



**HAL**  
open science

# Shapley effects for sensitivity analysis with dependent inputs: comparisons with Sobol' indices, numerical estimation and applications

Bertrand Iooss, Clémentine Prieur

## ► To cite this version:

Bertrand Iooss, Clémentine Prieur. Shapley effects for sensitivity analysis with dependent inputs: comparisons with Sobol' indices, numerical estimation and applications. 2017. hal-01556303v1

**HAL Id: hal-01556303**

**<https://inria.hal.science/hal-01556303v1>**

Preprint submitted on 4 Jul 2017 (v1), last revised 23 Nov 2019 (v7)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Shapley effects for sensitivity analysis with dependent inputs: comparisons with Sobol' indices, numerical estimation and applications

Bertrand Iooss

*EDF Lab Chatou, 6 Quai Watier, 78401 Chatou, France*

*Institut de Mathématiques de Toulouse, 31062, Toulouse, France*

Clémentine Prieur

*Université Grenoble Alpes, CNRS, LJK, F-38000 Grenoble, France*

*Inria project/team AIRSEA*

---

## Abstract

The global sensitivity analysis of a numerical model consists in quantifying, by the way of sensitivity indices, the contributions of each of its input variables in the variability of its output. Based on the functional variance analysis, the popular Sobol' indices present a difficult interpretation in the presence of statistical dependence between inputs. Recently introduced, the Shapley effects, which consist of allocating a part of the variance of the output at each input, are promising to solve this problem. In this paper, from several analytical results, we deeply study the effects of linear correlation between some Gaussian input variables on Shapley effects, comparing them to classical first-order and total Sobol' indices. This illustrates the interest, in terms of interpretation, of the Shapley effects in the case of dependent inputs. We also numerically study the numerical convergence of estimating the Shapley effects. For the engineering practical issue of computationally expensive models, we show that the substitution of the model by a metamodel (here kriging) makes it possible to estimate these indices with precision.

*Keywords:* Computer experiments, Kriging, Sensitivity Analysis, Shapley effects, Sobol' indices, Uncertainty Quantification

---

## 1. Introduction

When constructing and using a numerical model simulating physical phenomena, global sensitivity analysis methods are valuable tools [1]. They make it possible to determine which model input variables contribute the most to the variability of the model outputs, which are, on the contrary, the less influential variables and which variables interact with which others. The standard quantitative methods are based on the measure of variance and lead to the so-called Sobol' indices. In the simplest framework of  $d$  scalar inputs, denoted  $\mathbf{X} = (X_1, \dots, X_d) \in \mathbb{R}^d$ , and a single scalar output  $Y \in \mathbb{R}$ , the numerical model is expressed as

$$Y = f(\mathbf{X}) . \tag{1}$$

The first-order Sobol' indices of the model  $f$  give the contribution percentage of each single variable  $X_i$  in the variance of the model output  $Y$ . The Sobol' indices of order two ( $S_{ij}$ ) express the sensitivity of the interactions of the pairs of variables  $X_i$  and  $X_j$ , and so on for the higher orders. In the case of independent inputs, the interpretation of these indices is natural thanks to the decomposition of the variance of  $Y$  [2]. Indeed, the sum of all the Sobol' indices is equal to one and the indices are interpreted as a percentage of explanation of the variance.

However, in many applications, it is common that the input variables have a known statistical dependence structure [3] or that the input space is constrained to a non-rectangular region [4, 5]. As shown in the preliminary studies of [6] and [7], addressing dependence in sensitivity analysis is not an easy task. Many propositions appear in the literature, that are not always easy to interpret. The main issue relies on the difficulty to discriminate between correlation and interaction effects. Several attempts have tried to tackle this issue. One strategy proposed in Jacques et al. [8] is to define grouped indices for groups of correlated variables. However, it might be required for some applications to go below the scale of groups of correlated variables. Xu and Gertner [9] proposed

to decompose the partial variance of an input into a correlated part and an uncorrelated one. Such an approach allows to exhibit inputs that have an impact on the output only through their strong correlation with other incomes. However, they only investigated linear models with linear dependencies. Later, [10] extended this approach to more general models, using the concept of High Dimensional Model Representation (HDMR) [11]. HDMR is based on a hierarchy of component functions of increasing dimensions (truncation of ANOVA-Sobol decomposition in the case of independent variables). The component functions are then approximated by expansions in terms of some suitable basis functions (e.g., polynomials [12], splines, ...). This metamodeling approach allows the splitting of the response variance into a correlative contribution and a structural one of a set of inputs. [13] proposed, in the setting of linear dependencies, a first step of decorrelation of the inputs with the Gram-Schmidt procedure, and then to perform the ANOVA-HDMR on these new inputs.

In the same time, [14] proposed a non-parametric procedure to estimate first-order and total indices in presence of dependencies, non necessarily of linear type. Their methodology requires the knowledge of the conditional probability densities and the ability of sampling from those. Later, in [15], the authors established a link between the approaches in [14] and [13], allowing distinguishing the independent contributions of the inputs to the response variance from their mutual dependent contributions, via the estimation of four sensitivity indices (full and independent first-order indices, full and independent total indices) for each input. They proposed two sampling strategies, for dependent continuous inputs. The first one is based on the Rosenblatt transform [16]. The second one is a simpler method that estimates the sensitivity indices without requiring the knowledge of conditional probability densities. In [17], the authors proposed an alternative approach, mainly motivated to circumvent the issue of computational demands induced by the sampling of conditional probability densities. Their approach is based on a truncated Polynomial Chaos Expansion of the response whose coefficients are estimated by Partial Least Squares Regression, whereas the classical method (see [18] for example) used Ordinary Least Squares

Regression. It leads to a new kind of sensitivity indices.

As a different approach, [19, 20] initiated the construction of a new generalized moment free sensitivity index. Based on geometrical consideration, these indices measure the shift area between the outcome density and this same density conditionally to a parameter. Thanks to the properties of these new indices, a methodology is given to obtain them analytically through test cases.

Notice that none of these works has given an exact and unambiguous definition of the functional ANOVA for correlated inputs as the one provided by Hoeffding-Sobol' decomposition [21, 2] when inputs are independent. Consequently, the exact form of the model has neither been exploited to provide a general variance-based sensitivity measures in the dependent frame. A common way to address this last point is to build on work by Stone [22] and Hooker [23] who define an ANOVA for dependent inputs and then define variable importance through that generalization of ANOVA. This is the method taken by [24] for computer experiments. However, the dependent-variable ANOVA leads to importance measures with two conceptual problems, underlined in [25]: the inputs can get negative influences and the approach places strong restrictions on the joint probability distribution of the inputs.

To carry out a sensitivity analysis in the case where the model entries have dependencies between them, recent works have demonstrated the great interest of the Shapley values [26, 25, 27]. The Shapley values [28], well known in game theory and economics, have been proposed by [29] in the framework of sensitivity analysis of model outputs. Their use and interpretation do not require the continuity of the inputs. Instead of passing by the model output variance decomposition, it consists in a direct allocation of a part of the variance of the output at each input. Then, the two main properties and advantages of the Shapley values are that they cannot be negative and their sum is worth the variance of the output. The allocation rule is based on an equitable principle. For example, an interaction effect is equally distributed between each input involved in the interaction. From a conceptual point of view, the main remaining issue is to understand what is the effect of the dependence between inputs on

the Shapley values.

Several analytical cases where the Shapley values can be analytically computed have been given in [25]. It allows to discover several properties about the Shapley values (equality for two functionally equivalent variables, preservation under invertible transformations). Moreover, [25] gives the general analytical form of Shapley values in the case of Gaussian inputs  $\mathbf{X}$  and a linear function  $f(\mathbf{X}) = \beta_0 + \beta^\top \mathbf{X}$  (with  $\beta \in \mathbb{R}^d$ ). However, from this general form, it is difficult to understand the effects of the dependence between the inputs in the Shapley values. Therefore, it seems important to provide a deep study of several particular cases, sufficiently simple to provide some interpretation.

For the practical viewpoint, [26] has proposed two estimation algorithms of the Shapley effects (that we define as the normalized Shapley values), just illustrating them on two application cases. In the present paper, we develop the analytical solutions of Shapley effects on several particular test functions. Then, we perform a preliminary numerical convergence study and make the first comparisons between theoretical and numerical results of the Shapley effects. Finally, as for the Sobol' indices, the most important issue in practice will be the cost (in terms of number of required model evaluations) of estimating the Shapley effects. In this case, a classical solution is to use a metamodel which is a mathematical model approximating the numerical model (1) from an initial and limited set of runs [30, 31]. The metamodel solution is a current engineering practice for estimating sensitivity indices [32]. We introduce this technique to estimate Shapley effects on an analytical test case and a real engineering problem.

In the following section, we recall the general mathematical formulation of Sobol' indices and Shapley effects when the inputs are dependent. In Section 3, we develop the analytical formulas that one can obtain in several particular cases: linear models with Gaussian inputs in various dimensions, block-additive structure, Ishigami model with various dependencies between the three inputs. We specially focus on inequalities that can be developed between Sobol' indices and Shapley effects. Numerical estimations algorithms of Shapley effects are

studied in Section 4. Section 5 presents an industrial example which requires the use of a metamodel to estimate the sensitivity indices. A conclusion synthesizes this work.

## 2. General formulation of sensitivity indices

### 2.1. Sobol' indices

The Sobol' indices are defined in the following way [2]:

$$S_i = \frac{\text{Var}(\mathbb{E}[Y|X_i])}{\text{Var}(Y)}, \quad S_{ij} = \frac{\text{Var}(\mathbb{E}[Y|X_i, X_j])}{\text{Var}(Y)} - S_i - S_j, \quad \dots \quad (2)$$

The closed Sobol' index for a set of inputs indexed by  $u$  ( $u \subseteq \{1, \dots, d\}$ ) is also defined by [33]

$$S_u^{\text{clo}} = \text{Var}(\mathbb{E}[Y|\mathbf{X}_u]) / \text{Var}(Y). \quad (3)$$

This sensitivity index contains the effects of variables and interactions which are included in  $\mathbf{X}_u$ . For the mathematical developments of the following sections, we also denote the numerator of  $S_u^{\text{clo}}$  as

$$\tau_u^2 = \text{Var}(\mathbb{E}[Y|\mathbf{X}_u]). \quad (4)$$

In addition to the Sobol' indices in Eq. (2), total sensitivity indices have also been defined in order to express the “total” sensitivity to the variance of  $Y$  of an input variable  $X_i$  [34]:

$$S_{T_i} = S_i + \sum_{j \neq i} S_{ij} + \sum_{j \neq i, k \neq i, j < k} S_{ijk} + \dots \quad (5)$$

According to the law of total variance, the total Sobol' indices can be written as

$$S_{T_i} = \frac{\mathbb{E}(\text{Var}[Y|\mathbf{X}_{-i}])}{\text{Var}(Y)}, \quad (6)$$

where  $\mathbf{X}_{-i}$  is the vector  $(X_1, \dots, X_d)$  not containing  $X_i$ .

Recall that this paper is not restricted to independent inputs, so that the knowledge of first-order and total Sobol' indices does not give a complete information on the way an input  $X_i$  influences the output  $Y$ . In [15], the authors

propose a strategy, based on the estimation of four sensitivity indices per input, and described hereafter. We assume that  $\mathbf{X} = (X_1, \dots, X_n)$  is a set of continuous dependent random variables, with joint probability density function  $p(\mathbf{X}) = p(X_1)p(X_2|X_1)p(X_3|X_1, X_2) \dots p(X_n|\mathbf{X}_{-n})$ . Let  $F_i(x_i|\mathbf{V})$  be the cumulative distribution function of  $X_i$  conditionally to  $\mathbf{V}$ , with  $\mathbf{V} \subseteq \mathbf{X}_{-i}$ . The Rosenblatt transformation [16] of  $\mathbf{X}$  provides with a set of independent random variables  $\mathbf{U}^1$  uniformly distributed over the unit hypercube  $[0, 1]^n$ . That is,  $U_1^1 = F_1(X_1)$ ,  $U_2^1 = F_2(X_2|X_1)$ ,  $U_3^1 = F_3(X_3|X_1, X_2)$ ,  $\dots$ ,  $U_n^1 = F_n(X_n|\mathbf{X}_{-n})$ . In case of dependent variables, the Rosenblatt transformation is not unique and there are  $n!$  possibilities depending on how the random variables are ordered in the vector  $\mathbf{X}$ . We denote by  $\mathbf{U}^i$  the Rosenblatt transformation of the ordered vector  $(X_i, \dots, X_n, X_1, \dots, X_{i-1})$ . Such a mapping is bijective and we have  $Y = f(X_1, \dots, X_n) = g_i(\mathbf{U}^i)$ . We now define, as in [15],

$$S_{(i)} = \frac{\text{Var}\mathbb{E}(\mathbb{E}[g_i(\mathbf{U}^i)|U_1^i])}{\text{Var}(g_i(\mathbf{U}^i))} = \frac{\text{Var}(\mathbb{E}[Y|X_i])}{\text{Var}(Y)} = S_i, \quad (7)$$

$$S_{T_{(i)}} = \frac{\mathbb{E}(\text{Var}[g_i(\mathbf{U}^i)|\mathbf{U}_{-1}^i])}{\text{Var}[g_i(\mathbf{U}^i)]}, \quad (8)$$

$$S_{(i)}^{\text{ind}} = \frac{\text{Var}(\mathbb{E}[g_{i+1}(\mathbf{U}^{i+1})|U_n^{i+1}])}{\text{Var}[g_{i+1}(\mathbf{U}^{i+1})]}, \quad (9)$$

$$S_{T_{(i)}}^{\text{ind}} = \frac{\mathbb{E}(\text{Var}[g_{i+1}(\mathbf{U}^{i+1})|\mathbf{U}_{-n}^{i+1}])}{\text{Var}[g_{i+1}(\mathbf{U}^{i+1})]} = S_{T_i}. \quad (10)$$

The indices  $S_{(i)}$  and  $S_{T_{(i)}}$  include the effects of the dependence of  $X_i$  with other inputs, and are referred to as full sensitivity indices in [13]. The indices  $S_{(i)}^{\text{ind}}$  and  $S_{T_{(i)}}^{\text{ind}}$  measure the effect of an input  $X_i$ , that is not due to its dependence with other variables  $\mathbf{X}_{-i}$ . Such indices have also been introduced as uncorrelated effects in [13] and further discussed in [15] which refers them as the independent Sobol' indices. In [15], the authors propose to estimate the four indices  $S_{(i)}$ ,  $S_{T_{(i)}}$ ,  $S_{(i)}^{\text{ind}}$  and  $S_{T_{(i)}}^{\text{ind}}$  for the full set of inputs ( $i = 1, \dots, d$ ). Note that the full first-order indices coincide with the associated classical first-order



Sobol' indices, and that the independent total indices coincide with the associated classical total Sobol' indices. Note also that  $S_{(i)}^{\text{ind}} \leq S_{T_{(i)}}^{\text{ind}} = S_{T_{(i)}}$  and that  $S_{(i)} = S_i \leq S_{T_{(i)}}$ .

## 2.2. Shapley effects

In this section, continuity of the probability distribution of the inputs is not required anymore. An alternative to the estimation of the four indices  $S_{(i)}$ ,  $S_{T_{(i)}}$ ,  $S_{(i)}^{\text{ind}}$  and  $S_{T_{(i)}}^{\text{ind}}$  is the one of Shapley effects  $Sh_i$  defined as

$$Sh_i = \sum_{u \subseteq -\{i\}} \frac{(d - |u| - 1)! |u|!}{d!} [c(u \cup \{i\}) - c(u)] , \quad (11)$$

where  $c(\cdot)$  is a cost function and  $-\{i\}$  is the set of indices  $\{1 \dots d\}$  not containing  $i$ . The Shapley values [28], well known in game theory and economics, have been proposed by [29] in the framework of sensitivity analysis of model outputs. By using  $c(u) = S_u^{\text{clo}}$  (see Eq (3)), [29] proposes to use the corresponding Shapley values as new importance measures for sensitivity analysis of model output. These indices have been called Shapley effects by [26] which also shows that it is equivalent to use  $\mathbb{E}[\text{Var}(Y|\mathbf{X}_{-u})]/\text{Var}(Y)$  than  $\text{Var}(\mathbb{E}[Y|\mathbf{X}_u])/\text{Var}(Y)$  as the cost function.

More precisely, we consider in this paper a re-normalized version (by the variance of  $Y$ ) of the indices of Shapley proposed by [29]. The Shapley effects consist in a equitable allocation of a part of the variance of the output at each input. As the Shapley effects are non negative and their sum is one, each Shapley effect contains a part of the variance of  $Y$ .

In the following section, we discuss the difference of interpretation and behavior of the Shapley effects and the Sobol' indices.

## 3. Relations and inequalities between Sobol' indices and Shapley effects

In the case of independent variables, each Shapley effect  $Sh_i$  is framed by the corresponding first-order  $S_i = S_{(i)}$  and total indices  $S_{T_i} = S_{T_{(i)}}^{\text{ind}}$ :

$$S_i \leq Sh_i \leq S_{T_{(i)}} . \quad (12)$$

In fact, in addition to the individual effect of the variable  $X_i$ , the Shapley effects take into account the effects of interactions by distributing them equally in the index of each input that plays in the interaction [29]. In the form of Eq. (5), this writes

$$Sh_i = S_i + \frac{1}{2} \sum_{j \neq i} S_{ij} + \frac{1}{3} \sum_{j \neq i, k \neq i, j < k} S_{ijk} + \dots \quad (13)$$

However, no relation such as the one in Eq. (12) can be directly deduced in the case of dependent inputs. The goal of this section is to study particular cases where analytical deduction can be made. We concentrate the analysis on the full first-order indices (corresponding to classical first-order Sobol' indices)  $S_i = S_{(i)}$  (called ‘‘First-order Sobol’’ in the figures) and the independent total indices (corresponding to classical total Sobol' indices)  $S_{T_i} = S_{T_{(i)}}^{\text{ind}}$  (called ‘‘Ind total Sobol’’ in the figures). Indeed these are the indices mainly studied and discussed in the previous works on this subject, in particular in [14], whose numerical tests inspired the ones proposed in this section. Moreover, these indices are easily and directly provided by the Shapley effects estimation algorithms ([26], see Section 4), during the first and last iterations of the algorithm. Finally, the estimation of the independent first-order Sobol'  $S_{(i)}$  and of the full total Sobol' indices  $S_{T_{(i)}}$  is based on a rather cumbersome process, based on the use of  $d$  Rosenblatt transforms (see more details in Section 2.1). Comparisons with these complementary indices will be made in a further study.

### 3.1. Gaussian framework and linear model

Let us consider

$$Y = \beta_0 + \beta^\top \mathbf{X} \quad (14)$$

with  $\mathbf{X} \sim \mathcal{N}(\mu, \Sigma)$  and  $\Sigma \in \mathbb{R}^{d \times d}$  a full-rank matrix. We have  $\sigma^2 = \text{Var}(Y) = \beta_{\{1, \dots, d\}}^\top \Sigma_{\{1, \dots, d\}, \{1, \dots, d\}} \beta_{\{1, \dots, d\}}$ . We get from [25]:

$$Sh_i = \frac{1}{d} \sum_{u \subseteq -i} \binom{d-1}{|u|}^{-1} \frac{\text{Cov}(X_i, \mathbf{X}_{-u}^\top \beta_{-u} \mid \mathbf{X}_u)^2}{\sigma^2 \text{Var}(X_i \mid \mathbf{X}_u)}. \quad (15)$$

Recall now the following classical formula:

$$\text{Var}(\mathbf{X} \mid \mathbf{X}_{-j}) = \Sigma_{\{1, \dots, d\}, -j} \Sigma_{-j, -j}^{-1} \Sigma_{-j, \{1, \dots, d\}}. \quad (16)$$

From (16) and according to the law of total variance, we easily obtain for the Sobol' indices:

$$S_j = \frac{\text{Var}(\mathbb{E}[\beta_0 + \beta^\top \mathbf{X} | X_j])}{\sigma^2} = 1 - \frac{\mathbb{E}(\text{Var}[\beta^\top \mathbf{X} | \mathbf{X}_{-j}])}{\sigma^2} \quad (17)$$

$$= \frac{\beta_{\{1, \dots, d\}}^\top (\Sigma_{\{1, \dots, d\}, \{1, \dots, d\}} - \Sigma_{\{1, \dots, d\}, -j} \Sigma_{-j, -j}^{-1} \Sigma_{-j, \{1, \dots, d\}})}{\sigma^2} \beta_{\{1, \dots, d\}}, \quad (18)$$

$$S_{T_j} = \frac{\mathbb{E}(\text{Var}[\beta_0 + \beta^\top \mathbf{X} | \mathbf{X}_{-j}])}{\sigma^2} = \frac{\mathbb{E}(\text{Var}[\beta^\top \mathbf{X} | \mathbf{X}_{-j}])}{\sigma^2} \quad (19)$$

$$= \frac{\beta_{\{1, \dots, d\}}^\top \Sigma_{\{1, \dots, d\}, -j} \Sigma_{-j, -j}^{-1} \Sigma_{-j, \{1, \dots, d\}} \beta_{\{1, \dots, d\}}}{\sigma^2}. \quad (20)$$

Note that  $\beta_0$  do not play any role as a translation parameter in variance-based sensitivity analysis.

### 3.2. Gaussian framework, linear model with two inputs

We focus now on the case  $d = 2$ . We consider  $\mathbf{X} \sim \mathcal{N}_2(\mu, \Sigma)$  and  $Y = \beta^\top \mathbf{X}$ , with

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}, \quad -1 \leq \rho \leq 1, \sigma_1 > 0, \sigma_2 > 0.$$

Let  $\sigma^2 = \text{Var}(Y) = \beta_1^2\sigma_1^2 + 2\rho\beta_1\beta_2\sigma_1\sigma_2 + \beta_2^2\sigma_2^2$ . From Eq. (4),  $\tau_u^2 = \text{Var}(\mathbb{E}[Y | \mathbf{X}_u])$  ( $u \subseteq \{1, \dots, d\}$ ) and we obtain  $\tau_\emptyset^2 = 0$ ,  $\tau_1^2 = (\beta_1\sigma_1 + \rho\beta_2\sigma_2)^2$ ,  $\tau_2^2 = (\beta_2\sigma_2 + \rho\beta_1\sigma_1)^2$  and  $\tau_{12}^2 = \sigma^2$ . For  $j = 1, 2$ ,  $d = 2$ , the definitions (Eq. (11)) of Shapley effects give

$$\sigma^2 Sh_j = \frac{1}{d} \sum_{u \subseteq -\{j\}} \binom{d-1}{|u|}^{-1} (\tau_{u+\{j\}}^2 - \tau_u^2). \quad (21)$$

From that, we get

$$\begin{aligned} \sigma^2 Sh_1 &= \beta_1^2\sigma_1^2 \left(1 - \frac{\rho^2}{2}\right) + \rho\beta_1\beta_2\sigma_1\sigma_2 + \beta_2^2\sigma_2^2 \frac{\rho^2}{2}, \\ \sigma^2 Sh_2 &= \beta_2^2\sigma_2^2 \left(1 - \frac{\rho^2}{2}\right) + \rho\beta_1\beta_2\sigma_1\sigma_2 + \beta_1^2\sigma_1^2 \frac{\rho^2}{2}. \end{aligned} \quad (22)$$

For  $j = 1, 2$ , from the definition of first-order Sobol indices,  $S_j = \text{Var}(\mathbb{E}[Y | X_j]) / \text{Var}(Y)$ , we have

$$\begin{aligned} \sigma^2 S_1 &= \beta_1^2\sigma_1^2 + 2\rho\beta_1\beta_2\sigma_1\sigma_2 + \rho^2\beta_2^2\sigma_2^2, \\ \sigma^2 S_2 &= \beta_2^2\sigma_2^2 + 2\rho\beta_1\beta_2\sigma_1\sigma_2 + \rho^2\beta_1^2\sigma_1^2. \end{aligned} \quad (23)$$

Moreover, for  $j = 1, 2$ , from the definition of total Sobol indices,  $S_{T_j} = \mathbb{E}(\text{Var}[Y|\mathbf{X}_{-j}])/\text{Var}(Y) = 1 - \text{Var}(\mathbb{E}[Y|\mathbf{X}_{-j}])/\text{Var}(Y)$ , we have

$$\begin{aligned}\sigma^2 S_{T_1} &= \beta_1^2 \sigma_1^2 (1 - \rho^2), \\ \sigma^2 S_{T_2} &= \beta_2^2 \sigma_2^2 (1 - \rho^2).\end{aligned}\tag{24}$$

From Equations (22), (23) and (24) we get that the four following assertions are equivalent

$$\begin{aligned}Sh_j &\leq S_{T_j}, \\ S_j &\leq Sh_j, \\ \rho \left( \rho \frac{\beta_1^2 \sigma_1^2 + \beta_2^2 \sigma_2^2}{2} + \beta_1 \beta_2 \sigma_1 \sigma_2 \right) &\leq 0, \\ |\rho| &\leq \frac{2|\beta_1 \beta_2| \sigma_1 \sigma_2}{\beta_1^2 \sigma_1^2 + \beta_2^2 \sigma_2^2}.\end{aligned}\tag{25}$$

The equality of the three first assertions is obtained in the absence of correlation ( $\rho = 0$ ). In this case, the Shapley effects are equal to the first-order and total Sobol' indices in the absence of correlation. In the presence of non-zero correlation, the Shapley effects lie between the full first-order indices and the independent total indices: when  $Sh_j \leq S_{T_j}$  then  $S_j \leq Sh_j$ , and when  $Sh_j > S_{T_j}$  then  $S_j > Sh_j$ . We call this the sandwich effect. We remark that the effects of the dependence that we can see in the independent total indices (*e.g.*  $-\rho^2 \beta_1^2 \sigma_1^2$  for  $X_1$ ) and in the full first-order indices (*e.g.*  $2\rho \beta_1 \beta_2 \sigma_1 \sigma_2 + \rho^2 \beta_2^2 \sigma_2^2$  for  $X_1$ ) is allocated in half to the Shapley effect, in addition to the elementary effect (*e.g.*  $\beta_1^2 \sigma_1^2$  for  $X_1$ ).

These results are also illustrated in Figure 1. In the left figure, as the standard deviations of each variable are equal, the different sensitivity indices are superimposed and the Shapley effects are constant. In the right figure,  $X_2$  has more variability than  $X_1$ , and the corresponding sensitivity indices of  $X_2$  are logically larger than the ones of  $X_1$ . The effect of the dependence between the inputs is clearly shared on each input variable. The dependence between the two inputs lead to a rebalancing of their corresponding Shapley effects, while a full Sobol' index of an input adds to its own effect the effect of an input of which it is dependent. We also see on the right figure that the Shapley effects of two perfectly correlated variables are equal. Finally, the sandwich effect is

respected for each input: From Eq (25), we know that  $S_j \leq Sh_j \leq S_{T_j}$  when  $\rho \in [-0.8; 0]$  and that  $S_j > Sh_j > S_{T_j}$  else.

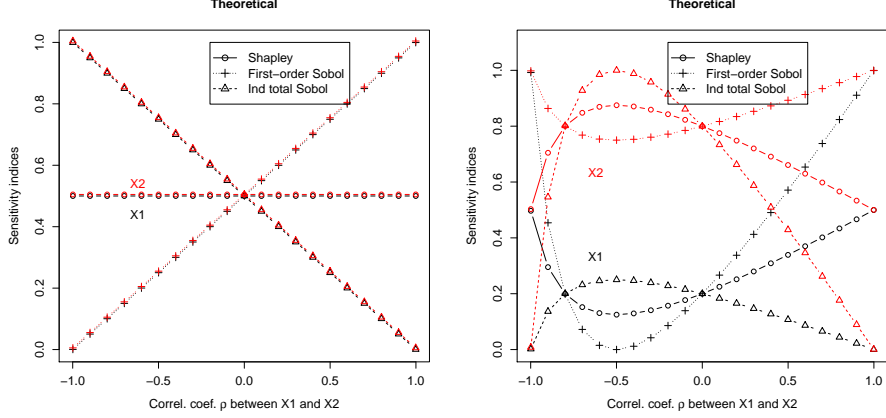


Figure 1: Sensitivity indices on the linear model ( $\beta_1 = 1, \beta_2 = 1$ ) with two Gaussian inputs. Left:  $(\sigma_1, \sigma_2) = (1, 1)$ . Right:  $(\sigma_1, \sigma_2) = (1, 2)$ .

### 3.3. Gaussian framework, dimension $d = 2$ , some insights into connections between correlation and interaction effects

The way Shapley effects handle the connections between correlation and interaction effects is not trivial. We propose to underline this issue on two particular cases:

- The model  $Y = f(X_1, X_2) = X_1$  with  $(X_1, X_2)$  two dependent standard Gaussian variables with a correlation coefficient  $\rho$ . It corresponds to the case  $\beta_1 = 1, \beta_2 = 0, \mu_1 = \mu_2 = 0, \sigma_1^2 = \sigma_2^2 = 1$  in the model introduced in Section 3.2. We then have:

$$Sh_1 = 1 - \frac{\rho^2}{2} \text{ and } Sh_2 = \frac{\rho^2}{2}. \quad (26)$$

- The model  $Y = f(X_1, X_2) = X_1 + \alpha X_1 X_2$  with  $(X_1, X_2)$  two independent standard Gaussian variables. We then have:

$$Sh_1 = \frac{1 + \alpha^2/2}{1 + \alpha^2} \text{ and } Sh_2 = \frac{\alpha^2/2}{1 + \alpha^2}. \quad (27)$$

Equalizing the Shapley effects of these two models gives the following equality:

$$\alpha^2 = \frac{\rho^2}{1 - \rho^2}. \quad (28)$$

An important remark is that, from (26), we see that an input not involved in the numerical model can have a non-zero effect if it is correlated with an influent input of the model. If the two inputs are perfectly correlated, their Shapley effects are equal. This extreme case corresponds to the pure-interaction model  $f(X_1, X_2) = X_1 X_2$ .

#### 3.4. Gaussian framework, linear model with three inputs

We consider a linear model with  $\mathbf{X} = (X_1, X_2, X_3)^\top$  a Gaussian random vector  $\mathbf{X} \sim \mathcal{N}(\mu, \Sigma)$  with  $\mu = (0, 0, 0)^\top$ . We assume that  $X_1$  is independent from both  $X_2$  and  $X_3$ , and that  $X_2$  and  $X_3$  may be correlated. The covariance matrix writes

$$\Sigma = \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2^2 & \rho\sigma_2\sigma_3 \\ 0 & \rho\sigma_2\sigma_3 & \sigma_3^2 \end{pmatrix}, \quad -1 \leq \rho \leq 1.$$

We obtained the following analytical results.

$$\begin{aligned} V &= \text{Var}[f(\mathbf{X})] = \sum_{j=1}^3 \beta_j^2 \sigma_j^2 + 2\rho\beta_2\beta_3\sigma_2\sigma_3, \\ Sh_1 &= (\beta_1^2 \sigma_1^2)/V, \\ Sh_2 &= [\beta_2^2 \sigma_2^2 + \rho\beta_2\beta_3\sigma_2\sigma_3 + \frac{\rho^2}{2}(\beta_3^2 \sigma_3^2 - \beta_2^2 \sigma_2^2)]/V, \\ Sh_3 &= [\beta_3^2 \sigma_3^2 + \rho\beta_2\beta_3\sigma_2\sigma_3 + \frac{\rho^2}{2}(\beta_2^2 \sigma_2^2 - \beta_3^2 \sigma_3^2)]/V, \end{aligned} \quad (29)$$

As expected, we have  $\sum_{j=1}^3 Sh_j = 1$  and we see in  $Sh_2$  and  $Sh_3$  how the correlation effect is distributed in each index. In the case of perfectly correlated variables, we obtain  $Sh_2 = Sh_3 = (\beta_2^2 \sigma_2^2 + \beta_3^2 \sigma_3^2 + 2\rho\beta_2\beta_3\sigma_2\sigma_3)/(2V)$  with  $\rho = \pm 1$ .

We study the particular case  $\beta_1 = \beta_2 = \beta_3 = 1$ ,  $\sigma_1 = \sigma_2 = 1$  and  $\sigma_3 = 2$ , for which in [14], the authors provide the formulas of full first-order and

independent total Sobol' indices. The results for that model are illustrated on Figure 2. The Shapley effects are equal to the Sobol' indices in the absence of correlation, and then lie between the associated full first-order and independent total indices in the presence of correlation. The sandwich effect is respected. The effect of an increasing correlation (in absolute value) can be interpreted as an attractive effect both for the full Sobol' indices (here the first-order one) and the Shapley effect. However, for the Shapley effects, the effect of the correlation is shared on each correlated variable, while. This leads to an increase of one Shapley effect and a decrease for the other. The Shapley effects show then an ease interpretation. As before, we see that the Shapley effects of two perfectly correlated variables are equal.

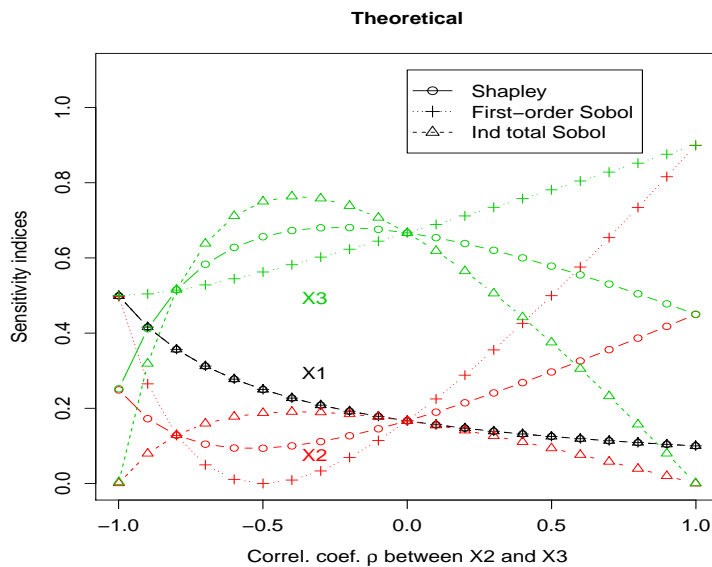


Figure 2: Sensitivity indices on the linear model with three Gaussian inputs.

### 3.5. Gaussian framework, linear model with an interaction and three inputs

In Sections 3.2 and 3.4, we have presented models for which the Shapley effect is between the associated full first-order and independent total indices. In the present section, we see that it is not always the case. Let us define the

model

$$Y = X_1 + X_2 X_3 \quad (30)$$

with

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} \sim \mathcal{N}_3 \left( \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \Sigma \right) \text{ and } \Sigma = \begin{pmatrix} \sigma_1^2 & 0 & \rho\sigma_1\sigma_3 \\ 0 & \sigma_2^2 & 0 \\ \rho\sigma_1\sigma_3 & 0 & \sigma_3^2 \end{pmatrix}, \quad -1 \leq \rho \leq 1.$$

By simple computations, we get  $\sigma^2 S_1 = \sigma_1^2$  and  $\sigma^2 S_{T_1} = \sigma^2 - (\sigma_2^2 \sigma_3^2 + \rho^2 \sigma_1^2) = (1 - \rho^2) \sigma_1^2$ . Recall that

$$\sigma^2 Sh_1 = \frac{1}{3} \left( \tau_1^2 - \tau_0^2 + \frac{1}{2} (\tau_{12}^2 - \tau_2^2 + \tau_{13}^2 - \tau_3^2) + \sigma^2 - \tau_{23}^2 \right).$$

We thus get

$$\sigma^2 Sh_1 = \sigma_1^2 \left( 1 - \frac{\rho^2}{2} \right) + \frac{\sigma_2^2 \sigma_3^2}{6} \rho^2. \quad (31)$$

A straightforward computation yields

$$S_{T_1} \leq Sh_1 \leq S_1.$$

We also get  $S_2 = 0$ ,  $\sigma^2 S_{T_2} = \sigma_2^2 \sigma_3^2$  and  $\sigma^2 Sh_2 = \frac{\sigma_2^2 \sigma_3^2}{6} (3 + \rho^2)$ . Thus

$$S_2 \leq Sh_2 \leq S_{T_2}.$$

Concerning the third input variable  $X_3$ , one gets  $\sigma^2 S_3 = \rho^2 \sigma_1^2$ ,  $\sigma^2 S_{T_3} = (1 - \rho^2) \sigma_2^2 \sigma_3^2$  and  $\sigma^2 Sh_3 = \frac{\rho^2 \sigma_1^2}{2} + \frac{\sigma_2^2 \sigma_3^2}{6} (3 - 2\rho^2)$ . Thus the two following assertions are equivalent:

$$\begin{aligned} S_3 &\leq Sh_3 \leq S_{T_3}, \\ \rho^2 \sigma_1^2 &\leq \frac{\sigma_2^2 \sigma_3^2}{3} (3 - 4\rho^2). \end{aligned}$$

The two following assertions are also equivalent:

$$\begin{aligned} T_3^{\text{ind}} &\leq \frac{\phi_3}{\sigma^2} \leq S_3, \\ \rho^2 \sigma_1^2 &\geq \frac{\sigma_2^2 \sigma_3^2}{3} (3 - 2\rho^2). \end{aligned}$$



It also happens that  $Sh_3$  is not comprised between  $S_3$  and  $S_{T_3}$ . The two following assertions are equivalent:

$$Sh_3 \geq \max(S_3, S_{T_3}),$$

$$\frac{3}{7} \leq \rho^2 \leq \frac{3}{5}.$$

Figure 3 illustrates the Sobol' indices and Shapley effects for this model. As expected, we see two intervals (approximately  $[-0.775; -0.655]$  and  $[0.655; 0.775]$ ) where the Shapley effects of  $X_3$  are larger than the full first-order Sobol' indices and the independent total Sobol' indices.

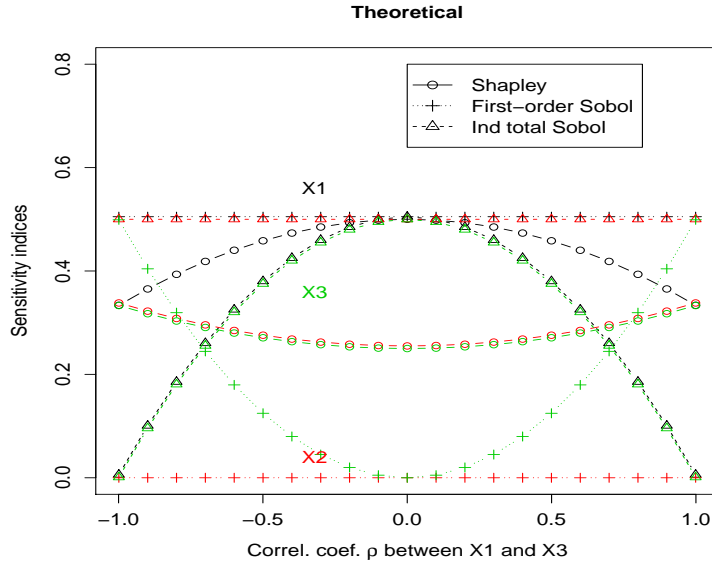


Figure 3: Sensitivity indices on the linear model with three Gaussian inputs and an interaction between  $X_2$  and  $X_3$ .

### 3.6. General model with three inputs and a block-additive structure

We consider the following model:

$$Y = g(X_1, X_2) + h(X_3), \quad (32)$$

which is called a “block-additive” structure. We do not assume anymore that the vector  $(X_1, X_2, X_3)^T$  is a Gaussian vector. We only assume that the three

inputs have finite variances and that  $X_3$  is independent from  $(X_1, X_2)$ . From independence properties one has:

$$\sigma^2 = \tau_{12}^2 + \tau_3^2, \quad \tau_{13}^2 = \tau_1^2 + \tau_3^2 \quad \text{and} \quad \tau_{23}^2 = \tau_2^2 + \tau_3^2. \quad (33)$$

From Equations (7), (10), (11) and (33) we get:

$$S_3 = Sh_3 = T_3^{\text{ind}}.$$

We also get that, for  $j = 1, 2$ , the three following assertions are equivalent

$$\begin{aligned} S_j &\leq Sh_j, \\ Sh_j &\leq S_{T_j}, \\ \frac{\tau_1^2 + \tau_2^2}{2} &\leq \frac{\tau_{12}^2}{2}. \end{aligned}$$

We study a test case with a block-additive structure, and choose, as in [14], the Ishigami function which is a non linear model involving interaction effects.

It writes:

$$f(\mathbf{X}) = \sin(X_1) + 7 \sin(X_2)^2 + 0.1 X_3^4 \sin(X_1) \quad (34)$$

where  $X_i \sim \mathcal{U}[-\pi, \pi] \forall i = 1, \dots, 3$  with a non-zero correlation  $\rho$  between a pair of variables.

Our first study considers some correlation between  $X_1$  and  $X_3$ , moreover  $X_2$  is assumed to be independent from  $X_1$  and  $X_3$ . This model has the structure of Eq. (32) with independence between the two blocks, up to a permutation on the inputs. These results are illustrated on Figure 4 obtained with a numerical estimation procedure that will be explained in the next section. We observe that the sandwich effect is respected for  $X_1$  and  $X_3$ . As  $X_2$  is independent from the group  $(X_1, X_3)$ , and as  $X_2$  has no interactions with that group, we observe that the Shapley index of  $X_2$  equals both its full first-order and independent total indices. Moreover, we observe some strong effects of the correlation between  $X_1$  and  $X_3$  which leads to an attraction between the Shapley effects of  $X_1$  and  $X_3$ , even for small correlation values. In the case of strong correlations, the Shapley effects of  $X_1$  and  $X_3$ , in which the effects of their interaction and their

correlation are distributed, have an easier interpretation than that of the Sobol' indices.

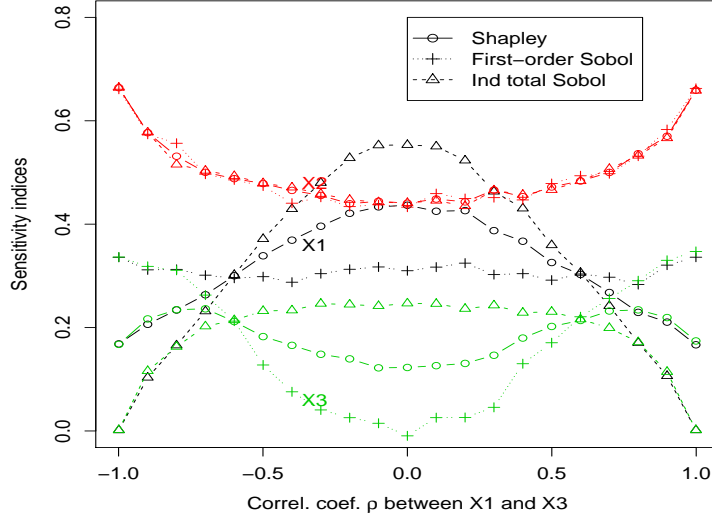


Figure 4: Sensitivity indices on the Ishigami function. Exact permutation method with  $N_o = 5 \times 10^3$ ,  $N_i = 3$ ,  $N_v = 10^4$ .

Our second and third studies consider respectively some correlation between  $X_1$  and  $X_2$ , then between  $X_2$  and  $X_3$ . Let us note that the model structure of Eq. (32) with independence between the two blocks is not conserved: The model cannot be decomposed in the sum of two terms which are independent. Figure 5 shows the results of these studies. Surprisingly, the sandwich effect seems still respected. However, it seems difficult to deduce general results on that more general kind of models. In these two studies, we also observe that the correlation effects are smaller on the Shapley effects than in the previous case (correlation between  $X_1$  and  $X_3$ ), except when the absolute value of the correlation is close to one. Clearly, the strong effects observed in the previous case are due to a conjunction between the correlation and the interaction between  $X_1$  and  $X_3$  present in the Ishigami function.

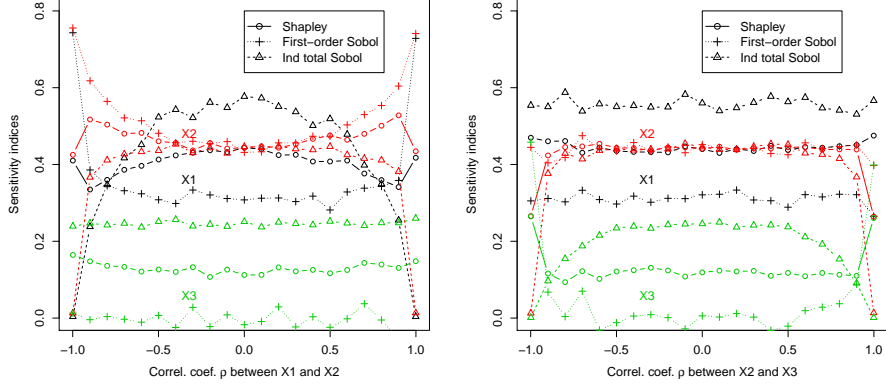


Figure 5: Sensitivity indices on the Ishigami function. Exact permutation method with  $N_o = 5 \times 10^3$ ,  $N_i = 3$ ,  $N_v = 10^4$ .

## 4. Numerical estimation issues

### 4.1. Estimation by direct sampling

In resuming previous work, [26] propose two algorithms for estimating the Shapley effects from formula (11) with

$$c(u) = \frac{\mathbb{E}(\text{Var}[Y|\mathbf{X}_u])}{\text{Var}(Y)} \quad (35)$$

as the cost function (which has been shown to be a more efficient cost function than the variance of the conditional expectation). We do not explain them in details but just give their main principles. The first must traverse all possible permutations between the inputs and is called the “Exact permutation method”. The second consists in randomly sample some permutations of the inputs and is called the “Random permutation method”. For each iteration of this loop on the inputs’ permutations, a conditional variance expectation must be computed. The cost  $C$ , in terms of number of model  $f$  evaluations, of these algorithms are then the following [26]:

1. Exact permutation method:  $C = N_i N_o d!(d - 1) + N_v$ , with  $N_i$  the inner loop size (conditional variance) in (35),  $N_o$  the outer loop size (ex-

pectation) in (35) and  $N_v$  the sample size for the variance computation (denominator in (35));

2. Random permutation method:  $C = N_i N_o m (d - 1) + N_v$ , with  $m$  the number of random permutations for discretizing the principal sum in (11).

Note that the full first-order Sobol indices (Eq. (7)) and the independent total Sobol' indices (Eq. (10)) are also estimated by applying these algorithms, each one during only one loop iteration.

From theoretical arguments, [26] have shown that the optimal values to use for the sizes of the different loops are the following:

- $N_i = 3$  and  $N_o = N_o^{\text{exa}}$  as large as possible for the exact permutation method,
- $N_i = 3$ ,  $N_o = N_o^{\text{rand}} = 1$  and  $m$  as large as possible for the random permutation method.

We consider these values in all our numerical tests and applications. In order to use the same cost for the two algorithms, we then find that

$$m = N_o^{\text{exa}} d! . \tag{36}$$

Finally, the choice of the value of  $N_v$  is independent from these values and we take  $N_v = 10^4$  in the following.

As test case, we consider the linear model with 3 Gaussian inputs of Section 3.4. Let us now take  $\beta_1 = \beta_2 = \beta_3 = 1$ ,  $\sigma_1 = \sigma_2 = 1$ ,  $\sigma_3 = 2$ . Figure 6 left (resp. right) gives the results of the exact permutation method (resp. random permutation method) in function of  $N_o$  (resp.  $m$ ) on this model with  $\rho = 0.9$ . By Eq (36) with  $d = 3$ , the same values of costs have been taken for two methods in order to be compared. The error bar in Figure 6 (right) is obtained from central limit theorem on the permutation loop (Monte Carlo sample of size  $m$ ) and then by taking two times the standard deviation of the estimates (95% confidence intervals). No such confidence interval can be obtained for the other method. In the left and right figures, we observe that the indices seem to

converge towards their theoretical values. Further analysis for the choice of  $m$ ,  $N_o$  and  $N_i$  will be nevertheless necessary.

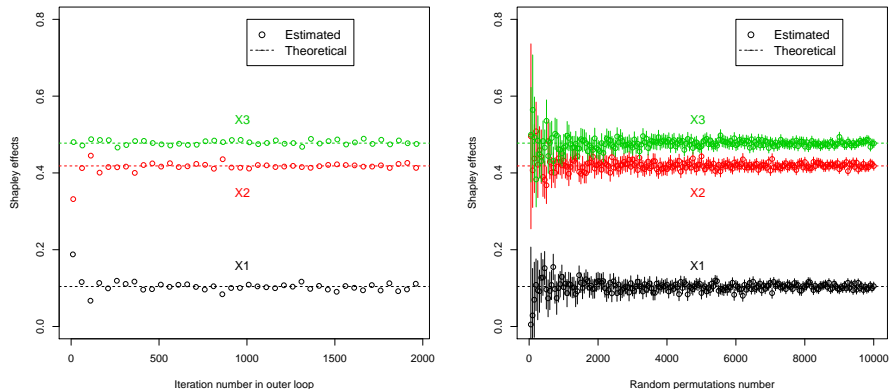


Figure 6: Numerical estimates by the exact permutation method (left) and the random permutation method (right) of the Shapley effects on the linear model.

By varying  $\rho$  between  $-0.9$  and  $0.9$ , Figure 7 shows the position of the Shapley effects and the Sobol' indices by using the methods of estimation (using the same cost by Eq. (36) with  $d = 3$ ). The theoretical results have been obtained in Section 3.4. First of all, the good agreement between the numerical results and the theoretical values (see Figure 2) is found. Second, the precisions of the numerical results of the two methods clearly seem equivalent.

#### 4.2. Metamodel-based estimation

In this section, we consider the case, relatively common in industrial applications, where the computational code is expensive in computation time, so that it cannot be evaluated intensively (e.g. only several hundred calculations are possible). It is therefore not possible to estimate the sensitivity indices with direct use of the model. Indeed, the estimate of Sobol' indices requires for each input several hundreds or thousands of evaluations of the model [1]). For the Shapley effects, the estimation cost is even more important, an additional loop being required.

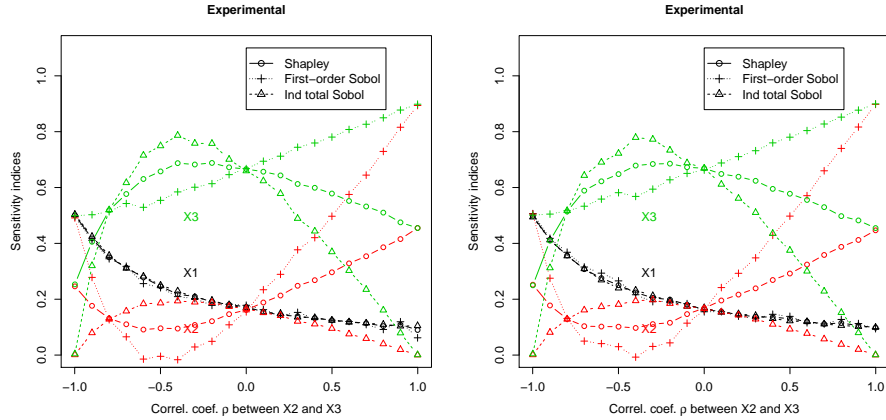


Figure 7: Numerical estimation of sensitivity indices on the linear model. Left: exact permutation method ( $N_o = 5 \times 10^3$ ,  $N_i = 3$ ,  $N_v = 10^4$ ). Right: random permutation method ( $m = 3 \times 10^4$ ,  $N_o = 1$ ,  $N_i = 3$ ,  $N_v = 10^4$ ).

We propose to use a metamodel instead of the model in the estimation procedure. A metamodel is a mathematical model of approximation of the numerical model, built on a learning basis [30]. The metamodel solution is a current engineering practice for estimating sensitivity indices [32]. We use the Gaussian process metamodel (also called kriging) [35, 36] which has been demonstrated in many practical situations with good predictive capacities (see [37] for example). The Gaussian process model is defined as follows:

$$Y(\mathbf{X}) = h(\mathbf{X}) + Z(\mathbf{X}), \quad (37)$$

where  $h(\cdot)$  is a deterministic trend function (typically a polynomial) and  $Z(\cdot)$  is a centered Gaussian process. We make the assumption that  $Z$  is second order stationary with variance  $\sigma^2$  and covariance Matérn 5/2 parametrized by its correlation lengths  $\theta \in \mathbb{R}^d$ . The hyperparameters  $\sigma^2$  and  $\theta$  are classically estimated by the maximum likelihood method on a learning sample of input/output simulation runs. Kriging provides an estimator of  $Y(\mathbf{X})$  which is called the kriging predictor and written  $\hat{Y}(\mathbf{X})$ . To quantify the prediction capabilities of the metamodel and to validate the predictor, the metamodel predictivity coefficient  $Q^2$  is estimated by cross-validation or on a test sample [37]. More precisely, the

Gaussian process model gives the following predictive distribution:

$$\forall \mathbf{X}^*, \quad (Y(\mathbf{X}^*) | y^N) \sim \mathcal{N}(\widehat{Y}(\mathbf{X}^*), \sigma_Y^2(\mathbf{X}^*)) \quad (38)$$

where  $\mathbf{X}^*$  is a point of the input space not contained in the learning sample,  $y^N$  is the output vector of the learning sample of size  $N$  and  $\sigma_Y^2(\mathbf{X})$  is the kriging variance that can also be explicitly estimated. In particular, the kriging variance  $\sigma_Y^2(\mathbf{X})$  quantifies the uncertainty induced by estimating  $Y(\mathbf{X})$  with  $\widehat{Y}(\mathbf{X})$ .

We study the Ishigami function (Eq. (34)) with a correlation coefficient  $\rho$  between  $X_1$  and  $X_3$ , on which [14] studied the Sobol' indices (see Section 3.6). When constructing the models, three different sizes  $N$  for the learning sample (50, 100 and 200) respectively give three predictive coefficients ( $Q^2$  which is equivalent to the  $R^2$  in prediction) different for the kriging predictor: 0.78, 0.88 and 0.98. Figure 8 shows that with a strong predictive metamodel ( $Q^2 = 0.98$  with  $N = 200$ ), the estimations of the Shapley effects by the metamodel have much precision. The results deteriorate rapidly with the decrease of the metamodel predictivity.

## 5. Industrial application

This application concerns a probabilistic analysis of an ultrasonic non-destructive control of a weld containing manufacturing defect. Complex phenomena occur in such heterogeneous medium during the ultrasonic wave propagation and a fine analysis to understand the effect of uncertain parameters is important. The simulation of this configuration is performed via the finite element code ATHENA2D, developed by EDF (Electricité de France). This code is dedicated to the simulation of elastic wave propagation in heterogeneous and anisotropic materials like welds.

A first study [38] has been realized with an inspection configuration aiming to detect a manufactured defect located in a 40 mm thick V grooveweld made of 316L steel (Figure 9). The weld material reveals a heterogeneous and anisotropic



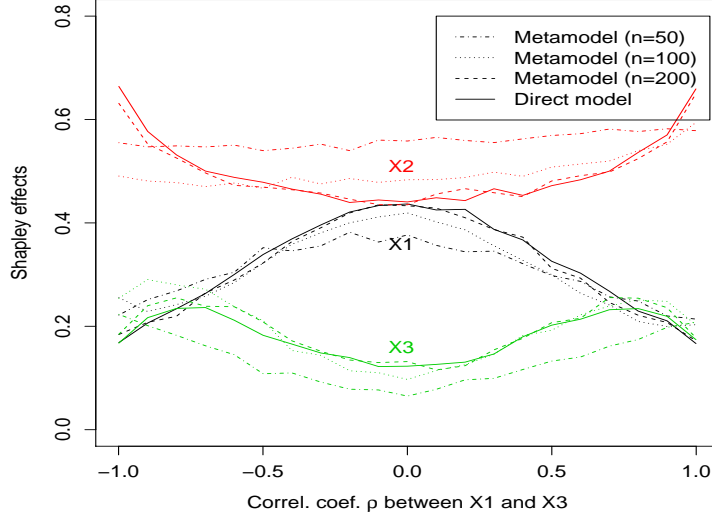


Figure 8: Sensitivity indices on the Ishigami function estimated by 3 metamodels built on 3 different learning bases. The exact permutation method (with  $N_o = 2 \times 10^3$ ,  $N_i = 3$ ,  $N_v = 10^4$ ) is applied on the metamodel predictor.

structure. It was represented by a simplified model consisting of a partition of 7 equivalent homogeneous regions with a specific grain orientation. 11 scalar input variables (4 elastic coefficients and 7 orientations of the columnar grains of the weld inspections) have been considered as uncertain and modeled by independent random variables, each one associated to a probability density function. The scalar output variable of the model is the amplitude of the defect echoes resulting from an ultrasonic inspection (maximum value on a so-called Bscan). Uncertainty and sensitivity analysis (based on polynomial chaos expansion [32]) have then been applied from 6000 Monte Carlo simulations of ATHENA2D in [38]. The sensitivity analysis has shown that all the inputs are influent, that the interaction effects are non negligible and that the orientations play a major role for explaining the amplitude variability. It confirms that an accurate determination of the microstructure is essential in these simulation studies. Finally, as a perspective of their work, [38] explains that the real configuration has been

strongly simplified by considering independent input random variables. Indeed, due to the welding physical process, dependence relationships exist between the orientations, in particular between two neighbor domains (see Figure 9 right).

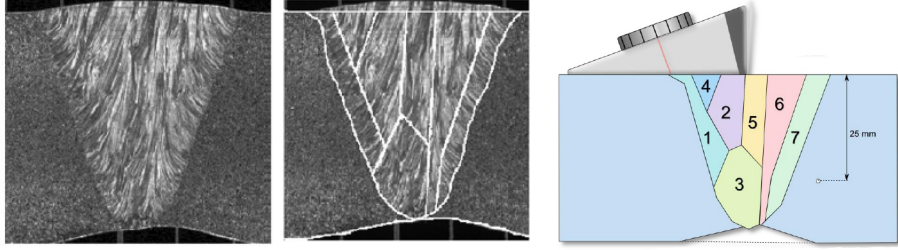


Figure 9: Metallographic picture (left), description of the weld in 7 homogeneous domains (middle) and inspection configuration (right).

The objective of the present study is then to perform a sensitivity analysis by using a more realistic probabilistic model for the input random variables. In our study, the probability distributions of all the inputs are considered Gaussian, with the same mean and standard deviation than in [38]. From physical models of welding process and solidification [39], engineers have been able to compute the following correlation matrix between the 7 orientations  $(\theta_1, \dots, \theta_7)$  of Figure 9 (right). They obtain:

$$\Sigma = \begin{pmatrix} 1 & 0.80 & 0.74 & 0.69 & 0.31 & 0.23 & 0.20 \\ 0.80 & 1 & 0.64 & 0.53 & 0.59 & 0.51 & 0.46 \\ 0.74 & 0.64 & 1 & 0.25 & 0.60 & 0.57 & 0.54 \\ 0.69 & 0.53 & 0.25 & 1 & -0.25 & -0.35 & -0.33 \\ 0.31 & 0.59 & 0.60 & -0.25 & 1 & 0.96 & 0.84 \\ 0.23 & 0.51 & 0.57 & -0.35 & 0.96 & 1 & 0.95 \\ 0.20 & 0.46 & 0.54 & -0.33 & 0.84 & 0.95 & 1 \end{pmatrix} .$$

As only several hundreds of numerical simulations of ATHENA2D can be performed in the schedule time of the study, our strategy consists in building a space filling designs in order to have a “good” learning sample for a metamodel building process. A Sobol’ sequence of  $N = 500$  points has then been generated for the  $d = 11$  input variables on  $[0, 1]^d$ . After transformations to physical input

values, while keeping the independent and uniform distributions for the inputs, the corresponding 500 runs of ATHENA2D have been computed. From the resulting  $N$ -size learning sample, a Gaussian process metamodel (parametrized as explained in Section 4.2) has then be fitted. We obtain a predictivity coefficient of  $Q^2 = 86.5\%$ . this result is rather satisfactory, specially when it is compared to the predictivity coefficient obtained by a simple linear model ( $Q^2 = 25\%$ ). Moreover, the test on Ishigami function (Section 4.2) has shown that the estimation of Shapley effects with a metamodel of predictivity close to 90% gives results rather close to the exact values.

We estimate the Shapley effects by using the metamodel predictor instead of ATHENA2D (Section 4.2). Due to the input dimension ( $d = 11$ ), the random permutation method is used with  $m = 10^4$ ,  $N_i = 3$ ,  $N_o = 1$  and  $N_v = 10^4$ . The cost is then  $3 \times 10^5$  in terms of required metamodel evaluations. It would be infeasible with the “true” computer code ATHENA2D, but it is feasible by using the metamodel predictor. Figure 10 gives the results of Shapley effects and Sobol’ indices of the elasticity coefficients ( $C11$ ,  $C13$ ,  $C55$ ,  $C13$ ) and orientations ( $Or1$ ,  $Or2$ ,  $Or3$ ,  $Or4$ ,  $Or5$ ,  $Or6$ ,  $Or7$ ). For the correlated inputs (the orientations), Shapley effects clearly succeed to discriminate influent inputs ( $Or3$ , then  $Or1$ , then  $Or2$  whose effects are larger than 12%) to less influent ones (the others whose effects are smaller than 6%). It seems that Sobol’ indices, by the way of full first order ones, are more difficult to interpret and allow a poorer discrimination on the influences. Indeed, from these indices, one can conclude that all the orientations are more influent than the elastic constants. However, the Sobol’ indices of  $Or4$ ,  $Or5$ ,  $Or6$  and  $Or7$  are mainly due to their dependence with other inputs, as Shapley effects show that the effects of  $Or4$ ,  $Or5$ ,  $Or6$  and  $Or7$  and the elastic constants are of the same order. Moreover, due to the cost  $c(\cdot)$  used in the estimation method, the estimation of the full first-order indices are imprecise. The negative estimated indices obtained for  $C33$ ,  $C55$  and  $C13$  come from this imprecision (theoretically these indices are positive) and one can conclude that these are indices are close to zero.

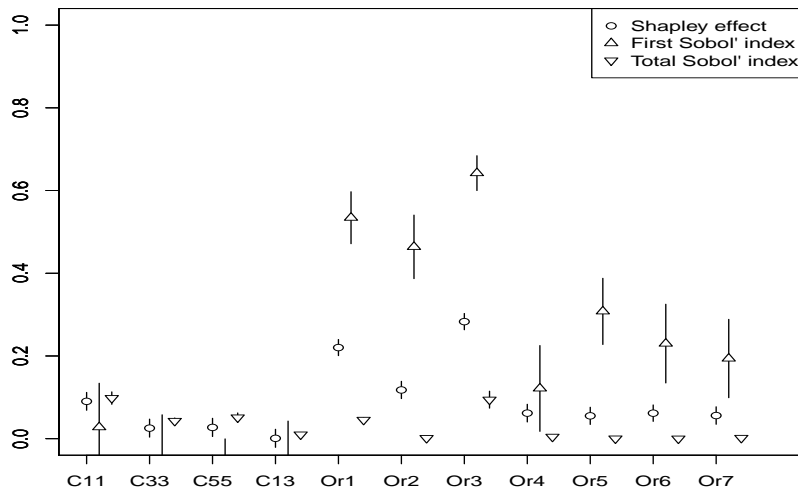


Figure 10: Shapley effects and Sobol' indices for the ultrasonic propagation application.

## 6. Conclusion

In many applications of global sensitivity analysis methods, it is common that the input variables have a known statistical dependence structure or that the input space is constrained to a non-rectangular region. In this paper we considered two answers to that issue: the Shapley effects (a normalized version of the Shapley values proposed in [29] in the framework of sensitivity analysis) and the methodology developed in [15]. This last methodology suggests the joint analysis of full and independent first-order and total indices, to analyze the sensitivity of a model to dependent inputs. In the present paper, we have lead a comparative analysis between Shapley effects on one side and full first-order and independent total indices on the other side. From analytical solutions obtained with linear models and Gaussian variables, we have shown that the dependence between inputs lead to a rebalancing of the corresponding Shapley effects, while a full Sobol' index of an input adds to its own effect the effect of an input of which it is dependent. Comparisons of Shapley effects with the complementary independent first-order and full total indices will be made in a

further study.

We have also illustrated the convergence of two numerical algorithms of estimating Shapley effects. These algorithms depend on various parameters:  $N_i$ ,  $N_o$ ,  $N_v$  and  $m$ . It would be interesting to investigate further the response of the algorithms to these different parameters and to derive empirical and asymptotic confidence intervals for the Shapley effects estimates. Finally, we have shown the relevance of using a metamodel (here the Gaussian process predictor) in the practical situations where the computer model is too time consuming to be evaluated thousands of times to apply the previous algorithms. Future work will consist in developing an algorithm exploiting the complete structure of the Gaussian process allowing to infer the error due to this approximation (see [40], [41] and [42] for the Sobol' indices and [43] for the Derivative-based Global Sensitivity Measures).

### Acknowledgments

We are grateful to Géraud Blatman who has performed the computations on ATHENA2D model and Roman Sueur for helpful discussions about Shapley effects. Numerical estimation of Shapley effects and Sobol' indices have been realized using the sensitivity package of the R software. Thanks to Eunhye Song, Barry L. Nelson and Jeremy Staum for providing preliminary versions of the Shapley effects estimation functions.

### References

- [1] B. Iooss, P. Lemaître, A review on global sensitivity analysis methods, in: C. Meloni, G. Dellino (Eds.), *Uncertainty management in Simulation-Optimization of Complex Systems: Algorithms and Applications*, Springer, 2015, pp. 101–122.
- [2] I. Sobol, Sensitivity estimates for non linear mathematical models, *Mathematical Modelling and Computational Experiments* 1 (1993) 407–414.

- [3] D. Kurowicka, R. Cooke, *Uncertainty analysis with high dimensional dependence modelling*, Wiley, 2006.
- [4] M. Petelet, B. Iooss, O. Asserin, A. Loredo, Latin hypercube sampling with inequality constraints, *Advances in Statistical Analysis* 94 (2010) 325–339.
- [5] S. Kucherenko, O. Klymenko, N. Shah, Sobol’ indices for problems defined in non-rectangular domains, *Reliability Engineering and System Safety* 167 (2017) 218–231.
- [6] A. Saltelli, S. Tarantola, On the relative importance of input factors in mathematical models: Safety assessment for nuclear waste disposal, *Journal of American Statistical Association* 97 (2002) 702–709.
- [7] S. Da Veiga, F. Wahl, F. Gamboa, Local polynomial estimation for sensitivity analysis on models with correlated inputs, *Technometrics* 51 (4) (2009) 452–463.
- [8] J. Jacques, C. Lavergne, N. Devictor, Sensitivity analysis in presence of model uncertainty and correlated inputs, *Reliability Engineering and System Safety* 91 (2006) 1126–1134.
- [9] C. Xu, G. Gertner, Uncertainty and sensitivity analysis for models with correlated parameters, *Reliability Engineering and System Safety* 93 (2008) 1563–1573.
- [10] G. Li, H. Rabitz, P. Yelvington, O. Oluwole, F. Bacon, C. Kolb, J. Schoendorf, Global sensitivity analysis for systems with independent and/or correlated inputs, *Journal of Physical Chemistry* 114 (2010) 6022–6032.
- [11] G. Li, H. Rabitz, J. Hu, Z. Chen, Y. Ju, Regularized random-sampling high dimensional model representation, *Journal of Mathematical Chemistry* 43 (3) (2008) 6022–6032.
- [12] B. Sudret, Y. Caniou, Analysis of covariance (ANCOVA) using polynomial chaos expansions, in: G. Deodatis (Ed.), *Proceedings of 11th Int. Conf. Struct. Safety and Reliability (ICOSSAR2013)*, New-York, USA, 2013.

- [13] T. Mara, S. Tarantola, Variance-based sensitivity indices for models with dependent inputs, *Reliability Engineering and System Safety* 107 (2012) 115–121.
- [14] S. Kucherenko, S. Tarantola, P. Annoni, Estimation of global sensitivity indices for models with dependent variables, *Computer Physics Communications* 183 (2012) 937–946.
- [15] T. Mara, S. Tarantola, P. Annoni, Non-parametric methods for global sensitivity analysis of model output with dependent inputs, *Environmental Modeling & Software* 72 (2015) 173–183.
- [16] M. Rosenblatt, Remarks on a multivariate transformation, *The Annals of Mathematical Statistics* 23 (3) (1952) 470–472.
- [17] J.-P. Gauchi, A. Bensadoun, F. Colas, N. Colbach, Metamodeling and global sensitivity analysis for computer models with correlated inputs: A practical approach with a 3D light interception computer model, *Environmental Modeling & Software* 92 (2017) 40–56.
- [18] B. Sudret, Global sensitivity analysis using polynomial chaos expansion, *Reliability Engineering and System Safety* 93 (2008) 964–979.
- [19] E. Borgonovo, A new uncertainty importance measure, *Reliability Engineering and System Safety* 92 (2007) 771–784.
- [20] E. Borgonovo, W. Castaings, S. Tarantola, Moment independent importance measures: new results and analytical test cases, *Risk Analysis* 31 (2011) 404–428.
- [21] W. Hoeffding, A class of statistics with asymptotically normal distributions, *Annals of Mathematical Statistics* 19 (1948) 293–325.
- [22] C. J. Stone, The use of polynomial splines and their tensor products in multivariate function estimation, *The Annals of Statistics* 22 (1) (1994) 118–184.

- [23] G. Hooker, Generalized functional anova diagnostics for high-dimensional functions of dependent variables, *Journal of Computational and Graphical Statistics*.
- [24] G. Chastaing, F. Gamboa, C. Prieur, Generalized Hoeffding-Sobol decomposition for dependent variables - Application to sensitivity analysis, *Electronic Journal of Statistics* 6 (2012) 2420–2448.
- [25] A. Owen, C. Prieur, On Shapley value for measuring importance of dependent inputs, *SIAM/ASA Journal on Uncertainty Quantification*, In press. URL <https://arxiv.org/abs/1610.02080>
- [26] E. Song, B. Nelson, J. Staum, Shapley effects for global sensitivity analysis: Theory and computation, *SIAM/ASA Journal of Uncertainty Quantification* 4 (2016) 1060–1083.
- [27] B. Iooss, C. Prieur, Analyse de sensibilité avec entrées dépendantes : estimation par échantillonnage et par métamodèles des indices de Shapley, in: *Proceedings of 49èmes Journées de Statistique*, Avignon, France, 2017.
- [28] L. Shapley, A value for n-persons game, in: H. Kuhn, A. Tucker (Eds.), *Contributions to the theory of games II*, *Annals of mathematic studies*, Princeton University Press, Princeton, NJ, 1953.
- [29] A. Owen, Sobol’ indices and Shapley value, *SIAM/ASA Journal of Uncertainty Quantification* 2 (2014) 245–251.
- [30] K.-T. Fang, R. Li, A. Sudjianto, *Design and modeling for computer experiments*, Chapman & Hall/CRC, 2006.
- [31] B. Iooss, F. Van Dorpe, N. Devictor, Response surfaces and sensitivity analyses for an environmental model of dose calculations, *Reliability Engineering and System Safety* 91 (2006) 1241–1251.
- [32] L. Le Gratiet, S. Marelli, B. Sudret, Metamodel-based sensitivity analysis: Polynomial chaos expansions and Gaussian processes, in: R. Ghanem,



- D. Higdon, H. Owhadi (Eds.), Springer Handbook on Uncertainty Quantification, Springer, 2017, pp. 1289–1325.
- [33] C. Prieur, S. Tarantola, Variance-based sensitivity analysis: Theory and estimation algorithms, in: R. Ghanem, D. Higdon, H. Owhadi (Eds.), Springer Handbook on Uncertainty Quantification, Springer, 2017, pp. 1217–1239.
- [34] T. Homma, A. Saltelli, Importance measures in global sensitivity analysis of non linear models, Reliability Engineering and System Safety 52 (1996) 1–17.
- [35] J. Sacks, W. Welch, T. Mitchell, H. Wynn, Design and analysis of computer experiments, Statistical Science 4 (1989) 409–435.
- [36] T. Santner, B. Williams, W. Notz, The design and analysis of computer experiments, Springer, 2003.
- [37] A. Marrel, B. Iooss, F. Van Dorpe, E. Volkova, An efficient methodology for modeling complex computer codes with Gaussian processes, Computational Statistics and Data Analysis 52 (2008) 4731–4744.
- [38] F. Rupin, G. Blatman, S. Lacaze, T. Fouquet, B. Chassignole, Probabilistic approaches to compute uncertainty intervals and sensitivity factors of ultrasonic simulations of a weld inspection, Ultrasonics 54 (2014) 1037–1046.
- [39] J. Moysan, A. Apfel, C. Corneloup, B. Chassignole, Modelling the grain orientation of austenitic stainless steel multipass welds to improve ultrasonic assessment of structural integrity, Int. J. Pressure Ves. Piping 80 (2003) 77–85.
- [40] A. Janon, M. Nodet, C. Prieur, Uncertainties assessment in global sensitivity indices estimation from metamodels, International Journal for Uncertainty Quantification 4 (1).

- [41] A. Janon, T. Klein, A. Lagnoux, M. Nodet, C. Prieur, Asymptotic normality and efficiency of two sobol index estimators, *ESAIM: Probability and Statistics* 18 (2014) 342–364.
- [42] L. Le Gratiet, C. Cannamela, B. Iooss, A Bayesian approach for global sensitivity analysis of (multifidelity) computer codes, *SIAM/ASA Journal of Uncertainty Quantification* 2 (2014) 336–363.
- [43] M. De Lozzo, A. Marrel, Estimation of the derivative-based global sensitivity measures using a Gaussian process metamodel, *Journal on Uncertainty Quantification* 4 (2016) 708–738.