ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/231066223

Global and local sensitivity analysis methods for a physical system

Article *in* European Journal of Physics · October 2011

DOI: 10.1088/0143-0807/32/6/011

CITATIONS
14

READS
3,782

1 author:

Jérôme Morio
The French Aerospace Lab ONERA

79 PUBLICATIONS **536** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

PhD Project: "Rare event probability estimation & Reliability-oriented sensitivity analysis under distribution parameter uncertainty -- Application to aerospace systems" View project

All content following this page was uploaded by Jérôme Morio on 23 May 2014.

The user has requested enhancement of the downloaded file.

Global and local sensitivity analysis methods for a physical system

Jérôme Morio

Onera - The French Aerospace Lab, F-91761 PALAISEAU Tel.: +33 1 80 38 66 54 jerome.morio@onera.fr

Abstract. Sensitivity analysis is the study of how the different input variations of a mathematical model influence the variability of its output. In this article, we review the principle of global and local sensitivity analysis of a complex black-box system. A simulated case of application is given at the end of this article to compare both approaches.

1. Introduction

A physical black-box model is characterized by a certain number of input parameters and a few equations that use those inputs to give a set of outputs. This type of model is usually deterministic, meaning that you always get the same outputs for a given set of inputs, no matter how many times you re-calculate. Figure 1 presents an example of air-to-air missile simulation code that estimates the missile target distance. Different inputs are considered (plane, missile, design, meteorology,...) and influenced the simulation results. An interesting question is: what are the most influent input parameters on the target distance? Is it possible to rank the inputs in function of their output influence? This type of study is called sensitivity analysis and can be run in rigorous manners that are not well known by the physics community. However, sensitivity analysis has several applications in the world of physics. Indeed, physicists can use sensitivity analysis methods to simplify their models, to investigate the robustness of their model predictions, to explore the impact of varying input assumptions. For that purpose, we review the principle of local and global sensitivity and detail the results for both methods on a simulated case.

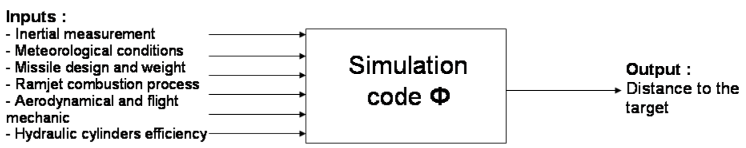

Figure 1. Air-to-air missile simulation code.

2. Context

In this article, we consider a black-box system modelled by a continuous scalar function $\phi : \mathbb{R}^p \to \mathbb{R}$. The input of this function is a p-dimensional random variable $X = \left(X^{(1)}, X^{(2)}, ..., X^{(p)}\right)$ and the output is defined by the variable Y such as $Y = \phi(X)$. The input components $X^{(i)}$ are supposed to be independent.

Sensitivity analysis describes how the uncertainty on the model inputs influences the model output Y. There are two main classes of sensitivity analysis method based on local or global definitions. Both approaches will be described in the following of the article.

3. Global sensitivity analysis

3.1. Definition

Global sensitivity analysis focuses on the variance of model output Y and more precisely on how the input variability influences the variance output [1, 2, 3, 4, 5, 6, 7]. It enables to determine which parts of the output variance are due to the different inputs with the estimation of Sobol indices. They are a central tool in sensitivity analysis since they give a quantitative and a rigorous overview of how the different inputs influence the output.

To apply global sensitivity analysis methods, it is necessary to assume in the following that ϕ is defined on the interval $[0, 1]^p$. A simple variable change is required if the input variable are not defined on this interval. In the considered framework, it is thus possible to show that ϕ can be decomposed in elementary functions [1] :

$$Y = \phi(X) = \phi_0 + \sum_{i=1}^{p} \phi_i(X^{(i)}) + \sum_{1 \leq i < j \leq p} \phi_{ij}(X^{(i)}, X^{(j)}) + ... + \phi_{1...p}(X^{(1)}, ..., X^{(p)}) \quad (1)$$

where ϕ is assumed to be integrable, ϕ_0 is a constant and $\phi_{i_1,...i_s}$ verifies for $\{i_1, ...i_s\} \subseteq \{1, ..., p\}$

$$\forall k = 1...s, \quad \int_0^1 \phi_{i_1,...i_s}(X^{(i_1)}, ..., X^{(i_s)}) dX^{(i_k)} = 0$$

A demonstration of this decomposition can be found in [8]. Since the inputs $X^{(i)}$ are random and independent, one can obtain the well-known ANOVA decomposition by applying the variance operator to equation 1:

$$Var(Y) = V = \sum_{i=1}^{p} V_i + \sum_{1 \leq i < j \leq p} V_{ij} + ... + V_{1...p}$$

with Var the variance operator and

$$\begin{aligned}
V_i &= Var(E(Y|X^{(i)})) \\
V_{ij} &= Var(E(Y|X^{(i)}, X^{(j)})) - V_i - V_j \\
V_{ijk} &= Var(E(Y|X^{(i)}, X^{(j)}, X^{(k)})) - V_i - V_j - V_k - V_{ij} - V_{ik} - V_{jk} \\
&\ldots
\end{aligned}$$

where E describes the mathematical expectation. Sobol sensitivity indices at first order S_i for the variable $X^{(i)}$ are then defined with

$$S_i = \frac{V_i}{V} = \frac{Var(E(Y|X^{(i)}))}{Var(Y)}$$

Sensitivity indices at second order S_{ij} can also be derived relatively to the variables $X^{(i)}$ and $X^{(j)}$:

$$S_{ij} = \frac{V_{ij}}{V}$$

Sensitivity indices at higher order can be defined in the same way. The interpretation of the sensitivity indices is easy since they vary between 0 and 1 and their sum is equal to 1. If S_i is close to 1, then the variable $X^{(i)}$ has a great influence on Y. When the input dimension p increases, the number of Sobol indices increases exponentially and thus, the estimation of all these indices is impossible. For that purpose, total sensitivity indices S_{T_i} are then introduced for each variable $X^{(i)}$ [3]:

$$S_{T_i} = \sum_{k\#i} S_k$$

where $\#i$ represents all the sets of indices that contain i. For instance, if $p = 3$, one has then:

$$S_{T_1} = S_1 + S_{12} + S_{13} + S_{123}$$

The Sobol indices are then often estimated with Monte-Carlo methods [2, 4].

3.2. Monte-Carlo Estimation

In this article, we propose to review how Sobol sensitivity indices can be estimated with Monte-Carlo methods. We denote the estimation of any parameter B with Monte-Carlo methods with the term $\hat{B}$ in the following. Let us consider two N sample size realizations of X:

$$\tilde{X}_k^i = (x_{k1}^i, ..., x_{kp}^i)_{k=1...N,\ i=1,2}$$

Each sample corresponds to a single point in the input space with p dimensions. From these two N sample size realizations of X, first order sensitivity indice $\hat{S}_i$ for the input $X^{(i)}$ is estimated in the following way [2]:

$$\hat{S}_i = \frac{\hat{V}_i}{\hat{V}} = \frac{\hat{U}_i - \hat{\phi}_0^2}{\hat{V}}$$

where the mean is estimated with

$$\hat{\phi}_0 = \frac{1}{N}\sum_{k=1}^{N} \phi(x_{k1}^1, ..., x_{kp}^1)$$

the variance is estimated by definition with

$$\hat{V} = \frac{1}{N}\sum_{k=1}^{N} \phi^2(x_{k1}^1, ..., x_{kp}^1) - \hat{\phi}_0{}^2$$

and finally, the term $\hat{U}_i$ is obtained using the two N sample size realizations of X with [2]

$$\hat{U}_i = \frac{1}{N}\sum_{k=1}^{N} \phi(x_{k1}^1, ..., x_{k(i-1)}^1, x_{ki}^1, x_{k(i+1)}^1, ..., x_{kp}^1) \times \phi(x_{k1}^2, ..., x_{k(i-1)}^2, x_{ki}^1, x_{k(i+1)}^2, ..., x_{kp}^2)$$

The estimation of $\hat{U}_i$ enables to evaluate the influence of the samples on dimension i. Second order sensitivity indices $\hat{S}_{ij}$ are estimated in the following manner [2]:

$$\hat{S}_{ij} = \frac{\hat{U}_{ij} - \hat{\phi_0}^2 - \hat{V}_i - \hat{V}_j}{\hat{V}}$$

where

$$\hat{U}_{ij} = \frac{1}{N}\sum_{k=1}^{N} \phi(x^1_{k1}, ..., x^1_{k(i-1)}, x^1_{ki}, x^1_{k(i+1)}, ..., x^1_{k(j-1)}, x^1_{kj}, x^1_{k(j+1)} ..., x^1_{kp})$$
$$\phi(x^2_{k1}, ..., x^2_{k(i-1)}, x^1_{ki}, x^2_{k(i+1)}, ..., x^2_{k(j-1)}, x^1_{kj}, x^2_{k(j+1)} ..., x^2_{kp})$$

Sensitivity indices of superior orders are then derived from theses relations. Total sensitivity indices can be estimated with the following relations [4]:

$$\hat{S}_{T_i} = 1 - \frac{\hat{U}_{\tilde{i}} - \hat{\phi_0}^2}{\hat{V}}$$

with

$$\hat{U}_{\tilde{i}} = \frac{1}{N}\sum_{k=1}^{N} \phi(x^1_{k1}, ..., x^1_{k(i-1)}, x^1_{ki}, x^1_{k(i+1)}, ..., x^1_{kp}) \times \phi(x^1_{k1}, ..., x^1_{k(i-1)}, x^2_{ki}, x^1_{k(i+1)}, ..., x^1_{kp})$$

One has thus described in this section how to estimate sensitivity Sobol indices for different orders with Monte-Carlo methods. These indices enable to analyse the influence of the different inputs on the variance of Y.

3.3. Alternatives to Monte-Carlo Estimation

The Sobol indices are most of the time estimated with Monte-Carlo methods [2, 4]. The main difficulty of Sobol indice estimation with Monte-Carlo methods is the too high number of generated samples to obtain an accurate estimation (convergence rate in $\sqrt{N}$ where N is the sample size). It is notably the case when the simulation code is time consuming. 10000 samples are often necessary for the estimation of one Sobol indice with a relative deviation of about 10%. The use of deterministic quasi Monte-Carlo sequence (for instance LPτ Sobol sequences) can reduce the number of required samples to estimate Sobol indices [9]. Quasi Monte-Carlo method is used for the computation of an integral that is based on low-discrepancy sequence. They are notably described as a Monte-Carlo alternative in [8, 10]. Another method of Sobol indice estimation is to consider FAST algorithm [11, 12] based on multidimensional Fourier transform which is relatively more accurate than Monte-Carlo and can be applied to total sensitivity indice estimation. However, this technique is not adapted when the problem dimension is too high. These alternatives can be applied when Monte-Carlo methods are not efficient on the studied case.

4. Local sensitivity analysis

4.1. Definition

Local sensitivity [13, 14, 9] analyses how a small perturbation near an input space value $x^0 = (x_1^0, ..., x_p^0)$ influences the value of $y = \phi(x^0)$. It consists in estimating

$$A_i = \frac{\partial y}{\partial x_i}(x_1^0, ..., x_p^0)$$

that characterizes the effect on the random value Y of a perturbation on input $X^{(i)}$ near a nominal value x_i^0. A classical approach to derive this quantity is to consider the OAT (One factor At Time) method.

4.2. Morris method

To proceed a local sensitivity analysis of a complex system, one often considers the Morris method. For that purpose, one builds trajectories in the input space in the following way:

$$x^i = x^{i-1} + \Delta^i e^i$$

with $i = 1...p$ and Δ^i is a step between two consecutive input space points of the trajectory. The vector $(e_1, ..., e_p)$ is the canonical basis of the input space. The vector x^i only differs from x^{i-1} by one unique factor (OAT approach). The parameter x^0 is randomly defined in the input space. The vector $(x^0, x^1, ..., x^p)$ defines a vector of dimension $p+1$ in the input space. R trajectories in the input space are determined to estimate the following indices. Elementary effect for each inputs and for one given trajectory is given by :

$$d_i = \frac{\phi(x^i) - \phi(x^{i-1})}{\Delta_i}$$

with $i = 1...p$. A mean of the elementary effects over the complete set of trajectories is then obtained with

$$\mu_i = \frac{1}{R}\sum_{i=1}^{R} d_i$$

The parameter μ_i characterizes the effect of input $X^{(i)}$ on the output. Since elementary effects can compensate each other, one also uses

$$\mu_i^* = \frac{1}{R}\sum_{i=1}^{R} |d_i|$$

The last statistics that is often estimated corresponds to a standard deviation:

$$\sigma_i = \sqrt{\frac{1}{R-1}\sum_{i=1}^{R}(d_i - \mu_i)^2}$$

The parameter σ_i is in the same way a measure of non-linearity for the input $X^{(i)}$ and also a measure of the input interactions implied by $X^{(i)}$. In general, these different parameters are interpreted in three categories:

Sobol indices	value
$\hat{S}_1$	0.313
$\hat{S}_2$	0.434
$\hat{S}_3$	0.001
$\hat{S}_{12}$	0
$\hat{S}_{13}$	0.253
$\hat{S}_{23}$	0
$\hat{S}_{T_1}$	0.576
$\hat{S}_{T_2}$	0.438
$\hat{S}_{T_3}$	0.254

Table 1. Mean estimation with 10000 Monte-Carlo samples over 20 retrials of partial and total Sobol indices.

- negligible inputs (weak value of μ_i^*).
- important inputs with linear effects without interaction (strong value of μ_i^* and weak value of σ_i).
- important inputs with non-linear effects without/with interactions (strong value of μ_i^* and strong value of σ_i).

The value of μ_i^* and σ_i for each input enables to analyse the influence of $X^{(i)}$ on the ouput Y.

5. Application of sensitivity analysis methods on a simple case

In this last section, we propose to apply global and local sensitivity analysis methods on a simple case. Let us define the following model :

$$Y = \phi(X) = \sin(X^{(1)}) + 7\sin^2(X^{(2)}) + \frac{{X^{(3)}}^4}{10}\sin(X^{(1)})$$

with $(X^{(1)}, X^{(2)}, X^{(3)})$ following uniform law between $-\pi$ and π. Using global sensitivity analysis method and more precisely Sobol indice estimation, we can then obtain the table 1 in mean over 20 retrials of 10000 samples with Monte-Carlo methods. The most influent parameters on the variance of Y are $X^{(1)}$ and $X^{(2)}$ at first order. The parameter $X^{(3)}$ has an interaction with $X^{(1)}$ in the model sensitivity.
With local sensitivity analysis method, one obtains the table 2 with 10000 input space trajectories. Input parameter $X^{(2)}$ has a great influence on the output value with non-linar effect due to the sinus function. Input $X^{(1)}$ and $X^{(3)}$ have equal importance with non-linear effect due to the sinus function for $X^{(1)}$ and due to interaction effect for $X^{(3)}$.

The same kind of results has thus been obtained with local and global sensitivity analysis method in a very simple case. In general, local sensitivity analysis is often less used that global sensitivity analysis because non-linear effect and interaction can affect too

parameter	value
μ_1^*	0.0188
σ_1	0.0206
μ_2^*	0.0447
σ_2	0.0216
μ_3^*	0.0198
σ_3	0.0268

Table 2. Mean estimation of Morris method parameter with 10000 input space trajectories.

much the sensitivity parameter estimation. Moreover, global sensitivity method enables to obtain more detailed results than local ones. Nevertheless, when it is very time consuming to run the model ϕ, local sensitivity analysis method can be preferred because a fewer number of simulations is required to get valuable sensitivity indice estimation.

[1] W. Hoeffding, "A class of statistics with asymptotically normal distributions," *Annals of Mathematical Statistics*, vol. 19, pp. 293–325, 1948.
[2] I. Sobol and S. Kuchereko, "Sensitivity estimates for non linear mathematical models," *Mathematical Modelling and Computationnal Experiments*, vol. 1, pp. 407–414, 1993.
[3] T. Homma and A. Saltelli, "Importance measures in global sensitivity analysis of non linear models," *Reliability Engineering and System Safety*, vol. 52, pp. 1–17, 1996.
[4] A. Saltelli, "Making best use of model evaluation to compute sensitivity indices," *Computer Physics Communication*, vol. 145, pp. 280–297, 2002.
[5] I. M. Sobol, *A Primer for the Monte Carlo Method*. Boca Raton, Fl.: CRC Press, 1994.
[6] A. Saltelli, S. Tarantola, and K.-S. Chan, "A quantitative model-independent method for global sensitivity analysis of model output," *Technometrics*, vol. 41, pp. 39–56, 1999.
[7] J. W. Hall, "Uncertainty-based sensitivity indices for imprecise probability distributions," *Reliability Engineering & System Safety*, vol. 91, no. 10-11, pp. 1443 – 1451, 2006.
[8] B. Efron and C. Stein, "The jackknife estimate of variance," *The Annals of Statistics*, vol. 9, no. 3, p. 586 596, 1981.
[9] A. Saltelli, K. Chan, and E. Scott, *Global sensitivity analysis-The primer*. Wiley, 2000.
[10] W. Morokoff and R. Caflisch, "Quasi-random sequences and their discrepancies," *SIAM J. Sci. Comput.*, vol. 15, no. 6, p. 1251 1279, 1994.
[11] H. Cukier, R. Levine, and K. Schuler, "Nonlinear sensitivity analysis of multiparameter model systems," *Journal of Computationnal Physics*, vol. 26, pp. 1 – 42, 1978.
[12] A. Saltelli, S. Tarantola, and K. Chan, "A quantitative model-independent method for global sensitivity analysis of model output," *Technometrics*, vol. 41, no. 1, pp. 39 – 56, 1999.
[13] F. Alam, K. McNaught, and T. Ringrose, "Using Morris'randomized OAT design as a factor screening method for developing simulation metamodels." Proceedings of the 2004 Winter Simulation Conference, 2004.
[14] G. Pujol, "Simplex-based screening designs for estimating metamodels," *Reliability Engineering & System Safety*, vol. 94, pp. 1156–1160, 2009.

View publication stats