An effective strategy for combining variance- and distribution-based global sensitivity analysis

Published Version:
Baroni G., Francke T. (2020). An effective strategy for combining variance- and distribution-based global sensitivity analysis. ENVIRONMENTAL MODELLING & SOFTWARE, 134, 1-14 [10.1016/j.envsoft.2020.104851].

Availability:
This version is available at: https://hdl.handle.net/11585/772812 since: 2020-09-28

Published:
DOI: http://doi.org/10.1016/j.envsoft.2020.104851

An effective strategy for combining variance- and distribution-based global sensitivity analysis.

Gabriele Baroni[1]*, Till Francke[2]

(1) Department of Agricultural and Food Sciences, University of Bologna, Viale Fanin 50, 40127 Bologna, Italy

(2) Institute of Environmental Science and Geography, University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany

*Corresponding author: Email: g.baroni@unibo.it; Tel: +39 051 20 9 6322

Abstract

We present a new strategy for performing global sensitivity analysis capable to estimate main and interaction effects from a generic sampling design. The new strategy is based on a meaningful combination of variance- and distribution-based approaches. The strategy is tested on four analytic functions and on a hydrological model. Results show that the analysis is consistent with the state-of-the-art Saltelli/Jansen formula but to better quantify the interaction effect between the input factors when the output distribution is skewed. Moreover, the estimation of the sensitivity indices is much more robust requiring a smaller number of simulations runs. Specific settings and alternative methods that can be integrated in the new strategy are also discussed. Overall, the strategy is considered as a new simple and effective tool for performing global sensitivity analysis that can be easily integrated in any environmental modelling framework.

Key words

global sensitivity analysis, variance, distribution, generic sampling design

Highlights

- A new strategy to perform global sensitivity analysis is developed
- Variance- and distribution-based approaches are combined in a meaningful way
- Main and interactions effect are estimated from a generic sampling design
- The new strategy converges faster than Saltelli/Jansen formula

Software availability

The analysis has been performed with the statistical software R 3.3.x (R Core Team, 2019) and the Saltelli/Jansen analysis using the package "sensitivity" (Iooss et al., 2019). The simulations with the hydrological model have been conducted using the package "hydromad" (Andrews et al., 2011). An example script on how to implement the new Combined Variance- and Distribution-based CVD strategy is freely available under https://github.com/baronig/GSA-cvd.

1 Introduction

Global sensitivity analysis (GSA) refers to a group of diagnostic modelling tools developed to study how the uncertainty in the output of a mathematical model can be apportioned to the uncertainty in the input factors (Saltelli et al., 2000). In this context, the term factor is interpreted in a broad sense of anything that can be subject to some degree of uncertainty in the model, e.g., parameters, input, boundary conditions or model structure (Ratto et al., 2007).

In contrast to the local one-at-the-time approach, where one single factor is perturbed while keeping the other fixed, GSA approaches are applicable independently of the characteristics of the input-output response function, they cover the entire input space and they identify non-linearity and interactions between the factors (Saltelli et al., 2008). For these reasons, they have been recognized as a fundamental analysis to support model understanding and improvements in many applications (Baroni et al., 2018; Borgonovo et al., 2017; Demirel et al., 2018; Guse et al., 2016; Haghnegahdar et al., 2017; Pianosi and Wagener, 2016; Reusser et al., 2011; Rosolem et al., 2012; Savage et al., 2016; Schürz et al., 2019; Tang et al., 2007; Xie et al., 2017) and they should be preferred to local approaches to avoid a perfunctory analysis (Saltelli et al., 2019; Saltelli and Annoni, 2010).

Several GSA methods have been developed and we refer to review papers and text books for an overview (Iooss and Lemaître, 2015; Pianosi et al., 2016; Razavi and Gupta, 2015; Saltelli et al., 2008; Song et al., 2015; Wagener and Pianosi, 2019; Wei et al., 2015). Here, we narrow the discussion to two probabilistic approaches based on Monte Carlo simulations: the variance-based approaches and the distribution-based (also called 'density-based' or 'moment independent') approaches.

Sensitivity analysis based on variance measures has been introduced in the early 70ies (Bier, 1982; Cukier et al., 1973; Iman, 1987). Major contributions on the approach, however, are attributed to the Russian mathematician I.M. Sobol’ who generalized the approach and provided a straightforward Monte Carlo-based implementation (Sobolʹ, 2001). At present, the most widely used variance-based measures are the so-called Sobol’ indices, and in particular Sobol’ first order sensitivity measure (or main effect), together with the total sensitivity indices (or total effect), introduced by Homma and Saltelli (1996).

The use of these indices enjoyed success among practitioners probably due to a clear interpretation of their meaning (Saltelli et al., 2008). Specifically, as described in Ratto et al. (2007), high values in the *main effect* represent the factors that, when constrained at their true value, would reduce the uncertainty in the model output the most and, therefore, make the model inference more robust. This type of analysis supports the prioritization in model improvements. Thus, it has been identified in literature with the term “factor prioritization”. On the contrary, low values for *total effect* identify those factors that have an irrelevant contribution to the uncertainty in the output and, therefore, can be constrained to an arbitrary value within their range of uncertainty, e.g., supporting model simplification. This type of analysis has been identified with the term “factor fixing”.

It should also be noted that the two analyses defined above can be performed independently, i.e., either estimating the main effect for factor prioritization or the total effect for factor fixing. However, it has been underlined that the differences between total and main effect reveals interactions among the factors (Saltelli et al., 2008). This additional information is very important when the modelers are interested to know if the specific factor is identifiable by, e.g., calibration (Ghasemizade et al., 2017; Guillaume et al., 2019). Namely, a factor that shows high interactions is likely not identifiable. So, for sake of clarity, we term this third type of analysis “factor identification” that, in contrast to “factor prioritization” and “factor fixing”, can be achieved only by estimating both indices.

Despite its wide use, it has been argued that variance-based approaches rely too heavily on the assumption that the variance is sufficient to describe the uncertainties and sensitivities encountered (Park and Ahn, 1994). The variance measure is ill-suited particularly to measure the dispersion of a variable with a heavy-tail or a multimodal distribution, or which contains some outliers (Auder and Iooss, 2009). To overcome this limitation, some approaches have been developed based on the idea of comparing the entire probabilistic distributions of the model output rather than the variance. The first methods have been derived based on the entropy measures (Liu et al., 2006; Park and Ahn, 1994). Borgonovo (2007) developed an important measure based on density functions. More recently, Pianosi and Wagener (2015) suggested the use of the Kolmogorov-Smirnov test (Kolmogorov, 1933). Generalization of the methods have been also presented for comparing the entire distributions (Veiga, 2015) or targeting some of their statistical moments (Dell'Oca et al., 2017).

These distribution-based methods have also been used in several studies (Borgonovo et al., 2011; Castaings et al., 2012; Fox et al., 2016; Gillies et al., 2016; Hosseini et al., 2018; Pianosi and Wagener, 2016; Pilz et al., 2017; Schürz et al., 2019; Sedighian et al., 2015) and they have showed to converge faster than variance-based methods (Pianosi and Wagener, 2015; Zadeh et al., 2017). However, these methods focus on estimating only one single effect, and separating main, total effect and interactions is not targeted. For this reason, the modelers remain without the identification of important features that can be achieved in global sensitivity analysis (Saltelli and Tarantola, 2002).

To overcome this limitation and exploit the advantages of both variance- and distribution-based approaches, their combined use has been suggested and tested in literature (Borgonovo et al., 2017; Massmann and Holzmann, 2012; Pappenberger et al., 2008). Most of the studies emphasized the complementarity of the different methods. However, they also highlighted the difficulties to directly compare the different indices because they are based on different quantities, they explore different ranges in the input-output space and, thus, they carry different information (Borgonovo and Tarantola, 2008; Mora et al., 2019).

In this study we develop and test an effective strategy to combine variance- and distribution-based sensitivity analysis. We refer to this combined strategy with the term CVD strategy. This strategy has been developed to take advantages of the two approaches and to allow a meaningful combined use of the different indices. The paper is structured as follows: in section 2, we first revise variance- and distribution-based approaches to better identify similarities and complementarities. This provides the basis for the development of the new CVD strategy. This new strategy is tested with four analytic functions (section 3) and on a hydrological model (section 4). In all tests, the strategy is compared to the state-of-the-art Saltelli/Jansen formula for estimating both main and total effect. The discussion is enriched based on specific settings and possible improvements (section 5). Conclusions are reported in section 6.

2 Methods

2.1 Variance-based approaches

We start considering a numerical deterministic model that can be written in the form:

$$y = g(x_i) \tag{1}$$

where x_i are the input factors with $i = 1 \ldots k$, and k the number of factors, g is the generic numerical model and y the output of the model. x_i can be regarded as any type of input factor (parameters, model structure, input and boundary conditions). Considering that all types of factors can be associated to a scalar discrete value (Baroni and Tarantola, 2014; Lilburne and Tarantola, 2009; Plischke et al., 2013), we proceed also here assuming that the factors x_i are a scalar input (e.g., parameters).

We now consider that the values of x_i are not known and their uncertainty can be described by their probability distributions P(x_i), e.g., fully defined by mean and standard deviation in case of Gaussian distributions. We underline that this first task (distribution assignment) is a crucial step in ensuring the quality and consistency of the results (Baroni et al., 2017; Haghnegahdar and Razavi, 2017; Plischke et al., 2013; Shin et al., 2013). Then, the distributions are sampled N times creating a N x k matrix X. The model can be now run in a Monte Carlo approach as follows:

$$Y = g(X) \tag{2}$$

The probabilistic distribution of the vector Y can be now defined as P(Y) and can be seen as the uncertainty in the model output by propagating the uncertainty in k input factors.

Variance-based approaches implicitly assume that the variance (the second-moment), here identified with the term V[·], is sufficient to describe the probabilistic distribution P(Y). Thus, a generic uncertainty importance measure of the factor i in explaining the uncertainty in the model output can be introduced as the reduced variance one would achieve by fixing one source of uncertainty as follow, expressed in relative terms:

$$\frac{V\left(Y|X_{i=E(X_i)}\right)}{V(Y)} \tag{3}$$

where E[·] indicates the mean operator and $V\left(Y|X_{i=E(X_i)}\right)$ indicates the conditional variance in the model output Y when the factor i is fixed to its mean E(X_i). Dividing this quantity over the unconditional variance V(Y) leads to an index that quantifies the fraction of remaining variance in the model output when the correct value of factor i is known.

Besides possibly having the problem of an index bigger than one (conditional variance > unconditional variance), one should recognize that E(X_i) could be not necessarily the true (i.e., optimal) value. Thus, this measure would not be a good estimation of the importance of the factor (Iman, 1987). For this reason, it has been suggested to calculate the conditional variance V($Y|X_i$) based on x_i fixed to r values in the range P(X_i) and to calculate their mean as follow:

$$\frac{E[V(Y|X_i)]}{V(Y)} \qquad (4)$$

In the simplest case, the number of conditional values r could be limited to the two extreme values within the range P(X_i) (Borgonovo, 2010; Francke et al., 2018; Iman, 1987). Most preferable, however, the values should cover the entire range in the distributions. Now, when this importance measure is *small*, x_i is an important factor. Considering that the model output variance can be decomposed as the sum of the variance of the conditional expectations E[V($Y|X_i$)] and the residual term V[E($Y|X_i$)] (Mood et al., 1974):

$$V(Y) = \mathrm{E}[\mathrm{V}(Y|X_i)] + \mathrm{V}[\mathrm{E}(Y|X_i)] \qquad (5)$$

the importance measure can be formulated in a complementary way as:

$$S_i = \frac{V[E(Y|X_i)]}{V(Y)} \qquad (6)$$

Now when this importance measure is *high*, x_i is an important factor. This expression is usually referred to as Sobol' first order sensitivity measure (or main effect) and it is indicated in literature with the symbol "S_i". As discussed in the introduction, this index represents the mean variance that one would reduce if the factor is known. For this reason, high values in the main effect identify the most important factors that should be considered to decrease the uncertainty in the model (or improve the model performance). Thus, this index supports the so-called factor prioritization.

It has been noted that the factor i could have a direct effect on Y or its importance could be related to the effect that the factor has due to interactions with other factors. For this reason, another index has been proposed as follows:

$$T_i = \frac{E[V(Y|X_{\sim i})]}{V(Y)} \qquad (7)$$

where $V(Y|X_{\sim i})$ indicates the variance of Y fixing all factors but not i. This terms is usually referred as total sensitivity index (or total effect) as introduced by Homma and Saltelli (1996). Please note that the index is generally denoted with the symbol "S_i^T", but for simplicity we adopt the notation T_i as suggested by Glen and Isaacs (2012). As discussed in the introduction, this index represents the main effect of the factor and its contribution based on interactions with other factors. Low values of T_i identify factors that are not important and can be fixed. Thus, this index can serve for factor fixing.

Finally, we explicitly derive the interaction term as a simple difference between the two indices:

$$I_i = T_i - S_i \tag{8}$$

This index is used to quantify the effect of the factor i due to interactions with the other factors and it is used in combination of the main effect to understand a factors' identifiability, i.e., if I_i is small and S_i is high, the factor is likely identifiable.

2.2 Estimation of first and total sensitivity index

In principle, the computational cost to estimate S_i, assuming the same number of samples n for approximating the unconditional and conditional distributions is:

$$N = (k+1) \cdot r \cdot n \tag{9}$$

where N is the required number of model evaluations, k is the number of factors i, r is the number of conditional values and n the sample size to approximate the distributions (Saltelli et al., 2008). This tailored sampling design is usually referred to as 'brute force double loop sampling'. In the first loop, for the current factor i, x_i is set to r prescribed values, while in the second loop n realizations of the remaining inputs $x_{\sim i}$ are generated. When, additionally, T_i is estimated, the cost of the tailored sampling design increases further.

Obviously, such an estimation is impracticable when the computational cost of running the model is high (e.g., more than minutes for one single run). For this reason, several methods have been developed to reduce the computational burden. Specifically, there are several computationally affordable methods proposed in literature to compute the first-order sensitivity indices (Cukier et al., 1973; Lewandowski et al., 2007; Mara et al., 2017; Mara and Joseph, 2008; McKay et al., 1999; Oakley and O'Hagan, 2004; Plischke, 2010; Ratto et al., 2007; Strong et al., 2012; Tarantola et al., 2006). In most of these cases, a dedicated sampling design is used (e.g., based on Fourier transformation or resampling). However, the estimation can be also directly performed based on a generic sampling design (Kucherenko et al., 2012; Wainwright et al., 2014; Plischke, 2010; Li and Mahadevan, 2016; Strong and Oakley, 2013). In the simplest case, Kucherenko et al. (2012) noted that there is no the need of brute-force double-loop sampling (eq. 9) but the analysis can be directly performed based on a filtering process over a generic sampling design of the cost N. The strategy is illustrated in Figure 1 (upper row) where the input-output space of a three parameter function is shown based on scatterplots. Based on this strategy, the ranges of X_i are partitioned into m conditioning intervals of equal size. In each interval the conditional mean (i.e., $\mathrm{E}(Y|X_i)$) is calculated. These values are represented in Figure 1 (upper row) by the red dots. Then, the main effect is estimated as the variance of these conditional means (variance of the red dots). These indices are shown in the bar plot on the right side. The indices indicate, for this case, high importance of the first factor x_1 and no main effect for x_2 and x_3.

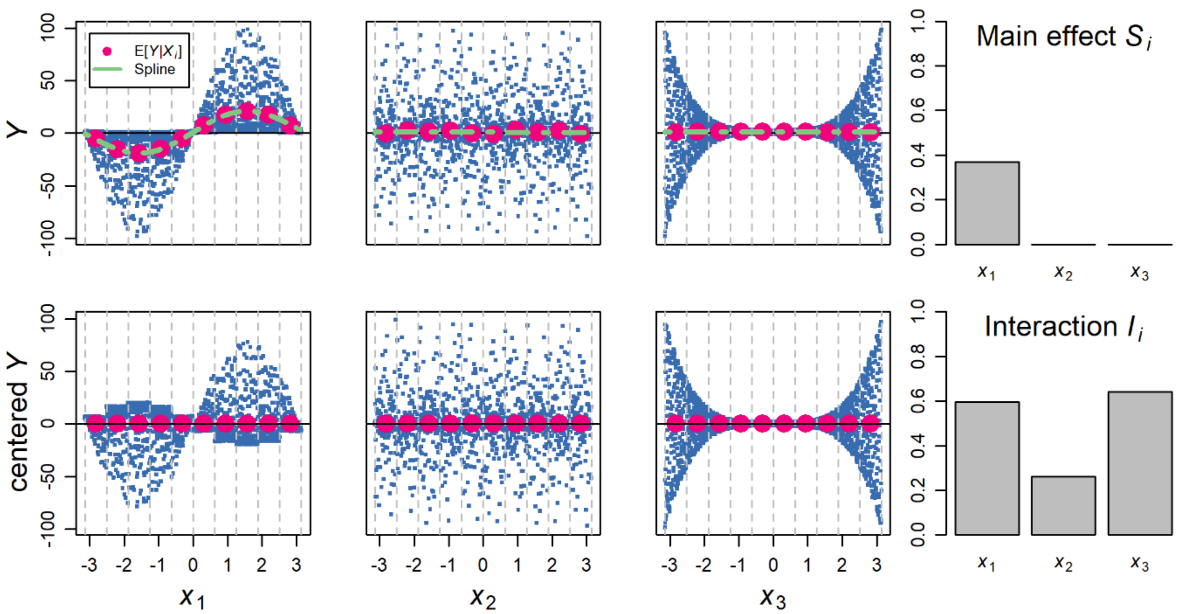

Figure 1. Scatterplots showing the input-output space for the Ishigami-Homma function (eq. 19). The input-output space is divided in m = 10 intervals. In each interval, the mean of the conditional distribution $\mathrm{E}[Y|X_i]$ is plotted as red dot. The variance of these conditional means can be used to estimate the main effect S_i as shown in the bar plot on the top-right corner. The estimation can be improved by interpolating the $\mathrm{E}[Y|X_i]$ (dashed green line) and calculating the variance over the interpolated values. In the lower row, the centered conditional distributions are plotted (i.e., removing the conditional mean). The interactions term I_i derived based on eq. 18 are shown in the bar plot in the lower-right corner.

This method has been found to perform well in estimating S_i in comparison to other approaches (Kucherenko and Song, 2017; Li and Mahadevan, 2016). Similar simple approaches have been not found, however, for estimating the total effect. In contrast, the estimation of this index relies on specific sampling designs (Fourier or resampling) and it is much more expensive to compute (Glen and Isaacs, 2012; Saltelli et al., 2010, 1999). Currently, one of the most applied and effective method is the one discussed by Saltelli et al. (2010). It has shown to perform well in several applications with the advantage of being extendable to the analysis of any sources of uncertainty (Baroni and Tarantola, 2014; Lilburne and Tarantola, 2009; Savage et al., 2016). We briefly introduce this approach also within the present study as reference.

First, two independent sets of input sample matrices A and B, each of which is an $n \times k$ matrix containing n sets of k-dimensional parameter vectors from Monte-Carlo sampling are generated. Sobol' quasi-random sampling (Sobol, 1976) is usually suggested for this purpose because it showed to increase the rate of convergence of the estimators (Becker et al., 2018; Kucherenko et al., 2011; Tarantola et al., 2012) but other techniques can be used as well, e.g., Latin hypercube sampling. From A and B, a matrix C_i (i = 1, 2, ….k) is created for each factor such that the i-th column of C_i is the same as the i-th column of A, and the other columns of C_i are the same as B. The number of required simulations is then $N = n \cdot (k + 2)$. Based on these matrices, different estimators of the main effect (eq. 6) and total effect (eq. 7) can be applied (Saltelli et al., 2010). For ease of interpretation, we report here the estimation first presented by Jansen (1999):

$$\hat{S}_i = 1 - \frac{\frac{1}{2N}\sum_{i=1}^{k}[g(A)-g(C_i)]^2}{V[g(A)]} \qquad (10)$$

$$\hat{T}_i = \frac{\frac{1}{2\mathrm{N}}\sum_{\mathrm{i}=1}^{\mathrm{k}}[g(B)-g(C_i)]^2}{\mathrm{V}[g(A)]} \qquad (11)$$

$$\hat{I}_i = \hat{T}_i - \hat{S}_i \qquad (12)$$

The "hat" is here used to indicate that these terms are estimation of the indices expressed by Equations 6-8. As discussed by Wainwright et al. (2014), these estimations offer an intuitive way to understand the meaning of the Sobol' indices that we report also here for the sake of clarity. The equations represent in fact a correlation-based measure between two matrices. In the first estimation (eq. 10), the parameter sets in A and C_i share the same values only for parameter i. Perturbing all the other factors except for i includes the total effects involving all the factors except for the first-order effect of i. If i is very influential, its value mainly determines the results so that $g(A)$ and $g(C_i)$ should be similar and the differences $[g(A) - g(C_i)]$ small. In the second index (eq. 11), B and C_i have the same values except for i. When we perturb i with the other parameters fixed, the difference $[g(B) - g(C_i)]$, hence T_i, accounts not only for the impact of i as a single factor, but also for the interaction effects with the other factors. If the parameter i is very influential, this factor determines the results so that $g(B)$ and $g(C_i)$ should now be different and the differences $[g(B) - g(C_i)]$ become high. Finally, the difference I_i between total effect and main effect (eq. 12) quantifies only the contribution of the factor i due to interactions.

2.3 Distribution-based approaches

Distribution-based important measures follow the same derivation as described in section 2.1 for variance-based approach. They vary only in the statistical operator to compare the conditional and the unconditional probabilistic distributions (Plischke et al., 2013; Veiga, 2015).

Specifically, some authors (Liu et al., 2006; Park and Ahn, 1994) compared the distributions based on the Kullback-Leibler entropy metric (Kullback and Leibler, 1951). In both cases, however, the analysis has been performed only by fixing x_i to its mean value (i.e., as in eq. 3 for the case of variance-based approach):

$$KL_i = \int P\left(Y|X_{i=E(X_i)}\right) \cdot \log \frac{P\left(Y|X_{i=E(X_i)}\right)}{P(Y)} dy \qquad (13)$$

Alternative distance measures have also been proposed (Chun et al., 2000). Later, Borgonovo (2007) developed the so-called δ-measure based on the absolute difference between the density functions $f(\cdot)$. In addition, he extended the previous methods by taking a statistic (half of the average E) over r conditional values to eliminate the dependency of the conditional point (the mean of X_i) as it has been developed for the variance-based approach (see eq. 4):

$$\delta_i = \frac{1}{2} E \int |f(Y|X_i) - f(Y)| dy \qquad (14)$$

More recently, Pianosi and Wagener (2015) introduced the PAWN method which uses r conditional values as well but the comparison between conditional and unconditional distributions is performed on the cumulative probabilistic distribution based on the Kolmogorov-Smirnov test KS (Kolmogorov, 1933).

$$PAWN_i = median\{KS[P(Y|X_i), P(Y)]\} \qquad (15)$$

where *median* is often used as statistical operator even if other operators (e.g., mean, maximum) can also be considered (Pianosi and Wagener, 2015). Additional measures have also been introduced underlining the strong analogy between them (Veiga, 2015). However, a comprehensive comparison of the different operators has not been carried out so far.

All these distribution-based approaches have been proposed based on the same tailored sampling design discussed for estimating the main effect in variance-based approach (eq. 9). However, in the same way that it has been tested for estimating the main effect (Kucherenko et al., 2012), later it has also been noted that the estimation of the distribution measures can also be performed based on filtering a generic sampling of the cost N without the need of the double-loop brute-force sampling design (Pianosi and Wagener, 2018; Plischke et al., 2013).

It should be noted, however, that extending the distribution-based approach to quantify both main and total effects as performed for variance-based approach (i.e., also comparing conditional distributions when all factors except i are fixed as defined in eq. 7) is not practical as the number of simulation greatly increases (Kucherenko et al., 2012; Saltelli et al., 2010, 1999). For this reason, the distribution-based approach has been used mainly to estimate one index and it has been extended to quantify main and total effects when the number of conditional points r have been limited to one, i.e., the mean of X_i (Liu et al., 2006).

Finally, it is interesting to note that despite the difficulties to directly compare variance- and distribution-based measures, it has been underlined that the distribution-based index contains global information and should be compared to the total effect index derived in variance-based approach (Auder and Iooss, 2009). This can be seen by noting that the estimation of the variance-based total effect T_i in eq. 11 (i.e., $[g(B) - g(C_i)]$) is equivalent to assessing the difference in Y when perturbing X_i with the other parameters fixed. This is like what happens also for the distribution-based comparison (see eq. 13-15).

2.4 An effective combined strategy

The basis for an effective strategy for combining variance- and distribution-based global sensitivity analysis starts by recognizing two key aspects of the approaches. First, both main effect (Kucherenko et al., 2012; Kucherenko and Song, 2017; Li and Mahadevan, 2016; Wainwright et al., 2014) and distribution indices (Pianosi and Wagener, 2018; Plischke et al., 2013) can be estimated based on a generic sampling design, i.e., N samples are filtered into m

intervals to approximate the conditional distributions. This reduces the computational cost in comparison to the double loop sampling design. In addition, it facilitates the analysis because it can be used after any generic Monte Carlo simulations that has been previously conducted for, e.g., uncertainty propagation assessment. Second, the two indices calculated based on variance- and distribution-based approaches carry different meanings. In the variance-based approach, the index refers to the main effect while in the distribution-based approach it carries global information associated also to interactions. For these reasons, the combined use has the potential to address, based on a generic sampling design, all the three purposes of a global sensitivity analysis: factor prioritization based on main effect, factor fixing based on total effect and factor identifiability based on interaction effect.

At this stage, however, the combined use of main effect and distribution-based index is not meaningful because the actual values are not comparable. To overcome this limitation, we proceed as follow.

First, we estimate the main effect S_i (eq. 6) based on a generic sampling design. Specifically, the estimation can be simply performed based on the filtering approach previously described (Kucherenko et al., 2012; Kucherenko and Song, 2017), i.e., based on the variance of the conditional mean $E(Y|X_i)$ estimated in m intervals (Figure 1, upper row, red dots). The estimation, however, can be boosted by interpolating the input-output model response (green dashed lines in Figure 1) and calculating the variance over the interpolated values as suggested in other studies (Saltelli et al., 2008; Wainwright et al., 2014). In the following, we applied this approach using a smooth spline for interpolating the one-dimensional input-output space. We refer to this estimation with the term $\tilde{S}_i$ to distinguish it from the term $\hat{S}_i$ that is used to denote the index estimated based on Saltelli/Jansen approach (eq. 10). A comparison between calculating the variance over the interpolated values (green dashed line) versus using the conditional means (red dots) is presented in the discussion section.

We now remove the main effect from the conditional distributions by centralizing each conditional distribution (i.e., removing the mean):

$$\tilde{P}(Y|X_i) = P\big(Y|X_i - E(Y|X_i)\big) \tag{16}$$

These *m* centralized conditional distributions are shown in Figure 1 (lower row). Differences to the upper row result only for factor x_1, for which now the mean of the conditional distributions (red dots) are equal to zero, too.

By removing the main effect, we now define an index based on these centered conditional distributions to have an estimation of the interaction effects alone (Figure 1, lower right corner). Namely, we can compare the conditional and unconditional distribution as performed by the distribution-based approaches (see eqs. 13, 14 and 15). However, we rather define the interaction index by comparing only all the combinations $c = \binom{m}{2}$ of the centered conditional distributions $\tilde{P}(\cdot)$ to better isolate the effect to the output of changing X_i as follow:

$$\tilde{I}_i = stat\left\{O\left[\tilde{P}\left(Y|X_{i=R_j}\right), \tilde{P}\left(Y|X_{i=R_{j+1}}\right)\right]\right\} \quad (17)$$

where $R_1 \ldots R_m$ represent the *m* intervals, *stat* is a statistical operator (e.g., mean, median) and *O* is the distribution-based measure. We use here the term $\tilde{I}_i$ to distinguish from the term $\hat{I}_i$ calculated as the difference between total and main effect estimated using variance-based approach (eq. 12). In the following we refer to this general Combined Variance- and Distribution-based strategy with the term CVD strategy.

In principle, any operator *O* described above (e.g., entropy-based, δ-measure or Kolmogorov-Smirnov test) can be used within the CVD strategy. Although it is shown that these measures have some analogies (Veiga, 2015), however, a comprehensive inter-comparison in their use within the distribution-based approaches has not been performed so far. In the following, we proceed with the use of the δ-measure of Borgonovo (2007) (eq. 14). Specifically, we fit the centered conditional distributions $\tilde{P}(\cdot)$ via a kernel density function (Parzen, 1962) and we quantify the difference between all the combination $c = \binom{m}{2}$ of the centered conditional density functions $\tilde{f}(\cdot)$ as follow:

$$\tilde{I}_i = \frac{1}{2c}\sum_{j=1}^{m-1}\sum_{p=2}^{m}\left|\tilde{f}\left(Y|X_{i=R_j}\right) - \tilde{f}\left(Y|X_{i=R_{p,p>j}}\right)\right| \quad (18)$$

A comparison between this operator and the alternative Kolmogorov-Smirnov test used in the PAWN method is presented in the discussion section.

3 Test functions

3.1 Functions and settings

We test the CVD strategy described in section 2.4 for estimating main effect $\tilde{S}_i$ and interaction effect $\tilde{I}_i$ on four analytic functions often used as a benchmark for sensitivity analyses studies (Borgonovo, 2007; Cuntz et al., 2015; Kucherenko et al., 2009; Kucherenko and Song, 2017; Mai and Tolson, 2019; Pianosi and Wagener, 2015; Plischke et al., 2013; Saltelli et al., 2008). These functions represent model of increasing complexity (number of factor k between 3 and 15) and with different output distributions (increasing skewness). For comparison, we also apply the method of Saltelli/Jansen for estimating main effect $\hat{S}_i$ and interaction effect $\hat{I}_i$ based on eqs. 10-11-12 and the distribution-based measure $PAWN_i$ (eq. 15). Thus, for each function we calculate and compare five indices.

The first function is the Ishigami-Homma function:

$$y = \sin(x_1) + a_1 \sin(x_2)^2 + a_2\, x_3^4 \sin(x_1) \qquad (19)$$

where all x_i follow a uniform distribution over $[-\pi, +\pi]$. Different values of the parameters a and b are encountered in literatures. In the following we use $a_1 = 2$ and $a_2 = 1$ as used by Pianosi and Wagener (2015) to allow for a direct comparison. The function exhibits strong non-linearity and non-monotonicity with interactions between the two terms x_1 and x_3. The input-output response is also shown in the scatter plots of Figure 1.

The second function has been introduced by Oakley and O'Hagan (2004):

$$y = \boldsymbol{a}_1^T X + \boldsymbol{a}_2^T \sin(X) + \boldsymbol{a}_3^T \cos(X) + X^T M X \qquad (20)$$

where X is a vector of k =15 input factors sampled over a standard normal distribution (i.e., mean = 0 and standard deviation = 1) and $\boldsymbol{a}_j$ (j = 1, 2, 3) and $\boldsymbol{M}$ are three k-vectors and a matrix $k \times k$ of parameters, respectively. The weights $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$ are chosen so that one group of five input factors ($x_{i = 10 \dots 15}$) accounts for most of the output variance while the remaining factors have smaller effect. In contrast, interaction contribution is almost equally distributed to all the factors. The values of these weights and of the matrix M can be downloaded from www.sheffield.ac.uk/st1jeo and are also reported in Saltelli et al. (2008).

The third function is the so-called G-function:

$$y = \prod_{i=1}^{k} \frac{|4x_i - 2| + a_i}{1 + a_i} \tag{21}$$

where the terms x_i are k independent variables uniformly distributed in the unit hypercube [0, 1] and the terms a_i are positive integers that condition the importance of the factor x_i: the smaller a_i, the higher main and total effects of x_i on the function are (Mara, 2009). This function is a more complicated nonlinear and non-additive function that produces a right-skewed output distribution. It is characterized by the presence of interactions among the model inputs, generated by their multiplication. We use a six-dimensional version of the function (k = *6)*, with coefficients a_i equal to [0, 0.5, 3, 9, 99, 99] as used in other studies (Glen and Isaacs, 2012; Saltelli et al., 2010), resulting in decreasing sensitivity of the output to the six factors.

Finally, we consider the function introduced by Bratley et al. (1992):

$$y = \sum_{i=1}^{k} (-1)^i \prod_{j=1}^{i} x_j \tag{22}$$

where k = 10 and all x_i follow a uniform distribution over [0, 1]. Like the G-function, the input factors also have a wide range of main and interaction effect, decreasing with j, but the function produces a left-skewed model output distribution.

For each function, the five sensitivity indices are estimated based on Sobol' quasi-random sampling (Sobol, 1976). Saltelli/Jansen estimations are performed based on $N = (k+2) \cdot 2^{15}$ simulation runs. The *PAWN* and CVD indices are estimated based on $N = 2^{17}$ simulation runs. All these indices are calculated by increasing sample size N to assess the convergence of the estimation. Moreover, the analyses are repeated 100 times to assess the robustness of the estimation and the approach of Owen (1995) is used to introduce randomness to the deterministic Sobol' sequence. The accuracy of the estimation is quantified based on the absolute mean error *MAE* as follow:

$$MAE = \frac{1}{k} \sum_{i=1}^{k} |\mathcal{E}_i - \mathcal{R}_i| \tag{23}$$

where k is the respective number of input factors i, and $\mathcal{E}$ and $\mathcal{R}$ represent the estimated and reference sensitivity index, respectively. For the main effect ($\hat{S}_i$ and $\tilde{S}_i$), the analytic values are used as references. For the interaction terms ($\hat{I}_i$ and $\tilde{I}_i$), however, the comparison is not straightforward as the values represent different quantities and the analytic values of the distribution-based indices are not available. For this reason, the assessment of the interaction terms is performed using as references in eq 23 the average indices over the 100 repetitions obtained with the maximum sample size *N*.

3.2 Results of the test functions

Figure 2 shows the five indices obtained for the four test functions. The average indices over the 100 repetitions achieved with the maximum number of simulation *N* are plotted. The indices are combined as conducted for other approaches (Morris, 1991) to visualize their relation and their different information content. Please also note that error bars (± one standard deviation based on the 100 repetitions) are also plotted but they are not always visible, indicating the robustness of the estimation achieved with the maximum sample size *N*.

The analysis conducted based on Saltelli/Jansen method on the Ishigami-Homma function (Figure 2, first column, second row) identifies x_1 as an important factor based on the main effect $\hat{S}_i$ (see also scatter plots of Figure 1). The same interaction effects $\hat{I}_i$ are quantified for the factors x_1 and x_3. The results suggest that x_2 is not relevant and can be fixed to any value (e.g., mean), x_1 and x_3 have interactions and are not identifiable.

The $PAWN_i$ index (Figure 2, first column, third row) ranks the factors consistently ($x_1 > x_3 > x_2$). However, based on this index alone, it is not possible to identify whether the factors have interactions or not. The comparison between $PAWN_i$ index and the main effect $\hat{S}_i$ suggests that the importance of the parameters x_2 and x_3 result from interactions. But it is not possible to conclude if these interaction effects are similar or not, indicating the difficulties to compare variance- and distribution-based indices as underlined in other studies (Borgonovo and Tarantola, 2008; Mora et al., 2019).

The main effect $\tilde{S}_i$ estimated based on the spline interpolation implemented in the CVD strategy (Figure 2, first column, last row) is consistent with Saltelli/Jansen method. In addition, the new interaction term $\tilde{I}_i$ (eq. 18) correctly identifies the same interaction effects on x_1 and x_3. Thus, the results show how the CVD strategy correctly removes the main effect and the distribution measure is now able to discriminate specifically the contribution due to the interaction effect.

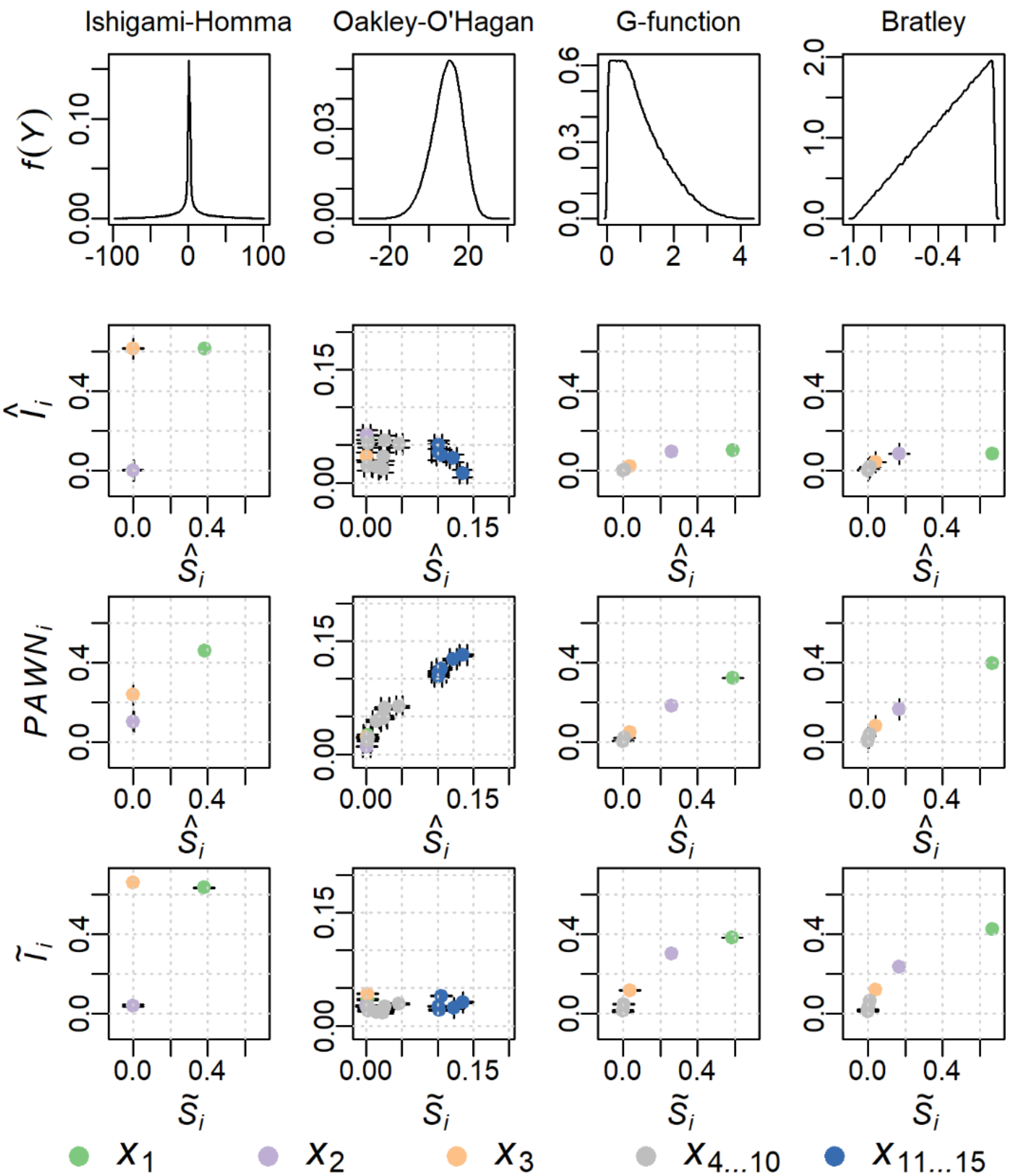

Figure 2. In the upper row, the output distribution of each analytic function is shown. The other plots show the sensitivity indices estimated based on the different methods: from top, main effect $\hat{\boldsymbol{S}}_{\boldsymbol{i}}$ and interaction effect $\hat{\boldsymbol{I}}_{\boldsymbol{i}}$ based on Saltelli/Jansen method; $PAWN_i$ index; and main effect $\tilde{\boldsymbol{S}}_{\boldsymbol{i}}$ and interaction term $\tilde{\boldsymbol{I}}_{\boldsymbol{i}}$ based on the CVD strategy. The points indicate the average indices over the 100 repetitions obtained with the maximum sample size *N*. Please also note that error bars (± one standard deviation based on the 100 repetitions) are also plotted for each index.

The same conclusions are derived by looking at the results of the Oakley and O'Hagan function (Figure 2, column 2). The Saltelli/Jansen method correctly identifies the most important five input factors based on the main effect ($x_{i\,=\,10\ldots15}$). Moreover, it quantifies interaction as evenly distributed over all the factors. In contrast, the $PAWN_i$ indices show strong correlation with the main effect but we are not able to conclude if this contribution is due to direct effect of the input factors to the model output or based on interactions. In contrast, the CVD strategy estimates well the main effect and it quantifies the interaction terms as detected by Satelli/Jansen method.

Some additional considerations arise by looking at the results of the other two functions for which model outputs show strong skewed distributions (Figure 2, column 3 and 4). All the three approaches rank the factors consistently. However, $PAWN_i$ indices and the interactions terms $\tilde{I}_i$ of the CVD strategy show higher values for the input factors x_1. In contrast, the interaction terms $\hat{I}_i$ estimated based on Saltelli/Jansen method rank x_1 and x_2 equally. While it cannot be concluded by looking at $PAWN_i$ indices what is the reason of these differences, by removing the main effect with the CVD strategy (eq. 16), the results suggest that the difference is due to stronger interactions of the factor x_1 with the other factors that cannot be detected using the variance as summery statistic.

We further explore this hypothesis by looking at the input-output space of the first two input factors of the Bratley function (Figure 3). The variability of the conditional means calculated in the m intervals (red dots in the upper row) is larger for the factor x_1 than for x_2, confirming also by visual inspection that the first factor is more important in explaining the model output variability (higher main effect). We now remove these main effects by centralizing each conditional distribution (eq. 16) to isolate the interaction effect (Figure 3, lower row). When removing these conditional means, the obtained centered conditional distributions of each input factor are very different. In the case of x_1, the distribution in the first left interval m is very narrow and then the spread strongly increases. In contrast, the differences between the centered conditional distributions of the factor x_2 are smaller. Despite these differences, when looking at the two factors, it is interesting to note that the variances calculated in each centered distributions are comparable (see green square in Figure 3, lower row). Thus, the results show how the variance is not an adequate summary statistic in this specific case. For this reason, the variance-based sensitivity analysis fails to reveal the difference between the factors and assigns the same interaction terms to both. In contrast, these differences can only be distinguished by looking at their entire distributions and for this reason, the distribution-based measures are better suited to quantify these interaction terms. Same conclusions are supported by looking at the input/output space of the G-function (data not shown).

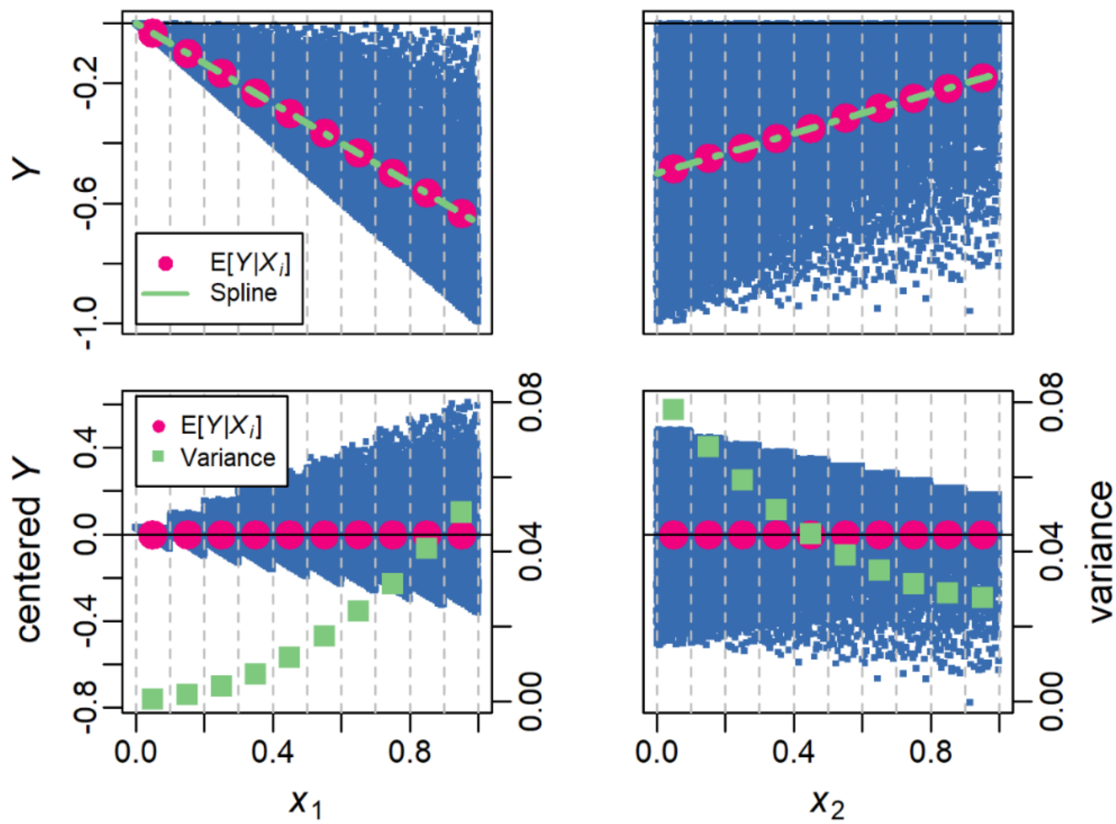

Figure 3. Scatterplots showing the input-output space for the first two input factors (x_1, x_2) of the Bratley function (eq. 22). The input-output space is divided in m = 10 intervals. In each interval, the mean of the conditional distribution $E[Y|X_i]$ is plotted as a red dot. The spline interpolation of these values is shown as a dashed green line. In the lower row, the centered conditional distributions are plotted (i.e., removing the conditional mean). In each interval m the variance of each conditional centered distribution is also plotted as a green square.

3.3 Convergence and robustness of the estimations

The mean absolute error *MAE* (eq. 23) is shown in Figure 4 for increasing sample size N to assess the convergence of the estimation of the indices. In addition, we visualize the width of the 95% confidence intervals of the performance metric distribution obtained with the 100 repetitions to assess the robustness of the estimation (variability over the repetitions). Please also note that the x-axis is plotted in $\log_{10}$ base to illustrate the behavior for small N.

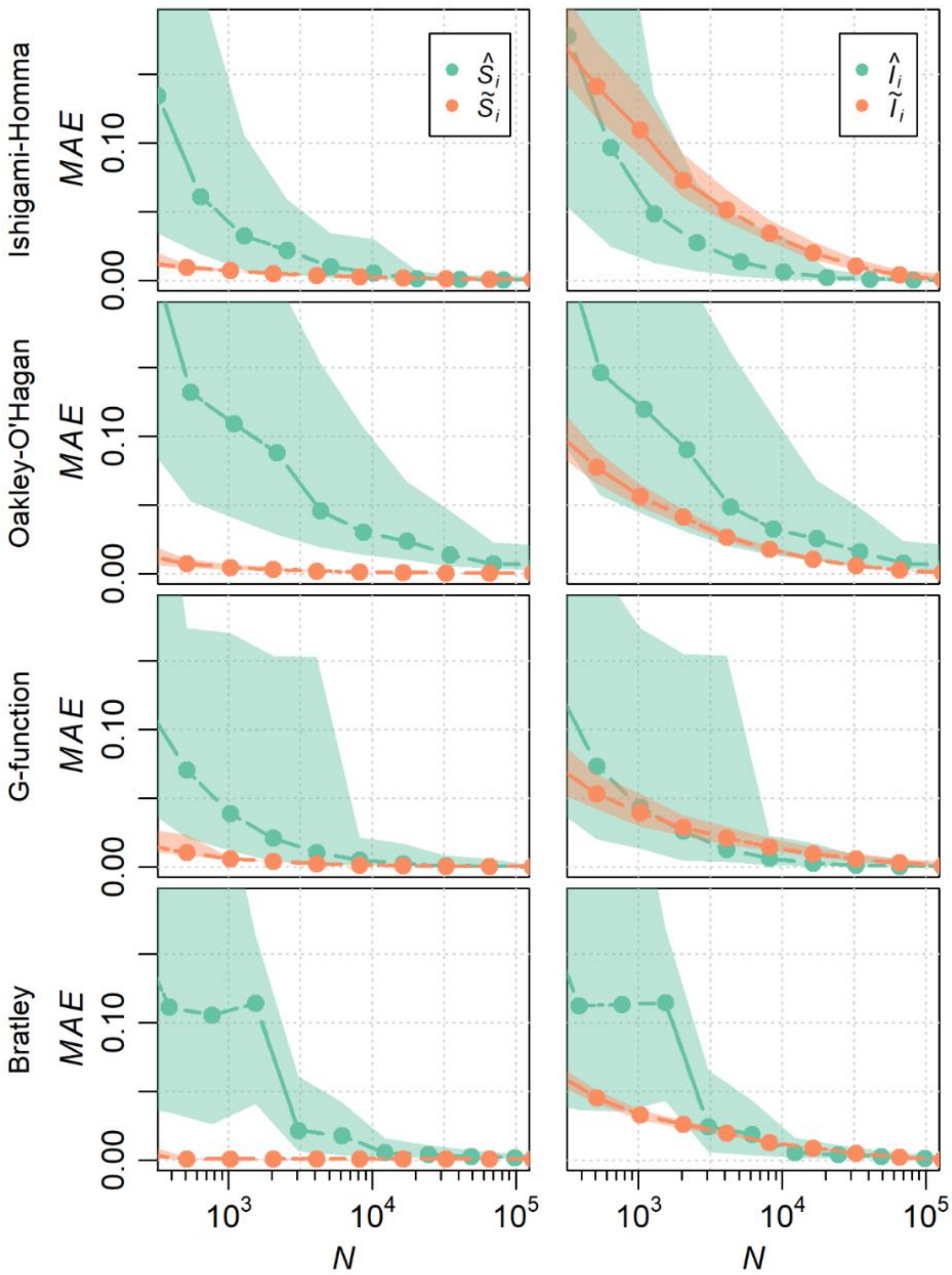

Figure 4. Sensitivity indices estimated with different methods and increasing sample size N for the four test functions: (from top) Ishigami-Homma function, Oakley and O'Hagan function, G-function function and Bratley function. In the left column, main effect estimated by Saltelli/Jansen formula ($\hat{S}_i$) and based on the spline interpolation integrated in the CVD strategy ($\widetilde{S_i}$). In the right column, the interaction effect estimated by Saltelli/Jansen formula ($\hat{I}_i$) and based on the CVD strategy ($\tilde{I}_i$). The colored envelopes demark the inner 95% confidence interval of the 100 repetitions.

Results show that the numerical estimates converge towards the reference values and the dispersion is constantly reduced as the sample size increases. The results confirm the accuracy of Saltelli/Jansen approach in estimating references indices but the need of a relatively high number of simulations ($N \sim 10^4$). In contrast, indices estimated based on the CVD strategy are more robust (lower spread over the replicates) and converge faster than Jansen/Saltelli method ($N \sim 10^3$). Specifically, the spline interpolation method used to estimate the main effect integrated in the CVD strategy outperforms Jansen/Saltelli method in all the test functions. Therefore, the results confirm the availability of efficient alternative methods when we are interested to quantify only the main effect (Kucherenko and Song, 2017; Plischke et al., 2013; Strong and Oakley, 2013; Wainwright et al., 2014). In contrast, the estimation of the interaction terms requires a relative larger sample size also with the CVD strategy. Still, the CVD strategy shows to converge faster to the reference index and to be much more robust (smaller spread over the replicates) than the Saltelli/Jansen method.

The only exception is detected for the interaction terms of the Ishigami-Homma function (Figure 4, second column, first row) for which the robustness is still much stronger than Saltelli/Jansen method (smaller colored envelopes) but the convergence to the reference values requires a larger sample size N. This behavior is explored in detail by looking at the estimation of the indices of each input factor (Figure 5). The results confirm that the interaction terms estimated based on Saltelli/Jansen method are well estimated (upper plot), but the estimation is not robust at a relative low sample size ($N < 10^4$). When looking at the CVD strategy (bottom plot), the factors are well ranked at relative low samples size too. Noteworthy, however, the convergence of the estimation of the interaction term for the factor x_2 is much slower than for the other two factors. This low performance can be explained considering that x_2 has a negligible interaction contribution ($I \sim 0$). When a relative low sample size N is used, however, the distribution-based measure (eq. 18) is affected by the numerical approximation of the conditional distributions and a higher value is quantified. Some possible improvements to address the estimation of very low values of the interaction indices are discussed in section 5.

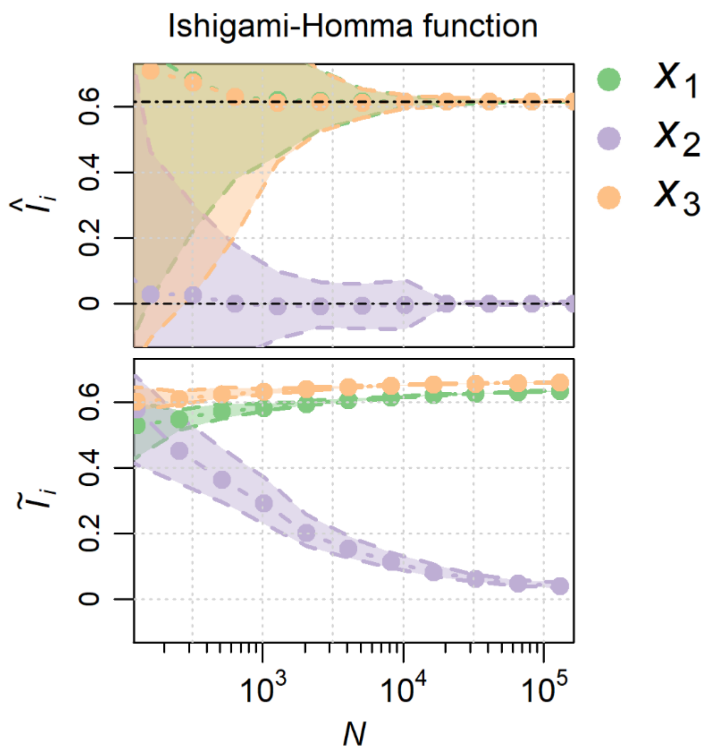

Figure 5. Interactions terms of the Ishigami-Homma function estimated based on Saltelli/Jansen method (upper plot) and based on the CVD strategy (bottom). The indices are estimated for increasing sample size *N*. The colored envelopes demark the inner 95% confidence interval of the 100 repetitions. In the upper plot the analytic indices are also plotted as dash-dotted horizonal lines for comparison.

4 A practical workflow based on a hydrological model

4.1 SAC-SMA model and setting analysis

We performed the global sensitivity analysis based on the CVD strategy and on the variance-based approach of Saltelli/Jansen to the Sacramento Soil Moisture Accounting Model (SAC-SMA). This is an intermediate-complexity conceptual rainfall-runoff model that represents the soil column by an upper and lower zone of multiple storages (Burnash, 1995). It has been used extensively in both research and operational applications. The model has also been used in the context of GSA and parameter identifiability (Blasone et al., 2008; Shin et al., 2013; van Werkhoven et al., 2009, 2008). The input, parametrizations and observations of Shin et al. (2013) is used within the present study to allow a comparison. The 13 parameters and ranges are listed in Table A1 in the appendix. The model runs for 10 years. The first year was not included in the sensitivity calculations to allow the model states to warm up and remove any impact of uncertain initial conditions. The ranges of the parameters are sampled based on Sobol' sampling design with $N = (13 + 2) \cdot 2^{14} = 122880$ model evaluations for the Saltelli/Jansen estimation and $N = 2^{17} =$ 131072 for the CVD strategy. These sample sizes have also been used in other studies (Shin et al., 2013; van Werkhoven et al., 2009). Simulations and analyses are repeated 100 times and the approach of Owen (1995) is used to introduce randomness to the deterministic Sobol' sequence. The modelling results are evaluated based on the following performance metrics, for which the sensitivities are assessed:

$$NSE = 1 - \frac{\sum_{t=1}^{T}(o_t - y_t)^2}{\sum_{t=1}^{T}(o_t - E[o])^2} \tag{22}$$

$$NSE^* = \frac{1 - \frac{\sum_{t=1}^{T}(o_t - y_t)^2}{\sum_{t=1}^{T}(o_t - E[o])^2}}{1 + \frac{\sum_{t=1}^{T}(o_t - y_t)^2}{\sum_{t=1}^{T}(o_t - E[o])^2}} \tag{23}$$

where T is the number of time steps t, y and o are simulated and measured river discharge, respectively, and E[o] indicates the mean of observation over the time series. Equation 23 is the Nash-Sutcliffe-index and has been commonly used for river discharge assessment. Its value is within in the range $[-\infty, +1]$. The modified version (eq. 23) has been introduced by Mathevet et al. (2006) and it is bounded between $[-1, +1]$. Thus, it reduces the influence of large negative values without otherwise changing the interpretation of the objective function (Shin et al., 2013). The two metrices have been selected as example to show the differences in the sensitivity results when model output present different distributions.

4.2 Results of the hydrological model

The results of the sensitivity analysis are shown in Figure 6. The indices are plotted to visualize the relation between main and interaction effect of each parameter as conducted for other approaches (Morris, 1991). Each plot is also arbitrary divided (dashed gray line) to better highlight low values in the indices.

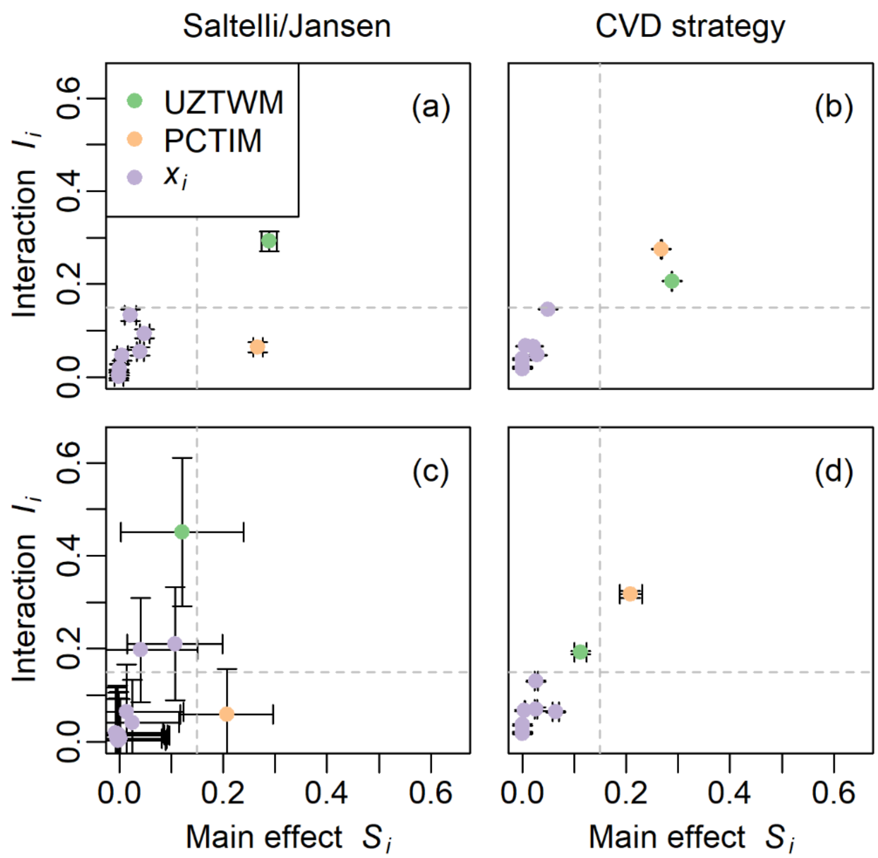

Figure 6. Results of the sensitivity analysis based on the modified NSE* performance metric (upper row) and based on the standard NSE (lower row). The analysis is performed based on Saltelli/Jansen method (a and c) and based on the CVD strategy (b and d).

For the case of the modified NSE*, the Saltelli/Jansen approach estimates low main effect and low interaction for most of the parameters (Figure 6a, gray dots). Thus, the values of these parameters can be fixed to any arbitrary values within their ranges and they can be omitted during further analysis. In contrast, the main effect of the parameters UZTWM and PCTIM, related to the soil water capacity and the land surface characteristic, respectively (see appendix), is relatively large (~0.26). For this reason, these parameters can be targeted for further model improvements by, e.g., calibration. The estimated interaction term, however, is low for PCTIM while is large for UZTWM. Thus, the analysis suggests that PCTIM is likely identifiable while the optimum value of UZTWM can have strong dependencies with the values of the other parameters. For this reason, this parameter could be not identifiable.

These conclusions are also supported by the corresponding indices estimated based on the CVD strategy (Figure 6a, gray dots). The main difference is only detected for the parameter PCTIM for which a large interaction term is quantified (even larger than the interaction index estimated for the parameter UZTWM). This difference to the results of the Saltelli/Jansen method can be explained by looking at the input-output space for these parameters (Figure 7). The variability of the conditional means (red dots) of the two parameters is similar in terms of range and smoothness (Figure 7a and b). However, when removing these conditional means to isolate the interaction effect, the obtained centered conditional distributions are very different. In the case of UZTWM (Figure 7c), the distributions are very similar and the main difference is in the variance (see green squares in Figure 7c). For this reason, the variance-based approach can capture the interaction term. In contrast, the centered conditional distributions for the parameter PCTIM are very similar when looking at the variances (green squares in Figure 7d) and they can only be distinguished by looking at their entire distributions. For this reason, the distribution-measure can capture the interaction term for this parameter while the variance-based approach does not quantify any interaction.

These results are particularly interesting in the light of the parameter identifiability. As an example, the best 1% simulations are filtered and represented in the scatterplots (orange color, Figure 7a and b). In the case of the parameter UZTWM, the best 1% simulations narrow the prior distribution to the lower value of its range. In contrast, the same best 1% simulation are evenly distributed in the range of the parameter PCTIM. Thus, the results show how PCTIM is less identifiable than UZTWM. For this reason, these results underline how the CVD strategy can better capture the interaction effects between the parameters and it is more consistent with the identifiability analysis when the differences in the distributions are not well represented by their variance.

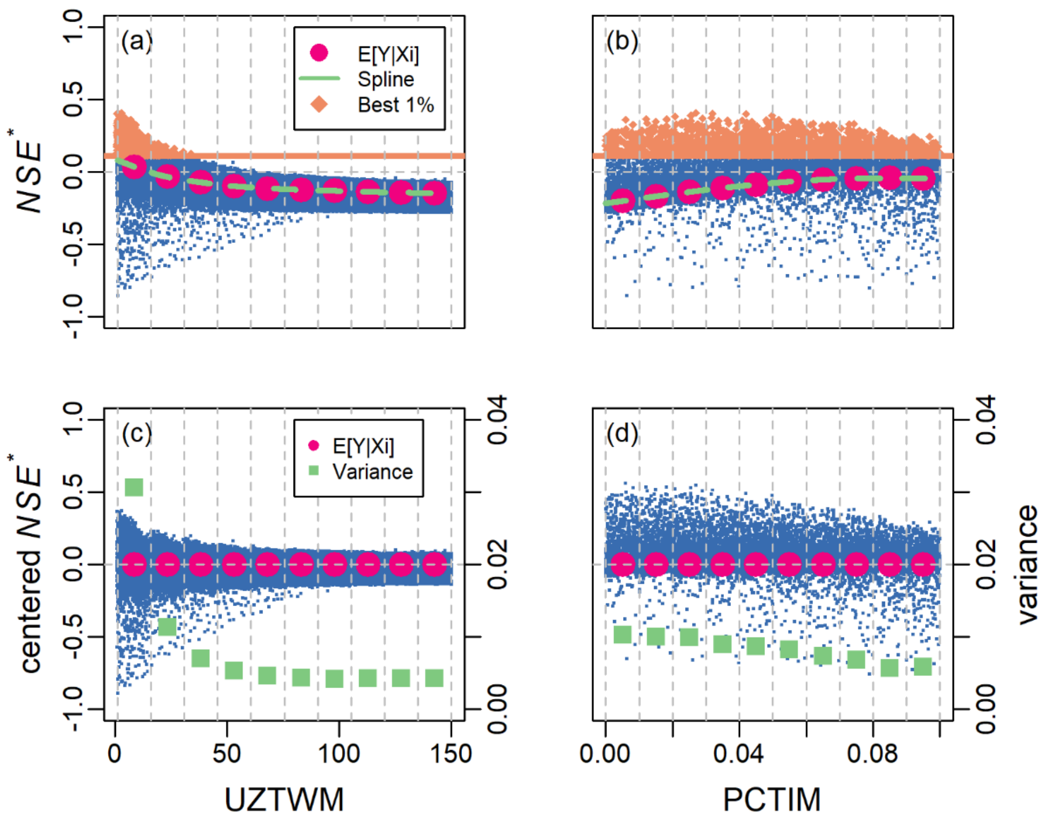

Figure 7. Scatterplots showing the response function of two parameters (UZTWM and PCTIM) of the SAC-SMA model and the model output with regard to the modified NSE*. The input-output space is divided in m = 10 intervals. The mean of the conditional distribution E[Y|X_i] is plotted as red dot in each interval m and the spline interpolation as dashed green line. The best 1% of the simulations are colored in orange. In the lower row, the centered conditional distributions are plotted (i.e., removing the conditional mean). In each interval m the variance of each conditional centered distribution is also plotted as green square.

The same results are obtained when looking at the interaction terms obtained with the standard NSE (Figure 6, lower row). In contrast, however, a lower main effect is estimated for the parameters UZTWM and PCTIM. In addition, the parameter PCTIM becomes slightly more influential than UZTWM. These differences can be explained considering that the standard NSE yields strongly skewed distribution with very low performances obtained on a few samples. In this case, it is hard to compute meaningful statistics to summarize the whole distribution (i.e., mean and variance can be biased by few outliers). Thus, the result confirms the sensitivity of the variance metric to measure the dispersion of a variable with a heavy-tail or which contains some outliers (Auder and Iooss, 2009) and the need to properly defined the performance matric for the model output. It is noteworthy, moreover, that the confidence intervals calculated based on 100 replicates are very large in the case of Saltelli/Jansen estimations. Thus, these results show the difficulties to apply this method when the output distributions are strongly skewed. For this reason, we recommend checking for the normality of the output distribution to understand the reliability in the use of this method. In contrast, the indices estimated based on the CVD strategy are very well estimated even in the face of the skewed distributions in the output response. For this reason, they represent a more general and robust estimation.

5 Discussion

5.1 Sample size N and the number of intervals m

The results have shown that the analysis conducted based on the CVD strategy provides the same (or even more detailed) information as the state-of-the-art of variance-based approach (Saltelli et al., 2010) with also the advantage of using a generic sampling design and converging with a lower sample size. Still, the number of samples N and the number of intervals m used to calculate the conditional distributions are two free parameters of the strategy and they should be selected with caution.

The sensitivity of the results to these free parameters has been explored for instance for the case of distribution-based approaches. It has been shown that in the case of PAWN, results are quite independent of the chosen value of m when N is relatively high and $m > 5$ (Mora et al., 2019; Pianosi and Wagener, 2018). Similarly, Plischke et al. (2013) underlined that increasing the number of m beyond 50 classes has negligible effect on the estimation accuracy. More recently, however, a more detailed analysis conducted on these free parameters have shown the risk of achieving perfunctory results (Puy et al., 2020). For this reason, assuming the number of simulations N being the maximum achievable based on the specific model run-time, we advise to test the robustness of the indices and of the factor ranking by modifying the number of intervals m, as performed in other studies (Li and Mahadevan, 2016; Puy et al., 2020).

We tested this approach by repeating the analyses varying the intervals m in the range [5 - 45] by increments of 5 and we found negligible differences in the indices (results not shown). This robustness is explained considering that the spline interpolation well represents the input-output space independently from the number of interval m. In contrast, the interaction index is calculated over the combination of all the conditional distributions, leading to a relative high number of pairs even when m is relatively low (e.g., when $m = 10$, the combinations $c = \binom{m}{2} = 45$). Thus, the results of the CVD strategy show to be largely independent from the number of intervals m and all the conclusions reported on the role of the different factors of the tests are considered well supported by the analyses.

5.2 Alternative methods in the CVD strategy

The CVD strategy has been applied using the spline interpolation for the estimation of the main effect and the δ-measure for the interaction index. In principle, however, other methods can be applied as well.

In the present study, we repeated the analyses using two different methods: the main effect is estimated based on the variance of the conditional mean $E[Y|X_i]$ (Kucherenko and Song, 2017) without the interpolation step and the interaction index is estimated based on Kolmogorov-Smirnov test, instead of the δ-measure.

Differences between the use of these approaches were negligible in most of the cases. Some differences in the estimation of the main effect have been identified only in the case the model response was highly skewed. For example, the main effects of the skewed functions (eq. 21-22) did not reach the analytic references but they show some differences also at large sample size. For the case of the distribution approach, the ranking was consistent in all the cases. However, it has been noted how δ-measure proposed by Borgonovo (2007) better discriminate the factors in comparison to the Kolmogorov-Smirnov test.

Thus, we suggest applying the CVD strategy based on the spline interpolation and the δ-measure. Still, further settings can be also tested. For instance, the spline interpolation can be optimized based on an iterative step as proposed by Ratto and Pagano (2010). Similar optimization could be also envisioned for the kernel density estimation. For this reason, while we consider the CVD strategy as a new effective strategy for combining variance- and distribution-based approaches, we leave to further studies comparing different settings or other alternative specific methods that can be integrated in the strategy and can perform better in specific conditions (Cukier et al., 1973; Lewandowski et al., 2007; Mara et al., 2017; Mara and Joseph, 2008; McKay et al., 1999; Oakley and O'Hagan, 2004; Plischke, 2010; Ratto et al., 2007; Tarantola et al., 2006; Veiga, 2015).

5.3 Correlated factors

In principle, it is straightforward to apply the CVD strategy also in the presence of correlated input factors. The only difference would be to sample from joint probability distributions before estimating the sensitivity indices. However, it should be noted that these indices lose their interpretability when factors are correlated (Saltelli and Tarantola, 2002). The main effect still is used in the context of identifying the model inputs that, when fixed, lead to the greatest reduction in output variance. However, it contains also interactions information "carried over" by correlation. For this reason, removing this main contribution by centralizing the conditional distributions (eq. 16) does not isolate anymore the interaction effects, and the distribution-based index (eq. 17) loses its information content. Thus, the main effect can be estimated as suggested by Kucherenko and Song (2017). However, we are not able to estimate the interaction effect with eq. 18. For this reason, we advise to work with uncorrelated samples whenever possible, e.g., by treating dependencies as explicit relationships with a noise term (Saltelli et al., 2008). We leave possible improvements to future studies and we refer to the following references for a deeper discussion on global sensitivity analysis applied to correlated input factors (Borgonovo and Tarantola, 2008; Kucherenko et al., 2017; Mara et al., 2015; Tarantola and Mara, 2017; Zhao et al., 2015).

5.4 Integrating good practices in the CVD strategy

Most of the GSA relies on Monte-Carlo simulations. For this reason, a good practice is to repeat the analysis to assess the robustness of the estimation as conducted within the present study (i.e., 100 replicates). When this is not possible, most likely due to computational burden to run the model, complementary approaches can be integrated in the CVD strategy to assist the interpretation of the results as it has been performed for other methods.

Bootstrapping (Efron and Tibshirani, 1994) is a widely applied approach to provide confidence intervals based on resampling the original sample set with replacement. This approach has extensively been used in global sensitivity analysis (Nossent et al., 2011; Sarrazin et al., 2016). This method is, however, inappropriate with small sample sizes. For this reason, the use of dummy variables (Mai and Tolson, 2019; Plischke et al., 2013; Zadeh et al., 2017) or bias-controlling statistical test (Plischke et al., 2013) has been introduced to support the assessment of the indices and the ranking of the different factors. Further improvements can also be performed by iteratively decreasing the input space to be sampled by discarding the factors that are well identified in an iterative screening approach (Cuntz et al., 2015; Lo Piano et al., 2017). Working with groups by perturbing all factors of the same group simultaneously is also very advantageous for models containing a high number of factors (hundreds or thousands). This method allows for the reduction of the number of model executions required, at the cost of losing information on the relative strength of the inputs belonging to the same group (Campolongo et al., 2007). The groups are generally defined a priori introducing some subjectivities into the analysis. More recently, however, an automatic selection of the groups has also been presented to overcome this limitation (Sheikholeslami et al., 2019). Comparison of all these different auxiliary methods should be performed in future studies to identify their advantages and to better guide their use in specific applications.

6 Conclusions

We developed a new strategy called CVD that combines the strength of variance- and distribution-based global sensitivity analysis in a meaningful and effective way. This new strategy enables to estimate main and interaction effects directly from a generic sampling design (random, Latin hypercube, quasi-Monte Carlo, etc.). For these reasons, it provides a comprehensive analysis of the model response that can be easily implemented in any modelling framework and assessment (Baroni and Tarantola, 2014; Uusitalo et al., 2015).

The new combined strategy has been tested on four analytical functions and on a hydrological model. The strategy has been implemented based on a spline interpolation (Saltelli et al., 2008) and the δ-measure (Borgonovo, 2007) for the estimation of the main and interaction term, respectively. However, other methods can be easily integrated and tested in future studies (Kucherenko and Song, 2017; Liu et al., 2006; Pianosi and Wagener, 2015; Ratto and Pagano, 2010). The results are compared to the state-of-the-art of variance-based approach for global sensitivity analysis (Saltelli et al., 2010).

The results showed that the new CVD strategy quantifies main and interaction effects correctly and with a lower sample size. The strategy is also better able to capture the interactions term when distributions are not Gaussian (i.e., the variance does not well represent the distributions). Thus, the strategy combines the strength of variance- and distribution-based approaches to explore the input-output space and the role of the different factors. Overall, the new strategy provides a new simple and comprehensive basis for performing a global sensitivity analysis that can be useful to improve and to facilitate the use of these diagnostic tools for environmental models and to avoid perfunctory analysis that are still very common in many modelling studies (Saltelli et al., 2019).

7 Appendix

Table A1: parameters description and ranges (as taken from Shin et al., 2013). All the parameters follow a uniform distribution.

Parameter name	Unit	Range	Description
UZTWM	[mm]	1–150	Upper zone tension water maximum capacity
UZFWM	[mm]	1–150	Upper zone free water maximum capacity
UZK	[1/day]	0.1–0.5	Upper zone free water lateral depletion rate
PCTIM	[–]	0.000001–0.1	Fraction of the impervious area
ADIMP	[–]	0–0.4	Fraction of the additional impervious area
ZPERC	[–]	1–250	Maximum percolation rate coefficient
REXP	[–]	0–5	Exponent of the percolation equation
LZTWM	[mm]	1–500	Lower zone tension water maximum capacity
LZFSM	[mm]	1–1000	Lower zone supplementary free water maximum capacity
LZFPM	[mm]	1–1000	Lower zone primary free water maximum capacity
LZSK	[1/day]	0.01–0.25	Lower zone supplementary free water depletion rate
LZPK	[1/day]	0.0001–0.25	Lower zone primary free water depletion rate
PFREE	[–]	0–0.6	Direct percolation fraction from upper to lower zone free water storage
SIDE	[–]	0.0 (fixed)	Fraction of base flow that is draining to areas other than the observed channel
RSERV	[–]	0.3 (fixed)	Fraction of the lower zone free water that is unavailable for transpiration purposes
RIVA	[–]	0.0 (fixed)	Fraction of the riparian vegetation area

8 Acknowledgments

We thank Razi Sheikholeslami and two anonymous Reviewers for their constructive comments on previous version of this manuscript. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. G.Baroni is indebted to the students of the graduate course “Uncertainty and Sensitivity analysis” at Environmental Science and Geography at the University of Potsdam (Germany) who provided always good questions and motivations to better study and explain the topic.

9 References

Andrews, F.T., Croke, B.F.W., Jakeman, A.J., 2011. An open software environment for hydrological model assessment and development. Environ. Model. Softw. 26, 1171–1185. https://doi.org/10.1016/j.envsoft.2011.04.006

Auder, B., Iooss, B., 2009. Global sensitivity analysis based on entropy, in: Safety, Reliability and Risk Analysis: Theory, Methods and Applications – Martorell et al. (Eds). Taylor & Francis Group, London, pp. 2107–2115.

Baroni, G., Scheiffele, L.M., Schrön, M., Ingwersen, J., Oswald, S.E., 2018. Uncertainty, sensitivity and improvements in soil moisture estimation with cosmic-ray neutron sensing. J. Hydrol. 564, 873–887. https://doi.org/10.1016/j.jhydrol.2018.07.053

Baroni, G., Tarantola, S., 2014. A General Probabilistic Framework for uncertainty and global sensitivity analysis of deterministic models: A hydrological case study. Environ. Model. Softw. 51, 26–34. https://doi.org/10.1016/j.envsoft.2013.09.022

Baroni, G., Zink, M., Kumar, R., Samaniego, L., Attinger, S., 2017. Effects of uncertainty in soil properties on simulated hydrological states and fluxes at different spatio-temporal scales. Hydrol Earth Syst Sci 21, 2301–2320. https://doi.org/10.5194/hess-21-2301-2017

Becker, W.E., Tarantola, S., Deman, G., 2018. Sensitivity analysis approaches to high-dimensional screening problems at low sample size. J. Stat. Comput. Simul. 88, 2089–2110. https://doi.org/10.1080/00949655.2018.1450876

Bier, V.M., 1982. A measure of uncertainty importance for components in fault trees (Thesis). Massachusetts Institute of Technology.

Blasone, R.-S., Vrugt, J.A., Madsen, H., Rosbjerg, D., Robinson, B.A., Zyvoloski, G.A., 2008. Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. Adv. Water Resour. 31, 630–648. https://doi.org/10.1016/j.advwatres.2007.12.003

Borgonovo, E., 2010. A Methodology for Determining Interactions in Probabilistic Safety Assessment Models by Varying One Parameter at a Time. Risk Anal. 30, 385–399. https://doi.org/10.1111/j.1539-6924.2010.01372.x

Borgonovo, E., 2007. A new uncertainty importance measure. Reliab. Eng. Syst. Saf. 92, 771–784. https://doi.org/10.1016/j.ress.2006.04.015

Borgonovo, E., Castaings, W., Tarantola, S., 2011. Moment Independent Importance Measures: New Results and Analytical Test Cases: Moment Independent Importance Measures. Risk Anal. 31, 404–428. https://doi.org/10.1111/j.1539-6924.2010.01519.x

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

Borgonovo, E., Tarantola, S., 2008. Moment independent and variance-based sensitivity analysis with correlations: An application to the stability of a chemical reactor. Int. J. Chem. Kinet. 40, 687–698. https://doi.org/10.1002/kin.20368

Bratley, P., Fox, B.L., Niederreiter, H., 1992. Implementation and tests of low-discrepancy sequences. ACM Trans. Model. Comput. Simul. 2, 195–213. https://doi.org/10.1145/146382.146385

Burnash, R.J.C., 1995. The NWS River Forecast System - Catchment model, in: Computer Models of Watershed Hydrology. Edited by P. Singh. Water Resources Publications, Highlands Ranch, Colo, pp. 311 – 366.

Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. Environ. Model. Softw. 22, 1509–1518. https://doi.org/10.1016/j.envsoft.2006.10.004

Castaings, W., Borgonovo, E., Morris, M.D., Tarantola, S., 2012. Sampling strategies in density-based sensitivity analysis. Environ. Model. Softw. 38, 13–26. https://doi.org/10.1016/j.envsoft.2012.04.017

Chun, M.-H., Han, S.-J., Tak, N.-I., 2000. An uncertainty importance measure using a distance metric for the change in a cumulative distribution function. Reliab. Eng. Syst. Saf. 70, 313–321. https://doi.org/10.1016/S0951-8320(00)00068-5

Cukier, R.I., Fortuin, C.M., Shuler, K.E., Petschek, A.G., Schaibly, J.H., 1973. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. J. Chem. Phys. 59, 3873–3878. https://doi.org/10.1063/1.1680571

Cuntz, M., Mai, J., Zink, M., Thober, S., Kumar, R., Schäfer, D., Schrön, M., Craven, J., Rakovec, O., Spieler, D., Prykhodko, V., Dalmasso, G., Musuuza, J., Langenberg, B., Attinger, S., Samaniego, L., 2015. Computationally inexpensive identification of noninformative model parameters by sequential screening. Water Resour. Res. 51, 6417–6441. https://doi.org/10.1002/2015WR016907

Dell'Oca, A., Riva, M., Guadagnini, A., 2017. Moment-based metrics for global sensitivity analysis of hydrological systems. Hydrol. Earth Syst. Sci. 21, 6219–6234. https://doi.org/10.5194/hess-21-6219-2017

Demirel, M., Koch, J., Mendiguren, G., Stisen, S., Demirel, M.C., Koch, J., Mendiguren, G., Stisen, S., 2018. Spatial Pattern Oriented Multicriteria Sensitivity Analysis of a Distributed Hydrologic Model. Water 10, 1188. https://doi.org/10.3390/w10091188

Efron, B., Tibshirani, R.J., 1994. An Introduction to the Bootstrap. CRC Press.

Fox, D.B., Koch, D.L., Tester, J.W., 2016. An analytical thermohydraulic model for discretely fractured geothermal reservoirs. Water Resour. Res. 52, 6792–6817. https://doi.org/10.1002/2016WR018666

Francke, T., Baroni, G., Brosinsky, A., Foerster, S., López-Tarazón, J.A., Sommerer, E., Bronstert, A., 2018. What Did Really Improve Our Mesoscale Hydrological Model? A Multidimensional Analysis Based on Real Observations. Water Resour. Res. https://doi.org/10.1029/2018WR022813

Ghasemizade, M., Baroni, G., Abbaspour, K., Schirmer, M., 2017. Combined analysis of time-varying sensitivity and identifiability indices to diagnose the response of a complex

environmental model. Environ. Model. Softw. 88, 22–34. https://doi.org/10.1016/j.envsoft.2016.10.011
Gillies, K., Krone, S.M., Nagler, J.J., Schultz, I.R., 2016. A Computational Model of the Rainbow Trout Hypothalamus-Pituitary-Ovary-Liver Axis. PLOS Comput. Biol. 12, e1004874. https://doi.org/10.1371/journal.pcbi.1004874
Glen, G., Isaacs, K., 2012. Estimating Sobol sensitivity indices using correlations. Environ. Model. Softw. 37, 157–166. https://doi.org/10.1016/j.envsoft.2012.03.014
Guillaume, J.H.A., Jakeman, J.D., Marsili-Libelli, S., Asher, M., Brunner, P., Croke, B., Hill, M.C., Jakeman, A.J., Keesman, K.J., Razavi, S., Stigter, J.D., 2019. Introductory overview of identifiability analysis: A guide to evaluating whether you have the right type of data for your modeling purpose. Environ. Model. Softw. 119, 418–432. https://doi.org/10.1016/j.envsoft.2019.07.007
Guse, B., Pfannerstill, M., Gafurov, A., Fohrer, N., Gupta, H., 2016. Demasking the integrated information of discharge: Advancing sensitivity analysis to consider different hydrological components and their rates of change. Water Resour. Res. https://doi.org/10.1002/2016WR018894
Haghnegahdar, A., Razavi, S., 2017. Insights into sensitivity analysis of Earth and environmental systems models: On the impact of parameter perturbation scale. Environ. Model. Softw. 95, 115–131. https://doi.org/10.1016/j.envsoft.2017.03.031
Haghnegahdar, A., Razavi, S., Yassin, F., Wheater, H., 2017. Multicriteria sensitivity analysis as a diagnostic tool for understanding model behaviour and characterizing model uncertainty. Hydrol. Process. n/a-n/a. https://doi.org/10.1002/hyp.11358
Homma, T., Saltelli, A., 1996. Importance measures in global sensitivity analysis of nonlinear models. Reliab. Eng. Syst. Saf. 52, 1–17. https://doi.org/10.1016/0951-8320(96)00002-6
Hosseini, S.M., Ataie-Ashtiani, B., Simmons, C.T., 2018. Density-based Global Sensitivity Analysis of Sheet-Flow Travel Time: Kinematic Wave-Based Formulations. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2018.02.052
Iman, R.L., 1987. A Matrix-Based Approach to Uncertainty and Sensitivity Analysis for Fault Trees1. Risk Anal. 7, 21–33. https://doi.org/10.1111/j.1539-6924.1987.tb00966.x
Iooss, B., Lemaître, P., 2015. A Review on Global Sensitivity Analysis Methods, in: Dellino, G., Meloni, C. (Eds.), Uncertainty Management in Simulation-Optimization of Complex Systems: Algorithms and Applications, Operations Research/Computer Science Interfaces Series. Springer US, Boston, MA, pp. 101–122. https://doi.org/10.1007/978-1-4899-7547-8_5
Iooss, B., Pujol, A.J. and G., Broto, with contributions from B., Boumhaout, K., Veiga, S.D., Delage, T., Fruth, J., Gilquin, L., Guillaume, J., Gratiet, L.L., Lemaitre, P., Marrel, A., Meynaoui, A., Nelson, B.L., Monari, F., Oomen, R., Rakovec, O., Ramos, B., Roustant, O., Song, E., Staum, J., Sueur, R., Touati, T., Weber, F., 2019. Sensitivity package: Global Sensitivity Analysis of Model Outputs.
Jansen, M.J.W., 1999. Analysis of variance designs for model output. Comput. Phys. Commun. 117, 35–43. https://doi.org/10.1016/S0010-4655(98)00154-4
Kolmogorov, A., 1933. Sulla determinazione empirica di una lgge di distribuzione. Inst Ital Attuari Giorn 4, 83–91.
Kucherenko, S., Feil, B., Shah, N., Mauntz, W., 2011. The identification of model effective dimensions using global sensitivity analysis. Reliab. Eng. Syst. Saf. 96, 440–449. https://doi.org/10.1016/j.ress.2010.11.003

Kucherenko, S., Klymenko, O.V., Shah, N., 2017. Sobol' indices for problems defined in non-rectangular domains. Reliab. Eng. Syst. Saf. 167, 218–231. https://doi.org/10.1016/j.ress.2017.06.001

Kucherenko, S., Rodriguez-Fernandez, M., Pantelides, C., Shah, N., 2009. Monte Carlo evaluation of derivative-based global sensitivity measures. Reliab. Eng. Syst. Saf., Special Issue on Sensitivity Analysis 94, 1135–1148. https://doi.org/10.1016/j.ress.2008.05.006

Kucherenko, S., Song, S., 2017. Different numerical estimators for main effect global sensitivity indices. Reliab. Eng. Syst. Saf. 165, 222–238. https://doi.org/10.1016/j.ress.2017.04.003

Kucherenko, S., Tarantola, S., Annoni, P., 2012. Estimation of global sensitivity indices for models with dependent variables. Comput. Phys. Commun. 183, 937–946. https://doi.org/10.1016/j.cpc.2011.12.020

Kullback, S., Leibler, R.A., 1951. On Information and Sufficiency. Ann. Math. Stat. 22, 79–86. https://doi.org/10.1214/aoms/1177729694

Lewandowski, D., Cooke, R.M., Tebbens, R.J.D., 2007. Sample-based estimation of correlation ratio with polynomial approximation. ACM Trans. Model. Comput. Simul. 18, 1–17. https://doi.org/10.1145/1315575.1315578

Li, C., Mahadevan, S., 2016. An efficient modularized sample-based method to estimate the first-order Sobol′ index. Reliab. Eng. Syst. Saf. 153, 110–121. https://doi.org/10.1016/j.ress.2016.04.012

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

Liu, H., Chen, W., Sudjianto, A., 2006. Relative Entropy Based Method for Probabilistic Sensitivity Analysis in Engineering Design. J. Mech. Des. 128, 326. https://doi.org/10.1115/1.2159025

Lo Piano, S., Ferretti, F., Puy, A., Albrecht, D., Tarantola, S., Saltelli, A., 2017. A new sample-based algorithms to compute the total sensitivity index. ArXiv170305799 Stat.

Mai, J., Tolson, B.A., 2019. Model Variable Augmentation (MVA) for Diagnostic Assessment of Sensitivity Analysis Results. Water Resour. Res. 55, 2631–2651. https://doi.org/10.1029/2018WR023382

Mara, T.A., 2009. Extension of the RBD-FAST method to the computation of global sensitivity indices. Reliab. Eng. Syst. Saf. 94, 1274–1281. https://doi.org/10.1016/j.ress.2009.01.012

Mara, T.A., Belfort, B., Fontaine, V., Younes, A., 2017. Addressing factors fixing setting from given data: A comparison of different methods. Environ. Model. Softw. 87, 29–38. https://doi.org/10.1016/j.envsoft.2016.10.004

Mara, T.A., Joseph, O.R., 2008. Comparison of some efficient methods to evaluate the main effect of computer model factors. J. Stat. Comput. Simul. 78, 167–178. https://doi.org/10.1080/10629360600964454

Mara, T.A., Tarantola, S., Annoni, P., 2015. Non-parametric methods for global sensitivity analysis of model output with dependent inputs. Environ. Model. Softw. 72, 173–183. https://doi.org/10.1016/j.envsoft.2015.07.010

Massmann, C., Holzmann, H., 2012. Analysis of the behavior of a rainfall–runoff model using three global sensitivity analysis methods evaluated at different temporal scales. J. Hydrol. 475, 97–110. https://doi.org/10.1016/j.jhydrol.2012.09.026

Mathevet, T., Michel, C., Andreassian, V., Perrin, C., 2006. A bounded version of the Nash-Sutcliffe criterion for better model assessment on large sets of basins, in: Large Sample

Basin Experiments for Hydrological Model Parameterization: Results of the Model Parameter ExperimentMOPEX. IAHS Publ. 307. AISH, pp. 211–219.
McKay, M.D., Morrison, J.D., Upton, S.C., 1999. Evaluating prediction uncertainty in simulation models. Comput. Phys. Commun. 117, 44–51. https://doi.org/10.1016/S0010-4655(98)00155-6
Mood, A.M., Graybill, F.A., Boes, D.C., 1974. Introduction to the theory of statistics, 3. ed. ed, McGraw-Hill series in probability and statistics. McGraw-Hill, Auckland.
Mora, E.B., Spelling, J., van der Weijde, A.H., 2019. Benchmarking the PAWN distribution-based method against the variance-based method in global sensitivity analysis: Empirical results. Environ. Model. Softw. 122, 104556. https://doi.org/10.1016/j.envsoft.2019.104556
Morris, M.D., 1991. Factorial Sampling Plans for Preliminary Computational Experiments. Technometrics 33, 161. https://doi.org/10.2307/1269043
Nossent, J., Elsen, P., Bauwens, W., 2011. Sobol' sensitivity analysis of a complex environmental model. Environ. Model. Softw. 26, 1515–1525. https://doi.org/10.1016/j.envsoft.2011.08.010
Oakley, J.E., O'Hagan, A., 2004. Probabilistic sensitivity analysis of complex models: a Bayesian approach. J. R. Stat. Soc. Ser. B Stat. Methodol. 66, 751–769. https://doi.org/10.1111/j.1467-9868.2004.05304.x
Owen, A.B., 1995. Randomly Permuted (t,m,s)-Nets and (t, s)-Sequences. Monte Carlo Quasi-Monte Carlo Methods Sci. Comput. 299–317. https://doi.org/10.1007/978-1-4612-2552-2_19
Pappenberger, F., Beven, K.J., Ratto, M., Matgen, P., 2008. Multi-method global sensitivity analysis of flood inundation models. Adv. Water Resour. 31, 1–14. https://doi.org/10.1016/j.advwatres.2007.04.009
Park, C.K., Ahn, K.-I., 1994. A new approach for measuring uncertainty importance and distributional sensitivity in probabilistic safety assessment. Reliab. Eng. Syst. Saf. 46, 253–261. https://doi.org/10.1016/0951-8320(94)90119-8
Parzen, E., 1962. On Estimation of a Probability Density Function and Mode. Ann. Math. Stat. 33, 1065–1076. https://doi.org/10.1214/aoms/1177704472
Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B., Wagener, T., 2016. Sensitivity analysis of environmental models: A systematic review with practical workflow. Environ. Model. Softw. 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. Environ. Model. Softw. 108, 197–207. https://doi.org/10.1016/j.envsoft.2018.07.019
Pianosi, F., Wagener, T., 2016. Understanding the time-varying importance of different uncertainty sources in hydrological modelling using global sensitivity analysis. Hydrol. Process. 30, 3991–4003. https://doi.org/10.1002/hyp.10968
Pianosi, F., Wagener, T., 2015. A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. Environ. Model. Softw. 67, 1–11. https://doi.org/10.1016/j.envsoft.2015.01.004
Pilz, T., Francke, T., Bronstert, A., 2017. lumpR 2.0.0: an R package facilitating landscape discretisation for hillslope-based hydrological models. Geosci. Model Dev. 10, 3001–3023. https://doi.org/10.5194/gmd-10-3001-2017

Plischke, E., 2010. An effective algorithm for computing global sensitivity indices (EASI). Reliab. Eng. Syst. Saf. 95, 354–360. https://doi.org/10.1016/j.ress.2009.11.005
Plischke, E., Borgonovo, E., Smith, C.L., 2013. Global sensitivity measures from given data. Eur. J. Oper. Res. 226, 536–550. https://doi.org/10.1016/j.ejor.2012.11.047
Puy, A., Lo Piano, S., Saltelli, A., 2020. A sensitivity analysis of the PAWN sensitivity index. Environ. Model. Softw. 127, 104679. https://doi.org/10.1016/j.envsoft.2020.104679
R Core Team, 2019. R: A language and environment for statistical computing. [WWW Document]. URL https://www.r-project.org/ (accessed 1.10.20).
Ratto, M., Pagano, A., 2010. Using recursive algorithms for the efficient identification of smoothing spline ANOVA models. AStA Adv. Stat. Anal. 94, 367–388. https://doi.org/10.1007/s10182-010-0148-8
Ratto, M., Young, P.C., Romanowicz, R., Pappenberger, F., Saltelli, A., Pagano, A., 2007. Uncertainty, sensitivity analysis and the role of data based mechanistic modeling in hydrology. Hydrol. Earth Syst. Sci. 11, 1249–1266. https://doi.org/10.5194/hess-11-1249-2007
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: A Critical Look at Sensitivity Analysis. Water Resour. Res. 51, 3070–3092. https://doi.org/10.1002/2014WR016527
Reusser, D.E., Buytaert, W., Zehe, E., 2011. Temporal dynamics of model parameter sensitivity for computationally expensive models with the Fourier amplitude sensitivity test. Water Resour. Res. 47. https://doi.org/10.1029/2010WR009947
Rosolem, R., Gupta, H.V., Shuttleworth, W.J., Zeng, X., de Gonçalves, L.G.G., 2012. A fully multiple-criteria implementation of the Sobolʹ method for parameter sensitivity analysis. J. Geophys. Res. 117. https://doi.org/10.1029/2011JD016355
Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., Wu, Q., 2019. Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. Environ. Model. Softw. 114, 29–39. https://doi.org/10.1016/j.envsoft.2019.01.012
Saltelli, A., Annoni, P., 2010. How to avoid a perfunctory sensitivity analysis. Environ. Model. Softw. 25, 1508–1517. https://doi.org/10.1016/j.envsoft.2010.04.012
Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., Tarantola, S., 2010. Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. Comput. Phys. Commun. 181, 259–270. https://doi.org/10.1016/j.cpc.2009.09.018
Saltelli, A., Chan, K., Scott, E.M., 2000. Sensitivity Analysis, 1st ed. Wiley.
Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S., 2008. Global Sensitivity Analysis. The Primer. John Wiley & Sons, Ltd, Chichester, UK.
Saltelli, A., Tarantola, S., 2002. On the Relative Importance of Input Factors in Mathematical Models: Safety Assessment for Nuclear Waste Disposal. J. Am. Stat. Assoc. 97, 702–709. https://doi.org/10.1198/016214502388618447
Saltelli, A., Tarantola, S., Chan, K.P.-S., 1999. A Quantitative Model-Independent Method for Global Sensitivity Analysis of Model Output. Technometrics 41, 39–56. https://doi.org/10.1080/00401706.1999.10485594

Sarrazin, F., Pianosi, F., Wagener, T., 2016. Global Sensitivity Analysis of environmental models: Convergence and validation. Environ. Model. Softw. 79, 135–152. https://doi.org/10.1016/j.envsoft.2016.02.005

Savage, J.T.S., Pianosi, F., Bates, P., Freer, J., Wagener, T., 2016. Quantifying the importance of spatial resolution and other factors through global sensitivity analysis of a flood inundation model. Water Resour. Res. 52, 9146–9163. https://doi.org/10.1002/2015WR018198

Schürz, C., Hollosi, B., Matulla, C., Pressl, A., Ertl, T., Schulz, K., Mehdi, B., 2019. A comprehensive sensitivity and uncertainty analysis for discharge and nitrate-nitrogen loads involving multiple discrete model inputs under future changing conditions. Hydrol. Earth Syst. Sci. 23, 1211–1244. https://doi.org/10.5194/hess-23-1211-2019

Sedighian, S., Kim, S.H., Cho, S.Y., Kim, M.J., Kim, D.H., Cha, D.W., 2015. Parameter Ranking System of Indoor Radon Concentration in South Korea, Case Studies: Dokdo Island, Yang Pyeong and NaeGi. Int J Env. Res 9, 1233–1236.

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

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

Sobol′, I., 2001. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. Math. Comput. Simul. 55, 271–280. https://doi.org/10.1016/S0378-4754(00)00270-6

Sobol, I.M., 1976. Uniformly distributed sequences with an additional uniform property. USSR Comput. Math. Math. Phys. 16, 236–242. https://doi.org/10.1016/0041-5553(76)90154-3

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

Strong, M., Oakley, J.E., 2013. An Efficient Method for Computing Single-Parameter Partial Expected Value of Perfect Information. Med. Decis. Making 33, 755–766. https://doi.org/10.1177/0272989X12465123

Strong, M., Oakley, J.E., Chilcott, J., 2012. Managing structural uncertainty in health economic decision models: a discrepancy approach. J. R. Stat. Soc. Ser. C Appl. Stat. 61, 25–45. https://doi.org/10.1111/j.1467-9876.2011.01014.x

Tang, Y., Reed, P., van Werkhoven, K., Wagener, T., 2007. Advancing the identification and evaluation of distributed rainfall-runoff models using global sensitivity analysis. Water Resour. Res. 43. https://doi.org/10.1029/2006WR005813

Tarantola, S., Becker, W., Zeitz, D., 2012. A comparison of two sampling methods for global sensitivity analysis. Comput. Phys. Commun. 183, 1061–1072. https://doi.org/10.1016/j.cpc.2011.12.015

Tarantola, S., Gatelli, D., Mara, T.A., 2006. Random balance designs for the estimation of first order global sensitivity indices. Reliab. Eng. Syst. Saf. 91, 717–727. https://doi.org/10.1016/j.ress.2005.06.003

Tarantola, S., Mara, T.A., 2017. Variance-based sensitivity indices of computer models with dependent inputs: the fourier amplitude sensitivity test. Int. J. Uncertain. Quantif. 7, 511–523. https://doi.org/10.1615/Int.J.UncertaintyQuantification.2017020291

Uusitalo, L., Lehikoinen, A., Helle, I., Myrberg, K., 2015. An overview of methods to evaluate uncertainty of deterministic models in decision support. Environ. Model. Softw. 63, 24–31. https://doi.org/10.1016/j.envsoft.2014.09.017

van Werkhoven, K., Wagener, T., Reed, P., Tang, Y., 2009. Sensitivity-guided reduction of parametric dimensionality for multi-objective calibration of watershed models. Adv. Water Resour. 32, 1154–1169. https://doi.org/10.1016/j.advwatres.2009.03.002

van Werkhoven, K., Wagener, T., Reed, P., Tang, Y., 2008. Characterization of watershed model behavior across a hydroclimatic gradient. Water Resour. Res. 44. https://doi.org/10.1029/2007WR006271

Veiga, S.D., 2015. Global sensitivity analysis with dependence measures. J. Stat. Comput. Simul. 85, 1283–1305. https://doi.org/10.1080/00949655.2014.945932

Wagener, T., Pianosi, F., 2019. What has Global Sensitivity Analysis ever done for us? A systematic review to support scientific advancement and to inform policy-making in earth system modelling. Earth-Sci. Rev. 194, 1–18. https://doi.org/10.1016/j.earscirev.2019.04.006

Wainwright, H.M., Finsterle, S., Jung, Y., Zhou, Q., Birkholzer, J.T., 2014. Making sense of global sensitivity analyses. Comput. Geosci. 65, 84–94. https://doi.org/10.1016/j.cageo.2013.06.006

Wei, P., Lu, Z., Song, J., 2015. Variable importance analysis: A comprehensive review. Reliab. Eng. Syst. Saf. 142, 399–432. https://doi.org/10.1016/j.ress.2015.05.018

Xie, H., Shen, Z., Chen, L., Qiu, J., Dong, J., 2017. Time-varying sensitivity analysis of hydrologic and sediment parameters at multiple timescales: Implications for conservation practices. Sci. Total Environ. 598, 353–364. https://doi.org/10.1016/j.scitotenv.2017.04.074

Zadeh, F.K., Nossent, J., Sarrazin, F., Pianosi, F., van Griensven, A., Wagener, T., Bauwens, W., 2017. Comparison of variance-based and moment-independent global sensitivity analysis approaches by application to the SWAT model. Environ. Model. Softw. 91, 210–222. https://doi.org/10.1016/j.envsoft.2017.02.001

Zhao, S., Zhang, W., Lin, X., Li, J., Song, D., He, X., 2015. Effect of parameters correlation on uncertainty and sensitivity in dynamic thermal analysis of thermal protection blanket in service. Int. J. Therm. Sci. 87, 158–168. https://doi.org/10.1016/j.ijthermalsci.2014.08.016