Reliability Engineering and System Safety 94 (2009) 237–243

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Non-parametric estimation of conditional moments for sensitivity analysis

M. Ratto [a,*], A. Pagano [a], P.C. Young [b,c,d]

[a] Joint Research Centre, European Commission, Ispra, Italy
[b] Institute of Environmental and Natural Sciences, Lancaster University, UK
[c] School of Environment and Society, Australian National University, Canberra, Australia
[d] School of Electrical and Communication Engineering, University of New South Wales, Sydney, Australia

ARTICLE INFO

Article history:
Received 29 October 2007
Received in revised form
9 March 2008
Accepted 9 March 2008
Available online 14 March 2008

Keywords:
Sensitivity analysis
Non-parametric methods
Conditional moments
State-dependent parameter models

ABSTRACT

In this paper, we consider the non-parametric estimation of conditional moments, which is useful for applications in global sensitivity analysis (GSA) and in the more general emulation framework. The estimation is based on the state-dependent parameter (SDP) estimation approach and allows for the estimation of conditional moments of order larger than unity. This allows one to identify a wider spectrum of parameter sensitivities with respect to the variance-based main effects, like shifts in the variance, skewness or kurtosis of the model output, so adding valuable information for the analyst, at a small computational cost.

1. Introduction

In global sensitivity analysis (GSA), the mapping $Y = f(\mathbf{X})$ between an output Y of a computational model and a set of uncertain input factors $\mathbf{X} = (X_1, \ldots, X_k)$ is analyzed in order to quantify the relative contribution of each input factor to the uncertainty of Y. Variance-based analysis is the most popular method in GSA. Variance-based sensitivity indices of single factors or of groups of them are defined as [1,24]

$$S_{\mathbf{I}} = \frac{\mathrm{Var}(E(Y|\mathbf{X}_{\mathbf{I}}))}{\mathrm{Var}(Y)} \tag{1}$$

where $\mathbf{X}_{\mathbf{I}}$ denotes a group of factors indexed by $\mathbf{I} = (i_1, \ldots, i_g)_{1 \leqslant g \leqslant k}$, and they tell the portion of variance of Y that is explained by $\mathbf{X}_{\mathbf{I}}$.

The two most popular variance-based sensitivity measures are the main effect

$$S_i = \frac{\mathrm{Var}(E(Y|X_i))}{\mathrm{Var}(Y)} \tag{2}$$

and the total effect

$$S_{Ti} = \frac{E(\mathrm{Var}(Y|\mathbf{X}_{-i}))}{\mathrm{Var}(Y)} \tag{3}$$

where $\mathbf{X}_{-i}$ indicates all input factors except X_i.

The main effect measures the singular contribution of the input factor X_i to the uncertainty (variance) of the output Y, while the total effect measures the overall contribution of X_i on Y, including all interaction terms of X_i with all other input factors.

There are clear links between variance-based sensitivity analysis and model emulation. First, a statistical approximation (the emulator) $\hat{f}(\mathbf{X})$ can be used to compute sensitivity indices in place of the original computational mapping $f(\mathbf{X})$. Second, the variance-based sensitivity measures can be interpreted as the non-parametric R^2 or correlation ratio, used in statistics to measure the explanatory power of covariates in regression [2,3]. In fact, it is well known that the inner argument $E(Y|\mathbf{X}_{\mathbf{I}})$ of (1) is the function of the subset of input factors that approximates $f(\mathbf{X})$, by minimizing a quadratic loss (i.e. maximizing the R^2). Therefore, estimating $E(Y|\mathbf{X}_{\mathbf{I}})$ provides a route for both a model approximation and sensitivity estimation. Smoothing methods that provide more or less accurate and efficient estimations of $E(Y|\mathbf{X}_{\mathbf{I}})$ are becoming a popular approach to sensitivity analysis [4–8]. State-dependent parameter (SDP) modelling is one class of non-parametric smoothing approach first suggested by Young [9,10]. The estimation is performed with the help of the 'classical' recursive (numerically non-intensive) Kalman filter (KF) and associated fixed interval smoothing (FIS) algorithms: it has been applied for sensitivity analysis by Ratto et al. in [11,12].

Variance-based techniques have a quite general applicability, since they apply to a very wide range of non-linear mappings $f(\cdot)$ and rely on only a few assumptions, namely Y has to be square integrable and the variance is an adequate measure of the uncertainty of Y. Nonetheless, these techniques are sometimes criticized, since all kinds of sensitivity patterns that cannot be

* Corresponding author. Tel.: +39 0332 785639; fax: +39 0332 785733.
E-mail addresses: marco.ratto@jrc.it (M. Ratto), andrea.pagano@jrc.it (A. Pagano), p.young@lancaster.ac.uk (P.C. Young).

0951-8320/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ress.2008.02.023

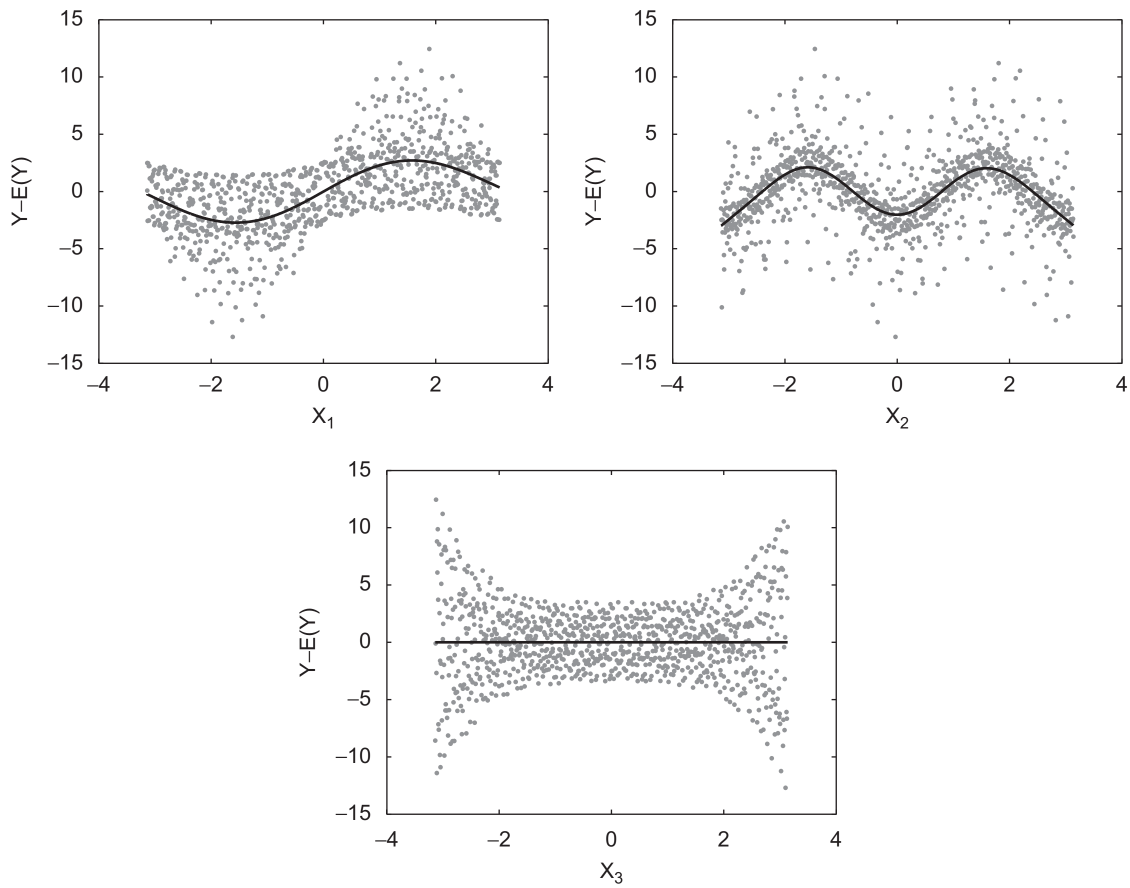

Fig. 1. Scatterplots of the de-meaned Ishigami test function (grey dots) and non-parametric estimation of $E(Y|X_i)$ (black lines).

attributed to shifts in the mean (the first moment—see factor X_3 in Fig. 1), are not accounted for by $E(Y|X_i)$ and the related variance-based sensitivity index. Such sensitivity patterns can be characterized by a shift in higher order moments: the simplest example of which is the heteroscedastic process, where the *variance* of Y changes along the conditioning term X_i. This lead to the development of a number of sensitivity techniques, such as entropy-based sensitivity measures [13,14] or moment independent sensitivity measures [15,16], that provide 'main effects' that are able to account for such phenomena.

In this paper, we show how non-parametric techniques can be applied to estimate conditional moments of order larger than one, allowing us to add valuable information to the standard variance-based analysis and, at the same time, avoid the computational load characterizing the latter class of sensitivity measures. In fact, the analysis does not require any additional model evaluation with respect to any standard smoothing method that may be applied to estimate the $E(Y|X_i)$ terms.

2. The method

Readers can refer to [12] for a discussion of the SDP approach to sensitivity analysis and to [10] for a more comprehensive discussion of SDP modelling and its algorithms. Here we synthesize some key concepts regarding the estimation of main effects.

Summarizing considerably, a state-dependent model approximating $E(Y|X_i)$, based on a Monte Carlo sample of dimension N, can be written as

$$Y_t = E(Y|X_{i,t}) + e_{i,t} = p_{i,t}(s_{i,t}) + e_{i,t} \qquad (4)$$

where $e_{i,t}$ is the observation noise (i.e. what is not explained by $E(Y|X_i)$), $p_{i,t}$ is a state-dependent parameter, depending on the state variable s_i that moves in a sorted order $t = 1, \ldots, N$ along the co-ordinates of $X_{i,t}$ and Y_t is the 'sorted' output. According to this sorting strategy, input factors of interest X_i's are characterized by a low frequency spectrum, while the remaining ones present a white spectrum. In this way, the estimation of $E(Y|X_i)$ reduces to the extraction of the low frequency component of the sorted output Y_t (scatterplot smoothing). To do so, the SDP's are modelled by one member of the generalized random walk (GRW) class of non-stationary processes. For instance, the integrated random walk (IRW) process turns out to provide good results, since it ensures that the estimated SDP relationship has the smooth properties of a cubic spline. Given the IRW characterization, model (4) can be put into state space form as

$$\text{Observation equation:} \quad Y_t = p_{i,t} + e_{i,t} \qquad (5a)$$

$$\text{SDP model:} \quad \begin{aligned} p_{i,t} &= p_{i,t-1} + d_{i,t-1} \\ d_{i,t} &= d_{i,t-1} + \eta_{i,t} \end{aligned} \qquad (5b)$$

where $e_{i,t}$ (observation noise) and $\eta_{i,t}$ (system disturbances) are zero mean white noise inputs with variance σ_i^2 and $\sigma_{\eta_i}^2$, respectively.

Given a SDP relationship such as (5), the SDP's are estimated using the recursive KF and associated recursive FIS algorithm. Under the Gaussian assumption for the distributions of $e_{i,t}$ and $\eta_{i,t}$, the *hyper-parameters* of such a relationship, in the form of the various noise variances and any other parameters associated with the GRW model for the SDPs, are optimized by maximum likelihood (ML), using prediction error decomposition. In fact, by a simple reformulation of the KF and FIS algorithms in the case of