Reliability Engineering and System Safety 121 (2014) 263–275

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

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

Construction of bootstrap confidence intervals on sensitivity indices computed by polynomial chaos expansion

S. Dubreuil [a,b,*], M. Berveiller [c], F. Petitjean [b], M. Salaün [a]

[a] Institut Clément Ader (ICA), ISAE, F-31055 Toulouse, France
[b] Institut Catholique d'Arts et Métiers (ICAM), F-31300 Toulouse, France
[c] EDF R&D-Département MMC, Site des Renardières, F-77818 Moret-sur-Loing, France

ARTICLE INFO

Article history:
Received 11 January 2013
Received in revised form
9 September 2013
Accepted 18 September 2013
Available online 26 September 2013

Keywords:
Sensitivity analysis
Polynomial chaos expansion
Bootstrap re-sampling

ABSTRACT

Sensitivity analysis aims at quantifying influence of input parameters dispersion on the output dispersion of a numerical model. When the model evaluation is time consuming, the computation of Sobol' indices based on Monte Carlo method is not applicable and a surrogate model has to be used. Among all approximation methods, polynomial chaos expansion is one of the most efficient to calculate variance-based sensitivity indices. Indeed, their computation is analytically derived from the expansion coefficients but without error estimators of the meta-model approximation. In order to evaluate the reliability of these indices, we propose to build confidence intervals by bootstrap re-sampling on the experimental design used to estimate the polynomial chaos approximation. Since the evaluation of the sensitivity indices is obtained with confidence intervals, it is possible to find a design of experiments allowing the computation of sensitivity indices with a given accuracy.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Performing global sensitivity analysis is often a major step in uncertainties propagation studies. It helps to understand how uncertainties of a quantity of interest could be explained and reduced. Different types of sensitivity analysis can be performed (see [21]). This paper focuses on variance-based ones, computed by polynomial chaos expansion. Sensitivity indices, coming from variance decomposition (ANOVA), are relevant informations as they allow one to quantify effect of a variable (alone or in interaction with one or more variables) but require estimate of many partial variances (see [21] for a description of sensitivity indices). When these partial variances cannot be expressed analytically, which is often the case in industrial applications, a Monte Carlo based method, developed in [23], leads to an approximation of these indices. If the computation of the model is time consuming (finite element models for example), Monte Carlo simulations become unrealistic and a common way to tackle this problem is the use of a meta-model. In this case, the idea is to replace the true model by an analytical one as precise as possible and then to use it in the Monte Carlo methodology. This involves two types of error: a meta-modelling error coming from the difference between the true model and its approximation and a sampling error due to the Monte Carlo methodology, used for the sensitivity indices computation.

An efficient way to compute sensitivity indices is to use an approximation of the model by polynomial chaos expansions (PCEs). Indeed, [26] shows that sensitivity indices are analytically calculated with the coefficients of that expansion. Then, the two types of error, given earlier, are reduced to the meta-modelling error only. Hence, it is of great importance to quantify and control it. The quality of meta-models is usually defined as the difference between the true model and the meta-model. This difference can be expressed using several error criteria like coefficient of determination, Mallows Cp, cross-validation, etc. Numerous methods propose iterative constructions of meta-models based on one (or several) of these criteria. For example, in [20], a quadratic surface response is built based, first, on the minimization of the sum of squared error, then in four different error criteria (Mallows Cp, AIC, BIC, adjusted coefficient of determination) and finally on leave-one-out validation. Concerning sensitivity analysis, [4] proposes an innovative construction of sparse PCE and selects the best PCE thanks to a corrected leave-one-out error. All these methodologies are efficient but do not take into account the aim of the meta-model. Moreover, it is difficult to link a global criterion error to the error on sensitivity indices computed from the meta-model. Finally, it is difficult to target a global error criterion value that allows a level of confidence on sensitivity indices. This problem also arises in reliability analysis and many authors propose error measurements and adaptive algorithms based on the probability

* Corresponding author at: Institut Clément Ader (ICA), ISAE, F-31055 Toulouse, France. Tel.: +33 561339245.
E-mail addresses: sylvain.dubreuil@isae.fr (S. Dubreuil), marc.berveiller@edf.fr (M. Berveiller), frank.petitjean@icam.fr (F. Petitjean), michel.salaun@isae.fr (M. Salaün).

0951-8320/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ress.2013.09.011

of failure obtained by meta-models and not only on global meta-modelling error. For example, in [11], the authors use a complete quadratic response surface as a meta-model and build confidence intervals by Jack-knife re-sampling around the design point. In [16], a bootstrap re-sampling on the failure probability is used to construct an optimized PCE. Confidence interval constructed by bootstrap re-sampling is also widely used in global sensitivity analysis. Plischke et al. [17] use bias-reducing bootstrap on variance-based and density-based sensitivity indices computed by sampling. Castaings et al. [5] use bootstrap on density-based sensitivity indices computed by different sampling strategies. In the field of sensitivity analysis performed by meta-models, [13] uses reduced-basis meta-models to estimate variance-based sensitivity indices and combine the property of reduced-basis meta-model and bootstrap re-sampling to compute confidence intervals. Storlie et al. [25] compares several types of meta-models and also uses bootstrap re-sampling on this meta-models to obtain confidence intervals.

This paper proposes to take advantage of the PCE in the estimation of variance based sensitivity indices. Then, in order to know if this approximation is accurate enough to estimate partial variances, a way to construct confidence intervals by bootstrap re-sampling is presented. In the first part of this paper, some important features about PCE and the determination of sensitivity indices are recalled. One important point deals with the method used to construct the polynomial basis. A quite recent methodology based on the *Least Angle Regression* (LAR) algorithm and developed in [3] is used. The second part presents an application of bootstrap re-sampling [10] to the computation of sensitivity indices, when they are estimated by PCE. Some results about the determination of confidence intervals are recalled and an algorithm is presented which is set up to build a design of experiments allowing one to obtain sensitivity indices with a given level of confidence. Finally, this methodology is tested, first, on academic cases (Ishigami and g-Sobol' functions) and, second, is used for a sensitivity analysis on a finite element model of satellite TARANIS, designed by the Centre National d'Etudes Spatiales (France).

2. Determination of sensitivity indices by polynomial chaos expansion

2.1. Approximation of a stochastic model by a PCE

Let us consider a numerical model $y(\mathbf{X})$, that depends on a random vector $\mathbf{X}=\{X_1,\ldots,X_n\}$ of n independent random variables, defined by the joint probability density function (PDF), say $f_{\mathbf{X}}(\mathbf{x})=\prod_{i=1}^{n} f_{X_i}(x_i)$. It is shown [24] that any second-order random variable can be expanded into a polynomial decomposition as

$$y(\mathbf{X})=\sum_{i=0}^{\infty} C_i\phi_i(X_1,\ldots,X_n) \tag{1}$$

where $\{\phi_i\}_{i\in\mathbb{N}}$ is an adequate orthogonal polynomial basis, with respect to the joint PDF, and $\{C_i\}_{i\in\mathbb{N}}$ are unknown coefficients. In practice, decomposition Eq. (1) is truncated to a finite number of terms, say P, according to

$$y(\mathbf{X})\approx\widehat{y}(\mathbf{X})=\sum_{i=0}^{P-1} C_i\phi_i(X_1,\ldots,X_n)=\sum_{i=0}^{P-1} C_i\phi_i(\mathbf{X}) \tag{2}$$

This paper only deals with the so called *non-intrusive* methods which do not need a modification of the numerical code computing the output $\mathbf{Y}$. They are simple to implement and do not ask special form of $\mathbf{Y}$, except that $E[Y^2]<\infty$.

The next subsections present the construction of a basis $\{\phi_i\}$ and the computation of the coefficients C_i.

Table 1
Univariate orthogonal polynomials for usual random variables.

Random variable	Orthogonal polynomials
Gaussian	Hermite
Uniform	Legendre
Beta	Jacobi
Gamma	Laguerre

2.2. Construction of the candidate basis

It is shown in [27] that classical univariate polynomial bases should be used for usual distributions (see Table 1). Then the orthogonal multivariate polynomial basis is obtained from the product of each univariate polynomial. This approach is chosen because only usual distributions are used. In other cases, the simplest solution consists in an iso-probabilistic transformation of the input variables into standard normal ones [14].

The multivariate polynomial basis in Eq. (1) is composed of an infinity of terms. As seen in Eq. (2), this basis is truncated to a finite number of terms, say P. In the following, polynomials are ranked by order (first polynomials are univariate of degree one, then multivariate using two variables of degree one, then the univariate at degree two, etc.).

The simplest way to truncate the basis is then to choose the P first polynomials. For example, the number P of polynomials necessary to reach a maximal order p is $P=(n+p)!/(n!p!)$, where n is the number of random variables. This strategy is efficient for problems of small dimension and responses that can be approximated by low degree polynomials. When it is not the case, the number of terms becomes important and leads to conditioning problems. Considering this issue, considerable research efforts were done during the last years to create efficient selection algorithms, leading to sparse bases in regression area and particularly in PCE area [2]. They will be detailed in Section 2.3.2.

2.3. Computation of the coefficients

2.3.1. Ordinary least square

Coefficients C_i are determined by minimizing the quadratic norm of the error $\varepsilon_y=\mathbf{Y}-\underline{\Phi}\mathbf{C}$, between some exact values $y(\mathbf{X})$ estimated at N different points (experimental design of size N) concatenated into vector $\mathbf{Y}$, and their estimation by the truncated polynomial expansion, concatenated into vector $\underline{\Phi}\mathbf{C}$, where $\mathbf{C}$ is the vector of unknown coefficients C_i in Eq. (2) and $\underline{\Phi}\in\mathbb{M}^{N,P}$ is the matrix of regressors. Column vectors of matrix $\underline{\Phi}$ are evaluations of polynomials ϕ_i, $i\in[0,P-1]$, at the N points of the experimental design. The least-square minimization criterion leads to

$$\mathbf{C}=(\underline{\Phi}^t\underline{\Phi})^{-1}\underline{\Phi}^t\mathbf{Y}. \tag{3}$$

2.3.2. LAR

Let us now introduce classical notations for sparse basis. First, the multi-index is $\boldsymbol{\alpha}=\{\alpha_1\cdots\alpha_i\cdots\alpha_n\}$, and $\mathcal{A}$ is the family of multi-indices $\boldsymbol{\alpha}$. From now, polynomial $\phi_{\boldsymbol{\alpha}}$ is the one acting on variables X_i at power α_i, for $i\in[1,n]$. Its total degree is $|\boldsymbol{\alpha}|=\sum_{i=1}^{n}\alpha_i$. With such notations, the polynomial chaos expansion of a stochastic model $y(\mathbf{X})$ (see Eq. (2)) reads

$$y(\mathbf{X})\approx\widehat{y}(\mathbf{X})=\sum_{\boldsymbol{\alpha}\in\mathcal{A}} C_{\boldsymbol{\alpha}}\phi_{\boldsymbol{\alpha}}(\mathbf{X}). \tag{4}$$

Given a full candidate basis $\mathcal{B}$ of maximal degree p, with $p=\max|\boldsymbol{\alpha}|$ and $card(\mathcal{B})=(n+p)!/(n!p!)$, a polynomial chaos expansion is said sparse if $card(\mathcal{A})<card(\mathcal{B})$. As the expansion coefficients are determined by regression, several tools, initially set-up

Download English Version:

https://daneshyari.com/en/article/7195889

Download Persian Version:

https://daneshyari.com/article/7195889

Daneshyari.com