Invited Review

Sensitivity analysis: A review of recent advances

Emanuele Borgonovo [a,*], Elmar Plischke [b]

[a] *Department of Decision Sciences, Bocconi University, Milan, Italy*
[b] *Clausthal University of Technology, Clausthal-Zellerfeld, Germany*

ABSTRACT

The solution of several operations research problems requires the creation of a quantitative model. Sensitivity analysis is a crucial step in the model building and result communication process. Through sensitivity analysis we gain essential insights on model behavior, on its structure and on its response to changes in the model inputs. Several interrogations are possible and several sensitivity analysis methods have been developed, giving rise to a vast and growing literature. We present an overview of available methods, structuring them into local and global methods. For local methods, we discuss Tornado diagrams, one way sensitivity functions, differentiation-based methods and scenario decomposition through finite change sensitivity indices, providing a unified view of the associated sensitivity measures. We then analyze global sensitivity methods, first discussing screening methods such as sequential bifurcation and the Morris method. We then address variance-based, moment-independent and value of information-based sensitivity methods. We discuss their formalization in a common rationale and present recent results that permit the estimation of global sensitivity measures by post-processing the sample generated by a traditional Monte Carlo simulation. We then investigate in detail the methodological issues concerning the crucial step of correctly interpreting the results of a sensitivity analysis. A classical example is worked out to illustrate some of the approaches.

1. Introduction

The solution of several managerial problems involves the creation of dedicated decision support models (models, henceforth). In supporting the solution of decision analysis problem, the model is typically an influence diagram, or a decision tree (Bielza & Shenoy, 1999). In the risk assessment of complex operational systems, the model takes the form of a Bayesian network, of a binary decision diagram or of a fault tree. Convex, linear or dynamic programming support decision making in production, inventory management, logistics (Bertsimas & Sim, 2004; Tan & Hartman, 2011; Wendell, 1985). Business plans help analysts in determining the financial viability of joint ventures and investment projects (Baucells & Borgonovo, 2013). The perspective becomes even broader if one regards this exercise in the context of scientific modeling. In this case, a scientific model reproduces the behavior of a certain system or portion of the real world. The model might be based on a theory which explains the phenomenon at hand. An important account of the role of sensitivity analysis in the broad scientific literature is given in Rabitz (1989). In the operational literature history, an outstanding (and early) exam-

ple is represented by Ford W. Harris' economic order quantity model (Harris EOQ, henceforth) (Harris, 1990). Harris applies the scientific method to the problem of determining the optimal lot size in a production system, obtaining the well known expression of the total system cost

$$T = \frac{1}{240M}(CX + S) + \frac{S}{X} + C \qquad (1)$$

where S is the setup cost of an order, M is the number of units used per month (*the movement of the stock* (Harris, 1990, p. 947)), and C is the unit price of items in the stock, the number 240 is the interest charge per unit, per month, corresponding to a yearly interest rate of 10 percent (Harris, 1990, p. 948)). By manipulation, the well known economic order quantity (EOQ) is given by

$$X^* = \sqrt{\frac{240MS}{C}} \qquad (2)$$

which leads to the optimized cost

$$T^* = \left(\sqrt{\frac{S}{240M}} + \sqrt{C} \right)^2. \qquad (3)$$

We shall use Harris EOQ model throughout this work as an illustrative example. Harris EOQ has been generalized to include alternative

aspects of the problem, in an important and still active research stream of subsequent operations research literature (Cárdenas-Barrón, Chung, & Trevi no-Garza, 2014). Harris himself was, however, well aware that the model does not exhaust the decision process: *But in deciding on the best size of order, the man responsible should consider all the factors that are mentioned [...] Hence, using the formula as a check, is at least warranted* (Harris, 1990, p. 947).

Little (1970) in addressing the relationship between managers and models observes that, of course, it is not the task of the manager to build the model. However, Little notes that it is important for a manager to be able to correctly interrogate the model. In fact, it is through this questioning process that the manager becomes acquainted with the model behavior and confident in using the model for decision making. Conversely, using the words of Little (1970, p. B469), a manager "*wants to know why. A process starts of finding out what it was about the inputs that made the outputs come out as they did* ".

From a general, scientific perspective, sensitivity analysis is the set of methods that allow us to understand key insights of scientific codes: "*The judicious application of sensitivity analysis techniques appears to be the key ingredient needed to draw out the maximum capabilities of mathematical modeling*" (Rabitz, 1989, p. 221). The need to use sensitivity analysis methods emerges even more sharply in recent applications.

The increase in computing power allows analysts to create mathematical models of increased sophistication. Given the complexity of the codes, it is not possible for the analyst to grasp the response of the model output to variations in the model inputs based on the sole intuition. Fürbringer (1996) asks rhetorically: "*Sensitivity Analysis for Modelers: Would you go to an orthopedist who did not use X-Ray?* ", highlighting the need to look into the model black box. (Samuelson, 1941, p. 97) suggests: "*In order for the analysis to be useful it must provide information concerning the way in which our equilibrium quantities will change as a result of changes in the parameters*".

Aware of these issues, several institutions and agencies such as the (EPA, 2009, p. 19), the Florida Commission on Hurricane Loss Projection Methodology (Iman, Johnson, & Watson, 2005), the NRC (2009), health organizations as CADTH (2006), Cochrane (Higgins & Green, 2011) and NICE (2013) recommend the use of sensitivity and uncertainty analysis methods as part of best practices to ensure the overall quality of the modeling process.

The variety of problems that have been addressed through scientific modeling and the wide set of interrogations which are possible for these models has led, over the past years, to the development of several sensitivity analysis methods. Some methods are *quantitative and model free* (Saltelli, 2002a), some methods are specifically tailored to the type of mathematical problem implied by the model (e.g., the tolerance sensitivity approach for linear programming).

Here, it is worth to point the reader to previous reviews that either precede or complement our work. Earlier reviews of sensitivity methods in interdisciplinary contexts are offered also by Saltelli, Ratto, Tarantola, and Campolongo (2005) and Saltelli, Ratto, Tarantola, and Campolongo (2012). Aven and Nøkland (2010) review the use of global sensitivity methods in reliability applications. Filippi (2010), Tan and Hartman (2011), Kleijnen (2010), Borgonovo (2013) offer earlier reviews of sensitivity analysis in linear programming, discrete dynamic programming, simulation and decision-analysis respectively. A review on robustness analysis is offered in Rosenhead (2011), and on sensitivity analysis in dynamic programming in Gal and Greenberg (1997).

We highlight that a crucial step for a meaningful sensitivity analysis is the clear statement of the insights that we wish to obtain from the model. We then carry out a systematic investigation about the questions that have been most widely asked in the literature. We review old sensitivity settings and propose new ones that help to make crisp the link between sensitivity methods and sensitivity insights.

The remainder of this article is organized as follows. Section 2 discusses the sensitivity analysis jargon and notation. This section serves as a general reference for the remaining sections. Section 3 discusses local and derivative-based sensitivity methods. Section 4 offers an overview of different sensitivity strategies under a fully probabilistic setting. Section 5 discusses computational aspects and presents a common framework for several of the global sensitivity measures introduced in this work. Section 6 offers a view on the philosophy of sensitivity analysis. Section 7 provides a discussion on additional relevant aspects and uses of sensitivity analysis. Section 8 offers final considerations.

2. Premise on notation and terminology

The standard setup of sensitivity analysis is as follows. The model is regarded as a mapping between the model inputs and the model outputs. These model outputs are the quantities of interest for the decision-maker. Let us say that there are n model inputs and D model outputs. We denote them by $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$ and $\mathbf{y} = \{y_1, y_2, \ldots, y_D\}$. Formally, we write

$$\mathbf{y} = g(\mathbf{x}) \tag{4}$$

where $g : \mathcal{X} \to \mathcal{Y}$ with $\mathcal{X} \subseteq \mathbb{R}^n$ and $\mathcal{Y} \subseteq \mathbb{R}^D$. We will call $\mathcal{X}$ the model input space.

As for terminology, we find alternative names for $\mathbf{x}$ and $\mathbf{y}$ depending on the field of application. For instance, in Economics they are called exogenous and endogenous variables, respectively. There, comparative static is a widely used synonym for sensitivity analysis, after the seminal article of Samuelson (1941). In Engineering studies, $\mathbf{x}$ is often called the vector of parameters. In Statistics, the term factors or covariates might be used for $\mathbf{x}$. The model is the mapping that binds $\mathbf{x}$ and $\mathbf{y}$. To illustrate, we consider Harris EOQ in a sensitivity analysis framework. The model output, y, is the optimal order quantity (X^* in Harris EOQ), the model inputs are the units per month (x_1, M in Harris EOQ), the unit price of items in the stock (x_2, C in Harris EOQ), and the setup cost of an order (x_3, S in Harris EOQ). We write:

$$y = \sqrt{\frac{240 x_1 x_3}{x_2}}. \tag{5}$$

In this case, it is $m = 1$ and $n = 3$, because we have one model output and three model inputs. However, if we consider also the total order cost at the optimum (T^*) as an output, we have $m = 2$ and $n = 3$. In the reminder, we shall let $\alpha = \{i_1, i_2, \ldots, i_k\}$ denote the subset of indices of a group of k model inputs ($k \leq n$). $\mathbf{x}_\alpha = \{x_{i_1}, x_{i_2}, \ldots, x_{i_k}\}$ denotes the vector of k model inputs with indices in α. To denote the complementary set of model inputs, we shall use the symbol $\sim \alpha$. The sensitivity of a model output with respect to x_i is often synthesized in a number, which is referred to as the sensitivity measure of x_i. When the sensitivity measure is used to rank model inputs based on their influence on then model output, it is called an *importance measure*.

3. Local sensitivity analysis methods

If one is interested in performing the analysis around a point of interest in the model input space, then one is performing a local sensitivity analysis. We call $\mathbf{x}^0$ the base case (or reference value) of the model inputs. To illustrate, we consider again Harris EOQ, with the reference point $\mathbf{x}^0$ of the copper connector example in Harris (1990). We have $x_1 = 1230$ pieces per month, a unit price $x_2 = \$0.0135$ per unit, a setup cost of \$2.15 leading to the model output $y^0 = g(\mathbf{x}^0) = 6856$. Clearly, for consistency $\mathbf{x}^0 \in \mathcal{X}$. A local sensitivity analysis is carried out in a deterministic framework, that is, no probability distributions are assigned to the model inputs.
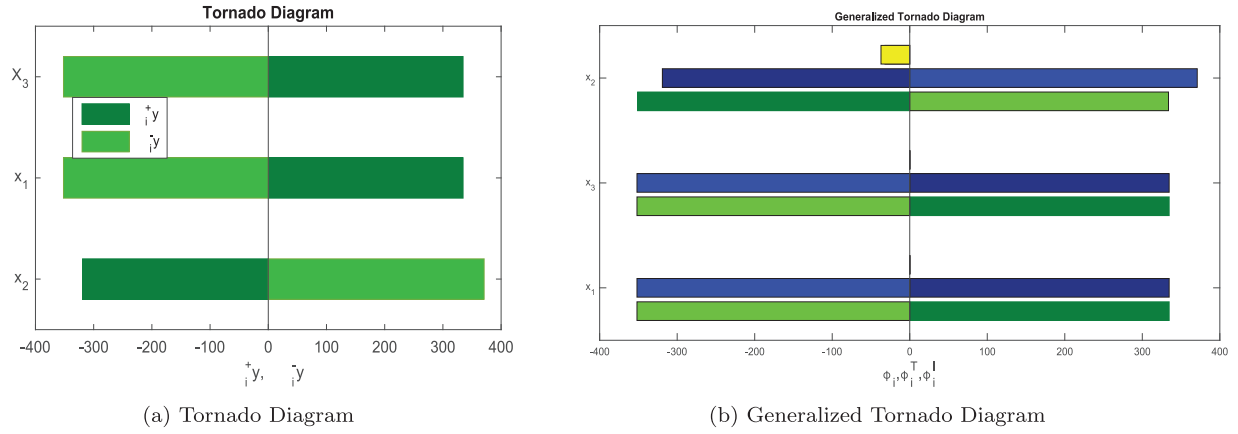
(a) Tornado Diagram

(b) Generalized Tornado Diagram

Fig. 1. Tornado diagram and generalized Tornado diagram for 10 percent model input changes in Harris EOQ. In graph (a), the sensitivity measures are individual effects. Darker green is used for the shift $\mathbf{x}^0 \to \mathbf{x}^+$. In graph (b), for each model input the first sensitivity measure (yellow) is the interaction effect ϕ_i^I, the second (blue) is the total effect ϕ_i^T and the third (green) is the individual effect ϕ_i^T. Darker colors are used for the $\mathbf{x}^0 \to \mathbf{x}^+$ shift. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.1. Sensitivity analysis with finite differences: one at a time methods

The simplest way to interrogate a model is to study the model output when we vary one model input at a time. This approach originates the so-called one-at-a-time (OAT) methods. The procedure is straightforward. We assign a base case ($\mathbf{x}^0$) and a sensitivity case ($\mathbf{x}^+$) to the model inputs. The corresponding vector of model input changes $\Delta^+\mathbf{x} = \mathbf{x}^+ - \mathbf{x}^0 = (\Delta^+x_1, \Delta^+x_2, \ldots, \Delta^+x_n)$. Then, we calculate the model by varying one model input at a time from the base case to the sensitivity case. The sensitivity measures are then the quantities

$$\Delta_i^+y = g(x_i + \Delta x_i^+, \mathbf{x}_{\sim i}^0) - g(\mathbf{x}^0) \qquad (6)$$

where $(x_i + \Delta x_i^+, \mathbf{x}_{\sim i}^0)$ is the point in $\mathcal{X}$ obtained shifting x_i alone to the sensitivity case. From a general perspective, Δ_i^+y are finite changes in y provoked by individual changes in the model inputs.

The sensitivity measures in Eq. (6) quantify the effect of the shift in x_i from the base case to the sensitivity case. The insights delivered by these sensitivity measures concern both the magnitude of the impact and the direction of change. To illustrate, in our example, the base case is $\mathbf{x}^0 = (1230, 0.0135, 2.15)$. If we suppose that, in our sensitivity case, the model inputs undergo a 10 percent increase, the sensitivity case becomes $\mathbf{x}^+ = (1353, 0.01485, 2.365)$. If we consider the variation in x_1 alone, we obtain the corresponding model output change $\Delta^+y_1 = y(1353, 0.0135, 2.15) - y^0 = 335$. That is, a 10 percent increase in x_1 generates an increase of 335 units in the optimal order quantity. Similarly, for the remaining model inputs, we obtain $\Delta^+y_2 = -319$ and $\Delta^+y_3 = y(1230, 0.01485, 2.15) - y^0 = 335$ meaning, respectively, that a 10 percent increase in x_2 and x_3 provokes a decrease of 319 units and an increase of 335 units in EOQ, respectively.

The most successful graphical representation of one-at-a-time (OAT) results is represented by Tornado diagrams. In a Tornado diagram, we represent the sensitivity measures in Eq. (6) as horizontal bars, sorted from the largest to the smallest. The introduction of Tornado diagrams is due to Howard (1988). A detailed description of their implementation is provided for by Eschenbach (1992). Actually, in a Tornado diagram we represent two series of OAT sensitivities, hypothesizing a second sensitivity case (worst case), $\mathbf{x}^-$. We then have a second set of model input changes, $\Delta^-\mathbf{x} = \mathbf{x}^0 - \mathbf{x}^- = [\Delta^-x_1, \Delta^-x_2, \ldots, \Delta^-x_n]$. The sensitivity measures represented in a Tornado diagram are, then, Δ_i^+y in Eq. (29) and $\Delta_i^-y = g(x_i - \Delta x_i^-, \mathbf{x}_{\sim i}^0) - g(\mathbf{x}^0)$. The computational cost of a complete Tornado diagram is, then, $C = 2n + 1$ model runs.

Graph (a) in Fig. 1 displays the sensitivity measures obtained with the two series of OAT variations for the Harris EOQ. We observe that the direction of change is reversed in the shift from $\mathbf{x}^0 \to \mathbf{x}^-$, i.e., the shift from the base case to the worst case [the sensitivity measures are Δ_i^-y; bars in lighter color in graph (a) of Fig. 1], with respect to the shift from the base case to the best case [the sensitivity measures are Δ_i^+y; bars in darker color in graph (a) of Fig. 1]. Not only, but the effects are not symmetric and in the shift $\mathbf{x}^0 \to \mathbf{x}^-$, the key driver becomes x_2, which ranked after x_1 and x_3 in the shift from $\mathbf{x}^0 \to \mathbf{x}^+$. This result shows that the conclusions about sign of change and key-drivers that we may infer from an OAT approach are limited to the specific pair of points under consideration.

We conclude with the following observation. The overall value in Harris EOQ for the change $\mathbf{x}^0 \to \mathbf{x}^+$ is $\Delta y^+ = 335$. However, if we sum the sensitivity measures of the Tornado diagram we obtain $\Delta^+y_1 + \Delta^+y_2 + \Delta^+y_3 = 350 \neq 335$. The difference is due to the fact that the sensitivity measures of a Tornado diagram neglect interaction effects. These sensitivity measures have been generalized in subsequent literature, as we discuss in the next subsection.

3.2. Scenario decomposition and generalized tornado diagrams

In several practical applications, managers (and scientists) explore scientific codes on alternative values of the model inputs, i.e., on alternative scenarios. It is common to find reports in which the results produced by a model are presented with the model inputs fixed at a base case, a worst case and a best case scenario. In assessing the model over these scenarios, the manager wishes to understand the response of the model to extreme variations of the inputs. However, scenario analysis in the operational research literature has also been studied from a broader perspective to guide strategic planning. According to Jungermann and Thuring (1988, p. 117), *Scenarios are descriptions of alternative hypothetical futures: They do not describe what the future will be (i.e., they are not prognoses) but which alternative futures we might reasonably imagine.* According to O'Brien (2004), they are *different states of the world in presence of uncertainty.* Scenarios are *different possible future states of a system* in Tietje (2005). Over time, scenario analysis has become a major decision tool in economics and strategic management. For general perspectives on scenario analysis, we refer the interested reader to the monograph of Schwartz (1996). We refer to O'Brien (2004) and Tietje (2005) for the methodological perspectives. In particular, according to Tietje (2005), scenarios must be *consistent*, *diverse*, in a *small number*, *reliable* and *efficient*.

When we come to quantitative modeling, scenario analysis implies to fix the model inputs at two or more predetermined points

in the model input space. To illustrate, the points $\mathbf{x}^0$, $\mathbf{x}^+$ and $\mathbf{x}^-$ can be seen as three possible scenarios. Then, in a scenario analysis we obtain the three values $y(\mathbf{x}^+)$, $y(\mathbf{x}^0)$ and $y(\mathbf{x}^-)$. By this quantitative information, we get to know the new values of the model outputs and we can appreciate whether the model output has increased or decreased. However, we are not able to quantitatively explain what has caused the change.

The methods we are to review in this section address this problem. Consider the generic change in model output $\Delta y = g(\mathbf{x}^+) - g(\mathbf{x}^0)$. A possible way to decompose this change in an exact fashion and through a finite number of terms is represented by using a finite differences approach (see Borgonovo 2010b for greater details on the mathematical aspects). One then writes

$$\Delta y = \sum_{i=1}^{n} \phi_i + \sum_{i<j} \phi_{i,j} + \cdots + \phi_{1,2,\ldots,n} \qquad (7)$$

where

$$\begin{cases} \phi_i = g(x_i^+; \mathbf{x}_{\sim i}^0) - g(\mathbf{x}^0) \\ \phi_{i,j} = g(x_i^+, x_j^+; \mathbf{x}_{\sim i,j}^0) - \phi_i - \phi_j - g(\mathbf{x}^0) \\ \phi_{i,j,k} = g(x_i^+, x_j^+, x_k^+; \mathbf{x}_{\sim i,j,k}^0) - \phi_{i,j} - \phi_{i,k} - \phi_{j,k} - \phi_i \\ \qquad - \phi_j - \phi_k - g(\mathbf{x}^0) \\ \cdots \end{cases} \qquad (8)$$

Eq. (7) decomposes the finite change in model output into the 2^n terms given in Eq. (8). For the terms $\phi_{i,j,\ldots,k}$ one can use the term finite change sensitivity indices (Borgonovo, 2010b). They have the following meaning: ϕ_i represents the individual effect of varying one model input at a time from the base case to the sensitivity case. Hence, we immediately have that

$$\phi_i = \Delta_i^+ y \qquad (9)$$

that is, first order finite change sensitivity indices are the sensitivity measures of Tornado diagrams. $\phi_{i,j}$ represents the residual interaction effect due to the simultaneous variation of x_i and x_j. In fact, $g(x_i^+, x_j^+; \mathbf{x}_{\sim i,j}^0) - g(\mathbf{x}^0)$ is the change in model output due to the simultaneous variation of x_i and x_j, from which we subtract the changes provoked by the individual variations. The higher order sensitivity indices possess a similar interpretation. To illustrate, let us compute $\phi_{1,3}$ for Harris EOQ. We have $g(x_1^+, x_3^+; x_2^0) - g(\mathbf{x}^0) = 686$ from which we need to subtract the individual effects of x_1 and x_3, obtaining $\phi_{1,3} = 686 - 335 - 335 = 16$. Thus, there is a positive interaction between x_1 and x_3. This means that the simultaneous change in these two model inputs amplifies their individual effects. Proceeding in a similar way, we obtain the remaining second order effects: $\phi_{1,2} = \phi_{2,3} = -16$. The residual effect $\phi_{1,2,3}$ can be computed from

$$\begin{aligned} \phi_{1,2,3} &= g(\mathbf{x}^+) - g(\mathbf{x}^0) - \phi_{1,2} - \phi_{1,3} - \phi_{2,3} - \phi_1 - \phi_2 - \phi_3 \\ &= 335 + 16 - 16 + 16 - 335 - 335 + 319 = 0. \end{aligned} \qquad (10)$$

Thus, there is no interaction term of order 3. One can define the total finite change indices for a generic model input as

$$\phi_i^T = \phi_i + \sum_{j=1}^{n} \phi_{i,j} + \sum_{k=1}^{n} \sum_{j=1}^{n} \phi_{i,j,k} + \cdots + \phi_{1,2,\ldots,n}. \qquad (11)$$

To illustrate, for Harris EOQ, we have

$$\phi_1^T = 335; \ \phi_i^T = 350; \ \phi_i^T = 335. \qquad (12)$$

If, from the total order indices, we subtract the indices of order 1, we obtain the interaction indices

$$\phi_i^I = \phi_i^T - \phi_i = \sum_{j=1}^{n} \phi_{i,j} + \sum_{k=1}^{n} \sum_{j=1}^{n} \phi_{i,j,k} + \cdots + \phi_{1,2,\ldots,n}. \qquad (13)$$

The interaction indices condense in one indicator the interaction effects of model input x_i with all the remaining model inputs and model input groups.

One can conveniently represent this information in a so-called generalized Tornado diagram (Borgonovo & Smith, 2011). In the diagram, for each model input, we represent a triplet of bars, displaying ϕ_i, ϕ_i^T and ϕ_i^I, that is, the first, total and interaction effects. That is, in each triplet, the first bar is the same as in a traditional Tornado diagram. By a generalized Tornado diagram, an analyst can rank model inputs based on their total effects and has a visualization of the relevance of interactions. The sign of interactions let us grasp whether the individual effects are amplified or dampened by interactions.

To illustrate, we report the generalized Tornado diagram for the Harris EOQ in Fig. 1. In this case, we sort the model inputs based on the absolute value of the total order indices. The second model input, x_2, now ranks as most important, followed by x_1 and x_3. The reason is that the indices now include also interaction and not only individual effects. In fact, Graph b) in Fig. 1 shows that the only model input for which interaction effects matter is x_2, which is associated with a non-null $\phi_i^I \approx -32$. This interaction effect is the same for both model input shifts, from $\mathbf{x}^0$ to $\mathbf{x}^+$ and from $\mathbf{x}^0$ to $\mathbf{x}^-$. Consider now the $\mathbf{x}^0 \to \mathbf{x}^+$ shift. In this case, the individual effect of x_2 is a negative $\phi_2 = -319$, that, summed to interaction effects leads to a total effect of $\phi_2^T = -350$, making x_2 the most important input for the shift $\mathbf{x}^0 \to \mathbf{x}^+$.

Let us now come to computational cost. Computing the total and interaction effects by calculating the sensitivity indices of all orders is equivalent to performing a full factorial design with two levels. The associated computational cost is 2^n model runs and makes the estimation infeasible for models with a large number of inputs. However, it is possible to estimate the first, interaction and total order finite change sensitivity indices at a cost of $2n + 2$ model runs [see Borgonovo (2010b) for the proof]. In particular, we obtain the total order finite change sensitivity indices starting at $\mathbf{x}^+$ and shifting the model inputs one at a time from $\mathbf{x}^+$ to $\mathbf{x}^0$. Thus, (finite change) total effects are obtained at a cost of $n + 1$ model runs. Because first order effects are obtained at the same cost, and the interaction effects ϕ_i^I are obtained by subtraction, the cost of computing the triplet of sensitivity measures ϕ_i, ϕ_i^I and ϕ_i^T is $2n + 2$ model runs. Then, if this operation is repeated twice (once for the $\mathbf{x}^0 \to \mathbf{x}^+$ and once for the $\mathbf{x}^0 \to \mathbf{x}^-$ shift), the overall cost is $4n + 3$ model runs.

3.3. Spiderplots and one way sensitivity functions

The sensitivity measures of Tornado diagrams foresee to register the values of the model output at the extremes of the variation ranges of the model inputs. An alternative approach is to register the values of the model output while the input varies over the entire predetermined range. In this case, the "sensitivity measure" is the function

$$h_i(x_i) = g(x_i; \mathbf{x}_{\sim i}^0) \qquad (14)$$

$h_i(x_i)$ is called a one-way sensitivity function. Sensitivity functions can be graphed either individually or condensed in the same plot. In this second case, we are dealing with a so-called spiderplot graph (Eschenbach, 1992). In a spiderplot, one displays the functions:

$$h_i^*(x_i) = h_i(x_i) - g(\mathbf{x}^0) = g(x_i; \mathbf{x}_{\sim i}^0) - g(\mathbf{x}^0) \qquad (15)$$

over normalized ranges of the model inputs. Observe that the sensitivity functions of spiderplots are the sensitivity measures of Tornado diagrams evaluated at alternative values of x_i, that is $h_i^*(x_i) = \phi_i = \Delta^+ y_1(x_i)$. Thus, a spiderplot can be also seen as providing the values of first order sensitivity indices as the model inputs vary in the predetermined ranges.

One-way sensitivity functions are applied in several fields, but they have been particularly studied in association with Bayesian networks, where they convey information about *systematically varying one of the network's parameter probabilities while keeping all other parameters fixed* (van der Gaag, Renooij, & Coupé, 2007, p. 104). Castillo, Gutiérrez, and Hadi (1996) and Castillo, Gutiérrez, and Hadi (1997)

obtain analytic expressions for the sensitivity functions of Bayesian network polynomials, exploiting the inherent multilinear structure. One way sensitivity plots for influence diagrams solved through decision circuits are discussed in Bhattacharjya and Shachter (2008). The concept of one-way sensitivity functions can be extended to multiple ways sensitivity plots (called *n*-way sensitivity plots), which are a multidimensional version of one-way sensitivity plots. We refer to Kjaerulff and van der Gaag (2000) for further discussion on *n*-way sensitivity functions.

3.4. Differentiation-based methods

An important class of sensitivity analysis methods is represented by differentiation methods. As stated in Kleijnen (2010, p. 1), *local changes enable the estimation of the gradient* which plays an important role in simulation optimization and, more generally, to the derivation of managerial insights. The rationale for differentiation methods consists in the Taylor series expansion of the model output. This is clearly formalized in Helton (1993). Here, Δy can be decomposed as:

$$\Delta y = g(\mathbf{x}^+) - g(\mathbf{x}^0) = \sum_{i=1}^{n} \frac{\partial g(\mathbf{x}^0)}{\partial x_i}(x_i^+ - x_i^0)$$
$$+ \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{n} \frac{\partial^2 g(\mathbf{x}^0)}{\partial x_i \partial x_k}(x_i^+ - x_i^0)(x_k^+ - x_k^0) + o(||\mathbf{x}^+ - \mathbf{x}^0||^2). \tag{16}$$

If the variations $x_i^+ - x_i^0$ are small enough so that second order terms can be neglected, then $\Delta y \approx \sum_{i=1}^{n} \frac{\partial g(\mathbf{x}^0)}{\partial x_i}(x_i - x_i^0)$ and $\frac{\partial g(\mathbf{x}^0)}{\partial x_i}$ then become natural sensitivity measures. Partial derivatives are the sensitivity measures we are probably most familiar with. These are the sensitivity measures of Samuelson's comparative statics (Samuelson, 1947). They play an extremely important role in reliability analysis where they are given the name of importance measures. Reliability importance measures support decision makers in several operational settings: regulators in risk-informed decision-making, and plant managers in decisions concerning graded quality assurance programs, equipment prioritization for maintenance and inspections, etc. (Brewer & Canady, 1999; Cheok, Parry, & Sherry, 1998). Unfortunately, there is no room in this review to enter into a full fledged description of system reliability notation. Then, we will follow an intuitive approach. Consider that y represents the probability of system failure and that x_i is the failure probability of a given component. Then, Birnbaum introduces the concept of a component being critical to a system as the probability that the component is in such a state that its failure causes the whole system to fail. Then, Birnbaum proves that, if failures are statistically independent, then the probability that component i is critical is equal to the partial derivative of y with respect to x_i (Birnbaum, 1969). The Birnbaum importance measure is, historically, the first reliability importance measure. We recall the Fussell–Vesely, the risk achievement worth and the risk reduction worth importance measures, and refer to Fussell (1975), Vesely, Kurth, and Scalzo (1990), Cheok et al. (1998) for a thorough discussion. The work of Nøkland and Aven (2013) presents a systematic overview about the selection of reliability importance measures best suited to the application at hand. We also refer the reader to the monograph on importance measures by Kuo and Zhu (2012). Let us apply partial derivatives to Harris EOQ. We obtain:

$$\frac{\partial g(\mathbf{x}^0)}{\partial x_1} = 2.787, \ \frac{\partial g(\mathbf{x}^0)}{\partial x_2} = -2.539 \cdot 10^5, \ \frac{\partial g(\mathbf{x}^0)}{\partial x_3} = 1.595 \cdot 10^3. \tag{17}$$

Note that if we were to consider the magnitudes of partial derivatives and utilize them to rank model inputs, we would conclude that x_2 is

by far the most important model input. However, the partial derivatives are not comparable, because they are denominated in different units. Thus, they cannot be used to rank the model inputs. A way to circumvent this limitation is to use the fraction of the differential (i.e., the entire term $\frac{\partial g(\mathbf{x}^0)}{\partial x_i}(x_i - x_i^0)$) instead of just the rate of change $\frac{\partial g(\mathbf{x}^0)}{\partial x_i}$. A sensitivity measure that accounts for these facts is the differential importance measure (Borgonovo & Apostolakis, 2001):

$$D_i = \frac{\frac{\partial g(\mathbf{x}^0)}{\partial x_i}dx_i}{\sum_{j=1}^{n} \frac{\partial g(\mathbf{x}^0)}{\partial x_j}dx_j} \tag{18}$$

which is the fraction of the differential change in the model output. Under the assumption of uniform perturbations [$dx_i = dx_j$ for all i and j], the differential importance measure produces equivalent ranking as partial derivatives [Borgonovo and Apostolakis (2001)]. Under the assumption of proportional perturbations, the differential importance measure produces the same ranking as elasticity Borgonovo and Peccati (2004). The elasticity of y with respect to x_i is defined as

$$E_i = \frac{\partial g(\mathbf{x}^0)}{\partial x_i} \frac{x_i^0}{g(\mathbf{x}^0)} \tag{19}$$

Then, if we assume proportional model input changes, we have $\frac{dx_i}{x_i^0} = \frac{dx_j}{x_j^0}$ for all i and j. Multiplying and dividing the numerator and denominator of Eq. (18) by x_i^0 and x_j^0, respectively, we obtain

$$D_i = \frac{\frac{\partial g(\mathbf{x}^0)}{\partial x_i}\frac{dx_i}{x_i^0}x_i^0}{\sum_{j=1}^{n} \frac{\partial g(\mathbf{x}^0)}{\partial x_j}\frac{dx_j}{x_j^0}x_j^0} = \frac{\frac{\partial g(\mathbf{x}^0)}{\partial x_i}\frac{x_i^0}{g(\mathbf{x}^0)}}{\sum_{j=1}^{n} \frac{\partial g(\mathbf{x}^0)}{\partial x_j}\frac{x_j^0}{g(\mathbf{x}^0)}} = \frac{E_i}{\sum_{j=1}^{n} E_j}. \tag{20}$$

In reliability analysis, the elasticity E_i is called criticality importance measure. To illustrate, in our running example, assuming proportional perturbations, we obtain the following differential importance measures for the model inputs:

$$D_1 = 1, \ D_2 = -1, \ D_3 = 1. \tag{21}$$

The identical magnitudes of the sensitivity measures suggest that the model inputs are equally important, with x_2 having, however, an opposite effects on Harris EOQ with respect to x_1 and x_3.

A convenient property of the differential importance measure is additivity. One retrieves the importance of any group of variables as sum of the individual sensitivity measures of the variables in the group. We refer to Borgonovo and Apostolakis (2001) for further details.

Now, consider the ratio between ϕ_i and Δx_i and take the limit for $x_i^+ \to x_i^0$, we obtain

$$\lim_{x_i^+ \to x_i^0} \frac{\phi_i}{\Delta x_i} = \lim_{x_i^+ \to x_i^0} \frac{\Delta_i^+ y}{\Delta x_i} = \frac{\partial y(\mathbf{x}^0)}{\partial x_i}. \tag{22}$$

Similarly, if we consider the ratio $\frac{\phi_i}{\sum_{j=1}^{n}\phi_j}$ we obtain

$$\lim_{x_i^+ \to x_i^0} \frac{\phi_i}{\sum_{j=1}^{n}\phi_j} = \lim_{x_i^+ \to x_i^0} \frac{\Delta_i^+ y}{\sum_{j=1}^{n}\Delta_j^+ y} = D_i. \tag{23}$$

The above equations state the relationships among differentiation-based importance measures and finite change sensitivity indices for infinitesimal perturbations of the model inputs. They also show that

partial derivatives and the differential importance measure do not account for interactions. The literature has then proposed several ways of extending differentiation based sensitivity measures to account for interactions, computing higher order derivatives. One is referred to the works of Hong and Lie (1993), Armstrong (1995) , Lu and Jiang (2007), Gao, Cui, and Li (2007), for extensions of partial derivatives to higher order sensitivity measures. In particular, the joint differential sensitivity measure of order k (called joint reliability measure of order k in reliability studies) is defined as $J_\alpha^k = \frac{\partial^k g(\mathbf{x}^0)}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}}$ and provides that strength of the local interaction between model inputs $x_{i_1}, x_{i_2}, \dots, x_{i_k}$. Zio and Podofillini (2006), Do Van, Barros, and Berenguer (2008), Borgonovo (2010a) propose extensions of the differential importance measure to higher order. D_i is extended to order 2 in Zio and Podofillini (2006) and to order k in Do Van et al. (2008). Borgonovo (2010a) introduces the total order importance measure (D_i^T) which synthesizes all joint sensitivity measures of all orders in a unique index. For infinitesimal changes, $D_l^T \to D_l$, that is, the total order importance tends to the differential importance. If the model is infinitely many times differentiable, then it can be proven that D_l^T and total order finite change sensitivity indices provide identical information. Conversely, when the model is not smooth, information about the effects of finite changes is provided by finite change sensitivity indices alone, which remain defined also when the model presents discontinuities. If the model is differentiable once and changes are small, then finite change sensitivity indices and total order sensitivity provide identical information, and this information coincides with the information provided for by the differential importance measure. As for estimation, differentiation can be either performed through brute force algorithms or through automated procedures (Bischof and Bücker, 2000; Griewank and Walther, 2008, see also www.autodiff.org). In this respect, the availability of automatic differentiation codes allows us to obtain partial derivatives at a relatively low computational cost, making the analysis feasible in most practical applications for which the model output is smooth.

3.5. Screening methods

The methods illustrated in Sections 3.1, 3.2 and 3.4 address the variation of a model around one point in the model input space. Thus, they do not permit a thorough exploration of the model behavior as the inputs vary at several locations in $\mathcal{X}$. Screening methods, instead, foresee the evaluation of the model at a limited number of locations in $\mathcal{X}$. The goal of a screening design is to identify the least important model inputs, while maintaining a parsimonious number of model evaluations. Several reviews on screening methods are available in the literature. We refer to Campolongo, Kleijnen, and Andres (2000, chap. 4); Kleijnen (2009); 2010); Kleijnen, Sanchez, Lucas, and Cioppa (2005). These reviews also show the strict link between screening methods and design of experiment (DOE). An application of DOE in the sensitivity analysis for investment evaluation models used in the energy sector is represented by van Groenendaal and Kleijnen (1997). It is not possible to enter into the details of the numerous screening designs that have been proposed. However, in the management science literature, particular attention has been devoted to sequential bifurcation which has been introduced in Bettonvil (1990) as a supersaturated sequential design of experiments (see also Bettonvil and Kleijnen, 1997). Sequential bifurcation works efficiently under the hypothesis listed in Kleijnen (2010). In particular, the main assumption in the original version of sequential bifurcation is that the model response is well fitted by a first order degree response surface, possibly containing interaction terms. A second assumption is that *the direction of the influence that a factor has on the output is known* (Bettonvil and Kleijnen, 1997, p. 183). Then, sequential bifurcation proceeds by first varying all model inputs from the base case to the sensitivity case and computing the corresponding effects. The model inputs are then split into two groups. One then repeats the analysis varying the model inputs in the first group from the base case to the sensitivity case. If the results of the analysis show that the group is unimportant, then the group is discarded from further analysis and attention is restricted to the second group. The works of Bettonvil (1990) and Bettonvil and Kleijnen (1997) have generated a stream of subsequent literature, making sequential bifurcation a successful method for sensitivity analysis in simulation. In particular, we recall the works of Wan, Ankenman, and Nelson (2006) and Wan, Ankenman, and Nelson (2010), which introduce the controlled sequential bifurcation procedure.

A further successful screening method is the method of Morris (1991) that considers segmenting the model input ranges $[x_i^-, x_i^+]$ in l levels. This segmentation generates a grid into the model input space which, for simplicity, is assumed to be the unit hypercube. Given l levels with n model inputs, we have a total of l^n points in the grid, from which a subsample of r points is drawn at random. For each of the r points, the model is evaluated performing a series of OAT sensitivities. The sensitivity measures, called elementary effects in Morris (1991), are the difference quotients in Eq. (22). The elementary effects are then averaged over the r points in the trajectory. This discussion illustrates the main intuition of the method. Additional technical details about the design are found in Morris (1991), Campolongo et al. (2000), Campolongo and Saltelli (1997). The computational cost of a complete Morris design (in which model inputs are increased and decreased with respect to the base case value) is $C = 2n \cdot r$ where r is the number of replicates. We estimated the elementary effects based on Morris design using a four level design with 32 replicates and the open source software SimLab (2011). One obtains sensitivity measures very much in agreement with partial derivatives, as in the original spirit of the method (see the discussion about the meaning of elementary effect at Morris (1991, p. 163)).

Campolongo, Cariboni, and Saltelli (2007) propose an amelioration of the Morris method using a normalized version of the elementary effects. The normalization avoids cancellation effects making the identification of the most important model inputs more robust.

The screening methods presented so far are using a deterministic framework. Adding a probabilistic setup, the randomized evaluation of partial derivatives is studied for derivative-based global sensitivity methods in Sobol' and Kucherenko (2010) and Rakovec, Hill, Clark, Weerts, Teuling, and Uijlenhoet (2014).

4. Global sensitivity methods

The methods that we have discussed upto now are intrinsically deterministic. The underlying mindset precludes the assignment of a probability distribution to the model inputs. In this respect, it is worth noting that several authors have argued both in favor and against the assignment of probability distributions for sensitivity analysis. van Groenendaal and Kleijnen (1997) maintain the usefulness of using a deterministic approach when the state of knowledge does not allow the analyst to confidently assign distributions to the model inputs. The trade-offs between using a deterministic method that has the advantage of not requiring the assessment of model input distributions, against the use of methods that require a full fledged probabilistic distribution assignment are discussed thoroughly in Wallace (2000). A comparison between deterministic design of experiments and probabilistic methods is offered in van Groenendaal and Kleijnen (2002). In all those situations where it is possible to assign a probability distribution to the model inputs, one can resort to the rich set of probabilistic (or global) sensitivity methods.

4.1. Setup for a probabilistic analysis

The fundamental assumption of a probabilistic sensitivity analysis is that *we assume to have information about the model inputs'*
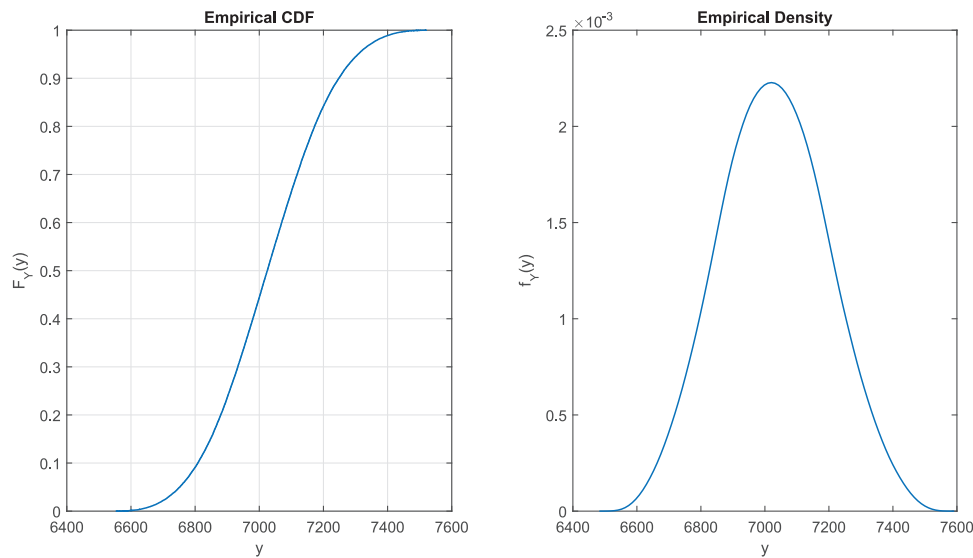
Fig. 2. Empirical cumulative distribution function and density for Harris EOQ.

probability distribution, either joint or marginal, with or without correlation (Saltelli and Tarantola, 2002, p. 704). Then, we introduce the probability space $(\mathcal{X}, \mathcal{B}(\mathcal{X}), \mathbb{P}_{\mathbf{X}})$, where $\mathbb{P}_{\mathbf{X}}$ is the probability distribution of the model inputs. Here, the model input space $\mathcal{X}$ coincides with the support of the distribution $\mathbb{P}_{\mathbf{X}}$. The model inputs become a random vector which is denoted by $\mathbf{X} = (X_1, \ldots, X_n)$, so that $\mathbf{x} = (x_1, \ldots, x_n)$ denotes one of the possible realizations. $F_{\mathbf{X}}(\mathbf{x})$ denotes the joint cumulative distribution function of the model inputs. If the model inputs are absolutely continuous, $f_{\mathbf{X}}(\mathbf{x})$ is the corresponding joint density. $F_{X_i}(x_i)$ and $f_{X_i}(x_i)$ denote the corresponding marginal cumulative distribution function and density, respectively. We denote by σ_i the standard deviation of model input X_i. If the inputs are independent, we can decompose the distribution via $F_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^{n} F_i(x_i)$.

Uncertainty in the model inputs makes y a random variable, which we denote by $Y = g(\mathbf{X})$. Its distribution and moments are obtained by propagating uncertainty in $\mathbf{X}$ through the map g, $\mathbb{P}_Y(\cdot) = \mathbb{P}_{\mathbf{X}}(g^{-1}(\cdot))$. Ideally, this propagation can be performed analytically. The corresponding cumulative distribution function is $F_Y(y)$ and $f_Y(y)$ denotes the probability density, if Y is continuous. We denote the model output standard deviation by σ_Y the variance of the model output. However, in most practical situations, and especially when the closed form expression for $g(\cdot)$ is not available, uncertainty propagation is performed numerically through Monte Carlo simulation. We generate a random sample of size $N \times n$ that follows the distribution of the model inputs and evaluate the model output in correspondence of each realization of $\mathbf{X}$. We then obtain a sample of size $N \times D$, representing N realizations of the model outputs. By this sample, it is possible to obtain the empirical cumulative distribution function (density) of the model output.

To illustrate, for Harris EOQ with a support of plus/minus 10 percent and uniform distributions $X_i \sim U[0.9 \cdot x_i^0, 1.1 \cdot x_i^0]$, with $N = 10,000$ Monte Carlo runs, we obtain the distribution and density in Fig. 2.

If the probabilities assigned to the model inputs are representative of our degree of belief about $\mathbf{X}$, then the distribution and density in Fig. 2 represent our degree of belief about Harris EOQ. The mean value of Y is 7027 and its standard deviation is $\sigma_Y = 167$.

However, several methods have been developed to answer additional questions that help both decision makers and analysts to dig deeper into the obtained results, to extract insights on key uncertainty drivers, direction of change, relevance of interactions,

etc., also in the presence of uncertainty. We shall detail the most used ones: non-parametric methods, variance-based, moment independent, value of information-based methods and Monte Carlo filtering.

4.2. Regression based methods

Non-parametric methods using linear regression techniques are, chronologically, the first class of global sensitivity methods which have been intensively investigated (Helton, 1993; Saltelli & Marivoet, 1990). One of the reasons of the success of non-parametric methods is the fact that these sensitivity measures can be directly estimated from a given Monte Carlo sample (Helton & Davis, 2003). Hence, the analyst can proceed with a Monte Carlo simulation and then estimate non-parametric sensitivity measures by post-processing the obtained input–output sample. No particular design is required in the estimation of non-parametric techniques. Their cost is equal to $C = N$ model evaluations, where N is the size of a Monte Carlo sample. The rationale for non-parametric methods is linear regression.

Assume that the input–output sample is well fitted by a linear regression model. Then, the input–output mapping can be approximated by the response surface (see the monograph by Montgomery (2000)) $g(\mathbf{X}) \approx b_0 + \sum_{i=1}^{n} b_i X_i$. Then, as Helton (1993) and Kleijnen and Helton (1999b) underline, the standardized regression coefficient (SRC) becomes a natural sensitivity measure:

$$\mathrm{SRC}_i = b_i \frac{\sigma_i}{\sigma_Y} \tag{24}$$

where σ_i and σ_Y denote the standard deviations of the model input X_i and the model output Y, respectively. A second widely used non-parametric sensitivity measure is Pearson's product moment correlation coefficient

$$\begin{aligned} \mathrm{PEAR}_i = \varrho(Y, X_i) &= \frac{\mathrm{Cov}(Y, X_i)}{\sigma_i \sigma_Y} \\ &= \frac{\sum_{j=1}^{N} (x_{ji} - \bar{x}_i)(y_j - \bar{y})}{\sqrt{\sum_{j=1}^{N} (x_{ji} - \bar{x}_i)^2} \sqrt{\sum_{j=1}^{N} (y_j - \bar{y})^2}} \end{aligned} \tag{25}$$

for a given sample $\{(x_{j1} \ldots, x_{jn}, y_j) | j = 1, \ldots, N\}$. If we assume that model inputs are independent then SRC_i and Pearson's correlation coefficient coincide. A related sensitivity measure is given by the explanatory power of the linear fit which is expressed by the

coefficient of determination or goodness-of-fit (R_Y^2). R_Y^2 is the fraction of the model output variance explained by the regression.

For Harris EOQ with an input uncertainty of plus/minus 10 percent and uniform distributions $X_i \sim U[0.9 \cdot x_i^0, 1.1 \cdot x_i^0]$ the regression hyperplane is given by $\varphi(x_1, x_2, x_3) = 3.43 \cdot 10^3 + 2.79x_1 - 2.55 \cdot 10^5 x_2 + 1.60 \cdot 10^3 x_3$. This yields standardized regression coefficients of $\mathrm{SRC}_1 = \mathrm{SRC}_3 = 0.58$ and $\mathrm{SRC}_2 = -0.58$. The goodness-of-fit here amounts to $R_Y^2 = 0.998$. Let us now modify the example by assuming that x_2 has an uncertainty range of ± 30 percent. The coefficients of the regression model change only slightly, however, the regression coefficients now yield $\mathrm{SRC}_1 = \mathrm{SRC}_3 = 0.29$, $\mathrm{SRC}_2 = -0.90$ with $R_Y^2 = 0.985$.

Further members of the non-parametric family are discussed in Helton (1993); Helton, Johnson, Sallaberry, and Storlie (2006b); Saltelli and Marivoet (1990); Storlie, Reich, Helton, Swiler, and Sallaberry (2013); Storlie, Swiler, Helton, and Sallaberry (2009). A comparison between non-parametric and screening methods is offered in van Groenendaal and Kleijnen (2002).

In general, the ability of non-parametric methods to capture variability can be measured by R_Y^2, the model coefficient of determination. If a too low value of R_Y^2 is registered, then we have a poor regression fit. Under these conditions, Campolongo and Saltelli (1997) suggest that it might become *unrealistic to assess influence of the input variables based on SRC's*. In general, when the input–output mapping is characterized by the presence of non-linearities and interactions (Saltelli & Marivoet, 1990), the regression fit tends to be weak.

The use of rank transformation can improve the reliability of non-parametric techniques (see Iman and Conover (1979) for a discussion of the use of rank transformation in regression; see also Saltelli and Sobol' (1995) on a more general perspective on the use of rank transformation in global sensitivity analysis). On the use of rank transformation for improving the performance of non-parametric methods, we refer to Helton et al. (2006b). If we replace the model output with the corresponding ranks and compute the Pearson correlation, we obtain the Spearman correlation coefficient (Spearman, 1904). This sensitivity measure allows for detecting nonlinear monotonic dependencies. In fact, the sign of the Spearman correlation coefficient suggests whether Y is (statistically) increasing in X_i. Moreover, if $Y = g(X)$ is a univariate monotonically increasing function of the X, then Spearman correlation coefficient is equal to unity. More generally, one can consider linear regression coefficients derived from transformed variables. Optimizing over all suitable transformation leads to the maximum correlation coefficient, $\sup_{\varphi, \psi} \varrho(\varphi(Y), \psi(X_i))$ where φ, ψ run over functions with finite and positive variance (Rényi, 1959).

The intuition that the robustness of the ranking induced by non-parametric methods is measured by the model coefficient of determination, can be seen as linking non-parametric methods with the class of variance-based sensitivity measures. In fact, R_Y^2 is a ratio of variances. Then, if one introduces sensitivity measures that are able to apportion variance to the contribution of each model input, one has, intuitively speaking, a class of sensitivity measures whose coefficient of model determination is unity. In the next section, we discuss variance-based sensitivity methods.

4.3. Variance-based methods

Variance-based sensitivity methods assess the importance of a model input based on the expected reduction in the model output variance provoked by knowing the value of the model input with certainty. From a formal viewpoint, variance-based sensitivity methods are based upon the well known variance decomposition formula $\mathbb{V}[Y] = \mathbb{V}[\mathbb{E}[Y|X_\alpha]] + \mathbb{E}[\mathbb{V}[Y|X_\alpha]]$ which separates the output variance into the explained variance due to the dependence on a parameter group X_α and the residual variance. These methods employ a non-

linear non-parametric regression curve or surface $\varphi_\alpha(x_\alpha) = \mathbb{E}[Y|X_\alpha = x_\alpha]$. The corresponding goodness-of-fit measure is given by $S_\alpha^{\mathrm{group}} = \frac{\mathbb{V}[\varphi_\alpha(X_\alpha)]}{\mathbb{V}[Y]}$. As in the linear case, $S_\alpha^{\mathrm{group}}$ corresponds to the fraction of output variance which is explained by a functional dependence on x_α. If $\alpha = \{i\}$, then $S_i = S_{\{i\}}^{\mathrm{group}}$ is called the main effect or first order index (Saltelli, Ratto, Andres, Campolongo, Cariboni, Gatelli, Saisana, & Tarantola, 2008). Compared to a linear regression, we have $S_i \geq \beta_i^2$. Indeed, Fréchet (1934) has shown that $S_i = \varrho^2(Y, \mathbb{E}[Y|X_i])$. Higher order effects are defined by the sole contribution of interactions between the factors of an index set α defined recursively by $S_\alpha = S_\alpha^{\mathrm{group}} - \sum_{\beta \subsetneq \alpha} S_\beta$. The index $S_i^T = 1 - S_{\sim i}^{\mathrm{group}}$ is called the total effect. This describes the fraction of output variance that is left unexplained if X_i and only X_i is allowed to vary in its range. The total effect supplies sufficient information to decide whether a factor is unimportant. For uncorrelated inputs, we have $S_i^T \geq S_i$, $\sum_{i=1}^n S_i \leq 1$. In order to cope with further group effects, Liu and Owen (2006) introduce the notion of subset and superset importance measures. Group effects are used to determine the effective dimension of the model (Caflisch, Morokoff, & Owen, 1997; Owen, 2003). This information can be used to reduce the model complexity (Wang, 2006; Wang & Sloan, 2011).

Historically, variance-based sensitivity measures have been defined in alternative disciplines. Variance-based sensitivity measures of order 1 (S_i) have been defined by Pearson (1905) and are known under the name of correlation ratio. Pearson introduces them to overcome limitations of the linear regression coefficient in the presence of non-linearities. However, the estimation of Pearson's correlation ratio is difficult without a computer and only with the rise of computing machinery and with the simulation revolution in applied sciences this measure has regained interest. The idea of using variance reduction to define sensitivity measures, in fact, becomes predominant in the early 1990s. In the field of Risk Analysis, Iman and Hora (1990) define variance-based sensitivity measures as $IH_i = \mathbb{V}[\mathbb{E}[Y|X_i]]$. IH_i is the numerator of Pearson's correlation ratio. In the same years, Sobol' (1993) defines variance-based sensitivity indices in the context of high-dimensional integration. The definition of Sobol' assumes a product measure, i.e., independence among the model inputs. In Operations Research, Wagner (1995) introduces first and total order variance-based sensitivity measures, without relying on the independence assumption. We observe that Sobol' (1993), Iman and Hora (1990) and Wagner (1995) arrive at similar definitions almost simultaneously but independently, as there are no cross-citations and, moreover these works do not cite Pearson's original work. Nonetheless, if we wish to be even more historically accurate, all these works are preceded by the seminal work of Cukier, Fortuin, Shuler, Petschek, and Schaibly (1973), where variance-based sensitivity measures are used in the context of scientific modeling for the first time.

Revisiting Harris EOQ with an uncertainty range of ± 30 percent for x_2^0 and of ± 10 percent for the other two factors, we estimate first order effects by computing the goodness of fit for the regression curves, see Fig. 3. We obtain $S_1 = 0.087$, $S_2 = 0.826$, $S_3 = 0.087$. As $S_i \approx \beta_i^2$ and $\sum_{i=1}^3 S_i \approx 1$ the model is essentially additive and linear. This is also supported by the scatterplots in Fig. 3. However, visual inspection of scatterplots might become cumbersome if the number of input factors is large. Kleijnen and Helton (1999a, 1999b) suggest the use of scatterplots to as an initial step (as a visual sensitivity screening exercise) of a broader sensitivity exercise that goes beyond simple visual examination, and foresees the use of non-parametric regression and correlation-based techniques [see also Helton et al. (2006b)].

An all-in-one-method is the use of curves which show the contribution to the sample mean (csm) or the cumulative sum for reordered normalized output (cusunoro). The output permutation is given by increasingly ordering the input of interest (Bolado-Lavin, Castaings, & Tarantola, 2009; Plischke, 2012a). An example is the lower right plot in Fig. 3 showing curves for all three parameters at once. The curves for x_1 and x_3 coincide, again supporting the finding that their
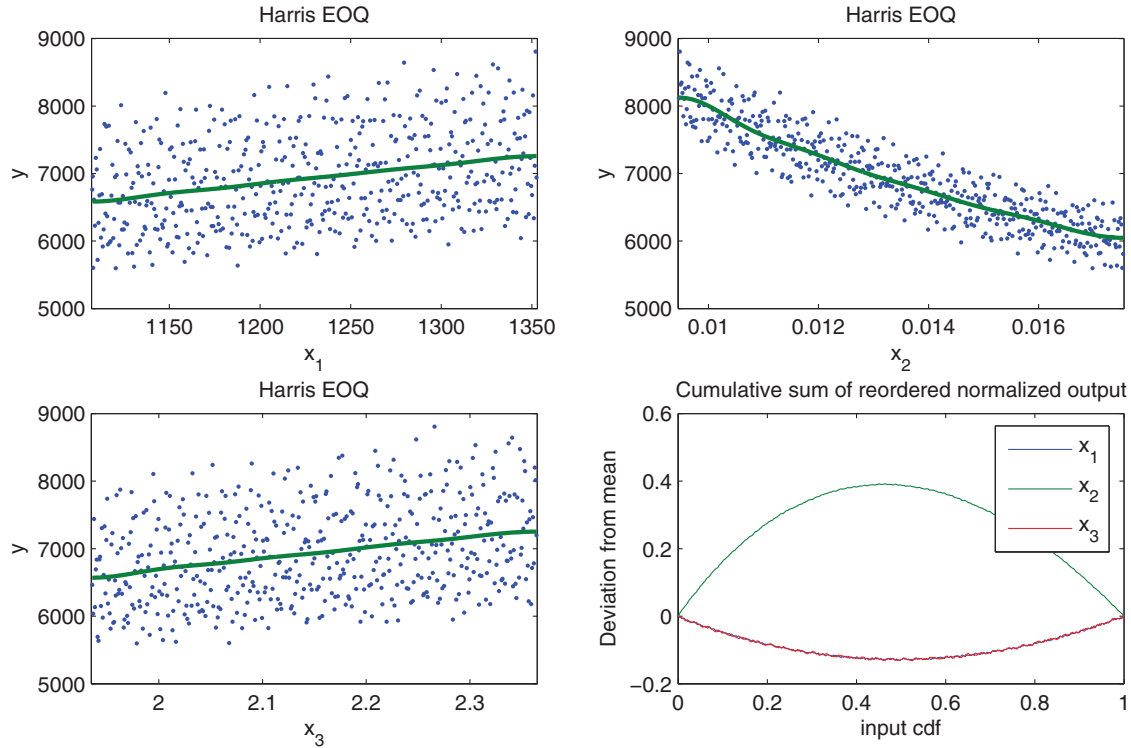
Fig. 3. Scatterplots and cusunoro sensitivity curves for Harris EOQ.

sensitivity is the same. A link to first order effects can be established via the averaged squared gradient of these curves. Variance-based sensitivity measures are intertwined with the functional ANOVA decomposition. The seminal result of Efron and Stein (1981) proves that a multivariate mapping g, integrable on $(\mathcal{X}, \mathcal{B}(\mathcal{X}), \mu)$, can be decomposed as follows:

$$g(\mathbf{x}) = g_0 + \sum_{i=1}^{n} g_i(x_i) + \sum_{i<j} g_{i,j}(x_i, x_j)$$
$$+ \cdots + g_{1,2,\ldots,n}(x_1, x_2, \ldots, x_n) \quad (26)$$

where

$$\begin{cases} g_0 = \int \cdots \int g(\mathbf{x}) d\mu \\ g_i(x_i) = \int \cdots \int g(\mathbf{x}) \prod_{k\neq i} d\mu_k - g_0 \\ g_{i,j}(x_i, x_j) = \int \cdots \int g(\mathbf{x}) \prod_{k\neq i,j} d\mu_k - g_i(x_i) - g_j(x_j) - g_0 \\ \cdots \end{cases} \quad (27)$$

In the broad context of scientific modeling, functional ANOVA is at the basis of the high dimensional model representation theory, which plays a fundamental role in global sensitivity analysis. This theory is developed in the works of Rabitz and Alış (1999), Li, Wang, Rosenthal, and Rabitz (2001), Alış and Rabitz (2001), Sobol' (2003). Under independence, the $g_\alpha(\mathbf{x}_\alpha)$ functions in Eq. (26) are orthogonal (Sobol', 1969). By orthogonality, $\mathbb{V}[Y] = \mathbb{E}[(g(\mathbf{X}) - g_0)^2]$ may be written as

$$\mathbb{V}[Y] = \sum_\alpha V_\alpha \quad \text{with} \quad V_\alpha = \int_{\mathcal{X}_\alpha} \left[g_\alpha(\mathbf{x}_\alpha)\right]^2 f_\alpha(\mathbf{x}_\alpha) \, d\mathbf{x}_\alpha. \quad (28)$$

By Eq. (28), the terms in the variance decomposition are the second moments of the g_α functions. Thus, the ANOVA decomposition of the model input–output mapping, $g(\cdot)$, is in one-to-one correspondence with the decomposition of its variance. As properly stated in Oakley and O'Hagan (2004, p.753), under independence, variance decomposition *reflects the structure of the model itself.* In the presence of correlations, however, the tidy correspondence between functional decomposition and variance-decomposition is lost. Technically, if the integration measure is not a product measure, orthogonality cannot be exploited to eliminate cross products, and we cannot obtain

the variance decomposition from the functional ANOVA decomposition. Moreover, Bedford (1998) shows that, if correlations are present, Eq. (26) is no more unique. Oakley and O'Hagan (2004) and Saltelli and Tarantola (2002) offer examples in which correlations generate spurious terms in the functional ANOVA expansion that are not associated with a corresponding functional dependence in the model input–output mapping.

In particular, Saltelli and Tarantola (2002) propose the following input–output mapping, $y = x_1 + a_{23}x_2x_3$ where a_{23} is a nonvanishing coefficient. Then, under the assumption that X_1 and X_2 are correlated, the first order variance-based sensitivity measure of X_1 would register a contribution from the term $a_{23}x_2x_3$. This dependence does not pertain to the model structure (and would not be registered if X_1 and X_2 were uncorrelated), in so far as the input–output mapping is additive with respect to X_1. The study of variance-based sensitivity measures in the presence of dependencies among the model inputs is attracting increasing attention. Several authors have recently proposed alternative approaches and research is actively ongoing (Chastaing, Gamboa, & Prieur, 2012; Kucherenko, Tarantola, & Annoni, 2012; Li & Rabitz, 2012; Mara & Tarantola, 2012). The literature has also observed that a null value of S_i (or V_i) does not reassure the analyst that the model output is independent of X_i. A family of functions for which this is observed is presented in Plischke, Borgonovo, and Smith (2013). This family generalizes the well known Ishigami test function, widely used in sensitivity analysis studies (Ishigami & Homma, 1990). We finally point out an interpretation issue that has appeared in association with variance-based sensitivity measures. Their popularity has led authors to interpret variance-based sensitivity measures as *the expected reduction in uncertainty due to observing x_i* (Oakley & O'Hagan, 2004, p. 753). However, it is well known that uncertainty in a random variable cannot be summarized by one of its moments. In particular, variance is sufficient to characterize variability or risk either under the assumption of normal distributions or under the assumption that the utility function of the decision-maker is quadratic (Huang & Litzenberger, 1988). In general, electing a moment as representative of uncertainty
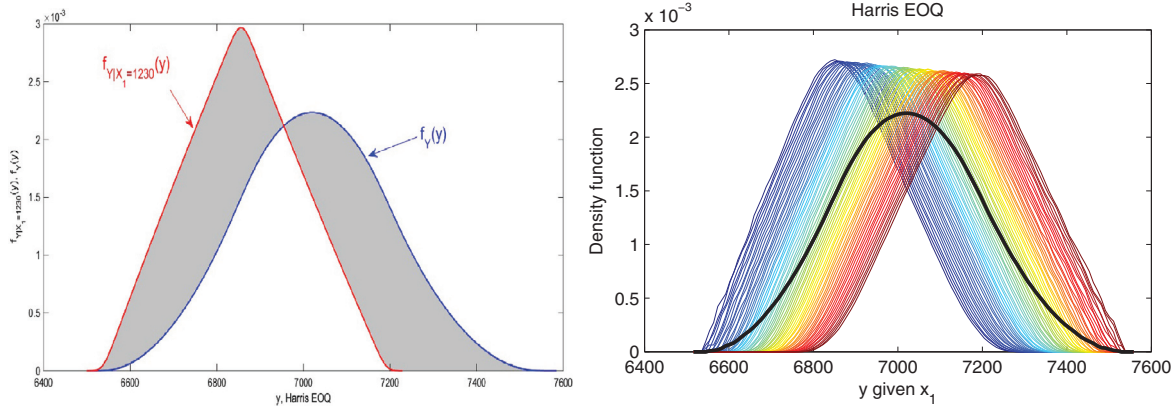
Fig. 4. Left: conditional density given that $X_1 = 1230$ and unconditional density for Harris EOQ. Right: conditioning on all $x_1 \in \mathcal{X}_1$ by binning the sample data. (For interpretation of the references to colour in the text, the reader is referred to the web version of this article.)

may lead to misleading conclusions also in sensitivity analysis (see Borgonovo, 2006, for an example).

4.4. Density-based methods

Sensitivity measures that consider the entire distribution without reference to a particular moment are called moment-independent methods. The intuition is as follows. If we let all model inputs free to vary in accordance with the assigned distributions, we obtain the unconditional model output density $f_Y(y)$. For the EOQ model, assuming uniform distribution over the plus 10 percent increase, we obtain the unconditional density in the left plot of Fig. 4 (continuous line; red). Suppose then, that we are informed that model input X_i is fixed at its base value, $x_1^0 = 1230$. We obtain the conditional model output density represented by the dotted line in the left graph of Fig. 4. Then, one can measure the effect of fixing X_1 at $x_1 = 1230$ by introducing a distance between these two densities. The works of Chun, Han, and Tak (2000); Park and Ahn (1994) use, respectively, the Kullback–Leibler distance and a Minkowski distance of order 2 to measure the separation between the unconditional model output density and the density conditional on fixing a model input at a sensitivity case. Alternatively, if one uses the L^1 norm, one would measure the gray area enclosed between the two curves in Fig. 4 (Borgonovo, 2007). However, the value $x_1 = 1230$ is not the only value that X_1 can assume. Then, the definitions in Park and Ahn (1994), Chun et al. (2000) would lead to sensitivity measures whose value depends on the chosen sensitivity case. Conversely, one has as many conditional densities as there are possible values of X_i. To illustrate, the right graph in Fig. 4 plots the conditional densities of Harris EOQ obtained by fixing X_1 at 10,000 randomly sampled values. This graph provides a visual intuition of the impact on the decision-maker's degree of belief about Y provoked by fixed X_1. Then, the sensitivity measure can be made independent of the sensitivity case by considering the expected distance over all possible values of the model input. This is achieved by averaging according to the marginal distribution of X_1. The first representative of the class of density-based sensitivity measures is the δ-sensitivity measure, defined as (Borgonovo, 2007)

$$\delta_i = \frac{1}{2}\mathbb{E}\left[\int_{\mathcal{Y}} \left| f_Y(y) - f_{Y|X_i}(y) \right| dy\right]. \tag{29}$$

The definition of δ in Eq. (29) does not require independence between the model inputs. In Eq. (29), the quantity $\gamma_i(x_i) = \int_{\mathcal{Y}} \left| f_Y(y) - f_{Y|X_i=x_i}(y) \right| dy$ is the area enclosed between the conditional and unconditional model output densities obtained for a particular value of X_i (left graph in Fig. 4). $\gamma_i(x_i)$ is called inner statistic (or inner separation). Eq. (29) suggests that δ_i is non-negative and normalized between zero and unity ($0 \leq \delta_i \leq 1$). Moreover, the joint

δ of all model inputs is equal to unity. That is, removing uncertainty in all model inputs leads to the highest possible effect.

Eq. (29) can be rewritten in symmetric form as (Plischke et al., 2013)

$$\delta_i = \frac{1}{2}\int_{\mathcal{X}_i}\int_{\mathcal{Y}} \left| f_{Y,X_i}(y, x_i) - f_Y(y)f_{X_i}(x_i) \right| dy\,dx_i. \tag{30}$$

Eq. (30) shows that δ measures the distance between the joint distribution of X_i and Y ($f_{Y,X_i}(y, x_i)$) and the product of their marginal distributions ($f_Y(y)f_{X_i}(x_i)$). In this respect, it has been ascertained that δ_i is a stronger measure of statistical dependence than Pearson's linear correlation coefficient or correlation ratio (first order variance-based sensitivity measures). In fact, one can prove that $\delta_i = 0$ if and only if Y is independent of X_i. That is, δ_i is null if and only if Y (the model output) is independent of X_i [see Plischke et al. (2013)]. Thus, a null value of δ_i reassures the analyst that the model output is independent of X_i.

By its properties, δ_i is attracting increasing research attention, and especially its estimation has been the subject of investigation in several works. As for estimation, we recall the works of Castaings, Borgonovo, Morris, and Tarantola (2012); Liu and Homma (2010); Luo, Lu, and Xu (2014); Zhang, Lu, Cheng, and Fan (2014). Recently, Zhai, Yang, Xie, and Zhao (2014a) propose a generalization of δ_i based on the Minkowski distance of order p among cumulative distribution functions.

Finally, δ_i is monotonic transformation invariant. This is a convenient property, especially in estimation, as we are to discuss in the next subsection.

A second popular choice to obtain a transformation invariant global sensitivity measure is represented by using the Kullback–Leibler divergence to measure the discrepancy between the conditional and unconditional model output densities. In this case, one defines the inner separation as

$$\gamma_i^{KL}(X_i) = \int_{\mathcal{Y}} f_{Y|X_i}(y)\big(\log f_{Y|X_i}(y) - \log f_Y(y)\big)dy \tag{31}$$

to obtain the sensitivity measure

$$\theta_i^{KL} = \mathbb{E}\left[\int_{\mathcal{Y}} f_{Y|X_i}(y)\big(\log f_{Y|X_i}(y) - \log f_Y(y)\big)dy\right] \tag{32}$$

To our knowledge, the first work proposing a global sensitivity measure based on the Kullback–Leibler divergence is Critchfield and Willard (1986). The work discusses a medical decision making application. Later works, such as Auder and Iooss (2009); Krzykacz-Hausmann (2001); Liu, Chen, and Sudjianto (2006); Park and Ahn (1994); Tang, Lu, Jiang, Pan, and Zhang (2013), discuss the use of the sensitivity measure in Eq.(32) in different fields of application.

The sensitivity measure in Eq. (32) possesses similar properties as the δ_i importance measure. In particular, $\theta_i^{KL} = 0$ if and only if the

model output Y is statistically independent of X_i, and θ_i^{KL} is monotonic transformation invariant.

The reason for this last result is that, as shown in Borgonovo, Tarantola, Plischke, and Morris (2014) (see also Da Veiga, 2015), both θ_i^{KL} and δ_i are based on inner statistics which are members of the family of Csiszar divergences:

$$\gamma_i^{\text{Csiszar}}(X_i) = \int_{\mathcal{Y}} f_{Y|X_i}(y) t \left\{ \frac{f_Y(y)}{f_{Y|X_i}(y)} \right\} dy \qquad (33)$$

where t is, by definition, convex and such that $t(0) = 1$, giving rise to the class of sensitivity measures

$$\theta_i^{\text{Csiszar}} = \mathbb{E}\left[\int_{\mathcal{Y}} f_{Y|X_i}(y) t \left\{ \frac{f_Y(y)}{f_{Y|X_i}(y)} \right\} dy \right] \qquad (34)$$

Specifying alternative forms for the convex function t, one obtains alternative metrics which are all monotonic transformation invariant. For instance, setting $t(s) = |s - 1|$ (total variational distance), one finds the δ_i importance measure (Borgonovo et al., 2014) or setting $t(s) = \left(\sqrt{s} - 1\right)^2$, one finds the Hellinger distance (Da Veiga, 2015). Other specifications lead to the Pearson's $\mathcal{X}^2$ divergence, and the Neyman $\mathcal{X}^2$ divergence (Da Veiga, 2015).

A general viewpoint on transformation invariance is offered in the next section.

4.5. Transformation invariant methods

The literature has shown that transformation invariance is a convenient property for a global sensitivity measure. In realistic applications, it is not infrequent that the model output is sparse or ranges across several orders of magnitude. In these situations, numerical convergence of global sensitivity estimators might not be at reach for reasonable sample sizes. Analysts then resort to monotonic transformations. As Iman and Hora (1990, p. 402) observe, *The scaling problem most often can be overcome by performing uncertainty importance calculations based on a logarithmic scale....* Rank transformations (see Iman and Conover, 1979; Saltelli and Sobol', 1995) are also typically used. In fact, transformations lead to an accelerated numerical convergence and improve estimation accuracy at a given sample size.

We say that a function $U = u(Y)$ is a monotonic transformation of the model output if $u(\cdot)$ is a monotonically increasing (or decreasing) function of its argument. For instance, in a logarithmic transformation, we have $U = \log(Y)$. Clearly, this transformation applies for all random model outputs with support included in $\mathbb{R}^+$.

From a sensitivity analysis viewpoint, performing calculations on a transformed model output (or, in general on a transformed random variable) introduces the problem of reinterpreting results back on the original scale, a problem already identified in Iman and Hora (1990). First, a transformation changes the structure of the input–output mapping. For instance, consider a product model

$$Y = X_1 \cdot X_2 \ldots X_n \qquad (35)$$

The model is clearly non-additive (it is, actually made of a unique interaction term). Then, consider the log-transformation of this model:

$$U = \log(Y) = \log(X_1) + \log(X_2) + \cdots + \log(X_n), \qquad (36)$$

The input–output mapping is now additive and no interactions are present. Clearly, sensitivity analysis insights obtained for the transformed model are not directly transferable to the original model. Moreover, numerical experiments in Borgonovo et al. (2014) show that we are not guaranteed that the ranking of a model input is the same on the original and the transformed scales. However, in the case the adopted sensitivity measure were transformation invariant, then its results would be independent of the transformation. An analysts could then benefit from the accelerated numerical convergence,

while avoiding the problems of transferring results back to the original scale.

Let $\xi_i(Y)$ $[\xi_i(U)]$ denote the generic sensitivity measure of model input X_i when the model output Y $[U]$. As in Baucells and Borgonovo (2013), we say that a sensitivity measure is monotonic transformation invariant if

$$\xi_i(U) = \xi_i(Y) \qquad (37)$$

for all monotonic transformations $U = u(Y)$.

Transformation invariant global sensitivity measures are obtained if one adopts an inner statistic which is transformation invariant. Baucells and Borgonovo (2013) introduce two sensitivity measures based on the Kolmogorov–Smirnov and Kuiper distances between cumulative distribution functions, defined as

$$\beta_i^{KS} = \mathbb{E}\left[\sup_y \left| F_Y(y) - F_{Y|X_i}(y) \right| \right] \text{ and}$$

$$\beta_i^{Ku} = \mathbb{E}\left[\sup_y \left\{ F_Y(y) - F_{Y|X_i}(y) \right\} + \sup_y \left\{ F_{Y|X_i}(y) - F_Y(y) \right\} \right] \qquad (38)$$

These sensitivity measures possess a managerial interpretation in terms of probability of success and failure, as discussed in Baucells and Borgonovo (2013). Alternative metrics lead to transformation invariant sensitivity measures. For instance, if one uses the Anderson and Darling (1952) and Mason and Schuenemeyer (1983) modifications of the Kolmogorov–Smirnov metric, or the modification proposed in Crnkovic and Drachman (1996) of the Kuiper metric, one obtains sensitivity measures which are transformation invariant (see Baucells and Borgonovo (2013) for further discussion on metric choice). In this respect, a general family of monotonic transformation invariant metrics is defined in Borgonovo et al. (2014) as:

$$z(\mathbb{P}, \mathbb{Q}) = \sup_{A \in \mathcal{A}} h\left(\left| \mathbb{P}(A) - \mathbb{Q}(A) \right| \right) \qquad (39)$$

where $\mathbb{P}$ and $\mathbb{Q}$ are two probability measures on measure space $(\Omega, \mathcal{A})$ and $h : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a continuous and non-decreasing mapping such that $h(0) = 0$ and $\sup_t \frac{h(2t)}{h(t)} < \infty$. The family in Eq. (39) encompasses the variational distance (Strasser, 1985), $z(\mathbb{P}, \mathbb{Q}) = \sup_{A \in \mathcal{A}} |\mathbb{P}(A) - \mathbb{Q}(A)|$, and the set of Birnbaum–Orlicz metrics whose main representative is the Kolmogorov–Smirnov metric (see Borgonovo et al. (2014) for the proof). One also obtains transformation invariant sensitivity measures using inner statistics based on divergences between densities, as we have discussed in the previous section. Historically, the first transformation-invariant sensitivity measure is the Spearman coefficient (Spearman, 1904). Clearly, ranks or empirical cumulative distribution functions are natural monotonic transformations. This idea directly leads to bivariate copulas. Now, copula theory is sometimes described as the study of transformation-invariant properties. In particular, the δ measure is discussed in a copula framework in Wei, Lu, and Song (2014).

Monotonic transformation invariance has also notable decision analysis implications. In several applications, the preferences of a decision maker should be expressed not through model output Y itself, but through a utility function on Y, $u(Y)$. It is well known that assessing the exact functional form of $u(Y)$ might become a cumbersome task in complex problems. However, the traditional assumption that $u(\cdot)$ is a monotonically increasing function of its arguments, can be exploited to obtain a notable simplification of sensitivity analysis. If we employ a monotonic transformation invariant sensitivity measure, the obtained results hold identical for Y and $u(Y)$, provided that $u(\cdot)$ is a monotonic function. In particular, results are unaffected by errors in the specification of $u(Y)$ (Baucells & Borgonovo, 2013). We conclude this section observing that, to our knowledge, the literature has focused on transformation invariance within a global sensitivity analysis context. However, transformation invariance might be useful in a local sensitivity as well. Nonetheless, none of the local sensitivity measures discussed in this review possesses the transformation invariance property.

4.6. Value of information based sensitivity measures

The operational research literature has pointed out the distinction between *value* and *decision* sensitivity. The distinction is discussed in Felli and Hazen (1998). In a value sensitivity, the analyst is interested in quantifying the change in model output due to a change in the model inputs. In a decision sensitivity, the analyst is also interested in assessing whether the optimal plan (or strategy) changes. Thus, sensitivity measures such as variance-based, density-based and distribution-based are value sensitivity measures. However, in several decision-analysis problems the decision-support models explicitly consider the selection among a set of D possible choices. D can be a continuous or discrete set, however, we shall consider a discrete set of alternatives, for notation simplicity. Suppose that we are considering alternative inventory management policies. Then, we estimate the payoff for each policy through the dedicated model. We then select the policy that maximizes the payoff. In general, let y_d ($d = 1, 2, \ldots, D$) be the estimate produced by the model of the quantity of interest (payoff, loss, utility or whatever criterion the decision maker chooses to support her analysis) under policy d. Then, the model input–output mapping becomes $Y_d = g_d(\mathbf{X})$, $d = 1, 2, \ldots, D$. Then, the problem solved by the decision-maker is

$$\max_{d=1,2,\ldots,D}\{\mathbb{E}[Y_d]\} = \max_{d=1,2,\ldots,D}\{\mathbb{E}[g_d(\mathbf{X})]\}. \tag{40}$$

where $\max_{d=1,2,\ldots,D}\{\mathbb{E}[Y_d]\}$ is the expected payoff.

In this context, the effect of getting to know X_i is not only a change in the distribution of each of the Y_d's, but, also, a change in the optimal plan. The sensitivity measures we have discussed so far do not capture this change. The sensitivity measure that captures both decision and value sensitivity is the partial expected value of perfect information (Felli & Hazen, 1998; Oakley, 2009; Strong & Oakley, 2013). Value of information has been introduced in Howard (1966), as *the difference between the expected profit that we shall obtain as a result of the clairvoyance and the expected profit that we would obtain without the clairvoyance* on X_i. Formally, one writes

$$\epsilon_i = \mathbb{E}\big[\max_j \{Y_j\} \mid X_i\big] - \max_j \big\{\mathbb{E}[Y_j(\mathbf{X})]\big\}. \tag{41}$$

Several properties of value of information have been discussed in the literature. We refer the interested reader to the monographs of Bernardo and Smith (2000) and Pratt, Raiffa, and Schlaifer (1995) and to the work of Hazen and Sounderpandian (1999). A notable property proven in Howard (1988) is that the joint value of information of two random variables is given by:

$$\epsilon_{i,j} = \epsilon_i + \epsilon_{j|i} = \epsilon_j + \epsilon_{i|j} \tag{42}$$

Eq. (42) suggests that the joint value of information of X_i and X_j is the sum of the value of information of X_i (X_j) and the value of information of X_j (X_i) conditional on getting to know X_i (X_j). Value of information has been applied in the most diverse contexts. Properties of value of information in stochastic programming are discussed in Avriel and Williams (1970), and sharp bounds are obtained in Huang, Vertinsky, and Ziemba (1977). In these works, the total value of information for resolving completely uncertainty in the model inputs is addressed. Partial value of information is proposed as a sensitivity measure in Felli and Hazen (1998) in a medical decision making context. It is also suggested as a sensitivity measure for decision-making in reliability by Pörn (1997). Recently, the works of Oakley (2009), Strong and Oakley (2013), Strong and Oakley (2014) address the estimation of value of information for complex computational codes.

4.7. Monte Carlo filtering techniques

Monte Carlo filtering (Rose, Smith, Gardner, Brenkert, & Bartell, 1991) is a method which allows us to sort out model runs that fail to meet some predefined performance criteria. In the engineering literature such an approach is discussed under the label of reliability

Table 1
Examples of inner statistics and associated sensitivity measures.

Inner statistic	Sensitivity measure
$\max_{d=1,\ldots,D}\{\mathbb{E}[Y_d\|X_i]\} - \max_{d=1,\ldots,D}\{\mathbb{E}[Y_d]\}$	ϵ_i
$\mathbb{V}[Y]^{-1}(\mathbb{E}[Y\|X_i] - \mathbb{E}[Y])^2$	S_i
$\sup \|F_Y - F_{Y\|X_i}\|$	β_i^{KS}
$\sup (F_Y - F_{Y\|X_i}) - \inf (F_Y - F_{Y\|X_i})$	β_i^{Ku}
$\frac{1}{2} \int \|f_{Y\|X_i}(y) - f_Y(y)\|dy$	δ_i

methods for structural safety (Madsen, Krenk, & Lind, 1986; Rackwitz, 2001). Here, the failure domain is given by $\mathcal{X}_{\text{fail}} = \{\mathbf{x} \in \mathcal{X} : g(\mathbf{x}) \leq 0\}$ and one is interested in the failure probability $\mathbb{P}_{\mathbf{X}}(\mathcal{X}_{\text{fail}}) = \mathbb{E}[g(\mathbf{x}) \leq 0] = F_Y(0)$. In applications, the limit state function is composed of load and resistance parts, $g(\mathbf{x}) = R(\mathbf{x}) - L(\mathbf{x})$ where the load function models external stresses and the resistance function is derived from internal material and design parameters. The FORM/SORM methods compute a most probable point of failure on the limit state surface and approximate the limit state function by a linear or quadratic curve through this point to estimate the failure probability. The coordinates of the most probable point of failure are used as parameter sensitivity measures. A link to value-of-information based sensitivity is established in Straub (2014).

The regionalized sensitivity analysis of Hornberger and Spear (1981) is comparing the conditional input distributions $F_i(x_i|Y \leq 0)$ and $F_i(x_i|Y > 0)$ using a Kolmogorov–Smirnov test. If instead of 0, the failure region is described by an output quantile, then a graphical tool to study Monte Carlo filtering is the conditional cobweb plot. Due to the symmetry property in Eq. (30), we can also interpret δ as MC filtering method reporting the average change in the input distribution when conditioning on the output.

5. Estimation: review and a common rationale

In this section, we discuss the computation of sensitivity measures. Variance-based, density-based and value of information-based sensitivity measures have a common trait. They can be seen as operators between probability distributions Borgonovo, Hazen, and Plischke (2015). Let $\mathbb{P}$ be a generic distribution and let $\zeta(\mathbb{P}, \mathbb{P})$ an operator between distributions that satisfies the condition $\zeta(\mathbb{P}, \mathbb{P}) = 0$. Then, one can represent a global sensitivity measure as composed of an inner statistic

$$\gamma_i(x_i) = \zeta(\mathbb{P}_{\mathbf{Y}}, \mathbb{P}_{\mathbf{Y}|X_i=x_i}) \tag{43}$$

whose value depends on the value assumed by X_i and an outer expectation. The inner statistic as a function of $x_i \in \mathcal{X}_i$ provides regional sensitivity information. To obtain a global sensitivity measure, one then needs to carry out an outer expectation over the marginal distribution of model input X_i:

$$\xi_i = \mathbb{E}[\gamma_i(X_i)]. \tag{44}$$

Eq. (44) is the global sensitivity measure of X_i based on operator $\zeta(\cdot, \cdot)$. To illustrate, Table 1 shows the inner statistics associated with several global sensitivity measures resting on this common rationale.

This table reads as follows. If we select as inner operator $\gamma_i^{\text{EI}}(X_i) = \max\{\mu_{Y|X_i}, 0\} - \max\{\mu_Y, 0\}$ and average, we obtain the value of information global sensitivity measure. If we select $\gamma_i^{\text{CR}}(X_i) = \sigma_Y^{-2}(\mu_Y - \mu_{Y|X_i})^2$, and average, we obtain variance-based sensitivity measures, etc. Based on this common rationale, it is possible to address the properties of global sensitivity measures from a general perspective. For instance, can prove that a sensitivity measure is transformation invariant if and only if the inner operator is transformation invariant (Borgonovo et al., 2015).

A first implication of the common rationale is that it opens the way to considering global sensitivity measures as measures of statistical dependence. Recently, Da Veiga (2015) proposes measures

of statistical dependence based on the distance correlation and on the Hilbert–Schmidt independence criterion. These measures are also discussed in Lozzo and Marrel (2014). Fort, Klein, and Rachdi (2014) discuss the definition of sensitivity measures as subordinated to the machine learning notion of contrast. A second implication is that all global sensitivity measures falling under the common rationale can be estimated through the same computational design. This observation then leads us to discuss computational aspects.

Sensitivity analysis and design of experiment regard the computational code at hand as a black box. Then, sensitivity measures need to be estimated numerically. By Eq. (44), an algorithm following the brute force definition of global importance measures requires a double loop of Monte Carlo evaluations. In fact, we need to take an "outer" expectation according to the marginal distribution of X_i of the inner statistic. The inner statistic in Eq. (43), in turn, requires a set of Monte Carlo runs to obtain numerical estimates of the unconditional distribution of the model output and of the conditional distribution given that X_i is fixed at x_i. Numerically, we first sample N_{ext} values from the marginal distribution of model input X_i. Then, for each sampled value of X_i, we run the model N_{int} times to estimate the inner statistic. Thus, the computational cost for a brute force estimation is, in principle, $C = n N_{int} N_{ext}$ where n is the number of model inputs and N_{int}, N_{ext} are the sample sizes needed for the inner and outer loops, respectively. Assuming that $n = 100$, and at least 500 Monte Carlo runs are needed to obtain accurate estimates of the inner statistic and outer expectation, one obtains a total cost of $C = 25{,}000{,}000$. Such cost would make the estimation of probabilistic sensitivity measures impossible. The literature has then addressed the problem proposing a series of ameliorations.

Much of the focus has been on the estimation of variance-based sensitivity measures. Here, we need to distinguish estimation strategies based on whether they require a specific design or not. The reason is practical. In the latter case, the analyst can use the results of a simple Monte Carlo propagation or the dataset of the model input–output generated through any scheme. In the former case, she needs to run the model according to a specific procedure and cannot profit or reuse model runs generated using a different scheme. In the first category, for variance-based sensitivity measures, we find the design of Sobol' (1993), Jansen (1999), the design of Owen (2013), the FAST design Cukier et al. (1973); Saltelli, Chan, and Scott (2000a) and the random balance design (RBD) (Tarantola, Gatelli, & Mara, 2006). The design of Sobol' and Jansen, improved in Homma and Saltelli (1996); Saltelli (2002a); Saltelli, Annoni, Azzini, Campolongo, Ratto, and Tarantola (2010), reduces the computational cost for first and total order sensitivity measures to $C = (k + 2)N$.

The Fourier Amplitude Sensitivity Test (FAST) design requires a particular generation of the inputs, based on a space filling curve that needs the assignment of different frequencies to the model inputs. The resulting output signal is then screened for resonances of these frequencies. The magnitude of these resonances accounts for the sensitivity of the associated model input. The FAST method allows for the estimation of variance based sensitivity indices at a cost of $C = 1 + 2M \max_{i=1,\dots,n} \omega_i$ where M is the number of harmonics to consider (usually 4 to 8) and ω_i are the frequencies of the curve. For the choice $\omega_i = \omega_0^{i-1}$ with $\omega_0 = 2M + 1$, contributions from all interaction terms can be identified. There are algorithms available that produce suitable sets of incommensurable frequencies which ignore higher interaction terms so that the computational costs can be reduced further.

RBD uses $\omega_i \equiv 1$ but employs different permutations for different input factors. Applying the appropriate permutation on the output, one is able to detect the resonances due to the associate input factor. Hence this method allows for the estimation of first order effects at the cost of $C = N$, which is the minimum computational cost for a probabilistic sensitivity measure, that is, the size of a single Monte Carlo sample.

Recent findings have shown that it is possible to achieve this estimation cost without recurring to a specific design, and using a so-called given data approach. For variance-based sensitivity, the EASI method can be thought of a reverse-logic RBD: The permutations are constructed from the given data (Plischke, 2010). The COSI method (Plischke, 2012b) simplifies this approach by replacing the fast Fourier transformation with a discrete cosine transformation.

The intuition behind a given data approach can be illustrated as follows. We consider the dataset $(X, Y) \in \mathbb{R}^{N \times (n+1)}$ generated through a single Monte Carlo run. As we mentioned in Section 2, this is the dataset produced in an uncertainty analysis and from which we obtain the unconditional distribution of the model output. Then, from this dataset, we can produce the scatterplot of X_i and Y. The horizontal plane in Fig. 5 displays the scatterplot obtained by projecting Y against X_1 for the Harris EOQ model. The dataset is produced using a quasi-random sequence with size $N = 500$.

Then, we partition the X_1-axis of the scatterplot, on the horizontal plane. Formally, we partition the support of X_i in M mutually exclusive subsets $\mathcal{X}_i^m$ ($m = 1, 2, \dots, M$), such that $\bigcup_{m=1}^{M} \mathcal{X}_i^m = \mathcal{X}_i$, $\mathcal{X}_i^m \cap \mathcal{X}_i^q = \emptyset$ ($m \neq q$). Then, we substitute the point condition $X_i = x_i$ with the bin condition $X_i \in \mathcal{X}_i^m$. Intuitively, instead of imposing that X_i is exactly equal to x_i, we are asking that X_i is in a partition bin that surrounds x_i. Formally, we consider the bin conditional distribution $F_{Y|X_i \in \mathcal{X}_i^m}(y)$, instead of $F_{Y|X_i=x_i}(y)$ (vertical axis in Fig. 5). One obtains the generic given data estimator:

$$\widehat{\xi}_i = \sum_{m=1}^{M(N)} \frac{N_m}{N} \zeta \left(\widehat{F}_Y, \widehat{F}_{Y|X_i \in \mathcal{X}_i^m(N)} \right) \tag{45}$$

where the symbol $M(N)$ denotes the fact that the choice of the partition size might vary with the sample size and N_m denotes the number of realizations of X_i falling into $\mathcal{X}_i^m$. To illustrate, a given data estimator for value of information is written as (Strong & Oakley, 2013):

$$\widehat{\epsilon}_i = \sum_{m=1}^{M(N)} \frac{N_m}{N} \max_{a=1,2,\dots,A} \left\{ \frac{1}{N_m} \sum_{\ell:x_{\ell i} \in \mathcal{X}_i^m} y_{\ell,a} \right\} - \max_{a=1,2,\dots,A} \left\{ \frac{1}{N} \sum_{i=1}^{N} y_{i,a} \right\} \tag{46}$$

where the expression $\max_{a=1,2,\dots,A} \left\{ \frac{1}{N_m} \sum_{\ell:x_{\ell i} \in \mathcal{X}_i^m} y_{\ell,a} \right\}$ is an estimator of the posterior value of action given that X_i is contained in the mth bin, namely, an estimator of $\max_{a=1,2,\dots,A} \mathbb{E}[Y_a | X_i \in \mathcal{X}_i^m]$. Several properties of this estimator are discussed in Strong and Oakley (2013). For variance-based sensitivity measures, the given data estimator is the one defined by Pearson (1905):

$$\widehat{v}_i = \sum_{m=1}^{M(N)} \frac{N_m}{N} \left(\frac{1}{N_m} \sum_{\ell:x_{\ell i} \in \mathcal{X}_i^m} y_\ell - \frac{1}{N} \sum_{i=1}^{N} y_i \right)^2. \tag{47}$$

This estimator is used in Strong, Oakley, and Chilcott (2012) and Plischke et al. (2013) for the estimation of variance-based sensitivity measures. Recently, Zhai, Yang, and Zhao (2014b) discuss additional analytic properties of this estimator and discuss heuristics for the optimal partition size.

Plischke et al. (2013) discuss the given data estimation of density-based sensitivity measures for a space mission risk assessment model with 827 inputs. Baucells and Borgonovo (2013) use a given data estimation approach to conduct the global sensitivity analysis of the well known Genzyme–Geltex pharmaceutical case study (Bruner, Bodily, & Jacquet, 1999), specializing Eq. (45) for transformation invariant inner statistics based on the Kuiper metric. As these recent references show, given data approaches are attracting increasing attention, because of the possibility of delivering sensitivity measures without a specific design and using directly the Monte Carlo sample produced in an uncertainty analysis.

Indeed, recently, Borgonovo et al. (2015) show that, under the requirements that the inner statistic $\gamma(X_i)$ is a Riemann–Stieltjes
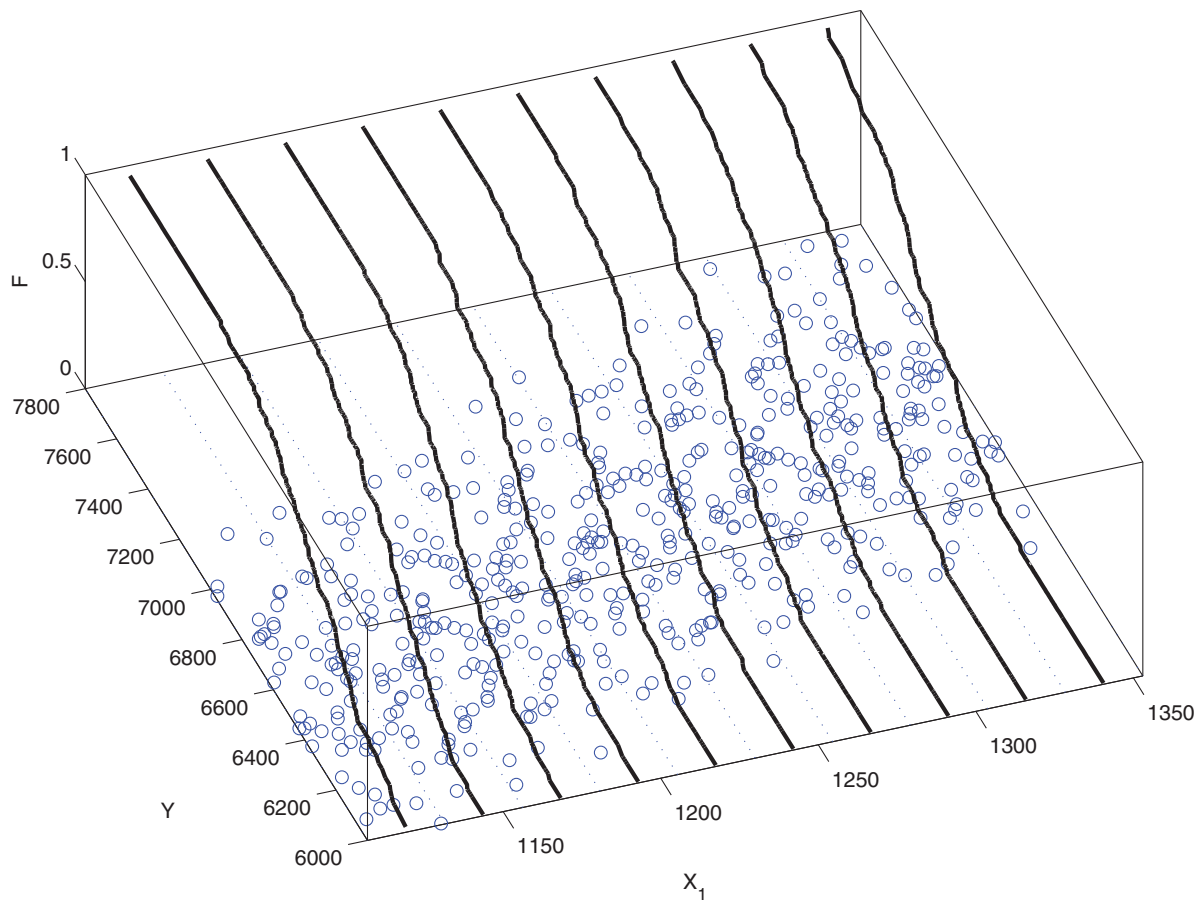
Fig. 5. Scatterplot of X_1 vs. Y for the Harris EOQ overlaid with conditional empirical CDFs from a partition of size 10.

integrable function and that the inner operator $\zeta(\cdot, \cdot)$ is a continuous function of its arguments, as N tends to infinity, a given data estimator in Eq. (45) tends to the corresponding value of the probabilistic sensitivity measure. This result provides a general envelope for the estimation of probabilistic sensitivity measures using given data approaches.

A further way to estimate global sensitivity measures at the minimal cost of N model runs is to fit a meta-model (or emulator) on the input–output sample and then use the emulator to replace the original model. The realm of model emulation is vast and outside the scope of our article. However, we refer the interested reader to the monograph of Santner, Williams, and Notz (2003), to the 2009 Technometrics special issue edited by Bayarri, Berger, and Steinberg (2009), and to the reviews in Oakley and O'Hagan (2004) for Gaussian process meta-modeling, in Ratto, Pagano, and Young (2007) for smoothing spline ANOVA meta-models, to Kleijnen (2008); 2009) for Kriging meta-models, to Sudret (2008) for polynomial chaos expansion, and to Buzzard (2012) for sparse grid interpolation. Provided that an accurate fit is obtained, given the very short running time of an emulator, one can then apply a brute force algorithm or can use specific designs. This approach is used extensively in the estimation of variance-based sensitivity measures, with works such as (Blatman & Sudret, 2010; Doksum & Samarov, 1995; Lewandowski, Cooke, & Duintjer Tebbens, 2007; Sudret, 2008) . Recently, it is used in Strong and Oakley (2014) for the estimation of value of information sensitivity measures using Gaussian Process meta-modeling. It is used in Borgonovo, Castaings, and Tarantola (2012) to estimate density-based sensitivity measures.

6. Sensitivity insights

Sensitivity analysis is essential to bridge the gap between decision makers (that often do not directly work or develop the model) and analysts (who work directly on the model). However, as (French, 2003, p. 229) states, *... sensitivity methods are used in many different circumstances for many different purposes; and the manner of their use in one context may be inappropriate in another.*

The literature has demonstrated that it is crucial to state up front the questions that we wish to answer through a sensitivity analysis (Saltelli, 2002b). If this is the case, then one can confidently identify the appropriate sensitivity analysis method for answering the stated questions. To illustrate, consider an analyst wishing to identify key-drivers of model response. If the question is answered under uncertainty, then a global sensitivity method is appropriate for the purpose. If the question is asked at a particular point in the model input space, then a local method is needed. The works of Saltelli and Tarantola (2002), Saltelli (2002b) formalize this aspect introducing the concept of sensitivity analysis setting. A setting is a formulation of the sensitivity analysis quest which is stated before the sensitivity analysis is carried out. Several insights can be sought by an analyst and sensitivity analysis settings have been introduced in the literature.

The *factor prioritization* setting is associated with obtaining insights about the most important model inputs. Best practices recommend that, under uncertainty, these insights are obtained using a global sensitivity method. If the goal of the analysis if variance-reduction, then variance-based sensitivity measures should be used.

If the goal is the determination of the effect of a model input on the entire distribution (density), then distribution (density) based method are appropriate for the task. Conversely, in a deterministic framework (i.e., in the absence of uncertainty), and if is considering finite changes in the model inputs, Tornado diagrams are the appropriate tools. Conversely, if infinitesimal perturbations are of concern, then differentiation-based sensitivity measures are appropriate.

A second well known setting is *factor fixing*. In this case, the analyst is interested in receiving information about the least influential factors, which can be confidently fixed at their nominal value. Screening methods are the appropriate tools for this task (Saltelli et al., 2008). A further setting concerns *model structure* (Borgonovo, 2010b). This setting applies in all those situations in which analysts are interested in determining whether the model response is the sum of the individual model input changes or whether interactions play a relevant role. The sensitivity measures appropriate for this setting are finite change sensitivity indices ϕ_i^l, if we are in a deterministic framework, or high-order variance based sensitivity indices, if we are in a probabilistic framework.

Direction (or sign) *of change* is a crucial insight in several applications. As Samuelson (1941) states: *This is a typical problem confronting the economist: in the absence of precise quantitative data he must infer analytically the qualitative direction of movement of a complex system.* That is, a direction of change setting envelops all those applications in which analysts are interested to know whether an increase (decrease) in a model input provokes an increase (decrease) in model output. The question may be asked deterministically or under uncertainty. If infinitesimal perturbations are of concern—as in the original work of Samuelson—the sign of partial derivatives provides the desired answer. If finite changes are of concerns, then the sign of the finite change sensitivity indices provide indication about direction of change. If the analyst is interested in the behavior of the model output when one of the inputs varies in a predetermined range, then one-way sensitivity functions are of interest. Finally, the sign of change question can be also asked in a global way, i.e., under uncertainty. Then, the answer to this question is the plot of the first order terms of the functional ANOVA expansion, namely the functions $g_i(x_i)$ in Eq. (26), provides the answer. In fact, if $g(\mathbf{x})$ is monotonic, the first order effect functions $g_i(x_i)$ retain the monotonicity of $g(\mathbf{x})$, as proven in Beccacece and Borgonovo (2011). Anderson, Borgonovo, Galeotti, and Roson (2014) apply these findings in a climate change application.

Besides these classical settings, additional questions have been stated in the literature, although possibly not formalized as a setting. In particular, we propose here to group several sensitivity studies in a *stability*[1] *setting.* This setting indeed comprises all those problems in which decision-makers are interested in determining the region of the model input space over which the optimal plan (or preferred alternative) does not change. It applies in sensitivity analysis of linear programming problems, where the analyst is interested in determining the ranges in the model inputs which leave the optimal plan unaltered. The method that delivers a consistent answer to this question is Wendell's tolerance sensitivity approach (Ravi & Wendell, 1989; Wendell, 1985). We refer to Filippi (2005) and Filippi (2010) for thorough reviews. It applies as well in the sensitivity analysis of decision support models expressed in the form of decision-trees (influence diagram), that help analysts in selecting among a discrete set of alternatives. The work of Nielsen and Jensen (2003) addresses the problem of finding the variation ranges in utilities and probabilities (the model inputs) that reassure us that the preferred alterna-

tive is stable, and the same sensitivity question is answered also in the works of Bhattacharjya and Shachter (2010) and Bhattacharjya and Shachter (2008). Stability analysis for Bayesian networks is discussed in Chan and Darwiche (2004), Chan and Darwiche (2005) and Brosnan (2006). A further important class of problems to which a *stability setting* applies is dynamic programming. Tan and Hartman (2011) provide a thorough review of sensitivity methods used in discrete dynamic programming to address the sensitivity of the optimal solution to changes in the model inputs. Tan and Hartman (2011, Section 5) observe that several discrete dynamic models can be represented as linear programs. One can then exploit tolerance sensitivity approaches to obtain stability insights.

7. Discussion

In this section, we wish to enter a bit further into the discussion of topics which are associated with sensitivity analysis. The discussion definitely deserves more space than a section, so, please consider this as a synthetic exposition of the main concepts.

The first aspect concerns the interpretation of probability in a sensitivity analysis. We have seen in Section 4 that the some authors favor the use of deterministic methods when state of knowledge is not enough to warrant the assessment of probabilities. Other authors, have started exploring the use of sensitivity analysis in conjunction with non-probabilistic approaches (Helton, Johnson, Oberkampf, & Sallaberry, 2006a). Nonetheless, the majority of the works in global sensitivity analysis rely on the assumption stated in Section 4.1, i.e., analysts are capable of assigning probability distributions. Then, one can question whether these distributions are objective or subjective, i.e., whether they reflect the analyst's personal degree of belief or whether the amount of data is such that a frequentist approach can be used. This question links strictly to the separation between aleatory and epistemic uncertainty. The literature has investigated this separation extensively, and we refer to works such as Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio (2013); Ellsberg (1961); Klibanoff, Marinacci, and Mukerji (2005) discussing the axiomatic implications of removing or substituting the axioms in Savage (1954). The works of Apostolakis (1990), Hoffman and Hammonds (1994), Paté-Cornell (1996), Winkler (1996), Parry (1996), Aven (2010), Helton and Johnson (2011) have examined the implications of distinguishing aleatory and epistemic uncertainty for risk-informed decision making. We try and illustrate the distinction. Aleatory uncertainty concerns propositions of the type "event E will occur" (de Finetti, 1937). Then, we write $Pr[E] = p$. If the process of calculating p does not lead to a unique p, then the "mental process of computing p" becomes in itself a random event [(Pearl, 1988)]. We then have epistemic uncertainty and a second order distribution that concerns events of the type "the probability of event E is p". In Economics, lack of knowledge in the probabilities of a problem is called ambiguity. Then, aleatory uncertainty is represented by a first order distribution and epistemic uncertainty by a second order distribution (Hansen, 2007).

In the context of global sensitivity analysis, this distinction is usually omitted. The reason is a practical one: distributions over plausible uncertainty ranges are often assigned with the goal of investigating how the model behaves and not necessarily of accurately reproducing the decision-maker's (or analyst's) degree of belief among the model inputs. Because the primary goal of a global sensitivity analysis is, in a sense, "mathematical", then whether uncertainty is aleatory or epistemic becomes a secondary problem. Moreover, the problem in sensitivity analysis of model output is made even more complicated by the fact that, as Der Kiureghian and Ditlevsen (2009) rightly observe, uncertainty is assigned to inputs of a mathematical model and these inputs might relate to alternative types of uncertainty. To illustrate, consider using value of information as a sensitivity measure. If we are carrying out a value of information analysis

---

[1] The setting could also be named *robustness setting.* However, this might create overlap with the use of the term robustness in optimization, where robust optimization refers to the solution of optimization problems characterized by uncertainty in distribution. The setting has not been officially named in the literature, yet, but it unifies several sensitivity questions and models of operations research.

in a decision-tree that represents a decision-problem under risk (i.e., probabilities are known) embedded in a Savage expected utility theory framework, then we are performing a sensitivity analysis on propositions of the type "how much I am willing to spend to know whether event E will occur". This type of sensitivity touches upon an aleatory uncertainty question, with the model embedded in a subjective framework. Conversely, suppose we are performing a value of information analysis on a model input of a complex code model. In this case, our statements are of the type "the value of X_i is in between $\bar{X}_i - \Delta$ and $\bar{X}_i + \Delta$". The interpretation here is less crisp. The statement of concern is not about an event as we mean it in common linguistic sense. However, if we are willing to assign a unique distribution, then we have a unique layer of uncertainty and then only first order (or aleatory) uncertainty appears. That this uncertainty is aleatory, however, might seem counterintuitive, because X_i is indeed a parameter and we usually regard parametric uncertainty as epistemic.

Thus, the characterization of uncertainty in a sensitivity analysis exercise is dependent both on the state of knowledge but also on the type of uncertainty addressed in the model. Thus, a thorough analysis would require studying the alternative types of models and the uncertainties that these models capture. Such a study is, however, outside the scope of the present work.

A second relevant aspect is the role of sensitivity analysis in model building. Felli and Hazen (2004, p. 93) suggest that *"sensitivity analysis can be used to guide model development: If decisions are insensitive to changes in some aspect of the model, then there is no need to model that particular aspect in more detail."* Saltelli, Tarantola, and Campolongo (2000b) maintain that sensitivity analysis can be used in the various phases of the modeling process, in model identification, model calibration and model corroboration and offer an example of the use of sensitivity analysis in a model calibration exercise. French (2003) offers a systematic discussion. In particular, French (2003) argues that sensitivity analysis can play a role in the initial construction phases, when analysts are interested in understanding the internal working of the model and also in the validation phase of the model, when analysts are interested in validating the model against observed data for fitting. Bayarri, Berger, Paulo, Sacks, Cafeo, Cavendish, Lin, and Tu (2007) present a thorough discussion of model validation and evidence the role that sensitivity analysis can play in the process, with particular emphasis on model simplification. Recently, Strong and Oakley (2014) address specifically how sensitivity analysis can be used to rank the internal sources of discrepancy in model validation. The approach in Strong and Oakley (2014) is based on value of information.

8. Conclusions

The development of a simulation or decision support model aids the solution of several problems in operational research. Numerical results of the codes inform decision makers in applications which might be crucial for the life of a corporation or might have strong societal and environmental impacts, especially when complex technological systems are of concern. We are then called to a systematic utilization of sensitivity analysis methods for: *(a)* making model builders fully aware of the response of the model to variations and uncertainty in the model inputs and, *(b)* providing decision makers with an enriched set of insights about the managerial problem at hand. We have focused on quantitative and model free methods, ranging from the widely used local techniques (from Tornado diagrams to finite change sensitivity indices, to differential operators), to more recently introduced global sensitivity methods, such as non-parametric, variance-based, moment independent and value of information-based sensitivity measures. We have presented recent advances in the theoretical framework and estimation procedures. These advances are capable of notably reducing computational burden, making the use of global methods at reach in an increasing number of applications. It

then becomes crucial to improve communication of sensitivity analysis results to managers. Our second last section has then been devoted to an approach for systematically obtaining managerial insights. Our discussion section has concerned the interpretation of probabilities in a global sensitivity analysis and the role of sensitivity analysis in model building.

References

Alış, Ö. F., & Rabitz, H. (2001). Efficient implementation of high-dimensional model representations. *Journal of Mathematical Chemistry, 29*(2), 127–142.

Anderson, B., Borgonovo, E., Galeotti, M., & Roson, R. (2014). Uncertainty in climate change modelling: Can global sensitivity analysis be of help? *Risk Analysis, 34*(2), 271–293.

Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *The Annals of Mathematical Statistics, 23*(2), 193–212.

Apostolakis, G. E. (1990). The concept of probability in safety assessment of technological systems. *Science, 250*, 1359–1364.

Armstrong, M. J. (1995). Joint reliability-importance of elements. *IEEE Transactions on Reliability, 44*(3), 408–412.

Auder, B., & Iooss, B. (2009). Global sensitivity analysis based on entropy. In *Safety, reliability and risk analysis: Theory, methods and applications* (pp. 2107–2115). London: Taylor & Francis.

Aven, T. (2010). On the need for restricting the probabilistic analysis in risk assessments to variability. *Risk Analysis, 30*, 354–360.

Aven, T., & Nøkland, T. E. (2010). On the use of uncertainty importance measures in reliability and risk analysis. *Reliability Engineering & System Safety, 95*(2), 127–133.

Avriel, M., & Williams, A. C. (1970). The value of information and stochastic programming. *Operation Research, 18*(5), 947–954.

Baucells, M., & Borgonovo, E. (2013). Invariant probabilistic sensitivity analysis. *Management Science, 59*(11), 2536–2549.

Bayarri, M. J., Berger, J., & Steinberg, D. M. (2009). Special issue on computer modeling. *Technometrics, 51*(4), 353.

Bayarri, M. J., Berger, J. O., Paulo, R., Sacks, J., Cafeo, J. A., Cavendish, J., Lin, C.-H., & Tu, J. (2007). A framework for validation of computer models. *Technometrics, 49*(2), 138–154.

Beccacece, F., & Borgonovo, E. (2011). Functional ANOVA, ultramodularity and monotonicity: Applications in multiattribute utility theory. *European Journal of Operational Research, 210*(2), 326–335.

Bedford, T. (1998). Sensitivity indices for (tree)-dependent variables. In *Proceedings of the second international symposium on sensitivity analysis of model output* (pp. 17–20). Venice (Italy).

Bernardo, J. M., & Smith, A. F. M. (2000). *Bayesian theory*. Chichester: John Wiley&Sons.

Bertsimas, D., & Sim, M. (2004). The price of robustness. *Operations Research, 52*(1), 35–53.

Bettonvil, B. (1990). *Detection of important factors by sequential bifurcation*. Tilburg: Tilburg University Press.

Bettonvil, B., & Kleijnen, J. P. C. (1997). Searching for important factors in simulation models with many factors: Sequential bifurcation. *European Journal of Operational Research, 96*(1), 180–194.

Bhattacharjya, D., & Shachter, R. (2008). Sensitivity analysis in decision circuits. In *Proceedings of the twenty-fourth annual conference on uncertainty in artificial intelligence (UAI-08)* (pp. 34–42). Corvallis, OR: AUAI Press.

Bhattacharjya, D., & Shachter, R. (2010). Three new sensitivity analysis methods for influence diagrams. In *Proceedings of the twenty-sixth annual conference on uncertainty in artificial intelligence (UAI-10)* (pp. 56–64). Corvallis, OR: AUAI Press.

Bielza, C., & Shenoy, P. P. (1999). A comparison of graphical techniques for asymmetric decision problems. *Management Science, 45*(11), 1552–1569.

Birnbaum, Z. W. (1969). On the importance of different components in a multicomponent system. In *Multivariate analysis: II* (pp. 581–592). Academic Press.

Bischof, C. H., & Bücker, H. M. (2000). Computing derivatives of computer programs. In *Proceedings modern methods and algorithms of quantum chemistry* (pp. 1–13). Argonne National Laboratory.

Blatman, G., & Sudret, B. (2010). Efficient computation of global sensitivity indices using sparse polynomials. *Reliability Engineering & System Safety, 95*(11), 1216–1229.

Bolado-Lavin, R., Castaings, W., & Tarantola, S. (2009). Contribution to the sample mean plot for graphical and numerical sensitivity analysis. *Reliability Engineering & System Safety, 94*(6), 1041–1049.

Borgonovo, E. (2006). Measuring uncertainty importance: Investigation and comparison of alternative approaches. *Risk Analysis, 26*(5), 1349–1362.

Borgonovo, E. (2007). A new uncertainty importance measure. *Reliability Engineering & System Safety, 92*(6), 771–784.

Borgonovo, E. (2010). The reliability importance of components and prime implicants in coherent and non-coherent systems including total-order interactions. *European Journal of Operational Research, 204*(3), 485–495.

Borgonovo, E. (2010). Sensitivity analysis with finite changes: An application to modified EOQ models. *European Journal of Operational Research, 200*(1), 127–138.

Borgonovo, E. (2013). Sensitivity analysis in decision-making. Encyclopedia of operations research and management science (pp. 1–12).

Borgonovo, E., & Apostolakis, G. E. (2001). A new importance measure for risk-informed decision making. *Reliability Engineering & System Safety, 72*(2), 193–212.

Borgonovo, E., Castaings, W., & Tarantola, S. (2012). Model emulation and moment-independent sensitivity analysis: An application to environmental modelling. *Environmental Modelling & Software, 34*, 105–115.

Borgonovo, E., Hazen, G. B., & Plischke, E. (2015). A common rationale for globalsensitivity measures and their estimation. *Risk Analysis* (submitted for publication).

Borgonovo, E., & Peccati, L. (2004). Sensitivity analysis in investment project evaluation. *International Journal of Production Economics, 90*(1), 17–25.

Borgonovo, E., & Smith, C. L. (2011). A study of interactions in the risk assessment of complex engineering systems: An application to space PSA. *Operations Research, 59*(6), 1461–1476.

Borgonovo, E., Tarantola, S., Plischke, E., & Morris, M. D. (2014). Transformations and invariance in the sensitivity analysis of computer experiments. *Journal of the Royal Statistical Society, Series B, 76*, 925–947.

Brewer, H. D., & Canady, K. S. (1999). Probabilistic safety assessment support for the maintenance rule at duke power company. *Reliability Engineering & System Safety, 63*(3), 243–249.

Brosnan, A. J. (2006). Sensitivity analysisof a Bayesian belief network in a tactical intelligence application. *Journal of Battlefield Technology, 9*(2), 33–40.

Bruner, R. F., Bodily, S. E., & Jacquet, P. (1999). Genzyme/Geltex pharmaceuticals joint venture. Technical report, case no. F-1254. Charlottesville, VA: University of Virginia Darden School Foundation.

Buzzard, G. T. (2012). Global sensitivity analysis using sparse grid interpolation and polynomial chaos. *Reliability Engineering & System Safety, 107*, 82–89.

CADTH (2006). Guidelines for the economic evaluation of health technologies: Canada. Technical report. Ottawa: Canadian Agency for Drugs and Technologies in Health. www.cadth.ca/media/pdf/186_EconomicGuidelines_e.pdf

Caflisch, R. E., Morokoff, W., & Owen, A. B. (1997). Valuation of mortgage backed securities using Brownian bridges to reduce effective dimension. *Journal of Computational Finance, 1*, 27–46.

Campolongo, F., Cariboni, J., & Saltelli, A. (2007). An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software, 22*, 1509–1518.

Campolongo, F., Kleijnen, J. P. C., & Andres, T. (2000). Screening methods (pp. 65–80). In Saltelli et al. (2000a).

Campolongo, F., & Saltelli, A. (1997). Sensitivity analysis of an environmental model: An application of different analysis methods. *Reliability Engineering & System Safety, 57*(1), 49–69.

Cárdenas-Barrón, L. E., Chung, K.-J., & Trevi no-Garza, G. (2014). Editorial: Celebrating a century of the economic order quantity model in honor of Ford Whitman Harris. *International Journal of Production Economics, 155*, 1–7.

Castaings, W., Borgonovo, E., Morris, M. D., & Tarantola, S. (2012). Sampling strategies in density-based sensitivity analysis. *Environmental Modelling & Software, 38*, 13–26.

Castillo, E., Gutiérrez, J. M., & Hadi, A. S. (1996). Goal oriented symbolic propagation in Bayesian networks. In *Proceedings AAAI-96 national conference* (pp. 1263–1268).

Castillo, E., Gutiérrez, J. M., & Hadi, A. S. (1997). Sensitivity analysis in discrete Bayesian networks. *IEEE Transactions on Systems, Man, and Cybernetics, 27*, 412–423.

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., & Montrucchio, L. (2013). Classical subjective expected utility. *PNAS, 110*(17), 6754–6759.

Chan, H., & Darwiche, A. (2004). Sensitivity analysis in Bayesian networks: From single to multiple parameters. In *Proceedings of the twentieth annual conference on uncertainty in artificial intelligence (UAI-04)* (pp. 67–75). Arlington, VA: AUAI Press.

Chan, H., & Darwiche, A. (2005). A distance measure for bounding probabilistic belief change. *International Journal of Approximate Reasoning, 38*, 149–174.

Chastaing, G., Gamboa, F., & Prieur, C. (2012). Generalized Hoeffding–Sobol decomposition for dependent variables—Application to sensitivity analysis. *Electronic Journal of Statistics, 6*, 2420–2448.

Cheok, M. C., Parry, G. W., & Sherry, R. R. (1998). Use of importance measures in risk-informed regulatory applications. *Reliability Engineering & System Safety, 60*(3), 213–226.

Chun, M.-H., Han, S.-J., & Tak, N.-I. (2000). An uncertainty importance measure using a distance metric for the change in a cumulative distribution function. *Reliability Engineering & System Safety, 70*(3), 313–321.

Critchfield, G. C., & Willard, K. E. (1986). Probabilistic analysis of decision trees using Monte Carlo simulation. *Medical Decision Making, 6*, 85–92.

Crnkovic, C., & Drachman, J. (1996). Quality control. *RISK, 9*(9), 139–143.

Cukier, R. I., Fortuin, C. M., Shuler, K. E., Petschek, A. G., & Schaibly, J. H. (1973). Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I. Theory. *Journal of Chemical Physics, 59*, 3873–3878.

Da Veiga, S. (2015). Global sensitivity analysis with dependence measures. *Journal of Statistical Computation and Simulation, 85*(7), 1283–1305.

de Finetti, B. (1937). Foresight: Its logical laws, its subjective sources. *Annales de l'Institut Henri Poincaré, 7*(1), 1–68 (in French).

Der Kiureghian, A., & Ditlevsen, O. (2009). Aleatory or epistemic? Does it matter? *Structural Safety, 31*(2), 105–112.

Doksum, K., & Samarov, A. (1995). Nonparametric estimation of global functionals and a measure of the explanatory power of covariates in regression. *The Annals of Statistics, 23*(5), 1443–1473.

Efron, B., & Stein, C. (1981). The jackknife estimate of variance. *The Annals of Statistics, 9*(3), 586–596.

Ellsberg, D. (1961). Risk, ambiguity and the Savage axioms. *Quarterly Journal of Economics, 75*(4), 643–669.

EPA, U. (2009). Guidance on the development, evaluation, and application of environmental models. http://www.epa.gov/crem/cremlib.html.

Eschenbach, T. G. (1992). Spiderplots versus Tornado diagrams for sensitivity analysis. *Interfaces, 22*(6), 40–46.

Felli, J. C., & Hazen, G. B. (1998). Sensitivity analysis and the expected value of perfect information. *Medical Decision Making, 18*, 95–109.

Felli, J. C., & Hazen, G. B. (2004). Javelin diagrams: A graphical tool for probabilistic sensitivity analysis. *Decision Analysis, 1*(2), 93–107.

Filippi, C. (2005). A fresh view on the tolerance approach to sensitivity analysis in linear programming. *European Journal of Operational Research, 167*, 1–19.

Filippi, C. (2010). Sensitivity analysis in linear programming. In *Encyclopedia of operations research and management science* (pp. 1–10). Wiley.

Fort, J.-C., Klein, T., & Rachdi, N. (2014). New sensitivity indices subordinated to a contrast. *Communications in Statistics-Theory and Methods* (in press).

Fréchet, M. (1934). Sur le coefficient, dit de corrélation et sur la corrélation en géneral. *Revue de l'Institut International de Statistique, 1*(4), 16–23.

French, S. (2003). Modelling, making inferences and making decisions: The roles of sensitivity analysis. *Sociedad de Estadística e Investigación Operativa Top, 11*(2), 229–251.

Fürbringer, J.-M. (1996). Sensitivity analysis for modellers. *Air Infiltration Review, 17*(4), 8–10.

Fussell, J. (1975). How to calculate system reliability and safety characteristics. *IEEE Transactions on Reliability, 24*(3), 169–174.

Gal, T., & Greenberg, H. J. E. (1997). Advances in sensitivity analysis and parametric programming. no. 6. In *International series in operations research & management science: ISOR*. Boston, MA: Kluwer Academic.

Gao, X., Cui, L., & Li, J. (2007). Analysis for joint importance of components in a coherent system. *European Journal of Operational Research, 182*, 282–299.

Griewank, A., & Walther, A. (2008). *Evaluating derivatives: Principles and techniques of algorithmic differentiation* (2nd ed.). Philadelphia, PA: SIAM.

Hansen, L. P. (2007). Beliefs, doubts and learning: Valuing macroeconomic risk. *American Economic Review, 97*(2), 1–30.

Harris, F. W. (1990). How many parts to make at once. *Factory, The Magazine of Management, 10*(2), 135–136,152 (reprinted in *Operations Research* 38(6), 947–950).

Hazen, G. B., & Sounderpandian, J. (1999). Lottery acquisition versus information acquisition: Prices and preference reversals. *Journal of Risk and Uncertainty, 18*(2), 125–136.

Helton, J. C. (1993). Uncertainty and sensitivity analyses techniques for use in performance assessment for radioactive waste disposal. *Reliability Engineering & System Safety, 42*(2–3), 327–367.

Helton, J. C., & Davis, F. J. (2003). Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliability Engineering & System Safety, 81*(1), 23–69.

Helton, J. C., & Johnson, J. D. (2011). Quantification of margins and uncertainties: Alternative representations of epistemic uncertainty. *Reliability Engineering & System Safety, 96*, 1034–1052.

Helton, J. C., Johnson, J. D., Oberkampf, W. L., & Sallaberry, C. J. (2006). Sensitivity analysis in conjunction with evidence theory representations of epistemic uncertainty. *Reliability Engineering & System Safety, 91*(10–11), 1414–1434.

Helton, J. C., Johnson, J. D., Sallaberry, C. J., & Storlie, C. B. (2006). Survey of sampling-based methods for uncertainty and sensitivity analysis. *Reliability Engineering & System Safety, 91*(10–11), 1175–1209.

Higgins, J. P. T., & Green, S. E. (2011). *Cochrane handbook for systematic reviews of interventions. The Cochrane Collaboration* cochrane.handbook.org.

Hoffman, F. O., & Hammonds, J. S. (1994). Propagation of uncertainty in risk assessments: The need to distinguish between uncertainty due to lack of knowledge and uncertainty due to variability. *Risk Analysis, 14*(5), 707–712.

Homma, T., & Saltelli, A. (1996). Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering & System Safety, 52*(1), 1–17.

Hong, J. S., & Lie, C. H. (1993). Joint reliability-importance of two edges in an undirected network. *IEEE Transactions on Reliability, 42*(1), 17–23.

Hornberger, G. M., & Spear, R. C. (1981). An approach to the preliminary analysis of environmental systems. *Journal of Environmental Management, 12*, 7–18.

Howard, R. (1966). Information value theory. *IEEE Transactions on Systems Science and Cybernetics, 2*(1), 22–26.

Howard, R. A. (1988). Uncertainty about probability: A decision analysis perspective. *Risk Analysis, 8*(1), 91–98.

Huang, C. C., Vertinsky, I., & Ziemba, W. T. (1977). Sharp bounds on the value of perfect information. *Operations Research, 25*(1), 128–139.

Huang, C.-F., & Litzenberger, R. H. (1988). *Foundations for financial economics*. Upper Saddle River, NJ: Prentice Hall.

Iman, R. L., & Conover, W. J. (1979). The use of the rank transform in regression. *Technometrics, 21*(4), 499–509.

Iman, R. L., & Hora, S. C. (1990). A robust measure of uncertainty importance for use in fault tree system analysis. *Risk Analysis, 10*(3), 401–406.

Iman, R. L., Johnson, M. E., & Watson, C. C. Jr. (2005). Uncertainty analysis for computer model projections of hurricane losses. *Risk Analysis, 25*(5), 1299–1312.

Ishigami, T., & Homma, T. (1990). An importance quantification technique in uncertainty analysis for computer models. In *Proceedings of the first international symposium on uncertainty modeling and analysis* (pp. 398–403). College, Park, MD.

Jansen, M. J. W. (1999). Analysis of variance designs for model output. *Computer Physics Communications, 117*(1–2), 35–43.

Jungermann, H., & Thuring, M. (1988). The labyrinth of experts' minds: Some reasoning strategies and their pitfalls. *Annals of Operations Research, 16*, 117–130.

Kjaerulff, U., & van der Gaag, L. C. (2000). Making sensitivity analysis computationally efficient. In *Proceedings of the sixteenth conference on uncertainty in artificial intelligence (UAI-2000)* (pp. 317–325). Stanford, CA: AUAI Press.

Kleijnen, J. P. C. (2008). *Design and analysis of simulation experiments.* Berlin: Springer Verlag.

Kleijnen, J. P. C. (2009). Kriging metamodeling in simulation: A review. *European Journal of Operational Research, 192*(3), 707–716.

Kleijnen, J. P. C. (2010). Sensitivity analysis of simulation models. In *Encyclopedia of operations research and management science* (pp. 1–10). Wiley.

Kleijnen, J. P. C., & Helton, J. C. (1999). Statistical analyses of scatterplots to identify important factors in large-scale simulations, 1: Review and comparison of techniques. *Reliability Engineering & System Safety, 65*(2), 147–185.

Kleijnen, J. P. C., & Helton, J. C. (1999). Statistical analyses of scatterplots to identify important factors in large-scale simulations, 2: Robustness of techniques. *Reliability Engineering & System Safety, 65*(2), 187–197.

Kleijnen, J. P. C., Sanchez, S. M., Lucas, T. W., & Cioppa, T. M. (2005). A user's guide to the brave new world of designing simulation experiments. *INFORMS Journal on Computing, 17*(3), 263–289.

Klibanoff, P., Marinacci, M., & Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica, 73*(6), 1849–1892.

Krzykacz-Hausmann, B. (2001). Epistemic sensitivity analysis based on the concept of entropy. In *Third international symposium on sensitivity analysis of model output (SAMO 2001)* (pp. 31–35). Madrid: CIEMAT.

Kucherenko, S., Tarantola, S., & Annoni, P. (2012). Estimation of global sensitivity indices for models with dependent variables. *Computer Physics Communications, 183*(4), 937–946.

Kuo, W., & Zhu, X. (2012). *Importance measures in reliability, risk, and optimization: Principles and applications.* Chichester: Wiley.

Lewandowski, D., Cooke, R. M., & Duintjer Tebbens, R. J. (2007). Sample-based estimation of correlation ratio with polynomial approximation. *ACM Transactions on Modeling and Computer Simulation, 18*(1), 3:1–3:16.

Li, G., & Rabitz, H. (2012). General formulation of HDMR component functions with independent and correlated variables. *Journal of Mathematical Chemistry, 50*(1), 99–130.

Li, G., Wang, S.-W., Rosenthal, C., & Rabitz, H. (2001). High dimensional model representations generated from low dimensional data samples. I. mp-Cut-HDMR. *Journal of Mathematical Chemistry, 30*(1), 1–30.

Little, J. D. C. (1970). Models and managers: The concept of a decision calculus. *Management Science, 16*(8), B466–B485.

Liu, H., Chen, W., & Sudjianto, A. (2006). Relative entropy based method for probabilistic sensitivity analysis in engineering design. *ASME Journal of Mechanical Design, 128*, 326–336.

Liu, Q., & Homma, T. (2010). A new importance measure for sensitivity analysis. *Journal of Nuclear Science and Technology, 47*(1), 53–61.

Liu, R., & Owen, A. B. (2006). Estimating mean dimensionality of analysis of variance decompositions. *Journal of the American Statistical Association, 101*(474), 712–721.

Lozzo, M. D., & Marrel, A. (2014). New improvements in the use of dependence measures for sensitivity analysis and screening. arXiv:1412.1414v1.

Lu, L., & Jiang, J. (2007). Joint failure importance for noncoherent fault trees. *IEEE Transactions on Reliability, 56*(3), 435–443.

Luo, X., Lu, Z., & Xu, X. (2014). A fast computational method for moment-independent uncertainty importance measure. *Computer Physics Communications, 185*, 19–27.

Madsen, H. O., Krenk, S., & Lind, N. C. (1986). *Methods of structural safety.* Englewood Cliffs, NJ: Prentice-Hall.

Mara, T. A., & Tarantola, S. (2012). Variance-based sensitivity indices for models with dependent inputs. *Reliability Engineering & System Safety, 107*, 115–121.

Mason, D. M., & Schuenemeyer, J. H. (1983). A modified Kolmogorov–Smirnov test sensitive to tail alternatives. *The Annals of Statistics, 11*(3), 933–946.

Montgomery, D. C. (2000). *Design and analysis of experiments* (5th ed.). John Wiley & Sons.

Morris, M. D. (1991). Factorial sampling plans for preliminary computational experiments. *Technometrics, 33*(2), 161–174.

NICE (2013). Guide to the methods of technology appraisal. Technical report. London and Manchester: National Institute for Health and Care Excellence. publications.nice.org.uk/pmg9

Nielsen, T. D., & Jensen, F. V. (2003). Sensitivity analysis in influence diagrams. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans, 33*(2), 223–234.

Nøkland, T. E., & Aven, T. (2013). On selection of importance measures in risk and reliability analysis. *International Journal of Performability Engineering, 9*(2), 133–147.

NRC, U. (2009). An approach for using probabilistic risk assessment in risk-informed decisions on plant-specific changes to the licensing basis. Technical report. Draft regulatory guide DG-1226, proposed revision 2 of regulatory guide 1.174. Office of Nuclear Regulatory Research.

Oakley, J. E. (2009). Decision-theoretic sensitivity analysis for complex computer models. *Technometrics, 51*(2), 121–129.

Oakley, J. E., & O'Hagan, A. (2004). Probabilistic sensitivity analysis of complex models: A Bayesian approach. *Journal of the Royal Statistical Society, Series B, 66*(3), 751–769.

O'Brien, F. A. (2004). Scenario planning—Lessons for practice from teaching and learning. *European Journal of Operational Research, 152*, 709–722.

Owen, A. B. (2003). The dimension distribution and quadrature test functions. *Statistica Sinica, 13*(1), 1–17.

Owen, A. B. (2013). Better estimation of small Sobol' sensitivity indices. *ACM Transactions on Modeling and Computer Simulation (TOMACS), 23*(2), 11.1–17.

Park, C. K., & Ahn, K. I. (1994). A new approach for measuring uncertainty importance and distributional sensitivity in probabilistic safety assessment. *Reliability Engineering&System Safety, 46*(3), 253–261.

Parry, G. W. (1996). The characterization of uncertainty in probabilistic risk assessments of complex systems. *Reliability Engineering & System Safety, 54*(2–3), 119–126.

Paté-Cornell, M. E. (1996). Uncertainties in risk analysis: Six levels of treatment. *Reliability Engineering & System Safety, 54*(2–3), 95–111.

Pearl, J. (1988). On probability intervals. *International Journal of Approximate Reasoning, 2*, 211–216.

Pearson, K. (1905). On the general theory of skew correlation and non-linear regression. In *Vol. XIV of mathematical contributions to the theory of evolution, Drapers' Company Research Memoirs.* London: Dulau & Co.

Plischke, E. (2010). An effective algorithm for computing global sensitivity indices (EASI). *Reliability Engineering & System Safety, 95*(4), 354–360.

Plischke, E. (2012a). An adaptive correlation ratio method using the cumulative sum of the reordered output. *Reliability Engineering & System Safety, 107*, 149–156.

Plischke, E. (2012b). How to compute variance-based sensitivity indicators with your spreadsheet software. *Environmental Modelling & Software, 35*, 188–191.

Plischke, E., Borgonovo, E., & Smith, C. L. (2013). Global sensitivity measures from given data. *European Journal of Operational Research, 226*(3), 536–550.

Pörn, K. (1997). A decision-oriented measure of uncertainty importance for use in PSA. *Reliability Engineering & System Safety, 56*(1), 17–27.

Pratt, J., Raiffa, H., & Schlaifer, R. (1995). *Introduction to statistical decision theory.* The MIT Press.

Rabitz, H. (1989). Systems analysis at the molecular scale. *Science, 246*, 221–226.

Rabitz, H., & Alış, Ö. F. (1999). General foundations of high-dimensional model representations. *Journal of Mathematical Chemistry, 25*(2–3), 197–233.

Rackwitz, R. (2001). Reliability analysis—A review and some perspectives. *Structural Safety, 23*(4), 365–395.

Rakovec, O., Hill, M. C., Clark, M. P., Weerts, A. H., Teuling, A. J., & Uijlenhoet, R. (2014). Distributed evaluation of local sensitivity analysis (DELSA), with application to hydrologic models. *Water Resources Research, 50*, 1–18.

Ratto, M., Pagano, A., & Young, P. C. (2007). State dependent parameter metamodelling and sensitivity analysis. *Computer Physics Communications, 177*(11), 863–876.

Ravi, N., & Wendell, R. (1989). The tolerance approach to sensitivity analysis of matrix coefficients in linear programming. *Management Science, 35*(9), 1106–1119.

Rényi, A. (1959). On measures of dependence. *Acta Mathematica Academiae Scientiarum Hungaricae, 10*, 441–451.

Rose, K. A., Smith, E. P., Gardner, R. H., Brenkert, A. L., & Bartell, S. M. (1991). Parameter sensitivities, Monte Carlo filtering, and model forecasting under uncertainty. *Journal of Forecasting, 10*, 117–133.

Rosenhead, J. (2011). Robustness analysis. In *Encyclopedia of operations research and management science* (pp. 1–10). Wiley.

Saltelli, A. (2002). Making best use of model evaluations to compute sensitivity indices. *Computer Physics Communications, 145*(2), 280–297.

Saltelli, A. (2002). Sensitivity analysis for importance assessment. *Risk Analysis, 22*(3), 579–590.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Design and estimator for the total sensitivity index. *Computer Physics Communications, 181*(2), 259–270.

Saltelli, A., Chan, K., & Scott, E. M. (2000). *Sensitivity analysis.* Chichester: John Wiley&Sons.

Saltelli, A., & Marivoet, J. (1990). Non-parametric statistics in sensitivity analysis for model output: A comparison of selected techniques. *Reliability Engineering & System Safety, 28*(2), 229–253.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., & Tarantola, S. (2008). *Global sensitivity analysis – the primer.* Chichester: John Wiley & Sons.

Saltelli, A., Ratto, M., Tarantola, S., & Campolongo, F. (2005). Sensitivity analysis for chemical models. *Chemical Reviews, 105*(7), 2811–2828.

Saltelli, A., Ratto, M., Tarantola, S., & Campolongo, F. (2012). Update 1 of: Sensitivity analysis for chemical models. *Chemical Reviews, 112*, PR1–PR21 Perennial Reviews, DOI:10.1021/cr200301u.

Saltelli, A., & Sobol', I. M. (1995). About the use of rank transformation in the sensitivity analysis of model output. *Reliability Engineering & System Safety, 50*(3), 225–239.

Saltelli, A., & 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), 702–709.

Saltelli, A., Tarantola, S., & Campolongo, F. (2000). Sensitivity analysis as an ingredient of modelling. *Statistical Science, 19*(4), 377–395.

Samuelson, P. (1947). *Foundations of economic analysis.* Cambridge, MA: Harvard University Press.

Samuelson, P. A. (1941). The stability of equilibrium: Comparativestatics and dynamics. *Econometrica, 9*(2), 97–120.

Santner, T. J., Williams, B. J., & Notz, W. I. (2003). The design and analysis of computer experiments. In *Springer series in statistics.* New York: Springer Verlag.

Savage, L. J. (1954). *The foundations of statistics.* John Wiley & Sons.

Schwartz, P. (1996). *The art of the long view.* New York: Double Day.

SimLab (2011). SimLab: Software package for uncertainty and sensitivity analysis. http://simlab.jrc.ec.europa.eu/.

Sobol', I. M. (1969). *Multidimensional quadrature formulas and Haar functions.* Moscow: Nauka (in Russian).

Sobol', I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical Modelling & Computational Experiments, 1*, 407–414.

Sobol', I. M. (2003). Theorems and examples on high dimensional model representation. *Reliability Engineering & System Safety, 79*, 187–193.

Sobol', I. M., & 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*(7), 1212–1217.

Spearman, C. (1904). The proof and measurement of association between two things. *American Journal of Psychology, 15*, 72–101.

Storlie, C. B., Reich, B. J., Helton, J. C., Swiler, L. P., & Sallaberry, C. J. (2013). Analysis of computationally demanding models with continuous and categorical inputs. *Reliability Engineering & System Safety, 113*, 30–41.

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

Strasser, H. (1985). *Mathematical theory of statistics. Vol. 7 of studies in mathematics*. Berlin: de Gruyter.

Straub, D. (2014). Value of information analysis with structural reliability methods. *Structural Safety, 49*, 75–85.

Strong, M., & Oakley, J. E. (2013). An efficient method for computing single-parameter partial expected value of perfect information. *Medical Decision Making, 33*(6), 755–766.

Strong, M., & Oakley, J. E. (2014). When is a model good enough? Deriving the expected value of model improvement via specifying internal model discrepancies. *SIAM/ASA Journal of Uncertainty Quantification, 2*, 106–125.

Strong, M., Oakley, J. E., & Chilcott, J. (2012). Managing structural uncertainty in health economic decision models: A discrepancy approach. *Journal of the Royal Statistical Society, Series C, 61*(1), 25–45.

Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansion. *Reliability Engineering & System Safety, 93*, 964–979.

Tan, C. H., & Hartman, J. C. (2011). Sensitivity analysis and dynamic programming. In *Encyclopedia of operations research and management science* (pp. 1–10). Wiley.

Tang, Z., Lu, Z., Jiang, B., Pan, W., & Zhang, F. (2013). Entropy-based importance measure for uncertain model inputs. *AIAA Journal, 51*(10), 2319–2334.

Tarantola, S., Gatelli, D., & Mara, T. A. (2006). Random balance designs for the estimation of first order global sensitivity indices. *Reliability Engineering & System Safety, 91*, 717–727.

Tietje, O. (2005). Identification of a small reliable and efficient set of consistent scenarios. *European Journal of Operational Research, 162*(4), 418–432.

van der Gaag, L. C., Renooij, S., & Coupé, V. M. H. (2007). Sensitivity analysis of probabilistic networks. In *Advances in probabilistic graphical models. Vol. 214 of studies in fuzziness and soft computing* (pp. 103–124). Springer Verlag.

van Groenendaal, W. J. H., & Kleijnen, J. P. C. (1997). On the assessment of economic risk: Factorial design versus Monte Carlo methods. *Reliability Engineering & System Safety, 57*(1), 91–102.

van Groenendaal, W. J. H., & Kleijnen, J. P. C. (2002). Deterministic versus stochastic sensitivity analysis in investment problems: An environmental case study. *European Journal of Operational Research, 141*(1), 8–20.

Vesely, W. E., Kurth, R. E. Scalzo, S. M. (1990). Evaluations of core melt frequency effects due to component aging and maintenance. Technical report. NUREG/CR-5510. Washington, DC: U.S. Nuclear Regulatory Commission.

Wagner, H. M. (1995). Global sensitivity analysis. *Operations Research, 43*(6), 948–969.

Wallace, S. M. (2000). Decision making under uncertainty: Is sensitivity analysis of any use? *Operations Research, 48*(1), 20–25.

Wan, H., Ankenman, B. E., & Nelson, B. L. (2006). Controlled sequential bifurcation: A new factor-screening method for discrete-event simulation. *Operations Research, 54*(4), 743–755.

Wan, H., Ankenman, B. E., & Nelson, B. L. (2010). Improving the efficiency and efficacy of controlled sequential bifurcation for simulation factor screening. *INFORMS Journal of Computing, 22*(3), 482–492.

Wang, X. (2006). On the effects of dimension reduction techniques on some high-dimensional problems in finance. *Operations Research, 54*(6), 1063–1078.

Wang, X., & Sloan, I. H. (2011). Quasi-Monte Carlo methods in financial engineering: An equivalence principle and dimension reduction. *Operations Research, 59*(1), 80–95.

Wei, P., Lu, Z., & Song, J. (2014). Moment-independent sensitivity analysis using copula. *Risk Analysis, 34*(2), 210–222.

Wendell, R. (1985). The tolerance approach to sensitivity analysis in linear programming. *Management Science, 31*(5), 564–578.

Winkler, R. L. (1996). Uncertainty in probabilistic risk assessment. *Reliability Engineering & System Safety, 54*(2–3), 127–132.

Zhai, Q., Yang, J., Xie, M., & Zhao, Y. (2014a). Generalized moment-independent importance measures based on Minkowski distance. *European Journal of Operational Research, 239*(2), 449–455.

Zhai, Q., Yang, J., & Zhao, Y. (2014b). Space-partition method for the variance-based sensitivity analysis: Optimal partition scheme and comparative study. *Reliability Engineering & System Safety, 131*, 66–82.

Zhang, L., Lu, Z., Cheng, L., & Fan, C. (2014). A new method for evaluating Borgonovo moment-independent importance measure with its application in an aircraft structure. *Reliability Engineering & System Safety, 132*, 163–175.

Zio, E., & Podofillini, L. (2006). Accounting for components interactions in the differential importance measure. *Reliability Engineering & System Safety, 91*(10–11), 1163–1174.