Sensitivity Analysis: an operational picture

A. Dell'Oca[1]

[1]Institute of Environmental Assessment and Water Research, IDAEA-CSIC, Carrer de Jordi Girona, 18-26, 08304, Barcelona, Spain

Abstract

Modelling is crucial to understand the behavior of environmental systems. A deeper comprehension of a model can be aided by global sensitivity analysis. *Variability* ascribed to model variables could have a stochastic (i.e., lack of knowledge) or an operational (i.e., possible design values) origin. Despite the possible different nature in the *variability*, current global sensitivity analysis strategies do not distinguish the latter in their formal derivations/developments. We propose to disentangle the *variability* in the operational and stochastic variables while assessing the model output sensitivity with respect to the former. Two operational sensitivity indices are introduced that serve to characterize the sensitivity of a model output of interest with respect to an operational variable in terms of (*a*) its average (with respect to the stochastic variables) intensity and (*b*) its degree of fluctuation (across the set of possible realizations of the stochastic variables), respectively. We exemplify our developments considering two scenarios. Results highlight the relevance of employing an operational global sensitivity analysis when the focus is on the influence of operational variables on model output.

Key Points

- Distinct natures of model variables are recognized within a global sensitivity analysis strategy
- Two global sensitivity indices for operational variables are proposed under uncertainty in the stochastic model variables
- The novel global sensitivity analysis strategy could help management and risk analysis for operational variables

Keywords: Global sensitivity analysis; Stochastic variables; Operational variables; Management; Uncertainty.

1. Introduction

Sensitivity analysis (SA) is establishing as a fundamental discipline to support environmental systems modelling (Razavi et al., 2021; Saltelli et al., 2021 and reference therein).

In particular, global SA (GSA) is gaining attention due to the possibility of characterizing the sensitivity of selected model output(s) with respect to model variables (or parameters) across the entirety

of their space of definition (e.g., Pianosi et al., 2016; Razavi & Gupta, 2015). From a mathematical point of view, the global character of the SA is accommodated by treating the model variables as random ones (either discrete or continuous). The latter choice appears to be natural to code and treat quantitatively *variability* and led to the emergence of diverse GSA strategies over the last decades. The most well-known strategy is the variance-based GSA of Sobol (Sobol, 1993; Owen 1994; Homma & Saltelli, 1996) according to which the degree of sensitivity of a model output is proportional to the amount of variance of the latter that is ascribable to the *variability* in one (or groups) variable(s). Recognition that the variance is insufficient to describe the model output uncertainty in case that its probability density function (*pdf*) deviates from Gaussianity, diverse strategies can be adopted by grounding the model output sensitivity onto the variations of: (*i*) distinct statistical moments of the output *pdf* (Dell'Oca et al., 2017, 2020); (*ii*) the whole output *pdf* (Borgonovo, 2007); (*iii*) the cumulative distribution function of the output (Pianosi and Wagner, 2015, 2018). A different picture is adopted in derivative-based GSA in which the model output sensitivity is linked to average (in the model variables space) measures of the variations in the model output due to variations in the model variables (Morris, 1991; Campolongo, et al 2007; Sobol & Kurechenco, 2009; Campolongo et al., 2011; Haghnegahdar & Razavi, 2017). Recently, variograms-based GSA has been advanced to characterize the sensitivity of the output as a function of the scale of perturbation of the model variables (Razavi & Gupta, 2016a; 2016b; Sheikholeslami & Razavi, 2020). Note that, variance-based and derivative-based GSA can be obtained from the variogram-based GSA. Correlation-based GSA have also been proposed by leveraging on diverse measures of the output and variable(s) degree of correlation spanning Pearosn, Spearman, partial correlation coefficients (Pastres et al.,1999), mutual information (Krzykacz-Hausmann, 2001) and copula density (Dell'Oca et al., 2020). More complete surveys of GSA can be found in Razavi & Gupta (2015), Pianosi et al., (2016) and Razavi et al., (2021).

Modelling of a system generally requires to deal with variables whose *variability* can be of an (*a*) operational or (*a*) a stochastic nature. In this context, the former relates with those system elements which values can be prescribed (either as a constants or spatially and time varying) following the prescriptions of an operator(s) 'external' to the system. This type of variables can also be understood as system design/management variables (also known as controllable, e.g., Saltelli et al., 2008 pp. 257). For example, in a subsurface hydrology setting an operational variable could be pumping rate of a well designed to mitigate the risk of water pollution. On the other hand, model variables of a stochastic nature relate with those system elements upon which we have a lack of knowledge that prevents the prescription of their values exactly. For example, in a subsurface hydrology setting the spatial distribution of the hydraulic conductivity of an aquifer is a stochastic system factor. Generally, even though our degree of knowledge about stochastic variables could evolve (e.g., through data acquisition or through a better understanding of the dynamics that determine their values), it is less likely that we can modify (or design) stochastic variables

(also known as uncontrollable, e.g., Saltelli et al., 2008 pp. 257). Yet, the distinction between operational and stochastic variables can be dependent on the specific context (e.g., a river flowrate can pass from being a stochastic variable to an operational one considering unmanaged and managed catchments). Nevertheless, despite their different nature we can treat both the operational and stochastic variables as random variables, in order to reflect (*i*) the fan of possible values that the operational variables could assume and (*ii*) the degree of uncertainty about the stochastic variables (given our current level of knowledge about them).

Operational and stochastic variables are commonly encountered jointly in a variety of water-realted setting, e.g., eutrophication of shallow water systems (Pastres et al., 1999), hydrological and financial management of river systems (Hamilton et al., 2022), protection of diked wetlands (Alminagorta, Rosenberg & Kettenring et al., 2016), water management within the socio-hydrological perspective (Elshafei et al., 2016), management of sewer overflows in a urban river (Riechel et al., 2016), risk assessment of drinking water supply (Cantoni et al., 2021), urban flood scenarios (Wu et al., 2021), live cycle of small water resource recovery facilities (Thompson et al., 2022) and tomato production in urban environments (Peña et al., 2022), regulation of rivers under climate change (Patil et al., 2022), management of grape harvest (Lo Piano et al., 2022), impact of coastal shrimp ponds in saltwater intrusion (Hou et al., 2022), crop yields under climate change (Karimi et al., 2022), analysis of water networks (Chen et al., 2022), investigation of riparian freshwater lenses (Jazayeri et al., 2021), impact of partially penetrating barriers on island freshwater lenses (Yan et al., 2021), impact of water withdrawals on waterfalls features (Schalko & Boes, 2021), functioning of sewer networks (Dobson et al., 2022), sediment management for dams (Niu & Shah, 2021), wave propagation in pressurized pipe (Wang, 2021) and algal growth dynamics (Hariz, Lawton & Craggs, 2023), just to name a few.

In this context, we are not aware of any GSA strategies in which the specific nature of the model variables is explicitly accounted for in terms of the GSA-framework and subsequent interpretation of the results. According to common GSA the stochastic and operational *variabilities* are blended (e.g., jointly sampled) during the evaluation of the output sensitivity with respect to the system variables. We believe that the above mentioned distinction is of tantamount relevance in the light of typical GSA purposes, i.e., scientific discovery, dimensionality reduction, data worth assessment and decision supports (Razavi et al., 2021). For instance, in a decision-making context our primarily objective could be to understand which operational variables are the most relevant (since it is just on those type of variables that we can intervene) while we have to account for the stochasticity in the other variables. Furthermore, inspection of the sensitivity of a model output(s) with respect to the operational variables (which values could be controlled), by explicitly recognizing that the sensitivity could vary across the possible ensemble of realizations of the stochastic variables, could enhanced the understanding of the model functioning.

For example, we are interested in evaluating the sensitivity of a well contamination with respect to (*i*) the imposed pumping rate and (*ii*) the duration of the latter. From our GSA we would like to support decisions (e.g., determine the relevance of the two operational variables) and grasp a better understanding of our system. The hydraulic properties of the aquifer are partially known (e.g., pointwise measurements of the hydraulic conductivity are available). This lack of knowledge leads to the adoption of a stochastic model for the generation of an ensemble of plausible hydraulic conductivity distributions that are compatible with the investigated aquifer (see e.g., Bear & Cheng, 2010). In this context, the common GSA strategies would suggest to explore the *variability* of operational and stochastic variables jointly (i.e., without making any distinction on their nature). As such, we will determine the sensitivity of the well contamination with respect to (*i*) the selected operational variables and (at the same time) (*ii*) the hydraulic conductivity distributions (considering interactions), i.e., the sensitivity is evaluated by implicitly blending the *variability* of the former with that of the latter despite their different nature. This approach is useful to answer questions like 'Is the *variability* in the hydraulic distribution (i.e., uncertainty) or that on the operational variables (i.e., fan of design values) the most relevant factor(s) for the well contamination?'. Questions of the latter kind are those typically posed within an uncertainty quantification analysis. On the other hand, following an operational point of view, questions of the kind 'What is the sensitivity of the well contamination to the operational variables, given the uncertainty about the hydraulic conductivity distribution?' are more relevant from an operational perspective. Note that, the uncertainty aspect is present also in the latter question. We propose to address this kind of questions by (*a*) disentangling the *variability* in the operational variables from that of the stochastic variables at the first stage of the GSA, i.e., we evaluate the GSA index (selecting the one that better suits our context) for diverse specific realizations of the stochastic variables (e.g., for diverse hydraulic conductivity fields in the example above). This leads to have a *pdf* for the GSA index. Afterwards, (*b*) we accommodate for the uncertainty in the stochastic variables by analyzing salient features of the GSA index *pdf*. We name this approach as operational GSA.

The paper is organized as follow. In Section 2 we recall the details of the moment-based GSA that we employ here to exemplify our approach. Then we present the operational GSA strategy. Section 3 presents two applications and conclusions are drawn in Section 4.

2. Methods

As a GSA technique, we focus on the moment-based approach of Dell'Oca et al. (2017). Nevertheless, the concepts and formulations associate with the operational GSA (Section 2.3) can be applied to a variety of GSA techniques that better suit needs and purposes depending on the specific context (e.g., Baroni & Francke, 2020; Razavi & Gupta, 2019; Pianosi & Wagner, 2015; Sobol, 1993).

2.1 Moment Based Sensitivity Index

2.1.2 Definition

In this section we briefly recall the definition of the moment-based sensitivity indices proposed by Dell'Oca et al. (2017). We consider a target system quantity, *y*, which depends on a set of random variables $\boldsymbol{\eta}$ associated with the system under investigation, i.e., $y(\boldsymbol{\eta})$. The sensitivity of *y* with respect to the *i*-th variable η_i, can be quantifies according with the following metric

$$AMASM_{\eta_i} = \frac{E_{\eta_i}\left[\left|\mathrm{SM}\left[y\right] - \mathrm{SM}\left[y\,|\,\eta_i\right]\right|\right]}{\mathrm{SM}\left[y\right]} \tag{1}$$

where $\mathrm{SM}[y]$ is a given unconditional statistical moment (SM) of *y*, $\mathrm{SM}\left[y\,|\,\eta_i\right]$ is a given SM of *y* conditional to η_i and $E_{\eta_i}\left[-\right]$ is the average operator with respect to η_i. In case of $\mathrm{SM}[y] = 0$, we drop the denominator in (1). According to (1), the sensitivity of *y* with respect to η_i is identified as the (average) variation of a SM of *y* due to the variability in η_i.

Note that, it is possible to consider diverse SM of *y* to characterize the sensitivity of the latter with respect to η_i (see also Dell'Oca et al., 2017 for further details). In this context, the quantification of the influence of each model variable η_i on the diverse SM of *y* (e.g., η_i could affect the first SM while not the second one) allows for a comprehensive understanding of the model functioning and can support factor screening and fixing procedures.

2.2.2 Evaluation workflow

Figure 1 depicts a sketch of the workflow employed to evaluate the GSA index $AMASM_{\eta_i}$ in Equation (1). For the purpose of illustration, we consider a model output *y* as a function of an independent coordinate *x* (e.g., space or time) and three random variables $\boldsymbol{\eta} = (\eta_1; \eta_2; \eta_3)$. The variables space $\boldsymbol{\eta}$ is explored by drawing a set of Monte Carlo of realizations of the stochastic variables $\boldsymbol{\eta}$ (e.g., according with a regular binning scheme). Leveraging on the latter it is possible to estimate the unconditional SM[*y*] and conditional $\mathrm{SM}[y\,|\,\eta_i]$ statistical moment of *y* and apply Equation (1) to obtain $AMASM_{\eta_i}$.

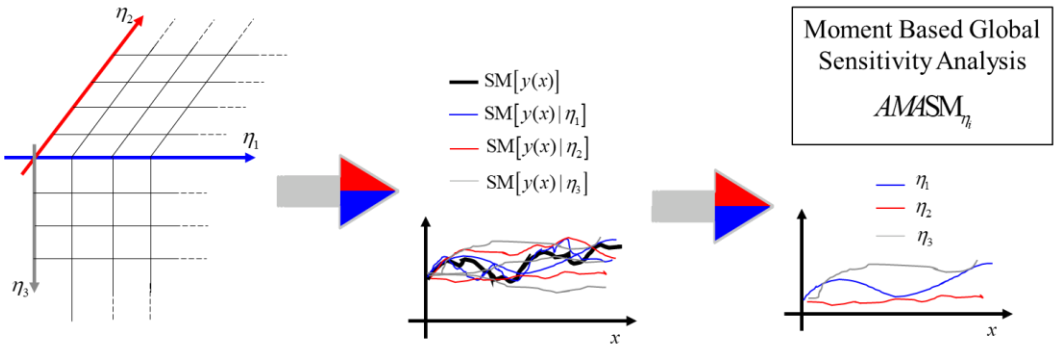

Figure 1. Sketch of the evaluation workflow leading to the moment-based GSA index $AMA\text{SM}_{\eta_i}$ for the output of interest *y*. For illustrative purpose, we consider three random variables $\boldsymbol{\eta} = (\eta_1; \eta_2; \eta_3)$ and we introduce the independent coordinate *x* (e.g. time or space).

2.2 Operational and Stochastic system variables

As highlighted in Section 1, in many environmental settings, it is possible to discern the set of system variables $\boldsymbol{\eta}$ in (*i*) a set of operational variables, $\boldsymbol{\theta}$, and (ii) a set of stochastic variables, $\boldsymbol{\omega}$. Thus, we specify that the target system quantity depends on both the operational and the stochastic variables, i.e., $y(\boldsymbol{\eta}) = y(\boldsymbol{\theta}, \boldsymbol{\omega})$. At this stage, we highlight that it is possible to evaluate the sensitivity of *y* according with Equation (1) considering an operational variable θ_i, i.e., we can quantify the impact that θ_i has on a given SM of *y* considering that θ_i can potentially assume diverse design values. The same can be done considering a stochastic variable ω_i, i.e., we can quantify the impact of ω_i on a given SM of *y* considering the uncertainty that affect ω_i. In this context, we point out that during the evaluation of $AMA\text{SM}_{\eta_i}$ no distinctions about the nature of η_i is done, as such proper cautions on the conclusions drawn upon $AMA\text{SM}_{\eta_i}$ must be taken especially if the focus of the study is tied to the characterization of the impact of the operational variables.

Additionally, according to Equation (1) the impact of an operational parameter $\eta_i = \theta_i$ on a SM of *y* is evaluated by focusing on the discrepancy between SM[*y*] and SM[*y*|θ_i], i.e., fixed the value of θ_i we evaluate a SM of *y* by averaging with respect to the other operational variables and with respect to the stochastic variables. This averaging is propaedeutic to the filtering of the impact of the other variables rather than θ_i and to account for possible interactions of θ_i. From an operational perspective, the issue with this way of proceeding lays in the fact that we evaluate the sensitivity of *y* with respect to θ_i, considering interaction with the other operational variables (a deemed aspect), but also as its interacts with the lack of

knowledge that we have about the stochastic variables of the system before the selected metric for sensitivity (i.e., the normalized average discrepancy between conditional and unconditional SM of y according to Equation (1)) is evaluated. From an operational perspective, we are more interested in the evaluation of the sensitivity of y with respect to θ_i by recognizing that the selected metric for sensitivity might varies across diverse realizations of $\boldsymbol{\omega}$. We address this aspect in the following Section 2.3.

In this context, we introduce two averaging operators, considering $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$ as independent of each other. Firstly, we consider the average with respect to the operational variables, i.e.,

$$E[\psi(\boldsymbol{\omega})] = \int_{\Gamma_{\boldsymbol{\theta}}} \psi(\boldsymbol{\omega},\boldsymbol{\theta})\, p_{\boldsymbol{\theta}}\, d\boldsymbol{\theta} \tag{2}$$

where p_θ represents the probability density function (*pdf*) of $\boldsymbol{\theta}$ and Γ_θ its support of definition. Note that, $E[\psi(\boldsymbol{\theta})]$ is still a random quantity since it is a function of $\boldsymbol{\omega}$. The second average type of average concerns the stochastic variables, i.e.,

$$<\psi(\boldsymbol{\theta})>= \int_{\Gamma_{\boldsymbol{\omega}}} \psi(\boldsymbol{\omega},\boldsymbol{\theta})\, p_{\boldsymbol{\omega}}\, d\boldsymbol{\omega} \tag{3}$$

where p_ω and Γ_ω are the *pdf* and support space of $\boldsymbol{\omega}$, respectively. Note that, in our context, $<\psi(\boldsymbol{\theta})>$ is a random quantity since we treat the possible variability in the operational variables by viewing them as random variables. Note that, in case all or some of the variables collected in $\boldsymbol{\theta}$ are discrete random variables we employ corresponding probability distribution (e.g., P_θ) and we substitute integrals with summations when needed in expressions in (2)-(3). We highlight this aspect since in many contexts it might be possible that some operational variables can only assume some well-defined values (e.g., due to practical constrains or design choices). The same holds in case of discrete variables in $\boldsymbol{\omega}$.

2.3 Moment Based Operational Sensitivity Indices

2.3.1 Definition

In this section we focus on evaluating the sensitivity of y with respect to the operational variable θ_i by explicitly recognizing that the metric selected to characterize the sensitivity can vary depending on the value of $\boldsymbol{\omega}$. Following Dell'Oca et al. (2017), we assess the impact of θ_i on a given SM of y as

$$AMASM_{\theta_i}(\omega) = \frac{E_{\theta_i}\left[\left|\mathrm{SM}[y(\omega)] - \mathrm{SM}\left[y(\omega)\,|\,\theta_i\right]\right|\right]}{\mathrm{SM}[y(\omega)]} \tag{4}$$

where SM[$y(\omega)$] is a given unconditional (with respect to $\boldsymbol{\theta}$) statistical moment of the system target quantity $y(\omega)$ conditional to a realization of $\boldsymbol{\omega}$, and SM[$y(\omega)|\,\theta_i$] is the counterpart conditional also to a value of θ_i;

the average $E_{\theta_i}[-]$ is the counterpart of Equation (2) conducted only with respect to θ_i. In case of SM[$y(\omega)$] = 0, we drop the denominator in Equation (4). The rationale behind Equation (4) is to assess the sensitivity of y with respect to θ_i, as the average variability of a given SM of y due to the possible variations in the operational variable θ_i, considering a generic realization of $\boldsymbol{\omega}$. In the light of the uncertainty about $\boldsymbol{\omega}$, the index $AMA\mathrm{SM}_{\theta_i}(\omega)$ is also a stochastic variable.

Thus, as two summary indices to characterize the operational sensitivity of y with respect to the operational variable θ_i we propose to evaluate

$$AMA\mathrm{SM}_{\theta_i}^{Op} = < AMA\mathrm{SM}_{\theta_i}(\omega) > \tag{5}$$

and

$$\sigma_{AMA\mathrm{SM}_{\theta_i}} = \sqrt{< (AMA\mathrm{SM}_{\theta_i}(\omega) - < AMA\mathrm{SM}_{\theta_i}(\omega) >)^2 >} \tag{6}$$

The index $AMA\mathrm{SM}_{\theta_i}^{Op}$ quantifies the sensitivity of y with respect to the operational variable θ_i, as the averaged value of the random $AMA\mathrm{SM}_{\theta_i}(\omega)$. In other words, $AMA\mathrm{SM}_{\theta_i}^{Op}$ provides a representative value (in the light of the uncertainty in $\boldsymbol{\omega}$) of the operational influence that θ_i has on the system target quantity y (considering a given SM of the latter). At the same time, we need to quantify the variability of the operational sensitivity of y with respect to θ_i, due to the uncertainty that plague $\boldsymbol{\omega}$. This aspect is (at least partially) quantified by $\sigma_{AMA\mathrm{SM}_{\theta_i}}$, whereas the larger it gets the more is variable (across the possible realizations of $\boldsymbol{\omega}$) the response of y to variations of θ_i. Note that, $\sigma_{AMA\mathrm{SM}_{\theta_i}}$ serves also to highlight possible interaction between θ_i and $\boldsymbol{\omega}$, whereas to a null value of $\sigma_{AMA\mathrm{SM}_{\theta_i}}$ corresponds the situation in which the sensitivity of y with respect to θ_i is equals in all the realizations of $\boldsymbol{\omega}$. Furthermore, depending on the context, it might be useful to investigate other features of the distribution of $AMA\mathrm{SM}_{\theta_i}(\omega)$ in addition to those quantified in Equations (5)-(6) (e.g., skewness and kurtosis of the distribution of $AMA\mathrm{SM}_{\theta_i}(\omega)$).

2.3.2 Evaluation Workflow

Figure 2 depicts a sketch of the workflow here employed to evaluate the operational GSA indices introduced in Equations (5)-(6). For the purpose of illustration, we consider a model output y that is a function of an independent coordinate x (e.g., space or time). The first step (S1) consists in exploring the stochastic space by drawing a set of Monte Carlo realizations of $\boldsymbol{\omega}$ (here we assume the latter to be a function

of x). Secondly (S2), each single realization of $\boldsymbol{\omega}$ interacts with the variability of the operational variables $\boldsymbol{\theta}$ (e.g., we employ a regular binning to discretize the variability of the latter; note that here $\boldsymbol{\theta}$ is independent from x) in order to obtain the unconditional SM[$y(\omega)$] and conditional SM[$y(\omega)|\theta_i$] statistical moment of interest of y given a realization of $\boldsymbol{\omega}$. Afterwards (S3), the stochastic sensitivity index $AMASM_{\theta_i}(\omega)$ in (4) is evaluated. Then steps (S2-S3) are repeated for whole set of Monte Carlo realizations of $\boldsymbol{\omega}$ and then (S5) (*i*) the average value and (*ii*) standard deviation of $AMASM_{\theta_i}(\omega)$ are evalauted, i.e., the operational GSA indices $AMASM_{\theta_i}^{Op}$ and $\sigma_{AMASM_{\theta_i}}$, respectively. A quick comparison of Figure 1 and Figure 2 helps to grasp the difference on the nature of the moment-based GSA index $AMASM_{\eta_i}$ (see Section 2.1.1) and the operational moment-based GSA indices $AMASM_{\theta_i}^{Op}$ and $\sigma_{AMASM_{\theta_i}}$ (see Section 2.3.1), especially with regard to the disentangling of the stochastic and operational variables during the evaluation of the latter.

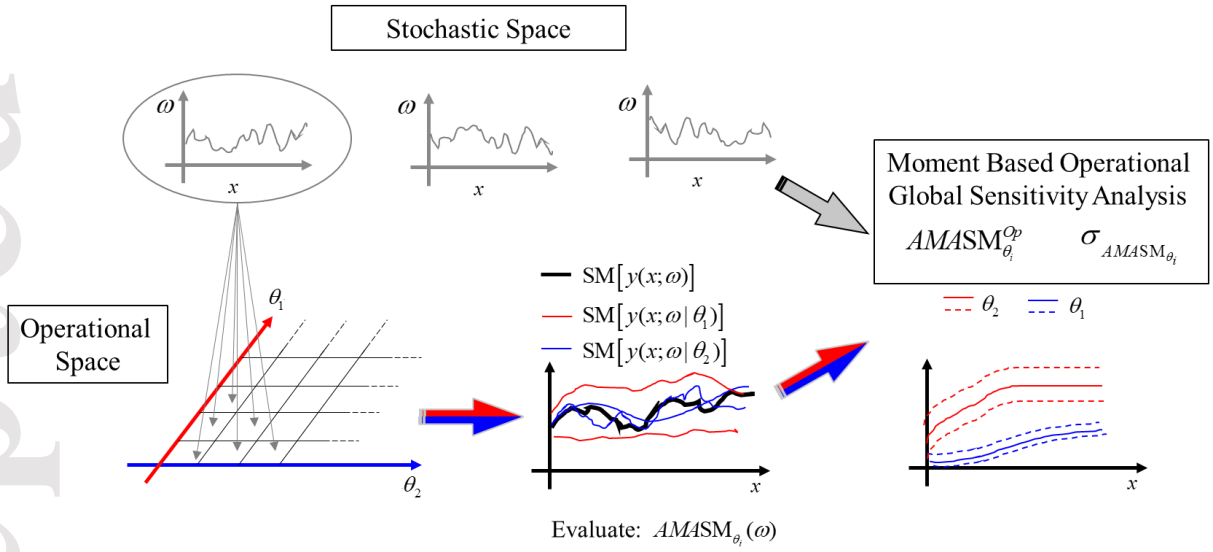

Figure 2. Sketch of the evaluation workflow to determine the operational moment-based GSA indices $AMASM_{\theta_i}^{Op}$ and $\sigma_{AMASM_{\theta_i}}$ for the output of interest y. For illustrative purpose, we consider one stochastic $\boldsymbol{\omega}$ and two operational $\boldsymbol{\theta} = (\theta_1, \theta_2)$ variables and we introduce the independent coordinate x (e.g. time or space). For each realization ω we evaluate the random moment-based SA index associated with the *i*-th operational variable, i.e., $AMASM_{\theta_i}(\omega)$. This allows to disantagle the variability in $\boldsymbol{\theta}$ from that in $\boldsymbol{\omega}$ during the evaluation of the sensitivity of y. Afterwards, the expected value and standard deviation of

$AMASM_{\theta_i}(\omega)$ are evaluated to obtain the operational moment-based GSA indices $AMASM_{\theta_i}^{Op}$ and $\sigma_{AMASM_{\theta_i}}$, respectively. In the latter step we accommodate for the variability in the stochastic variable $\boldsymbol{\omega}$.

3 Applications

In this section we exemplify the moment-based GSA strategies detailed in Section 2, considering (*i*) the Ornstein-Uhlenbeck (OU) stochastic process and (*ii*) a solute transport scenario taking place within a heterogeneous porous formation. For both example we consider the first, i.e., $\mathrm{SM} = E$ (or expected value) and the second (centered), i.e., $\mathrm{SM} = V$ (or variance), statistical moment of the output quantities of interest.

3.1 Ornstein-Uhlenbeck Process

The OU process is a well know stochastic process with application in a variety of environmental disciplines (e.g., Risken, 1989). According to the OU, a stochastic variable $y(s)$, with s being an independent variable, obeys to the following Langevin equation

$$\frac{dy(s)}{ds} = (\mu - y(s))/\gamma + \xi(s) \tag{7}$$

where μ is the long-term average value of the process, γ is the relaxation constant and $\xi(s)$ is a white noise (we consider a unitary variance). As initial condition we assume $y(0)=1e^3$. Here, we treat μ and γ as the operational variables of the system, i.e., $\boldsymbol{\theta} = (\mu, \gamma)$. Table 1 lists the support of definitions of the two operational variables, that are treated as independent uniformly distributed random variables. At the same time, the noise term $\xi(s)$ is the stochastic variable, i.e., $\boldsymbol{\omega} = (\xi)$. The OU process is solved numerically according with a Euler scheme.

Table 1. Support of definitions of the operational variables of the Ornstein-Uhlenbeck process.

Operational factor	Support
μ	$\Gamma_\mu = [1000 - 1010]$
γ	$\Gamma_\gamma = [10:100]$

In this context, our goal is the quantification of the moment-based operational global sensitivity of $y(s)$ with respect to μ and γ by means of the indices in Equations (5)-(6). We accomplish the latter by

employing the evaluation workflow of Section 2.3.2 considering 10^5 realizations of the noise series $\xi(s)$ and by discretizing both Γ_μ and Γ_γ in 10^4 uniform bins (we consider all pairwise combinations). This allows us to evaluate the necessary unconditional, i.e., SM[$y(s;\xi)$], and conditional, i.e., SM[$y(s;\xi)|\theta_i$], statistical moment(s) of interest of the OU process. Furthermore, with the same set of simulations and discretizations we follow the evaluation workflow in Section 2.1.2 to obtain the moment-based GSA detailed Section 2.1.1.

3.1.2 Global Sensitivity Analysis of the First Statistical Moment of the Ornstein-Uhlenbeck

In this section we investigate the sensitivity of the OU process by focusing on its first SM.

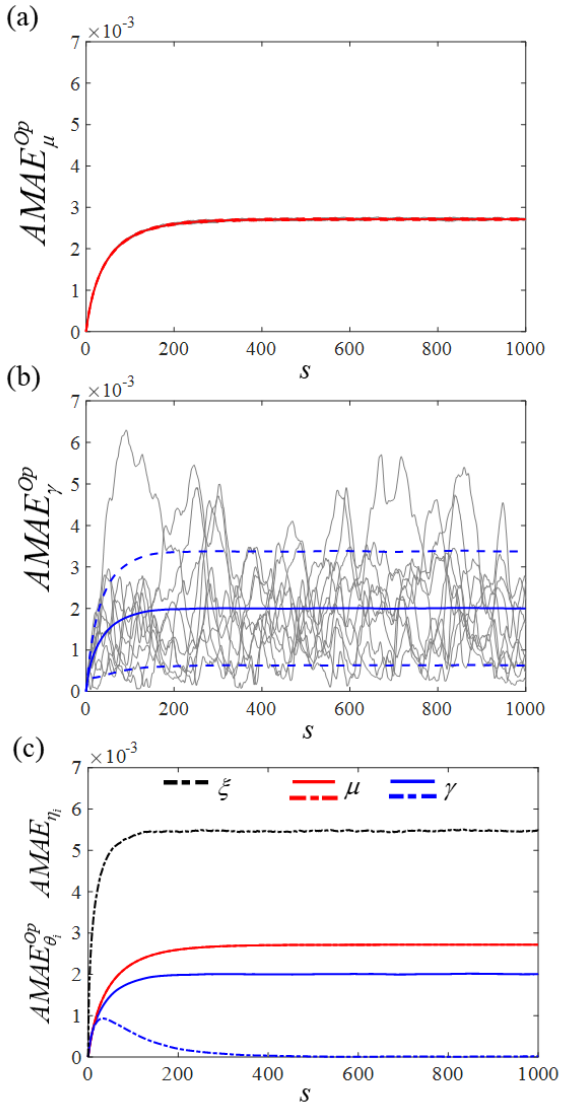

Figure 3. Moment-based GSA for the Ornstein-Uhlenbeck process considering the first SM. Operational indices $AMAE_{\theta_i}^{Op}$ (continuous curves) and $AMAE_{\theta_i}^{Op} \pm \sigma_{AMAE_{\theta_i}}$ (dashed curves) are depicted for θ_i = ((a) μ (red); (b) γ (blue)), versus *s*. Sets of twenty randomly chosen $AMAE_{\theta_i}(\xi)$ are also depicted (note that, the latter overlaps with $AMAE_{\theta_i}^{Op}$ in panel (a)). (c) Comparison of $AMAE_{\eta_i}$ (dash-dotted curves) for η_i = (ξ (black); μ (red); γ (blue)) and $AMAE_{\theta_i}^{Op}$ (continuous curves) for θ_i = (μ (red); γ (blue)) (note that, $AMAE_{\mu}$ and $AMAE_{\mu}^{Op}$ practically coincide).

Figure 3 depicts (a) $AMAE_{\mu}^{Op}$ (red continuous curve) jointly with the confidence intervals $AMAE_{\mu}^{Op} \pm \sigma_{AMAE_{\mu}}$ (red dashed curves) versus *s*. Few randomly chosen realizations of $AMAE_{\mu}(\xi)$ (grey curves) are also depicted. Inspection of Figure 3a highlights that the operational sensitivity of the first SM of *y* with respect to μ (as summarized through $AMAE_{\mu}^{Op}$) increase with *s*, i.e., as expected, variations in the value of the long-term average μ of the OU process majorly affect the first SM of *y* as it tends to relax around μ. At the same time, the first SM of *y* for small *s* is not very sensitive to μ, i.e., the behaviour of first SM of *y* for small *s* is strongly influenced by the (deterministic) initial condition. Meanwhile, the variability of the operational sensitivity of the first SM of *y* with respect to μ (as quantified by $AMAE_{\mu}^{Op} \pm \sigma_{AMAE_{\mu}}$) is always small, i.e., the impact of variations in μ on the first SM of *y* is more or less independent from the specificity of the random noise $\xi(s)$ series.

Figure 3b is the counterpart of Figure 3a considering the relaxation constant γ. Inspection of Figure 3b suggests that also the operational sensitivity of the first SM of *y* with respect to γ (as summarized through $AMAE_{\gamma}^{Op}$) is low for small values of *s* (see discussion of $AMAE_{\mu}^{Op}$ above). At the same time, both $AMAE_{\gamma}^{Op}$ and $AMAE_{\gamma}^{Op}$ aproach constant values for $s > 300$ (approximately), i.e., the operational sensitivity of the first SM of *y* does not vary when the OU reaches its long-term/asymptotic behavior. Meanwhile, $AMAE_{\gamma}^{Op}$ tends to be smaller than $AMAE_{\mu}^{Op}$ while the variability of the operational sensitivity of the first SM of *y* with respect to γ (as quantified by $AMAE_{\gamma}^{Op} \pm \sigma_{AMAE_{\gamma}}$) is larger than that associated with μ, i.e., variations in the relaxation variable γ lead to diverse values of the first SM of *y* depending on the intensity of the series of $\xi(s)$ since γ interacts with the current state of *y* (see Equation 7).

Figure 3c compares $AMAE_{\eta_i}$ (dash-dotted curve) for η_i = (μ (red); γ (blue); ξ (black)) with $AMAE_{\theta_i}^{Op}$ (continous curve) for θ_i = (μ (red); γ (blue)), as a function of *s*. We note that, $AMAE_{\mu}$ and $AMAE_{\mu}^{Op}$ practically coincide, i.e., the sensitivity of the first SM of *y* with respect to μ is independent from disentangling or not the variability in $\xi(s)$ from that in (μ, γ). We explain this result by considering that μ determines the value around which the OU relaxes (see Equation 7) and thus the influence of μ on the first SM of *y* is detached from the specificity of the noise $\xi(s)$. On the contrary, $AMAE_{\gamma}$ and $AMAE_{\gamma}^{Op}$ behave differently: as *s* increases $AMAE_{\gamma}$ approaches a null value after a peak, while $AMAE_{\gamma}^{Op}$ reaches a constant. The behavior of $AMAE_{\gamma}$ can be explained by noticing that $E[y(s)|\gamma]$ approaches the same value independently of γ (yet, at diverse 'speed' depending on γ resulting in the non-zero values of $AMAE_{\gamma}$ over small *s*) leading to $AMAE_{\gamma} = 0$ for $s > 300$, i.e., considering $E[y(s)|\gamma]$ (rather than $E[y(s;\xi)|\gamma]$) we do not keep track of the specificity of the series of $\xi(s)$ and the interaction between γ and the current state of $y(s)$ (this in contrast with $AMAE_{\gamma}^{Op}$, see previous discussion). Finally, for the specific values of the ranges listed in Table 1, $AMAE_{\xi}$ is larger than $AMAE_{\mu}$ and $AMAE_{\gamma}$ for all *s*: the lack of knowledge about $\xi(s)$ leads to stronger discrepancies between $E[y(s)]$ and $E[y(s)|\xi]$ that those ascribable to the variability in the long term average and in the relaxation constant.

3.1.2 Global Sensitivity Analysis of the Second Statistical Moment of the Ornstein-Uhlenbeck

In this section we investigate the sensitivity of the OU process by focusing on its second (centered) SM.

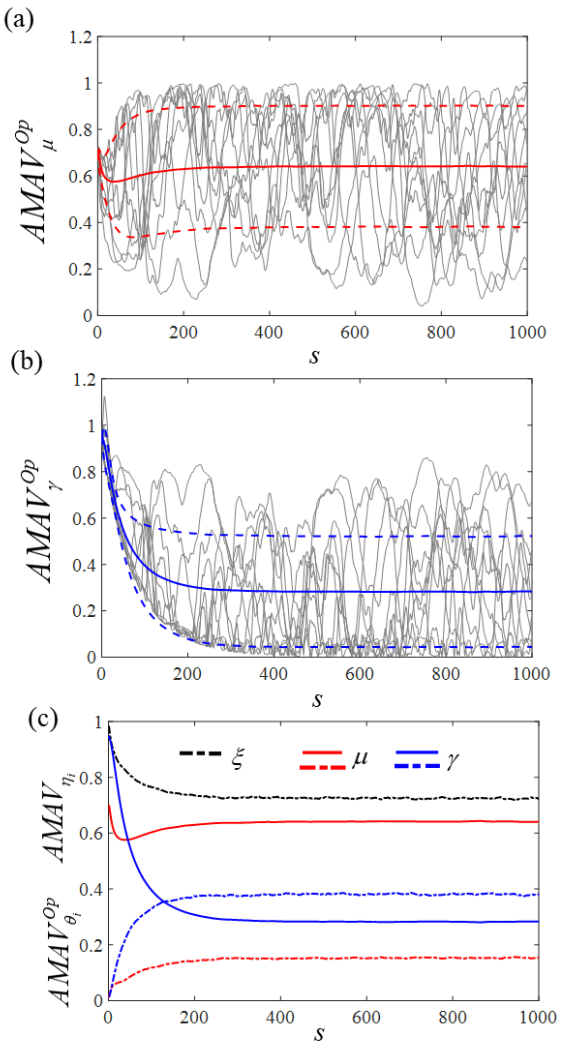

Figure 4. Moment-based GSA for the Ornstein-Uhlenbeck process considering the second (centered) SM. Operational indices $AMAV_{\theta_i}^{Op}$ (continuous curves) and $AMAV_{\theta_i}^{Op} \pm \sigma_{AMAV_{\theta_i}}$ (dashed curves) are depicted for θ_i = ((a) μ (red); (b) γ (blue)), versus s. Sets of twenty randomly chosen $AMAV_{\theta_i}(\xi)$ are also depicted. (c) Comparison of $AMAV_{\eta_i}$ (dash-dotted curves) for η_i = (ξ (black); μ (red); γ (blue)) and $AMAV_{\theta_i}^{Op}$ (continuous curves) for θ_i = (μ (red); γ (blue)).

Figure 4 depicts $AMAV_{\theta_i}^{Op}$ (continous curve) jointly with the confidence intervals $AMAV_{\theta_i}^{Op} \pm \sigma_{AMAV_{\theta_i}}$ (dashed curve) for θ_i = ((a) μ (red); (b) γ (blue)) and twenty corresponding randomly

chosen realizations of $AMAV_{\theta_i}(\xi)$ (grey curves). Comparison of $AMAE_{\theta_i}^{Op}$ (Figure 3a-b) and $AMAV_{\theta_i}^{Op}$ (Figure 3a-b) reveals different trends in the operational sensitivity of the OU depending on the SM in correspondece of s < 300 (approximately). At the same time, considering s > 300 (approximately) both $AMAE_{\theta_i}^{Op}$ and $AMAV_{\theta_i}^{Op}$ become constants. The latter observation suggests that once the OU approaches the long-term/asymptotic behaviour the variability in each operational variable influences the first and second (centered) SM of the OU in a stationary manner, i.e., there is no influence of the pre-asymptotic behaviour. Furthermore, the bounds $AMAV_{\mu}^{Op} \pm \sigma_{AMAV_{\mu}}$ are not negligebles (this is in contrast with values of $AMAE_{\mu}^{Op} \pm \sigma_{AMAE_{\mu}}$ in Figure 3a), i.e., the way in which variations in μ influence the second (centered) SM of the OU do depend on the specificity of $\xi(s)$.

Figure 4c compares $AMAV_{\eta_i}$ (dash-dotted curve) for η_i = (ξ (black); μ (red); γ (blue)) and $AMAV_{\theta_i}^{Op}$ (continous curve) for θ_i = ((a) μ (red); (b) γ (blue)), as a function of s. Inspection of Figure 4c reveals that also the diverse $AMAV_{\eta_i}$ approach constant values for s > 300 and that the stochasticity in ξ is the dominant factor. At the same time, we note that the relative importance of μ and γ vary when considering $AMAV_{\eta_i}$ or $AMAV_{\theta_i}^{Op}$. Furthermore, the comparison of $AMAV_{\gamma}$ and $AMAE_{\gamma}$ (see Figure 3c) reveals a persisten influence of γ on the OU process when we consider the second (centered) SM rather than the first SM, i.e., the relaxation variable continously influences the variability (as measured by the second (centered) SM) across diverse realizations (associated with variations of μ and ξ) of the OU process even when the long-term behaviour is approched.

3.2 Solute Transport within Heterogeneous Porous Formation

In this section we focus on the spreading of a solute body that travels through a porous heterogeneous formation under the action of four pumping wells, i.e., W1, W2, W3 and W4. (see Figure 5 for a sketch of the problem setting).

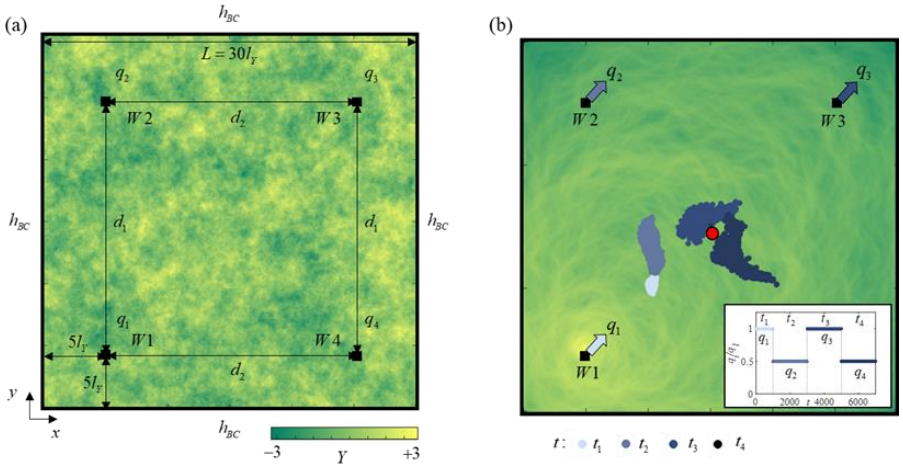

Figure 5. Sketch of the solute transport scenario within randomly heterogeneous porous formation. (a) The system stochastic variable is the spatial distribution of the logarithm of the hydraulic conductivity, i.e., $Y(\boldsymbol{x})$ (see colormap). The operational variables are the wells W1-4 distances (d_1, d_2) and the ratio of the wells pumping rates $r = q_2/q_1$ (with $q_3 = q_1$ and $q_4 = q_2$). A constant hydraulic head h_{BC} is imposed along all the domain boundaries. (b) Solute body is initially at the domain center (red circle). The magnitude of the Darcy's velocity field during the activation of W1 is depicted as a colormap (green low, yellow high), note the increase in the flow perturbations in proximity of W1. Inset depicts an example of the pumping wells activation schedule considering a ratio of pumping $r = 0.5$. The time window t_i represents the time of activation of the i-th well. We depict four snapshots of the solute plume at the end of each t_i interval.

Considering heterogeneous geological formations, the typical data scarcity and the erratic nature of the hydraulic conductivity lead to a lack of knowledge about the exact spatial arrangement of the latter. Thus hydraulic conductivity is treated as a random field (e.g., Bear and Cheng, 2010). As a common assumption, the logarithm of the hydraulic conductivity (divided by its geometric mean, here set equal to one for convenience), i.e., $Y(\boldsymbol{x})$ (where $\boldsymbol{x} = (x, y)$ is the space location coordinate, see Figure 5a), is considered as a stationary multi-Gaussian field, i.e., $Y(\boldsymbol{x}) \sim N(0, \sigma^2_Y)$ and we set $\sigma^2_Y = 1$. Regarding the spatial correlation of $Y(\boldsymbol{x})$, we assume an isotropic exponential spatial covariance, i.e., $C_Y(\boldsymbol{z}) = \sigma^2_Y \exp(-\boldsymbol{z}/l_Y)$, where $\boldsymbol{z}$ is a spatial lag vector and l_Y is the correlation length scale of $Y(\boldsymbol{x})$. Therefore, the random field $Y(\boldsymbol{x})$. represents the stochastic variable of the system under investigation, i.e., $\boldsymbol{\omega} = Y(\boldsymbol{x})$. The stochastic dimension is explored by drawing 5000 Monte Carlo realizations of $Y(\boldsymbol{x})$ (we employ a regular grid of square elements of edge size equals to $l_Y/10$). The physical domain of interest is a two dimensional square with edge length $L = 30l_Y$ (see Figure 5a).

Regarding the set of operational variables of the problem, we consider the distance between (*i*) W1-W2 and W3-W4, i.e., d_1, and between (*ii*) W1-W4 and W3-W2, i.e., d_2 (see the wells arrangement in Figure 6a). Additionally, we consider (*iii*) the ratio between the pumping rate in W2 and W1, i.e., $r = q_2/q_1$, given that $q_3 = q_1$ and $q_4 = q_2$. Thus, the vector of operational variable is $\boldsymbol{\theta} = (d_1; d_2; r)$. We treat the operational variables as independent and discrete random variables with probabilities listed in Table 2. This choice reflects a possible design phase in which the operators decide to focus on few equally likely values of the operational variables. Regarding the pumping schedule of the wells we firstly activate W1 within the time window $t_1 = [0\text{-}t_p/2]$ and subsequently W2 during $t_2 = [t_p/2\text{-}3t_p/2]$, W3 during $t_3 = [3t_p/2\text{-}5t_p/2]$ and, lastly, W4 during $t_4 = [5t_p/2\text{-}7t_p/2]$, with t_p=2000 (in consistent unit) (see Figure 5b for an example of a pumping schedule).

Table 2. Discrete probabilities, *P*, assigned to the diverse values of the operational variables for the solute transport scenario.

Distance d_1	Distance d_2	Pumping ratio *r*
$P(d_1 = 20) = 1/3$	$P(d_2 = 20) = 1/3$	$P(r = 0) = 1/3$
$P(d_1 = 21) = 1/3$	$P(d_2 = 21) = 1/3$	$P(r = 0.5) = 1/3$
$P(d_1 = 22) = 1/3$	$P(d_2 = 22) = 1/3$	$P(r = 1) = 1/3$

We simplify the flow problem by neglecting storage effects due to matrix and fluid compressibility. Thus, conservation of mass translates into the divergence free of the Darcy' flow (see Bear & Cheng, 2010). Regarding the boundary conditions, we impose a constant value of the hydraulic head h_{BC} along the domain edges in order to avoid any background flow. The flow problem is solved by leveraging on the Matlab Reservoir Simulation Toolbox (Lie, 2019), employing a regular grid of square elements of size $l_Y/10$. Transport of solute within the ensuing Darcy's scale flow field is grounded on the advection and diffusion equation (ADE) (see Bear and Cheng, 2010) Here we consider a spatially constant diffusion coefficient *D*. The ADE is solved using a standard particle tracking method employing 10^5 solute particles per each realization Darcy's flow field. As initial condition, solute particles are released at the center of the physical domain (see Figure 5b).

Our system output of interest is the spreading scale of the solute plume. Considering a solute plume transported within a Darcy's flow field, we evaluate the standard deviation of the plume particles positions along both principal coordinates, i.e.,

$$\sigma_x(t) = \sqrt{\left\{\left(x_p(t) - \left\{x_p(t)\right\}\right)^2\right\}}$$
$$\sigma_y(t) = \sqrt{\left\{\left(y_p(t) - \left\{y_p(t)\right\}\right)^2\right\}} \quad (8)$$

where x_p and y_p are the particle spatial coordinates at time t and {-} represents the average over the ensemble of particles released in each single Darcy's flow field within the Monte Carlo set. The spreading scales $\sigma_x(t)$ and $\sigma_y(t)$ are typically employed in setting characterized by the presence of a preferential mean flow direction aligned with one of the principal coordinates (e.g., given that x is aligned with the main flow direction, $\sigma_x(t)$ and $\sigma_y(t)$ characterize the longitudinal and transverse plume spreading, respectively, see e.g., Dentz et al., 2004). The latter condition does not hold in our setting and we characterize the solute spreading through the following spreading scale

$$\sigma(t;\theta,\omega) = \sqrt{\sigma_x(t;\theta,\omega)\sigma_y(t;\theta,\omega)} / \sqrt{2Dt} \quad (9)$$

Note that, we normalize the spreading scale by the purely diffusive counterpart.

Thus, given the system output of interest in Equation (9) and we proceed to evaluate its operational sensitivity with respect to the operational parameters $\theta_i = (d_1; d_2; r)$. The latter is conducted by following the evaluation workflow in Section 2.3.2 in which the variability of $Y(\boldsymbol{x})$ is sampled through the Monte Carlo realizations while that of the operational variables is sampled according the discrete probabilities in Table 1. On the base of the same set of simulations we follow the evaluation workflow in Section 2.1.2 to conduct the GSA presented in Section 2.1.1.

3.2.1 Global Sensitivity Analysis of the First Statistical Moment of Solute Spreading

In this Section we conduct the sensitivity analysis of $\sigma(t)$ by focusing on its first SM. Figure 6 depicts $AMAE_{\theta_i}^{Op}$ (continous curve) jointly with the confidence intervals $AMAE_{\theta_i}^{Op} \pm \sigma_{AMAE_{\theta_i}}$ (dashed curve) for θ_i =((a) d_1 (red); (b) d_2 (blue); r (green)) and few corresponding randomly chosen realizations of $AMAE_{\theta_i}(Y)$ (grey curves). Additionally, Figures 6d-e depict the unconditional first SM of the solute spreading $E_{\sigma(t;Y)}$ (black curve) and the conditional counterpart $E_{\sigma(t;Y|\theta_i)}$ for θ_i =(d_1 (red curves); (b) d_2 (blue curves); (c) r (green curves)) considering two randomly chosen realizations of $Y(\boldsymbol{x})$.

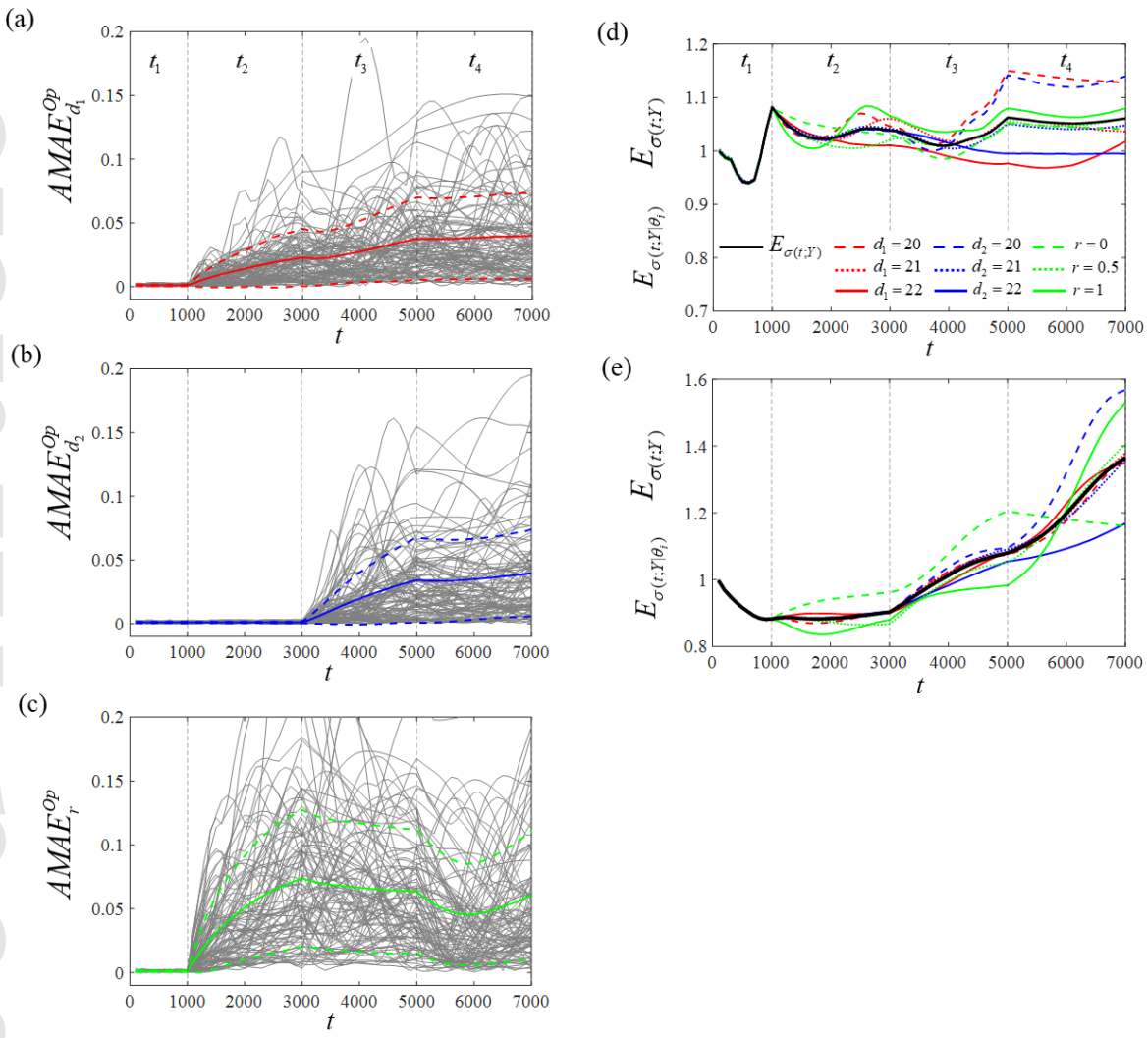

Figure 6. Moment-based operational GSA for the solute spreading considering the first SM, $AMAE^{Op}_{\theta_i}$ (continuous curves) and $AMAE^{Op}_{\theta_i} \pm \sigma_{AMAE_{\theta_i}}$ (dashed curves) for θ_i =((a) d_1 (red); (b) d_2 (blue); (c) r (green)), as a function of t. Sets of randomly chosen $AMAE_{\theta_i}(Y)$ are also depicted (grey curves). Vertical dashed lines delineate the period of wells activations. (d-e) Unconditional first SM of solute spreading $E_{\sigma(t;Y)}$ (black curve) and the conditional counterpart $E_{\sigma(t;Y|\theta_i)}$ for θ_i = (d_1 (red curves); d_2 (blue curves); r (green curves)) (see figure legend) considering two randomly chosen realizations of $Y(\boldsymbol{x})$.

Inspection of Figure 6a-c reveals that during t_1 the $AMAE_{\theta_i}^{Op}$ for the diverse θ_i (*i*) coincide and (*ii*) are slightly larger than zero. The same holds for the diverse $AMAE_{\theta_i}^{Op} \pm \sigma_{AMAE_{\theta_i}}$. Inspection of Figure 6d-e during t_1 highlights, for diverse realizations of $Y(\boldsymbol{x})$, the similarity between $E_{\sigma(t;Y)}$ and $E_{\sigma(t;Y|\theta_i)}$ for all the operational variables. During t_1 none of the latter impact the flow field that is experienced by the solute cloud as it is conveyed towards W1 and the only detected (small) discrepancies are those due to the difference in the (random) diffusive component of the solute motion across realizations of $Y(\boldsymbol{x})$ (i.e., in each single realization of $Y(\boldsymbol{x})$ the diffusive component of the particles motion is modelled as an independent white noise leading to very small discrepancies in the plume spreading scale across the realizations of $Y(\boldsymbol{x})$ during t_1, even though none of the θ_i is influencing the solute spreading).

Considering t_2, we note that $AMAE_{d_1}^{Op}$ and $\sigma_{AMAE_{d_1}}$ increase with time, see Figure 6a. The solute spreading is obviously impacted by the distance between W2 and the solute cloud during t_2. Thus we expect d_1 to be influential. In particular, solute spreading is dictated by the flow perturbations triggered by W2 that are experienced by the plume as it travels towards the latter (see also Figure 5b). Inspection of Figure 6d-c highlights how $E_{\sigma(t;Y|d_1)}$ tends to deviates from $E_{\sigma(t;Y)}$ as time passes within t_2 and in particular $E_{\sigma(t;Y|d_1)}$ decreases with d_1. Concurrently, the intensity of the flow perturbations dictating the plume spreading vary across the $Y(\boldsymbol{x})$ realizations (e.g., compare Figure 6d and Figure 6e), leading to increasing degree of variability of the first SM of solute spreading sensitivity with respect to d_1, i.e., $\sigma_{AMAE_{d_1}}$ increases during t_2. Considering d_2, the indices $AMAE_{d_2}^{Op}$ and $\sigma_{AMAE_{d_2}}$ have the same values recorded during t_1 since d_2 does not influence the W2-flow perturbations and thus solute spreading. The latter observation is corroborated by the equivalence of $E_{\sigma(t;Y|d_2)}$ and $E_{\sigma(t;Y)}$ in Figure 6d-e during t_2. Regarding r, inspection of Figure 6c highlights that $AMAE_r^{Op}$ and σ_{AMAE_r} increase during t_2, i.e., similarly to d_1, at diverse values of r correspond diverse intensity of the flow perturbations experienced by the solute cloud as it is conveyed towards W2 within each realization of $Y(\boldsymbol{x})$. Furthermore, inspection of Figures 6d-e reveals that $E_{\sigma(t;Y|r)}$ for $r = 0$ tends towards unity (see normalization in Equation 9) since diffusion is the only active transport mechanism. The variability in the dominant transport mechanism during t_2 ascribable to variations in r concur to the increase of $AMAE_r^{Op}$ while the growth of σ_{AMAE_r} is rooted in the influence of the heterogeneous arrangement of $Y(\boldsymbol{x})$ on the intensity of the solute spreading due to the advective component of transport (e.g., compare the sets of $E_{\sigma(t;Y|r)}$ in Figure 6d and Figure 6e).

Focusing on the time window of activation of W3, Figure 6a highlights that $AMAE_{d_1}^{Op}$ and $\sigma_{AMAE_{d_1}}$ slightly decrease at the beginning of t_3. After the activation of W3 the intensity of the flow perturbations experienced by the plume drastically diminish with respect to those at the end of t_2 leading to the decrease in the sensitivity of the first SM of spreading with respect to d_1 in comparison with the relevance of the latter during the ending of t_2. Moreover, in some realizations of $Y(\boldsymbol{x})$ the decrease in the flow perturbations experienced by the plume could promote the dominance of the diffusive transport (e.g., $E_{\sigma(t;Y|d_1)}$ for d_1 = 22 during t_3 in Figure 6d). The latter underpins the reduction of $\sigma_{AMAE_{d_1}}$ during the early stages of t_3. As time passes within t_3, the operational sensitivity of d_1 increases, i.e., as the plume is conveyed towards W3 the distance d_1 gains relevance due to its influence on the position of W3 with respect to the solute cloud and thus on the intensity of the flow perturbations associated with the activation of W3 that are experienced by the solute plume. Regarding $AMAE_{d_2}^{Op}$ and $\sigma_{AMAE_{d_2}}$ Figure 6b highlights that both increase during t_3, as expected (see discussion about $AMAE_{d_1}^{Op}$ and $\sigma_{AMAE_{d_1}}$ during t_1). At the same time, Figure 6c reveals that $AMAE_r^{Op}$ and σ_{AMAE_r} decrease within t_3 since the flow perturbations associated with W3 are not influenced by r. Nevertheless, the influence of r persist during t_3 due to its relevance in determining the spreading state at the beginning of t_3 (i.e., during t_3 the indices $AMAE_r^{Op}$ and σ_{AMAE_r} do not immediately drop to zero).

Lastly, Figures 6a-c reveal that all $AMAE_{\theta_i}^{Op}$ and $\sigma_{AMAE_{\theta_i}}$ decrease at the beginning of t_4 to then increase thereafter. Right after the activation of W4 the flow perturbations experienced by the solute plume drastically drop leading to the decrease in the sensitivity of the first SM of spreading with respect to all the operational parameters (see previous discussion of $AMAE_{d_1}^{Op}$ during t_3). At the same time, as the intensity of the flow perturbations acting on the plume decrease the relevance of diffusion increases (e.g., see $E_{\sigma(t;Y|d_1)}$ for d_1 = 22 and $E_{\sigma(t;Y|d_2)}$ for d_2 = 22 in Figure 6d and $E_{\sigma(t;Y|r)}$ for r = 0 in Figure 6d-e, during t_4) leading to the temporary drop in $\sigma_{AMAE_{\theta_i}}$ (see also previous discussion of $\sigma_{AMAE_{d_1}}$ during t_3). As the plume is conveyed towards W4 the solute spreading evolves differently depending on the value of the operational parameters (see the dispersion of $E_{\sigma(t;Y|\theta_i)}$ in Figure 6d-e during t_4) leading to the increase of $AMAE_{\theta_i}^{Op}$ as time passes during t_4 for all θ_i. Moreover, as the plume approaches W4 advection becomes the main transport mechanism and thus the differences in the heterogeneous arrangement of $Y(\boldsymbol{x})$ acquire relevance. This sustain the variability of the sensitivity of the first SM of the solute spreading across realizations of

$Y(\boldsymbol{x})$ (compare the relative dispersion of $E_{\sigma(t;Y|\theta_i)}$ around $E_{\sigma(t;Y)}$ during t_4 for the two realizations of $Y(\boldsymbol{x})$ in Figure 6d and Figure 6e), i.e., $\sigma_{AMAE_{\theta_i}}$ increases while time passes during t_4 for all θ_i.

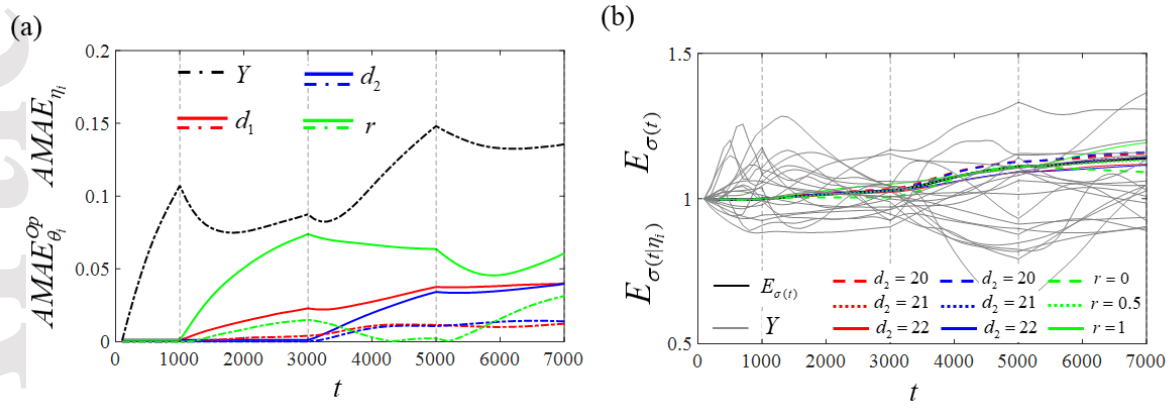

Figure 7. (a) Comparison of the GSA based on the first SM of the solute spreading. $AMAE_{\eta_i}$ (dash-dotted curve) for η_i = (Y (black); d_1 (red); d_2 (blue); r (green)), and $AMAE_{\theta_i}^{Op}$ (continuous curves) for θ_i = (d_1 (red); d_2 (blue); r (green)). (b) Unconditional $E_{\sigma(t)}$ (black curve) and conditional $E_{\sigma(t|\eta_i)}$ first SM of solute spreading η_i (d_1 (red curves); d_2 (blue curves); r (green curve); Y (grey curves)) for the diverse values of the operational variables (d_1; d_2; r) (see Figure legend and Table 1) and twenty random realizations of $Y(\boldsymbol{x})$.

Figure 7a depicts the results of the GSA grounded on $AMAE_{\eta_i}$ (dash-dotted curves) for η_i = (d_1 (red); d_2 (blue); r (green); Y (black)) versus t. For comparison, we also depict $AMAE_{\theta_i}^{Op}$ (continuous curves) for θ_i = (d_1 (red); d_2 (blue); r (green)). Moreover, Figure 7b depicts $E_{\sigma(t)}$ (black curve), $E_{\sigma(t|d_1)}$ (red curves), $E_{\sigma(t|d_2)}$ (blue curves), $E_{\sigma(t|r)}$ (green curves) and $E_{\sigma(t|Y)}$ (grey curves, twenty random realizations of $Y(\boldsymbol{x})$).

Inspection of Figure 7a reveals that, for the current setting, the first SM of solute spreading is less sensitive to the envisioned variability in the wells placements (d_1, d_2) and in the pumping rates (r) than to the lack of knowledge in the spatial arrangement of the hydraulic conductivity. Inspection of Figure 7b highlights that the variations in the wells characteristics η_i = (d_1, d_2, r) result in $E_{\sigma(t|\eta_i)}$ that are close to

$E_{\sigma(t)}$, i.e., the first SM of solute spreading conditional to the operational variables and the unconditional counterpart are dominated by the uncertainty in $Y(\boldsymbol{x})$. At the same time, the first SM of solute spreading is strongly impacted by the knowledge of $Y(\boldsymbol{x})$, i.e., the diverse $E_{\sigma(t|Y)}$ are strongly dispersed around $E_{\sigma(t)}$. In particular, the latter holds during t_1 and t_3 since the heterogeneity-driven advective spreading of solute is surely active (i.e., wells W1 and W3 are activated despite the value of r). This results in the increase of $AMAE_Y$ during t_1 and t_3, see Figure 7a. On the contrary, during t_2 and t_4 the wells W2 and W4 could be inactive in case of $r = 0$ allowing diffusion to homogenize the state of solute spreading across the ensemble of realizations of $Y(x)$. The latter results in the tendency of $AMAE_Y$ to decrease during the initial stages of t_2 and t_4. At the same time, in the case of $r = 0$ the corresponding $E_{\sigma(t|r)}$ tends towards unity during t_2 and t_4. This is in contrast with the tendency of $E_{\sigma(t|r)}$ to increase in case of $r > 0$. Thus, the possible prevalence of advective or diffusive solute transport (regulated by r during t_2 and t_4) enhances the influence of the ratio of pumping rate on the first SM of solute spreading during t_2 and t_4.

3.2.2 Global Sensitivity Analysis of the Second Statistical Moment of Solute Spreading

In this Section we investigate the sensitivity of $\sigma(t)$ by focusing on its second (centered) SM. Figure 8 depicts $AMAV_{\theta_i}^{Op}$ (continous curve) jointly with the confidence intervals $AMAV_{\theta_i}^{Op} \pm \sigma_{AMAV_{\theta_i}}$ (dashed curve) for θ_i =((a) d_1 (red); (b) d_2 (blue); (c) r (green)) and few corresponding randomly chosen realizations of $AMAV_{\theta_i}(Y)$ (grey curves). Additionally, Figure 8e depicts the unconditional second (centered) SM of solute spreading $V_{\sigma(t;Y)}$ (black curve) and the conditional counterpart $V_{\sigma(t;Y|\theta_i)}$ for θ_i = (d_1 (red curves); d_2 (blue curves); r (green curves)) considering one random realization of $Y(\boldsymbol{x})$.

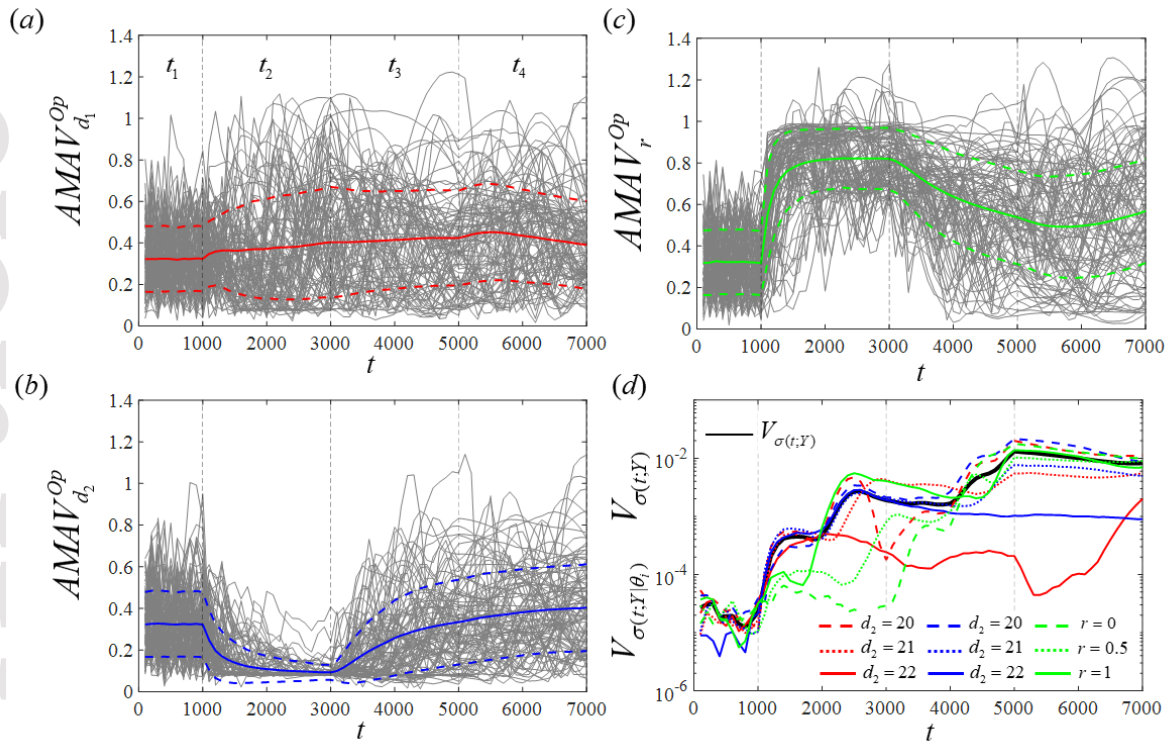

Figure 8. Moment-based operational GSA for the solute spreading considering the second (centered) SM, $AMAV^{Op}_{\theta_i}$ (continuous curves) and $AMAV^{Op}_{\theta_i} \pm \sigma_{AMAV_{\theta_i}}$ (dashed curves) for θ_i = ((a) d_1 (red); (b) d_2 (blue); (c) r (green)), as a function of t. Sets of randomly chosen $AMAV_{\theta_i}(Y)$ are also depicted (grey curves). Vertical dashed lines delineate the period of wells activations. (d-e) Unconditional second (centered) SM of solute spreading $V_{\sigma(t;Y)}$ (black curve) and the conditional counterpart $V_{\sigma(t;Y|\theta_i)}$ for θ_i = (d_1 (red curves); d_2 (blue curves); r (green curves)) (see figure legend) considering two randomly chosen realizations of $Y(\boldsymbol{x})$.

The comparison of results in Figures 8a-c and Figure 6a-c reveals an overall qualitative agreement of the operational GSA grounded on $AMAV_{\theta_i}(Y) \pm \sigma_{AMAV_{\theta_i}}$ and $AMAE_{\theta_i}(Y) \pm \sigma_{AMAE_{\theta_i}}$ for the diverse θ_i during t_1, t_3 and t_4. In particular, during t_1 the values of $AMAV^{Op}_{\theta_i}$ and $AMAV_{\theta_i}(Y) \pm \sigma_{AMAV_{\theta_i}}$ associated with the diverse θ_i coincide and are larger than zero. These results confirm the role of the discrepancies in $\sigma(t)$ due to the randomness in the diffusive component of solute motion (see previous discussion of $AMAE_{\theta_i}(Y) \pm \sigma_{AMAE_{\theta_i}}$ during t_1 in Section 3.2.1). Note that, Figure 8*d* reveals that during t_1 $V_{\sigma(t;Y|\theta_i)}$ (for

all θ_i) and $V_{\sigma(t;Y)}$ are (*i*) small (discrepancies in the plume spreading due to diffusive motion are small, this is expected given the high number of solute particles employed to resolve each plume) and (*ii*) their discpepancies are of a similar order of magnitude of $V_{\sigma(t;Y)}$, which leads not negligibles values of $AMAV_{\theta_i}(Y) \pm \sigma_{AMAV_{\theta_i}}$.

The most striking qualitative difference between Figure 8 and Figure 6 is noted during t_2 when $AMAV_{d_2}(Y) \pm \sigma_{AMAV_{d_2}}$ decreases while $AMAE_{d_2}(Y) \pm \sigma_{AMAE_{d_2}}$ remains more or less constant. Inspection of Figure 8d reveals that, $V_{\sigma(t;Y)}$ increases during t_2 under the influence of variations in d_1 and r which affect the intensity of the flow perturbations induced by W2. Meanwhile, variations in d_2 do not affect the behavior of W2 and thus the conditional second (centered) SM $V_{\sigma(t;Y|d_2)}$ shows a decreasing degree of (relative) dispersion around $V_{\sigma(t;Y)}$ during t_2, i.e., $AMAV_{d_2}(Y) \pm \sigma_{AMAV_{d_2}}$ decreases during t_2.

Figure 9a depicts the results of the GSA as grounded on $AMAV_{\eta_i}$ (dash-dotted curves) for η_i = (d_1 (red); d_2 (blue); r (green); Y (black)) versus t. For comparison, we also depict $AMAV_{\theta_i}^{Op}$ (continuous curves) for θ_i = (d_1 (red); d_2 (blue); r (green)). Moreover, Figure 9b depicts $V_{\sigma(t)}$ (black curve), $V_{\sigma(t|d_1)}$ (red curves), $V_{\sigma(t|d_2)}$ (blue curves), $V_{\sigma(t|r)}$ (green curves) and $V_{\sigma(t|Y)}$ (grey curves, twenty random realizations of $Y(\mathbf{x})$).

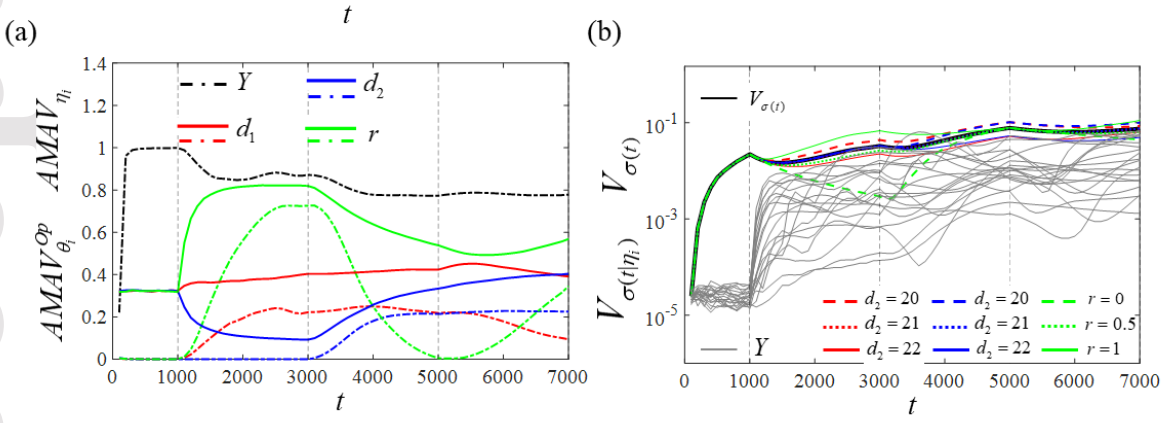

Figure 9. (a) Comparison of the GSA based on the second (centered) SM of the solute spreading. $AMAV_{\eta_i}$ (dash-dotted curve) for η_i = (Y (black); d_1 (red); d_2 (blue); r (green)), and $AMAV_{\theta_i}^{Op}$ (continuous curves)

for θ_i = (d_1 (red); d_2 (blue); r (green)). (b) Unconditional $V_{\sigma(t)}$ (black curve) and conditional $V_{\sigma(t|\eta_i)}$ second (centered) SM of solute spreading η_i (d_1 (red curves); d_2 (blue curves); r (green curve); Y (grey curves)) for the diverse values of the operational variables (d_1; d_2; r) (see Figure legend and Table 1) and twenty random realizations of $Y(\boldsymbol{x})$.

Inspection of Figure 9a reveals (*i*) the dominance of $AMAV_Y$ and (*ii*) an high degree of similarity in the trends of $AMAV_{\theta_i}^{Op}$ and $AMAV_{\eta_i}$ for $\eta_i = \theta_i = (d_1; d_2; r)$. whereas $AMAV_{\eta_i}$ is generally lower than its $AMAV_{\theta_i}^{Op}$ counterpart. These observations are in agreement with the analysis of $AMAE_{\theta_i}^{Op}$ and $AMAE_{\eta_i}$ in Section 3.2.1.

Comparison of $AMAV_{\eta_i}$ (see Figure 9a) and $AMAE_{\eta_i}$ (see Figure 7a) reveals a general higher degree of sensitivity of the solute spreading with respect to $\eta_i = (d_1; d_2; r)$ when we consider the second (centered) SM. At the same time, we note an overall similarity in the trends of $AMAV_{\eta_i}$ and $AMAE_{\eta_i}$ for $\eta_i = (d_1; d_2; r)$ over time suggesting a coherence in the influence of the latter on the diverse SM of the solute spreading. We note the general similarity in the dispersion of $V_{\sigma(t|\eta_i)}$ around $V_{\sigma(t)}$ in Figure 9b and the counterpart for $E_{\sigma(t|\eta_i)}$ around $E_{\sigma(t)}$ in Figure 7b while considering $\eta_i = (d_1; d_2; r)$. Furthermore, we note that during t_2 and t_4 the value of $V_{\sigma(t|r)}$ for $r = 0$ (dashed green curve) decreases since diffusion tends to homogenize the solute spreading state across the diverse combinations of the remaining variables $\eta_i = (d_1; d_2; r)$ given $Y(\boldsymbol{x})$ (see also $E_{\sigma(t|r)}$ for $r = 0$ in Figure 7b). This results in the increase of $AMAV_r$ during t_2 and less markedly during t_4.

At the same time, Figure 9a reveals that the second (centered) SM of the solute spreading becomes less sensitive to the lack of knowledge about $Y(\boldsymbol{x})$ as time passes, i.e., $AMAV_Y$ decreases over time. This is in contrast with the behavior of $AMAE_Y$ (see Figure 7a and previous discussion). Figure 9b reveals that the knowledge of $Y(\boldsymbol{x})$ corresponds to a drastic reduction of the second (centered) SM of the solute spreading during t_1, i.e., $V_{\sigma(t|Y)}$ is order of magnitude smaller than the unconditional counterpart $V_{\sigma(t)}$. The latter results in $AMAV_Y \approx 1$ during t_1. Concurrently, $V_{\sigma(t)}$ practically coincides with $V_{\sigma(t|\eta_i)}$ for $\eta_i = (d_1; d_2; r)$ during t_1 since variations in none of these variables affect the spreading of solute, i.e., $AMAV_{\eta_i} \approx 0$ for η_i

= (d_1; d_2; r) during t_1. As time passes, $V_{\sigma(t|Y)}$ increases since it reflects the variability in the solute spreading due to the variations in η_i = (d_1; d_2; r) given the knowledge of $Y(\boldsymbol{x})$. Yet, over time the knowledge of $Y(\boldsymbol{x})$ becomes less relevant to the reduction of the uncertainty about the spreading of solute, i.e., $V_{\sigma(t)}$ and $V_{\sigma(t|Y)}$ become similar since the $V_{\sigma(t)}$ is more influenced by the assumed variability in (d_1; d_2; r) rather than the lack of knowledge about hydraulic conductivity field.

4. Conclusions

We highlight the role of recognizing the different nature, i.e., stochastic (due to a lack of knowledge) or operational (due to diverse design/interventional choices), of the variability ascribed to model variables in the context of global sensitivity analysis (GSA). Contrary to standard GSA methodologies, we propose to disentangle the variability of the stochastic variables from that of the operational ones when assessing the sensitivity of a model output of interest with respect to the latter. We accomplish the latter by evaluating the sensitivity of a model output with respect to the operational variables for a random realization of the stochastic variables. Due to the lack of knowledge about the stochastic variables the latter evaluation is repeated for diverse realizations. This leads to the probability distribution of the (random) model output sensitivity with respect to an operational variable for which we evaluate (*a*) the expected value and (*b*) the standard deviation (both evaluated with respect to the stochastic variables) in order to characterize the GSA of a model output with respect to operational variables in the presence of stochastic ones. In this context, (*a*) and (*b*) quantify the average degree of sensitivity of a model output with respect to an operational variable and its degree of variability across the diverse realizations of the stochastic variables, respectively. We term this GSA approach as the operational GSA. We recall here that in standard GSA the evaluation of the model output sensitivity is conducted by blending the stochastic and operational variabilities during the evaluation of selected GSA indices. Thus, while the former addresses questions like 'Is the lack of knowledge about the stochastic variables or the variations in the operational parameters the most influential aspect to the output?' the second focuses on 'What is the relevance of the diverse operational variables on the output, in the presence of a lack of knowledge about stochastic system variables?'. This means that the two strategies are complementary, as it is typically the case for different GSA.

The moment-based GSA of Dell'Oca et al., (2017) is adopted here as the reference GSA strategy, while the idea and developments proposed are fully compatibles with other GSA techniques. A comparison of the results grounded on the standard moment-based GSA and its operational counterpart for two exemplifying models highlights different degrees of sensitivity depending on: (*a*) the statistical moment of

interest of the model output and (*b*) the adoption of the standard moment-based GSA or the operational counterpart. Our set of results highlights the importance of adopting the operational GSA perspective to fully characterize the sensitivity of model output to operational variables in the presence of stochastic variables.

Author contributions

A.D.: Conceptualization, Formal Analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing, Software.

Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

A.D. acknowledges the funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie Grant Agreement No. 895152 (MixUQ).

Open Research

The software to conduct the presented operational global sensitivity analysis is available at doi: 10.17632/2bkc7gp2sd.3. The software can be run on a standard PC within the Matlab environment.

References

Alminagorta, O., Rosenberg, D. E., & Kettenring, K. M. (2016). Systems modeling to improve the hydro-ecological performance of dikedwetlands. Water Resources Research. 52, 7070–7085, doi:10.1002/2015WR018105

Baroni, G., & Francke, T. (2020). An effective strategy for combining variance- and distribution-based global sensitivity analysis. Environmental Modeing & Software, 134, 104851. 10.1016/j.envsoft.2020.104851

Bear, J., & Cheng, A.H.D. (2010). Modelling groundwater flow and contaminant transport. Springer Science & Business Media. Berlin. https://doi.org/10.1007/978-1-4020-6682-5

Borgonovo, E. (2007). A new uncertainty importance measure. Reliability Engineering & System Safety. 92, 771-784. https://doi.org/10.1016/j.ress.2006.04.015.

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

Campolongo, F., Saltelli, A., & Cariboni, J. (2011). From screening to quantitative sensitivity analysis. A unified approach. Computer Physics Communication. 182, 978–988. https://doi.org/10.1016/j.cpc.2010.12.039

Cantoni, B., Penserini, L., Vries, D., Dingemans, M. M. L., Bokkers, B. G. H., Turolla, A., Smeets, P. W. M. H., & Antonelli, M. (2021). Development of a quantitative chemical risk assessment (QCRA) procedure for contaminants of emerging concern in drinking water supply. Water Research, 194, 116911. https://doi.org/10.1016/j.watres.2021.116911

Chen, X., Zhou, X., Xin, K., Liao, Z., Yan, H., Wang, J., & Tao, T. (2022). Sensitivity-oriented clustering method for parameter grouping in water network model calibration. Water Resources Research, 58, e2021WR031206. https://doi.org/10.1029/2021WR031206

Dell'Oca, A., Riva, M., & Guadagnini, A. (2017). Moment-based Metrics for Global Sensitivity Analysis of Hydrological Systems. Hydrological and Earth Systems Science. 21, 6219–6234. https://doi.org/10.5194/hess-21-6219-2017

Dell'Oca, A., Riva, M., & Guadagnini, A. (2020). Copula density-driven metrics for sensitivity analysis: theory and application to flow and transport in porous media. Advances in Water Research. 145, 103714, 1-11. https://doi.org/10.1016/j.advwatres.2020.103714

Dentz, M., Cortis, A., Scher, H., & Berkowitz, B. (2004). Time behavior of solute transport in heterogeneous media: transition from anomalous to normal transport. Advances in Water Resources. 27, 155-173. https://doi.org/10.1016/j.advwatres.2003.11.002

Dobson, B., Watson-Hill, H., Muhandes, S., Borup, M., & Mijic, A. (2022). A reduced complexity model with graph partitioning for rapid hydraulic assessment of sewer networks. Water Resources Research, 58, e2021WR030778. https://doi.org/10.1029/2021WR030778

Elshafei, Y., Tonts, M., Sivapalan, M., & Hipsey, M. R. (2016). Sensitivity of emergent sociohydrologic dynamics to internal system properties and external sociopolitical variables: Implications for water management. Water Resources Research. 52, 4944–4966. doi:10.1002/2015WR017944.

Haghnegahdar, A., Razavi, S. (2017). Insights into sensitivity analysis of Earth and environmental systems models: on the impact of parameter perturbation scale. Environmental Modeling & Software. 95, 115–131. https://doi.org/10.1016/j.envsoft.2017.03.031.

Hamilton, A. L., Characklis, G. W., & Reed, P. M. (2022). From stream flows to cash flows: Leveraging Evolutionary Multi-Objective Direct Policy Search to manage hydrologic financial risks. Water Resources Research, 58, e2021WR029747. https://doi.org/10.1029/2021WR029747

Hariz, H.B., Lawton, R.J., & Craggs, R.J. (2023). Effects of operational parameters on the performance of unialgal Oedogonium sp. filamentous algae nutrient scrubbers under controlled environmental conditions. Journal of Environmental Management, 326. https://doi.org/10.1016/j.jenvman.2022.116705.

Homma, T., & Saltelli, A. (1996). Importance measures in global sensitivity analysis of nonlinear models. Reliability Engineering & System Safety. 52, 1–17. https://doi.org/10.1016/09518320(96)00002-6

Hou, Y., Yang, J., Russoniello, C. J., Zheng, T., Wu, M.-l., & Yu, X. (2022). Impacts of coastal shrimp ponds on saltwater intrusion and submarine groundwater discharge. Water Resources Research, 58, e2021WR031866. https://doi.org/10.1029/2021WR031866

Jazayeri, A., Werner, A. D., Wu, H., & Lu, C. (2021). Effects of river partial penetration on the occurrence of riparian freshwater lenses: Experimental investigation. Water Resources Research, 57, e2021WR029728. https://doi.org/10.1029/2021WR029728

Karimi, T., Reed, P., Malek, K., & Adam, J. (2022). Diagnostic framework for evaluating how parametric uncertainty influences agro-hydrologic model projections of crop yields under climate change. Water Resources Research, 58, e2021WR031249. https://doi.org/10.1029/2021WR031249

Krzykacz-Hausmann (2001). Epistemic sensitivity analysis based on the concept of entropy. Proceedings of SAMO 2001, CIEMAT, Madrid (2001), pp. 31-35.

Lie, K.-A. (2019). An Introduction to Reservoir Simulation Using MATLAB/GNU Octave, User Guide for the MATLAB Reservoir Simulation Toolbox (MRST). Cambridge University Press. https://doi.org/10.1017/9781108591416

Lo Piano, S., Parenti, A., & Guerrini, L (2022). Uncertainty appraisal provides useful information for the management of a manual grape harvest. Biosystems Engineering. 219, 259-267. https://doi.org/10.1016/j.biosystemseng.2022.05.006

Morris, M.D. (1991). Factorial sampling plans for preliminary computational experiments. Technometrics. 33, 161-174. https://doi.org/10.1080/00401706.1991.10484804.

Niu, Y., & Shah, F. A. (2021). Economics of optimal reservoir capacity determination, sediment management, and dam decommissioning. Water Resources Research, 57, e2020WR028198. https://doi.org/10.1029/2020WR028198

Owen, A. (1994). Lattice sampling revisited: monte carlo variance of means over randomized orthogonal arrays. The Annual Statistics. 22, 930–945. https://doi.org/10.1214/aos/1176325504.

Pastres, R., Chan, K., Solidoro, C., & Dejak, C. (1999). Global sensitivity analysis of a shallow-water 3D eutrophication model. Computer Physics Communication. 117, 62e74. https://doi.org/10.1016/S0010-4655(98)00164-7

Patil, R., Wei, Y., Pullar, D., & Shulmeister, J. (2022). Sensitivity of streamflow patterns to river regulation and climate change and its implications for ecological and environmental management. Journal of Environmental Management, 319, 115680. https://doi.org/10.1016/j.jenvman.2022.115680

Peña, A., Rovira-Val, M. R., & Mendoza, J. M. F. (2022). Life cycle cost analysis of tomato production in innovative urban agriculture systems. Journal of Cleaner Production, 367, 133037. https://doi.org/10.1016/j.jclepro.2022.133037

Pianosi, F., Beven, K., Freer, J., Wall, J.H., Rougier J., Stephenson, D.B., & Wagener, T. (2016). Sensitivity analysis of environmental models: A systematic review with practical workflow. Environmental Modeling & Software, 79, 214-232. https://doi.org/10.1016/j.envsoft.2016.02.008

Pianosi, F., & Wagener, T. (2018). Distribution-based sensitivity analysis from a generic input-output sample. Environmental Modeing & Software, 108, 197-207. https://doi.org/10.1016/j.envsoft.2018.07.019

Razavi, S., & Gupta, H.V. (2015). What do we mean by sensitivity analysis? The need for comprehensive characterization of ‘‘global’’ sensitivity in Earth and Environmental systems models. Water Resources Research, 51, 3070-3092. doi:10.1002/2014WR016527

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

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

Razavi, S., & Gupta, H.V. (2019). A multi-method Generalized Global Sensitivity Matrix approach to accounting for the dynamical nature of earth and environmental systems models. Environmental Modeling & Software, 114, 1-11. https://doi.org/10.1016/j.envsoft.2018.12.002

Razavi, S., Jakeman, A., Saltelli, A., Prieur, C., Iooss, B., Borgonovo, E., Plischke, E., Lo Piano, S., Iwanaga, T., Becker, W., Tarantola, S., Guillaume, J.H.A., Jakeman, J., Gupta, H., Melillo, N., Rabitti, G., Chabridon, V., Duan, Q., Sun, X., Smith, S., Sheikholeslami, R., Hosseini, N., Asadzadeh, M., Puy, A., Kucherenko, S., & Maier, H.R. (2021). The future of sensitivity analysis: an essential discipline for systems modeling and policy support. Environmental Modeling & Software, 137, 104954. 10.1016/j.envsoft.2020.104954

Riechel, M., Matzinger, A., Pawlowsky-Reusing, E., Sonnenberg, H., Uldack, M., Heinzmann, B., Caradot, N., von Segger, D., & Rouault, P. (2016). Impacts of combined sewer overflows on a large urban river e Understanding the effect of different management strategies. Water Research, 105, 264-273. https://doi.org/10.1016/j.watres.2016.08.017

Risken, H., 1989. The Fokker-Planck Equation: Methods of Solution and Applications. Springer. Berlin. https://doi.org/10.1007/978-3-642-61544-3

Saltelli, A., Jakeman, A., Razavi, S., & Wu, Q. (2021). Sensitivity analysis: A discipline coming of age. Environmental Modeling & Software, 146, 105226. https://doi.org/10.1016/j.envsoft.2021.105226

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S. (2008). Global sensitivity analysis. The primer. Wiley.

Schalko, I., & Boes, R. M. (2021). Effect of water withdrawal on the appearance and sound level of waterfalls. Water Resources Research, 57, e2021WR030980. https://doi.org/10.1029/2021WR030980

Sheikholeslami, R., & 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.

Sobol', I.M. (1993). Sensitivity analysis for non-linear mathematical models. Mathematical Modelling and Computational Experiment, Translated from Russian: I. M. Sobol', Sensitivity estimates for nonlinear mathematical models. Matematicheskoe Modelirovanie 2, 407–414 (1990) 112–118 1.

Sobol', I.M., & Kucherenko, S. (2009). Derivative based global sensitivity measures and their link with global sensitivity indices. Mathematics and Computers in Simulation. 79, 3009-3017. https://doi.org/10.1016/j.matcom.2009.01.023

Thompson, M., Moussavi, S., Li, S., Barutha, P., & Dvorak, B. (2022). Environmental life cycle assessment of small water resource recovery facilities: comparison of mechanical and lagoon systems. Water Research, 215, 118234. https://doi.org/10.1016/j.watres.2022.118234

Wang, X. (2021). Uncertainty quantification and global sensitivity analysis for transient wave propagation in pressurized pipes. Water Resources Research, 57, e2020WR028975. https://doi.org/10.1029/2020WR028975

Wu, Z., Ma, B., Wang, H., Hu, C., Lv, H., & Zhang, Z. (2021). Identification of sensitive parameters of urban flood model based on artificial neural network. Water Resource Mangement, 35, 2115-2128. https://doi.org/10.1007/s11269-021-02825-3

Yan, M., Lu, C., Werner, A. D., & Luo, J. (2021). Analytical, experimental, and numerical investigation of partially penetrating barriers for expanding island freshwater lenses. Water Resources Research, 57, e2020WR028386. https://doi.org/10.1029/2020WR028386

Pianosi, F., & Wagener, T. (2015). A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. Environmental Modelling & Software, 67, 1-11. https://doi.org/10.1016/j.envsoft.2015.01.004