Model Emulation and Moment Independent Sensitivity Analysis: An Application to Environmental Modelling

E. Borgonovo[1], W. Castaings[2] and S. Tarantola[3]

[1]Dept. of Decision Sciences and ELEUSI, Bocconi University, Via Roentgen 1, 20136, Milano, Italy

[2]EDYTEM (Université de Savoie, CNRS), Campus Scientifique, Bâtiment Pôle Montagne, F-73376 Le Bourget du Lac cedex

[3]Statistical Indicators for Policy Assessment, Joint Research Centre of the European Commission, TP 361, Via E. Fermi, 2749 - 21027 Ispra (Italy)

Abstract

Moment-independent sensitivity methods are attracting increasing attention among practitioners, since they provide a thorough way of investigating the sensitivity of model output under uncertainty. However, their estimation is challenging, especially in the presence of computationally intensive models. We argue that replacement of the original model by a metamodel can contribute in lowering the computation burden. A numerical estimation procedure is set forth. The procedure is first tested on analytical cases with increased structural complexity. We utilize the emulator proposed in Ratto and Pagano (2010). Results show that the emulator allows an accurate estimation of density-based sensitivity measures, when the main structural features of the original model are captured. However, performance deteriorates for a model with interactions of order higher than 2. For this test case, also a Kriging emulator is investigated, but no gain in performance is registered. However, an accurate estimation is obtained by applying a logarithmic transformation of the model output for both the Kriging and Ratto and Pagano (2010) emulators. These findings are then applied to the investigation of a benchmark environmental case study, the LevelE model. Results show that use of the metamodel allows an efficient estimation of moment-independent sensitivity measures while leading to a drastic reduction in computational burden.

Keywords: *Global Sensitivity Analysis; Uncertainty; Meta-Modelling; Environmental Models.*

1 Introduction

In environmental sciences, computer models play an essential role in guiding analysts' understanding of natural systems behaviour. When the phenomena under investigation cover large temporal and spatial scales, quite elaborate calculations interpose themselves between the model input and output [Yu (2010)]. In addition, uncertainty characterizes both the model building and utilization phases [Risbey et al. (2005)]. *"To help determine when a model, despite its uncertainties, can be appropriately used to inform a decision,"* the US Environmental Protection Agency explicitly suggests sensitivity analysis as part of best practices [US EPA, 2009; p. vii.]. Specifically, it recommends that *"model developers and users ... perform sensitivity and uncertainty analyses. Sensitivity analysis evaluates the effect of changes in input values or assumptions on a model's*

results [US EPA, 2009; p. vii.]" Sensitivity and uncertainty analyses are essential ingredients of the model quality assurance checklist [Risbey et al. (2005)]. Without a proper sensitivity analysis, one is exposed to the so-called black-box effect, namely the risk of not fully understanding the behavior of the model on which the inference is based. Also, from the point of view of exploiting a model's information content, Rabitz (1989) observes that sensitivity analysis (SA) techniques "*appear to be the key ingredient needed to draw out the maximum capabilities of mathematical modelling.*"

The utilization of SA methods in the environmental literature is increasing. In the analysis of the behavior of a landscape based sediment source and transport model, Newham et al. (2003) utilize differentiation methods. Castaings et al. (2009) show that the use of adjoint sensitivity analysis — implemented using the reverse mode of algorithmic differentiation [Cacuci (1981), Griewank (2000)] — enables a local but extensive analysis (spatio-temporal sensitivity analysis) of the input-output mapping operated by a rainfall-runoff model. In Norton (2008), analytical expressions for sensitivity measures with finite factor[1] changes for use in environmental modelling are proposed. Norton (2008)'s algebraic approach allows one to estimate both first and second-order sensitivity measures which are independent of the magnitudes of the changes. The approach also provides one with a way for detecting interactions via the Hessian matrix. The identification of interactions is the natural bridge towards screening methods. Examples of works in the environmental literature using screening methods are Campolongo and Braddock (1999), where the Morris method is applied in the SA of the IMAGE model for greenhouse emissions, and Cryer and Havens (1999), where a fractional factorial scheme based on Plackett and Burman (PB) design is used in the study of an air pollution model. The Morris method is then refined in Campolongo et al. (2007) and Saltelli and Annoni (2010) to better account for interactions.

When the goal of the SA exercise becomes assessing "*the contribution of* factors *to uncertainties in the model output* [Kohler and Wirtz (2002), p. 613]," then global SA methods become the appropriate ones, because exploring the input factor space at a few points is of limited value in understanding the consequences of uncertainty [Oakley and O'Hagan (2004)]. Conversely, global SA methods grant one with a thorough exploration of the factor uncertainty ranges [Saltelli et al. (1993)]. Among the first developed and mostly employed global SA methods are non-parametric techniques [Saltelli and Marivoet (1990), Helton (1993)]. In environmental modelling, they are applied by Manache and Melching (2008), in the context of the uncertainty and global SA of water quality models. Variance-based methods are used by Varella et al. (2010) in the global SA of a crop model, by Estrada and Diaz (2010) in the global SA of eutrophication models. Confalonieri et al. (2010) apply variance-based and screening methods in order to analyze the WARM rice growth model. In Confalonieri et al. (2010), variance-based sensitivity measures are estimated by the Sobol' method [Sobol' (1993)]. Confalonieri et al. (2010) compare variance-based results to screening results obtained by the Morris method. Discrepancies are revealed in the factor ranking, in accordance with the fact that the two types of techniques have different scopes and screening

---

[1]By factors we mean all those numerical values (parameters, coefficients, exogenous variables) over which the analyst wishes to carry out a sensitivity analysis.

methods become only partially informative in the presence of uncertainty [Oakley and O'Hagan (2004)].

A further way for exploring the sensitivity of a model output in the presence of uncertainty is represented by distribution-based techniques (also called moment-independent). A technique is distribution-based (or moment-independent) if the associated global sensitivity measures explicitly consider the entire model output distribution (either the cumulative distribution function or the density) instead of relying on a particular moment of this distribution (e.g., variance). They are well-suited to address the decision-maker's state of knowledge in the model output, especially when distributions are skewed or multimodal. In fact, in these cases, uncertainty cannot be completely captured by variance [Huang and Litzenberger (1988)]. Figure 1 presents the distribution of the output of the environmental model utilized in this work.
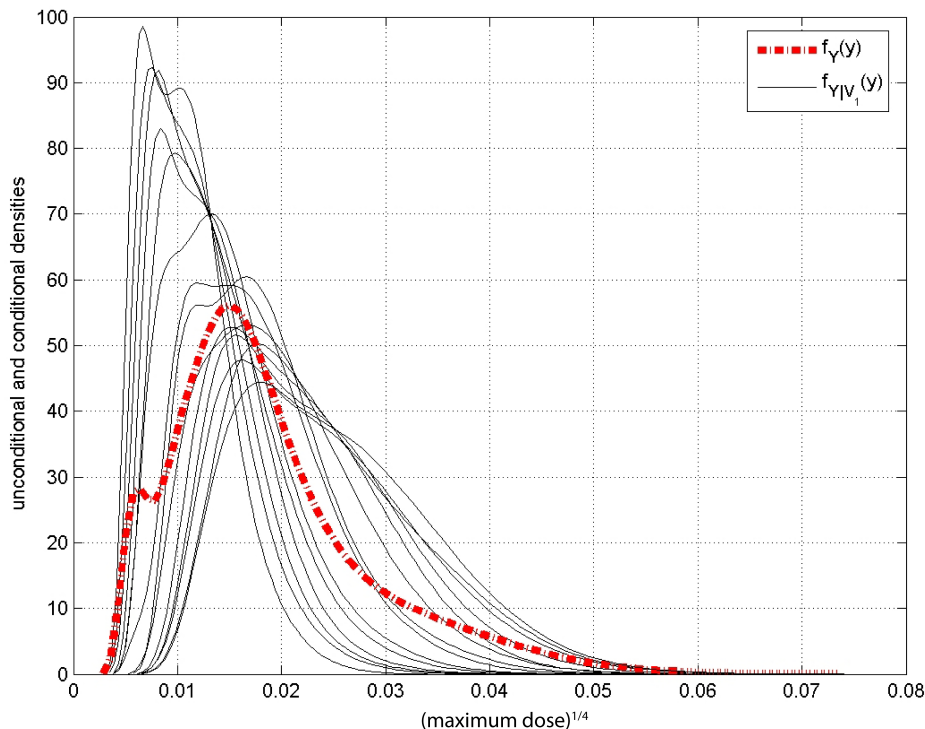
Figure 1: Conditional distributions of the LevelE model output obtained by conditioning on different values of one of the model inputs (in this case $V^{(1)}$, namely water speed in geosphere's layer 1). Due to the large range of the model output (maximum dose) a monotonic transformation $[y^{1/4}]$ was applied to generate this plot.

In Figure 1, the dotted line represents the model output density $[f_Y(y)]^2$. The continuous lines represent the set of conditional densities $[f_{Y|V^{(1)}=v^{(1)}}(y)]$ obtained when one is informed that factor

---

[2]In the remainder, we shall refer to $f_Y(y)$ as "unconditional model output density", to distinguish it from the conditional $f_{Y|X_i=x_i}(y)$, when necessary.

$V^{(1)}$ — namely water speed in geosphere's layer 1 — has assumed a certain value. Figure 1 shows that these distributions are skewed and can be multi-modal. Figure 1 reveals that the effect of fixing $V^{(1)}$ at $v^{(1)}$ is a change in the entire distribution of the model output. Sensitivity methods that quantify a factor's influence by measuring the separation between the model output density $f_Y(y)$ and the conditional densities $f_{Y|X_i=x_i}(y)$ belong to the category of distributional or moment-independent sensitivity analysis methods [Park and Ahn (1994), Chun et al. (2000), Borgonovo (2006), Borgonovo (2007), Liu and Homma (2009)]. These methods have recently attracted the attention of analysts and practitioners, because they allow a thorough consideration of uncertainty and are well-posed in the presence of correlated model inputs [Borgonovo and Tarantola (2008)]. However, their estimation is a challenging task. An accurate inspection of the factor uncertainty ranges might require large samples and several thousands of model runs. In the presence of computationally intensive models, this might render moment-independent methods inapplicable. Two ways might be envisaged to solve the problem: one is the application of a screening method, followed by the calculation of moment-independent importance measures on a reduced number of factors. The second consists of the substitution for the model of a metamodel (or emulator) that drastically reduces the computational time, while allowing estimation of moment-independent measures for all factors considered in the analysis. In this work, we investigate this second approach. In this respect, we employ the implementation proposed by Ratto and Pagano (2010), which is related to the ANOVA smoothing spline context [Gu (2002)].

To obtain a structured methodology for model validation and audit, based on the combination of distribution-based sensitivity measures and emulators, we proceed as follows. First, the combination moment-independent techniques and emulator is assessed on a set of analytical test cases: a multiplicative model with lognormally distributed model inputs, a non-additive model with gamma-distributed model inputs, and the Ishigami test function. For the first two test cases analytical estimates of the moment-independent sensitivity measures are available [Borgonovo et al. (2011)]. They allow us to assess the performance of our method based on the joint utilization of moment-independent techniques and emulators in a variety of model structures and factor distributions. We then discuss the application to a complex environmental model developed by OECD [OECD (1989), OECD (1993)] for predicting the radiologic release to humans due to the underground migration of radionuclides from a nuclear waste disposal site. The model is known as LevelE [OECD (1989), OECD (1993)], and, with time, has become the benchmark model in global sensitivity analysis studies [Saltelli and Marivoet (1990), Saltelli and Tarantola (2002), Ratto et al. (2007)]

The remainder of the paper is organized as follows. Section 2 reviews the principles of emulators and the definitions and properties of distribution-based sensitivity analysis. Section 3 presents sampling plans and the systematic procedure. Section 4 discusses numerical experiments and results for three analytical test cases. Section 5 illustrates the application to the numerical test case. Section 6 offers conclusions and future research perspectives.

2 Moment-independent importance measures and Model Emulation: A Review

2.1 Distributional Sensitivity Analysis: a brief Overview

In this section, we sketch the principles of distribution-based (or moment-independent) sensitivity analysis with the $\delta-$importance measure.

Let $\Omega_X \subseteq \mathbb{R}^n$ denote the set of all possible values that the factors can assume. Let $\mathbf{x} \in \Omega_X$ be a possible value of the model inputs. If $\mathbf{x}$ is not known with certainty, we let $f_{\mathbf{X}}(\mathbf{x})$ denote the distribution characterizing one's state of knowledge about the model inputs. The value $\mathbf{x}$ is, then, one of the possible realizations of $\mathbf{X}$. $\mathbf{X}$ is a random vector and the marginal distribution of one of its components is denoted by $f_{X_i}(x_i)$. We denote by

$$y = g(\mathbf{x}), \, g : \Omega_X \subseteq \mathbb{R}^n \to \Omega_Y \subseteq \mathbb{R} \tag{1}$$

the relationship that links the model inputs to the model output. Because $\mathbf{X}$ is uncertain, y becomes a random variable, denoted by Y. The image of g, Ω_Y, then coincides with the support of Y.

In order to assess the sensitivity of y to $\mathbf{x}$, when $\mathbf{x}$ is uncertain, a global SA method needs to be used. If one utilizes a variance-based method [Saltelli and Tarantola (2002), Oakley and O'Hagan (2004)], then one is framing the SA quest within the following SA setting: "*We are asked to bet on the factor that, if determined (i.e., fixed to its true value), would lead to the greatest expected reduction in the variance of Y* [Saltelli and Tarantola (2002), p. 705]." The appropriate sensitivity measures are [Iman and Hora (1990), Saltelli and Tarantola (2002)]

$$S_i = \frac{V[Y] - \mathbb{E}_{X_i}[V\{Y|X_i\}]}{V[Y]} = \frac{V_{X_i}[\mathbb{E}\{Y|X_i\}]}{V[Y]} \tag{2}$$

where $V[Y]$ is the model output variance. A value $S_i > S_j$ implies that fixing X_i leads to a greater expected reduction in $V[Y]$ than fixing X_j. X_i is, therefore, deemed as more important than X_j.

In a distribution-based sensitivity analysis, a different setting is proposed. Namely, the SA setting is: "*We are asked to bet on the factor that, if determined, would lead to the greatest expected modification in the distribution of Y* [Borgonovo and Tarantola (2008).]"

Distribution-based importance measures can be originally found in Park and Ahn (1994), where the Kullback-Leibler divergence measure is used to quantify the shift. Limitations connected with this metric are then overcome in Chun et al. (2000) and generalized in Borgonovo (2006) and Borgonovo (2007) — see Borgonovo et al. (2011) for further details. — The sensitivity measure of interest in this work is defined as
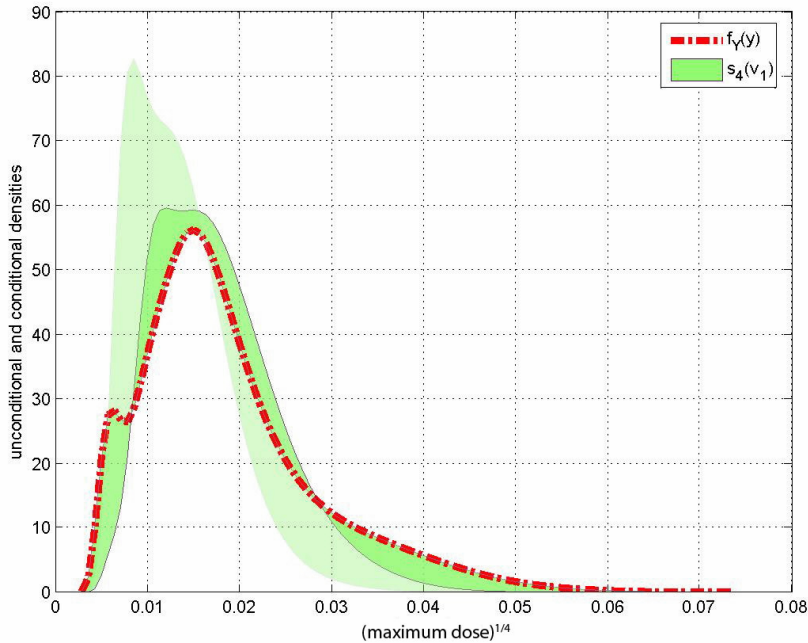
$$\delta_i := \frac{1}{2}\mathbb{E}_{X_i}[s_i(X_i)] \tag{3}$$

where

$$s_i(x_i) := \int_{\Omega_Y} \left| f_Y(y) - f_{Y|X_i = x_i}(y) \right| \mathrm{d}y \tag{4}$$

$s_i(x_i)$ [eq. (4)] is the shift between the model output distribution [$f_Y(y)$] and the conditional model

output distribution given that factor X_i is fixed at x_i $[f_{Y|X_i=x_i}(y)]$. A visual representation of this concept is offered in Figure 2.

Figure 2: $s_{V^{(1)}}(v^{(1)})$ between $f_{Y|V^{(1)}=v^{(1)}}(y)$ and $f_Y(y)$ the output of the model of interest in this work. The two shaded areas represent the separation between two conditional densities obtained by fixing a model input (the same as in Figure 1) at two different values.

Figure 2 portrays $f_Y(y)$ and two of the conditional distributions $[f_{Y|V^{(1)}=v^{(1)}}(y)]$ obtained by fixing one factor ($V^{(1)}$ as in Figure 1) at two different values. It also shows the areas (shaded regions) between the conditional and unconditional densities. Geometrically, in fact, $s_i(x_i)$ coincides with the area enclosed between $f_Y(y)$ and $f_{Y|X_i=x_i}(y)$. From a technical viewpoint, $s_i(x_i)$ is a separation measurement in the sense of Glick (1975). Because the area between $f_Y(y)$ and $f_{Y|X_i=x_i}(y)$ depends on the value at which X_i is fixed, the separation needs to be averaged over the possible values of X_i. This is captured by the expectation operator $\mathbb{E}_{X_i}[s_i(x_i)]$ in eq. (3).

The properties of δ_i of particular interest in this work are: *i)* individual and joint normalization [Borgonovo (2007)]; and *ii)* invariance for monotonic transformation [Borgonovo et al. (2011)].

As far as *i)* is concerned, it can be shown that $0 \leq \delta_i \leq 1$. In particular, if Y and X_i are independent, $\delta_i = 0$. Consider then the joint importance of all factors. Let us denote it by $\delta_{1,2,\dots,n}$. By definition,

$$\delta_{1,2,\dots,n} = \frac{1}{2}\mathbb{E}_{X_i}\Big[\int_{\Omega_Y}\big|f_Y(y) - f_{Y|X_1=x_1,X_2=x_2,\dots,X_n=x_n}(y)\big|\,\mathrm{d}y\Big] \tag{5}$$

6

In eq. (5), $f_{Y|X_1=x_1,X_2=x_2,...,X_n=x_n}(y)$ is a Dirac $\delta-$function[3] centered at $(x_1, x_2, ..., x_n)$. $\delta_{1,2,...,n}$ is, therefore, the expected distance between the current state of uncertainty and the state in which all factors are known. It can be shown [Borgonovo (2007)] that $\delta_{1,2,...,n} = 1$ (independently of the point $(x_1, x_2, ..., x_n)$). This results indicates that the distance one travels between the current state and the state in which all model inputs are known is unity, when measured with the L^1 norm.

As far as property *ii)* is concerned, consider the model output y and a monotonic transformation $t(y)$. $t(y)$ can be, for instance, the result of a rescaling of y. Then, it can be shown that δ_i is invariant if estimated on y or on their transformed $t(y)$. As underlined by Iman and Hora (1990), in the presence of long tail input/output distributions robustness problems might emerge in the estimation of statistical quantities as, for example, variance-based sensitivity measures. To circumvent this problem, Iman and Hora (1990) propose to use a logarithmic transformation, which indeed leads to an improvement in the computational robustness of the analysis. However, as underlined by Saltelli and Sobol' (1995) and investigated in detail for rank transformation, the outcomes of a sensitivity analysis drawn using data transformation cannot be easily transferred back to the original model. Moment independent importance measures are not affected by this limitation because they are scale invariant.

2.2 Model Emulation

The $\delta-$importance measure can be a very effective tool in situations that involve nonlinear and non-monotonic relationships between inputs and outputs. However, its implementation requires a large number of model evaluations. Therefore, its use risks to be restricted to models where thousands of model evaluations are possible in a reasonable computer time.

Sometimes the relationship between input and output can only be evaluated numerically, solving systems of non-linear differential equations, and its form remains usually unknown to the analyst, while some sort of direct representation would make the model's properties more transparent [Saltelli et al. (2008)].

Metamodelling techniques can mitigate the potential problems indicated above. In the principles of metamodelling, the relationship [g in eq. (1)] that binds the model inputs to the model output is considered as a black box [Sacks et al. (1989), Santner et al. (2003), Oakley and O'Hagan (2004); see Bayarri et al. (2009) for a recent overview]. A metamodel, then, is a mapping that links the model inputs to the model output through a known relationship. The mapping reproduces (or emulates) the original model but is less computationally demanding. A vast literature exists on this subject. Besides local approximation methods, which make use of the Taylor expansion formula to construct a simple function that fits the model in the neighborhood of a nominal point in the input space, parametric regression techniques (either linear or non-linear, stepwise or based on ranks) are also able to produce a metamodel, though it is necessary to provide a prior specification of its exact algebraic form. This specification is required to hold across the entire mapping from inputs

---

[3]Given a function $t(y)$, the Dirac-δ function centered at y^0 is a function $\delta(y)$ such that $\int\limits_{-\infty}^{+\infty} t(y)\delta(y - y^0)\mathrm{d}y = y^0$. In the theory of distributions, $\delta(y)$ is the derivative of the Heaviside function.

7

to outputs, making the representation of the local behavior difficult [Storlie and Helton (2008)].

With non-parametric regression techniques (often referred to as smoothing techniques) an iterative construction procedure is used to build an approximating function that can represent local patterns [Storlie and Helton (2008), Storlie et al. (2009).] Popular smoothing procedures include locally weighted regression (LOESS) [Simonoff (1996)], generalized additive models (GAM) [Hastie and Tibshirani (1992)], projection pursuit regression [Friedman and Stuetzle (1981)], recursive partitioning regression (also known as regression tree) [Breiman et al. (1984)], multivariate adaptive regression splines [Friedman (1991)], gradient boosting regressions [Friedman (2001)] and the adaptive component selection and shrinkage operator (ACOSSO) [Storlie et al. (2010)].

In all the aforementioned works, as well as in Ratto et al. (2007), Sudret (2008) and Ziehn and Tomlin (2009), the starting point of the metamodel building process is the representation of g in eq. (1) in terms of the ANOVA model:

$$g(\mathbf{x}) = g_0 + \sum_{i=1}^{n} g_i(x_i) + \sum_{i<j} g_{i,j}(x_i, x_j) + ... + g_{1,2,...n}(x_1, x_2, ..., x_n) \tag{6}$$

where

$$\begin{cases} g_0 = \mathbb{E}[g] = \int \cdots \int g(\mathbf{x})d\mathbf{x} \\ g_i(x_i) = \mathbb{E}[g|x_i] - g_0 = \int \cdots \int g(\mathbf{x}) \prod_{k \neq i} dx_k - g_0 \\ g_{i,j}(x_i,x_j) = \mathbb{E}[g|x_i, x_j] - g_i(x_i) - g_j(x_j) - g_0 \\ ... \end{cases} \tag{7}$$

In eq. (6), g_0 is the expected value of g, the functions $g_i(x_i)$ account for the factors individual effects, the second-order terms $[g_{i,j}(x_{i,j})]$ account for the residual cooperation of all factor pairs, the third-order terms for the cooperation of all triplets, etc. Then, one looks for the proper truncation order at which the original model can be accurately mapped by the metamodel. One speaks about a first, second, third-order metamodel, depending on the order at which eq. (6) is truncated.

This representation, based on smoothing splines ANOVA models Gu (2002), is at the basis of the metamodels of Ratto et al. (2007), Ziehn and Tomlin (2009), Sudret (2008) [see also Blatman and Sudret (2010)]. The metamodels of Ziehn and Tomlin (2009), Ratto et al. (2007) and Sudret (2008) differ in the way the component functions in eq. (6) are approximated. Sudret (2008) and Ziehn and Tomlin (2009) use orthonormal polynomials, while the emulator by Ratto et al (2007) employs a different approach based on State-Dependent Regression (SDR), and recursive Kalman smoothing [Young (2001)]. SDR utilizes piecewise-defined polynomials (cubic splines) to construct approximating functions that are only piecewise smooth. The simplest example of smoothing spline estimation of an input/output mapping $z(x)$ is the additive model:

$$g(\mathbf{x}) = g_0 + \sum_{j=1}^{n} g_j(x_j) \tag{8}$$

which can be estimated using a multivariate cubic smoothing spline minimization problem with a

penalized residual sum of squares

$$1/N \sum_{i=1}^{N}[z_i - g(x_i)]^{\mathbf{2}} + \sum_{j=1}^{p} \lambda_j \int_0^1 g''(x_j)^2 \mathrm{d}x_j \qquad (9)$$

where $\{z\}$ is the data set obtained running the simulation model N times and λ_j are the smoothing hyper-parameters to be estimated. This minimization problem has a unique solution. The λ_j can be estimated in various ways; in particular, in the SDR recursive form, the Maximum Likelihood estimator is applied, whose nice properties are reviewed in Ratto and Pagano (2010).

The recursive least-squares implementation of Ratto et al. (2007) provides great flexibility in adapting to discontinuities, heavy nonlinearity and heteroskedastic error terms. Recently, Ratto and Pagano (2010) propose a unified approach for a more accurate and efficient identification of smoothing spline ANOVA models that combines the best of SDR and ACOSSO. The routines for the emulator of Ratto and Pagano (2010) are available at: http://eemc.jrc.ec.europa.eu/Software.htm. This approach provides direct estimation of ANOVA-based sensitivity indices up to order 2, i.e., including second-order terms in eq. (8), namely

$$g(\mathbf{x}) \simeq g_0 + \sum_{i=1}^{n} g_i(x_i) + \sum_{i<j} g_{i,j}(x_i, x_j) \qquad (10)$$

As far as eq. (10) is concern, Li et al. (2001) note that a cut in the functional ANOVA expansion is necessary to obtain quantitative information in real-life problems, because the size of the expansion is exponential in the number of factors — for $n = 20$, one obtains more than 1 million terms in the full functional ANOVA expansion [eq. (6)]. — Li et al. (2001) also observe that in many real-life studies the first or the second order approximations are enough to capture the model behavior [see also Ziehn and Tomlin (2009) and Borgonovo and Smith (2001), where a 393-factor-model describing a complex system is analyzed and no interactions of order higher than 2 emerge.] In fact, the terms $g_{i,j,k}(x_i, x_j, x_k)$ [we consider a third order term for simplicity] account for the residual interaction of order three, after the individual and second order effects have been removed

The metamodel building phase requires running the model on N training points (obtaining $\{z\}$). Once identified, estimated and parametrized the metamodel provides a direct, though approximated, algebraic expression of the input/output relationship, which incorporates direct estimation of sensitivity indices. Alternatively, these latter can be estimated more precisely running the nearly-costless metamodel on a set of points much larger than N.

All the above mentioned methods can be classified as regression techniques. Complementary to this class is the class of interpolation procedures. Interpolation procedures look at functions that pass through a set of data points spanning the whole domain of the input/output mapping. The approximation is obtained by estimating the p parameters of the interpolator (e.g. the coefficients of the polynomials) using p data points. Examples of interpolating metamodels are the Gaussian emulator [see for instance Oakley and O'Hagan (2004)] and the kriging procedure [Krige (1951),

Kleijnen (2008), and Kleijnen (2009)]. As opposed to smoothing-splines, which operate on truncated ANOVA terms, Gaussian emulators try to interpolate and predict the input/output mapping using a Gaussian kernel that has the same dimensionality n as the factor space. For large n, the number of hyper-parameters linked to the covariance structure of the Gaussian kernel grows rapidly, leading to problems with their identification.

3 Sampling plans and numerical estimation procedure

In this section, we discuss sampling designs at the basis of the estimation of δ_i. Approaches to the numerical estimation of δ_i can be found in Borgonovo and Tarantola (2008) and Liu and Homma (2009). However, the first work offering a systematic view to sampling strategies in distributional sensitivity analysis is Castaings et al. (2010).

Eq. (3) is composed of two integrals, one internal, namely $\int_{\Omega_Y} \left| f_Y(y) - f_{Y|X_i=x_i}(y) \right| \mathrm{d}y = s_i(x_i)$, and one external. The internal integral assesses the distance between the unconditional density of Y and the conditional density of Y when X_i is fixed at x_i. The external integral is the average of $s_i(x_i)$ and can be seen as $\int_{\Omega_Y} f_{X_i}(x_i) s_i(x_i) \mathrm{d}x_i$, where $f_{X_i}(x_i)$ is the marginal density of X_i. The ingredients needed to compute $s_i(x_i)$ are $f_Y(y)$ and $f_{Y|X_i=x_i}(y)$. To obtain them numerically, one proceeds as follows. First step is the generation of an unconditional sample $\widehat{\mathbf{X}}$ of size N. By construction, $\widehat{\mathbf{X}}$ is a matrix of size $N \times n$, where N is the number of Monte Carlo runs and n the number of model inputs. $\widehat{\mathbf{X}}$ is a set of possible realizations of the random vector $\mathbf{X}$. The model is then evaluated at $\widehat{\mathbf{X}}$ with N model runs. One obtains a corresponding vector of model output of size N, which we denote by $\widehat{\mathbf{Y}}$. By $\widehat{\mathbf{Y}}$ the unconditional model output density is estimated via kernel-density [Parzen (1962)].

The second ingredient, namely the conditional density, can be found by forming a conditional sample $\widehat{\mathbf{X}_i^j}$ utilizing elements of $\widehat{\mathbf{X}}$. Consider a generic factor, X_i. The element $\widehat{x}_i^j$ of $\widehat{\mathbf{X}}$ represents the value assumed by X_i in the jth Monte Carlo run. Substituting $\widehat{x}_i^j$ in the ith column of $\widehat{\mathbf{X}}$ is equivalent to conditioning on $X_i = \widehat{x}_i^j$. At this stage, if the factors are dependent one re-samples the values of the remaining model inputs given $X_i = \widehat{x}_i^j$, and obtains the conditional sample $\widehat{\mathbf{X}_i^j}$. If the factors are independent resampling is not necessary [Borgonovo et al. (2011)]. By evaluating the model at $\widehat{\mathbf{X}_i^j}$ one obtains the conditional model output vector $\widehat{\mathbf{Y}}_{X_i=\widehat{x}_i^j}$. By this vector, the density $f_{Y|X_i=\widehat{x}_i^j}(y)$ is estimated via kernel density [Silverman (1986)]. Once both the conditional and unconditional densities of Y are obtained, one estimates $s_i(\widehat{x}_i^j)$, by eq. (4). The procedure can be repeated for all columns of $\widehat{\mathbf{X}}$, leading to nN^2 model runs (N blocks of size N for X_i, $i = 1, \cdots, n$).

Using the above procedure (brute force approach), the number of points over a conditional value and the number of conditional values (for the external integral) explored are both fixed at N. Overall, the cost is $(N + nN^2)$ model runs. However, given that the external integral is one-dimensional, a quadrature formula can be applied. Castaings et al. (2010) have shown that this substituted column sampling plan combined with a Gauss-Legendre quadrature method [Davis and Rabinowitz (1984)] is efficient and reliable. In this contribution, the authors also show that using

Table 1: Steps for the estimation of the delta-importance measure via emulators (steps which are specific to the use of meta-modelling are marked by a star)

Nr.	Step
1	Generation of the base sample of size N (denoted $\widehat{\mathbf{X}}$)
2	Evaluation of original model output at $\widehat{\mathbf{X}}$ (denoted $\widehat{\mathbf{Y}}$)
3*	Calibration of metamodel using $\left(\widehat{\mathbf{X}}, \widehat{\mathbf{Y}}\right)$
4*	Evaluation of metamodel output at $\widetilde{\mathbf{X}}$ (denoted $\widetilde{\mathbf{Y}}$)
5*	Estimation of the metamodel unconditional model output distribution (using $\widetilde{\mathbf{Y}}$)
6	Estimation of quadrature points $\widehat{\mathbf{Q}}$ for all X_i $(i = 1, \cdots, n)$
7	Generation of conditional samples (grouped in $\widetilde{X}_c$) by column substitution (using $\widetilde{\mathbf{X}}$ and $\widehat{\mathbf{Q}}$)
8*	Evaluation of metamodel output at $\widetilde{\mathbf{X}}_c$ (model output denoted $\widetilde{\mathbf{Y}_c}$)
9*	Estimation of the metamodel conditional model output distributions (using $\widetilde{\mathbf{Y}_c}$)
10	Estimation of $s_i(x_i)$ and δ_i $(i = 1, \cdots, n)$

permuted column sampling plans, performances are comparable (or slightly better) for ranking but produce biased estimates for the moment-independent sensitivity measures. If the number of points used for the quadrature is denoted by N_{ext}, N_{ext} blocks of size N are necessary for the estimation of conditional densities for each X_i $(i = 1, \cdots, n)$. Therefore, the overall number of model runs is reduced to $N(1 + nN_{ext})$ with $N_{ext} < 10$.

Using real models with a long computational time (T), the overall time required for estimating the sensitivity measures is $T \cdot [N(1 + nN_{ext})]$. This figure can still be considerably high and make an accurate estimation of δ_i impractical. However, if the model is replaced by the emulator, then the computational time can be notably reduced. To appreciate the reduction, it is first necessary to introduce a systematic procedure for the estimation of δ via emulator. The steps are reported in Table 1.

The first step consists in generating the base sample $\widehat{\mathbf{X}}$. The second step is the evaluation of the model output at $\widehat{\mathbf{X}}$. The model output vector $\widehat{Y}$ is obtained. In this third step, the metamodel model is calibrated on $\widehat{\mathbf{X}} \rightarrow \widehat{\mathbf{Y}}$. This metamodel is subsequently evaluated at a second sample $\widetilde{X}$. Because the emulator running time is negligible, the size of $\widetilde{\mathbf{X}}$ can be arbitrarily large. One can then accurately estimate the unconditional density (steps 4 and 5). In steps 6 and 7, the conditional samples are generated by column substitution using quadrature points. Then the metamodel is evaluated on conditional samples in order to estimate the required conditional probability density functions (steps 8 and 9). The final step consists in using all those ingredients for the estimation of $s_i(x_i)$ and δ_i.

Let us now denote by τ the time required for a metamodel evaluation. Then, the overall time for implementing the steps in Table 1 is given by:

$$T^{\delta}_{emul} = N \cdot T + nN_{ext}N \cdot \tau \tag{11}$$

This time needs to be compared against the computational time required for direct estimation of

δ_i with the original model, which equals

$$T_{direct}^{\delta} = N \cdot T + nN_{ext}N \cdot T. \tag{12}$$

Therefore, the gain (or reduction) is

$$G = \frac{N \cdot T + nN_{ext}N \cdot T}{N \cdot T + nN_{ext}N \cdot \tau} \tag{13}$$

which can be notable if $\tau << T$. In particular, if $\tau << \dfrac{T}{nN_{ext}}$ (i.e., if the cost of the metamodel is much lower than the model's one), then $N \cdot T >> nN_{ext}N \cdot \tau$. In this case, eq. (13) reduces to $G \simeq 1 + nN_{ext}$. Thus, $(1 + nN_{ext})$ represents an upper bound for the potential gain. Because the same sample N is used for training the metamodel and obtaining the unconditional distribution of Y, then the cost of the analysis is potentially reduced from $N(1 + nN_{ext})$ to N model runs.

In the next section, we address the application of the procedure in Table 1 to three analytical test cases.

4 Analytical test cases

In this section, the performance of metamodelling for moment-independent sensitivity analysis is assessed for a set of analytical test cases. These models allow us to explore a variety of structures and factor distributions.

For the first and third case studies moment-independent sensitivity measures can be computed analytically using the approach proposed by Borgonovo et al. (2011). The analytical values represent a direct benchmark for the comparison of results obtained numerically. Their knowledge allows us to utilize the root mean square error (RMSE) as a measure of convergence. The RMSE at sample size N obtained by r replicates of the calculations is given by:

$$RMSE_i(N) = \sqrt{\frac{\sum_{l=1}^{r}[\widehat{\delta}_{i,l}(N) - \delta_i]^2}{r}} \tag{14}$$

where $\widehat{\delta}_{i,l}(N)$ is the estimate of δ_i at replicate l. If $RMSE(N)$ tends to 0 as N tends to infinity, the estimation strategy is correct. Conversely, some error (or bias) is present.

In all test cases of this section, the steps in Table 1 are followed. The Improved Substituted Column Sampling plan [Castaings et al. (2010)] is utilized for estimating δ_i, with $N_{ext} = 4$. An SDR second-order metamodel is constructed based on the approach by Ratto and Pagano (2010) using N evaluations of the original model. The metamodel is then employed in the estimation of δ_i in further $nN_{ext}N$ metamodel runs. For each test case, results are compared to the estimates obtained using the original model in all $(N + nN_{ext}N)$ runs.

For the first test case, the model is

$$Y = \frac{X_1}{X_1 + X_2} \tag{15}$$

with $X_1 \sim \Gamma(\alpha, \theta)$ and $X_2 \sim \Gamma(\beta, \theta)$, independently distributed. By a classical result of statistics [see Kotz et al. (2000)], Y is Beta distributed with density

$$f_Y(y) = \frac{y^{\alpha-1}(1-y)^{\beta-1}}{\int_0^1 s^{\alpha-1}(1-s)^{\beta-1} \mathrm{d}s} \tag{16}$$

The conditional distributions given X_1 and X_2 are given by [Borgonovo et al. (2011)]

$$f_{Y|X_1=x_1}(y) = \frac{(1-y)^{\beta-1}}{y^{\beta+1}} x_1^\beta \frac{e^{-\frac{x_1}{\theta}(\frac{1-y}{y})}}{\theta^\beta \Gamma(\beta)} \text{ and } f_{Y|X_2=x_2}(y) = \frac{(y)^{\alpha-1}}{(y-1)^{\beta+1}} x_1^\beta \frac{e^{-\frac{x_2}{\theta}(\frac{1-y}{y})}}{\theta^\beta \Gamma(\beta)}, \text{ respectively.} \tag{17}$$

Then, by eq. (17)

$$s_1(x_1) = \int_0^1 \left| \frac{y^{\alpha-1}(1-y)^{\beta-1}}{\int_0^1 s^{\alpha-1}(1-s)^{\beta-1}\mathrm{d}s} - \frac{(1-y)^{\beta-1}}{y^{\beta+1}} x_1^\beta \frac{e^{-\frac{x_1}{\theta}(\frac{1-y}{y})}}{\theta^\beta \Gamma(\beta)} \right| \mathrm{d}y \tag{18}$$

δ_1 follows by taking the expectation of $s_1(x_1)$ over the marginal density of X_1. One obtains

$$\delta_1 = \int_0^\infty x_1^{\alpha-1} \frac{e^{-\frac{x_1}{\theta}}}{\theta^\alpha \Gamma(\alpha)} \int_0^1 \left| \frac{y^{\alpha-1}(1-y)^{\beta-1}}{\int_0^1 s^{\alpha-1}(1-s)^{\beta-1}\mathrm{d}s} - \frac{(1-y)^{\beta-1}}{y^{\beta+1}} x_1^\beta \frac{e^{-\frac{x_1}{\theta}(\frac{1-y}{y})}}{\theta^\beta \Gamma(\beta)} \right| \mathrm{d}y \tag{19}$$

A similar procedure allows one to find δ_2. In Borgonovo et al. (2011), it is shown that $\delta_1 = \delta_2$ if $\alpha = \beta$. In fact, the marginal densities of X_1 and X_2 and the conditional model output densities given X_1 or X_2 become equal if $\alpha = \beta$. Thus, even if the model does not depend on them in the same form, if $\alpha = \beta$ the two factors are equally important (also using variance-based importance measures). Given $\alpha = \beta = 3$ and $\theta = 1$, one obtains $\delta_1 = \delta_2 = 0.315$.

Figure 3 reports the convergence results obtained using the steps described in Table 1.

Figure 3 displays the RMSE in the estimation of δ using the original model and the emulator (square and triangles, respectively.) In both cases, the average root mean square error (average across 20 replicates) declines rapidly with the number of model runs. The advantage of using model emulation leads to a gain G [eq. (13)] around 8 or 9. The accuracy of the convergence shows that the model structure is well captured by the emulator. We note that, while the model in eq. (15) is structurally non-additive, the model response is additive. In fact, in Borgonovo et al. (2011), it is shown that the sum of the first order variance-based indices covers around 99% of the model output variance. Thus, interactions play a minor role in eq. (15).

We then investigate the emulator performance further by considering a numerical test case,
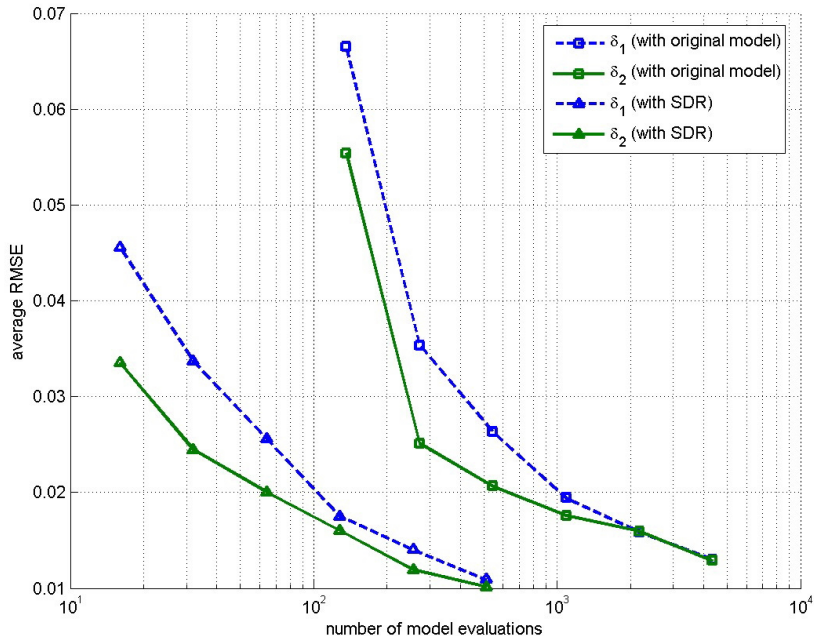
Figure 3: RMSE ($r = 20$) as a function of N for the non-additive model with Gamma inputs in eq. (15). Emulator (triangles), original model (squares).

widely utilized in SA, namely the Ishigami function:

$$Y = \sin(X_1) + 7\sin(X_2)^2 + 0.1X_3^4 \sin(X_1) \tag{20}$$

Each factor is uniformly distributed in $[-\pi, \pi]$. This test case is well known for its non-linear and non-monotonic behavior but also for the presence of a strong interaction of order 2 between X_3 and X_1. For this test case, an explicit expression of the distribution-based importance measures is not available. We therefore cannot base convergence analysis on the RMSE as defined in eq. (14). Instead, the average estimates (across the 20 replicates) for δ_i ($i = 1, 2, 3$) as produced by the direct utilization of the model or by the emulator are considered.

Figure 4 shows the convergence of the estimation via emulator (triangles) and with direct estimation (squares). The result shows that the emulator proivides an accurate estimation of δ. In turn, this implies that the model structure is correctly reproduced by the emulator. In fact, we are in the presence of an interaction of order 2.

As observed in Section 2.2, a cut of the functional ANOVA expansion at order 2 is often sufficient for accurately reproducing the behavior of realistic systems [Li et al. (2001), Ziehn and Tomlin (2009); see also a recent study by Borgonovo and Smith (2011)]. However, let us challenge the emulator performance by considering a model with a strong interaction of order 3. In particular,

14
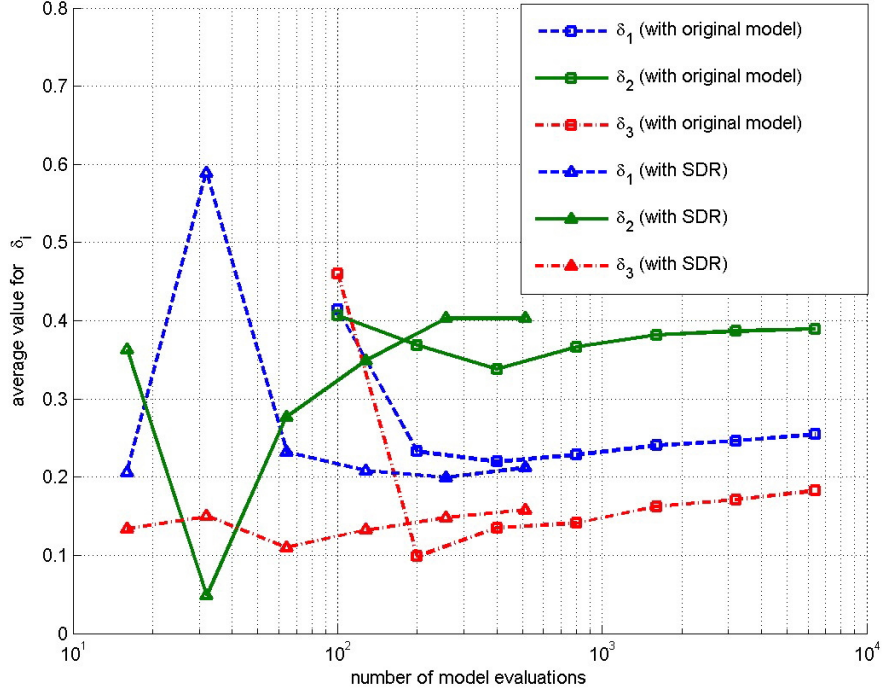
Figure 4: Average value of δ_i ($i = 1, 2, 3$), average across 20 replicates) for Ishigami function in eq. (20). Emulator (triangles), original model (squares).

let

$$Y = \prod_{i=1}^{3} X_i^{a_i} \qquad (21)$$

with $a_i \in \mathbb{R} - \{0\}$, with X_i ($i = 1, 2, 3$) lognormally independently distributed factors, with density

$$f_\mathbf{x}(\mathbf{x}; \boldsymbol{\eta}, \boldsymbol{\xi}) = \prod_{i=1}^{n} f_i(x_i; \eta_i, \xi_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\xi_i x_i} e^{\dfrac{-1}{2}\left[\dfrac{\ln(x_i) - \eta_i}{\xi_i}\right]^2}. \qquad (22)$$

By the model structure and the choice of the factor distributions, it is possible to obtain the importance δ_i ($i = 1, 2, 3$) analytically [see also Proposition 5 in Borgonovo et al. (2011)]:

$$\delta_i =$$
$$= \mathbb{E}_{X_i}[LN(e^{y_1}; \eta_Y, \xi_Y^2) - LN(e^{y_2}; \eta_Y, \xi_Y^2) + LN(e^{y_2}; \eta_{Y|x_i}, \xi_{Y|x_i}^2) - LN(e^{y_1}; \eta_{Y|x_i}, \xi_{Y|x_i}^2)] \qquad (23)$$
$$= \mathbb{E}_{X_i}[N(y_1; \eta_Y, \xi_Y^2) - N(y_2; \eta_Y, \xi_Y^2) + N(y_2; \eta_{Y|x_i}, \xi_{Y|x_i}^2) - N(y_1; \eta_{Y|x_i}, \xi_{Y|x_i}^2)]$$

15

where $LN(\cdot)$ denotes the lognormal distribution,

$$\xi_Y^2 = \sum_{i=1}^{n} a_i^2 \xi_i^2$$

$$\xi_{Y|x_i}^2 = \sum_{\substack{s=1 \\ s \neq i}}^{n} a_s^2 \xi_s^2 \tag{24}$$

$$\eta_Y = \mathbf{a}^T \boldsymbol{\eta}$$

$$\eta_{Y|x_i} = \eta_Y - a_i \eta_i + a_i \ln x_i$$

and where

$$y_{1,2} = \frac{1}{\xi_Y^2 - \xi_{Y|x_i}^2} \left(\xi_Y^2 \eta_Y - \xi_{Y|x_i}^2 \eta_{Y|x_i} \pm \sqrt{\xi_Y^2 \xi_{Y|x_i}^2 \left[(a_i x_i)^2 + (\xi_Y^2 - \xi_{Y|x_i}^2) \ln(\frac{\xi_Y^2}{\xi_{Y|x_i}^2}) \right]} \right) \tag{25}$$

Eq. (23) is readily implemented via a numerical software as Mathcad (used by the authors) or Matlab. Let us assign $a_1 = a_2 = a_3 = 1$, $\eta_1 = \eta_2 = \eta_3 = 1$ and $\xi_1 = 4, \xi_2 = 2$, and $\xi_3 = 1$. The output probability distribution is lognormal, with parameters $\eta_Y = 3$ and $\xi_Y^2 = 21$. This distribution is highly skewed. By eq. (23) one obtains the values of the importance measures: $\{\delta_1, \delta_2, \delta_3\} = \{0.47236, 0.15567, 0.07166\}$.

These values allow us to evaluate the performance of the emulator. Numerical results are plotted in Figure 5.

Figure 5 displays the RMSE as a function of the number of models evaluation. Figure 5 shows that convergence is obtained using the original model, while the estimation carried using the SDR metamodel fails to converge. In fact, this test case is unfavorable to the SDR metamodel of Ratto and Pagano (2010), because the output variability is driven by a third-order interaction, while the metamodel is designed for interactions up to order 2.

To explore this issue further, we construct a kriging emulator using the DACE Matlab toolbox [Lophaven et al. (2002)]. Results for the DACE emulator are reported in Figure 5 (diamonds). As one notes, also the DACE-based estimates fail to converge.

We recall that an accurate estimation of δ_i implies that the emulator correctly reproduces both the conditional and unconditional distributions of the model output. Conversely, an imprecise estimation of δ_i means that some structural feature of the original model is not captured by the emulator. This aspect is encountered here, because neither SDR nor DACE are able to emulate the behavior of a model exclusively consisting of a third-order interaction.

However, performance can be improved using the following strategy. In this specific case, the mapping to be emulated [eq. (21)] is a multiplicative function that can be transformed into an additive linear model by applying a logarithmic transformation during the construction of the metamodel. We recall that such a transformation leaves the estimates of δ_i unchanged, by the scale invariance property. We then follow the steps in Table 1, applying a log-transformation to the original model output when building the metamodel. The obtained results are plotted in Figure 6.
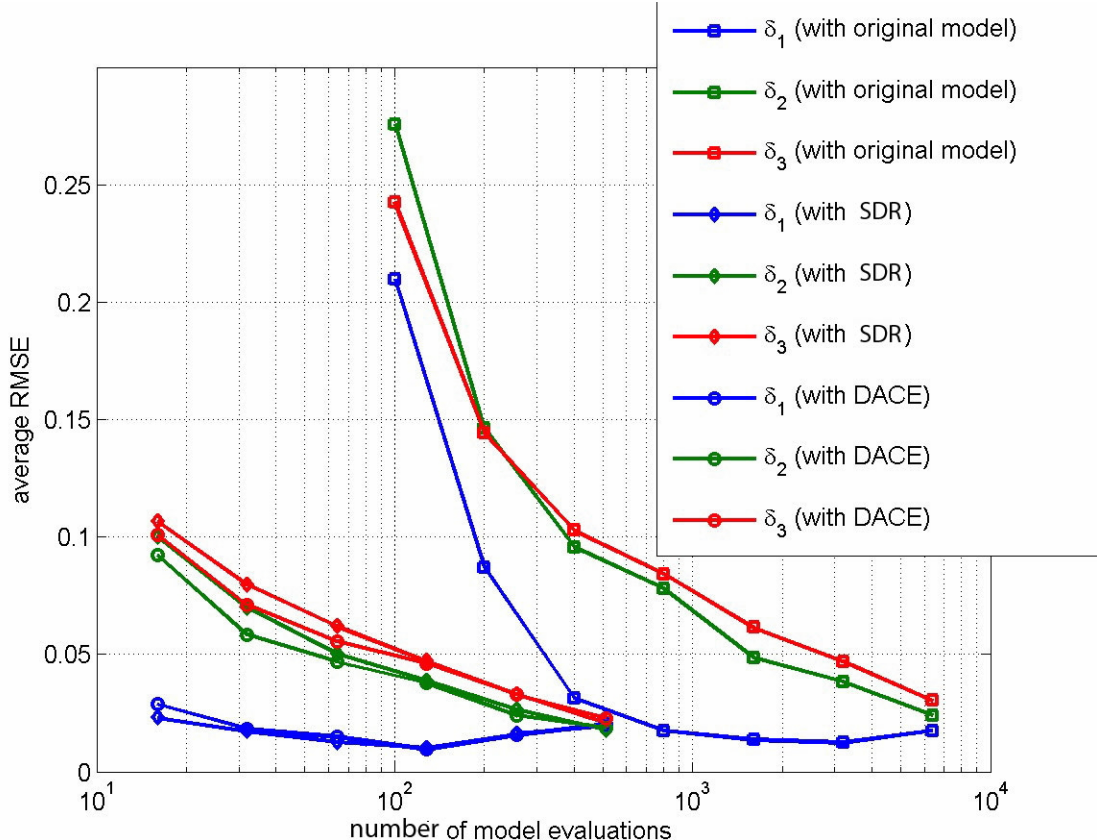
16
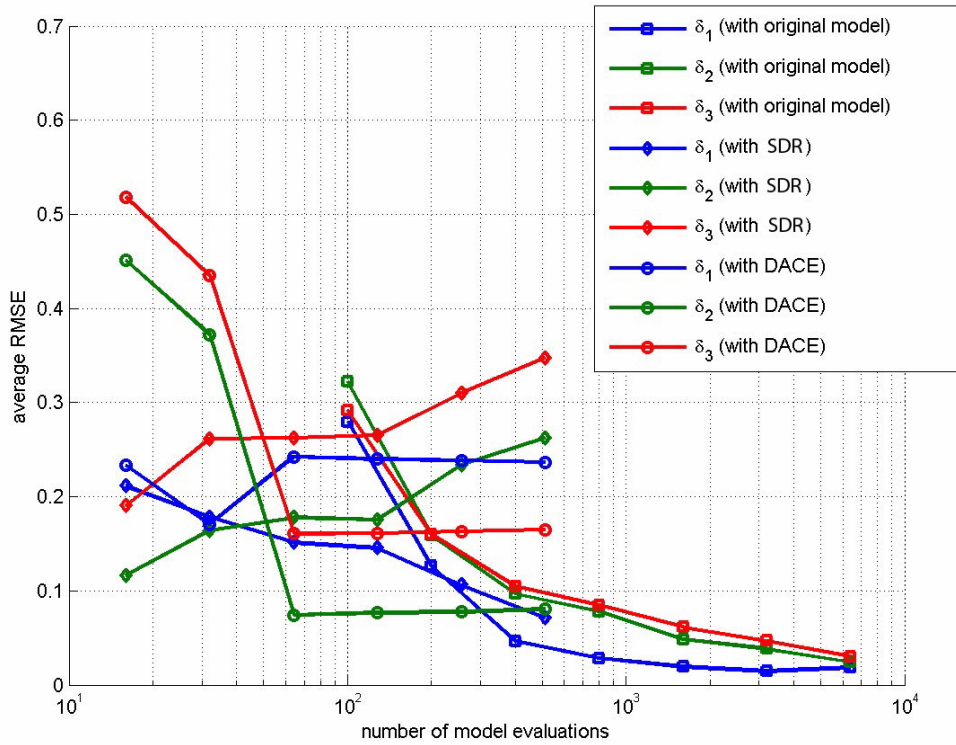
Figure 5: RMSE (average across 20 replicates) as a function of N for the multiplicative model with Lognormal input variables, and logarithmic transformation applied in moment independent sensitivity analysis [for the model of eq. 21]. SDR Emulator (Triangles), DACE Emulator (Diamonds), original model (Squares).

Figure 6 shows that the RMSE is tending to zero with both the DACE and SDR metamodels, as well as with the direct model calculation. Furthermore, the use of the metamodel produces a gain of around 1 order of magnitude in terms of the number of model evaluations (N) necessary to achieve a given accuracy. $[(1 + nN_{ext}) = 13$ represents an upper bound for G (eq.(13))]. As a reference, consider the estimation of δ_3. The combination model+metamodel makes it necessary to evaluate the model around $N = 80$ times to lower the RMSE below 0.005. To obtain the same accuracy, one needs around $N = 10^3$ direct model evaluations.

The three test cases of this section have allowed us to understand how different strategies can be used for improving the SDR emulator performance for estimating moment-independent sensitivity measures. We are going to apply these insights in the context of an environmental case study known in the literature for its structural complexity.

5 Application to an Environmental Case Study: LevelE

In this section, we apply the procedure of Table 1 and the insights obtained in Section 4, to the study of an environmental model, namely LevelE [OECD (1989), OECD (1993)].

17

Figure 6: RMSE across 20 replicates as a function of N for multiplicative model in eq. (21), with lognormal input variables, when a logarithmic transformation is applied in the construction of the metamodel. Emulator (triangles), direct model computation (squares).

LevelE simulates the radiological dose released from a nuclear waste disposal site to humans. The dose is due to the underground migration of radionuclides. LevelE has been widely utilized in the literature. We recall its utilization as a benchmark for Monte Carlo calculations in [OECD (1989), OECD (1993)], for variance-based techniques in Saltelli and Tarantola (2002), for emulators in Ratto et al. (2007) and recently, for moment-independent methods, in Castaings et al. (2010). While we refer to OECD (1989) for a detailed description of the model, a succinct illustration is proposed here. The repository is represented as a point source and the one-dimensional dispersion is tracked over geological time scales (up to 10^7 years). The model describes the transport of iodine (^{129}I), neptunium, uranium and thorium ($^{237}Np \rightarrow ^{233}U \rightarrow ^{229}Th$) through two geosphere layers characterized by specific hydro-geological properties. The governing equations account for radioactive decay, dispersion, advection and chemical reaction between the migrating nuclides and the porous medium. Epistemic uncertainty is driven by twelve uncertain model inputs whose probability distributions were assigned on the basis of expert judgment (see Table 2 and OECD (1993)).

Table 2: Model Inputs list for the LeveE Model

Notation	Definition	Distribution	Range	Units
T	Containment time	Uniform	$[100, 1000]$	yr
k_I	Leach rate for Iodine	Log-uniform	$[10^{-3}, 10^{-2}]$	$mols/yr$
k_c	Leach rate for Np chain	Log-uniform	$[10^{-6}, 10^{-5}]$	$mols/yr$
$V^{(1)}$	Water speed in geosphere's layer 1	Log-uniform	$[10^{-3}, 10^{-1}]$	m/yr
$l^{(1)}$	Length of geosphere's layer 1	Uniform	$[100, 500]$	m
$R^{(1)}$	Retention factor for I (first layer)	Uniform	$[1, 5]$	$-$
$R_C^{(1)}$	Retention coeff. for Np chain layer 1	Uniform	$[3, 30]$	$-$
$V^{(2)}$	Water speed in geosphere's layer 2	Log-uniform	$[10^{-2}, 10^{-1}]$	m/yr
$l^{(2)}$	Length of geosphere's layer 2	Uniform	$[50, 200]$	m
$R^{(2)}$	Retention factor for I (layer 2)	Uniform	$[1, 5]$	$-$
$R_C^{(2)}$	Retention coeff. for Np chain layer 2	Uniform	$[3, 30]$	$-$
W	Stream flow rate	Log-uniform	$[10^5, 10^7]$	m^2/yr

Two output of this model are analyzed in the literature. The maximum radiological dose simulated over the time period up to 10^7 years and the radiological dose at given times. For the maximum dose, the unconditional and conditional density functions have been previously reported in Figure 1. One notes that the unconditional distribution is highly skewed ($\gamma = 7.7841$).

In order to compare the estimates provided by the moment-independent sensitivity analysis calculation method using the original model or the SDR emulator, the experiments carried out for the analytical test cases (Section 4) are repeated for LevelE. Given the skewness of the maximum dose distribution, the results obtained in Section 4 suggest application of a logarithmic transformation during the construction of the metamodel.[4]

The Steps in Table 1 are applied to LevelE and produce the results in Figure 7.

The estimates and convergence paths are quite similar for the emulator and the original model (dotted line and continuous lines in Figure 7). However, by looking at Figure 7, one notes that stable factor ranking is identified much earlier using the emulator. For instance, in order to reverse the ranking between factors $V^{(1)}$ and W, the number of model evaluations is reduced by approximately a factor 50. Also, the ranking of the factors with low and moderate importance is correctly identified by the emulator.

As far as the computational gain [eq. (13)] is concerned, because $n = 12$, G is potentially at $(1 + 12 \cdot N_{ext})$ model runs, in this case. Setting $N_{ext} = 4$, (minimum suggested value Castaings et al. (2010)), the potential gain is 49. At $N = 1000$ and with a model run of 2 secs, one needs 98000 secs (around 28hours) for the analysis. By replacing the model via the emulator, one has the following computational time: 2000secs for building the emulator; 1000secs for performing the additional 120000 evaluations. Thus, the total time required by the model emulation strategy is 50mins, as opposed to the initial 27 hours. From a more general standpoint, the gain becomes even

---

[4]We performed the analysis both with and without the log-transformation. The log-transformation lead to a notable improvement of the numerical performance of the emulator.
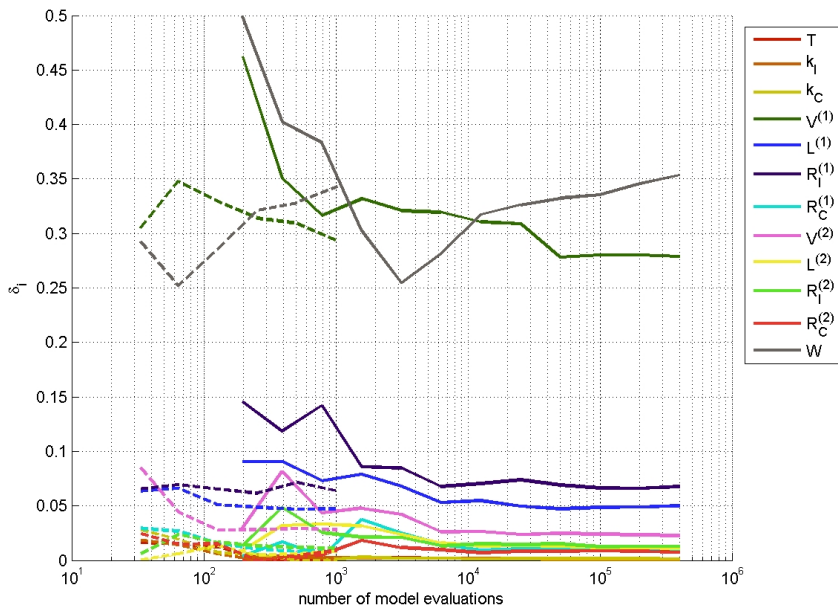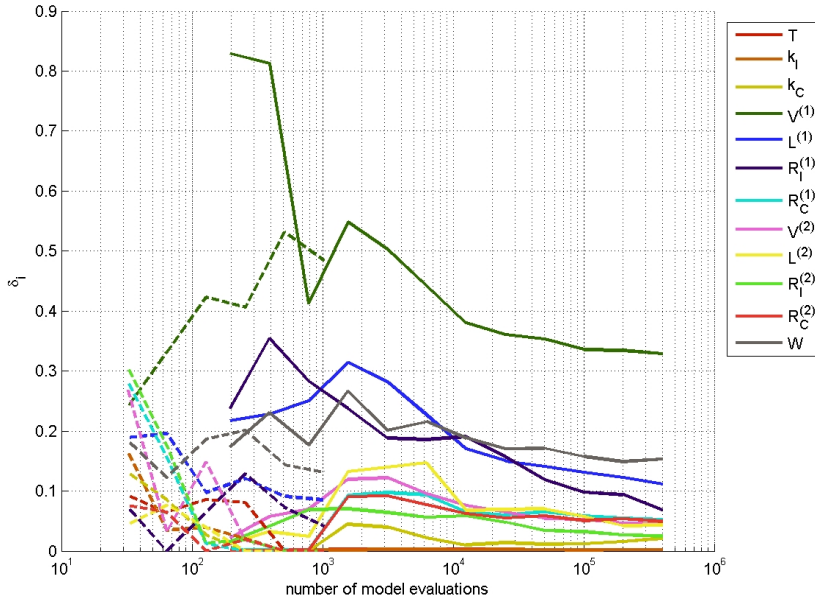
Figure 7: δ_i, $i = 1, 2, .., 12$, for the maximum dose output of the LevelE model. Dotted lines (left) represent results obtained with the metamodel. Continuous lines (right) represent results obtained by running the original model.

more evident for models with larger computational times or in the presence of a larger number of factors.

Let us come to the insights one obtains by Figure 7 from an environmental viewpoint. The maximum dose is of interest for environmental safety [OECD (1989)]. The most important factors are W and $V^{(1)}$, namely, stream flow rate and water speed in geosphere layer 1. Thus, these are the factors on which we should focus our efforts to eliminate the main sources of shifts in the distribution of the maximum dose. Conversely, low ranked factors, like k_C and k_I, can be fixed at any value within their range of uncertainty without altering the analyst's state of knowledge on the maximum dose.

In the literature, as mentioned, not only the maximum dose has been studied as an output of LevelE, but also the time evolution of the released dose. Of particular interest is the study of the dose simulated at $t = 2 \cdot 10^5$ years, during the transition zone between fast and slow dynamics (approximately between 10^5 and 10^6 years). Previous studies on LevelE [Saltelli et al. (1993), Saltelli and Sobol' (1995), Saltelli and Tarantola (2002)] have shown that during this specific time range, the mapping is non-monotonic for some model inputs (e.g. $V^{(1)}$) and the variance is mainly driven by the contribution of interactions [Homma and Saltelli (1996)]. Although the situation is less favorable than in the case of the maximum dose, the analysis of Figure 8 reveals that the emulator performs well in the estimation of moment-independent importance measures, even for this very complex mapping. The values of the δ−importance obtained with around $N = 10^3$ model

runs lead to the same ranking obtained through $N = 4 \cdot 10^4$ original model evaluations.

Figure 8: δ_i, $i = 1, 2, .., 12$, for the dose at time $2 \cdot 10^5$, output of the LevelE model. Dotted lines (left) represent results obtained with the metamodel. Continuous lines (right) represent results obtained by running the original model.

Let us now come to the identification of the most relevant factors. By comparing Figure 8 to Figure 7, one observes that $V^{(1)}$ is now the factor that influences the model output distribution the most. Moreover, $W, L^{(1)}$ and $R_I^{(1)}$ have an intermediate influence on the dose at $2 \cdot 10^5$ years, while the other factors have minor influence.

6 Conclusions and Future Research

In this work, we have analyzed the joint application of emulators and moment-independent sensitivity analysis methods in the context of uncertainty analysis of environmental models.

Distribution-based (or moment-independent) SA techniques convey the importance of a factor without relying on a particular moment of the output distribution. The estimation of moment-independent importance measures is, however, a challenging task, because of the double integration their definition implies. Recent works have proposed and evaluated sampling designs which reduce the computational burden. However, the number of model runs required for an accurate estimation of the importance measures can still be too high for computationally complex models.

We have then argued that, because emulators have the ability of drastically reducing computational burden, they can enable the estimation of moment-independent importance measures in computationally intensive models. In particular, we have made use of the combination of the SDR emulator developed by Ratto et al. (2007) and the improved substituted columns sampling plans

developed in Castaings et al. (2010). We have set forth an eight-step procedure, which makes the approach systematic. This has allowed us to compute the gain (in terms of reduction in computational time; eq. (13)) that the metamodel can generate. The gain can potentially be of the order of $1 + nN_{ext}$, if the computational time for running the emulator is negligible with respect to that of the original model. We have then tested the performance of the combination emulator-moment-independent importance measures for three test cases. The first two test cases (double gamma and Ishigami) consisted of models with no interactions and interactions of order 2 respectively. Numerical experiments confirmed the ability of the SDR emulator in reproducing the model structure and, therefore, leading to an accurate estimation of the density-based importance measures. The third case study is a 3-factor, completely interactive model, with skewed distributions, for which the analytical expression of δ_i is available. Results have shown a poor performance of the SDR emulator in reproducing the model behavior related to the cut at order 2 of the functional ANOVA expansion. However, we have seen that applying a logarithmic transformation to the model output when building the emulator allows one to obtain convergence also for this case. This strategy has then been replicated in the estimation of moment-independent importance measures for the LevelE model replaced by the SDR emulator. Results confirm that the emulator considerably reduces the run time, while producing accurate estimates of the importance measures.

Results of our paper can then be interpreted as follows. Model emulation has the potential to achieving a significant reduction in the computational burden for estimating moment-independent sensitivity measures. This achievement is, however, subject to a careful calibration of the metamodel (this aspect is crucial to all metamodelling applications). At the same time, the ability (or inability) of the emulator to estimate density-based sensitivity measures can be seen as an additional validation (or invalidation) of the constructed metamodel. In this respect, this work paves the way for future research in using density-based sensitivity measures as a metamodel validation tool, to accompany traditional model emulation performance measures for case studies in which δ_i is analytically known. In fact, for a precise estimation of δ, the emulator needs to accurately reproduce the distribution of the model output, what, in turn, implies capturing all main structural features of the original model.

References

BAYARRI M. J. , BERGER J. AND STEINBERG D.M., 2009: "Special Issue on Computer Modeling," *Technometrics*, 51(4), p. 353-353.

BLATMAN G. AND SUDRET B., 2010: "Efficient computation of global sensitivity indices using sparse polynomials," *Reliability Engineering & System Safety*, 95, pp. 1216–1229.

BORGONOVO E., 2006: "Measuring uncertainty importance: investigation and comparison of alternative approaches," *Risk Analysis*, 2006, 26(5), pp. 1349-1362.

BORGONOVO E., 2007: "A new uncertainty importance measure," *Reliability Engineering and System Safety*, 92, pp. 771-784.

BORGONOVO E. AND TARANTOLA S., 2008: "Moment independent and variance-based sensitivity analysis with correlations: an application to the stability of a chemical reactor," *International Journal of Chemical Kinetics*, 2008, 40 (11), pp. 687-698.

BORGONOVO E. AND SMITH C.L., 2011: "A Study of Interactions in the Risk Assessment of Complex Engineering Systems: An Application to Space PSA," *Operations Research*, forthcoming.

BORGONOVO E., CASTAINGS W. AND TARANTOLA S., 2011: "Moment Independent Importance Measures: New Results and Analytical Test Cases," *Risk Analysis*, 31 (3), pp. 404-428.

BREIMAN L., FRIEDMAN J.H., OLSHEN R.A., STONE C.J.. 1984: "Classification and regression trees," Chapman&Hall, New York, NY, USA, ISBN 0-412-04841-8.

CACUCI D.G., 1981: "Sensitivity theory for nonlinear systems. I. Nonlinear functional analysis approach," *Journal Mathematical Physics*, 22(12), pp. 2794-2802.

CAMPOLONGO F. AND BRADDOCK R., 1999: "Sensitivity analysis of the IMAGE Greenhouse model," *Environmental Modelling and Software*, 14, pp. 275–282.

CAMPOLONGO F., CARIBONI J. AND SALTELLI A., 2007: "An effective screening design for sensitivity analysis of large models," *Environmental Modelling and Software*, 22, pp. 1509-1518.

CASTAINGS, W., DARTUS, D., LE DIMET, F.-X., AND SAULNIER, G.-M.: Sensitivity analysis and parameter estimation for distributed hydrological modeling: potential of variational methods, *Hydrological Earth System Science*, 13, pp. 503-517.

CASTAINGS W., BORGONOVO E., AND TARANTOLA S., 2010: "Sampling plans for the estimation of moment-independent importance measures," *Procedia - Social and Behavioral Sciences*, 2(6), pp. 7629-7630.

CHUN M-H., HAN S-J. AND TAK N-IL., 2000: "An uncertainty importance measure using a distance metric for the change in a cumulative distribution function," *Reliability Engineering and System Safety*, 70, pp. 313-321.

CONFALONIERI R., BELLOCCHI G., TARANTOLA S., ACUTIS M., DONATELLI M., GENOVESE G., 2010: "Sensitivity analysis of the rice model WARM in Europe: Exploring the effects of different locations, climates and methods of analysis on model sensitivity to crop parameters," *Environmental Modelling and Software*, 25, pp. 479–488.

CRYER S.A. AMD HAVENS P.L., 1999:"Application of ADIFOR for air pollution model sensitivity studies,"*Environmental Modelling and Software*, 14, pp. 613-624.

DAVIS P.J. AND RABINOWITZ P., 1984: "Methods of Numerical Integration," Academic Press, ISBN-0122063600.

ESTRADA V. AND DIAZ M.S., 2010: "Global sensitivity analysis in the development of first principle-based eutrophication models," *Environmental Modelling & Software*, doi:10.1016/j.envsoft.2010.06.009.

FRIEDMAN J.H., W. STUETZLE 1981: "Projection pursuit regression," *Journal of the American Statistical Association*, 76, (376), 817–23.

FRIEDMAN J.H., 1991, "Multivariate adaptive regression splines (with discussion)," *Annals of Statistics*, 19(1), pp. 1-67.

FRIEDMAN J.H., 2001: "Greedy function approximation: a gradient boosting machine," *Annals of Statistics*, 29(5), pp. 1189-232.

GLICK N., 1975: "Measurements of separation among probability densities or random variables," *The Canadian Journal of Statistics*, 3 (2), pp. 267-276.

GRIEWANK A., 2000: "Evaluating Derivatives, Principles and Techniques of Algorithmic Differentiation," Frontiers in Appl. Math, No 19, SIAM, Philadelphia (USA).

GU C., 2002: "Smoothing Spline ANOVA Models," Springer, Berlin (GER).

HASTIE T.J. AND R.J. TIBSHIRANI, 1992: "Generalized Additive Models," Chapman and Hall, New York (NY, USA), ISBN 0-412-34390-8.

HELTON J.C., 1993: "Uncertainty and sensitivity analyses techniques for use in performance assessment for radioactive waste disposal," *Reliability Engineering & System Safety*, 42, 327-367.

HE S., CARMICHAEL G.R., SANDU A., HOTCHKISS B. AND DAMIAN-IORDACHE V., 2000: "Application of ADIFOR for air pollution model sensitivity studies," *Environmental Modelling and Software*, 15, pp. 549–557.

HUANG C.-F. AND LITZENBERGER R.H., 1988: "Foundations for Financial Economics," Prentice Hall, Upper Saddle River, NJ, USA, ISBN 0-13-500653-8.

HOMMA T. AND SALTELLI A., 1996: "Importance Measures in Global Sensitivity Analysis of Nonlinear Models," *Reliability Engineering and System Safety*, 52, 1-17.

KLEIJNEN J., 2008: "Design and Analysis of Simulation Experiments," *Springer*, 216 pages.

KLEIJNEN J. 2009: "Kriging metamodeling in simulation: a review," *European Journal of Operational Research,* 192, no. 3, pp. 707-716

KOHLER P. AND WIRTZ K.W., 2002: "Linear understanding of a huge aquatic ecosystem model using a group-collecting sensitivity analysis," *Environmental Modelling and Software*, 17, pp. 613–625.

KOTZ S., BALAKRISHNAN N. AND JOHNSON N.L., 2000: "Continuous Multivariate Distributions," Volume 1, Models and Applications, 2nd Edition [Hardcover], Wiley&Sons NY, USA, ISBN-10: 0471183873.

KRIGE D.G., 1951: "A statistical approach to some basic mine valuation problems on the Witwatersrand," *Journal of the Chemical, Metallurgial and Mining Society of South Africa*, 52 (6), pp. 119-139.

IMAN R.L. AND HORA S.C., 1990: "A Robust Measure of Uncertainty Importance for Use in Fault Tree System Analysis," *Risk Analysis*, 10 (3), pp. 401-406.

LI G., WANG S.-W., ROSENTHAL C. AND RABITZ H., 2001: " High dimensional model representations generated from low dimensional data samples. I. mp-Cut-HDMR," *Journal of Mathematical Chemistry*, 30(1), pp. 1-30

LIU Q. AND HOMMA T., 2009: "A new computational method of a moment-independent importance measure," *Reliability Engineering and System Safety*, 2009, 94, pp. 1205–1211.

LOPHAVEN, S., NIELSEN, H., SONDERGAARD, J., 2002: "DACE: A Matlab kriging toolbox," Version 2.0. Technical Report IMM-TR-2002-12, Informatics and Mathematical Modelling, Technical University of Denmark (2002). http://www.immm.dtu.dk/hbn/dace.

MANACHE G. AND MELCHING C.S., 2008: "Identification of reliable regression- and correlation-based sensitivity measures for importance ranking of water-quality model parameters," *Environmental Modelling & Software*, 23, pp. 549-562.

NEWHAM L.T.H., NORTON J.P., PROSSER I.P., CROKE B.F.W. AND JAKEMAN A.J., 2003: "Sensitivity analysis for assessing the behaviour of a landscapebased sediment source and transport model," *Environmental Modelling and Software*, 18, 741–751.

NORTON J.P., 2008: "Algebraic sensitivity analysis of environmental models," *Environmental Modelling & Software*, 23, pp. 963-972.

OAKLEY J. AND O'HAGAN A., 2004: "Probabilistic sensitivity analysis of complex models: a Bayesian approach", *Journal of the Royal Statistical Society B*, 66, Part 3, pp. 751–769.

OECD, 1989: "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," technical report, OECD—NEA, Paris.

OECD, 1993: "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," technical report, OECD—NEA, Paris.

PAPPENBERGER F., BEVENA K.J., RATTO M. AND MATGEN P., 2006: "Multi-method global sensitivity analysis of flood inundation models", *Advances in Water Resources*, 31, pp. 1–14.

PARK C.K. AND AHN K.I., 1994: "A new approach for measuring uncertainty importance and distributional sensitivity in probabilistic safety assessment," *Reliability Engineering and System Safety*, 46, pp. 253-61.

PARZEN E., 1962: "On Estimation of a Probability Density Function and Mode," *The Annals of Mathematical Statistics*, 33 (3), pp.1065-1076.

RABITZ H., 1989: "System Analysis at the Molecular Scale," *Science*, 246, pp. 221-226.

RABITZ H. AND ALIS ÖF., 1999: "General foundations of high-dimensional model representations," *Journal of Mathematical Chemistry*, 25 (2-3), pp. 197–233.

RATTO M., PAGANO A. AND YOUNG P., 2007: "State Dependent Parameter metamodelling and sensitivity analysis," *Computer Physics Communications*, 177 (11), pp. 863-876.

RATTO M. AND PAGANO A. 2010: "Using recursive algorithms for the efficient identification of smoothing spline ANOVA models," *Advances in Statistical Analysis*, 94: 367-388.

RISBEY J., VAN DER SLUIJS J., KLOPROGGE P., RAVETZ J., FUNTOWICZ S. AND QUINTANA S.C., 2005: "Application of a checklist for quality assistance in environmental modelling to an energy model," *Environmental Modeling and Assessment*, 10 (1), pp. 63-79.

SACKS J., WELCH W.J., MITCHELL T.J. AND WYNN H.P., 1989: "Design and analysis of computer experiments," *Statistical science*, 4, pp. 409–423.

SALTELLI A. AND MARIVOET J., 1990: "Non-parametric Statistics in Sensitivity Analysis for Model Output: a Comparison of Selected Techniques," Reliability Engineering and System Safety, 28, 229-253.

SALTELLI A., RATTO M., ANDRES T., CAMPOLONGO F., CARIBONI J., GATELLI D., SAISANA M. AND TARANTOLA S., 2008: "Global Sensitivity Analysis: the Primer," John Wiley and Sons, Chichester, UK, ISBN 978-0-470-05997-S.

SALTELLI A., ANDRES T.H. AND HOMMA T., 1993: " Sensitivity analysis of model output : An investigation of new techniques, " *Computational Statistics & Data Analysis*, 15 (2), pp. 211 – 238.

SALTELLI A. AND SOBOL' I.M., 1995: "About the use of rank transformation in sensitivity analysis of model output," *Reliability Engineering & System Safety*, 50 (3), pp. 225–239, 1995.

SALTELLI A. AND TARANTOLA S., 2002: "On the Relative Importance of Input Factors in Mathematical Models: Safety Assessment for Nuclear Waste Disposal," *Journal of the American Statistical Association*, 97 (459), p. 702-709.

SALTELLI A. AND ANNONI P., 2010: "How to avoid a perfunctory sensitivity analysis," *Environmental Modeling and Software*, forthcoming, doi:10.1016/j.envsoft.2010.04.012.

SALTELLI A. AND D'HOMBRES B., 2010: "Sensitivity analysis didn't help. A practitioner's critique of the Stern review," *Global Environmental Change*, forthcoming, doi:10.1016/j.gloenvcha.2009.12.003.

SANTNER T. WILLIAMS B. AND NOTZ W., 2003: "The Design and Analysis of Computer Experiments," Springer Series in Statistics, Springer Verlag, New York (USA), ISBN 978-4419-2992-1.

SILVERMAN B.W., 1986: "Density Estimation For Statistics and Data Analysis," Monographs on Statistics and Applied Probability, London: Chapman and Hall, Uk.

SIMONOFF J.S., 1996: Smoothing Methods in Statistics, New York, Springer-Verlag

SOBOL', I.M., 1993: "Sensitivity estimates for nonlinear mathematical models," *Mathematical Modelling & Computational Experiments*, 1, 407–414.

STORLIE C.B. AND HELTON J., 2008: "Multiple predictor smoothing methods for sensitivity analysis: description of techniques," *Reliability Engineering and System Safety*, 93, 28-54.

Storlie C.B., Swiler L.P., Helton J.C., and Sallaberry J., 2009: "Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models," *Reliability Engineering & System Safety*, 94 (11), pp. 1735-1763.

Storlie C.B., Bondell H.D., Reich B.J. and Zhang, H.H., 2010: "Surface estimation, variable selection, and the nonparametric oracle property," *Statistica Sinica*, 1, 1-2.

Sudret B., 2008: "Global sensitivity analysis using polynomial chaos expansions," *Reliability Engineering & System Safety*, 93, pp. 964–979.

US EPA, 2009: "Guidance on the Development, Evaluation, and Application of Environmental Models," EPA/100/K-09/003, March 2009, www.epa.gov/crem.

Varella H., Gue M. and Buis S., 2010: "Global sensitivity analysis measures the quality of parameter estimation: The case of soil parameters and a crop model, *Environmental Modelling & Software*, 25, pp. 310–319.

Young P.C., 2001: "The identification and estimation of nonlinear stochastic systems," In: Mees Aiea (ed), Nonlinear Dynamics and Statistics, Birkhäuser, Boston, MA (USA).

Yu D., 2010: "Parallelization of a two-dimensional flood inundation model based on domain decomposition," *Environmental Modelling & Software*, doi:10.1016/j.envsoft.2010.03.003.

Ziehn T. and Tomlin A.S., 2010: "GUI–HDMR – A software tool for global sensitivity analysis of complex models," *Environmental Modelling & Software*, 24, pp. 775–785.