ELSEVIER

0951-8320(95)00099-2

Reliability Engineering and System Safety **50** (1995) 225–239

0951-8320/95/$9.50

About the use of rank transformation in sensitivity analysis of model output

Andrea Saltelli[a] & Ilya M. Sobol'[b]

[a]*Atmospheric Chemistry Unit of the Environment Institute, Joint Research Centre of the EC Ispra I-21020 Italy*
[b]*National Centre for Mathematical Modelling of the Russian Academy of Science 4A Miusskaya Square, 125047 Moscow (CIS)*

(Received 3 May 1995; accepted 9 September 1995)

Rank transformations are frequently employed in numerical experiments involving a computational model, especially in the context of sensitivity and uncertainty analyses. Response surface replacement and parameter screening are tasks which may benefit from a rank transformation. Ranks can cope with nonlinear (albeit monotonic) input-output distributions, allowing the use of linear regression techniques. Rank transformed statistics are more robust, and provide a useful solution in the presence of long tailed input and output distributions.

As is known to practitioners, care must be employed when interpreting the results of such analyses, as any conclusion drawn using ranks does not translate easily to the original model. In the present note an heuristic approach is taken, to explore, by way of practical examples, the effect of a rank transformation on the outcome of a sensitivity analysis. An attempt is made to identify trends, and to correlate these effects to a model taxonomy.

Employing sensitivity indices, whereby the total variance of the model output is decomposed into a sum of terms of increasing dimensionality, we show that the main effect of the rank transformation is to increase the relative weight of the first order terms (the 'main effects'), at the expense of the 'interactions' and 'higher order interactions'.

As a result the influence of those parameters which influence the output mostly by way of interactions may be overlooked in an analysis based on the ranks. This difficulty increases with the dimensionality of the problem, and may lead to the failure of a rank based sensitivity analysis.

We suggest that the models can be ranked, with respect to the complexity of their input-output relationship, by mean of an 'Association' index I_y. I_y may complement the usual model coefficient of determination R_y^2 as a measure of model complexity for the purpose of uncertainty and sensitivity analysis.

1 INTRODUCTION

Numerical experiments meant to assess the uncertainty in model prediction belong to the domain of Uncertainty Analysis (UA). Although UA is not restricted to the uncertainty originating from the input data (see, for instance, the discussion on structural uncertainty in Ref. 13), uncertainty in the value of the model input parameters is surely a major concern in modelling. On this note we focus on those numerical experiments—often implemented together with (or as part of) a UA—which aim to apportion the output uncertainty to the uncertainty in the input parameters. This is usually referred to as Sensitivity Analysis (SA).

In fact the definition 'Sensitivity analysis' may be applied to a number of very different problem settings, such as:

1. Parameter screening, where the task is to identify active factors in a system with many parameters. Examples of this approach are Refs 8, 39, 35, 2, 52, 1.
2. Global SA, where the emphasis is on apportioning the output uncertainty to the uncertainty in the input parameters. In this approach the uncertainty range given in input reflects our imperfect knowledge of those parameters. Often this kind of analysis is implemented by means of regression, correlation techniques (see eg Refs 32, 25, 23, 26, 28, 22, 17, 16, 18). One particular

case of global SA is a variance analysis such as that produced by the Fourier Amplitude Sensitivity Test (FAST,[10,11,46,34,31]) or by Sobol' sensitivity indices.[49]

3. System analysis by way of local sensitivities, where the emphasis is on the impact of the parameters not of the model variance but of the model itself. This is usually achieved by computing partial derivatives of the output functions with respect to the input variables (Jacobian). To this effect, in order to numerically compute the derivative, the input parameters are given a small interval of fractional variation around the nominal value. The interval is usually the same for all variables and has no relation to our degree of knowledge of the variables.[21,4,6,38]

A recent review of UA techniques was published in this journal;[15] strategies for SA are compared in Refs 24, 16, 42. Rank transformation is a fairly common procedure in uncertainty and sensitivity analysis of model output, and has found applications especially in settings (1) and (2) above. This derives mostly from the use of the rank transform in regression. A regression based sensitivity analysis aims to screen model input parameters based on the absolute value of the regression coefficients of the parameters. For instance, standardised regression coefficients (SRC) can be used, as well as correlation measures such as the Pearson coefficient, and the partial correlation coefficients (PCC).[12] This approach may become impractical for nonlinear models, especially when the model output has a long tailed distribution. Insofar as the input output relationship is monotonic, a rank transformation of the problem may allow a successful sensitivity analysis at no extra computational cost, replacing, for instance, the SRC by their rank equivalent (the standardised rank regression coefficients SRRC). A discussion of this approach is given in the appendix of Ref. 17.

The problem setting for the present investigation can be stated as follows:

- A model $\mathbf{y} = f(x)$ is given, where $\mathbf{y} \equiv (Y_1, \ldots, Y_r)$ is a vector of output variables linked through the model f to a set of input variables $x \equiv (X_1, \ldots, X_n)$. Throughout this work we shall assume that the X_j's are independent. Each X_j is characterised by its probability density function which is assumed as given.

  The model can be seen as a set of mathematical functions $Y_s \equiv f_s\ (X_1, \ldots, X_n)$, $s = 1, \ldots, k$, but often the form of the function is not available, Y_s being the result of a numerical integration, or, more generally, the prediction of a computer code. The expression 'computational model' has been proposed for such a problem setting.[35] For each output variable Y_s, a vector of N values $(y_s1, \ldots, y_{sN})$ was generated by repeatedly evaluating the model f for a set of N sampled vectors $(\mathbf{x}_{11}, \ldots, \mathbf{x}_{1n}), \ldots, (\mathbf{x}_{N1}, \ldots, \mathbf{x}_{Nn})$
  where n is the number of variables. Each column vector $\mathbf{x}_{ij}$, $(i = 1, 2, \ldots, N)$ is a sample from the probability density function of X_j.

  Subsequently the Y_{si}'s values assumed in the N simulations are replaced by their ranks, $R(Y_s)$, i.e.:

$$R(y_{sr}) = 1 \text{ if } y_{sr} = \text{Max}\,(y_{si} \mid i = 1, 2, \ldots, N);$$

$$R(y_{st}) = N \text{ if } y_{st} = \text{Min}\,(y_{st} \mid i = 1, 2, \ldots, N)$$

- For clarity we shall refer in the following to the 'Y-Rank transformed model', which is the one obtained when $y_s = f_s\ (X_1, \ldots, X_n)$ is replaced by $R(y_s) \equiv f_s^*\ (X_1, \ldots, X_n)$. Note that the transformed model is also a function of the sample, (size, scheme used, random seed, ...). The rank of the output obtained in run number 'i', $R(y_{si})$ depends, in fact, on the entire sample matrix $\underline{\underline{X}}$, and not on the x_{ij}, $j = 1, 2, , , n$ alone. This dependency will be neglected in the formalism, though it will be kept in mind in the following discussion.
- In some instances another kind of transformed model is used in sensitivity analysis, in which not only the output, but also the input x_{ij}'s are ranked, to yield a model of the type $R(y_s) \equiv f_s^{**}\ (R(X_1), \ldots, R(X_n))$ (again neglecting sample dependence). This model f_s^{**} (which could be called a XY-Rank transformed model) is the one used, for instance, when computing the Spearman correlation coefficient of the Standardised Rank Regression Coefficients. Here the effect of X_j on y_s is estimated by looking at the effect of $R(X_j)$ on $R\ (y_s)$. Note that also, in this setting, the model is always evaluated on the original (raw) data x_{js}'s, and that the ranks are only taken afterward.
- The outcome of a sensitivity analysis based on the ranks is hence compared with that obtained without the rank transformation.

Rank transformations are associated with the 'multiple evaluation' of the model, which entails a sampling strategy (eg pure Monte Carlo, Latin hypercube, importance, stratified, quasi-Monte Carlo, factorial, ...), whose aim is to scan the space of the input parameters. Rank transformations are hence used for global sensitivity analysis rather than for local SA.

In many of the applications quoted above a correlation/regression based SA is performed using techniques such as the Standardised Rank Regression Coefficients (SRRC) and the Partial Rank Correlation Coefficients (PRCC). The linear version of these statistics, ie the standardised regression coefficients SRC or the partial correlation coefficients PCC are used in practice only when the problem is proven to

be linear in all the influential variables in **x**, using, for instance, the model coefficient of determination R_Y^2.[12]

In spite of the widespread use of linear techniques in SA, there are instances in which the rank transformation cannot linearise the model. This happens, for instance, for nonmonotonic models. As shown in Refs 42 and 40, this leads to a failure of SRRC/PRCC based SA. It should be stressed that although automated non-linear regression techniques are available, these are not frequently seen in the literature, the linear rank-based techniques being by far the most popular.

Another reason for taking the ranks may be linked to the limited robustness of SA statistics when applied to the raw values. Robustness of an SA estimator here means its capability to reproduce its prediction (eg on the order of importance of the input parameters) when repeating the analysis on a different sample of the same population. Evidently reproducibility can be achieved by increasing the sample size, but this is not always possible.

Iman and Hora[27] investigate the performance of a sensitivity measure based on the percentage variance in f explained by any variable X_j. This technique is known as measure of importance,[20] and its use is associated with the estimation of the quantity

$$\frac{\mathrm{Var}_{X_j}[E(f|X_j)]}{\mathrm{Var}(f)} \qquad (1)$$

where $E(f|X_j)$ indicates the expectation value of f when the j^{th} variable is fixed to the value Xj, $\mathrm{Var}_{X_j}[\bullet]$ stands for the variance of the argument over all the possible values of X_j and $\mathrm{Var}(f)$ is the unconditional (total) variance of f.

Iman and Hora[27] observe that, although mathematically correct, this importance measure lacks robustness, and can be highly influenced by outliers associated with long tailed input distributions. They suggest an alternative measure based on replacing f with $\log(f)$ and estimating $E(\log(f|X_j))$ using linear regression. This solution has the advantage of robustness, but—as observed by the authors—the conclusions based on $\log(f)$ are not easily converted back to f. Similar consideration apply to the rank transformation. As shown in the examples of Section 3, a given input parameter could be the main contributor to the variance of the rank (or the log) of Y, but not to the variance of Y itself.

A discussion of the increased robustness of the rank version of SA statistics can be found in Refs 44, 42. Rank versions of the importance measure were tested in Refs 40, 19 and 33.

The use of rank with the importance measure is conceptually different from the use of rank in a regression based sensitivity analysis technique. As far as—say—the SRRC's are concerned, the rank transformation is essential to the analysis. It allows the detection of parameters non linearly correlated with the output and which could otherwise be overlooked. In the same spirit the use of the rank version of the importance measure finds its justification in its ability to detect nonlinear *and* nonmonotonic relationships, which could escape detection even using the SRRC's. Yet the importance measure could yield the same result even without the rank transformation. The reason for the transformation lies in this case in the limited robustness of the raw value version of the test, which would make the sample size needed for a reliable SA unacceptably high.

For the reasons discussed above, a rank transformation may be the only viable choice for a given problem. Hence it would be desirable to have an idea of the error involved in replacing f with either f^* (as in the importance measure) or f^{**} (as in the rank regression/correlation techniques). In the following we shall always use the f^*, f^{**} symbols to denote the rank-transformed versions of the model f, although this notation is somewhat improper; as mentioned before, f^* and f^{**} are in fact also influenced by the particular sample of input values taken. In order to guard against possible fluctuations, all the examples discussed in this work were computed at large (>1,000) sample size (the sample dependence decreases with increasing sample size). Furthermore random variation associated with the sample selection would not alter our conclusions.

In order to compare f with f^*, f^{**} two elements are needed. An investigator should first have an analytical tool to quantify the differences between f and f^*, f^{**}. Then, some kind of taxonomy of model linearity or monotonicity should be taken in order to relate the above difference to an accepted scale. The sensitivity indices developed by Sobol'[49] and further elaborated by other investigators[19,43] are an ideal tool to investigate the difference between f and f^*, f^{**}.

The sensitivity indices are discussed in Section 2. A model taxonomy is indeed missing, and our work will proceed mostly by way of examples. In Section 3, two families of functions are taken as reference scales, whereby models are ordered by their linearity, and by the influence of synergistic terms among variables. A practical test case will be discussed as well. The worked examples will show that the difference between f and f^*, although not predictable on a systematic basis, displays a clear typology. In the ANOVA terminology, the variance of a model can be thought of as decomposed in 'main effects', 'interaction effect' and 'higher order interactions' (see eg the factorial design chapters in Ref. 3; see also Ref. 14). Adopting this terminology we can say that the rank transformation tends to make the model more additive, reducing the relative contribution of the 'interaction' and 'higher order interactions' effects. Given that the possibility of important parameter

interaction increases with the number of model input parameters, the differences between f and f^* are likely to increase with the dimensionality of the problem. This may lead to significant errors when using ranks in sensitivity analysis and in a response surface replacement.

We suggest in Section 4 a new statistic to characterise this aspect of model complexity, an 'Association index' I_y, which will be higher for models where interaction effects among influential parameters plays an important role. Low values of R_y^2 and high values of I_y characterise 'difficult' models.

2 SENSITIVITY INDICES

The sensitivity indices were developed by Sobol',[49] and have conceptual similarities with another well known SA method, the Fourier Amplitude Sensitivity Test (FAST, see references in Section 1). Both methods allow the total model variance D to be written as the sum of terms of different dimension.[11]

Furthermore the sensitivity indices have much in common with the importance measure discussed in Section 1.[20,27] Especially after the modification to the importance measure made by other investigators[29,30,40] the measure of importance was seen to coincide—even from the computational point of view—with the Sobol' sensitivity index of the first order.[19] Sensitivity measures similar to the sensitivity indices have also been used by Cotter[8] and by Sacks *et al.*[39]

A derivation of the Sobol' global sensitivity estimates is given in Ref. 49; computational improvements are given in Ref. 19. Some essential features are repeated here for reader's convenience (Ref. 49 was in Russian; a translation became available in 1993, see references).

We assume that, without loss of generality, each of the X_i's lies with uniform probability between 0 and 1, so that the sample space of $f(\mathbf{x}) = f(x_1, \ldots, x_n)$ is the n-dimensional unit cube:

$$K^n = \{\mathbf{x} \mid 0 \leq x_i \leq = 1; i = 1, \ldots, n\} \tag{2}$$

Under assumptions described in Ref. 49 it is possible to decompose $f(\mathbf{x})$ into summands of different dimensions, eg:

$$f(x_1, \ldots, x_n) = f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum\sum_{1 \leq = i \leq = j \leq = n} f_{ij}(x_i, x_j) + \ldots + f_{12\ldots n}(x_1, x_2, \ldots x_n) \tag{3}$$

In this formula f_0 is a constant, a generic term $f_i(x_i)$ is a function only of variable x_i, the term $f_{ij}(x_i, x_j)$ is a function of only variables x_i, x_j and so on. It can be proven[49] that the expansion (2) is unique whenever $f(x_1, \ldots, x_n)$ is integrable over K^n.

As an example, $f_i(x_i)$ can be computed as

$$f_i(x_i) = \int_0^1 \ldots \int_0^1 f(\mathbf{x}) \mathrm{d}x_{-i} - f_0, \tag{4}$$

where the notation $\mathrm{d}x_{-i}$ means integration over all the variables $x_1, \ldots, x_n$ except for variable x_i. Analogous formulae hold for the higher order terms. The terms in the series development are orthogonal to each other.

At this point the sensitivity index $S_{i_1} \ldots S_{i_s}$ can be introduced:

$$S_{i_1} \ldots S_{i_s} = \frac{D_{i1 \cdots i_s}}{D}, \quad \text{where } D = \int_{K^n} f^2(\mathbf{x}) \, \mathrm{d}\mathbf{x} - f_0^2 \tag{5}$$

is the total variance of $f(\mathbf{x})$ and

$$D_{i1 \ldots i_s} = \int_0^1 \ldots \int_0^1 f_{i1 \ldots i_s}^2 \mathrm{d}x_{i1} \ldots \mathrm{d}x_{is} \tag{6}$$

where $f_{i1 \ldots i_s}$ denotes a generic term of the series development (3). As shown in Ref. 49:

$$D = \sum_{i=1}^{n} D_i + \sum\sum_{1 \leq = i \leq = j \leq = n} D_{ij} + \ldots + D_{12 \ldots n} \tag{7}$$

An analogous expression was derived by Cukier *et al.*[11] in the framework of the FAST method for sensitivity analysis. Similar decompositions are discussed in Refs 8, 9, 14 and 39. A consequence of (7) is that:

$$\overset{\#}{\sum} S_{i_1 \ldots i_s} = 1$$

where the $\overset{\#}{\sum}{}^n$ notation indicates the sum over all the combinations of indices. $S_{i1 \ldots i_s}$ can be considered as true global sensitivity estimates, as they give the fraction of the total variance of $f(\mathbf{x})$ which is due to any individual parameter or combination of parameters. This approach can be compared to that of Sacks *et al.*[39] where inspection of plots of the functions $f_{i1 \ldots i_s}$ in the development (3) is used for SA.

The applicability of the sensitivity indices $S_{i1 \ldots i_s}$ to a large class of functions $f(\mathbf{x})$ is linked to the possibility of evaluating the multidimensional integrals above via Monte Carlo methods. This is detailed in Ref. 49. As an illustration, for a given sample size N the following estimates are possible:

$$\hat{f}_0 = \frac{1}{N} \sum_{m=1}^{N} f(\mathbf{x}_m) \tag{9}$$

where x_m is a sampled point in the space K_n and

$$\hat{D} - \hat{f}^2{}_0 = \frac{1}{N} \sum_{m=1}^{N} f^2(\mathbf{x}_m) \tag{10}$$

The one-indexed terms S_i can be obtained from D_i, which is estimated as

$$\hat{D}_i + \hat{f}^2{}_0 = \frac{1}{N} \sum_{m=1}^{N} f(\mathbf{u}_{im}, x_{im}) f(\mathbf{v}_{im}, x_{im}) \tag{11}$$

where $\mathbf{u}$ and $\mathbf{v}$ denote projections of $\mathbf{x}$ on K^{n-1}, ie the

original space excluding the variable X_i. Eqn (11) indicates that in estimating $f(\mathbf{u}_{im}, x_{im})$ all the X_j's are sampled, while $f(\mathbf{v}_{im}, x_{im})$ is estimated by re-sampling all the X_j's but X_i . The higher order $S_{i1\ldots i_s}$ are estimated in the same manner.

Unfortunately one separate sample (of size N) is needed to compute each of the $S_{i1\ldots i_s}$. Given that the number of terms in the development (7) is $2^n - 1$, and that one sample is needed for f_0, then $N \times (2^n)$ model evaluations are to be computed.

In applications with a large number of variables this number can be prohibitive. For this reason a new statistic S_{Ti} was proposed in Refs 19, 43:

$$S_{Ti} = 1 - S_{ci} \tag{12}$$

where S_{ci} equals the sum of all the $S_{i1\ldots i_s}$ terms where the index 'i' is excluded. Consequently S_{Ti} equals the total effect of variable X_i. Imagine a system with just three input variables; then for variable X_1:

$$S_{T1} = S_1 + S_{12} + S_{13} + S_{123} \tag{13}$$

S_{ci} can be computed with just one Monte Carlo integral, thus reducing the number of model evaluations to $N \times (n + 1)$ i.e. one sample for f_0 plus one sample for each variable. This approach may be classified as a 'one-factor-at-a-time' (OAT) treatment.

Both $S_{i1\ldots i_s}$ and S_{Ti} can be computed on ranks as well, with a net gain in robustness.[40,19] These new measures are indicated with the symbols $S^*_{i1\ldots i_s}$ and S^*_{Ti}. To make an example, the steps needed to compute the first order terms S^*_j are as follows:

- The $N \times (n + 1)$ row vectors $x_{ij}, j = 1, \ldots, n$ are sampled (see details in Refs 40 and 49).
- The corresponding $N \times (n + 1)$ output values y_{si} are computed.
- In order to compute the D^*_j (eqn (11)) the two N-dimensional vectors containing $f(\mathbf{u}_{jm}, x_{jm})$, $f(\mathbf{v}_{jm}, x_{jm})$ are replaced by their ranks. Note that, unlike the case of the rank based correlation measures (eg Spearman) we do not need to rank the input vectors x_{ij}, $j = 1, \ldots, n$.

For the present work the Monte Carlo integrals are computed with large sample sizes, using Sobol' LP_τ sequences for the sampling.[47,48,50] As discussed in Ref. 50, quasi random numbers are characterised by an enhanced convergence, i.e. the $N^{-1/2}$ stochastic convergence rate of the crude Monte Carlo can—in some cases and depending on the nature of the function under investigation—become as large as $N^{-1+\epsilon}$ with an arbitrary small $\epsilon > 0$. Subroutines to generate $LP\tau$ sequences are available.[7,51] The use of $LP\tau$ sequences in multidimensional integration is discussed in Ref. 5. In Section 3 both analytical and numerical formula have been used to estimate $S_{i1\ldots i_s}$; $S^*_{i1\ldots i_s}$ has always been estimated numerically. The error on the numerical estimates was evaluated using as a yardstick for our computations the probable error associated with the crude Monte Carlo estimate; for example, the probable error on $\hat{S}_i$ is:

$$\delta\hat{S}_i = \frac{\delta\hat{D}_i}{\hat{D}} + \frac{\hat{D}_i \delta\hat{D}}{\hat{D}^2} \tag{14}$$

where the errors on $\hat{D}_i$ and $\hat{D}$ are the probable errors, eg:

$$\delta\hat{D}_i \approx \frac{0{\cdot}6745}{\sqrt{N}}\sqrt{F - I^2} \text{ and } Pr\{|D_i - \hat{D}_i| \leq \delta\hat{D}_i\} = 0.5 \tag{15}$$

with

$$F = \frac{1}{N}\sum_{m=1}^{N} [f(\mathbf{u}_{im}, x_{im}) f(\mathbf{v}_{im}, x_{im})]^2 \tag{16}$$

$$I = \frac{1}{N}\sum_{m=1}^{N} f(\mathbf{u}_{im}, x_{im}) f(\mathbf{v}_{im}, x_{im}) \tag{17}$$

In (14, 15) $\hat{f}_0^2$ and its error were neglected. This is justified as the problem is usually scaled before computing the variances, so that $\hat{f}_0^2$ is small. A similar expression can be written for $\delta\hat{D}$.

In the following section we shall use the difference between $S_{i1\ldots i_s}$ and $S^*_{i1\ldots i_s}$ (or their estimates) as a measure of the differences between f and f^*. It is important to note that such an exercise would have been hardly possible without the sensitivity indices. Although the FAST method can in principle yield the same information, its computational complexity would make the evaluation of the higher order indices a major obstacle to the exercise. On this subject see also Ref. 31.

In spite of the simplification imposed by the rank transformation discussed in this article, the rank-based global sensitivity indices S^*_{Ti} are superior to the SRRC's. These are, in fact, unable to screen parameters in the presence of nonmonotonic relationships. On the other hand, the S^*_{Ti} are more expensive to compute (being an OAT test) than the SRRC. For other examples of OAT treatments capable of capturing high dimension effects see also Refs 8, 39 and 35. Those authors are reviewed in Ref. 19. An application of the global sensitivity indices to a chemical reactions system is given in Ref. 41.

OAT techniques are often compared unfavourably against alternative approaches such as factorial design.[3] Nevertheless, the computation of the higher order terms (the interaction effects) using the sensitivity indices differs substantially from the analogous computations done with factorial design. To make an example, the computation of an 'interaction effect' among two variables computed with a factorial design at two level[3] is based on the assumption that the relation between those variables and the output is linear, or at least monotonic. Such assumptions are not needed for the sensitivity indices.

Having introduced the sensitivity indices we are now able to give a working definition of an 'influential', or 'important' parameter, at least in a relative sense.

It is clear that the relative importance of input parameters depends on the sensitivity measure adopted. In local sensitivity analysis X_j is more important than X_k if $|\frac{\partial Ys}{\partial X_j}|$ is greater than $|\frac{\partial Y_s}{\partial X_k}|$ at the selected point $\mathbf{x}^0 = x_1^0, x_2^0, \ldots, x_n^0$. When using a regression method X_j is more important than X_k if it is given a larger weight in the regression model for Y_s. A natural definition of importance, which descends from the sensitivity indices, is the following: X_j is more important (or influential) than X_k if X_j —either by itself or in combination with other parameters— accounts for a larger fraction of the variance of Y_s than that accounted by X_k, i.e. if $S_T(X_j) > S_T(X_k)$.

This also allows us to make precise the meaning of a sensitivity analysis 'error' in the present context. This is an incorrect ranking of the order of importance of the input parameter, the extreme case being the non identification of an influential parameter. As shown in Section 3, the error arises when the 'importance' of X_j to Y_s is markedly different from that of X_j(or $R(X_j)$) to $R(Y_s)$.

Furthermore, given that the contribution of X_j to the variance of $R(y_s)$ is the same if X_j is varied between 0 and 1, or if $R(X_j)$ varies between 1 and the sample size N, then within the scope of the present analysis the difference between f and f^* can be taken as identical to the difference between f and f^{**}. Note that this would not be the case if the sensitivity analysis were to be performed using a regression method, where the input values (or their rank) enter directly into the computation of the sensitivity coefficients (eg Pearson, Spearman, S(R)RC).

3 THE NUMERICAL EXPERIMENT

In this section we explore, through a practical application and some test functions, the nature of the differences between f and f^*, trying to evaluate the error associated with a rank transformation. It may be important to recall at this point that in all our examples the input variables are assumed independent, and that the effect of the sample on f^* is neglected. This is justified as this effect decreases asymptotically with the sample size, and large samples ($N > 1000$) have been used here.

3.1 Level E test case.

The Level E test case[36,40] displays interesting nonmonotonic features that are suitable for testing SA strategies. For this reason it has also been used for an international exercise on sensitivity analysis.[37] Level E simulates the transport of radionuclides from an underground disposal vault containing nuclear waste up to man, by way of migration through a system of natural and engineered barriers. The main barrier considered in the model is the geosphere itself, which includes a two-layer pathway, where nuclide dispersion, advection, retention and radioactive decay are considered. The isotope ^{129}I and the decay chain ^{237}Np-^{233}U-^{229}Th are the migrating species. The system includes twelve uncertain parameters whose input (Table 1) is in the form of probability distributions. The core of the model is the set of partial differential equations which describe the nuclide migration in the geosphere; to give an example, for ^{233}U, the equation is:

$$R^{(k)}_{\mathrm{U}} \frac{\partial F^{(k)}_{\mathrm{U}}}{\partial t} + v^{(k)} \frac{\partial F^{(k)}_{\mathrm{U}}}{\partial x} - v^{(k)} d^{(k)} \frac{\partial^2 F^{(k)}_{\mathrm{U}}}{\partial x^2} = - \lambda_{\mathrm{U}} R^{(k)}_{\mathrm{U}} F^{(k)}_{\mathrm{U}} + \lambda_{\mathrm{N}} R^{(k)}_{\mathrm{N}} F^{(k)}_{\mathrm{N}} \quad (18)$$

where:

U stands for the ^{233}U isotope
N for ^{237}Np
the superscript (k) refers to geosphere layer number k (1 or 2)
R is the nuclide retention (adimensional)
f is the nuclide flux (mol/a)
t is time (a)
v is the water travel velocity in the geosphere layer (m/a)
x is space (m)
d is the dispersion length in the geosphere layer (m)
λ is the nuclide decay constant (1/a)

The geosphere path segments have length $l^{(1)}$ and $l^{(2)}$

The model and its predictions are discussed at length in the references quoted above. Here we show (Fig. 1, triangles) the model coefficients of determination R_y^2 computed on the raw values and on the ranks for a sample size $N = 1024$ using Sobol' quasi random numbers. The R_y^2 coefficient[12] provides a measure of how well the linear regression model based on either SRC's or SRRC's can reproduce the actual output vector $\mathbf{Y}$. In particular:

$$R_y^2 = \frac{\sum_N i(\hat{y}_i - \bar{y})^2}{\sum_N i(y_i - \bar{y})^2} \quad (19)$$

where $\bar{y}$ is the mean of the output values y_i and the $\hat{y}_i$ are the model predictions based on the S(R)RC's, so that R_y^2 represents the fraction of the variance of the output vector explained by the regression. The difference between the SRC's (and the related R_y^2) on

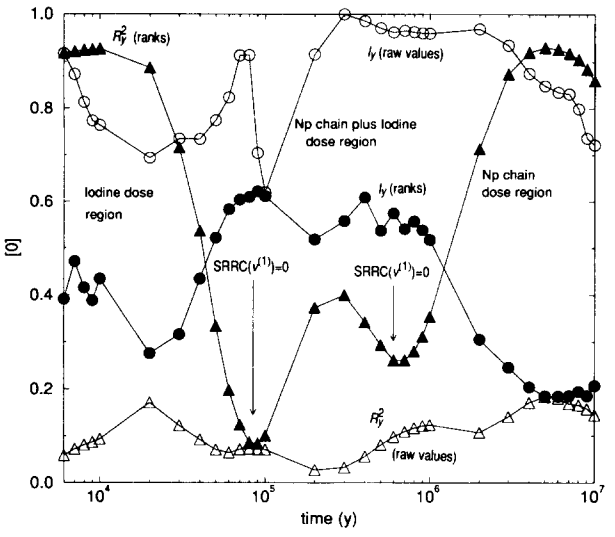

Fig. 1. Model coefficient of determination R_y^2 (triangles) and association index I_y (circles) for the Level E case study as function of time (y). Empty symbols refer to the original model, full symbols to the ranked one. The output variable is always the total annual dose, and R_y^2 always refers to the regression model based on all the 12 variables given in Table 1. The arrows indicate the points where the SRRC of variable $v^{(1)}$ changes its sign. Also highlighted is the region where the total dose is dominated by Iodine(left), by Neptunium decay chain (right) or by both (centre).

one hand and the SRRC's on the other is that the former are based on the raw values of input and output, while the latter are based upon ranks. The closer R_y^2 is to unity the better is the regression model performance and the quality of the associated SA. The R_y^2 values associated with the SRC's are generally lower than that associated with the SRRC's, especially for non-linear models. This is also the case in Fig. 1. The low values of the SRC-based R_y^2 indicates that a SA method based on linear SA estimators (like the SRC, the Pearson test, etc.) are ineffective. The multimodal shape of the R_y^2 curve based on the SRRC similarly indicates that for the low R_y^2 points even the nonparametric estimators (SRRC, Spearman, etc.) are inappropriate. The value of the SRRC's as a function of time for the most important variables is given in Fig. 2. As discussed in Ref. 40 when the fraction of variance accounted for by the regression coefficients is low, it is not possible to rank the parameters based on these coefficients.

In the same article it was shown, by means of input-output scatterplots, that the most influential parameter for most of the time points is the water flow velocity in the first layer of the geosphere ($v^{(1)}$), which is linked to total dose in a nonmonotonic fashion. The local minima of the R_y^2 curve correspond to those points where the total dose vs $v^{(1)}$ scatter plot (on ranks) is almost symmetrically bell-shaped and the corresponding SRRC for this variable passes through zero (Fig. 2, same sample of $N = 1024$, quasi random). Although the influence of $v^{(1)}$ is evident from the scatter plots (not shown here), SRRC predict zero sensitivity for $v^{(1)}$ at those time points, drawing a horizontal line across the scatterplots. Deprived of its most influential parameter, sensitivity analysis fails (R_y^2 is very low on both the raw values and the ranks).

In Fig. 3 the first order sensitivity index ($\hat{S}_i$, empty symbol) and the total effect ($\hat{S}_{Ti}$, full symbol) have been plotted for the three most important variables of the Level E exercise. Each index for each variable has been computed using a different sample of size 1024 simulations using Sobol' LP_τ sequences.

A few computational details can be given here. The constructive dimension (cd) of a Monte Carlo algorithm was defined by Sobol'[50] as the number of

Table 1. Input data for the Level E exercise. Note: for uniform and log-uniform distributions the attributes are the interval endpoints. For the normal distributions these are the mean and the standard deviation. For the log-normal mean and standard deviation refer to the logarithm (base 10) of the variable

Notation	Definition	Distribution	Attributes (endpoints)	Units
T	containment time	uniform	/100,1000/	a
k_I	leach rate for Iodine	log-uniform	/10^{-3},10^{-2}/,	a^{-1}
k_C	leach rate for Np chain nuclides	log-uniform	/10^{-6},10^{-5}/	a^{-1}
$v^{(1)}$	water velocity in geosphere's first layer	log-uniform,	/10^{-3},10^{-1}/	m/a
$l^{(1)}$	length of geosphere's first layer	uniform	/100,500/	m
$R_I^{(1)}$	geosphere retardation coeff. for Iodine (first layer)	uniform	/1,5/	-
$R_C^{(1)}$	factor to compute geosphere retardation coeff. for Np chain nuclides (first layer)	uniform	/3,30/	-
$v^{(2)}$	water velocity in geosphere's second layer	log-uniform	/10^{-2},10^{-1}/	m/a
$l^{(2)}$	length of geosphere's second layer	uniform	/50,200/	m
$R_I^{(2)}$	geosphere retardation coeff. for Iodine (second layer)	uniform	/1,5/	-
$R_C^{(2)}$	factor to compute geosphere retardation coeff. for Np chain nuclides (second layer)	uniform	/3,30/	-
W	stream flow rate	log-uniform	/10^5,10^7/	m^3/a

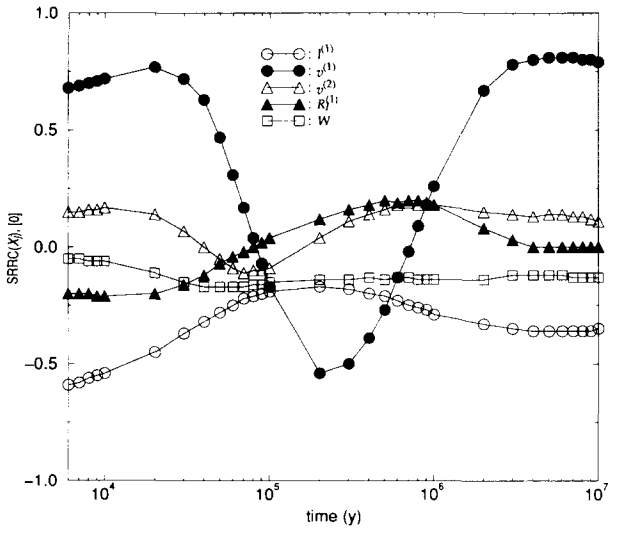

Fig. 2. Standardised rank regression coefficients for the input parameters of the Level E case study as function of time (y). The output considered is the total annual dose.

pseudo or quasi random numbers which must be generated to compute all the random variables needed for one single trial. For this test case the cd is 2×12, given that there are 12 variables, and that we need two data matrices for the resampling procedure described in Section 2. Hence, in order to compute the $\hat{S}_i$, a LP_τ matrix of row dimension 24 and column dimension 2^{10} was generated. The test model was then executed $(12+1) \times 1024$ times, ie once for the base sample plus once for each resampling. Unfortunately different resamples are needed for $\hat{S}_i$ and $\hat{S}_{Ti}$, and the entire procedure has to be repeated to compute $\hat{S}_{Ti}$. Values of $\hat{S}_i$, $\hat{S}_{Ti}$ smaller than their probable errors δS_i, δS_{Ti} computed as in eqn (14) are discarded.

The predominance of the higher order terms is evident from the Level E test case (Fig. 3). The $\hat{S}_i$'s stay below 0·1 for almost the entire time span, with local maxima at about 0·2. The $\hat{S}_{Ti}$'s are much higher, and often close to one. $v^{(1)}$ is more important than W at times below 10^5 y, then the situation is reversed between 10^5 and 10^7 y, although the two variables are almost equally important. Variable $l^{(1)}$ is almost always the third most important variable.

In Fig. 4 the same quantities have been plotted for the ranked model. The new statistics, $\hat{S}_i^*$ and $\hat{S}_{Ti}^*$, are computed from the same two batches discussed previously, i.e. they do not need new model evaluations.

We see that although two important variables ($v^{(1)}$, $l^{(1)}$) are correctly identified, W, which was seen to be the most important variable at $t > 10^5$ y, now disappears. The first order term for $v^{(1)}$ is much higher for the rank based measures. This has happened at the expense of higher order terms, in such a way that the total effect for $l^{(1)}$ is considerably reduced, and that of W disappears altogether. Further computations[19] where all the higher order terms have been computed, show that W's influence is mostly due to association terms, like $v^{(1)}W$, $v^{(1)}WR$, $v^{(1)}Wl^{(1)}$ and even $v^{(1)}WRl^{(1)}$. When the relative weight of these is cut by the rank transformation, then W becomes non influential.

The fact that for $t > 10^5$ y the 'actual' most important variable (in the sense discussed in Section 2

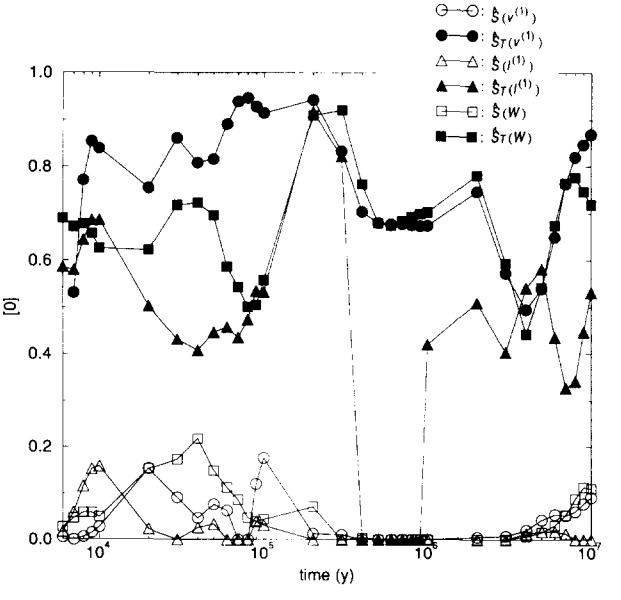

Fig. 3. $\hat{S}_i$ (empty symbols) and $\hat{S}_{Ti}$ (full symbols) for the input parameters of the Level E case study as a function of time (y). The output considered is the total annual dose.

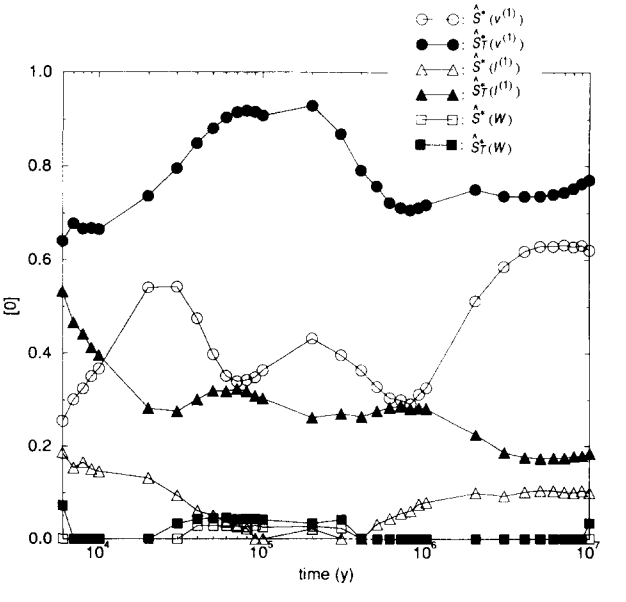

Fig. 4. $\hat{S}_i^*$ (empty symbols) and $\hat{S}_{Ti}^*$ (full symbols) for the input parameters of the Level E case study as a function of time (y). The output considered is the total annual dose.

is not identified by the rank based analysis (either using the $\hat{S}^*_{Ti}$ or the SRRC) gives a measure of the severity of the problem.

We may now return to the difference between the Y-ranked and XY-ranked models (f^* and f^{**}) discussed in Section 1. Although a comparison of $\hat{S}^*_{Ti}$ (using the former model) with SRRC (using the latter) is hindered by the difference of the two estimators (they even use different sample sizes, $N \times (n+1)$ against N), it seems reasonable to hypothisise that the effect of the rank transformation has similar effects, ie both using ranked sensitivity measures and ranked regression coefficents we tend to flatten the higher order effects. What we learn about $\hat{S}^*_{Ti}$ has relevance to SRRC as well.

3.2 Some analytic test cases

The effect of the rank transformation on models can be also investigated by simpler test functions.

We have selected a function of three indices i,j,k:

$$f_{ijk} = \sum_{r=0}^{i} \frac{(ax)^r}{r!} + \sum_{s=0}^{j} \frac{(by)^s}{s!} + \sum_{t=0}^{k} \frac{(cxy)^t}{t!} \qquad (20)$$

where a,b,c are arbitrary constants and x,y are independent variables uniformly distributed in [0,1]. The nonlinearity of f_{ijk} with respect to x,y and their cross product will depend upon the value of the indices i,j,k respectively. Also:

$$\lim_{i,j,k\to\infty} f_{ijk} = e^{(ax)} + e^{(by)} + e^{(cxy)} \qquad (21)$$

We can write the sensitivity indices of f_{ijk} and f^*_{ijk} as

$$1 = S_x + S_y + S_{xy} = S^*_x + S^*_y + S^*_{xy} \qquad (22)$$

In Fig. 5 we have plotted the values of $(\hat{S}_{xy} - \hat{S}xy^*)$ and of $\dfrac{(\hat{S}xy - \hat{S}^*_{xy})}{\hat{S}_{xy}}$ vs the ratio $k/Max(i,j)$ for all the combinations of $i,j,k \in [1,5]$. For this figure $a = b = 1$; $c = 2$, and a $LP\tau$ sample of base

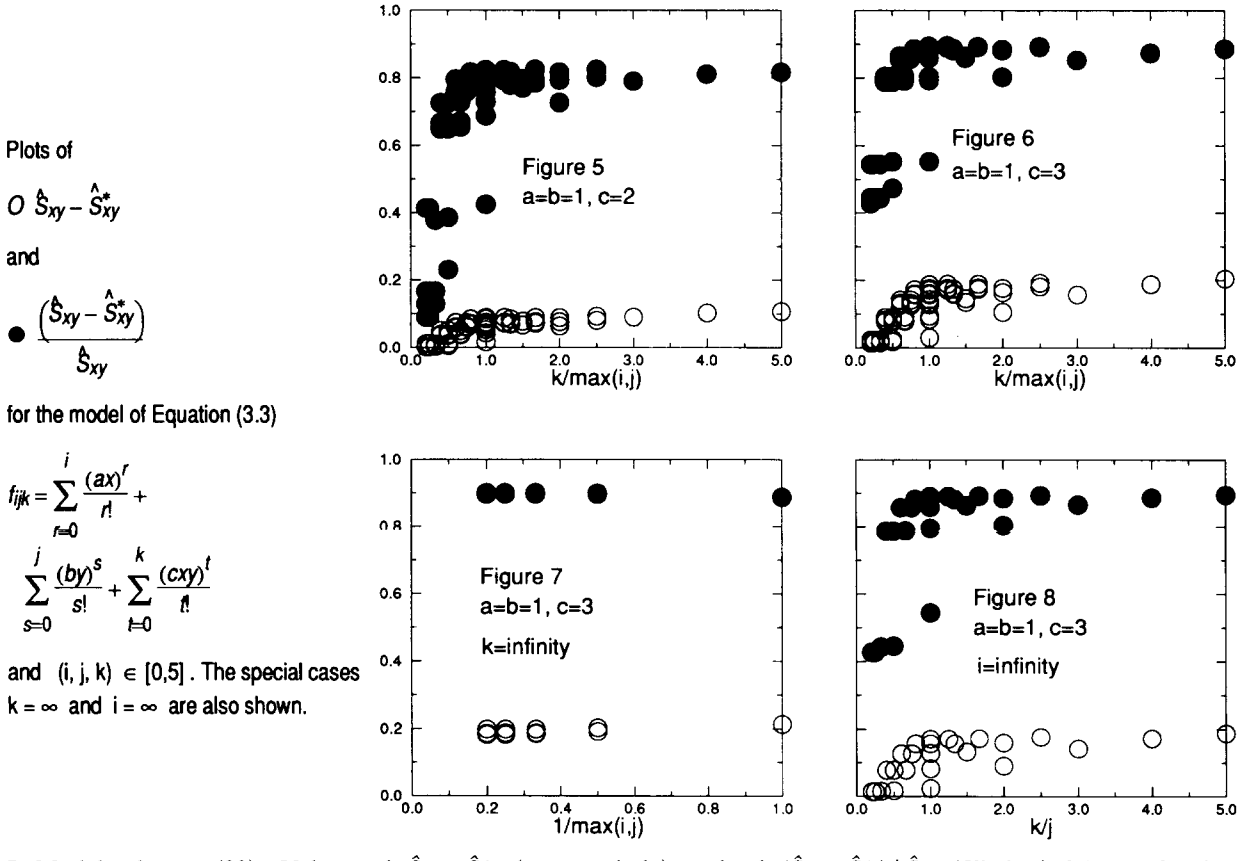

Fig. 5. Model of eqn (20). Values of $\hat{S}_{xy} - \hat{S}^*_{xy}$ (empty circle) and of $(\hat{S}_{xy} - \hat{S}^*_{xy})/\hat{S}_{xy}$ (filled circle). $a = b = 1$, $c = 2$, abscissa = $k/_{max(i,j)}$.

Fig. 6. Model of eqn (20). Values of $\hat{S}_{xy} - \hat{S}^*_{xy}$ (empty circle) and of $(\hat{S}_{xy} - \hat{S}^*_{xy})/\hat{S}_{xy}$ (filled circle). $a = b = 1$, $c = 3$, abscissa = $k/_{max(i,j)}$.

Fig. 7. Model of eqn (20). Values of $\hat{S}_{xy} - \hat{S}^*_{xy}$ (empty circle) and of $(\hat{S}_{xy} - \hat{S}^*_{xy})/\hat{S}_{xy}$ (filled circle). $a = b = 1$, $c = 3$, $k = \infty$, abscissa = $1/_{max(i,j)}$.

Fig. 8. Model of eqn (20). Values of $\hat{S}_{xy} - \hat{S}^*_{xy}$ (empty circle) and of $(\hat{S}_{xy} - \hat{S}^*_{xy})/\hat{S}_{xy}$ (filled circle). $a = 1$, $b = 1$, $c = 3$, $i = \infty$, abscissa = $k/_j$.

size $N = 2^{15}$ was used. The new dimension of the LP_τ matrix is, in this case, 4. The number model evaluations is $(2^{15} \times (2 + 1))$, as the cross terms ($\hat{S}xy$ and $\hat{S}xy^*$) can be computed by eqn (22).

The same quantities have been plotted in Fig. 6 for $a = b = 1$; $c = 3$. The limiting cases $f_{ij\infty}$ and $f_{\infty jk}$ are shown in Fig. 7 and Fig. 8 (only for $a = b = 1$; $c = 3$). Those figures indicate that there is indeed a marked difference between f_{ijk} and f^*_{ijk}. They show that for this test case the second order term, due to the x, y interaction, can be completely overlooked (90% error) by working on the rank transformed data. The error increases with the ratio of the importance of the second order term relative to the first order ones, i.e. with $k/Max(i,j)$ as well as with $c/Max(a,b)$.

Is this trend general? One might wonder if the decrease of the relative weight of the higher order terms is always associated to a rank transformation, regardless of the function f at hand. In reality, this is not the case, i.e. for some functions there might be an opposite effect, the relative weight of the cross order term being moderately higher for f^* than for f. Yet this occurrence seems to be very mild compared to the opposite effect. In Fig. 9 we have plotted the cross order terms for a set of different symmetric functions of two variables $f(x,y)$, with x,y again independent and uniformly distributed in [0,1]. The values plotted are the cross order terms S_{xy} for raw values against S^*_{xy} (on ranks). The first order terms can be computed by differencing, as $2S_x = 2S_y = 1 - S_{xy}$. A large LP_τ sample size (2^{15}) has been used, so that the variances from which the S_{xy}, S^*_{xy} are computed are practically the expectation values.

It can be seen that even choosing 'bad' functions (which would disprove our main hypothesis), the points show a distinct preference for the portion of the plane above the main diagonal, i.e. the rank transformation 'consumes' higher order terms much more easily than it creates them.

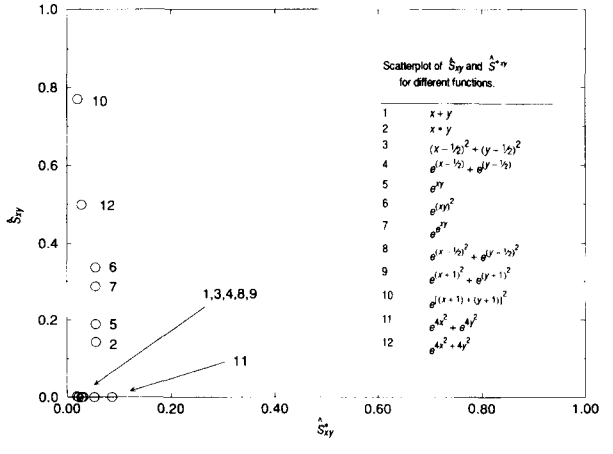

Fig. 9. Scatterplot of S_{xy} (raw values) against S^*_{xy} (ranks) for various symmetric functions of two variables $f(x,y)$.

An important aspect of the problem is that the differences between the f and f^* are likely to increase with the dimensionality n of the space of the parameters, and increasingly so when a consistent fraction of the n input variables are influential. In order to illustrate this aspect we have considered the following function, defined in the the n-dimensional unit cube:

$$f = \prod_{i=1}^{n} g_i(x_i) \tag{23}$$

where

$$g_i(x) = \frac{|4x - 2| + a_i}{1 + a_i}, \quad a_i \geq 0 \text{ a parameter} \tag{24}$$

Plots of $g_i(x)$ for various values of a are given in Fig. 10. For the case where all the a_i's are zero, eqn (23) and (24) reduce to a test case discussed in Ref. 5 in the context of numerical estimation of multidimensional integrals. An analysis of the performances of the sensitivity indices on this test function was conducted in Ref. 45.

1. For all these functions $\int_0^1 gi(x)dx = 1$ and therefore

$$\int_0^1 \ldots \int_0^1 f \, dx_1 \ldots dx_n = 1 \tag{25}$$

2. The variation of the function $g_i(x)$ is

$$1 - \frac{1}{1 + a_i} \leq g_i(x) \leq 1 + \frac{1}{1 + a_i} \tag{26}$$

Therefore, the parameter a_i can be used for specifying the role of the corresponding variable x_i. For example:

1. If $a_i = 0$ the variable x_i is 'important': $0 \leq g_i(x) \leq 2$.

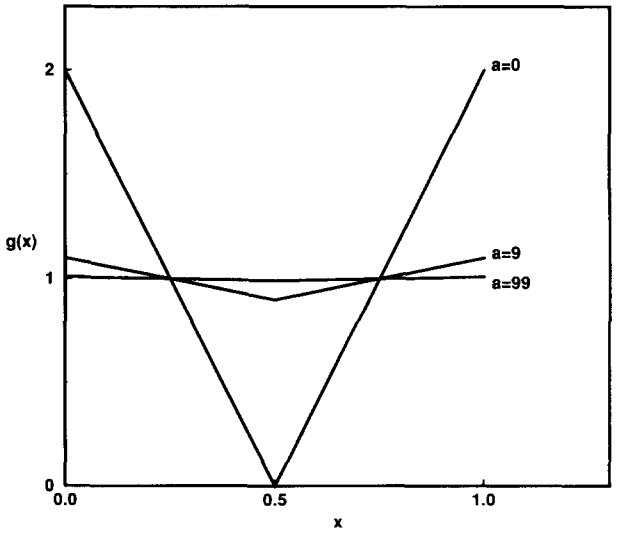

Fig. 10. Function $g(x)$, eqn (24), for different values of a.

2. If $a_i = 9$ the variable x_i is 'non important': $0{\cdot}90 \leq g_i(x) \leq 1{\cdot}10$.
3. If $a_i = 99$ the variable x_i is 'non significant': $0{\cdot}99 \leq g_i(x) \leq 1{\cdot}01$.

For this function the sensitivity indices can be computed analytically.[45] It can be observed that the derivative of g_i changes sign within the interval of variation and is undefined in its midpoint, so that a local sensitivity analysis would be impractical for this function. Furthermore, g_i is nonmonotonic in all its input variables, and linear correlation or regression based techniques would offer little help. If we were to compute the R_y^2 for this function, both for the raw values and the ranks, we would obtain zero, regardless of the value of n and of the a_i's. A linear regression technique (even if applied on ranks) would draw an horizontal line on Fig. 10. For the same reason a two level factorial design would be useless.

The indices for the ranked model were computed numerically. For a simulation with $n = 4$, $N = 2^{13}$ (LP_τ sample) and all the a_i's $=0$ the sensitivity indices yield:

$$S_i = 0{\cdot}154 \quad \hat{S}_i^* = 0{\cdot}253$$

$$S_{ij} = 0{\cdot}051 \quad \hat{S}_{ij}^* = 0{\cdot}009$$

$$S_{Ti} = 0{\cdot}36 \quad \hat{S}_{Ti}^* = 0{\cdot}26$$

As observed in the previous example the rank transformation tends to subtract importance from the higher order terms and add it to the first order one. As a result the real global effect term for—say—X_1, which is

$$S_{T1} = S_1 + S_{12} + S_{13} + S_{14} + S_{123} + \ldots + S_{1234} \quad (27)$$

is much larger than the corresponding S_{T1}^*. Taking as a measure of the error of the ranked model the ratio $R = \dfrac{S_{Ti} - \hat{S}_{Ti}^*}{S_{Ti}}$ we get $R \approx 0{\cdot}28$. For the eight-dimensional case ($n = 8$, $N = 2^{13}$ ($LP\tau$)) and all the a_i's $=0$, the corresponding results are

$$S_i = 0{\cdot}037 \quad \hat{S}_i^* = 0{\cdot}120$$

$$S_{ij} = 0{\cdot}012 \quad \hat{S}_{ij}^* = 0{\cdot}002$$

$$S_{Ti} = 0{\cdot}310 \quad \hat{S}_{Ti}^* = 0{\cdot}135$$

and $R \approx 0{\cdot}56$. This effect is hardly surprising, as the number of terms in the series development (8) grows as 2^n, and consequently also the number of potentially important higher order terms grows.

The previous results have been obtained for the extreme case where all the variables are equally important (all the a_i's $=0$). Similar conclusions apply when the a_i's are different. An instructive example is the case ($a_1 = a_2 = 0$; $a_3 = 3$; $a_5 = \ldots = a_8 = 9$), with 2 important variables, a moderately important and 5 non important ones. Here $n = 8$, $N = 2^{13}$ (LP_τ), and the global effect terms are

$$S_{T1} = S_{T2} = 0{\cdot}55 \quad \hat{S}_{T1}^* = \hat{S}_{T2}^* = 0{\cdot}51$$

$$S_{T3} = 0{\cdot}05 \quad \hat{S}_{T3}^* = 0{\cdot}02$$

$$S_{Ti} = 0{\cdot}007 \quad \hat{S}_{Ti}^* = 0{\cdot}003 \text{ for } i \neq 1,2,3$$

It can be seen that the relative error on X_1, X_2 is mild, while that on the other variables, whose importance mostly arises from higher order terms including X_1, X_2, is very large.

As a final illustration of the effect of dimensionality on the test case defined by eqns (23, 24) consider the case in which $n = 20$, and the a_i's display a range of values computed as $a_i = (i - 1)/2$, so that for the most important variable $a_1 = 0$ and for the least important variable $a_{20} = 9{\cdot}5$. In Fig. 11 we have plotted the ratio $R = \dfrac{\hat{S}_{Ti} - \hat{S}_{Ti}^*}{\hat{S}_{Ti}}$ as function of the index 'i'. Also for this simulation $N = 2^{13}$ (LP_τ), and the input matrix row dimension is 40. All the indices have been computed numerically and only those values are plotted for which both $\hat{S}_{Ti}$ and $\hat{S}_{Ti}^*$ are greater than $\approx 1/2\,\delta\hat{S}$ computed as in (14). It can be seen that the error associated with the rank transformation grows with a declining weight of the parameter.

4 A MEASURE OF ASSOCIATION

We mentioned in the introduction that the present exercise suffers from the lack of an established model

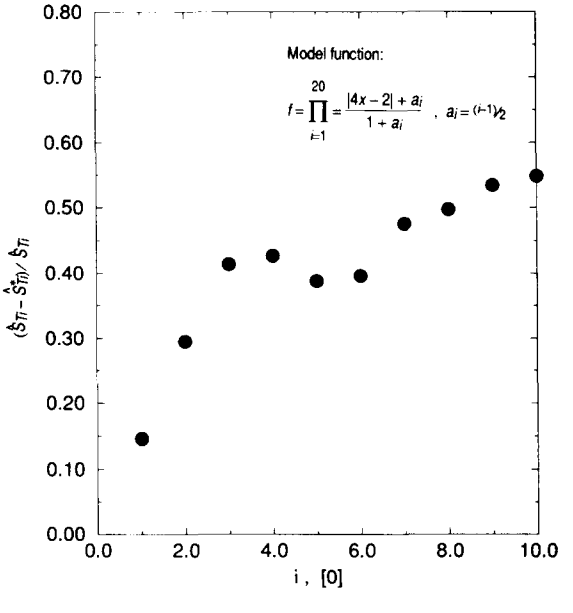

Fig. 11. Model of eqns (23, 24), $n = 20$ and $a_i = (i - 1)/2$. Numerical estimate. Value of $(\hat{S}_{Ti} - \hat{S}_{Ti}^*)/\hat{S}_{Ti}$ vs i. The figure is truncated at $n = 10$, as for higher values the probable error is too large.

taxonomy, capable of ordering models as functions of their nonlinearity, nonmonotonicity, and of their potential difficulty for SA. If such a scale existed, we could, in principle, estimate the potential error associated to a rank transformation, quantified using the sensitivity indices, for different class of models.

We could at this point turn the argument upside down, and suggest the use of some function of the sensitivity indices themselves as a measure of model difficulty.

A natural choice would be, for instance, a model association index I_y defined as:

$$I_y = 1 - \sum_{i=1}^{n} S_i \tag{28}$$

For a model which is perfectly additive in all its n variables $I_y = 0$; we could call such a model 'perfectly additive'. The worst possible model would be the one for which $I_y = 1$. For the model described by eqn (20) and the $(a = b = 1;\ c = 2)$ case, the maximum value of $\hat{I}_y$ when i, j, k are varied in the interval [1,5] is 0·13, and corresponds—as expected—to the $(i = j = 1;\ k = 5)$ point. For the case $(a = b = 1;\ c = 3)$ the maximum $\hat{I}_y$ value is 0·23 for the same point (a LP_τ sample size of 2^{15} was used for these estimates).

For the model defined by eqns (23, 24), the following indices are computed for cases of different dimensionality (Table 2)

We can call I_y an 'Association' index. While R_y^2, computed on the raw values and ranks, is a measure of model nonlinearity and of nonmonotonicity, I_y gives information on the relevance of the interaction terms in the model. Both statistics measure model's complexity.

The estimated association index $\hat{I}_y$ for the Level E model is shown in Fig. 1, together with the model coefficient of determination R_y^2. Also indicated in this figure are the region of radionuclide predominance, and the points where the nuclide dose has a nonmonotonic dependence from $v^{(1)}$. The open symbols refer to the original model, the full ones to the rank version. This model seems to confirm the fact that real applications are always more complex than the analytical functions, even when these latter are especially designed to be difficult. The association of Level E can be very close to 1, and does not seem to be especially correlated to model nonmonotonicity. In computing this statistic we have discarded first order sensitivity terms smaller than their respective error (eqn (14)). The maximum of the $\hat{I}_y$ curve corresponds to that region of the output where doses from Iodine (predominating at shorter times) and doses from the Np chain (predominating at later times) are comparable. This is consistent with the previous findings, ie the model's complexity (association) grows with the number of influential parameters. The ranked model behaves better, with lower $\hat{I}_y$ values.

Table 2. Values of I_y (analytical value) for the model defined by eqns (23, 24) at different dimensions and a_i's values

dimension	a_i's	I_y
4	0000	0·383
4	0009	0·272
4	0099	0·147
4	0999	0·00974
5	00000	0·481
5	00009	0·384
5	00099	0·274
5	00999	0·149
5	09999	0·0129
6	000000	0·567
6	000009	0·483
6	000099	0·386
6	000999	0·276
6	009999	0·152
6	099999	0·016

It may be worth stressing that both R_y^2 and I_y are global measures, and look at that part of a model's complexity that is due to those parameters that contribute appreciably to the output variation. In other words, a model could be highly nonmonotonic and 'associated' with respect to a subset of variables, but if these do not make an appreciable contribution to the model variance they will not be reflected in either R_y^2 nor I_y. Similar considerations apply when contrasting local vs global sensitivity analyses. A given parameter X_i may have a very high $\frac{\partial f}{\partial X_i}$ (and many non negligible $\frac{\partial^2 f}{\partial X_i \partial X_j}$ terms, $j \neq i$) in proximity of the nominal value of these parameters.

Yet it may happen that X_i is reasonably well determined, and that it is given a small range of variation in the analysis. As a result X_i may not contribute appreciably to the output variance, and hence it will not be identified by either SRC(X_i), SRRC(X_i), S_{Ti} or S_{Ti}^*.

5 CONCLUSIONS

The problem of Sensitivity Analysis is to investigate how X_j influences Y_s. Often robustness problems force the investigator to re formulate the problem, and to look instead at how X_j influences $R(Y_s)$, or at how $R(X_j)$ influences $R(Y_s)$. The present article describes qualitatively the impact of this transformation.

Rank transformation is a useful tool for uncertainty and sensitivity analysis, especially when regression techniques are used. The rank transformation can cope with nonlinear (albeit monotonic) models and mitigate the detrimental effect of long tailed output

distributions. Rank transformation may help in building a model response surface replacement in uncertainty analysis and screening parameters in sensitivity analysis.

In practical SA applications both linear and rank regression coefficients are to be used, as recommended for instance by Ref. 22. Whenever the model coefficient of determination computed on the raw values is low, the analyst can use the rank version in search for a better fit to the data at no extra model computation cost.

It is important that the limitations linked to this simplification of the original model are kept in mind. One limitation is that rank transformations cannot linearise nonmonotonic models. This problem is well known to SA practitioners, and is flagged by a low value of the rank based R_y^2. Another drawback in the use of rank transformation was highlighted in the present work. The main effect of the rank transformation was shown to be a forced linearisation of the system, by an artificial increase in the relative weight of the first order terms. This is done at the expense of the (possible) higher order terms, which are nonlinear. As a result the influence of those parameters whose 'total effect' mostly arises from parameter interaction may be overlooked in an analysis based on the ranks. Depending on the model at hand this may or may not result in an unacceptable error in a rank based sensitivity or uncertainty analysis. The Level E test case discussed in this article is not an extreme case. In a chemical kinetics model described in Ref. 19 we have also found significant differences between the parameters ranking obtained with $\hat{S}_{Ti}$ and $\hat{S}^*_{Ti}$. A crucial point is that a deficiency in the ranking provided by the $\hat{S}^*_{Ti}$ implies *a fortiori* a deficiency in the widely used SRRC and PRCC when applied to the same problem, as shown in the level E analysis.

The examples discussed in this note provide a clear qualitative description of the typology of the error for the transformed models f^* and f^{**}, as far as sensitivity is defined in terms of contribution to the output variance. As mentioned in Section 2 the f^* and f^{**} models are no longer equivalent if sensitivity is measured in terms of regression coefficients. On the other hand, as discussed in this article and in previous ones (eg Ref. 40), we believe that the f^* - based techniques, such as the importance measures and the sensitivity indices, should be preferred over the f^{**} - based rank regression/correlation techniques for the purpose of automated SA. The reason of this choice is that it is difficult to know *a priori* if a model is monotonic in all its variables, and the rank regression methods are known to fail with nonmonotonic cases.

The association index suggested in this article appears capable of capturing that part of model complexity which comes from the level of interaction among model input parameters. I_y is a global index, being based on the fractional contribution to the output variance that arises from (interactions among) parameters, rather than on the local effect of that parameter. This association index, as well as the model coefficient of determination R_y^2, reflects the model as well as the range of variation assigned to its input parameters, and cannot be easily related to the terms in the Taylor expansion of the model around some central point of the parameters' space.

The estimation of the I_y is more computationally expensive than that of R_y^2. On the other hand an investigation of I_y for a class of models may provide a useful insight on the possibility of severe error in a rank based UA/SA for a model of that class. The Level E test case exemplifies this kind of situation; were we to use ranks in SA for a similar applications involving mass transfer in a complex medium we would expect that parameters which influence the output linearly (such as W in Level E) could be overlooked.

ACKNOWLEDGEMENT

The authors are grateful to Marian Scott of the Department of Statistics of the University of Glasgow, to Graeme Archer of the Environment Institute at the JRC, and to two anonymous reviewers, for providing useful comments, references and review.

REFERENCES

1. Andres, T. H. & Hajas, W. C., Using iterated fractional factorial design to screen parameters in sensitivity analysis of a probabilistic risk assessment model. In *Proc. Joint Int. Conf. on Mathematical Methods and Supercomputing in Nuclear Application*, Karlsruhe, Germany, 19–23 April 1993, pp. 328–340.
2. Bettonvil B. & Kleijnen, J. P. C., Identifying the important factors in simulation models with many factors. Tilburg University (NL), publication of the Dept. of Economics (1991).
3. Box, G. E. P., Hunter, W. G. & Hunter, J. S., *Statistics for Experimenters,* John Wiley and Sons, New York (1978).
4. Cacuci, D. G., Sensitivity theory for nonlinear systems. Parts I and II, *J. Math. Phys.* **22** (1981) 2794–2802 and 2803–2812.
5. Davis, P. J. & Rabinowitz, P., *Methods of Numerical Integration,* 2nd edition, Academic Press, New York, 1984.
6. Dunker, A. M., The decoupled direct method for calculating sensitivity coefficients in chemical kinetics. *J. Chem. Phy.* **81** (1984) 2385–2393.
7. Bratley, P. & Fox, B. L., Algorithm 659. Implementing Sobol's quasirandom sequence generator. *ACM Trans. Mathematical Software,* **14** (1988) 88–100.
8. Cotter, S. C., A screening design for factorial experiments with interactions. *Biometrika,* **66** (1979) 317–320.

9. Cox, D. C., An analytical method for uncertainty analysis of nonlinear output functions, with applications to fault-tree analysis. *IEEE transaction on reliability*, **R-31** (1982) 465–468.
10. Cukier, R. I., Fortuin, C. M., Schuler, K. E., Petschek, A. G., & Schaibly, J. K., Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory. *J. Chem. Phys.*, **59** (1973) 3873–3878.
11. Cukier, R. I., Levine, H. B., & Schuler, K. E., Nonlinear sensitivity analysis of multiparameter model systems. *J. Comput. Phys.*, **26** (1978) 1–42.
12. Draper N. R., & Smith, H., *Applied Regression Analysis*, John Wiley and Sons, New York, 1981.
13. Draper, D., Assessment and propagation of model uncertainty (with discussion). *J. R. Stat. Soc* **B57** (1995) 45–97.
14. Efron, B. & Stein, C., The jackknife estimate of variance. *The Annals of Statistics*, **9** (1981) 586–596.
15. Helton, J. C., Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal. *Reliab. Engng System Safety*, **42** (1993) 327–367.
16. Helton, J. C., Garner, J. W., McCurley, R. D., & Rudeen, D. K., Sensitivity analysis techniques and results for the performance assessment at the waste isolation pilot plant. *SAND90-7103*, Sandia National Laboratories, 1991.
17. Helton, J. C., Iman, R. L., Johnson, J. D., & Leigh, C. D., Uncertainty and sensitivity analysis of a dry containment test problem for the MAEROS aerosol model. *Nucl. Sci. and Engng*, **102** (1989) 22–42.
18. Helton, J. C., Rollstin, J. A., Sprung, J. L., & Johnson, J. D., An exploratory sensitivity study with the MACCS reactor accident consequence model. *Reliab. Engng System Safety*, **36** (1992) 137–164.
19. Homma, T., & Saltelli, A., Importance measures in global sensitivity analysis of nonlinear models. Accepted for publication in *Reliab. Engng System Safety*.
20. Hora, S. C. & Iman, R. L., A comparison of Maximum/Bounding and Bayesian/Monte Carlo for fault tree uncertainty analysis. *SAND*85-2839, Sandia National Laboratories, 1986.
21. Hwang, J.-T., Dougherty, E. P., Rabitz, S., & Rabitz H., The Green's function method of sensitivity analysis in chemical kinetics. *J. Chem. Phys.*, **69** (1978) 5180–5191.
22. Iman, R. L., & Conover, W. J., The use of rank transform in regression. *Technometrics*, **21** (1979) 499–509.
23. Iman, R. L., Davenport, J. M., Frost, E. L., & Shortencarier, M. J., Stepwise regression with PRESS and rank regression (Program user's guide). *SAND*79-1472, Sandia National Laboratories, 1980.
24. Iman, R. L., & Helton, J. C., An investigation of uncertainty and sensitivity analysis techniques for computer models. *Risk Analysis*, **8** (1988) 71–90.
25. Iman, R. L., Helton, J. C., & Campbell, J. E., Risk methodology for geological disposal of radioactive waste: sensitivity analysis techniques. *SAND*78-0912, Sandia National Laboratories, 1978.
26. Iman, R. L., Helton, J. C., & Campbell, J. E., An approach to sensitivity analysis of computer models, Parts I and II. *J. of Quality Techn.*, **13** (1981) 174–183 and 232–240.
27. Iman, R. L., & Hora, S. C., A robust measure of uncertainty importance for use in fault tree system analysis. *Risk Analysis*, **10** (1990) 401–406.
28. Iman, R. L., Shortencarier, M. J., & Johnson, J. D., A FORTRAN 77 program and user's guide for the calculation of partial correlation and standardized regression coefficients. *SAND*85-0044, Sandia National Laboratories, 1985.
29. Ishigami, T. & Homma, T., An importance quantification technique in uncertainty analysis. *JAERI-M* 89-111, Japan Atomic Energy Research Institute, 1989.
30. Ishigami, T. & Homma, T., An importance quantification technique in uncertainty analysis for computer models. In *Proc. ISUMA '90 First International Symposium on Uncertainty Modelling and Analysis*, University of Maryland, USA, 3–5 December 1990, pp.398–403.
31. Liepman, D. & Stephanopoulos, G., Development and global sensitivity analysis of a closed ecosystem model. *Ecological Modelling*, **30** (1985) 13–47.
32. McKay, M. D., Conover, W. J. & Whiteman, D. E., On the application of statistical techniques to the analysis of computer codes. *LA-6479-MS*, Los Alamos Scientific Laboratory (informal report), 1976.
33. McKay, M. D., & Beckman, R. J., A procedure for assessing uncertainty in models. In *Proc. PSAM-II*, San Diego, CA, 20–25 March 1994, (eds G. E. Apostolakis & J. S. Wu) vol. 1, pp. 009-13–009-18.
34. Mc Rae, G. J., Tilden, J. W., & Seinfeld, J. H., Global sensitivity analysis — a computational implementation of the Fourier amplitude sensitivity test (FAST). *Computers & Chem. Engng*, **6** (1982) 15–25.
35. Morris, M. D., Factorial sampling plans for preliminary computational experiments. *Technometrics*, **33** (1991) 161–174.
36. OECD/NEA PSAC User Group, *PSACOIN Level E Intercomparison, An international code intercomparison exercise on a hypothetical safety assessment case study for radioactive waste disposal systems*. OECD/NEA publication, Paris, 1989.
37. OECD/NEA PSAG User Group, *PSACOIN Level S Intercomparison, An international code intercomparison exercise on a hypothetical safety assessment case study for radioactive waste disposal systems*. OECD/NEA publication, Paris, 1993.
38. Rabitz, H., System analysis at molecular scale, *Science*, **246** (1989) 221–226.
39. Sacks, J., Welch, W. J., Mitchell, T. J., & Wynn, H. P., Design and analysis of computer experiments. *Stat. Sci.*, **4** (1989) 409,435.
40. Saltelli, A., Andres, T. H. & Homma, T., Sensitivity analysis of model output: An investigation of new techniques. *Comp. Stat. and Data Analysis*, **15** (1993) 211–238.
41. Saltelli, A., & Hjorth, J., Uncertainty and sensitivity analyses of OH-initiated dimethylsulphide (DMS) oxidation kinetics. *J. Atmos. Chem.* **21** (1995) 187–221.
42. Saltelli, A. & Homma, T., Sensitivity analysis for model output; Performance of black box techniques on three international benchmark exercises. *Comp. Stat. and Data Analysis* **13** (1992) 73–94.
43. Saltelli, A., Homma, T., & Andres, T. H., A new measure of importance in global sensitivity analysis of model output. In *Proc. 6th Joint EPS/APS International Conference on Physics and Computing*, (eds R. Gruber & M. Tommasini), Lugano, August 1994, pp.511–514.
44. Saltelli, A. & Marivoet, J., Nonparametric statistics in sensitivity analysis for model output; A comparison of selected techniques. *Reliab. Engng System Safety*, **28** (1990) 229–253.
45. Saltelli, A. & Sobol', I. M., Sensitivity analysis for nonlinear mathematical models: numerical experience.

Matematicheskoe Modelirovanie, **7** (1994).
46. Schaibly, J. K., & Schuler, K. E., Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. Part II, Applications. *J. Chem. Phys.*, **59** (1973) 3879–3888.
47. Sobol', I. M., On the distribution of points in a cube and the approximate evaluation of integrals. *USSR Computational Mathematics and Mathematical Physics*, **7** (1967) 86–112.
48. Sobol', I. M., Uniformly distributed sequences with an additional uniform property. *USSR Computational Mathematics and Mathematical Physics,* **16** (1976) 236–242.
49. Sobol', I. M., Sensitivity estimates for nonlinear mathematical models. *Matematicheskoe Modelirovanie*, **2** (1990) 112–118 (in Russian), translated in *Mathematical Modelling and Computational Experiments,* **1** (1993) 407–414.
50. Sobol', I. M., Quasi-Monte Carlo methods. *Progress in Nuclear Energy,* **24** (1990) 55–61.
51. Sobol' I. M., Turchaninov, V. I., Levitan, Yu. L., & Shukhman, B. V., *Quasi random sequence generators.* Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, Moscow, 1992. (The LPTAU subroutine is distributed by OECD/NEA Data Bank, 12 bd. des Iles 92130 Issy les Moulineaux, France.)
52. Welch, W. J., Buck, R. J., Sacks, J., Wynn, H. P., Mitchell T. J., & Morris M. D., Screening, predicting, and computer experiments. *Technometrics,* **34** (1992) 15–47.