Pick and Freeze estimation of sensitivity index for static and dynamic models with dependent inputs

Mathilde Grandjacques, Benoît Delinchant, Olivier Adrot

Pick and Freeze estimation of sensitivity index for static and dynamic models with dependent inputs

Mathilde Grandjacques[1], Benoît Delinchant[1], Olivier Adrot[2]

Abstract

This article addresses the estimation of the Sobol index for static and dynamic inputs. We study transformations in the input, whose image is an input with independent components. They have the basic property to give the equality of the σ-algebra between a subset of inputs and their image that allows to compute the Pick and Freeze method.
We first focus on the static case. The Gaussian and non Gaussian cases are detailed. In the Gaussian case the dependent variables are separated into two groups of independent variables. In the non Gaussian case we apply the conditional quantile function generally used to simulate random vectors.
In the dynamic case the definition of the index has been slightly modified in order to take into account the two dimensions of dependence (temporal and spatial). For Gaussian processes the same method as previously is used. For non Gaussian processes, we propose to use a meta-model copula to get back to Gaussian inputs. Different meta-models are studied in order to focus on the limit, in sensitivity studies, of correlations taken as measures of dependence.

Introduction

Global sensitivity analysis (GSA) aims to pick out, in a input-output system, the variables that contribute the most to the uncertainty on the output.

GSA is popular for systems such as non linear regression or more complex systems, which are studied mainly by stochastic tools for independent inputs.

Many methods exist in the literature (see for example [23] and references therein). The most used one is the Sobol index which is defined if the variables are assumed to be independent random variables. Their probability distributions account for the practitioner's belief in the input uncertainty. This turns the model output into a random variable, whose total variance can be split down into different partial variances (this is the so-called Hoeffding decomposition,

[1]G2ELab, Université de Grenoble, 38402 St-Martin d'Hères, France
[2]GSCOP,46 avenue Félix Viallet, 38031 Grenoble Cedex 1, France

also known as functional ANOVA, see [18]). Each partial variance is defined as the variance of the conditional expectation of the output with respect to each input variable. By considering the ratio of each partial variance to the total variance, we obtain the Sobol sensitivity index of the variable [25, 26]. This index quantifies the impact of the variability of the factor on the output. Its value is between 0 and 1, allowing to prioritize the variables according to their influence.

However, in most applications, the parameters or the inputs are dependent due to physical constraints. The interpretation in this case is not easy. If we want to study the sensitivity with respect to a component say X^1, it is of course not sufficient to study only X^1 as it appears explicitly in the model. It contributes to uncertainty through the other components linked with it. So, conventional methods of sensitivity analysis cannot be used with dependent inputs. The classical orthogonal Hoeffding decomposition must be used with precautions. In particular the notion of interaction between two components, valid for independent cases, is here meaningless.

Several studies have been conducted in the case of dependent parameters. Mara and al.[20], Xu and al. [30] give approaches that are most often used for linear models and specific forms of dependence. Da Veiga and al. [9] use a very natural idea : when we have a sample $(X^{(i)}, Y^{(i)})_{i=1,\dots,N}$ of the system, even if f is unknown, we can estimate $\mathbf{E}(Y|X^1)$ and the conditional moments of the output with respect to some factor of interest : X^1. The use of non parametric statistics for this kind of problem is common in other fields such as econometrics, for instance LOESS method is quite easy in this framework. Kutcherenko and al. [17] calculate a sensitivity index analogous to Sobol's formula, from a priori knowledge of probability distribution functions. To obtain it they propose to transform the input into a Gaussian copula. Then this copula is used as a metamodel. If we take an input with uniform marginals and given correlation, there are a lot of models with various properties, so the misspecification when one chooses a metamodel can be important.

A deep work on dependent inputs, starting from ideas of Stone [27] and Hooker [13], is the work of Chastaing and al. ([5], [7]),[6]) perhaps limited by computation problems but giving a clear formulation of the dependence role. This work is based on the existence of an Hoeffding representation under light conditions on the density of the input:

$$f(X^1,\dots,X^p) = \sum_{j=1}^{p} f_j(X^j) + \sum_{\substack{k,j=1\\k\neq j}}^{p} f_{k,j}(X^j,X^k) + \dots + f_{1,\dots,p}(X^1,\dots,X^p) \tag{1}$$

The classical orthogonality due to independence which allows easy computations of the Sobol index is lost but a useful other form of orthogonality extends the classical one. Hierarchical orthogonal decomposition means that a term indexed by $k_1,\dots,k_p$ is orthogonal to any term indexed by a subset of $\{k_1,\dots,k_p\}$ and this property is sufficient to obtain (1).

In a time related framework such that :

$$Y_t = f_t\left((\boldsymbol{X}_s)_{s=0,\dots,t}\right) \tag{2}$$

where $f(\cdot) : \mathbb{R}^{p(t+1)} \to \mathbb{R}$, and the input variable is (not necessarily independent) a vectorial process such that $(\boldsymbol{X}_s)_{s\in\mathbb{N}} \subset \mathbb{R}^p$. Few studies propose to study the sensitivity for dynamic inputs. The sensitivity is calculated at each time step t without taking into account the dynamic behaviour of the input. Indeed, the impact of the variability is not always instantaneous. Therefore it seems necessary to develop a new method for dynamic dependent inputs. The Sobol index definition has been modified in order to take into account the dynamic behaviour of the inputs. Each partial variance is defined as the variance of the conditional expectation of the output with respect to a certain time of observation (called memory) : $(\boldsymbol{X}_t, \dots, \boldsymbol{X}_{t-k})$ of the input vector variable. So the index is defined for each $k \in [0, t]$ and each t. For stationary processes we prove that the sensitivity is independent of k and converges as $t \to \infty$ at least for very general situations but probably not for all (long memory inputs in hydrology are perhaps a counter example).

Thus the main goal of this paper is to find an efficient method of estimation of Sobol indices when the independence assumption on the variables is relaxed and to study the case of time dependence.

Popular methods to calculate Sobol indices are for instance :

- Fourier methods ([8],[19]) used in a different setting, with the aim to simplify computations
- Orthogonal polynomials for example polynomial chaos [28]
- Random balance design [29]

but they require independent input components and (or) a known precise analytical form of f.

A quite different approach, suggested by Sobol (see [26] and [10]) is the Sobol Pick-Freeze (SPF) scheme. It is also based on the independence of components but it is more flexible on the form of the inputs and does not take into account the shape of the input-output model. In SPF, a Sobol index is viewed as the regression coefficient between the output of the model and its pick-freezed replication. This replication is obtained by holding the value of the variable of interest (frozen variable) and by sampling the other variables (picked variables). The sampled replications are then combined to produce an estimator of the Sobol index. There is no requirement about the knowledge of f, except the possibility to simulate the system which is of course a severe constraint. Janon and al. ([14], [15]) give asymptotic results when the sample size tends to infinity. Estimators are convergent, satisfy a central limit theorem and have robustness properties.

Our proposition is to find a transformation that turns dependent inputs into independent inputs and that keeps the sensitivity invariant. It is then possible to apply the Pick and Freeze method to the independent variables.

The first part of the paper is dedicated to the static part. For Gaussian inputs, a general multivariate regression gives the transformation. For non Gausian inputs a method based on the conditional quantile method is proposed. This method is the most general to simulate random vectors. We detail a copula type metamodel and to complete Kutcherenko and al.'s [17] study, we show that the choice of the copula is important, concerning sensitivity analysis. This metamodel will be used in the dynamic part.

The second part is focused on dynamic models. We first consider Gaussian process inputs and we can apply the same method as in the static Gaussian case. For the non Gaussian case we consider a copula type metamodel to get back to Gaussian inputs. It is chosen starting from each input marginal function and starting from the correlation between $\boldsymbol{X}_t$ and $\boldsymbol{X}_{t-1}, \boldsymbol{X}_{t-2}, \ldots$. The sensitivities are computed for different simple models and we deal with, specifically for sensitivity values, the difficulties arising from the use of correlation as a dependence measure to build convenient metamodels.

1 Sensitivity for dependent inputs : vectorial case

1.1 Sobol index and Pick and Freeze method : independent case

Let an input-output system given by :

$$Y = f(\boldsymbol{X}) \tag{3}$$

with $Y \in \mathbb{R},\ \boldsymbol{X} = (X^1, \ldots, X^p) \in \mathbb{R}^p$

The Sobol index is defined as :

$$S^{X^J} = \frac{\mathbf{Var}\left(\mathbf{E}(Y|\boldsymbol{X}^J)\right)}{\mathbf{Var}Y}. \tag{4}$$

with $\boldsymbol{X}^J = (X^{j_1}, \ldots, X^{j_q})$, $J = \{j_1, \ldots, j_q\} \subset \{1, \ldots, p\}$.

Lemma 1. *Sobol [26]: Let $\boldsymbol{X} = (\boldsymbol{X}^J, \boldsymbol{X}^{\bar{J}})$ and $Y = f(\boldsymbol{X}^J, \boldsymbol{X}^{\bar{J}})$ with $\bar{J} = \{1, \ldots, p\} \setminus J$. If $\boldsymbol{X}^J$ and $\boldsymbol{X}^{\bar{J}}$ are independent :*

$$\mathbf{Var}(\mathbf{E}(Y|\boldsymbol{X}^J)) = \mathbf{Cov}(Y, Y^X)$$

with $Y^{X^J} = f(\boldsymbol{X}^J, (\boldsymbol{X}^{\bar{J}})')$ where $(\boldsymbol{X}^{\bar{J}})'$ is an independent copy of $\boldsymbol{X}^{\bar{J}}$. Copy meaning a random vector independent of $\boldsymbol{X}^{\bar{J}}$ with the same distribution.

We can deduce the expression of the index S^{X^J} when $\boldsymbol{X}^J$ and $\boldsymbol{X}^{\bar{J}}$ are independent :

$$S^{X^J} = \frac{\mathbf{Cov}(Y, Y^{X^J})}{\mathbf{Var}Y} \tag{5}$$

A natural estimator consists in taking the empirical estimators of the covariance and of the variance. Let a $N-$sample $\{(Y^{(1)}, Y^{X,(1)}), \ldots, (Y^{(N)}, Y^{X,(N)})\}$ a natural estimator of S^{X^J} is :

$$\widehat{S}^{X^J} = \frac{\frac{1}{N}\sum_{i=1}^N Y^{(i)} Y^{X^J,(i)} - (\frac{1}{N}\sum_{i=1}^N Y^{(i)})(\frac{1}{N}\sum_{i=1}^N Y^{X^J,(i)})}{\frac{1}{N}\sum_{i=1}^N (Y^{(i)})^2 - (\frac{1}{N}\sum_{i=1}^N Y^{(i)})^2} \tag{6}$$

Janon and al. [14] suggest an improvement using a symmetric form:

$$\tilde{S}^{X^J} = \frac{\frac{1}{N}\sum_{i=1}^N Y^{(i)} Y^{X^J,(i)} - (\frac{1}{2N}\sum_{i=1}^N Y^{(i)} + Y^{X^J,(i)})^2}{\frac{1}{N}\sum_{i=1}^N \frac{(Y^{(i)})^2 + (Y^{X^J,(i)})^2}{2} - (\frac{1}{N}\sum_{i=1}^N \frac{Y^{(i)} + Y^{X^J,(i)}}{2})^2} \tag{7}$$

In [14] these estimators are shown to be consistent, and if $\mathbf{E}(|Y|^4) < \infty$, they satisfy a central limit theorem. The variance of $\widehat{S}^{X^J}$ is the lowest possible one.

1.2 General framework to reduce the dependent case to the independent case

1.2.1 General framework

Let $Y = f(\boldsymbol{X})$ an input-output system and $\boldsymbol{X} = (X^1, \ldots, X^p)$. The Pick and Freeze method is based on the Sobol lemma which requires independent inputs. If $\boldsymbol{X}^J = (X^{j_1}, \ldots, X^{j_q})$ is a set of components on which sensitivity is computed, so $\boldsymbol{X}^J$ has to be independent of $\boldsymbol{X}^{\bar{J}}$ the set of components not in $\boldsymbol{X}^J$. To apply the Pick and Freeze method to dependent inputs we look for a transformation $T : \mathbb{R}^p \to \mathbb{R}^p$ so that there is a pair of independent vectors $(\boldsymbol{V}, \boldsymbol{W})$ such that :

$$\boldsymbol{X}^J = \phi(\boldsymbol{V}) \tag{8}$$
$$\boldsymbol{X}^{\bar{J}} = \psi(\boldsymbol{V}, \boldsymbol{W}) \tag{9}$$

(ϕ, ψ) measurable functions giving T with the basic property of equality of $\sigma-$algebras :

$$\sigma(\boldsymbol{X}^J) = \sigma(\boldsymbol{V}) \tag{10}$$

The equation (10) is equivalent to the existence of ϕ^{-1} a measurable function.

Under these conditions, Y has a new expression :

$$\begin{aligned} Y &= f(\boldsymbol{X}^J, \boldsymbol{X}^{\bar{J}}) \\ &= f\left(\phi(\boldsymbol{V}), \psi(\boldsymbol{V}, \boldsymbol{W})\right) \\ &= \tilde{f}(\boldsymbol{V}, \boldsymbol{W}) \end{aligned}$$

So the conditional expectation of Y given the $\sigma-$algebra of $\boldsymbol{X}^J$ can be rewritten as :

$$\mathbf{E}\left(f(\boldsymbol{X}^J, \boldsymbol{X}^{\bar{J}})|\sigma(\boldsymbol{X}^J)\right) = \mathbf{E}\left(f(\boldsymbol{X}^J, \boldsymbol{X}^{\bar{J}})|\sigma(\boldsymbol{V})\right) \tag{11}$$
$$= \mathbf{E}\left(\tilde{f}(\boldsymbol{V}, \boldsymbol{W})|\sigma(\boldsymbol{V})\right) \tag{12}$$

Thus we can compute S^{X^J} computing S^V with the new expression $\tilde{f}(\boldsymbol{V}, \boldsymbol{W})$:

$$S^{X^J} = S^V \tag{13}$$

To apply the Pick and Freeze method to calculate the Sobol index, we apply the Sobol Lemma 1 to the input-output system :

$$Y = \tilde{f}(\boldsymbol{V}, \boldsymbol{W})$$

So, to estimate the index we need to be able to simulate large samples of $(\boldsymbol{V}, \boldsymbol{W})$. The simplest example is the case of Gaussian inputs.

1.2.2 The Gaussian case

Let $\boldsymbol{X}$ a Gaussian vector and $Y = f(\boldsymbol{X})$ a input-output model. The multi-regression of $\boldsymbol{X}^{\bar{J}}$ onto $\boldsymbol{X}^J$ is written as :

$$\boldsymbol{X}^{\bar{J}} = \Lambda \boldsymbol{X}^J + \boldsymbol{W} \tag{14}$$

where $\boldsymbol{W}$ is a Gaussian vector independent of $\boldsymbol{X}^J$ and Λ a $(p-q) \times q$ matrix given by :

$$\Lambda = \Gamma^{\bar{J}J}(\Gamma^{JJ})^{-1} \tag{15}$$

with $\Gamma^{\bar{J}J} = \mathbf{E}(\boldsymbol{X}^{\bar{J}}(\boldsymbol{X}^J)^*)$ and $\Gamma^{JJ} = \mathbf{E}(\boldsymbol{X}^J(\boldsymbol{X}^J)^*)$.

Thus :

$$\begin{cases} \boldsymbol{V} & = \boldsymbol{X}^J \\ \boldsymbol{W} & = \boldsymbol{X}^{\bar{J}} - \Lambda \boldsymbol{V} \end{cases} \tag{16}$$

The Pick and Freeze method can be applied to the variables $(\boldsymbol{V}, \boldsymbol{W})$ and the Gaussian vector $\boldsymbol{W}$ whose covariance matrix is :

$$\mathbf{E}((\boldsymbol{X}^{\bar{J}} - \Lambda \boldsymbol{X}^J)(\boldsymbol{X}^{\bar{J}} - \Lambda \boldsymbol{X}^J)^*) \tag{17}$$

can be simulated.

1.3 The conditional quantile method

We now consider a model $Y = f(X^1, \dots, X^p)$, $\boldsymbol{X} = (X^1, \dots, X^p)$. We suppose that $\boldsymbol{X}$ has a density g with respect to the Lebesgue measure.

Let:

$$\begin{aligned} G(x^k | X^1 = x^1, \dots, X^{k-1} = x^{k-1}) & = G_{k|1,\dots,(k-1)}(x^k, x^1, \dots, x^{k-1}) \\ & = \mathbf{P}(X^k < x^k | X^1 = x^1, \dots, X^{k-1} = x^{k-1}) \end{aligned} \tag{18}$$

the conditional distribution of X^k when $(X^1, \dots, X^{k-1})$ are fixed. It results from the existence of g that all these conditional distributions are well defined.

Lemma 2. *Lévy-Rosenblatt [22] : Let $(U^1, \dots, U^p)$ the random variables defined for $1 \le k \le p$ such that :*

$$U^k = G_{k|1,\dots,(k-1)}(X^k, X^1, \dots, X^{k-1}) \tag{19}$$

so $(U^1, \dots, U^p)$ are uniform and independent random variables.

Proof. The proof is quite obvious

$$\begin{aligned} \mathbf{P}(U^i \le u^1, i = 1, \dots, p) & = \int_{\{U, U^i \le u^i\}} \dots \int G_{p|1,\dots,(p-1)}(x^p, x^1, \dots, x^{p-1})\, \mathrm{d}x^p \dots \int G_1(x^1) \mathrm{d}x^1 \\ & = \int_0^{u^k} \dots \int_0^{u^1} \mathrm{d}u^k \dots \mathrm{d}u^1 \end{aligned}$$

by definition of conditional distributions and chain property. □

The formula (19) can be read as : there is a not unique transformation T such that :

$$(U^1, \ldots, U^p) \xrightarrow{\mathrm{T}} (X^1, \ldots, X^p)$$

Every permutation $(\tau(1), \ldots, \tau(p))$ of $(1, \ldots, p)$ gives a transformation $T \circ \tau$ with the same properties.

We now study T^{-1}, which can be seen as a vector of conditional quantile function.

In order to simplify the definition of inverse functions, we make a (weak) assumption on g.

Let $\mathcal{C} = closure\{\boldsymbol{x}, \;\; g(\boldsymbol{x}) > 0\}$ and assume:

$$g(\boldsymbol{x}) > 0 \text{ if } \boldsymbol{x} \in interior(\mathcal{C}) \tag{20}$$

From (20), $G_{k|1,\ldots,(k-1)}(x^k, x^1, \ldots, x^{k-1})$ is a strictly increasing continuous function from $\mathbb{R}$ to $[0, 1]$ for every $(x^1, \ldots, x^{k-1})$. Thus $G^{-1}_{k|1,\ldots,(k-1)}$ is well defined.

So from (20) we get by induction :

$$\begin{aligned} X^k &= G^{-1}_{k|1,\ldots,(k-1)}(U^k, X^1, \ldots, X^{k-1}) \\ &= G^{-1}_{k|1,\ldots,(k-1)}(U^k, X^1(U^1), \ldots, X^{k-1}(U^1, \ldots, U^{k-1})) \end{aligned}$$

noted for simplicity :

$$G^{-1}_{k|1,\ldots,(k-1)}(U^k, U^1, \ldots, U^{k-1}) \tag{21}$$

The following lemma is a consequence of (19) and (21)

Lemma 3. *The $\sigma-$fields $\sigma(U^1, \ldots, U^k)$ and $\sigma(X^1, \ldots, X^k)$ are equal for every k and we have the equality of conditional expectations as operators onto $L^2(g(x)dx)$:*

$$\mathbf{E}(\cdot|X^1, \ldots, X^k) = \mathbf{E}(\cdot|U^1, \ldots, U^k)$$

Now :

$$\begin{aligned} Y &= f(X^1, \ldots, X^p) && (22) \\ &= f(X^1(U^1), \ldots, X^p(U^1, \ldots, U^p)) && (23) \\ &= \tilde{f}(U^1, \ldots, U^p) && (24) \end{aligned}$$

and by lemma (3):

$$S^{U^1} = S^{X^1} \tag{25}$$

where S^{U^1} is given by $\dfrac{\mathbf{VarE}\left(\tilde{f}(U^1, \ldots, U^p)|U^1\right)}{\mathbf{Var}(Y)}$, $\tilde{f}$ given by (24),

When we use a transformation T associated with a specific ordering, say $X^1, \ldots, X^p$ we can only compute sensibilities as S^{X^1}, $S^{X^1X^2}, \ldots,$ $S^{X^1 \ldots X^p}$.

Thus if we want to compute all first order sensitivity indices S^{X^k} with $k = 1, \dots, p$, we choose an order among the $(p-1)!$ beginning by X^k.

If we want all second order sensitivity indices we need exactly $\dfrac{p(p-1)}{2}$ different orders and so on, if we want $S^{X^{j_1}, \dots, X^{j_q}}$ we have to take an order beginning by $(j_1, \dots, j_q)$.

The Sobol index S^{X^J} is defined as $\dfrac{\mathbf{Var}\left(\mathbf{E}(Y|X^J)\right)}{\mathbf{Var}(Y)}$. For every q, $J = j_1, \dots, j_q$:

$$S^{U^J} = S^{X^J} \tag{26}$$

Before studying the estimation of S^{X^1} let us take the following example :

Example : Let $p = 2$. The input vector is assumed to be : $\boldsymbol{X} = (X^1, X^2)$ defined by a uniform distribution on the triangle:

$$D = \begin{cases} 0 \le x^1, x^2 \le 1 \\ x^1 + x^2 \le 1 \end{cases}$$

and the input-output model is :

$$f(X^1, X^2) = X^1 + X^2$$

We compare the index S^{X^1} calculated directly and the index S^{U^1} using the transformation T such that : $(U^1, U^2) \xrightarrow{\text{T}} (X^1, X^2)$.

Let us calculate U^1 and U^2 and deduce X^1 and X^2 :

$$U^1 = G(X^1) = 2X^1 - (X^1)^2 \text{ implies } X^1 = 1 - \sqrt{1 - U^1}$$
$$U^2 = G_{X^2|X^1}(X^2) = \frac{X^2}{1 - X^1} \mathbb{1}_{[0, 1-X^1]} \text{ so } X^2 = U^2\sqrt{1 - U^1}$$

The density of f is $2v\mathbb{1}_{0<v<1}$ and its variance $\dfrac{1}{18}$.

$\mathbf{E}(X^1 + X^2|X^1) = \dfrac{1 + X^1}{2}$ and $\mathbf{Var}(\dfrac{1 + X^1}{2}) = \dfrac{1}{72}$, thus the indices values are :

$$S^{X^1} = S^{X^2} = 1/4$$

Now if we use the function $\tilde{f}(U^1, U^2) = 1 - \sqrt{1 - U^1} + U^2\sqrt{1 - U^1}$:

$$\mathbf{E}\left(\tilde{f}(U^1, U^2)|U^1\right) = 1 - \sqrt{1 - U^1} + \frac{1}{2}\sqrt{1 - U^1} = 1 - \frac{1}{2}\sqrt{1 - U^1}$$
$$\mathbf{Var}\left(1 - \frac{1}{2}\sqrt{1 - U^1}\right) = 1/72$$

So : $S^{U^1} = S^{X^1}$.
But we can notice that $S^{U^2} \neq S^{X^2}$. To calculate S^{X^2} we need to reorder the variables. It means that :

$$U^1 = G(X^2)$$
$$U^2 = G_{X^1|X^2}(X^1)$$

Then $S^{X^2} = S^{U^1}$ with this notation.

The Hoeffding decomposition of $\tilde{f}$ is obtained by centering $\sqrt{1-U^1}$ and U^2 :

$$\tilde{f}(U^1, U^2) = \frac{1}{3} - \frac{1}{2}\left(\sqrt{1-U^1} - \frac{2}{3}\right) - \frac{2}{3}\left(U^2 - \frac{1}{2}\right) + \left(U^2 - \frac{1}{2}\right)\left(\sqrt{1-U^1} - \frac{2}{3}\right)$$

S^{U^1,U^2} can be interpreted as an interaction sensibility between U^1 and U^2. The sensitivity with respect to X^2 depends on U^1 and U^2 so there is no obvious interpretation in (X^1, X^2).

1.3.1 Pick and Freeze estimation and conditional quantile method

Starting from $Y = f(X^1, \ldots, X^p)$ we have built a model $Y = \tilde{f}(U^1, \ldots, U^p)$ using a transformation $T : \boldsymbol{U} \to T(\boldsymbol{X})$ from $\mathbb{R}^p$ to $\mathbb{R}^p$ which uses a specific order on $\{1, \ldots, p\}$.

If τ is a permutation of $1, \ldots, p$ then $T \circ \tau$ gives the same $\tilde{f}$ which is permutation invariant.

$\tilde{f}$ can be considered as the intrinsic model associated to f. It is *the only model with i.i.d uniform variable inputs giving the same output* Y. Now this model has a Hoedfding form which is *the intrinsic Hoeffding form.*
In fact we use this form to define the Pick and Freeze method.

Thus the algorithm to estimate S^{X^1} is as follows:

1. Simulate $(p-1)-$samples $(U^{2,(i)})', \ldots, (U^{p,(i)})'$), $i = 1, \ldots, N$ of uniform independent variables.

2. **Main step:** Solve equations recursively :

$$G_{k|1,\ldots,k-1}((x^{k,(i)})', (x^{1,(i)})', \ldots, (x^{k-1,(i)})') = (U^{k,(i)})' \quad i = 1, \ldots, N$$

   Let $(X^{k,(i)})'$ be the solution, $k = 2, \ldots, p$; $i = 1, \ldots, N$ and for
   $k = 1$, $G_1(X^{1,(i)}) = U^{1,(i)}$

   The Newton or Quasi Newton method is easy to apply here to solve these one dimensional equations, for $G_{k|1,\ldots,k-1}$ are continuous, strictly increasing functions of x^k ([12]).

3. Compute $Y^{(i)}$ and $Y^{X^J,(i)}$ using as inputs $(X^{1,(i)}, \ldots, X^{p,(i)})$ or
   $(X^{1,(i)}, (X^{2,(i)})', \ldots, (X^{p,(i)})')$

4. With these outputs we can estimate $\hat{S}^{X^1}$ using formula (6) or (7).

Algorithm 1 Quantile method

Require: N,$G_{i|1,\ldots,(i-1)}$ for all $i = 1, \ldots, p$

$U = matrix(0, ncol = N, nrow = p)$; $U' = matrix(0, ncol = N, nrow = p)$
$X = matrix(0, ncol = N, nrow = p)$; $X' = matrix(0, ncol = N, nrow = p)$

$\boldsymbol{U} \sim \mathcal{U}$ {Simulation of a sample of Uniform variables of size N}
$\boldsymbol{U'} \sim \mathcal{U}$ {Simulation of a second sample of Uniform variables of size N}

$X[1,] =$ Solve $(G_1(X) = U[1,])$
$X'[1,] = X[1,]$ {$X[1,]$ is frozen}
for $i = 2$ to p **do**
  for $j = 1$ to N **do**
    $X[i,j] =$ Solve $\big(G_{i|1,\ldots,(i-1)}(X) = U[i,j]\big)$
    $X'[i,j] =$ Solve $\big(G_{i|1,\ldots,(i-1)}(X) = U'[i,j]\big)$
  end for
end for

$Y = \eta(X)$ {Sample of size N with the variable $X^1[1,]$ frozen}
$Y^X = \eta(X')$ {Sample of size N with the variable $X^1[1,]$ frozen}
return Y, Y^X

Remark 1. *The use of conditional quantile functions is the most general key to simulate any random vector. If $G_{k|1,\ldots,k-1}$ is known for every k, it can be possible to simulate $\boldsymbol{X}^{(i)}$ using the simulation of $\boldsymbol{U}^{(i)} = (U^{1,(i)}, \ldots, U^{p,(i)})$ then solving this equation recursively :*

$$G_{k|1,\ldots,k-1}(x^{k,(i)}, X^{1,(i)}, \ldots, X^{k-1,(i)}) = U^{k,(i)}$$

when the solution is $X^{1,(i)}$.

Remark 2. *In the Gaussian case, the independence is obtained by using the properties of the Gaussian distribution as said. In fact, the choice of uniform distribution is arbitrary, one can choose any fixed by advance distribution, for instance Gaussian. The functions $\phi_1, \ldots, \phi_p$ are in the Gaussian case linear and T is simply a linear transformation given as follows.*

For Gaussian distribution, we define sequentially [22] :

$$G(x^k | x^1, \ldots, x^{k-1}) = \phi\left(\frac{x^k - m_k + \sum_{j=1}^{k-1}(C_{kj}/C_{kk})(x^j - m_j)}{\sqrt{C/C_{kk}}}\right) \tag{27}$$

where C^r_{kj} is the cofactor of C_{kj} in C^r when C^r is the restriction of the covariance matrix $C = (C_{kj})_{k=1,\ldots,p;\ j=1,\ldots,p}$ to $1 \le j, k \le r$

Remark 3. *We have detailed two methods to get the application :*

$$T(\boldsymbol{X}^J, \boldsymbol{X}^{\bar{J}}) \to (\boldsymbol{V}, \boldsymbol{W})$$

with $(\boldsymbol{V}, \boldsymbol{W})$ independent and with : $\sigma(\boldsymbol{V}) = \sigma(\boldsymbol{X}^J)$. Other methods can be of course used, for instance Mara and al. [20].

1.4 Applications of the two methods to meta models associated to constrained copulas

Copulas are used to build metamodels when the information extracted from the data is incomplete. The information is the p repartitions $F^1, \dots, F^p$ of the p random variables of interest (such as the input in an input-output system) ([21],[24]). The basic representation is :

Proposition 1. *[24] If $F^1, \dots, F^p$ are the distribution functions of $X^1, \dots, X^p$, then if F is the distribution function of $\boldsymbol{X}$, there is a function C such that :*

$$F(x^1, \dots, x^p) = C(F^1(x^1), \dots, F^p(x^p)) \tag{28}$$

If F^i admits a probability density f^i then C admits on $\mathbb{R}^p$ the probability density $c = \dfrac{\partial^p C}{\partial x^1 \dots \partial x^p}$

Gaussian copulas associated to uniform distributions (see for instance [2]) are the most popular .

Let a Gaussian vector $\boldsymbol{Z} = (Z^1, \dots, Z^p)$ whose correlation matrix is Υ.
Let Φ the standard Normal cumulative distribution :

$$X^i = \Phi(Z^i) \tag{29}$$

$\boldsymbol{X} = (X^1, \dots, X^p)$ is a random vector with Uniform components defining completely the uniform Gaussian copula.
Of course it can be an arbitrary choice in the set of probabilities defined on $\mathbb{R}^p$ with p uniform marginals, but the copula can be constrained by a specific choice of Υ. We suppose from now on that the $p \times p$ matrix $R = (R^{i,j})_{\substack{i=1,\dots,p \\ j=1,\dots,p}}$ defined as :

$$R^{i,j} = \mathbf{Cor}(X^i, X^j) = \frac{\mathbf{E}(X^i X^j) - \mathbf{E}(X^i)\mathbf{E}(X^j)}{\sqrt{\mathbf{Var}(X^i)\mathbf{Var}(X^j)}} \tag{30}$$

is given and thus we want to choose $\Upsilon = (\Upsilon^{i,j})_{\substack{i=1,\dots,p \\ j=1,\dots,p}}$ (the correlation matrix of the Gaussian vector $\boldsymbol{Z}$) so that R is the correlation matrix of $\boldsymbol{X}$.

By a calculation of the type :

$$\mathbf{E}(X^i X^j) = \frac{1}{2\pi\sqrt{1-(\Upsilon^{i,j})^2}} \int \Phi(Z^i)\Phi(Z^j) \exp^{-\frac{1}{2(1-(\Upsilon^{i,j})^2)}((Z^i)^2 - 2\Upsilon^{i,j}Z^iZ^j + (Z^j)^2} \mathrm{d}Z^i \mathrm{d}Z^j \tag{31}$$

R and Υ are linked by a simple form [1]:

$$\Upsilon^{i,j} = 2\sin\left(\frac{\pi R^{i,j}}{6}\right) \tag{32}$$

Note that $|\Upsilon| = 1 \Leftrightarrow |R| = 1$, $\Upsilon = 0 \Leftrightarrow |R| = 0$.

Numerically :

$$\Upsilon^{i,j} = 1.047R^{i,j} - 0.047(R^{i,j})^3 \tag{33}$$

is a good approximation. Thus this correspondence between $R^{i,j}$ and $\Upsilon^{i,j}$ is well defined. R is given as a correlation matrix but nothing says that (in p dimensions) $\Upsilon = (\Upsilon^{i,j})_{\substack{i=1,\dots,p \\ j=1,\dots,p}}$ is a correlation matrix (positive type matrix). This point is related to the following definition.

Definition 1. *Let $(F^i)_{i=1,\ldots,p}$ a family of marginal distributions and R a correlation matrix. We say that $((F^i)_{i=1,\ldots,p}, R)$ is feasible as a Gaussian copula if and only if there is a Gaussian vector $\boldsymbol{Z} = (Z^i)_{i=1,\ldots,p}$ whose correlation matrix is Υ, satisfying :*

$$X^i = (F^i)^{-1}(\Phi(Z^i)) \text{ for } i = 1, \ldots, p \tag{34}$$

where X has R as correlation.

We don't discuss here the problem of feasibility. If Υ is not positive it is possible to find correlation matrices "close to" Υ ([11]).

We now want to illustrate what happens for sensitivity values for different choices of constrained copulas.

Sensitivity is estimated by the Pick and Freeze method in all the cases. We have selected *the model of Ishigami*, a classical toy model in sensitivity and optimisation studies defined as follows :

$$Y = \sin(X^1) + 7\sin(X^2) + 0.1(X^3)^4 \sin(X^1) \tag{35}$$

(X^1, X^2, X^3) have a uniform distribution with support $[-\pi, \pi]$.

The correlation matrix given is : $\begin{pmatrix} 1 & 0 & \rho \\ 0 & 1 & 0 \\ \rho & 0 & 1 \end{pmatrix}$ and X^2 is supposed (to simplify) to be independent of the pair (X^1, X^3). We consider two distributions :

- case 1 :the Gaussian copula
- case 2 : the copula given by this density probability :

$$f_\alpha(x^1, x^2, x^3) = \frac{1}{8\pi^3}(\mathbb{1}_{[-\pi,\pi]^2}(x^1, x^3) + \alpha x^1 x^3 \tag{36}$$

  f_α is a density probability if $|\alpha| \leq \dfrac{1}{4\pi^2}$. This condition implies that :
  $\rho = \mathbf{E}(X^1 X^2) = \dfrac{4\pi^3\alpha}{3}$ thus $|\rho| \leq \dfrac{\pi}{9}$.

The sensitivity values are calculated by applying the method of conditional quantile and the results are discussed with respect to different ρ values.

Case 1 : Gaussian copula :

$Z^i, i = 1, 2, 3$ are defined by :

$$X^i = \pi(2\Phi(Z^i) - 1)$$

where X^i is uniform on $[-\pi, \pi]$. The correlation ρ' of Z^1, Z^3 is given by :

$$\rho' = 2\sin(\frac{\pi\rho}{6})$$

Following our previous results in section 1.2.2 we write the Ishigami model with independent Gaussian variables Z^1, Z^2, W.
Z^1 is defined by : (29) and W by regression :

$$Z^3 = \rho' Z^1 + \sqrt{1-(\rho')^2}W \ \text{ with } W \sim \mathcal{N}(0,1)$$

Thus the input-output system is now :

$$Y = \sin(\pi(2\Phi(Z^1)-1)) + 7\sin(\pi(2\Phi(Z^2)-1)) + 0.1\left(\pi(2\Phi(\rho' Z^1 + \sqrt{1-\rho'^2}W)-1)\right)^4 \sin(\pi(2\Phi(Z^1)-1)) \tag{37}$$

As X^1 X^2 are independent we know that $S^{X^1} = S^{Z^1}$ and $S^{X^2} = S^{Z^2}$.
To compute S^{X^3} we need to use the regression $Z^1 = \rho' Z^3 + \sqrt{1-\rho'^2}W$ in the Ishigami output Y.

Results related to ρ are plotted in figure : 1.

$\rho = 0$ corresponds to the case of independent variables. If $\rho = 1$, $S^{X^1} = S^{X^3}$.

Case 2 : f_α copula :
The conditional quantile method is used to calculate the indices. First as X^1 and X^2 are independent variables we have :

$$U^1 = \frac{X^1/\pi + 1}{2}$$
$$U^2 = \frac{X^2/\pi + 1}{2}$$

U^1 and U^2 are uniform independent variables.
U^3 is defined such that :

$$U^3 = F_{X^3|X^1}(X^3) = \frac{1}{2\pi}(X^3+\pi) + \pi\alpha X^1 \frac{(X^3)^2 - \pi^2}{2}$$

U^3 is a uniform variable independent from U^1 and U^2.
Thus :

$$X^3 = \frac{-1 + \sqrt{1 - 8\pi^2\alpha X^1(1 - \pi^2\alpha X^1 - 2U^3)}}{4\alpha\pi X^1} \tag{38}$$

The canonical formula for the Ishigami and the input given by the $\alpha-$copula is obtained by the substitution in (35). With this order $1,2,3$ we can calculate the indices $S^{X^1}, S^{X^2}, S^{X^1X^2}$ (X^1 and X^2 are independent in this case). If we want to compute S^{X^3} we have to resume our work choosing the order $(3,2,1)$ for example :

$$U^1 = \frac{X^3/\pi + 1}{2}$$
$$U^2 = \frac{X^2/\pi + 1}{2}$$

and thus U^3 is defined by :

$$U^3 = \frac{1}{2\pi}(X^1 + \pi) + \pi\alpha X^3 \frac{(X^1)^2 - \pi^2}{2}$$

The results are given in figure : 1. We can only compare the results for $0 \leq |\rho| \leq \frac{\pi}{9}$, for instance in table : 1

These results show that the practitioner has to be cautious with the use of models with incomplete information when sensitivities are computed. Correlation does not give, in any case, a very good information on dependences when we compute the sensitivity of different inputs. For a same correlation we get different copulas, which gives very different sensitivity results.

	ρ	S^{X^1}	S^{X^2}	S^{X^3}
copula f_α	10^{-7}	0.31	0.44	0
copula f_α	$\pi/9$	0.36	0.40	0.46
Gaussian copula	0	0.31	0.44	0
Gaussian copula	$\pi/9$	0.30	0.50	0.08

Table 1: Sensitivity for Gaussian copula and f_α copula, for different values of ρ

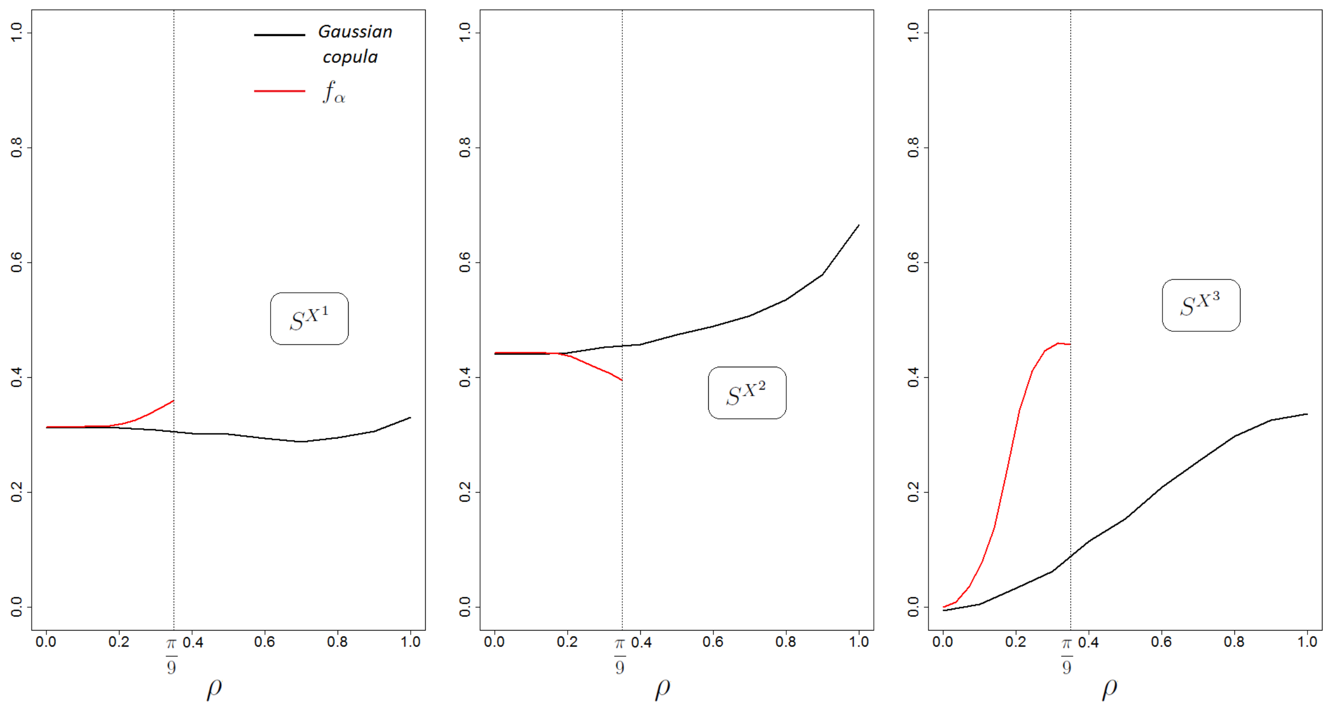

Figure 1: Sensitivity indices for different values of ρ' applied to the Ishigami model for the Gaussian copula and the f_α copula

2 Sensitivity for vectorial stochastic process inputs

2.1 Sensitivity for vectorial stochastic process inputs and memories

Suppose that we consider an input-output system :

$$Y_t = f_t(\boldsymbol{X}_t, \dots, \boldsymbol{X}_{t-k}, \dots, \boldsymbol{X}_0), \;\; t \in \mathbb{N} \tag{39}$$

In the following section the following notations are chosen :

- $(\boldsymbol{X}_t)_{t\in\mathbb{Z}} = (X_t^1, \dots, X_t^p)_{t\in\mathbb{Z}}$ a stochastic vectorial process of size p.
- If $J = \{j_1, \dots, j_q\} \subset \{1, \dots, p\}$ and $\bar{J} = \{1, \dots, p\} \setminus J$, $\boldsymbol{X}_t^J = (\boldsymbol{X}_t^{j_1}, \dots, \boldsymbol{X}_t^{j_q})$ is a process of dimension q
- $\mathbb{X}_{t,t-k} = (\boldsymbol{X}_t, \boldsymbol{X}_{t-1}, \dots, \boldsymbol{X}_{t-k})$ a $p \times (k+1)$ matrix.
- $\Gamma_{s,v}^{i,j} = \mathbf{E}(X_s^i X_v^j)$
- $\Gamma_{\lfloor s,t \rfloor,v}^{i,j} = \{\Gamma_{u,v}^{i,j}, s \leq u \leq t\}$ a vector of dimension $(t-s+2)$ whose generic term is $\Gamma_{u,v}^{i,j}$.

The output process at time t depends on its past instants Y_{t-h} and also the past instant of the input process $\boldsymbol{X}_{t-k}$. Due to this phenomenon of memory, it is not wise to calculate the sensitivity of the instant at time t in relation to the input at time t but to calculate the sensitivity with respect to $\mathbb{X}_{t,t-k}$.

Definition 2. $k-$sensitivity
The $k-$sensitivity *is the Sobol index of* Y_t *with respect to* $\mathbb{X}_{t,t-k}^J$ *for* $0 \leq k < t$.
It is defined by :

$$S_{t,k}^{X^J} = \frac{\mathbf{Var}\left(\mathbf{E}(Y_t | \mathbb{X}_{t,t-k}^J)\right)}{\mathbf{Var}(Y_t)} \tag{40}$$

The index is measured as the ratio of the conditional expectation of Y_t when $(\boldsymbol{X}_t^J, \dots, \boldsymbol{X}_{t-k}^J)$ is fixed on the total variance Y_t.

The $\sigma-$algebra of functions $\sigma(\boldsymbol{X}_t^J, \dots, \boldsymbol{X}_{t-k}^J)$ increases when k increases. So :

$$0 < S_{t,k}^{X^J} < 1 \;\;\text{ and } S_{t,k-1}^{X^J} \leq S_{t,k}^{X^J}$$

The instantaneous sensitivity corresponds to $k = 0$.

Definition 3. *Total sensitivity :*
The total sensitivity is the sensitivity taking into account the whole past of the input $\boldsymbol{X}_t^J$.
It is thus defined as :

$$S_t^{X^J} = \frac{\mathbf{Var}(\mathbf{E}(Y_t | \mathbb{X}_t^J))}{\mathbf{Var}(Y_t)}. \tag{41}$$

It corresponds to $k = 0$.

So, we have :

$$S_t^{X^J} \geq S_{t,k}^{X^J} \quad 0 \leq k \leq t$$

$S_{t,k}^{X^J}$ is an increasing function of k. When k tends towards infinity $S_{t,k}^{X^J}$ converges to $S_t^{X^J}$. In practice we choose k as the value from which the index $S_{t,k}^{X^J}$ does not increase in a significant manner. This heuristic value k is called *useful memory* in terms of sensitivity. So the definition is :

Definition 4. *Let $\varepsilon > 0$ fixed. The $\varepsilon-$useful memory is defined as :*

$$k_\varepsilon = inf\left\{k, \ \ S_{t,h}^{X^J} - S_{t,k}^{X^J} \leq \varepsilon, \ \ h > k\right\} \tag{42}$$

In applications, ε is of course chosen considering the fit quality of the input and also the statistical errors made when estimating $S_k^{X^J}$.

2.2 Stationary case

The input-output system $(\boldsymbol{X}_t, Y_t)$ defines a stochastic process with values in $\mathbb{R}^p \times \mathbb{R}$. We consider now the case where this process $(\boldsymbol{X}_t, Y_t)_{t\in\mathbb{N}}$ is stationary. This implies that $(Y_t)_{t\in\mathbb{N}}$ and $(\boldsymbol{X}_t)_{t\in\mathbb{N}}$ are stationary stochastic processes.

We distinguish two special cases :

- $Y_t = f(\boldsymbol{X}_t, \ldots, \boldsymbol{X}_{t-h})$, h is the memory and $(\boldsymbol{X}_t)_{t\in\mathbb{N}}$ is a stationary system. h is fixed and f non depending on t
- There is a stationary process $Y_t^\star$ (Bernoulli shift process) such that $Y_t^\star = f(\boldsymbol{U}_t, \ldots, \boldsymbol{U}_0, \boldsymbol{U}_{-1}, \ldots)$ is a stationary process and we consider the process associated :
  $Y_t = f_t(\boldsymbol{U}_t, \ldots, \boldsymbol{U}_0) = f(\boldsymbol{U}_t, \ldots, \boldsymbol{U}_0, 0, \ldots)$

  For example :

  $$Y_t = \alpha Y_{t-1} + \sum_{k=1}^{p} \beta_k \boldsymbol{X}_{t-k}$$

  and :

  $$Y_t = \sum_{h=0}^{\infty} \alpha^h \beta_k \boldsymbol{X}_{t-k-h}$$

  and thus it satisfies (39).

For the h memory stationary process, the invariance by translation implies that $S_{t,k}^{X^J}$ is independent of t for every k and thus the total sensitivity $S_t^{X^J}$ is an increasing sequence :

$$\lim_{t\to+\infty} S_t^{X^J} = \lim_{t\to+\infty} \lim_{t\to+k} S_{t,k}^{X^J} \tag{43}$$

For the $Y_t^\star$ process associated to a linear Bernoulli shift we prove in [] that $\lim_{t\to+\infty} S_t^{X^J} = S_\infty^{X^J}$ where $S_\infty^{X^J}$ is a constant. We conjecture that this result is true for a large class of non linear Bernoulli shifts.

2.3 Stochastic Gaussian processes

We can apply to Gaussian processes the Pick and Freeze method introduced in section 1.2.2. To compute for instance $S_{t,k}^{X^J}$ we use the decomposition :

$$\mathbb{X}_{t,t-k}^{\bar{J}} = \Lambda_{\lfloor t-k,t \rfloor, \lfloor t-k,t \rfloor} \mathbb{X}_{t,t-k}^{J} + \mathbb{W}_{t,t-k} \tag{44}$$

with $\Lambda_{t,t}$ given by :

$$\Lambda_{\lfloor t-k,t \rfloor, \lfloor t-k,t \rfloor} = \left(\Gamma_{\lfloor t-k,t \rfloor, \lfloor t-k,t \rfloor}^{JJ}\right)^{-1} \Gamma_{\lfloor t-k,t \rfloor, \lfloor t-k,t \rfloor}^{J\bar{J}} \tag{45}$$

$\Gamma_{\lfloor t-k,t \rfloor, \lfloor t-k,t \rfloor}^{JJ}$ is invertible.
For each t we apply the Pick and Freeze method to :

$$\begin{aligned} Y_t &= f_t(\mathbb{X}_t^J, \Lambda_{\lfloor t-k,t \rfloor, \lfloor t-k,t \rfloor} \mathbb{X}_{t,t-k}^J + \mathbb{W}_t) && (46) \\ &= g_t(\mathbb{X}_t^J, \mathbb{W}_t) && (47) \end{aligned}$$

with $\mathbb{X}_t^J$ and $\mathbb{W}_t$ independent vectors.

2.3.1 Example of toy models for Gaussian inputs

We study two stationary non linear toy models given by:

$$\begin{aligned} Y_t &= 0.5Y_{t-1} + 0.3X_t^1 X_t^2 && (48) \\ Y_t &= X_t^1 X_t^2 - \arctan(X_t^2) && (49) \end{aligned}$$

X_t^1, X_t^2 is a $VAR(1)$ stationary process given by:

$$\begin{pmatrix} X_t^1 \\ X_t^2 \end{pmatrix} = \begin{pmatrix} 0.1 & 0.4 \\ 0.8 & 0.2 \end{pmatrix} \begin{pmatrix} X_{t-1}^1 \\ X_{t-1}^2 \end{pmatrix} + \omega_t \tag{50}$$

where ω_t is a stationary Gaussian noise of covariance matrix $\Theta = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}$.
Indices are estimated with samples of size $N = 10000$. Results are given in figures : 2.3.1 and 2.3.1.

All the indices converge quickly to a constant. The useful memory, that is to say the time when the index does not change significantly, is different according to the model and the variable. It is $k = 4$ for the model (48) and $k = 2$ (variable X^1) or $k = 3$ (variable X^2) for the model (49).

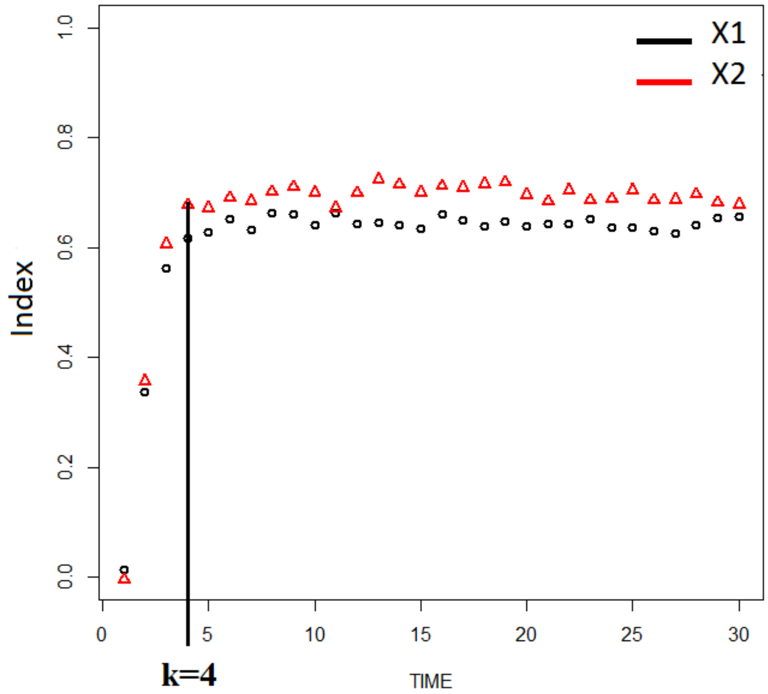

Figure 2: Plot of Sobol indices applied to model (48) in function of k .

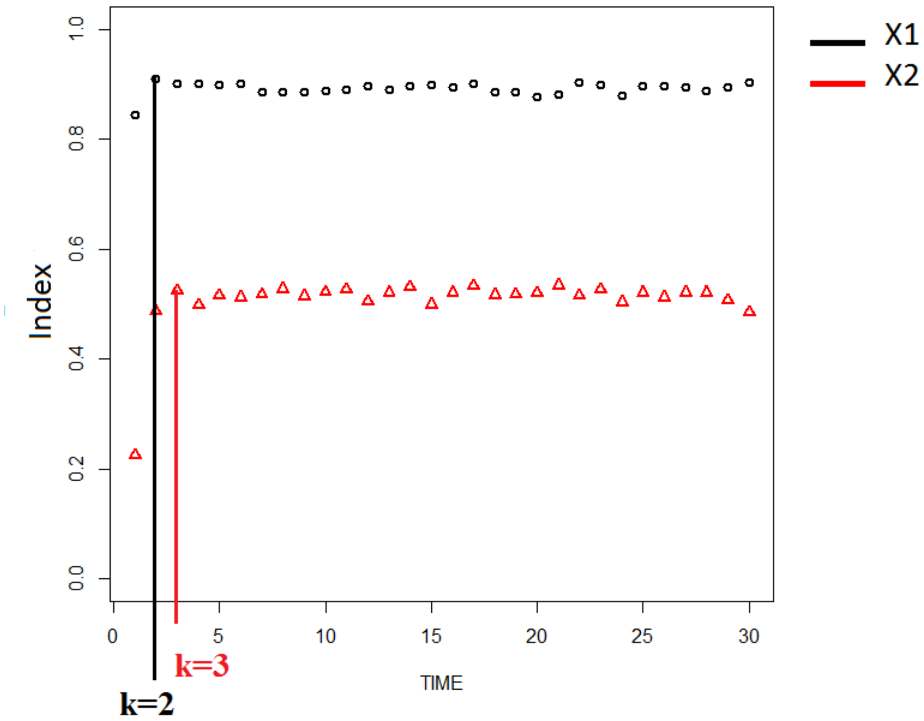

Figure 3: Plot of Sobol indices applied to model (49) in function of k.

2.4 Sensitivity and meta-models for non Gaussian vectorial process inputs

2.4.1 Non Gaussian model

A lot of variables have probability laws that are not Gaussian. For example climatic variables (temperature, wind) heating or energy source variables are usually bounded. If we want to study the impact of extreme cold or even a wave of heat on the indoor temperature, we cannot use Gaussian variables because they have too heavy tails. It is the same for phenomena which present two main values. The density of the variable, in this case, is not Gaussian but bimodal.

Starting from data, the construction of a non Gaussian stochastic process is difficult and all the more in a multivariate context. Thus, as often in these situations, a metamodel must be chosen. It must take into account some of the information which can be extracted from the data and which seem the most important to the practitioner. These informations can be qualitative or quantitative or mixed. For instance these informations concern :

- the marginal distribution of the inputs
- the time dependence structure

For marginal distributions, qualitative information is, for instance, the number of modes (important regimes). The semi-qualitative information is for instance the boundedness of the support of the distribution. The quantitative informations can be the mean, the variance,the skewness, the kurtosis. Information on dependence can be translated in terms of some correlation coefficients or in terms of Markovian properties. Once these properties extracted or estimated from the data we have to choose the input model and to be sure concerning our goals that it allows to compute sensitivities with a quite good approximation. This last point is of course an important constraint to build a metamodel of input, when we want to study sensitivity.

The most classical problem is the following, which can be set in terms of constrained copulas : suppose we want to build a stationary input model $\boldsymbol{X}_t$ with fixed marginals $(F^1, \ldots, F^p)$ and some fixed correlations for instance : $\mathbf{Cor}(\boldsymbol{X}_t, \boldsymbol{X}_t)$ and $\mathbf{Cor}(\boldsymbol{X}_t, \boldsymbol{X}_{t-1})$. These correlations are in fact the correlations estimated with the data. The fixed marginals can be estimated using the data in a parametric family, large enough to take into account qualitative and quantitative properties according to the practitioner's experience on sensitivity.

Correlations are estimated empirically and require less data. Thus we need to choose a family parametrized for instance by the first four moments of the distribution, flexible enough to allow properties as : boundedness, bimodality, light and heavy tails. This is the case of some families such as Pearson or Johnson [16]. Their properties in relation with our work are detailed in the appendix.

Let the correlation matrices R_q defined by :

$$R_q = \mathbf{Cor}(\boldsymbol{X}_t, \boldsymbol{X}_{t-q}) \text{ for } 0 \leq q \leq Q$$

Definition 5. *Let $\boldsymbol{R} = \{R_q, 0 \leq q \leq Q\}$ a correlation matrix given. We say that it is a problem $(\boldsymbol{F}, \boldsymbol{R})$ feasible if there is a stationary stochastic process $\boldsymbol{X}_t$ such that its p marginals are given by $\boldsymbol{F}$ and the $Q+1$ first correlations are given by $\boldsymbol{R}$.*

Until to day there are no complete results on this problem [4]. The most usual way is to try to build a metamodel associated to a Gaussian one (modified to be feasible) and which moreover allows to compute sensitivity indices.

The given information on $\boldsymbol{X}_t$ is $(F^1, \ldots, F^p)$ the p marginal distributions of the stationary process $\boldsymbol{X}_t$ and the correlation matrix $\boldsymbol{R} = \{R_q, 0 \leq q \leq Q\}$ with $R_q = (R_q^{i,j})_{1 \leq i,j \leq p}$.
Let $\boldsymbol{Z}_{t \in \mathbb{Z}}$ a Gaussian stationary process defined by :

$$Z_t^i = (\Phi^{-1} \circ F^i)(X_t^i) \tag{51}$$

We look for the correlation matrix $\Upsilon = (\Upsilon_0, \ldots, \Upsilon_Q)$ of Z_t^i in order that :

$$\Upsilon_q^{i,j} = \mathbf{Cor}\left(((F^i)^{-1} \circ \Phi)(Z_t^i), ((F^j)^{-1} \circ \Phi)(Z_{t-q}^j)\right) = R_q^{i,j} \tag{52}$$

This can be easily done by computing integrals analogous to (31), taking $\Upsilon_q^{i,j}$ instead of $\Upsilon^{i,j}$.

Thus Υ is now fixed. If Υ is a positive definitive matrix, it gives the first Q correlation of the process $\boldsymbol{Z}_t$. We discuss later the case when Υ is not positive definite.

The class of stationary Gaussian $VAR(Q)$ processes can be associated to Υ. This class has the property to allow easy computations of sensitivities by the Pick and Freeze method.

Let

$$\boldsymbol{Z}_t = A_1 \boldsymbol{Z}_{t-1} + \cdots + A_Q \boldsymbol{Z}_{t-Q} + \omega_t$$

with $\mathbf{E}(\omega_t \omega_t^*) = \Theta$, ω_t being a Gaussian white noise.

$A_1, \ldots, A_Q$ can be quite easily computed from $(\Upsilon_0, \ldots, \Upsilon_Q)$ and Θ. $\boldsymbol{Z}_t$ defines a $VAR(Q)$ Gaussian process, which is the $VAR(Q)$ Gaussian copula associated to $\boldsymbol{X}_t$ by :

$$X_t^i = (F^i)^{-1} \circ \Phi(\boldsymbol{Z}_t^i) \tag{53}$$

It may happen that Υ is not a correlation matrix (matrix not positive definite) leading to a stationary process $\boldsymbol{Z}_t$. There are, until to day, only empirical methods ([3],[2]) to overpass this obstacle. The most efficient is to take a smaller Q (in general Q is chosen by an Akaike criterion) and to change R slightly.

Now to compute the sensitivity, we use the following basic facts : $(F^i)^{-1} \circ \Phi$ is a monotone function, for every $(t_1, \ldots, t_q)$ and $(j_1, \ldots, j_q)$ for every q.

Thus we have the equality of $\sigma-$algebra :

$$\sigma(Z_{t_1}^{j_1}, \ldots, Z_{t_n}^{j_q}) = \sigma(X_{t_1}^{j_1}, \ldots, X_{t_n}^{j_q}) \tag{54}$$

Let Y_t the output of the system :

$$Y_t = f(\boldsymbol{X}_t, \boldsymbol{X}_{t-1}, \ldots, \boldsymbol{X}_{t-k})$$

for instance then :

$$\begin{aligned} Y_t &= f(T^{-1}(\boldsymbol{Z}_t), \dots, T^{-1}(\boldsymbol{Z}_{t-k})) \\ &= \tilde{f}(\boldsymbol{Z}_t, \dots, \boldsymbol{Z}_{t-k}) \end{aligned}$$

where $T(\boldsymbol{Z}_t) = \left((F^i)^{-1} \circ \Phi(Z_t^i)\right)_{i=1,\dots,p}$ Thus :

$$\mathbf{E}\left(\mathbf{E}\left(Y_t | \sigma(X_t^1, \dots, X_{t-s}^1)\right)^2\right) = \mathbf{E}\left(\mathbf{E}\left(Y_t | \sigma(Z_t^1, \dots, Z_{t-s}^1)\right)^2\right) \tag{55}$$

for every s.

We can compute : $\mathbf{E}\left(Y_t | \sigma(X_t^1, \dots, X_{t-s}^1)\right)$ using the Pick and Freeze method already defined for the Gaussian process $(\boldsymbol{Z}_t)_{t\in\mathbb{Z}}$ in section 2.3.

Let us give an example.

2.4.2 Example

We study a non linear stationary model given by:

$$Y_t = 0.5Y_{t-1} - 0.2\sin(U_t^2) + 0.2U_t^1 \tag{56}$$

where $\boldsymbol{U}_t$ is a stationary process with Uniform components. We suppose that the correlation matrices are :

$$\begin{aligned} R_0 &= \begin{pmatrix} 1 & R_0^{12} \\ R_0^{12} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0.39 \\ 0.39 & 1 \end{pmatrix} \\ R_1^U &= \begin{pmatrix} 1 & R_1^{12} \\ R_1^{12} & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1.37 \\ 1.37 & 1 \end{pmatrix} \end{aligned}$$

The correlation matrices of the Gaussian process must verify : $\Upsilon_k^{ij} = 2\sin(\frac{\pi R_k^{ij}}{6})$, where Υ is the correlation of the process such as : $Z_t^i = \Phi^{-1}(U_t^i)$

So :

$$\begin{aligned} \Upsilon_0 &= \begin{pmatrix} 1 & 0.41 \\ 0.41 & 1 \end{pmatrix} \\ \Upsilon_1 &= \begin{pmatrix} 1 & 1.32 \\ 1.32 & 1 \end{pmatrix} \end{aligned}$$

One of the corresponding Gaussian processes might be :

$$\boldsymbol{Z}_t = \begin{pmatrix} 0.1 & 0.4 \\ 0.8 & 0.2 \end{pmatrix} \boldsymbol{Z}_{t-1} + \omega_t$$

where ω_t is a Gaussian noise of covariance matrix $\Theta = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}$

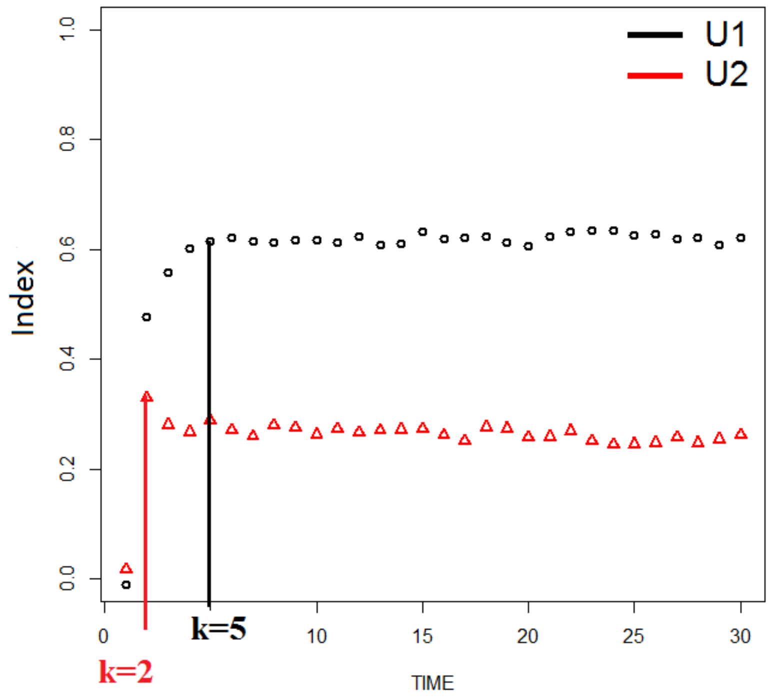

Figure 4: Plot of Sobol indices applied to model (56)

So to apply the Pick and Freeze method, the model used is :

$$Y_t = 0.5Y_{t-1} - 0.2\sin(\Phi(Z_t^2)) + 0.2\Phi(Z_t^1)$$

where Φ is the distribution function and $\boldsymbol{Z}$ a Gaussian process defined as previously. To separate the variables, we use the method developed in section 2.3.

The results are present in figure : 2.

3 Conclusion

We give a general framework for computing sensitivities for dependent inputs. In the static case, for dependent inputs the definition of the sensitivity indices remains the same as for independent inputs. But in the dynamic case, in order to take in consideration the temporal dependence, the definition is slightly modified. We take into account the notion of memory related to sensitivity, "memory" being used in the Physical sense. In the stationary case, the index converges to a constant. The useful memory is the instant when the index does not change.

We study transformations in the input whose image is an input with independent components. These transformations have been detailed. They have the basic property to give the equality of the $\sigma-$algebra between a subset of inputs and their image. This property allows to use the Pick and Freeze method to get the sensitivities.

When the inputs are Gaussian we first consider the static case, then the dynamic case. In both cases the transformation consists in separating the inputs into tow groups of independent variables.

When the inputs are not Gaussian, in the static case we use the conditional quantile functions. They are a nodal point for the sensitivity studies and simulations. This method is the same as the one used to simulate random vectors in general. The output Y takes the canonical form $Y = \tilde{f}(U^1, \ldots, U^p)$ where $(U^1, \ldots, U^p)$ are p uniform independent random variables. This canonical form allows to apply the Pick and Freeze method but also all the more or less classical methods to compute sensitivity starting from Hoeffding formula.
We have to take precautions with the order in which we calculate the index. When we want to calculate the index of each variable we have to start with the variable listed first and then reorder the list and so on for the other variables.
In the dynamic case we use the metamodel copula to go back to the Gaussian case. The metamodel chosen is an extension of the Gaussian copula applied to stochastic processes. The correlations used for the metamodel of the non Gaussian process $\boldsymbol{X}_t$ are those between $\boldsymbol{X}_t$ and $\boldsymbol{X}_{t-1}, \boldsymbol{X}_{t-2}, \ldots$. These correlations define the dynamics of the process. A formula links the correlations of the non Gaussian process and the correlations of the Gaussian process on which we can apply the Pick and Freeze method.
In practical situations the notion of metamodel copula has to be managed carefully. For the Ishigami example we have shown that the correlation used to represent the dependence between variables can be very weak for sensitivity studies. In the case of stochastic process inputs and sensitivity estimation the same caution is required. The specification of inputs becomes in application very important and difficult. Metamodelisation can be improved using some quantitative and qualitative information, essential for the practical problem and which can be extracted from the input data. For instance, for a practitioner, instead of using the complete probability distribution of the inputs, the information can be summarized by the mean, the variance, the skewness and the kurtosis of all the marginal functions of the inputs. This is the case for the copulas with Johnson (or Pearson) distribution.
Finally, we could apply the conditional quantile function method to a dynamic case but if the processes have a too important memory the computation is heavy.

4 Appendix

Processes having marginal distributions from the Johnson translation system are defined by a cumulative distribution function F_X such as:

$$F_X(x) = \Phi(\gamma + \delta f|(x - \xi)/\lambda|) \tag{57}$$

where γ and δ are shape parameters, ξ location parameter, λ a scale parameter and $f(\dot{)}$ is one of the following function :

$$f(y) = \begin{cases} \log(y) & \text{lognormal family} \\ \log(y + \sqrt{y^2+1}) & \text{unbounded law} \\ \log(\dfrac{y}{1-y}) & \text{bounded law} \\ y & \text{normal family} \end{cases} \tag{58}$$

Φ being the Gaussian repartition. $(\gamma, \delta, \xi, \lambda)$ system is equivalent to the mean, variance, skewness, kurtosis one. The maximal number of modes is 2.

Let X a random variables with distribution F and Z a Gaussian normal variable such that :

$$F(X) = \Phi(Z) \tag{59}$$

equality between uniform variables.

If X is a Johnson distribution thus $Z = \gamma + \delta f(\dfrac{X-\xi}{\lambda})$ or $X = \xi + \lambda f^{-1}(\dfrac{Z-\gamma}{\delta})$ well defined for f is a strictly increasing function. These formula are of course simpler than (59).

Thus the construction of the metamodel is done estimating from the data for every j ; $(f_j, \xi_j, \gamma_j, \lambda_j, \delta_j)$. We have at this stage taken into account the main qualitative features of every $F^j, j = 1, \ldots, p$ (maximum likelihood can be the tool for estimation).

References

[1] Bahar Biller and Soumyadip Ghosh. Multivariate input processes. *Handbooks in Operations Research and Management Science*, 13:123–153, 2006.

[2] Bahar Biller and Barry L Nelson. Modeling and generating multivariate time-series input processes using a vector autoregressive technique. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 13(3):211–237, 2003.

[3] Bahar Biller and Barry L Nelson. Fitting time-series input processes for simulation. *Operations Research*, 53(3):549–559, 2005.

[4] Marne C Cario and Barry L Nelson. Autoregressive to anything: Time-series input processes for simulation. *Operations Research Letters*, 19(2):51–58, 1996.

[5] Gael Chastaing, Fabrice Gamboa, and Clementine Prieur. Generalized hoeffding-sobol decomposition for dependent variables-application to sensitivity analysis. *Arxiv preprint arXiv:1112.1788*, 2011.

[6] Gaëlle Chastaing. *Indices de Sobol généralisés pour variables dépendantes*. PhD thesis, Université de Grenoble, 2013.

[7] Gaëlle Chastaing and Loic Le Gratiet. Anova decomposition of conditional gaussian processes for sensitivity analysis with dependent inputs. *Journal of Statistical Computation and Simulation*, 85(11):2164–2186, 2015.

[8] R.I. Cukier, C.M. Fortuin, K.E. Shuler, A.G. Petschek, and J.H. Schaibly. Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. i. theory. *J. Chemical Physics*, 59:3873–3878, 1973.

[9] Sebastien Da Veiga, Francois Wahl, and Fabrice Gamboa. Local polynomial estimation for sensitivity analysis on models with correlated inputs. *Technometrics*, 51(4):452–463, 2009.

[10] Fabrice Gamboa, Alexandre Janon, Thierry Klein, Agnes Lagnoux-Renaudie, and Clémentine Prieur. Statistical inference for sobol pick freeze monte carlo method. *arXiv preprint arXiv:1303.6447*, 2013.

[11] Soumyadip Ghosh and Shane G Henderson. Corrigendum: Behavior of the norta method for correlated random vector generation as the dimension increases. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 16(1):93–94, 2006.

[12] Andreas Habegger, Andreas Stahel, Josef Goette, and Marcel Jacomet. An efficient hardware implementation for a reciprocal unit. In *Electronic Design, Test and Application, 2010. DELTA'10. Fifth IEEE International Symposium on*, pages 183–187. IEEE, 2010.

[13] Giles Hooker. Generalized functional anova diagnostics for high-dimensional functions of dependent variables. *Journal of Computational and Graphical Statistics*, 16(3), 2007.

[14] Alexandre Janon, Thierry Klein, Agnes Lagnoux, Maëlle Nodet, and Clémentine Prieur. Asymptotic normality and efficiency of two sobol index estimators. *ESAIM: Probability and Statistics*, 2013.

[15] Alexandre Janon, Thierry Klein, Agnes Lagnoux, Maëlle Nodet, and Clémentine Prieur. Asymptotic normality and efficiency of two sobol index estimators. *ESAIM: Probability and Statistics*, 18:342–364, 2014.

[16] Norman L Johnson. Bivariate distributions based on simple translation systems. *Biometrika*, pages 297–304, 1949.

[17] Sergei Kucherenko, Stefano Tarantola, and Paola Annoni. Estimation of global sensitivity indices for models with dependent variables. *Computer Physics Communications*, 183(4):937–946, 2012.

[18] Ruixue Liu and Art B Owen. Estimating mean dimensionality of analysis of variance decompositions. *Journal of the American Statistical Association*, 101(474):712–721, 2006.

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

[20] Thierry A Mara and Stefano Tarantola. Variance-based sensitivity indices for models with dependent inputs. *Reliability Engineering & System Safety*, 107:115–121, 2012.

[21] Roger B Nelsen. *An introduction to copulas*, volume 139. Springer Science & Business Media, 2013.

[22] Murray Rosenblatt. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, pages 832–837, 1956.

[23] Andrea Saltelli, Karen Chan, E Marian Scott, et al. *Sensitivity analysis*, volume 1. Wiley New York, 2000.

[24] M Sklar. *Fonctions de répartition à n dimensions et leurs marges*. Université Paris 8, 1959.

[25] Il'ya M. Sobol. Sensitivity estimates for nonlinear mathematical models. *Math. Modeling Comput. Experiment*, 1(4):407–414 (1995), 1993.

[26] Ilya M Sobol. Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates. *Mathematics and computers in simulation*, 55(1-3):271–280, 2001.

[27] Charles J Stone, Mark H Hansen, Charles Kooperberg, Young K Truong, et al. Polynomial splines and their tensor products in extended linear modeling: 1994 wald memorial lecture. *The Annals of Statistics*, 25(4):1371–1470, 1997.

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

[29] Stefano Tarantola, Debora Gatelli, and Thierry Alex Mara. Random balance designs for the estimation of first order global sensitivity indices. *Reliability Engineering & System Safety*, 91(6):717–727, 2006.

[30] Chonggang Xu and George Zdzislaw Gertner. Reliability of global sensitivity indices. *Journal of Statistical Computation and Simulation*, 81(12):1939–1969, 2011.