VISCOUS: A Variance-Based Sensitivity Analysis Using Copulas for Efficient Identification of Dominant Hydrological Processes

R. Sheikholeslami[1], S. Gharari[2], S. M. Papalexiou[3] and M. P. Clark[2]

[1] Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, UK
[2] Centre for Hydrology, University of Saskatchewan, Saskatoon, Saskatchewan, Canada
[3] Department of Civil, Geological, and Environmental Engineering, University of Saskatchewan, Saskatoon, Saskatchewan, Canada

Key Points:

- VISCOUS bridges analysis of variance, copula theory, and Gaussian mixture models for global sensitivity analysis from given data
- VISCOUS identifies dominant hydrological processes, with much lower computational cost than the conventional sampling-based estimator
- VISCOUS yields robust estimates of the factor sensitivities and importance ranking for computationally expensive hydrologic models

Corresponding author: Razi Sheikholeslami (razi.sheikholeslami@ouce.ox.ac.uk)

Abstract

Global Sensitivity Analysis (GSA) has long been recognized as an indispensable tool for model analysis. GSA has been extensively used for model simplification, identifiability analysis, and diagnostic tests, among others. Nevertheless, computationally efficient methodologies are sorely needed for GSA, not only to reduce the computational overhead, but also to improve the quality and robustness of the results. This is especially the case for process-based hydrologic models, as their simulation time is often too high and is typically beyond the availability for a comprehensive GSA. We overcome this computational barrier by developing an efficient variance-based sensitivity analysis using copulas. Our data-driven method, called VISCOUS, approximates the joint probability density function of the given set of input-output pairs using Gaussian mixture copula to provide a given-data estimation of the sensitivity indices. This enables our method to identify dominant hydrologic factors by recycling pre-computed set of model evaluations or existing input-output data, and thus avoids augmenting the computational cost. We used two hydrologic models of increasing complexity (HBV and VIC) to assess the performance of the proposed method. Our results confirm that VISCOUS and the original variance-based method can detect similar important and unimportant factors. However, while being robust, our method can substantially reduce the computational cost. The results here are particularly significant for, though not limited to, process-based models with many uncertain parameters, large domain size, and high spatial and temporal resolution.

Plain Language Summary

Unraveling how various uncertain and interacting factors influence hydrologic models' behavior underscores the need for continued development of the effective tools for model analysis. Methodologies such as sensitivity analysis (SA) are powerful methods in this regard, as they provide information on how a variable of interest in the model changes over time or in space by varying its uncertain driving factors. However, many such methods are sampling-based techniques, for which two major issues preclude their efficient application. First, the need for an *ad-hoc* experimental design (i.e., sampling method) makes it impossible to reuse an existing ensemble of model runs or a generic sample of input-output data for SA. Second, sampling-based methods often require many model evaluations, which makes the computational burden unmanageable for complex modelling problems. These unique obstacles can hamper obtaining stable and robust results for computationally expensive models. We address this issue by proposing an efficient data-driven method, which can perform SA on fixed data sets, i.e., without the need for further sampling. Our method is based on copula theory and can successfully identify key drivers of uncertainty, when only a (small) sample of the input-output space is available.

1. Introduction

1.1. On the high computational cost incurred by the sampling-based global sensitivity analysis

With the rapid development of the computing capability and speed of processors, an increasing number of distributed and semi-distributed process-based hydrologic models are introduced to

simulate the quantity and quality of water on a range of spatiotemporal scales (Fatichi et al., 2016; Clark et al., 2017; Baroni et al., 2019). The ever-growing complexity of these models is driven by the need to represent Nature, leading to a greater sophistication of the modeled processes and a higher number of processes included, though, some observers suggest a sociological element linked to modelers' hubris (Saltelli et al., 2020). The increased complexity ultimately results in modification of the model structure through embedding new components into the model, for example, by adding extra parameters, new feedbacks, or changing the boundary conditions.

This high level of complexity in hydrological models can inevitably cause a high level of uncertainty, which needs to be quantified in order to glean useful information about the system's behavior and make robust decisions (Hall, 2007; Refsgaard et al., 2007; Clark et al., 2008; Gupta et al., 2008), notably in the face of deep uncertainty (Maier et al., 2016). Therefore, a comprehensive analysis of the behavior of such models can shed light on how an output of interest represents a distinct aspect of the hydrologic cycle, and can provide an informative view of the full spectrum of the system's behavior under different conditions. This is of vital importance, particularly when coping with a range of decision-making problems, such as investigating the system's response to a diverse range of plausible scenarios (e.g., future hydroclimatic conditions).

Global Sensitivity Analysis (GSA) (Saltelli et al., 2008; Slatelli et al., 1993) is one of the powerful tools for model analysis and has been broadly applied to a wide range of modelling problems, including model calibration, identifiability analysis, model simplification, and diagnostic testing (see, e.g., Rakovec et al., 2014; Guse et al., 2016; Markstorm et al., 2016; Huo et al., 2019; Quinn et al., 2019; Dell'Oca et al., 2020). GSA is well recognized to be an essential means of solving some of the outstanding challenges associated with complexity of the hydrologic models, including but not limited to (Razavi and Gupta, 2015): (1) determining importance of the input factors, which is helpful for factor prioritization and data acquisition; (2) detecting non-influential factors, whose variations have no impact on the behavior of the model and can be fixed to reduce model complexity; and (3) apportioning the total uncertainty of the output of interest to multiple sources of uncertainty.

However, a comprehensive GSA of the hydrologic models typically requires many model runs, and accordingly a tremendous amount of computational resources. This is mainly due to the high number of interacting, uncertain input factors causing the "curse of dimensionality" problem. This means that the number of samples required for GSA grows exponentially with the number of input factors/dimensionality of the model to maintain a desired level of stability and robustness (Saltelli et al., 2019). When using a sampling-based strategy to perform GSA, the curse of dimensionality along with the highly nonlinear nature of the response surfaces in hydrologic models become two major obstacles for efficient GSA. In other words, adequate exploration of the factor space and characterizing nonlinearity of the response surface necessitate a large sample size in traditional sampling-based GSA, which results in a considerable number of model evaluations.

The problem of high computational demand incurred by the sampling-based GSA manifests itself in the fact that a majority of the recent GSA studies in environmental modelling and hydrology (approximately 70%) have used relatively low-dimensional models (Sheikholeslami et al., 2019). In contrast, the number of factors in complex, state-of-the-art hydrologic models is relatively high. More importantly, due to the high number of model evaluations required by existing GSA techniques and the computationally expensive nature of the hydrologic models, analysts usually

tend to conduct GSA without evaluating its stability and convergence (for more discussion see, e.g., Nossent et al. 2011; Cosenza et al., 2013; Ferretti et al., 2016; Sarrazin et al. 2016; Mai and Tolson, 2019). It is therefore common to choose the sample size (number of model runs) only based on the available computational budget, which in turn can result in lack of robustness, and consequently contaminate the assessment of the sensitivities (Hill et al., 2016). Therefore, devising strategies to minimize the computational cost of the GSA techniques is one of the major challenges in the quest to understand hydrologic models' behavior. In the next subsection, we critically review three cost-effective strategies that have been adopted in the literature to accelerate GSA of the computationally expensive models.

1.2. Strategies for breaking down the barriers that deter an efficient GSA

1.2.1. Improved sampling strategies for GSA

Sampling from the input/factor space is a cornerstone of the sampling-based model analysis methods. An initial set of sample points needs to be drawn from the factor space in sampling-based GSA to extract the sensitivity-related information. However, the high dimensionality of the factor space, along with the nonlinearity of the model response surface, requires many properly distributed sample points. This issue can make sampling-based GSA very time-consuming, if not unattainable, particularly when the computationally expensive models are used. The number of model runs needed for a comprehensive GSA can be effectively reduced if the utilized sampling strategy conveys the maximum amount of information from the output space with minimum sample size (Andres, 1997). A number of studies have shown that utilizing an appropriate sampling plan for GSA can decrease the number of model runs several orders of magnitude (see, e.g., Castaings et al., 2012; Gan et al., 2014; Gong et al., 2015).

Therefore, any sampling-based GSA must be equipped with an improved sampling strategy (also known as an optimal design) that attempts to limit the computational load by optimizing some basic sample quality merits (Crombecq et al., 2011). A proper sample set should contain sample points that are uniformly distributed over the entire factor space such that all regions of the factor space are equally explored (this criterion is known as *space-fillingness*) (Pronzato and Müller, 2012; Sheikholeslami et al., 2019). As shown by Janouchová and Kučerová (2013), a sample set with poor space-filling properties can cause large errors when estimating sensitivity measures. Additionally, there are alternative sampling methods that were only designed for specific GSA methods (see, e.g., Chan et al., 2000; Saltelli 2002; Campolongo et al., 2007; Ruano et al., 2012; Razavi and Gupta, 2016). In a comprehensive study, Lo Piano et al. (2021) investigated computational efficiency of the several sample-based estimators for variance-based sensitivity indices. One of the main drawbacks of these sampling strategies is that the sample size should be specified a priori. In fact, the entire sample set is generated at one stage. Therefore, there is a risk of using larger sample sizes, which may unnecessarily increase the computational cost (Sheikholeslami and Razavi, 2017).

1.2.2. Emulation techniques for GSA

Constructing a cheaper-to-run emulator to replace the original computationally expensive model is another common approach to lower the computational cost in GSA (Ratto et al., 2012). Given a set of pre-existing model runs as a "training dataset", this strategy first fits a statistical model (called emulator or surrogate model) to approximate the relationship between input factors and model output. Then, sensitivity measures are estimated using the learned emulator via either Monte Carlo–based methods or analytical approaches. Examples of the Monte Carlo-based emulators for GSA include neural networks (Marseguerra et al., 2003) and polynomial regression (Iooss et al., 2006) (for a comprehensive comparison of various Monte Carlo-based emulation techniques see Storlie et al. (2009) and Verrelst et al. (2016)). On the contrary, there are emulators that allow to estimate sensitivity measures analytically from the fitted surrogate model, including Gaussian radial basis functions (Chen et al., 2005), polynomial chaos expansions (Sudret, 2008), Gaussian process model (Bounceur et al., 2015), complex linear model (Jourdan, 2012), and sparse polynomial chaos (Wu et al., 2020). Since variance-based method is an established good GSA practice, many of these emulation-based algorithms have been only developed for estimating variance-based sensitivity indices. Note that unlike the above-mentioned emulators which fit the (static) model output, dynamic emulation modelling provides a low-order emulator that preserves the dynamical nature of the original model. Monte Carlo-based approach is often used to identify dominant modes of dynamic behavior in dynamic emulation modelling (Young 1999; Young and Ratto, 2009).

Because emulators are only an approximation of the original model, the emulation-based GSA results are typically prone to two types of uncertainties (O'Hagan, 2006): (i) the code uncertainty and (ii) the uncertainty due to finite sample size. The former is associated with the emulator's degree of accuracy, while the latter is due to not knowing the output of the model at unsampled regions outside of the training dataset. Therefore, quantifying uncertainty of the emulation-based GSA results is necessary and can function as a quality check. Another eminent challenge is that these emulation techniques usually work well when the problem at hand has a relatively low number of factors. But when the dimensionality of the problem is reasonably high, they become practically unsuitable for finding influential factors due to the curse of dimensionality and over-fitting issues (Becker et al., 2018).

1.2.3. A given-data approach to GSA

A disadvantage of most of the efficient GSA algorithms cited so far is the requirement of a special sampling design. This means that available data from previous model runs (generated for other modelling purposes, e.g., for uncertainty propagation, model emulation, or calibration) cannot be reused for GSA purposes. To overcome this issue, the given-data approach (otherwise known as data-driven method) has been introduced. The given-data principle states that the sensitivity-related information can be extracted from a finite (generic) sample of existent dataset, independent of having a model and/or without re-running the model (Borgonovo et al. 2017). In this context, Pliscke (2010); Strong et al. (2012); Strong and Oakley (2013); Wainwright et al. (2014); Li and Mahadevan (2016) presented computationally frugal algorithms for calculating variance-based sensitivity indices. Plischke et al. (2013) also proposed a class of data-driven techniques, known as moment-independent methods, which do not rely on any specific moment. More recently,

Sheikholeslami and Razavi (2020) developed a new data-driven method based on theory of variogram analysis of response surfaces for GSA from given data. To the authors' knowledge, the given-data approach was first applied to hydrologic modelling by Borgonovo et al. (2017) and Sheikholeslami and Razavi (2020).

The rationale for given-data approach is as follows: If the sample of input-output data contains enough information to determine the dominant factors that exert strong influence on system's behavior, then the given-data estimator can be utilized as a post-processing GSA module for the pre-computed model evaluations. Using this approach, given a sample of input-output data representing the behavior of the system, which might be obtained from observations, laboratory or field experiments, we can directly calculate the sensitivity measures from a generic sample set, without having a model. In a burgeoning era of big data, this becomes more profitable than ever because data collection is becoming much easier and faster. Due to a large factor space of the high-dimensional models, generating '*representative*' input-output data for GSA may require many properly distributed sample points (or observations) to explore a full spectrum of the system's behavior. Given this, regardless of the chosen method for GSA from given data, it is beneficial to spend some time up front to find an optimal sample set by prudently using sampling algorithms before performing GSA.

Compared to aforementioned strategies that have been utilized to enable a computationally tractable GSA, the given-data approach has several advantages. First, the use of improved sampling techniques (**section 1.2.1**) does not eliminate the need for running the computationally expensive models. On the other hand, the given-data approach mainly focuses on situations where analysts want to conduct GSA retrospectively, i.e., using data from previous experiments, such as calibration or uncertainty analysis. Second, emulation-based GSA techniques (**section 1.2.2**) have major issues that render them less useful compared to given-data approach. Emulation-based GSA mainly suffers from two major drawbacks (i) choosing the best emulator is not easy due to the existing of a plethora of different emulators, and (ii) some emulators also need an ad-hoc experimental designs and specific arrangement of sample points for the sake of efficiency, while the given-data approach assumes that the sample is arbitrary.

1.3. Objectives and scope

The primary goal of this paper is to develop a data-driven variance-based GSA method that is based on the copula functions. The method is mainly designed to alleviate the high-computational demand associated with GSA of the process-based hydrologic models. At present, unlike the sampling-based GSA methods, most of the given-data GSA methods only focus on estimating the variance-based first-order sensitivity index (i.e., main effect) and cannot estimate the total effect index and interactions (see, e.g., Strong et al., 2012; Strong and Oakley, 2013; Sparkman et al., 2016). Hence, the secondary objective of this paper is to develop an efficient data-driven method to simultaneously compute first-order and total effect indices.

Our proposed GSA method (hereafter called VISCOUS) utilizes a class of functions known as Gaussian mixture copula (Tewari et al., 2011) to model the joint probability distribution of the inputs and outputs, which might have multimodal densities and nonlinear dependencies. By approximating the conditional distribution of the output, when the inputs are conditioned at

particular values, VISCOUS extracts sensitivity-related information directly from the given data, without augmenting the computational cost. This avoids the need for additional model runs and will help modelers to efficiently identify dominant factors, particularly when dealing with computationally expensive models. Furthermore, VISCOUS can be effectively used to rank the strength of relationships in multivariate data sets (not shown in this paper). We demonstrate how well VISCOUS approximates the variance-based sensitivity indices, when only a fixed amount of input-output data is available, using two real-world case studies. These include a conceptual rainfall-runoff model (HBV) and a physically-based hydrologic model (VIC) configured for simulating the streamflow values of a head water basin in the Canadian Rockies.

The reminder of this paper is organized as follows: in **Section 2** we briefly review the conventional variance-based GSA algorithm, summarize the concepts of copulas and mixture models, present our copula-based data-driven GSA method, and discuss its implementation details. In **Section 3**, we describe two hydrologic modelling case studies followed by explaining the setup for computational experiments. **Section 4** applies the proposed GSA method to these case studies and reports numerical results and analyses. Finally, we conclude the paper in **Section 5** and provide recommendations for future work.

2. Methodology

2.1. Original variance-based GSA method

We denote the output of a model $y = g(\boldsymbol{X})$ as a function of an m-dimensional input vector $\boldsymbol{X} = [x_1, x_2, \dots, x_m]^T$, where x_i $(i = 1,2, \dots, m)$ is an uncertain factor of the model, such as model parameters, initial conditions, or boundary conditions. Here, we also assume $g(.)$ is a deterministic, scalar function of input variables. In the context of variance-based GSA, uncertainty of the input vector $\boldsymbol{X}$ is often modeled in a probabilistic framework, i.e., $\boldsymbol{X}$ is deemed to be a random vector and uncertainty of the $\boldsymbol{X}$ propagates through $g(\boldsymbol{X})$. Therefore, the model output uncertainty can be expressed in terms of its variance. Basically, in variance-based GSA the main goal is to apportion the model output variance into fractions attributed to each input factor and their interactions. As a result, we can quantify how much the variability of an individual (uncertain) input factor x_i (and its interaction with other factors) contributes to the variability of the output. To do this, Sobol method (Sobol, 2001) uses a well-known variance decomposition formula (known as generalized Hoeffding-Sobol decomposition, Hoeffding, 1948), as follows:

$$V_y = var\big(g(\boldsymbol{X})\big) = \sum_{i=1}^{m} V_i + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} V_{ij} + \cdots + V_{1\dots m} \qquad (1)$$

In **Eq. (1)**, variance of the $g(\boldsymbol{X})$ is decomposed into the partial variances, where the first-order partial variance, V_i, represents the contribution of an individual input factor x_i to the output variance V_y, i.e.,

$$V_i = var\big[\mathrm{E}_{\boldsymbol{X}_{\sim i}}(y|x_i)\big] \qquad (2)$$

In other words, if we fix x_i, the expected reduction in output variance will be V_i. Moreover, V_{ij} describes the portion of the output variance explained by the interaction between x_i and x_j, and so on for higher order interactions (V_{ijk}, ...). Variance-based GSA uses a set of indices to measure the importance of each input factor. By normalizing **Eq. (1)** by the total variance, V_y, we obtain:

$$\sum_{i=1}^{m} S_i + \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} S_{ij} + \cdots + S_{1\ldots m} = 1 \tag{3}$$

where $S_i = {V_i}/{V_y}$, $S_{ij} = {V_{ij}}/{V_y}$, ... , $S_{1\ldots m} = {V_{1\ldots m}}/{V_y}$

S_i is the first-order sensitivity index (henceforth called main effect), which corresponds to x_i and measures main effect of the x_i on model output. S_{ij} is the second-order sensitivity index and accounts for the interaction between x_i and x_j. Similarly, the higher order indices can be defined.

Because the total number of terms in **Eq. (3)** is equal to $2^m - 1$, quantifying all higher order indices becomes computationally prohibitive for models with many factors (large values of *m*). In such situations, an alternative variance-based measure, known as total effect index (Homma and Saltelli, 1996; Saltelli et al., 2000), is typically used to assess the higher order interaction effects. This index measures the first and all higher order effects of an individual input factor x_i by:

$$S_{Ti} = S_i + \sum_{j \neq i}^{m} S_{ij} + \cdots + S_{ij\ldots m} = 1 - \frac{var\left[\mathrm{E}_{x_i}(y|\boldsymbol{X}_{\sim i})\right]}{V_y} = \frac{\mathrm{E}_{\boldsymbol{X}_{\sim i}}[var(y|\boldsymbol{X}_{\sim i})]}{V_y} \tag{4}$$

where $\boldsymbol{X}_{\sim i} = [x_1, \ldots, x_{i-1}, x_{i+1}, \ldots x_m]^T$ is the vector of all input factors except x_i. **Eq. (4)** can be interpreted as follows: if we fix all input factors, except x_i, then the expected variance that is left can be expressed by S_{Ti}. In fact, it is easy to show that the interaction effect between x_i and all other factors can be calculated by subtracting S_i from S_{Ti} (Saltelli, 2002).

The main and total effect variance-based sensitivity indices are often estimated using a sampling-based strategy, wherein the Monte Carlo integration technique is applied to **Eq. (2)** and **(4)**. Several numerical schemes for sampling-based strategy have been developed to compute theses indices (Tarantola et al., 2006; Kucherenko and Song, 2017). In this study, a widely-used and computationally affordable algorithm introduced by Saltelli (2002) and Saltelli et al. (2010) is utilized. We briefly describe this algorithm in the following as the basis for the comparison of our method with conventional variance-based GSA.

The sampling-based algorithm starts by randomly generating two independent sample matrices: $\boldsymbol{A}$ and $\boldsymbol{B}$, each of size $N \times m$. Then, for calculating the *i*-th sensitivity index, an additional sample matrix $\boldsymbol{C}^{(i)}$ $(i = 1,2, \ldots, m)$ of size $N \times m$ is constructed by combining the columns of $\boldsymbol{A}$ and $\boldsymbol{B}$, such that $\boldsymbol{C}^{(i)}$ contains all columns of $\boldsymbol{B}$ except the *i*-th column which is taken from $\boldsymbol{A}$. Using the Jansen (1999)'s estimator, the sensitivity indices can be obtained from:

$$var\left[\mathrm{E}_{\boldsymbol{X}_{\sim i}}(y|x_i)\right] = V_y - \frac{1}{2N}\sum_{k=1}^{N}\left(g(\boldsymbol{B}_k) - g\left(\boldsymbol{C}^{(i)}{}_k\right)\right)^2 \tag{5}$$

$$\mathrm{E}_{\boldsymbol{X}_{\sim i}}[var(y|\boldsymbol{X}_{\sim i})] = \frac{1}{2N}\sum_{k=1}^{N}\left(g(\boldsymbol{A}_k) - g\left(\boldsymbol{C}^{(i)}{}_k\right)\right)^2 \quad (6)$$

where $\boldsymbol{A}_k$ denotes the k-th row of the matrix $\boldsymbol{A}$.

Estimating both S_i and S_{Ti} using **Eq. (5)** and **(6)** results in $N(m+2)$ model evaluations for which $2N$ simulations are needed for evaluating model output values corresponding to $\boldsymbol{A}$ and $\boldsymbol{B}$ (i.e., $g(\boldsymbol{A})$ and $g(\boldsymbol{B})$) and $N \times m$ simulations are required to obtain $g\left(\boldsymbol{C}^{(i)}\right)$ for all $i = 1,2,\ldots,m$. Note that achieving convergence and stability when computing variance-based sensitivity indices requires a sufficiently large samples size, i.e., N is typically selected to be some arbitrarily large number.

2.2. The proposed copula-based method for GSA from given data

Our proposed given-data approach (VISCOUS) is a probabilistic framework that estimates variance-based GSA indices from a given input-output sample. VISCOUS is a non-parametric approach and does not require generating new sample set for GSA, instead it recycles the existing data for determining factor importance. An essential building block of our proposed method is to infer the joint probability density function of the given data using copula-based models. The theoretical background and implementation details of VISCOUS are provided in the following subsections.

2.2.1. Linking copula theory to variance-based GSA

Estimating the expected value of the model output conditioned on a factor, $\mathrm{E}_{\boldsymbol{X}_{\sim i}}(y|x_i)$, is the cornerstone of the variance-based method (see **Eq. (2)** and **(4)**). For a given set of model output values ($\boldsymbol{Y}$) obtained from running a simulation model using a sample of input factors ($\boldsymbol{X}$), let $f_{\boldsymbol{X},\boldsymbol{Y}}$ denote the multivariate joint probability density function (PDF) of $\boldsymbol{Y}$ and $\boldsymbol{X}$. Assume the marginal PDF along the i-th direction is f_{x_i} and the conditional distribution of $\boldsymbol{Y}$ on $\boldsymbol{X}$ is $f_{\boldsymbol{Y}|\boldsymbol{X}}$. Then, the conditional expectation can be written as (Beckman and McKAy, 1987):

$$\mathrm{E}_{\boldsymbol{X}_{\sim i}}(y|x_i) = \int y f_{\boldsymbol{Y}|\boldsymbol{X}}(y|x_i)\,dy = \int y \frac{f_{\boldsymbol{X},\boldsymbol{y}}(x_i,y)}{f_{x_i}}dy \quad (7)$$

Therefore, if $f_{\boldsymbol{X},\boldsymbol{Y}}$ accurately describes the joint PDF of the input-output sample $(\boldsymbol{X},\boldsymbol{Y})$, we can simply compute various sensitivity indices based on **Eq. (7)**. This has been motivated a number of studies to investigate several strategies for characterizing the joint PDF of the input-output data for the purpose of variance-based GSA (see, e.g., Cheng et al., 2015; Wei et al., 2015; Jia and Taflanidis, 2016; Sparkman et al., 2016). In this paper, we rely on the copula theory which bridges one-dimensional marginal distributions to multidimensional joint distribution. Copula theory provides a powerful tool to model dependence structure between random variables from their associated marginals. Based on Sklar (1959)'s theorem, the joint cumulative distribution function

(CDF) of the input-output sample $F(\boldsymbol{X}, \boldsymbol{Y})$, can be expressed as a unique function of the marginal distributions:

$$F(\boldsymbol{X}, \boldsymbol{Y}) = \mathcal{C}\big(F(\boldsymbol{X}), F_y(\boldsymbol{Y});\ \lambda\big) \tag{8}$$

where $F(\boldsymbol{X}) = [u_1, u_2, \dots, u_m]$ is the vector of marginal CDFs of $\boldsymbol{X}$, $F_y(\boldsymbol{Y}) = u_{\boldsymbol{Y}}$ is the CDF of the model output, and $\mathcal{C}$ is a copula function parameterized by λ. The kernel density function can be used to transform the input-output sample into $[u_1, u_2, \dots, u_m, u_{\boldsymbol{Y}}]$.

Therefore, if the copula function and the marginal distributions are differentiable, the joint PDF of the $(\boldsymbol{X}, \boldsymbol{Y})$ can be obtained as product of the individual marginal densities and the copula density:

$$f_{\boldsymbol{X},\boldsymbol{Y}} = c\big(F(\boldsymbol{X}), F_y(\boldsymbol{Y});\ \lambda\big) \times f_{x_1} \times f_{x_2} \dots \times f_{x_m} \times f_{\boldsymbol{Y}} \tag{9}$$

where c is the PDF of the copula function, $f_{\boldsymbol{Y}}$ and f_{x_i} $(i = 1,2, \dots, m)$ are the PDFs of the model output and input variables; respectively. If the forms of the marginal PDFs and the copula density are known, **Eq. (9)** is utilized for estimating the unknown multivariate joint PDF. On the other hand, in case of a known multivariate joint PDF, the copula density can be derived by rearranging **Eq. (9)** as:

$$c\big(F(\boldsymbol{X}), F_y(Y); \lambda\big) = \frac{f_{\boldsymbol{X},\boldsymbol{Y}}}{f_{x_1} \times f_{x_2} \dots \times f_{x_m} \times f_y} \tag{10}$$

For example, in case of standard normal multivariate Gaussian, having marginal densities with zero mean and unit variance along all dimensions, **Eq. (10)** can be solved analytically to obtain *Gaussian copulas*. The Gaussian copula has been previously used in conjunction with various GSA algorithms, including method of Morris (Ţene et al., 2018), variance-based method (Sainte-Marie and Cournède, 2019), and moment-independent method (Wei et al., 2014; Plischke and Borgonovo, 2019). However, the usefulness of the formal Gaussian copula can be severely limited, when the input-output sample has multiple modes with nonlinear relationships. This issue is crucial for GSA of the hydrologic models, as it is not uncommon to encounter a highly complex and nonlinear response surfaces with multimodality, sharp discontinuities, and small-scale roughness in hydrology (Duan et al., 1992; Kavetski and Clark, 2010).

Alternatively, it has been shown that other types of copula functions such as empirical checkerboard copula (Deheuvels, 1979) and Gaussian mixture copula model (Tewari et al., 2011; Li et al., 2011) are more effective in modeling nonlinear relationships, particularly of non-Gaussian data (see, e.g., Genest et al., 2017; Bilgrau et al., 2016; Fan et al., 2016; Kasa et al., 2020). In the present work, to capture the complex behavior of the hydrologic models, we developed an estimator based on Gaussian mixture copula model for constructing the joint PDFs, as presented in the following sub-section. Note that our proposed framework for GSA is not limited to the Gaussian mixture copula models and can be extended to any types of copula functions.

2.2.2. Gaussian mixture copula-based estimator for sensitivity indices

Gaussian mixture copula model (GMCM) combines advantages of the copula theory with capability of the Gaussian mixture models (GMM) in approximating an arbitrary PDF, wherein GMM is defined as a mixture statistical model of a finite number of Gaussian distributions. In fact, GMCM infers the dependence structure of the multivariate data through coupling GMM into copulas. Using **Eq. (10)**, GMCM estimates the PDF of the copula function by (Tewari et al., 2011):

$$c_{GMCM}\left(F(\boldsymbol{X}), F_y(Y); \boldsymbol{\Theta}\right) = \frac{\psi(\boldsymbol{\Psi}^{-1}(u_1), \boldsymbol{\Psi}^{-1}(u_2), \dots, \boldsymbol{\Psi}^{-1}(u_{\boldsymbol{Y}}); \boldsymbol{\Theta})}{\phi\left(\boldsymbol{\Psi}^{-1}(u_1)\right) \times \phi\left(\boldsymbol{\Psi}^{-1}(u_2)\right) \times \dots \times \phi\left(\boldsymbol{\Psi}^{-1}(u_{\boldsymbol{Y}})\right)} \tag{11}$$

where $\boldsymbol{\Psi}^{-1}$denotes the inverse function of the standard normal distribution, and ψ is the joint PDF of GMM, which can be expressed by the weighted sum of η-component Gaussian densities ϕ, i.e.,

$$\psi(\boldsymbol{z}; \boldsymbol{\Theta}) = \sum_{k=1}^{\eta} \omega_k \times \phi\left(\boldsymbol{z}; \boldsymbol{\mu}^{(k)}, \boldsymbol{\Sigma}^{(k)}\right) \tag{12}$$

where $\boldsymbol{z} = [\boldsymbol{\Psi}^{-1}(u_1), \boldsymbol{\Psi}^{-1}(u_2), \dots, \boldsymbol{\Psi}^{-1}(u_{\boldsymbol{Y}})]$, ω_k is the weight for the *k*-th GMM component (note that all elements of the ω_k sum to unity), $\boldsymbol{\mu}^{(k)}$ is the mean vector, $\boldsymbol{\Sigma}^{(k)}$ represents the covariance matrix for the *k*-th component, and $\boldsymbol{\Theta} = \left\{\omega_k, \boldsymbol{\mu}^{(k)}, \boldsymbol{\Sigma}^{(k)}\right\}$ comprises all the weights, mean vectors, and covariance matrices for $k = 1,2, \dots, \eta$.

After building GMCM, the conditional expectation can be calculated by combining **Eq. (9)** and **(11)** with **(7)**. Importantly, the integration in **Eq. (7)** can be performed without re-running the original computer model (or physical model), which makes VISCOUS a computationally frugal method. We use a Monte Carlo-based approximation to evaluate the integral, as below:

$$\mathrm{E}_{\boldsymbol{X}_{\sim c}}(y|\boldsymbol{X} = \boldsymbol{x}_c) \approx \frac{1}{N_{MC}} \sum_{j=1}^{N_{MC}} {F_y}^{-1}\left(u_{\boldsymbol{Y}}{}^{(j)}\right) c_{GMCM}\left(\boldsymbol{u}_c, u_{\boldsymbol{Y}}{}^{(j)}; \boldsymbol{\Theta}\right) \tag{13}$$

$$\begin{aligned} var\left[\mathrm{E}_{\boldsymbol{X}_{\sim c}}(y|\boldsymbol{X} = \boldsymbol{x}_c)\right] \\ \approx \frac{1}{2{N_{MC}}^3} \sum_{i=1}^{N_{MC}} \sum_{j=1}^{N_{MC}} \left(\sum_{k=1}^{N_{MC}} {F_y}^{-1}\left(u_{\boldsymbol{Y}}{}^{(k)}\right) c_{GMCM}\left(\boldsymbol{u}_c{}^{(i)}, u_{\boldsymbol{Y}}{}^{(k)}; \boldsymbol{\Theta}\right) \right. \\ \left. - \sum_{k=1}^{N_{MC}} {F_y}^{-1}\left(u_{\boldsymbol{Y}}{}^{(k)}\right) c_{GMCM}\left(\boldsymbol{u}_c{}^{(j)}, u_{\boldsymbol{Y}}{}^{(k)}; \boldsymbol{\Theta}\right) \right)^2 \end{aligned} \tag{14}$$

where $\boldsymbol{X} = \boldsymbol{x}_c$ is the given random input factor of interest for which the Sobol indices should be estimated, N_{MC} is the number of Monte Carlo samples for integration, ${F_y}^{-1}(.)$ is the inverse CDF of $\boldsymbol{Y}$, and $\boldsymbol{u}_c = F_c(\boldsymbol{X} = \boldsymbol{x}_c)$ is the marginal CDF of the input factor of interest.

2.2.3. Steps for performing VISCOUS

There are six main steps involved in applying the proposed VISCOUS framework. In the following we outline the algorithmic implementation of these steps in detail:

Step 1. From the given data extract $\boldsymbol{X} = \boldsymbol{x}_c$ (and the corresponding output values $\boldsymbol{Y}$) for which the variance-based sensitivity indices should be estimated.

Step 2. Use kernel density function to convert the given data matrix $(\boldsymbol{X}, \boldsymbol{Y})$ into marginal CDFs, i.e., $F(\boldsymbol{X}) = [u_1, u_2, \dots, u_m]$ and $F_y(\boldsymbol{Y}) = u_{\boldsymbol{Y}}$,.

Step 3. Transform marginal CDFs of **Step 2** into data of standard normal distribution, i.e., $\boldsymbol{\Psi}^{-1}(u_1), \boldsymbol{\Psi}^{-1}(u_2), \dots, \boldsymbol{\Psi}^{-1}(u_{\boldsymbol{Y}})$.

Step 4. Build a mixture model (i.e., GMM) by fitting η Gaussian density components to the data obtained from **Step 3**. If p is the number of indices that GSA needs to be performed, we build p copula models for calculating each index. For example, estimating $\{S_1, S_2, S_{12}\}$ requires constructing $p = 3$ copula models. A critical step here is choosing optimal number of the components, η, in **Eq. (12)**. To find the optimal value of η, we utilized a likelihood-based goodness-of-fit statistic known as Bayes Information Criterion (BIC). BIC compares multiple models with varying values of η and chooses the optimal number of components associated with the lowest BIC value (Steele and Raftery, 2010).

Step 5. Fit GMCM copula (**Eq. (11)**) using the data of **Step 3** and GMM obtained from **Step 4**. We employed the expectation-maximization (EM) algorithm of Tewari et al. (2011) for estimating GMCM's parameters. Starting from initial values for parameters $\boldsymbol{\Theta}$, the EM algorithm iteratively attempts to maximize the likelihood function. To avoid the risk of trapping into local optima., we used several starting points within an iterative loop for global optimum investigation. To do this, we applied *k*-means clustering strategy for initialization of the EM algorithm (for more details see McLachlan and Peel (2000)). Note that a similar strategy was implemented in the R-package of Bilgrau et al. (2016) and *fitgmdist* function of the MATLAB for parameter estimation.

Step 6. Compute the variance-based sensitivity indices using **Eq. (13)** and **(14)** based on the GMCM model learned from **Step 5**. We arbitrarily set $N_{MC} = 10^4$ to reduce the numerical error of integration.

In **Appendix**, we exemplify the use of VISCOUS and demonstrate its efficiency and effectiveness by focusing on an analytical test function adopted from Kucherenko et al. (2011).

We conclude this section by highlighting the key distinctions between our proposed given-data approach and emulation techniques for GSA. First, the latter is mainly designed to approximate an output value corresponding to an individual realization of the uncertain input factors (see **section 1.2.2**), while the former approximates a distribution of the output given probability distributions of the input factors. Particularly, for a given set of input-output data, VISCOUS trains GMCM in the probability space, but emulators are always constructed in the output space. Second, to perform GSA and assess the output uncertainty associated with uncertainty of a given input factor, emulators typically need to be evaluated many times for multiple realizations of that uncertain

factor. On the other hand, one run of our non-parametric approach yields the conditional probability distributions of the output, without repeated runs of the model (or surrogate model).

Another major difference is that the dimensionality of the joint PDF obtained by GMCM in VISCOUS is independent of the dimensionality of the original model. This makes VISCOUS a suitable candidate in case of high-dimensional models for which fitting an emulator often suffers from the curse of dimensionality (see **section 1.2.2)**. In VISCOUS, the dimensionality of the joint PDF is $s + 1$, where s is the order of the variance-based GSA indices. By way of example, the joint PDF is only two-dimensional when computing the first-order indices.

3. Numerical Experiments

3.1. Study site

In this study, we set up two hydrologic models for simulating the streamflow values of the Bow River at Banff, Alberta, Canada (**Figure 1(a)**) with a basin area of approximately 2,210 km^2. The river is in the eastern Front Range of the Rocky Mountains. The temperature in this part of the Rockies can dip below –40°C in winter on mountain tops and rise up to +40°C in summer in Bow Valley (basin average temperature are depicted in **Figure 2**). Most parts of the basin receive a large proportion of their precipitation as snowfall (~50%), which results in substantial water storage as snowpack, and thus snowmelt (beginning in mid-April) is the dominant process that strongly controls the dynamics and generation of the runoff in the Bow River Basin (Nivo-glacial regime). The average basin elevation is 2,130 m ranging from 1,380 m above mean sea level at the outlet (town of Banff) to above 3,000 m at the highest elevations (**Figure 1(b)**). The basin annual precipitation is approximately 1,000 mm with range of 500 mm for the Bow Valley up to 2,000 mm for the mountain peaks (based on weather research and forecasting model, WRF, simulation for continental United States, Rasmussen and Liu (2017)). The predominant land cover (**Figure 1(c)**) is conifer forest in the Bow Valley, and rocks and gravels for mountain peaks above the tree line.

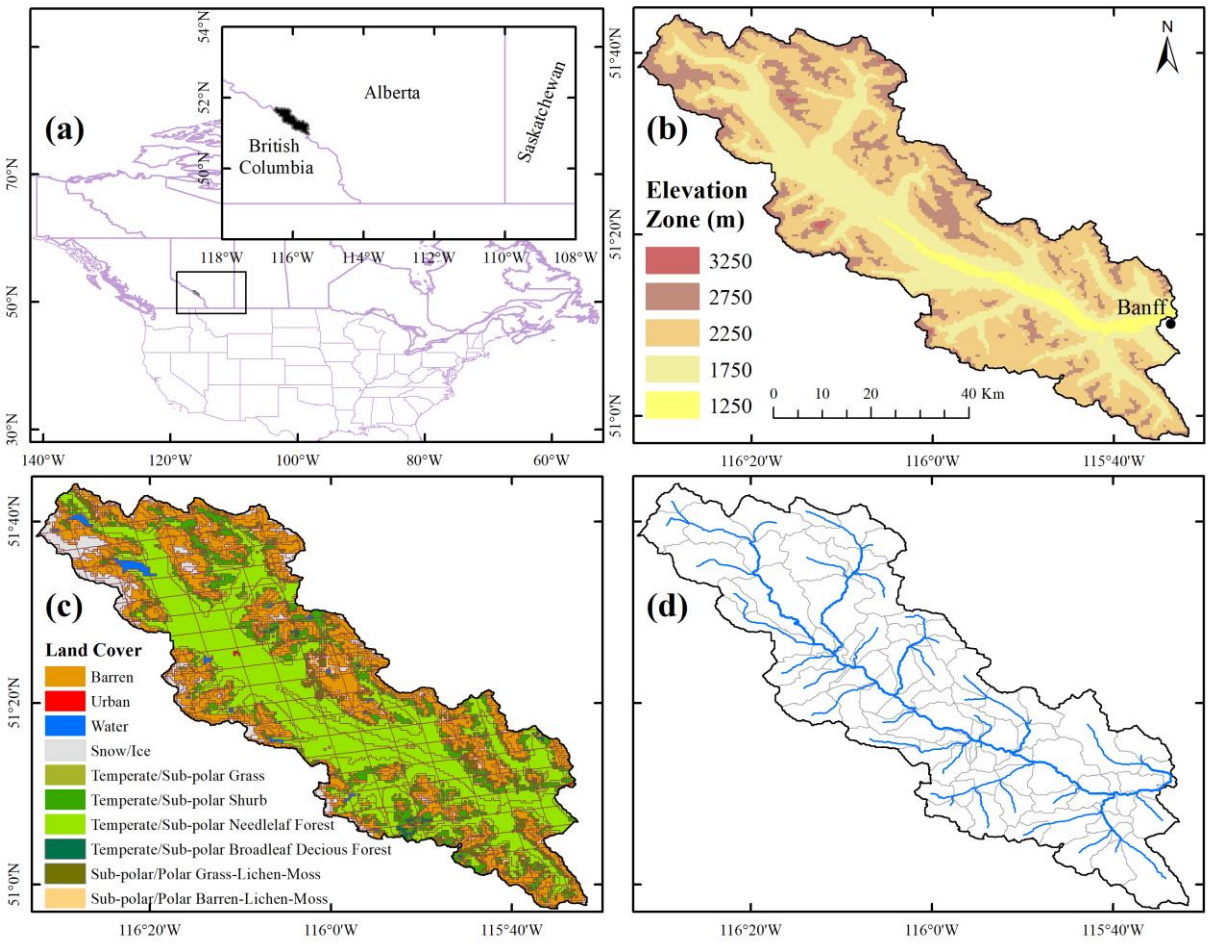

Figure 1. The upper panel shows **(a)** the location of the Bow River Basin at Banff, and **(b)** Bow River Basin elevation. The lower panel shows **(c)** computational units forced at WRF original resolution at 4 km color coded based on land cover type, and **(d)** river network topology and associated sub-basins that is used for the vector-based routing.

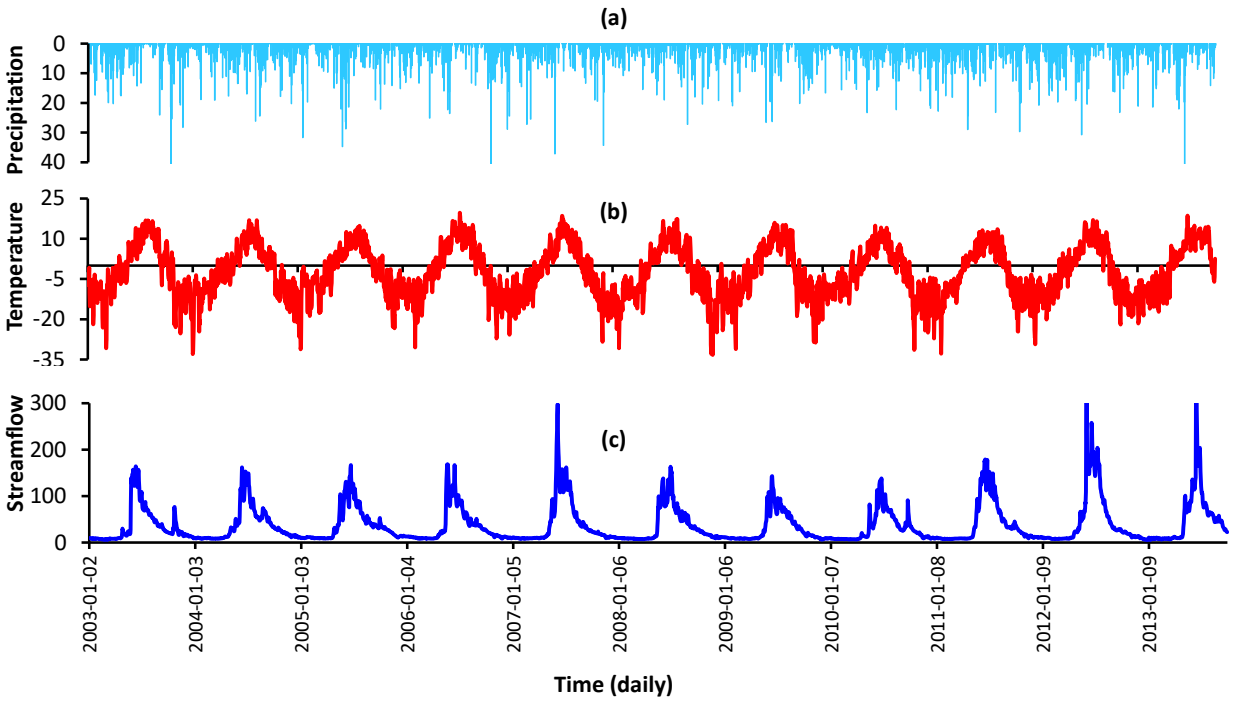

Figure 2. Time series plots illustrating **(a)** daily precipitation (upper panel), **(b)** temperature (middle panel), and **(c)** observed streamflow (lower panel) for the Bow River Basin at Banff, Alberta, Canada, from 2003 to 2013.

3.2. A conceptual rainfall-runoff model: HBV

We first illustrate the utility of the proposed GSA algorithm using the HBV model, which is a lumped, conceptual hydrologic model (Bergström,1995; Seibert, 1997). HBV simulates daily streamflow timeseries using input data of mean temperature, precipitation (daily values) and potential evapotranspiration (monthly estimates). The model consists of three main components: (i) snow module, (ii) soil moisture module, and (iii) flow routing module. The latter is composed of two reservoirs for quick flow (non-linear reservoir) and slow flow generations (linear reservoir). HBV transforms the simulated outflow using a triangular weighting function for channel routing. In this study, we implemented a slightly modified version of the HBV model developed by Razavi et al. (2019). The total number of calibration parameters was set to 10. **Table 1** summarizes parameter description, units of measurements, and their feasible ranges of variation. Note that the feasible ranges of these parameters were determined using a combination of expert knowledge and previous studies.

Table 1. Parameters of the HBV model used in this study

Parameter[unit]	Range	Definition
TT [°C]	[-4,4]	Air temperature threshold for melting/freezing
C0 [mm/°C per day]	[0,10]	Base melt factor
ETF [1/°C]	[0,1]	Temperature anomaly correction of potential evapotranspiration
LP [mm]	[0,1]	Soil moisture below which evaporation becomes supply limited
FC [mm]	[50,500]	Field capacity of soil
beta [-]	[0,7]	Shape parameter for soil release equation
FRAC [-]	[0.05,0.9]	Fraction of soil release entering fast reservoir
K1 [-]	[0.05,1]	Fast reservoir coefficient
alpha [-]	[1,3]	Shape parameter for fast reservoir equation
K2 [-]	[0,0.05]	Slow reservoir coefficient

3.3. A distributed physically-based hydrologic model: VIC

As the second case study we utilized a more computationally expensive, process-based model, known as Variable Infiltration Capacity (VIC). The VIC model is developed as a simple land model (Liang et al., 1994) which simulates both mass and energy fluxes. The model setup used in this study has 18 parameters and is based on the concept of flexible vector-based configuration for land models (for more details on the model set up refer to configuration Case-2-4km in Gharari et al., 2020). Description of the model parameters are presented in **Table 2**. Parameter ranges were specified using a combination of expert knowledge and previous studies, and recommendations in the manual.

In our implementation, the VIC model considers soil type, vegetation, and is forced with 7 meteorological variables from WRF simulations with spatial resolution of 4km (Rasmussen and Liu 2017). The soil column is modeled using three layers, wherein the upper two layers are responsible for the surface flow generation, and the third layer generates the slow response runoff. Furthermore, the traditional 4-parameter baseflow layer is replaced with a simplified 1-parameter layer instead (for further implementation details refer to Gharari et al., 2020). In the VIC model, the saturation excess mechanism generates all fast response runoff, and a gravity-driven process controls the drainage between the three soil layers of each grid cell (Liang et al., 1994).

Furthermore, the streamflow was simulated using a vector-based routing tool named mizuRoute (Mizukami et al., 2016). mizuRoute uses a Gamma distribution-based unit hydrograph to route runoff from hillslope to the river channel (read in-basin routing). Moreover, mizuRoute has an in-channel impulse response function routing scheme based on diffusive wave along the river

segment. Note that four mizuRoute parameters (i.e., parameters 15-18 in **Table 2**) were also considered for perturbation in our experiment.

Table 2. Parameters of the VIC and mizuRoute routing model

Parameter [unit]	Range	Definition
b_{inf} [-]	[0.01,0.50]	The variable infiltration parameter
E_{exp} [-]	[3,12]	The slope of water retention curve
K_{sat} [mm/day]	[5,1000]	Saturated hydraulic conductivity
d_{1f} [m]	[0.05,0.20]	The depth of top soil layer for forested area
d_{1nf} [m]	[0.05,0.2]	The depth of top soil layer for non-forested area
d_{2f} [m]	[0.20,2.0]	The depth of the second soil layer for forested area
d_{2nf} [m]	[0.20,2.0]	The depth of the second soil layer for non-forested area
d_{1rf} [-]	[0.20,0.80]	The root distribution for top soil layer for forested area
d_{1rnf} [-]	[0.20,0.80]	The root distribution for top soil layer for non-forested area
d_{2rf} [-]	[0.20,0.80]	The root distribution for the second soil layer for forested area
d_{2rnf} [-]	[0.20,0.80]	The root distribution for the second soil layer for non-forested area
K_{slow} [1/day]	[0.001,0.90]	Slow reservoir coefficient for simplified baseflow representation
S_{LAI} [-]	[0.5-1.5]	Scale factor to scale the leaf area index (LAI)
$S_{stomatal}$ [-]	[0.5-1.5]	Scale factor to scale the stomal resistance
G_{shape} [-]	[0.5,5.0]	Shape factor for in-basin routing for Gamma function
G_{scale} [s]	[3600,172800]	Scale factor for in basin routing for Gamma function
V [m/s]	[0.5, 1.5]	Velocity of the diffusive wave
D_{diff} [m/s^2]	[500, 10000]	Diffusivity for the diffusive wave

3.4. Configuration of the computational experiments

3.4.1. Selecting output function

Residual-based summary statistics, such as Nash-Sutcliffe efficiency criterion (E_{NS}), are common objective functions in hydrology and have been widely used for sensitivity analysis to investigate which parameters significantly influence the model predictions. This provides valuable information for parameter estimation and model calibration methods. Therefore, we calculated the E_{NS} in our experiments and considered it as the response function for GSA, as follows:

$$E_{NS} = 1 - \frac{\sum_{t=1}^{T}\left({Q_t}^{obs} - {Q_t}^{sim}\right)^2}{\sum_{t=1}^{T}\left({Q_t}^{obs} - \bar{Q}\right)^2} \quad (15)$$

where T is the number of time steps, ${Q_t}^{obs}$ is the observed streamflow values at time t, ${Q_t}^{sim}$ is the simulated value of the streamflow at time t, and $\bar{Q}$ is the mean value of the observed streamflow values. Here, we compared the simulated flows with the observed streamflow data of the Water Survey Canada (gauge ID of 05BB001).

3.4.2. Generating synthetic datasets for VISCOUS runs

To evaluate the effectiveness of the proposed VISCOUS method and ensure a meaningful comparison, the GSA results obtained by the sampling-based algorithm were deemed to be the '*true*' sensitivities. Therefore, we first ran the conventional sampling-based GSA algorithm for both case studies using the method described in **Section 2.1.3**. For the first case study, the HBV model was run using a large sample size to ensure the stability and convergence of the sampling-based algorithm, resulting in 1,000,008 evaluations. The computational efficiency of the HBV model made it feasible to select this large sample size, and consequently enabled us to benchmark our results against the true sensitivities. Note that, in practice, the bootstrap method along with sensitivity of sensitivity analysis can be used to avoid a large number of simulations, when assessing the robustness and convergence of the GSA results in computationally expensive models. For the second case study, we focused on the GSA of the VIC model using a dataset consisting of 21,000 realizations of the input parameters and the corresponding E_{NS} values. A high-performance computing system operating parallelly on 250 cores was used to execute 21,000 realizations of the VIC model requiring about four days to complete the task.

In the next step, we produced synthetic input-output datasets by randomly sampling (with progressively increasing sample size) from the available model evaluations (1,000,008 and 21,000 input-output pairs of HBV and VIC, respectively). Each input-output sample set was considered to be a set of 'given data'. Then, we applied our VISCOUS method to all the given sample sets and compared the results with true sensitivities.

3.4.3. Accounting for randomness in sampling variability

Because of the randomness in our experiments, the performance of the GSA methods can be quite sensitive to the choice of the sample points, i.e., the results can be afflicted by the randomness in the selection of sample points drawn from the input-output space. This necessitates the use of several trials with different random seeds for each experiment to obtain a more reliable measure of the algorithm performance. Therefore, we replicated all experiments 50 times in this study, each replication using different random seed. This allowed us to see a range of possible performances for the proposed algorithm, and thus evaluate its robustness. Note that the progressive Latin hypercube sampling strategy (Sheikholeslami and Razavi, 2017) was used to randomly generate parameter sets.

4. Results and Discussion

4.1. Comparison of the proposed VISCOUS and original variance-based methods

4.1.1. Results for the main effects

Figure 3 and **4** show the main effect indices (S_i) for the 10-parameter HBV model and the 18-parameter VIC model estimated by the original variance-based and VISCOUS methods. For the original variance-based method, the sampling-based algorithm described in **Section 2.1.3** was applied; for the VISCOUS method, **Eq. (13)** was applied to the given set of input-output data. In **Figure 3**, calculation of the final/true sensitivity indices (the dashed horizontal lines) required about 1,000,000 model evaluations using the original variance-based method, while the proposed VISCOUS method has been applied by increasing the size of input-output data incrementally up to 50,000 (i.e., 1,000, 2,000, …, 50,000). The VISCOUS-based estimates of the main effect indices for VIC depicted in **Figure 4** were also obtained by randomly sampling with progressively increasing size (i.e., 1,000, 1,800, 1,600, …, 21,000) from the available pre-computed model evaluations. In addition, **Figure 3** and **4** show the 90% intervals of the estimated sensitivity indices obtained from 50 replicates of the experiments. Overall, these results suggest that both VISCOUS and the original variance-based methods yielded similar main effect indices across the parameter space, thereby confirming the accuracy of our proposed method in estimating the sensitivity indices for HBV and VIC.

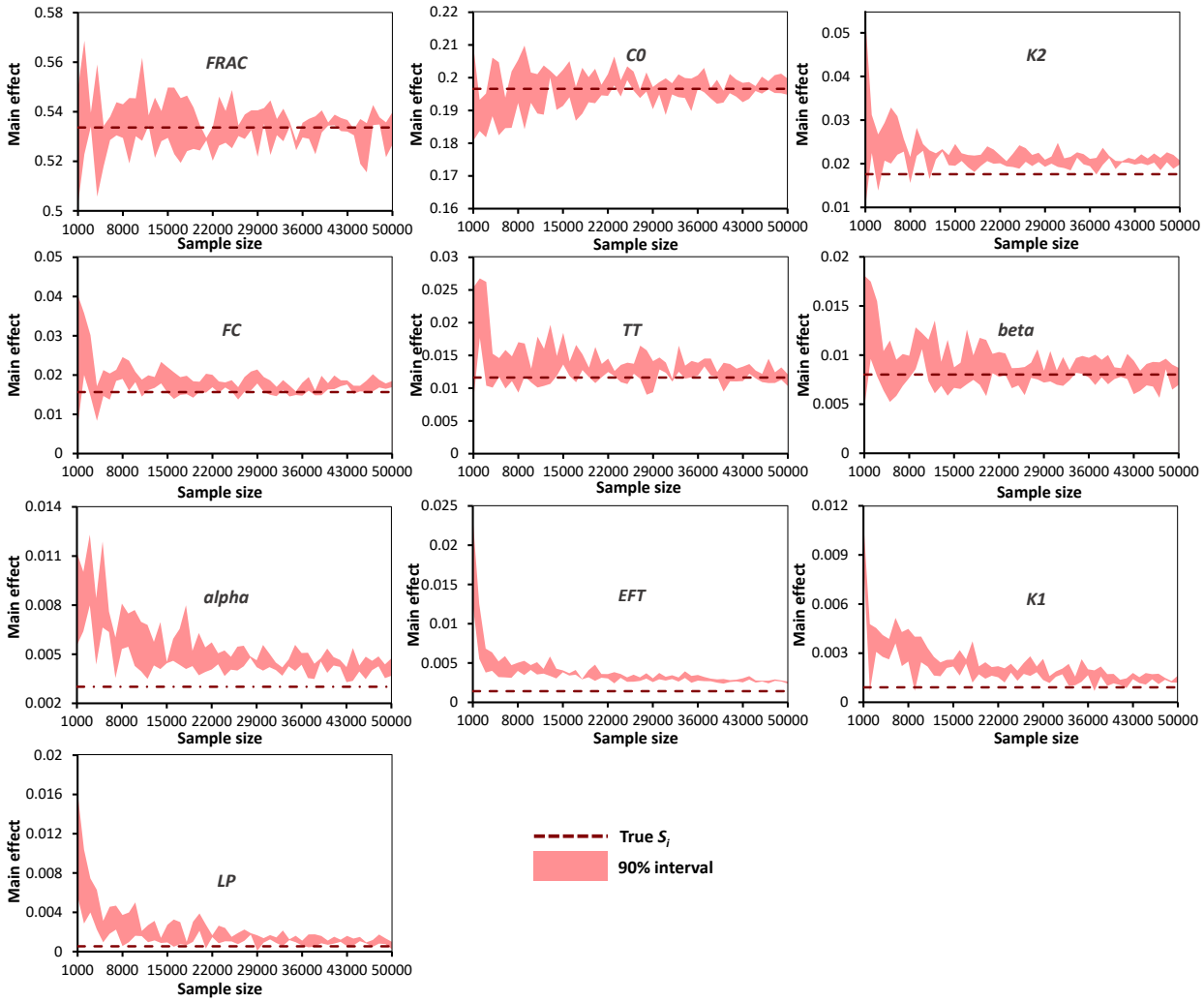

Figure 3. Convergence plots of the main effect (first-order) sensitivity indices for the HBV model. The main effect indices were estimated using the proposed VISCOUS method. The (light red) envelopes depict the 90% intervals of the 50 independent trials of this experiment. The dashed (dark red) horizontal lines show the true values of the sensitivity indices for the HBV parameters obtained by the conventional variance-based algorithm.

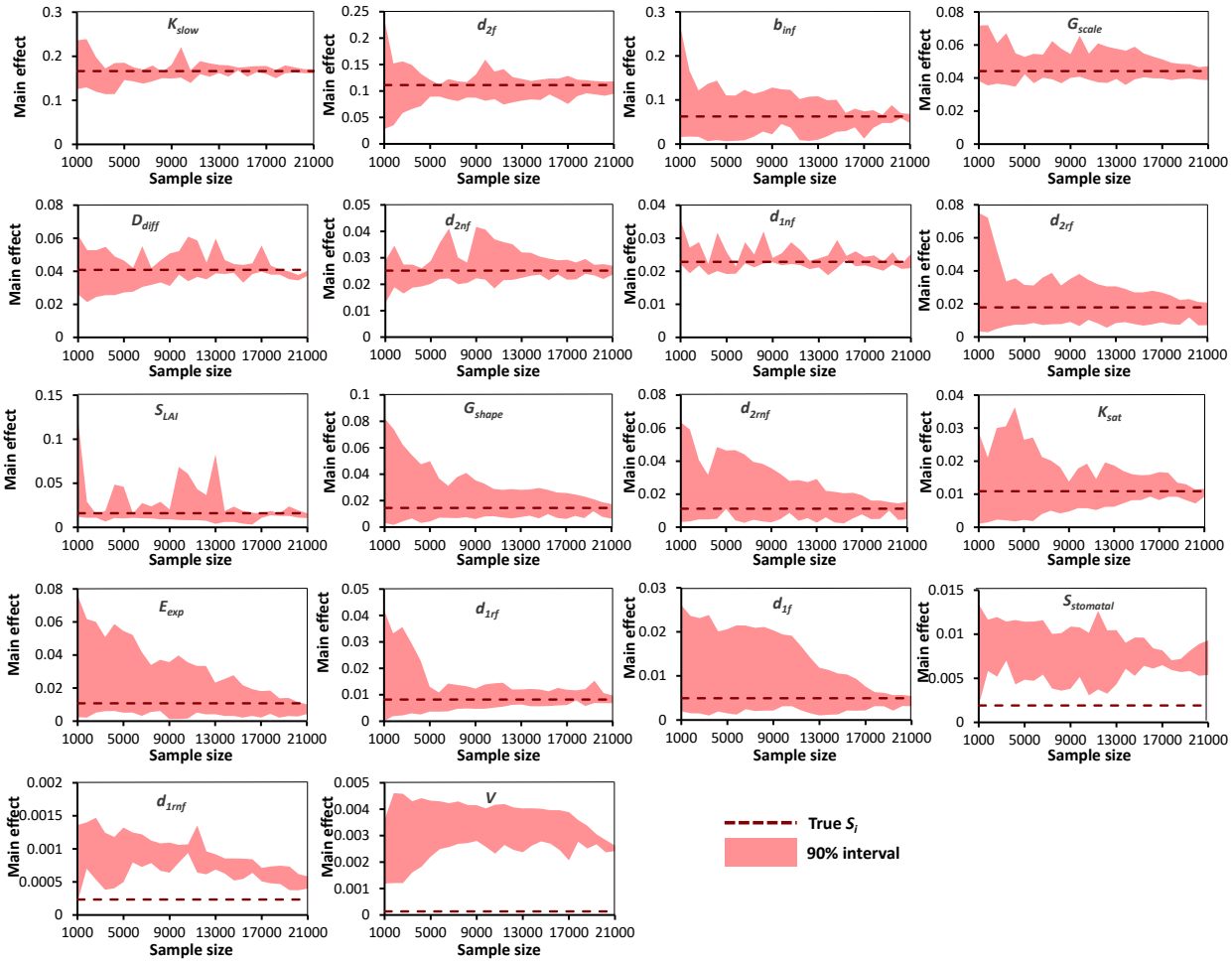

Figure 4. Convergence plots of the main effect (first-order) sensitivity indices for the VIC model. The main effect indices were estimated using the proposed VISCOUS method. The (light red) envelopes depict the 90% intervals of the 50 independent trials of this experiment. The dashed (dark red) horizontal lines show the true values of the sensitivity indices for the VIC parameters obtained by the conventional variance-based algorithm.

To assess the effect of sampling uncertainty and evaluate the robustness of VISCOUS to sampling variability, the 5th and 95th percentiles (90% intervals) of the 50 realizations of each experiment are presented in **Figure 3** and **4**. For HBV (see **Figure 3**), the results show that our proposed method converged to the true sensitivity indices within the 90% interval when the samples size was larger than about 8,000. In addition, by analyzing **Figure 3** we see that sensitivity indices associated with three important parameters of HBV (*FRAC*, *C0*, and *beta* in **Figure 3**) show a high variability compared to other parameters, when the sample size was small (less than 8,000). For the VIC model (see **Figure 4**), the 90% intervals indicate that most of the sensitivity indices reached a reasonable convergence when the sample size was larger than about 13,000. Note that width of the 90% intervals of the sensitivity indices in the case of VIC are quite narrow for the sample seizes greater than 17,000.

It is worth mentioning that three HBV parameters with lower sensitivity (*alpha*, *EFT*, and *K1* in **Figure 3**) have not converged to true indices (slightly overestimated) even when the sample size was 50,000. The main effect values associated with these less important parameters are very small (close to zero), and therefore overestimation of the sensitivity indices is mainly because of the numerical errors in the Monte Carlo integration step (**Eq**. (**13-14**). Like HBV, some of the less important parameters of VIC exhibit poor convergence behavior in **Figure 4** (i.e., d_{1rnf}, $S_{stomatal}$, and *V*). If the goal of conducting GSA is finding the accurate estimates of these sensitivity indices, a larger sample size should be used to achieve a higher degree of accuracy. However, when dealing with computationally expensive models, estimating accurate values of the sensitivity indices is often not of interest. In practice, one might focus on determining if a parameter belongs to a high, medium, or low-influence group rather than on its exact ranking, particularly taking into account that parameter ranking usually converges faster than sensitivity indices (Sarrazin et al., 2016; Sheikholeslami et al., 2019).

4.1.2. Results for the interaction effects

One straightforward approach to quantify the interaction effects in variance-based GSA is to calculate the sum of main effects (Borgonovo et al., 2017). The sum of main effects explains to what extent model parameters are individually important, while the remainder to one implies the degree to which the interaction among parameters is influential, i.e., the remainder $(1 - \sum_{i=1}^{m} S_i)$ is fraction of the output variance because of the interactions among all model parameters (see **Eq. (3)**). Based on our results for HBV, the sum of main effect indices is about 0.80 (80% of the variance), suggesting that interactions (expressed by total effect index, S_{Ti} in **Eq. (4)**) have very limited relevance for the this case study. Hence, the HBV response (i.e., E_{NS} values) can be approximately considered as an additive function of the parameters over their feasible ranges.

However, for the VIC model the sum of main effect indices is about 0.50 (50% of the variance), indicating that interactions play a significant role in the variability of the VIC model output. As a result, in this section, we opted for presenting the total effect results only for VIC. To investigate the interaction effects, we calculated total effect sensitivity indices, S_{Ti}, using the proposed VISCOUS algorithm. **Figure 5** compares the total effect and main effect indices of an arbitrarily chosen realization out of 50 independent realizations of this experiment for the VIC model. As can be seen, most of the parameters show a significant total effect compared to the main effect.

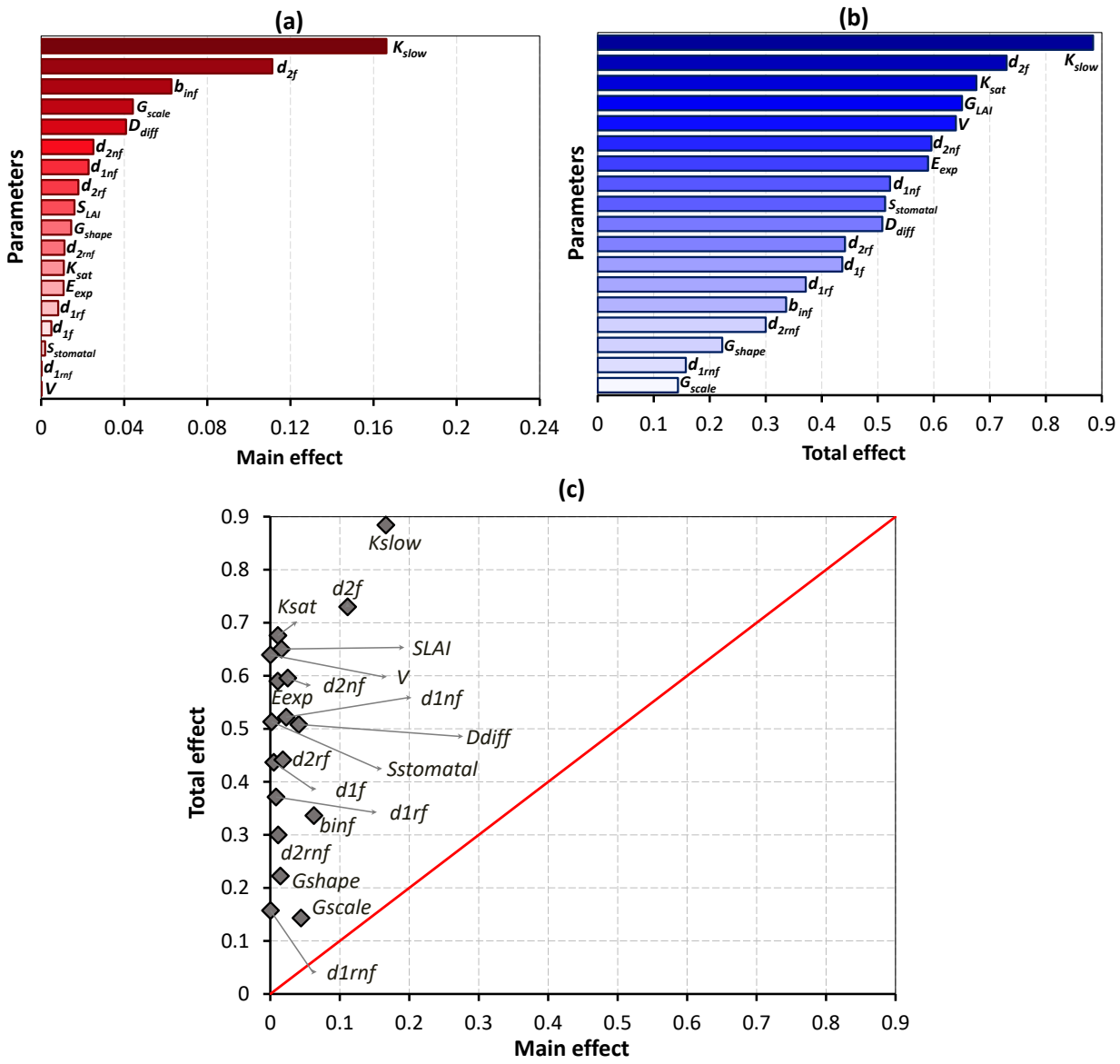

Figure 5. Parameter importance measured by **(a)** main effect and **(b)** total effect sensitivity indices estimated using the proposed VISCOUS method for the VIC model. Subplot **(c)** visualizes the relation between the main effect indices and corresponding total effect indices. Each diamond in subplot **(c)** represents one parameter of the model. Note that these results belong to an arbitrarily chosen realization out of 50 independent trials of the experiment.

Figure 5(a) and **(b)** highlights a clear difference between total effect and main effect indices. We closely examine this difference in **Figure 5(c)**. As shown, most of the parameters lie further toward the top-left corner of **Figure 5(c)**, confirming the strong dependencies/interactions between parameters of the VIC. This is an important sign of a possible non-identifiability issue in our model. In fact, it has been asserted that parameters associated with low main effects but having large total effects are typically less identifiable (see, e.g., Borgonovo et al. (2017); Hill et al. (2016); Ratto et al. (2007)). For example, Hill and Tiedeman (2007) reported the same observation

for their groundwater problem with dominant interaction effects. On the other hand, if parameters lie close to the ideal (1:1) line (red line in **Figure 5(c)**), they are likely identifiable. Note that having small main effects but high total effects does not necessarily indicate that parameters are non-identifiable (i.e., insensitivity is not a sufficient condition for non-identifiability).

Figure 5(c) may also be an indication of a badly defined calibration problem, which typically corresponds to a highly over-parameterized model. In such models, non-identifiability problem can be mainly attributed to the negligible influence of most of the parameters and their associated physical processes on the E_{NS} values (parameters with low sensitivity in **Figure 5(a)**). The non-identifiability can also occur due to the compensation among parameters, wherein changes in one parameter may be compensated for by changes in other parameters, as evident by strong interaction between VIC parameters in **Figure 5(b)**. Overall, our results follow the identifiability analysis of Gharari et al. (2020). For example, they have also reported that two soil parameters (K_{sat} and E_{exp}) were non-identifiable for the utilized configuration of VIC. As shown in **Figure 5(c)**, these parameters have very small main effects but high total effects, implying lack of identification.

4.2. High importance hydrological factors identified by the VISCOUS method

To characterize dominant processes affecting model responses, we implemented the recently developed grouping-based importance ranking approach of Sheikholeslami et al. (2019). This approach uses agglomerative hierarchical clustering to categorize the parameters into distinct groups based on similarities between their sensitivity indices, and then ranks parameters according to importance group rather than individually. **Figure 6** shows parameter grouping results of the HBV and the VIC models obtained from VISCOUS-based estimations of the main effect indices from 50 independent trials of the experiments.

As shown in **Figure 6**, we can categorize the HBV parameters into three importance groups according to their sensitivities:{*FRAC*, *C0*, *K2*} is the strongly influential group; {*FC*, *TT*, *beta*} is the moderately influential group; and {*alpha*, *ETF*, *K1*, *LP*} with the least influence on E_{NS}. For VIC, parameters can be also allocated into three groups as well: strongly influential parameters{K_{slow}, d_{2f}}; moderately influential parameters {b_{inf}, G_{scale}, D_{diff}}; and the rest of the parameters were ranked as least influential parameters.

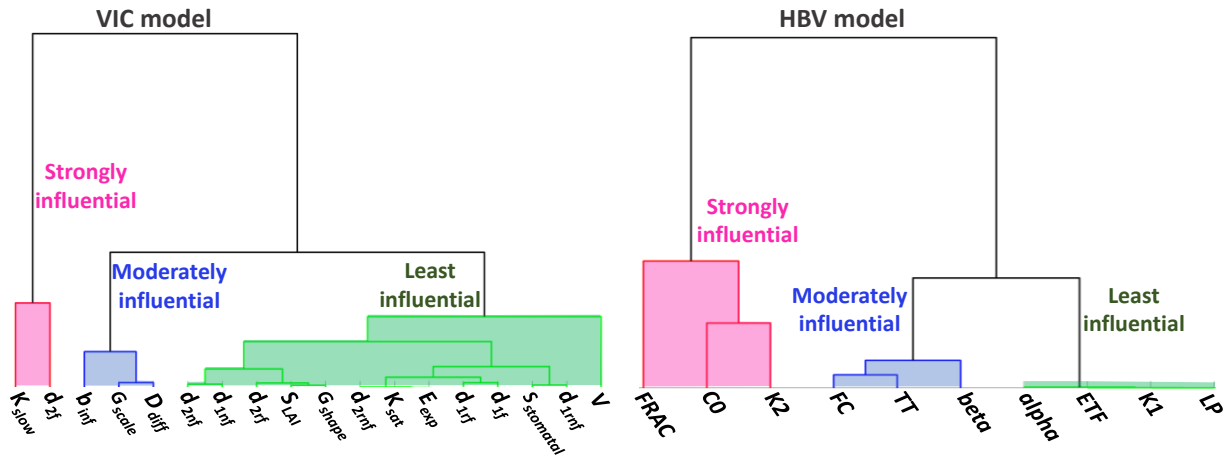

Figure 6. Parameter grouping for the 18-parameter VIC and the 10-parameter HBV models using the VISCOUS results. The parameters in the horizontal axis are sorted from the most important (to the left) to the unimportant (to the right) ones. The height of the vertical lines representant how similar/different parameters are from each other with respect to their influence on the model output. Note that this grouping is based on the main effect indices.

Factor importance ranking and grouping results in **Figure 6** point at dominant hydrological processes that govern the streamflow generation mechanisms in the Bow River basin. **Figure 6** reveals that slow reservoir coefficient in both models (K_{slow} in VIC and *K2* in HBV) is among the most important parameters influencing streamflow simulations. Note that we observe a strong yearly cycle for the Bow River due to the presence of snow and glaciers (see time series in **Figure 2**). In cold or mountainous regions, such as the Bow River where its headwaters are in the glaciated eastern slopes of the Canadian Rockies, snow melting and release of water from glacier storage greatly contribute to streamflow generation as opposed to rainfall. For example, Hopkinson and Young (1998) reported that the glacial contributions to streamflow in extremely dry years can be responsible for up to 50% of the late-summer flow in the Bow River at Banff. A significant contribution of these parameters to simulation of the streamflow can be due to the fact that K_{slow} in VIC and *K2* in HBV are responsible for the basin storage and timing of the release for getting the recession limb of the hydrograph correctly.

Soil moisture parameters in both models (b_{inf} in VIC and *beta* in HBV) also play a key role in streamflow simulation (see moderately influential parameters in **Figure 6**). These parameters have significant effect on the partitioning of the precipitation into infiltration and runoff. An increase in b_{inf} results in a lower infiltration capacity, and consequently yields higher runoff values simulated by VIC. For HBV, however, a higher *beta* reduces runoff from soil, thereby increasing evaporation and reducing flow. These findings are consistent with previous studies conducted by Gou et al. (2020); Melsen et al. (2016); Demaria et al. (2007).

Overall, based on our parameter grouping in **Figure 6**, it can be concluded that dominant factors of the HBV model mainly control the soil (*FRAC* and *FC*) and snowmelt (*C0*) processes. *FRAC* determines the fraction of soil release entering fast reservoir, as a result a higher *FRAC* value allows high flow of water from soil to the fast reservoir. *FC* accounts for partitioning of the precipitation into runoff and soil moisture. Hence, both *FRAC* and *FC* have a direct role on runoff generation in the basin. The high importance of *C0* can be justified because the hydrological regime of the Bow River is highly influenced by the snow accumulation and melt, and the amount of snowmelt is linearly proportional to temperature via *C0*.

In the VIC model, d_{2f} is one of the most important parameters. An explanation might be the fact that d_{2f} together with S_{LAI} are controlling the transpiration from plants. Also, notice the high total effect value associated with S_{LAI} in **Figure 5b**. The transpiration process in VIC is based on the Jarvis (1976)'s formulation, which is very sensitive to the leaf area index, S_{LAI}, assigned to each vegetation type. This parameter together with d_{2f} can significantly affect the simulated evaporation from the forested areas (dominant vegetation type), and therefore adjust the transpiration as one of the major water balance components to get a closer streamflow to observation. Lastly, from **Figure 6** we observe that two mizuRout routing parameters (G_{scale} and D_{diff}) are highly important

parameters of the model. These parameters are mainly responsible for peak time and flashiness of the hydrograph.

5. Conclusions and Future Research

Achieving robust and stable sensitivity analysis results while minimizing the computational cost has greatly increased the need for development of the computationally efficient algorithms for global sensitivity analysis (GSA). In particular, a high computational cost incurred by the sampling-based GSA methods limits their application to complex process-based hydrologic models (Clark et al., 2017). To tackle this problem, we developed an effective data-driven algorithm for variance-based sensitivity analysis using copulas (VISCOUS). In VISCOUS, there is no assumption on the underlying structure of the input-output data because distributions, correlations, and interactions between input factors and model output are learned only from available data. Motivations behind developing VISCOUS include:

- VISCOUS is a given-data approach. This means that it makes use of any collected datasets or pre-existing model runs, and accordingly prevents augmenting computational burden. Hence, VISCOUS enables modelers to efficiently conduct GSA for cases in which properties of the input-output distributions and of the underlying response surface are unknown and only a (sub-)sample of the input-output space is available.

- VISCOUS opens the possibility to incorporate information about the dependence structure between model parameters into the variance-based GSA. The dependencies can be specified by modelling their joint distribution. VISCOUS uses Gaussian mixture copula model (GMCM) to approximate the joint probability density function of the input-output sample for estimating several variance-based sensitivity indices.

We discussed numerical implementation and algorithmic details of our proposed given-data estimator, and examined its performance using two popular hydrologic models of increasing complexity (HBV and VIC). VISCOUS efficiently identified dominant parameters (and processes) of the VIC and HBV models that significantly influence prediction of the streamflow values in the Bow River Basin at Banff, Alberta, Canada. Overall, the numerical experiments indicate that:

- Parameter sensitivities obtained by VISCOUS are on par with the original sampling-based method of Sobol, while our method requires lower computational demand.

- VISCOUS is robust to sampling variability and randomness. The level of accuracy achieved by our proposed method is promising for practical purposes, such as for factor prioritization and factor ranking.

- Parameter identifiability of the physically-based distributed hydrologic models (VIC in present study) can be troublesome and efficient GSA methods such as VISCOUS can serve as a computationally tractable tool for diagnosing lack of identification.

- VISCOUS provides physically sound results regarding the sensitivity behavior of the parameters, even in cases where the sample size (or number of simulations) is limited. Our

sensitivity analysis results correspond well with the expected behavior and dynamics of runoff generation in cold regions.

Future studies should include an extension of the proposed method to multi-criteria sensitivity analysis of the hydrologic models. Our data-driven approach, in principle, can be used in the setting of multi-criteria sensitivity analysis to efficiently estimate the sensitivity indices for models with multiple dependent outputs. One possible option is to combine VISCOUS with multiple-response emulators of Liu et al. (2019), which integrates the copula functions with multiple-response Gaussian process emulator. In addition, since the GMCM used in VISCOUS has a set of design parameters, a "sensitivity analysis of sensitivity analysis" (SA of SA; Puy et al., 2020) should be tried in future studies to investigate VISCOUS' sensitivity to its own design parameters.

Furthermore, our proposed method has a great potential to be applied to GSA of the new generation of complex hydrological models, such as the SUMMA framework (Clark et al., 2015a,b). VISCOUS can facilitate fast and effective GSA for computationally demanding models and can help achieve accurate and robust results. Finally, testing our method on real-world datasets obtained from field observations, remote sensing, or laboratory experiments is an interesting avenue for further research. This can help gain an in-depth insight into the underlying system's behavior, thereby better supporting model development and understanding.

Data and Code Availability

The WRF simulations are available from Rasmussen and Liu (2017). The HBV and VIC models are available through Razavi et al. (2019) and Gharari et al. (2020), respectively. The MATLAB codes for the proposed VISCOUS method will be available on GitHub upon publication of the manuscript (https://github.com/Razi-Sheikh/VISCOUS).

Acknowledgement

This paper was written during the COVID-19 pandemic which has transformed the world, from personal adjustments to global shift. We would like to take this opportunity to acknowledge a deep sense of gratitude to the workers on the frontline against COVID-19 around the world – healthcare workers, including doctors, nurses, paramedics, and others. We pay tribute to their courage, dedication, and sacrifice and wish them rest and quiet from the storm. The authors would also like to thank editor D. Scott Mackay and associate editor Thorsten Wagener for handling this paper. We thank reviewers, Andrea Saltelli and two anonymous reviewers, for their constructive comments that helped us improve this paper.

Appendix A. Illustration Using an Analytical Test Function

We consider a type C benchmark problem for GSA introduced by Kucherenko et al. (2011), which has the following mathematical form:

$$Y(x_1, x_2, \dots, x_d) = \prod_{j=1}^{d} |4x_j - 2| \tag{A1}$$

where factors x_j $(j = 1,2, \dots, d)$ are uniformly distributed over a $[0,1]^d$ hypercube.

This function has dominant higher-order interaction terms with no unimportant subset of factors. Kucherenko et al. (2011) mathematically proved that for the function defined in **Eq. (A1)**:

$$\begin{cases} \dfrac{S_1}{S_{T1}} = \left(\dfrac{4}{3}\right)^{1-d} \\ \displaystyle\sum_{j=1}^{d} S_j = \dfrac{d}{3\left(\left(\dfrac{4}{3}\right)^d - 1\right)} \end{cases} \tag{A2}$$

where S_j and S_{Tj} are the first order and total order sensitivity indices for the j-th factor $(j = 1,2, \dots, d)$, respectively. The non-linearity and non-monotonicity of this function along with the availability of analytical results (**Eq. (A2)**) make it a suitable problem for the study of GSA methods. As can be seen from **Eq. (A1)**, this function is the product of contributions from each input factor, it is non-additive and features interactions of all orders.

In this study, for illustrative purposes, the number of factors was set to $d = 10$, and the true values of S_1/S_{T1} and $\sum S_j$ were used to evaluate the performance of our proposed approach. To implement VISCOUS, 100,000 points were randomly sampled from $[0,1]^{10}$ using Latin hypercube sampling strategy. In the next step, we produced synthetic input-output datasets by randomly sampling with progressively increasing sample size (i.e., 1,000, 2,000, …, 100,000) from the available 100,000 evaluations of the function. Each input-output sample set was considered to be a set of 'given data'. Then, we applied our VISCOUS method to all the given sample sets and compared the results with analytical values. In addition, we replicated this experiment 50 times, each replication using different random seed.

Figure A1 compares the VISCOUS-based estimates of the S_1/S_{T1} and $\sum S_j$ with their true values. Overall, these results suggest that VISCOUS results are in good agreement with theoretical values, thereby confirming the accuracy of our proposed method in estimating the sensitivity indices. Figure S1 also depicts the 5th and 95th percentiles (90% intervals) of the 50 realizations of each. Note that width of the 90% intervals rapidly become narrower when the sample size increases.

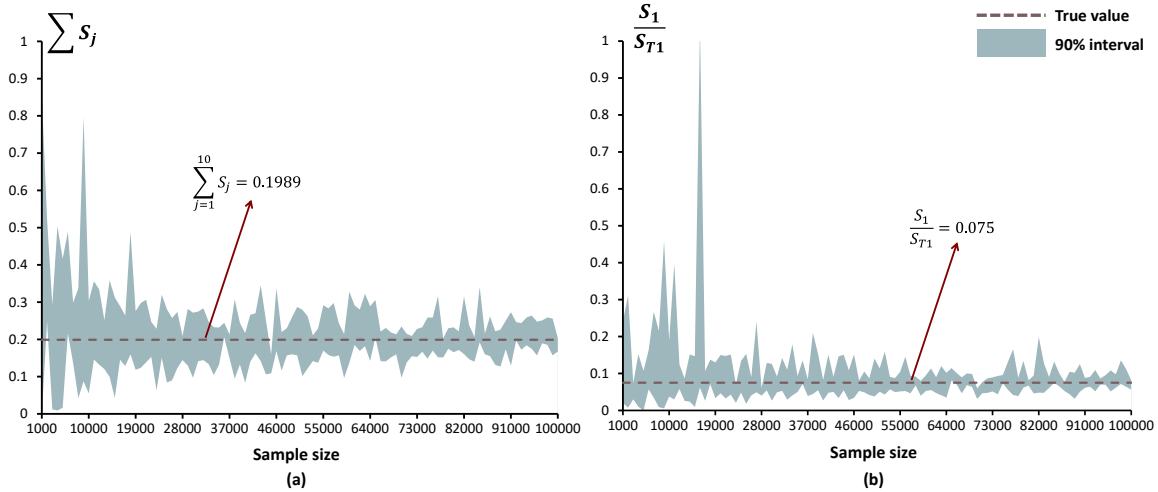

Figure A1. Convergence results for the type C function defined in Eq. (A1). The main and total effect indices were estimated using the proposed VISCOUS method. The (light blue) envelopes depict the 90% intervals of the 50 independent trials of this experiment. The dashed (dark red) horizontal lines show the true values of $\sum S_j$ (subplot a) and S_1/S_{T1} (subplot b).

References

Andres, T.H., 1997. Sampling methods and sensitivity analysis for large parameter sets. *Journal of Statistical Computation and Simulation*, *57*(1-4), 77-110.

Baroni, G., Schalge, B., Rakovec, O., Kumar, R., Schüler, L., Samaniego, L., Simmer, C. and Attinger, S., 2019. A Comprehensive Distributed Hydrological Modeling Intercomparison to Support Process Representation and Data Collection Strategies. *Water Resources Research*, *55*(2), 990-1010. https://doi.org/10.1029/2018WR023941

Becker, W.E., Tarantola, S. and Deman, G., 2018. Sensitivity analysis approaches to high-dimensional screening problems at low sample size. *Journal of Statistical Computation and Simulation*, *88*(11), 2089-2110. https://doi.org/10.1080/00949655.2018.1450876

Beckman, R.J. and McKay, M.D., 1987. Monte Carlo estimation under different distributions using the same simulation. *Technometrics*, *29*(2), 153-160. https://doi.org/10.1080/00401706.1987.10488206

Bergström, S., 1995. *The HBV model*, Water Resources Publications, Colorado, 443-476.

Bilgrau, A.E., Eriksen, P.S., Rasmussen, J.G., Johnsen, H.E., Dybkær, K. and Bøgsted, M., 2016. GMCM: Unsupervised clustering and meta-analysis using gaussian mixture copula models. *Journal of Statistical Software*, *70*(2), 1-23. https://doi.org/10.18637/jss.v070.i02

Borgonovo, E., Lu, X., Plischke, E., Rakovec, O. and Hill, M.C., 2017. Making the most out of a hydrological model data set: Sensitivity analyses to open the model black-box. *Water Resources Research*, *53*(9), 7933-7950. https://doi.org/10.1002/2017WR020767

Bounceur, N., Crucifix, M. and Wilkinson, R.D., 2015. Global sensitivity analysis of the climate–vegetation system to astronomical forcing: an emulator-based approach. *Earth System Dynamics*, *6*(1), 205-224. http://dx.doi.org/10.5194/esd-6-205-2015

Campolongo, F., Cariboni, J. and Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software*, *22*(10), 1509-1518. https://doi.org/10.1016/j.envsoft.2006.10.004

Castaings, W., Borgonovo, E., Morris, M.D. and Tarantola, S., 2012. Sampling strategies in density-based sensitivity analysis. *Environmental Modelling & Software*, *38*, 13-26. https://doi.org/10.1016/j.envsoft.2012.04.017

Chan, K., Saltelli, A. and Tarantola, S., 2000. Winding stairs: A sampling tool to compute sensitivity indices. *Statistics and Computing*, *10*(3), 187-196. https://doi.org/10.1023/A:1008950625967

Chen, W., Jin, R. and Sudjianto, A., 2005. Analytical variance-based global sensitivity analysis in simulation-based design under uncertainty. *Journal of Mechanical Design*, *127*(5), 875-886. https://doi.org/10.1115/1.1904642

Cheng, L., Lu, Z. and Zhang, L., 2015. Application of Rejection Sampling based methodology to variance based parametric sensitivity analysis. *Reliability Engineering & System Safety*, *142*, 9-18. https://doi.org/10.1016/j.ress.2015.04.020

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

Clark, M.P., Nijssen, B., Lundquist, J.D., Kavetski, D., Rupp, D.E., Woods, R.A., Freer, J.E., Gutmann, E.D., Wood, A.W., Brekke, L.D. and Arnold, J.R., 2015a. A unified approach for process-based hydrologic modeling: 1. Modeling concept. *Water Resources Research*, *51*(4), 2498-2514. https://doi.org/10.1002/2015WR017198

Clark, M.P., Nijssen, B., Lundquist, J.D., Kavetski, D., Rupp, D.E., Woods, R.A., Freer, J.E., Gutmann, E.D., Wood, A.W., Gochis, D.J. and Rasmussen, R.M., 2015b. A unified approach for process-based hydrologic modeling: 2. Model implementation and case studies. *Water Resources Research*, *51*(4), 2515-2542. https://doi.org/10.1002/2015WR017200

Clark, M.P., Slater, A.G., Rupp, D.E., Woods, R.A., Vrugt, J.A., Gupta, H.V., Wagener, T. and Hay, L.E., 2008. Framework for understanding structural errors (FUSE): A modular framework to diagnose differences between hydrological models. *Water Resources Research*, *44*(12). https://doi.org/10.1029/2007WR006735

Cosenza, A., Mannina, G., Vanrolleghem, P.A. and Neumann, M.B., 2013. Global sensitivity analysis in wastewater applications: A comprehensive comparison of different methods. *Environmental Modelling & Software*, *49*, 40-52. https://doi.org/10.1016/j.envsoft.2013.07.009

Crombecq, K., Laermans, E. and Dhaene, T., 2011. Efficient space-filling and non-collapsing sequential design strategies for simulation-based modeling. *European Journal of Operational Research*, *214*(3), 683-696. https://doi.org/10.1016/j.ejor.2011.05.032

Deheuvels, P. 1979. La fonction de dépendance empirique et ses propriétés. *Bulletins of the Royal Academy of Belgium*. *65*(5), 274-292.

Dell'Oca, A., Riva, M. and Guadagnini, A., 2020. Global sensitivity analysis for multiple interpretive models with uncertain parameters. *Water Resources Research*, *56*(2), e2019WR025754. https://doi.org/10.1029/2019WR025754

Demaria, E.M., Nijssen, B. and Wagener, T., 2007. Monte Carlo sensitivity analysis of land surface parameters using the Variable Infiltration Capacity model. *Journal of Geophysical Research*, *112*(D11113). https://doi.org/10.1029/2006JD007534

Duan, Q., Sorooshian, S. and Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. *Water resources research*, *28*(4), 1015-1031. https://doi.org/10.1029/91WR02985

Fan, Y.R., Huang, W.W., Huang, G.H., Li, Y.P., Huang, K. and Li, Z., 2016. Hydrologic risk analysis in the Yangtze River basin through coupling Gaussian mixtures into copulas. *Advances in Water Resources*, *88*, 170-185. https://doi.org/10.1016/j.advwatres.2015.12.017

Fatichi, S., Vivoni, E.R., Ogden, F.L., Ivanov, V.Y., Mirus, B., Gochis, D., Downer, C.W., Camporese, M., Davison, J.H., Ebel, B. and Jones, N., 2016. An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. *Journal of Hydrology*, *537*, 45-60. https://doi.org/10.1016/j.jhydrol.2016.03.026

Ferretti, F., Saltelli A. and Tarantola, S., 2016, Trends in sensitivity analysis practice in the last decade. *Science of the Total Environment*, *568*, 666-670. https://doi.org/10.1016/j.scitotenv.2016.02.133

Gan, Y., Duan, Q., Gong, W., Tong, C., Sun, Y., Chu, W., Ye, A., Miao, C. and Di, Z., 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

Genest, C., Nešlehová, J.G. and Rémillard, B., 2017. Asymptotic behavior of the empirical multilinear copula process under broad conditions. *Journal of Multivariate Analysis*, *159*, 82-110. https://doi.org/10.1016/j.jmva.2017.04.002

Gharari, S., Clark, M. P., Mizukami, N., Knoben, W. J. M., Wong, J. S., and Pietroniro, A., 2020. Flexible vector-based spatial configurations in land models, *Hydrology and Earth System Sciences, Discussion*. https://doi.org/10.5194/hess-2020-111 (in review)

Gong, W., Duan, Q.Y., Li, J.D., Wang, C., Di, Z.H., Ye, A.Z., Miao, C.Y. and Dai, Y.J., 2016. An intercomparison of sampling methods for uncertainty quantification of environmental dynamic models. *Journal of Environmental Informatics*, *28*(1), 11-24. https://doi.org/10.3808/jei.201500310

Gou, J., Miao, C., Duan, Q., Tang, Q., Di, Z., Liao, W., Wu, J. and Zhou, R., 2019. Sensitivity analysis-based automatic parameter calibration of the variable infiltration capacity (VIC) model for streamflow simulations over China. *Water Resources Research*, *56*, e2019WR025968. https://doi.org/10.1029/2019WR025968

Gupta, H.V., Wagener, T. and Liu, Y., 2008. Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *Hydrological Processes*, *22*(18), 3802-3813. https://doi.org/10.1002/hyp.6989

Guse, B., Pfannerstill, M., Gafurov, A., Fohrer, N., and Gupta, H. 2016. Demasking the integrated information of discharge: Advancing sensitivity analysis to consider different hydrological components and their rates of change, *Water Resources Research, 52*, 8724-8743, https://doi.org/10.1002/2016WR018894

Hall, J., 2007. Probabilistic climate scenarios may misrepresent uncertainty and lead to bad adaptation decisions. *Hydrological Processes,* 21(8), 1127-1129. https://doi.org/10.1002/hyp.6573

Hill, M.C., Kavetski, D., Clark, M., Ye, M., Arabi, M., Lu, D., Foglia, L. and Mehl, S., 2016. Practical use of computationally frugal model analysis methods. *Groundwater*, *54*(2), 159-170. https://doi.org/10.1111/gwat.12330

Hoeffding W. 1992. A class of statistics with asymptotically normal distribution. In: Kotz S., Johnson N.L. (eds) *Breakthroughs in Statistics*. Springer Series in Statistics (Perspectives in Statistics). Springer, New York, NY. https://doi.org/10.1007/978-1-4612-0919-5_20Homma, T. and Saltelli, A., 1996. Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety*, *52*(1), 1-17. https://doi.org/10.1016/0951-8320(96)00002-6

Hopkinson, C. and Young, G.J., 1998. The effect of glacier wastage on the flow of the Bow River at Banff, Alberta, 1951-1993. *Hydrological Processes*, 12(10-11), pp.1745-1762. https://doi.org/10.1002/(SICI)1099-1085(199808/09)12:10/11<1745::AID-HYP692>3.0.CO;2-S

Huo, X., Gupta, H., Niu, G.Y., Gong, W. and Duan, Q., 2019. Parameter sensitivity analysis for computationally intensive spatially distributed dynamical environmental systems models. *Journal of Advances in Modeling Earth Systems*, *11*(9), 2896-2909. https://doi.org/10.1029/2018MS001573

Iooss, B., Van Dorpe, F. and Devictor, N., 2006. Response surfaces and sensitivity analyses for an environmental model of dose calculations. *Reliability Engineering & System Safety*, *91*(10-11), 1241-1251. https://doi.org/10.1016/j.ress.2005.11.021

Janouchová, E. and Kučerová, A., 2013. Competitive comparison of optimal designs of experiments for sampling-based sensitivity analysis. *Computers & Structures*, *124*, 47-60. https://doi.org/10.1016/j.compstruc.2013.04.009

Jansen, M.J., 1999. Analysis of variance designs for model output. *Computer Physics Communications*, *117*(1-2), 35-43. https://doi.org/10.1016/S0010-4655(98)00154-4

Jarvis, P.G., 1976. The interpretation of the variations in leaf water potential and stomatal conductance found in canopies in the field. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, *273*(927), 593-610. https://doi.org/10.1098/rstb.1976.0035

Jia, G. and Taflanidis, A.A., 2016. Efficient evaluation of Sobol' indices utilizing samples from an auxiliary probability density function. *Journal of Engineering Mechanics*, *142*(5), p.04016012. https://doi.org/10.1061/(ASCE)EM.1943-7889.0001061

Jourdan, A., 2012. Global sensitivity analysis using complex linear models. *Statistics and Computing*, *22*(3), 823-831. https://doi.org/10.1007/s11222-011-9239-y

Kasa, S.R., Bhattacharya, S. and Rajan, V., 2020. Gaussian mixture copulas for high-dimensional clustering and dependency-based subtyping. *Bioinformatics*, *36*(2), 621-628. https://doi.org/10.1093/bioinformatics/btz599

Kavetski, D. and Clark, M.P., 2010. Ancient numerical daemons of conceptual hydrological modeling: 2. Impact of time stepping schemes on model analysis and prediction. *Water Resources Research*, *46*(10). https://doi.org/10.1029/2009WR008896

Kucherenko, S. and Song, S., 2017. Different numerical estimators for main effect global sensitivity indices. *Reliability Engineering & System Safety*, *165*, 222-238. https://doi.org/10.1016/j.ress.2017.04.003

Kucherenko, S., Feil, B., Shah, N. and Mauntz, W., 2011. The identification of model effective dimensions using global sensitivity analysis. *Reliability Engineering & System Safety*, 96, 440-449. https://doi.org/10.1016/j.ress.2010.11.003

Li, C. and Mahadevan, S., 2016. An efficient modularized sample-based method to estimate the first-order Sobol ′index. *Reliability Engineering & System Safety*, *153*, 110-121. https://doi.org/10.1016/j.ress.2016.04.012

Li, Q., Brown, J.B., Huang, H. and Bickel, P.J., 2011. Measuring reproducibility of high-throughput experiments. *The annals of applied statistics*, *5*(3), 1752-1779. https://doi.org/10.1214/11-AOAS466

Liang, X., Lettenmaier, D.P., Wood, E.F. and Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research: Atmospheres*, *99*(D7), 14415-14428. https://doi.org/10.1029/94JD00483

Liu, F., Wei, P., Tang, C., Wang, P. and Yue, Z., 2019. Global sensitivity analysis for multivariate outputs based on multiple response Gaussian process model. *Reliability Engineering and System Safety*, *189*(C), 287-298. https://doi.org/10.1016/j.ress.2019.04.039

Piano, S.L., Ferretti, F., Puy, A., Albrecht, D. and Saltelli, A., 2021. Variance-based sensitivity analysis: The quest for better estimators and designs between explorativity and economy. *Reliability Engineering & System Safety*, *206*. https://doi.org/10.1016/j.ress.2020.107300

Mai, J. and Tolson, B.A., 2019. Model variable augmentation (MVA) for diagnostic assessment of sensitivity analysis results. *Water Resources Research*, *55*(4), 2631-2651. https://doi.org/10.1029/2018WR023382

Maier, H.R., Guillaume, J.H., van Delden, H., Riddell, G.A., Haasnoot, M. and Kwakkel, J.H., 2016. An uncertain future, deep uncertainty, scenarios, robustness and adaptation: How do they fit together?. *Environmental Modelling & Software*, *81*, 154-164. https://doi.org/10.1016/j.envsoft.2016.03.014

Markstrom, S.L., Hay, L.E. and Clark, M.P., 2016. Towards simplification of hydrologic modeling: identification of dominant processes. *Hydrology & Earth System Sciences*, *20*(11). https://doi.org/10.5194/hess-20-4655-2016

Marseguerra, M., Masini, R., Zio, E. and Cojazzi, G., 2003. Variance decomposition-based sensitivity analysis via neural networks. *Reliability Engineering & System Safety*, *79*(2), 229-238. https://doi.org/10.1016/S0951-8320(02)00234-X

McLachlan, G., and Peel. D., 2000. *Finite Mixture Models*. Hoboken, NJ: John Wiley & Sons, Inc.

Melsen, L., Teuling, A., Torfs, P., Zappa, M., Mizukami, N., Clark, M. and Uijlenhoet, R., 2016. Representation of spatial and temporal variability in large-domain hydrological models: case study for a mesoscale pre-Alpine basin. *Hydrology and Earth System Sciences*, 20(6), 2207-2226. https://doi.org/10.5194/hess-20-2207-2016

Mizukami, N., Clark, M.P., Sampson, K., Nijssen, B., Mao, Y., McMillan, H., Viger, R.J., Markstrom, S.L., Hay, L.E., Woods, R. and Arnold, J.R., 2016. mizuRoute version 1: a river network routing tool for a continental domain water resources applications. *Geoscientific Model Development*, *9*(6), 2223-2238. https://doi.org/10.5194/gmd-9-2223-2016

Nossent, J., Elsen, P. and Bauwens, W., 2011. Sobol'sensitivity analysis of a complex environmental model. *Environmental Modelling & Software*, *26*(12), 1515-1525. https://doi.org/10.1016/j.envsoft.2011.08.010

O'Hagan, A., 2006. Bayesian analysis of computer code outputs: A tutorial. *Reliability Engineering & System Safety*, *91*(10-11), 1290-1300. https://doi.org/10.1016/j.ress.2005.11.025

Plischke, E. and Borgonovo, E., 2019. Copula theory and probabilistic sensitivity analysis: Is there a connection?. *European Journal of Operational Research*, *277*(3), 1046-1059. https://doi.org/10.1016/j.ejor.2019.03.034

Plischke, E., Borgonovo, E. and Smith, C.L., 2013. Global sensitivity measures from given data. *European Journal of Operational Research*, *226*(3), 536-550. https://doi.org/10.1016/j.ejor.2012.11.047

Plischke, E., 2010. An effective algorithm for computing global sensitivity indices (EASI). *Reliability Engineering & System Safety*, *95*(4), 354-360. https://doi.org/10.1016/j.ress.2009.11.005

Pronzato, L. and Müller, W.G., 2012. Design of computer experiments: space filling and beyond. *Statistics and Computing*, *22*(3), 681-701. https://doi.org/10.1007/s11222-011-9242-3

Quinn, J.D., Reed, P.M., Giuliani, M. and Castelletti, A., 2019. What Is Controlling Our Control Rules? Opening the Black Box of Multireservoir Operating Policies Using Time-Varying Sensitivity Analysis. *Water Resources Research*, *55*(7), 5962-5984. https://doi.org/10.1029/2018WR024177

Rakovec, O., Hill, M.C., Clark, M.P., Weerts, A.H., Teuling, A.J. and Uijlenhoet, R., 2014. Distributed evaluation of local sensitivity analysis (DELSA), with application to hydrologic models. *Water Resources Research*, *50*(1), 409-426. https://doi.org/10.1002/2013WR014063

Rasmussen, R., and Liu, C., 2017. High resolution WRF simulations of the current and future climate of North America. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/D6V40SXP

Ratto, M., P. C. Young, R. Romanowicz, F. Pappenberger, A. Saltelli, and Pagano A., 2007. Uncertainty, sensitivity analysis and the role of data based mechanistic modeling in hydrology. *Hydrology and Earth System Sciences*, *11*(4), 1249–1266. https://doi.org/10.5194/hess-11-1249-2007

Ratto, M., Castelletti, A. and Pagano, A., 2012. Emulation techniques for the reduction and sensitivity analysis of complex environmental models. *Environmental Modelling & Software*, *34*, 1-4 https://doi.org/10.1016/j.envsoft.2011.11.003

Razavi, S. and Gupta, H.V., 2015. What do we mean by sensitivity analysis? The need for comprehensive characterization of "global" sensitivity in E arth and E nvironmental systems models. *Water Resources Research*, 51(5), 3070-3092. https://doi.org/10.1002/2014WR016527

Razavi, S. and Gupta, H.V., 2016. A new framework for comprehensive, robust, and efficient global sensitivity analysis: 2. Application. *Water Resources Research*, *52*(1), 440-455. https://doi.org/10.1002/2015WR017559

Razavi, S., Sheikholeslami, R., Gupta, H.V. and Haghnegahdar, A., 2019. VARS-TOOL: A toolbox for comprehensive, efficient, and robust sensitivity and uncertainty analysis. *Environmental modelling & software*, *112*, 95-107. https://doi.org/10.1016/j.envsoft.2018.10.005

Refsgaard, J.C., van der Sluijs, J.P., Højberg, A.L. and Vanrolleghem, P.A., 2007. Uncertainty in the environmental modelling process–a framework and guidance. *Environmental modelling & software*, *22*(11), 1543-1556. https://doi.org/10.1016/j.envsoft.2007.02.004

Ruano, M.V., Ribes, J., Seco, A. and Ferrer, J., 2012. An improved sampling strategy based on trajectory design for application of the Morris method to systems with many input factors. *Environmental Modelling & Software*, *37*, 103-109. https://doi.org/10.1016/j.envsoft.2012.03.008

Sainte-Marie, J. and Cournède, P.H., 2019. Insights of global sensitivity analysis in biological models with dependent parameters. *Journal of Agricultural, Biological and Environmental Statistics*, *24*(1), 92-111. https://doi.org/10.1007/s13253-018-00343-1

Saltelli, A., 2002. Making best use of model evaluations to compute sensitivity indices. *Computer physics communications*, *145*(2), 280-297. https://doi.org/10.1016/S0010-4655(02)00280-1

Saltelli, A., Bammer, G., Bruno, I., Charters, E., Di Fiore, M., Didier, E., Espeland, W.N., Kay, J., Piano, S.L., Mayo, D. and Pielke Jr. R., 2020. Five ways to ensure that models serve society: a manifesto. *Nature*, *582*, 482-484. https://doi.org/10.1038/d41586-020-01812-9

Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S. and Wu, Q., 2019. Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling & Software*, *114*, 29-39. https://doi.org/10.1016/j.envsoft.2019.01.012

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M. and 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. and Tarantola, S., 2008. *Global sensitivity analysis: the primer*. John Wiley & Sons.

Saltelli A., Andres T. H. and Homma T., 1993. Some new techniques in sensitivity analysis of model output. *Computational Statistics and Data Analysis*, *15*, 211-238. https://doi.org/10.1016/0167-9473(93)90193-W

Sarrazin, F., Pianosi, F. and Wagener, T., 2016. Global Sensitivity Analysis of environmental models: convergence and validation. *Environmental Modelling & Software*, *79*, 135-152. https://doi.org/10.1016/j.envsoft.2016.02.005

Seibert, J., 1997. Estimation of parameter uncertainty in the HBV model. *Hydrology Research*, *28*(4-5), 247-262. https://doi.org/10.2166/nh.1998.15

Sheikholeslami, R. and Razavi, S., 2017. Progressive Latin Hypercube Sampling: An efficient approach for robust sampling-based analysis of environmental models. *Environmental Modelling & Software*, *93*, 109-126. https://doi.org/10.1016/j.envsoft.2017.03.010

Sheikholeslami, R., Razavi, S., Gupta, H.V., Becker, W. and Haghnegahdar, A., 2019. Global sensitivity analysis for high-dimensional problems: How to objectively group factors and measure robustness and convergence while reducing computational cost. *Environmental Modelling & Software*, *111*, 282-299. https://doi.org/10.1016/j.envsoft.2018.09.002

Sheikholeslami, R., Razavi, S. and Haghnegahdar, A., 2019. What should we do when a model crashes? recommendations for global sensitivity analysis of Earth and environmental systems models. *Geoscientific Model Development*, *12*(10), 4275-4296. https://doi.org/10.5194/gmd-12-4275-2019

Sheikholeslami, R. and Razavi, S., 2020. A fresh look at variography: measuring dependence and possible sensitivities across geophysical systems from any given data. *Geophysical Research Letters*, *47*, e2020GL089829. https://doi.org/10.1029/2020GL089829

Sklar, A. (1959), "Fonctions de répartition à n dimensions et leurs marges", Publ. Inst. Statist. Univ. Paris, 8: 229–231.

Sobol, I.M., 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and computers in simulation*, *55*(1-3), 271-280. https://doi.org/10.1016/S0378-4754(00)00270-6

Sparkman, D., Millwater, H.R., Garza, J. and Smarslok, B.P., 2016. Importance sampling-based post-processing method for global sensitivity analysis. In *18th AIAA Non-Deterministic Approaches Conference* (p. 1440). https://doi.org/10.2514/6.2016-1440

Steele, R.J. and Raftery, A.E., 2010. Performance of Bayesian model selection criteria for Gaussian mixture models, In *Frontiers of Statistical Decision Making and Bayesian Analysis*, 2, New York: Springer. 113-130.

Storlie, C.B., Swiler, L.P., Helton, J.C. and Sallaberry, C.J., 2009. Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. *Reliability Engineering & System Safety*, *94*(11), 1735-1763. https://doi.org/10.1016/j.ress.2009.05.007

Strong, M. and Oakley, J.E., 2013. An efficient method for computing single-parameter partial expected value of perfect information. *Medical Decision Making*, *33*(6), 755-766. https://doi.org/10.1177/0272989X12465123

Strong, M., Oakley, J.E. and Chilcott, J., 2012. Managing structural uncertainty in health economic decision models: a discrepancy approach. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *61*(1), 25-45. https://doi.org/10.1111/j.1467-9876.2011.01014.x

Sudret, B., 2008. Global sensitivity analysis using polynomial chaos expansions. *Reliability engineering & system safety*, *93*(7), 964-979. https://doi.org/10.1016/j.ress.2007.04.002

Tarantola, S., Gatelli, D. and Mara, T. A., 2006. Random balance designs for the estimation of first order global sensitivity indices. *Reliability Engineering & System Safety*, *91*(6), 717-27. https://doi.org/10.1016/j.ress.2005.06.003

Ţene, M., Stuparu, D.E., Kurowicka, D. and El Serafy, G.Y., 2018. A copula-based sensitivity analysis method and its application to a North Sea sediment transport model. *Environmental modelling & software*, *104*, 1-12. https://doi.org/10.1016/j.envsoft.2018.03.002

Tewari, A., Giering, M.J. and Raghunathan, A., 2011, December. Parametric characterization of multimodal distributions with non-Gaussian modes. In *2011 IEEE 11th International Conference on Data Mining Workshops* (286-292). IEEE. https://doi.org/10.1109/ICDMW.2011.135

Verrelst, J., Sabater, N., Rivera, J.P., Muñoz-Marí, J., Vicent, J., Camps-Valls, G. and Moreno, J., 2016. Emulation of leaf, canopy and atmosphere radiative transfer models for fast global sensitivity analysis. *Remote Sensing*, *8*(8), p.673. https://doi.org/10.3390/rs8080673

Wainwright, H.M., Finsterle, S., Jung, Y., Zhou, Q. and Birkholzer, J.T., 2014. Making sense of global sensitivity analyses. *Computers & Geosciences*, *65*, 84-94. https://doi.org/10.1016/j.cageo.2013.06.006

Wei, P., Lu, Z. and Song, J., 2014. Moment-Independent Sensitivity Analysis Using Copula. *Risk Analysis*, *34*(2), 210-222. https://doi.org/10.1111/risa.12110

Wei, P., Lu, Z. and Song, J., 2015. Regional and parametric sensitivity analysis of Sobol ′indices. *Reliability Engineering & System Safety*, *137*, 87-100. https://doi.org/10.1016/j.ress.2014.12.012

Wu, Z., Wang, D., Wang, W., Zhao, K., Zhou, H. and Zhang, W., 2020. Hybrid metamodel of radial basis function and polynomial chaos expansions with orthogonal constraints for global sensitivity analysis. *Structural and Multidisciplinary Optimization*, 1-21. https://doi.org/10.1007/s00158-020-02516-4

Young, P.C. and Ratto, M., 2009. A unified approach to environmental systems modeling. *Stochastic Environmental Research and Risk Assessment*, *23*(7), 1037-1057. https://doi.org/10.1007/s00477-008-0271-1

Young, P., 1999. Data-based mechanistic modelling, generalised sensitivity and dominant mode analysis. *Computer Physics Communications*, *117*(1-2), 113-129. https://doi.org/10.1016/S0010-4655(98)00168-4