Comparison of some efficient methods to evaluate the main effect of computer model factors

Thierry Alex Mara [a] & Onimihamina Rakoto Joseph [b]

[a] Laboratory of Building Physics and Systems, Faculty of Sciences and Technologies , University of Reunion Island , 15 Avenue Rene Cassin, BP 7151, 97715, Saint-Denis, Messag Cedex 09, Reunion Island, France

[b] Institute of Energy Management, University of Antananarivo , Ambohitsaina, BP 566, 101, Antananarivo, Madagascar
Published online: 15 Feb 2008.

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis
Taylor & Francis Group

Comparison of some efficient methods to evaluate the main effect of computer model factors

THIERRY ALEX MARA*† and ONIMIHAMINA RAKOTO JOSEPH‡

†Laboratory of Building Physics and Systems, Faculty of Sciences and Technologies,
University of Reunion Island, 15 Avenue Rene Cassin, BP 7151, 97715 Saint-Denis,
Messag Cedex 09, Reunion Island, France
‡Institute of Energy Management, University of Antananarivo, Ambohitsaina,
BP 566, 101 Antananarivo, Madagascar

In this paper, three methods to compute the factors main effect on model responses are compared. The first one is the improved Fourier amplitude sensitivity test [Tarantola, S., Gatelli, D. and Mara, T.A., 2006, Random balance designs for the estimation of first-order global sensitivity indices. *Reliability Engineering and System Safety*, **91**(6), 717–727.]. The second one is the extension of the previous technique to the method of Sobol. At last, an original approach that combines a sampling-based method (Monte Carlo) with iterated one-dimensional fittings is also investigated. In our works, we show that the three methods are able to estimate the factors main effect. On the one hand, the improved FAST and the proposed sampling-based strategy are the less expensive methods as they only require one single sample set of simulation runs. On the other hand, when factors are correlated or an uncertainty analysis is also investigated, the former is the most suited. Besides, it is also demonstrated that the estimates of the main effects with the proposed sampling-based strategy is more robust and is computationally efficient.

Keywords: Uncertainty analysis; Sensitivity analysis; Main effect; FAST; Monte Carlo; LHS

1. Introduction

This paper addresses the problem of uncertainty analysis (UA) and sensitivity analysis (SA) of model output. In the context of numerical experiments, UA involves the determination of the distribution of the output Y of a given computer model that results from the distributions of its uncertain input factors $X = \{X_1, X_2, \ldots, X_k\}$, whereas SA intends to estimate how much the variability of Y is dependent on each of its input factors. Although SA is closely tied to UA, it tends to be a more complex undertaking due to the variety of possible measures of sensitivity. The sensitivity measure of interest is generally related to the aim of the survey. In particular, the estimation of main effects (also named first-order sensitivity index) is the aim of a particular problem setting called factors prioritisation [1]. Indeed, the uncertain factor that contributes the most to the uncertainty of a model response has the greater main effect.

*Corresponding author. Email: mara@univ-reunion.fr

Consequently, it is of high interest to be able to efficiently compute all the factors' main effects, at low computational cost and so, for any model. This sensitivity index is defined by

$$S_i = \frac{\mathrm{Var}(E(Y|X_i))}{\mathrm{Var}(Y)} \tag{1}$$

where $\mathrm{Var}(\cdot)$ denotes the variance and $E(Y|X_i)$ is the expectation of Y conditional on X_i.

In global SA, two classes of methods are available: the deterministic and the stochastic approaches. The latter bring the analysis in the spectral domain and are so-called FAST methods [2, 3] whereas the former, based on Monte Carlo sampling, is the Sobol's method [4] or the replicated Latin hypercube sampling (r-LHS) for instance [5].

Computing the entire factors' main effects with the existing stochastic methods requires as much as sampling sets than the number of input factors. On the opposite, with the classical FAST method only one sample set is sufficient to evaluate the entire main effects but it has other drawbacks such as the choice of the frequency set and a number of simulation runs related to the number of factors. Recently, Tarantola *et al.* [6] have proposed an improved FAST approach that avoids the problem of the choice of the frequency set. In this work, the improvement is extended to the stochastic methods.

Unfortunately, to the authors' knowledge, deterministic approaches do not deal with correlated factors. Besides, SA (computation of the main effects) and UA (assessment of the probability density function of the model response) are performed separately by using different simulation sets [7, 8]. In the present paper, we propose an approach to compute both the uncertainty of the model response and the factors main effect with only one sample set, and we demonstrate that it is more efficient than the improved deterministic and stochastic methods.

2.   Computation of main effects

2.1   *The improved FAST*

In the classical approach, the parametric curve used in FAST is defined as

$$X_i(s) = G_i(\sin \omega_i s), \quad \forall i = 1, 2, \ldots, k$$

where G_i are functions to be chosen by the analyst to get the desired probability density function for X_i, s is the parametric variable varying in $\{-\pi, \pi\}$, and the frequencies ω_i are selected such that they are free of interferences up to a given order M. The selection of frequencies is made using the algorithm described in ref. [9] as function of the input dimension k, and the number of simulation runs must be greater than $N = 2 \max(\omega_i) + 1$. However, N is an increasing function of k and for large k, the computational cost can be too high to be acceptable.

The improvement of the FAST method goes back to the idea of Satterthwaite [10], who uses random balance designs in regression problems and extends it to the case of SA of model output. In the improved FAST approach, all the factors are sampled using the same frequency as follows

$$X_i, (s) = G_i(\sin \omega s_i), \quad \forall i = l, 2, \ldots, k$$

where ω is an arbitrary integer, which is set to 1 for simplicity, s_i are random permutations of uniform sampled values over $\{-\pi, \pi\}$. Then, the model $Y = f(X)$ is evaluated N times.

To compute S_i, the main effect of X_i, the values of model output $Y(s)$ are then arranged such that the corresponding values s_i, are ranked in increasing order. With such an arrangement, it is like the simulation runs have been performed with only the factor X_i sampling along with

a periodic curve and the others being randomly sampled. The main effect is computed in the same way as for the classical FAST approach by computing the Fourier spectrum. Note that, the method requires only N simulation runs as the same set of model output is just re-arranged for the computation of each main effect.

2.2 *The improved Sobol's method*

The method originally proposed by Sobol is based on Monte Carlo sampling [4]. To compute the main effect of X_i, the model is first evaluated N times with factors values randomly sampled in their range of variation and according to their respective probability function. Let $Y^{(0)}$ be the response of the first set of model evaluations. Then, a second set of model evaluations is performed by changing all the factors values except for the factor X_i. Let $Y^{(i)}$ be the corresponding response. Then, X_i's main effect can be computed as follows

$$S_i = 1 - \frac{\text{Var}((Y^{(i)} - Y^{(0)})/\sqrt{2})}{\text{Var}(Y)} \tag{2}$$

In this approach, $k + 1$ sample sets are necessary to compute the entire main effects. The simulation runs are performed according to the following designs

$$Y^{(0)} = f(X_1^{(0)}, X_2^{(0)}, \ldots, X_k^{(0)}) \quad \text{and} \quad Y^{(i)} = f(X_1^{(1)}, X_2^{(1)}, \ldots, X_{i-1}^{(1)}, X_i^{(0)}, X_{i+1}^{(1)} \ldots, X_k^{(1)})$$
$$\forall i = 1, 2, \ldots, k$$

To reduce the simulation cost to only two sets of simulation runs, a random balance design can also be derived. Indeed, the first simulation remains unchanged but the second set is built from the first one by permuting the values of each column. We then obtain the following sample sets

$$Y^{(0)} = f(X_1^{(0)}, X_2^{(0)}, \ldots, X_k^{(0)}) \quad \text{and} \quad Y^{(1)} = f(X_1^{(1)}, X_2^{(1)}, \ldots, X_i^{(1)}, \ldots, X_k^{(1)})$$

where $X_i^{(1)} = \text{RP}_i(X_i^{(0)})$ and RP_i is the random permutation function of X_i's values. In the same way as for the improved FAST, the values of model output $Y^{(0)}$ are rearranged with the corresponding permutation RP_i to compute S_i. With such an arrangement, it is like $Y^{(1)}$ has been obtained by varying all the factors except the factor X_i as compared to the first sample set. X_i's main effect is then computed with the following formula

$$S_i = 1 - \frac{\text{Var}((Y^{(1)} - \text{RP}_i(Y^{(0)}))/\sqrt{2})}{\text{Var}(Y)} \tag{3}$$

2.3 *Coupling sampling-based methods with iterated one-dimensional regressions*

The improvements proposed above are based upon random permutations of factors values. But, such a methodology is inapplicable to models with dependent input factors. Indeed, whatever the class of methods employed, the permutations do not preserve the factors correlations. Besides, factors prioritisation setting is closely related to UA, but the methods discussed above are only suited for SA and are usually not employed to perform UA.

Such an analysis involves the exploration of a mapping from uncertain factors to uncertain outcomes. Several sampling strategies are available, including random sampling, importance sampling and LHS. The former is very popular because its efficient stratification properties allows for the extraction of a large amount of UA and SA information with a relatively small

sample size [11]. The attractive features of Monte Carlo analysis are the extensive sampling from the ranges of uncertain factors and that correlated factors can be propagated through the model. Unfortunately, with one single-simulation set, the quantitative SA tools usually employed (standardized regression coefficient, standardized rank regression coefficient, stepwise rank regression, etc.) are restricted and subject to model assumptions such as linearity or monotonicity. Given that LHS is the suitable approach for UA, we describe, in the following, a strategy to compute all the factors' main effects from a single sample set.

Let us consider the ANOVA representation proposed by Sobol [4]

$$Y = f(X) = f_0 + \sum_{i=1}^{k} f_i(X_i) + \sum_{j>i}^{k} f_{ij}(X_i, X_j) + \cdots + f_{1,2,\dots,k}(X_1, X_2, \dots, X_k)$$

where the summands in the expansion are orthogonal. As a consequence, it is straightforward to prove that $E(Y) = f_0$ by integrating the previous expansion over all variables. In the same way, integrating over all variables but X_i, we obtain

$$E(Y|X_i) = f_0 + f_i(X_i)$$

that yields to

$$\mathrm{Var}(E(Y|X_i)) = \mathrm{Var}(f_i(X_i)).$$

The former is the numerator in the definition of the sensitivity index defined in equation (1). As a consequence, if one has an estimate of $f_i(X_i)$ then he also has an estimate of S_i. We propose to approximate $f_i(X_i)$ by fitting the model response (in a least-square sense) to the input factors individually. Such a one-dimensional regression may be polynomial or more sophisticated. So, once the simulation runs are achieved from the Monte Carlo sampling (LHS), k one-dimensional regressions are built from the model response by successively considering one factor at a time. In the case of polynomial regression, the model response is fit on a function of the form

$$P_i(X_i) = a_{i0} + a_{i1}X_i + a_{i2}X_i^2 + \cdots + a_{iM}X_i^M \quad \forall i = 1, 2, \dots, k$$

where M is an integer and the a_{ij}'s are the polynomial coefficients to determine. The factors main effect is then computed as follows

$$S_i = \frac{\mathrm{Var}(P_i(X_i))}{\mathrm{Var}(Y)} \tag{4}$$

It has to be noted that because $\mathrm{Var}(P_i(X_i))$ is evaluated using the sampled values of X_i, there are neither interpolation nor extrapolation problems encountered. The sampling method employed in the next sections is LHS.

3. Assessment of the methods on an international benchmark: the g-function of Sobol's

3.1 *Model with eight factors*

We test the performances of the three methods on the g-function of Sobol's. This function has been widely used as benchmark for SA [see 12]. The function is defined by

$$Y = \prod_{i=1}^{k} g(X_i)$$

where X_i and $g_i(X_i)$ are given by

$$g(X_i) = \frac{|4X_i - 2| + a_i}{1 + a_i}$$

for X_i uniformly distributed between [0,1] and $a_i \geq 0$.

The values of a_i's determine the relative importance of the X_i's. For instance, $a_i = 0$ implies that the factor X_i is very important; for $a_i = 1$ it is considered important, while for $a_i = 9$ it becomes non important and for $a_i = 99$ non significant. The partial variances of the first order and the total unconditional variance of the model output are given by

$$\text{Var}(E(Y|X_i)) = V_i = \frac{1}{3(1 + a_i)^2} \quad \text{Var}(Y) = \prod_{i=1}^{k}(V_i + 1) - 1$$

from which the main effect can be calculated analytically.

In a first survey, we test the methods on the model with eight factors. Four tests cases are selected : test case A, a very difficult test case for which each factor has the same importance with a very weak main effect (0.04), test case B, is also a case where all the factors have the same importance but the model is additive ($S_i = 0.125$), in case C, we select the coefficients a_i as {0, 1, 4.5, 9, 99, 99, 99, 99}, so that only the two first factors are the most important and at last and, for test case D, we choose factors in random order of importance by setting a_i = {99, 0, 9, 0, 99, 4.5, 1, 99}. The analytic main effects S_i's and their estimates S_i^*'s are computed with the three methods described in section 2. The estimation procedures are replicated $r = 100$ times for increasing sample sizes N to test the robustness of the methods. Their performances are assessed by comparing the sum of absolute error defined by

$$\text{SAE} = \sum_{i=1}^{k} |S_i^* - S_i|$$

The arithmetic means and the standard deviations of the SAE's are plotted against the sample size in figure 1; the standard deviations are presented as error bars. Whatever the method employed, the estimated indices converge towards the analytic values. But it also highlights that the iterated polynomial regressions based method gives better estimates of the main effects especially at low sample sizes. Indeed, with this method the SAE's values are significantly lower and the error bars narrower, at low sample sizes, in all cases. In general, the improved FAST approach performs poorly at low sample sizes. This is due to the fact that the estimates for the non important factors are generally biased as reported by Tarantola *et al.* [6]. The improved Sobol's method performs better than the spectral approach at low sample sizes but is less accurate at large sample sizes. One inconvenient feature of the improved Sobol's method is that negative indices can be obtained (not shown). This is inherent to the method itself and can be easily understood if one refers to equation (3).

3.2 *Performance for large dimensional models*

In a second exercise, we test the ability of the three methods to detect the most influential factors of a high dimensional model. In this purpose, we use the g-function by fixing four very important factors ($ai = 0$), four important factors ($a_i = 1$), two non-important ones ($a_i = 9$) and 90 non-significant ones. For each method, 10 replicates of the estimates of S_i's have been obtained at seven different sample sizes (see figure 2). Once again, the efficiency of the proposed method is highlighted. Its performances remain unchanged in case of high
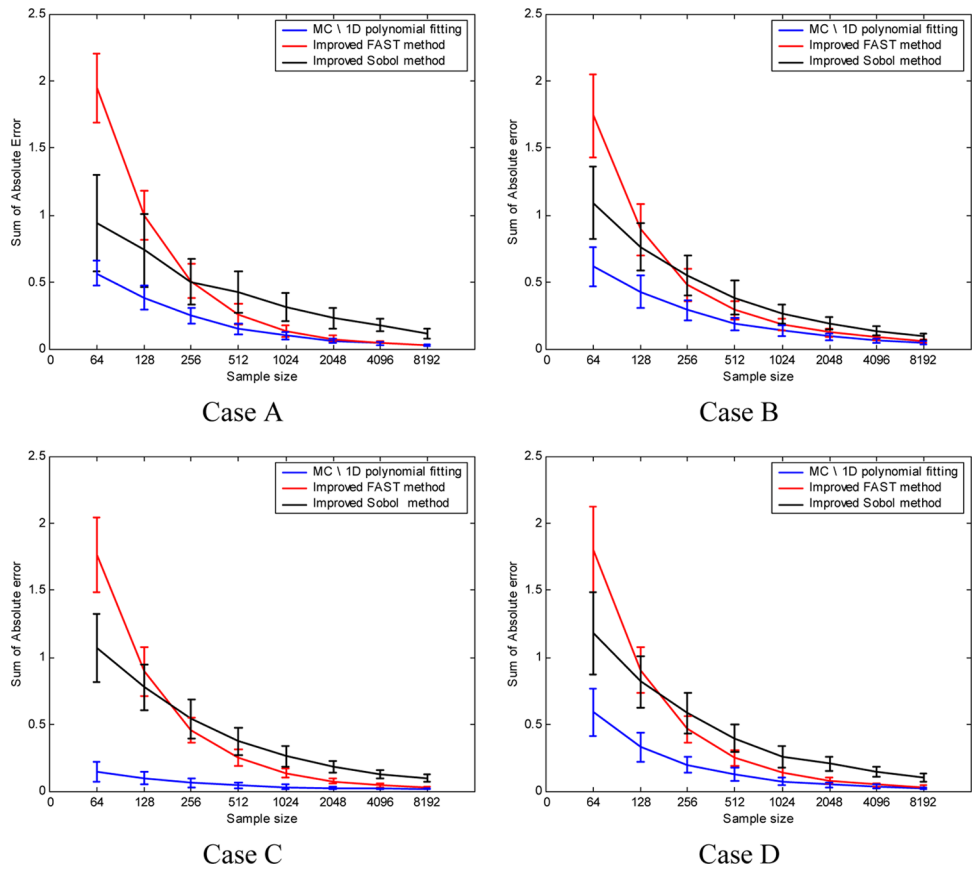
Figure 1.   Assessment of the three methods by application to four analytical test cases. The calculation of the eight factors main effect is more precise with the iterated one-dimensional polynomial regressions approach than with the random balance methods.
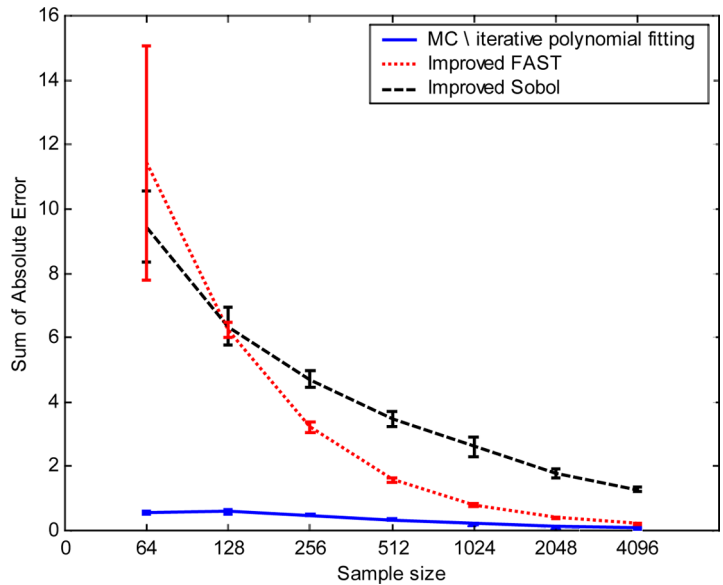

Figure 2.   Performance of the three methods on a high dimensional model (100 input factors). The calculation of the 100 main effects are replicated ten times. The errorbar represents the standard deviation of the SAE's.
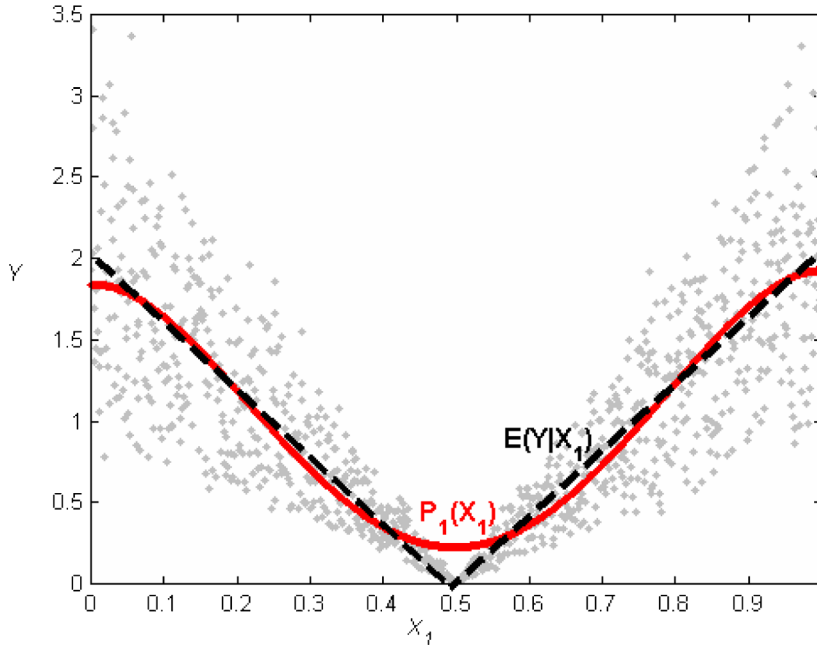
Figure 3.   Scatterplots of the g-function of Sobol versus factor $X_1 (a_1 = 0)$. The theorical marginal relationship between factor X_1 and the model response is in dashed line. In solid line is the estimated marginal relationship obtained by regressing the model response on a one-dimensional fourth-order polynomial. With a polynomial approximation some lack of accuracy appears at the points of discontinuity of the derivative ($X_1 = 0.5$) and at the limits of the interval of variations (0, 1).

dimensionality problems contrary to the other methods. In particular, it is able to estimate accurately the factors main effect at low sample sizes.

It is also interesting to compare the estimated and the analytical marginal relationship between the factor X_i and the model response. The first one is represented by the one-dimensional polynomial obtained by regression $P_i(X_i)$. Figure 3 shows that the polynomial identified for factor X_1 performs poorly where the relationship (or its derivative) is discontinuous and at the limits of the factor's range of variation. In spite of this lack of accuracy, the estimation of the S_i's is good. The use of non-polynomial regressions (likewise an artificial neural network) or piecewise polynomial regressions should lead to better accuracy.

4.   UA and SA of a building thermal model

4.1   *Aim of the study*

The survey concerns the thermal model of a real test cell that was set at the University of Reunion Island [13]. The studied test cell is a cubic shape building with a single window on the south wall and a door to the north. All vertical walls are structurally identical and are composed of cement fibre panel and polyurethane. The roof and the floor are also insulated. The thermal model is capable to predict the indoor air temperature given the building thermal properties and the weather solicitations (outdoor air temperature, solar radiation, wind speed and direction). Figure 4 shows a picture of the test cell and on the left, the weather station that provides the solicitations to the computer model.

Figure 4.   Picture of the test cell and in the back, the weather station that provides solicitations to the model.

The uncertain factors are the thermo-physical properties of the building materials. From the knowledge of the factors uncertainty, the uncertainty of the model response (indoor air temperature) is estimated. In the same time, to plan future investigations, we also search for the factors that mainly contribute to the variance of the model response. We use LHS to evaluate
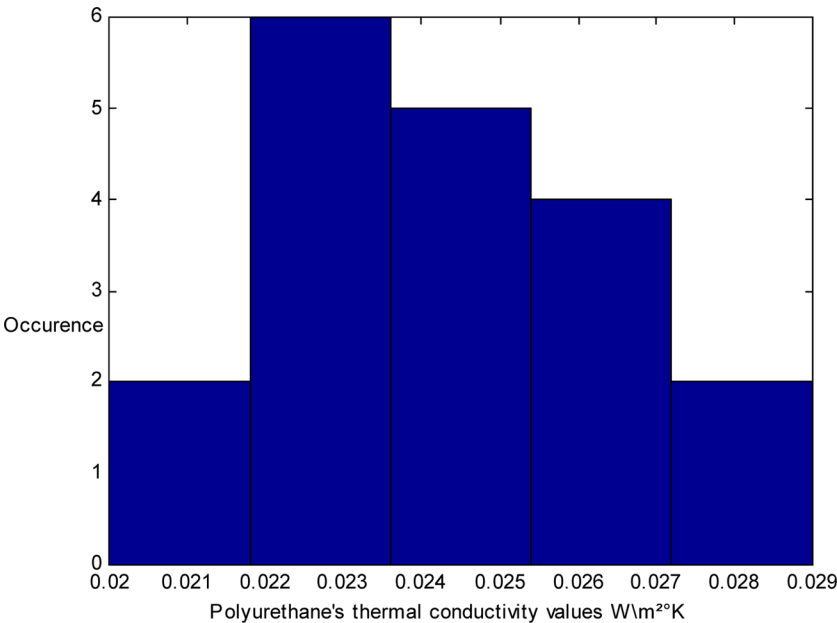

Figure 5.   Pdf of the thermal conductivity of the polyurethane extracted from the BEPAC [14].

the uncertainty of the predicted indoor air temperature and the iterated one-dimensional polynomial regressions to identify the important factors.

4.2 *The uncertainty analysis*

The uncertainty assigned to the materials thermophysical properties are assessed by analysing the reference document of the Building Environmental Performance Analysis Club (BEPAC [14]) which provides a list of materials properties commonly used in buildings construction. We proceed as follows: for each thermophysical property all the values are taken from the database (different values are provided as they come from different sources and measurements) and an histogram is plotted to deduce the associated probability distribution function (pdf). An example of such a pdf is shown in figure 5. Then, the factors values are resampled according to the empirical distribution. A uniform distribution is assumed for the factors that are not reported in the document and for which no information is available. The factors and their associated distribution are listed in table 1.

Five hundred simulation runs were performed by sampling the 21 factors from a LHSs and according their pdfs. The predicted indoor air temperature, a time-dependent output, is stored after each simulation run. Generally, to analyse the uncertainty, the cumulative distribution function is used. Because the output is time-dependent, such an analysis may be cumbersome. So, in this study, the uncertainty of the model response is summarised by the 95% two-sided confidence interval. The former is delimited by the 0.025 and 0.975 quantiles of the model response. In figure 6 are plotted the 95% confidence bounds as well as the mean of the predicted indoor air temperature for the third day. In the daytime, the amplitude of the interval reaches 3 °C whereas in the nighttime it is lower (2 °C).

Table 1. List of the factors and their assigned pdf.

Factors	Distribution
Thermal conductivity of polyurethane (W/m °K)	Empirical distribution, [0.02, 0.024, 0.029]
Density of polyurethane (Kg/m^3)	Empirical distribution, [24, 31, 37]
Thermal conductivity of cement fibre (W/m °K)	Empirical distribution, [0.08, 0.287, 0.722]
Density of cement fibre (Kg/m^3)	Empirical distribution, [350, 965, 1856]
Thermal conductivity of heavyweight concrete (W/m °K)	Empirical distribution, [0.6, 0.85, 1.1]
Density of heavyweight concrete (Kg/m^3)	Empirical distribution, [1649, 1850, 2100]
Thermal conductivity of polystyrene (W/m °K)	Empirical distribution, [0.025, 0.036, 0.044]
Density of polystyrene (Kg/m^3)	Empirical distribution, [11, 23.2, 41.6]
Thermal conductivity of aluminium (W/m °K)	Empirical distribution, [200, 209, 230]
Density of aluminium (Kg/m^3)	Empirical distribution, [2700, 2728, 2800]
Thermal conductivity of wood panel (W/m °K)	Empirical distribution, [0.10, 0.15, 0.23]
Density of wood panel (Kg/m^3)	Empirical distribution, [530, 624, 800]
Windows transmittance (−)	N(0.8, 0.05)
Air infiltration rate (vol h^{-1})	U(0, 4)
Outdoor albedo (−)	U(0.2, 0.4)
Absorptivity of outdoor surfaces (−)	U(0.2, 0.5)
Absorptivity of indoor surfaces except the floor (−)	U(0.6, 0.9)
Absorptivity of the floor (−)	U(0.6, 0.9)
Indoor vertical surfaces convective heat transfer coefficient (W/m^2 °K)	U(3, 6)
Indoor convective heat transfer coefficient of the roof (W/m^2 °K)	U(5, 7.5)
Indoor convective heat transfer coefficient of the floor (W/m^2 °K)	U(0.5, 2.5)

Notes: 'Empirical distribution' concerns factors that are listed in ref. [14], in brackets are mentioned respectively the minimal, the mean and the maximal values of the variable. *U* stands for uniform pdf and *N* for normal pdf.
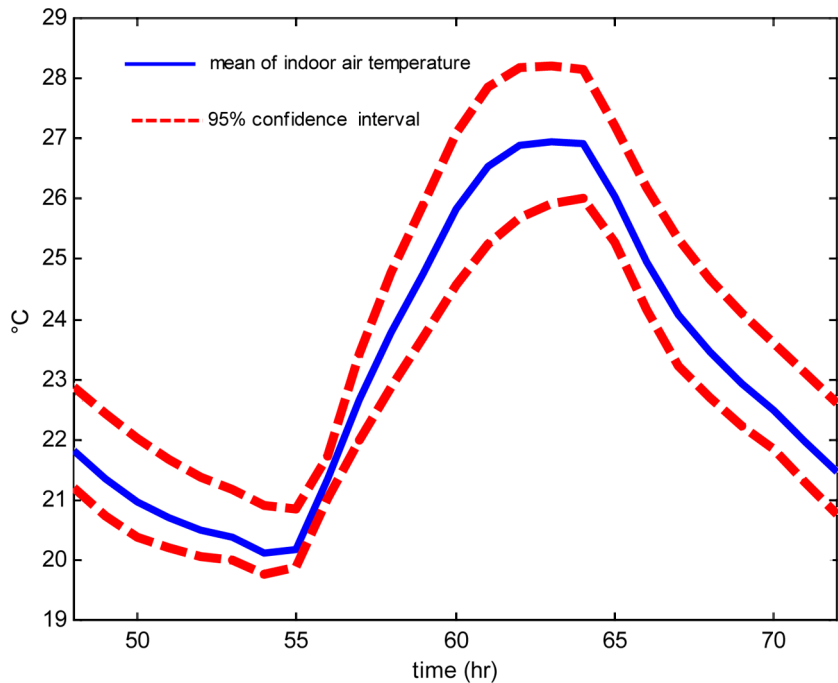
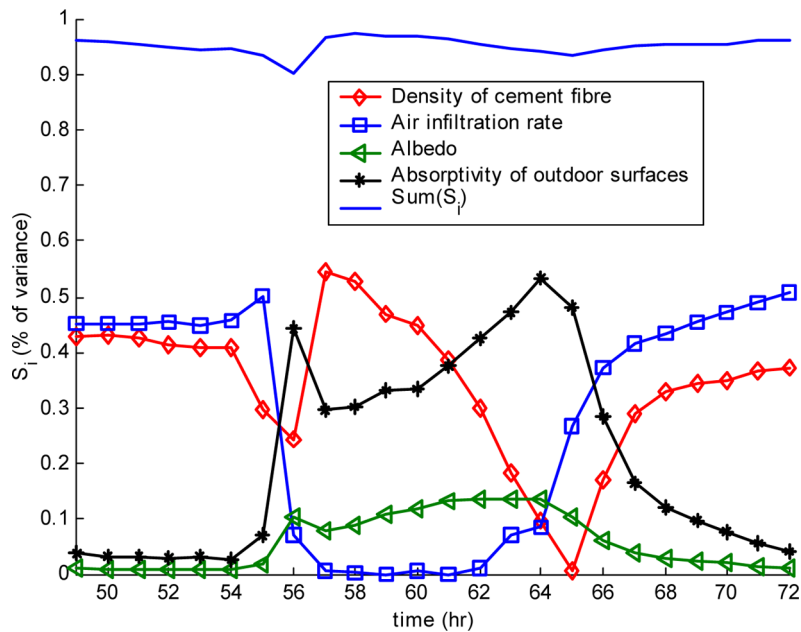Figure 6.   The predicted indoor air temperature and its 95% two-sided confidence bounds.

Figure 7.   Hourly variation of the four most important factors main effect. Some factors are important in the nighttime whereas others are preponderant in the daytime. The sum of the S_i's is close to one meaning that the relationship between the factors and the output is additive.

4.3 *SA of the indoor air temperature*

Given that the uncertainty interval associated to the output is too large, we investigated the factors that contribute the most to it. Such an investigation requires to estimate the factors main effect. From the same sample set, we calculate the S_i's with the method previously described in subsection 2.3. For all the factors, the regression is based on a fourth-order polynomial. The effects of the most important factors are plotted in figure 7 as well as the sum of the S_i's. The former is close to one meaning that the factors interactions are negligible. Only two factors are preponderant during the daytime: the absorptivity of the outdoor surfaces and the albedo. This is explained by the fact that they are related to the solar radiation. The infiltration rate has a strong importance during the nighttime because of its refreshing capacity. The uncertainty of the cement fibre's density has a great impact all along the day. Consequently, to increase the reliability of the test cell model, some specific measurements should be carried out to improve the knowledge of these four factors value.

5. Conclusion

In this article, we proposed a comparison of three efficient methods to evaluate the factors main effect of computer models. The methods described only require one or two sample sets. Two of them are based on random balance designs whereas the last one is based on iterated one-dimensional regressions. Applications to analytical test cases demonstrated the superiority of the former even though some improvements seem possible regarding the nature of the regression (only polynomial regressions were considered in this work). One of its attractive features is that coupling with an appropriate sampling strategy, one can perform both UA and SA with the same sample set. Another compelling feature is that correlations between factors can be accounted for (not shown in the present article).

An application to a real problem allows to highlight the factors that contribute the most to the variation of the model response. Such an investigation helps the modellers planning future experiments. In particular, to reduce the uncertainty of the model response, an effort must be made to obtain accurate measure of four input factors value: the cement fibre's density, the absorptivity of the outdoor surfaces, the infiltration rate in the test cell and the albedo of the outdoor ground.

References

[1] Saltelli, A. and Tarantola, S., 2002, On the relative importance of input factors in mathematical models: safety assessment for nuclear waste disposal. *Journal of American Statistical Association*, **97**, 702–709.
[2] Cukier, R.I., Fortuin, C.M., Shuler, K.E., Petschek, A.G. and Schaibly, J.H., 1973, Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I. Theory. *Journal of Chemical Physics*, **59**, 3873–3878.
[3] Saltelli, A., Tarantola, S. and Chan, K., 1999, A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics*, **41**(1), 39–56.
[4] Sobol, I.M., 1993, Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling and Computational Experiments*, **1**(4), 407–414.
[5] McKay, M.D., 1995, Evaluation prediction uncertainty. Report NUREG/CR-6311, LA-12915-MS, Los Alamos National Laboratory.
[6] Tarantola, S., Gatelli, D. and Mara, T.A., 2006, Random balance designs for the estimation of first-order global sensitivity indices. *Reliability Engineering and System Safety*, **91**(6), 717–727.
[7] Crosetto, M., Tarantola, S. and Saltelli, A., 2000, Sensitivity and uncertainty analysis in spatial modelling based on GIS. *Agriculture, Ecosystems and Environment*, **81**, 71–79.
[8] Kioutsioukis, I., Tarantola, S., Saltelli, A. and Gatelli, D., 2004, Uncertainty and global sensitivity analysis of road transport emission estimates. *Atmospheric Environment*, **38**, 6609–6620.
[9] Cukier, R.I., Levine, H.B. and Shuler, K.E., 1978, Nonlinear sensitivity analysis of multiparameter model systems. *Journal of Computational Physics*, **26**, 1–42.

[10] Satterthwaite, F.E., 1959, Random balance experimentation. *Technometrics*, **1**, 111–137.

[11] Helton, J.C., Davis, F.J. and Johnson, J.D., 2005, A comparison of uncertainty and sensitivity analysis results obtained with random and latin hypercube sampling. *Reliability Engineering and System Safety*, **89**(3), 305–330.

[12] Saltelli, A. and Sobol, I.M., 1995, About the use of rank transformation in sensitivity analysis model. *Reliability Engineering and System Safety*, **50**, 225–239.

[13] Mara, T.A., Garde, F., Boyer, H. and Mamode, M., 2001, Empirical validation of the thermal model of a passive solar test cell. *Energy and Buildings*, **33**(6), 589–599.

[14] Clarke, J.A., Yaneske, P.P. and Pinney, A.A., 1991, The harmonisation of thermal properties of building materials. Technical Note 91/6, BEPAC.