Computer Physics Communications 182 (2011) 978–988

Contents lists available at ScienceDirect

Computer Physics Communications

www.elsevier.com/locate/cpc

From screening to quantitative sensitivity analysis. A unified approach

Francesca Campolongo *, Andrea Saltelli, Jessica Cariboni

Joint Research Centre, Institute for the Protection and Security of the Citizen, Ispra, Italy

ARTICLE INFO

Article history:
Received 24 June 2010
Received in revised form 22 December 2010
Accepted 27 December 2010
Available online 6 January 2011

Keywords:
Sensitivity analysis
Radial design
Variance based measures
Elementary effect method

ABSTRACT

The present work is a sequel to a recent one published on this journal where the superiority of 'radial design' to compute the 'total sensitivity index' was ascertained. Both concepts belong to sensitivity analysis of model output. A radial design is the one whereby starting from a random point in the hyperspace of the input factors one step in turn is taken for each factor. The procedure is iterated a number of times with a different starting random point as to collect a sample of elementary shifts for each factor. The total sensitivity index is a powerful sensitivity measure which can be estimated based on such a sample. Given the similarity between the total sensitivity index and a screening test known as method of the elementary effects (or method of Morris), we test the radial design on this method. Both methods are best practices: the total sensitivity index in the class of the quantitative measures and the elementary effects in that of the screening methods. We find that the radial design is indeed superior even for the computation of the elementary effects method. This opens the door to a sensitivity analysis strategy whereby the analyst can start with a small number of points (screening-wise) and then – depending on the results – possibly increase the numeral of points up to compute a fully quantitative measure. Also of interest to practitioners is that a radial design is nothing else than an iterated 'One factor At a Time' (OAT) approach. OAT is a radial design of size one. While OAT *is not* a good practice, modelers in all domains keep using it for sensitivity analysis for reasons discussed elsewhere (Saltelli and Annoni, 2010) [23]. With the present approach modelers are offered a straightforward and economic upgrade of their OAT which maintain OAT's appeal of having just one factor moved at each step.

1. Introduction

Uncertainty and sensitivity analyses study how the uncertainties in the model inputs $(X_1, X_2, \ldots, X_k)$ affect the model's response Y, which for simplicity we assume to be a scalar:

$$Y = Y(X_1, X_2, \ldots, X_k).$$

Uncertainty analysis quantifies the output variability while Sensitivity Analysis (SA) describes the relative importance of each input in determining this variability. The importance of SA in the model building process has been extensively demonstrated in many scientific areas and has been also recognized in official guidelines of international institutions [6,8,20].

Sensitivity analysis methods can be broadly classified into local and global methods. Local sensitivity measures, often referred to as One At a Time (OAT) measures, are based on the estimation of partial derivatives, and assess how uncertainty in one factor affects the model output keeping the other factors fixed to a nominal value. The main drawback of this approach is that interactions among factors can not be detected, since they become evident when the inputs are changed simultaneously. A possible way to overcome this problem is to include multi-dimensional averaging of local measures [18,16,38,39].

Global measures offer instead a comprehensive approach to model analysis, since they evaluate the effect of a factor while all others are varying as well, exploring efficiently the multi-dimensional input space. A wide range of global SA methods is available (for reviews see [25,28,31,27,11,21]), ranging from qualitative screening methods [18,3,29,15] to quantitative techniques based on variance decomposition [34,12,24,19].

Using sensitivity techniques capable of detecting interactions among factors is of key importance to completely assess the model prediction [35,36]. Examples of application of SA tools to test the role of interactions in real models range in very diverse fields, from chemistry (see e.g. [27]), ecological modeling (see e.g. [4]), to hydraulic modeling (e.g. [9]).

Among variance-based measures, the total order sensitivity index ST [12] measures the total effect of a factor, including its first order effect and interactions of any order:

$$S_{Ti} = \frac{V_{Ti}(Y)}{V(Y)} = \frac{E_{\mathbf{X}_{\sim i}}(V_{X_i}(Y \mid \mathbf{X}_{\sim i}))}{V(Y)} = 1 - \frac{V_{\mathbf{X}_{\sim i}}(E_{X_i}(Y \mid \mathbf{X}_{\sim i}))}{V(Y)}, \quad (1)$$

* Corresponding author.
E-mail address: francesca.campolongo@jrc.it (F. Campolongo).

0010-4655/$ – see front matter
doi:10.1016/j.cpc.2010.12.039

where $\mathbf{X}_{\sim i}$ denotes the matrix of all variables but X_i. In $E_{\mathbf{X}_{\sim i}}(V_{X_i}(Y \mid \mathbf{X}_{\sim i}))$ the inner variance V of Y is taken over all possible values of X_i while keeping $\mathbf{X}_{\sim i}$ fixed, while the output expectation E is taken over all possible values $\mathbf{X}_{\sim i}$ [12]. Different designs have been developed in order to compute efficiently sensitivity indices, spanning from the well known approach of the Latin hypercube [17,10] to the recently design developed in [30]. One efficient possibility is to use a radial design where factors are moved of one step starting from the same random point in the hyperspace of the input factors.

In a different class of SA methods, the class of screening tests, the method of the elementary effects, introduced by Morris [18] and refined by Campolongo and co-workers [3], is – in the opinion of the authors – a good practice. This method investigates the model answer to a change in the inputs by varying one factor at a time, while keeping all the others fixed. The method associates to each factor its so-called elementary effect, defined as the ratio between the variation in the model output and the variation in the input factor itself (see also Section 2). In order to obtain a global measure, different elementary effects for each factor are estimated and averaged. Here, adopting the terminology proposed by Campolongo and co-workers, we refer to the Elementary Effect method as to the EE method, and to its sensitivity measure, which is the average of the elementary effects in absolute values, as to μ^*. As described next, the average is customarily taken over points generated via a trajectory-based design [18,3]. A trajectory based sampling is very similar to the so-called winding stairs sampling [5]. In both methods one factor at a time is moved. In winding stairs there is a single trajectory linking all sampled points while in the trajectory-based, as used e.g. in [18], a new trajectory is started each time all factors have been moved once. A variant of this latter is used in [3] as discussed below.

Which SA method to use depends upon the model to be analyzed and goal of the experiment [26]. By quantitative methods, we mean those methods which estimate a sensitivity index whose meaning has a clear interpretation in terms of variance decomposition. In other words, quantitative sensitivity indices estimate which percentage of output variance each factor is responsible for, due to its first order component and/or its interactions with the other factors. However, to obtain good estimates of quantitative sensitivity indices, a high number of model evaluations is needed. For complicated models, where the number of involved factors is large and/or the model is time consuming, the estimation of quantitative sensitivity measures could be unfeasible and screening experiments need to be developed instead. Screening experiments provide sensitivity measures that do not have a straightforward interpretation in terms of variance decomposition. They only provide a ranking saying that a factor is more important than another, and allow to identifying the subset of non-important factors in a model with few model runs.

This paper proposes a novel sampling strategy based on Sobol' numbers [32,33] which allows performing screening experiments first and then, if the computational cost of the model allows it, moving towards a quantitative experiment without the need to discard any model run. The logic of the work runs as follows:

- Campolongo and co-workers [3] empirically demonstrated that there exist similarities in sensitivity results obtained using μ^* and the variance-based total sensitivity index. For screening purposes, where the aim is to isolate the set of non-influential factors with few model runs, μ^* has proven to be provide effective results.[1]
- Saltelli and co-workers [30] compared the estimates for the total sensitivity index obtained using a radial sampling strategy with those obtained via a trajectory-based sampling and showed that the former outperforms the latter.
- Elsewhere [23] it has been noted that, although it can be proven to be inefficient, purely OAT (a radial of size one) is the most widespread practice not only for sensitivity analysis purposes but also for model verification purposes (e.g. verification of results, verification of the sign of the effects, ...).
- Hence, if it could be shown that by iterating OAT one can compute efficiently the EE statistics, such an approach should be received with favor by modelers.
- This approach could be made even more appealing by the possibility of seamlessly computing together with the EE the total sensitivity index, by merely increasing the sample size used for EE, within the same batch of computer simulations.

As discussed in [23] modelers have several reasons to like OAT, including the fact that if only one factor has been moved then the change in the model response can be safely attributed to that factor. Thus we can imagine a modeler starting with a pure OAT, one point surrounded by k additional points (one per factor). The modeler can add then a second 'star' of $(k+1)$ points, and then another. Already with few such stars one can compute μ^* according to the original developer [18]. If convenient and computationally affordable, the modeler can then decide to continue the procedure generating a few hundred of stars and estimate the total sensitivity index, without discarding any of the previous simulations.

To check whether this approach is convenient we test a radial sampling strategy – the one we use normally to compute the total sensitivity index [22] – in a screening experiment for the estimation of μ^* using EE. The strategy, which makes use of Sobol' quasi-random sequences [32,33] to build the sample, is tested on a series of mathematical functions. This is a batch of 'difficult' test functions. The same was previously used to identify the best sampling algorithm for the total sensitivity index [30]. The batch is enriched here of the test function used by Morris in [18], the standard benchmark for the EE method.

We compare the results obtained via a radial design with existing strategies used to compute the EE measure: the strategy of the trajectories [18], the strategy of the optimized trajectories [3], and a new strategy recently proposed by Saltelli and co-workers in [29].

The paper is organized as follows. The next Section briefly reviews the best practice for screening experiments, which is the EE method and its sampling strategies. Section 3 describes how to use Sobol' quasi-random sequences to build a radial OAT design for screening experiments. Section 4 focuses on numerical experiments and the last section concludes.

2. The elementary effect method

Screening methods aim at identifying the subset of non-influent factors in a model using a small number of model evaluations. Screening is used for the investigation of large and/or computationally expensive models, where the use of more demanding quantitative techniques is not affordable due to simulation time. The EE method has proven to be an effective screening method since its origin. It is based on a sample where factors are moved

[1] The EE method has been recently compared to derivative-based global sensitivity measures [16,38] and it has been demonstrated that the latter approach can be sometimes more accurate. However, as also stressed in [16,38], high accuracy of estimates may not be required for screening purposes, where the aim is the detection of non-important factors with few model runs.

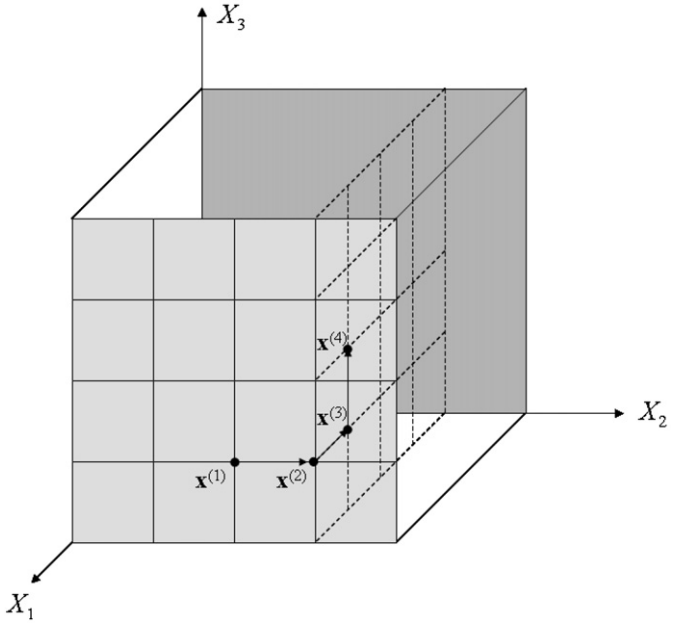

Fig. 1. Example of trajectory in 3 dimensions for the original EE method.

OAT of a step Δ_i in the inputs' domain. The sensitivity test is based on the elementary effect, from which the name of the method:

$$EE_i = \frac{Y(x_1, \ldots, x_{i-1}, x_i + \Delta_i, x_{i+1}, \ldots, x_k) - Y(x_1, \ldots, x_k)}{\Delta_i}. \quad (2)$$

r different elementary effects are estimated by randomly sampling r different trajectories; the final sensitivity measure adopted is the average of these effects:

$$\mu_i = \frac{\sum_{i=1}^{r} EE_i}{r}.$$

The original version [18] of the EE method is based on a sample of r trajectories where factors are moved OAT on a grid of levels covering the inputs' domain, see Fig. 1 for a 3-dimensional case. The number of points of each trajectory is $(k+1)$, where k is the number of factors of the model. Along a trajectory each input is increased or decreased by the same step Δ.[2] In Fig. 1 the trajectory is composed of the four points $\{\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(4)}\}$ and $\Delta = 0.25$. If we consider, for instance, the first two consecutive points $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$, they differ only in the second component, which is increased by Δ:

$$EE_2 = \frac{Y(\mathbf{x}^{(2)}) - Y(\mathbf{x}^{(1)})}{0.25}.$$

A second measure is proposed in [18], which assesses the extent of interactions and non-linear effects of each factor:

$$\sigma_i = \frac{\sum_{i=1}^{r}(EE_i - \mu_i)^2}{r}.$$

In [3], Campolongo and co-workers presented an improved version of the EE method, where:

- The sampling strategy is improved so to better scan the input domain using the same number of points. This is achieved by first generating a high number of different trajectories (e.g. $M \sim 500$–1000) and then selecting the r trajectories with the largest dispersion in the input space.[3] We call this improved strategy that of 'optimized trajectories'.
- A new measure μ^* is defined which alone assesses the factors' importance:

$$\mu_i^* = \frac{\sum_{i=1}^{r} |EE_i|}{r}. \quad (3)$$

The use of the absolute value of the EE allows to solve the problem of the effects of opposite signs which occurs when the model is non-monotonic. In [3] it was also empirically demonstrated that for screening purposes there exists similarities between results obtained via μ^* and via the variance-based total index ST.

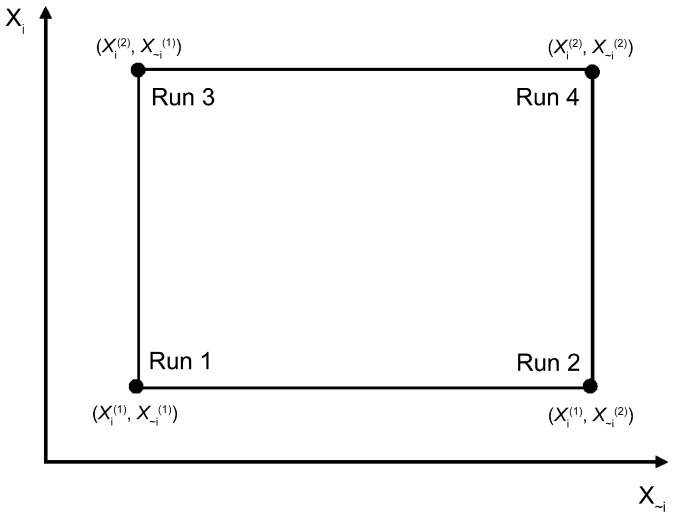

Fig. 2. Scheme of the cell-based sampling strategy.

A different design, that we call 'cell-based', is presented in [29], where the input space is explored combining steps along the X_i axis with steps along the $X_{\sim i}$ *axis*, an axis where all inputs but X_i change their values.[4] Fig. 2 exemplifies how the sampling strategy works. Starting from the first run $x_i^{(1)} x_{\sim i}^{(1)}$, the strategy first moves in the direction $\sim i$, i.e. to the second run $x_i^{(1)} x_{\sim i}^{(2)}$ and then along the i axis to the third run $x_i^{(2)} x_{\sim i}^{(1)}$. The last run, where all factors are changed respect to the first run, completes the cell.

With respect to the optimized trajectories strategy, this design has the advantage of providing, at no extra computational cost, an additional sensitivity measure (on top of the classic ones based on the elementary effects) to assess interaction effects:

$$EI = \frac{1}{2}\left|Y^{(1)} - Y^{(2)} + Y^{(4)} - Y^{(3)}\right|, \quad (4)$$

where the super-indices refer to the corresponding input points in Fig. 2. This second measure can be useful for non-additive models characterized by pure interactions terms. However tests on two relevant benchmarks [29], the typically non-additive model introduced by Saltelli et al. in [28] (see also Section 4), and a 10-inputs version of the Morris function [18], showed that results obtained via the cell-based strategy only slightly outperforms the ones obtained via the design based on trajectories. Such findings are confirmed in the present study based on a larger set of test functions.

[2] We assume that each input is uniformly distributed in $[0, 1]$. For non-uniform factors, we sample quantiles in $[0, 1]$ and derive the input values through the inverse distribution function.

[3] In [3] a dispersion measure is defined in terms of distance between couples of trajectories.

[4] Despite $X_{\sim i}$ refers to $(k-1)$ variables, we will point to a variation of these $(k-1)$ variables as a *step* taken onto the X_i *axis*.

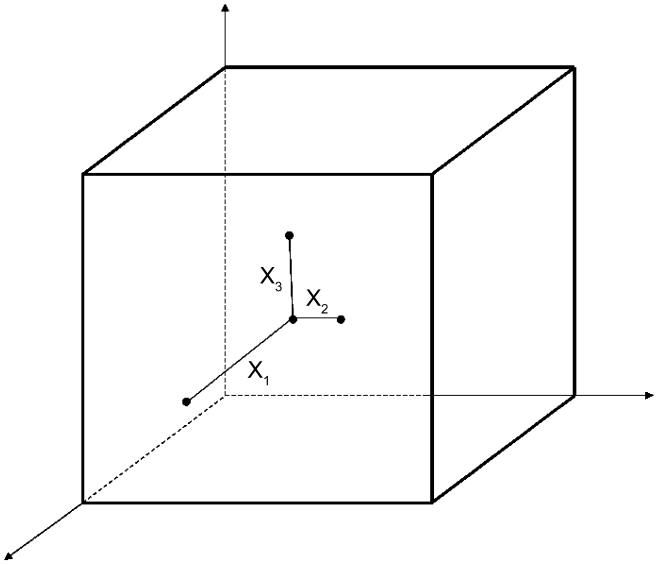

Fig. 3. Example of a radial sample in three dimensions.

Table 1
Block of size $(k+1)$ runs for a Monte Carlo simulation for sensitivity analysis. Radial (left-hand) and trajectory (right-hand) schemes are compared. The left-hand block as such – i.e. if not complemented by other blocks – corresponds to an OAT design as often seen in the literature. For a proper screening, a number $r > 1$ of such blocks is needed, for a total computation cost of $N = r(k+1)$.

Radial sampling	Trajectory sampling
$a_1, a_2, a_3, \ldots, a_k$	$a_1, a_2, a_3, \ldots, a_k$
$b_1, a_2, a_3, \ldots, a_k$	$b_1, a_2, a_3, \ldots, a_k$
$a_1, b_2, a_3, \ldots, a_k$	$b_1, b_2, a_3, \ldots, a_k$
$a_1, a_2, b_3, \ldots, a_k$	$b_1, b_2, b_3, \ldots, a_k$
$\ldots$	$\ldots$
$a_1, a_2, a_3, \ldots, b_k$	$b_1, b_2, b_3, \ldots, b_k$

3. Using Sobol' numbers to build a radial OAT experiment

The alternative design to compute the elementary effects presented here is based on a radial-like configuration. The radial-like configuration is presented in Table 1 for k factors and exemplified in three dimensions in Fig. 3. Table 1 also reports for comparison the trajectory-like configuration discussed above.[5] **a** and **b** are two different k-dimensional random vectors which can be used to realize the so-called X_i steps [30]. An X_i step is made of two points which are apart only for one coordinate, i.e. only for factor X_i, all others being the same. An X_i step is used for the computation of an elementary effect for that factor and a series of such steps allows an estimate of the factor's importance μ^*. In radial design (left-hand side in Table 1) one goes back to the first point $(a_1, a_2, a_3, \ldots, a_k)$ after each step, while in trajectory sampling one keeps moving away after each step, so that starting from $(a_1, a_2, a_3, \ldots, a_k)$ one ends in $(b_1, b_2, b_3, \ldots, b_k)$. Following [14] the trajectory design is also known as 'winding stairs', although in the proper winding stairs design one keeps moving away from $(b_1, b_2, b_3, \ldots, b_k)$ in order to start the second block [5], while in trajectory sampling one starts over a new trajectory with a fresh random number, see e.g. the original design of Morris [18]. One can call **a** the baseline point and **b** the auxiliary point. Note that compared to the Morris sampling strategy presented in the previous section and exemplified in Fig. 1, the random points **a** and **b** are such that the steps taken by different factors can be different.

[5] The cell-based configuration is not reported.

In [30] we have extensively tested the two designs to compute the total sensitivity index, and found the radial design superior. Given the similarity between *ST* and the elementary effect measure [3], we have done here a similar comparison using the same set of test functions, augmented by the twenty-factor function of Morris which was not included in [30]. In keeping with [30], in place of using plain random numbers to generate the random vectors **a** and **b**, we use Sobol's quasi-random numbers [32,33] – the so-called LP_τ sequences. As discussed in [30], Sobol's sequences outperform crude Monte Carlo sampling in the estimation of multi-dimensional integrals [37].

The first eight points of a ten-dimensional Sobol's quasi-random sequence are given, as an example, in Table 2. This matrix could be used for a ten-dimensional test case with $r = 4$ repetitions, using the first four rows of Table 2 for the four baseline points **a**'s, and the next four for the auxiliary points **b**'s. In other words the first row of Sobol' numbers in Table 2 would be used for point **a** in Table 1 (left-hand side), and the fifth row of the Sobol' matrix for point **b**. Next, moving from $r = 1$ to $r = 2$, the second and sixth rows of Table 2 would be used for the next couple of **a**, **b** numbers and so on. Although the approach above would be the most natural to build a pure screening analysis, we adopt a slightly different strategy – in line with the findings of [30] – to ensure that the same points can be used to compute the total effect indices. The idea is to generate a matrix of quasi-random Sobol' numbers of dimensions $(r, 2k)$, r being the number of points/trajectories and k being the number of inputs, and use the left half of the matrix for the points **a** and the right half of the matrix for the points **b**. As in the first points of the Sobol' sequence the values of the coordinates tend to repeat (i.e. for the first point they are all equal to 0.5, for the second they are alternates couples of 0.25 and 0.75 and so on, see Table 2), in order to achieve different coordinates' values for the points **a** and **b**, we need to generate a quasi-random matrix of Sobol' numbers of size $(R, 2k)$, with $R > r$, and discard the first few points for the auxiliary points, i.e. shifting the k rightmost columns downward.[6] We obtain good results by systematically discarding four points $(R = r + 4)$. Thus, denoting as $x_{i,j}$ the element of the matrix of quasi-random Sobol' numbers in row i and column j, the first base point is $\mathbf{a} = (x_{1,1}, x_{1,2}, \ldots, x_{1,k})$ and its auxiliary point is $\mathbf{b} = (x_{4,k+1}, x_{4,k+2}, \ldots, x_{4,2k})$.

Note that for each couple of **a**, **b** points one need to estimate the model Y a total of $(k+1)$ times, so that the total cost of the analysis is $r(k+1)$, with r the number of repetitions.[7] We have also tested 'higher economy' configurations, such as the one described by the following two steps:

1. Using LP_τ rows one and two for the first couple **a**, **b**. We have to compute the output for the point[8] $x_i^{(1)} x_{\sim i}^{(1)}$ as well as for the k points $x_i^{(2)} x_{\sim i}^{(1)}$. Thus k effects for $(k+1)$ simulations.
2. Using LP_τ rows one and three for the second couple **a**, **b**. As above, but we have already point $x_i^{(1)} x_{\sim i}^{(1)}$, so we only need to compute the k points $x_i^{(3)} x_{\sim i}^{(1)}$. Now we obtain k effects from the couples $x_i^{(1)} x_{\sim i}^{(1)}$, $x_i^{(3)} x_{\sim i}^{(1)}$ as well as an additional k effects from the couples $x_i^{(2)} x_{\sim i}^{(1)}$, $x_i^{(3)} x_{\sim i}^{(1)}$.

[6] In the following we refer to this strategy as the shifted LP_τ sequence.

[7] Note that, in the radial strategy, the model is estimated for the baseline point $y(a_1, a_2, a_3, \ldots, a_k)$ and for the k additional points $y(a_1, a_2, \ldots, b_i, \ldots, a_k)$, though it is not estimated for the auxiliary point $y(b_1, b_2, b_3, \ldots, b_k)$.

[8] We indicate as $x_i^{(1)} x_{\sim i}^{(1)}$ the point where all coordinates come from the first row of the sampling matrix, and as $x_i^{(2)} x_{\sim i}^{(1)}$, with $i = 1, 2, \ldots, k$, the k points where all coordinates are from the first row of the sampling matrix but the coordinate ith, which is from the second row of the same matrix.

Table 2
Ten-dimensional Sobol' quasi-random sequence, first eight points. The support of each dimension is [0, 1].

0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000	0.5000
0.2500	0.7500	0.2500	0.7500	0.2500	0.7500	0.2500	0.7500	0.7500	0.2500
0.7500	0.2500	0.7500	0.2500	0.7500	0.2500	0.7500	0.2500	0.2500	0.7500
0.1250	0.6250	0.8750	0.8750	0.6250	0.1250	0.3750	0.3750	0.8750	0.6250
0.6250	0.1250	0.3750	0.3750	0.1250	0.6250	0.8750	0.8750	0.3750	0.1250
0.3750	0.3750	0.6250	0.1250	0.8750	0.8750	0.1250	0.6250	0.1250	0.8750
0.8750	0.8750	0.1250	0.6250	0.3750	0.3750	0.6250	0.1250	0.6250	0.3750
0.0625	0.9375	0.6875	0.3125	0.1875	0.0625	0.4375	0.5625	0.8125	0.6875

The configuration above is economic as it gets $3k$ elementary effects from $(1+2k)$ simulations, against the k effects per $(k+1)$ simulations of the basic radial design (Table 1). As discussed in [30], the enhanced efficiency is achieved at the expenses of the exploration of the input factors' space. This was confirmed in the present work, as the 'higher economy' results – not reported here – were not better than the plain economy ones.

As already stressed above, unlike the classic EE sampling strategies described in the previous section, when using a radial design each effect is computed over a different step size, equal to the distance between e.g. $x_i^{(u)}x_{\sim i}^{(u)}$ and $x_i^{(v)}x_{\sim i}^{(u)}$ which is to say the difference between $x_i^{(u)}$ and $x_i^{(v)}$, where u and v denote two rows of the sampling matrix chosen as described above. Under this notation, the absolute value of the elementary effect has been computed as:

$$EE_i = \left| \frac{y(x_i^{(u)}x_{\sim i}^{(u)}) - y(x_i^{(v)}x_{\sim i}^{(u)})}{x_i^{(u)} - x_i^{(v)}} \right| \tag{5}$$

and the screening test μ^* has been taken as the average of r such effects for each factor, as described in formula (3) of the previous section.

Via the screening experiment the modeler gets a first insight of the model behavior and can for instance understand if a revision is needed, due for instance to the fact that some factors intended as important are actually not influencing the model response. In alternative, using the proposed design, if the computational cost of the model allows it, the modeler can increase r, the number of repetitions, up to achieve a sample size compatible with the estimation of the global sensitivity index ST. In this case one must replace the estimator (5) for the EE method with an estimator for ST (Eq. (1)). As discussed in [30] we recommend the estimator of Jansen [14,13]:

$$E_{\mathbf{X}_{\sim i}}\left(V_{X_i}(Y \mid \mathbf{X}_{\sim i})\right) = \frac{1}{2r}\sum_{j=1}^{r}\left(y\left(a_1^{(j)}, a_2^{(j)}, \ldots, a_k^{(j)}\right) - y\left(a_1^{(j)}, a_2^{(j)}, \ldots, b_i^{(j)}, \ldots, a_k^{(j)}\right)\right)^2,$$

where $\mathbf{a}^{(j)}$ is the jth baseline point and $\mathbf{b}^{(j)}$ is the jth auxiliary point. An example of convergence towards the total sensitivity indices is presented in Section 4.

As already mentioned, another advantage of the radial design is that it is a natural extension of the OAT approach preferred by most modelers. While OAT *per se* is inefficient [23], its iterated form here described is a good practice.

4. Numerical experiments

4.1. Test cases

The radial design presented in the previous section is compared with the optimized trajectories strategy [3] and with the cell-based design [29] by performing SA experiments on the following set of functions with $k = 20$ input factors:

1. A modified version of the Sobol' G function [1] introduced in [30]:

$$G^*(X_1, \ldots, X_k; a_1, \ldots, a_k, \delta_1, \ldots, \delta_k, \alpha_1, \ldots, \alpha_k) = \prod_{i=1}^{k} g_i^*, \tag{6}$$

$$g_i^* = \frac{(1+\alpha_i)\cdot|2(X_i+\delta_i - I[X_i+\delta_i]) - 1|^{\alpha_i} + a_i}{1+a_i}, \tag{7}$$

where X_i are the input factors, uniformly distributed in the interval $[0, 1]$. $a_i > 0$ are the traditional G function parameters, $\delta_i \in [0, 1]$, $\alpha_i > 0$ are shift and curvature parameters respectively, while $I[X_i + \delta_i]$ is the integer part of $(X_i + \delta_i)$. The relative importance of the factors depends on the choice of the parameters. Two different sets of parameters are investigated (see Table 3) where the number of important factors is either 4 (G_4^* function) or 10 (G_{10}^* function).
2. The function introduced by Bratley et al. [2] and also used in [16]:

$$K = \sum_{i=1}^{k}(-1)^i \prod_{j=1}^{i} X_j, \tag{8}$$

where X_i are uniformly distributed in the interval $[0, 1]$. In this function there are few dominant variables (basically X_1 and X_2) and interaction increases for variables with higher indices.
3. The non-additive function by Saltelli et al. in [28]:

$$B = \sum_{i=1}^{m} X_i \cdot X_{m+i}, \tag{9}$$

where $m = k/2$, $X_i \sim N(\overline{X_i}, \sigma_{X_i})$, $i = 1, 2, \ldots, k$ and $N(X, \sigma)$ denotes the normal distribution with mean X and standard deviation σ. Depending on the choice of the normal distribution parameters, the importance of the factors and the role of the interactions can be calibrated. We choose a setting resulting in a moderate number of important factors. However it has to be stressed that, opposite to the case of the G^* and K functions, the non-important factors have a non-nihil effect.
4. The Morris function [18]:

$$M = \beta_0 + \sum_{i=1}^{20}\beta_i w_i + \sum_{i<j}^{20}\beta_{i,j} w_i w_j + \sum_{i<j<l}^{20}\beta_{i,j,l} w_i w_j w_l \tag{10}$$

$$+ \sum_{i<j<l<s}^{20}\beta_{i,j,l,s} w_i w_j w_l w_s, \tag{11}$$

where $w_i = 2(X_i - 0.5)$ for all i except for $i = 3, 5, 7$ where $w_i = 2(1.1X_i/(X_i + 0.1) - 0.5)$, $X_i \sim U[0, 1]$, and

Table 3
Parameters and total order sensitivity indices for the test functions used in the SA experiments. *ST* corresponds to analytic values while $\widehat{ST}$ to estimated values.

	G_4^*			G_{10}^*			K	B			M
	a	α	ST (%)	a	α	ST (%)	ST (%)	$\bar{X}$	σ_X	ST (%)	$\widehat{ST}$ (%)
1	100	1	0.00	100	1	0.00	75.00	0	0.5	0.39	26.00
2	0	4	69.32	0	4	65.57	25.00	0	0.5	0.62	25.60
3	100	1	0.00	100	1	0.00	8.33	0	1	1.55	10.10
4	100	1	0.00	100	1	0.00	2.78	0	1	1.55	25.70
5	100	1	0.00	100	1	0.00	0.93	0	2	12.41	10.80
6	100	1	0.00	100	1	0.00	0.31	0	2	22.34	9.50
7	1	0.5	3.28	1	0.4	8.21	0.10	0	1	3.49	5.80
8	0	3	55.80	10	3	52.78	0.03	0	0.5	1.40	12.30
9	100	1	0.00	0	0.8	20.24	0.01	0	1.5	20.25	13.80
10	100	1	0.00	0	0.7	17.37	0.00	0	2	36.00	12.40
11	0	2	38.68	9	2	41.23	0.00	1	2	0.31	0.20
12	100	1	0.00	0	1.3	32.73	0.00	2	2	0.31	0.20
13	100	1	0.00	100	1	0.00	0.00	2	1	0.31	0.10
14	100	1	0.00	100	1	0.00	0.00	2	1	0.31	0.30
15	1	0.5	3.28	4	0.3	5.06	0.00	3	1	1.24	0.30
16	100	1	0.00	100	1	0.00	0.00	3	3	11.17	0.20
17	100	1	0.00	100	1	0.00	0.00	1.5	3	2.79	0.20
18	0	1.5	27.05	7	1.5	33.75	0.00	3	3	0.70	0.10
19	100	1	0.00	100	1	0.00	0.00	2	5	17.46	0.20
20	1	0.5	3.28	2	0.6	13.92	0.00	2	5	31.03	0.20

$$\beta_i = 20, \quad i = 1, 2, \ldots, 10,$$

$$\beta_{i,j} = -15, \quad i = 1, 2, \ldots, 6,$$

$$\beta_{i,j,l} = -10, \quad i = 1, 2, \ldots, 5,$$

$$\beta_{i,j,l,s} = 5, \quad i = 1, 2, \ldots, 4.$$

The remaining first- and second-order coefficients for this function are independently generated from a standard normal distribution. The remaining third-order coefficients are set to zero. In this function the first 10 factors are important either because of their main effect (X_8, X_9, X_{10}) or because of combination of main effect and interaction effects. The remaining factors ($X_{11}, \ldots, X_{20}$) are non-relevant.

The values of the parameters and the corresponding total order sensitivity indices are reported in Table 3 as follows:

- Function G^*. We consider two different cases, with 4 and 10 important factors, labeled as G_4^* and G_{10}^* respectively. In the table we present analytic values for the total order sensitivity indices for these functions. Note that the sensitivity indices do not depend on the coefficients δ_i [30] and thus the values of the δ_i are randomly assigned.
- Functions K and B: analytic values [30].
- Function of Morris: the *ST* are estimated via an experiment performed using the standard method [22] with sample size $N = 172{,}032$.

4.2. Simulation results

For each function, the measure μ^* is estimated using the radial OAT sampling, the optimized trajectories design, and the cell-based strategy. For the radial sampling we present the results obtained using the shifted LP_τ sequences as described in Section 3. For the optimized trajectories design, we use for each factor a grid of $p = 4$ levels and, following [18], we set $\Delta = p/[2(p-1)]$. It is worth noting that, while in the trajectories and in the radial sampling for each r we derive one EE per input, when using the cell method we obtain two EEs per input, as these correspond to the two sides of the square. For this reason, when comparing SA results, we build r trajectories for the optimized trajectories method, r points for the radial strategy and $s = r/2$ squares for the cell sampling. r is varied in the set $\{2, 4, 8\}$, in order to investigate the performance of the methods at very low sample size ($r = 2$) or at sample sizes typical of the EE method ($r = 4$–8).

The computational cost of each SA exercise for the optimized trajectories and the radial OAT strategies varies between $N = 42$ (i.e. $N = r(k+1) = 2 \cdot 21$) and $N = 168$ (i.e. $N = r(k+1) = 8 \cdot 21$), corresponding respectively to $r = 2$ and $r = 8$ points/trajectories. For the cell-based method the cost is higher and ranges between $N = 49$ and $N = 196$ model evaluations (for details see [29]). A Monte Carlo experiment is performed where the SA exercise is repeated 1000 times for each method.

We use μ^* values to ascertain relevant factors from non-relevant ones and define the following criterion to compare results:

- Based on the total order sensitivity indices (analytic *ST* or estimated $\widehat{ST}$), a relevant factor is defined as one whose *ST* is greater than $1/k$, where k is the number of input factors. Let us label as φ the set of important factors and let I be the cardinality of this set. For all our functions $k = 20$ and thus only factors with $ST > 5\%$ are considered to be relevant.
- Using Table 3, this implies that the number of relevant factors for the functions G_4^*, G_{10}^*, K, B and M are respectively $I = 4$ (i.e. factors X_2, X_8, X_{11}, X_{18}), $I = 10$ (i.e. factors X_2, X_7, X_8, X_9, X_{10}, X_{11}, X_{12}, X_{15}, X_{18}, X_{20}), $I = 3$ (i.e. factors X_1, X_2, X_3), $I = 7$ (i.e. factors X_5, X_6, X_9, X_{10}, X_{16}, X_{19}, X_{20}) and $I = 10$ (i.e. factors $X_1, X_2, \ldots, X_{10}$).
- In the screening experiments factors are ranked according to μ^*. Let us label as ξ the group of the first I important factors according to μ^*.
- The factors in φ and ξ are compared and the following score g is assigned:

$$g = \frac{card(\varphi \bigcap \xi)}{I}, \tag{12}$$

where *card* is the cardinality of the set. For instance, in the case of the G_4^* function ($I = 4$), if one relevant factor is not identified by the method, the intersection between φ and ξ will be composed of 3 elements and g will be equal to 0.75 ($= 3/4$).[9]

[9] Note that from the definition of the score g, it follows that whenever an important factor is not recognized as relevant, a non-important factor is recognized to be relevant.

Table 4
Results of the SA experiments: statistics for the g score defined in formula (12). Statistics are computed using 1,000 replicas.

Function	r	Radial OAT		Optimized trajectories		Cell	
		Average	StDev	Average	StDev	Average	StDev
G_4^*	2	0.8852	0.1411	0.7933	0.1720	0.6935	0.1893
	4	0.9147	0.1306	0.8023	0.1608	0.7628	0.1762
	8	0.9530	0.1015	0.9043	0.1276	0.8227	0.1602
G_{10}^*	2	0.9956	0.0210	0.8758	0.1097	0.8086	0.1197
	4	0.9973	0.0168	0.8995	0.0993	0.8884	0.1037
	8	0.9998	0.0045	0.9885	0.0337	0.9483	0.0786
K	2	0.9907	0.0550	0.8587	0.1797	0.8553	0.1732
	4	0.9920	0.0510	0.8833	0.1652	0.9007	0.1525
	8	0.9970	0.0105	0.9197	0.1426	0.9347	0.1324
B	2	0.8614	0.0684	0.9053	0.0915	0.8871	0.0982
	4	0.9790	0.0510	0.9696	0.0592	0.9520	0.0713
	8	0.997	0.006	0.9949	0.0274	0.9846	0.0444
M	2	0.9266	0.0700	0.9628	0.0559	0.9000	0.0781
	4	0.9799	0.0428	0.9972	0.0165	0.9463	0.0627
	8	0.9992	0.0089	1	0	0.9886	0.0321

Since the peculiarity of the cell design is to provide the additional measure of the model's interactions defined in Eq. (4), for this method relevant factors are identified by making use of both μ^* and this second measure, i.e. one factor is important if at least one of the tests (μ^* or Eq. (4)) captures it.

Table 4 shows the statistics for g (mean and standard deviation), for each function and each design, using the 1000 MC simulations. Results demonstrate that the radial design visibly outperforms the other methods for the G^* and the K. For the B function the radial design outperforms the other two methods only slightly for $r = 4, 10$ while for $r = 2$ the optimized EE design provides the best results. For the M function, the best results are obtained with the optimized EE design; only for high sample size ($r = 8$) the methods give similar results. On average over all results the radial method is best. Also in light of its convergence properties, the radial design can be considered the best practice.

To further investigate the relative performance of the methods, the distribution of the number of unrecognized important factors is shown in Figs. 4–6. Within each figure, different columns of graphs refer to different functions and each of the three rows of plots show results for one of the selected strategy. The bars in each graph count the number of MC runs where the strategy fails in recognizing one or more important factors, as defined by the criteria presented above. All the sample sizes tested are presented, i.e. $r = 2$ (black bars), $r = 4$ (grey bars) and $r = 8$ (white bars). The graphs show that for the radial design the most populated bucket is always the 0 one (except for the B and M functions and $r = 2$), i.e. the design allow recognizing in most of the runs all relevant factors. Moreover, except in two cases (B and M functions, $r = 2$), the distribution for the radial OAT is always decreasing and the last bucket is rarely populated. As expected, for all strategies the size of the first bar is increasing with increasing the number of points/trajectories/cells while the size of the other bars is decreasing, i.e. the strategies improve their performance with increasing the sample size.

Fig. 7 shows for the three significant factors of the K function the convergence of the estimated variances V_{Ti} obtained using the radial design, towards its analytical value. We show on the x axis the logarithm of the number of considered trajectories/points. The left vertical axis shows the values of the output conditional variances V_{Ti} as defined in Eq. (1). The solid lines with the squares plot the estimated conditional variances which converge towards the analytical values represented by the dashed lines. The right axis present for exemplification only the value of μ^* for the three factors.

This convergence exercise thus confirms that the proposed design represents a unified approach valuable both for screening purposes and in quantitative SA settings. Note that the proposed strategy also allows to estimate error bounds for the effect of each factor. This can be done using bootstrap as proposed in [7]. As an example we use again the K function. Table 5 shows for its three most significant factors the analytic variances, the estimated variances at $r = 512$, and the respective standard errors based on 25 boot replicas.

5. Conclusions

The present paper suggests a possible unified practice covering both screening and quantitative sensitivity analysis, using the same design and the same sample to move seamlessly from the former to the latter.

The practice involves:

1. Determine a plausible space of variation for the uncertain input factors and mapping those to the unit hyper-cube.
2. Start with a set of few different baseline points and use a radial design to complete the sample: for each point build a 'star' with factors randomly shifted around the baseline. Quasi-random Sobol' numbers are used to build the radial OAT sample.
3. Estimate the effect of each factor by using existing screening measures.
4. If computationally affordable, using the same design and increasing the sample size, estimate total order sensitivity indices based e.g. on Jansen's formula [14,13].
5. Use bootstrap to compute error bounds for the effect of each factor [1,30].
6. Use a stopping rule based on (3) above to decide if increasing the number of data points.

This approach allows modelers to change one factor at a time, within each 'star', as well as to increase the number of stars without wasting the points already explored. The approach is straightforward to implement based on the *formulae* given in the present paper, and the quasi-random sequences are freely available.[10]

[10] http://sensitivity-analysis.jrc.ec.europa.eu/software/index.htm.

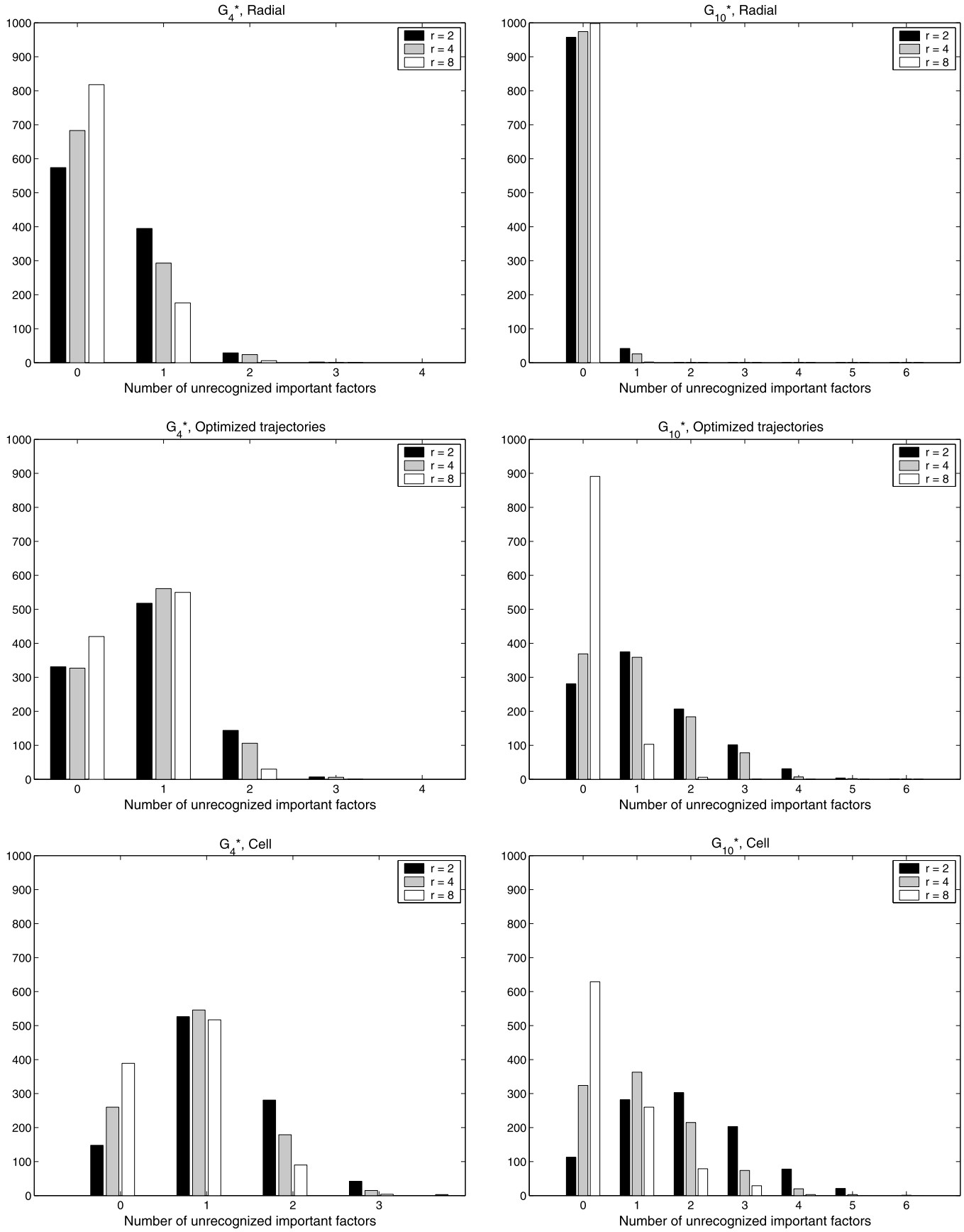

Fig. 4. Distributions of the number of unrecognized relevant factors for the two G^* functions. The distributions are built using results of the 1000 Monte Carlo replicas of the SA experiment.

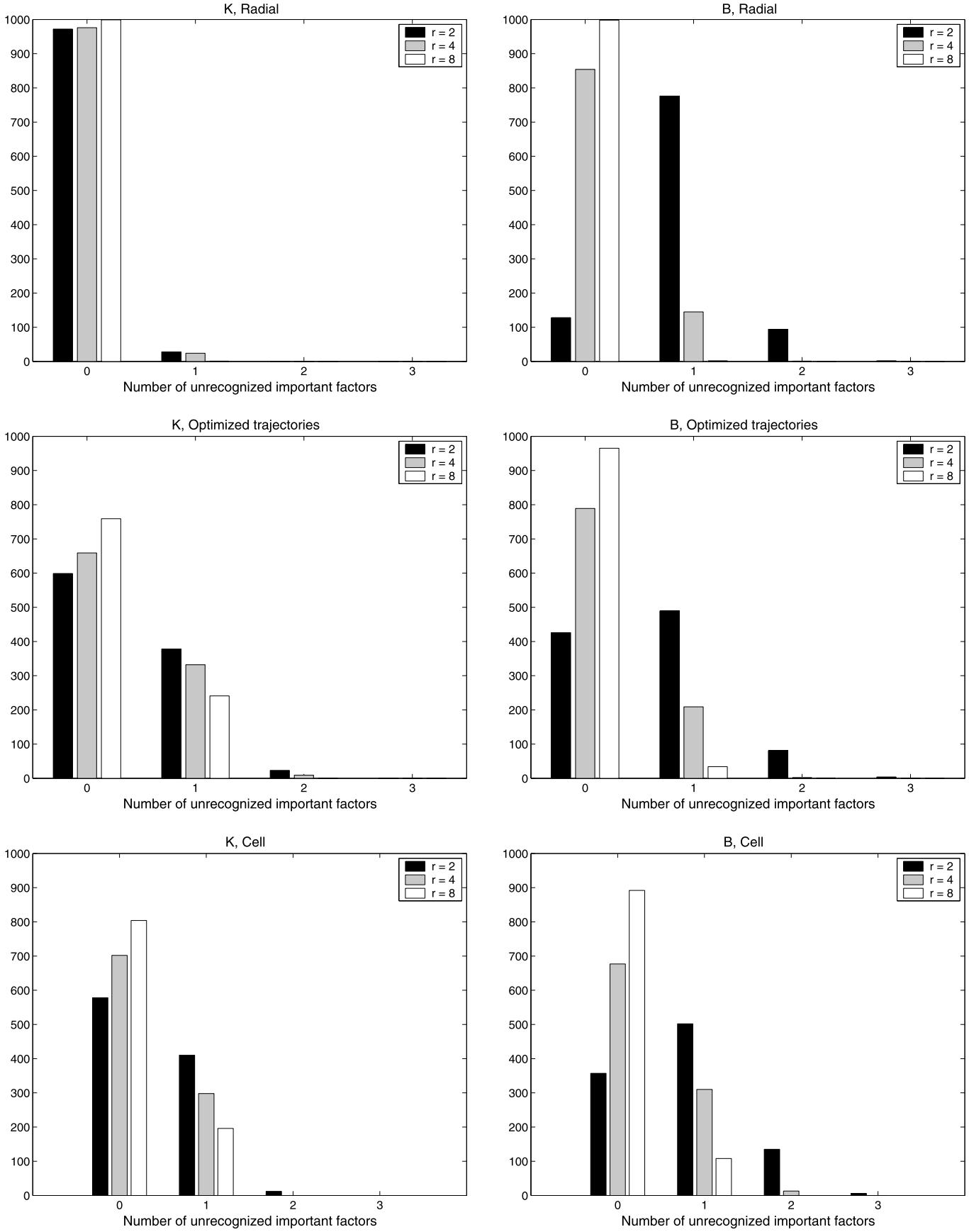

Fig. 5. Distributions of the number of unrecognized relevant factors for functions K and B. The distributions are built using results of the 1000 Monte Carlo replicas of the SA experiment.

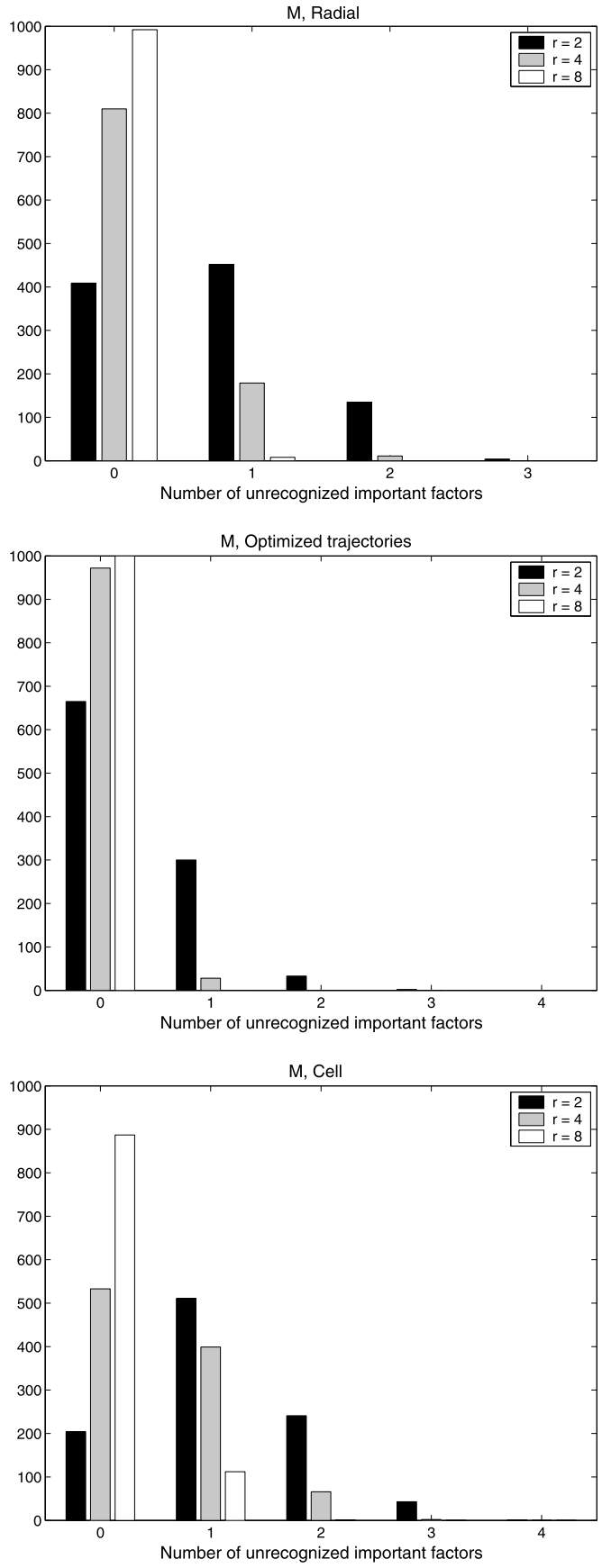

Fig. 6. Distribution of the number of unrecognized relevant factors for function M. The distribution is built using results of the 1000 Monte Carlo replicas of the SA experiment.

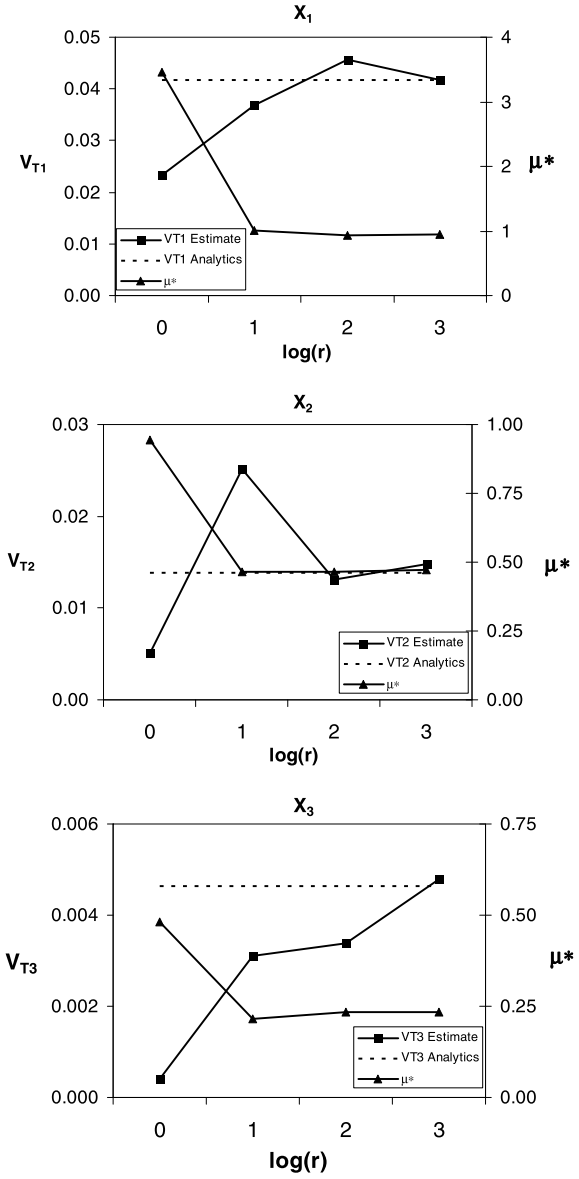

Fig. 7. Left vertical axis: convergence of the estimated variances for the three significant factors of the K function with increasing the sample size r. The right vertical axis shows the value of the screening measure μ^*.

Table 5
Analytic variances for factors one to three of the K function together with the estimated variances at $r = 512$ and the respective standard errors based on 25 boot replicas.

Factor	V_{Ti} analytic	$\hat{V}_{Ti}$ estimated	Standard error
X_1	0.04166	0.0427	0.0026
X_2	0.01389	0.0138	0.0009
X_3	0.00463	0.0036	0.0004

References

[1] G. Archer, A. Saltelli, I. Sobol, Sensitivity measures, anova-like techniques and the use of bootstrap, Journal of Statistical Computation and Simulation 58 (1997) 99–120.

[2] B. Bratley, B.L. Fox, H. Niederreiter, Implementation and test of low discrepancy sequences, ACM Transactions on Modeling and Computer Simulation 2 (3) (1992) 195–213.

[3] F. Campolongo, J. Cariboni, A. Saltelli, An effective screening design for sensitivity analysis of large models, Environmental Modelling and Software 22 (2007) 1509–1518.

[4] J. Cariboni, D. Gatelli, R. Liska, A. Saltelli, The role of sensitivity analysis in ecological modelling, Ecological Modelling 203 (2007) 167–182.
[5] K. Chan, A. Saltelli, S. Tarantola, Winding stairs: A sampling tool to compute sensitivity indices, Statistics and Computing 10 (2000) 187–196.
[6] EC. Impact assessment guidelines, 15 January 2009. Technical Report 92, SEC, 2009, p. 24, accessed April 10, http://ec.europa.eu/governance/impact/docs/key_docs/iag_2009_en.pdf, 2009.
[7] B. Efron, R.J. Tibshirani, An Introduction to Bootstrap, Chapman and Hall/CRC, 1998.
[8] EPA. Guidance on the development, evaluation, and application of environmental models. Technical report, Office of the Science Advisor, Council for Regulatory Environmental Modeling, March 2009. EPA /100/K-09/003, p. 26, accessed April 10, http://www.epa.gov/crem/library/cred_guidance_0309.pdf, 2009.
[9] J. Hall, S. Boyce, R. Dawson, S. Tarantola, Y. Wang, A. Saltelli, Sensitivity analysis for hydraulic models, Journal of Hydraulic Engineering 135 (11) (2009) 959–969.
[10] J.C. Helton, F.J. Davis, Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems, Reliability Engineering and System Safety 81 (1) (2003) 23–69.
[11] J.C. Helton, J.D. Johnson, C.J. Salaberry, C.B. Storlie, Survey of sampling-based methods for uncertainty and sensitivity analysis, Reliability Engineering and System Safety 91 (1996) 1175–1209.
[12] T. Homma, A. Saltelli, Importance measures in global sensitivity analysis of model output, Reliability Engineering and System Safety 52 (1) (1996) 1–17.
[13] M.J.W. Jansen, Analysis of variance designs for model output, Computer Physics Communication 117 (1999) 35–43.
[14] M.J.W. Jansen, W.A.H. Rossing, R.A. Daamen, Monte Carlo estimation of uncertainty contributions from several independent multivariate sources, in: J. Gasmanand, G. van Straten (Eds.), Predictability and Nonlinear Modelling in Natural Sciences and Economics, Kluwer Academic Publishers, Dordrecht, 1994, pp. 334–343.
[15] J.P.C. Kleijnen, Design and Analysis of Simulation Experiments, Springer, 2008.
[16] S. Kucherenko, M. Rodriguez-Fernandez, C. Pantelides, N. Shah, Monte Carlo evaluation of derivative-based global sensitivity measures, Reliability Engineering and System Safety 94 (2009) 1135–1148.
[17] M.D. McKay, R.J. Beckman, W.J. Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics 21 (2) (1979) 239–245.
[18] M.D. Morris, Factorial sampling plans for preliminary computational experiments, Technometrics 33 (2) (1991) 161–174.
[19] J.E. Oakley, A. O'Hagan, Probabilistic sensitivity analysis of complex models: A Bayesian approach, Journal of the Royal Statistical Society Series B 66 (2004) 751–769.
[20] OMB. Proposed risk assessment bulletin. Technical report, The Office of Management and Budget's – Office of Information and Regulatory Affairs (OIRA), January 2006, 010906.pdf, p. 16–17, accessed April 10, 2009, http://www.whitehouse.gov/omb/inforeg/proposed_risk_assessment_bulletin_010906.pdf, 2009.
[21] E. De Rocquigny, N. Devictor, S. Tarantola (Eds.), Uncertainty in Industrial Practice, Wiley, 2008.
[22] A. Saltelli, Making best use of model valuations to compute sensitivity indices, Computer Physics Communications 145 (2002) 280–297.
[23] A. Saltelli, P. Annoni, How to avoid a perfunctory sensitivity analysis, Environmental Modelling and Software 25 (2010) 1508–1517.
[24] A. Saltelli, S. Tarantola, K. Chan, Quantitative model-independent method for global sensitivity analysis of model output, Technometrics 41 (1) (1999) 39–56.
[25] A. Saltelli, K. Chan, M. Scott (Eds.), Sensitivity Analysis Probability and Statistics Series, John Wiley and Sons, 2000.
[26] A. Saltelli, S. Tarantola, F. Campolongo, M. Ratto, Sensitivity Analysis in Practice. A Guide to Assessing Scientific Models, Wiley, 2004.
[27] A. Saltelli, M. Ratto, S. Tarantola, F. Campolongo, Sensitivity analysis for chemical models, Chemical Reviews 105 (7) (2005) 2811–2828.
[28] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, Global Sensitivity Analysis. The Primer, John Wiley and Sons, 2008.
[29] A. Saltelli, F. Campolongo, J. Cariboni, Screening important inputs in models with strong interaction properties, Reliability Engineering and System Safety 94 (7) (2009) 1149–1155.
[30] A. Saltelli, P. Annoni, I. Azzini, F. Campolongo, M. Ratto, S. Tarantola, Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, Computer Physics Communication 181 (2) (2010) 259–270.
[31] T.J. Santner, B.J. Williams, W.I. Notz, Design and Analysis of Computer Experiments, Springer-Verlag, 2003.
[32] I.M. Sobol', On the distribution of points in a cube and the approximate evaluation of integrals, USSR Computational Mathematics and Mathematical Physics 7 (1967) 86–112.
[33] I.M. Sobol', Uniformly distributed sequences with an addition uniform property, USSR Computational Mathematics and Mathematical Physics 16 (1976) 236–242.
[34] I.M. Sobol', Sensitivity analysis for non-linear mathematical models, Mathematical Modelling and Computational Experiment 1 (1993) 407–414;
Translated from Russian: I.M. Sobol', Sensitivity estimates for nonlinear mathematical models, Matematicheskoe Modelirovanie 2 (1990) 112–118.
[35] I.M. Sobol', Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Mathematics and Computers in Simulation 55 (2001) 271–280.
[36] I.M. Sobol', Global sensitivity analysis indices for the investigation of nonlinear mathematical models, Matematicheskoe Modelirovanie 19 (11) (2007) 23–24, in Russian.
[37] I.M. Sobol', S. Kucherenko, On global sensitivity analysis of quasi-Monte Carlo algorithms, Monte Carlo Methods and Applications 11 (1) (2005) 1–9.
[38] I.M. Sobol', S. Kucherenko, Derivative based global sensitivity measures and their link with global sensitivity indices, Mathematics and Computers Simulation 79 (10) (2009) 3009–3017.
[39] I.M. Sobol, S. Kucherenko, A new derivative based importance criterion for groups of variables and its link with the global sensitivity indices, Computer Physics Communications 181 (7) (2010) 1212–1217.