Variance-based sensitivity indices of computer models with dependent inputs: The Fourier Amplitude Sensitivity Test

Stefano Tarantola, Thierry A. Mara

VARIANCE-BASED SENSITIVITY INDICES OF COMPUTER MODELS WITH DEPENDENT INPUTS: THE FOURIER AMPLITUDE SENSITIVITY TEST

S. Tarantola[1] & T.A. Mara[2,3,*]

[1]*Directorate for Energy, Transport and Climate, Joint Research Centre, European Commission, Ispra (VA), Italy, 21027*

[2]*PIMENT, EA 4518, Université de La Réunion, FST, 15 Avenue René Cassin, 97715 Saint-Denis, Réunion*

[3]*Directorate for Modelling, Indicators and Impact Evaluation, Joint Research Centre, European Commission, Ispra (VA), Italy, 21027*

*Address all correspondence to: T.A. Mara, PIMENT, EA 4518, Université de La Réunion, FST, 15 Avenue René Cassin, 97715 Saint-Denis, Réunion, E-mail: thierry.mara@univ-reunion.fr

Several methods are proposed in the literature to perform the global sensitivity analysis of computer models with independent inputs. Only a few allow for treating the case of dependent inputs. In the present work, we investigate how to compute variance-based sensitivity indices with the Fourier amplitude sensitivity test. This can be achieved with the help of the inverse Rosenblatt transformation or the inverse Nataf transformation. We illustrate so on two distinct benchmarks. As compared to the recent Monte Carlo based approaches recently proposed by the same authors in [1], the new approaches allow to divide by two the computational effort to assess the entire set of first-order and total-order variance-based sensitivity indices.

KEY WORDS: *Fourier amplitude sensitivity test; inverse Rosenblatt transformation; inverse Nataf transformation; variance-based sensitivity indices; dependent contributions; independent contributions*

1. INTRODUCTION

Good practice in computer model simulations requires that the uncertainties in the process under study be acknowledged. This is achieved by treating the model scalar inputs like random variables and functional inputs (temporally or spatially dependent) like stochastic fields [2,3]. Subsequently, the model responses are also random and their uncertainties can be assessed, among others, via Monte Carlo simulations. Global Sensitivity Analysis (GSA) of computer model responses usually accompanies the uncertainty assessment. It aims to point out the set of input factors that mainly contributes to the model responses uncertainty. Such an information is essential to indicate the modellers what is the subset of inputs upon which they should concentrate their effort in future works in order to obtain more accurate and relevant predictions.

Several methods exist to perform GSA of model outputs when the inputs are independent [4]. The choice of the method to employ depends on the importance measure to be estimated. Two types of quantitative sensitivity measures are of particular interest: the variance-based sensitivity measures [5] and the moment-independent sensitivity measures [6,7]. Numerical methods to assess these sensitivity measures can be classified as: non-parametric Monte Carlo approaches (among others, [8,9]), parametric spectral methods (likewise the Fourier amplitude sensitivity test [10,11]) or emulator-based approaches (e.g. [12,13]). Amongst the variance-based sensitivity measures, practitioners mostly focus on the first-order sensitivity index (also called correlation ratio [14,15]), that measures the marginal effect of one single input, and the total sensitivity index that accounts for the both marginal effects and interaction effects of the input of interest with the other ones [16]. The interested readers are referred to [17] for more details about the usefulness of these sensitivity measures in some GSA settings.

Performing sensitivity analysis of computer models with dependent inputs is more challenging. Just like models can be of different natures (linear, non-linear, additive and non-additive), the dependence structure amongst the inputs can be of different natures too (linear, non-linear, pairwise and non-pairwise). In case of linear dependence structure, the inputs are said correlated. But in many cases, the dependence structure can be more complex. The latter is embedded in the joint probability distribution of the inputs and in their copula.

The authors in [18] propose to distinguish the influence of an input due to its correlation with the other variables, and the influence that is not due to its possible correlations. The idea is that if an input is not influential on its own (without accounting for its correlations with the other inputs) then it can be concluded that it is a spurious input, only influential because of its correlations. This concept was later on extended to the variance-based sensitivity measures in [1,19] and to the density-based sensitivity measure in [20].

Kucherenko et al. [21] generalize the first-order and total Sobol' indices for the case of computer models with dependent inputs. A non-parametric Monte Carlo approach is proposed to evaluate them based on the theory of Gaussian copula. Simultaneously, in [19] four new variance-based sensitivity measures are defined: two that account for the dependencies of an individual input with the other ones (the authors called them the first-order and total full sensitivity indices respectively) and two that do not account for the dependencies (called uncorrelated/independent first-order and total sensitivity indices respectively). Moreover, the authors show that the first-order sensitivity index generalized in [21] is the same as their *full first-order sensitivity index* (that accounts for correlations) while the generalized total sensitivity index of [21] is the same as their *independent total sensitivity index* (that does not account for correlations). Furthermore, two new variance-based indices namely the *full total sensitivity index* and *independent first-order sensitivity index* are also introduced in [19]. These four different sensitivity measures are discussed in the next Section.

In [1], the new variance-based sensitivity measures are formally defined (see Eqs.(1-4)) and two non-parametric methods are proposed to estimate them. The first method uses sampling strategies in conjunction with the inverse Rosenblatt transformation [22] while the second method employs the inverse Nataf transformation [23]. The latter corresponds to the procedure of Iman and Conover when the target dependence structure is the correlation matrix instead of the rank correlation matrix as in their original paper [24] and the Gaussian-copula-based approach employed in [21] for generating correlated samples. In [19], the polynomial chaos expansion method is employed to evaluate the four variance-based sensitivity measures. The latter is very efficient because it only requires one single input/output random sample. But it also requires a procedure to make the input sample at hand independent. For this purpose, the authors derive a specific procedure only valid for some specific correlation structures (like pairwise linear and non-linear correlations).

Another idea is proposed in [25] that consists of distinguishing the *correlative* effect of a given input onto the computer model response from its *structural* effect. This can be achieved, for instance, by first identifying the model structure via an ANalysis Of VAriance decomposition in the Sobol' sense (for ANOVA see [8]). For this purpose independence of the inputs is mandatory. Then, the structural and correlative effects can be inferred by analyzing the covariance structure of the ANOVA decomposition stemming from the correlation structure amongst the inputs. Caniou names this approach the ANCOVA (acronym for ANalysis of COVAriance) decomposition and the author uses the polynomial chaos expansion for prior ANOVA decomposition [26]. There are other approaches proposed in the literature that are not discussed here (the interested reader can refer to [27,28], among others).

So far, no numerical approach has been proposed in the literature to compute the four variance-based sensitivity indices introduced in [19] with the Fourier Amplitude Sensitivity Test (FAST). One can cite [29] in which the author derives a cheap FAST-based approach to evaluate the independent first-order sensitivity index and the full first-order sensitivity index. One can also mention the early work of Xu and Gertner [30] that allows to assess the full first-order sensitivity index. Therefore, the proposed approaches are unable to account for interactions in computer models with dependent inputs. The present work aims to fill this gap. To this end, we show that the extended FAST [11] in conjunction with the inverse Rosenblatt transformation [22] or the inverse Nataf transformation [23] allows to compute the four sensitivity indices.

The paper is organized as follows: we start by recalling the definitions of the variance-based sensitivity indices in the case of dependent inputs in Section 2. The link with the law of total variance is made in Section 3. In Section 4, we discuss the two transformations, namely, the inverse Rosenblatt transformation and the inverse Nataf transformation. Section 5 recalls the classical FAST method for models with independent inputs and then its extension to the case of dependent inputs is described. In Section 6, the new approaches are tested before concluding.

2. DEFINITION OF THE SENSITIVITY INDICES

Let $f(\boldsymbol{x})$ be a square integrable function over an n-dimensional space where $\boldsymbol{x} = \{x_1, \cdots, x_n\}$ a continuous random vector defined by a joint probability density function $p(\boldsymbol{x})$. The scalar $f(\boldsymbol{x})$ can be regarded, without loss of generality, as the scalar response of a computer model to the input set. In the sequel, we set $\boldsymbol{x} = (\boldsymbol{x}_1, \boldsymbol{x}_2)$ with $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ two non-empty subsets of $\boldsymbol{x}$. The importance of $\boldsymbol{x}_1$ for $f(\boldsymbol{x})$ can be measured with the variance-based sensitivity indices (also called Sobol' indices,[5]). The Sobol' indices can either measure the amount of variance of $f(\boldsymbol{x})$ due to $\boldsymbol{x}_1$ alone, or measure the amount of variance that also includes its interactions with $\boldsymbol{x}_2$. Besides, when $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are dependent, it is possible to distinguish the two possible types of contribution of $\boldsymbol{x}_1$ to the variance of $f(\boldsymbol{x})$: i) the independent contribution that does not account for the dependence of $\boldsymbol{x}_1$ with $\boldsymbol{x}_2$ and, ii) the full contribution that accounts for the dependence between $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$. To the authors' best knowledge, this concept was first introduced in [18] although the partial correlation coefficient of [31] is related to this concept. The variance-based sensitivity measures were defined in [19] for correlated variables and recently generalized in [1]. They are defined as follows:

$$
\begin{aligned}
S_{\boldsymbol{x}_1} &= \frac{\mathbb{V}\left[\mathbb{E}\left[f(\boldsymbol{x})|\boldsymbol{x}_1\right]\right]}{\mathbb{V}\left[f(\boldsymbol{x})\right]}, && (1)\\
ST^{ind}_{\boldsymbol{x}_1} &= \frac{\mathbb{E}\left[\mathbb{V}\left[f(\boldsymbol{x})|\boldsymbol{x}_2\right]\right]}{\mathbb{V}\left[f(\boldsymbol{x})\right]}, && (2)\\
S^{ind}_{\boldsymbol{x}_1} &= \frac{\mathbb{V}\left[\mathbb{E}\left[f(\boldsymbol{x})|\bar{\boldsymbol{x}}_1)\right]\right]}{\mathbb{V}\left[f(\boldsymbol{x})\right]}, && (3)\\
ST_{\boldsymbol{x}_1} &= \frac{\mathbb{E}\left[\mathbb{V}\left[\left[f(\boldsymbol{x})|\bar{\boldsymbol{x}}_2\right]\right]\right.}{\mathbb{V}\left[f(\boldsymbol{x})\right]}, && (4)
\end{aligned}
$$

where $\mathbb{V}\left[\cdot\right]$ is the variance operator, $\mathbb{E}\left[\cdot\right]$ is the mathematical expectation while $\mathbb{V}\left[\cdot|\cdot\right]$ and $\mathbb{E}\left[\cdot|\cdot\right]$ are the conditional variance and expectation respectively.

The variables with an overbar are conditional variables, therefore: $\bar{\boldsymbol{x}}_1 \sim p_{\boldsymbol{x}_1|\boldsymbol{x}_2}(\boldsymbol{x}_1|\boldsymbol{x}_2)$ and $\bar{\boldsymbol{x}}_2 \sim p_{\boldsymbol{x}_2|\boldsymbol{x}_1}(\boldsymbol{x}_2|\boldsymbol{x}_1)$. While the first two sensitivity indices Eqs.(1-2) are the classical definitions of the Sobol' indices [21], the last two are only defined for dependent input variables. All these indices are scaled within $[0, 1]$ and we have $S_{\boldsymbol{x}_1} \leq ST_{\boldsymbol{x}_1}$, $S^{ind}_{\boldsymbol{x}_1} \leq ST^{ind}_{\boldsymbol{x}_1}$.

The full first-order sensitivity index $S_{\boldsymbol{x}_1}$ measures the amount of variance of $f(\boldsymbol{x})$ due to $\boldsymbol{x}_1$ and its dependence with $\boldsymbol{x}_2$ but does not include the interactions of $\boldsymbol{x}_1$ with $\boldsymbol{x}_2$. The full total sensitivity index $ST_{\boldsymbol{x}_1}$ does account for these two types of contributions (dependence and interaction). The independent first-order sensitivity index $S^{ind}_{\boldsymbol{x}_1}$ measures the contribution of $\boldsymbol{x}_1$ by ignoring its correlations and interactions with $\boldsymbol{x}_2$ while $ST^{ind}_{\boldsymbol{x}_1}$ accounts for interactions and ignores correlations. An input x_i can contribute to the model response variance only because of its strong correlations with the other inputs. In this case, we shall find $ST_{x_i} \geq 0$ and $ST^{ind}_{x_i} = 0$.

The authors in [1] propose two non-parametric methods to evaluate these sensitivity indices. The first approach consists of generating random samples of the dependent variables from the inverse Rosenblatt transformation while the second one uses the sampling technique of Iman and Conover [24]. The first approach is advisable when the conditional densities $p_{\boldsymbol{x}_1|\boldsymbol{x}_2}$ and $p_{\boldsymbol{x}_2|\boldsymbol{x}_1}$ are known while the procedure of Iman and Conover (IC) is to be preferred when the marginal densities and the rank correlation structure of the input variables are known. We note that, when the target dependence structure is the correlation matrix, the IC procedure is equivalent to the Gaussian copula approach used in [21] and the inverse Nataf transformation described in the present work.

3. LINK WITH THE LAW OF TOTAL VARIANCE

It is usual to define the Sobol' indices from the law of total variance, namely

$$\mathbb{V}\left[f(\boldsymbol{x})\right] = \mathbb{V}\left[\mathbb{E}\left[f(\boldsymbol{x})|\boldsymbol{x}_1\right]\right] + \mathbb{E}\left[\mathbb{V}\left[f(\boldsymbol{x})|\boldsymbol{x}_1\right]\right]. \tag{5}$$

Dividing this equation by the left-hand side term yields,

$$1 = S_{\boldsymbol{x}_1} + ST^{ind}_{\boldsymbol{x}_2}. \tag{6}$$

In principal, Eq. (5) should be written as follows,

$$\mathbb{V}_{\boldsymbol{x}}\left[f(\boldsymbol{x})\right] = \mathbb{V}_{\boldsymbol{x}_1}\left[\mathbb{E}_{\bar{\boldsymbol{x}}_2}\left[f(\boldsymbol{x})|\boldsymbol{x}_1\right]\right] + \mathbb{E}_{\boldsymbol{x}_1}\left[\mathbb{V}_{\bar{\boldsymbol{x}}_2}\left[f(\boldsymbol{x})|\boldsymbol{x}_1\right]\right] \tag{7}$$

with the variables over which the conditional operators are applied indicated in subscript. But in the case of independent variables, it is not necessary to indicate so and Eq. (5) (without subscript) is adopted for the sake of simplicity. However, such a precision is necessary when the variables are dependent because of the axiom of conditional probabilities: $p_{\boldsymbol{x}}(\boldsymbol{x}) = p_{\boldsymbol{x}_1}(\boldsymbol{x}_1)p_{\boldsymbol{x}_2|\boldsymbol{x}_1}(\boldsymbol{x}_2|\boldsymbol{x}_1) = p_{\boldsymbol{x}_2}(\boldsymbol{x}_2)p_{\boldsymbol{x}_1|\boldsymbol{x}_2}(\boldsymbol{x}_1|\boldsymbol{x}_2)$. In effect, setting $\bar{\boldsymbol{x}}_1 = \boldsymbol{x}_1|\boldsymbol{x}_2$ and $\bar{\boldsymbol{x}}_2 = \boldsymbol{x}_2|\boldsymbol{x}_1$, one can also write the law of total variance as follows,

$$\mathbb{V}_{\boldsymbol{x}}\left[f(\boldsymbol{x})\right] = \mathbb{V}_{\bar{\boldsymbol{x}}_1}\left[\mathbb{E}_{\boldsymbol{x}_2}\left[f(\boldsymbol{x})|\bar{\boldsymbol{x}}_1\right]\right] + \mathbb{E}_{\bar{\boldsymbol{x}}_1}\left[\mathbb{V}_{\boldsymbol{x}_2}\left[f(\boldsymbol{x})|\bar{\boldsymbol{x}}_1\right]\right]. \tag{8}$$

Normalizing the latter equation, yields,

$$1 = S^{ind}_{\boldsymbol{x}_1} + ST_{\boldsymbol{x}_2}. \tag{9}$$

The two different versions of the law of total variance hold because according to the axiom of conditional probabilities, $\boldsymbol{x}_1$ and $\bar{\boldsymbol{x}}_2$ (resp. $\bar{\boldsymbol{x}}_1$ and $\boldsymbol{x}_2$) are independent random vectors (i.e. $p(\boldsymbol{x}) = p_{\boldsymbol{x}_1}(\boldsymbol{x}_1)p_{\bar{\boldsymbol{x}}_2}(\bar{\boldsymbol{x}}_2)$). Therefore, for the sake of clarity, the definitions of the Sobol' indices in the case of dependent input variables should be written as follows,

$$\begin{aligned}
S_{\boldsymbol{x}_1} &= \frac{\mathbb{V}_{\boldsymbol{x}_1}\left[\mathbb{E}_{\bar{\boldsymbol{x}}_2}\left[f(\boldsymbol{x})|\boldsymbol{x}_1\right]\right]}{\mathbb{V}\left[f(\boldsymbol{x})\right]}, &(10)\\
ST^{ind}_{\boldsymbol{x}_1} &= \frac{\mathbb{E}_{\boldsymbol{x}_2}\left[\mathbb{V}_{\bar{\boldsymbol{x}}_1}\left[f(\boldsymbol{x})|\boldsymbol{x}_2\right]\right]}{\mathbb{V}\left[f(\boldsymbol{x})\right]}, &(11)\\
S^{ind}_{\boldsymbol{x}_1} &= \frac{\mathbb{V}_{\bar{\boldsymbol{x}}_1}\left[\mathbb{E}_{\boldsymbol{x}_2}\left[f(\boldsymbol{x})|\bar{\boldsymbol{x}}_1)\right]\right]}{\mathbb{V}\left[f(\boldsymbol{x})\right]}, &(12)\\
ST_{\boldsymbol{x}_1} &= \frac{\mathbb{E}_{\bar{\boldsymbol{x}}_2}\left[\mathbb{V}_{\boldsymbol{x}_1}\left[\left[f(\boldsymbol{x})|\bar{\boldsymbol{x}}_2\right]\right]\right]}{\mathbb{V}\left[f(\boldsymbol{x})\right]}. &(13)
\end{aligned}$$

4. TWO PROBABILISTIC TRANSFORMATIONS

4.1 The inverse Rosenblatt transformation

It is shown in [1] that the Rosenblatt transformation [22] is the key for estimating the four sensitivity indices defined in Eqs.(1-4). Indeed, the Inverse Rosenblatt Transformation (IRT) provides a set of dependent variables $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ from a set of independent random vectors $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ uniformly distributed over the unit hypercube $\mathbb{K}^n = [0, 1]^n$. Assuming that $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is a vector of continuous random variables, the inverse Rosenblatt transformation writes,

$$\begin{cases} \boldsymbol{x}_2 = F_{\boldsymbol{x}_2}^{-1}(\boldsymbol{u}_2) \\ \bar{\boldsymbol{x}}_1 = F_{\boldsymbol{x}_1|\boldsymbol{x}_2}^{-1}(\boldsymbol{u}_1|\boldsymbol{u}_2) \end{cases} \tag{14}$$

where, $F_{\boldsymbol{x}_2}^{-1}$ is the inverse cumulative density function of $\boldsymbol{x}_2$ (that is, $p_{\boldsymbol{x}_2} = \mathrm{d}F_{\boldsymbol{x}_2}/\mathrm{d}\boldsymbol{x}_2$) and $F_{\boldsymbol{x}_1|\boldsymbol{x}_2}^{-1}$ is the one of $\bar{\boldsymbol{x}}_1$ (that is, $p_{\boldsymbol{x}_1|\boldsymbol{x}_2} = \partial F_{\boldsymbol{x}_1|\boldsymbol{x}_2}/\partial \boldsymbol{x}_1$). IRT simply exploits the axiom of conditional probabilities: $p(\boldsymbol{x}_1, \boldsymbol{x}_2) = p_{\boldsymbol{x}_1|\boldsymbol{x}_2}(\boldsymbol{x}_1|\boldsymbol{x}_2)p_{\boldsymbol{x}_2}(\boldsymbol{x}_2)$ but is not unique. Indeed, as already aforementioned, one can also write $p(\boldsymbol{x}_1, \boldsymbol{x}_2) = p_{\boldsymbol{x}_2|\boldsymbol{x}_1}(\boldsymbol{x}_2|\boldsymbol{x}_1)$ $p_{\boldsymbol{x}_1}(\boldsymbol{x}_1)$, which yields the following transformation,

$$\begin{cases} \boldsymbol{x}_1 = F_{\boldsymbol{x}_1}^{-1}(\boldsymbol{u}_1) \\ \bar{\boldsymbol{x}}_2 = F_{\boldsymbol{x}_2|\boldsymbol{x}_1}^{-1}(\boldsymbol{u}_2|\boldsymbol{u}_1) \end{cases} . \tag{15}$$

To generate the sets $(\bar{\boldsymbol{x}}_1, \boldsymbol{x}_2)$ and $(\boldsymbol{x}_1, \bar{\boldsymbol{x}}_2)$ from these two transformations, an independent set $(\boldsymbol{u}_1, \boldsymbol{u}_2)$ uniformly distributed over the unit hypercube is required. This is efficiently performed with, for instance the LP_τ sequences of [32]. Using $(\bar{\boldsymbol{x}}_1, \boldsymbol{x}_2)$ and $(\boldsymbol{x}_1, \bar{\boldsymbol{x}}_2)$, one can estimate the sensitivity measures defined in Eqs.(1-4) as shown in [1].

4.2 The inverse Nataf transformation

The application of IRT requires the knowledge of the conditional probability densities. In many situations, only the individual cumulative densities (i.e. $F_{x_j}(x_j), \forall j \in [\![1, n]\!]$) and the correlation matrix $\mathbf{R}_{\boldsymbol{x}}$ are known. In this case, the inverse Nataf transformation (INT) is more suitable than IRT to generate samples with correlation approaching the desired matrix $\mathbf{R}_{\boldsymbol{x}}$. It is worth mentioning that the procedure of Iman and Conover [24] and the Gaussian copula-based approach employed in [21] are tantamount to INT. The latter uses a set of correlated standard normal variables $\boldsymbol{z}^{\boldsymbol{c}} = (z_1^c, \ldots, z_n^c)$ with correlation matrix $\mathbf{R}_{\boldsymbol{z}}$ and generates the desired correlated random variables as follows,

$$x_j = F_{x_j}^{-1}(\Phi(z_j^c)), \forall j = 1, \ldots, n \tag{16}$$

where Φ is the cumulative density function of the standard normal variable. We note that transformation (16) implies that x_j and z_j^c have identical ranking. Therefore, INT is equivalent to the procedure of [24]. However, INT and the original IC procedure differ in the fact that the former generates samples of $\boldsymbol{x}$ w.r.t. the correlation matrix $\mathbf{R}_{\boldsymbol{x}}$ whereas the latter generates samples w.r.t. the rank correlation matrix of $\boldsymbol{x}$. Consequently, INT is a bit more complicated than the original IC procedure. We note that Eq. (16) is also employed in the theory of Gaussian copula (see [21]).

Sampling $\boldsymbol{x}$ with INT requires to generate $\boldsymbol{z}^c$ with the desired correlation matrix $\mathbf{R}_{\boldsymbol{z}}$. This is achieved with the Cholesky transformation. Let us denote by $\mathbf{L}$ the lower triangular matrix such that $\mathbf{R}_{\boldsymbol{z}} = \mathbf{L}\mathbf{L}^T$, the superscript T stands for the transpose operator. This decomposition is possible because $\mathbf{R}_{\boldsymbol{z}}$ is positive-semidefinite. Then, from an independent standard normal vector $\boldsymbol{z}$, the correlated standard normal vector $\boldsymbol{z}^c$ is obtained as follows,

$$\boldsymbol{z}^c = \boldsymbol{z}\mathbf{L}^T. \tag{17}$$

The issue with this approach is to find $\mathbf{R}_{\boldsymbol{z}}$ such that $\boldsymbol{x}$ has the desired correlation matrix $\mathbf{R}_{\boldsymbol{x}}$. This can be achieved with an optimization scheme in which $\mathbf{R}_{\boldsymbol{z}}$ is iteratively adjusted until the correlation matrix of $\boldsymbol{x}$ satisfactorily matches $\mathbf{R}_{\boldsymbol{x}}$ [33]. The relationship between $\mathbf{R}_{\boldsymbol{z}}$ and $\mathbf{R}_{\boldsymbol{x}}$ is discussed for some densities in [34].

5. THE FOURIER AMPLITUDE SENSITIVITY TEST

5.1 The classical FAST for independent variables

FAST was introduced in [10] to compute the individual first-order sensitivity index for models with independent inputs (Eq. (1) with $\boldsymbol{x}_1 = x_i$). In FAST input values are sampled over a periodic curve that explores the input space. Each input is associated with a distinct integer frequency. The periodic sampled values are propagated through the model. Then, the Fourier transform of the model output is computed. The Parseval-Plancherel theorem allows to compute the variance-based sensitivity indices via the Fourier coefficients evaluated at specific frequencies. The individual first-order sensitivity index of a given input uses the Fourier coefficients of the associated frequency and its higher harmonics (see Algorithm 1).

The set $\boldsymbol{\omega} = \{\omega_1, \ldots, \omega_n\}$ of integer frequencies must be chosen in order to avoid interferences between higher harmonics. This is possible up to a given interference order M. Cukier et al. [35] provides an algorithm to generate such a set of frequencies for prescribed M and n. In practice, N draws of the input values are generated by discretizing s as follows: $s_k = \frac{2k\pi}{N}$, $k = 1, \ldots, N$. With classical FAST, all first-order sensitivity indices can be obtained with only one set of model runs N. The Nyquist criterion imposes that $N \geq 2M \times \max(\boldsymbol{\omega}) + 1$. Therefore, the dimension of the model n and the choice of the interference factor M, considerably impact the number of model runs N and also complicate the choice of an interference-free set of frequencies. To circumvent this problem, the random balance design trick of [36] was extended to FAST in [37]. Algorithm 1 describes the steps to perform the classical FAST approach.

Algorithm 1: The classical FAST procedure

1. Set $u_j(s) = \frac{1}{2} + \frac{1}{\pi}\arcsin(\sin(\omega_j s + \varphi_j))\ \forall j \in [\![1, n]\!]$, with s varying uniformly over $(0, 2\pi]$, $\varphi_j \in (0, 2\pi]$ is a randomly chosen shift parameter and ω_j is an integer frequency
2. Perform the transformation $x_j = F^{-1}(u_j), \forall j = 1, \ldots, n$ to get $x_j \sim p_{x_j}(x_j)$
3. Evaluate $f(\boldsymbol{x}(s))$ and compute the Fourier coefficients,
$$c_\omega = \frac{1}{\pi}\int_0^{2\pi} f(\boldsymbol{x}(s))\mathrm{e}^{-i\omega s}\mathrm{d}s, \forall \omega \in \mathbb{N}^* \tag{18}$$
4. Compute the first-order sensitivity indices,
$$S_{x_i}^{FAST} = \frac{\sum_{l=1}^{+\infty} |c_{l\omega_i}|^2}{\sum_{\omega=1}^{+\infty} |c_\omega|^2}, \forall i = 1, \ldots, n. \tag{19}$$

To compute the total sensitivity index of x_i, that is ST_{x_i}, Saltelli et al. [11] propose to assign a high frequency to x_i (typically $\omega_i = 2M \times \max(\boldsymbol{\omega}_{\sim i})$, where $\boldsymbol{\omega}_{\sim i} = \boldsymbol{\omega}/\omega_i$) and small values to the other frequencies $\boldsymbol{\omega}_{\sim i}$. These latter do not need to be free of interferences although recommended. Consequently, in the Fourier spectrum of the model response the amount of variance attributed to x_i (including its marginal effects and its interactions with the other inputs) is localized in the high frequency range ($\omega > M \times \max(\boldsymbol{\omega}_{\sim i})$). The total sensitivity index is estimated as follows:
$$\hat{ST}_{x_i} = \frac{\sum_{\omega=\frac{\omega_i}{2}+1}^{N/2} |\hat{c}_\omega|^2}{\sum_{\omega=1}^{N/2} |\hat{c}_\omega|^2} \tag{20}$$
with $\hat{c}_\omega = \frac{2}{N}\sum_{k=1}^{N} f(\boldsymbol{x}(s_k))\mathrm{e}^{-i\omega s_k}$ an estimator of Eq. (18).

The drawback of the proposed approach (called EFAST) is that the computational effort to estimate (S_{x_i}, ST_{x_i}) is high since the Nyquist criterion imposes that $N > 2M\omega_i$. But, S_{x_i} can be estimated simultaneously with ST_{x_i} (i.e. with no extra cost). Hence, $n \times N$ model runs are necessary to compute all individual first-order and total sensitivity indices. The author in [38] proposes a slight different version of EFAST that do not alleviate much the computational burden, but allows for the calculation of the total sensitivity indices for groups of inputs.

5.2 EFAST and the inverse Rosenblatt transformation

Using EFAST and IRT it is possible to derive an algorithm to compute the variance-based sensitivity indices in the case of dependent input variables (from now on, this procedure is named EFAST-IRT). Indeed, in the both EFAST and IRT algorithms the vector $\boldsymbol{x}$ is generated from uniformly and independently distributed variables. Consequently,

Algorithm 2: EFAST with the inverse Rosenblatt transformation (EFAST-IRT)

1. Set M (usually 4 or 6 but sometimes even 10 if we know a priori that the model has strong non linearities). Select a set of $n-1$ integer frequencies $\boldsymbol{\omega}_{\sim i}$ and infer $\omega_i = 2M\max(\boldsymbol{\omega}_{\sim i})$
2. Generate $u_j = \frac{1}{2} + \frac{1}{\pi}\arcsin(\sin(\omega_j s + \varphi_j)), \forall j = 1, \ldots, n$
3. Generate the vector $\boldsymbol{x}$ from IRT (Eq. (15)) with $\boldsymbol{x}_1 = x_i$ and $\bar{\boldsymbol{x}}_2 = \boldsymbol{x}_{\sim i}$
4. Run the model and save the model responses of interest $f(\boldsymbol{x})$
5. Compute the Fourier coefficients and deduce the variance-based sensitivity indices as follows,

$$\hat{S}_{x_i} = \frac{\sum_{l=1}^{M} |\hat{c}_{l\omega_i}|^2}{\sum_{\omega=1}^{N/2} |\hat{c}_\omega|^2} \tag{22a}$$

$$\hat{ST}_{x_i} = \frac{\sum_{\omega=\frac{\omega_i}{2}+1}^{N/2} |\hat{c}_\omega|^2}{\sum_{\omega=1}^{N/2} |\hat{c}_\omega|^2}. \tag{22b}$$

one can take advantage of the (deterministic) periodical sampling of EFAST to generate the dependent variables in conjunction with IRT. The new algorithm to compute (S_{x_i}, ST_{x_i}) is given by algorithm 2.

It can be noticed that one sample of size $N > 2M\omega_i$ is necessary to compute both indices for a given x_i and $n \times N$ model runs are necessary to compute all first order and total sensitivity indices. To evaluate the independent sensitivity indices ($S_{x_i}^{ind}$, $ST_{x_i}^{ind}$) one must proceed as just shown by operating the inverse Rosenblatt transformation from Eq. (14) with $\bar{\boldsymbol{x}}_1 = x_i$ and $\boldsymbol{x}_2 = \boldsymbol{x}_{\sim i}$. The sensitivity indices are estimated as previously, namely,

$$\hat{S}_{x_i}^{ind} = \frac{\sum_{l=1}^{M} |\hat{c}_{l\omega_i}|^2}{\sum_{\omega=1}^{N/2} |\hat{c}_\omega|^2} \tag{21a}$$

$$\hat{ST}_{x_i}^{ind} = \frac{\sum_{\omega=\frac{\omega_i}{2}+1}^{N/2} |\hat{c}_\omega|^2}{\sum_{\omega=1}^{N/2} |\hat{c}_\omega|^2}. \tag{21b}$$

Using samples of size N, an overall of $2N \times n$ model runs are necessary to compute the four sensitivity indices $(S_{x_i}, ST_{x_i}, S_{x_i}^{ind}, ST_{x_i}^{ind}), \forall i = 1, \ldots, n.$ As compared to the non-parametric methods described in [1], which require $4N \times n$ samples to compute the same set of sensitivity indices, the computational effort is halved.

5.3 EFAST and the inverse Nataf transformation

The algorithm to perform GSA via the inverse Nataf transformation (from now on named EFAST-INT) is more subtle than with IRT. Indeed, the calculation of (S_{x_i}, ST_{x_i}) or $(S_{x_i}^{ind}, ST_{x_i}^{ind})$ depends on the position of z_i in the vector $\mathbf{z}$

Algorithm 3: EFAST with the inverse Nataf transformation (EFAST-INT)

1. Choose a value for M and set $i = 1$. Select a set of $n-1$ integer frequencies $\boldsymbol{\omega}_{\sim i}$ and set $\omega_i = 2M\max(\boldsymbol{\omega}_{\sim i})$
2. Generate $u_j = \frac{1}{2} + \frac{1}{\pi}\arcsin(\sin(\omega_j s + \varphi_j)), \forall j = 1, \ldots, n$ with s regularly sampled over $(0, 2\pi]$ using N points
3. Deduce the independent standard normal variables, $z_j = \Phi^{-1}(u_j), \forall j = 1, \ldots, n$. Consider the following vector ordering $\boldsymbol{z} = (z_i, \boldsymbol{z}_{\sim i})$ to estimate S_{x_i} and ST_{x_i} (resp. $\boldsymbol{z} = (\boldsymbol{z}_{\sim i}, z_i)$ to estimate $S^{ind}_{x_i}$ and $ST^{ind}_{x_i}$)
4. Set $\mathbf{R}_{\boldsymbol{z}} = \mathbf{R}_{\boldsymbol{x}}$
5. Find $\mathbf{L}$, get $\boldsymbol{z}^c$ from Eq. (17) and generate $\boldsymbol{x}$ from Eq. (16)
6. Calculate the sample correlation matrix $\hat{\mathbf{R}}_{\boldsymbol{x}}$ from the sample $\boldsymbol{x}$. If $\hat{\mathbf{R}}_{\boldsymbol{x}}$ is not satisfactory, in the sense that it is not close enough to the desired correlation structure $\mathbf{R}_{\boldsymbol{x}}$, modify $\mathbf{R}_{\boldsymbol{z}}$ and resume from 5, otherwise, continue
7. Run the model on $\boldsymbol{x}$ and save the model responses of interest $f(\boldsymbol{x})$. Finally, compute the variance-based sensitivity indices $(\hat{S}_{x_i}, \hat{ST}_{x_i})$ (resp. $(\hat{S}^{ind}_{x_i}, \hat{ST}^{ind}_{x_i})$) as in Eq. (22) (resp. Eq. (21))
8. if $i = n$ then stop. Otherwise, set $i = i + 1$, $\mathbf{R}_{\boldsymbol{x}} = \mathbf{P}\mathbf{R}_{\boldsymbol{x}}\mathbf{P}^T$ and resume from 3.

of standard normal variables used in the Cholesky transformation (Eq. (17)). As explained in [1], if we consider the set $(z_i, \boldsymbol{z}_{\sim i})$ and apply the Cholesky transformation, (S_{x_i}, ST_{x_i}) can be computed. If the Cholesky transformation is applied to the set $(\boldsymbol{z}_{\sim i}, z_i)$, then $(S^{ind}_{x_i}, ST^{ind}_{x_i})$ can be obtained. In both cases, EFAST-INT assigns the highest frequency ω_i to z_i.

The procedure must also include an algorithm to find the optimal $\mathbf{R}_{\boldsymbol{z}}$ that produces a sample with the desired correlation matrix $\mathbf{R}_{\boldsymbol{x}}$. Given $\mathbf{R}_{\boldsymbol{x}}$ the correlation matrix and the marginal densities of each input variables, EFAST-INT proceeds as in Algorithm 3 in which the following permutation matrix is employed,

$$\mathbf{P} = \begin{bmatrix} 0 & \mathbf{e}^T_{n-1} \\ \mathbf{e}_{n-1} & \mathbf{I}_{n-1} \end{bmatrix} \tag{23}$$

with $\mathbf{e}^T_{n-1} = (0, \ldots, 1)$ and $\mathbf{I}_{n-1}$ the $(n-1) \times (n-1)$ identity matrix.

The entire set of sensitivity indices $(\hat{S}_{x_i}, \hat{ST}_{x_i}, \hat{S}^{ind}_{x_i}, \hat{ST}^{ind}_{x_i}), \forall i = 1 \ldots, n$ are obtained with $2n$ samples of size N by considering circular permutations of the set $\boldsymbol{z} = (z_1, \ldots, z_n)$. More specifically, considering the set $(z_1, z_2, \ldots, z_n)$, with one sample $(\hat{S}_{x_1}, \hat{ST}_{x_1})$ can be evaluated by assigning the highest frequency to u_1, and with another sample $(\hat{S}^{ind}_{x_n}, \hat{ST}^{ind}_{x_n})$ are obtained by assigning the highest frequency to u_n. By considering the set $(z_2, \ldots, z_n, z_1)$ and the correlation matrix associated to $(x_2, \ldots, x_n, x_1)$, $(\hat{S}_{x_2}, \hat{ST}_{x_2}, \hat{S}^{ind}_{x_1}, \hat{ST}^{ind}_{x_1})$ can be estimated with other two samples, and so on. At the i-th iteration, step 8) transforms the correlation matrix $\mathbf{R}_{\boldsymbol{x}}$ so that the latter corresponds to the correlation matrix of the circularly permuted vector $(x_i, x_{i+1}, \dots, x_n, x_1, \dots, x_{i-1})$.

6. NUMERICAL TEST CASES

6.1 Example with EFAST-IRT

To illustrate the EFAST-IRT approach, let us consider the following non-linear function: $f(\boldsymbol{x}) = x_1x_2 + x_3x_4$ where $(x_1, x_2) \in [0, 1]^2$ is uniformly distributed within the triangle $x_1+x_2 \leq 1$ and $(x_3, x_4) \in [0, 1]^2$ is uniformly distributed within the triangle $x_3 + x_4 \geq 1$. This function was studied in [1] with an non-parametric approach based on Quasi-Monte Carlo sampling. The inverse Rosenblatt transformation that yields (x_1, x_2) in the domain $x_1 + x_2 \leq 1$ from the independent variables $(u_1, u_2) \in [0, 1]^2$ is (see details in [1]),

$$\begin{cases} x_1 &= 1 - \sqrt{1-u_1} \\ x_2 &= u_2\sqrt{1-u_1} \end{cases} . \tag{24}$$

Because of the symmetry, the IRT of (x_2, x_1) is obtained by simply inverting x_1 and x_2 in Eq. (24). In the same way, the IRT of (x_3, x_4) writes,

$$\begin{cases} x_3 &= \sqrt{u_3} \\ x_4 &= (u_4 - 1)\sqrt{u_3} + 1 \end{cases} . \tag{25}$$

The symmetry of the problem implies that the sensitivity indices of x_1 and x_2 are equal, as well as those of x_3 and x_4. We have computed 100 replicate estimates of these indices with the EFAST-IRT approach. This has been achieved by randomly drawing the four shift parameters $\{\varphi_i, i = 1, \dots, 4\}$ at each replication. We imposed a sample size $N = 4\ 095$, which corresponded to a total number of function calls per replicate equal to $4 \times 4,095 = 16\ 380$, and we selected $M = 8$. Consequently, $\omega_i = (N-1)/M = 255$. Finally, we chose the following set of low frequencies $\boldsymbol{\omega}_{\sim i} = (5, 9, 14)$, although other choices were equally suitable.

Fig. 1 depicts a few draws obtained with EFAST-IRT. We note that the input values are sampled accordingly with the desired constraints. The sample fills the input space quite well. Tab. 1 shows good agreement between the mean estimates of the sensitivity indices and the analytical values. The results indicate that (x_3, x_4) are the most relevant inputs for the response variance.

By referring to the work of [1], it can be inferred that one possible ANOVA decomposition (in the Sobol' sense

[5]) of $f(\boldsymbol{x})$ is:

$$f(\boldsymbol{x}) = f_0 + f_1(x_1) + f_{\bar{2}}(\bar{x}_2) + f_{1\bar{2}}(x_1, \bar{x}_2) + f_3(x_3) + f_{\bar{4}}(\bar{x}_4) + f_{3\bar{4}}(x_3, \bar{x}_4) \tag{26}$$

where the functions in Eq. (26) have the same properties than those of the Sobol's ANOVA decomposition [5]. In particular, they are orthogonal. This property of orthogonality allows to cast the variance of $f(\boldsymbol{x})$, denoted V, as follows:

$$V = V_1 + V_2^{ind} + V_{12} + V_3 + V_4^{ind} + V_{34} \tag{27}$$

with $V_i = E[f_i^2(x_i)]$, $V_i^{ind} = E[f_{\bar{i}}^2(\bar{x}_i)]$ and $V_{ij} = E[f_{ij}^2(x_i, \bar{x}_j)]$. By denoting $VT_i = V_i + V_{ij}$ yields the following variance decomposition,

$$V_y = VT_1 + V_2^{ind} + VT_3 + V_4^{ind}. \tag{28}$$

The normalization of the latter equation by V yields the following relationship between the variance-based sensitivity indices,

$$ST_{x_1} + S_{x_2}^{ind} + ST_{x_3} + S_{x_4}^{ind} = 1.$$

The analytical variance-based sensitivity indices reported in Tab. 1 satisfy this relationship. Moreover, we can infer that $ST_{x_1} + S_{x_2}^{ind} = 0.10$, which indicates that the pair (x_1, x_2) explains 10% of the response variance since the pairs (x_1, x_2) and (x_3, x_4) are independent.

TABLE 1: One hundred replicate estimates of the sensitivity indices computed with EFAST-IRT.

		S_{x_1}	ST_{x_1}	$S_{x_2}^{ind}$	$ST_{x_2}^{ind}$	S_{x_3}	ST_{x_3}	$S_{x_4}^{ind}$	$ST_{x_4}^{ind}$
Analytical		0.033	0.044	0.056	0.067	0.233	0.400	0.500	0.666
EFAST-IRT	2.5^{th} perc.	0.028	0.043	0.051	0.066	0.209	0.361	0.475	0.639
	mean	0.032	0.048	0.055	0.071	0.226	0.395	0.497	0.669
	97.5^{th} perc.	0.037	0.056	0.061	0.077	0.248	0.448	0.523	0.705

6.2 The Ishigami function with EFAST-INT

The Ishigami function is one of the benchmarks model for assessing the efficiency of GSA methods [39]. It has been intensively used by statisticians to test their sensitivity analysis approaches in the case of independent inputs (e.g. [40,41], among others). Recently, the authors in [21] introduced a new method for computing S_{x_i} and $ST_{x_i}^{ind}$ in the case of models with correlated inputs and tested it on the Ishigami function. We repeat here the same example with

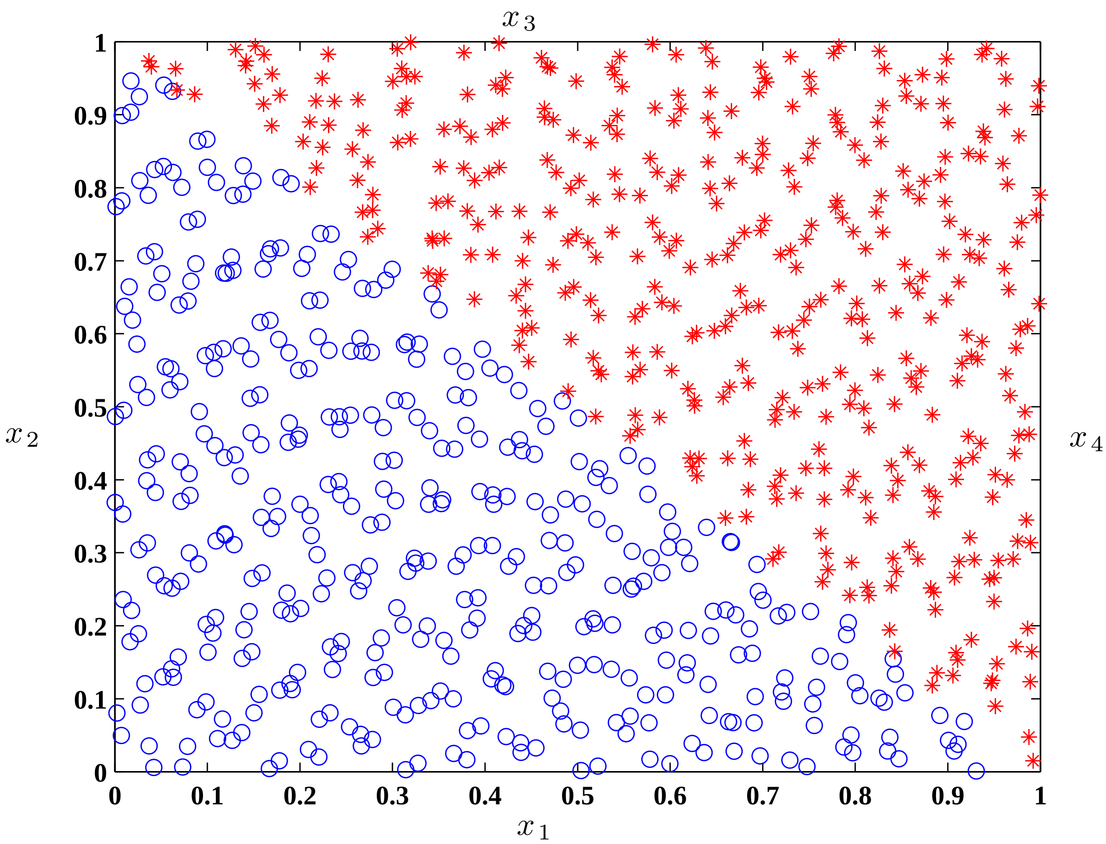

FIG. 1: Some draws obtained with FAST after the inverse Rosenblatt transformation. We have $x_1 + x_2 \leq 1$ (blue circles) and $x_3 + x_4 \geq 1$ (red stars).

EFAST-INT but compute $(S_{x_i}, ST^{ind}_{x_i}, S^{ind}_{x_i}, ST_{x_i}), \forall i = 1, 2, 3$. The Ishigami function writes,

$$f(\boldsymbol{x}) = \sin x_1 + 7\sin^2 x_2 + 0.1 x_3^4 \sin x_1 \tag{29}$$

with input variables being uniformly distributed: $-\pi \leq x_i \leq \pi$, $i = 1, 2, 3$. Note that this function is non-linear in all inputs and x_2 does not interact with the other two variables. Although this function has a low dimension (three inputs only), it is a challenging function for EFAST because the last term, which represents the interaction between x_1 and x_3, is strongly non-linear and its Fourier coefficients have considerable amplitudes at high frequencies. This strong non-linearity led us to choose a considerably high value of M. The difficulty of the test was emphasized because a correlation $r_{13} \in]-1, 1[$ was imposed between x_1 and x_3.

Likewise the previous exercise, 100 replication estimates of the sensitivity indices were performed. The design of the EFAST-INT was the following: $M = 15$, $N = 8\,191$, $\boldsymbol{\omega}_{\sim i} = (5, 8)$ and $\omega_i = 273$. We chose higher N here than in the previous example because of the strong non-linearities mentioned above. Following the work of [21], the exercise was conducted by varying the correlation coefficient r_{13} within $]-1, 1[$.

The results are depicted in Fig. 2. To our best knowledge, no analytical sensitivity indices are available for this test function. We can infer that, because x_2 is not correlated to the other two inputs, its full and independent sensitivity indices are equal. Besides, because x_2 does not interact with the other two inputs, its total and first-order indices are also equal. Moreover, we note that the full index S_{x_i} is always greater or equal to the independent index $S_{x_i}^{ind}$ (resp. $ST_{x_i} \geq ST_{x_i}^{ind}$). However, we may also find $S_{x_i}^{ind} > S_{x_i}$ (or $ST_{x_i}^{ind} > ST_{x_i}$) as also shown in the previous exercise (see Tab. 1).

As far as x_1 and x_3 are concerned we note that, when the correlation coefficient is zero, as expected the full and independent sensitivity indices are equal. When r_{13} is close to ± 1, the independent sensitivity indices (both first and total order) of x_1 and x_3 are null. This makes sense because if $r_{13} = \pm 1$, then all the information in $f(\boldsymbol{x})$ is captured by only one of the pairs (x_1, x_2) or (x_2, x_3). Indeed, in this case, x_1 and x_3 contain the same information and it is not possible to distinguish their individual contributions in the model response. This finding is peculiarly important for the modeller as it indicates that the output uncertainty is explained by one of these two pairs only, thus, allowing some kind of dimensionality reduction. Indeed, the modeller now knows that to obtain narrower output uncertainty, he/she should pay some further effort to reduce the uncertainty either in the pair (x_1, x_2) or in (x_2, x_3).

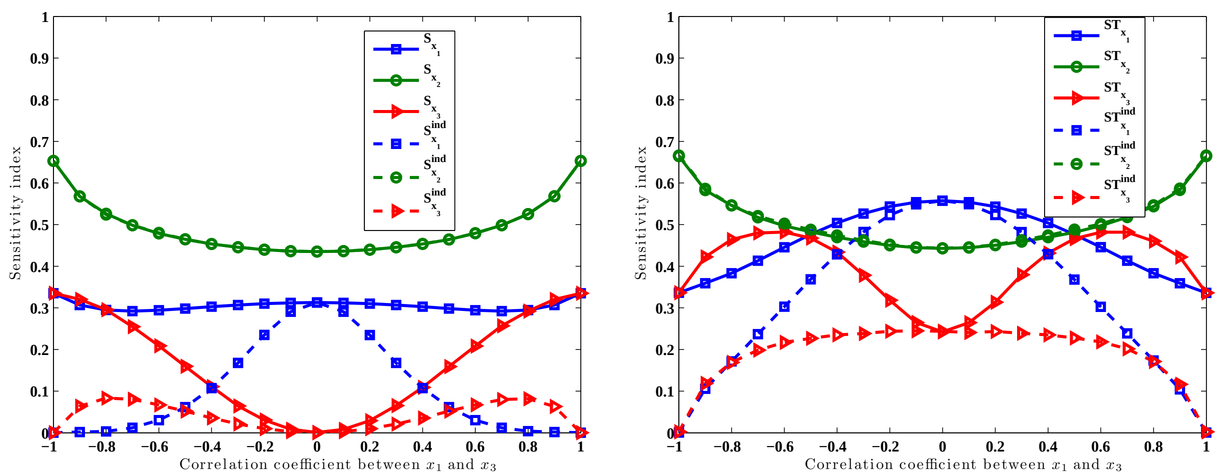

FIG. 2: The variance-based sensitivity indices estimated with EFAST-INT versus the correlation coefficient between x_1 and x_3.

7. CONCLUSION

Performing global sensitivity analysis of model output with dependent inputs is a challenging issue. Variance-based sensitivity indices have been defined in [1,19,21] and different approaches have been proposed to estimate them. Four types of sensitivity indices can be of interest: (i) the full first-order (resp. full total) sensitivity index that measures the amount of model response variance explained by an input factor which takes into account its dependence with the other inputs; (ii) the full total-order sensitivity index that measures the amount of model response variance explained by an input factor which takes into account the both its dependence and interactions with the other inputs; (iii) the independent first-order sensitivity index of an input that measures its relative contribution alone to the response variance by ignoring its dependence with the other inputs; and finally (iv) the independent total-order sensitivity index of an input that measures its relative contribution to the response variance by ignoring its dependence with the other input variables but by accounting for its interactions with the latter.

The Fourier amplitude sensitivity test is one of the first methods for variance-based global sensitivity analysis [10]. Since that, the method has been extended by several authors [11,29,30,37,38]. Specifically, in [30] FAST has been adapted to account for correlations among inputs by using the sampling technique of Iman and Conover [24]. In this work, we extend FAST to compute the four sensitivity indices defined above. The main idea of our approach is to impose either a dependence structure amongst the inputs with the inverse Rosenblatt transformation [22], with Algorithm 2 denoted EFAST-IRT, or a correlation structure with the inverse Nataf transformation [23], denoted EFAST-INT (see Algorithm 3). The numerical tests shown in the paper confirm the suitability of both EFAST-IRT and EFAST-INT.

The sampling strategy proposed by [21] allows for estimating the overall full first-order sensitivity indices and the independent total sensitivity indices with $(2n+2)$ (quasi) Monte Carlo samples. The sampling strategies proposed in [1] allows for assessing the four sensitivity indices of all the input variables with $4n$ (quasi) Monte Carlo samples. In the present work, we show that $2n$ samples are sufficient to compute the four sensitivity indices of all the inputs with either EFAST-IRT or EFAST-INT.

ACKNOWLEDGMENTS

The corresponding author would like to thank the French National Research Agency for its financial support (Research project RESAIN n$^\circ$ANR-12-BS06-0010-02).

REFERENCES

1. Mara, T.A., Tarantola, S., and Annoni, P., Non-parametric methods for global sensitivity analysis of model output with dependent inputs, *Environmental Modelling and Software*, 72:173–183, 2015.
2. Blatman, G. and Sudret, B., Adaptive sparse polynomial chaos expansion based on least angle regression, *Journal of Computational Physics*, 230(6):2345–2367, 2011.
3. Anstett-Collin, F., Goffart, J., Mara, T., and Denis-Vidal, L., Sensitivity analysis of complex models: Coping with dynamic and static inputs, *Reliability Engineering and System Safety*, 134:268–275, 2015.
4. Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., and Tarantola, S., *Global Sensitivity Analysis: The Primer*, Probability and Statistics, John Wiley and Sons, Chichester, 2008.
5. Sobol', I.M., Sensitivity estimates for nonlinear mathematical models, *Math. Mod. and Comput. Exp.*, 1:407–414, 1993.
6. Chun, M.H., Han, S.J., and Tak, N.I., An uncertainty importance measure using distance metric for the change in a cumulative distribution function, *Reliability Engineering and System Safety*, 70:313–321, 2000.
7. Borgonovo, E., Measuring uncertainty importance: Investigation and comparison of alternative approaches, *Risk Analysis*, 26(5):1349–1361, 2006.
8. Sobol', I.M., Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates, *Mathematics and Computers in Simulation*, 55:271–280, 2001.
9. Saltelli, A., Making best use of model evaluations to compute sensitivity indices, *Computational Physics Communications*, 145:280–297, 2002.
10. Cukier, R.I., Fortuin, C.M., Shuler, K.E., Petschek, A.G., and Schaibly, J.H., Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I. theory, *J. Chemical Physics*, 59:3873–3878, 1973.
11. Saltelli, A., Tarantola, S., and Chan, K., A quantitative model independent method for global sensitivity analysis of model output, *Technometrics*, 41:39–56, 1999.
12. Oakley, J.E. and O'Hagan, A., Probabilistic sensitivity analysis of complex models: a Bayesian approach, *J. Royal Statist. Soc. B*, 66:751–769, 2004.
13. Ratto, M., Pagano, A., and Young, P., State dependent parameter metamodelling and sensitivity analysis, *Computer Physics Communications*, 117(11):863–876, 2007.
14. Pearson, K. On the general theory of skew correlation and non-linear regression. In *Mathematical contributions to the theory of evolution*, Vol. XIV. Drapers's Company Research Memoirs, 1905.
15. McKay, M.D., Variance-based methods for assessing uncertainty importance, Tech. Rep., Technical Report NUREG-1150,UR-1996-2695, Los Alamos National Laboratory, 1996.

16. Homma, T. and Saltelli, A., Importance measures in global sensitivity analysis of nonlinear models, *Reliability Engineering and System Safety*, 52:1–17, 1996.

17. Saltelli, A. and Tarantola, S., On the relative importance of input factors in mathematical models: Safety assessment for nuclear waste disposal, *Journal of the American Statistical Association*, 97:702–709, 2002.

18. Xu, C. and Gertner, G.Z., Uncertainty and sensitivity analysis for models with correlated parameters, *Reliability Engineering and System Safety*, 93:1563–1573, 2008b.

19. Mara, T.A. and Tarantola, S., Variance-based sensitivity indices for models with dependent inputs, *Reliability Engineering and System Safety*, 107:115–121, 2012.

20. Zhou, C., Lu, Z., Zhang, L., and Hu, J., Moment independent sensitivity analysis with correlations, *Applied Mathematical Modelling*, 38:4885–4896, 2014.

21. Kucherenko, S., Tarantola, S., and Annoni, P., Estimation of global sensitivity indices for models with dependent variables, *Computer Physics Communications*, 183:937–946, 2012.

22. Rosenblatt, M., Remarks on the multivariate transformation, *Annals of Mathematics and Statistics*, 43:470–472, 1952.

23. Nataf, A., Détermination des distributions dont les marges sont données, *Comptes Rendus de l'Académie des Sciences*, 225:42–43, 1962.

24. Iman, R.I. and Conover, W.J., A distribution-free approach to inducing rank correlation among input variables, *Communications in Statistics: Simulation and Computation*, 11:311–334, 1982.

25. Li, G., Rabitz, H., Yelvington, P.E., Oluwole, O.O., Bacon, F., Kolb, C.E., and Schoendorf, J., Global sensitivity analysis for systems with independent and/or correlated inputs, *J. Physical Chemistry*, 114:6022–6032, 2010.

26. Caniou, Y., Analyse de sensibilité globale pour les modèles imbriqués et multiéchelles, PhD thesis, University Blaise Pascal, Clermont-Ferrand, 2012.

27. Chastaing, G., Gamboa, F., and Prieur, C., Generalized Hoeffding-Sobol decomposition for dependent variables – application to sensitivity analysis, *Electronic Journal of Statistics*, 6:2420–2448, 2012.

28. Da Veiga, S., Wahl, F., and Gamboa, F., Local polynomial estimation for sensitvity analysis of models with correlated inputs, *Technometrics*, 51(4):452–463, 2009.

29. Xu, C., Decoupling correlated and uncorrelated uncertainty contributions for nonlinear models, *Applied Mathematical Modelling*, 37:9950–9969, 2013.

30. Xu, C. and Gertner, G.Z., A general first-order global sensitivity analysis method, *Reliability Engineering and System Safety*, 93:1060–1071, 2008a.

31. Fisher, R.A., The distribution of the partial correlation coefficient, *Metron*, pp. 329–332, 1924.

32. Sobol', I.M., Turchaninov, V.I., Levitan, Y.L., and Shukman, B.V. *Quasi-random sequence generator (routine LPTAU51)*. Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, 1992.

33. Li, H., Lu, Z., and Yuan, X., Nataf transformation based point estimate method, *Chinese Science Bulletin*, 53(17):2586–2592, 2008.

34. Ditlevsen, O. and Madsen, H.O., *Structural Reliability Methods*, John Wiley and Sons, New York, 1996.

35. Cukier, R.I., Schaibly, J.H., and Shuler, K.E., Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. iii. analysis of the approximations, *J. Chemical Physics*, 63:1140–1149, 1975.

36. Satterthwaite, F.E., Random balance experimentation, *Technometrics*, 1:111–137, 1959.

37. Tarantola, S., Gatelli, D., and Mara, T.A., Random balance designs for the estimation of first-order global sensitivity indices, *Reliability Engineering and System Safety*, 91:717–727, 2006.

38. Mara, T.A., Extension of the rbd-fast method to the computation of global sensitivity indices, *Reliability Engineering and System Safety*, 94:1274–1281, 2009.

39. Saltelli, A., Chan, K., and Scott, E.M., *Sensitivity analysis*, John Wiley and Sons, Chichester, 2000.

40. Sudret, B., Global sensitivity analysis using polynomial chaos expansions, *Reliability Engineering and System Safety*, 93:964–979, 2008.

41. Marrel, A., Iooss, B., Laurent, B., and Roustant, O., Calculations of Sobol indices for the Gaussian process metamodel, *Reliability Engineering and System Safety*, 94:742–751, 2009.