Moment-based Metrics for Global Sensitivity Analysis of Hydrological Systems

A. Dell'Oca[1], M. Riva[1,2], A. Guadagnini[1,2]

[1]Dipartimento di Ingegneria Civile e Ambientale (DICA), Politecnico di Milano, Piazza L. Da Vinci, 32, 20133 Milano, Italy

[2]Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, Arizona, USA

Abstract

We propose new metrics to assist global sensitivity analysis, GSA, of hydrological and Earth systems. Our approach allows assessing the impact of uncertain parameters on main features of the probability density function, *pdf*, of a target model output, *y*. These include the expected value of *y*, the spread around the mean and the degree of symmetry and tailedness of the *pdf* of *y*. Since reliable assessment of higher order statistical moments can be computationally demanding, we couple our GSA approach with a surrogate model, approximating the full model response at a reduced computational cost. Here, we consider the generalized Polynomial Chaos Expansion (gPCE), other model reduction techniques being fully compatible with our theoretical framework. We demonstrate our approach through three test cases, including an analytical benchmark, a simplified scenario mimicking pumping in a coastal aquifer, and a laboratory-scale conservative transport experiment. Our results allow ascertaining which parameters can impact some moments of the model output *pdf* while being uninfluential to others. We also investigate the error associated with the evaluation of our sensitivity metrics by replacing the original system model through a gPCE. Our results indicate that the construction of a surrogate model with increasing level of accuracy might be required depending on the statistical moment considered in the GSA. Our approach is fully compatible with (and can assist the development of) analysis techniques employed in the context of reduction of model complexity, model calibration, design of experiment, uncertainty quantification and risk assessment.

1. Introduction

Our improved understanding of physical-chemical mechanisms governing hydrological processes at multiple space and time scales and the ever increasing power of modern computational resources are at the heart of the formulation of conceptual models which are frequently characterized by marked levels of sophistication and complexity. This is evident when one considers the spectrum of mathematical formulations and ensuing level of model parametrization rendering our conceptual understanding of given environmental scenarios (Willmann et al., 2006; Grauso et al., 2007; Koutsoyiannis, 2010; Wagener et al., 2010; Elshorbagy et al., 2010a,b; Wagener and Montanari, 2011; Hartmann et al., 2013; Herman et al., 2013; Förster et al., 2014; Paniconi and Putti, 2015). Model complexity can in turn exacerbate challenges associated with the need to quantify the way uncertainties associated with parameters of a given model propagate to target state variables.

In this context, approaches based on rigorous sensitivity analysis are valuable tools to improve our ability to (*i*) quantify uncertainty, (*ii*) enhance our understanding of the relationships between model input and outputs, and (*iii*) tackle the challenges of model- and data- driven design of experiments. These also offer insights to guide model simplification, e.g., by identifying model input parameters that have negligible effects on a target output. The variety of available sensitivity methodologies can be roughly subdivided into two broad categories, i.e., local and global approaches. Local sensitivity analyses consider the variation of a model output against variations of model input solely in the neighbourhood of a given set of parameters values. Otherwise, global sensitivity analysis (GSA) quantifies model sensitivity across the complete support within which model parameters can vary. Error measurements and/or lack of knowledge about parameters can be naturally accommodated in a GSA by specifying appropriate parameter intervals and evaluating sensitivity over the complete parameter space. Recent studies and reviews on available sensitivity analysis and approaches are offered by, e.g., Pianosi et al. (2016), Sarrazin et al. (2016), and Razavi and Gupta (2015).

Our study is framed in the context of GSA methods. A broadly recognized strategy to quantify global sensitivity of uncertain model parameters to model outputs relies on the evaluation of the

Sobol' indices (Sobol, 1993). These are typically referred to as variance-based sensitivity measures because the output variance is taken as the metric upon which sensitivity is quantified. A key limitation of a variance-based GSA is that the uncertainty of the output is implicitly considered to be fully characterized by its variance. Relying solely on this criterion can provide an incomplete picture of a system response to model parameters, also considering that probability densities of typical hydrological quantities can be characterized by highly skewed and tailed distributions (e.g., Borgonovo et al., 2011). Recent studies (e.g., Krykacz-Hausmann, 2001; Borgonovo, 2007; Borgonovo et al., 2011) introduced a sensitivity metric grounded on the complete probability density function, *pdf,* of the model output. These so-called moment-independent analyses may suffer from operational constraints, because a robust evaluation of the complete *pdf* may require a number of model runs which is computationally unaffordable. The PAWN method developed by Pianosi and Wagener (2015) attempts to overcome this limitation introducing a sensitivity metric based on the cumulative density function, which can potentially be estimated more robustly than its associated *pdf* for a given sample size.

It is clear that while a variance-based GSA can be favored for its conceptual simplicity and ease of implementation and variance can be considered in some cases as an adequate proxy of the spread around the mean, it does not yield a forthright quantification of the way variations of a parameter can affect the structure of the *pdf* of a target model output. Otherwise, moment-independent methodologies condense sensitivity of the entire *pdf* in only one index, somehow shadowing our understanding of how the structure of the *pdf* is affected by variations of each uncertain model parameter. Here, our distinctive objective is to contribute to bridge the gap between these two types of GSA. We do so by introducing theoretical elements and an implementation strategy which enable us to appraise parameter sensitivity through the joint use of sensitivity indices based on four (statistical) moments of the *pdf* of the model output: expected value, variance, skewness and kurtosis. The key idea at the basis of this strategy is that linking parameter sensitivity to multiple statistical moments leads to improved understanding of the way a given uncertain parameter can govern key features of the shape of the *pdf* of desired model outputs, which is of interest in modern applications of hydrological and Earth sciences.

Variance-based GSA has also been applied (*a*) to guide reduction of model complexity, e.g., by setting the value of a parameter which is deemed as uninfluential to the variance of a target model output (e.g., Fu et al., 2012; Chu et al., 2015; Punzo et al., 2015), and (*b*) in the context of uncertainty quantification (Saltelli et al., 2008; Pianosi et al., 2016; Colombo et al., 2016). Only limited attention has been devoted to assess the relative effects of uncertain model parameters to the first four statistical moment of the target model output. This information would complement a model complexity analysis by introducing a quantification of the impact that conditioning the process on prescribed parameter values would have on the first four statistical moment of the output. Our approach is based on the joint use of multiple (statistical) moments for GSA. It enables us to address the following critical questions: When can the variance be considered as a reliable proxy for characterizing model output uncertainty? Which model parameter mostly affects asymmetry and/or the tailing behavior of a model output *pdf*? Does a given model parameter have a marked role in controlling some of the first four statistical moments of the model output, while being uninfluential to others? Addressing these questions would contribute to prioritizing our efforts to characterize model parameters that are most relevant in affecting important aspects of model prediction uncertainty.

Even as the richness of information content that a GSA grounded on the first four statistical moments might carry can be a significant added value to our system understanding, it may sometimes be challenging to obtain robust and stable evaluation of the proposed metrics for complex and computationally demanding models. This can be especially true when considering higher-order moments such as skewness and kurtosis. To overcome this difficulty, we cast the problem within a computationally tractable framework and rely on the use of surrogate models that mimic the full model response with a reduced computational burden. Amongst the diverse available techniques to construct a surrogate model (see, e.g., Razavi et al., 2012a,b), we exemplify our approach by considering the generalized Polynomial Chaos Expansion (gPCE) that has been successfully applied to a variety complex environmental problems (Sudret, 2008; Ciriello et al., 2013; Formaggia et al., 2013; Riva et al., 2015; Gläser et al., 2016), other model reduction techniques being fully compatible with our GSA framework. In this context, we also investigate the error associated with the evaluation of the sensitivity metrics we propose by replacing the original (full) system model through the selected surrogate model. We consider three test cases in our analysis. These include a widely employed analytical benchmark, a pumping scenario in a coastal aquifers, and a laboratory-scale transport setting. The remainder of the work is organized as follows. Section 2 presents our theoretical framework and developments. Section 3 illustrates our results for the three test cases indicated above and conclusions are drawn in Section 4.

2. Theoretical framework

We start by recalling the widely used variance-based GSA metrics in Section 2.1. These allow quantifying the contribution of each uncertain parameter to the total variance of a state variable of interest. We also provide a brief overview of the generalized Polynomial Chaos Expansion (gPCE) technique, which we use to construct a surrogate of the full system model. We then illustrate in Section 2.2 the theoretical developments underlying our approach and introduce novel GSA indices.

2.1 Sobol' indices for variance-based GSA and generalized Polynomial Chaos Expansion

We consider a target system state variable, y, which depends on N random parameters. These are collected in vector $\boldsymbol{x} = (x_1,\ x_2,\ \ldots,\ x_N)$ and defined in the parameter space $\Gamma = \Gamma_1 \times \Gamma_2 \times \ldots \Gamma_N$, $\Gamma_i = [x_{i,\min}, x_{i,\max}]$ being the support of the i-th random variable x_i. Variance-based GSA approaches consider variance as the sole metric to quantify the contribution of each uncertain parameter to the uncertainty of y. Iman and Hora (1990) introduce the following index

$$HI_{x_i} = V[y] - E\left[V[y \mid x_i]\right] = V\left[E[y \mid x_i]\right], \tag{1}$$

E[-] and V[-] respectively denoting expectation and variance operators. Index HI_{x_i} quantifies the expected reduction of variance due to knowledge of x_i (the notation $\mid x_i$ in Eq. (1) indicates conditioning on x_i). A similar measure is offered by the widely used Sobol' indices (Sobol, 1993). These have been defined starting from the Hoeffendig/Sobol decomposition (see, e.g., Sobol, 1993, Le Maître and Knio, 2010) of $y(\boldsymbol{x})$ when $\boldsymbol{x}$ is a collection of independent random variables as

$$y(\boldsymbol{x}) = y_0 + \sum_{x_i=1}^{N} y_{x_i}(x_i) + \sum_{x_i<x_j} y_{x_i,x_j}(x_i,x_j) + \ldots + y_{x_1,x_2,\ldots,x_N}(x_1,x_2,\ldots,x_N), \tag{2}$$

where

$$\begin{aligned}
y_0 &= \int_{\Gamma} y(\boldsymbol{x})\rho_{\Gamma\boldsymbol{x}}\, d\boldsymbol{x}, \\
y_{x_i}(x_i) &= \int_{\Gamma\sim x_i} y(\boldsymbol{x})\rho_{\Gamma\sim x_i} d\boldsymbol{x}_{\sim x_i} - y_0, \\
y_{x_i,x_j}(x_i,x_j) &= \int_{\Gamma\sim x_i,x_j} y(\boldsymbol{x})\rho_{\Gamma\sim x_i,x_j} d\boldsymbol{x}_{\sim x_i,x_j} - y_{x_i}(x_i) - y_{x_j}(x_j) - y_0,
\end{aligned} \tag{3}$$

and so on, $\rho_{\Gamma\boldsymbol{x}}$ being the *pdf* of $\boldsymbol{x}$. The integral $\int_{\Gamma\sim x_i} y(\boldsymbol{x})\rho_{\Gamma\sim x_i} d\boldsymbol{x}_{\sim x_i}$ in Eq. (3) represents integration of $y(\boldsymbol{x})$ over the space of all entries of vector $\boldsymbol{x}$ excluding x_i, $\rho_{\Gamma\sim x_i}$ being the corresponding *pdf*. The Sobol' index $S_{x_{i_1},x_{i_2},\ldots,x_{i_s}}$ is associated with the mixed effect of $x_{i_1},x_{i_2},\ldots,x_{i_s}$ on the variance of $y(\boldsymbol{x})$, $V[y]$, and can be computed as

$$S_{x_{i_1},x_{i_2},\ldots,x_{i_s}} = \frac{1}{V[y]} \int_{\Gamma_{x_{i_1},x_{i_2},\ldots,x_{i_s}}} y_{x_{i_1},x_{i_2},\ldots,x_{i_s}}(x_{i_1},x_{i_2},\ldots,x_{i_s})\, \rho_{\Gamma x_{i1},x_{i2},\ldots x_{is}} dx_{i_1}\ldots dx_{i_s}. \tag{4}$$

The principal and total Sobol' indices are respectively defined as

$$S_{x_i} = \frac{1}{V[y]} \int_{\Gamma_{x_i}} \left[y_{x_i}(x_i)\right]^2 \rho_{\Gamma x_i} dx_i, \tag{5}$$

$$S_{x_i}^T = S_{x_i} + \sum_{x_j} S_{x_i,x_j} + \sum_{x_j,x_k} S_{x_i,x_j,x_k} + \ldots \tag{6}$$

Note that S_{x_i} describes the relative contribution to $V[y]$ due to variability of only x_i. Otherwise, $S_{x_i}^T$ quantifies the total contribution of x_i to $V[y]$, including all terms where x_i appears. In other words,

$S_{x_i}^T$ also includes interactions between x_i and the remaining uncertain parameters, collected in vector $\boldsymbol{x}_{\sim x_i}$. Note that according to Eq.s (1)-(2) and Eq. (5)

$$S_{x_i} = \frac{V\left[E[y \mid x_i]\right]}{V\left[y\right]} = \frac{HI_{x_i}}{V\left[y\right]}, \tag{7}$$

i.e., the principal Sobol' index represents the relative expected reduction of process variance due to knowledge of (or conditioning on) a parameter. Sobol' indices are commonly evaluated via Monte Carlo quadrature schemes that can be markedly demanding in terms of computational time, especially for complex and highly non-linear settings. Relying on a generalized Polynomial Chaos Expansion, gPCE, as a surrogate of the full mathematical model of the system (Ghanem and Spanos, 1991; Xiu and Karniadakis, 2002; Le Maitre and Knio, 2010; Formaggia et al., 2013; Ciriello et al., 2013; Riva et al., 2015) allows reducing the computational burden associated with GSA techniques. The process $y(\boldsymbol{x})$ is represented as a linear combination of multivariate polynomials, $\psi_{\boldsymbol{p}}(\boldsymbol{x})$, i.e.,

$$\begin{gathered} y(\boldsymbol{x}) \cong \beta_0 + \sum_{i=1}^{N} \sum_{\boldsymbol{p} \in \Im_i} \beta_{\boldsymbol{p}} \psi_{\boldsymbol{p}}(\boldsymbol{x}) + \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{\boldsymbol{p} \in \Im_{i,j}} \beta_{\boldsymbol{p}} \psi_{\boldsymbol{p}}(\boldsymbol{x}) + \ldots, \\ \psi_{\boldsymbol{p}}(\boldsymbol{x}) = \prod_{i=1}^{N} \psi_{i,p_i}(x_i), \quad \beta_{\boldsymbol{p}} = \int_{\Gamma} y(\boldsymbol{x}) \psi_{\boldsymbol{p}}(\boldsymbol{x}) \rho_{\Gamma \boldsymbol{x}} \, d\boldsymbol{x}, \end{gathered} \tag{8}$$

where $\boldsymbol{p} = \left\{ p_1, \ldots, p_N \right\} \in \mathrm{N}^N$ is a multi-index expressing the degree of each univariate polynomial, $\psi_{i,p_i}(x_i)$; $\beta_{\boldsymbol{p}}$ are the gPCE coefficients; $\Im_i$ contains all indices such that only the *i*-th component does not vanish; $\Im_{i,j}$ contains all indices such that only the *i*-th and *j*-th components are not zero, and so on. Note that $\beta_0 \equiv y_0$, i.e., β_0 is the unconditional mean of $y(\boldsymbol{x})$. Finally, the Sobol' indices Eq.s (4)-(5) and the variance of $y(\boldsymbol{x})$ can be computed from Eq. (8) as

$$S_{x_{i_1},\ldots,x_{i_s}} = \frac{1}{V\left[y\right]} \sum_{\boldsymbol{p} \in \Im_{i_1,\ldots,i_s}} \beta_{\boldsymbol{p}}^2, \qquad S_{x_i} = \frac{1}{V\left[y\right]} \sum_{\boldsymbol{p} \in \Im_i} \beta_{\boldsymbol{p}}^2, \qquad V\left[y\right] = \sum_{\boldsymbol{p} \in \mathrm{N}^N} \beta_{\boldsymbol{p}}^2 - \beta_0^2. \tag{9}$$

2.2 New metrics for multiple-moment GSA

We introduce new metrics to quantify the expected relative change of main features of the *pdf* of *y* due to variability of model input parameters. In contrast with traditional variance-based GSA techniques of the kind described in Section 2.1, we quantify changes in the *pdf* of *y* through its first four statistical moments, i.e., mean, $E[y]$, variance, $V[y]$, skewness, $\gamma[y]$, and kurtosis, $k[y]$. The latter is an indicator of the behavior of the tails of the *pdf* of *y* and is particularly useful in the context of risk analysis, $\gamma[y]$ quantifying the asymmetry of the *pdf* of *y*.

The effect of changes of $\boldsymbol{x}$ on the mean of *y* cannot be systematically analyzed by the metrics currently available in the literature. We therefore introduce the following quantity

$$AMAE_{x_i} = \begin{cases} \dfrac{1}{|y_0|} \displaystyle\int_{\Gamma_{x_i}} \left| y_0 - E[y \mid x_i] \right| \rho_{\Gamma x_i} dx_i \; = \; \dfrac{1}{|y_0|} E\left[\left| y_0 - E[y \mid x_i] \right| \right] & if \quad y_0 \neq 0 \\ \displaystyle\int_{\Gamma_{x_i}} \left| E[y \mid x_i] \right| \rho_{\Gamma x_i} dx_i \; = \; E\left[\left| E[y \mid x_i] \right| \right] & if \quad y_0 = 0 \end{cases}, \tag{10}$$

y_0 being defined in Eq. (3). Extension of Eq. (10) to consider the joint effect of $x_{i_1}, x_{i_2}, ..., x_{i_s}$ on the mean of *y* is straightforward, leading to the following index

$$AMAE_{x_{i_1},...,x_{i_s}} = \begin{cases} \dfrac{1}{|y_0|} \displaystyle\int_{\Gamma_{x_{i1},...,x_{is}}} \left| y_0 - E[y \mid x_{i_1}, ..., x_{i_s}] \right| \rho_{\Gamma x_{i1},...,x_{is}} dx_{i_1} ... dx_{i_s} & \\ \quad = \dfrac{1}{|y_0|} E\left[\left| y_0 - E[y \mid x_{i_1}, ..., x_{i_s}] \right| \right] & if \quad y_0 \neq 0 \\ \displaystyle\int_{\Gamma_{x_{i_1},...,x_{i_s}}} \left| E[y \mid x_{i_1}, ..., x_{i_s}] \right| \rho_{\Gamma x_{i1},...,x_{is}} dx_{i_1} ... dx_{i_s} \; = \; E\left[\left| E[y \mid x_{i_1}, ..., x_{i_s}] \right| \right] & if \quad y_0 = 0 \end{cases}. \tag{11}$$

Note that index $AMAE_{x_i}$ quantifies the expected relative variation of the mean of *y* due to variations of only x_i, while $AMAE_{x_{i_1},...,x_{i_s}}$ also includes all interactions amongst parameters $x_{i_1}, x_{i_2}, ..., x_{i_s}$.

Along the same lines, we introduce the following index

$$AMAV_{x_i} = \frac{1}{V[y]} \int_{\Gamma_{x_i}} \left| V[y] - V[y \mid x_i] \right| \rho_{\Gamma x_i} dx_i \; = \; \frac{E\left[\left| V[y] - V[y \mid x_i] \right| \right]}{V[y]}, \tag{12}$$

quantifying the relative expected discrepancy between unconditional and conditional (on x_i) process variance. Note that Eq. (12) does not generally coincide with the principal Sobol' index S_{x_i} in Eq. (7) that quantifies the expected relative reduction of the variance due to knowledge of x_i (or, in other words, the relative contribution to the variance arising from uncertainty in x_i). Index $AMAV_{x_i}$ reduces to S_{x_i} only if the conditional variance, $V[y \mid x_i]$, is always (i.e., for each value of x_i) smaller than (or equal to) its unconditional counterpart $V\left[y\right]$. The difference between $AMAV_{x_i}$ and S_{x_i}, as well as advantages of using $AMAV_{x_i}$, will be elucidated through the numerical examples illustrated in Section 3. Extension of Eq. (12) to consider the joint effect of $x_{i_1}, x_{i_2}, ..., x_{i_s}$ reads

$$\begin{aligned} AMAV_{x_{i_1} ..., x_{i_s}} &= \frac{1}{V\left[y\right]} \int\limits_{\Gamma_{x_{i_1} ..., x_{i_s}}} \left| V\left[y\right] - V[y \mid x_{i_1}, ..., x_{i_s}] \right| \rho_{\Gamma x_{i1},...,x_{is}} dx_{i_1} ... dx_{i_s} \\ &= \frac{1}{V\left[y\right]} E\left[\left| V\left[y\right] - V\left[y \mid x_{i_1}, ..., x_{i_s} \right] \right| \right] \end{aligned} . \tag{13}$$

Index $AMAV_{x_{i_1}, ..., x_{i_s}}$ quantifies the expected relative discrepancy between $V\left[y\right]$ and the variance of the process conditional to joint knowledge of $x_{i_1}, x_{i_2}, ..., x_{i_s}$.

We then quantify the relative expected discrepancy between unconditional, $\gamma[y]$, and conditional, $\gamma[y \mid x_i]$, skewness through the index

$$AMA\gamma_{x_i} = \begin{cases} \dfrac{1}{\left|\gamma\left[y\right]\right|} \int\limits_{\Gamma_{x_i}} \left| \gamma\left[y\right] - \gamma[y \mid x_i] \right| \rho_{\Gamma x_i} dx_i = \dfrac{1}{\left|\gamma\left[y\right]\right|} E\left[\left| \gamma_y - \gamma[y \mid x_i] \right| \right] & if \;\; \gamma_y \neq 0 \\ \int\limits_{\Gamma_{x_i}} \left| \gamma[y \mid x_i] \right| \rho_{\Gamma x_i} dx_i = E\left[\left| \gamma[y \mid x_i] \right| \right] & if \;\; \gamma_y = 0 \end{cases} . \tag{14}$$

Extension of Eq. (14) to consider the joint effect of $x_{i_1}, x_{i_2}, ..., x_{i_s}$ gives

$$AMA\gamma_{x_{i_1}, ..., x_{i_s}} = \begin{cases} \dfrac{1}{\left|\gamma\left[y\right]\right|} \int\limits_{\Gamma_{x_{i_1}, ..., x_{i_s}}} \left| \gamma\left[y\right] - \gamma[y \mid x_{i_1}, ..., x_{i_s}] \right| \rho_{\Gamma x_{i1},...,x_{is}} dx_{i_1} ... dx_{i_s} & \\ \quad = \dfrac{1}{\left|\gamma\left[y\right]\right|} E\left[\left| \gamma\left[y\right] - \gamma\left[y \mid x_{i_1}, ..., x_{i_s} \right] \right| \right] & if \;\; \gamma\left[y\right] \neq 0 \\ \int\limits_{\Gamma_{x_{i_1}, ..., x_{i_s}}} \left| \gamma[y \mid x_{i_1}, ..., x_{i_s}] \right| \rho_{\Gamma x_{i1},...,x_{is}} dx_{i_1} ... dx_{i_s} = E\left[\left| E[y \mid x_{i_1}, ..., x_{i_s}] \right| \right] & if \;\; \gamma\left[y\right] = 0 \end{cases} \tag{15}$$

The relative variation of the kurtosis of y due to variations of a parameter x_i or of the parameter set $x_{i_1}, x_{i_2}, ..., x_{i_s}$ can be respectively quantified through

$$AMAk_{x_i} = \frac{1}{k[y]} \int_{\Gamma_{x_i}} \left| k[y] - k[y \mid x_i] \right| \rho_{\Gamma x_i} dx_i = \frac{1}{k[y]} E\left[\left| k[y] - k[y \mid x_i] \right| \right], \tag{16}$$

$$\begin{aligned} AMAk_{x_{i_1}, ..., x_{i_s}} &= \frac{1}{k[y]} \int_{\Gamma_{x_{i_1}, ..., x_{i_s}}} \left| k[y] - k[y \mid x_{i_1}, ..., x_{i_s}] \right| \rho_{\Gamma x_{i1}, ..., x_{is}} dx_{i_1} ... dx_{i_s} \\ &= \frac{1}{k[y]} E\left[\left| k[y] - k\left[y \mid x_{i_1}, ..., x_{i_s} \right] \right| \right] \end{aligned} \tag{17}$$

Relying jointly on Eq.s (10)-(17) enables one to perform a comprehensive GSA of the target process $y(\boldsymbol{x})$ quantifying the impact of $\boldsymbol{x}$ on the first four (statistical) moments of the *pdf* of $y(\boldsymbol{x})$. This strategy yields information about the way important elements of the distribution of $y(\boldsymbol{x})$, such as mean, spread around the mean, symmetry, and tailedness, are affected by uncertain model parameters collected in the parameter vector $\boldsymbol{x}$. This analysis is not feasible through a classical variance- based GSA.

Calculation of the indices we propose entails evaluation of conditional moments of $y(\boldsymbol{x})$. This step can be computationally very demanding. Along the lines of our discussion about Sobol' indices in Section 2.1, the new metrics Eq.s (10)-(17) can be evaluated via a surrogate model, as we illustrate through our examples in Section 3.

3. Illustrative Examples

The theoretical framework introduced in Section 2 is here applied to three diverse testbeds: (*a*) the Ishigami function, which constitutes an analytical benchmark typically employed in GSA studies; (*b*) a pumping scenario in a coastal aquifer, where the state variable of interest is the critical pumping rate, i.e. the largest admissible pumping rate to ensure that the extraction well is still not contaminated by seawater; and (*c*) a laboratory-scale setting associated with non-reactive transport in porous media. In the first two examples the relatively low computational costs associated with the complete mathematical description of the target outputs enables us to assess also the error associated with the evaluation of indices Eq. (10), Eq. (12), Eq. (14) and Eq. (16) through a gPCE representation of the output. In the third case, due to the complexity of the problem and the associated computational costs, we relay on the gPCE representation for the target quantity of interest. We emphasize that the use of a gPCE as a surrogate model is here considered only as an example, our GSA approach being fully compatible with any full model and/or model order reduction technique. A critical limiting factor to our and any GSA approach could be the associated computational burden. The latter is expected to increase according to the following two features, which are mainly associated with the conceptual and mathematical model used to describe the target variables of interest: (*a*) the complexity of the hydrological system (in terms of, e.g., hydrogeological heterogeneity, non-linearity and/or transient effects), and/or (*b*) the number of uncertain model input parameters considered. According to the relative weight of these features, some computational constraints might arise limiting our ability to (*i*) perform the analysis by relying exclusively on the full system model, or (*ii*) construct a sufficiently accurate surrogate model through a number of full model runs that can be affordable in terms of available computational resources. Application of our GSA methodology to scenarios of increased level of complexity will be the subject of a future study.

In all of the above scenarios, uncertain parameters x_i collected in $\boldsymbol{x}$ are considered as independent and identically distributed, *i.i.d.*, random variables, each characterized by a uniform distribution within the interval $\Gamma_i = \left[x_{i,\min}, x_{i,\max}\right]$. Note that varying the *pdf* of the uncertain model input parameters does not impact the definition of the GSA indices proposed in Section 2. Otherwise, it may affect the actual results, depending on the test case considered. All results are grounded on 5×10^5 Monte Carlo realizations, enabling convergence of all statistical moments analyzed. Series appearing in the gPCE Eq. (8) are evaluated up to a given order of truncation in all three examples. Here, we apply the total-degree rule and construct a polynomial of order *w* through a sparse grid technique (see, e.g., Formaggia et al., 2013 and references therein). We then analyze the way the selected order *w* influences the results. Note that the optimal choice of the polynomial $\psi_{\boldsymbol{p}}(\boldsymbol{x})$ in Eq.

(8) depends on the *pdf* of the random variables collected in $\boldsymbol{x}$ (Xiu and Karniadakis, 2002). In our exemplary settings we use the multidimensional Legendre polynomials which are orthonormal with respect to the uniform *pdf*.

3.1 Ishigami function

The non-linear and non-monotonic Ishigami function

$$y(\boldsymbol{x}) = ISH(\boldsymbol{x}) = \sin(2\pi x_1 - \pi) + a\sin^2(2\pi x_2 - \pi) + b\left(2\pi x_3 - \pi\right)^4 \sin(2\pi x_1 - \pi) \tag{18}$$

is widely used in the literature (e.g., Homma and Saltelli, 1996; Chun et al., 2000; Borgonovo, 2007; Sudret, 2008; Crestaux et al., 2009; Borgonovo et al. 2011) to benchmark GSA methods. Here, x_i (i = 1, 2, 3) are *i.i.d.* random variables uniformly distributed within the interval [0, 1]. Unconditional mean $E[ISH]$, variance, $V[ISH]$, skewness, $\gamma[ISH]$, and kurtosis, $k[ISH]$, of Eq. (18) can be evaluated analytically as

$$E[ISH] = \frac{a}{2}, \qquad V[ISH] = \frac{1}{2} + \frac{a^2}{8} + b\pi^4\left(\frac{1}{5} + \frac{b\pi^4}{18}\right), \qquad \gamma[ISH] = 0, \tag{19a}$$

$$k[ISH] = \frac{1}{2V^2[ISH]}\left\{\frac{3}{4} + b\pi^4\left[\frac{3}{5} + b\pi^4\left(\frac{1}{2} + 3b\pi^4\left(\frac{1}{13} + \frac{\pi^4 b}{68}\right)\right)\right] + \frac{3}{2}a^2\left[\frac{1}{2} + \frac{a^2}{32} + \pi^4 b\left(\frac{1}{5} + \frac{\pi^4 b}{18}\right)\right]\right\}. \tag{19b}$$

Equation (19) reveals that the unconditional *pdf* of *ISH* is symmetric with tails that increase with |*b*| and decrease with |*a*|, as quantified by $k[ISH]$. The conditional mean $E[ISH \mid x_i]$, variance $V[ISH \mid x_i]$, skewness $\gamma[ISH \mid x_i]$ and kurtosis $k[ISH \mid x_i]$ can be evaluated analytically as

$$E[ISH \mid x_1] = \frac{a}{2} - \frac{1}{5}\left(5 + b\pi^4\right)\sin\left(2\pi x_1\right), \qquad E[ISH \mid x_2] = a\sin^2\left(2\pi x_2\right), \qquad E[ISH \mid x_3] = \frac{a}{2}, \tag{20}$$

$$\begin{aligned} &V[ISH \mid x_1] = \frac{a^2}{8} + \frac{8b^2\pi^8}{225}\left(1 - \cos\left(4\pi x_1\right)\right), \qquad V[ISH \mid x_2] = \frac{1}{2} + b\pi^4\left(\frac{1}{5} + \frac{b}{18}\pi^4\right), \\ &V[ISH \mid x_3] = \frac{a^2}{8} + \frac{1}{2}\left(1 + b\pi^4\left(1 - 2x_3\right)^4\right)^2, \end{aligned} \tag{21}$$

$$\gamma\left[ISH \mid x_1\right] = -\frac{128 b^3 \pi^{12} \sin^3\left(2\pi x_1\right)}{4875\left(V\left[ISH \mid x_1\right]\right)^{3/2}}, \qquad \gamma\left[ISH \mid x_2\right] = 0, \qquad ISH\left[y \mid x_3\right] = 0, \quad (22)$$

$$\begin{aligned}
k\left[ISH \mid x_1\right] &= \frac{1}{V^2\left[ISH \mid x_1\right]}\left\{\frac{3}{128}a^4 + \frac{4}{75}b^2\pi^8\sin^2\left(2\pi x_1\right)\left[a^2 + \frac{1849}{5525}b^2\pi^8\sin^2\left(2\pi x_1\right)\right]\right\}, \\
k\left[ISH \mid x_2\right] &= \frac{1}{2V^2\left[ISH \mid x_2\right]}\left\{\frac{3}{4} + b\pi^4\left[\frac{3}{5} + b\pi^4\left(\frac{1}{2} + 3b\pi^4\left(\frac{1}{13} + \frac{1}{68}b\pi^4\right)\right)\right]\right\}, \\
k\left[ISH \mid x_3\right] &= \frac{3}{128V^2\left[ISH \mid x_3\right]}\left\{a^4 + 16\left(1 + b\pi^4\left(1 - 2x_3\right)^4\right)^2\left[a^2 + \left(1 + b\pi^4\left(1 - 2x_3\right)^4\right)^2\right]\right\}.
\end{aligned} \quad (23)$$

For the sole purpose of illustrating our approach, here and in the following we set *a* = 5 and *b* = 0.1, which corresponds to $E\left[ISH\right] = 2.50$, $V\left[ISH\right] = 10.84$ and $k\left[ISH\right] = 4.18$. Figure 1 depicts the first four moments of *ISH* conditional to values of x_1 (blue curves), x_2 (red curves) and x_3 (green curves) within the parameter space. The corresponding unconditional moments (black curves) are also depicted for completeness.

Comparing Eq. (19a) and Eq. (20), it is seen that $E\left[ISH \mid x_3\right]$ coincides with its unconditional counterpart $E\left[ISH\right]$, indicating that conditioning on any value of x_3 does not impact the mean of *ISH*. Otherwise, setting x_1 or x_2 to a given value clearly affects the mean of *ISH* in a way which is governed by Eq. (20) and shown in Fig. 1a. It is clear from Eq. (20) that $E\left[ISH \mid x_2\right]$ has a higher frequency of oscillation within Γ_{x_2} than has $E\left[ISH \mid x_1\right]$ within Γ_{x_1}. The global index in Eq. (10) can be evaluated analytically as

$$AMAE_{x_1} = \frac{4}{a\pi}\left|1 + \frac{b}{5}\pi^4\right|, \qquad AMAE_{x_2} = \frac{2}{\pi}\frac{|a|}{a}, \qquad AMAE_{x_3} = 0. \quad (24)$$

Note that $AMAE_{x_2}$ does not depend on specific values of *a* and *b*.

Equation (21) shows that all random model parameters influence the variance of *ISH*, albeit to different extents, as also illustrated in Fig. 1b. Note that $V\left[ISH \mid x_2\right]$ is always smaller than $V\left[ISH\right]$ (compare Eq. (19a) and Eq. (21)) and does not depend on x_2, i.e., conditioning *ISH* on x_2

reduces the process variance regardless the conditioning value. Otherwise, $V\left[ISH \mid x_3\right]$ can be significantly larger or smaller than its unconditional counterpart. Table 1 lists values of $AMAV_{x_i}$ (x_i = x_1, x_2, x_3) computed via Eq. (12) with the *a* and *b* values selected for our demonstration. The principal Sobol' indices (Sudret, 2008)

$$S_{x_1} = \frac{\left(5 + b\pi^4\right)^2}{50\,V\left[ISH\right]}, \qquad S_{x_2} = \frac{a^2}{8\,V\left[ISH\right]}, \qquad S_{x_3} = 0, \tag{25}$$

are also listed for completeness. As expected, values of $AMAV_{x_i}$ listed in Table 1 suggest that conditioning on x_3 has the strongest impact on the variance of *ISH*, followed by x_1 and x_2. Note that S_{x_3} = 0, a result which might be interpreted as a symptom that *ISH* is insensitive to x_3. The apparent inconsistency between the conclusions which could be drawn by analysing $AMAV_{x_3}$ and S_{x_3} is reconciled by the observation that the function $V\left[ISH\right] - V\left[ISH \mid x_3\right]$ can be positive and negative in a way that its integration over Γ_{x_3} vanishes (see also Fig. 1b). Therefore, the mean reduction of the variance of *ISH* due to knowledge of (or conditioning on) x_3 is zero. It is remarked that this observation does not imply that the variance of *ISH* does not vary with x_3, as clearly highlighted by Fig. 1b and quantified by $AMAV_{x_3}$.

The symmetry of the *pdf* of *ISH* is not affected by conditioning on x_2 or x_3, as demonstrated by Eq. (22). Otherwise, $\gamma\left[ISH \mid x_1\right]$ is left (or right) skewed when x_1 is smaller (or larger) than 0.5, as dictated by Eq. (22) and shown in Fig. 1c.

The conditional kurtosis $k\left[ISH \mid x_2\right]$ does not depend on the conditioning value x_2 (see Eq. (23)). We then note that this conditional moment is always larger than (or equal to) its unconditional counterpart $k\left[ISH\right]$, regardless the particular values assigned to *a* and *b*, as we verified through extensive numerical tests. This result implies that the *pdf* of *ISH* conditional on x_2 is characterized by tails which are heavier than those of its unconditional counterpart. Figure 1d reveals that $k\left[ISH \mid x_1\right]$ and $k\left[ISH \mid x_3\right]$ are smaller than $k\left[ISH\right]$ for the values of *a* and *b* implemented in this example. Table 1 lists the resulting values of $AMAk_{x_i}$ (x_i = x_1, x_2, x_3) for the selected *a* and *b* values.

We close this part of the study by investigating the error which would arise when one evaluates our GSA indices by replacing *ISH* through a gPCE surrogate model. We do so on the basis of the absolute relative error

$$e_j = \begin{cases} \left| \dfrac{j_{gPCE} - j_{full\,model}}{j_{full\,model}} \right| & if \quad j_{full\,model} \neq 0 \\ \left| j_{gPCE} - j_{full\,model} \right| & if \quad j_{full\,model} = 0 \end{cases}, \tag{26}$$

where j = $AMAE_{x_i}$, $AMAV_{x_i}$, $AMA\gamma_{x_i}$ or $AMAk_{x_i}$ (x_i = x_1, x_2, x_3); the subscripts *full model* and *gPCE* respectively indicate that quantity j is evaluated via Eq. (18) or through a gPCE surrogate model, constructed as outlined in Section 2.1. Figure 2 depicts Eq. (26) versus the total degree *w* of the gPCE. Note that the lower limit of the vertical axis of Fig. 2 is set to 0.001% for convenience of graphical representation. Approximation errors associated with GSA indices related to the mean, $AMAE_{x_i}$, rapidly approach zero as *w* increases. Note that $e_{AMAE_{x_3}}$ is smaller than 0.001% for all tested values of *w* and it is therefore not included in Fig. 2a. Values of e_j linked to $AMAV_{x_i}$, $AMA\gamma_{x_i}$ and $AMAk_{x_i}$ do not show a consistently decreasing trend until $w \geq 5$. Values of e_j associated with the variance, skewness and kurtosis decrease with approximately the same average linear rate (in log-log scale) for the largest *w* considered (Fig.s 2b, 2c and 2d). This example reinforces the need for reliably testing the accuracy of a gPCE-based model approximation as a function of the total degree desired, depending on the statistical moment of interest. Note that a generalization of our findings about the error (26) is outside the scope of the current study. This would require the derivation of (*a*) the analytical format of the *pdf* of a target model output through its gPCE based approximation at a given order w (see, e.g., Riva et al., 2015), and (*b*) the corresponding *pdf* resulting from the full system model (e.g., by formulating and solving exact equations for the target *pdf*, or its moments, typically invoking problem specific assumptions).

3.2 Critical Pumping Rate in Coastal Aquifers

The example we consider here is taken from the study of Pool and Carrera (2011) related to the analysis of salt water contamination of a pumping well operating in a homogenous confined coastal aquifer of uniform thickness b'. The setting is sketched in Fig. 3. A constant discharge, Q_w' [L^3 T^{-1}], is pumped from a fully penetrating well located at a distance x_w' [L] from the coastline and a constant freshwater flux, q_f' [L T^{-1}], flowing from the inland to the coastline, is set. Pool and Carrera (2011) introduced a dimensionless well discharge $Q_w = Q_w' / (b' x_w' q_f')$ and defined the critical pumping rate Q_c as the value of Q_w at which a normalized solute concentration monitored at the well exceeds 0.1%. A key result of the study of Pool and Carrera (2011) is that Q_c can be approximated through the following implicit equation

$$\lambda_D = 2\left[1 - \frac{Q_c}{\pi}\right]^{1/2} + \frac{Q_c}{\pi} \ln \frac{1-\left(1-Q_c/\pi\right)^{1/2}}{1+\left(1-Q_c/\pi\right)^{1/2}} \qquad \text{with} \qquad \lambda_D = \frac{\Delta\rho'}{\rho_f'} \frac{1-\left(Pe_T\right)^{-1/6}}{x_w J}. \tag{27}$$

Here, $x_w = x_w' / b'$; $J = q_f' / K$; $Pe_T = b' / \alpha_T'$; K [L T^{-1}] is the uniform hydraulic conductivity; α_T' [L] is transverse dispersivity; $\Delta\rho' = \rho_s' - \rho_f'$, ρ_f' and ρ_s' being fresh- and salt-water densities, respectively. The quantity Pe_T is a measure of the intensity of dispersive effects, J is the natural head gradient of the incoming freshwater, and x_w is the dimensionless distance of the well from the coastline. Pool and Carrera (2011) demonstrated the accuracy of Eq. (27) in predicting the critical pumping rate when $\lambda_D \in (0-10]$. Additional details about the problem setting, boundary and initial conditions, as well as geometrical configuration of the system can be found in Pool and Carrera

(2011). Here, we focus on the main result of Eq. (27) which represents the complete mathematical description of the problem we analyze. We perform a sensitivity analysis of Q_c with respect to Pe_T , J, and x_w. While the first two quantities are difficult to assess experimentally in practical applications, the well location can be considered as an operational/design variable. Table 2 lists the intervals of variation we consider for Pe_T , J and x_w. These are designed to (*a*) resemble realistic field values and (*b*) obey the above mentioned constraint about λ_D .

Numerical evaluation of the first four unconditional statistical moment of Q_c yields a mean value $E\left[Q_c\right]=1.65$, variance $V\left[Q_c\right]=0.17$, skewness $\gamma\left[Q_c\right]=-0.30$ (which indicates a light asymmetry in the *pdf*), and kurtosis $k\left[Q_c\right]=2.51$ (i.e., *pdf* tails decrease faster than those of a Gaussian distribution). Figure 4 depicts the first four moments of Q_c conditional to values of Pe_T (blue curves), J (green curves), and x_w (red curves) within the parameter space. The corresponding unconditional moments (black curves) are also depicted for completeness. Note that each parameter interval of variation has been normalized to span the range [0, 1] for graphical representation purposes. Table 3 lists the values of indices $AMAE_{x_i}$ $AMAV_{x_i}$, S_{x_i}, $AMA\gamma_{x_i}$ and $AMAk_{x_i}$ (x_i = Pe_T , J, x_w) associated with Q_c . As in our first example, it is clear that sensitivity of Q_c with respect to Pe_T , J, x_w depends on the statistical moment of interest.

Inspection of Fig. 4a reveals that the mean of Q_c is more sensitive to conditioning on J or x_w than to conditioning on Pe_T . Note that increasing Pe_T , i.e., considering advection-dominated scenarios, leads to an increase of the mean value of Q_c . This is so because the dispersion of the intruding saltwater wedge is diminished and the travel time of solutes to the well tends to increase. High values of the natural head gradient of the incoming freshwater, J, are associated with high mean values of Q_c . This is consistent with the observation that the inland penetration of the wedge is contrasted by the effect of freshwater which flows in the opposite direction. As expected, decreasing

x_w (moving the pumping well towards the coast) leads to a reduction of the mean value of Q_c. Figure 4a shows that mean Q_c varies with x_w and J in a similar way. This outcome is consistent with Eq. (27) where Q_c depends on the product $x_w J$, i.e., increasing x_w or J has the same effect on Q_c.

It can be noted (see Tab. 3) that $AMAE_{Pe_T}$ is smaller than $AMAE_J$ and $AMAE_{x_w}$, consistent with Fig. 4a. Figure 4b shows that the variance of Q_c decreases as Pe_T, J, or x_w increase. This trend suggests that the uncertainty on Q_c, as quantified by the variance, decreases as (*i*) the intruding wedge sharpens or is pushed toward the seaside boundary by the incoming freshwater or (*ii*) the well is placed at increasing distance from the coastline. Inspection of Fig. 4c and 4d shows that conditioning on Pe_T, J, or x_w causes the *pdf* of Q_c to become less asymmetric and less tailed than its unconditional counterpart. This behavior suggests that the relative frequency of occurrence of (high or low) extreme values of Q_c tends to decrease as additional information about the model parameters become available.

Figure 5 depicts error, e_j, Eq. (26) versus total degree, *w*, of the gPCE representation of Q_c, for *j* = (a) $AMAE_{x_i}$, (b) $AMAV_{x_i}$, (c) $AMA\gamma_{x_i}$ and (d) $AMAk_{x_i}$ ($x_i = Pe_T$ (blue curves), J (red curves), x_w (green curves)). These results indicate that: (*i*) e_j associated with $AMAE_{x_i}$ is negligible (≈ 1%) even for low *w*; (*ii*) $e_{AMAV_{Pe_T}}$ ≈ 10% for *w* = 2 and rapidly decreases to values below 1% for increasing *w*; (*iii*) e_{AMAV_J} and $e_{AMAV_{x_w}}$ are always smaller than 1%; and (*iv*) the trend of $e_{AMA\gamma_{x_i}}$ is similar to that of $e_{AMAk_{x_i}}$ for all x_i, with values of the order of 10% or higher for *w* = 2 and displaying a decrease with increasing *w* to then stabilize around values smaller than 1% when *w* ≈ 4 or 5. We note that the absolute relative error (26) for $AMAE_{x_i}$ with a given value of *w* is always lower than errors associated with higher order moments. Similar to our results in Section 3.1, it is clear from Fig. 5 that attaining a given level of accuracy for the gPCE based indices for Q_c requires considering a diverse total order w of the gPCE depending on the order of the statistical moment considered. As such, following the typical practice of assessing the reliability of a gPCE surrogate model solely on the basis of the variance or of a few random model realizations does not guarantee a satisfactory accuracy of the uncertainty analysis of a target model output which should consider higher-order statistical moments.

3.3 Solute transport in a laboratory-scale porous medium with zoned heterogeneity

As a last exemplary showcase, we consider the laboratory-scale experimental analysis of nonreactive chemical transport illustrated by Esfandiar et al. (2015). These authors consider tracer transport within a rectangular flow cell filled with two types of uniform sands. These were characterized by diverse porosity and permeability values, which were measured through separate, standard laboratory tests. A sketch of the experimental set-up displaying the geometry of the two uniform zones respectively formed by coarse and fine sand is illustrated in Fig. 6.

After establishing fully saturated steady-state flow, a solution containing a constant tracer concentration is injected as a step input at the cell inlet. The tracer breakthrough curve is then defined in terms of the temporal variation of the spatial mean of the concentration detected along the flow cell outlet. Esfandiar et al. (2015) modeled the temporal evolution of normalized (with respect to the solute concentration of the injected fluid) concentration at the outlet, $\overline{C}(t)$ (t denoting time), by numerically solving within the flow domain the classical Advection-Dispersion Equation implementing an original and accurate space-time grid adaptation technique. Unknown longitudinal dispersivities of the two sands ($a_{L,i}$, i = 1, 2 respectively denoting the coarse and fine sand) were considered as uncertain system parameters to be estimated against the available experimental solute breakthrough data. To minimize the computational costs in the model calibration process, Esfandiar et al. (2015) relied on a gPCE approximation of $\overline{C}(t)$. The authors constructed a gPCE of total degree w = 3 by considering $\log_{10}\left(a_{L,i}\right)$ to be two *i.i.d.* random variables uniformly distributed within

$\Gamma_{\log_{10}(a_{L,i})}$ = [-6, -2], $a_{L,i}$ being expressed in [m]. Further details about the problem set-up, numerical discretization and grid adaptation technique as well of the construction of the gPCE representation can be found in Esfandiar et al. (2015). Here, we ground the application of our new GSA metrics on the gPCE surrogate model already constructed by Esfandiar et al. (2015) to approximate $\overline{C}(t)$.

Figure 7 depicts the temporal evolution of the unconditional expected value, $E\left[\overline{C}(t)\right]$, variance, $V\left[\overline{C}(t)\right]$, skewness, $\gamma\left[\overline{C}(t)\right]$, and kurtosis, $k\left[\overline{C}(t)\right]$, of normalized $\overline{C}(t)$. Time steps $t_{0.02}$, $t_{0.4}$, and $t_{0.96}$, i.e., the times at which $E\left[\overline{C}(t)\right]$ = 0.02, 0.4, and 0.96, respectively, are highlighted in Fig. 7a. Figure 7a reveals a pronounced tailing of $E\left[\overline{C}(t)\right]$ at late times, the short time mean breakthrough being associated with a rapid temporal increase of $E\left[\overline{C}(t)\right]$. A local minimum at $t_{0.4}$ and two local peaks and are recognized in $V\left[\overline{C}(t)\right]$ (Fig. 7b). The variance peaks at times approximately corresponding to the largest values of $\partial^2 E\left[\overline{C}(t)\right]/\partial t^2$. This outcome is consistent with the results of numerical Monte Carlo (MC) simulations depicted in Fig. 8 of Esfandiar et al. (2015) where the largest spread of the MC results is observed around these locations. The local minimum displayed by $V\left[\overline{C}(t)\right]$ suggests that $\overline{C}(t)$ at observation times close to $t_{0.4}$ is mainly driven by advection, consistent with the observation that advective transport components are the main driver of the displacement of the center of mass of a solute plume. The late time variance $V\left[\overline{C}(t)\right]$ tends to vanish because the normalized breakthrough curve is always very close to unity irrespective of the values of $a_{L,1}$ and $a_{L,2}$. Joint inspection of Fig.s 7c and 7d reveals that the *pdf* of $\overline{C}(t)$ tends to be symmetric around the mean (Fig. 7c) and characterized by light tails (Fig. 7d) at about $t_{0.4}$. Otherwise, the *pdf*s of $\overline{C}(t)$ tends to display heavy right or left tails, respectively for observation times shorter or longer than $t_{0.4}$. These observations suggest that the relative frequency of rare events (i.e., very low or high solute concentrations, which can be of some concern in the context of risk assessment) is lowest at intermediate observation times across the duration of the experiment.

Figure 8 depicts the temporal evolution of (a) $AMAE_{x_i}$, (b) $AMAV_{x_i}$, (c) $AMA\gamma_{x_i}$, and (d) $AMAk_{x_i}$ (x_i = $\log_{10}(a_{L,1})$, $\log_{10}(a_{L,2})$) of $\overline{C}(t)$. Results embedded in Fig. 8 show that statistical moments of $\overline{C}(t)$ are more sensitive to $\log_{10}(a_{L,1})$ than to $\log_{10}(a_{L,2})$ at early times. The opposite occurs when $t > t_{0.4}$. Our set of results suggests that the overall early time pattern of solute breakthrough is mainly dictated by the value of $a_{L,1}$, the late time behavior being chiefly influenced by $a_{L,2}$. These conclusions are supported by the results of Fig.s 9-11, where we depict the expected value, variance, skewness, and kurtosis of $\overline{C}(t)$ conditional to $\log_{10}(a_{L,1})$ (blue curves) and $\log_{10}(a_{L,2})$ (red curves), at times $t = t_{0.02}$ (Fig. 9), $t_{0.4}$ (Fig. 10), and $t_{0.96}$ (Fig. 11). The corresponding unconditional moments are also depicted (black curves) for ease of comparison. Figure 9 shows that the first four statistical moments of $\overline{C}\left(t_{0.02}\right)$ are practically insensitive to the value of the fine sand dispersivity, $a_{L,2}$. As one could expect by considering the relative size and geometrical pattern of the two sand zones, Fig. 9a shows that the average amount of solute reaching the cell outlet at early times increases with $a_{L,1}$, because dispersion of solute increases through the coarse sand which resides in the largest portion of the domain. Figure 9b shows $V\left[\overline{C}(t_{0.02})\right]$ is negligible when $a_{L,1}$ is known. Consistent with this result, Fig.s 9c and 9d respectively show a reduction in the asymmetry and in the tailing behavior of the *pdf* of $\overline{C}\left(t_{0.02}\right)$ when $a_{L,1}$ is fixed. These results are a symptom of a reduced process uncertainty, which is in line with the observation that the bulk of the domain is filled with the coarse sand whose dispersive properties become deterministic when $a_{L,1}$ is known.

Inspection of the first four unconditional statistical moments of $\overline{C}\left(t_{0.4}\right)$ (black curves in Fig. 10) indicates that the unconditional *pdf* of $\overline{C}$ at this intermediate time is closely resembling a

Gaussian distribution. Conditioning $\overline{C}\left(t_{0.4}\right)$ on dispersivity causes a variance reduction, an increase of the tailing and the appearance of a negative (left) or positive (right) skewness, respectively when conditioning is performed on $a_{L,1}$ or $a_{L,2}$. The latter behavior suggests that in the type of experimental setting analyzed the variability of $a_{L,1}$ promotes the appearance of values of $\overline{C}\left(t_{0.4}\right)$ larger than the mean, the opposite occurring when solely $a_{L,2}$ is considered as uncertain.

Figure 11 shows that all four statistical moment of $\overline{C}\left(t_{0.96}\right)$ are chiefly sensitive to the dispersivity of the fine sand box, which is placed near the cell outlet. One can note that knowledge of $a_{L,2}$ yields a diminished variance of $\overline{C}\left(t_{0.96}\right)$, which drops almost to zero, an increased degree of symmetry and a reduce tailing of the *pdf* of $\overline{C}\left(t_{0.96}\right)$, all these evidences being symptoms of uncertainty reduction.

Results depicted in Fig.s 9-11 and our earlier observations about Fig. 7 are consistent with the expected behavior of transport in the system and the relative role of the dispersivities of the two sand regions. The high level of sensitivity of $\overline{C}(t)$ to $a_{L,1}$ at the early times of solute breakthrough is in line with the observation that solute particles are mainly advected and dispersed through the coarse sand. Both dispersivities affect the behavior of $\overline{C}(t)$ at intermediate times, when solute is traveling through both sands. The dispersivity of the coarse sand plays a minor role at late times, because virtually no concentration gradients arise in this portion of the domain. Otherwise, concentration gradients persist in the fine sand zone close to the outlet and the solute breakthrough is clearly controlled by the dispersive properties of the fine sand.

4. Conclusions

We introduce a set of new indices to be employed in the context of global sensitivity analysis, GSA, of hydrological and Earth systems. These indices consider the first four (statistical) moments of the probability density function, *pdf*, of a desired model output, *y*. As such, they quantify the expected relative variation, due to the variability in one (or more) model input parameter(s), of the expected value, variance, skewness and kurtosis of *y*. When viewed in the current research trend, our work is intended to bridge the gap between variance-based and *pdf*-based GSA approaches since it embeds the simplicity of the former while allowing for a higher-order description of how the structure of the *pdf* of *y* is affected by variations of uncertain model parameters. We cope with computational costs, which might be high when evaluating higher-order moments, by coupling our GSA approach with techniques approximating the full model response through a surrogate model. For the sake of our study, we consider the generalized Polynomial Chaos Expansion (gPCE), other model reduction techniques being fully compatible with our approach. Our new indices can be of interest in applications in the context of current practices and evolution trends in factor fixing procedures (i.e., assessment of the possibility of fixing a parameter value on the basis of the associated output sensitivity), design of experiment, uncertainty quantification and environmental risk assessment, due to the role of the key features of a model output *pdf* in such analyses.

We exemplify our methodology on three testbeds: (*a*) the Ishigami function, which is widely employed to test sensitivity analysis techniques, (*b*) the evaluation of the critical pumping rate to avoid salinization of a pumping well in a coastal aquifer, and (*c*) a laboratory-scale nonreactive transport experiment. Our theoretical analyses and application examples lead to the following major conclusions.

1. The calculated sensitivity of a model output, *y*, with respect to a parameter depends on the selected global sensitivity index, i.e., variability of a model parameter affects statistical moments of *y* in different ways and with different relative importance, depending on the statistical moment considered. Relying on the indices we propose allows enhancing our ability to quantify how model parameters affect features of the model output *pdf*, such as mean, degree of spread, symmetry and tailedness, in a straightforward and easily transferrable way.
2. Joint inspection of our moment-based global sensitivity indices and of the first four statistical conditional and unconditional moments of *y* increases our ability to understand the way the

structure of the model output *pdf* is controlled by model parameters. As demonstrated in our examples, classical variance-based GSA methods cannot be used for this purpose, leading, in some cases, to the unwarranted conclusion that a given parameter have a limited impact on a target output.

3. Analysis of the errors associated with the use of a surrogate model for the evaluation of our moment-based sensitivity indices suggests that: (*a*) attaining a given level of accuracy for the gPCE based indices associated with a target variable, *y*, might require considering a diverse total order *w* of the gPCE, depending on the target statistical moment considered in the GSA of *y*; and (*b*) in our examples, the absolute relative error (26) for $AMAE_{x_i}$ based on a given total degree *w* of the gPCE approximation is always lower than its counterpart associated with higher order moments (see Fig. 2 and 5).

Acknowledgements

Funding from MIUR (Italian Ministry of Education, University and Research, Water JPI, WaterWorks 2014, Project: WE-NEED- Water NEEDs, availability, quality and sustainability) and from the European Union's Horizon 2020 Research and Innovation program (Project: Furthering the knowledge Base for Reducing the Environmental Footprint of Shale Gas Development - FRACRISK, grant agreement 640979) is acknowledged.

References

Borgonovo, E.: A new uncertainty importance measure, Reliability Eng. Syst. Safety, 92, 771-784, 2007.

Borgonovo, E., Castaings, W., and Tarantola, S.: Moment Independent Importance Measures: New Results and Analytical Test Cases, Risk Anal., 31, 404-428, 2011.

Chu, J., Zhang, C., Fu, G., Li, Y. and Zhou, H.: Improving multi-objective reservoir operation optimization with sensitivity-informed dimension reduction, Hydrol. Earth Syst. Sci., 19, 3557-3570, doi:10.5194/hess-19-3557-2015, 2015.

Chun, M. H., Han, S. J. and Tak, N. I. L.: An uncertainty importance measure using a distance metric for the change in a cumulative distribution function, Reliab. Eng. Syst. Saf., 70, 313-321, 2000.

Ciriello, V., Di Federico, V., Riva, M., Cadini, F., De Sanctis, J., Zio, E. and Guadagnini, A.: Polynomial chaos expansion for global sensitivity analysis applied to a model of radionuclide migration in a randomly heterogeneous aquifer, Stoch. Environ. Res. Risk. Assess., 27, 945-954, doi: 10.1007/s00477-012-0616-7, 2013.

Colombo, I., Porta, G.M., Ruffo, P. and Guadagnini, A.: Uncertainty quantification of overpressure buildup through inverse modeling of compaction processes in sedimentary basins, Hydrogeol. J., doi:10.1007/s10040-016-1493-9, 2016.

Crestaux, T., Le Maître, O. and Martinez, J. M.: Polynomial chaos expansion for sensitivity analysis, Reliab. Eng. Syst. Safety, 94(7), 1161-1172, doi: org/10.1016/j.ress.2008.10.008, 2009.

Elshorbagy, A., Corzo, G., Srinivasulu, S., and Solomatine, D. P.: Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology - Part 1: Concepts and methodology, Hydrol. Earth Syst. Sci., 14, 1931-1941, doi:10.5194/hess-14-1-2010, 2010a.

Elshorbagy, A., Corzo, G., Srinivasulu, S. and Solomatine, D. P.: Experimental investigation of the predictive capabilities of data driven modeling techniques in hydrology - Part 2: Application, Hydrol. Earth Syst. Sci., 14, 1943-1961, doi:10.5194/hess-14-1943-2010, 2010b.

Esfandiar, B., Porta, G., Perotto, S. and Guadagnini, A.: Impact of space-time mesh adaptation on solute transport modeling in porous media, Water Resour. Res., 51, 1315-1332, doi:10.1002/2014WR016569, 2015.

Formaggia, L., Guadagnini, A., Imperiali, I., Lever, V., Porta, G., Riva, M., Scotti, A. and Tamellini, L.: Global sensitivity analysis through polynomial chaos expansion of a basin-scale geochemical compaction model, Comput. Geosci., 17, 25-42, doi:org/10.1007/s10596-012-9311-5, 2013.

Förster, K., Meon, G., Marke, T. and Strasser, U.: Effect of meteorological forcing and snow model complexity on hydrological simulations in the Sieber catchment (Harz Mountains, Germany), Hydrol. Earth Syst. Sci., 18, 4703-4720, doi:10.5194/hess-18-4703-2014, 2014.

Fu, G., Kapelan, Z. and Reed, P.: Reducing the complexity of multiobjective water distribution system optimization through global sensitivity analysis, J. Water Resour. Plann. Manage., 38(3), 196-207, doi:0.1061/(ASCE)WR.1943-5452.0000171, 2012.

Ghanem, R. G and Spanos, P. D.: Stochastic finite elements: a spectral approach, Berlin: Springer; 1991.

Gläser, D., Dell'Oca, A., Tatomir, A., Bensabat, J., Class, H., Guadagnini, A., Helmig, R., McDermott, C., Riva, M., and Sauter, M.: An approach towards a FEP-based model risk assessment for hydraulic fracturing operations, Energ. Procedia, 97, 387-394, 2016.

Grauso, G., Fattoruso, G., Crocetti C. and Montanari, A.: A spatially distributed analysis of erosion susceptibility and sediment yield in a river basin by means of geomorphic parameters and regression relationships, Hydrol. Earth Syst. Sci. Discussions, 4, 627-654, 2007.

Hartmann, A., Weiler, M., Wagener, T., Lange, J., Kralik, M., Humer, F., Mizyed, N., Rimmer, A., Barbera, J. A., Andreo, B., Butscher, C. and Huggenberger, P.: Process-based karst modelling to relate hydrodynamic and hydrochemical characteristics to system properties, Hydrol. Earth Syst. Sci., 17, 3305-3321, doi:10.5194/hess-17-3305-2013, 2013.

Herman, J. D., Kollat, J. B., Reed, P. M. and Wagener, T.: From maps to movies: high-resolution time-varying sensitivity analysis for spatially distributed watershed models, Hydrol. Earth Syst. Sci., 17, 5109-5125, doi:10.5194/hess-17-5109-2013, 2013.

Homma, T. and Saltelli, A.: Importance measures in global sensitivity analysis of nonlinear models, Reliab. Eng. Syst. Saf., 52, 1-17, 1996.

Iman, R.L. and Hora, S. C.: A robust measure of uncertainty importance for use in fault tree system analysis, Risk Anal., 10(3), 401-406, 1990.

Koutsoyiannis, D.: "A random walk on water", Hydrol. Earth Syst. Sci., 14, 585-601, 2010.

Krykacz-Hausmann, B.: Epistemic sensitivity analysis based on the concept of entropy, In: Prado, P., Bolado, R. (Eds.), Proceedings of SAMO2001, Madrid, pp. 31-35, 2001.

Le Maître, O. P. and Knio, O. M.: Spectral methods for uncertainty quantification, Scientific computation, Springer, 2010.

Paniconi, C. and Putti, M.: Physically based modeling in catchment hydrology at 50: survey and outlook, Water Resour. Res., 51, 7090-7129, doi:10.1002/ 2015WR017780, 2015.

Pianosi, F. and Wagener, T.: A simple and efficient method for global sensitivity analysis based on cumulative distribution functions, Environ. Model. Softw., 67, 1-11, doi: 10.1016/j.envsoft.2015.01.004, 2015.

Pianosi, F., Wagener, T., Beven, K., Freer, J., Hall, J.W., Rougier, J. and Stephenson, D.B.: Sensitivity Analysis of Environmental Models: a Systematic Review with Practical Workflow, Environmental Modelling & Software, 79, 214-232, doi: 10.1016/j.envsoft.2016.02.008, 2016.

Pool, M. and Carrera, J.: A correction factor to account for mixing in Ghyben-Herzberg and critical pumping rate approximations of seawater intrusion in coastal aquifers, Water Resour. Res., 47, W05506, doi:10.1029/2010WR010256, 2011.

Punzo, V., Marcello, M. and Biagio, C.: Do we really need to calibrate all the parameters? Variance-based sensitivity analysis to simplify microscopic traffic flow models, Intel. Trans. Sys. IEEE Trans., 16(1), 184-193, 2015.

Razavi, S. and Gupta H. V.: What do we mean by sensitivity analysis? The need for comprehensive characterization of "global" sensitivity in Earth and Environmental systems models, Water Resour. Res., 51, doi:10.1002/2014WR016527, 2015.

Razavi, S., Tolson, B.A. and Burn, D.H.: Numerical assessment of metamodelling strategies in computationally intensive optimization, Environ. Model. Softw., 34(0), 67-86, doi: 10.1016/j.envsoft.2011.09.010, 2012a.

Razavi, S., Tolson, B.A. and Burn, D.H.: Review of surrogate modeling in water resources. Water Resour. Res. 48 (7), W07401, doi: 10.1029/ 2011WR011527, 2012b.

Riva, M., Guadagnini, A. and Dell'Oca, A.: Probabilistic assessment of seawater intrusion under multiple sources of uncertainty, Adv. Water Resour., 75, 93-104, doi: 10.1016/j.advwatres.2014.11.002, 2015.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S.: Global Sensitivity Analysis. The Primer. Wiley, 2008.

Sarrazin, F., Pianosi, F. and Wagener, T.: Global sensitivity analysis of enviromental models: convergence and validation, Environ. Model. Softw., 79, 135-152, doi: 10.1016/j.envsoft.2016.02.005, 2016.

Sobol, I. M.: Sensitivity estimates for nonlinear mathematical models, Math. Model. Comput. Exp., 1, 407-417, 1993.

Sudret, B.: Global sensitivity analysis using polynomial chaos expansions, Reliab. Eng. & Syst. Safety, 93, 964-979, doi:10.1016/j.ress.2007.04.002, 2008.

Wagener, T. and Montanari, A.: Convergence of approaches toward reducing uncertainty in predictions in ungauged basins, Water Resour. Res., 47, W060301, doi:10.1029/2010WR009469, 2011.

Wagener, T., Sivapalan, M., Troch, P. A., McGlynn, B. L., Harman, C. J., Gupta, H. V., Kumar, P., Rao, P. S. C., Basu, N. B., and Wilson, J. S.: The future of hydrology: An evolving science for a changing world, Water Resour. Res., 46, W05301, doi:10.1029/2009WR008906, 2010.

Willmann, M., Sanchez-Vila, X., Carrera, J. and Guadagnini, A.: Block-upscaling of transport in heterogeneous aquifers, Calibration and Reliability in Groundwater Modelling: From Uncertainty to Decision Making, Proceedings of ModelCARE 2005, The Hague, The Netherlands, June 2005, IAHS Publ. 304, 2006.

Xiu, D. and Karniadakis, G. E. M.: The Wiener-Askey polynomial chaos for stochastic differential equations, SIAM J. Sci. Comput., 24(2), 619-644, doi: 10.1137/S1064827501387826, 2002.

Table 1. Global sensitivity index $AMAE_{x_i}$ Eq. (10), $AMAV_{x_i}$ Eq. (12), $AMA\gamma_{x_i}$ Eq. (14), and $AMAk_{x_i}$ Eq. (16) associated with the Ishigami function Eq. (18). Principal Sobol' indices, S_{x_i} Eq. (7), are also listed; $x_i = x_1, x_2, x_3$.

	$AMAE_{x_i}$	$AMAV_{x_i}$	S_{x_i}	$AMA\gamma_{x_i}$	$AMAk_{x_i}$
x_1	0.75	0.40	0.40	0.45	0.37
x_2	0.64	0.29	0.29	0.00	0.33
x_3	0.00	0.84	0.00	0.00	0.53

Table 2. Intervals of variations of Pe_T , J, x_w .

	$\Gamma_n = [x_{n,\min} - x_{n,\max}]$
Γ_{Pe_T}	$[0.01 - 0.1]$
Γ_J	$[8e^{-4} - 2.5e^{-3}]$
Γ_{x_w}	$[10 - 33]$

Table 3. Global sensitivity index $AMAE_{x_i}$ Eq. (10), $AMAV_{x_i}$ Eq. (12), $AMA\gamma_{x_i}$ Eq. (14), and $AMAk_{x_i}$ Eq. (16) associated with the critical pumping rate Q_c (25). Principal Sobol' indices, S_{x_i} Eq. (7) , are also listed; $x_i = Pe_T$, J, x_w .

	$AMAE_{x_i}$	$AMAV_{x_i}$	S_{x_i}	$AMA\gamma_{x_i}$	$AMAk_{x_i}$
Pe_T	0.07	0.14	0.09	0.35	0.09
J	0.14	0.41	0.41	0.88	0.12
x_w	0.15	0.48	0.48	0.78	0.11

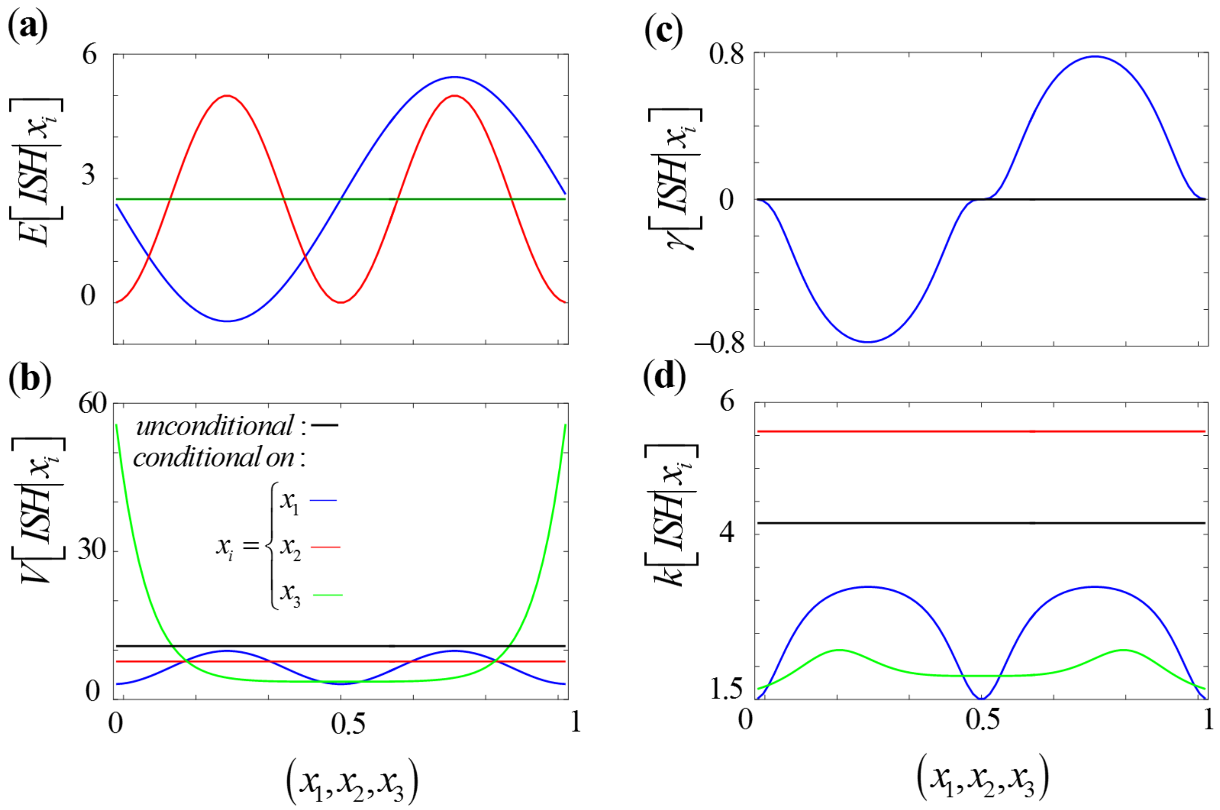

Figure 1. Variation of the first four moments of *ISH* Eq. (18) conditional to values of x_1 (blue curves), x_2 (red curves) and x_3 (green curves) within the parameter space: (**a**) expected value, $E[ISH \mid x_i]$, (**b**) variance, $V[ISH \mid x_i]$, (**c**) skewness, $\gamma[ISH \mid x_i]$, and (**d**) kurtosis, $k[ISH \mid x_i]$, (i = 1, 2, 3). The corresponding unconditional moments (black curves) are also depicted.

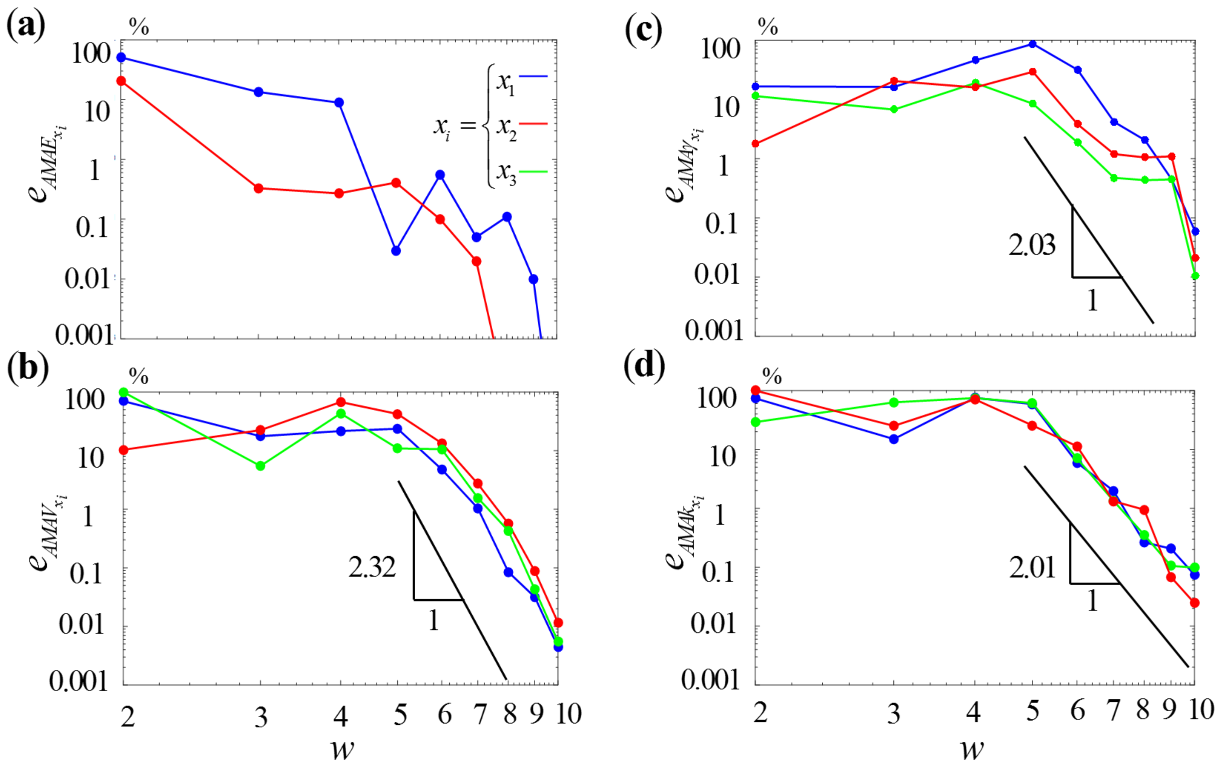

Figure 2. Error e_j Eq. (26) versus the total degree w of the gPCE representation of *IS*H for j = (**a**) $AMAE_{x_i}$, (**b**) $AMAV_{x_i}$, (**c**) $AMA\gamma_{x_i}$ and (**d**) $AMAk_{x_i}$, with $x_i = x_1$ (blue curves), x_2 (red curves), x_3 (green curves). Note that $AMAE_{x_3}$ is always smaller than 0.001%. Average slope of the rate of decrease of e_j for the largest w values considered are indicated as a reference in (**a**)-(**d**).

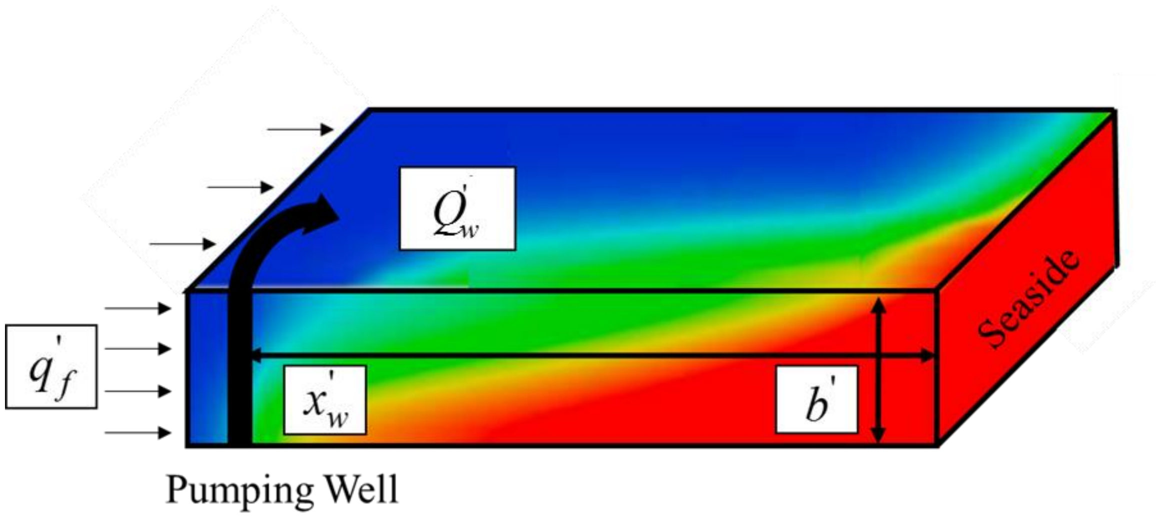

Figure 3. Sketch of the critical pumping scenario taking place within a coastal aquifer of thickness $b^{'}$. A constant freshwater (in blue) flux, $q_f^{'}$, flows from the inland to the coastline (saltwater in red). A constant discharge, $Q_w^{'}$, is pumped from a fully penetrating well located at a distance $x_w^{'}$ from the coastline. Color scale indicating variable concentration of salt is only qualitative for illustration purposes.

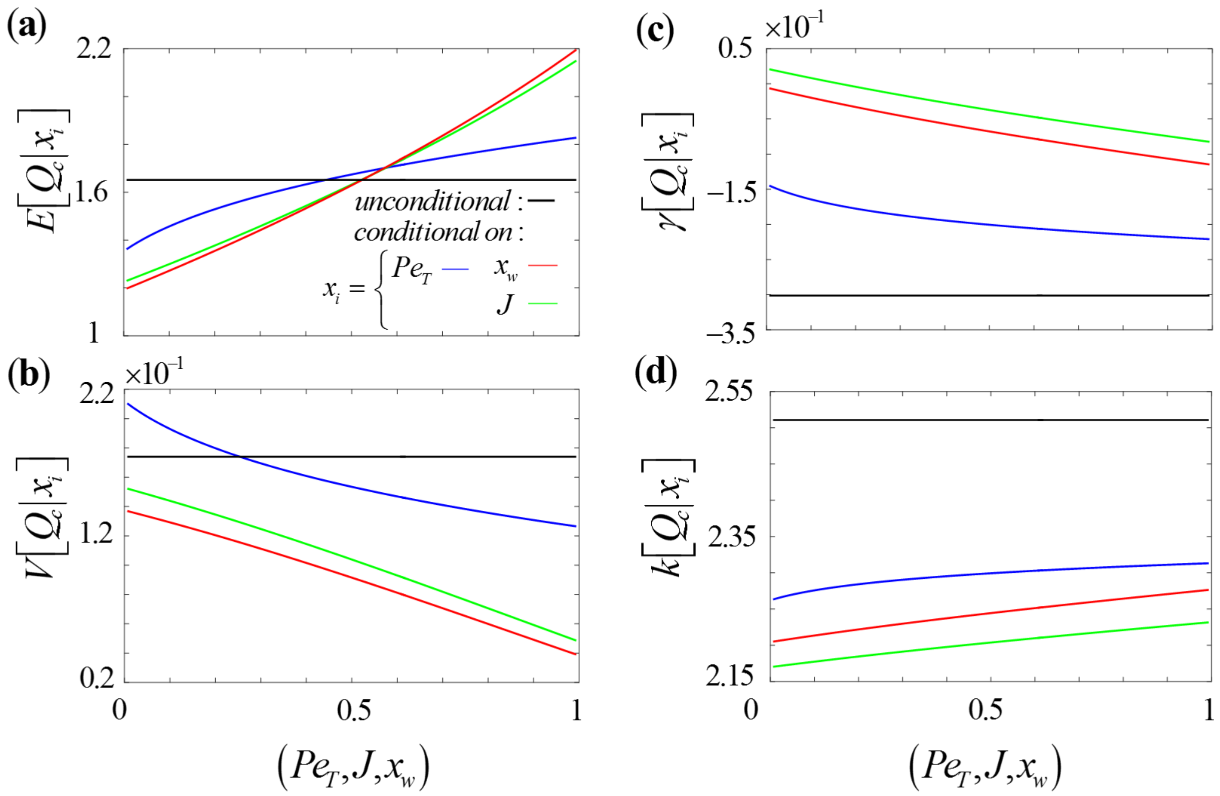

Figure 4. First four moments of Q_c Eq. (27) conditional to values of Pe_T (blue curves), J (green curves), and x_w (red curves) within the parameter space: (**a**) expected value, $E[Q_c \mid x_i]$, (**b**) variance, $V[Q_c \mid x_i]$, (**c**) skewness, $\gamma[Q_c \mid x_i]$, and (**d**) kurtosis, $k[Q_c \mid x_i]$, (x_i = Pe_T, J, x_w). The corresponding unconditional moments (black curves) are also depicted. Intervals of variation of Pe_T, J and x_w has been rescaled between zero and one for graphical representation purposes.

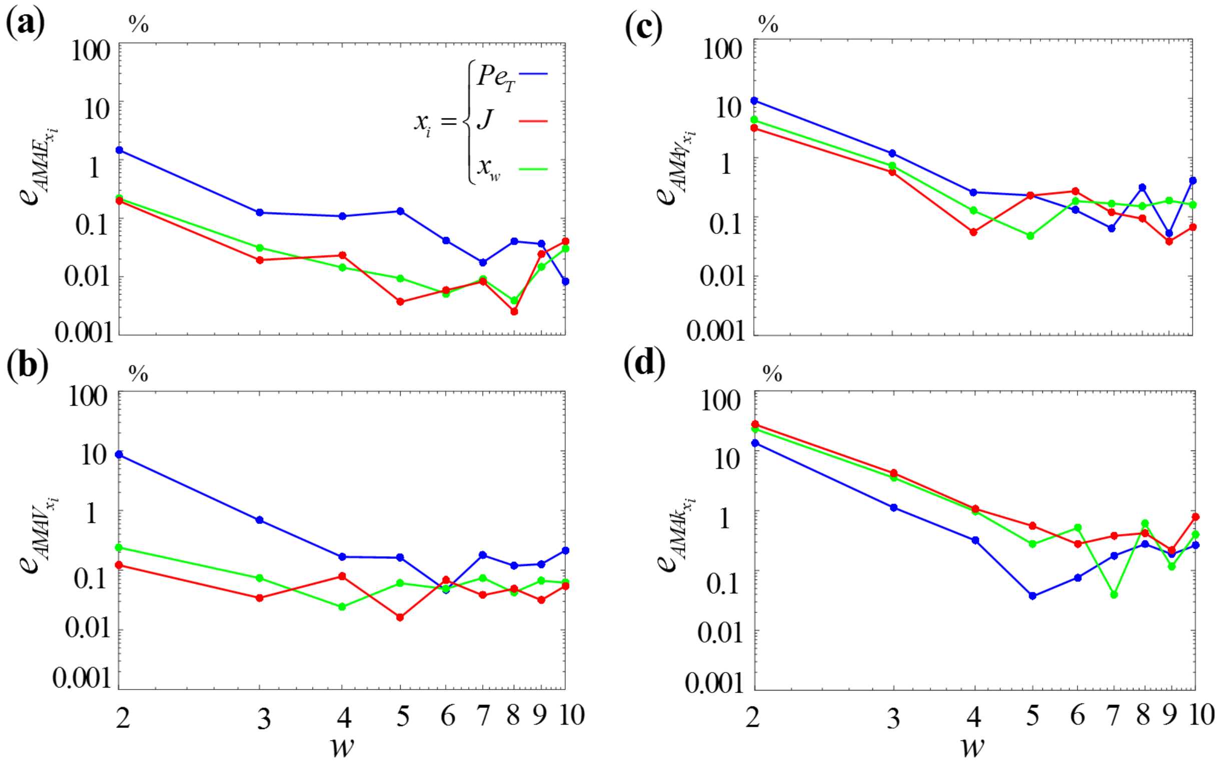

Figure 5. Error e_j Eq. (26) versus total degree w of the gPCE representation of Q_c, for j = (**a**) $AMAE_{x_i}$, (**b**) $AMAV_{x_i}$, (**c**) $AMA\gamma_{x_i}$ and (**d**) $AMAk_{x_i}$, $x_i = Pe_T$ (blue curves), J (red curves), x_w (green curves).

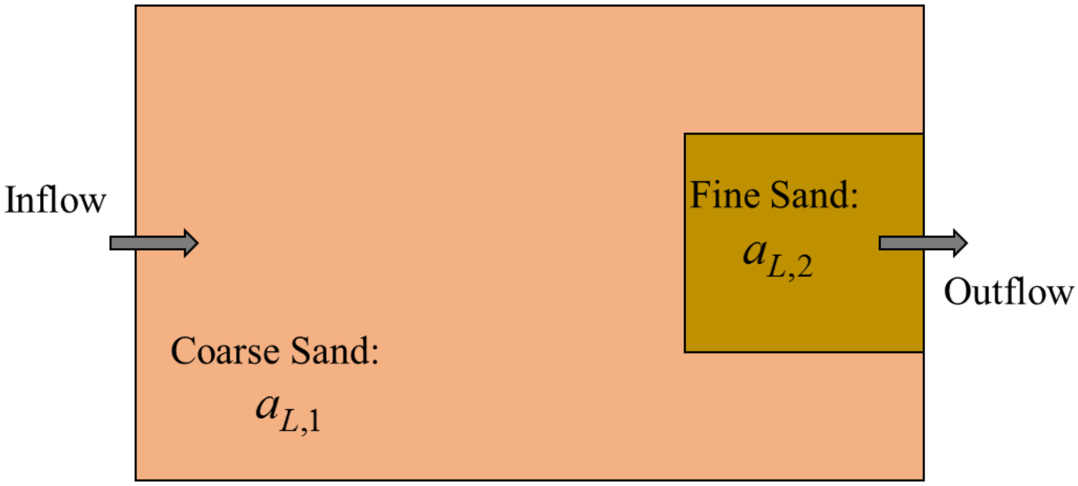

Figure 6. Sketch of the solute transport setting considered by Esfandiar et al. (2015).

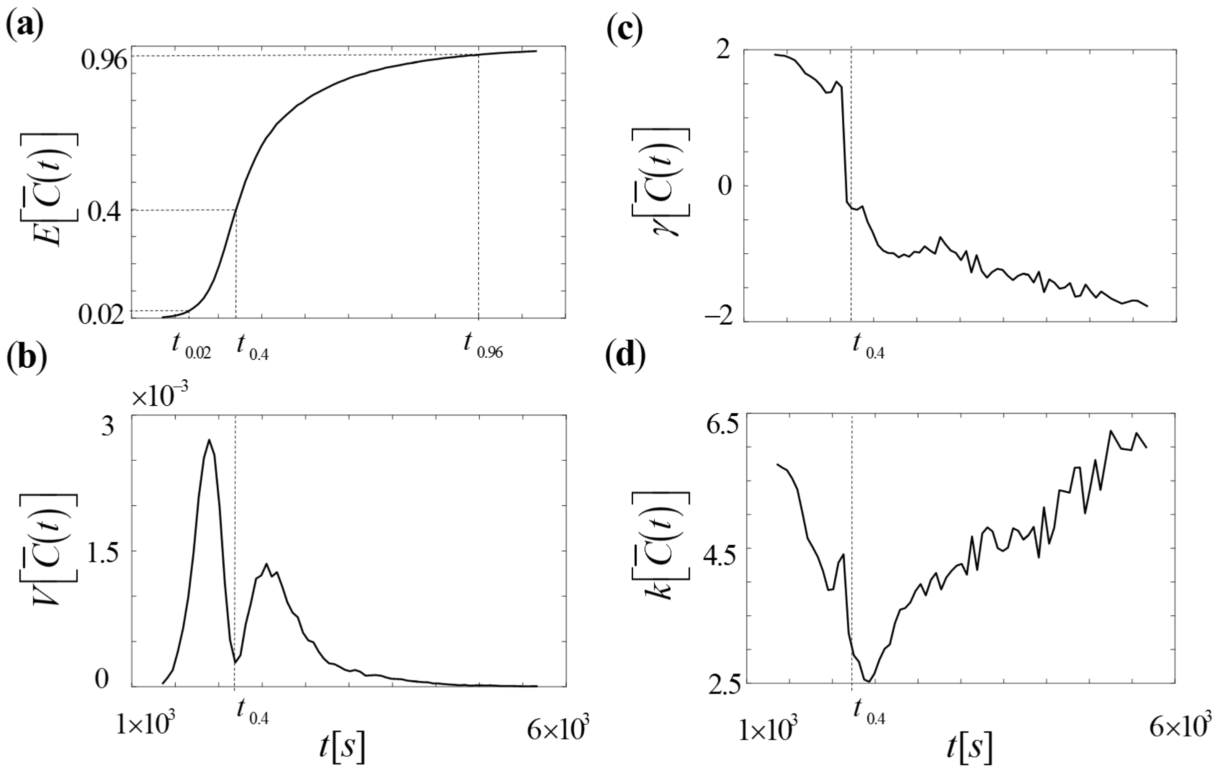

Figure 7. Temporal evolution of the unconditional (***a***) expected value, $E\left[\overline{C}(t)\right]$, (**b**) variance, $V\left[\overline{C}(t)\right]$, (**c**) skewness, $\gamma\left[\overline{C}(t)\right]$, and (**d**) kurtosis, $k\left[\overline{C}(t)\right]$, of normalized $\overline{C}(t)$. Vertical lines in (***a***) correspond to time steps $t_{0.4}$, $t_{0.02}$ and $t_{0.96}$, i.e., the times at which $E\left[\overline{C}(t)\right]$ = 0.02, 0.4, and 0.96, respectively.

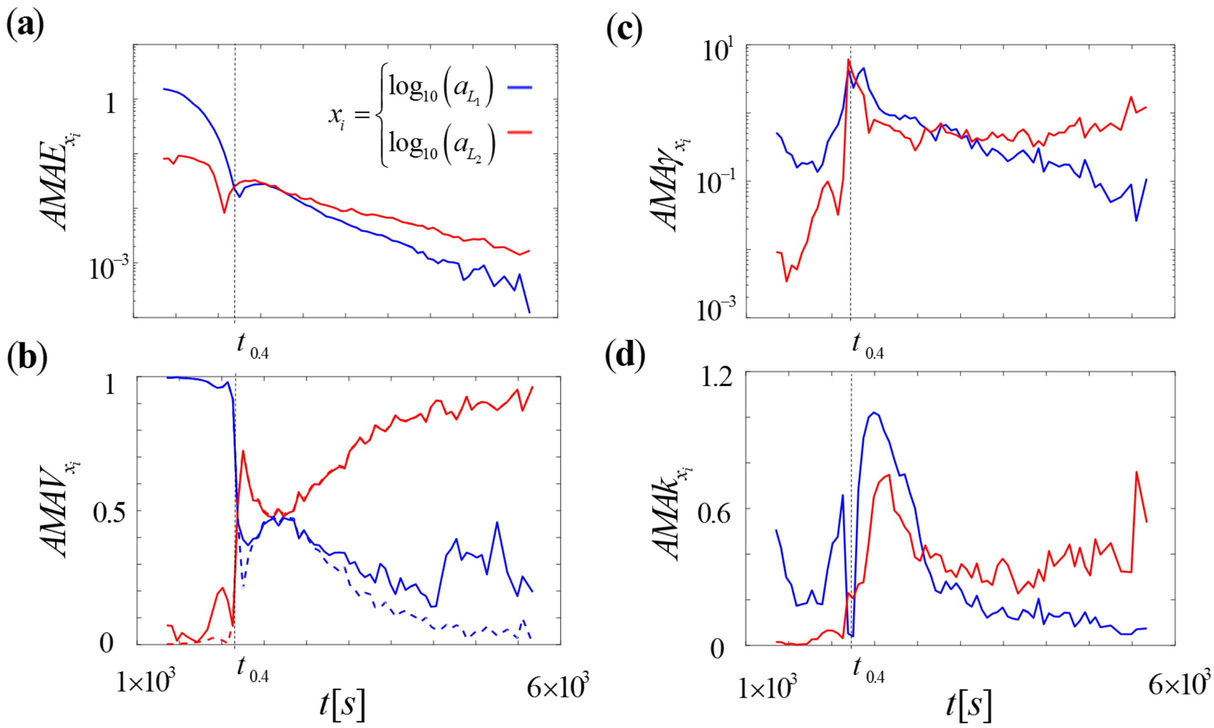

Figure 8. Time evolution of the global sensitivity index (**a**) $AMAE_{x_i}$, (**b**) $AMAV_{x_i}$ and S_{x_i} (dashed curves), (**c**) $AMA\gamma_{x_i}$, and (**d**) $AMAk_{x_i}$ of $\overline{C}(t)$ (x_i = $\log_{10}(a_{L,1})$ (blue), or $\log_{10}(a_{L,2})$ (red)).

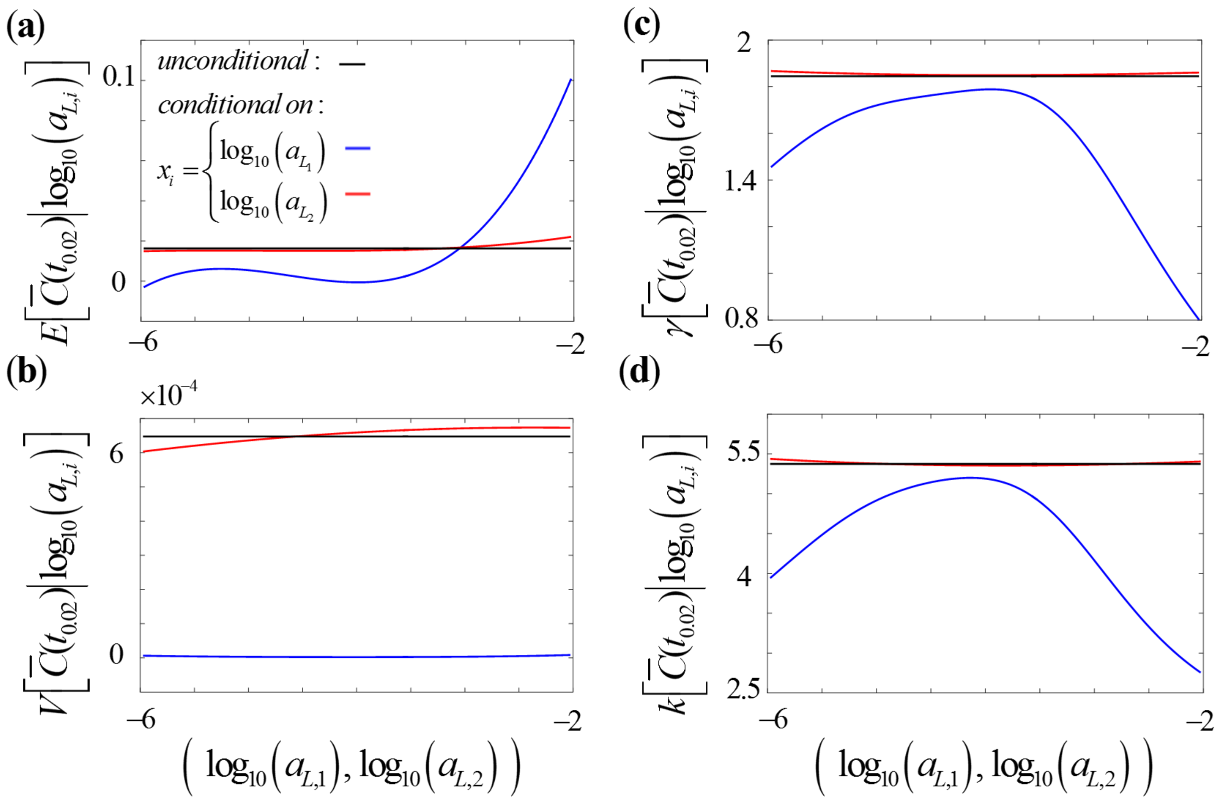

Figure 9. First four moments of $\overline{C}(t = t_{0.02})$ conditional on $\log_{10}(a_{L,1})$ (blue curves) and $\log_{10}(a_{L,2})$ (red curves), at time $t = t_{0.02}$: (**a**) expected value, $E\left[\overline{C}(t_{0.02})\middle|\log_{10}\left(a_{L,i}\right)\right]$, (**b**) variance, $V\left[\overline{C}(t_{0.02})\middle|\log_{10}\left(a_{L,i}\right)\right]$, (**c**) skewness, $\gamma\left[\overline{C}(t_{0.02})\middle|\log_{10}\left(a_{L,i}\right)\right]$, and (**d**) kurtosis, $k\left[\overline{C}(t_{0.02})\middle|\log_{10}\left(a_{L,i}\right)\right]$ ($i = 1, 2$). The corresponding unconditional moments are also depicted (black curves).

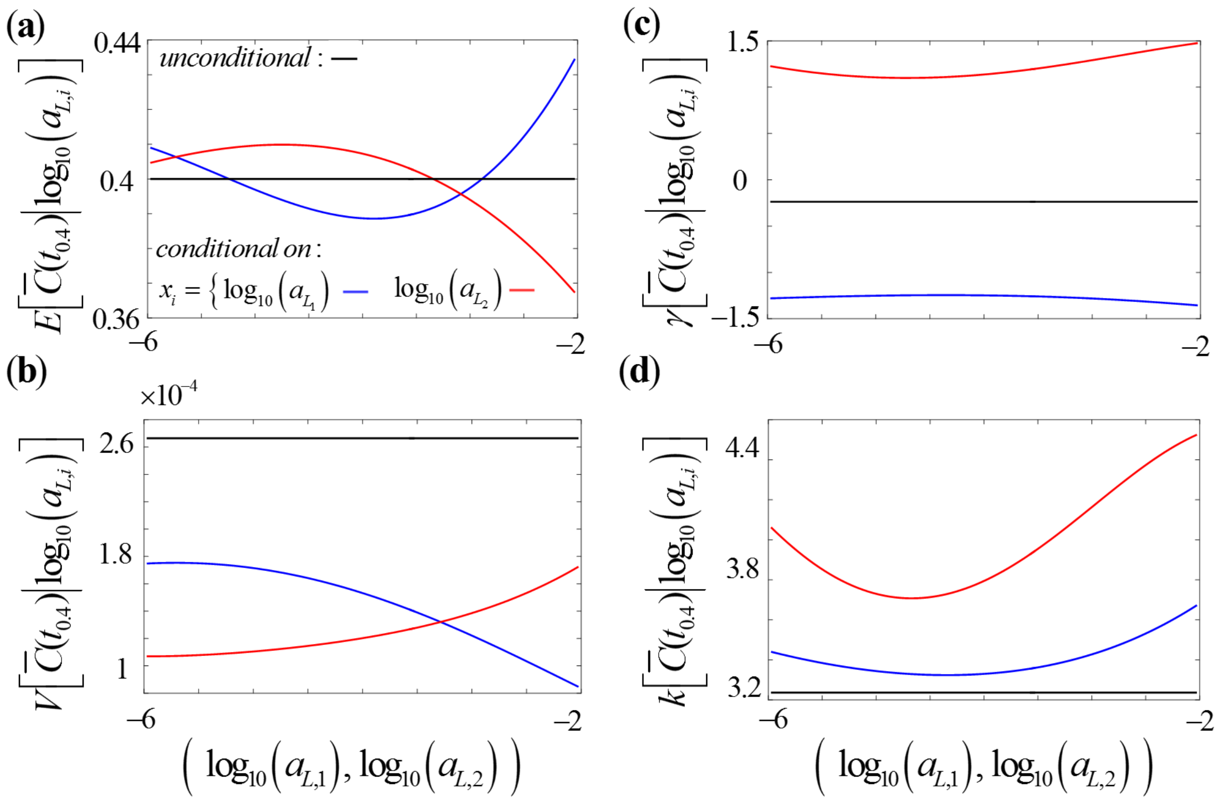

Figure 10. First four moments of $\overline{C}(t=t_{0.4})$ conditional on $\log_{10}(a_{L,1})$ (blue curves) and $\log_{10}(a_{L,2})$ (red curves), at time $t=t_{0.4}$: (**a**) expected value, $E\left[\overline{C}(t_{0.4})\middle|\log_{10}\left(a_{L,i}\right)\right]$, (**b**) variance, $V\left[\overline{C}(t_{0.4})\middle|\log_{10}\left(a_{L,i}\right)\right]$, (**c**) skewness, $\gamma\left[\overline{C}(t_{0.4})\middle|\log_{10}\left(a_{L,i}\right)\right]$, and (**d**) kurtosis, $k\left[\overline{C}(t_{0.4})\middle|\log_{10}\left(a_{L,i}\right)\right]$ (i = 1, 2). The corresponding unconditional moments are also depicted (black curves).

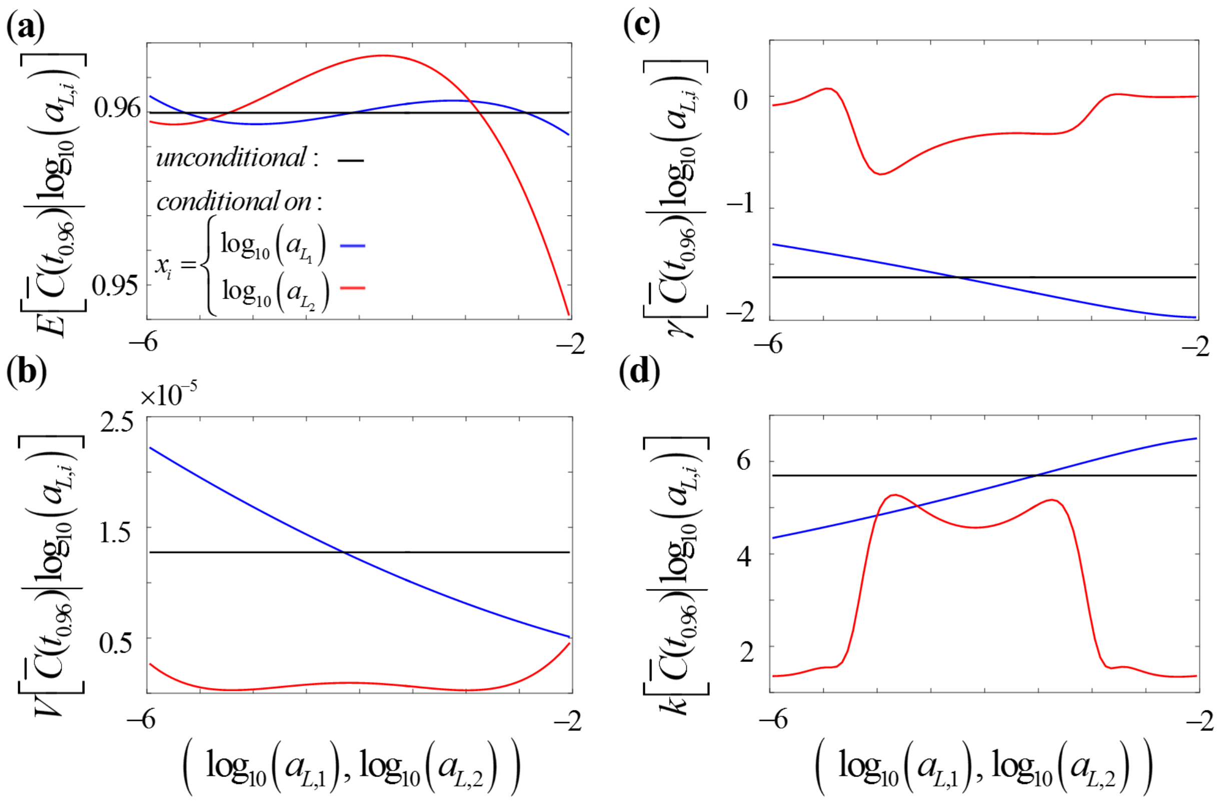

Figure 11. First four moments of $\overline{C}(t = t_{0.96})$ conditional on $\log_{10}(a_{L,1})$ (blue curves) and $\log_{10}(a_{L,2})$ (red curves), at time t = $t_{0.96}$: (**a**) expected value, $E\left[\overline{C}(t_{0.96})\middle|\log_{10}\left(a_{L,i}\right)\right]$, (**b**) variance, $V\left[\overline{C}(t_{0.96})\middle|\log_{10}\left(a_{L,i}\right)\right]$, (**c**) skewness, $\gamma\left[\overline{C}(t_{0.96})\middle|\log_{10}\left(a_{L,i}\right)\right]$, and (**d**) kurtosis, $k\left[\overline{C}(t_{0.96})\middle|\log_{10}\left(a_{L,i}\right)\right]$ (i = 1, 2). The corresponding unconditional moments are also depicted (black curves).