Safety, Reliability and Risk Analysis: Theory, Methods and Applications – Martorell et al. (eds)
© 2009 Taylor & Francis Group, London, ISBN 978-0-415-48513-5

Global sensitivity analysis based on entropy

Benjamin Auder & Bertrand Iooss
CEA Cadarache, Saint-Paul-lez-Durance, France

ABSTRACT: This paper deals with the sensitivity analysis of model output, using entropy. By the past, variance has been used, and some measures directly related to data distributions have been performed. We expect entropy to give another point of view, refining the other sensitivity analysis, and may be to solve some situations where the other methods were useless. Two different indices based on entropy are presented in this work: those defined by Krzykacz-Hausmann 2001,, based on conditional entropy which use directly the definition of Shannon's entropy, and those later suggested by Liu, Chen, and Sudjianto 2006, based on the Kullbak-Leibler entropy, which measure the difference between two probability distributions. After a brief presentation of the two entropy-based indices, two application cases are studied, one analytical function and one industrial computer code. Our goal is to show that some indices are better suited for some specific cases, and that they all enable different but relevant sensitivity analyses to be performed.

1 INTRODUCTION

Simulation of complex phenomena often leads to physical models which depend on many input variables, and some of them can be redondant or have negligible effects. It is therefore interesting to highlight the most influent factors. In order to do so, we have first to define what an influent factor is, by quantifying its contribution to the model. Our subject here is the global sensitivity analysis: effects of variables are studied in their whole range of variations, in opposition to the restricted scope of local sensitivity analysis.

Nowadays, almost all the sensitivity analysis studies are based on the variance measure (Saltelli, Chan, and Scott 2000), with its well known limitations (Middleton 2005). For example, the variance measure is ill-suited to measure the dispersion of a variable with an heavy-tail or a multimodal distribution, or which contains some outliers (Lejeune 2004). In this paper, we focus on two alternative measures of sensitivity supposed to take into account the situations where variance is not well adapted. First, we study the entropy-based sensitivity analysis of Krzykacz-Hausmann 2001. These indices are well suited in the case of epistemic uncertainty, i.e. when the variables are deterministic but not known exactly. Second, we look at the indices defined by Liu, Chen, and Sudjianto 2006,, who use the relative entropy definition.

The following section of this paper is devoted to the mathematical definitions of the sensitivity indices. Then, we extract two situations to illustrate these definitions. In the fourth section, we give a brief review of the computation algorithms of the sensitivity measures. In a last section, comparisons of the three methods are performed throughout an analytic function and an industrial computer code.

2 SENSITIVITY INDICES

This section focuses on the useful definitions to introduce sensitivity indices.

2.1 *Variance-based sensitivity indices*

Sobol's sensitivity indices are the most popular variance-based indices, introduced in Sobol 1993. Let Y be the output of a physical model, and ψ the function linking the independent input variables $X_1, \ldots, X_s$ to Y: $Y = \psi(X_1, \ldots, X_s)$. ψ is actually an approximation of a more complex physical relation between X and Y. We can now decompose the model variance as follows:

$$\mathrm{Var}(Y) = \sum_{k=1}^{s} \sum_{1 \le i_1 < \ldots < i_k \le s} V_{i_1,\ldots,i_k}(Y)$$

with $V_{i_1,\ldots,i_k}(Y)$ recursively defined by

$$V_{i_1,\ldots,i_k}(Y) = \mathrm{Var}(\mathbb{E}[Y|X_{i_1} \cdots X_{i_k}]) - \sum_{1 \le j_1 < \cdots < j_{k-1} \le s} \mathrm{Var}(\mathbb{E}[Y|X_{j_1} \cdots X_{j_{k-1}}])$$

for $k = 2 \dots s$, and $V_i(Y) = \mathrm{Var}(\mathbb{E}[\mathrm{Y}|\mathrm{X_i}])$.

Sobol's variance-based sensitivity indices can now be defined by:

$$S_{i_1,\dots,i_k} = \frac{V_{i_1,\dots,i_k}(Y)}{\mathrm{Var(Y)}} .$$

First order indices S_i indicates the part of the variability of Y explained by the input X_i. Second order sensitivity indices S_{ij} stand for the Y variance sensitivity to the interaction of couples of variables (X_i,X_j), and so on. Total sensitivity indices have also been defined to take into account all the effects a variable has on the output variance. The definition for an output Y, focusing on the input X_i writes

$$S_{T_i} = \sum_{l=1}^{s} \sum_{1 \le i_1 < \dots < i_k = i < \dots \le i_l} S_{i_1,\dots,i_l} ,$$

In practical computations, when the model has a lot of input variables, we often restrict to estimate the first order and total sensitivity indices.

Finally, these indices are well suited when variability of the output around its mean is meaningful, that is to say when the output distribution is not too much scattered around the space: the Sobol's indices are interesting when this last distribution has only one non-negligible mode.

2.2 *Entropy-based sensitivity analysis*

Entropy is a well known function in the theory of information, which indicates the loss of information within a system - then, by opposition, the amount of information. The *entropy* of a discrete random variable X ranging in $x_1,\dots,x_n$ with respective probabilities $p_1,\dots,p_n$ writes

$$H(X) = -\sum_{k=1}^{n} p_k \ln(p_k) ,$$

with the convention $0\ln(0) = 0$. This quantity does not depend on the values that X takes, but only on their probabilities. This is a considerable difference with the variance, which is calculated by the mean of the gaps to the mean, those gaps depending on the values taken by the random variable. We limit our study to the discrete case, the only one arisen in computer applications. Before giving some important definitions, let us notice that the entropy $H(X)$ is maximal for a uniform distribution, and minimal (equals zero) when X is deterministic. Indeed, these two extreme cases are respectively the one which gives the less information on X, and the one which gives the most information when $X = a$, $a \in \mathbb{R}$.

The *conditional entropy* $H(Y|X)$ indicates the average loss of information on a random variable Y when the behavior of X (which is also a random variable) is known. Intuitively, the more Y "depends" on X, weaker will be the conditional entropy, and conversely. The conditional entropy of Y given X writes

$$H(Y|X) = \mathbb{E}_X[H(Y|x)] ,$$

where Y takes values in $\Upsilon \subset \mathbb{R}^d$ and $H(Y|x) = -\sum_{y\in\Upsilon} p_Y^{X=x}(y)\ln(p_Y^{X=x}(y))$.

The *mutual information* between two random variables X and Y indicates the information explained by X in Y (resp. by Y in X):

$$I(X,Y) = H(X) + H(Y) - H(X,Y) \quad \text{or}$$
$$I(X,Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) .$$

Using these definitions, Krzykacz-Hausmann 2001 defines an entropy-based sensitivity indice as

$$\eta_i = \frac{I(X_i,Y)}{H(Y)} = 1 - \frac{H(Y|X_i)}{H(Y)},$$

which is a representation of the information learnt on Y by the knowledge of X. It can be shown that $I(X,Y)$ is non negative.

In order to define the sensitivity indice used by Liu, Chen, and Sudjianto 2006, we must introduce the *relative entropy* notion, called Kullback-Leibler entropy. The relative entropy $D(p:q)$ of a probability measure p in regard of another probability measure q is

$$D(p:q) = \int p(x)\ln\frac{p(x)}{q(x)}d\mu(x) .$$

We can show that this last quantity is non negative. Then, sensitivity indices (called KL-entropy based sensitivity indices) can be derived:

$$KL_i(p_1|p_0) = \int_{-\infty}^{+\infty} p_1(y(x_1,\dots,\overline{x_i},\dots,x_n)) \left|\log\frac{p_1(y(x_1,\dots,\overline{x}_i,\dots,x_n))}{p_0(y(x_1,\dots,x_i,\dots,x_n))}\right| dy$$

with p_1 and p_0 respectively the probability distributions of the model output, depending if X_i become known or not, and $\overline{x_i} = \mathbb{E}[X_i]$. These indices thus indicate how much the Y probability distribution changes by fixing a model input to its mean.

3 NON-DISCRIMINANT CASES

We highlight two elementary situations where the entropy and variance sensitivity analyses lead to opposite conclusions. In the first example, the entropy-based sensitivity indices bring the right interpretation,

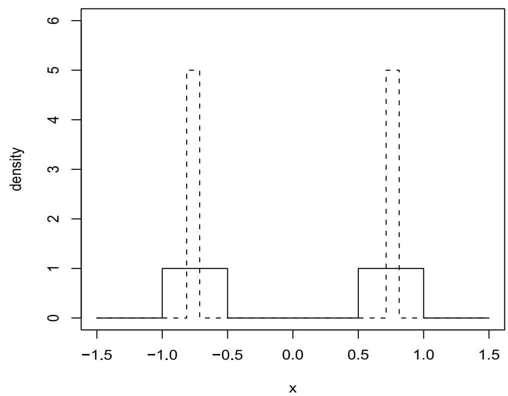

Figure 1. Density functions $f_1(x)$ (solid line) and $f_2(x)$ (dashed line).

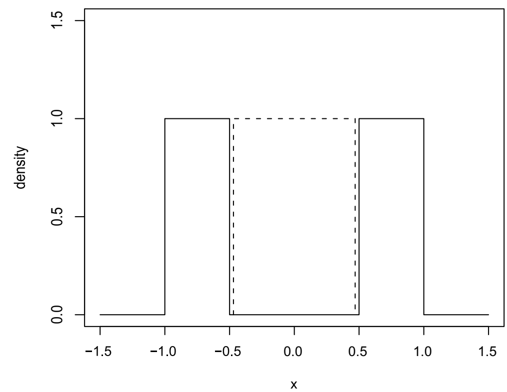

Figure 2. Density functions $g_1(x)$ (solid line) and $g_2(x)$ (dashed line).

while in the second example, the variance-based sensitivity indices lead to a correct interpretation. This justifies the introduction of entropy to compute sensitivity analysis, as a complementary tool.

3.1 *Example where variance does not separate variables*

Let X_1 and X_2 be two random variables on $[-1,1]$, with respective density functions $f_1(x) = 1_{[-1,-\frac{1}{2}]}(x) + 1_{[\frac{1}{2},1]}(x)$ and $f_2(x) = 5 \times 1_{[-\alpha,-\alpha+\frac{1}{10}]}(x) + 5 \times 1_{[\alpha-\frac{1}{10},\alpha]}(x)$, with $\alpha \simeq 0.8132$ (figure 1). Let the output Y verify $Y = \psi(X_1,X_2)$,

$$\begin{aligned}\psi : \mathbb{R}^2 &\longrightarrow \mathbb{R}\\ (x,y) &\mapsto x+y\,.\end{aligned}$$

After few elementary calculations, we get:

$$\begin{aligned}&\mathbb{E}[Y|X_1] = X_1\\ &\mathbb{E}[Y|X_2] = X_2\,,\end{aligned}$$

as expected, because X_1 and X_2 are independent.

Thus, variance-based sensitivity analysis provides equal influence for both variables X_1 and X_2. Reversely, we get the following conditional entropies:

$$\begin{aligned}&H(Y|X_1) = -\ln(5)\\ &H(Y|X_2) = 0\,.\end{aligned}$$

Finally, in this example $\eta_1 > \eta_2$ as intuitively stated: X_1 has more impact than X_2 because its probability distribution is closer to the uniform one. X_1 has a larger entropy than X_2, and so contains more uncertainty.

3.2 *Example where entropy does not separate variables*

We proceed as in the last paragraph, creating a model where entropy cannot distinguish two variables.

Let X_1 and X_2 be two random variables on $[-1,1]$, with respective density functions $g_1(x) = 1_{[-1,-\frac{1}{2}]}(x) + 1_{[\frac{1}{2},1]}(x)$ and $g_2(x) = 1_{[-\frac{1}{2},\frac{1}{2}]}(x)$ (figure 2). The output $Y = \psi(X_1,X_2)$ verifies:

$$\begin{aligned}\psi : \mathbb{R}^2 &\longrightarrow \mathbb{R}\\ (x,y) &\mapsto x+y\,.\end{aligned}$$

As before, we compute:

$$\begin{aligned}&\mathbb{E}[Y|X_1] = X_1\\ &\mathbb{E}[Y|X_2] = X_2\,.\end{aligned}$$

Finally, we observe that the sensitivity indice is higher for X_1 than for X_2: X_1 induces more variation for Y than X_2 does. Concerning the indices η_i, after few calculations we get $H(Y|X_1) = H(Y|X_2) = 0$, so conditional entropy is not relevant in this case.

4 INDICES COMPUTATION ALGORITHM

In this section we give some algorithms to compute sensitivity indices and some convergence results of η_i indices. Concerning variance-based sensitivity indices, we use the linear algorithm in the number of sample points of (Saltelli 2002). Less expensive algorithms exist but this is not the purpose of this paper to perform these comparisons.

4.1 *Practical computation of entropy-based sensitivity indices*

The discrete definition of the η_i indices for a mono-dimensional output is, assuming each interval of variation splitted into n sub-intervals:

$$\eta_i = -\left[\sum_{j,k=1}^{n} F^i_{j,k} \ln\left(\frac{F^i_{j,k}}{\int_{x=x^i_j}^{x^i_{j+1}} f_{X_i}(x)dx \int_{y=y_k}^{y_{k+1}} f_Y(y)dy}\right)\right] {H_0}^{-1},$$

with $H_0 = -\sum_{k=1}^{n} \int_{y=y_k}^{y_{k+1}} f_Y(y)dy \ln\left(\int_{y=y_k}^{y_{k+1}} f_Y(y)dy\right)$ and $F^i_{j,k} = \int_{x=x^i_j}^{x^i_{j+1}} \int_{y=y_k}^{y_{k+1}} f_{(X_i,Y)}(x,y)dxdy$.

We usually cannot explicit the function ψ which links the model inputs to the output. So, we assume this situation for the algorithm computing the indices. The method simply consists on a Monte-Carlo sampling to evaluate the quantities $F^i_{j,k}$. Then we can calculate the sums inside the formula giving η_i.

Algorithm's time complexity is then $\Theta(sSn^{d+1})$, with s being the number of inputs, S the number of sampling points and d the output space dimension (the formula above corresponds to the common case $d = 1$). There is also a multiplicative constant depending on the amount of intermediate results stored in memory.

A similar study concerning the indices KL_i hasn't been completed, but given the similarity between the definitions of η_i and KL_i, we assume that algorithm and complexity are very close, as verified in experiments.

4.2 *Convergence properties*

Let n be the number of sub-intervals in the discretization of each random variable's distribution. We can then state the following result.
Property: For all integrable real random variables X and Y on the real segment $[a,b]$,

$$\lim_{n\to+\infty} H_n(X) - H_n(Y) = 0\,.$$

This property indicates that when the subdivision step decreases near zero, all the distributions get the same discrete entropy (on a given interval $[a,b]$). We therefore have to choose n judiciously, and not to change it. We could normalize all the indices by a quantity like $\ln(n)$ and get a convergence with n. However, it is useless in practice because computations are very greedy when n increases (exponential increasing needed to get a few digits gain on the precision).

Let S be the number of sample points in the Monte-Carlo simulation. Then S should be greater than βn^2 where β lays in $[0,1]$ - but often close to one in practice - to observe a clear convergence of the indices. β depends on the inputs distributions and the hidden function ψ, so we have no control on it. Finally, we will often not be able to get a proper convergence in small time.

Nevertheless, we can derive a practical methodology to compute entropy-based sensitivity indices:

- set n as big as possible to run simulations with $n^{d+\alpha}$ loops in limited time, $\alpha \in \{1,2\}$ depending on the indices order (we never go further than two, because of exponential complexity);
- start running the algorithm with $S = \beta n^{d+\alpha}$, $\beta \in [0,1]$ as close to one as the random variables distributions have a large entropy. For instance we can choose $\beta = \frac{1}{n\ln(n)} \sum_{k=1}^{n} F^n_k$, where F^n_k is the discretized entropy of the kth input variable;
- increase S until time consumption is too big or indices clearly stabilize.

5 APPLICATIONS

Here are two application cases for the previously described algorithms.

5.1 *Ishigami analytical function*

This function is a classic benchmark for algorithms calculating the variance-based sensitivity indices. It writes

$$is : \mathbb{R}^3 \longrightarrow \mathbb{R},$$
$$(x_1,x_2,x_3) \longrightarrow \sin(x_1) + 7\sin^2(x_2) + 0.1x_3^4 \sin(x_1)$$

where x_1,x_2,x_3 are three independent realizations of uniform random variables on $[-\pi,\pi]$. The approximative distribution of $Y = is(X_1,X_2,X_3)$ is drawn on the figure 3.

We set the number of sub-intervals at $n = 100$ for the discretization. The results for the indices η_i are in table 1. Time consumption is very small because we use a determinist sampling available when dealing with entropy-based indices and uniform distributions. Computations were run with a C++ version of the algorithm, with S ranging from 100 to 10^6. As we can check on figure 4, the relative indices converge faster than the sensitivity indices. We see that $S \geq 10000$ allows to obtain reliable relative entropy-based sensitivity indices.

The table 2 gives the results for variance-based total sensitivity indices. Indeed, first-order indices based on variance do not take into account the interactions between a variable and the others, whereas entropy-based indices contain a global information brought by the variable. Thus, indices η_i are more global: they

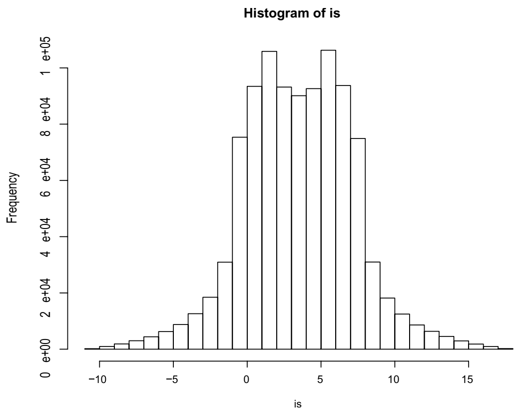

Figure 3. Histogram of the Ishigami function output.

Table 1. Entropy-based sensitivity indices.

S	η_1	η_2	η_3
100 (<1s)	0.5289	0.5624	0.5115
500 (<1s)	0.4565	0.4457	0.3693
1000 (<1s)	0.5180	0.3996	0.4100
5000 (<1s)	0.1324	0.1537	$7.924e{-2}$
10000 (<1s)	0.1220	0.1591	$8.609e{-2}$
50000 (<1s)	$9.815e{-2}$	0.1332	$6.409e{-2}$
100000 (<1s)	$9.254e{-2}$	0.1238	$6.093e{-2}$
500000 (1s)	$8.967e{-2}$	0.1198	$5.821e{-2}$
1000000 (3s)	$9.189e{-2}$	0.1210	$5.778e{-2}$
	(33.9%)	(44.7%)	(21.4%)

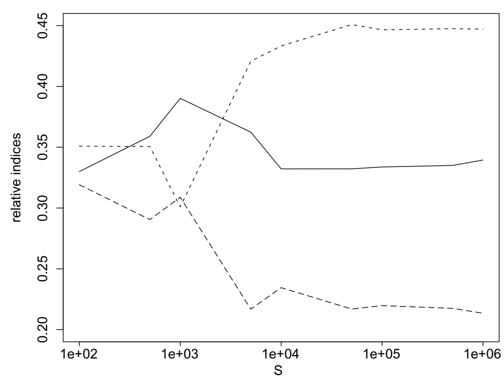

Figure 4. Relative η indices in function of S. η_1: solid, η_2: short dashed, η_3: long dashed.

have to be compared to the total S_{T_i} variance-based indices. Computation time for 100 repetitions are between parenthesis, and standard deviations can be found just below the indices values.

Table 2. Variance-based total sensitivity indices.

S	S_{T_1}	S_{T_2}	S_{T_3}
100 (<1s)	0.5779	0.4663	0.2715
	0.165	0.110	0.113
500 (<1s)	0.5555	0.4451	0.2380
	$8.18e{-2}$	$4.92e{-2}$	$4.91e{-2}$
1000 (1s)	0.5459	0.4462	0.2436
	$4.48e{-2}$	$4.27e{-2}$	$3.49e{-2}$
5000 (2s)	0.5470	0.4361	0.2447
	$1.95e{-2}$	$1.67e{-2}$	$1.62e{-2}$
10000 (4s)	0.5615	0.4460	0.2426
	$1.80e{-2}$	$1.20e{-2}$	$1.09e{-2}$
50000 (20s)	0.5604	0.4429	0.2432
	$6.95e{-3}$	$4.89e{-3}$	$4.62e{-3}$
100000 (40s)	0.5549	0.4427	0.2439
	$4.58e{-3}$	$3.80e{-3}$	$2.78e{-3}$
500000 (3min10s)	0.5578	0.4418	0.2437
	$2.22e{-3}$	$1.56e{-3}$	$1.57e{-3}$
$S \Rightarrow +\infty$ (theory)	0.5574	0.4442	0.2410
	(44.9%)	(35.7%)	(19.4%)

We notice that the variance-based indices converge faster than entropy-based indices. As the computation times are longer (because of the 100 repetitions), advantages and drawbacks compensate.

The ranks of X_1 and X_2 according to the sensitivity analysis are reversed, but X_3 stays with a small influence in both systems. Nevertheless, we notice that the proportions remain similar. It is certainly due to the near-Gaussian distribution of the Ishigami function output (Fig. 3). Indeed, entropy and variance brings the same information for a Gaussian random variable Y. In this case, we have $H(Y) = 0.5[1 + \ln(2\pi) + \ln(\mathbb{R}mVar(Y))]$. We can conclude that X_1 and X_2 have the same global influence on the output Y.

Finally, the results of the KL-entropy based method are given in the table 3, with the standard deviations just below the indices' values.

The convergence is similar to the one observed for the former case involving entropy, slow but clear convergence. We have stopped at $S = 10^7$ sampling points to keep running time relatively small, and for each value of S we repeat the algorithm 10 times because the (basic) deterministic sampling used previously provides weird results for this method: the absolute values of the indices seem to converge till $S = 5.10^5$, but then we observe a divergence before a new convergence around $S = 10^7$. We see that $S \geq 500000$ allows to obtain reliable KL-entropy based sensitivity indices. The figure 5 shows the faster convergence of the relative indices compared to the absolute indices.

We notice that the ranking is completely reversed between these indices and the first entropy-based indices studied. To understand why this method

Table 3. Kullback-Leibler entropy-based sensitivity indices.

S	KL_1	KL_2	KL_3
500 (<1s)	0.8698	0.7733	0.8255
	0.1695	$8.888e{-2}$	$5.1507e{-2}$
1000 (<1s)	1.1769	0.7130	1.1141
	0.207345	$4.685e{-2}$	$7.422e{-2}$
5000 (1s)	2.0630	0.6699	1.5563
	0.1267	$2.756e{-2}$	$2.103e{-2}$
10000 (1s)	2.4042	0.6746	1.6192
	0.1286	$1.454e{-2}$	$4.148e{-2}$
50000 (3s)	2.6568	0.6548	1.6880
	$9.635e{-2}$	$5.569e{-3}$	$1.291e{-2}$
100000 (6s)	2.7344	0.6479	1.7003
	$4.449e{-2}$	$3.780e{-3}$	$1.284e{-2}$
500000 (32s)	2.8545	0.6484	1.7203
	$2.551e{-2}$	$1.526e{-3}$	$4.156e{-3}$
1000000 (1min04s)	2.8763	0.6469	1.7212
	$2.683e{-2}$	$1.843e{-3}$	$4.334e{-3}$
5000000 (5min20s)	2.9198	0.6470	1.7289
	$9.346e{-3}$	$4.603e{-4}$	$1.954e{-3}$
10000000 (25min)	2.9209	0.6472	1.7299
	$8.311e{-3}$	$4.008e{-4}$	$1.884e{-3}$
	(55%)	(12%)	(33%)

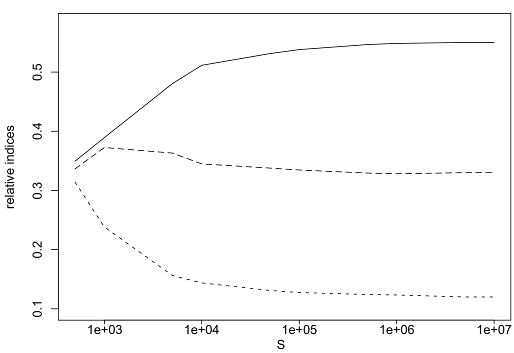

Figure 5. Relative KL-entropy based sensitivity indices in function of S. KL_1: solid, KL_2: short dashed, KL_3: long dashed.

gives this new ranking, we draw in figures 6, 7 and 8, the three conditional distributions of ($Y =$ is $(X_1, X_2, X_3)|X_i = \bar{x}_i$), $i = 1, \ldots, 3$ (remind that the KL-entropy based indices indicates how much the conditional distribution differs from the a priori distribution), where $\bar{x}_i$ is the mean of X_i.

The ranking is now clearly explained: the histogram of figure 6 is extremely different of the original histogram of the figure 3, so the indice KL_1 gets the largest value. The distribution of figure 7 looks almost identical to the original distribution, except that there is only one mode. Then, KL_2 gets the smallest value. Finally,

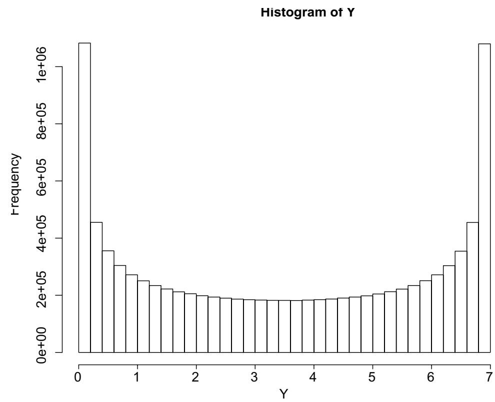

Figure 6. Conditional output distribution of the Ishigami function with $X_1 = 0$.

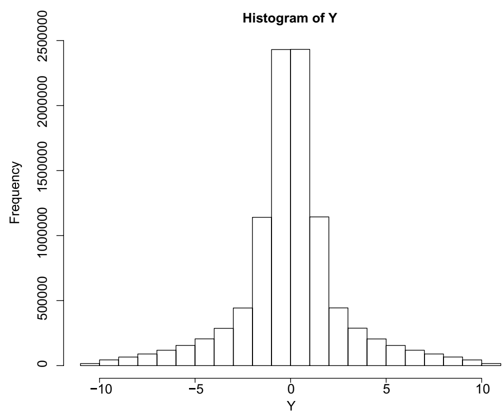

Figure 7. Conditional output distribution of the Ishigami function with $X_2 = 0$.

the histogram of figure 8 has the same shape than the original histogram, with two modes but without the low tail on each side. Then, the indice KL_3 has an intermediate value.

In conclusion of this analytical exercise, we have seen that variance-based and entropy-based sensitivity indices are similar when the output distribution is close to a Gaussian. Those two measures are global while the KL-entropy based sensitivity measure gives a more-local information: it studies the departure from the output distribution by fixing the inputs at their mean values.

5.2 *An industrial application*

We study at present a physical computer code simulating severe nuclear accident on a pressurized water

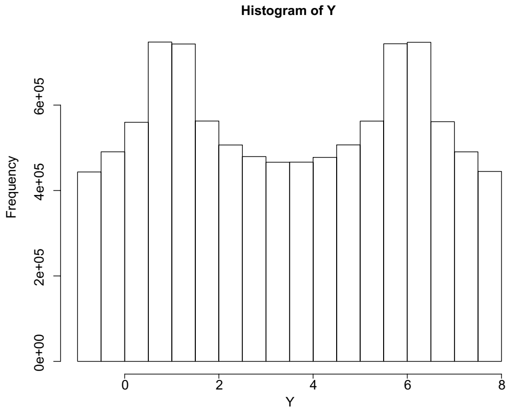

Figure 8. Conditional output distribution of the Ishigami function with $X_3 = 0$.

Table 4. Morris' screening results for Y_2.

input	$\mu*$	σ
X_5	0.76630	0.69778
X_{18}	0.29496	0.52215
X_{17}	0.09264	0.29294
X_9	0.11254	0.18079
X_6	0.04666	0.14621
X_{15}	0.01591	0.03603
X_3	0.00208	0.00656
X_{10}	0.00200	0.00632
X_1	0.00184	0.00582
X_{20}	0.00113	0.00243
X_{13}	0.00035	0.00110

Table 5. Sensitivity analysis on the computer code.

Y_1 SA	X_5	X_6	X_9	X_{15}	X_{17}	X_{18}
S_{T_i}	0.816	1.31*e*-2	0.529	0.109	0.000	1.88*e*-2
η_i	0.396	0.110	0.226	0.110	0.104	0.109
Y_2 SA	X_5	X_6	X_9	X_{15}	X_{17}	X_{18}
S_{T_i}	0.822	4.04*e*-2	0.389	3.31*e*-2	1.52*e*-2	0.346
η_i	0.241	0.130	0.140	0.128	0.130	0.328

reactor. It simulates the complex physical and chemical processes (core melt progression, fission-product release and transport, etc) governing the phenomena that determine the potential radiological releases outside the nuclear facilities. For confidentiality reasons, we do not present rthe technical details of this scenario.

The physical code has 32 random input variables (uniform distributions) and we study 2 of its output variables: Y_1 and Y_2. It would be computationally untractable to calculate the indices for each input relatively to each output. To adress this issue, we apply first a screening design using the method of Morris 1991, which allows to identify the variables which have the the most influence on the outputs. This method builds a multi-dimensional regular grid on which it evaluates the values of the function. It then sums up the means ($\mu*$) and variance (σ^2) observed due to each variable, when fixing all the others and moving along one dimension. Ten levels are used for the Morris' screening, with ten elementary effects computed per factor. The results for the second output are given in the table 4 (all non mentioned variables have zero values for $\mu*$ and σ), as example. The inputs are sorted according to their $\mu * + \sigma$ values. We conclude after this preliminary study that the parameters $X_5, X_6, X_9, X_{15}, X_{17}$ and X_{18} are the most influent, and we will consequently concentrate the analysis on them.

At present, we study 6 input variables and we can perform a precise sensitivity analysis on the 2 ouput variables using quantitative methods. We restrict our study to the Sobol and η indices. The results for 10000 sample points are given in table 5. A simple Monte-Carlo sampling was used for both types of sensitivity analysis.

Concerning the output Y_1, the two methods give the same ranking, except the case of X_{15} which is not stated as more important than X_6 and X_{18} for the η indices, whereas Sobol indice is higher. For the output Y_2, rankings are quite different : X_5, X_9 then X_{18} for variance-based indices, X_{18}, X_5 then X_9 for entropy-based indices. This shows that those two methods don't compare the inputs in the same way, but don't give so different results.

To understand why the indices are different, let us draw the two histograms corresponding to the two outputs on figures 9 and 10. The first distribution seems to distinguish two modes which correspond to two different phenomena involved in the physics of the computer code. We know that variance is not well suited in case of a multimodal distribution, then entropy-based sensitivity indices are useful in this case. However, in our case, the bimodality is not clear beacuse the second mode is noticeably smaller than the first mode. Therefore it explains why variance and entropy give the same influent input variables. The second output histogram includes just one mode, but a heavy right tail, which is another case where variance is not a well-adapted measure. The predominant input variable with the variance is X_5 ($S_{T_5} = 0.82$), X_9 and X_{18} having a smaller influence ($S_{T_9} = 0.39$ and $S_{T_{18}} = 0.35$). The most influent input variables with the entropy are X_{18} ($\eta_{18} = 0.33$) and X_5 ($\eta_9 = 0.24$). As the entropy is influenced

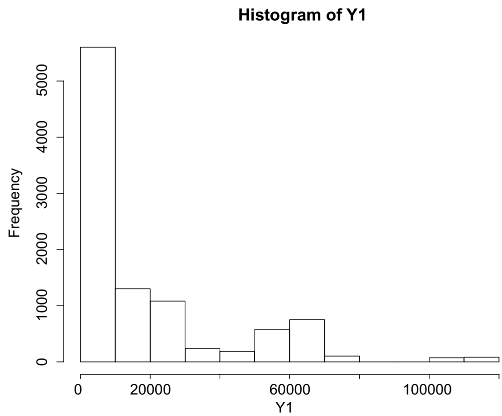

Figure 9. Y_1 distribution for the industrial application.

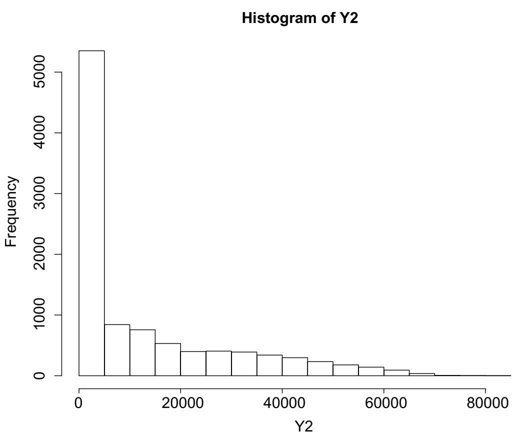

Figure 10. Y_2 distribution for the industrial application.

by the entire distribution (including the heavy tail) and the variance is only influenced by the central part of the distribution, we suspect that the input variable X_{18} has a strong influence on this distribution tail, then on the extreme values of the output variable Y_2. This kind of information could be useful for the analyst engineer.

As a conclusion, variance is in general more discriminant than entropy, clearly giving a few very important indices and ranking all parameters without ex-aequo. By contrast, entropy seems to rank several indices around the same level, due to its logarithmic nature. This not means that variance should be prefered: entropy does separate variables clearly, it would only need an appropriate exponential transformation, and, more important, a larger number of sample points.

As a conclusion, in our examples variance is in general more discriminant than entropy, clearly giving a few very important indices and ranking all parameters without ex-aequo. By contrast, entropy seems to rank several indices around the same level, due to its logarithmic nature. This last limitation makes the convergence too slow to observe stable results with only 10000 points in the input sampling. However, this not means that variance should be prefered: entropy does separate variables clearly, it would only need an appropriate exponential transformation, and, more important, a larger number of sample points. To address this issue, a future direction of research will be to replace the computer code by a metamodel (mathematical function with negligible cpu time, approximating the computer code), then computing the sensitivity indices using this metamodel (Volkova, Iooss, and Van Dorpe 2008).

6 CONCLUSIONS

The use of the entropy measures brings new information in sensitivity analysis of model outputs, as for instance the symetrization of the roles of input and output, focusing on the random variables distributions and so on. However, we should be aware that the entropy-based indices don't fill the interval $[0, 1]$ uniformly, because of the logarithmic nature of entropy. In addition, in some situations entropy cannot discriminate two inputs when variance can (and vice-versa).

As entropy stands for a global measure of influence, whereas variance only takes into account second-order moments, we can think entropy as a complement to the variance measure: entropy-based indices will more likely be used to complete or precede an analysis using variance. Moreover in this paper, we have not mention an interesting additional property: entropy naturally deals with multi-output sensitivity analysis by extending the definition of the η (resp. KL) indices to multi-sums (resp. multi-integrals), so we can avoid to compute sensitivity indices for each output as we do when using variance.

Finally, we conclude that the Sobol indices work with the measure of variance, whereas the η indices use conditional entropy and the KL indices measure differences between conditional and a priori output distribution. Those three methods therefore often produce significantly different results, as we have checked throughout the examples.

REFERENCES

Krzykacz-Hausmann, B. (2001). Epistemic sensitivity analysis based on the concept of entropy. In *Proceedings of SAMO 2001*, Madrid, pp. 31–35. CIEMAT.

Lejeune, M. (2004). *Statistique - La th´eorie et ses applications*. Springer.

Liu, H., W. Chen, and A. Sudjianto (2006). Relative entropy based method for probabilistic sensitivity analysis in engineering design. *ASME Journal of Mechanical Design 128*, 326–336.

Middleton, B. (2005). *Use of information theory techniques with system dynamics models*. Thesis of the Massachussets Institute of Technology.

Morris, M. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics 33*, 161–174.

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

Saltelli, A., K. Chan, and E. Scott (Eds.) (2000). *Sensitivity analysis*. Wiley Series in Probability and Statistics. Wiley.

Sobol, I. (1993). Sensitivity estimates for non linear mathematical models. *Mathematical Modelling and Computational Experiments 1*, 407–414.

Volkova, E., B. Iooss, and F. Van Dorpe (2008). Global sensitivity analysis for a numerical model of radionuclide migration from the RRC "Kurchatov Institute" radwaste disposal site. *Stochastic Environmental Research and Risk Assesment 22*, 17–31.