# Shapley effects for sensitivity analysis with dependent inputs: bootstrap and kriging-based algorithms

Nazih Benoumechiara, Kevin Elie-Dit-Cosaque

HAL Id: hal-01677501

https://hal.sorbonne-universite.fr/hal-01677501v1

Submitted on 9 Jan 2018

# Shapley effects for sensitivity analysis with dependent inputs: bootstrap and kriging-based algorithms

Nazih Benoumechiara[1, 3] and Kevin Elie-Dit-Cosaque[2, 4]

[1]Laboratoire de Probabilités, Statistique et Modélisation, Sorbonne Universités, Paris, FRANCE
[2]Institut Camille Jordan, Université Claude Bernard Lyon 1, Lyon, FRANCE
[3]EDF Lab, Chatou, FRANCE
[4]Group Actuarial Department, SCOR SE, Paris, FRANCE

**Abstract**

In global sensitivity analysis, the well known Sobol' sensitivity indices aim to quantify how the variance in the output of a mathematical model can be apportioned to the different variances of its input random variables. These indices are based on the functional variance decomposition and their interpretation become difficult in the presence of statistical dependence between the inputs. However, as there is dependence in many application studies, that enhances the development of interpretable sensitivity indices. Recently, the Shapley values developed in the field of cooperative games theory have been connected to global sensitivity analysis and present good properties in the presence of dependencies. Nevertheless, the available estimation methods don't always provide confidence intervals and require a large number of model evaluation. In this paper, we implement a bootstrap sampling in the existing algorithms to estimate confidence intervals of the indice estimations. We also proposed to consider a metamodel in substitution of a costly numerical model. The estimation error from the Monte-Carlo sampling is combined with the metamodel error in order to have confidence intervals on the Shapley effects. Besides, we compare for different examples with dependent random variables the results of the Shapley effects with existing extensions of the Sobol' indices.

## 1 Introduction

In the last decades, computational models have been increasingly used to approximate physical phenomenons. The steady improvement of computational means, lead to the use of very complex numerical codes involving an increasing number of parameters. In many situations, the model inputs are uncertain resulting in uncertain outputs. In this case it is necessary to understand the global impact of inputs uncertainties on the output to validate the computer code and use it properly. Sensitivity Analysis methods aim at solving this range of issues by characterizing input-output relationships of computer codes.

Within Sensitivity Analysis, three kinds of methods can be distinguished. First, Screening methods aim to discriminate influential inputs from non influential ones, especially when the inputs are numerous and the problem should be simplified. Secondly, local methods, based on partial derivatives, are used to assess the influence of input variables for small perturbations. Finally, Global Sensitivity Analysis (GSA) methods aim at ranking input random variables according to their importance in the output uncertainty, or even quantify the global influence of a particular input on the output. In this paper we are specifically interested in global sensitivity analysis. One can refer to Iooss and Lemaître (2015) for a comprehensive review of sensitivity

1

analysis methods.

Among GSA methods, variance-based approaches are a class of probabilistic ones that measure the part of variance of the model output which is due to the variance of a particular input. Theses methods were popularized by Sobol (1993) who introduced the well known first order Sobol' indices. Shortly after, the total Sobol' indices have been introduced by Homma and Saltelli (1996) also taking advantage of Jansen et al. (1994). These sensitivity indices are based on the functional ANalyse Of VAriance (ANOVA) which is unique only making the assumption of independence between the input random variables. However, this hypothesis is sometimes not verified in practice making their interpretation much harder. Several works have been carried to deal with this difficulty and extend Sobol' indices to the case of a stochastic dependence between the input variables, as Chastaing et al. (2012); Mara and Tarantola (2012); Mara et al. (2015); Kucherenko et al. (2012). But the practical estimation of these sensitivity measures and their interpretation remain difficult.

Recently, Owen (2014) established a relation between the Shapley values (Shapley and Shubik, 1954) coming from the field of game theory and Sobol' indices. Song et al. (2016) proposed an algorithm to estimate these indices. Some studies also highlighted the potential of this kind of index in the case of correlated input, as Owen and Prieur (2017); Iooss and Prieur (2017). In this last case the Shapley effects can be a good alternative to the existing extensions of Sobol' indices mentioned above. Indeed, Shapley effects allows an apportionment of the interaction and dependences contributions between the input involved, making them condensed and easy-to-interpret indices.

Most estimation procedures of the Sobol' indices and Shapley effects are based on Monte-Carlo sampling. These methods require large sample sizes in order to have a sufficiently low estimation error. When dealing with costly computational models, a precise estimation of these indices can be difficult to achieve or even unfeasible. Therefore, the use of a surrogate model (or metamodel) instead of the actual one can be a good alternative and dramatically decrease the computational cost of the estimation. Various kinds of surrogate model exists in the literature, such as Fang et al. (2005). In this paper we get interested in the use of kriging as metamodels (see for example Martin and Simpson (2004)). The approach developed here is based on Le Gratiet et al. (2014) who provides an estimation algorithm of Sobol' indices using kriging models which allows computing the meta-model and Monte-Carlo errors.

In this paper, we draw a comparison between the Shapley effects and the *independent* and *full* Sobol' indices defined in Mara et al. (2015). We also establish an extension of the Shapley estimation algorithm proposed in Song et al. (2016) by implementing a bootstrap sampling to catch the Monte-Carlo error. Inspired by the work of Le Gratiet et al. (2014), we used a kriging model in substitution of the true model for the estimation of these indices. Thus, the kriging model error is associated to the Monte-Carlo error in order to correctly catch the overall estimation error.

The paper's outline is as follow: Section 2 recalls the basic concept of Sobol' indices in the independent and dependent configuration; Section 3 introduces the Shapley values and their links with sensitivity analysis; Section 4 theoretically compares the Sobol' indices and the Shapley effects for two toy examples; Section 5 studies the quality of the estimated Shapley effects and their confidence intervals; Section 6 introduces the kriging model and how the kriging and Monte-Carlo errors can be separated from the overall error; Section 7 compares the indice performances using a kriging model on two toy examples; finally, Section 8 synthesizes this work and suggests some perspectives.

2

# 2 Sobol' sensitivity indices

## 2.1 Sobol' indices with independent inputs

Consider a model $Y = \eta(\mathbf{X})$ with $d$ random inputs denoted by $\mathbf{X}_\mathcal{D} = \{X_1, X_2, \ldots, X_d\}$, where $\mathcal{D} = \{1, 2, \ldots, d\}$, and $\mathbf{X}_\mathcal{J}$ indicates the vector of inputs corresponding to the index set $\mathcal{J} \subseteq \mathcal{D}$. $\eta : \mathbb{R}^d \to \mathbb{R}$ is a deterministic squared integrable function and $Y \in \mathbb{R}$ the model output random variable. The random vector $\mathbf{X}$ follows a distribution $p_\mathbf{X}$ and we suppose, in this section, that $p_\mathbf{X}$ follows a $d$-dimensional uniform distribution $\mathcal{U}([0,1]^d)$. However, these results can be extended to any marginal distributions. In particular, all inputs are independent and the distribution of $\mathbf{X}$ is only defined by its margins.

The Hoeffding decomposition introduced in Hoeffding (1948), also known as high dimensional model representation (HDMR) (Li et al., 2001), allows writing $\eta(\mathbf{X})$ in the following way:

$$\eta(\mathbf{X}) = \eta_\emptyset + \sum_{i=1}^{d} \eta_i(X_i) + \sum_{1 \leqslant i < j \leqslant d} \eta_{i,j}(X_i, X_j) + \cdots + \eta_{1,\ldots,d}(\mathbf{X}), \tag{1}$$

for some $\eta_\emptyset, \eta_i, \ldots, \eta_{1,\ldots,d}$ set of functions. In this formula, $\eta$ is decomposed into $2^d$ terms such as $\eta_\emptyset$ is a constant and the other terms are square integrable functions.

The decomposition (1) is not unique due to the infinite possible choice for these terms. The unicity condition is granted by the following orthogonality constraint:

$$\int_0^1 \eta_{i_1, i_2, \ldots, i_s}(x_{i_1}, x_{i_2}, \ldots, x_{i_s}) \mathrm{d}x_{i_w} = 0, \tag{2}$$

where $1 \leq i_1 < i_2 < \cdots < i_s \leq d$ and $i_w \in \{i_1, i_2, \ldots, i_s\}$. The consequence of this condition is that the terms of (1) are orthogonal to one another. This property implies the independence of the random variables $X_i$ in the stochastic configuration and allow to obtain the following expressions for the functions $\eta_{i_1, i_2, \ldots, i_s}$ of (1) :

$$\eta_\emptyset = \mathbb{E}(Y), \tag{3}$$
$$\eta_i(X_i) = \mathbb{E}_{\mathbf{X}_{\sim i}}(Y|X_i) - \mathbb{E}(Y), \tag{4}$$
$$\eta_{i,j}(X_i, X_j) = \mathbb{E}_{\mathbf{X}_{\sim ij}}(Y|X_i, X_j) - \eta_i - \eta_j - \mathbb{E}(Y) \tag{5}$$

where $\mathbf{X}_{\sim i} = \mathbf{X}_{\mathcal{D} \setminus \{i\}}$ and similarly for higher orders. Thus, the functions $\{\eta_i\}_{i=1}^{d}$ are the main effects, the $\eta_{i,j}$ for $i < j = 1, \ldots, d$ are the second-order interaction effects, and so on.

The representation (1) leads to the functional ANAlyse Of VAriance (ANOVA) which decompose the global variance into a sum of partial variances such as

$$\mathrm{Var}(Y) = \sum_{i=1}^{d} \mathrm{Var}[\eta_i(X_i)] + \sum_{i=1}^{d} \sum_{i<j}^{d} \mathrm{Var}[\eta_{i,j}(X_i, X_j)] + \cdots + \mathrm{Var}[\eta_{1,\ldots,d}(\mathbf{X})]. \tag{6}$$

The so-called Sobol' indices (Sobol, 1993) can be derived from (6) by dividing both sides with $\mathrm{Var}(Y)$. This operation results in the following property:

$$\sum_{i=1}^{d} S_i + \sum_{i=1}^{d} \sum_{i<j}^{d} S_{ij} + \cdots + S_{1,\ldots,d} = 1, \tag{7}$$

3

where $S_i$ is a first-order sensitivity index, $S_{ij}$ is a second-order sensitivity index and so on. Thus, sensitivity indices are defined as

$$S_i = \frac{\text{Var}[\eta_i(X_i)]}{\text{Var}(Y)}, \qquad S_{ij} = \frac{\text{Var}[\eta_{i,j}(X_i, X_j)]}{\text{Var}(Y)}, \qquad \ldots \tag{8}$$

The first-order index $S_i$ measures the part of variance of the model output that is due to the variable $X_i$, the second-order $S_{ij}$ measure the part of variance of the model output that is due to the interaction of $X_i$ and $X_j$ and so on for higher interaction orders.

Another popular variance based coefficient, called total Sobol' index by Homma and Saltelli (1996), gathers the first-order effect of a variable with all its interactions. This index is defined by

$$ST_i = S_i + \sum_{i \neq j} S_{ij} + \cdots + S_{1,\ldots,d} = 1 - \frac{\text{Var}_{\mathbf{X}\setminus i}[\mathbb{E}_{X_i}(Y|\mathbf{X}_{\sim i})]}{\text{Var}(Y)} = \frac{\mathbb{E}_{\mathbf{X}_{\sim i}}[\text{Var}_{X_i}(Y|\mathbf{X}_{\sim i})]}{\text{Var}(Y)}. \tag{9}$$

The property (7) does not always hold for the total indices as summing total indices for all variables introduces redundant interactions terms appearing only once in (7). Thus, in most cases $\sum_i^d ST_i \geq 1$. Note that both the first order and total Sobol' indices are normalized measures.

As mentioned in the introduction (6) holds only if the random variables are independent. Different approaches exist to treat the case of dependent input and one of them is explained in Section 2.2.

## 2.2 Sobol' indices with dependent inputs

In this section, we suppose that $\mathbf{X} \sim p_{\mathbf{X}}$ with correlations between the components. Thanks to the Rosenblatt Transformation (RT) (Rosenblatt, 1952), it is possible to transform $\mathbf{X} \sim p_{\mathbf{X}}$ into an uniform and independent random vector $\mathbf{U} \sim \mathcal{U}^d(0,1)$. However, due to the possible permutations of elements of $\mathbf{X}$, this transformation is not unique and has $d!$ possibilities. Note that in this procedure, only the $d$ RT obtained after left circularly reordering the elements $(X_1, \ldots, X_d)$ are considered. We denote as $\mathbf{U}^i = (U_1^i, \ldots, U_d^i)$ the RT of the set $(X_i, X_{i+1}, \ldots, X_d, X_1, \ldots, X_{i-1})$ such as

$$(X_i, X_{i+1}, \ldots, X_d, X_1, \ldots, X_{i-1}) \sim p_{\mathbf{X}} \xrightarrow{RT} (U_1^i, \ldots, U_d^i) \sim \mathcal{U}^d(0,1). \tag{10}$$

It is important to note that this RT corresponds to a particular ordering $i$. Changing the ordering results in another RT. Such a mapping is bijective and we can consider a function $g_i$ such as $Y = \eta(\mathbf{X}) = g_i(\mathbf{U}^i)$. Because the elements of $\mathbf{U}^i$ are independent, the ANOVA decomposition is unique and can be established to compute sensitivity indices. Thus, we can write

$$g_i(\mathbf{U}^i) = g_\emptyset + \sum_{j_1=1}^{d} g_{j_1}(U_{j_1}^i) + \sum_{j_1=1}^{d} \sum_{j_1 < j_2}^{d} g_{j_1 j_2}(U_{j_1}^i, U_{j_2}^i) + \cdots + g_{1\cdots d}(U_1^i, \ldots, U_d^i) \tag{11}$$

where $g_\emptyset = \mathbb{E}[g_i(\mathbf{U}^i)]$. Because the summands in (11) are orthogonal, the variance based decomposition can be derived, such that

$$\text{Var}(Y) = \sum_{j_1=1}^{d} V_{j_1} + \sum_{j_1=1}^{d} \sum_{j_1 < j_2}^{d} V_{j_1, j_2} + \cdots + V_{1\cdots d} \tag{12}$$

where $V_{j_1} = \text{Var}[\mathbb{E}(Y|U_{j_1}^i)]$, $V_{j_1,j_2} = \text{Var}[\mathbb{E}(Y|U_{j_1}, U_{j_2})] - V_{j_1} - V_{j_2}$ and so on for higher orders. The so-called Sobol' indices are defined by dividing (12) with the total variance $\text{Var}(Y)$ such that,

$$S_{u_k}^i = \frac{\text{Var}[\mathbb{E}[g_i(\mathbf{U}^i)|U_k^i]]}{\text{Var}[g_i(\mathbf{U}^i)]} = \frac{V_k}{\text{Var}(Y)}. \tag{13}$$

We also consider the total sensitivity indice introduced by Saltelli (2002) which is the overall contribution of $U_k^i$ on the model output including the marginal and interaction effects. They can be written as

$$ST_{u_k}^i = \frac{\mathbb{E}[\text{Var}[g_i(\mathbf{U}^i)|U_{\sim k}^i]]}{\text{Var}[g_i(\mathbf{U}^i)]} = \frac{\sum_{s=1}^d \sum_{\{j_1,\ldots,j_s\} \ni k} V_{j_1 \cdots j_s}}{\text{Var}(Y)}, \tag{14}$$

where $U_{\sim k}^i$ is the vector $\mathbf{U}^i$ not containing $U_k^i$. We refer to Iooss and Lemaître (2015) for a review on the sensitivity indices and their properties.

Mara et al. (2015) derived equations (13) and (14), to introduce two types of indices which deal with correlated variables:

- the *full* Sobol' indices which describe the influence of a variable including its dependencies with other variables,

- the *independent* Sobol' indices which describe the influence of variables without its dependencies with other variables.

For a given ordering $(X_i, X_{i+1}, \ldots, X_d, X_1, \ldots, X_{i-1})$, the joint density of $\mathbf{X}$ can be written as

$$p(x) = p_i(x_i)p_{i+1}(x_{i+1}|x_i)\ldots p_d(x_d|x_i, x_{i+1}, \ldots, x_{d-1})p_1(x_1|x_i, \ldots, x_d), \ldots, p_{i-1}(x_{i-1}|\mathbf{x}_{\sim(i-1)}). \tag{15}$$

From this ordering and a given RT, the full and independent Sobol' indices can be described by

$$S_i = \frac{\text{Var}[\mathbb{E}[g_i(\mathbf{U}^i)|U_1^i]]}{\text{Var}[g_i(\mathbf{U}^i)]} \tag{16}$$

$$ST_i = \frac{\mathbb{E}[\text{Var}[g_i(\mathbf{U}^i)|U_{\sim 1}^i]]}{\text{Var}[g_i(\mathbf{U}^i)]} \tag{17}$$

$$S_i^{ind} = \frac{\text{Var}[\mathbb{E}[g_{i+1}(\mathbf{U}^{i+1})|U_d^{i+1}]]}{\text{Var}[g_{i+1}(\mathbf{U}^{i+1})]} \tag{18}$$

$$ST_i^{ind} = \frac{\mathbb{E}[\text{Var}[g_{i+1}(\mathbf{U}^{i+1})|U_{\sim d}^{i+1}]]}{\text{Var}[g_{i+1}(\mathbf{U}^{i+1})]} \tag{19}$$

with the convention that $\mathbf{U}^{d+1} = \mathbf{U}^1$.

Thanks to RT, we can also define the sensitivity indices of $(X_i|X_u), i = 1, \ldots, d$ and $u \subset \mathcal{D}\backslash\{i\}, u \neq \emptyset$ via $U_u^i$ which represent the effect of $X_i$ without its mutual dependent contribution with $X_u$. These indices can be estimated with a Monte Carlo algorithm and the procedure is describe in the next section.

## 2.3 Estimation

The estimation of $(S_i, ST_i, S_{i-1}^{ind}, ST_{i-1}^{ind})$ can be done with four samples using the "pick and freeze" strategy (see Saltelli et al. (2010)). The procedure is divided in two steps:

- 
  - Two independent sampling matrices $\mathbf{A}$ and $\mathbf{B}$ of size $N \times d$ are created from $\mathcal{U}(0,1)^d$

– $\mathbf{B}_{\mathbf{A}}^{(1)}\left(\mathbf{B}_{\mathbf{A}}^{(d)}\right)$ : all columns from $\mathbf{B}$ except the 1-th ($d$-th) column which is from $\mathbf{A}$

- Compute the indices with a given estimator:

$$\widehat{S}_i = \frac{\frac{1}{N}\sum_{j=1}^{N} g_i(A)_j g_i(\mathbf{B}_{\mathbf{A}}^{(1)})_j - g_{i_0}^2}{\widehat{V}} \tag{20}$$

$$\widehat{ST}_i = 1 - \frac{\frac{1}{N}\sum_{j=1}^{N} g_i(B)_j g_i(\mathbf{B}_{\mathbf{A}}^{(1)})_j - g_{i_0}^2}{\widehat{V}} \tag{21}$$

$$\widehat{S}_{i-1}^{ind} = \frac{\frac{1}{N}\sum_{j=1}^{N} g_i(A)_j g_i(\mathbf{B}_{\mathbf{A}}^{(d)})_j - g_{i_0}^2}{\widehat{V}} \tag{22}$$

$$\widehat{ST}_{i-1}^{ind} = 1 - \frac{\frac{1}{N}\sum_{j=1}^{N} g_i(B)_j g_i(\mathbf{B}_{\mathbf{A}}^{(d)})_j - g_{i_0}^2}{\widehat{V}} \tag{23}$$

where $g_{i_0}$ is the estimate of the mean and $\widehat{V} = \frac{1}{N}\sum_{j=1}^{N}(g_i(A)_j)^2 - g_{i_0}^2$

This procedure considers the estimator from Janon et al. (2014) and the overall cost is $4dN$ with $N$ the number of samples. However, another estimator can be used to estimate the indices. See Saltelli et al. (2010) for a review of various estimators of sensitivity indices.

# 3 Shapley effects

The purpose of the Sobol' indices is to decompose $\mathrm{Var}(Y)$ and allocate it to *each subset* $\mathcal{J}$ whereas the Shapley effects decompose $\mathrm{Var}(Y)$ and allocate it to *each input* $X_i$. This difference allows to consider any variables regardless of their dependence with other inputs.

## 3.1 Definition

One of the main issues in cooperative games theory is to define a relevant way to allocate the earnings between players. A fair share of earnings of a $d$ players coalition has been proposed in Shapley (1953). Formally, in Song et al. (2016) a $d$-player game with the set of players $\mathcal{D} = \{1, 2, \ldots, d\}$ is defined as a real-valued function that maps a subset of $\mathcal{D}$ to its corresponding cost, i.e., $c : 2^{\mathcal{D}} \mapsto \mathbb{R}$ with $c(\emptyset) = 0$. Hence, $c(\mathcal{J})$ represents the cost that arises when the players in the subset $\mathcal{J}$ of $\mathcal{D}$ participate in the game. The Shapley value of player $i$ with respect to $c(\cdot)$ is defined as

$$v^i = \sum_{\mathcal{J} \subseteq \mathcal{D}\backslash\{i\}} \frac{(k - |\mathcal{J}| - 1)!|\mathcal{J}|!}{d!} \left(c\left(\mathcal{J} \cup \{i\}\right) - c\left(\mathcal{J}\right)\right) \tag{24}$$

where $|\mathcal{J}|$ indicates the size of $\mathcal{J}$. In other words, $v_i$ is the incremental cost of including player $i$ in set $\mathcal{J}$ averaged over all sets $\mathcal{J} \subseteq \mathcal{D} \{i\}$.

This formula can be transposed to the field of global sensitivity analysis (Owen, 2014) if we consider the set of inputs of $\eta(\cdot)$ as the set of players $\mathcal{D}$. We then need to define a $c(\cdot)$ function such that for $\mathcal{J} \subseteq \mathcal{D}$, $c(\mathcal{J})$ measures the part of variance of $Y$ caused by the uncertainty of the inputs in $\mathcal{J}$. To this aim, we want a cost function that verifies $c(\emptyset) = 0$ and $c(\mathcal{D}) = \mathrm{Var}(Y)$.

Functions $\tilde{c}(\mathcal{J}) = \mathrm{Var}\left[\mathbb{E}\left[Y|\mathbf{X}_{\mathcal{J}}\right]\right]/\mathrm{Var}(Y)$ and $c(\mathcal{J}) = \mathbb{E}\left[\mathrm{Var}\left[Y|\mathbf{X}_{-\mathcal{J}}\right]\right]/\mathrm{Var}(Y)$ satisfy the two conditions above. Besides, Song et al. (2016) showed the Shapley values defined using cost functions $\tilde{c}(\mathcal{J})$ and $c(\mathcal{J})$ are equivalent.

However, for some reasons described at the end of the section 3.1 of the article Song et al. (2016), about the estimation of these two cost functions, it is better to define the Shapley effect of the $i$-th input, $Sh^i$, as the Shapley value obtained by applying the cost function $c$ instead of $\tilde{c}$. We denote in the sequel the Shapley effect by $Sh^i$ and a generic Shapley value by $v^i$. A valuable property of the Shapley effects defined in this way is that they are non-negative and they sum to one. Each one can therefore be interpreted as a measure of the part of the variance of $Y$ related to the $i$-th input of $\eta$.

## 3.2 Estimation of the Shapley effects

An issue with the Shapley value is its computational complexity as all possible subsets of the players need to be considered. Castro et al. (2009) proposed an estimation method based on an alternative definition of the Shapley value.

Indeed, the Shapley value can also be expressed in terms of all possible permutations of the players. Let us denote by $\Pi(\mathcal{D})$ the set of all possible permutations with player set $\mathcal{D}$. Given a permutation $\pi \in \Pi(\mathcal{D})$, define the set $P_i(\pi)$ as the players that precede player $i$ in $\pi$. Thus, the Shapley value can be rewritten in the following way :

$$v^i = \frac{1}{d!} \sum_{\pi \in \Pi(\mathcal{D})} \left[ c\left(P_i(\pi) \cup \{i\}\right) - c\left(P_i(\pi)\right) \right] \tag{25}$$

From this formula, Castro et al. (2009) proposed to estimate $v^i$ with $\widehat{v}^i$ by drawing randomly $m$ permutations in $\Pi(\mathcal{D})$ and thus we have :

$$\widehat{v}^i = \frac{1}{m} \sum_{l=1}^{m} \Delta_i c(\pi_l) \tag{26}$$

with $\Delta_i c(\pi_l) = c\left(P_i(\pi) \cup \{i\}\right) - c\left(P_i(\pi)\right)$ and $c(\cdot)$ the cost function.

Section 4 of Song et al. (2016) proposed some improvements on the Castro's algorithm by including the Monte-Carlo estimation $\widehat{c}$ of the cost function $c(\mathcal{J}) = \mathbb{E}\left[\mathrm{Var}\left[Y|\mathbf{X}_{-\mathcal{J}}\right]\right]/\mathrm{Var}(Y)$ to estimate the Shapley effects. The estimator writes:

$$\widehat{Sh}^i = \frac{1}{m} \sum_{l=1}^{m} \left[ \widehat{c}\left(P_i(\pi_l) \cup \{i\}\right) - \widehat{c}\left(P_i(\pi_l)\right) \right] \tag{27}$$

where $m$ refers to the number of permutations. Song et al. (2016) proposed the following two algorithms whose we just give the main features:

- The *exact permutation method* if $d$ is small, one does all possible permutations between the inputs (i.e. $m = d!$);

- The *random permutation method* which consists in randomly sampling $m$ permutations of the inputs in $\Pi(\mathcal{D})$.

For each iteration of this loop on the inputs' permutations, a conditional variance expectation must be computed. The cost $C$ of these algorithms is the following $C = N_v + m(d-1)N_o N_i$ with $N_v$ the sample size for the variance computation, $N_o$ the outer loop size for the expectation, $N_i$ the inner loop size for the conditional variance and $m$ the number of permutations according to the selected method.

Note that the full first-order Sobol' indices and the independent total Sobol' indices can be also estimated by applying these algorithms, each one during only one loop iteration.

Based on theoretical results, Song et al. (2016) recommends to fix parameters at the following values to obtain an accurate approximation of Shapley effects computationally affordable:

- The *exact permutation method*: $N_o$ as large as possible and $N_i = 3$;

- The *random permutation method*: $N_o = 1$, $N_i = 3$ and $m$ as large as possible.

The choice of $N_v$ is independent from these values and Iooss and Prieur (2017) have also illustrated the convergence of two numerical algorithms for estimating Shapley effects.

### 3.3 Confidence interval for the Shapley effects

In this part, we propose a methodology to compute confidence interval for the Shapley effects, which will allow us to quantify the Monte-Carlo error (sampling error).

**Exact permutation method: bootstrap**

Concerning this algorithm, we'll use the bias-corrected percentile method of the Bootstrap (Efron, 1981).

Let be $\widehat{\theta}(X_1, \ldots, X_n)$ an estimator of a unknown parameter $\theta$, function of n independent and identically distributed observations of law $\mathcal{F}$. In non-parametric Bootstrap, from a n-sample $(x_1, \ldots, x_n)$, we compute $\widehat{\theta}(x_1, \ldots, x_n)$. After, we draw with replacement a bootstrap sample $(x_1^*, \ldots, x_n^*)$ from the original sample $(x_1, \ldots, x_n)$ and compute $\theta^* = \widehat{\theta}(x_1^*, \ldots, x_n^*)$. We repeat this procedure B times and obtain B bootstrap replications $\theta_1^*, \ldots, \theta_B^*$ which allows the estimate of the following confidence interval of level $1 - \alpha$ for $\theta$:

$$\left[ \widehat{G}^{-1} \circ \Phi(2\widehat{z}_0 + z_{\alpha/2}) \quad ; \quad \widehat{G}^{-1} \circ \Phi(2\widehat{z}_0 - z_{\alpha/2}) \right] \tag{28}$$

where

- $\Phi$ is the cdf of a standard normal distribution;

- $z_{\alpha/2}$ percentile of level $\alpha/2$ of $\mathcal{N}(0,1)$;

- $\widehat{G}$ is the cdf of the bootstrap distribution for the estimator $\widehat{\theta}$;

- and $\widehat{z}_0 = \Phi^{-1} \circ \widehat{G}(\widehat{\theta})$ is a bias correction constant.

This confidence interval has been justified in Efron (1981) when there exists an increasing transformation $g(.)$ such that $g(\widehat{\theta}) - g(\theta) \sim \mathcal{N}(-z_0\sigma, \sigma^2)$ and $g(\widehat{\theta}^*) - g(\widehat{\theta}) \sim \mathcal{N}(-z_0\sigma, \sigma^2)$ for some constants $z_0 \in \mathbb{R}$ and $\sigma > 0$. In the sequel, we'll see in our examples that $g(.)$ can be considered as identity.

Thus, we need independent observations to obtain this interval but in our case as there is conditioning in the Shapley effects (more exactly in the cost function), it's not possible. To overcome this problem and estimate correctly the cdf $\widehat{G}(.)$, we make a bootstrap by bloc (on the $N_o$ blocs) in order to use independent observations and preserve the correlation within each one. This strategy allowed to develop the algorithm (1) in order to obtain the distribution of $\widehat{\widetilde{Sh}}^i$ to calculate the confidence interval for $Sh^i$.

---
**Algorithm 1:** Compute confidence intervals for $Sh^i$
---
**1** Generate a sample $\mathbf{x}^{(1)}$ of size $N_v$ from the random vector $\mathbf{X}$ ;

**2** Compute $\mathbf{y}^{(1)}$ from $\mathbf{x}^{(1)}$ to estimate $\mathrm{Var}(Y)$ ;

**3** Generate a sample $\mathbf{x}^{(2)}$ of size $m(d-1)N_oN_i$ from the different conditional laws necessary to estimate $\mathbb{E}\left[\mathrm{Var}\left[Y|\mathbf{X}_{-\mathcal{J}}\right]\right]$ ;

**4** Compute $\mathbf{y}^{(2)}$ from $\mathbf{x}^{(2)}$ ;

**5** Compute $\widehat{Sh}^i$ thanks to the equation (27) ;

**6 for** $b = 1, \dots, B$ **do**

**7** $\quad$ Sample with replacement a realization $\tilde{\mathbf{y}}^{(1)}$ of $\mathbf{y}^{(1)}$ to compute $\mathrm{Var}(Y)$ ;

**8** $\quad$ Sample by bloc with replacement a realization $\tilde{\mathbf{y}}^{(2)}$ of $\mathbf{y}^{(2)}$ ;

**9** $\quad$ Compute $\widehat{Sh}^i_b$ thanks to the equation (27). ;

**10 end**

**11** Compute confidence intervals for $Sh^i$ with 28.
---

It's suitable to remark that confidence intervals for the Shapley effects can also be calculated from the Central Limit Theorem (CLT) on the outer loop (Monte Carlo sample of size No) as Iooss and Prieur (2017) did it. But, it was necessary to establish a method based on the Bootstrap in order to be able to design in the sequel an algorithm allowing to distinguish correctly the metamodel and Monte-Carlo errors.

**Random permutation method: CLT**

For the random permutation method, we have two options to calculate confidence intervals.

● The first one is to use the CLT like Iooss and Prieur (2017). Indeed, in Castro et al. (2009) the CLT gives us:

$$\widehat{Sh}^i \xrightarrow[m\to\infty]{\mathcal{L}} \mathcal{N}\left(Sh^i, \frac{\sigma^2}{m}\right) \tag{29}$$

with $\sigma^2 = \dfrac{\mathrm{Var}\left(\Delta_i c(\pi_l)\right)}{\mathrm{Var}(Y)^2}$.

Thus, by estimating $\sigma$ by $\hat{\sigma}$ we have the following $1 - \alpha$ asymptotic confidence interval for the Shapley effects :

$$Sh^i \in \left[\widehat{Sh}^i + z_{\alpha/2}\frac{\hat{\sigma}}{\sqrt{m}} \quad ; \quad \widehat{Sh}^i - z_{\alpha/2}\frac{\hat{\sigma}}{\sqrt{m}}\right]$$

with $z_{\alpha/2}$ percentile of level $\alpha/2$ of $\mathcal{N}(0,1)$.

● The second one is we can estimate the confidence interval doing a bootstrap on the permutations. We describe in the algorithm (2) the procedure allowing to do that.

## 4 Examples in Gaussian framework: analytical results and relations between indices

In this section, we compare and interpret the analytic results of the studied indices for two different Gaussian models: an interactive and a linear model. We study the variation of the indices by varying the correlation between the input random variables.

---

**Algorithm 2:** Compute confidence intervals for $Sh^i$

---

**1** Generate a sample $\mathbf{x}^{(1)}$ of size $N_v$ from the random vector $\mathbf{X}$ ;

**2** Compute $\mathbf{y}^{(1)}$ from $\mathbf{x}^{(1)}$ to estimate $\mathrm{Var}(Y)$ ;

**3** Draw randomly $m$ permutations in $\Pi(\mathcal{D})$ ;

**4** Generate a sample $\mathbf{x}^{(2)}$ of size $m(d-1)N_oN_i$ from the different conditional laws necessary to estimate $\mathbb{E}\left[\mathrm{Var}\left[Y|\mathbf{X}_{-\mathcal{J}}\right]\right]$ ;

**5** Compute $\mathbf{y}^{(2)}$ from $\mathbf{x}^{(2)}$ ;

**6** Compute $\widehat{Sh}^i$ thanks to the equation (27) ;

**7** **for** $b = 1, \ldots, B$ **do**

**8**     Sample with replacement a realization $\tilde{\mathbf{y}}^{(1)}$ of $\mathbf{y}^{(1)}$ to compute $\mathrm{Var}(Y)$ ;

**9**     Sample with replacement $m$ permutations from the original sample and retrieve in $\mathbf{y}^{(2)}$ those corresponding to drawn bootstrap permutations ;

**10**     Compute $\widehat{Sh}_b^i$ thanks to the equation (27). ;

**11** **end**

**12** Compute confidence intervals for $Sh^i$ with 28.

---

## 4.1 Interactive model with two inputs

Let us consider a purely interactive model

$$Y = (\beta_1 X_1) \times (\beta_2 X_2) \tag{30}$$

with $\mathbf{X} \sim \mathcal{N}(0, \Sigma)$. We consider two cases: a model with independent variables and another with dependent variables. So we have the two following covariance matrices:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \qquad\qquad \Sigma = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

with $-1 \leq \rho \leq 1, \sigma_1 > 0, \sigma_2 > 0$.

From the definition of sensitivity indices, for $j = 1, 2$, we get for these models the results presented in Table 1.

In the independent model, the independent and full first-order Sobol indices are null because there is no dependence and the inputs have not marginal contribution. Thus, the independent and full total Sobol indices represent the variability in the model which is due to interactions only. These ones are each equal to the variance model, i.e. each input is fully responsible of the model uncertainty, due to its interaction with the other variable. In contrast, the Shapley effects award fairly the interaction effect to each input, which is more logical.

About the dependent model, $S_j^{ind} = 0, j = 1, 2$ are still null because the inputs have not uncorrelated marginal contribution. But now, $S_j^{full} \neq 0, j = 1, 2$ and represent marginal contribution due to the dependence. We see in these terms that the dependence effect $(\rho^2\beta_1^2\beta_2^2\sigma_1^2\sigma_2^2)$ is counted two times in comparison with the total variance. Concerning the independent and full total Sobol indices, the interaction effect $(\beta_1^2\beta_2^2\sigma_1^2\sigma_2^2)$ of these indices is still allocated half in Shapley effects. Besides, for the full total Sobol indices, each term is equal to the variance model whereas the interaction and dependence effects are fairly distributed for the Shapley effects which sum to the total variance.

| Independent model | Dependent model |
|---|---|
| **Model variance** | |
| $\sigma^2 = \text{Var(Y)} = \beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ | $\sigma^2 = \text{Var(Y)} = (1+\rho^2)\beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ |
| **Independent first-order Sobol'indices** | |
| $S_1^{ind} = 0$ | $S_1^{ind} = 0$ |
| $S_2^{ind} = 0$ | $S_2^{ind} = 0$ |
| **Independent total Sobol'indices** | |
| $\sigma^2 ST_1^{ind} = \beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ | $\sigma^2 ST_1^{ind} = (1-\rho^2)\beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ |
| $\sigma^2 ST_2^{ind} = \beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ | $\sigma^2 ST_2^{ind} = (1-\rho^2)\beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ |
| **Full first-order Sobol'indices** | |
| $S_1^{full} = 0$ | $\sigma^2 S_1^{full} = 2\rho^2\beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ |
| $S_2^{full} = 0$ | $\sigma^2 S_2^{full} = 2\rho^2\beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ |
| **Full total Sobol'indices** | |
| $\sigma^2 ST_1^{full} = \beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ | $\sigma^2 ST_1^{full} = (1+\rho^2)\beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ |
| $\sigma^2 ST_2^{full} = \beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ | $\sigma^2 ST_1^{full} = (1+\rho^2)\beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ |
| **Shapley effects** | |
| $\sigma^2 Sh^1 = \frac{1}{2}\beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ | $\sigma^2 Sh^1 = \frac{1}{2}(1+\rho^2)\beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ |
| $\sigma^2 Sh^2 = \frac{1}{2}\beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ | $\sigma^2 Sh^2 = \frac{1}{2}(1+\rho^2)\beta_1^2\beta_2^2\sigma_1^2\sigma_2^2$ |

Table 1: Sensitivity indices of independent and dependent Gaussian models

This example supports the idea mentioned in Iooss and Prieur (2017) whereby *a full Sobol index of an input comprises the effect of another input on which it is dependent.* We can add that the model is independent or not, the phenomenon is similar for the interaction effect about the independent and full total Sobol indices of an input, i.e. these indices comprise the effect of another input on which the input is interacting.

In their article, Iooss and Prieur (2017) tell which goals of the SA settings defined in Saltelli and Tarantola (2002) and Saltelli et al. (2004) the four Sobol indices as well as the Shapley effects apply.
According to them, a combined interpretation of the four Sobol indices would just allow to do the FP (Factor prioritization) setting. But we can add that these indices allow also to do the FF (Factor Fixing) setting only if a factor has both indices $ST_i^{ind}$ and $ST_i^{full}$ which are null. Indeed, if $ST_i^{ind} = \frac{\mathbb{E}[\text{Var}(Y|\mathbf{X}_{\sim i})]}{\text{Var}(Y)} = 0$ and $ST_i^{full} = \frac{\mathbb{E}[\text{Var}(Y|(\mathbf{X}_{\sim i}|X_i))]}{\text{Var}(Y)} = 0$ and as the variance is always a positive function, that implies $\text{Var}(Y|\mathbf{X}_{\sim i}) = 0$ and $\text{Var}(Y|(\mathbf{X}_{\sim i}|X_i)) = 0$. Thus, $Y$ can be expressed only as a function of $\mathbf{X}_{\sim i}$ or $\mathbf{X}_{\sim i}|X_i$,i.e. $X_i$ has not impact on $Y$ taking account the dependence or not.

About the Shapley effects, they would allow to do the VC (Variance Cutting) setting as the sum is equal to $\text{Var}(Y)$ and the FF setting. Sure enough, if $Sh^i = 0$, then we have $\forall \mathcal{J} \subseteq \mathcal{D}\backslash\{i\}, \text{Var}\left[Y|\mathbf{X}_{-(\mathcal{J}\cup\{i\})}\right] = \text{Var}[Y|\mathbf{X}_{-\mathcal{J}}]$ and so express $Y$ as a function of $\mathbf{X}_{-(\mathcal{J}\cup\{i\})}$ equates to express $Y$ as a function of $\mathbf{X}_{-\mathcal{J}}$. Hence, $X_i$ is not an influential input in the model and can be fixed. The FP setting is not achieved according to them because of the fair distribution of the interaction and dependence effects in the indice. However, this share allocation makes the Shapley effects easier to interpret than the Sobol' indices and might be a great alternative to the four Sobol' indices. Thus, in the sequel, we'll compare the Sobol indices' and the Shapley effects on an basic examples to see if they make correctly the factor prioritization.

## 4.2 Linear model with three inputs

Let us consider

$$Y = \beta_0 + \boldsymbol{\beta}^{\mathsf{T}}\mathbf{X} \tag{31}$$

with the constants $\beta_0 \in \mathbb{R}$, $\boldsymbol{\beta} \in \mathbb{R}^3$ and $\mathbf{X} \sim \mathcal{N}(0, \Sigma)$ with the following covariance matrix :

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \alpha\sigma_1\sigma_2 & \rho\sigma_1\sigma_3 \\ \alpha\sigma_1\sigma_2 & \sigma_2^2 & \gamma\sigma_2\sigma_3 \\ \rho\sigma_1\sigma_3 & \gamma\sigma_2\sigma_3 & \sigma_3^2 \end{pmatrix}, -1 \le \alpha, \rho, \gamma \le 1, \sigma_1 > 0, \sigma_2 > 0, \sigma_3 > 0.$$

We obtained the following analytical results.

$$\sigma^2 = Var(Y) = \beta_1^2\sigma_1^2 + \beta_2^2\sigma_2^2 + \beta_3^2\sigma_3^2 + 2\gamma\beta_2\beta_3\sigma_2\sigma_3 + 2\beta_1\sigma_1(\alpha\beta_2\sigma_2 + \rho\beta_3\sigma_3)$$

- For $j = 1, 2, 3$, from the definition of independent Sobol indices, we have:

$$\sigma^2 S_1^{ind} = \sigma^2 ST_1^{ind} = \frac{\beta_1^2\sigma_1^2\left(-1 + \alpha^2 + \gamma^2 + \rho^2 - 2\alpha\gamma\rho\right)}{\gamma^2 - 1}$$

$$\sigma^2 S_2^{ind} = \sigma^2 ST_2^{ind} = \frac{\beta_2^2\sigma_2^2\left(-1 + \alpha^2 + \gamma^2 + \rho^2 - 2\alpha\gamma\rho\right)}{\rho^2 - 1}$$

$$\sigma^2 S_3^{ind} = \sigma^2 ST_3^{ind} = \frac{\beta_3^2\sigma_3^2\left(-1 + \alpha^2 + \gamma^2 + \rho^2 - 2\alpha\gamma\rho\right)}{\alpha^2 - 1}$$

- For $j = 1, 2, 3$, from the definition of full Sobol indices, we have:

$$\sigma^2 S_1^{full} = \sigma^2 ST_1^{full} = (\beta_1\sigma_1 + \alpha\beta_2\sigma_2 + \rho\beta_3\sigma_3)^2$$

$$\sigma^2 S_2^{full} = \sigma^2 ST_2^{full} = (\alpha\beta_1\sigma_1 + \beta_2\sigma_2 + \gamma\beta_3\sigma_3)^2$$

$$\sigma^2 S_3^{full} = \sigma^2 ST_3^{full} = (\rho\beta_1\sigma_1 + \gamma\beta_2\sigma_2 + \beta_3\sigma_3)^2$$

In both cases, full and independent Sobol indices, the first order index is equal to the total order index because the model is linear, i.e., there is no interaction between the inputs.

- For $j = 1, 2, 3$, **in this example** we obtain the following decomposition for the Shapley effects :

$$Sh^1 = \frac{1}{3}\left(S_1^{full} + \frac{1}{2}ST_{1|2} + \frac{1}{2}ST_{1|3} + ST_1^{ind}\right)$$

$$Sh^2 = \frac{1}{3}\left(S_2^{full} + \frac{1}{2}ST_{2|1} + \frac{1}{2}ST_{2|3} + ST_2^{ind}\right)$$

$$Sh^3 = \frac{1}{3}\left(S_3^{full} + \frac{1}{2}ST_{3|1} + \frac{1}{2}ST_{3|2} + ST_3^{ind}\right)$$

So, for the **linear Gaussian model** we found a relation between the Shapley effects and the sensitivity indices obtained with the RT method. For more details about the calculation of Shapley effects, we refer the readers to the Appendix A.1.

About the results, as the formula is similar regardless the input, we analyse it with the first input. We observe that the Shapley effect $Sh^1$ is in some way the average of all possible contributions of the input $X_1$ in the model. Indeed, $S_1^{full}$ represents the full marginal contribution of $X_1$. Then, we have the total contributions of $X_1$ without its correlative contribution with each element of

the set $\mathcal{D} = \{1,2,3\}\backslash\{1\} = \{2,3\}$. Sure enough, $ST_{1|2}$ is the total contribution of $X_1$ without its correlative contribution with $X_2$, i.e. ones just look at the total effect with its dependence with $X_3$ ; $ST_{1|3}$ is the total contribution of $X_1$ without its correlative contribution with $X_3$, i.e. ones just look at the total effect with its dependence with $X_2$ and finally the uncorrelated total contribution of $X_1$ via $ST_1^{ind} = ST_{1|2,3}$. As in $\{2,3\}$, there are two elements of size one, we find the coefficients $1/2$ before $ST_{1|2}$ and $ST_{1|3}$ and $1$ for $ST_1^{ind}$. We then find the fair allocation of the Shapley effects.

## Particular cases

Now, we'll consider several particular cases of correlation in order to compare the prioritization obtained with the Sobol' indices and the Shapley effects. We'll take in the following examples $\beta_0 = 0$ ; $\beta_1 = \beta_2 = \beta_3 = 1$ and $\sigma_1 = 0.2, \sigma_2 = 0.6, \sigma_3 = 1$. By making this choice, we define implicitly the most influential variables and we want to observe how the correlation affects the indices. Besides, for each considered case, we verify that the covariance matrix is positive definite.

| | $\alpha = \rho = \gamma = 0$ | | | $\alpha = \rho = \gamma = 0.5$ | | | $\alpha = \rho = 0.75, \gamma = 0.15$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $X_1$ | $X_2$ | $X_3$ | $X_1$ | $X_2$ | $X_3$ | $X_1$ | $X_2$ | $X_3$ |
| $S_i^{ind}$ | 0.0286 | 0.2571 | 0.7143 | 0.0115 | 0.1034 | 0.2874 | 0.0004 | 0.0085 | 0.0236 |
| $ST_i^{ind}$ | 0.0286 | 0.2571 | 0.7143 | 0.0115 | 0.1034 | 0.2874 | 0.0004 | 0.0085 | 0.0236 |
| $S_i^{full}$ | 0.0286 | 0.2571 | 0.7143 | 0.4310 | 0.6207 | 0.8448 | 0.9515 | 0.3932 | 0.7464 |
| $ST_i^{full}$ | 0.0286 | 0.2571 | 0.7143 | 0.4310 | 0.6207 | 0.8448 | 0.9515 | 0.3932 | 0.7464 |
| $Sh_i$ | 0.0286 | 0.2571 | 0.7143 | 0.1715 | 0.3123 | 0.5163 | 0.4553 | 0.1803 | 0.3644 |

Table 2: Sensitivity indices of linear model with different configurations of correlation

As part of the independent linear model, the Shapley effects are equal to the Sobol' indices as proved in Iooss and Prieur (2017) and thus, all the indices carry out to the same ranking of the inputs.

In the second configuration with the symmetric case, we remark a decrease of the independent Sobol indices and an increase of the full Sobol indices with respect to the independent model ($\alpha = \rho = \gamma = 0$). As regards of the Shapley effects, it reduces for the third input, raises slightly for the second input and significantly for the first input. All these changes are due to the mutualization of uncertainties within the inputs because of the correlation but the individual contributions of the inputs are still well captured for all the indices. Indeed, in spite of the correlation, all the indices indicate the same ranking for the inputs.

In this last configuration, we have strongly correlated a non-influential variable ($X_1$ has a low variance) in the model with two very influential variables. The independent Sobol' indices give us as ranking: $X_3, X_2, X_1$. However, as the values of these indices are close to zero, we can suppose they are not significant and implicitly the ranking too. We obtain with the full indices the following ranking $X_1, X_3, X_2$. $X_1$ is supposed to be a non-influential variable and turns out to explain 95% of the model variance. Which is logical because being highly correlated with $X_2$ and $X_3$, $X_1$ has a strong impact on these variables. Then, $X_2$ and $X_3$ are correlated in the same way with $X_1$ and weakly between them. Regardless of the correlations, $X_3$ is more influential than $X_2$ in the model, hence this second position taking account the correlation. Lastly, we obtain the same ranking as the full Sobol' indices with the Shapley effects. FP (Factors Prioritization) setting aims to find which factors would allow to have the largest expected reduction in the variance of the model output. Thus, if we follow the previous ranking, we should reduce the

uncertainty on the first input. But we'll make several tests by reducing the uncertainty of 20% one by one on each input and we get:

| Setting $\alpha = \rho = 0.75, \gamma = 0.15$ | Model variance |
|---|---|
| $\sigma_1 = 0.2, \sigma_2 = 0.6, \sigma_3 = 1$ | 2.06 |
| $\sigma_1 = 0.16, \sigma_2 = 0.6, \sigma_3 = 1$ | 1.95 |
| $\sigma_1 = 0.2, \sigma_2 = 0.48, \sigma_3 = 1$ | 1.86 |
| $\sigma_1 = 0.2, \sigma_2 = 0.6, \sigma_3 = 0.8$ | 1.60 |

Table 3: Model variance by reducing the uncertainty on each input one by one

It is clearly observed that the largest expected reduction in the variance is obtained with the third input. These results conflict the obtained ranking with the full Sobol indices and the Shapley effects. Indeed, $X_1$ is an influential input only because of the strong correlation with $X_2$ and $X_3$, and these indices capture this trend. However, without this correlation $X_1$ is non-influential input and the independent Sobol indices are supposed to highlight **meaningfully** that these are the inputs $X_2$ and $X_3$ which are the most influential without take account the correlation. Nevertheless, these indices struggle to emphasize the uncorrelated marginal contributions of these inputs due to the small values we obtain.

Thus, on this basic example we can see that the combined interpretation of the four Sobol indices as well as the Shapley effects doesn't allow to answer correctly to the purpose of the Factor Prioritization (FP) setting, i.e. on which inputs the reduction of uncertainty leads to the largest reduction of the output uncertainty. We can make a factor prioritization with these indices but not for the goal defined at the outset.

# 5    Numerical studies

Optimal values for the parameters of the exact and random permutation methods were given by Song et al. (2016). Using a toy example, we empirically study how the algorithm settings can influence the estimation of the indices. We compare the accuracy of the estimations of the Sobol' indices from the Shapley algorithm and the RT method.

## 5.1    Parametrization of the Shapley algorithms

As defined in Section 3.2, three parameters of the Shapley algorithm govern the estimation accuracy: $N_v$, $N_o$ and $N_i$. The first one, is the sample-size for the output variance estimation of $Y$. The second, is the number of outer loop for the sample-size of the expectation and the third one is the number of inner loop which controls the sample-size for the variance estimation of each conditioned distribution.

Theses variances are estimated through Monte-Carlo procedures. The output variance $\text{Var}[Y]$ is computed from a sample $\{Y_j = \eta(\mathbf{X}^{(j)})\}_{j=1\ldots,N_v}$. Because $N_v$ is a small proportion of the overall cost $C = N_v + m(d-1)N_o N_i$, especially when the $d$ is large, we can select $N_v$ as large as possible in order to reach the smallest possible estimation error of $\text{Var}[Y]$. However, it is more difficult to chose $N_o$ and $N_i$ to estimate the conditioned variances. These choices also depend on the used algorithm: exact or random permutations.

Therefore, we empirically show the influence of $N_o$ and $N_i$ on the estimation error and the coverage probability. The Probability Of Coverage (POC) is defined as the probability to have

the true indice value inside the confidence intervals of the estimation. We consider the three dimensional linear Gaussian model of Section 4.2 as a toy example with independent inputs, $\beta_1 = \beta_2 = \beta_3 = 1$, $\sigma_1 = \sigma_2 = 1$ and $\sigma_3 = 2$. The POC is estimated with 100 independent algorithm runs and for a 90 % confidence interval. When the bootstrap procedure is considered, the confidence intervals are estimated with 500 bootstrap sampling. We also set a large value of $N_v = 10000$ for all the experiments.

First experiments aim to show the influence of $N_o$ on the estimation accuracy and the POC for the exact permutation algorithm. The Figure 1 shows the variation of the POC (solid lines) and the absolute error (dashed lines), averaged over the three indices, in function of the product $N_o N_i m$, where only $N_o$ is varying and for three values of $N_i$ at 3, 9 and 18. Because the errors are computed for 100 independent runs, we show in color areas the 95% quantiles.
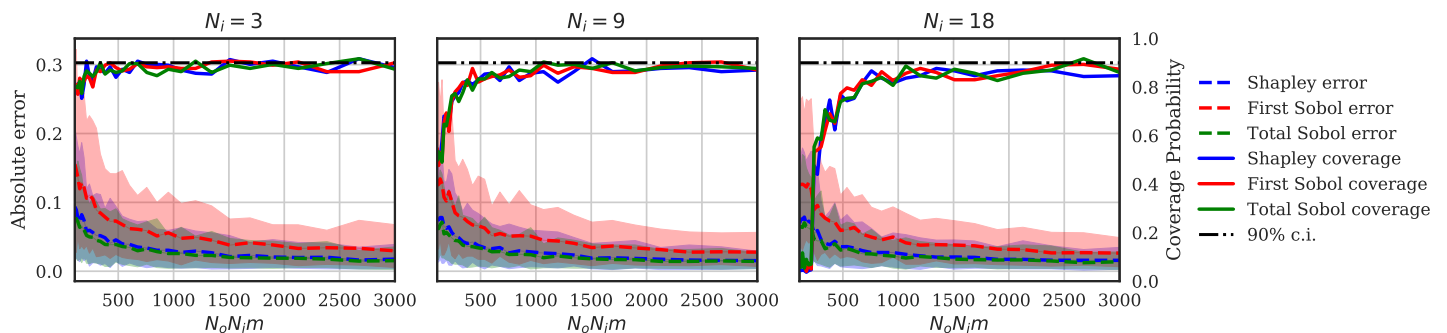


Figure 1: Variation of the absolute error and the POC with $N_o$ for three values of $N_i = 3, 9, 18$ for the exact permutation algorithm ($m = d! = 6$).

We observe that the estimation error is similar for the three different values of $N_i$ and decrease to 0 at the same rate. The true difference is for the POC which tends, at different rates, to the true probability: 90 %. For a same computation budget $N_o N_i m$, the smaller the value of $N_i$ and the larger the value of $N_o$. Thus, these results show that, in order to have a correct confidence interval it is more important to have a large value of $N_o$ instead of $N_i$. Indeed, exploring multiple conditioned variances with a lower precision (large $N_o$ and low $N_i$) is more important than having less conditioned variances with a good precision (low $N_o$ and large $N_i$).

The Figure 2 is similar to Figure 1 but for the random permutation algorithm and by fixing $N_o = 1$ and by varying the number of permutations.
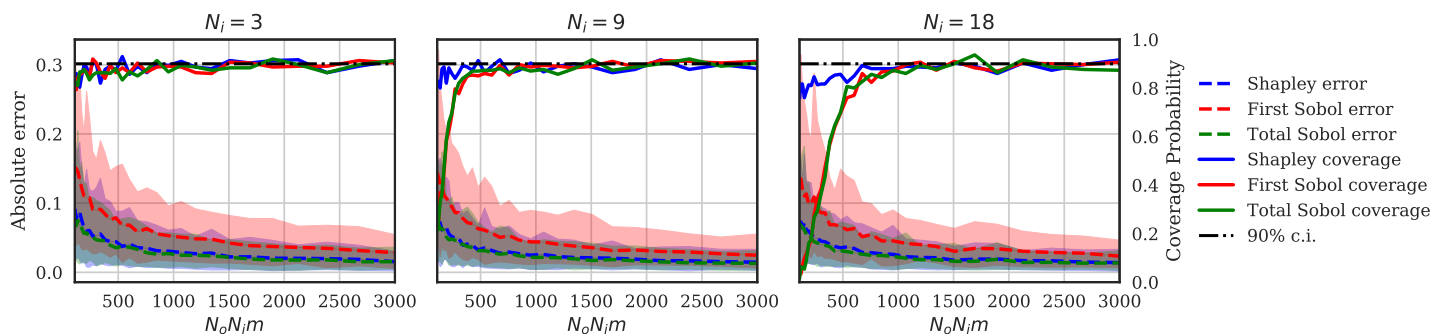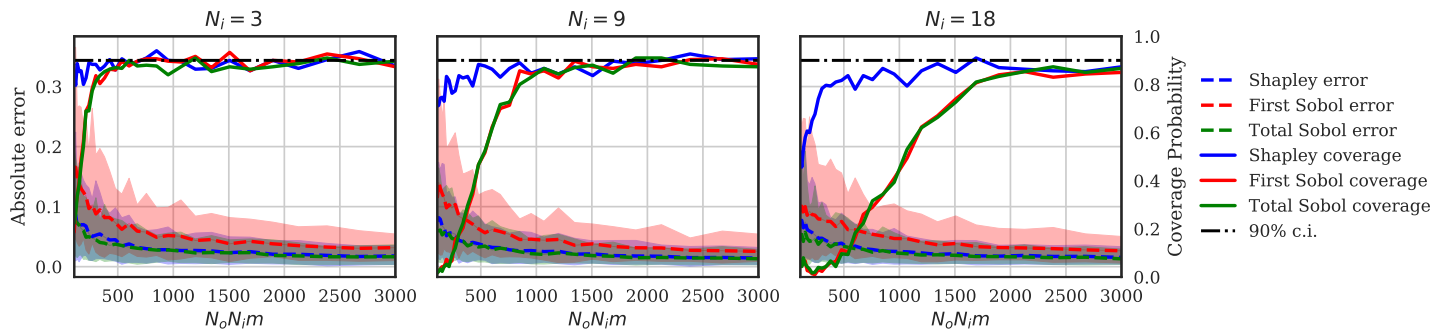


Figure 2: Variation of the absolute error and the POC with $m$ for three values of $N_i = 3, 9, 18$ and $N_o = 1$ for the random permutation algorithm.

As for the exact permutation algorithm, we can see that the estimation errors are similar for the three values of $N_i$ and the difference is shown for the POC. We observe that the lower $N_i$ and the faster the POC converges to the true probability. Indeed, for a same computational cost, the lower $N_i$ and the larger the number of permutations $m$ can be.

To show the influence of $N_o$ with the random permutation algorithm, the Figure 3 is the same as Figure 2 but with $N_o = 3$. We observe that the convergence rates of the POC are slower than the ones for $N_o = 1$. Thus, it shows that having a lower value of $N_o$ and a large value of $m$ is more important to have consistent confidence intervals.



Figure 3: Variation of the absolute error and the POC with $m$ for three values of $N_i = 3, 9, 18$ and $N_o = 3$ for the random permutation algorithm.

From these experiments, we can conclude that the parametrization does not significantly influence the estimation error but has a strong influence on the POC. Moreover, these experiments were established on different toy examples (Ishigami model defined in Section 7.2 and interactive model) and the same conclusion arises. Therefore, in order to have consistent confidence intervals, we can suggest:

- for the exact algorithm to consider $N_i = 3$ and to take $N_o$ as large as possible,

- for the random permutation algorithm to consider $N_i = 3$, $N_o = 1$ and take $m$ as large as possible.

This conclusion confirms the proposed parametrization of Song et al. (2016) explained in 3.2 and the suggestion analyzed in Iooss and Prieur (2017).

## 5.2 Minor bias observed

At the start of this section, we chose to establish these experiments for independent random variables. This choice was justified by unexpected results obtained for correlated variables. The Figure 4 illustrates the same experiment as Figure 1 but with a correlation of $\gamma = 0.9$ between $X_2$ and $X_3$. We observed that the POC of the total Sobol' indice starts to tend to the true probability (at 90%) before slowly decreasing. Thus, it seems that the confidence intervals are underestimated or the indice estimation is biased.

To verify this observation, Figure 5 shows the estimation of the total Sobol' indice for $N_v = 20000$, $N_o = 10000$, $N_i = 3$ with the histogram from the bootstrap sampling in blue, the estimated indice $ST_i$ in red line and the true indice in green line. It is clear that the true value for $X_2$ and $X_3$ is outside of estimated distribution. This explains why the coverage probability is decreasing in 4. Moreover, this phenomenon only happens to the indices of $X_2$ and $X_3$, which
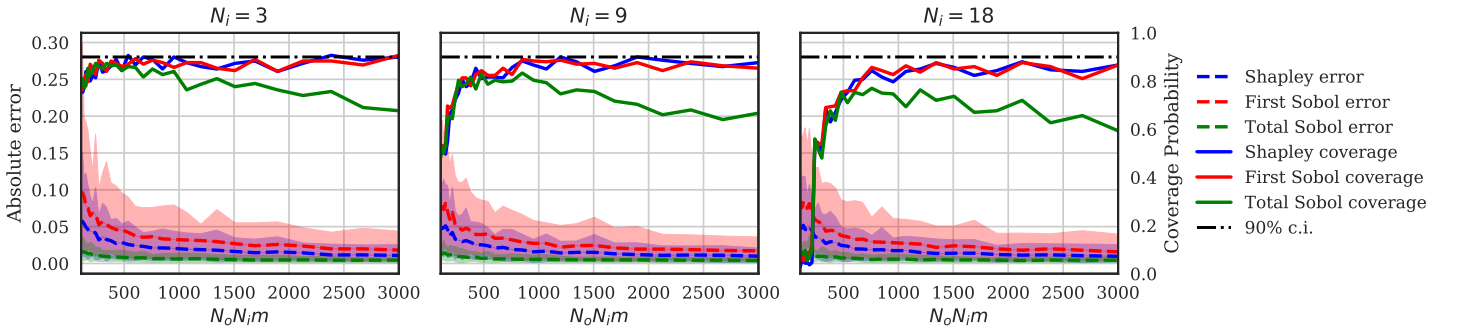
16

Figure 4: Variation of the absolute error and the POC with $N_o$ for three values of $N_i = 3, 9, 18$ for the exact permutation algorithm and a correlation $\gamma = 0.9$ between $X_2$ and $X_3$.

are correlated and it seems that this bias increases with the correlation strength for this example. Therefore, the reasons of this slight bias should be investigated in future works
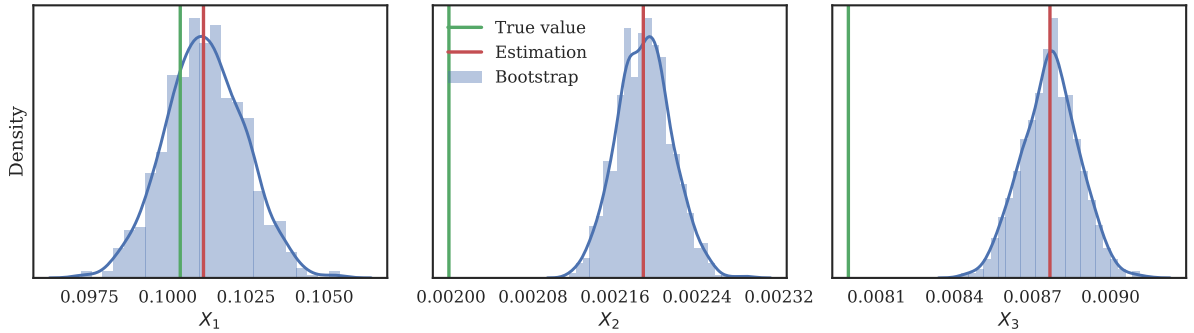


Figure 5: Estimated bootstrap estimations of the total Sobol' indices from the exact Shapley algorithm with a correlation of 0.99 between $X_2$ and $X_3$.

## 5.3 Comparing Sobol' index estimation using Shapley algorithm and RT method

An interesting result of the Shapley algorithm, is that it gives the full first-order Sobol' indices and the independent total Sobol' indices in addition to the Shapley effects. We compare the estimation accuracy of the Sobol' indices obtained from the Shapley algorithm and the ones from the RT method. We consider the same example as Section 5.1 but with dependent random variables. In this section, only the pair $X_2$-$X_3$ is correlated with parameter $\gamma$.

A first experiment aims to validate the confidence intervals estimated from the bootstrap sampling of the RT method by doing the same experiments as in Section 5.1 by increasing the sample-size $N$. The Figure 6 shows the absolute error and the POC with the computational budget $(4 \times N \times d)$ for the full first-order Sobol' indices and the independent total Sobol' indices for $\gamma = 0.5$. As we can see the error tends to zero and the POC converges quickly to the true probability. Thus, we can see that the confidence intervals correctly catch the Monte-Carlo error.

We recall from Section 3 that the full first-order Sobol' indices are equivalent to the classical

Figure 6: Variation of the absolute error and the POC with the computational budget for the RT method.

first-order Sobol' indices and the independent total indices are the classical total indices. The Figure 7 shows the estimated indices with $\gamma = 0.5$ from the Shapley algorithm and the RT method for similar computational costs. We observe that both algorithms seem to correctly estimate the Sobol' indices for a low computational cost. However, in this example, the estimation errors from the RT method is much larger than the ones from the Shapley algorithm.
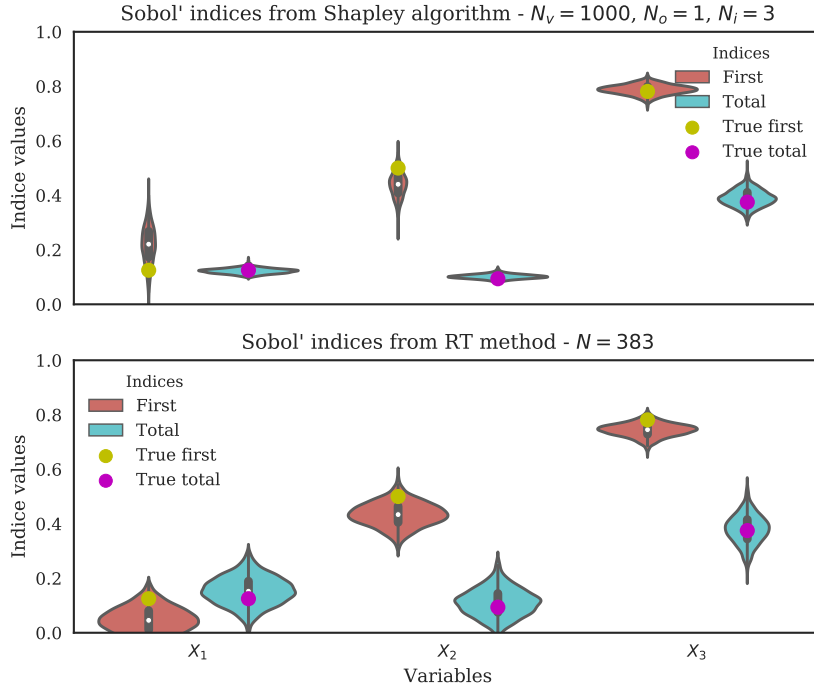


Figure 7: Sobol' indices estimation from the exact permutation method of the Shapley algorithm (top) and the RT method (bottom) using the Janon estimator for similar number of evaluation: $N_v + N_o N_i m(d-1) = 4Nd = 4600$.

We recall in Section 2.3 that RT method used the Janon estimator from Janon et al. (2014). The accuracy of the Sobol' estimator depends on the values of the target indices and the Janon estimator is less accurate for low value indices. Changing with another estimator, such as the one from Mara et al. (2015), can lead to another estimation variance as shown in Figure 8. We

observed that the estimation errors from the RT method depends of the used estimator and this error is lower using estimator from Figure 8 than the one from Figure 7.
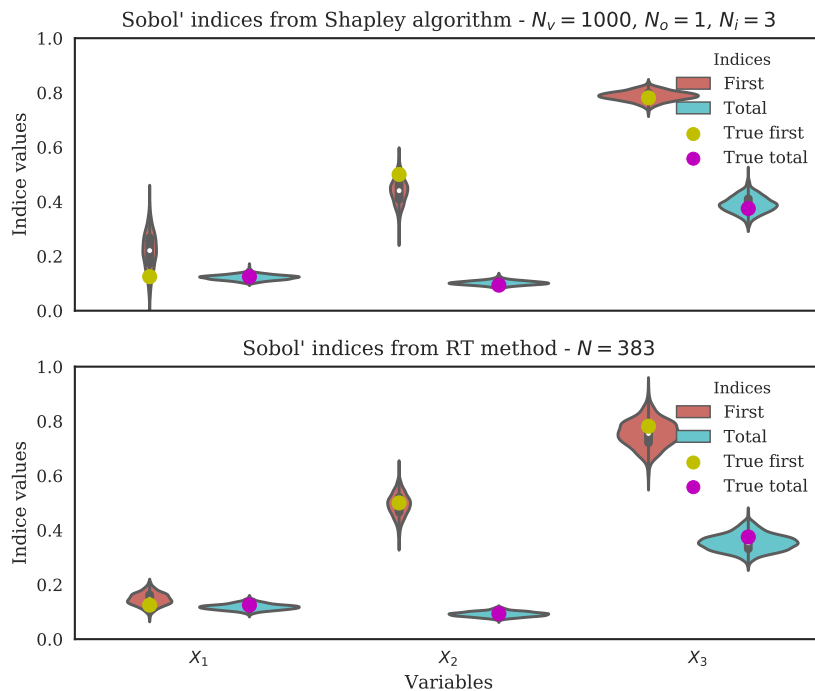


Figure 8: Sobol' indices estimation from the exact permutation method of the Shapley algorithm (top) and the RT method (bottom) using the estimator from Mara et al. (2015) for similar number of evaluation: $N_v + N_o N_i m(d-1) = 4Nd = 4600$.

The Figure 9 shows the Sobol' indices for the exact Shapley algorithm and the RT method in function of the correlation $\gamma$ between $X_2$ and $X_3$. The lines shows the true values of the indices and the areas are the 95% confidence intervals of the indices.

This experiment shows that the estimation of the Sobol' indices from the Shapley algorithm gives satisfying estimations of the first full and total ind Sobol' indices. Note that the error of estimation is similar for both the exact or random permutation algorithm if we consider the same computational budget.

## 6 Kriging metamodel with inclusion of errors

Shapley effects are a suitable tool for performing global sensitivity analysis. However, their estimates require an important number of simulations of the costly function $\eta(\mathbf{x})$ and often cannot be processed under reasonable time constraint. To handle this problem, we use $\tilde{\eta}(\mathbf{x})$ an approximating function of the numerical model under study $\eta(\mathbf{x})$ (Fang et al., 2005). Its main advantage is obviulsy to be much faster-to-calculate than the original one. In addition, if one uses a kriging method (Sacks et al., 1989) to build this $\tilde{\eta}(\mathbf{x})$ surrogate model, a quantification of the approximation uncertainty can be easily produced. The Shapley effects can then be calculated using the metamodel $\tilde{\eta}(\mathbf{x})$ instead of $\eta(\mathbf{x})$ with a control on the estimation error.
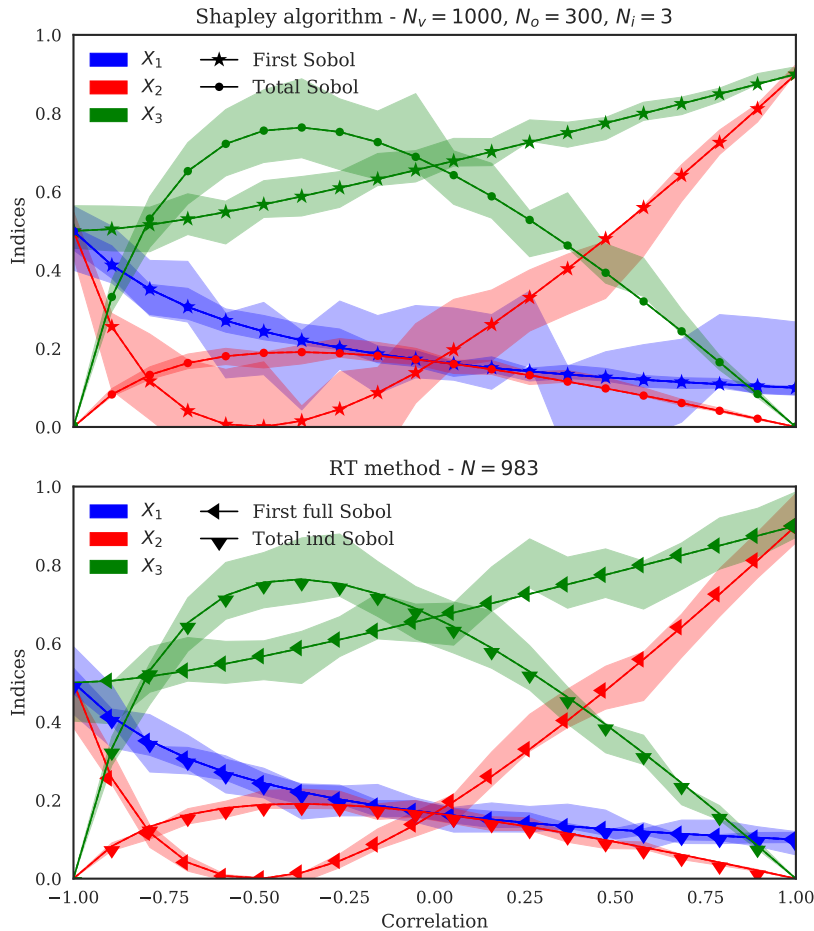
Figure 9: Sobol' indices estimations from the exact permutation method of the Shapley algorithm (top) and the RT method (bottom) in fonction of $\gamma$.

We present in this section a methodology for estimating the Shapley effects through a kriging surrogate model taking into account both the Monte Carlo error and the surrogate model error.

## 6.1 Introduction to the kriging model

*Kriging*, also called *metamodeling by Gaussian process*, is a method consisting in the use of an emulator of a costly computer code for which the interpolated values are modeled by a Gaussian process. More precisely, it is based on the assumption that the $\eta(x)$ function is the realization of a Gaussian random process. The data is then used to infer characteristics of this process, allowing a joint modelization of the code itself and the uncertainty about the interpolation on the domain. In general, one assumes a particular parametric model for the mean function of the process and for its covariance. The parameters of these two functions are called "hyperparameters" and are estimated using the data. The Gaussian hypothesis then provides an explicit formula for the law of the process conditionaly to the value taken by $\eta$ on a design of experiments $\mathbf{D}$.

Thus, we consider that our expensive function $\eta(x)$ can be modeled by a Gaussian process $H(x)$ which's mean and variance are such that $\mathbb{E}[H(x)] = \mathbf{f}'(x)\boldsymbol{\beta}$ and $Cov(H(x), H(\tilde{x})) = \sigma^2 r(x, \tilde{x})$, where $r(x, \tilde{x})$ is the covariance kernel (or the correlation function) of the process.

Then, $\eta(x)$ can be easily approximated by the conditional Gaussian process $H_n(x)$ having the predictive distribution $[H(x)|H(\mathbf{D}) = \boldsymbol{\eta}^n, \sigma^2]$ where $\boldsymbol{\eta}^n$ are the known values of $\eta(x)$ at points in the experimental design set $\mathbf{D} = \{x^1, \ldots, x^n\}$ and $\sigma^2$ is the variance parameter. Therefore, we have

$$H_n(x) \sim GP\left(m_n(x), s_n^2(x, \tilde{x})\right), \tag{32}$$

where the mean $m_n(x)$ is given by

$$m_n(x) = \mathbf{f}'(x)\widehat{\boldsymbol{\beta}} + \mathbf{r}'(x)\mathbf{R}^{-1}\left(\boldsymbol{\eta}^n - \mathbf{F}\widehat{\boldsymbol{\beta}}\right),$$

where $\mathbf{R} = [r(x_i, x_j)]_{i,j=1,\ldots,n}$, $\mathbf{r}'(x) = [r(x, x_i)]_{i=1,\ldots,n}$, $\mathbf{F} = [\mathbf{f}'(x_i)]_{i=1,\ldots,n}$, and

$$\widehat{\boldsymbol{\beta}} = \left(\mathbf{F}'\mathbf{R}^{-1}\mathbf{F}\right)^{-1}\mathbf{F}'\mathbf{R}^{-1}\boldsymbol{\eta}^n.$$

The variance $s_n^2(x, \tilde{x})$ is given by

$$s_n^2(x, \tilde{x}) = \sigma^2\left(1 - \begin{pmatrix} \mathbf{f}'(x) & \mathbf{r}'(x) \end{pmatrix}\begin{pmatrix} 0 & \mathbf{F}' \\ \mathbf{F} & \mathbf{R} \end{pmatrix}^{-1}\begin{pmatrix} \mathbf{f}(\tilde{x}) \\ \mathbf{r}(\tilde{x}) \end{pmatrix}\right)$$

The variance parameter $\sigma^2$ can be estimated with a restricted maximum likelihood method.

## 6.2 Kriging based Shapley effects and estimation

Inspired by the idea used in Le Gratiet et al. (2014) for the Sobol indices, we substitute the true function $\eta(\mathbf{x})$ with $H_n(\mathbf{x})$ in (25) which leads to

$$Sh_n^i = \frac{1}{d!}\sum_{\pi \in \Pi(\mathcal{D})}\left[c_n\left(P_i(\pi) \cup \{i\}\right) - c_n\left(P_i(\pi)\right)\right] \tag{33}$$

where the exact function $Y = \eta(\mathbf{X})$ is replaced by the Gaussian process $H_n(\mathbf{X})$ in the cost function such as $c_n(\mathcal{J}) = \mathbb{E}\left[\text{Var}\left[H_n(\mathbf{X})|\mathbf{X}_{-\mathcal{J}}\right]\right]$.

Therefore, if we denote by $(\Omega_H, \mathcal{F}_H, \mathbb{P}_H)$ the probability space where the Gaussian process $H(x)$ lies, then the index $Sh_n^i$ lies in $(\Omega_H, \mathcal{H}, \mathbb{P}_H)$ (it is hence random).

Then, for estimating $Sh_n^i$, we use the same estimator (27) developed by Song et al. (2016) in which we remplace $Y$ by the Gaussian process $H_n(\mathbf{X})$ in the cost function to obtain :

$$\widehat{Sh}_n^i = \frac{1}{m}\sum_{l=1}^m\left[\widehat{c}_n\left(P_i(\pi_l) \cup \{i\}\right) - \widehat{c}_n\left(P_i(\pi_l)\right)\right] \tag{34}$$

where $\widehat{c}_n$ is the Monte-Carlo estimator of $c_n$.

## 6.3 Estimation of errors : Monte Carlo and surrogate model

The estimator (34) above integrates two sources of uncertainty : the first one is related to the metamodel approximation, and the second one is related to the Monte Carlo integration. So, in this part, we quantify both by decomposing the variance of $\widehat{Sh}_n^i$ as follows :

$$\text{Var}(\widehat{Sh}_n^i) = \text{Var}_H\left(\mathbb{E}_X\left[\widehat{Sh}_n^i|H_n(x)\right]\right) + \text{Var}_X\left(\mathbb{E}_H\left[\widehat{Sh}_n^i|(\mathbf{X}_{\kappa_l})_{l=1,\ldots,B}\right]\right)$$

where $\text{Var}_H \left( \mathbb{E}_X \left[ \widehat{Sh}_n^i | H_n(x) \right] \right)$ is the contribution of the metamodel on the variability of $\widehat{Sh}_n^i$ and $\text{Var}_X \left( \mathbb{E}_H \left[ \widehat{Sh}_n^i | (\mathbf{X}_{\kappa_l})_{l=1,\ldots,B} \right] \right)$ is that of the Monte Carlo integration.

In section 4 of the article Le Gratiet et al. (2014), they proposed the algorithm (3) we adapted here to estimate each of these contributions.

---

**Algorithm 3:** Evaluation of the distribution of $\widehat{Sh}_{\kappa,n}^i$

---

**1** Build $H_n(x)$ from the $n$ observations $\boldsymbol{\eta}^n$ of $\eta(x)$ at points in $\mathbf{D}$ ;
**2** Generate a sample $\mathbf{x}^{(1)}$ of size $N_v$ from the random vector $\mathbf{X}$ ;
**3** Generate a sample $\mathbf{x}^{(2)}$ of size $m(d-1)N_oN_i$ from the different conditional laws necessary to estimate $\mathbb{E}\left[\text{Var}\left[Y|\mathbf{X}_{-\mathcal{J}}\right]\right]$ ;
**4** Set $N_H$ as the number of samples for $H_n(x)$ and $B$ the number of bootstrap samples for evaluating the uncertainty due to Monte Carlo integration ;
**5** **for** $k = 1, \ldots, N_H$ **do**
**6** $\quad$ Sample a realization $\{\mathbf{y}^{(1)}, \mathbf{y}^{(2)}\} = \eta_n(\mathbf{x})$ of $H_n(\mathbf{x})$ with $\mathbf{x} = \{\mathbf{x}^{(1)}, \mathbf{x}^{(2)}\}$ ;
**7** $\quad$ Compute $\widehat{Sh}_{n,k,1}^i$ thanks to (33) from $\eta_n(\mathbf{x})$ ;
**8** $\quad$ **for** $l = 2, \ldots, B$ **do**
**9** $\quad\quad$ Sample with replacement a realization $\tilde{\mathbf{y}}^{(1)}$ of $\mathbf{y}^{(1)}$ to compute $\text{Var}(Y)$ ;
**10** $\quad\quad$ Sample by bloc with replacement a realization $\tilde{\mathbf{y}}^{(2)}$ of $\mathbf{y}^{(2)}$;
**11** $\quad\quad$ Compute $\widehat{Sh}_{n,k,l}^i$ thanks to the equation (33) from $\{\tilde{\mathbf{y}}^{(1)}, \tilde{\mathbf{y}}^{(2)}\}$ ;
**12** $\quad$ **end**
**13** **end**
**14** **return** $\left( \widehat{Sh}_{n,k,l}^i \right)_{\substack{k = 1, \ldots, N_H \\ l = 1, \ldots, B}}$

---

The output $\left( \widehat{Sh}_{n,k,l}^i \right)_{\substack{k = 1, \ldots, N_H \\ l = 1, \ldots, B}}$ of the algorithm (3) is a sample of size $N_H \times B$ representative of the distribution of $\widehat{Sh}_n^i$ and takes into account both the uncertainty of the metamodel and that of the Monte Carlo integration.

From this algorithm and some theoretical results, Le Gratiet et al. (2014) proposed estimators in section 4.2 to estimate each of these contributions.

# 7  Numerical simulations with kriging model

This section aims at estimating the studied indices using a surrogate model in substitution of the true and costly computational code. The previous section explained the theory behind the Gaussian processes to emulate a function. The Section 6.3 explained that the kriging error can be estimating through a large number of realization of the Gaussian Process in addition to the Monte-Carlo error estimated through a bootstrap sampling. In this section, we illustrate the decomposition of the overall error from the estimation of the indices and we consider as examples the additive Gaussian framework and the Ishigami function. We also consider the

industrial application of introduced in Rupin et al. (2014) and also used in Iooss and Prieur (2017).

## 7.1 Gaussian framework

We use the same configuration as in the Section 5.3 with a correlation coefficient $\rho = 0.7$. To illustrate the influence of the kriging model in the estimation of the indices, we show in Figure 10 the distribution of the estimators of the indices with the procedure using the true function (top figure) and using the surrogate function (bottom figure). We took $N_v = 1000, N_o = 100$ and $N_i = 3$ for the two graphics.
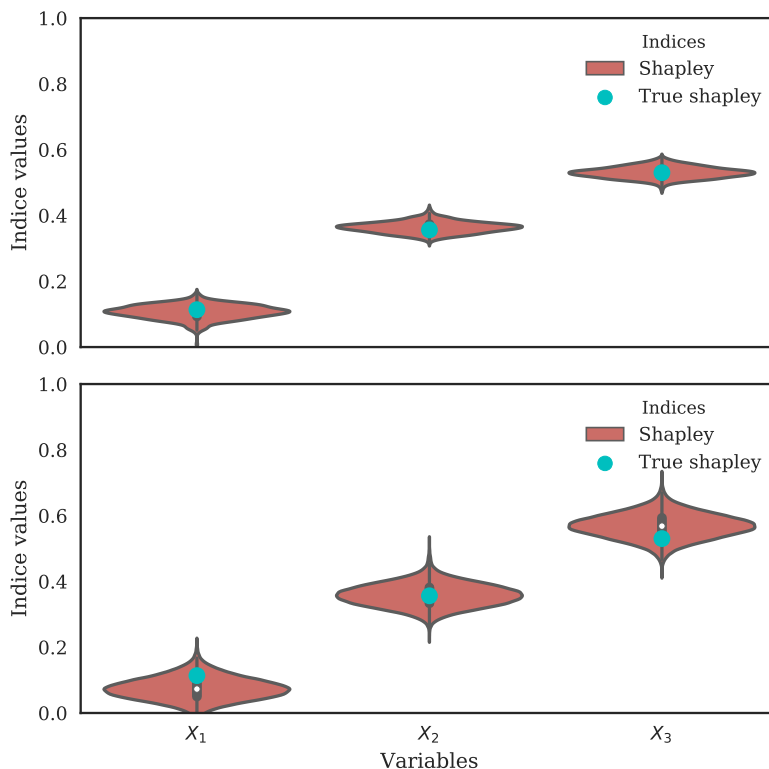


Figure 10: Estimation of the Shapley effects with the exact permutation algorithm. The top and bottom figures respectively show the estimation results with the true function and the kriging model with $Q^2 = 0.90$.

The kriging model is built with 10 points using a LHS sampling (at independence) and a Matern kernel with a linear basis, leading to a $Q^2$ of 0.90 and the kriging error is estimated with $N_H = 300$ realizations. We intentionally took low values for the algorithm parameters in order to have a relatively high variance. If we compare the violinplots of the two figures, we observe that the variance of the estimation is larger for the kriging configuration. This is due to the additional error from the kriging model. The Figure 11 allows to distinguish which part the overall error is due to the kriging. We see immediately what the kriging error is larger than the Monte-Carlo error and it is normal that this error feeds through to the quality of the estimations as observed in Figure 10.
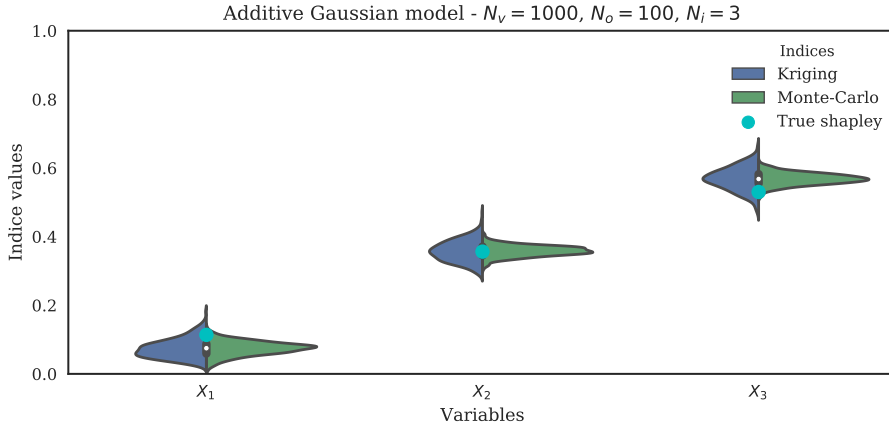
Figure 11: Separation of the uncertainty from the Monte-Carlo estimation and the kriging model approximation.

## 7.2 Ishigami Function

Introduced in Ishigami and Homma (1990), the Ishigami function is typically used as a benchmarking function for uncertainty and sensitivity analysis. It is interesting because it exhibits a strong non-linearity and has interactions between variables. For any variable $\mathbf{x} = (x_1, x_2, x_3) \in [-\pi, \pi]^3$, the model function can be written as

$$\eta(\mathbf{x}) = \sin(x_1) + 7\sin^2(x_2) + 0.1x_3^4\sin(x_1). \tag{35}$$

In this example, we consider that the random variable $\mathbf{X}$ follows a distribution $p_{\mathbf{X}}$ with uniform margins $\mathcal{U}[-\pi, \pi]$ and a multivariate Gaussian copula $c_{\boldsymbol{\rho}}$ with parameter $\boldsymbol{\rho} = (\rho_{12}, \rho_{13}, \rho_{23})$. Thanks to the Sklar Theorem (Sklar, 1959), the multivariate cumulative distribution function $F$ of $\mathbf{X}$ can be written as

$$F(x_1, x_2, x_3) = C_{\boldsymbol{\rho}}\left(F_1(x_1), F_2(x_2), F_3(x_3)\right) \tag{36}$$

where $F_1, F_2, F_3$ are the marginal cumulative distribution functions of $\mathbf{X}$. In the independent case, analytical full first order and independent total Sobol' indices are derived as well as the Shapley effects. Unfortunately, no analytical results are available for the other indices. Thus, we place in the sequel in the independent framework.

Remind that the main advantage of the metamodel is to be much faster-to-calculate than the original function. Thus, we can use this characteristic in order to decrease the Monte-Carlo error during the estimation of the indices by increasing the calculation budget.

In this example, the kriging model is built with 200 points using an optimized LHS sampling (at independence) and a Matern kernel with a linear basis, leading to a $Q^2$ of 0.98 and the kriging error will be estimated subsequently with $N_H = 300$ realizations.

To illustrate the influence of the kriging model in the estimation of the indices, we show in Figure 12 the distribution of the estimators of the indices obtained with the true function (top figure) for $N_v = 1000, N_o = 100, N_i = 3$ and using the surrogate function (bottom figure) with $N_v = 5000, N_o = 600$ and $N_i = 3$. We intentionally took high values for the estimation with the metamodel in order to decrease the overall variance.

If we compare the violinplots of the two figures, we observe that the variance of the estimations is higher with the true function. For the true function, the uncertainty is only due to the

24

Monte-Carlo estimation. For the surrogate function, as observed in Figure 13, in spite of a slight metamodel error, this same Monte-Carlo is obviously lower owing to a higher calculation budget. Hence, if the metamodel approximates correctly the true function, it is better to use it to estimate the sensitivity indices to gain accuracy on the distribution of the estimators.
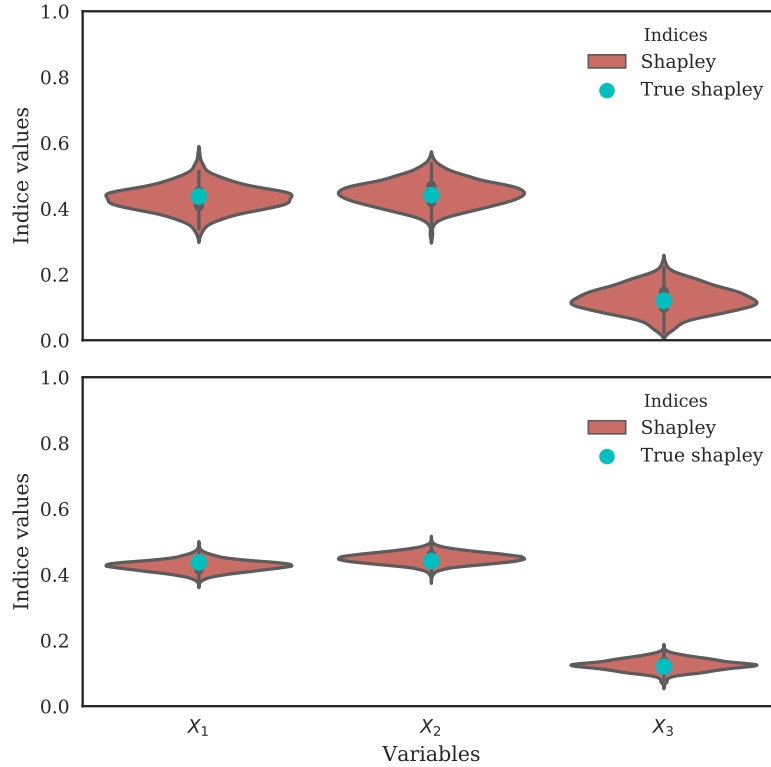


Figure 12: Estimation of the Shapley effects with the exact permutation algorithm. The top and bottom figures respectively show the estimation results with the true function and the kriging model with $Q^2 = 0.98$.

## 8 Conclusion

Throughout this article, we studied the Shapley effects and the *independent* and *full* Sobol' indices defined in Mara and Tarantola (2012) for the models with a dependence structure on the input variables. The comparison between these indices revealed that

- the full Sobol' index of an input includes the effect of another input on which it is dependent,

- the independent and full total Sobol' indices of an input includes the effect of another input on which it is interacting,

- the Shapley effects rationally allocate these different contributions for each input.

Each of these indices allows to answer certain objectives of the SA settings defined in Saltelli and Tarantola (2002) and Saltelli et al. (2004). But, it's important to pay attention about the FP setting. This one can be made with the Shapley effects but not for the goal defined at the outset, i.e. prioritize the input variables taking account the dependence but not to find which would
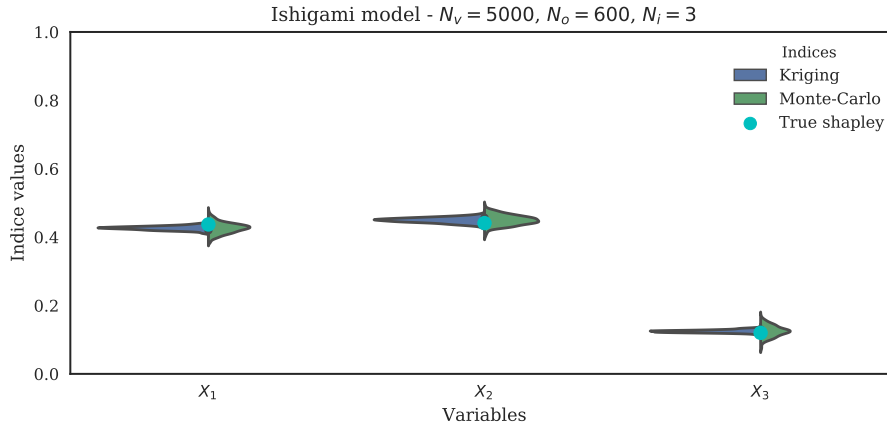
Figure 13: Separation of the uncertainty from the Monte-Carlo estimation and the kriging model approximation.

allow to have the largest expected reduction in the variance of the model output. Always about the FP setting, it was declared in conclusion of our example that the combined interpretation of the four Sobol' indices doesn't allow to answer correctly to the purpose of the FP setting due to the small values that have been obtained for the *independent* Sobol' indices. However, although these values were close to zero, the ranking that they had provided was correct to make FP setting. Hence, it could be investigated whether these values are significant or not.

A relation between the Shapley effects and the Sobol' indices obtained from the RT method was found for the linear Gaussian model. It would be interesting to see if this relation could be extended to a general linear model in the first instance and subsequently if a overall relation can be established between these indices for a global model.

About the estimation procedure of the Shapley effects, a major contribution of this article is the implementation of a bootstrap sampling to estimate the Monte-Carlo error. The CLT can give confidence intervals but require large sample size in order to be consistent, which is rarely possible in practice for expensive computer codes. We confirmed that the parametrization of the Shapley algorithms proposed by Song et al. (2016) and analyzed by Iooss and Prieur (2017) is correct and optimal in order to have consistent confidence intervals. The numerical comparison of the Sobol' indices estimated from the Shapley algorithm and the RT method for a toy example showed that the estimations from the Shapley algorithm are a bit less accurate than the ones from the RT method, but are very satisfying for an algorithm that is not design for their estimation.

A second contribution is the splitting of the metamodel and Monte-Carlo errors when using a kriging model to substitute the true model. The numerical results showed that for a reasonable number of evaluations of a kriging model, one can estimate the Shapley effects, as well as the Sobol' indices and still correctly catch estimation error due to the metamodel or the Monte-Carlo sampling. Unfortunately, the computational cost to generate a sample from a Gaussian Process realization increases significantly with the sample-size. Thus, because the Shapley algorithm becomes extremely costly in high dimension, the estimation of indices using this technique can be computationally difficult.

The Shapley algorithm from Song et al. (2016) is efficient, but is extremely costly in high dimension. The cost is mainly due to the estimation of the conditional variances. A valuable improvement of the algorithm would be the use of a Kernel estimation procedure in order to

significantly reduce the number of evaluation. The Polynomial Chaos Expension are good to compute the Sobol' indices analytically from the polynomial coefficients (Crestaux et al., 2009). It would be interesting to have such a decomposition for the Shapley effects.

# Acknowledgments

# A   Appendix

## A.1   Gaussian framework, linear model

Let us consider

$$Y = \beta_0 + \boldsymbol{\beta}^\mathsf{T} \mathbf{X} \tag{37}$$

with the constants $\beta_0 \in \mathbb{R}$, $\boldsymbol{\beta} \in \mathbb{R}^3$ and $\mathbf{X} \sim \mathcal{N}(0, \Sigma)$ with the following covariance matrix :

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

We obtained the following analytical results.

$$\sigma^2 = Var(Y) = \beta_1^2\sigma_1^2 + \beta_2^2\sigma_2^2 + \beta_3^2\sigma_3^2 + 2\gamma\beta_2\beta_3\sigma_2\sigma_3 + 2\beta_1\sigma_1(\alpha\beta_2\sigma_2 + \rho\beta_3\sigma_3)$$

- For $j = 1, 2, 3$, from the definition of full Sobol indices, we have:

$$\sigma^2 S_1^{full} = \sigma^2 ST_1^{full} = (\beta_1\sigma_1 + \alpha\beta_2\sigma_2 + \rho\beta_3\sigma_3)^2$$
$$\sigma^2 S_2^{full} = \sigma^2 ST_2^{full} = (\alpha\beta_1\sigma_1 + \beta_2\sigma_2 + \gamma\beta_3\sigma_3)^2$$
$$\sigma^2 S_3^{full} = \sigma^2 ST_3^{full} = (\rho\beta_1\sigma_1 + \gamma\beta_2\sigma_2 + \beta_3\sigma_3)^2$$

- We calculate also the full first order Sobol indices for the others subsets of $\mathcal{D}$ and we have :

$$\sigma^2 S_{1,2}^{full} = \beta_1^2\sigma_1^2 + \beta_2^2\sigma_2^2 + 2\gamma\beta_2\beta_3\sigma_2\sigma_3 + 2\beta_1\sigma_1(\alpha\beta_2\sigma_2 + \rho\beta_3\sigma_3) - \frac{\beta_3^2\sigma_3^2\left(\gamma^2 + \rho^2 - 2\alpha\gamma\rho\right)}{\alpha^2 - 1}$$

$$\sigma^2 S_{1,3}^{full} = \beta_1^2\sigma_1^2 + \beta_3^2\sigma_3^2 + 2\gamma\beta_2\beta_3\sigma_2\sigma_3 + 2\beta_1\sigma_1(\alpha\beta_2\sigma_2 + \rho\beta_3\sigma_3) - \frac{\beta_2^2\sigma_2^2\left(\alpha^2 + \gamma^2 - 2\alpha\gamma\rho\right)}{\rho^2 - 1}$$

$$\sigma^2 S_{2,3}^{full} = \beta_2^2\sigma_2^2 + \beta_3^2\sigma_3^2 + 2\gamma\beta_2\beta_3\sigma_2\sigma_3 + 2\beta_1\sigma_1\left(\alpha\beta_2\sigma_2 + \rho\beta_3\sigma_3\right) - \frac{\beta_1^2\sigma_1^2\left(\alpha^2 + \rho^2 - 2\alpha\gamma\rho\right)}{\gamma^2 - 1}$$

$$\sigma^2 S_{\mathcal{D}}^{full} = \sigma^2$$

- We calculate also the total Sobol indices for the variables $(X_i|X_u), i = 1, \ldots, 3$ and $u \subset \mathcal{D}\setminus\{i\}, u \neq \emptyset$ and we have :

$$\sigma^2 ST_{1|2} = -\frac{(\beta_1\sigma_1\,(\alpha^2 - 1) + \beta_3\sigma_3(\alpha\gamma - \rho))^2}{\alpha^2 - 1} \qquad \sigma^2 ST_{1|3} = -\frac{(\beta_1\sigma_1\,(\rho^2 - 1) + \beta_2\sigma_2(\gamma\rho - \alpha))^2}{\rho^2 - 1}$$

$$\sigma^2 ST_{2|1} = -\frac{(\beta_2\sigma_2\,(\alpha^2 - 1) + \beta_3\sigma_3(\alpha\rho - \gamma))^2}{\alpha^2 - 1} \qquad \sigma^2 ST_{2|3} = -\frac{(\beta_2\sigma_2\,(\gamma^2 - 1) + \beta_1\sigma_1(\gamma\rho - \alpha))^2}{\gamma^2 - 1}$$

$$\sigma^2 ST_{3|1} = -\frac{(\beta_3\sigma_3\,(\rho^2 - 1) + \beta_2\sigma_2(\alpha\rho - \gamma))^2}{\rho^2 - 1} \qquad \sigma^2 ST_{3|2} = -\frac{(\beta_3\sigma_3\,(\gamma^2 - 1) + \beta_1\sigma_1(\alpha\gamma - \rho))^2}{\gamma^2 - 1}$$

- For $j = 1, 2, 3$, from the definition of Shapley effects, we have:

$$
\begin{aligned}
Sh_1 &= \frac{1}{3}\left( (\tilde{c}(1) - \tilde{c}(\emptyset)) + \frac{1}{2}(\tilde{c}(1,2) - \tilde{c}(2)) + \frac{1}{2}(\tilde{c}(1,3) - \tilde{c}(3)) + (\tilde{c}(1,2,3) - \tilde{c}(2,3)) \right) \\
&= \frac{1}{3}\left( S_1^{full} + \frac{1}{2}\left( S_{1,2}^{full} - S_2^{full} \right) + \frac{1}{2}\left( S_{1,3}^{full} - S_3^{full} \right) + \left( S_{1,2,3}^{full} - S_{2,3}^{full} \right) \right) \\
&= \frac{1}{3}\left( S_1^{full} + \frac{1}{2}ST_{1|2} + \frac{1}{2}ST_{1|3} + ST_1^{ind} \right)
\end{aligned}
$$

$$
\begin{aligned}
Sh_2 &= \frac{1}{3}\left( (\tilde{c}(2) - \tilde{c}(\emptyset)) + \frac{1}{2}(\tilde{c}(1,2) - \tilde{c}(1)) + \frac{1}{2}(\tilde{c}(2,3) - \tilde{c}(3)) + (\tilde{c}(1,2,3) - \tilde{c}(1,3)) \right) \\
&= \frac{1}{3}\left( S_2^{full} + \frac{1}{2}\left( S_{1,2}^{full} - S_1^{full} \right) + \frac{1}{2}\left( S_{2,3}^{full} - S_3^{full} \right) + \left( S_{1,2,3}^{full} - S_{1,3}^{full} \right) \right) \\
&= \frac{1}{3}\left( S_2^{full} + \frac{1}{2}ST_{2|1} + \frac{1}{2}ST_{2|3} + ST_2^{ind} \right)
\end{aligned}
$$

$$
\begin{aligned}
Sh_3 &= \frac{1}{3}\left( (\tilde{c}(3) - \tilde{c}(\emptyset)) + \frac{1}{2}(\tilde{c}(1,3) - \tilde{c}(1)) + \frac{1}{2}(\tilde{c}(2,3) - \tilde{c}(2)) + (\tilde{c}(1,2,3) - \tilde{c}(1,2)) \right) \\
&= \frac{1}{3}\left( S_3^{full} + \frac{1}{2}\left( S_{1,3}^{full} - S_1^{full} \right) + \frac{1}{2}\left( S_{2,3}^{full} - S_2^{full} \right) + \left( S_{1,2,3}^{full} - S_{1,2}^{full} \right) \right) \\
&= \frac{1}{3}\left( S_3^{full} + \frac{1}{2}ST_{3|1} + \frac{1}{2}ST_{3|2} + ST_3^{ind} \right)
\end{aligned}
$$

# References

Castro, J., Gómez, D., and Tejada, J. (2009). Polynomial calculation of the shapley value based on sampling. *Computers & Operations Research*, 36(5):1726–1730.

Chastaing, G., Gamboa, F., Prieur, C., et al. (2012). Generalized hoeffding-sobol decomposition for dependent variables-application to sensitivity analysis. *Electronic Journal of Statistics*, 6:2420–2448.

Crestaux, T., Le Maı̂tre, O., and Martinez, J.-M. (2009). Polynomial chaos expansion for sensitivity analysis. *Reliability Engineering & System Safety*, 94(7):1161–1172.

Efron, B. (1981). Nonparametric standard errors and confidence intervals. *canadian Journal of Statistics*, 9(2):139–158.

Fang, K.-T., Li, R., and Sudjianto, A. (2005). *Design and modeling for computer experiments.* CRC Press.

Hoeffding, W. (1948). A class of statistics with asymptotically normal distribution. *Annals of Mathematical Statistics*, 19(3):293–325.

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

Iooss, B. and Lemaître, P. (2015). A review on global sensitivity analysis methods. In *Uncertainty Management in Simulation-Optimization of Complex Systems*, pages 101–122. Springer.

Iooss, B. and Prieur, C. (2017). Shapley effects for sensitivity analysis with dependent inputs: comparisons with sobol'indices, numerical estimation and applications. *arXiv preprint arXiv:1707.01334*.

Ishigami, T. and Homma, T. (1990). An importance quantification technique in uncertainty analysis for computer models. In *Uncertainty Modeling and Analysis, 1990. Proceedings., First International Symposium on*, pages 398–403. IEEE.

Janon, A., Klein, T., Lagnoux, A., Nodet, M., and Prieur, C. (2014). Asymptotic normality and efficiency of two sobol index estimators. *ESAIM: Probability and Statistics*, 18:342–364.

Jansen, M. J., Rossing, W. A., and Daamen, R. A. (1994). Monte carlo estimation of uncertainty contributions from several independent multivariate sources. In *Predictability and Nonlinear Modelling in Natural Sciences and Economics*, pages 334–343. Springer.

Kucherenko, S., Tarantola, S., and Annoni, P. (2012). Estimation of global sensitivity indices for models with dependent variables. *Computer Physics Communications*, 183(4):937–946.

Le Gratiet, L., Cannamela, C., and Iooss, B. (2014). A bayesian approach for global sensitivity analysis of (multifidelity) computer codes. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):336–363.

Li, G., Wang, S.-W., Rosenthal, C., and Rabitz, H. (2001). High dimensional model representations generated from low dimensional data samples. i. mp-cut-hdmr. *Journal of Mathematical Chemistry*, 30(1):1–30.

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

Mara, T. A., Tarantola, S., and Annoni, P. (2015). Non-parametric methods for global sensitivity analysis of model output with dependent inputs. *Environmental Modelling & Software*, 72:173–183.

Martin, J. D. and Simpson, T. W. (2004). On the use of kriging models to approximate deterministic computer models. In *ASME 2004 international design engineering technical conferences and computers and information in engineering conference*, pages 481–492. American Society of Mechanical Engineers.

Owen, A. B. (2014). Sobol'indices and shapley value. *SIAM/ASA Journal on Uncertainty Quantification*, 2(1):245–251.

Owen, A. B. and Prieur, C. (2017). On shapley value for measuring importance of dependent inputs. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):986–1002.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. *The annals of mathematical statistics*, 23(3):470–472.

Rupin, F., Blatman, G., Lacaze, S., Fouquet, T., and Chassignole, B. (2014). Probabilistic approaches to compute uncertainty intervals and sensitivity factors of ultrasonic simulations of a weld inspection. *Ultrasonics*, 54(4):1037–1046.

Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989). Design and analysis of computer experiments. *Statistical science*, pages 409–423.

Saltelli, A. (2002). Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications*, 145(2):280–297.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S. (2010). Variance based sensitivity analysis of model output. design and estimator for the total sensitivity index. *Computer Physics Communications*, 181(2):259–270.

Saltelli, A. and Tarantola, S. (2002). On the relative importance of input factors in mathematical models: safety assessment for nuclear waste disposal. *Journal of the American Statistical Association*, 97(459):702–709.

Saltelli, A., Tarantola, S., Campolongo, F., and Ratto, M. (2004). *Sensitivity analysis in practice: a guide to assessing scientific models*. John Wiley & Sons.

Shapley, L. S. (1953). A value for n-person games. *Contributions to the Theory of Games*, 2(28):307–317.

Shapley, L. S. and Shubik, M. (1954). A method for evaluating the distribution of power in a committee system. *American political science review*, 48(3):787–792.

Sklar, M. (1959). *Fonctions de répartition à n dimensions et leurs marges*. Institut de Statistique de l'Université de Paris.

Sobol, I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments*, 1(4):407–414.

Song, E., Nelson, B. L., and Staum, J. (2016). Shapley effects for global sensitivity analysis: Theory and computation. *SIAM/ASA Journal on Uncertainty Quantification*, 4(1):1060–1083.