Proceedings of the 2018 Winter Simulation Conference
M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, eds.

ESTIMATION OF MOMENT INDEPENDENT IMPORTANCE MEASURES USING A COPULA AND MAXIMUM ENTROPY FRAMEWORK

Pierre Derennes

Department IMT
Université de Toulouse
118 Route de Narbonne
Toulouse, 31062, FRANCE

Jérôme Morio

ONERA/DTIS
2 Avenue Edouard Belin
Toulouse, 31055, FRANCE

Florian Simatos

ISAE-SUPAERO
10 Avenue Edouard Belin
Toulouse, 31400, FRANCE

ABSTRACT

The moment-independent sensitivity analysis technique introduced by E. Borgonovo has gained increasing attention to characterize the uncertainty of complex systems and optimize their reliability. The estimation of corresponding indices is a challenging task. This paper aims at presenting a new estimation scheme valid for dependent model inputs. This scheme is build on the copula representation of indices and uses maximum entropy methods to estimate this copula. Accuracy of the proposed method is evaluated through numerical simulations and is compared to two competitive methods, namely an importance sampling based approach and a second one which uses the Nataf transformation.

1 INTRODUCTION

Safety analysis of complex systems is a subject of current interest in various fields such as aerospace, aeronautics, finance or nuclear domain. In this paper, we focus on the case where the system is represented by an input-output model where the observation is expressed as a deterministic function of external parameters assumed to be random. For this kind of model, a practical question is how to identify and rank inputs with respect to their impact on the output. This study is known as the *sensitivity analysis* (SA) and presents two main objectives: decrease the output uncertainty by reducing uncertainty of the most influential inputs, and simplify the model by omitting contribution of least ones. The influence criterion depends on the considered SA approach. There are various SA techniques in literature and essentially two families stand out: local and global SA methods, see Iooss and Lemaître (2015) and associated references for a review. Local methods aim at studying the behavior of the output locally around a nominal value of inputs. In contrast, global methods consider the whole variation range of inputs. Variance-based SA techniques (i.e., Sobol indices) (Sobol' 1990) look at the contribution of inputs to the variance of the output. However, this method focuses on the second-order moment of the output which may poorly reflect the entire variability of the output distribution. To overcome this drawback, Borgonovo (2007) proposed an alternative approach which takes the entire output distribution into account.

However, estimating those indices is a challenging problem because they involve L^1 norms of differences of conditional and unconditional output probability density functions (PDFs). Borgonovo (2007) originally proposed a PDF-based method involving a double loop Monte Carlo estimation combined with a kernel estimation procedure. This method is unfortunately not workable since too many calls to the model are needed. In order to improve this scheme, a pseudo-double loop design involving a partition of the input space is proposed in Plischke et al. (2013). Liu and Homma (2009) proposes to express the δ-sensitivity indices with respect to the unconditional and conditional output cumulative distributions functions (CDFs). However, similarly as the PDF-based method, the CDF-based approach is a double-loop method and the density estimation is replaced by the necessity to find the intersection points of unconditional and conditional PDFs of the output which leads to additional computational time and approximation errors. Zhang et al. (2014) describes a new method combining principle of fractional moment-based maximum entropy and use of Nataf transformation which substantially improves computational burden. However, this technique rests on various technical assumptions such as independence between inputs. To overcome this constraint, Wei et al. (2013) introduces a single-loop Monte Carlo simulation scheme which needs only one set of samples for computing all the δ-sensitivity indices. Derennes et al. (2018) shows via simulation that this method may be inaccurate and proposes a new estimation scheme which greatly improves accuracy of the single-loop and combines importance sampling and kernel estimation procedures. Nevertheless, estimates obtained with this last method may still be inaccurate, mostly due to the kernel approximation of the joint density of the output and the considered input.

The present paper proposes a new estimation scheme. The starting point is the observation made by Wei et al. (2014) that the moment independent indices can be expressed in terms of copula which paves the way for a simple Monte Carlo estimation. This approach needs an estimation of the copula which is performed in this article by maximum entropy method using fractional moments constraints.

The rest of this article is organized as follows. Section 2 briefly reviews the δ-sensitivity measures. Computational steps of the proposed scheme are described in Section 3.3. Some numerical examples are considered in Section 4 in order to highlight the gain of accuracy of the proposed method in comparison to the nonparametric importance sampling design (Derennes et al. 2018) and the method of Zhang et al. (2014).

2 BORGONOVO'S MOMENT INDEPENDENT IMPORTANCE MEASURES

We consider throughout a general input-output model $Y = \mathscr{M}(\mathbf{X})$ where the output Y depends on a d-dimensional real valued random variable $\mathbf{X} = (X_1, \dots, X_d)$ through a deterministic scalar function $\mathscr{M} : \mathbb{R}^d \longrightarrow \mathbb{R}$ called "black box". It is assumed throughout that for every $i \in \{1, \dots, d\}$, the pair (X_i, Y) is absolutely continuous with respect to the Lebesgue measure with PDF $f_{X_i,Y}$. This implies in particular that random variables X_i, Y and Y conditioned on $X_i = x_i$ for any $i \in \{1, \dots, d\}$ and $x_i \in \mathbb{R}$ are also absolutely continuous with respect to the Lebesgue measure, and we will denote by f_{X_i}, f_Y and $f_Y^{X_i=x_i}$ their respective PDFs.

The moment independent SA method introduced by Borgonovo (2007) focuses on finding inputs that, if fixed at their distribution ranges, will lead to the most significant modification of the entire output distribution. This difference between conditional and unconditional model output densities $f_Y^{X_i=x_i}$ and f_Y is quantified by the *shift* $s(x_i)$ defined as their L^1 distance which measures the area enclosed between their representative curve:

$$s(x_i) = \left\| f_Y - f_Y^{X_i=x_i} \right\|_{L^1(\mathbb{R})} = \int \left| f_Y(y) - f_Y^{X_i=x_i}(y) \right| \mathrm{d}y. \tag{1}$$

So as to consider the whole range of values the random variable X_i can take, the sensitivity of the output Y with respect to input X_i is defined as the normalized expectation of the shift over X_i, i.e., the δ-sensitivity measure is given by

$$\delta_i := \frac{1}{2} \mathbb{E}\left[s(X_i) \right]. \tag{2}$$

Owing to its convenient advantages, this importance measure has recently attracted attention of practitioners. Firstly, it is monotonic transformation invariant, meaning that δ_i equals to the δ-sensitivity measure of the model $\tilde{Y} := \varphi \circ \mathscr{M}(\mathbf{X})$ for any C^1 diffeomorphism φ. Moreover, this SA technique is suitable for the case of dependent inputs and no assumptions on the model are necessary, in particular the function $\mathscr{M}$ may be nonlinear. Finally, this approach does not focus on a particular moment as the variance-based SA methods that consider only the second-order moment which is not always sufficient to represent the entire variability of the output distribution.

One can mention that the definition (2) can be generalized to a strict group of inputs $\mathbf{X}_I = (X_i, i \in I)$ with $I \subset \{1,\dots,d\}$ by

$$\delta_I := \frac{1}{2}\mathbb{E}\left[s(\mathbf{X}_I)\right] \text{ with } s(\mathbf{x}_I) = \left\| f_Y - f_Y^{\mathbf{X}_I=\mathbf{x}_I} \right\|_{L^1(\mathbb{R})},$$

assuming that the pair $(\mathbf{X}_I, Y)$ is absolutely continuous. Throughout this paper, we restrict our attention to the case of the first-order indices δ_i, but all our results can be generalized to the higher order indices.

3 PROPOSED APPROACH FOR ESTIMATING THE δ_i INDICES

The δ-sensitivity measure δ_i may be reinterpreted using copula framework (Wei et al. 2014):

$$\delta_i = \frac{1}{2}\int_0^1 \int_0^1 |c(u,v) - 1|\, du dv, \tag{3}$$

where c is the PDF of the couple $(F_{X_i}(X_i), F_Y(Y))$. This representation suggests to estimate δ_i by classical Monte Carlo methods. Wei et al. (2014) raised the issue that "*the main drawback of this method is that the accuracy of the estimated delta indices depends on the precision of the copula density estimates*". However, there are various methods available in literature for estimating nonparametric bivariate copula densities. In Wei et al. (2014), authors propose a Gaussian kernel estimator using a diffusion method provided in Botev et al. (2010), which may perform poorly when data are bounded. To prevent this bias induced by the boundary, one may consider a *local linear version* of a bounded support kernel (Chen and Huang 2007). Also, a thresholding estimation using wavelet method framework is proposed in Autin et al. (2010). However, authors point out that this method may be inaccurate when the copula density presents strong peaks at the corner. An alternative to using the kernel estimation for approximating nonparametric bivariate copula densities is to use maximum entropy framework (AghaKouchak 2014; Hao and Singh 2015; Butucea et al. 2015; Hao and Singh 2013).

3.1 Maximum Entropy Estimation of the Copula

Maximum entropy principle, introduced by Jaynes (1957), makes it possible to estimate the PDF $f_{\mathbf{Z}}$ of an n-dimensional real valued random variable $\mathbf{Z}$ by the PDF that somehow bears the largest uncertainty given *available information* on $\mathbf{Z}$. The measure of uncertainty of $\mathbf{Z}$ is defined by the *Shannon entropy* (also called differential entropy) of the PDF $f_{\mathbf{Z}}$:

$$H(f_{\mathbf{Z}}) = -\int_{\mathbb{R}^n} f_{\mathbf{Z}}(\mathbf{z}) \ln(f_{\mathbf{Z}}(\mathbf{z}))\, d\mathbf{z}. \tag{4}$$

In addition, the information available on the sought density is of the form $\mathbb{E}\left[g(\mathbf{Z})\right] = \mathbf{b} \in \mathbb{R}^m$, where g is a given mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ with m the number of constraints. For instance, one may know the first two moments μ_1 and μ_2 and the support S of $\mathbf{Z}$, corresponding to $g(z) = (z, z^2, 1(z \in S))$ and $\mathbf{b} = (\mu_1, \mu_2, 1)$. Then, the *Maximum Entropy* (ME) *estimator* $\hat{f}_{\mathbf{Z}}$ of $f_{\mathbf{Z}}$ is defined as a solution of the following optimization problem

$$\begin{cases} \hat{f}_{\mathbf{Z}} = \arg\max\ H(f) \\ \text{s.t. } \int g(\mathbf{z}) f(\mathbf{z}) d\mathbf{z} = \mathbf{b} \text{ and } f \in L^1(S, \mathbb{R}^+) \end{cases}. \tag{5}$$

It is well known that (5) is a convex optimization problem that may be reformulated by using Lagrange Multipliers, see for instance Boyd and Vandenberghe (2004) or Kapur and Kesavan (1992). Indeed, strong duality holds, so that a feasible point of the maximum entropy problem (5) is the distribution

$$x \mapsto c_0(\Lambda^*)\mathbb{1}_S(\mathbf{z})e^{-\langle \Lambda^*, g(\mathbf{z})\rangle},$$

where $c_0(\Lambda^*)$ is a normalization constant and where Λ^* is a feasible solution of the dual optimization problem (here and elsewhere, $\langle \cdot, \cdot \rangle$ denotes the inner product)

$$\Lambda^* = \underset{\Lambda \in \mathbb{R}^m}{\text{Argmin}} \ \langle \Lambda, \mathbf{b} \rangle + \log \int e^{-\langle \Lambda, g(\mathbf{z})\rangle} d\mathbf{z}. \tag{6}$$

3.2 Choice of Constraints

As a density of the pair $(F_{X_i}(X_i), F_Y(Y))$, the density copula c appearing in (3) may be estimated by an ME estimator. Let us denote marginals of the copula by $U_i := F_{X_i}(X_i)$ and $V := F_Y(Y)$. In AghaKouchak (2014), Hao and Singh (2015) and Hao and Singh (2013), it is proposed to consider the support $S = [0,1]^2$ and constraints associated to the mapping

$$g(u,v) = (u, u^2, \dots, u^m, v, v^2, \dots, v^m, uv). \tag{7}$$

This corresponds to specifying the first m integer moments of the marginals U_i and V and also the moment of U_iV. In the present paper, we consider constraints associated to the mapping

$$g(u,v) = (u^{\alpha_k} v^{\alpha_l}, k, l = 1, \dots, m), \tag{8}$$

where $\alpha_1 < \cdots < \alpha_m$: this corresponds to a collection of m^2 *fractional moments* of the product U_iY. The reason for considering fractional moments rests on a numerical study of Zhang and Pandey (2013) showing that they may provide better estimates than integer moments.

In the following, we will denote by $\mu_{k,l}$ the expectation of $U_i^{\alpha_k} V^{\alpha_l}$ so that constraints in (5) will read $\int g(\mathbf{z}) f(\mathbf{z}) d\mathbf{z} = (\mu_{k,l})$. Since the law of a positive real valued random variable is uniquely characterized by an infinite sequence of fractional moments (Lin 1992), we may think that as m increases, the ME estimator should converge to the true distribution. In dimension one this was proved in Novi Inverardi and Tagliani (2003). It is reasonable to think that it should continue to hold in higher dimension although, to our knowledge, no proof is available. This expected result provides the (informal) justification of our method.

In our framework, the distribution of U_iV is unknown and so its fractional moments $\mu_{k,l}$ also. The constraints $\{\mu_{k,l}\}$ thus need to be estimated, which will be done by Monte Carlo methods from a sample $\{(U^k, V^k)\}_{k=1}^N$. This leads to constraints of the form $\int g(\mathbf{z}) f(\mathbf{z}) d\mathbf{z} = (\hat{\mu}_{k,l})$, see below for more details and Dudík et al. (2004) for relaxed optimization problems in this case of estimated constraints.

Note that in contrast to (7), we do not impose constraints on marginals U_i and V. At first sight, it is tempting to add the constraint $\int u^{\alpha_k} f(u,v) du dv = 1/(\alpha_k + 1)$ since U_i is uniformly distributed. However, we found numerically that fixing marginal constraints to theoretical values provides less accurate estimates showing high variability. We believe that this is due to an incompatibility with the empirical distribution $\frac{1}{N}\sum_{k=1}^N \delta_{U_i^k}$ which is not uniformly distributed. This further suggests to add an estimated constraint $\int u^{\alpha_k} f(u,v) du dv = \hat{\mu}_k = \frac{1}{N}\sum_k U_i^k$ but we found that this does not improve accuracy of the final estimator.

With this choice of estimated constraints, the dual problem (6) corresponds to minimization of the following function:

$$\Gamma(\Lambda) = \sum_{k=1}^m \sum_{l=1}^m \lambda_{k,l} \hat{\mu}_{k,l} + \log \left(\int_0^1 \int_0^1 \exp\left(- \sum_{k=1}^m \sum_{l=1}^m \lambda_{k,l} u^{\alpha_k} v^{\alpha_l} \right) du dv \right) \ , \ \Lambda = (\lambda_{k,l}) \in \mathbb{R}^{m^2}. \tag{9}$$

Then Γ is strictly convex on the set of feasible points and so admits a unique minimizer Λ^* which can been found using standard convex optimization techniques (Boyd and Vandenberghe 2004).

3.3 Implementation Steps of the Proposed Estimation Scheme

We can now present our estimation scheme of the δ-sensitivity measure δ_i:

Step 1. Generate $(\mathbf{X}^1,\dots,\mathbf{X}^N)$ i.i.d. with common distribution $\mathbf{X}$, and then obtain N observations of the model by $Y^n = \mathscr{M}(\mathbf{X}^n) = \mathscr{M}(X_1^n,\dots,X_d^n)$ for $n = 1,\dots,N$.

Step 2 - Estimation of the constraints. Consider m real numbers $\alpha_1 < \cdots < \alpha_m$ and then approximate moments $\{\mu_{k,l} = \mathbb{E}[U_i^{\alpha_k} V^{\alpha_l}]\}_{k,l=1}^m$ by Monte Carlo:

$$\mu_{k,l} \approx \hat{\mu}_{k,l} = \frac{1}{N}\sum_{n=1}^{N} \left(F_{X_i}(X_i^n)\right)^{\alpha_k} \left(\hat{F}_Y(Y^n)\right)^{\alpha_l},\ k,l = 1,\dots,m,$$

where $\hat{F}_Y$ is the empirical CDF of the output CDF F_Y.

Step 3 - Computation of Lagrange multipliers. Using interior-point algorithm, find the minimizer $\hat{\Lambda}^* = (\hat{\lambda}_{k,l}^*)$ of the following function:

$$\hat{\Gamma}(\Lambda) = \sum_{k=1}^{m}\sum_{l=1}^{m} \lambda_{k,l}\hat{\mu}_{k,l} + \log\left(\int_0^1\int_0^1 \exp\left(-\sum_{k=1}^{m}\sum_{l=1}^{m}\lambda_{k,l}u^{\alpha_k}v^{\alpha_l}\right)dudv\right)\ ,\ \Lambda = (\lambda_{k,l}) \in \mathbb{R}^{m^2}. \tag{10}$$

Step 4. Generate $\{(U_1^k, U_2^k)\}_{k=1}^{N'} \overset{\text{i.i.d.}}{\sim} U([0,1]^2)$ and estimate δ_i by:

$$\hat{\delta}_i^{\mathrm{ME}} = \frac{1}{2N'}\sum_{k=1}^{N'} \left|\hat{c}_m(U_1^k, U_2^k) - 1\right|, \tag{11}$$

where

$$\hat{c}_m(u,v) = \mathbb{1}_{[0,1]^2}(u,v)c_0(\hat{\Lambda}^*)\exp\left(-\sum_{k=1}^{m}\sum_{l=1}^{m}\hat{\lambda}_{k,l}^* u^{\alpha_k}v^{\alpha_l}\right), \tag{12}$$

is the estimation of c obtained by solving (5), with $c_0(\hat{\Lambda}^*)$ the normalization constant.

It has to be noted that only N evaluations of the black-box function $\mathscr{M}$ are needed for computing all the δ-sensitivity indices. In addition, the parameter N' of **Step 4.** can be chosen as large as desired without computational cost. Finally, as mentioned above this method does not assume independent inputs.

4 NUMERICAL EXAMPLES

In this section, several numerical examples are considered in order to illustrate advantages of the proposed estimation scheme. Comparisons are drawn with estimates obtained with the nonparametric importance sampling approach (Derennes et al. 2018) and the approach of Zhang et al. (2014). The first method greatly improves accuracy of the single-loop design introduced in Wei et al. (2013). The second method is very competitive since it necessitates only very few calls to model. For instance, Zhang et al. (2014) obtain accurate estimates on a Gaussian linear model with only 25 evaluations of the output function $\mathscr{M}$.

4.1 Nonparametric Importance Sampling Approach

The main idea behind the first method is to rewrite the δ-sensitivity measure in the following way:

$$\delta_i = \frac{1}{2}\left\|f_{X_i}f_Y - f_{X_i,Y}\right\|_{L^1(\mathbb{R}^2)}.$$

Then, δ_i can be estimated by the importance sampling estimator

$$\hat{\delta}_i^{\mathrm{IS}} = \frac{1}{2N'} \sum_{k=1}^{N'} \frac{\left| f_{X_i}(U_1^k) \hat{f}_Y(U_2^k) - \hat{f}_{X_i,Y}(\mathbf{U}^k) \right|}{h(\mathbf{U}^k)} \ , \text{ with } \mathbf{U}^k = (U_1^k, U_2^k) \overset{\text{i.i.d.}}{\sim} h,$$

where h is the sampling distribution chosen by the user and where $\hat{f}_Y$ and $\hat{f}_{X_i,Y}$ are the Gaussian kernel estimators of the output PDF f_Y and the joint PDF $f_{X_i,Y}$. In all our numerical studies, numerical values for this method are taken from Derennes et al. (2018).

4.2 Approach via Nataf Transformation

In the second method, the joint PDF $f_{X_i,Y}$ is estimated using the Nataf transformation. The Nataf transformation of (X_i, Y) is defined by $\mathbf{R}_i = (r_i(X_i), r_y(Y)) = \left(\Phi^{-1}(F_{X_i}(X_i)), \Phi^{-1}(F_Y(Y))\right)$ where Φ denotes the CDF of the standard Gaussian distribution $N(0,1)$. By construction, the marginals r_{X_i} and r_Y follow the standard Gaussian distribution $N(0,1)$ but without additional hypotheses, $\mathbf{R}_i$ is not in general a Gaussian vector. In Zhang et al. (2014) the following assumptions are enforced:

(A0) Inputs X_i are independent;
(A1) The following relation is verified:

$$f_{X_i,Y}(x,y) = f_{X_i}(x) f_Y(y) \frac{\phi_2(r_i(x), r_y(y))}{\phi(r_i(x))\phi(r_y(y))},$$

where ϕ and ϕ_2 are respective PDFs of distributions $N(0,1)$ and $N_2(0_{\mathbb{R}^2}, \rho_{0,i})$ where $\rho_{0,i}$ is the correlation matrix of the couple $(r_i(X_i), r_y(Y))$;
(A2) The model $\mathscr{M}$ is well approximated using cut-HDMR expansion (Zhang and Pandey 2013).

Actually, the Assumption (A1) is quite strong as it is equivalent to assuming that $\mathbf{R}_i$ is a Gaussian vector with correlation matrix $\rho_{0,i}$. It is also equivalent to approximating the copula density of (X_i, Y) by

$$c(u,v) \approx \frac{\phi_2(\Phi^{-1}(u), \Phi^{-1}(v))}{\phi(\Phi^{-1}(u))\phi(\Phi^{-1}(v))}. \tag{13}$$

Their estimator is then defined as

$$\hat{\delta}_i^{\mathrm{Nataf}} = \frac{1}{2N'} \sum_{k=1}^{N'} \left| \frac{\phi(r_i^k)\phi(r_y^k)}{\phi_2(r_i^k, r_y^k)} - 1 \right| \ , \text{ where } (r_i^k, r_y^k) \overset{\text{i.i.d.}}{\sim} N_2(0, \hat{\rho}_{0,i}),$$

where $\hat{\rho}_{0,i}$ is an approximation of $\rho_{0,i}$. In Zhang et al. (2014), the matrix $\rho_{0,i}$ is estimated from resolution of a nonlinear equation which connects $\rho_{0,i}$ with the correlation matrix of (X_i, Y) and from maximum entropy estimation of the output PDF f_Y associated to moment constraints approximated using the assumption (A2). For further details on implementation steps of these methods, we refer the reader to the associated references.

In the first example that we will consider, Assumptions (A0), (A1) and (A2) will be satisfied and this scheme yields very good results. However, we find that performance of this method deteriorates in our second and third examples where these assumptions are violated, whereas performance of our estimation scheme remains the same.

In the subsequent numerical studies, this method only makes very few calls to the model. This is due to the fact that the assumption (A2) allows to reduce computation of moment constraints used for the estimation of the output PDF to the deterministic numerical evaluation of one-dimensional integrals. Note finally that in this approach, the number of calls to the model is not tunable since it is automatically fixed by the numerical procedure used in the estimation of $\rho_{0,i}$. This is the reason why this method will be compared with the other two, but with a different simulation budget.

4.3 Performance Evaluation

An indicator of efficiency of an estimator $\tilde{\delta}_i$ of the importance measure δ_i is the coefficient of variation (cv)

$$\mathrm{cv}(\tilde{\delta}_i) = \frac{\sqrt{\mathbb{V}\mathrm{ar}(\tilde{\delta}_i)}}{\mathbb{E}\left(\tilde{\delta}_i\right)}.$$

For each estimator $\tilde{\delta}_i$ of the importance measure δ_i, we approximate its mean and its standard deviation using Monte Carlo methods. Considering M estimates $(\tilde{\delta}_i^1, \ldots, \tilde{\delta}_i^M)$, we compute the respective estimators of the mean and the standard deviation:

$$\bar{\delta}_i = \frac{1}{M}\sum_{k=1}^{M} \tilde{\delta}_i^k \quad \text{and} \quad \bar{\sigma}_{\tilde{\delta}_i} := \sqrt{\frac{1}{M-1}\sum_{k=1}^{M} (\tilde{\delta}_i^k - \bar{\delta}_i)^2}.$$

When unconditional and conditional output distributions are known, theoretical values of importance measures are available using numerical integration. In this case, the relative difference (RD)

$$\frac{\bar{\delta}_i - \delta_i}{\delta_i},$$

may be computed in order to appreciate the error of the estimator $\tilde{\delta}_i$.

4.4 Example 1: a Toy Case

In this section, the following toy case model output is considered

$$Y_1 = \prod_{i=1}^{4} X_i,$$

where the $d = 4$ inputs X_i are i.i.d. with common distribution the lognormal distribution $L(0,1)$. By symmetry each input has the same influence on the output Y. Thus, we only look at the importance measure δ_1 of the first input X_1. In the current model, unconditional and conditional output distributions are known and given by

$$Y_1 \sim L(0,4) \ ; \ (Y_1 | X_1 = x) \sim L(\ln(x), 3),$$

and theoretical value of δ_1 is available by performing numerical integration.

The three considered methods are applied on this model output. Respective budgets and computation times for computing δ_1 are $N = 5{,}000$ and 10s for $\hat{\delta}_1^{\mathrm{IS}}$, $N = 21$ and 62s for $\hat{\delta}_1^{\mathrm{Nataf}}$ and $N = 5{,}000$ and 8s for $\hat{\delta}_1^{\mathrm{ME}}$. Results are gathered in Table 1.

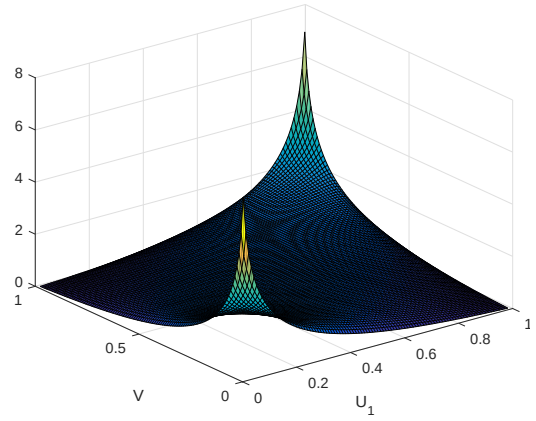
Table 1: Estimates of the δ_i indices of the example 1.

Input	**Theoretical**	$\hat{\delta}_1^{\mathrm{IS}}$			$\hat{\delta}_1^{\mathrm{Nataf}}$			$\hat{\delta}_1^{\mathrm{ME}}$		
	value δ_1	Mean	cv	RD	Mean	cv	RD	Mean	cv	RD
X_1	0.1846	0.4016	0.1017	1.1757	0.1946	0.0032	0.0541	0.1864	0.0277	0.0096

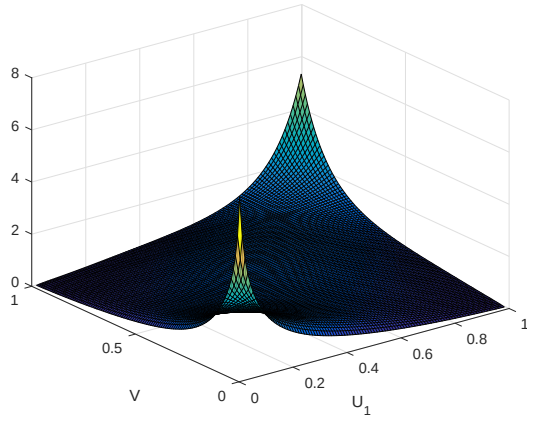
From Table 1, one can see that all considered methods converge since estimates are obtained with a coefficient of variation lower than 10%. Estimates of $\hat{\delta}_1^{\mathrm{IS}}$ show non satisfactory precision with relative differences greater than 100%. This inaccuracy is caused by the kernel estimation step, probably due to involvement of the lognormal distribution which is heavy-tailed (Derennes et al. 2018). On the other hand,

estimates of $\hat{\delta}_1^{\text{Nataf}}$ present a very low variability and a good precision with relative difference around 5%, which is all the more impressive as only 21 calls to the model are done. Estimates of $\hat{\delta}_1^{\text{ME}}$ are very close to the theoretical value of δ_1 with a relative difference around 1% and obtained with a relative standard deviation of the order of 2%. Then, accuracy of estimates is improved with a factor 5 by adopting the approach exposed in the present paper. In addition, it can be seen on Figure 1(b) that the copula of (X_1, Y_1) (see Figure 1(a)) is well approximated by the ME-based estimator $\hat{c}_m$ in comparison to the Gaussian kernel estimator (see Figure 1(c)).

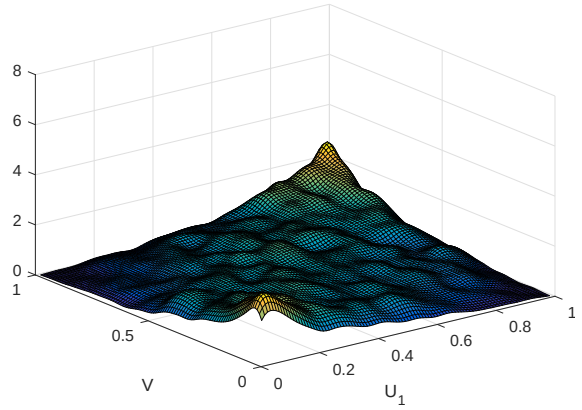
In the present example, inputs are independent and it can be seen that the output model is equal to its cut-HDMR approximation. Furthermore, one can see that the true density copula of (X_1, Y_1) (Fig 1(a)) is close to the approximation defined by Eq.(13) (see Figure 1(d)) which seems to show that the assumption $(A1)$ is licit. Thus, the proposed method presents an higher accuracy than the Nataf transformation based method on this use case for which all the assumptions (A0), (A1) and (A2) are verified. However, this must be put into perspective by the fact that the budget of the Nataf transformation based method is drastically lower than the budget of our approach in this use case.



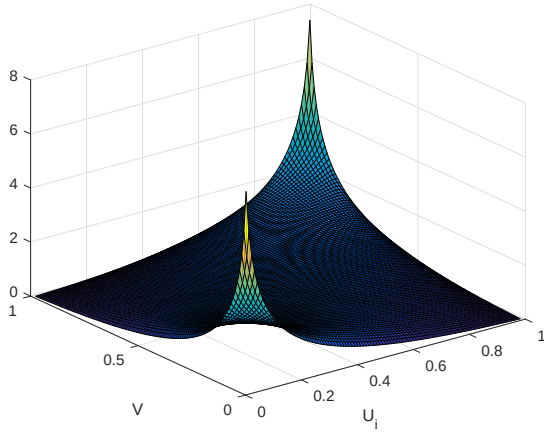
(a) Representation of the theoretical copula c of (X_1, Y_1).



(b) Representation of the ME estimator copula $\hat{c}_m$ obtained by setting parameters $N = 10^4$ and $\alpha_k \in \{\frac{2k}{3}\}_{k=1}^{3}$ $(m = 3)$.



(c) Representation of the Gaussian kernel estimator obtained with a sample of size $N = 10^4$.



(d) Representation of the approximated density of (U_1, V) defined by Eq.(13) (this figure is obtained assuming that the output distribution is known).

Figure 1: Display of the copula of (X_1, Y_1) and its estimates.

4.5 Example 2: a Risk Assessment Model

In this part we consider the probabilistic risk assessment model introduced in Iman (1987). Here, the probability of the top event is written as

$$Y_2 = X_1X_3X_5 + X_1X_3X_6 + X_1X_4X_5 + X_1X_4X_6 + X_2X_3X_4 + X_2X_3X_5 + X_2X_4X_5 + X_2X_5X_6 + X_2X_4X_7 + X_2X_6X_7,$$

where all inputs X_i are independent random variables following lognormal distributions, and theirs parameters (mean and standard deviation) are listed in Table 2.

The three considered methods are applied on this model output. Respective budgets and computation times are $N = 5,000$ and 41s for the $\hat{\delta}_i^{\text{IS}}$'s, $N = 36$ and 140s for the $\hat{\delta}_i^{\text{Nataf}}$'s and $N = 5,000$ and 56s for the $\hat{\delta}_i^{\text{ME}}$'s. Results are gathered in Table 3.

Table 2: Distribution of input variables of the example 2.

Input	**Distribution**	**Mean of** $\ln(X_i)$	**Variance of** $\ln(X_i)$
X_1	Lognormal	0.6044	0.1776
X_2	Lognormal	1.0098	0.1776
X_3	Lognormal	-6.9965	0.1776
X_4	Lognormal	-6.3034	0.1776
X_5	Lognormal	-5.6103	0.1776
X_6	Lognormal	-5.3871	0.1776
X_7	Lognormal	-5.89792	0.1776

Table 3: Estimates of indices of the example 2.

Input	$\hat{\delta}_i^{\text{IS}}$			$\hat{\delta}_i^{\text{Nataf}}$			$\hat{\delta}_i^{\text{ME}}$		
	Mean	STD (10^{-3})	cv	Mean	STD (10^{-3})	cv	Mean	STD (10^{-3})	cv
X_1	0.0828	5.2	0.0628	0.0570	0.0677	0.0012	0.0730	5.9	0.0806
X_2	0.2237	5.3	0.0237	0.1866	0.5943	0.0032	0.2328	4.9	0.0209
X_3	0.0647	4.6	0.0711	0.0379	0.0527	0.0014	0.0486	5.1	0.1050
X_4	0.1137	5.3	0.0466	0.0893	0.1269	0.0014	0.1122	6.1	0.0548
X_5	0.1532	6.3	0.0411	0.1302	0.2413	0.0019	0.1580	5.9	0.0377
X_6	0.1751	5.6	0.0320	0.1482	0.2840	0.0019	0.1809	5.6	0.0312
X_7	0.0855	5.1	0.0596	0.0665	0.0912	0.0014	0.0776	5.5	0.0713

From Table 3, one can see that the three considered methods converge since estimates are obtained with a coefficient of variation lower than 10%. In addition, coefficients of variation obtained with the Nataf transformation based method are very low, which is due to the deterministic evaluation of constraints used in the maximum entropy estimation of the output PDF. All the considered methods lead to the same importance ranking, mainly $X_2 > X_6 > X_5 > X_4 > X_7 > X_1 > X_3$. Nevertheless, one can point out that the importance sampling based approach and the proposed one provide approximately the same estimates than those provided by the PDF double loop method (Liu and Homma 2009). In comparison, estimates of ($\hat{\delta}_i^{\text{Nataf}}$) show noticeable differences.

4.6 Example 3: a Gaussian Linear Model

In this subsection we assume that the expression of the model output is defined in the following way:

$$Y_3 = \mathbf{AX},$$

where $\mathbf{A} = [1.7\ 1.8\ 1.9\ 2]$ and where the input $\mathbf{X}$ follows the Gaussian distribution $N(0,\Sigma)$ where the covariance matrix $\Sigma = (\Sigma_{ij})$ is defined as follows:

$$\Sigma = \begin{pmatrix} 1 & 1/2 & 1/3 & 1/4 \\ 1/2 & 1 & 1/2 & 1/3 \\ 1/3 & 1/2 & 1 & 1/2 \\ 1/4 & 1/3 & 1/2 & 1 \end{pmatrix}.$$

Classical results on Gaussian vectors enable to determine unconditional and conditional output distributions:

$$Y_3 \sim N(0, \mathbf{A}\Sigma\mathbf{A}^T) \text{ and } Y_3|X_i = x_i \sim N\left(m_i, \sigma_i^2\right)\ ,$$

where the mean m_i and the variance σ_i^2 are given by:

$$m_i = A_i x_i + \mathbf{A}_{-i}\mathbf{C}_i\Sigma_{ii}^{-1}x_i \text{ and } \sigma_i^2 = \mathbf{A}_{-i}(\Sigma_{-i} - \mathbf{C}_i\mathbf{C}_i^T\Sigma_{ii}^{-1})\mathbf{A}_{-i}^T,$$

where $\mathbf{A}_{-i}$ is the vector $\mathbf{A}$ private of its i-th component A_i, Σ_{-i} is the matrix Σ private of its i-th row and column and $\mathbf{C}_i$ is the column vector $[\Sigma_{ij}]_{j\neq i}$. Thus, theoretical values of indices are known and can be computed using numerical integration.

We compare those reference values with the estimates obtained with $M = 100$ runs of the three considered methods. Results are displayed in Table 4. Respective budgets and computation times for computing all the δ indices are $N = 5,000$ and 54s for the $\hat{\delta}_i^{\text{IS}}$'s, $N = 21$ and 65s for the $\hat{\delta}_i^{\text{Nataf}}$'s and $N = 5,000$ and 57s for the $\hat{\delta}_i^{\text{ME}}$'s.

The importance sampling based method and the proposed one provide the good ranking, mainly $X_3 > X_2 > X_4 > X_1$. Furthermore, relative differences of proposed estimates are slightly lower than those of the importance sampling estimates $(\delta_i^{\text{IS}})_{1\leq i\leq 4}$. In contrast, correlation between inputs implies that the estimates obtained with the Nataf transformation show notable inaccuracies with relative differences around 50% and do not respect the true importance ranking.

Table 4: Estimates of indices of the example 3.

Input	**Theoretical value** δ_i	$\hat{\delta}_i^{\text{IS}}$			$\hat{\delta}_i^{\text{Nataf}}$			$\hat{\delta}_i^{\text{ME}}$		
		Mean	cv	RD	Mean	cv	RD	Mean	cv	RD
X_1	0.2857	0.2707	0.0218	-0.0526	0.1654	0.0026	-0.4209	0.2840	0.0174	-0.0058
X_2	0.3620	0.3444	0.0157	-0.0486	0.1777	0.0028	-0.5090	0.3538	0.0138	-0.0225
X_3	0.3792	0.3607	0.0157	-0.0487	0.1906	0.0037	-0.4973	0.3688	0.0121	-0.0274
X_4	0.3176	0.3010	0.0199	-0.0522	0.2042	0.0058	-0.3571	0.3141	0.0180	-0.0110

5 CONCLUSION

This paper presents an estimation scheme for evaluating the δ-sensitivity importance measures which combines the expression of the δ-sensitivity indices in term of copula (Wei et al. 2014) and maximum entropy theory. Several numerical examples are performed in order to illustrate the accuracy and the robustness of the method in comparison to the nonparametric importance sampling approach (Derennes et al. 2018) and the method of reference Zhang et al. (2014) combining maximum entropy theory and Nataf transformation. The method of Zhang et al. (2014) is very competitive since it necessitates very few calls to the model output. However, it is valid only under some strong assumptions on the model output, like the independence of the inputs and particular models, which are not necessary with the proposed method.

For further investigations, it may be interesting to enhance the choice of constraints by optimizing the sequence $\alpha_1, \ldots, \alpha_m$ while minimizing the parameter m. This may increase computational burden but may be negligible when the black box function $\mathscr{M}$ is expensive.

ACKNOWLEDGMENTS

The first author is currently enrolled in a PhD program, funded by Université Toulouse III. His financial support is gratefully acknowledged.

REFERENCES

AghaKouchak, A. 2014. "Entropy–Copula in Hydrology and Climatology". *Journal of Hydrometeorology* 15(6):2176–2189.

Autin, F., E. Le Pennec, and K. Tribouley. 2010. "Thresholding Methods to Estimate Copula Density". *Journal of Multivariate Analysis* 101(1):200–222.

Borgonovo, E. 2007. "A New Uncertainty Importance Measure". *Reliability Engineering & System Safety* 92(6):771–784.

Botev, Z. I., J. F. Grotowski, and D. P. Kroese. 2010. "Kernel Density Estimation via Diffusion". *The Annals of Statistics* 38(5):2916–2957.

Boyd, S., and L. Vandenberghe. 2004. *Convex Optimization*. Cambridge: Cambridge University Press.

Butucea, C., J.-F. Delmas, A. Dutfoy, and R. Fischer. 2015. "Maximum Entropy Copula with Given Diagonal Section". *Journal of Multivariate Analysis* 137:61–81.

Chen, S. X., and T.-M. Huang. 2007. "Nonparametric Estimation of Copula Functions for Dependence Modelling". *Canadian Journal of Statistics* 35(2):265–282.

Derennes, P., J. Morio, and F. Simatos. 2018. "A Nonparametric Importance Sampling Estimator for Moment Independent Importance Measures". *Reliability Engineering & System Safety* (In press).

Dudík, M., S. J. Phillips, and R. E. Schapire. 2004. "Performance Guarantees for Regularized Maximum Entropy Density Estimation". In *Learning Theory*, edited by J. Shawe-Taylor and Y. Singer, 472–486. Berlin, Heidelberg: Springer Berlin Heidelberg.

Hao, Z., and V. P. Singh. 2013. "Modeling Multisite Streamflow Dependence with Maximum Entropy Copula". *Water Resources Research* 49(10):7139–7143.

Hao, Z., and V. P. Singh. 2015. "Integrating Entropy and Copula Theories for Hydrologic Modeling and Analysis". *Entropy* 17(4):2253–2280.

Iman, R. L. 1987. "A Matrix-Based Approach to Uncertainty and Sensitivity Analysis for Fault Trees". *Risk Analysis* 7(1):21–33.

Iooss, B., and P. Lemaître. 2015. "A Review on Global Sensitivity Analysis Methods". In *Uncertainty Management in Simulation-Optimization of Complex Systems: Algorithms and Applications*, edited by G. Dellino and C. Meloni, 101–122. Boston, Massachusetts: Springer US.

Jaynes, E. T. 1957. "Information Theory and Statistical Mechanics". *Physical Review* 106(4):620–630.

Kapur, J. N., and H. K. Kesavan. 1992. *Entropy Optimization Principles with Applications*. Boston, Massachusetts: Academic Press, Inc.

Lin, G. D. 1992. "Characterizations of Distributions via Moments". *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)* 54(1):128–132.

Liu, Q., and T. Homma. 2009. "A New Computational Method of a Moment-Independent Uncertainty Importance Measure". *Reliability Engineering & System Safety* 94(7):1205–1211.

Novi Inverardi, P., and A. Tagliani. 2003. "Maximum Entropy Density Estimation from Fractional Moments". *Communications in Statistics-Theory and Methods* 32(2):327–345.

Plischke, E., E. Borgonovo, and C. L. Smith. 2013. "Global Sensitivity Measures from Given Data". *European Journal of Operational Research* 226(3):536–550.

Sobol', I. M. 1990. "On Sensitivity Estimation for Nonlinear Mathematical Models". *Matematicheskoe Modelirovanie* 2(1):112–118.

Wei, P., Z. Lu, and J. Song. 2014. "Moment-Independent Sensitivity Analysis Using Copula". *Risk Analysis* 34(2):210–222.

Wei, P., Z. Lu, and X. Yuan. 2013. "Monte Carlo Simulation for Moment-Independent Sensitivity Analysis". *Reliability Engineering & System Safety* 110:60–67.

Zhang, L., Z. Lu, L. Cheng, and C. Fan. 2014. "A New Method for Evaluating Borgonovo Moment-Independent Importance Measure with its Application in an Aircraft Structure". *Reliability Engineering & System Safety* 132:163–175.

Zhang, X., and M. D. Pandey. 2013. "Structural Reliability Analysis Based on the Concepts of Entropy, Fractional Moment and Dimensional Reduction Method". *Structural Safety* 43:28–40.

AUTHOR BIOGRAPHIES

PIERRE DERENNES is a second year PhD student of applied mathematics at Universié Paul Sabatier in Toulouse. His home lab is ONERA (Toulouse). He obtained his master's degree at Université Claude-Bernard in Lyon. His e-mail address is pierre.derennes@onera.fr.

JÉRÔME MORIO is a senior scientist at ONERA and Université de Toulouse. He obtained his Ph.D. in physics from Aix-Marseille University, France, in 2007. His main research interests include uncertainty management, rare event probability estimation, sensitivity analysis. His e-mail address is jerome.morio@onera.fr.

FLORIAN SIMATOS is professor in Probability and Statistics at ISAE-SUPAERO and Université de Toulouse. He obtained his Ph.D. from Ecole Polytechnique, France, in 2009. His main research interests include performance analysis of communication networks, branching processes and sensitivity analysis. His e-mail address is florian.simatos@isae.fr.