Process interactions can change process ranking in a coupled complex system under process model and parametric uncertainty

Jing Yang[1,2], Ming Ye[1], Xingyuan Chen[3], Heng Dai[4], Anthony P. Walker[5]

[1]Department of Earth, Ocean, and Atmosphere Science and Department of Scientific Computing, Florida State University, Tallahassee, FL 32306, USA.

[2]School of Environmental Studies, China University of Geosciences, Wuhan 430074, China

[3]Pacific Northwest National Laboratory, Richland, WA 99354, USA

[4]State Key Laboratory of Biology and Environmental Geology, China University of Geosciences, Wuhan 43007, China

[5]Environmental Sciences Division and Climate Change Science Institute, Oak Ridge National Laboratory, Oak Ridge, TN 37831, USA

Corresponding author: Ming Ye (mye@fsu.edu)

Jing Yang (ORCID: 0000-0001-9838-5102)

Ming Ye (ORCID: 0000-0002-7080-0578)

Heng Dai (ORCID: 0000-0002-7721-9858)

Xingyuan Chen (ORCID: 0000-0003-1928-5555)

Anthony P. Walker (ORCID: 0000-0003-0557-5594)

Revised manuscript submitted for publication in Water Resources Research

September 2021

Key points

- A new total-effect process sensitivity index is derived to account for process interactions under process model and parameter uncertainty.
- The total-effect process sensitivity index has a first-order term for process importance and higher-order terms for process interactions.
- Accounting for process interactions allows for identifying influential system processes and/or screening non-influential system processes.

Abstract

For a complex hydrologic system with multiple processes and process interactions, global sensitivity analysis is often used to identify important or influential parameters for model development and improvement. The identification is complicated by process model uncertainty, when a system process can be represented by multiple process models. This study develops a new total-effect process sensitivity index to identify influential processes under model uncertainty. This is done by extending Sobol's total-effect *parameter* sensitivity index for one system model to total-effect *process* sensitivity index for multiple system models to account for uncertainty in process models and model parameters. The total-effect process sensitivity index includes not only the first-order process sensitivity index for measuring importance of individual processes but also higher-order indices that account for process interactions. The total-effect process sensitivity index can identify an influential process that itself and its interactions with other processes influence a model output. The total-effect process sensitivity index is applied to two numerical examples: (1) Sobol's G*-functions with analytical solutions of first-order and total-effect process sensitivity indices, and (2) groundwater flow models with interactions between recharge, geology, and snowmelt processes. The second evaluation shows that, due to second-order and higher-order process interactions, the first-order and total-effect process sensitivity indices give different process ranking. It is thus necessary to estimate both first-order and total-effect process sensitivity indices to appreciate the difference between the first-order impact of a process alone and the overall total-effect impact of the process itself and its interactions with other processes on a model output.

Plain Language Summary

When studying a complex hydrologic system, it is necessary to identify non-influential processes of the system so that limited resources are not spent on improving our understanding of these processes. On the other hand, it is important to identify influential processes of the system so that limited resources can be efficiently spent on better understanding the influential processes. Identification of the influential and non-influential processes is difficult when a process can be represented by several plausible process models because it is not always clear which process model to choose. To resolve this issue, we develop a new total-effect process sensitivity index that consider all the plausible process models without choosing one model and discarding other models. This is done by integrating the model averaging method with the Sobol's total-effect parameter sensitivity index. We use two numerical examples to verify computer codes and to demonstrate how to use the index to identify influential and non-influential processes. Applied to groundwater flow modeling, our new index demonstrates that accounting for interactions between recharge, geology, and snowmelt processes gives a ranking of process influence that is different from the ranking of process importance based on the first-order process sensitivity index.

1. Introduction

When developing and improving process-based models of hydrologic and other systems, how to represent sub-system physical, chemical, and biological processes and their interactions is a fundamental decision (Gupta et al., 2012; Clark et al., 2015a, 2015b, 2016, 2017; Mendoza et al., 2015; Markstrom et al., 2016; Chang et al., 2017; Beven, 2018; Walker et al., 2018; Bianchi Janetti et al., 2019; Spielet et al., 2019; Mai et al., 2020). Clark et al. (2015a) stated that "*[the decision] involves making choices on (1) model complexity, i.e., which physical processes should be represented explicitly, and, correspondingly, which processes can be ignored or greatly simplified; and (2) process representation, i.e., what modeling approaches should be used to represent the dominant biophysical and hydrologic processes*". The two choices depend on each other, and making the choices requires the identification of dominant, important, or influential processes. This identification is often done by expert judgement, but more formal methods have gained popularity in the last few decades as computational resources become more powerful. These more formal methods are primarily *parameter* sensitivity analyses, and a process is deemed to be important or influential if a system model output is sensitive to variation in the parameters of a given process (for hydrology-related review articles see Sivakumar, 2004; Shin et al., 2013; Gan et al., 2014; Song et al., 2015; Razavi and Gupta, 2015). However, these parameter-based methods commonly assume no uncertainty in conceptual and/or mathematical representations of a system process, and a single process model is employed to describe a system process.

The concept of single-model representation of a sub-system process is questionable when process model uncertainty exists, i.e., multiple plausible process models or hypotheses are available to represent a process. As a result, the structural uncertainty in the integrated system

model arises from multiple choices in various process models that together make up the system model. The system model uncertainty in surface and subsurface hydrology has been recognized for about two decades (Beven, 2002; Bredehoeft, 2003, 2005; Neuman, 2003; Poeter and Anderson, 2005). Since system model behaviors depend on the specific choice of process models that comprise the system model, the importance or influence of a system process is likely to be different when the process is simulated by different process models. Therefore, identifying important or influential processes without considering process model uncertainty may lead to biased or even wrong identification (Dai et al., 2017a; Walker et al., 2021). Dai et al. (2017a) showed that fully accounting for process model uncertainty can reorder process ranking compared with parameter sensitivity analysis methods that do not account for process model uncertainty.

Statistical methods have been developed recently to identify important or influential system processes by implicitly or explicitly addressing process model uncertainty (Mai et al., 2020). In an implicit approach, multiple models are developed, sensitivity analysis is performed for parameters of each system model, and importance or influence of the parameters is examined across the models (van Werkhoven et al., 2008, Herman et al., 2013, Günther et al., 2019; Bianchi Janetti et al., 2019). Taking the study of Bianchi Janetti et al. (2019) as an example, they considered three conceptual geological models with uncertain parameters, applied three different methods of global sensitivity analysis to each model, and evaluated intra-model and inter-model differences of the results of sensitivity analysis. They found that the choice of the conceptual model characterizing aquifer lithology affects parameter influence on model outputs.

The study of Tarantola et al. (2002) appears to be the first to explicitly consider model uncertainty in Sobol's framework of variance-based global sensitivity analysis, which was

further developed by Lilburne and Tarantola (2009) and by Baroni and Tarantola (2014) into a general probabilistic framework. The framework treats multiple uncertainty sources (e.g., parameter uncertainty and model uncertainty) as equal model input factors, and estimates the first-order and total-effect sensitivity indices for the factors. Savage et al. (2016) used the framework for a flood inundation model to quantify the importance of four input factors, i.e., spatial resolution, digital elevation model, Manning's friction coefficients, and inflow hydrograph as boundary conditions.

Without using the Sobol's framework of variance-based global sensitivity analysis, Van Hoey et al. (2014) conducted a qualitative model component sensitivity analysis, where a model component is defined as "a conceptual description of a subprocess of the entire model". The model component sensitivity analysis method also explicitly considers model uncertainty by varying model components one at a time to evaluate how model outputs change accordingly, which is similar to the one-factor-at-a-time method for parameter sensitivity analysis. Francke et al. (2018) also used the one-at-a-time approaches and global sensitivity analysis to examine the impacts of model enhancements on improving system model performance. The model enhancements include not only enhancing process formulations but also collecting more monitoring data and improving calibration data. With calibration data, the variation between models (process models, model components, or system models) can be optimized in a framework used by Spieler et al. (2019) for automatic model structure identification.

Dai et al. (2017a) were the first to integrate Sobol' variance-based global sensitivity analysis with model averaging approaches for identifying important processes. The model averaging approaches provide a general framework to quantify parametric and model uncertainty in a hierarchical structure, in that parametric uncertainty is quantified first for individual models and

147 model uncertainty is quantified subsequently for multiple models (Ye et al., 2004, 2008; Poeter
148 and Hill, 2007; Burnham and Anderson, 2002; Hoeting et al., 1999; Draper, 1995). Using the
149 hierarchical structure for quantifying parametric and model uncertainty is a major difference
150 between the method of Dai et al. (2017a) and the method of Tarantola et al. (2002). The first-
151 order process sensitivity index has been used not only in hydrologic research but also in
152 ecological studies (e.g. Walker et al., 2018, 2020). Dell'Oca et al. (2020) also integrated model
153 averaging and global sensitivity analysis, and they further extended model averaging to multiple
154 statistical moments (including mean, variance, skewness, and kurtosis). The process sensitivity
155 indices of Dell'Oca et al. (2020), to our knowledge, are the most comprehensive set in the
156 literature to consider model choice, parameter choice, and the model and parameter contributions
157 to the statistical moments of a quantity of interest.

158 The first-order process sensitivity index developed by Dai et al. (2017a) can only identify
159 *important* processes, and cannot account for interactions between processes to identify *influential*
160 processes. Similar to the characterization of factor interactions discussed in Saltelli et al. (2004),
161 process interactions may imply that extreme values of a model output are obtained by particular
162 combinations of model processes, not by a single process. In other words, two processes are said
163 to interact when their joint effects on the model output cannot be expressed as a sum of their
164 single effects. Following Saltelli et al. (2004), important and influential processes are
165 distinguished in the context of the following two different "settings" in which sensitivity analysis
166 is used:

167 (1) Process prioritization setting to identify important processes. Importance of a process is
168 measured in terms of reducing model output uncertainty achieved by fixing the process, i.e.,
169 fixing the process representation to a process model. In other words, fixing the most

important process leads to the greatest reduction of predictive uncertainty, and important processes corresponding to large uncertainty reduction should receive high priority when designing laboratory experiments and/or field campaigns.

(2) Process fixing setting to identify noninfluential processes. Influence of a process is measured by its contribution (through itself or its interactions with other processes) to predictive uncertainty. If a process has zero contribution to predictive uncertainty and thus is absolutely non-influential process, then one does not need to spend resources to further study the process and its representation can be fixed to any process model during model development and improvement.

While an important process is also an influential one, the most important process may not be the most influential one, if the process does not interact with other processes. On the other hand, a process may not be important, but may be influential due to its interactions with other processes. If a process itself is not important and does not interact with other processes, this process is non-influential. A key question to the process sensitivity analysis of this study is to investigate whether accounting for process interactions gives process influence ranking that is different from process importance ranking for a system model with a given set of process model representations and probabilistic parameter distributions.

To tacking the question on how process interactions affect process ranking, this study develops a new total-effect process sensitivity index. The difference between the total-effect process sensitivity index of this study and the first-order process sensitivity index of Dai et al. (2017a) is that the former identifies influential (or non-influential on the other hand) processes based on its first-order effect (also called main effect) and interactions with all other process and their respective uncertainties, while the latter identifies important processes based solely on their

first-order effect. As discussed below, the first-effect process sensitivity index is part of the total-effect process sensitivity index, which also includes second- and higher-order effects. Similar to the first-order process sensitivity index, the total-effect process sensitivity index is developed by integrating Sobol's total-effect parameter sensitivity index with the model-averaging approaches. Because of the integration, the total-effect process sensitivity index can identify influential (and non-influential) processes with consideration of uncertainty in process models and model parameters including all interactions. The derivation of the index and its numerical approximation using Monte Carlo simulation are given in Section 2 below. The numerical approximation is evaluated by a numerical experiment using the Sobol's G*-function. By treating each G*-function as a process and considering three G*-functions (each with two alternative expressions), we formulate a total of eight (2^3) system models, and derive analytical solutions of the first-order and total-effect process sensitivity indices for the eight system models. Another numerical experiment of groundwater flow modeling is used to evaluate the two indices in the hydrologic context. A total of eight system models are developed by considering three processes (recharge, geology, and snowmelt) and two alternative process models for each process. This modular modeling structure has been widely used in surface water and groundwater modeling (e.g., Clark et al., 2016; Ye et al., 2010, 2016; Lu et al., 2015). The numerical experiment of groundwater modeling is used to demonstrate the following: (1) how the first-order process sensitivity index is used to identify important processes, (2) how the total-effect process sensitivity index is used to identify influential and non-influential processes, and (3) why the first-order and total-effect process sensitivity indices give different process ranking due to process interactions. The two numerical experiments are presented in Section 3. Parameter interactions and process interactions considered in the total-effect parameter and process

sensitivity indices are discussed in Section 4, which also includes general guidelines on when and how to consider impacts of process interactions on identification of influential processes. Using the concept of within-model and between-model variance, it is explained in Section 5 that the first-order and total-effect *process* sensitivity indices of multiple system models are not simply weighted averages of the first-order and total-effect *parameter* sensitivity indices of individual system models. A discussion on limitations of this study is also given in Section 5, followed by Section 6 with the major conclusions of this study.

2. Methodology

Denote a quantity of interest as Δ and assume that it is a model output. Without process model uncertainty, Δ is simulated by a single system model, M, i.e., $\Delta = M(\theta) = M(\theta_1,...,\theta_d)$, where $\theta = \{\theta_1,...,\theta_d\}$ is a vector of d-dimensional system model parameters. Note that we here use θ to denote one realization of a set of parameter values and reserve $\boldsymbol{\theta}$ to denote an ensemble of θ. If the system of interest consists of multiple processes (denoted as A, B, …) and each process has its own process models (denoted as M_A, M_B, …) and associated parameters (denoted as θ_A, θ_B, …), the system model, $M(\theta)$, may be viewed as an integration of the process models, i.e., $M(\theta) = \bigcup\left(M_A(\theta_A), M_B(\theta_B),...\right)$, together with other system model components common to the processes (e.g., domain discretization, system initial conditions, and driving forces). The process model integration (i.e., the union symbol above) may need to consider nonlinear interactions between the process models. An example of such integration is the development of a system model of groundwater reactive transport, as it integrates processes of groundwater flow, thermal transport, solute transport, and biogeochemical reactions. In addition, a process may

have its sub-processes. For example, a groundwater flow process may consist of sub-processes such as recharge, geology, and evapotranspiration. More discussions on process-based modular modeling are referred to Clark et al. (2008, 2015a,b, 2016).

With the presence of process model uncertainty, a system process may be represented by several alternative process models. Taking process *A* as an example, it may be represented by multiple process models that form a set, $\mathbf{M}_A(\theta_A) = \left\{M_{A_1}(\theta_{A_1}), M_{A_2}(\theta_{A_2}),...\right\}$; each process model may have its own parameters or have parameters in common with other process models. The integration of the alternative process models leads to multiple alternative system models, i.e., $\mathbf{M}(\theta) = \bigcup\left(\mathbf{M}_A(\theta_A), \mathbf{M}_B(\theta_B),...\right)$. If the process model parameters are random, they are denoted as $\boldsymbol{\theta}$ to be differentiated from θ, which is a realization of $\boldsymbol{\theta}$. In this case, $\mathbf{M}(\theta) = \bigcup\left(\mathbf{M}_A(\theta_A), \mathbf{M}_B(\theta_B),...\right)$ becomes $\mathbf{M}(\boldsymbol{\theta}) = \bigcup\left(\mathbf{M}_A(\boldsymbol{\theta}_A), \mathbf{M}_B(\boldsymbol{\theta}_B),...\right)$.

2.1. First-Order and Total-Effect Process Sensitivity Indices

With the above formulation of process models and system models, we derive the total-effect process sensitivity index within the framework of variance-based global sensitivity analysis. The uncertainty of model output, Δ, is measured by its variance $V(\Delta)$, and decomposing the variance in two different ways leads to the first-order process sensitivity index defined by Dai et al. (2017a) and the total-effect process sensitivity index of this study. The two indices are discussed here to better understand the derivation of the total-effect process sensitivity index.

The first-order process sensitivity index is defined by examining how much the variance, $V(\Delta)$, can be reduced if the process model of process *K* can be fixed to the true model M_K^*, i.e., $V(\Delta)$ becomes $V_{\mathbf{M}_{\sim K}}[\Delta \mid M_K^*]$, where subscription $\mathbf{M}_{\sim K}$ indicates that the variance of Δ is over the

process models of all processes but process K, and $\Delta \mid M_K^*$ means that Δ is simulated conditioning on M_K^*. The variance reduction due to fixing the process model is $V(\Delta) - V_{\mathbf{M}_{\sim K}}[\Delta \mid M_K^*]$. Since the true process model, M_K^*, is unknown, it is practically sensible to examine the average reduction over all possible process models, i.e., over the set, $\mathbf{M}_K$, of the process models for process K. Therefore, the average reduction of variance is $V(\Delta) - E_{\mathbf{M}_K}(V_{\sim \mathbf{M}_K}[\Delta \mid M_K])$, where M_K is one plausible process model in the process model set, $\mathbf{M}_K$. Considering the law of total variance with respect to the models of process K, i.e.,

$$V(\Delta) = E_{\mathbf{M}_K}(V_{\mathbf{M}_{\sim K}}[\Delta \mid M_K]) + V_{\mathbf{M}_K}(E_{\mathbf{M}_{\sim K}}[\Delta \mid M_K]). \quad (1)$$

we have $V(\Delta) - E_{\mathbf{M}_K}(V_{\mathbf{M}_{\sim K}}[\Delta \mid M_K]) = V_{\mathbf{M}_K}(E_{\mathbf{M}_{\sim K}}[\Delta \mid M_K])$. Normalizing it by $V(\Delta)$ gives the first-order process sensitivity index (Dai et al., 2017a),

$$PS_K = \frac{V_{\mathbf{M}_K}(E_{\mathbf{M}_{\sim K}}[\Delta \mid M_K])}{V(\Delta)}, \quad (2)$$

which measures the importance of process K in terms of variance reduction by fixing this process. A process with larger value of PS_K is more important in the sense of larger variance reduction.

The total-effect process sensitivity index is defined by examining how much the variance, $V(\Delta)$, is left if the models of all processes but process K are fixed to the true model $M_{\sim K}^*$, i.e., $V(\Delta)$ becomes $V_{\mathbf{M}_K}(\Delta \mid M_{\sim K}^*)$ with the subscription $\mathbf{M}_K$ indicating that the variance of Δ is over the process models of process K. Since the true process model, $M_{\sim K}^*$, is unknown, it is practically sensible to examine the average variance, $E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$, over the process models, $\mathbf{M}_{\sim K}$, of all processes but process K. The average variance answers the following question: *Which process would leave the largest variance of* Δ*, if the process is left unfixed but*

all other processes are fixed? This average variance, $E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$, can be evaluated as $V(\Delta) - V_{\mathbf{M}_{\sim K}}(E_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ based on the following expression of the law of total variance with respect to the process models of all the processes but *K*, i.e.,

$$V(\Delta) = E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}]) + V_{\mathbf{M}_{\sim K}}(E_{\mathbf{M}_K}[\Delta \mid M_{\sim K}]). \quad (3)$$

Normalizing either $E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ or $V(\Delta) - V_{\mathbf{M}_{\sim K}}(E_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ by $V(\Delta)$ gives two equivalent expressions of the total-effect process sensitivity index for process K, i.e.,

$$PS_{TK} = \frac{E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])}{V(\Delta)} = 1 - \frac{V_{\mathbf{M}_{\sim K}}(E_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])}{V(\Delta)}. \quad (4)$$

A process with larger value of PS_{TK} is more influential to system model output Δ; influence of a process is discussed in the paragraph below. The total-effect process sensitivity index is symbolically similar to the Sobol's total-effect parameter sensitivity index for a single model, M_K,

$$S_{Ti} = \frac{E_{\boldsymbol{\theta}_{\sim i}}(V_{\boldsymbol{\theta}_i}[\Delta \mid \theta_{\sim i}])}{V_{M_K}(\Delta)} = 1 - \frac{V_{\boldsymbol{\theta}_{\sim i}}(E_{\boldsymbol{\theta}_i}[\Delta \mid \theta_{\sim i}])}{V_{M_K}(\Delta)}, \quad (5)$$

where $\boldsymbol{\theta}_i$ and $\boldsymbol{\theta}_{\sim i}$ denote the parameter sets of the *i-th* parameter and the other parameters, respectively, $\theta_{\sim i}$ denotes a realization of $\boldsymbol{\theta}_{\sim i}$, and $V_{M_K}(\Delta)$ denotes variance of Δ for model M_K due to parametric uncertainty. We explained in Section 5 that the total-effect process sensitivity index is not simply weighted average of the total-effect parameter sensitivity index. This is also true for the first-order parameter and process sensitivity indices.

One purpose of defining the total-effect process sensitivity index is to screen a non-influential process (or several processes) that can be set at any of its process models. Following Saltelli et al. (2004), a process is said to be influential to system model output Δ if the process

itself affects Δ and its interactions with other processes also affect Δ. Process interactions may imply that extreme values of Δ are obtained by particular combinations of model processes, not by a single process. In other words, two processes are said to interact when their joint effects on Δ cannot be expressed as a sum of their single effects on Δ. We can imagine an extreme case that a process has no interactions with other processes. In this case, $E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ becomes $V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}]$, and it measures the effects of process *K* only. We explain in Section 4 that the total-effect process sensitivity index includes the first-order process sensitivity index of a single process and high-order indices for interactions between multiple processes (e.g., second-order for two processes and third-order for three processes). Using the total-effect process sensitivity index avoids calculating the higher-order indices, because they are all included in the total-effect index.

2.2. Extension to Include Parametric Uncertainty

Equation (4) considers only process model uncertainty, and it is necessary to extend the definition to consider uncertainty in the parameters of each process model. The extension can be done in two ways for the $E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ and $V(\Delta) - V_{\mathbf{M}_{\sim K}}(E_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ terms used to define the total-effect process sensitivity index in equation (4). When working with $E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$, using the definition of variance gives

$$\begin{aligned} E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}]) &= E_{\mathbf{M}_{\sim K}}\left(E_{\mathbf{M}_K}\left([\Delta \mid M_{\sim K}]\right)^2 - \left(E_{\mathbf{M}_K}[\Delta \mid M_{\sim K}]\right)^2\right) \\ &= E_{\mathbf{M}_{\sim K}} E_{\mathbf{M}_K}\left([\Delta \mid M_{\sim K}]\right)^2 - E_{\mathbf{M}_{\sim K}}\left(E_{\mathbf{M}_K}[\Delta \mid M_{\sim K}]\right)^2 \end{aligned}. \quad (6)$$

Subsequently, applying the law of total expectation to the two expectations with respect to process model $M_{\sim K}$ at the right-hand side of equation (6) leads to

$$E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}]) = E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}} E_{\mathbf{M}_K}\left([\Delta \mid \theta_{\sim K}, M_{\sim K}]\right)^2 - E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}\left(E_{\mathbf{M}_K}[\Delta \mid \theta_{\sim K}, M_{\sim K}]\right)^2, \quad (7)$$

where subscript $\boldsymbol{\theta}_{\sim K} \mid M_{\sim K}$ of $E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}$ indicates that the expectation is with respect to random parameter $\boldsymbol{\theta}_{\sim K}$ specific to process model combination, $M_{\sim K}$, and the $\Delta \mid \theta_{\sim K}, M_{\sim K}$ term indicates that the system model output, Δ, is for a single combination $M_{\sim K}$ of $\mathbf{M}_{\sim K}$ and for a single realization $\theta_{\sim K}$ of $\boldsymbol{\theta}_{\sim K}$. Applying the law of total expectation again to process model M_K gives

$$\begin{aligned} E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}]) = {} & E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}} E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K|M_K}\left([\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]\right)^2 \\ & - E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}\left(E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K|M_K}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]\right)^2. \end{aligned} \quad (8)$$

If process model uncertainty does not exist, the two expectation terms, $E_{\boldsymbol{M}_{\sim K}}$ and $E_{\boldsymbol{M}_K}$, disappear, and $E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ becomes $E_{\boldsymbol{\theta}_{\sim K}}\left(V_{\boldsymbol{\theta}_K}[\Delta \mid \theta_{\sim K}]\right)$. Accordingly, the total-effect process sensitivity index defined in equation (4) becomes the total-effect parameter sensitivity index defined in equation (5).

When working with $V_{\mathbf{M}_{\sim K}}(E_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ in equation (4), using the definition of variance, $V_{\mathbf{M}_{\sim K}}(E_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ can be evaluated via

$$V_{\mathbf{M}_{\sim K}}(E_{\mathbf{M}_K}[\Delta \mid M_{\sim K}]) = E_{\mathbf{M}_{\sim K}}(E_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])^2 - (E_{\mathbf{M}_{\sim K}} E_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])^2. \quad (9)$$

Similar to the derivation of equations (7) and (8), applying the law of total expectation twice to M_K and $M_{\sim K}$, equation (9) becomes

$$\begin{aligned} V_{\mathbf{M}_{\sim K}}(E_{\mathbf{M}_K}[\Delta \mid M_{\sim K}]) = {} & E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}(E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K|M_K}[\Delta \mid M_{\sim K}, \theta_{\sim K}, \theta_K, M_K])^2 \\ & - (E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}} E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K|M_K}[\Delta \mid M_{\sim K}, \theta_{\sim K}, \theta_K, M_K])^2 \end{aligned} \quad (10)$$

If process model uncertainty does not exist, the two expectation terms, $E_{\boldsymbol{M}_{\sim K}}$ and $E_{\boldsymbol{M}_K}$, disappear, and $V_{\mathbf{M}_{\sim K}}(E_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ becomes $V_{\boldsymbol{\theta}_{\sim K}}\left(E_{\boldsymbol{\theta}_K}[\Delta \mid \theta_{\sim K}]\right)$. Accordingly, the total-effect process sensitivity index defined in equation (4) becomes the total-effect parameter sensitivity index defined in equation (5). Adding equations (8) and (10) together leads to

$$\begin{aligned} & E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}]) + V_{\mathbf{M}_{\sim K}}(E_{\mathbf{M}_K}[\Delta \mid M_{\sim K}]) = V(\Delta) \\ & = E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}} E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K|M_K}\left([\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]\right)^2 \\ & \quad -(E_{\boldsymbol{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}} E_{\boldsymbol{M}_K} E_{\boldsymbol{\theta}_K|M_K}[\Delta \mid M_K, \theta_K, \theta_{\sim K}, M_{\sim K}])^2 \end{aligned} \quad . \tag{11}$$

This is another way of evaluating $V(\Delta)$ based on the definition of variance, i.e., $V(\Delta) = E(\Delta^2) - (E\Delta)^2$, confirming that the deviation of equations (8) and (10) is correct.

2.3. Evaluation of Total-effect Process Sensitivity Index

Although equations (8) and (10) are theoretically equivalent, equation (10) is used in this study to numerically evaluate the total-effect process sensitivity index, because this equation is symbolically the same as equation (6) of Dai et al. (2017a) used to evaluate the first-order process sensitivity index. This enables us to use the same algorithm structure to evaluate the two process sensitivity indices with minor code revisions. For the two expectation terms, $E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}$ and $E_{\boldsymbol{\theta}_K|M_K}$, with respect to model parameters, they are evaluated using Monte Carlo methods. For the two expectation terms, $E_{\mathbf{M}_{\sim K}}$ and $E_{\mathbf{M}_K}$, with respect to process models, they are evaluated using the model averaging approach. Taking the $E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}$ term as an example, $E_{\mathbf{M}_{\sim K}}$ is evaluated using the model averaging approach via

$$E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}(\bullet) = \sum_{\mathbf{M}_{\sim K}} E_{\sim \boldsymbol{\theta}_K|\sim M_K}(\bullet) P(M_{\sim K}) \ , \tag{12}$$

where the dot ($\bullet$) denotes the two terms $(E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K|M_K}[\Delta \,|\, M_{\sim K}, \theta_{\sim K}, \theta_K, M_K])^2$ or $E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K|M_K}[\Delta \,|\, M_{\sim K}, \theta_{\sim K}, \theta_K, M_K]$ at the right-hand side of equation (10), and $P(M_{\sim K})$ is the process model weight of a process model combination $M_{\sim K}$. For the process model weights, it is required that, for process *K*, the weights, $P(M_K)$, of its process models satisfy $\sum P(M_K) = 1$. To evaluate $P(M_{\sim K})$ for a combination of process models, it is assumed that selecting one process model of a process is independent to selecting one process model of another process. With this assumption, $P(M_{\sim K})$ can be evaluated as $P(M_{\sim K}) = \prod_{i=1}^{N_{\sim K}} P(M_{\sim K,i})$, where $N_{\sim K}$ is the number of processes minus one (i.e., number of all processes but process *K*) and $P(M_{\sim K,i})$ is the model weight of a process model. Thinking a system with three processes *A*, *B*, and *C*, each process has two alterative models (i.e., A_1, A_2, B_1, B_2, C_1, and C_2), and the process models have the following weights, i.e., $P(A_1) = P(A_2) = 0.5$, $P(B_1) = 0.7$, $P(B_2) = 0.3$, $P(C_1) = 0.6$, and $P(C_2) = 0.4$. The combination, B_1C_1, of process models B_1 and C_1 has the probability of $P(B_1C_1) = P(B_1) \times P(C_1) = 0.7 \times 0.6 = 0.42$. The model weights can be evaluated in more sophisticated ways, e.g., using Bayesian network as shown in Dai et al. (2019).

Figure 1 is the pseudo code for evaluating the total-effect process sensitivity index based on equation (10). The implementation is computationally straightforward, and four loops are needed, i.e., two loops for the process models $\mathbf{M}_{\sim K}$ and $\mathbf{M}_K$, and two loops for the process model parameters $\boldsymbol{\theta}_{\sim K} \,|\, M_{\sim K}$ and $\boldsymbol{\theta}_K \,|\, M_K$ associated with model $M_{\sim K}$ and M_K, respectively. The total number of model executions is $(N_{\sim K} \times m_{\sim K}) \times (N_K \times m_K)$, where $N_{\sim K}$ is the number of process model combinations for the processes ~*K*, $m_{\sim K}$ is the number of realizations for parameters

372 $\boldsymbol{\theta}_{\sim K} \mid M_{\sim K}$, N_K is the number of process models for process *K*, and m_K is the number of
373 realizations for parameter $\boldsymbol{\theta}_K \mid M_K$. This numerical implementation is computational expensive,
374 and several ways of reducing the computational cost are discussed in Section 5.3 below.

375 **3. Two Numerical Experiments**

376 The first-order and total-effect process sensitivity indices are evaluated in two numerical
377 experiments. One experiment is based on the Sobol's G*-function (Saltelli et al., 2010). Since
378 anaytical solutions of the first-order and total-effect process sensitivity indices can be derived,
379 this experiment is used to verify the computer code that implements the pesuedo code shown in
380 Figure 1 for numerically evaluating the two indices. The other numerical experiment is for
381 groundwater flow modeling based on the synthetic case of Dai et al. (2017a). This experiment
382 shows that the first-order and total-effect process sensitivity indices may give different process
383 ranking based on process importance and influence. We explain, from a hydrologic viewpoint,
384 that the difference is caused by interactions between the recharge, geology, and snowmelt
385 processes involved in the groundwater modeling.

386 **3.1. Sobol's G*-functions**

387 The Sobol's G*-function is defined as (Saltelli et al., 2010)

388
$$G^*(X_1,...,X_k;a_1,...a_k;\delta_1,...,\delta_k;\alpha_1,...,\alpha_k)=\prod_{i=1}^{k} g_i^*$$
$$g_i^*=\frac{(1+\alpha_i)\,|\,2(X_i+\delta_i-I[X_i+\delta_i])-1|^{\alpha_i}+a_i}{1+a_i} \quad , \qquad (13)$$

389 where X_i is the dependent variable following the uniform distributed U[0,1], $a_i \in R^+$ is a
390 positive coefficient, $\delta_i \in [0,1]$ and $\alpha_i > 0$ are shift and curvature parameters, respectively,

391 $I[X_i + \delta_i]$ is the integer part of $X_i + \delta_i$. For this function, analytical expressions of the first-

392 order (S_i) and total-effect (S_{T_i}) parameter sensitivity indices of X_i are (Saltelli et al., 2010)

$$
\begin{aligned}
&S_i = V_i / V \\
&S_{Ti} = V_{Ti} / V \\
&V_i = V_{X_i}(E_{\mathbf{X}_{\sim i}}(Y \mid X_i)) = \frac{\alpha_i^{\,2}}{(1+2\alpha_i)(1+a_i)^2} \\
&V_{Ti} = V_i \prod_{j \neq i} \left(1 + V_j\right) \\
&V = \prod_{i=1}^{k} \left(1 + V_i\right) - 1
\end{aligned}
\quad , \tag{14}
$$

394 where V_i is the variance reduction of g_i^* (the mean of g_i^* is 1). Since the sensitivity indices do

395 not depend on the shift parameter, δ_i, we set $\delta_i = 0$ to simplify the derivation of the analytical

396 solutions of the first-order and total effect process sensitivity indices in Appendix A. The g_i^*

397 function becomes $g_i^* = \frac{(1+\alpha_i)\,|\,2X_i - 1|^{\alpha_i} + a_i}{1 + a_i}$.

398 For simplicity and without loss of generality, we set $k = 3$ to consider three product elements,

399 g_1^*, g_2^*, and g_3^*, and further assume that each element, g_i^*, represents a process. When there is

400 no process model uncertainty, since each element has only one parameter (X_i), equation (14)

401 derived for parameter sensitivity analysis can be used for process sensitivity analysis. For

402 example, if parameter X_1 is the most important parameter among X_1, X_2, and X_3, then process 1

403 (represented by g_1^*) is the most important process among the three processes represented by g_1^*,

404 g_2^*, and g_3^*.

405 **3.1.1. Process Models**

To consider process model uncertainty, we assumed that each process has two process models, g_i^{*1} and g_i^{*2}, with the same form but different *a* and α values. For the α values, we set $\alpha_i^1 = 1$ and $\alpha_i^2 = 2$ for all the three g_i^* elements, where the superscripts (1 and 2) are for the two process models. For each g_i^* element, we set the *a* values as follows $a_1^{\ 1} = 1.5$, $a_1^{\ 2} = 1.2$, $a_2^{\ 1} = 4.2$, $a_2^{\ 2} = 1.8$, $a_3^{\ 1} = 6.5$, and $a_3^{\ 2} = 2.3$. These settings give a total of six process models (two for each process). For the process represented by $g^*_{\ 1}$, its two process models are $g_1^{*1} = \frac{2|2X_1 - 1| + 1.5}{2.5}$ and $g_1^{*2} = \frac{3|2X_1 - 1|^2 + 1.2}{2.2}$. For the process represented by $g^*_{\ 2}$, its two process models are $g_2^{*1} = \frac{2|2X_2 - 1| + 4.2}{5.2}$ and $g_2^{*2} = \frac{3|2X_2 - 1|^2 + 1.8}{2.8}$. For the process represented by $g^*_{\ 3}$, its two process models are $g_3^{*1} = \frac{2|2X_3 - 1| + 6.5}{7.5}$ and $g_3^{*2} = \frac{3|2X_3 - 1|^2 + 2.3}{3.3}$. A combination of the process models gives a total of eight system models, i.e., $g_1^{*1}g_2^{*1}g_3^{*1}$, $g_1^{*1}g_2^{*1}g_3^{*2}$, $g_1^{*1}g_2^{*2}g_3^{*1}$, $g_1^{*1}g_2^{*2}g_3^{*2}$, $g_1^{*2}g_2^{*1}g_3^{*1}$, $g_1^{*2}g_2^{*1}g_3^{*2}$, $g_1^{*2}g_2^{*2}g_3^{*1}$, and $g_1^{*2}g_2^{*2}g_3^{*2}$.

3.1.2. Parameter Sensitivity Index

Table 1 lists the S_i and S_{Ti} values calculated by using the analytical expressions (equation 14) for the three parameters (X_1, X_2, and X_3) of each of the eight system models. Since g_i^* has only one parameter, X_i, the sensitivity for X_i is considered to be the sensitivity for process g_i^*. The parameter sensitivity indices were calculated using SALib (Herman and Usher 2017). The numerical results are almost identical to the analytical results with the maximum difference less than 0.1%, For each system model, S_{Ti} is larger than S_i for all the three processes, indicating the

existence of interactions between the processes. The interactions, however, are not substantial, because the differences between S_i and S_{Ti} are small. In Table 1, the three processes of each system model are ranked based on the S_i and S_{Ti} values, and the ranks based on the two indices are consistent. The rank changes substantially between the system models. Taking the first two system models, $g_1^{*1}g_2^{*1}g_3^{*1}$ and $g_1^{*1}g_2^{*1}g_3^{*2}$, as an example, the most important or influential process is X_1 for system model $g_1^{*1}g_2^{*1}g_3^{*1}$, but it is X_3 for system model $g_1^{*1}g_2^{*1}g_3^{*2}$. The rank change is attributed to the fact that different process models have different α_i and a_i values that determine the V_i values as shown in equation (14). The change in process importance between the system models makes it difficult to determine which process is important, and this manifests the needs of using process sensitivity indices to determine important or influential process under process model uncertainty.

3.1.3. Process Sensitivity Index

To evaluate the first-order and total-effect process sensitivity indices, for each process, we assumed that its two process models have equal weights, i.e., $P(g_i^{*1}) = P(g_i^{*2}) = 0.5$. For these process model weights, the analytical expressions of the first-order (PS_K) and total-effect (PS_{TK}) process sensitivity indices of the *i*-th process are derived in equations (A2), (A11), and (A14) of Appendix A, and they are

$$PS_i = \frac{\frac{1}{2}(V_i^1 + V_i^2)}{V(\Delta)}$$

$$PS_{Ti} = \frac{\frac{1}{8}\sum_{j=1}^{8} V_{Ti}{}^{j}}{V(\Delta)} \quad , \qquad (15)$$

$$V(\Delta) = \frac{1}{8}\sum_{j=1}^{8} V^{j}$$

where V^j is the V of equation (14) for the *j-th* (*j*=1, 2, …, 8) system model, V_i^1 and V_i^2 are V_i of equation (14) for the two process models, and $V_{Ti}{}^j$ is the V_{Ti} of equation (14) for the *j-th* system model. The V, V_i and V_{Ti} values of a single system model are calculated using equation (14) with the α_i^1, α_i^2, $a_i{}^1$, and $a_i{}^2$ values given above.

Table 1 lists the values of the first-order (PS_K) and total-effect (PS_{TK}) process sensitivity indices. Our numerical results are almost identical to the analytical results with the maximum difference less than 0.1%, which again verifies our computer code developed based on the pseudo code shown in Figure 1. PS_{TK} is larger than PS_K for all the three processes, indicating interactions between the processes. The interactions however is not substantial, because the differences between PS_K and PS_{TK} are small. Based on the two indices, process g_1^* is the most important or influential one followed by g_2^*, and g_3^* is the least important or influential one. This process ranking is not surprising, because g_1^* (g_3^*) is ranked as the most (least) important or influential process in five out of eight system models. Based on the values of PS_K and PS_{TK}, none of the three processes is non-influential. Table 1 also suggests that the process sensitivity indices of all the eight system models are not simply weighted average of the parameter sensitivity indices of the individual system models. The theoretical reasons are discussed in Section 5.

3.2. Groundwater Flow Models

This numerical experiment of groundwater flow modeling is revised after Dai et al. (2017a). As shown in Figure 2, the unconfined aquifer has a length of L=10,000 m, and is under a steady state condition. The aquifer is bounded by two constant-head boundaries, and subject to a uniform precipitation. Process model uncertainty exists in the following three system processes: (1) recharge process that converts precipitation to aquifer recharge, (2) geology process that

characterizes hydraulic conductivity of the aquifer, and (3) snow-melt process that estimates snow-melt rate to determine hydraulic head at the east boundary of the domain. The process models are described below.

3.2.1. Process Models and Estimation of Process Sensitivity Indices

Following Dai et al. (2017a), two recharge models (R_1 and R_2) are used to simulate the recharge process that converts precipitation [inch/yr] to groundwater recharge [m/d], and they are

$$\begin{aligned} R_1 &: w = a(P-14)^{0.5} \times 0.0254/365 \\ R_2 &: w = b(P-15.7) \times 0.0254/365 \end{aligned}, \quad (16)$$

where a and b are scaling parameters which are assumed to follow the normal distribution, $N(2.0, 0.4^2)$, and the uniform distribution, $U(0.2, 0.5)$, respectively, P is an annual precipitation which is set as 60 inch/yr. Hydraulic conductivity (K) [m/d] over the domain is parameterized by two geology process models (G_1 and G_2) as follows,

$$\begin{aligned} G_1 &: K \text{ for any } x \\ G_2 &: K = \begin{cases} K_1 \text{ for } x < 7000 \\ K_2 \text{ for } x \geq 7000 \end{cases} \end{aligned}. \quad (17)$$

In model G_1, the aquifer is assumed to be homogeneous, and the hydraulic conductivity follows the lognormal distribution, $LN(2.9, 0.5^2)$. Model G_2 has two zones of hydraulic conductivity separated at the location of x_0 = 7,000 m (Figure 2). The hydraulic conductivities K_1 of zone 1 (x < 7000 m) and K_2 of zone 2 ($x \geq$ 7000 m) are assumed to follow the lognormal distributions, $LN(2.6, 0.3^2)$ and $LN(3.2, 0.3^2)$, respectively.

For the lake boundary, its hydraulic head is set as h_1 = 300 m (Figure 2). For the river boundary, the river stage, h_2, is determined by using the empirical rating curve function below to covert river discharge, Q [m^3/s], into river stage h_2:

$$h_2 = 0.3Q^{0.6} + 289, \quad (18)$$

where 0.3, 0.6, and 289 are arbitrarily chosen coefficients. The river discharge depends on snowmelt runoff via

$$Q = C_{sn} \times M \times SVC \times A \times \frac{0.001}{86400}, \quad (19)$$

where C_{sn} is a runoff coefficient, M is the snow-melt rate [mm/d], SVC is the ratio of snow-covered area to watershed area, A [m^2], and 0.001/86400 is the conversion factor from mm/d to m^3/s. We set C_{sn} = 0.8, SVC = 0.7 and A = 2,000 km^2, and use two models to estimate the daily snow-melt rate based on the degree-day method and the restricted degree-day radiation balance method (Hock, 2003; Kustas et al., 1994; Martinec et al., 2007). The two models are

$$\begin{aligned} M_1 &: M = f_1(T_a - T_m) \\ M_2 &: M = f_2(T_a - T_m) + rR_n \end{aligned}, \quad (20)$$

where f_1 and f_2 [mm·°C^{-1}·d^{-1}) are snowmelt factors that are assumed to follow a normal distribution, $N(3.5, 0.75^2)$, and a normal distribution, $N(2.5, 0.3^2)$, respectively, T_a [°C] is the average temperature for a given day, and T_m [°C] is the temperature threshold (typically set to be 0) when snow melt occurs. The second model considers the effects of surface radiation budget, R_n [W/m^{-2}], and uses a conversion factor r [(mm/d)/(W/m^2)] to convert the energy flux to snowmelt rate. T_a and R_n are set as 7 °C and 80 W/m^{-2}, respectively, and r was assumed to follow a normal distribution, $N(0.3, 0.05^2)$.

A combination of the two recharge process models, two geology process models, and two snowmelt process models leads to a total of eight system models, and they are 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 $R_2G_2M_2$. All process models have equal weights, i.e., $P(R_1) = P(R_2) = 0.5$, $P(G_1) = P(G_2) = 0.5$, and $P(M_1) = P(M_2) = 0.5$. The groundwater discharge per unit width at the location x_0 = 7,000 m is the quantity of interest for the sensitivity analysis, and it is labeled as q_0 in Figure 2. The analytical solution of the discharge, $q(x)$, at any location x, was derived as (Appendix B)

$$q(x) = K_1 \frac{h_1^2 - h_2^2}{2\left(x_0 - \lambda x_0 + \lambda L\right)} - \frac{1}{2} w \frac{x_0^2 - \lambda x_0^2 + \lambda L^2}{x_0 - \lambda x_0 + \lambda L} + wx \ , \tag{21}$$

where $\lambda = K_1 / K_2$. For the geology model (G_1) with homogenous hydraulic conductivity, i.e., $K = K_1 = K_2$, equation (21) reduces to

$$q(x) = K \frac{h_1^2 - h_2^2}{2L} - \frac{1}{2} wL + wx \ , \tag{22}$$

which is the well-known solution available in hydrogeology textbooks. The analytical solutions help us understand the process interactions. For example, equation (22) indicates that the recharge process (w) does not interact with the geology (K) and snowmelt (h_2) processes. More discussion on process interactions in the sensitivity analysis is given in Section 4.

Using the analytical solutions of $q(x)$, the first-order and total-effect sensitivity indices for parameters and processes are evaluated using Monte Carlo simulations based on the pseudo code shown in Figure 1. For the simplicity of programming to simultaneously estimate all the sensitivity indices (parameter and process), 300 parameter realizations are generated for each process model. The total number of model executions is thus 216,000,000 = 2 × 300 × 2 × 300 × 2 × 300 for the two recharge process models, two geology process models, and two snowmelt

process models. This is equivalent to ($2^2 \times 300^2$) × (2 × 300) for the number of model executions, $(N_{\sim K} \times m_{\sim K}) \times (N_K \times m_K)$, given in Figure 1. It should be noted that the 216 millions of model execution is only for the convenience of programming. As shown in Figure C1 in Appendix C, the two process sensitivity indices converge after several millions of model executions.

3.2.2. Parameter Sensitivity Index

Table 2 lists the numerical results of the first-order (S_i) and total-effect (S_{Ti}) parameter sensitivity indices for the individual models. If a process model has more than one parameter (e.g., K_1 and K_2 associated with process model G_2, and f_2 and r associated with process model M_2), the indices are evaluated for the grouped parameters using the methods introduced by Saltelli et al. (2008). Since the parameter sensitivity indices are evaluated for all parameters of each process, within a system model, the indices measure process importance or influence of each process. Taking system model $R_2G_2M_2$ as an example, S_i of parameter b measures importance of the recharge process (equation (16)), S_i of parameters K_1 and K_2 measures importance of the geology process (equation (17)), and S_i of parameters f_2 and r measures importance of the snowmelt process (equation (20)).

Table 2 shows that both the first-order and total-effect parameter sensitivity indices change between the individual system models. As a result, any of the three processes can be ranked as the most important/influential or least important/influential process. For example, in system model $R_1G_1M_1$, the recharge process is identified as the least important/influential process and the snowmelt process as the most important/influential process, whereas it is the opposite in system model $R_2G_2M_2$. This indicates again the needs of using the process sensitivity indices to identify the important/influential processes for all system models to address process model uncertainty. For models $R_2G_1M_1$ and $R_2G_1M_2$, the two parameter sensitivity indices give different

process ranking for each system model, as shown in Table 2. This is different from the case of the Sobol's G*-function that the first-order and total-effect process sensitivity indices give the same process ranking. We discuss the two process sensitivity indices separately below.

3.2.3. First-Order Process Sensitivity Index

In Table 2, the first-order process sensitivity index, PS_K, suggests that the snowmelt process is the most important process followed by the recharge process and then the geology process. The reasons of the ranking can be explained by using Figure 3 that plots the probability density functions (PDFs) of the specific discharge q_0 for the six cases of fixing the snowmelt model to either M_1 or M_2, fixing the recharge model to either R_1 or R_2, and fixing the geology model to either G_1 or G_2. In each plot, the PDF in black is based on the Monte Carlo simulation for all the eight system models and their parameters, and the other four PDFs in other colors are for the four system models after one process model is fixed. Since the PDFs are of Gaussian shapes, it is reasonable to use variance to measure the uncertainty of q_0, i.e., the PDF spreading. Figure 3 shows that, whenever a process model is fixed, the PDFs of the remaining four system models are narrower than the original PDF in black, indicating uncertainty reduction. When the snowmelt model (M) is fixed to either M_1 (Figure 3a) or M_2 (Figure 3b), the PDFs of the remaining four models are more concentrated than the PDFs of the remaining four models shown in Figures 3c – 3f when the recharge and the geology process models are fixed. It suggests that the uncertainty reduction is the largest when the snowmelt process model is fixed, which explains why PS_K ranks the snowmelt process as the most important one.

According to equation (1), the remaining uncertainty is measured by $E_{\mathbf{M}_K}(V_{\sim\mathbf{M}_K}[\Delta \mid M_K])$, and it becomes $E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_k \mid M_K}(V_{\sim\mathbf{M}_K}[\Delta \mid \theta_K, M_K])$ to consider parametric uncertainty of each M_K. The

569 $E_{\boldsymbol{\theta}_k|\mathbf{M}_K}(V_{\sim\mathbf{M}_K}[\Delta|\theta_K, M_K])$ values are shown in Figure 3, and $E_{\mathbf{M}_K}E_{\boldsymbol{\theta}_k|\mathbf{M}_K}(V_{\sim\mathbf{M}_K}[\Delta|\theta_K, M_K])$ can be
570 calculated by using weighted average for each process. Figure 3 indicates that the variance
571 reduction is the largest when the snowmelt process model (M) is fixed to M_1 or M_2, because their
572 remaining variances of 0.61 and 0.8 are the smallest among the six plots. The corresponding
573 remaining variance $E_{\mathbf{M}_K}E_{\boldsymbol{\theta}_k|\mathbf{M}_K}(V_{\sim\mathbf{M}_K}[\Delta|\theta_K, M_K])$ is 0.5 × 0.61 + 0.5 × 0.88 = 0.75, where 0.5 is
574 the model weight of M_1 or M_2. For the recharge process, the remaining variance
575 $E_{\mathbf{M}_K}E_{\boldsymbol{\theta}_k|\mathbf{M}_K}(V_{\sim\mathbf{M}_K}[\Delta|\theta_K, M_K])$ is 0.5 × 1.75 + 0.5 × 1.79 = 1.77; for the geology process, the
576 remaining variance $E_{\mathbf{M}_K}E_{\boldsymbol{\theta}_k|\mathbf{M}_K}(V_{\sim\mathbf{M}_K}[\Delta|\theta_K, M_K])$ is 0.5 × 2.43 + 0.5 × 1.43 = 1.93. The
577 remaining variance heavily depends on model averaging weights. If the equal weights of G_1 and
578 G_2 are changed to unequal weights of $P(G_1)$ = 0.2 and $P(G_2)$ = 0.8, the remaining variance is
579 changed to 0.2 × 2.43 + 0.8 × 1.43 = 1.63. Since it is smaller than 1.77 for fixing the recharge
580 process model, the geological process will be considered more important than the recharge
581 process. Evaluating the impacts of model weights on process ranking is beyond the scope of this
582 study, but is warranted in a future study.

583 **3.2.4. Total-Effect Process Sensitivity Index**

584 The PS_{TK} values in Table 2 suggest that the snowmelt process is the most influential process,
585 followed by the geology process and then by the recharge process. The reasons for this ranking
586 can be explained by Figure 4 that plots the PDFs of q_0 for evaluating the total-effect process
587 sensitivity index. Each roll of Figure 4 corresponds to a situation of fixing the process models of
588 all processes but one. Taking Figures 4 (a1) – (a4) for the snowmelt process (M) as an example,
589 the process models of the recharge (R) and geology (G) processes are fixed to R_1G_1, R_1G_2, R_2G_1,
590 and R_2G_2 in each plot, but the snowmelt process model (M) varies between M_1 and M_2 in each
591 plot. The remaining variance is the largest for the snowmelt process, M, because the PDFs of M_1

592 and M_2 are separated from each other. The remaining variance is the smallest for the recharge
593 process, R, because the PDFs of R_1 and R_2 (the third row of Figure 4) have substantial overlaps.
594 These PDFs qualitatively explain why the snowmelt (recharge) process is considered as the most
595 (least) influential process by PS_{TK}.

596 A quantitative understanding on the process ranking given by PS_{TK} can be gained by
597 examining the $E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ term used to define PS_{TK} in equation (4). This term measures
598 the remaining variance caused by alternative models of a process when the rest of the process
599 models are fixed. This term is extended as $E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim k}|\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid \theta_{\sim K}, M_{\sim K}])$ to consider
600 parametric uncertainty, and the values of $E_{\boldsymbol{\theta}_{\sim k}|\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid \theta_{\sim K}, M_{\sim K}])$ are shown in Figure 4. For
601 the snowmelt process, the value of $E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ is 0.25 × 2.20 + 0.25 × 1.05 + 0.25 ×
602 2.20 + 0.25 × 1.05 = 1.63, where 0.25 = 0.5 × 0.5 is the model averaging weight of any
603 combination of two process models of the recharge and geology processes. For the geology
604 process, the value of $E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ is 0.25 × 0.30 + 0.25 × 0.55 + 0.25 × 0.33 + 0.25 ×
605 0.61 = 0.45; for the recharge process, the value of $E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ is 0.25 × 0.23 + 0.25 ×
606 0.23 + 0.25 × 0.37 + 0.25 × 0.37 = 0.30. These results quantitatively explain why the snowmelt
607 (recharge) process is ranked by PS_{TK} as the most (least) influential one.

608 **4. Parameter Interactions and Process Interactions**

609 **4.1 Parameter Interactions**

610 Since parameter interactions are discussed intensively in the books of Saltelli et al. (2004,
611 2008), discussing parameter interactions is not a focus of this study. We thus use system model
612 $R_2G_1M_1$ as an example to illustrate parameter interactions before discussing process interactions,
613 which is a focus of this study. According to Saltelli et al. (2004), for the three processes of the

614 system model, the variance of the groundwater discharge per unit width (q_0) at x_0 = 7,000 m can
615 be decomposed as $V = V_b + V_K + V_{f1} + V_{K,b} + V_{f1,b} + V_{f1,K} + V_{f1,b,K}$, where parameters b, K, and f_1
616 are associated with the recharge, geology, and snowmelt processes, respectively. Table 3 lists the
617 numerical results (based on Monte Carlo simulations) of the three fist-order variance terms (V_i),
618 the three second-order variance (V_{ij}) terms, and the only third-order variance (V_{ijk}) term. Table 3
619 shows that K and f_1 (or equivalently the geology and snowmelt processes) interact, whereas b
620 does not interact with either K or f_1 (i.e., the recharge process does not interact with the geology
621 process or the snowmelt process). This is straightforward according to equation (22),
622 $q(x) = K\frac{h_1^2 - h_2^2}{2L} - \frac{1}{2}wL + wx$, for model $R_2G_1M_1$ with homogeneous K, because the recharge (w)
623 does not interact with the geology (K) or snowmelt (h_2). Table 3 also lists the first-order, second-
624 order, third-order, and total-effect sensitivity indices. The total-effect sensitivity index of a
625 parameter is the sum of the other three indices. Taking parameter b of the recharge process as an
626 example, $S_{Tb} = S_b + S_{K,b} + S_{f_1,b} + S_{b,K,f_1}$. Because b does not interact with either K or f_1, $S_{Tb} = S_b$.
627 Considering the parameter interactions, we have $S_{TK} = S_K + S_{K,f_1}$ and $S_{Tf_1} = S_{f_1} + S_{K,f_1}$ (these
628 theoretical conditions are not fully satisfied in Table 3 due to the use of Monte Carlo simulations
629 for numerical calculation). Although the first-order index (S_i) indicates that the recharge process
630 is more important than the geology process, because of the interactions between K and f_1, the
631 total-effect index (S_{Ti}) indicates that the recharge process is less influential than the geology
632 process. The process interactions affect process importance/influence in a similar manner.

633 **4.2 Process Interactions**

Based on the variance decomposition of the total variance $V(\Delta)$, the total-effect process sensitivity index of process K is the sum of the first-order process sensitivity index, the second-order and higher-order process sensitivity indices, i.e.,

$$PS_{TK} = PS_K + \sum_K \sum_V PS_{K,V} + \sum_K \sum_V \sum_W PS_{K,V,W} + \ldots, \tag{23}$$

where PS_{KV} is the second-order process sensitivity index to consider the interaction between processes K and V, PS_{KVW} is the third-order process sensitivity index to consider the interaction between processes K, V and W. The second-order process sensitivity index is defined as:

$$PS_{K,V} = \frac{V_{K,V}(E_{\sim(K,V)}[\Delta \mid K,V])}{V(\Delta)} - \frac{V_K(E_{\sim K}[\Delta \mid K])}{V(\Delta)} - \frac{V_V(E_{\sim V}[\Delta \mid V])}{V(\Delta)}, \tag{24}$$

The $V_{K,V}(E_{\sim(K,V)}[\Delta \mid K,V])$ term at the right hand side of equation (24) considers the joint effect of variance reduction by simultaneously fixing processes K and V, and the $V_K(E_{\sim K}[\Delta \mid K])$ and $V_V(E_{\sim V}[\Delta \mid V])$ terms consider the variance reduction by fixing process K and process V, respectively. Extension of these terms to consider parametric uncertainty is straightforward and similar to the procedure given in Section 2.2. The extension of $V_{K,V}(E_{\sim(K,V)}[\Delta \mid K,V])$ is

$$\begin{aligned} &V_{K,V}(E_{\sim(K,V)}[\Delta \mid K,V]) = V_{\mathbf{M}_{K,V}}(E_{\mathbf{M}_{\sim(K,V)}}[\Delta \mid M_{K,V}]) \\ &= E_{\mathbf{M}_{K,V}} E_{\boldsymbol{\theta}_{K,V}|M_{K,V}} (E_{\mathbf{M}_{\sim(K,V)}} E_{\boldsymbol{\theta}_{\sim(K,V)}|M_{\sim(K,V)}}[\Delta \mid M_{K,V}, \theta_{K,V}, M_{\sim(K,V)}, \theta_{\sim(K,V)}])^2, \\ &-(E_{\mathbf{M}_{K,V}} E_{\boldsymbol{\theta}_{K,V}|M_{K,V}} E_{\mathbf{M}_{\sim(K,V)}} E_{\boldsymbol{\theta}_{\sim(K,V)}|M_{\sim(K,V)}}[\Delta \mid M_{K,V}, \theta_{K,V}, M_{\sim(K,V)}, \theta_{\sim(K,V)}])^2 \end{aligned} \tag{25}$$

where $\mathbf{M}_{K,V}$ is a set of combined process models for processes K and V, $M_{K,V}$ is a single model in $\mathbf{M}_{K,V}$, $\boldsymbol{\theta}_{K,V} \mid M_{K,V}$ denotes a set of parameter realizations associated with $M_{K,V}$, and $\theta_{K,V}$ is a realization in $\boldsymbol{\theta}_{K,V} \mid M_{K,V}$. Similarly, $\mathbf{M}_{\sim(K,V)}$, $M_{\sim(K,V)}$, $\boldsymbol{\theta}_{\sim(K,V)} \mid M_{\sim(K,V)}$, and $\theta_{\sim(K,V)}$ are the counterparts for the combined process models other than those for K and V. Equation (25) can

652 be numerically evaluated using the Monte Carlo and model averaging methods described in
653 Section 2.3. The third-order and higher-order process interaction effects can be defined and
654 evaluated in the same manner, and are not given here.

655 Table 4 lists the numerical results (based on Monte Carlo simulations) of the three first-order
656 variance terms (V_K), the three second-order variance ($V_{K,V}$) terms, and the only third-order
657 variance ($V_{K,V,W}$) term with consideration of both process model uncertainty and process
658 parametric uncertainty. The second-order variance term, $V_{M,R}$, for the snowmelt and recharge
659 processes is zero, because w does not interact with h_2 in any process model combinations. This is
660 straightforward by examining equation (21),
661 $q(x) = K_1 \frac{h_1^2 - h_2^2}{2(x_0 - \lambda x_0 + \lambda L)} - \frac{1}{2} w \frac{x_0^2 - \lambda x_0^2 + \lambda L^2}{x_0 - \lambda x_0 + \lambda L} + wx$, of the two-zone geologic model or
662 equation (22), $q(x) = K \frac{h_1^2 - h_2^2}{2L} - \frac{1}{2} wL + wx$, of the homogenous geology model. In the two
663 equations, recharge process (w) does not interact with snowmelt process (h_2). The $V_{M,G}$ term for
664 the snowmelt and geology processes is large, and this is attributed to the first term (i.e., the
665 multiplication of K and h_2) at the right hand side of equation (21) or (22). The third-order
666 variance, $V_{M,R,G}$, for the three processes is theoretically zero, because the analytical solution does
667 not have any terms involving all the three variables, h_2, w, and K, of the three processes. Because
668 of the interaction between the snowmelt and geology processes, the values of the total-effect
669 process sensitivity index are different from those of the first-order process sensitivity index, and
670 this is fully explained by Table 4. The snowmelt process is not only the most important process
671 (measured by PS_K) but also the most influential process (measured by PS_{TK}). While the recharge
672 process is the second most important process, it becomes the least influential process, due to its
673 negligible interactions with the geology and snowmelt processes.

Since the analytical solutions of groundwater discharge per unit width (equations 21 and 22) are for any location x in the model domain, the first-order, second-order, third-order, and total-effect process sensitivity indices are evaluated for the entire domain (the third-order process sensitivity index is theoretically zero). This helps understand process interactions throughout the model domain. Figure 5(a) plots the first-order (PS_K) and total-effect (PS_{TK}) process sensitivity indices, and Figure 5(b) plots the second-order (PS_{KV}) process sensitivity indices. Figure 5(b) shows that $PS_{R,M}$ (for the interaction between recharge and snowmelt processes) is zero and $PS_{R,G}$ (for the interaction between recharge and geology processes) is negligible over the entire domain, because of zero interaction between recharge and snowmelt processes and negligible interaction between recharge and geology processes, as discussed above. Accordingly, PS_R (the first-order process sensitivity index of recharge process) is visually identical to PS_{TR} (the total-effect process sensitivity index of recharge process), as shown in Figure 5(a). In Figure 5(b), $PS_{G,M}$ (for interaction between geology and snowmelt processes) is significant ranging between 10% and 17% (with the largest interaction at the center of the model domain), because of the strong interaction between the geology and snowmelt processes, as discussed above. The $PS_{G,M}$ values explain the difference between PS_G and PS_{TG} for the geology process and between PS_M and PS_{TM} for the snowmelt process shown in Figure 5(a). Due to the interaction between the geology and snowmelt processes, the rank of important process (measured by PS_K) may differ from the rank of influential process (measured by PS_{TK}), and the differences change over the domain. All the discussion above on process interactions is based on the analytical solutions of q_0 that directly reveal the process interactions.

4.3. General Guidelines on Impacts of Process Interactions

696 We need to distinguish between process interactions and impacts of process interactions on
697 identifying important or influential processes. Both the first-order process sensitivity index (PS_K)
698 of Dai et al. (2017) and the total-effect process sensitivity index (PS_{TK}) of this study explicitly
699 consider process interactions, because the two indices are derived for $\Delta | \theta_K, M_K, \theta_{\sim K}, M_{\sim K}$,
700 which is system model output, Δ, simulated by model M_K (and its parameters θ_K) of process K
701 interacting with model $M_{\sim K}$ (and its parameters $\theta_{\sim K}$) of other processes than K. However, for the
702 first-order process sensitivity index, PS_K, the process interactions do not have impacts on
703 identification of important processes, because PS_K is for the single process K based on the
704 $V_{\mathbf{M}_K}(E_{\mathbf{M}_{\sim K}}[\Delta | M_K])$ term used in Equation (6) of Dai et al. (2017). The total-effect process
705 sensitivity index, PS_{TK}, is different from PS_K, because
706 $PS_{TK} = PS_K + \sum_K \sum_V PS_{K,V} + \sum_K \sum_V \sum_W PS_{K,V,W} + \ldots$ of Equation (23) has terms of process
707 interactions, e.g., the interaction between processes K and V and the interaction between
708 processes K, V, and W. As a result, for PS_{TK}, process interactions have impacts on identification
709 of influential processes.

710 Based on our current understanding of the theoretical impacts of process interactions and on
711 the empirical results of the groundwater flow modeling discussed above, we provide the general
712 guidelines below on when process interactions impact influential process identification and how
713 to evaluate the impacts:

714 (1) If a study considers only two processes denoted as A and B, then $PS_{TA} = PS_A + PS_{A,B}$ and
715 $PS_{TB} = PS_B + PS_{A,B}$. Because the $PS_{A,B}$ term is used in both PS_{TA} and PS_{TB}, the process
716 influence ranking based on PS_{TA} and PS_{TB} is the same as the process importance ranking

based on PS_A and PS_B. As a result, there is no need to consider the interactions between processes *A* and *B*.

(2) If a study considers more than two processes, process interactions should be considered for identifying influential processes. One should not blindly rely on the process sensitivity indices, but starts with gaining physical understanding of process interactions. If a model output is determined by two processes in a nonlinear way, the interaction between the two processes impacts process influence ranking. An example of this is the $K_1 \frac{h_1^2 - h_2^2}{2(x_0 - \lambda x_0 + \lambda L)}$ or $K \frac{h_1^2 - h_2^2}{2L}$ term in the analytical solution of the groundwater flow problem given in Section 4.2 above.

(3) If an analytical solution or a physical understanding is unavailable for a complex hydrologic system, one may use the process sensitivity indices in the following procedure to investigate process interactions:

(3a) Calculate the first-order process sensitivity index, PS_K, for each process, and the sum of PS_K should be less than one, as shown in Table 2. If the sum is close to one, then it indicates that process interactions is weak.

(3b) Calculate the total-effect process sensitivity index, PS_{TK}, for each process, and the sum of PS_{TK} should be larger than one, as shown in Table 2. If the sum is close to one, then it indicates that process interactions is weak.

(3c) If PS_{TK} is larger than PS_K for a process, then it indicates that the process interacts with other processes. One can calculate the second-order process sensitivity index of the process to investigate the process interactions, as shown in Table 4. Calculating the

second-order process sensitivity index should start with the process whose PS_{TK} and PS_K values are significantly different. The results of (3c) can help advance our physical understanding of process interactions.

Steps (3a) – 3(c) suggest the importance of estimating both PS_K and PS_{TK} to appreciate the difference between the first-order impact (PS_K) of process K alone on model output, Δ, and the overall total-effect impact (PS_{TK}) of process K itself and interactions with other processes on Δ. More comprehensive and quantitative guidelines are beyond the scope of this study, but are warranted in a future study.

5. Discussion

5.1. Relation between Parameter and Process Sensitivity Indices

This section shows that the process sensitivity indices of multiple system models are not simply a weighted average of parameter sensitivity indices of individual system models for two reasons. The first reason is that the total variance, $V(\Delta)$ (the denominator in equations (2) and (4)) of multiple system models is more than a weighted average of the variance of the individual models. It is derived in Appendix A that, with consideration of both parametric and model uncertainty, the mean and variance of Δ are (equation (A7) of Appendix A)

$$E(\Delta) = \sum_{N_K}\sum_{N_{\sim K}} E_{\boldsymbol{\theta}_K|M_K} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]P(M_{\sim K})P(M_K)$$

$$V(\Delta) = \underbrace{\sum_{N_K}\sum_{N_{\sim K}} V_{\boldsymbol{\theta}_K|M_K, \boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]P(M_{\sim K})P(M_K)}_{\textit{within-model variance}} \quad , \qquad (26)$$

$$+ \underbrace{\sum_{N_K}\sum_{N_{\sim K}} \left(E_{\boldsymbol{\theta}_K|M_K} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] - E(\Delta)\right)^2 P(M_{\sim K})P(M_K)}_{\textit{between-model variance}}$$

where $E(\Delta)$ and $V(\Delta)$ are also called posterior mean and posterior variance, respectively, in the context of model averaging for multiple system models, and $E_{\boldsymbol{\theta}_K|M_K} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]$ $V_{\boldsymbol{\theta}_K|M_K, \boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]$ are the mean and variance, respectively, of a single system model consisting of process models M_K and $M_{\sim K}$. The expression of $V(\Delta)$ includes not only the weighted average of the variance of individual models (the within-model variance term) but also the variance between the means of individual models (the between-model variance term).

The between-model terms also exist for the variance-reduction term $V_{\mathbf{M}_K}(E_{\mathbf{M}_{\sim K}}[\Delta \mid M_K])$ used to define the first-order process sensitivity index in equation (2) and for the variance-left term $E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ used to define the total-effect process sensitivity index. This is the other reason that the process sensitivity indices of multiple system models are not simply a weighted average of parameter sensitivity indices of individual system models. For the variance-reduction term $V_{\mathbf{M}_K}(E_{\mathbf{M}_{\sim K}}[\Delta \mid M_K])$, it is expanded as (Equation (A9) in Appendix A)

$$
\begin{aligned}
&V_{\mathbf{M}_K}(E_{\mathbf{M}_{\sim K}}[\Delta \mid M_K]) = \\
&= \underbrace{\sum_{N_K}\sum_{N_{\sim K}} V_{\boldsymbol{\theta}_K|M_K} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] P(M_{\sim K}) P(M_K)}_{\textit{within-model variance reduction}} \\
&+ \underbrace{\sum_{N_K}\sum_{N_{\sim K}} \left(\begin{array}{l} \left(E_{\boldsymbol{\theta}_K|M_K} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] - E(\Delta) \right)^2 \\ -E_{\boldsymbol{\theta}_K|M_K}\left(E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] - E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] \right)^2 \end{array} \right) P(M_{\sim K}) P(M_K)}_{\textit{between-model variance reduction}}
\end{aligned}
\quad . \quad (27)
$$

The $V_{\boldsymbol{\theta}_K|M_K} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]$ term at the right-hand side term of equation (27) is the variance-reduction term of a single system model (consisting of M_K and $M_{\sim K}$) when its parameters, $\boldsymbol{\theta}_K \mid M_K$, of process model M_K are fixed. Therefore, the first item (called *within-model variance reduction*) at the right-hand side of equation (27) is a weighted average of the variance reduction for all the system models. The second term at the right-hand side of equation

(27) is called *between-model variance reduction*, and it measures the variance of the variance-reduction between multiple system models. Note that the between-model variance reduction could be negative. For the variance-left term $E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$, instead of using the equations in Section 2.2, it is expanded as equation (A13) in Appendix A, i.e.,

$$\begin{aligned} & E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}]) = E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}]) \\ & = \underbrace{\sum_{N_{\sim K}} \sum_{N_K} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}} V_{\boldsymbol{\theta}_K|M_K}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] P(M_K) P(M_{\sim K})}_{\textit{within-model variance left}} \\ & + \underbrace{\sum_{N_{\sim K}} \sum_{N_K} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}} \left(E_{\boldsymbol{\theta}_K|M_K}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] - E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K|M_K}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] \right)^2 P(M_K) P(M_{\sim K})}_{\textit{between-model variance left}} \end{aligned} \quad . \quad (28)$$

The $E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}} V_{\boldsymbol{\theta}_K|M_K}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]$ term is the variance-left term of a single system model (consisting of $M_{\sim K}$ and M_K) when its parameters, $\boldsymbol{\theta}_{\sim K} \mid M_{\sim K}$, for all process model but those of process K are fixed. Therefore, the first item (called *within-model variance left*) at the right-hand side of equation (28) is a weighted average of the variance left for all the system models. The second term is called *between-model variance left*, and it measures the variance of the variance-left between multiple system models. Because the three between-model terms for $V(\Delta)$, $V_{\mathbf{M}_K}(E_{\mathbf{M}_{\sim K}}[\Delta \mid M_K])$, and $E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ in equations (26) – (28), the first-order and total-effect process sensitivity indices for all the system models are not simply the weighted average of first-order and total-effect parameter sensitivity indices of individual system models.

5.2. Decomposition of Variance Reduction and Variance Left

Equation (27) for the first-order process sensitivity index is useful for understanding whether the variance reduction (after a process model is fixed) is from the within-model or the between-model variance reduction. This is further useful for understanding whether process ranking of a single system model is the same as process ranking of multiple system models, in that the two

792 sets of ranking are not likely to be the same when the between-model variance reduction is
793 substantially larger than the within-model variance reduction. In an extreme case that the
794 between-model variance reduction is zero, the two sets of ranking may still be different,
795 depending on process model weights. This is also true for the ranking given by the total-effect
796 parameter and process sensitivity indices, according to equation (28).

797 Figures 6(a) – 6(c) plot the total variance reduction and its within-model and between-model
798 terms (calculated using equation (27)) for the entire domain when recharge, geology, and
799 snowmelt process models are fixed. Figures 6(d) – 6(f) do the same for the variance left when
800 recharge, geology, and snowmelt process models are left unfixed. The variance reduction/left
801 terms are a constant for the geology and snowmelt processes, but vary in space for the recharge
802 process. This is not surprising, because equations (21) and (22) show that the recharge variable w
803 is multiplied with x to determine groundwater discharge q, but geology variable K and snowmelt
804 variable h_2 are not. For the recharge process, Figures 6(a) and 6(d) show that the within-model
805 terms are larger than the between-model terms over the entire domain except at the center of the
806 domain and its vicinity, where the two terms equal zero. The reason that the between-model
807 terms are small can be explained by examining the PDFs shown in Figures 3 and 4 for the
808 location of x = 7,000 m. Taking Figure 4 as an example, Figures 4(c1) – 4(c4) show that, in each
809 figure, the two PDFs are close, indicating that the between-model variance left is small for each
810 $M_{\sim K}$. Figures 6(c) and 6(f) show that the between-model terms are larger than the within-model
811 terms for the snowmelt process, and this can also be explained by examining Figures 3 and 4.
812 Taking Figure 4 as an example again, the two PDFs in each plot of Figures 4(a1) – 4(a4) are
813 separated, and this explains why the between-model term is larger than the within-model term
814 for the snowmelt process. It also should be noted that Figure 6(b) shows the between-model

variance reduction of geology process is negative, resulting in the total variance reduction is smaller than the within-model variance reduction. These analyses suggest that identifying important or influential processes under model and parametric uncertainty is complicated, as it depends on multiple factors discussed at the beginning of the section.

5.3. Limitations and Future Research

The total-effect process sensitivity index is subject to three limitations. One limitation is that the current numerical implementation is computationally expensive, because it requires a nested sampling structure for process models $M_{\sim K}$ and M_K, as shown in Figure 1. If N samples are used for $M_{\sim K}$ and N samples are used for M_K, the nested sampling structure requires N^2 model executions, which is computationally unaffordable for most of practical problems. There exist solutions to this problem, such as the binning approach used by Dai et al. (2017a,b) to break the nested sampling structure, the moment equation method used by Dell'Oca et al. (2020) to evaluate mean and variance without using MC simulations, and/or surrogate approaches that have been used widely to reduce computational cost of model simulations. Dai et al. (2021, under review) recently developed a quasi-MC method for estimating the first-order process sensitivity index of Dai et al. (2017), and the method reduces the number of model execution from the order of N^2 to the order of $2N$. It is relatively straightforward to apply the quasi-MC method for estimating the total-effect process sensitivity index of this study, and this is warranted in a future study.

The second limitation is that, although the total-effect process sensitivity index addresses process interactions in a physical sense, the model averaging method used in this study for deriving the index assumes that the process models are probabilistically independent. This is a drawback because it neglects process interactions and correlations when determining process

838 model weights. A related issue is the hierarchical relation between a process and its sub-
839 processes brought up by Martyn Clark in a personal communication. For example, a vadose zone
840 flow process may be represented by gravity drainage and Richards' equation at a coarse level; at
841 a fine level, there may be multiple soil retention curves associated with the Richards' equation.
842 The fine-level uncertainty depends on the coarse-level uncertainty, in that the probability of
843 choosing a soil retention curve depends on the probability of choosing the Richards' equation.
844 One possible solution to these problem may be using the Bayesian network to incorporate known
845 process interactions, as shown in Dai et al. (2019). If new process interactions are found by using
846 the first-order and total-effect process sensitivity indices, the new interactions may be
847 incorporated into the Bayesian network in an iterative manner. Resolving these problems also
848 depends on progress on developing a more comprehensive modular modeling framework for
849 process-based modeling.

850 A last limitation is that the current study uses prior parameter distributions and process model
851 weights without conditioning them on observations of system states. When the observations are
852 available, it is more realistic to use them for estimating posterior parameter distributions and
853 process model weights. Redefining the first-order and total-effect process sensitivity indices
854 using the posterior distributions and weights is straightforward, as discussed in Dai et al. (2017a).
855 The key challenge is to estimate the posterior parameter distributions and posterior process
856 weights due to high computational cost of Markov chain Monte Carlo simulations (Liu et al.,
857 2016; Elshall and Ye, 2019). Estimating the posterior process weights are particularly
858 challenging because it is different from estimating posterior system model weights and has not
859 been attempted. Conditioning sensitivity analysis on observations of system states is warranted
860 in a future study.

6. Conclusions

Within the framework of variance-based sensitivity analysis, this study derives the total-effect process sensitivity index by integrating Sobol's total-effect parameter sensitivity index and the model averaging method to consider uncertainty in both process models and process model parameters. Different from the first-order process sensitivity index defined by Dai et al. (2017a) that identifies important processes in terms of variance reduction, the total-effect process sensitivity index identifies influential processes in terms of not only process importance but also process interactions that jointly influence a system model output. Because of the process interactions, the process ranking given by the total-effect process sensitivity index may differ from that given by the first-order process sensitivity index. This is why we develop the total-effect process sensitivity index in this study, especially for identifying non-influential processes, as an unimportant process may be influential if its interactions with other processes are substantial. It is often necessary to estimate both first-order and total-effect process sensitivity indices for understanding the difference between the first-order impact of one process alone on a model output and the overall total-effect impact of the process itself and its interactions with other processes on the model output.

For the purpose of model development and improvement, the first-order process sensitivity index of Dai et al. (2017a) falls into the processes prioritization setting, and can be used to identify important processes. In other words, importance of a process is measured in terms of reducing model output uncertainty achieved by fixing the process, i.e., fixing the process representation to a process model. Identifying important processes is necessary to determine how to spend limited resources for model improvement, in that the important processes corresponding to large uncertainty reduction should receive high priority for more laboratory experiments or

field campaigns. The total-effect process sensitivity index of this study falls into the processes fixing setting that is to screen out non-influential processes so that limited resources are not spent on such processes. If a process itself is not important and does not interact with other processes, this process is considered to be non-influential, and its representation can be fixed to any process model during model development and improvement.

The numerical experiments using the Sobol's G*-function and a simple three-process groundwater model demonstrate that process importance or influence may change across system models, indicating the needs of using the process sensitivity indices to explicitly consider uncertainty in process models and process model parameters. The first-order and total-effect process sensitivity indices of multiple system models are not simply weighted averages of the first-order and total-effect parameter sensitivity indices of individual system models. This is demonstrated in the numerical experiment using the Sobol's G*-function, for which analytical solutions of all the sensitivity indices are derived so that numerical errors do not affect the demonstration. Because of the analytical solutions, the numerical experiment also serves the purpose of verifying the numerical code developed to evaluate all the sensitivity indices. The numerical experiment of groundwater modeling further demonstrates that the first-order and total-effect process sensitivity indices give different process ranking due to process interactions. For a complex hydrologic system that cannot be solved analytically, the two process sensitivity indices should be a useful tool for understanding process interactions by evaluating differences between the ranking of important processes and influential processes. In addition, examining the values of first-order, higher-order, and total-effect process sensitivity indices may be a tool to explore process interactions, as explained in the general procedure given in Section 4.3.

Acknowledgments

907 The first two authors were supported in part by U.S. Department of Energy grant DE-SC0019438
908 and National Science Foundation grant EAR-1552329. The third author was supported by the
909 River Corridor Science Focus Area (SFA) project at the Pacific Northwest National Laboratory
910 (PNNL), as part of the Subsurface Biogeochemical Research Program (SBR) of the U.S.
911 Department of Energy (DOE), Office of Biological and Environmental Research (BER). PNNL
912 is operated for the DOE by Battelle Memorial Institute under contract DE-AC05-76RL01830.
913 ORNL is managed by UT-Battelle, LLC, for the DOE under contract DE-AC05-1008
914 00OR22725. This paper describes objective technical results and analysis. Any subjective views
915 or opinions that might be expressed in the paper do not necessarily represent the views of the U.S.
916 Department of Energy or the United States Government. The authors would like to thank
917 Associate Editor Thorsten Wagener, Alberto Guadagnini and the other anonymous reviewer for
918 their insightful comments and suggestions, which helped improve this manuscript. All computer
919 codes and data used in this study are available online at https://github.com/jyangfsu/total-effect
920 and at Zenodo via https://zenodo.org/record/4542541#.YCtnm2gzZPY with DOI:
921 10.5281/zenodo.4542541.

Appendix A. Derivation of Analytical solutions PS_K and PS_{T_K} for Sobol's G*-function

For the first-order process sensitivity index defined in equation (2), following equations (4) and (5) of Ye et al. (2004), the variance term $V(\Delta)$ in the denominator is evaluated as

$$\begin{aligned} E(\Delta) &= \sum_{k=1}^{K} E(\Delta|M_k)P(M_k) \\ V(\Delta) &= \sum_{k=1}^{K} V(\Delta|M_k)P(M_k) + \sum_{k=1}^{K} \left(E(\Delta|M_k) - E(\Delta)\right)^2 P(M_k) \end{aligned}, \quad \text{(A1)}$$

where $E(\Delta)$ is model averaged mean, M_k (k=1, 2,…,8) is a system model, and its probability $P(M_k)$ = 1/8 because each process model has the probability of 1/2, and $E(\Delta|M_k)$ and $V(\Delta|M_k)$ are the mean and variance of each system model, respectively. Since the mean of each element, g_i^*, in the Sobol's G*-function is 1, $E(\Delta|M_k)$ = 1, and $E(\Delta)$ = 1. As a result, we have

$$V(\Delta) = \sum_{k=1}^{K} V(\Delta|M_k)P(M_k) = \frac{1}{8}\sum_{j=1}^{8} V^j, \quad \text{(A2)}$$

where V^j is the variance of each system model, i.e., the V term in equation (14).

For the numerator term $V_{\mathbf{M}_K}(E_{\mathbf{M}_{\sim K}}[\Delta \mid M_K])$ in the definition of the first-order process sensitivity index, according to the law of total variance, it can be expressed as

$$V_{\mathbf{M}_K}(E_{\mathbf{M}_{\sim K}}[\Delta \mid M_K]) = V(\Delta) - E_{\mathbf{M}_K}(V_{\mathbf{M}_{\sim K}}[\Delta \mid M_K]). \quad \text{(A3)}$$

Applying the law of total expectation to the second term at the right-hand side of equation (A3) gives

$$E_{\mathbf{M}_K}(V_{\mathbf{M}_{\sim K}}[\Delta \mid M_K]) = E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K|M_K}(V_{\mathbf{M}_{\sim K}}[\Delta \mid \theta_K, M_K]), \quad \text{(A4)}$$

which considers parametric uncertainty. For the $V_{\mathbf{M}_{\sim K}}[\Delta \mid M_K]$ term of equation (A4), it is for multiple models, $\mathbf{M}_{\sim K}$, and using the expression of model-averaged variance (equation 5 of Ye et al., 2004) gives

$$
\begin{aligned}
&E_{\mathbf{M}_K}(V_{\mathbf{M}_{\sim K}}[\Delta \mid M_K]) = E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K|M_K}(V_{\mathbf{M}_{\sim K}}[\Delta \mid M_K]) \\
&= E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K|M_K} \left(\begin{array}{l} E_{\mathbf{M}_{\sim K}} V_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] + \\ E_{\mathbf{M}_{\sim K}} \left(E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] - E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] \right)^2 \end{array} \right) \\
&= E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K|M_K} E_{\mathbf{M}_{\sim K}} V_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] \\
&+ E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K|M_K} E_{\mathbf{M}_{\sim K}} \left(E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] - E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] \right)^2
\end{aligned}
\quad . \quad \text{(A5)}
$$

Assuming independent processes and parameters, equation (A5) can be rearranged as:

$$
\begin{aligned}
&E_{\mathbf{M}_K}(V_{\mathbf{M}_{\sim K}}[\Delta \mid M_K]) \\
&= E_{\mathbf{M}_K} E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_K|M_K} V_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] \\
&+ E_{\mathbf{M}_K} E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_K|M_K} \left(E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] - E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] \right)^2 \\
&= \sum_{N_K} \sum_{N_{\sim K}} E_{\boldsymbol{\theta}_K|M_K} V_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] P(M_{\sim K}) P(M_K) \\
&+ \sum_{N_K} \sum_{N_{\sim K}} E_{\boldsymbol{\theta}_K|M_K} \left(E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] - E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] \right)^2 P(M_{\sim K}) P(M_K)
\end{aligned}
\quad . \quad \text{(A6)}
$$

Similarly, separating a system model into a combination of process models, equation (A1) becomes

$$
\begin{aligned}
E(\Delta) &= \sum_{N_K} \sum_{N_{\sim K}} E_{\boldsymbol{\theta}_K|M_K} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] P(M_{\sim K}) P(M_K) \\
V(\Delta) &= \sum_{N_K} \sum_{N_{\sim K}} V_{\boldsymbol{\theta}_K|M_K, \boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] P(M_{\sim K}) P(M_K) \\
&+ \sum_{N_K} \sum_{N_{\sim K}} \left(E_{\boldsymbol{\theta}_K|M_K} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] - E(\Delta) \right)^2 P(M_{\sim K}) P(M_K)
\end{aligned}
\quad . \quad \text{(A7)}
$$

Substituting equations (A6) and (A7) into equation (A3) leads to

948
$$
\begin{aligned}
&V_{\mathbf{M}_K}(E_{\mathbf{M}_{\sim K}}[\Delta \mid M_K]) \\
&= \sum_{N_K}\sum_{N_{\sim K}} \begin{pmatrix} V_{\boldsymbol{\theta}_{K}|M_K,\boldsymbol{\theta}_{\sim K}|M_{\sim K},}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]P(M_{\sim K})P(M_K) \\ -E_{\boldsymbol{\theta}_K|M_K} V_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]P(M_{\sim K})P(M_K) \end{pmatrix} P(M_{\sim K})P(M_K) \\
&+ \sum_{N_K}\sum_{N_{\sim K}} \begin{pmatrix} \left(E_{\boldsymbol{\theta}_K|M_K} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] - E(\Delta)\right)^2 \\ -E_{\boldsymbol{\theta}_K|M_K}\left(E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] - E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]\right)^2 \end{pmatrix} P(M_{\sim K})P(M_K) \cdot
\end{aligned}
\tag{A8}
$$

949 Applying the law of total variance to the term in the first parenthesis at the right-hand side of (A8)

950 gives

951
$$
\begin{aligned}
&V_{\mathbf{M}_K}(E_{\mathbf{M}_{\sim K}}[\Delta \mid M_K]) \\
&= \sum_{N_K}\sum_{N_{\sim K}} V_{\boldsymbol{\theta}_K|M_K} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]P(M_{\sim K})P(M_K) \\
&+ \sum_{N_K}\sum_{N_{\sim K}} \begin{pmatrix} \left(E_{\boldsymbol{\theta}_K|M_K} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] - E(\Delta)\right)^2 \\ -E_{\boldsymbol{\theta}_K|M_K}\left(E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] - E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]\right)^2 \end{pmatrix} P(M_{\sim K})P(M_K)
\end{aligned}
\tag{A9}
$$

952 Since the mean of each element, g_i^*, in the Sobol's G*-function is 1, the mean of each system

953 model is 1, and the model-averaged mean is also 1. As a result, the second term at the right-hand

954 side of (A9) is zero. This leads to

955
$$
\begin{aligned}
&V_{\mathbf{M}_K}(E_{\mathbf{M}_{\sim K}}[\Delta \mid M_K]) \\
&= \sum_{N_K}\sum_{N_{\sim K}} V_{\boldsymbol{\theta}_K|M_K} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]P(M_{\sim K})P(M_K) \quad \cdot
\end{aligned}
\tag{A10}
$$

956 Therefore, the first-order process sensitivity index becomes,

957
$$
PS_K = \frac{\sum_{N_K}\sum_{N_{\sim K}} V_{\boldsymbol{\theta}_K|M_K} E_{\boldsymbol{\theta}_{\sim K}|M_{\sim K}}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}]P(M_{\sim K})P(M_K)}{V(\Delta)} = \frac{\frac{1}{2}\left(V_K^1 + V_K^2\right)}{V(\Delta)}. \tag{A11}
$$

958 For the total-effect process sensitivity,

959
$$
PS_{TK} = \frac{E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])}{V(\Delta)}, \tag{A12}
$$

its $E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}])$ term is derived in a similar manner as

$$\begin{aligned} & E_{\mathbf{M}_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid M_{\sim K}]) \\ & = \sum_{N_{\sim K}} \sum_{N_K} E_{\boldsymbol{\theta}_{\sim K} \mid M_{\sim K}} V_{\boldsymbol{\theta}_K \mid M_K}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] P(M_K) P(M_{\sim K}) \\ & + \sum_{N_{\sim K}} \sum_{N_K} E_{\boldsymbol{\theta}_{\sim K} \mid M_{\sim K}} \left(E_{\boldsymbol{\theta}_K \mid M_K}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] - E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K \mid M_K}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] \right)^2 P(M_K) P(M_{\sim K}) \end{aligned} \quad \text{(A13)}$$

The second term at the right-hand side of (A13) is zero for the reason explained for (A8). Thus, the total-effect process sensitivity index is

$$PS_{TK} = \frac{\sum_{N_{\sim K}} \sum_{N_K} E_{\boldsymbol{\theta}_{\sim K} \mid M_{\sim K}} V_{\boldsymbol{\theta}_K \mid M_K}[\Delta \mid \theta_K, M_K, \theta_{\sim K}, M_{\sim K}] P(M_K) P(M_{\sim K})}{V(\Delta)} = \frac{\frac{1}{8} \sum_{j=1}^{8} V_{T_K}{}^{j}}{V(\Delta)} \quad \text{(A14)}$$

Appendix B. Derivation of analytical solution of groundwater discharge per unit width

For the groundwater flow model, its governing equations and boundary conditions are as follows:

$$\begin{aligned} & \frac{d}{dx}\left(h_x^1 \frac{h_x^1}{dx} \right) + \frac{w}{K_1} = 0 \qquad \text{for } x < 7000 \text{ m} \\ & \frac{d}{dx}\left(h_x^2 \frac{dh_x^2}{dx} \right) + \frac{w}{K_2} = 0 \qquad \text{for } x \geq 7000 \text{ m} \\ & h_x^1 \big|_{x=0} = h_1 \\ & h_x^2 \big|_{x=L} = h_2 \\ & h_x^1 \big|_{x=x_0} = h_x^2 \big|_{x=x_0} \\ & K_1 \frac{dh_x^1}{dx} \Big|_{x=x_0} = K_2 \frac{dh_x^2}{dx} \Big|_{x=x_0} \end{aligned} \quad \text{(B1)}$$

where h_x^1 and h_x^2 are hydraulic heads in the two hydraulic conductivity zones , respectively, h_1 and h_2 are the hydraulic heads at the left and right boundaries, respectively. The last two equations indicate the continuity of hydraulic head and groundwater flow at the location ($x = x_0$)

where the two zones of hydraulic conductivity are separated. The general solutions of the first two equations are as follows,

$$[h_x^1]^2 = -\frac{w}{K_1}x^2 + C_1 x + C_2 \quad , \qquad [h_x^2]^2 = -\frac{w}{K_2}x^2 + C_3 x + C_4 \tag{B2}$$

where C_1 – C_4 are four coefficients to be determined. Using the boundary conditions and the continuity conditions in equation (B1) gives

$$\begin{aligned} C_1 &= -\left(h_1^{\,2} - h_2^{\,2} - \frac{w}{K_1}x_0^{\,2} + \frac{w}{K_2}x_0^{\,2} - \frac{w}{K_2}L^2 \right) / \left(x_0 - \frac{K_1}{K_2}x_0 + \frac{K_1}{K_2}L\right) \\ C_2 &= h_1^{\,2} \\ C_3 &= \frac{K_1}{K_2}C_1 \\ C_4 &= h_2^{\,2} + \frac{w}{K_2}L^2 - C_3 L \end{aligned} \quad . \tag{B3}$$

The groundwater discharge per unit width at a given location $q(x)$ can be obtained at:

$$\begin{aligned} q(x) &= -K_1 h_x^1 \frac{dh_1}{dx} = wx - \frac{K_1 C_1}{2} \qquad \text{for } x < 7000 \text{ m} \\ q(x) &= -K_2 h_x^2 \frac{dh_2}{dx} = wx - \frac{K_2 C_3}{2} \qquad \text{for } x \geq 7000 \text{ m} \end{aligned} \quad . \tag{B4}$$

The third equation of equations (B3) indicates that $K_2C_3 = K_1C_1$, suggesting the flow continuity in the study domain. Thus equation (B4) can be rewritten as:

$$q(x) = -K_1 h_x^1 \frac{dh_1}{dx} = wx - \frac{K_1 C_1}{2} = K_1 \frac{h_1^{\,2} - h_2^{\,2}}{2\left(x_0 - \lambda x_0 + \lambda L\right)} - \frac{1}{2} w \frac{x_0^{\,2} - \lambda x_0^{\,2} + \lambda L^2}{x_0 - \lambda x_0 + \lambda L} + wx \tag{B5}$$

where $\lambda = K_1 / K_2$.

Appendix C. Convergence of PS_K and PS_{TK} for groundwater flow modeling

< Insert Figure C1 here>

References

Baroni, G. & Tarantola S. (2014). A General Probabilistic Framework for uncertainty and global sensitivity analysis of deterministic models: A hydrological case study. *Environmental Modelling & Software*, 51, 26-34. https://doi.org/10.1016/j.envsoft.2013.09.022

Beven, K. J. (2002). Chapter 12 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. J. (2018). On hypothesis testing in hydrology: Why falsification of models is still a really good idea. *Wiley Interdisciplinary Reviews: Water*, 5(3), e1278. https://doi.org/10.1002/wat2.1278

Bredehoeft, J. D. (2003). From models to performance assessment: the conceptualization problem. *Ground Water*, 41(5), 571-577. https://doi.org/10.1111/j.1745-6584.2003.tb02395.x

Bredehoeft, J. (2005). The conceptualization model problem-surprise. *Hydrogeology Journal*, 13, 37–46. https://doi.org/10.1007/s10040-004-0430-5

Burnham, K. P. & Anderson A. R. (2002). Model selection and multiple model inference: a practical information-theoretical approach, 2nd edition, New York, Springer

Chang, Y., Wu, J., Jiang, G., & Kang, Z. (2017). Identification of the dominant hydrological process and appropriate model structure of a karst catchment through stepwise simplification of a complex conceptual model. *Journal of Hydrology*, 548, 75–87. https://doi.org/10.1016/j.jhydrol.2017.02.050

1007 Clark, M. P., Bierkens M. F. P., Samaniego L., Woods R. Uijlenhoet A., R., Bennett K. E.,
1008 Pauwels V. R. N., Cai X., Wood A. W., & Peters-Lidard C.D. (2017). The evolution of
1009 process-based hydrologic models: historical challenges and the collective quest for
1010 physical realism. *Hydrology and Earth System Sciences*, 21, 3427–3440.
1011 https://doi.org/10.5194/hess-21-3427-2017

1012 Clark, M.P. et al. 2016. Improving the representation of hydrologic processes in Earth System
1013 Models. *Water Resources Research*, 51(8): 5929-5956.
1014 https://doi.org/10.1002/2015wr017096

1015 Clark, M.P. et al. 2015a. A unified approach for process-based hydrologic modeling: 1.
1016 Modeling concept. *Water Resources Research*, 51(4): 2498-2514.
1017 https://doi.org/10.1002/2015WR017198

1018 Clark, M.P. et al., 2015b. A unified approach for process-based hydrologic modeling: 2. Model
1019 implementation and case studies. *Water Resources Research*, 51(4): 2515-2542.
1020 https://doi.org/10.1002/2015wr017200

1021 Dai, H., Q. Liu, M. Ye, B. Hu, and S. Yuan (2021), A computationally efficient quasi-MC
1022 method for estimating multi-model process sensitivity index, Water Resoirces Research,
1023 Under Review.

1024 Dai, H., Chen X., Ye M., Song X., Hammond G., Hu B. X., & Zachara J. M. (2019). Using
1025 Bayesian networks for sensitivity analysis of complex biogeochemical models, *Water*
1026 *Resources Research*, 55, 3541-3555. https://doi.org/10.1029/2018WR023589

1027 Dai, H., Ye, M., Walker, A.P., Chen, X. (2017a). A new process sensitivity index to identify
1028 important system processes under process model and parametric uncertainty. *Water*
1029 *Resources Research*, 53(4): 3476-3490. https://doi.org/10.1002/2016wr019715

1030 Dai, H., Chen X., Ye M., Song X., & Zachara J. M. (2017b). A geostatistics-informed
1031 hierarchical sensitivity analysis method for complex groundwater flow and transport
1032 modeling. *Water Resources Research*, 53(5), 4327-4343.
1033 https://doi.org/10.1002/2016WR019756

1034 Dai, H. & Ye, M. (2015). Variance-based global sensitivity analysis for multiple scenarios and
1035 models with implementation using sparse grid collocation. *Journal of Hydrology*, 528:
1036 286-300. https://doi.org/10.1016/j.jhydrol.2015.06.034.

1037 Dell'Oca, A., Riva, M., & Guadagnini, A. 2020. Global Sensitivity Analysis for multiple
1038 interpretive models with uncertain parameters. *Water Resources Research*.
1039 https://doi.org/10.1029/2019wr025754

1040 Draper, D. (1995). Assessment and propagation of model uncertainty. *Journal of the Royal*
1041 *Statistical Society: Series B (Methodological)*, 57(1), 45-97.
1042 https://doi.org/10.1111/j.2517-6161.1995.tb02015.x

1043 Elshall, A.S. and M. Ye, Making steppingstones out of stumbling blocks: A Bayesian model
1044 evidence estimator with application to groundwater transport model selection, Water, 11,
1045 1579, doi:10.3390/w11081579.

1046 Francke, T., Baroni, G., Brosinsky, A., Foerster, S., López-Tarazón, J. A., Sommerer, E., &
1047 Bronstert, A. (2018). What did really improve our mesoscale hydrological model? A

multidimensional analysis based on real observations. *Water Resources Research*, 54, 8594–8612. https://doi.org/10.1029/2018WR022813

Gan, Y. et al., 2014. A comprehensive evaluation of various sensitivity analysis methods: A case study with a hydrological model. *Environmental Modelling & Software*, 51: 269-285. https://doi.org/10.1016/j.envsoft.2013.09.031

Günther, D., Marke, T., Essery, R., & Strasser, U. (2019). Uncertainties in snowpack simulations - Assessing the impact of model structure, parameter choice, and forcing data error on point-scale energy balance snow model performance. *Water Resources Research*, 55, 2779–2800. https://doi.org/10.1029/2018WR023403

Gupta, H. V., Clark M. P., Vrugt J. A., Abramowitz G., & Ye M. (2012). Towards a comprehensive assessment of model structural adequacy. *Water Resources Research*, 48, W08301, https://doi.org/10.1029/2011WR011044

Herman, J.D., Reed, P.M., & Wagener, T., 2013. Time-varying sensitivity analysis clarifies the effects of watershed model formulation on model behavior. *Water Resources Research*, 49, 1400-1414. https://doi.org/10.1002/wrcr.20124

Herman, J. & Usher W. 2017. SALib: An open-source python library for sensitivity analysis. *The Journal of Open Source Software*, 2(9): 771-784. https://doi.org/10.21105/joss.00097

Hock, R., 2003. Temperature index melt modelling in mountain areas. *Journal of Hydrology*, 282(1-4): 104-115. https://doi.org/10.1016/s0022-1694(03)00257-9

Hoeting, J.A., Madigan D., Raftery A.E., Volinsky C.T. (1999). Bayesian model averaging: A tutorial. *Statistical Science*. 14(4), 382-417

Bianchi Janetti, E., Guadagnini, L., Riva, M., & Guadagnini, A. 2019. Global sensitivity analyses of multiple conceptual models with uncertain parameters driving groundwater flow in a regional-scale sedimentary aquifer. *Journal of Hydrology*, 574: 544-556. https://doi.org/10.1016/j.jhydrol.2019.04.035

Kustas, W.P., Rango, A., & Uijlenhoet, R. 1994. A simple energy budget algorithm for the snowmelt runoff model. *Water Resources Research*, 30(5): 1515-1527. https://doi.org/10.1029/94wr00152

Lilburne, L., and S. Tarantola (2009), Sensitivity analysis of spatial models, Int. J. Geogr. Inf. Sci., 23, 151–168, doi:10.1080/13658810802094995.

Liu, P., A.S. Elshall, M. Ye, P. Beerli, X. Zeng, D. Lu, and Y. Tao (2016), Evaluate marginal likelihood with thermodynamic integration method and comparison with several other numerical methods, *Water Resources Research*, 52, doi:10.1002/2014WR016718.

Lu, D., Ye, M., & Curtis, G.P. 2015. Maximum likelihood Bayesian model averaging and its predictive analysis for groundwater reactive transport models. *Journal of Hydrology*, 529: 1859-1873. https://doi.org/10.1016/j.jhydrol.2015.07.029

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, 5835–5858. https://doi.org/10.5194/hess-24-5835-2020

Markstrom, S.L., Hay L.E., & Clark M. P. (2016). Toward simplication of hydrologic modeling: identification of dominant processes. *Hydrology and Earth System Sciences*, 20, 4655–4671. https://doi.org/10.5194/hess-20-4655-2016

Martinec, J., Rango A., & Roberts R. (2008). Snowmelt Runoff Model (SRM): User's Manual. New Mexico State University, Las Cruces.

Mendoza, P. A. et al. (2015). Are we unnecessarily constraining the agility of complex process-based models? *Water Resources Research*, 51(1): 716-728. https://doi.org/10.1002/2014wr015820

Neuman, S.P. (2003), Maximum likelihood Bayesian averaging of alternative conceptual-mathematical models. *Stochastic Environmental Research and Risk Assessment*, 17(5), 291-305, https://doi.org/10.1007/s00477-003-0151-7.

Poeter, E.P., & Anderson D. A. (2005). Multimodel ranking and inference in ground water modeling. *Ground Water*, 43(4), 597-605. https://doi.org/10.1111/j.1745-6584.2005.0061.x

Poeter, E.P. & Hill M.C. (2007). MMA: A computer code for multi-model analysis. U.S. Geological Survey Techniques and Methods 6-E3. Reston, Virginia: USGS.

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, 3070–3092, https://doi.org/10.1002/2014WR016527.

Razavi, S. & Gupta H. V. (2016a), A new framework for comprehensive, robust, and efficient global sensitivity analysis: 1. Theory. *Water Resources Research*, 52, 423–439, https://doi.org/10.1002/2015WR017558.

Razavi, S. & Gupta H. V. (2016b), A new framework for comprehensive, robust, and efficient global sensitivity analysis: 2. Application. *Water Resources Research*, 52, 440–455, https://doi.org/10.1002/2015WR017559.

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., Ratto M., Andres T., Campolongo F., Cariboni J., Gatelli D., Saisana M., & Tarantola S. (2008). Global Sensitivity Analysis: The Primer, *John Wiley & Sons*, Chichester, England.

Saltelli, A., Tarantola S., Campolongo F., & Ratto M. (2004). Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models, *John Wiley & Sons*, Chichester, England.

Savage, J. T. S., Pianosi F., Bates P., Freer J., & Wagener T. (2016). Quantifying the importance of spatial resolution and other factors through global sensitivity analysis of a flood inundation model. *Water Resources Research*, 52(11), 9146-9163. https://doi.org/10.1002/2015wr018198

Shen, Z.Y., Hong, Q., Yu, H. (2008). Parameter uncertainty analysis of the non-point source pollution in the Daning River watershed of the Three Gorges Reservoir Region, China. *Science of the Total Environment*, 405(1-3): 195-205. https://doi.org/10.1016/j.scitotenv.2008.06.009

Shi, X. et al. (2014). Assessment of parametric uncertainty for groundwater reactive transport modeling. *Water Resources Research*, 50(5): 4416-4439. https://doi.org/10.1002/2013wr013755

Shin, M. J., Guillaume, J. H. A., Croke, B .F. W., & Jakeman, A. J. 2013. Addressing ten questions about conceptual rainfall-runoff models with global sensitivity analyses in R. *Journal of Hydrology*, 503: 135-152. https://doi.org/10.1016/j.jhydrol.2013.08.047

Sivakumar, B. 2004. Dominant processes concept in hydrology: moving forward. *Hydrological Processes*, 18(12): 2349-2353. https://doi.org/10.1002/hyp.5606

Song, X., Zhang J., Zhan C., Xuan Y., Ye M., Xu C. (2015). Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications, *Journal of Hydrology*, 523(4), 739-757, https://doi.org/10.1016/j.jhydrol.2015.02.013.

Spieler, D., Mai, J., Craig, J. R., Tolson, B. A., & Schütze, N. (2020). Automatic model structure identification for conceptual hydrologic models. *Water Resources Research*, 56, e2019WR027009. https://doi.org/10.1029/2019WR027009

Tarantola, S., Giglioli N., Jesinghaus J., & Saltelli A. (2002). Can global sensitivity analysis steer the implementation of models for environmental assessments and decision-making?, *Stochastic Environmental Research and Risk Assessment (SERRA)*, 16(1), 63-76. https://doi.org/10.1007/s00477-001-0085-x

Van Hoey, S., Seuntjens, P., van der Kwast, J., & Nopens, I. (2014). A qualitative model structure sensitivity analysis method to support model selection. *Journal of Hydrology*, 519, 3426–3435. https://doi.org/10.1016/j.jhydrol.2014.09.052

Walker, A.P., Ye M., Lu D., De Kauwe M.G., Gu L., Medlyn B.E., Rogers A., & Serbin S.P. (2018). The Multi-Assumption Architecture and Testbed (MAAT v1.0): R code for generating ensembles with dynamics 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., Lu D., Ricciuto D.M., Serbin A.P., & Ye M. (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*. 2021, 27: 804-822. https://doi.org/10.1111/gcb.15366

van Werkhoven, K., Wagener, T., Reed, P., & Tang, Y. 2008. Characterization of watershed model behavior across a hydroclimatic gradient, *Water Resources Research*, 44, W01429, doi:10.1029/2007WR006271.

Ye, M., Neuman S.P., & Meyer P.D. (2004). Maximum likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff. *Water Resources Research*, 40, W05113, doi:10.1029/2003WR002557.

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. *Ground Water*, 48(5): 716-28. https://doi.org/10.1111/j.1745-6584.2009.00633.x

Ye, M., Wang L., Pohlmann K.F., & Chapman J.B. (2016). Estimate groundwater interbasin flow using multiple models and multiple types of calibration data. *Ground Water*, 54(6), 805-817. https://doi.org/10.1111/gwat.12422.

LIST OF TABLES

Table 1. Analytical solutions (%) of first-order (S_i) and total-effect (S_{Ti}) parameter sensitivity indices for each system model and of first-order (PS_K) and total-effect (PS_{TK}) process sensitivity indices of multiple system models developed based on Sobol's G*-function. Numerical solutions are almost identical to the analytical solutions with the maximum difference less than 0.1%.

	Considering parametric uncertainty but not process model uncertainty											
Model	$g_1^{*1}g_2^{*1}g_3^{*1}$			$g_1^{*1}g_2^{*1}g_3^{*2}$			$g_1^{*1}g_2^{*2}g_3^{*1}$			$g_1^{*1}g_2^{*2}g_3^{*2}$		
Para.	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3
Rank	1	2	3	2	3	1	2	1	3	3	1	2
S_i	73.42	16.97	8.16	36.87	8.52	50.79	31.80	60.85	3.53	21.67	41.46	29.85
S_{Ti}	74.77	17.98	8.70	40.07	9.64	54.15	35.26	64.47	4.10	25.64	46.88	34.65
Model	$g_1^{*2}g_2^{*1}g_3^{*1}$			$g_1^{*2}g_2^{*1}g_3^{*2}$			$g_1^{*2}g_2^{*2}g_3^{*1}$			$g_1^{*2}g_2^{*2}g_3^{*2}$		
Para.	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3	X_1	X_2	X_3
Rank	1	2	3	1	3	2	1	2	3	1	2	3
S_i	88.56	6.60	3.17	62.07	4.63	27.58	56.64	34.97	2.03	43.67	26.96	19.41
S_{Ti}	90.18	7.74	3.75	67.45	5.79	32.54	62.79	40.99	2.61	51.66	33.72	24.92
	Considering both process model uncertainty and parametric uncertainty											
Process	g_1^*			g_2^*			g_3^*					
Rank	1			2			3					
PS_K	49.85			26.08			18.10					
PS_{TK}	54.79			30.07			21.23					

Table 2. Numerical solutions (%) of first-order (S_i) and total-effect (S_{Ti}) parameter sensitivity indices for each system model and of first-order (PS_K) and total-effect (PS_{TK}) process sensitivity indices of multiple system models of groundwater flow model. Each system model has three processes: recharge (*R*), geology (*G*), and snowmelt (*M*).

	Considering parametric uncertainty but not process model uncertainty											
	$R_1G_1M_1$			$R_1G_1M_2$			$R_1G_2M_1$			$R_1G_2M_2$		
	a	*K*	f_1	*a*	*K*	f_2&*r*	*a*	K_1&K_2	f_1	*a*	K_1&K_2	f_2&*r*
S_i	12.41	22.47	50.78	15.08	47.17	29.75	33.43	11.35	51.48	35.20	36.36	26.35
Rank	3	2	1	3	1	2	2	3	1	2	1	3
S_{Ti}	12.37	36.87	66.06	14.98	55.74	37.95	34.08	15.78	54.65	35.83	38.66	27.97
Rank	3	2	1	3	1	2	2	3	1	2	1	3
	$R_2G_1M_1$			$R_2G_1M_2$			$R_2G_2M_1$			$R_2G_2M_2$		
	b	*K*	f_1	*b*	*K*	f_2&*r*	*b*	K_1&K_2	f_1	*b*	K_1&K_2	f_2&*r*
S_i	22.04	20.00	45.23	26.14	41.03	25.84	48.70	10.68	37.48	49.43	30.10	18.55
Rank	2	3	1	2	1	3	1	3	2	1	2	3
S_{Ti}	22.04	32.82	58.80	26.02	48.49	33.01	49.55	14.43	39.79	50.30	32.16	19.67
Rank	3	2	1	3	1	2	1	3	2	1	2	3

	Considering both process model uncertainty and parametric uncertainty		
	R	*G*	*M*
PS_K	14.32	6.79	63.99
Rank	2	3	1
PS_{TK}	14.67	21.68	78.54
Rank	3	2	1

Table 3. Variance decomposition and various sensitivity indices for groundwater discharge per unit width (q_0) at x_0 = 7,000 m simulated by system model $R_2G_1M_1$.

First-order effect*			Second-order effect*			Third-order effect*		
Parameter	V_i	Si (%)	Pair	V_{ij}	S_{ij} (%)	Pair	V_{ijk}	S_{ijk} (%)
b	0.2847	22.04	K, b	0.00	0.00	f_1, b, K	0.00	0.00
K	0.2584	20.00	f_1, b	0.00	0.00			
f_1	0.5884	45.23	f_1, K	0.1591	12.32			

Total effect

Parameter	V_{Ti}*	S_{TK}*(%)	V_{Ti}**
b	0.2847	22.04	$V_{Tb} = V_b + V_{K,b} + V_{f1,b} + V_{f1,b,K} = V_b = 0.2847$
K	0.4241	32.82	$V_{TK} = V_K + V_{K,b} + V_{f1,K} + V_{f1,b,K} = V_K + V_{f1,K} = 0.4175$
f_1	0.7597	58.80	$V_{Tf1} = V_{f1} + V_{f1,b} + V_{f1,K} + V_{f1,b,K} = V_K + V_{f1,K} = 0.7475$

Variance $V = V_b + V_K + V_{f1} + V_{K,b} + V_{f1,b} + V_{f1,K} + V_{f1,b,K} = 1.2906$**

* : Calculated based on Monte Carlo approximation

** : Calculated based on variance decomposition

Table 4. Variance decomposition and various sensitivity indices for groundwater discharge per unit width (q_0) at x_0 = 7,000 m simulated by all system model considering both model uncertainty and parametric uncertainty

First-order effect*			Second-order effect*			Third-order effect*		
Process	V_K	PS_K (%)	Pair	$V_{K,V}$	$PS_{K,V}$(%)	Pair	$V_{K,V,W}$	$PS_{K,V,W}$ (%)
R	0.2966	14.32	*G, R*	0.0071	0.34	*M, R, G*	0.00	0.00
G	0.1406	6.79	*M, R*	0.00	0.00			
M	1.3252	63.99	*M, G*	0.3013	14.55			

Total effects			
Process	V_{TK}*	PS_{TK}* (%)	V_{TK}**
R	0.3038	14.67	$V_{TR} = V_R + V_{G,R} + V_{M,R} + V_{M,R,G} = V_R + V_{G,R}$ =0.3037
G	0.4490	21.68	$V_{TG} = V_G + V_{G,R} + V_{M,G} + V_{M,R,G} = V_G + V_{G,R} + V_{M,G}$ =0.4490
M	1.6265	78.54	$V_{TM} = V_M + V_{M,R} + V_{M,G} + V_{M,R,G} = V_M + V_{M,G}$ =1.6265
Variance $V = V_R + V_G + V_M + V_{G,R} + V_{M,R} + V_{G,M} + V_{M,R,G}$ = 2.0708**			

* : Calculated based on Monte Carlo approximation

** : Calculated based on variance decomposition

FIGURE CAPTIONS

Figure 1. Pseudo code to compute the total-effect process sensitivity index using Monte Carlo and model averaging methods based on equation (10). The total number of model executions is $(N_{\sim K} \times m_{\sim K}) \times (N_K \times m_K)$, where $N_{\sim K}$ is the number of process model combinations for the processes ~K, $m_{\sim K}$ is the number of realizations for parameters $\boldsymbol{\theta}_{\sim K} \mid M_{\sim K}$, N_K is the number of process models for process K, and m_K is the number of realizations for parameter $\boldsymbol{\theta}_K \mid M_K$.

Figure 2. Sketch of the one-dimensional groundwater flow model domain. The groundwater discharge per unit width (q_0) at x_0 = 7,000 m is the quantity of interest for sensitivity analysis.

Figure 3. Probability density functions (PDFs) of the groundwater discharge per unit width at x_0 = 7,000 m simulated by individual system models when fixing (a) snowmelt process *M* to M_1, (b) snowmelt process *M* to M_2, (c) recharge process *R* to R_1, (d) recharge process *R* to R_2, (e) geology process *G* to G_1, and (f) geology process *G* to G_2. The $E_{\boldsymbol{\theta}_K \mid M_K}(V_{\mathbf{M}_{\sim K}}[\Delta \mid \theta_K, M_K])$ term is the average variance of the system output after fixing a process model. Taking Figure 3f as an example, M_K denotes process model G_2, $\boldsymbol{E}_{\boldsymbol{\theta}_K \mid M_K}$ denotes the expectation with respect to all parameter realizations of K_1 and K_2 associated with process model G_2, $V_{\mathbf{M}_{\sim K}}[\Delta \mid \theta_K, M_K]$ represents the output variance considering all other process model combinations (i.e., R_1M_1, R_1M_2, R_2M_1, and R_2M_2) and the corresponding parameter combinations under a single parameter realization of K_1 and K_2. The PDF in black is for q_0 values simulated by using all the system models and their random parameters; this PDF thus presents the original uncertainty measured by $V(\Delta)$.

Figure 4. Probability density functions (PDFs) of the groundwater discharge per unit width (q_0) at x_0 = 7,000 m simulated by individual system models when varying (a1 – a4) snowmelt process model from M_1 to M_2, (b1 – b4) geology process model from G_1 to G_2, and (c1 – c4) recharge process model from R_1 to R_2. The $E_{\boldsymbol{\theta}_{\sim K} \mid M_{\sim K}}(V_{\mathbf{M}_K}[\Delta \mid \theta_{\sim K}, M_{\sim K}])$ term is the average variance of q_0 after fixing the other two process models. Taking Figure 4(c4) as an example, $M_{\sim K}$ denotes the process model combination G_2M_2, $\boldsymbol{E}_{\boldsymbol{\theta}_{\sim K} \mid M_{\sim K}}$ is the expectation of all realizations of parameters K_1, K_2, f_2 and *r* associated with G_2M_2, $V_{\mathbf{M}_K}[\Delta \mid \theta_{\sim K}, M_{\sim K}]$ is the output variance of q_0 considering the uncertainty of process model *R* and parameter realizations under a single parameter combination of *K*, f_2 and *r*.

Figure 5. (a) First-order and total-effect process sensitivity indices and (b) second-order process sensitivity indices for the groundwater discharge per unit width throughout the entire domain.

Figure 6. Within- and between-model variance reduction for (a) recharge process (b) geology process, and (c) snowmelt process, and within- and between-model variance left for (d) recharge process (e) geology process, and (f) snowmelt process over the entire domain.

Figure C1. Convergence of (a) first-order process sensitivity index and (b) total-effect process sensitivity index of the recharge (*R*), geology (*G*), and snowmelt (*M*) processes for the groundwater flow modeling.

Figure1.

Loop [1] over process models $M_{\sim K}$ of process $\sim K$ in $\mathbf{M}_{\sim K}$

Loop [2] over model parameter realizations $\theta_{\sim K}$ of model $M_{\sim K}$ in $\boldsymbol{\theta}_{\sim \mathbf{K}}$

Loop [3] over process models M_K of process K in $\mathbf{M}_K$

Loop [4] over model parameter realizations θ_K of model M_K in $\boldsymbol{\theta}_{\mathbf{K}}$

Compute $\Delta | M_{\sim K}, \theta_{\sim K}, M_K, \theta_K$

End Loop [4]

Compute $E_{\boldsymbol{\theta}_K | M_K}[\Delta | M_{\sim K}, \theta_{\sim K}, M_K, \theta_K]$

End Loop [3]

Compute $E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K | M_K}[\Delta | M_{\sim K}, \theta_{\sim K}, M_K, \theta_K]$ and
$(E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K | M_K}[\Delta | M_{\sim K}, \theta_{\sim K}, M_K, \theta_K])^2$ using model averaging

End Loop [2]

Compute $E_{\boldsymbol{\theta}_{\sim K} | M_{\sim K}} E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K | M_K}[\Delta | M_{\sim K}, \theta_{\sim K}, M_K, \theta_K]$ and
$E_{\boldsymbol{\theta}_{\sim K} | M_{\sim K}} (E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K | M_K}[\Delta | M_{\sim K}, \theta_{\sim K}, M_K, \theta_K])^2$

End Loop [1]

Compute $(E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K} | M_{\sim K}} E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K | M_K}[\Delta | M_{\sim K}, \theta_{\sim K}, M_K, \theta_K])^2$ and
$E_{\mathbf{M}_{\sim K}} E_{\boldsymbol{\theta}_{\sim K} | M_{\sim K}} (E_{\mathbf{M}_K} E_{\boldsymbol{\theta}_K | M_K}[\Delta | M_{\sim K}, \theta_{\sim K}, M_K, \theta_K])^2$ using model averaging

Figure2.

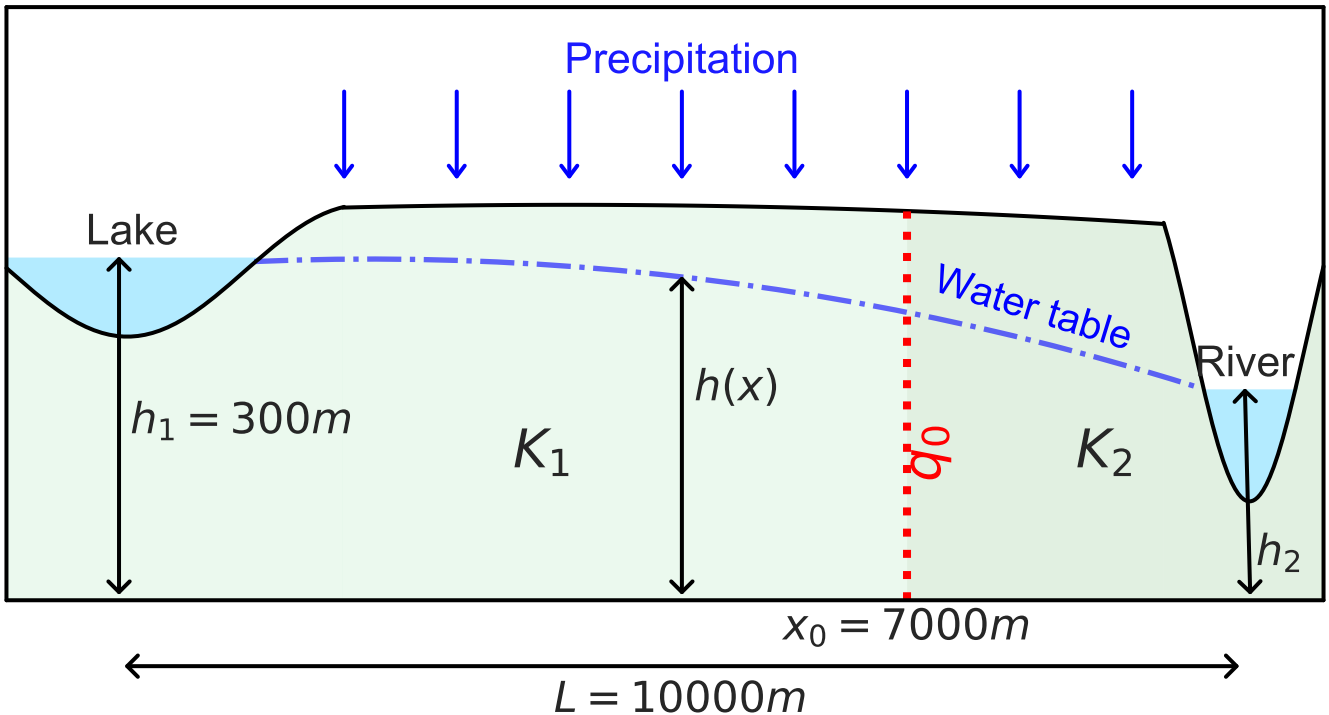

Precipitation
Lake
Water table
River
$h_1 = 300m$
$h(x)$
K_1
q_0
K_2
h_2
$x_0 = 7000m$
$L = 10000m$

Figure3.

(a) Fix M to M_1

$E_{\boldsymbol{\theta}_K|M_K} V_{\mathbf{M}_{-K}}(\Delta|M_K, \theta_K) = 0.61$

$R_1G_1M_1$
$R_1G_2M_1$
$R_2G_1M_1$
$R_2G_2M_1$

(b) Fix M to M_2

$E_{\boldsymbol{\theta}_K|M_K} V_{\mathbf{M}_{-K}}(\Delta|M_K, \theta_K) = 0.88$

$R_1G_1M_2$
$R_1G_2M_2$
$R_2G_1M_2$
$R_2G_2M_2$

(c) Fix R to R_1

$E_{\boldsymbol{\theta}_K|M_K} V_{\mathbf{M}_{-K}}(\Delta|M_K, \theta_K) = 1.75$

$R_1G_1M_1$
$R_1G_1M_2$
$R_1G_2M_1$
$R_1G_2M_2$

(d) Fix R to R_2

$E_{\boldsymbol{\theta}_K|M_K} V_{\mathbf{M}_{-K}}(\Delta|M_K, \theta_K) = 1.79$

$R_2G_1M_1$
$R_2G_1M_2$
$R_2G_2M_1$
$R_2G_2M_2$

(e) Fix G to G_1

$E_{\boldsymbol{\theta}_K|M_K} V_{\mathbf{M}_{-K}}(\Delta|M_K, \theta_K) = 2.43$

$R_1G_1M_1$
$R_1G_1M_2$
$R_2G_1M_1$
$R_2G_1M_2$

(f) Fix G to G_2

$E_{\boldsymbol{\theta}_K|M_K} V_{\mathbf{M}_{-K}}(\Delta|M_K, \theta_K) = 1.43$

$R_1G_2M_1$
$R_1G_2M_2$
$R_2G_2M_1$
$R_2G_2M_2$

PDF

q_0 $[m^2/d]$

Figure4.

(a1) Vary M
$R_1G_1M_1$
$R_1G_1M_2$
$E_{\boldsymbol{\theta}_{-K}|M_{-K}} V_{\mathbf{M}_K}(\Delta|M_{-K},\theta_{-K}) = 2.20$

(a2) Vary M
$R_1G_2M_1$
$R_1G_2M_2$
$E_{\boldsymbol{\theta}_{-K}|M_{-K}} V_{\mathbf{M}_K}(\Delta|M_{-K},\theta_{-K}) = 1.05$

(a3) Vary M
$R_2G_1M_1$
$R_2G_1M_2$
$E_{\boldsymbol{\theta}_{-K}|M_{-K}} V_{\mathbf{M}_K}(\Delta|M_{-K},\theta_{-K}) = 2.20$

(a4) Vary M
$R_2G_2M_1$
$R_2G_2M_2$
$E_{\boldsymbol{\theta}_{-K}|M_{-K}} V_{\mathbf{M}_K}(\Delta|M_{-K},\theta_{-K}) = 1.05$

(b1) Vary G
$R_1G_1M_1$
$R_1G_2M_1$
$E_{\boldsymbol{\theta}_{-K}|M_{-K}} V_{\mathbf{M}_K}(\Delta|M_{-K},\theta_{-K}) = 0.30$

(b2) Vary G
$R_1G_1M_2$
$R_1G_2M_2$
$E_{\boldsymbol{\theta}_{-K}|M_{-K}} V_{\mathbf{M}_K}(\Delta|M_{-K},\theta_{-K}) = 0.55$

(b3) Vary G
$R_2G_1M_1$
$R_2G_2M_1$
$E_{\boldsymbol{\theta}_{-K}|M_{-K}} V_{\mathbf{M}_K}(\Delta|M_{-K},\theta_{-K}) = 0.33$

(b4) Vary G
$R_2G_1M_2$
$R_2G_2M_2$
$E_{\boldsymbol{\theta}_{-K}|M_{-K}} V_{\mathbf{M}_K}(\Delta|M_{-K},\theta_{-K}) = 0.61$

(c1) Vary R
$R_1G_1M_1$
$R_2G_1M_1$
$E_{\boldsymbol{\theta}_{-K}|M_{-K}} V_{\mathbf{M}_K}(\Delta|M_{-K},\theta_{-K}) = 0.23$

(c2) Vary R
$R_1G_1M_2$
$R_2G_1M_2$
$E_{\boldsymbol{\theta}_{-K}|M_{-K}} V_{\mathbf{M}_K}(\Delta|M_{-K},\theta_{-K}) = 0.23$

(c3) Vary R
$R_1G_2M_1$
$R_2G_2M_1$
$E_{\boldsymbol{\theta}_{-K}|M_{-K}} V_{\mathbf{M}_K}(\Delta|M_{-K},\theta_{-K}) = 0.37$

(c4) Vary R
$R_1G_2M_2$
$R_2G_2M_2$
$E_{\boldsymbol{\theta}_{-K}|M_{-K}} V_{\mathbf{M}_K}(\Delta|M_{-K},\theta_{-K}) = 0.37$

PDF

$q_0\ [m^2/d]$

Figure5.

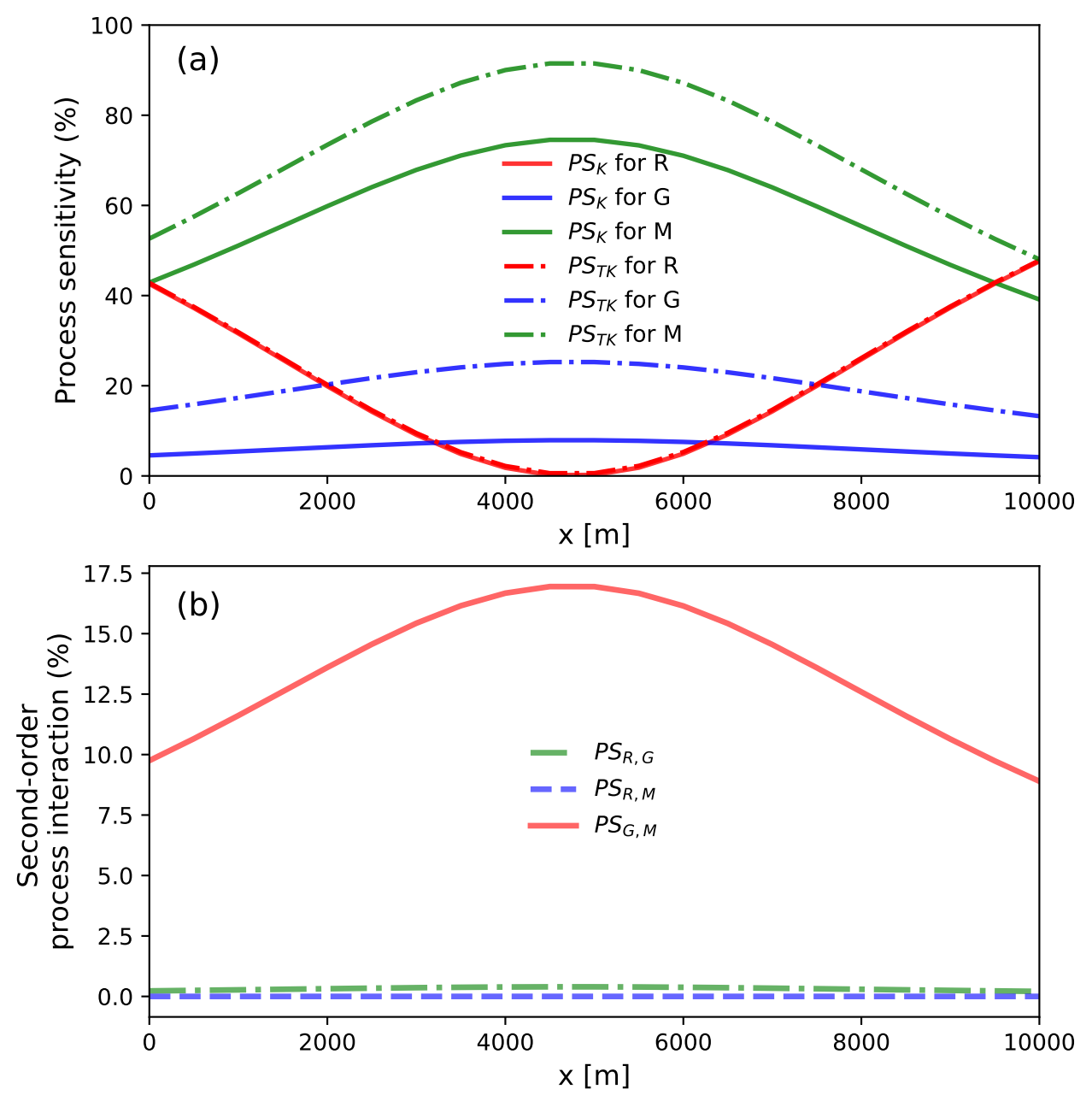
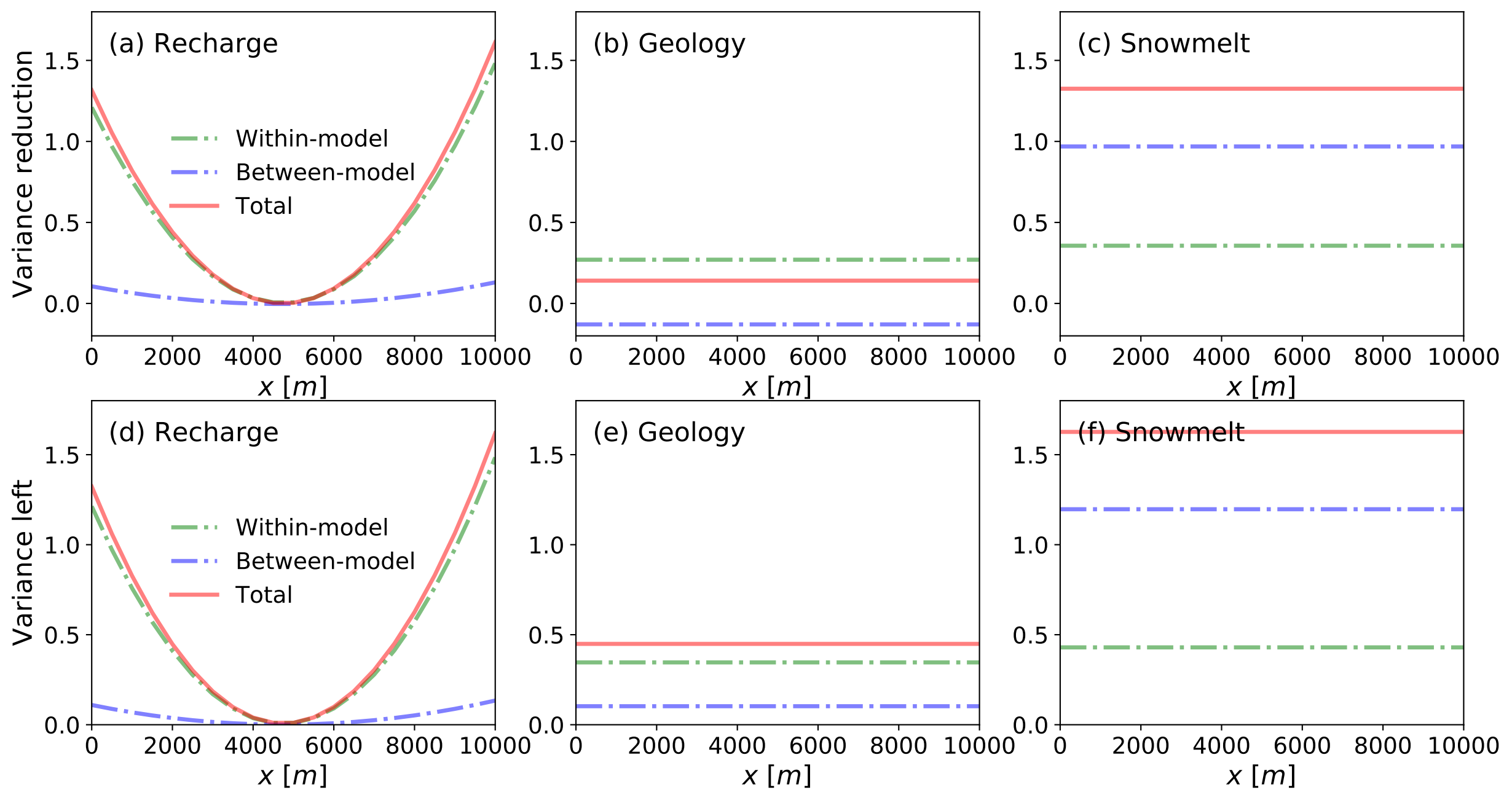
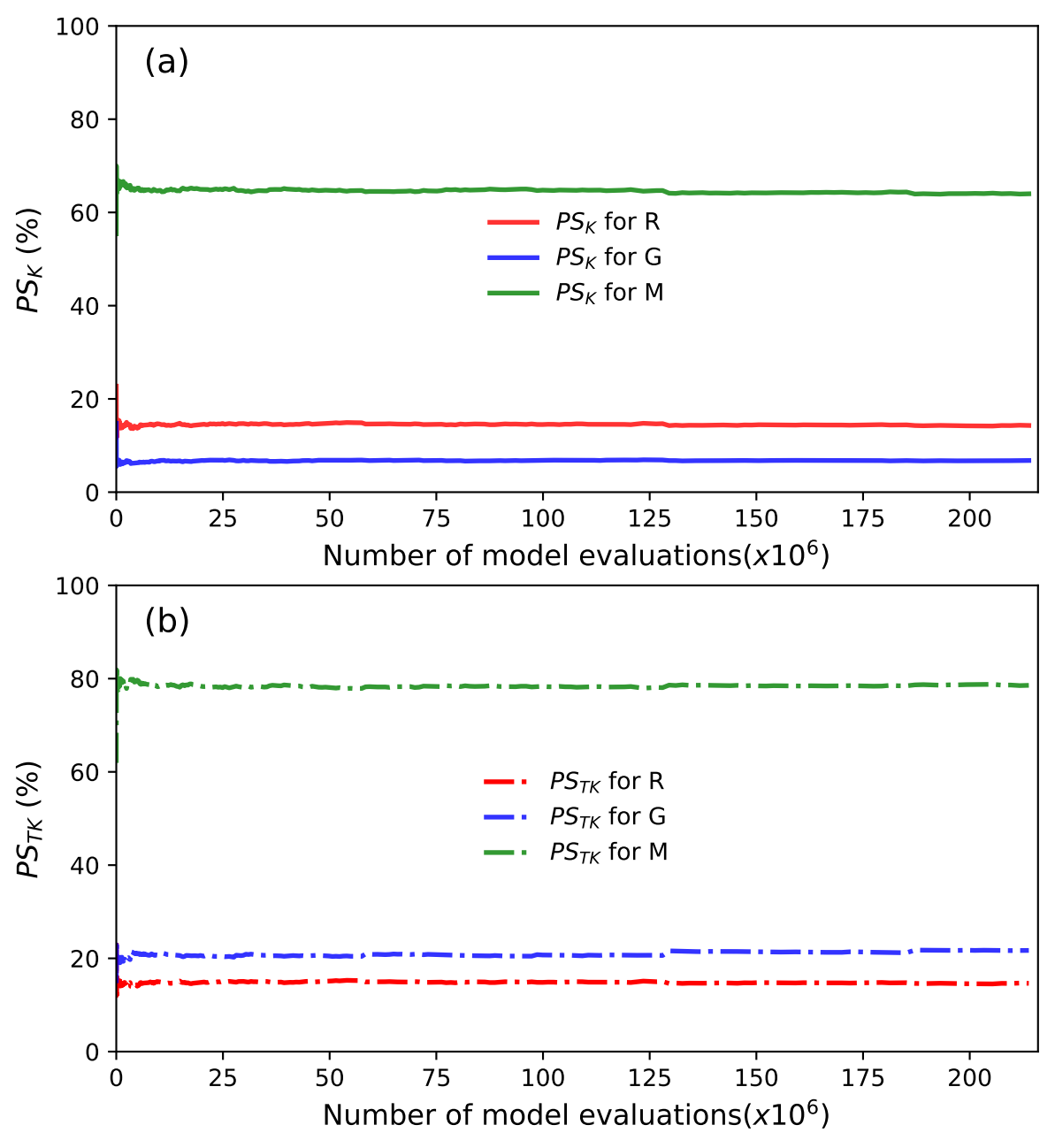

(a)
Process sensitivity (%)
100
80
60
40
20
0
0
2000
4000
6000
8000
10000
x [m]
PS_K for R
PS_K for G
PS_K for M
PS_{TK} for R
PS_{TK} for G
PS_{TK} for M
(b)
Second-order process interaction (%)
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
0
2000
4000
6000
8000
10000
x [m]
$PS_{R,G}$
$PS_{R,M}$
$PS_{G,M}$

Figure6.


(a) Recharge
(b) Geology
(c) Snowmelt
(d) Recharge
(e) Geology
(f) Snowmelt
Variance reduction
Variance left
Within-model
Between-model
Total
x [m]
0
2000
4000
6000
8000
10000
0.0
0.5
1.0
1.5

FigureC1.


(a)
PS_K (%)
PS_K for R
PS_K for G
PS_K for M
Number of model evaluations($x10^6$)
(b)
PS_{TK} (%)
PS_{TK} for R
PS_{TK} for G
PS_{TK} for M
Number of model evaluations($x10^6$)