers of model executions with the absolute relative error of 1% for each method.

with respect to the final MC-based value becomes less than 1% after about 28,000 model executions. Such a number corresponds to considering 2 recharge models × 2 geology models × 2 snowmelt models × 2 × 1,750 parameter samples generated for each of the eight process models. Figure 5 indicates that the new quasi-MC method outperforms (*i*) the MC method in terms of convergence rate and (*ii*) the binning method in terms of both accuracy and convergence rate for the evaluation of the process sensitivity index.

Table 3. Absolute relative error (%) linked to the process sensitivity indices associated with the geology, recharge, and snowmelt processes and corresponding to the binning (with equal width and equal depth strategies) and quasi-MC methods for the simulated ethene concentrations at the seven target locations in the domain.

	Recharge			Geology			Snow-melt		
Location(m)	Equi-Width	Equi-Depth	Quasi-MC	Equi-Width	Equi-Depth	Quasi-MC	Equi-Width	Equi-Depth	Quasi-MC
5400	0.2796	0.0778	0.7984	1.5456	0.3580	0.0086	0.7735	0.6782	1.0289
5500	0.2388	0.2106	0.4243	0.8651	0.4204	0.0096	0.6713	0.9621	0.9705
5600	0.3614	0.0622	0.2105	0.2426	0.4407	0.0368	0.3259	1.2965	0.2177
5700	0.1023	0.4645	0.1925	0.6884	0.4469	0.0878	0.0370	1.3867	0.5190
5800	0.6383	1.4938	0.9261	2.0012	0.5508	0.2523	0.3519	0.9014	0.8293
5900	1.8053	2.9334	0.9541	0.4282	0.8386	0.6702	0.8577	0.3819	0.5835
6000	2.7889	4.0617	1.9872	4.1174	1.0582	1.5017	1.9536	2.6813	0.2478
Maximum	2.7889	4.0617	1.9872	4.1174	1.0582	1.5017	1.9536	2.6813	1.0289
Minimum	0.1023	0.0622	0.1925	0.2426	0.3580	0.0086	0.0370	0.3819	0.2177
Mean	0.8878	1.3292	0.7847	1.4126	0.5877	0.3667	0.7101	1.1840	0.6281

A similar analysis is performed for several seven locations of the model domain considered in the system. These results are summarized in Table 3. The table lists the values of the absolute relative error (%) for the binning (considering both equal width and equal depth) and quasi-MC methods for the process sensitivity index related to the geology, recharge, and snowmelt processes at the seven locations. Figure 6 complements these results by depicting the process sensitivity index related to the (a) geology, (b) recharge, and (c)

469 snowmelt processes evaluated through the three approaches (using binning with equal with).
470 The analysis of the ensemble of these results suggests that, while differences are hardly
471 discernible from visual inspection of Figure 6, the quasi-MC method generally yields more
472 accurate results than their counterparts grounded on the binning method. Moreover, the equal
473 depth binning strategy yields slightly more accurate results than the quasi-MC method in
474 some of the process sensitivity indices. For example, considering the geology and snowmelt
475 processes, Table 3 shows that the quasi-MC method outperforms the equal width and equal
476 depth binning methods. The binning method based on an equal depth strategy performs better
477 than the quasi-MC method with reference to the recharge process sensitivity indices.

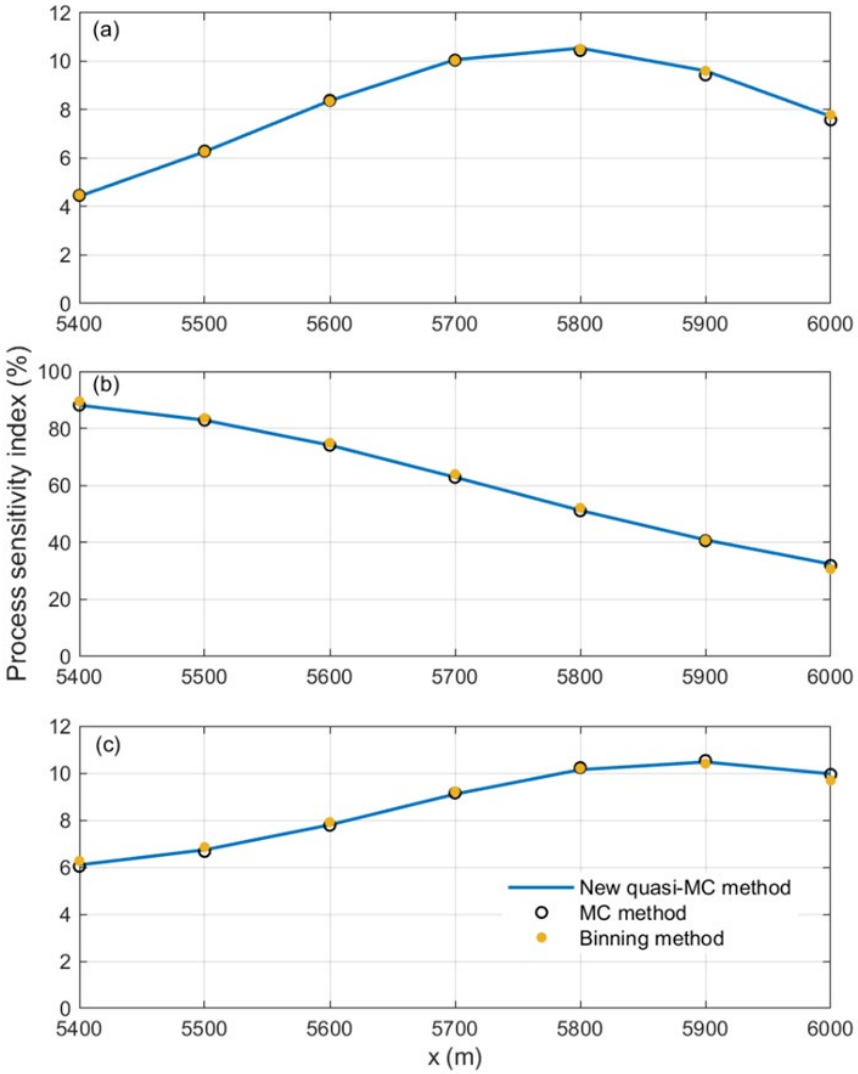


Figure 6. Process sensitivity indices associated with the (a) geology, (b) recharge, and (c) snowmelt processes and related to the simulated ethene concentrations at seven locations in the domain evaluated via the new quasi-MC, binning, and MC methods.

It should be noted that the results of the binning method depicted in Figure 6 and listed in Table 3 are associated with a carefully tuned number of (*i*) bins and (*ii*) random parameter samples within these. We remark that selection of the number of bins to obtain the results depicted in Figure 6 for the binning method is mostly empirical and time consuming. In this context, results of the binning method are markedly sensitive to the number of bins when considering both the equal width and equal depth strategies. This limitation is elucidated in Figure 7 that depicts the way that values of the absolute relative errors (between the results of

the binning method and the reference results obtained using 3,920,000 MC simulations) depend on the number of bins with reference to the snowmelt process at three selected locations where ethene concentrations are observed. Figure 7 illustrates the marked impact of the selection of the number of bins when the binning method is implemented through the equal width and equal depth strategies. An optimal (i.e., with zero absolute relative error) number of bins can be identified for each location at which the process sensitivity index is estimated. The error increases when the number of bins deviates from such an optimal number. It can be as large as 8.7% and 9% for the equal width and for the equal depth algorithm, respectively. It is further remarked that the selection of optimal number of bins is empirical, and it is only possible on the basis of our prior knowledge of the results of the MC method. In other words, if the MC results are not available, the binning method may not provide accurate results due to its empirical nature.

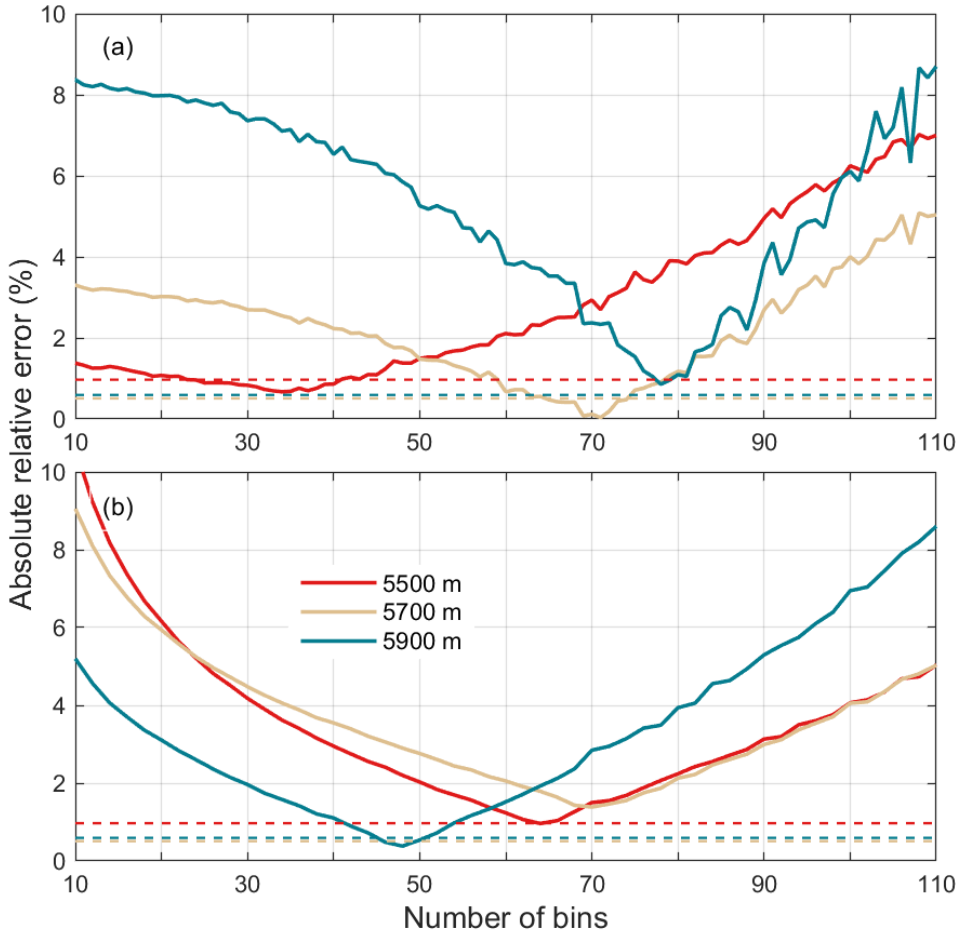
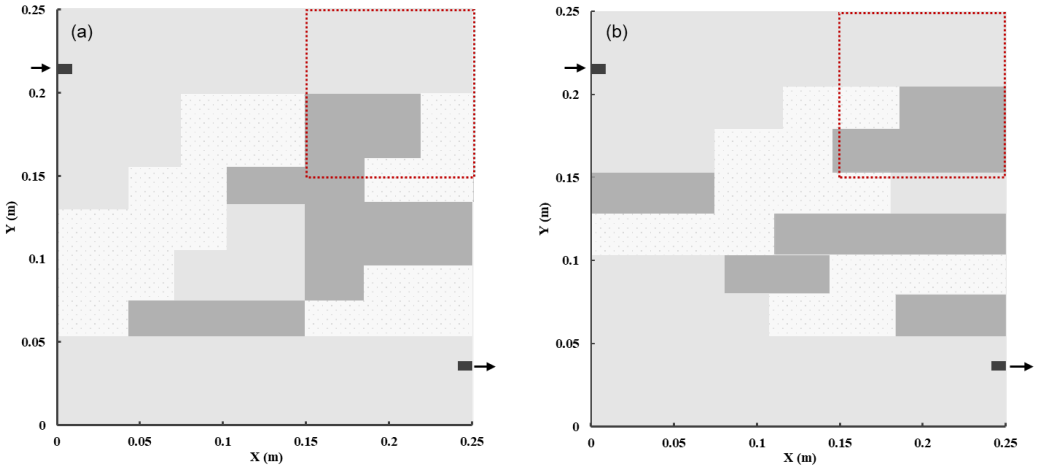


Figure 7. Absolute relative error versus the number of bins with reference to the snowmelt process at three selected locations where ethene concentrations are observed. Results correspond to the binning method implemented according to (a) the equal-width and (b) the equal-depth binning strategies. Reference values correspond to results obtained via 3,920,000 MC simulations. The dashed lines correspond to the absolute relative errors of the quasi-MC method.

3.2. Transport and sorption of zinc across a two-dimensional heterogeneous medium

This section illustrates the comparison of the binning and quasi-MC methods through the evaluation of the dynamics involved in a more complex scenario. The latter is inspired by the studies of Duan et al. (2020) and Maina et al. (2018) and involves transport and sorption of zinc through a laboratory scale heterogeneous porous medium. The domain of groundwater flow and solute transport is taken from the study of Maina et al. (2018). This numerical case considers a flow cell (internal size of $25 \times 25 \times 1.5\,\mathrm{cm}^3$) packed with three types of sand materials (corresponding to a fine, medium, and large grain size) and distributed according to

517 the heterogeneous patterns depicted in Figure 8. An inlet and an outlet are located at the
518 upper left and lower right corners of the system, respectively, as shown in Figure 8. Water
519 with dissolved zinc at a concentration of 0.5 mol/L is injected at the inlet according to a
520 constant flow rate of 4.0 L/d. Migration and sorption of zinc are simulated through
521 PFLOTRAN, a well established open-source software for subsurface flow and reactive
522 transport modeling (Steefel et al., 2015). The quantity of interest is the amount of absorbed
523 zinc across the area demarcated by red boxes in Figure 8. Transport of zinc in the domain is
524 rendered through a classical advection-dispersion equation coupled with a sorption model, as
525 described in the following.



526
527 Figure 8. Spatial distributions of three types of sand (sand 1 marked in light grey, sand 2 in
528 grey, and sand 3 in dark grey) across the flow and transport domain to mimic geological models
529 (a) G_1 and (b) G_2. Inlet and outlet sections are represented by black rectangles. Black arrows
530 represent the flow direction. Red boxes in the upper-right corners demarcate the area of interest
531 where the amount of adsorbed zinc is estimated.

532 This analysis considers model and parametric uncertainty related to geology and sorption
533 processes associated with zinc dynamics within the domain. With reference to the geology
534 process, models G_1 and G_2 represent two differing spatial distributions of the three sand
535 materials in the cell, as shown in Figure 8. Parameters K_1 [m/d], K_2 [m/d], and K_3[m/d]
536 represent hydraulic conductivity of sand 1, sand 2, and sand, 3 respectively. Parameters K_1

and K_2 are considered to be described through a normal distributions. The distributions of these two parameters and of the other parameters discussed below are listed in Table 4. The value of K_3 was deterministically set and is listed in Table 5 together with the values employed for the other deterministic parameters discussed below. For the purpose of our demonstration and to keep computational time manageable when performing the reference MC analysis (note that the total computational time associated with 60,000 simulations is about 1 week on a 10 system cores-based machine with Intel(R) Core(TM) i9-10900KF CPU and 64.0 GB RAM), we consider diffusion and dispersion to be characterized by deterministic parameter values. The effective aqueous diffusion coefficient is set to be constant across the entire modeling domain. Different values of longitudinal and transverse dispersivities are set for sands 1 - 3.

Table 4. Uncertain parameters and their distributions used in the second test scenario.

Process	Model	Parameter	Distribution	Unit
Geology	G_1 & G_2	K_1	$\mathcal{N}(691, 691)$	m/d
		K_2	$\mathcal{N}(20.7, 20.7)$	m/d
Sorption	S_1, S_2 & S_3	K_d	$U(0.426, 1.14)$	kg/cm^3
	S_2	k_1	$\mathcal{N}(0.664, 0.0664)$	h^{-1}
	S_3	$K_{d,\mathrm{Fe(III)}}$	$\mathcal{N}(1.28, 0.128)$	kg/cm^3

Table 5. Values of deterministic parameters used in the second test scenario.

Parameter	Value	Unit
Hydraulic conductivity of sand 3 (K_3)	0.23	m/d
Effective aqueuse diffusion coefficient	1.0×10^{-9}	m^2/s
Longitudinal dispersivity of sand 1	1.1×10^{-3}	m
Longitudinal dispersivity of sand 2	5.3×10^{-4}	m
Longitudinal dispersivity of sand 3	2.3×10^{-4}	m
Transverse dispersivity of sand 1	1.1×10^{-4}	m
Transverse dispersivity of sand 2	5.3×10^{-5}	m
Transverse dispersivity of sand 3	2.3×10^{-5}	m

552

553 The sorption process is represented by three alternative models (S_1, S_2, and S_3). Model

554 S_1 rests on a linear equilibrium formulation, i.e.,

555 $$S_1 : q_{Zn} = K_d \times C_{Zn} \tag{25}$$

556 where q_{Zn} [mol/cm^3] is the amount of adsorbed zinc, C_{Zn} [mol/kg] is the aqueous zinc

557 concentration, and K_d [kg/cm^3] is the linear equilibrium sorption constant. The latter is

558 assumed to be affected by uncertainty and to be characterized by a uniform distribution. Model

559 S_2 is based on describing zinc sorption through a dual first-order kinetic model with five

560 parameters(K_d, k_1, k_2, f_1 and f_2), i.e.,

561 $$\begin{aligned} S_2 : \frac{dq_{t,i}}{dt} &= k_i \left(q_{e,i} - q_{t,i} \right) \ (i = 1,2) \\ q_{e,i} &= K_d * C_{Zn} * f_i \\ q_t &= q_{t,1} + q_{t,2} \end{aligned} \tag{26}$$

562 where k_i [h^{-1}] is the first-order rate constant at sorption site i; f_i is the site fraction for

563 site i; $q_{t,i}$ [mol/cm^3] is the actual amount of adsorbed zinc at contact time t at site i; and

564 $q_{e,i}$ [mol/cm^3] is the amount of adsorbed zinc at time t at site i in equilibrium with aqueous

565 zinc concentration, C_{Zn} [mol/kg]. Model S_3 is based on the conceptual picture according to

566 which the oxidation process of mineral-bonded Fe(II) to Fe(III) in a porous medium is assumed

567 to improve zinc sorption capacity and an adjusted linear equilibrium sorption model can then

568 be employed, i.e.,

569 $$S_3 : q_{Zn} = (K_d + K_{d,\,\mathrm{Fe(III)}}) \times C_{Zn}, \tag{27}$$

570 where $K_{d,\mathrm{Fe(III)}}$ [kg/cm^3] is an additional term contributing to increase the linear equilibrium

571 sorption constant. We consider K_d and $K_{d,\mathrm{Fe(III)}}$ to be uncertain and characterized by a

572 uniform and a normal distribution, respectively.

573 Equal weights are employed for the different models included in the process sensitivity

analysis targeting the geology and sorption processes, i.e., $P(G_1) = P(G_2) = 0.5$ and $P(S_1) = P(S_2) = P(S_2) = 0.33$. The amount of adsorbed zinc in the upper-right corner of the domain (see Figure 8) at seven observation times within the range of $5\,\text{h} \le t \le 8\,\text{h}$ from the solute injection is the output of interest against which we evaluate the process sensitivity index via the MC, binning, and quasi-MC methods. The above mentioned time period is selected because preliminary simulations document that the amount of absorbed zinc is negligible for $t \le 5\,\text{h}$ and does not vary significantly with time after $t \ge 8\,\text{h}$. The brute force MC method is implemented upon relying on 60,000 model simulations with 10 seconds of computing time for each simulation. These correspond to 2 geology models × 100 parameter realizations for each geology model) × 3 sorption models × 100 parameter realizations for each sorption model. Similar to Section 3.1, we consider the brute force MC results at the largest number of realizations as a reference and evaluate the absolute relative errors (%) with respect to it for the three methods.

Figure 9 depicts the results of this analysis for time $t = 6.5$. Figure 9(a) depicts the relative error associated with the geology process sensitivity index versus the number of realizations evaluated by the MC method. Convergence to the reference value is noted after about 50,000 model executions (where the absolute relative error becomes less than 1%). A 1% relative error is attained after about 5,500 and 2,100 model executions for the binning and quasi-MC methods, respectively, as shown in Figures 9(b) and 9(c). These results suggest that (*i*) the binning and quasi-MC methods are significantly more efficient than the brute force MC for the evaluation of the process sensitivity index and (*ii*) the quasi-MC method is computationally more efficient than the binning approach.

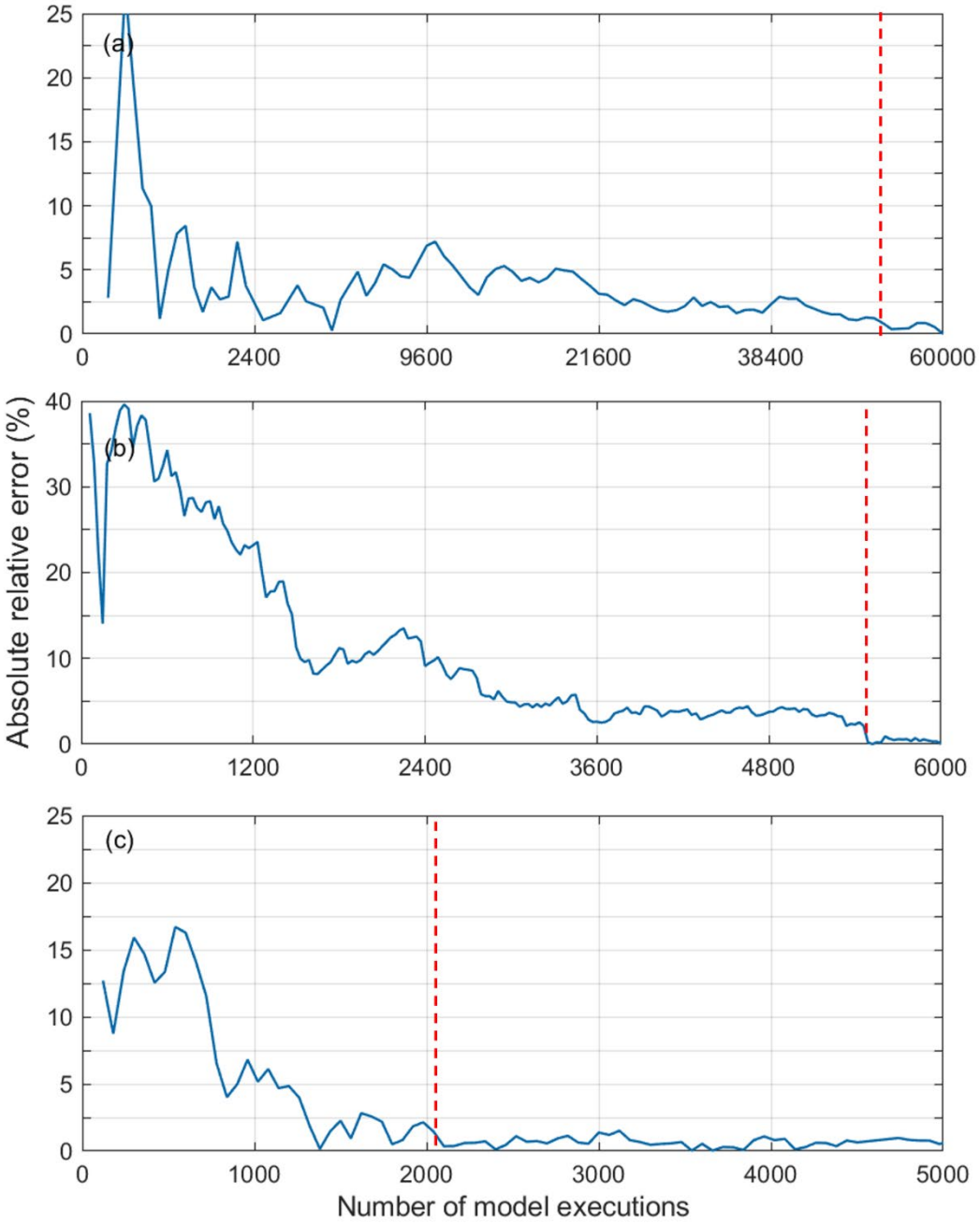


Figure 9. Absolute relative error (%) associated with the process sensitivity index estimated for the geology process using (a) MC, (b) binning and (c) the new quasi-MC method for the amount of Zn adsorbed in the upper right area of the system at time t = 6.5 h. The process sensitivity index value obtained through the MC method with 60,000 model simulations is used as the reference value for the evaluation of the absolute relative error. The red dashed lines mark the numbers of model executions with the absolute relative error of 1% for each method.

Figure 10 plots the process sensitivity index estimated for the geology and sorption processes using the brute force MC, binning, and quasi-MC methods at seven observation times within the range of $5h \le t \le 8h$. These results show that the binning and quasi-MC methods

607 can yield accurate results for all of the observation times. Table 6 complements these results
608 by listing the absolute relative errors (%) associated with the binning and quasi-MC methods
609 for the geology and sorption process sensitivity index at the seven observation times. While the
610 both methods yield accurate results, the quasi-MC method is generally more accurate than the
611 binning method in terms of the mean values of absolute relative errors, an exception being
612 noted for the results related to the sorption process and based on the equal depth binning method.

613 Table 6. Absolute relative error (%) linked to the process sensitivity indices associated with
614 the geology and sorption processes and corresponding to the binning (with equal width and
615 equal depth strategies) and quasi-MC methods for the simulated adsorbed Zn within the
616 target area.

	Geology			Sorption		
Time (h)	Equi-Width	Equi-Depth	Quasi-MC	Equi-Width	Equi-Depth	Quasi-MC
5.0	2.0379	1.9245	0.9028	1.4238	0.4066	1.1701
5.5	1.2610	1.5114	0.3454	0.8802	0.4538	0.9951
6.0	0.5760	1.0577	0.1382	0.6963	0.1781	0.5943
6.5	0.0182	0.6009	0.5175	0.7349	0.2709	0.0193
7.0	0.4131	0.1813	0.8231	0.8398	0.6857	0.6230
7.5	0.6621	0.1351	1.0288	0.9212	0.9887	1.1771
8.0	0.8113	0.3884	1.1770	0.9108	1.1129	1.5340
Maximum	2.0379	1.9245	1.1770	1.4238	1.1129	1.5340
Minimum	0.0182	0.1351	0.1382	0.6963	0.1781	0.0193
Mean	0.8257	0.8285	0.7047	0.9153	0.5853	0.8733

617 Similar to Figure 7, Figure 11 depicts the absolute relative errors versus the number of bins
618 for the binning method implemented through the equal width and equal depth strategies. These
619 results reinforce the observation that the absolute relative errors depend heavily on the number
620 of bins. We further emphasize that determining appropriate values of the binning variables is
621 empirical and time consuming and does not ensure attaining optimal results in the absence of
622 a reference value (which we have at our disposal in this demonstration).

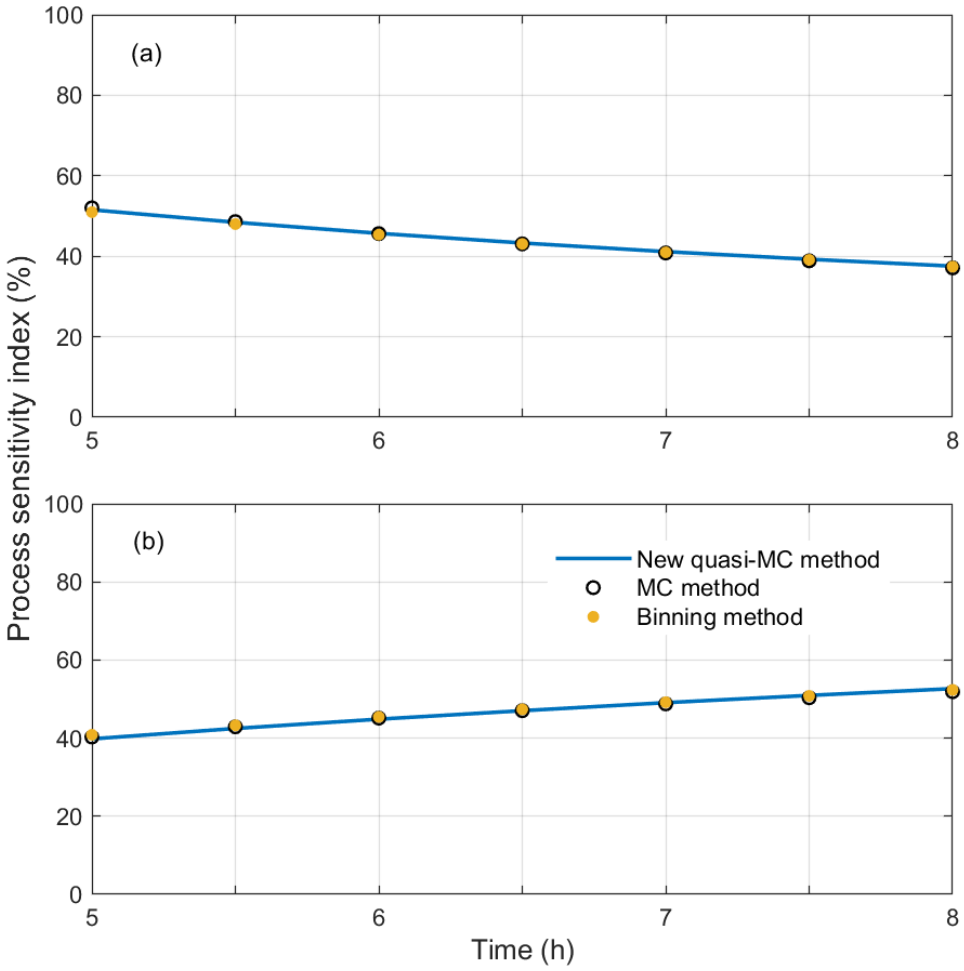


Figure 10. Process sensitivity indices associated with the (a) geology and (b) sorption processes and related to the simulated amount of adsorbed Zn in the target area of Figure 8 evaluated at seven observation times via the MC, binning, and quasi-MC methods.

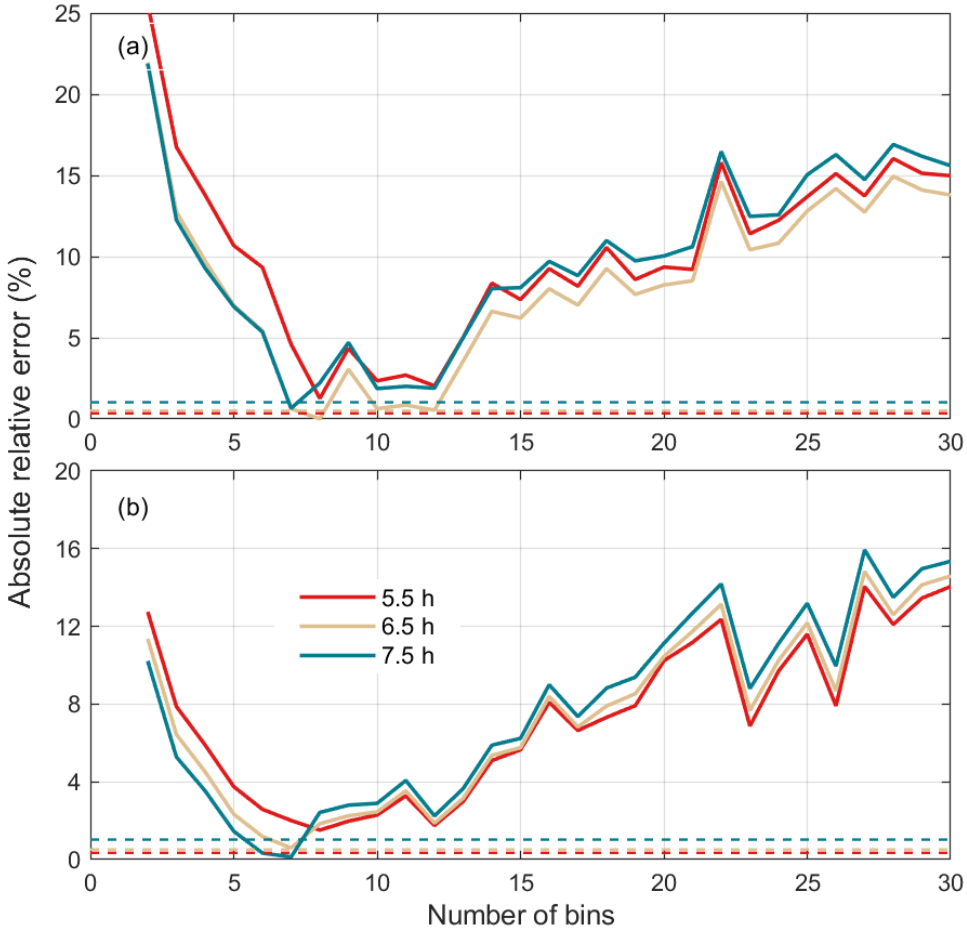


Figure 11. Absolute relative errors (%) versus the number of bins with reference to the geology process sensitivity index. Results correspond to the binning method implemented according to (a) the equal-width and (b) the equal-depth binning strategies. Reference values correspond to results obtained via 60,000 MC simulations. The dashed lines correspond to the absolute relative errors of the quasi-MC method.

4. Conclusions

We present a new theoretically robust quasi-MC method geared towards the evaluation of the multi-model process sensitivity index derived in Dai et al. (2017). Such an approach is here employed for the first time for process sensitivity analysis. The method is then assessed against the traditional Monte Carlo approach and the recent binning method proposed by Dai et al. (2017) through two synthetic cases associated with groundwater flow and transport scenarios with increasing level of complexity. These scenarios embed uncertainties associated

642 with process models and their parametrization. Our study leads to the following major
643 conclusions.
644 1. As a significant advancement, our quasi-MC method removes the need for the nested
645 structure of parameter sampling that is required by the typically used brute force MC
646 method. Hence, it is significantly more computationally efficient than it brute force MC
647 counterpart, because the number of model simulations is reduced from the order of N^2
648 for the MC method to the order of $2N$ for the quasi-MC method, N being the number
649 of parameter realizations of each process model.
650 2. Our quasi-MC method outperforms the binning method used in Dai et al. (2017) in terms
651 of theoretical aspects. The quasi-MC method is theoretically rigorous, thus removing the
652 barrier imposed by the inherently empirical nature of the binning approach. The latter is
653 plagued by the need of a time consuming procedure of bin selection/tuning which, by its
654 nature, is otherwise lacking the capability of ensuring results of optimal quality.
655 3. Results about the performance of our quasi-MC method against the binning and brute
656 force MC approaches document that the former yields highly accurate results in a
657 computationally more efficient manner. As such, the quasi-MC method is a remarkably
658 promising alternative for the evaluation of process sensitivity indices in the presence of
659 multiple sources of uncertainty.

660 ## Acknowledgement

661 The first author was supported by National Natural Science Foundation of China (Grant No.
662 42172280), and the third author was supported by NSF grant EAR-1552329. The source codes
663 of implementing the quasi-MC method and the two numerical case studies are available online

at https://github.com/qi667/Quasi_MC.git.

Reference

Antonetti. M., Buss, R., Scherrer, S., Margreth, M., & Zappa, M. (2016). Mapping dominant runoff processes: an evaluation of different approaches using similarity measures and synthetic runoff simulations. *Hydrology and Earth System Sciences*, 20(7), 2929-2945. https://doi.org/10.5194/hess-20-2929-2016

Antonetti, M., Scherrer S., Kienzler P.M., Margreth M., & Zappa M. (2017). Process-based hydrological modelling: The potential of a bottom-up approach for runoff predictionsin ungauged catchments. *Hydrological Processes*, 31(16), 2902–2920. https://doi.org/10.1002/hyp.11232

Aziz, C.E., Newell, C.J., Gonzales, J.R., Haas, P.E., Clement, T.P., & Sun, Y. (2000). BIOCHLOR Natural Attenuation Decision Support System: User's Manual, Version 1.0. U.S. EPA Office of Research and Development, Washington, DC.

Beven, K. J. (2002). Uncertainty and the detection of structural change in models of environmental systems. *Developments in Environmental Modelling*, 22, 227–250. https://doi.org/10.1016/S0167-8892(02)80013-6

Beven, K. (2006). A manifesto for the equifinality thesis. *Journal of hydrology*, 320(1-2), 18–36. https://doi.org/10.1016/j.jhydrol.2005.07.007

Bloschl, G. (2001). Scaling in hydrology. *Hydrological Processes*, 15(4), 709–711. https://doi.org/10.1002/hyp.432.

Ceriotti, G., Guadagnini, L., Porta, G., & Guadagnini, A. (2018). Local and global sensitivity analysis of Cr (VI) geogenic leakage under uncertain environmental conditions. *Water Resources Research*, 54(8), 5785–5802. https://doi.org/10.1029/ 2018WR022857.

Clark, M. P., Nijssen, B., Lundquist, J. D., Kavetski, D., Rupp, D.E., Woods, R. A., et al. (2015). A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resources Research*, 51(4), 2498–2514. https://doi.org/10.1002/2015WR017198.

Dai, H., Ye, M., Walker, A. P., & Chen, X. (2017). A new process sensitivity index to identify important system processes under process model and parametric uncertainty. *Water Resources Research*, 53(4),3476–3490. https://doi.org/10.1002/2016WR019715

Dai, H., & Ye, M. (2015). Variance-based global sensitivity analysis for multiple scenarios and models with implementation using sparse grid collocation. *Journal of Hydrology*, 528, 286–300. https://doi.org/10.1016/j.jhydrol.2015.06034

Dell'Oca, A., Riva, M., & Guadagnini, A. (2017). Moment-based metrics for global sensitivity analysis of hydrological systems. *Hydrology and Earth System Sciences*, 21(12), 6219–6234. https://doi.org/10.5194/hess-21-6219-2017

Draper, D. (1995), Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 45–97. https://doi.org/10.1111/j.2517-6161.1995.tb02015.x

Duan, Y., Li, R., Gan, Y., Yu, K., Tong, J., Zeng, G., et al. (2020). Impact of physico-chemical heterogeneity on arsenic sorption and reactive transport under water extraction. *Environ. Environmental Science & Technology*, 54(23), 14974–14983. https://doi.org/10.1021/acs.est.0c03587

Grayson, R. B., & G. Bloschl (2000). Spatial Patterns in Catchment Hydrology: Observations and

Modeling, Cambridge Univ. Press, Cambridge, U. K.
Guse, B., Pfannerstill, M., Gafurov, A., Fohrer, N., & Gupta, H. (2016). Demasking the integrated information of discharge: Advancing sensitivity analysis to consider different hydrological components and their rates of change. *Water Resources Research*, 52(11), 8724–8743. https://doi.org/10.1002/ 2016WR018894
Hou, T., Nuyens, D., Roels, S., & Janssen, H. (2019). Quasi-Monte Carlo based uncertainty analysis: sampling efficiency and error estimation in engineering applications. *Reliability Engineering & System Safety*, 191, 106549. https://doi.org/10.1016/j.ress.2019.106549
Ishigami, T., & Homma, T. (1990). An importance quantification technique in uncertainty analysis for computer models. *In [1990] Proceedings. First International Symposium on Uncertainty Modelling and Analysis*, December 3–6, University of Maryland.
Mai, J., Craig J. R., & Tolson B.A. (2020). Simultaneously determining global sensitivities of model parameters and model structure. *Hydrology and Earth System Sciences*, 24(12), 5835–5858. https://doi.org/10.5194/hess-24-5835-2020.
Maina, F. H., Ackerer, P., Younes, A., Guadagnini, A., & Berkowitz, B. (2018). Benchmarking numerical codes for tracer transport with the aid of laboratory-scale experiments in 2D heterogeneous porous media. *Journal of contaminant hydrology*, 212, 55–64. https://doi.org/10.5194/hess-24-5835-2020
Melsen, L. A., & Guse, B. (2019). Hydrological drought simulations: how climate and model structure control parameter sensitivity. *Water Resources Research*, 55(12), 10527-10547, https://doi.org/10.1029/2019WR025230
Montanari, A., & Koutsoyiannis, D. (2012). A blueprint for process-based modeling of uncertain hydrological systems. *Water Resources Research*, 48(9). https://doi.org/10.1029/2011WR011412
Razavi, S., & Gupta H. V. (2015). What do we mean by sensitivity analysis? The need for comprehensive characterization of "global" sensitivity in Earth and Environmental systems models. *Water Resources Research*, 51(5), 3070–3092. https://doi.org/10.1002/2014WR016527
Caflisch, R. E (1998), Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica*, 7,1–49.
Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer physics communications*, 181 (2), 259–270. https://doi.org/10.1016/j.cpc.2009.09.018
Saltelli, A., Tarantola, S., & Chan, K.P.-S. (1999). A quantitative model independent method for global sensitivity analysis of model output. *Technometrics*, 41(1), 39–56.
Sobol, I.M. (1993). Sensitivity analysis for nonlinear mathematical models. *Mathematical modelling and computational experiment,* 1 (4), 407–414.
Steefel, C. I., Appelo, C. A. J., Arora, B., Jacques, D., Kalbacher, T., Kolditz, O., et al (2015). Reactive transport codes for subsurface environmental simulation. *Computational Geosciences*, 19(3), 445-478.
Sun, Y., Petersen, J. N., & Clement, T. P. (1999). Analytical solutions for multiple species reactive transport in multiple dimensions. *Journal of contaminant hydrology*, 35(4), 429–440. https://doi.org/10.1016/S0169-7722(98)00105-3
Neuman, S.P., (2003). Maximum likelihood Bayesian averaging of alternative conceptual-mathematical models. *Stochastic Environ. Res. Risk Assess*, 17 (5), 291–305. https://doi.org/10.1007/s00477-003-0151-7
Wainwright, H. M., Finsterle, S., Jung, Y., Zhou, Q., & Birkholzer, J. T. (2014). Making sense of

global sensitivity analysis. *Computers & Geosciences*, 65, 84–94. https://doi.org/10.1016/j.cageo.2013.06.006

Walker, A. P., Ye, M., Lu, D., De Kauwe, M. G., Gu, L. H., Medlyn, B. E., et al. (2018). The multi-assumption architecture and testbed (MAAT v1.0): R code for generating ensembles with dynamic model structure and analysis of epistemic uncertainty from multiple sources. *Geoscientific Model Development*, 11(8), 3159-3185. https://doi.org/10.5194/gmd-11-3159-2018

Walker, A.P., Johnson, A.L., Rogers, A., Anderson, J., Bridges, R.A., Fisher, R.A., et al. (2021). Multi-hypothesis comparison of Farquhar and Collatz photosynthesis models reveals the unexpected influence of empirical assumptions at leaf and global scales. *Global Change Biology*, 27(4), 804-822. https://doi.org/10.1111/gcb.15366

Xu, K., Xu, B. B., Ju, J. L., Wu, C. H., Dai, H., & Hu, B. X. (2019). Projection and uncertainty of precipitation extremes in the CMIP5 multimodel ensembles over nine major basins in China. *Atmospheric Research*, 226, 122-137. https://doi.org/10.1016/j.atmosres.2019.04.018

Yang, J., Ye, M., Chen, X., Dai, H., & Walker A. P. (2022). Process interactions can change process ranking in a coupled complex system under process model and parametric uncertainty. *Water Resour. Res*., 58(3). https://doi.org/10.1029/2021WR029812

Yang, J. & M. Ye (2022). A new multi-model absolute difference-based sensitivity (MMADS) analysis method to screen non-influential process under process model and parametric uncertainty, Journal of Hydrology, 608, 127609. https://doi.org/10.1016/j.jhydrol.2022.127609

Ye, M., Meyer, P. D., & Neuman, S. P. (2008). On model selection criteria in multimodel analysis. *Water Resources Research*, 44(3). https://doi.org/10.1029/2008WR006803

Ye, M., Pohlmann, K.F., Chapman, J.B., Pohll, G.M., & Reeves, D.M. (2010). A model averaging method for assessing groundwater conceptual model uncertainty. *Groundwater*, 48(5), 716-728. https://doi.org/10.1111/j.1745-6584.2009.00633.x

Zhang, X. C. (2019). Determining and modeling dominant processes of interrill soil erosion. *Water Resources Research*, 55(1), 4–20. https://doi.org/10.1029/2018WR023217