



HAL
open science

Derivative based global sensitivity measures

Serge Kucherenko, Bertrand Iooss

► **To cite this version:**

Serge Kucherenko, Bertrand Iooss. Derivative based global sensitivity measures. 2014. hal-01079358v1

HAL Id: hal-01079358

<https://hal.science/hal-01079358v1>

Preprint submitted on 1 Nov 2014 (v1), last revised 22 Jul 2015 (v3)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Title:	Derivative based global sensitivity measures
Name:	Sergeï Kucherenko ¹ and Bertrand Iooss ^{2,3}
Affil./Addr. 1:	Imperial College London London, SW7 2AZ, UK E-mail: s.kucherenko@imperial.ac.uk
Affil./Addr. 2:	EDF R&D 6 quai Watier, 78401 Chatou, France E-mail: bertrand.iooss@edf.fr
Affil./Addr. 3:	Institut de Mathématiques de Toulouse Université Paul Sabatier 118 route de Narbonne, 31062 Toulouse, France
Keywords:	No Keywords Given
SumOriWork:	No Years and Authors of Summarized Original Work Given

Derivative based global sensitivity measures

SERGEÏ KUCHERENKO¹ AND BERTRAND IOOSS^{2,3}

¹ Imperial College London
London, SW7 2AZ, UK

E-mail: s.kucherenko@imperial.ac.uk

² EDF R&D

6 quai Watier, 78401 Chatou, France

E-mail: bertrand.iooss@edf.fr

³ Institut de Mathématiques de Toulouse

Université Paul Sabatier

118 route de Narbonne, 31062 Toulouse, France

Years and Authors of Summarized Original Work

No Years and Authors of Summarized Original Work Given

Keywords

No Keywords Given

Abstract

The method of derivative based global sensitivity measures (DGSM) has recently become popular among practitioners. It has a link with the Morris screening method and Sobol' sensitivity indices. DGSM are very easy to implement and evaluate numerically. The computational time required for numerical evaluation of DGSM is generally much lower than that for estimation of Sobol' sensitivity indices. We present a survey of

recent advances in DGSM concerning lower and upper bounds on the values of Sobol' total sensitivity indices S_i^{tot} . Using these bounds it is possible in most cases to get a good practical estimation of the values of S_i^{tot} . Several examples are used to illustrate an application of DGSM.

Introduction

Global sensitivity analysis (SA) offers a comprehensive approach to the model analysis. Unlike local SA, global SA methods evaluate the effect of a factor while all other factors are varied as well and thus they account for interactions between variables and do not depend on the choice of a nominal point. Reviews of different global SA methods can be found in Saltelli et al [21] and Sobol and Kucherenko [27]. The method of global sensitivity indices suggested by Sobol [23, 24], and then further developed by Homma and Saltelli [7] is one of the most efficient and popular global SA techniques. It belongs to the class of variance-based methods. These methods provide information on the importance of different subsets of input variables to the output variance. There are two types of Sobol' sensitivity indices: the main effect indices, which estimate the individual contribution of each input parameter to the output variance, and the total sensitivity indices, which measure the total contribution of a single input factor or a group of inputs. The total sensitivity indices are used to identify non-important variables which can then be fixed at their nominal values to reduce model complexity. This approach is known as "factors' fixing setting" [21]. For high-dimensional models the direct application of variance-based global SA measures can be extremely time-consuming and impractical.

A number of alternative SA techniques have been proposed. One of them is the screening method by Morris [16]. It can be regarded as global as the final measure is obtained by averaging local measures (the elementary effects). This method is considerably cheaper than the variance based methods in terms of computational time. The Morris method can be used for identifying unimportant variables. However, the Morris method has two main drawbacks. Firstly, it uses random sampling of points from the fixed grid (levels) for averaging elementary effects which are calculated as finite differences with the increment delta comparable with the range of uncertainty. For this reason it can not correctly account for the effects with characteristic dimensions much less than delta. Secondly, it lacks the ability of the Sobol' method to provide information about main effects (contribution of individual variables to uncertainty) and it can't distinguish between low and high order interactions.

In this paper we present a survey of derivative based global sensitivity measures (DGSM) and their link with Sobol' sensitivity indices. DGSM are based on averaging local derivatives using Monte Carlo or Quasi Monte Carlo sampling methods. This technique is much more accurate than the Morris method as the elementary effects are evaluated as strict local derivatives with small increments compared to the variable uncertainty ranges. Local derivatives are evaluated at randomly or quasi randomly selected points in the whole range of uncertainty and not at the points from a fixed grid.

The so-called alternative global sensitivity estimator defined as a normalized integral of partial derivatives was firstly introduced by Sobol and Gershman [26]. Kucherenko et al [13] introduced some other derivative-based global sensitivity measures (DGSM) and coined the acronym DGSM. They showed that DGSM can be seen as the generalization of the Morris method [16]. Kucherenko et al [13] also established empirically the link between DGSM and Sobol' sensitivity indices. They showed that

the computational cost of numerical evaluation of DGSM can be much lower than that for estimation of Sobol' sensitivity indices.

Sobol and Kucherenko [28] proved theoretically that, in the cases of uniformly and normally distributed input variables, there is a link between DGSM and the Sobol' total sensitivity index S_i^{tot} for the same input. They showed that DGSM can be used as an upper bound on total sensitivity index S_i^{tot} . Small values of DGSM imply small S_i^{tot} , and hence unessential factors x_i . However, ranking influential factors using DGSM can be similar to that based on S_i^{tot} only for the case of linear and quasi-linear models. For highly non-linear models two rankings can be very different. They also introduced modified DGSM which can be used for both a single input and groups of inputs [29]. Such measures can be applied for problems with a high number of input variables to reduce the computational time. From DGSM, Kucherenko and Song [12] have also derived lower bounds on total sensitivity index. Lamboni et al [14] extended results of Sobol' and Kucherenko for models with input variables belonging to the general class of continuous probability distributions. In the same framework, Roustant et al [20] have defined crossed-DGSM, based on second-order derivatives of model output, in order to bound the total Sobol' indices of an interaction between two inputs.

The numerical efficiency of the DGSM method can be improved by using the automatic differentiation algorithm for calculation DGSM as was shown in Kiparissides et al [11]. However, the number of required function evaluations still remains to be proportional to the number of inputs. This dependence can be greatly reduced using an approach based on algorithmic differentiation in the adjoint or reverse mode [5]. It allows estimating all derivatives at a cost at most 4-6 times of that for evaluating the original function [9].

This paper is organised as follows: the Morris method and Sobol' global sensitivity indices are firstly described in the two following sections. DGSM and lower and upper bounds on total Sobol' sensitivity indices for uniformly and normally distributed random variables are then presented, followed by DGSM for groups of variables and their link with total Sobol' sensitivity indices. The next section presents the upper bounds results in the general case of variables with continuous probability distributions. Then, test cases illustrate an application of DGSM and their links with total Sobol' sensitivity indices. Finally, conclusions are presented in the last section.

The Morris method

Consider a differentiable function $G(\mathbf{x})$, where $\mathbf{x} = (x_i)_{i=1\dots d}$ is a vector of input variables defined in the unit hypercube H^d ($0 \leq x_i \leq 1$, $i = 1, \dots, d$). Local sensitivity measures are based on partial derivatives

$$E_i(\mathbf{x}^*) = \frac{\partial G(\mathbf{x}^*)}{\partial x_i}. \quad (1)$$

The local sensitivity measure $E_i(\mathbf{x}^*)$ depends on a nominal point \mathbf{x}^* and it changes with a change of \mathbf{x}^* . This deficiency can be overcome by averaging $E_i(\mathbf{x}^*)$ over the parameter space H^d .

The Morris method is traditionally used as a screening method for problems with a high number of variables for which function evaluations can be CPU-time consuming. It is composed of individually randomized 'one-factor-at-a-time' experiments. Each input factor may assume a discrete number of values, called levels, which are chosen within the factor range of variation.

The sensitivity measures proposed in the original work of Morris [16] are based on what is called an elementary effect. It is defined as follows. The range of each input variable is divided into p levels. Then the elementary effect (incremental ratio) of the i -th input factor is defined as

$$EE_i(\mathbf{x}^*) = \frac{[G(x_1^*, \dots, x_{i-1}^*, x_i^* + \Delta, x_{i+1}^*, \dots, x_d^*) - G(\mathbf{x}^*)]}{\Delta}, \quad (2)$$

where Δ is a predetermined multiple of $1/(p-1)$ and point $x^* \in H^d$ is such that $x_i^* + \Delta \leq 1$. One can see that the elementary effect are finite difference approximations of local sensitivity measures (1) with large Δ .

The distribution of elementary effects EE_i is obtained by randomly sampling N points from H^d . Two sensitivity measures are evaluated for each factor: μ_i an estimate of the mean of the distribution EE_i , and σ_i an estimate of the standard deviation of EE_i . A high value of μ_i indicates an input variable with an important overall influence on the output. A high value of σ_i indicates a factor involved in interaction with other factors or whose effect is nonlinear. The computational cost of the Morris method is $N_F = N(d+1)$. The revised version of the $EE_i(\mathbf{x}^*)$ measure and a more effective sampling strategy, which allows a better exploration of the space of the uncertain input factors was proposed by Campolongo et al [2]. To avoid the canceling effect which appears in non-monotonic functions Campolongo et al [2] introduced another sensitivity measure μ_i^* based on the absolute value of $EE_i(\mathbf{x}^*)$: $|EE_i(\mathbf{x}^*)|$. It was also noticed that μ_i^* has similarities with the total sensitivity index S_i^{tot} in that it can give a ranking of the variables similar to that based on the S_i^{tot} but no formal proof of the link between μ_i^* and S_i^{tot} was given [2].

Sobol' global sensitivity indices

Definitions

The method of global sensitivity indices developed by Sobol' is based on ANOVA decomposition [6]. Consider a square integrable function $G(\mathbf{x})$ defined in the unit hypercube H^d . It can be expanded in the following form

$$G(\mathbf{x}) = g_0 + \sum_i g_i(x_i) + \sum_{i < j} g_{ij}(x_i, x_j) + \dots + g_{12\dots d}(x_1, x_2, \dots, x_d). \quad (3)$$

This decomposition is unique if conditions $\int_0^1 g_{i_1\dots i_s} dx_{i_k} = 0$ for $1 \leq k \leq s$, are satisfied. Here $1 = i_1 < \dots < i_s = d$.

The variances of the terms in the ANOVA decomposition add up to the total variance of the function

$$V = \sum_{s=1}^d \sum_{i_1 < \dots < i_s} V_{i_1\dots i_s},$$

where $V_{i_1\dots i_s} = \int_0^1 g_{i_1\dots i_s}^2(x_{i_1}, \dots, x_{i_s}) dx_{i_1}, \dots, x_{i_s}$ are called partial variances.

Sobol' defined the global sensitivity indices as the ratios

$$S_{i_1\dots i_s} = V_{i_1\dots i_s}/V.$$

All $S_{i_1\dots i_s}$ are non negative and add up to one:

$$\sum_{i=1}^d S_i + \sum_i \sum_j S_{ij} + \sum_i \sum_j \sum_k S_{ijk\dots} + S_{1,2,\dots,d} = 1.$$

Sobol' also defined sensitivity indices for subsets of variables. Consider two complementary subsets of variables \mathbf{y} and \mathbf{z} :

$$\mathbf{x} = (\mathbf{y}, \mathbf{z}).$$

Let $\mathbf{y} = (x_{i_1}, \dots, x_{i_m})$, $1 \leq i_1 < \dots < i_m \leq d$, $K = (i_1, \dots, i_m)$. The variance corresponding to a set $y = \mathbf{y}$ is defined as

$$V_y = \sum_{s=1}^m \sum_{(i_1 < \dots < i_s) \in K} V_{i_1 \dots i_s}.$$

V_y includes all partial variances $V_{i_1}, V_{i_2}, \dots, V_{i_1 \dots i_s}$ such that their subsets of indices $(i_1, \dots, i_s) \in K$.

The total sensitivity indices were introduced by Homma and Saltelli [7]. The total variance V_y^{tot} is defined as

$$V_y^{\text{tot}} = V - V_z.$$

V_y^{tot} consists of all $V_{i_1 \dots i_s}$ such that at least one index $i_p \in K$ while the remaining indices can belong to the complementary to K set \bar{K} . The corresponding global sensitivity indices are defined as

$$\begin{aligned} S_y &= V_y/V, \\ S_y^{\text{tot}} &= V_y^{\text{tot}}/V. \end{aligned} \quad (4)$$

The important indices in practice are S_i and S_i^{tot} , $i = 1, \dots, d$:

$$\begin{aligned} S_i &= V_i/V, \\ S_i^{\text{tot}} &= V_i^{\text{tot}}/V. \end{aligned} \quad (5)$$

Their values in most cases provide sufficient information to determine the sensitivity of the analyzed function to individual input variables. Variance-based methods generally require a large number of function evaluations to achieve reasonable convergence and can become impractical for large engineering problems.

Link between total Sobol' sensitivity indices and partial derivatives

Consider continuously differentiable function $G(\mathbf{x})$ defined in the unit hypercube $H^d = [0, 1]^d$. In this section we present a theorem that establishes links between the index S_i^{tot} and the limiting values of $|\partial G / \partial x_i|$.

We note that in the case when $\mathbf{y} = (x_i)$, Sobol'-Jansen formula [10][25][22] for D_i^{tot} can be rewritten as

$$D_i^{\text{tot}} = \frac{1}{2} \int_{H^d} \int_0^1 \left[G(\mathbf{x}) - G\left(\overset{\circ}{\mathbf{x}}\right) \right]^2 dx dx'_i, \quad (6)$$

where $\overset{\circ}{\mathbf{x}} = (x_1, \dots, x_{i-1}, x'_i, x_{i+1}, \dots, x_n)$.

Theorem 1. Assume that $c \leq \left| \frac{\partial G}{\partial x_i} \right| \leq C$, then

$$\frac{c^2}{12V} \leq S_i^{\text{tot}} \leq \frac{C^2}{12V}. \quad (7)$$

Proof: Consider the increment of $G(\mathbf{x})$ in (6):

$$G(\mathbf{x}) - G\left(\overset{\circ}{\mathbf{x}}\right) = \frac{\partial G(\hat{\mathbf{x}})}{\partial x_i} (x_i - x'_i), \quad (8)$$

where $\hat{\mathbf{x}}$ is a point between \mathbf{x} and $\overset{\circ}{\mathbf{x}}$. Substituting (8) into (6) we obtain

$$V_i^{\text{tot}} = \frac{1}{2} \int_{H^d} \int_0^1 \left(\frac{\partial G(\hat{\mathbf{x}})}{\partial x_i} \right)^2 (x_i - x'_i)^2 d\mathbf{x} dx'_i. \quad (9)$$

In (9) $c^2 \leq (\partial G/\partial x_i)^2 \leq C^2$ while the remaining integral is

$$\int_0^1 \int_0^1 (x'_i - x_i)^2 dx'_i dx_i = \frac{1}{6}.$$

Thus we obtain inequalities that are equivalent to (7). Consider the function $G = g_0 + c(x_i - 1/2)$. In this case $C = c$, $V = 1/12$ and $S_i^{\text{tot}} = 1$ and the inequalities in (7) become equalities. \blacksquare

DGSM for uniformly and normally distributed variables

Uniformly distributed variables

Assume that $\partial G/\partial x_i \in L_2$. We define three different DGSM measures:

$$\nu_i = \int_{H^d} \left(\frac{\partial G(\mathbf{x})}{\partial x_i} \right)^2 d\mathbf{x}, \quad (10)$$

$$w_i^{(m)} = \int_{H^d} x_i^m \frac{\partial G(\mathbf{x})}{\partial x_i} d\mathbf{x}, \quad (11)$$

where $m > 0$ is a constant, and

$$\varsigma_i = \frac{1}{2} \int_{H^d} x_i(1 - x_i) \left(\frac{\partial G(\mathbf{x})}{\partial x_i} \right)^2 d\mathbf{x}. \quad (12)$$

We note that ν_i is in fact the mean value of $(\partial G/\partial x_i)^2$.

Lower bounds on S_i^{tot} To present further results we make the following notations and present some useful relationships. Denote $u_i(\mathbf{x})$ the sum of all terms in ANOVA (3) that depend on x_i :

$$u_i(\mathbf{x}) = g_i(x_i) + \sum_{j=1, j \neq i}^d g_{ij}(x_i, x_j) + \cdots + g_{12 \dots d}(x_1, \dots, x_d). \quad (13)$$

From the definition of ANOVA decomposition it follows that

$$\int_{H^d} u_i(\mathbf{x}) d\mathbf{x} = 0. \quad (14)$$

It is obvious that

$$\frac{\partial G}{\partial x_i} = \frac{\partial u_i}{\partial x_i}. \quad (15)$$

Denote $\mathbf{z} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d)$ the vector of all variables but x_i , then $\mathbf{x} \equiv (x_i, \mathbf{z})$ and $G(\mathbf{x}) \equiv G(x_i, \mathbf{z})$. The ANOVA decomposition of $G(\mathbf{x})$ (3) can be presented in the following form

$$G(\mathbf{x}) = u_i(x_i, \mathbf{z}) + v(\mathbf{z}),$$

where $v(\mathbf{z})$ is the sum of terms independent of x_i . Because of (14) it is easy to show that $v(\mathbf{z}) = \int_{H^d} G(\mathbf{x}) dx_i$. Hence

$$u_i(x_i, \mathbf{z}) = G(\mathbf{x}) - \int_{H^d} G(\mathbf{x}) dx_i. \quad (16)$$

The total partial variance V_i^{tot} can be computed as

$$V_i^{\text{tot}} = \int_{H^d} u_i^2(\mathbf{x}) d\mathbf{x} = \int_{H^d} u_i^2(x_i, \mathbf{z}) dx_i d\mathbf{z}.$$

Then the total sensitivity index S_i^{tot} (5) is equal to

$$S_i^{\text{tot}} = \frac{\int_{H^d} u_i^2(\mathbf{x}) d\mathbf{x}}{V}. \quad (17)$$

Theorem 2. There exists the following lower bound between DGSM (10) and the Sobol' total sensitivity index:

$$\frac{(\int_{H^d} [G(1, \mathbf{z}) - G(0, \mathbf{z})] [G(1, \mathbf{z}) + G(0, \mathbf{z}) - 2G(\mathbf{x})] d\mathbf{x})^2}{4\nu_i V} < S_i^{\text{tot}} \quad (18)$$

The proof of this Theorem is given in Kucherenko and Song [12] and is based on a Cauchy-Schwartz inequality applied on $\int_{H^d} u_i(\mathbf{x}) \frac{\partial u_i(\mathbf{x})}{\partial x_i} d\mathbf{x}$.

We call

$$\frac{(\int_{H^d} [G(1, \mathbf{z}) - G(0, \mathbf{z})] [G(1, \mathbf{z}) + G(0, \mathbf{z}) - 2G(\mathbf{x})] d\mathbf{x})^2}{4\nu_i V}$$

the lower bound number one (LB1).

Theorem 3. There exists the following lower bound, denoted $\gamma(m)$, between DGSM (11) and the Sobol' total sensitivity index:

$$\gamma(m) = \frac{(2m+1) \left[\int_{H^d} (G(1, \mathbf{z}) - G(\mathbf{x})) d\mathbf{x} - w_i^{(m+1)} \right]^2}{(m+1)^2 V} < S_i^{\text{tot}}. \quad (19)$$

The proof of this Theorem is given in Kucherenko and Song [12] and is based on a Cauchy-Schwartz inequality applied on $\int_{H^d} x_i^m u_i(\mathbf{x}) d\mathbf{x}$.

In fact, Theorem 3 gives a set of lower bounds depending on parameter m . We are interested in the value of m at which $\gamma(m)$ attains its maximum. Further we use star to denote such a value m : $m^* = \arg \max(\gamma(m))$ and call

$$\gamma(m^*) = \frac{(2m^*+1) \left[\int_{H^d} (G(1, \mathbf{z}) - G(\mathbf{x})) d\mathbf{x} - w_i^{(m^*+1)} \right]^2}{(m^*+1)^2 V} \quad (20)$$

the lower bound number two (LB2).

We define the maximum lower bound LB^* as

$$\text{LB}^* = \max(\text{LB1}, \text{LB2}). \quad (21)$$

We note that both lower and upper bounds can be estimated by a set of derivative based measures:

$$\mathcal{Y}_i = \{\nu_i, w_i^{(m)}, \zeta_i\}, \quad m > 0. \quad (22)$$

Upper bounds on S_i^{tot}

Theorem 4. There exists the following upper bound between DGSM (10) and the Sobol' total sensitivity index:

$$S_i^{\text{tot}} \leq \frac{\nu_i}{\pi^2 V}. \quad (23)$$

The proof of this Theorem is given in Sobol and Kucherenko [28]. It is based on inequality:

$$\int_0^1 u^2(x) dx \leq \frac{1}{\pi^2} \int_0^1 \left(\frac{\partial u}{\partial x} \right)^2 dx$$

and relationships (15) and (17).

Consider the set of values ν_1, \dots, ν_n , $1 \leq i \leq n$. One can expect that smaller ν_i correspond to less influential variables x_i . This importance criterion is similar to the modified Morris importance measure μ^* , whose limiting values are

$$\mu_i^* = \int_{H^n} \left| \frac{\partial G(\mathbf{x})}{\partial x_i} \right| d\mathbf{x}.$$

From a practical point of view the criteria μ_i and ν_i are equivalent: they are evaluated by the same numerical algorithm and are linked by relations $\nu_i \leq C\mu_i$ and $\mu_i \leq \sqrt{\nu_i}$.

We further call (23) the upper bound number one (UB1).

Theorem 5. There exists the following upper bound between DGSM (12) and the Sobol' total sensitivity index:

$$S_i^{\text{tot}} \leq \frac{\varsigma_i}{V}. \quad (24)$$

Proof: We use the following inequality [6]:

$$0 \leq \int_0^1 u^2 dx - \left(\int_0^1 u dx \right)^2 \leq \frac{1}{2} \int_0^1 x(1-x)u'^2 dx. \quad (25)$$

The inequality is reduced to an equality only if u is constant. Assume that u is given by (13), then $\int_0^1 u dx = 0$, and from (25) we obtain (24). \blacksquare

Further we call ς_i/D the upper bound number two (UB2). We note that $\frac{1}{2}x_i(1-x_i)$ for $0 \leq x_i \leq 1$ is bounded: $0 \leq \frac{1}{2}x_i(1-x_i) \leq 1/8$. Therefore, $0 \leq \varsigma_i \leq \nu_i/8$.

Computational costs All DGSM can be computed using the same set of partial derivatives $\frac{\partial G(\mathbf{x})}{\partial x_i}$, $i = 1, \dots, d$. Evaluation of $\frac{\partial G(\mathbf{x})}{\partial x_i}$ can be done analytically for explicitly given easily-differentiable functions or numerically:

$$\frac{\partial G(\mathbf{x}^*)}{\partial x_i} = \frac{[G(x_1^*, \dots, x_{i-1}^*, x_i^* + \delta, x_{i+1}^*, \dots, x_n^*) - G(\mathbf{x}^*)]}{\delta}. \quad (26)$$

Here δ is a small increment (we note a similarity with the elementary effect formula (2) which is however computed with large Δ).

In the case of straightforward numerical estimations of all partial derivatives (26) and computation of integrals using MC or QMC methods, the number of required function evaluations for a set of all input variables is equal to $N(d+1)$, where N is a number of sampled points. Computing LB1 also requires values of $G(0, \mathbf{z})$, $G(1, \mathbf{z})$, while computing LB2 requires only values of $G(1, \mathbf{z})$. In total, numerical computation of LB* for all input variables would require $N_G^{\text{LB}^*} = N(d+1) + 2Nd = N(3d+1)$ function evaluations. Computation of all upper bounds require $N_G^{\text{UB}} = N(d+1)$ function evaluations. We recall that the number of function evaluations required for computation of S_i^{tot} is $N_G^S = N(d+1)$ [22]. The number of sampled points N needed to achieve numerical convergence can be different for DGSM and S_i^{tot} . It is generally lower for the case of DGSM. Moreover, the numerical efficiency of the DGSM method can be significantly increased by using algorithmic differentiation in the adjoint (reverse) mode [5]. This approach allows estimating all derivatives at a cost independent of d , at most 4-6 times of that for evaluating the original function $G(\mathbf{x})$ [9].

Randomly distributed variables

Consider a function $G(X_1, \dots, X_d)$, where X_1, \dots, X_d are independent random variables, defined in the Euclidian space R^d , with cumulative density functions (cdfs) $F_1(x_1), \dots, F_d(x_d)$. The following DGSM was introduced in Sobol and Kucherenko [28]:

$$\nu_i = \int_{R^d} \left(\frac{\partial G(\mathbf{x})}{\partial x_i} \right)^2 dF(\mathbf{x}). \quad (27)$$

We introduce a new measure

$$w_i = \int_{R^d} \frac{\partial G(\mathbf{x})}{\partial x_i} dF(\mathbf{x}). \quad (28)$$

The lower bound on S_i^{tot} **Theorem 6.** If X_i is normally distributed with a finite variance σ_i^2 , there exists the following lower bound between DGSM (28) and the Sobol' total sensitivity index:

$$\frac{\sigma_i^2 w_i^2}{V} \leq S_i^{\text{tot}}. \quad (29)$$

The proof of this Theorem is given in Kucherenko and Song [12] and is based on a Cauchy-Schwartz inequality applied on $\int_{R^d} x_i u_i(\mathbf{x}) dF(\mathbf{x})$.

The upper bounds on S_i^{tot} The following Theorem 7 is a generalization of Theorem 1.

Theorem 7. If X_i has a finite variance σ_i^2 and $c \leq \left| \frac{\partial G}{\partial x_i} \right| \leq C$, then

$$\frac{\sigma_i^2 c^2}{V} \leq S_i^{\text{tot}} \leq \frac{\sigma_i^2 C^2}{V}. \quad (30)$$

The constant factor σ_i^2 cannot be improved.

Theorem 8. If X_i is normally distributed with a finite variance σ_i^2 , there exists the following upper bound between DGSM (27) and the Sobol' total sensitivity index:

$$S_i^{\text{tot}} \leq \frac{\sigma_i^2}{V} \nu_i. \quad (31)$$

The constant factor σ_i^2 cannot be reduced.

Proofs are presented in Sobol and Kucherenko [28].

Derivative based importance measure for groups of variables

Let $\mathbf{x} = (x_1, \dots, x_d)$ be a point in the d -dimensional unit hypercube with Lebesgue measure $d\mathbf{x} = dx_1 \cdots dx_d$. Consider an arbitrary subset of the variables $\mathbf{y} = (x_{i_1}, \dots, x_{i_s})$, $1 \leq s < d$, and the set of remaining complementary variables \mathbf{z} , so that $\mathbf{x} = (\mathbf{y}, \mathbf{z})$, $d\mathbf{x} = d\mathbf{y}d\mathbf{z}$. In the following, we write y instead of \mathbf{y} . Further all the integrals are written without integration limits. We assume that each integration variable varies independently from 0 to 1.

Consider the following DGSM τ_y :

$$\tau_y = \sum_{p=1}^s \int \left(\frac{\partial G(\mathbf{x})}{\partial x_{i_p}} \right)^2 \frac{1 - 3x_{i_p} + 3x_{i_p}^2}{6} d\mathbf{x}. \quad (32)$$

Theorem 9. If $G(\mathbf{x})$ is linear with respect to x_{i_1}, \dots, x_{i_s} , then $V_y^{\text{tot}} = \tau_y$, or in other words $S_y^{\text{tot}} = \frac{\tau_y}{V}$.

Theorem 10. A general inequality holds: $V_y^{\text{tot}} \leq (24/\pi^2) \tau_y$ or in other words $S_y^{\text{tot}} \leq \frac{24}{\pi^2} \frac{\tau_y}{V}$.

The proofs of these Theorems are given in Sobol and Kucherenko [29]. The second theorem shows that small values of τ_y imply small values of S_y^{tot} and this allows identification of a set of unessential factors y (usually defined by a condition of the type $S_y^{\text{tot}} < \epsilon$, where ϵ is small).

Importance criterion τ_i

Consider the one dimensional case when the subset y consists of only one variable $y = (x_i)$, then measure $\tau_y = \tau_i$ has the form

$$\tau_i = \int \left(\frac{\partial G(\mathbf{x})}{\partial x_i} \right)^2 \frac{1 - 3x_i + 3x_i^2}{6} d\mathbf{x}. \quad (33)$$

It is easy to show that $\nu_i/24 \leq \tau_i \leq \nu_i/6$. From UB1 it follows that

$$S_i^{\text{tot}} \leq \frac{24}{\pi^2} \frac{\tau_i}{V}. \quad (34)$$

Thus small values of τ_i imply small values of S_i^{tot} , that are characteristic for non important variables x_i . At the same time from Theorem 9 we obtain a corollary: if $G(\mathbf{x})$ depends linearly on x_i , then $S_i^{\text{tot}} = \tau_i/V$. Thus τ_i is closer to V_i^{tot} than ν_i .

Note that the constant factor $1/\pi^2$ in (23) is the best possible. But in the general inequality for τ_i (34) the best possible constant factor is unknown.

There is a general link between importance measures τ_i , ς_i and ν_i :

$$\tau_i = -\varsigma_i + \frac{1}{6}\nu_i,$$

then

$$\varsigma_i = \frac{1}{6}\nu_i - \tau_i.$$

Normally distributed random variables

Consider independent normal random variables X_1, \dots, X_d with parameters $(\mu_i; \sigma_i)$. Define τ_i as

$$\tau_i = \frac{1}{2} \mathbf{E} \left[\left(\frac{\partial G(\mathbf{x})}{\partial x_i} \right)^2 (x'_i - x_i)^2 \right].$$

The expectation over x'_i can be computed analytically. Then

$$\tau_i = \frac{1}{2} \mathbf{E} \left[\left(\frac{\partial G(\mathbf{x})}{\partial x_i} \right)^2 \frac{(x_i - \mu_i)^2 + \sigma_i^2}{2} \right].$$

Theorem 11. If X_1, \dots, X_d are independent normal random variables, then for an arbitrary subset y of these variables, we have

$$S_y^{\text{tot}} \leq \frac{2\tau_y}{V}.$$

The proof is given in Sobol and Kucherenko [29].

DGSM upper bounds in the general case

As previously, we consider the function $G(X_1, \dots, X_d)$, where X_1, \dots, X_d are independent random variables with cdfs $F_1(x_1), \dots, F_d(x_d)$. We further assume that each X_i admits a probability density function (pdf), denoted by $f_i(x_i)$. The developments in this section are based on the classical L^2 -Poincaré inequality:

$$\int G(\mathbf{x})^2 dF(\mathbf{x}) \leq C(F) \int \|\nabla G(\mathbf{x})\|^2 dF(\mathbf{x}) \quad (35)$$

which is valid for all functions G in $L^2(F)$ such that $\int G(\mathbf{x}) dF(\mathbf{x}) = 0$ and $\|\nabla f\| \in L^2(F)$. The constant $C(F)$ in Eq. (35) is called a Poincaré constant of F and the optimal Poincaré constant $C_{\text{opt}}(F)$ is the best possible constant. In measure theory, the Poincaré constants are expressed as a function of so-called Cheeger constants [1] which are used in Lamboni et al [14] (see Roustant et al [20] for more details).

A connection between total indices and DGSM has been established by Lamboni et al [14] for variables with continuous distributions (called Boltzmann probability measures in their paper).

Theorem 12. Let F_i and f_i be respectively the cdf and the pdf of X_i , we have

$$V_i^{\text{tot}} \leq C(F_i) \nu_i, \quad (36)$$

with ν_i the DGSM defined in Eq. (27) and

$$C(F_i) = 4 \left[\sup_{x \in \mathbb{R}} \frac{\min(F_i(x), 1 - F_i(x))}{f_i(x)} \right]^2. \quad (37)$$

This result comes from the direct application of the L^2 -Poincaré inequality (35) on $u_i(\mathbf{x})$ (see Eq. (13)).

In Lamboni et al [14] and Roustant et al [20], the particular case of log-concave probability distribution has been developed. It includes classical distributions as for instance the normal, exponential, Beta, Gamma and Gumbel distributions. In this case, the constant writes

$$C(F_i) = \frac{1}{f_i(\tilde{m}_i)^2} \quad (38)$$

with \tilde{m}_i the median of the distribution F_i . This allows to obtain analytical expressions for $C(F_i)$ in several cases [14]. In the case of a log-concave truncated distribution on $[a, b]$, the constant writes [20]

$$(F_i(b) - F_i(a))^2 / f_i \left(q_i \left(\frac{F_i(a) + F_i(b)}{2} \right) \right)^2 \quad (39)$$

with $q_i(\cdot)$ the quantile function of X_i .

For studying second-order interactions, Roustant et al [20] have derived a similar to (36) inequality based on the squared crossed derivatives of the function. Assuming that second-order derivatives of G are in $L^2(F)$, it uses the so-called crossed-DGSM

$$\nu_{i,j} = \int \left(\frac{\partial^2 G(\mathbf{x})}{\partial x_i \partial x_j} \right)^2 dF(\mathbf{x}), \quad (40)$$

introduced by Friedman and Popescu [3]. An inequality link is made with an extension of the total Sobol' sensitivity indices to general sets of variables (called superset importance or total interaction index) proposed by Liu and Owen [15]. In the case of a pair of variables $\{X_i, X_j\}$, the superset importance is defined as

$$V_{i,j}^{\text{super}} = \sum_{I \supseteq \{i,j\}} V_I. \quad (41)$$

The estimation methods of this total interaction index have also been studied by Fruth et al [4].

Theorem 13. For all pairs $\{i, j\}$ ($1 \leq i, j \leq d$),

$$V_{i,j} \leq V_{i,j}^{\text{super}} \leq C(F_i)C(F_j)\nu_{i,j}. \quad (42)$$

Roustant et al [20] have shown on several examples how to apply this result in order to detect pairs of inputs that do not interact together [17], [4].

Test cases

In this section we consider two test cases which illustrate application of DGSM and their links with S_i^{tot} .

Example 1. Consider a linear with respect to x_i function:

$$G(x) = a(\mathbf{z})x_i + b(\mathbf{z}).$$

For this function $S_i = S_i^{\text{tot}}$, $V_i^{\text{tot}} = \frac{1}{12} \int_{H^{d-1}} a^2(\mathbf{z}) d\mathbf{z}$, $\nu_i = \int_{H^{d-1}} a^2(\mathbf{z}) d\mathbf{z}$, $\text{LB1} = \frac{(\int_{H^d} (a^2(\mathbf{z}) - 2a^2(\mathbf{z})x_i) d\mathbf{z} dx_i)^2}{4V \int_{H^{d-1}} a^2(\mathbf{z}) d\mathbf{z}} = 0$ and $\gamma(m) = \frac{(2m+1)m^2 (\int_{H^{d-1}} a(\mathbf{z}) d\mathbf{z})^2}{4(m+2)^2(m+1)^2V}$. A maxi-

mum value of $\gamma(m)$ is attained at $m^*=3.745$, while $\gamma^*(m^*) = \frac{0.0401}{V} \left(\int a(\mathbf{z}) d\mathbf{z} \right)^2$. The

lower and upper bounds are $\text{LB}^* \approx 0.48S_i^{\text{tot}}$, $\text{UB1} \approx 1.22S_i^{\text{tot}}$. $\text{UB2} = \frac{1}{12V} \int_0^1 a(\mathbf{z})^2 d\mathbf{z} = S_i^{\text{tot}}$.

For this test function $UB2 < UB1$.

Example 2. Consider the so-called g-function which is often used in global SA for illustration purposes:

$$G(x) = \prod_{i=1}^d v_i,$$

where $v_i = \frac{|4x_i - 2| + a_i}{1 + a_i}$, $a_i (i = 1, \dots, d)$ are constants. It is easy to see that for this function $g_i(x_i) = (v_i - 1)$, $u_i(x) = (v_i - 1) \prod_{j=1, j \neq i}^d v_j$ and as a result $LB1=0$. The total variance is $V = -1 + \prod_{j=1}^d \left(1 + \frac{1/3}{(1 + a_j)^2}\right)$. The analytical values of S_i , S_i^{tot} and $LB2$ are given in Table 1.

Table 1. The analytical expressions for S_i , S_i^{tot} and $LB2$ for g-function.

S_i	S_i^{tot}	$\gamma(m)$
$\frac{1/3}{(1 + a_i)^2 V}$	$\frac{1/3}{(1 + a_i)^2} \prod_{j=1, j \neq i}^d \left(1 + \frac{1/3}{(1 + a_j)^2}\right)$	$(2m + 1) \left[1 - \frac{4(1 - (1/2)^{m+1})}{m+2}\right]^2$
	V	$(1 + a_i)^2 (m + 1)^2 V$

By solving equation $\frac{d\gamma(m)}{dm} = 0$, we find that $m^* = 9.64$, $\gamma(m^*) = \frac{0.0772}{(1 + a_i)^2 V}$. It is interesting to note that m^* does not depend on a_i , $i = 1, 2, \dots, d$ and d . In the extreme cases: if $a_i \rightarrow \infty$ for all i , $\frac{\gamma(m^*)}{S_i^{\text{tot}}} \rightarrow 0.257$, $\frac{S_i}{S_i^{\text{tot}}} \rightarrow 1$, while if $a_i \rightarrow 0$ for all i , $\frac{\gamma(m^*)}{S_i^{\text{tot}}} \rightarrow \frac{0.257}{(4/3)^{d-1}}$, $\frac{S_i}{S_i^{\text{tot}}} \rightarrow \frac{1}{(4/3)^{d-1}}$. The analytical expression for S_i^{tot} , $UB1$ and $UB2$ are given in Table 2.

Table 2. The analytical expressions for S_i^{tot} , $UB1$ and $UB2$ for g-function.

S_i^{tot}	$UB1$	$UB2$
$\frac{1/3}{(1 + a_i)^2} \prod_{j=1, j \neq i}^d \left(1 + \frac{1/3}{(1 + a_j)^2}\right)$	$16 \prod_{j=1, j \neq i}^d \left(1 + \frac{1/3}{(1 + a_j)^2}\right)$	$4 \prod_{j=1, j \neq i}^d \left(1 + \frac{1/3}{(1 + a_j)^2}\right)$
V	$(1 + a_i)^2 \pi^2 V$	$3(1 + a_i)^2 V$

For this test function $\frac{S_i^{\text{tot}}}{UB1} = \frac{\pi^2}{48}$, $\frac{S_i^{\text{tot}}}{UB2} = \frac{1}{4}$, hence $\frac{UB2}{UB1} = \frac{\pi^2}{12} < 1$.

Values of S_i , S_i^{tot} , $UB1$, $UB2$ and $LB2$ for the case of $\mathbf{a} = [0, 1, 4.5, 9, 99, 99, 99, 99]$, $d=8$ are given in Table 3 and shown in Fig. 1. One can see that the knowledge of $LB2$ and $UB1$ allows to rank correctly all the variables in the order of their importance.

Table 3. Values of LB^* , S_i , S_i^{tot} , $UB1$ and $UB2$. Example 2, $\mathbf{a} = [0, 1, 4.5, 9, 99, 99, 99, 99]$, $d=8$.

	x_1	x_2	x_3	x_4	$x_5 \dots x_8$
LB^*	0.166	0.0416	0.00549	0.00166	0.000017
S_i	0.716	0.179	0.0237	0.00720	0.0000716
S_i^{tot}	0.788	0.242	0.0343	0.0105	0.000105
$UB1$	3.828	1.178	0.167	0.0509	0.00051
$UB2$	3.149	0.969	0.137	0.0418	0.00042

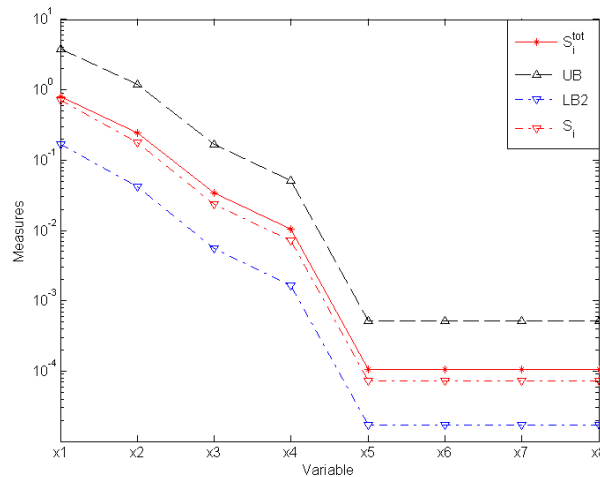


Fig. 1. Values of S_i , S_i^{tot} , LB2 and UB1 for all input variables. Example 2 with $\mathbf{a} = [0, 1, 4.5, 9, 99, 99, 99, 99]$, $d = 8$.

Conclusions

We can conclude that using lower and upper bounds based on DGSM it is possible in most cases to get a good practical estimation of the values of S_i^{tot} at a fraction of the CPU cost for estimating S_i^{tot} . Upper and lower bounds can be estimated using MC/QMC integration methods using the same set of partial derivative values. Most of the applications show that DGSM can be used for fixing unimportant variables and subsequent model reduction because small values of DGSM imply small values of S_i^{tot} . In a general case variable ranking can be different for DGSM and variance based methods but for linear function and product function, DGSM can give the same variable ranking as S_i^{tot} .

Applications of DGSM can be found for instance in Kiparissides et al [11] and Rodriguez-Fernandez et al [19] for biological systems modeling, Patelli et al [18] for structural mechanics, Iooss et al [8] for an aquatic prey-predator model and Roustant et al [20] for a simple river flood model. One of the main prospect in practical situations is to use algorithmic differentiation in the reverse (adjoint) mode on the numerical model, allowing to estimate efficiently all partial derivatives of this model. In this case, the cost of DGSM estimations would be independent of the number of input variables.

Acknowledgements

The authors would like to thank Prof. I. Sobol', Dr. S. Song, Dr. O. Roustant and Prof. F. Gamboa for their contributions to this work. One of the authors (SK) gratefully acknowledges the financial support by the EPSRC grant EP/H03126X/1.

Recommended Reading

1. Bobkov SG (1999) Isoperimetric and analytic inequalities for log-concave probability measures. *The Annals of Probability* 27(4):1903–1921
2. Campolongo F, Cariboni J, Saltelli A (2007) An effective screening design for sensitivity analysis of large models. *Environmental Modelling and Software* 22:1509–1518
3. Friedman J, Popescu B (2008) Predictive Learning via Rule Ensembles. *The Annals of Applied Statistics* 2(3):916–954

4. Fruth J, Roustant O, Kuhnt S (2014) Total interaction index: A variance-based sensitivity index for second-order interaction screening. *Journal of Statistical Planning and Inference* 147:212–223
5. Griewank A, Walther A (2008) Evaluating derivatives: Principles and techniques of automatic differentiation. SIAM Philadelphia
6. Hardy G, Littlewood J, Polya G (1973) Inequalities. Cambridge University Press, Second edition
7. Homma T, Saltelli A (1996) Importance measures in global sensitivity analysis of non linear models. *Reliability Engineering and System Safety* 52:1–17
8. Iooss B, Popelin AL, Blatman G, Ciric C, Gamboa F, Lacaze S, Lamboni M (2012) Some new insights in derivative-based global sensitivity measures. In: *Proceedings of the PSAM11 ESREL 2012 Conference*, Helsinki, Finland, pp 1094–1104
9. Jansen K, Leovey H, Nube A, Griewank A, Mueller-Preussker M (2014) A first look of quasi-Monte Carlo for lattice field theory problems. *Computer Physics Communication* 185:948–959
10. Jansen M (1999) Analysis of variance designs for model output. *Computer Physics Communication* 117:25–43
11. Kiparissides A, Kucherenko S, Mantalaris A, Pistikopoulos E (2009) Global sensitivity analysis challenges in biological systems modeling. *Journal of Industrial and Engineering Chemistry Research* 48:1135–1148
12. Kucherenko S, Song S (2014, submitted) Derivative-based global sensitivity measures as bounds on variance-based global sensitivity indices. *Mathematics and Computers in Simulation*
13. Kucherenko S, Rodriguez-Fernandez M, Pantelides C, Shah N (2009) Monte carlo evaluation of derivative-based global sensitivity measures. *Reliability Engineering and System Safety* 94:1135–1148
14. Lamboni M, Iooss B, Popelin AL, Gamboa F (2013) Derivative-based global sensitivity measures: general links with sobol' indices and numerical tests. *Mathematics and Computers in Simulation* 87:45–54
15. Liu R, Owen A (2006) Estimating mean dimensionality of analysis of variance decompositions. *Journal of the American Statistical Association* 101(474):712–721
16. Morris M (1991) Factorial sampling plans for preliminary computational experiments. *Technometrics* 33:161–174
17. Muehlenstaedt T, Roustant O, Carraro L, Kuhnt S (2012) Data-driven Kriging models based on FANOVA-decomposition. *Statistics & Computing* 22:723–738
18. Patelli E, Pradlwarter HJ, Schuëller GI (2010) Global sensitivity of structural variability by random sampling. *Computer Physics Communications* 181:2072–2081
19. Rodriguez-Fernandez M, Banga J, Doyle F (2012) Novel global sensitivity analysis methodology accounting for the crucial role of the distribution of input parameters: application to systems biology models. *International Journal of Robust Nonlinear Control* 22:1082–1102
20. Roustant O, Fruth J, Iooss B, Kuhnt S (2014) Crossed-derivative-based sensitivity measures for interaction screening. *Mathematics and Computers in Simulation* 105:105–118
21. Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Salsana M, Tarantola S (2008) *Global sensitivity analysis - The primer*. Wiley
22. Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M, Tarantola S (2010) Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communication* 181:259–270
23. Sobol I (1990) Sensitivity estimates for non linear mathematical models (in Russian). *Matematicheskoe Modelirovanie* 2:112–118
24. Sobol I (1993) Sensitivity estimates for non linear mathematical models. *Mathematical Modelling and Computational Experiments* 1:407–414
25. Sobol I (2001) Global sensitivity indices for non linear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation* 55:271–280
26. Sobol I, Gershman A (1995) On an alternative global sensitivity estimators. In: *Proceedings of SAMO 1995*, Belgirate, pp 40–42
27. Sobol I, Kucherenko S (2005) Global sensitivity indices for non linear mathematical models. Review. *Wilmott Magazine* 1:56–61
28. Sobol I, Kucherenko S (2009) Derivative based global sensitivity measures and their links with global sensitivity indices. *Mathematics and Computers in Simulation* 79:3009–3017
29. Sobol I, Kucherenko S (2010) A new derivative based importance criterion for groups of variables and its link with the global sensitivity indices. *Computer Physics Communications* 181:1212 – 1217

svcyclop-warning: **missing mandatory text elements**

No **Problem Definition** given - use `\ProbDef` heading

No **Key Results** given - use `\KeyRes` heading