



**HAL**  
open science

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

Sylvain Dubreuil, Marc Berveiller, Frank Petitjean, Michel Salaün

► **To cite this version:**

Sylvain Dubreuil, Marc Berveiller, Frank Petitjean, Michel Salaün. Construction of bootstrap confidence intervals on sensitivity indices computed by polynomial chaos expansion. *Reliability Engineering and System Safety*, 2014, vol. 121, pp. 263-275. 10.1016/j.ress.2013.09.011 . hal-00933583

**HAL Id: hal-00933583**

**<https://hal.science/hal-00933583>**

Submitted on 20 Jan 2014

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

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



## Open Archive Toulouse Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in: <http://oatao.univ-toulouse.fr/>  
Eprints ID: 9717

**To link to this article:** DOI: 10.1016/j.res.2013.09.011  
URL: <http://dx.doi.org/10.1016/j.res.2013.09.011>

**To cite this version:** Dubreuil, Sylvain and Berveiller, Marc and Petitjean, Frank and Salaün, Michel *Construction of bootstrap confidence intervals on sensitivity indices computed by polynomial chaos expansion*. (2014) Reliability Engineering & System Safety, vol. 121 . pp. 263-275. ISSN 0951-8320

Any correspondence concerning this service should be sent to the repository administrator: [staff-oatao@inp-toulouse.fr](mailto:staff-oatao@inp-toulouse.fr)

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

S. Dubreuil <sup>a,b,\*</sup>, M. Berveiller <sup>c</sup>, F. Petitjean <sup>b</sup>, M. Salaün <sup>a</sup>

<sup>a</sup> Institut Clément Ader (ICA), ISAE, F-31055 Toulouse, France

<sup>b</sup> Institut Catholique d'Arts et Métiers (ICAM), F-31300 Toulouse, France

<sup>c</sup> EDF R&D-Département MMC, Site des Renardières, F-77818 Moret-sur-Loing, France

## ARTICLE INFO

### 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.

## 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](mailto:sylvain.dubreuil@isae.fr) (S. Dubreuil), [marc.berveiller@edf.fr](mailto:marc.berveiller@edf.fr) (M. Berveiller), [frank.petitjean@icam.fr](mailto:frank.petitjean@icam.fr) (F. Petitjean), [michel.salaun@isae.fr](mailto:michel.salaun@isae.fr) (M. Salaün).

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  $\mathbf{X} = \{X_1, \dots, X_n\}$  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

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

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

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

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  $\phi_i$ ,  $i \in [0, P-1]$ , at the  $N$  points of the experimental design. The least-square minimization criterion leads to

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

#### 2.3.2. LAR

Let us now introduce classical notations for sparse basis. First, the multi-index is  $\boldsymbol{\alpha} = \{\alpha_1 \dots \alpha_i \dots \alpha_n\}$ , 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  $\alpha_i$ , 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

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

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

in this area, can be used to select relevant polynomials. In this study, we shall use the *Least Angle Regression* (LAR) algorithm, introduced in [9], and already used in [1] to polynomial basis selection. Let us now recall some important features of this algorithm. First of all, the LAR returns a collection of meta-models that are less and less sparse (the last iteration is the classical least-square solution) in  $\text{card}(\mathcal{B})$  iterations. The meta-model, which is selected in this collection, is determined by estimation of a given indicator. This point will be discussed later on. So the LAR algorithm proceeds as follows:

- Set all coefficients  $C_\alpha$  to 0.
- Find the polynomial that is the most correlated with the response, say  $\phi_{j_1}$ .
- Take the largest possible step in the direction of  $\phi_{j_1}$  until another polynomial, say  $\phi_{j_2}$ , has as much correlation with the current residual.
- Define the equiangular direction between  $\phi_{j_1}$  and  $\phi_{j_2}$  and take the largest step in this direction, until a new predictor has as much correlation with the residual, and so on.

Note that, in [1], the LAR is only used for variable selection and coefficients  $C_\alpha$  are calculated by least-square method for each sparse basis.

Here, once the LAR algorithm returned the  $P$  meta-models, one has to choose the *best* one according to an error criterion. Several regression error measurements have been tested in the literature, see for example [3] and the discussion in [9]. We recall two of them, used in next sections: the classical coefficient of determination  $R^2$  and the leave-one-out error  $Q^2$ .

## 2.4. Regression error measurement

### 2.4.1. Coefficient of determination $R^2$

The coefficient of determination is defined as

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2},$$

where  $y_i$  and  $\hat{y}_i$  are respectively evaluations of the *real* model and of the meta-model at points  $\mathbf{x}_i$ , while

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i.$$

This coefficient measures the generalization error, i.e. the part of variance of the real model explained by the meta-model [6]. A major drawback of this coefficient is that it only takes into account the points of the experimental design. Moreover, it tends to one when the number of polynomials in the meta-model increases, which makes it irrelevant to find over-fitting phenomena. To take into account the number of terms in the meta-model, other indicators are available, for example Mallows's Cp [15].

Considering these drawbacks, coefficient  $R^2$  has to be used with care and its value has to be compared with other indicators that are less sensitive to over-fitting, e.g. the leave-one-out error.

### 2.4.2. Leave-one-out error

Leave-one-out is a particular case of cross-validation, where the size of validation set is one. The idea is to leave one point out of the design of experiments, to create a meta-model on this new design, then to evaluate the residual at the left point and, finally, to loop on each point of the original design of experiments. This methodology can be quantified through the following formula, where  $\hat{y}^{(-i)}$  stands for the value at  $\mathbf{x}_i$  of the meta-model built on

the experimental design in which point  $\mathbf{x}_i$  has been removed

$$Q^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}^{(-i)})^2}{\sum_{i=1}^N (y_i - \bar{y})^2}.$$

As for the  $R^2$  coefficient, the leave-one-out coefficient can be penalized by the number of terms in the meta-model. Such a penalized  $Q^2$  is used for the selection of the meta-model within the collection obtained by the LAR algorithm (see [3] for more details).

## 2.5. Construction of the design of experiments

The construction of an experimental design consists in choosing a method to sample in the space of input variables in order to compute the  $N$  exact values  $\mathbf{y}(\mathbf{X})$  used in the least-square problem. In our study, we mainly used Monte Carlo sampling which is a random sampling in the joint PDF of the input variables. A comparison with *Latin Hypercube Sampling* (LHS) is performed Section 4.1.3. LHS design of experiments splits the input space into equiproportional subspaces and allows only one sample per subspace which leads to a better filling of the space.

## 2.6. Post processing

### 2.6.1. Statistical moments

Once the PCE is built (i.e. once the basis is chosen and the coefficients are computed) statistical moments are computed analytically. As the decomposition functions  $\phi_\alpha$  have nice properties of orthogonality and zero mean, it can be shown (see [26]) that

$$\mathbb{E}[\hat{y}(\mathbf{X})] = C_0,$$

and

$$\text{Var}[\hat{y}(\mathbf{X})] = \sum_{\alpha \in \mathcal{A}} C_\alpha^2 \mathbb{E}[\phi_\alpha^2(\mathbf{X})].$$

### 2.6.2. Sensitivity analysis

The idea, pointed out in [26], is to identify the polynomial chaos expansion with an ANOVA decomposition (see [18] for example). For this purpose, Eq. (4) is rewritten according to

$$\begin{aligned} \hat{y}(\mathbf{X}) = & y_0 + \sum_{i=1}^N \sum_{\alpha \in L_i} C_\alpha \phi_\alpha(X_i) + \sum_{i_1=1}^{N-1} \sum_{i_2=i_1+1}^N \sum_{\alpha \in L_{i_1, i_2}} C_\alpha \phi_\alpha(X_{i_1}, X_{i_2}) \\ & + \dots + \sum_{i_1=1}^{N-s+1} \dots \sum_{i_s=i_{s-1}+1}^N \sum_{\alpha \in L_{i_1, \dots, i_s}} C_\alpha \phi_\alpha(X_{i_1}, \dots, X_{i_s}) \\ & + \dots + \sum_{\alpha \in L_{1, \dots, N}} C_\alpha \phi_\alpha(X_1, \dots, X_N), \end{aligned} \quad (5)$$

where  $L_{i_1, \dots, i_s}$  represent sets

$$L_{i_1, \dots, i_s} = \left\{ \alpha = (\alpha_k)_{k=1, \dots, n} \left/ \begin{array}{l} \alpha_k \in \mathbb{N}^*, k \in (i_1, \dots, i_s) \\ \alpha_k = 0, k \notin (i_1, \dots, i_s) \end{array} \right. \right\}.$$

Then the identification with the ANOVA decomposition is straightforward. Moreover, sensitivity indices, or Sobol' indices [22], derived from this decomposition, can be computed with coefficients of the sparse polynomial chaos decomposition [4],

$$\hat{S}_{i_1, \dots, i_s} = \frac{\sum_{\alpha \in L_{i_1, \dots, i_s}} C_\alpha^2 \mathbb{E}[\phi_\alpha^2(X_{i_1}, \dots, X_{i_s})]}{\text{Var}[\hat{y}(\mathbf{X})]}. \quad (6)$$

The most common sensitivity indices are the first order ones and the total ones. The first order sensitivity index of variable  $X_i$ , called  $S_i$ , gives the part of the output variance explained by the randomness in the sole input variable  $X_i$ . According to the above notation,

its approximation reads

$$\widehat{S}_i = \frac{\sum_{\alpha \in L_i^+} C_{\alpha}^2 \mathbb{E}[\phi_{\alpha}^2(X_i)]}{\text{Var}[\widehat{y}(\mathbf{X})]}.$$

Total sensitivity indices of an input variable  $X_i$ , called  $S_{T_i}$ , takes into account all interactions between  $X_i$  and all other variables. Its estimate by polynomial chaos expansion is given by

$$\widehat{S}_{T_i} = \frac{\sum_{\alpha \in L_i^+} C_{\alpha}^2 \mathbb{E}[\phi_{\alpha}^2(X_i)]}{\text{Var}[\widehat{y}(\mathbf{X})]},$$

where  $L_i^+$  is the set  $L_i^+ = \{\alpha \in \mathcal{A} / \alpha_i \neq 0\}$ . Let us remark that, according to Eqs. (5) and (6), if there is no interaction in the model, the sum of first order indices is one and total index  $S_{T_i}$  is equal to first order index  $S_i$  for every variable  $X_i$ .

This section shows how the quality of sensitivity indices is linked to the quality of the approximation by the meta-model. Nevertheless, accuracy needed for sensitivity analysis is strongly dependent on its aim (variable selection, variable ranking, variance reduction). So it seems difficult to use a global quality criterion (like  $R^2$  or  $Q^2$  described before) to reach the level of accuracy needed for sensitivity analysis. For example, in [4], the authors propose to target  $Q^2 \in [0.990, 0.999]$  to have a correct approximation of sensitivity indices. In order to have a goal-oriented information, we propose here to use bootstrap re-sampling on the design of experiments to build confidence intervals on sensitivity indices approximated by PCE.

### 3. Bootstrap re-sampling applied to polynomial chaos expansion

#### 3.1. Construction of confidence intervals by bootstrap re-sampling

Bootstrap [10] is a re-sampling method which aims at determining confidence intervals on a quantity of interest using only one design of experiments. It is well suited when the computation of this quantity is time-consuming and use of replicates is then impossible. The main idea is to create several new designs of experiments, say  $B$ , by drawing with replacement in the first one (source design) and, then, to use these new designs to get an empirical distribution of the statistic variables calculated on these designs. This methodology has already been applied on several surrogate models, as in [13,11].

Here, we are interested in sensitivity indices, denoted as collection  $\mathbf{S}$ . PCE of the response of interest gives an estimator of these indices, denoted as  $\widehat{\mathbf{S}}$ . For every new design  $D_k$  ( $k = 1, \dots, B$ ), the methodology of sparse PCE described in Section 2 is used, leading to sensitivity indices collections  $\widehat{\mathbf{S}}_k$  for design  $D_k$  (it should be noted that in order to avoid ill-conditioning of regression matrix  $\underline{\Phi}$  defined in Eq. (3), the size of the design of experiments is three times higher than  $\text{card}(\mathcal{A})$ , see [19]). After the computation of the  $B$  sensitivity indices collections (one per re-sampled design), empirical confidence intervals are built. Let  $\widehat{S}_i^*$  and  $\text{Var}\widehat{S}_i$  be the estimators of the mean and variance of the empirical distribution of the  $B$  collections of indices  $\widehat{S}_{ik}$ . Different types of confidence intervals can be obtained from bootstrap re-sampling.

- The first one is called *standard interval*,

$$S_i \pm u_{1-\alpha/2} \sqrt{\text{Var}\widehat{S}_i}, \quad i \in [1, \dots, n],$$

where  $u_{1-\alpha/2}$  is the  $1-\alpha/2$  quantile of the standard normal distribution. This confidence interval is based on an asymptotic approximation of the bootstrap distribution by a normal one.

- The second one is the percentile bootstrap. Confidence interval is

$$\widehat{S}_{i[\alpha/2]} \leq S_i \leq \widehat{S}_{i[1-\alpha/2]}, \quad i \in [1, \dots, n],$$

where  $\widehat{S}_{i[\alpha/2]}$  and  $\widehat{S}_{i[1-\alpha/2]}$  are the  $\alpha/2$  and  $1-\alpha/2$  empirical quantiles.

This interval does not need any hypothesis on  $\widehat{S}_i$  distribution, but needs a lot of re-sampling  $B$  (higher than 500, see [16]) in order to approximate these quantiles with a sufficient precision. The setting of this parameter is going to be discussed in Section 4.

- The last one is the bias corrected and accelerated bootstrap ( $BC_a$ ), as first introduced in [8]. The main idea is to assume that normality of the bootstrap distribution can be achieved by some transformation. In the following we only describe the construction of the intervals. For details on the theoretical basis, the reader is referred to [8] and also to [7] for a review on bootstrap confidence intervals.  $BC_a$  confidence intervals are

$$\widehat{S}_{i[\alpha_1]} \leq S_i \leq \widehat{S}_{i[\alpha_2]}, \quad i \in [1, \dots, n],$$

where  $\widehat{S}_{i[\alpha_1]}$  and  $\widehat{S}_{i[\alpha_2]}$  are the  $\alpha_1$  and  $\alpha_2$  empirical quantiles of the bootstrap distribution.  $\alpha_1$  and  $\alpha_2$  are defined by

$$\alpha_1 = \Phi\left(u_0 + \frac{u_0 + u_{\alpha/2}}{1 - a(u_0 + u_{\alpha/2})}\right)$$

and

$$\alpha_2 = \Phi\left(u_0 + \frac{u_0 + u_{1-\alpha/2}}{1 - a(u_0 + u_{1-\alpha/2})}\right)$$

where  $\Phi$  is the standard normal cumulative distribution function,  $u_{\alpha/2} = \Phi^{-1}(\alpha/2)$  and  $u_{1-\alpha/2} = \Phi^{-1}(1-\alpha/2)$ .  $u_0$  is called the *bias-correction* and it is defined by  $u_0 = \Phi^{-1}(\text{card}(\widehat{S}_{ik} < \widehat{S}_i))$ ,  $k \in [1, B]$ .  $a$  is called the *acceleration* and it is linked to an idealized transformation of the bootstrap distribution to normality. This parameter is not known in practice but it is shown in [8] that a correct approximation of  $a$  is

$$a \approx \widehat{a} = \frac{\sum_{j=1}^N (\widehat{S}_{i0} - \widehat{S}_{i-j})^3}{6\{\sum_{j=1}^N (\widehat{S}_{i0} - \widehat{S}_{i-j})^2\}^{3/2}} \quad (7)$$

where  $\widehat{S}_{i-j}$  is the estimation of  $S_i$  removing the  $j$ th point of the design of experiments, and  $\widehat{S}_{i0}$  is the mean of the  $N$   $\widehat{S}_{i-j}$  values. This interval is supposed to be more accurate as it takes into account some characteristic of the empirical distribution. It should be noted that it is a correction of the percentile bootstrap. Indeed if the correction terms  $u_0$  and  $a$  are null, then  $BC_a$  is equal to percentile bootstrap.

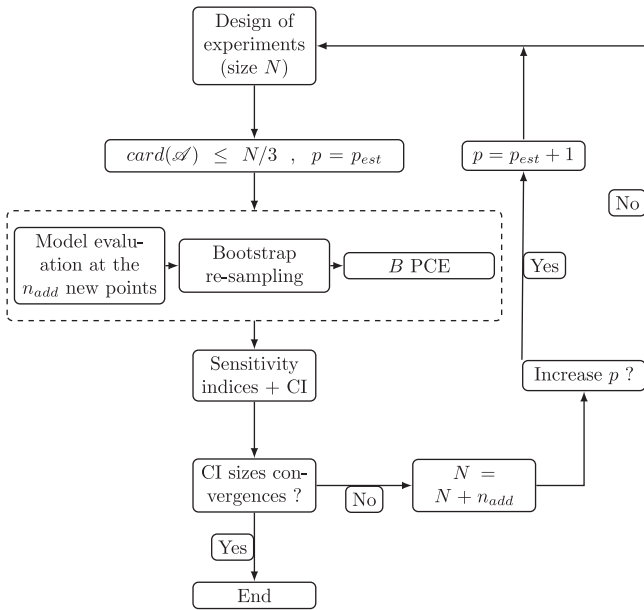
In the following numerical examples, we mainly used percentile bootstrap.  $BC_a$  procedure is illustrated but it should be noted that the computation of the acceleration is time consuming. Hypothesis of normality for *standard interval* is too restrictive and never verified in practice. This is the reason why we will not use this one.

Considering informations given by these confidence intervals, we now propose a sequential strategy to create a design of experiments which eventually provides a required accuracy on the sensitivity indices.

#### 3.2. Sequential construction of an optimized design of experiments

Here an application of bootstrap re-sampling is presented in order to minimize the number of points in the design of experiments, so as to get sensitivity indices from PCE with a fixed level of confidence. The algorithm described here is inspired from a previous work [16], using bootstrap re-sampling in a similar way on reliability indices. Our methodology, summarized in Fig. 1, is split in five main steps:





**Fig. 1.** Sequential construction of the experimental design leading to sensitivity indices with a fixed level of confidence.

1. An experimental design of size  $N$  is used to build a polynomial chaos meta-model. Parameters for PCE construction are the degree  $p$  of the candidate basis  $\mathcal{B}$  and the maximal size of the polynomial basis selected by the LAR algorithm, *i.e.*  $\text{card}(\mathcal{A}) \leq N/3$ . Note that this last condition is due to the sparse nature of the selected meta-model. This can be adjusted according to the penalization coefficient applied to leave-one-out error  $Q^2$ , used in selection process. In our studies, during bootstrap re-sampling, all the PCEs are sparse enough to reach this condition. If it is not the case, one can increase the penalization coefficient. At first iteration,  $N$  is initialized by the user as well as the estimation  $p_{est}$  of  $p$ .
2.  $B$  bootstrapped samples are used to construct the 95% confidence intervals (CIs) as presented in the previous section. This number  $B$  depends on the complexity of the response. Discussion about the choice of this parameter is presented later on. Three different confidence intervals being described in Section 3.1, the difference between them will be also discussed.
3. As far as convergence criterion is concerned, we choose to stop the algorithm at the iteration for which all CIs sizes have reached a given range  $[\hat{S}_{i[\alpha/2]}, \hat{S}_{i[1-\alpha/2]}]$  for  $\alpha = 0.05$ , which is formulated as a function of the maximum mean value of sensitivity indices (let us recall that, if there is no interaction, the sum of the first order sensitivity indices is equal to one). Once again, this convergence criterion will be discussed in the next section.
4. If the convergence criterion is not reached,  $n_{add}$  new points are added to the experimental design and a new iteration starts. The value of  $n_{add}$  is a compromise between the computation time of a single model evaluation and the computation time of one algorithm iteration ( $B$  PCE constructions). If  $n_{add}$  is too small, most of the time is spent during bootstrap iteration. If it is too large, the algorithm converges in less iterations but the final experimental design could be far from an optimal one. If the convergence criterion is reached, one obtains sensitivity indices with enough confidence, using an optimal experimental design. Sampling schemes that can be used to build and increase the experimental design are discussed in the following.
5. The last point deals with degree  $p$  of the candidate basis  $\mathcal{B}$ . If it is too low, the convergence will never be reached and, if it is too high,  $\text{card}(\mathcal{B})$  is large which makes the LAR algorithm time consuming. A way to increase it, if necessary, is described in the next section.

In conclusion this strategy leads to a PCE allowing to compute sensitivity indices with a given accuracy. It should be noted that this algorithm relies on the fact that PCE converges to the quantity of interest when the number of polynomials in the expansion and the degree of the candidate basis increase.

#### 4. Numerical examples

This section discusses the choice of parameters in the method presented above. Let us first recall that these parameters are:

- The number of bootstrap samples  $B$ .
- The methodology used to create the design of experiments and to increase its size.
- The degree  $p$  of the full candidate polynomial basis  $\mathcal{B}$  and, possibly, the way to increase it.
- The type of confidence interval ( $BC_a$  or percentile bootstrap).
- The convergence criterion based on the confidence intervals size.

All these points are enlightened through the example of the Ishigami function.

##### 4.1. Application to Ishigami function

###### 4.1.1. Presentation

The Ishigami function [21] is a well-known test case for sensitivity analysis because, among its three parameters, two have close first order sensitivity indices and one appears only in interaction. It is defined by

$$Y = \sin X_1 + 7 \sin^2 X_2 + 0.1X_3^4 \sin X_1,$$

where variables  $X_j$  have uniform distribution over the range  $[-\pi, \pi]$ . Sensitivity indices may be computed analytically as presented in Table 2.

In the following, all results are obtained by the methodology presented in Section 3.2.

###### 4.1.2. Number of bootstrap samples $B$

Here, our goal is to check classical recommendation on the number of bootstrap samples  $B$ . Usually, building a 95% confidence interval by percentile bootstrap or  $BC_a$  requires between 500 and 1000 bootstrap re-sampling [10]. As  $BC_a$  is only a correction of percentile bootstrap, all the following CI are constructed by percentile bootstrap. In order to fix the size of re-sampling, we propose to increase it from 100 to 1000 and to observe the variation of the lower and upper sensitivity indices CI boundaries for  $\alpha = 0.05$ . This is performed using a sparse PCE of the function with a full candidate basis  $\mathcal{B}$  of degree  $p = 10$ .

Fig. 2 shows the evolution of  $\hat{S}_{X_1[0.05/2]}$  (first order sensitivity index of variable  $X_1$ , lower CI boundary), and Fig. 3 shows  $\hat{S}_{X_1[1-0.05/2]}$  (the upper CI boundary). In both cases, a comparison is made for several sizes of design of experiments.

**Table 2**  
Ishigami function—sensitivity indices.

Variable	Analytical values	
	$S_i$	$S_{T_i}$
$X_1$	0.3138	0.5574
$X_2$	0.4424	0.4424
$X_3$	0.0	0.2436

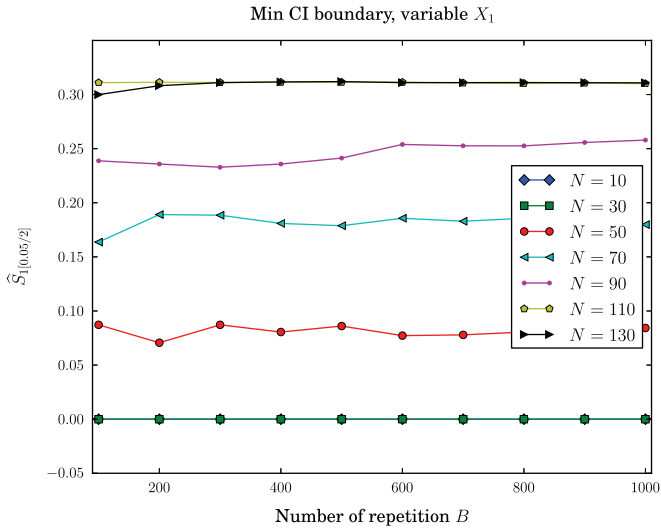


Fig. 2. Evolution of  $\widehat{S}_{X_1|[0.05/2]}$  vs. number of bootstrap samples  $B$  and size of the experimental design  $N$ .

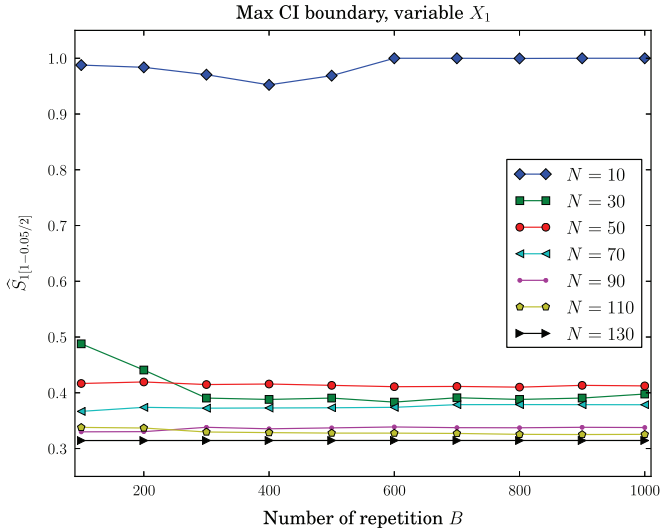


Fig. 3. Evolution of  $\widehat{S}_{X_1|[1-0.05/2]}$  vs. number of bootstrap samples  $B$  and size of the experimental design  $N$ .

Such a comparison was done for the three variables and the same two conclusions can be drawn:

1. The bigger the design of experiments is, the faster is the convergence as a function of  $B$ . This allows one to be confident in the fact that number  $B$  can be fixed, in a conservative way, at the first iteration.
2. Even for small design of experiments sizes, an admissible convergence in  $B$  is reached as soon as  $B=700$ . So, this value will be kept in the following.

#### 4.1.3. Choice of the methodology to create the design of experiments

As an iterative algorithm is proposed to increase the experimental design, the position of the samples in the input variables space has to be discussed. Several experimental design strategies are possible namely Monte Carlo sampling (random sampling using the joint PDF of the input variables) or *Latin hypercube sampling* (LHS).

Bootstrap re-sampling is usually performed on random experiments. Nevertheless, the next section gives a comparison between

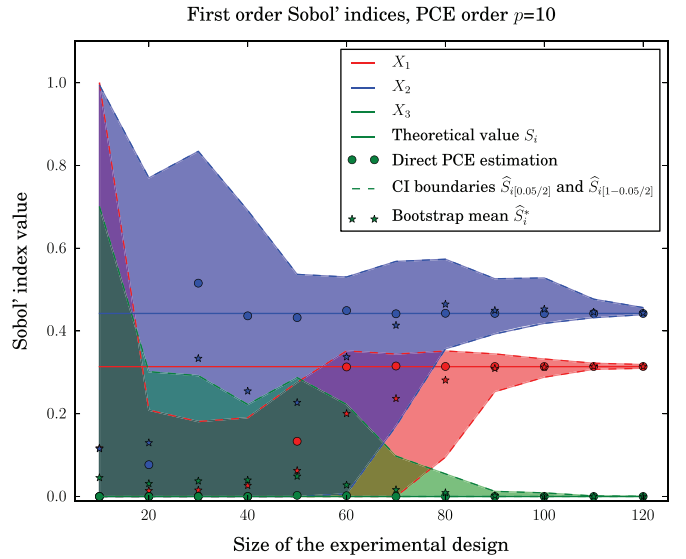


Fig. 4. Evolution of first order sensitivity indices—Monte Carlo experimental design.

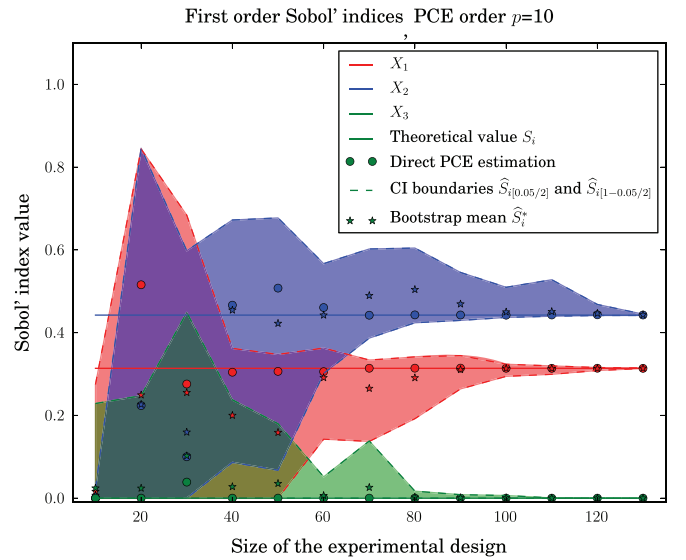


Fig. 5. Evolution of first order sensitivity indices—LHS experimental design.

results obtained with a Monte Carlo design of experiments and a LHS one. In this last case, note that a new LHS experimental design is built at each iteration. Like in the previous section, sensitivity indices are obtained by sparse polynomial chaos expansion with a full candidate basis of degree  $p=10$  and confidence intervals are built with 700 bootstrap repetitions with percentile bootstrap method.

Fig. 4 shows the evolution of the first order sensitivity indices ( $\widehat{S}_{X_1}, \widehat{S}_{X_2}, \widehat{S}_{X_3}$ ) versus the experimental design size for a Monte Carlo experimental design, whereas Fig. 5 presents the same results obtained with a LHS experimental design at each iteration. Legend *Direct PCE estimation* stands for sensitivity indices computed by PCE built on all the design of experiments. Finally, the values of  $R^2$  and  $Q^2$  (see Section 2.3) versus the size of the experimental designs are also given in Table 3.

A similar study was carried out on total sensitivity indices and led to the same conclusions. First, we can observe that the meta-models, built on the whole experimental design, are generally more accurate with less points when a LHS experimental design is used, as shown by  $R^2$  and  $Q^2$ . It should also be noted that, for



**Table 3**

Direct PCE estimation,  $R^2$  and  $Q^2$  values—first order sensitivity indices, Monte Carlo experimental design and LHS experimental design.

Monte Carlo			LHS		
ExD size	$R^2$	$Q^2$	ExD size	$R^2$	$Q^2$
10	0.982	0.955	10	0.987	0.943
20	0.957	0.906	20	0.889	0.657
30	0.800	0.651	30	0.996	0.953
40	0.712	0.564	40	1.000	0.997
50	0.971	0.867	50	1.000	0.996
60	0.987	0.972	60	1.000	0.994
70–120	1.000	1.000	70–130	1.000	1.000

$N > 30$ ,  $R^2$  and  $Q^2$  are quite different (especially for Monte Carlo design of experiment) which reveals the interest of the comparison between several indicators. Nevertheless, both design strategies lead to similar confidence interval convergence in a very close number of iterations. Finally, this numerical example does not allow to conclude on the interest of using LHS in our context.

#### 4.1.4. Polynomial basis degree

We recall (see Section 2.1) that the polynomial chaos expansion is based on a projection of the stochastic response onto a family of polynomials contained in a full candidate basis  $\mathcal{B}$ , made of all suitable polynomials up to a degree  $p$ . This section is devoted to the choice of this degree  $p$ . In the previous examples, a large value of  $p$  was chosen because:

1. It is known that degree 7 is necessary to properly approximate the Ishigami function (see [26]).
2. The function has only 3 input parameters. So the time spent by the LAR algorithm remains reasonable.

Considering these two points, it is clear that, in an industrial case with many variables (greater than 10 for example), the size of the candidate basis  $\mathcal{B}$  can make the LAR algorithm time consuming. Moreover choosing the degree a priori needs an expert judgment and, if it may be realistic in most of the cases, it could be sometimes difficult to evaluate the degree of the stochastic response. In order to tackle this problem, two different approaches are proposed and tested here.

1. If it is possible to have an accurate estimate of degree  $p$  (it is called  $p_{est}$ ) that is necessary to well approximate the response, and if the number of variables is not too important, one can choose  $p = p_{est} + \delta p$  where  $\delta p$  is equal to 1, 2 or 3.
2. If the number of input parameters is important, we propose to choose  $p = p_{est}$ , even if this value is hazardous. Then, the degree is increased following this empirical rule: if, after a few iterations (in practice four iterations appears as a good guess), the maximal confidence interval size is not divided by two, degree  $p$  of the candidate basis  $\mathcal{B}$  is increased by one.

As previously, these two strategies are illustrated on Ishigami function. For strategy #1,  $p$  is chosen equal to 10 and, for strategy #2,  $p_{est}$  is chosen equal to 5 and 3. Fig. 6 gives the results obtained by the first strategy for which  $R^2$  and  $Q^2$  values obtained with direct PCE construction are given by Table 3 (Monte Carlo column). Figs. 7 and 8 correspond to the second one for  $p_{est} = 5$  and  $p_{est} = 3$  respectively and  $R^2$  and  $Q^2$  values obtained with direct PCE construction are given by Table 4. In both cases, the experimental design is built by Monte Carlo sampling, confidence intervals are

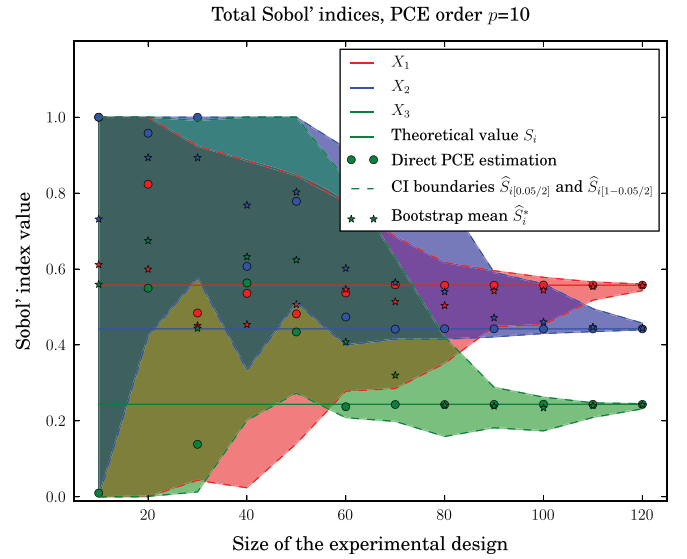


Fig. 6. Evolution of total sensitivity indices—first strategy— $p=10$ .

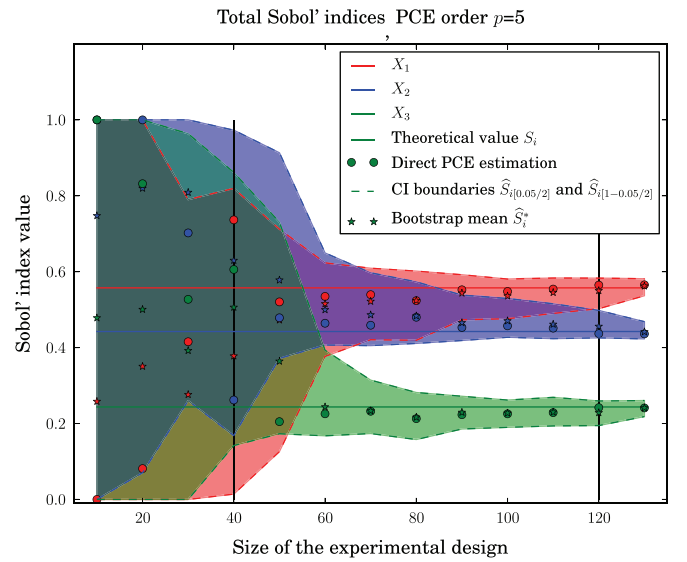


Fig. 7. Evolution of total sensitivity indices—second strategy—initial value of  $p_{est} = 5$ — $p$  increases by 1 at each vertical line.

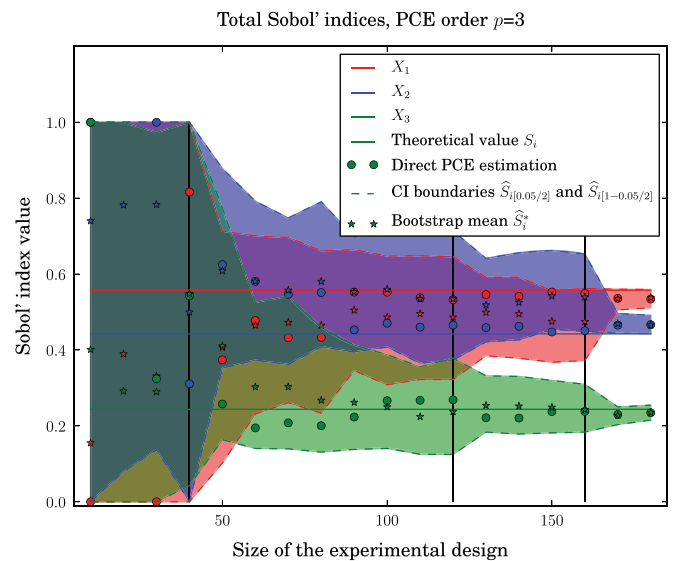


Fig. 8. Evolution of total sensitivity indices—second strategy—initial value of  $p_{est} = 3$ — $p$  increases by 1 at each vertical line.

**Table 4**  
Direct PCE estimation,  $R^2$  and  $Q^2$  values—total sensitivity indices— $p_{est} = 5$  and  $p_{est} = 3$ .

$p_{est} = 5$				$p_{est} = 3$			
ExD size	$p$	$R^2$	$Q^2$	ExD size	$p$	$R^2$	$Q^2$
10	5	0.480	0.171	10	3	0.480	0.171
20	5	0.714	0.543	20	3	0.000	-0.108
30	5	0.929	0.794	30	3	0.373	0.195
40	5	0.797	0.692	40	3	0.561	0.309
50	6	0.973	0.934	50	4	0.820	0.699
60	6	0.988	0.974	60	4	0.863	0.784
70	6	0.988	0.976	70	4	0.812	0.760
80	6	0.992	0.983	80	4	0.810	0.764
90	6	0.991	0.985	90	4	0.991	0.985
100	6	0.991	0.986	100	4	0.894	0.854
110	6	0.991	0.985	110	4	0.865	0.839
120	6	0.996	0.994	120	4	0.861	0.837
130	7	0.996	0.993	130	5	0.990	0.984
-	-	-	-	140	5	0.989	0.984
-	-	-	-	150	5	0.985	0.981
-	-	-	-	160	5	0.985	0.981
-	-	-	-	170	6	0.989	0.984
-	-	-	-	180	6	0.989	0.985

obtained by the percentile method with  $B=700$  and we look for the total sensitivity indices.

The comparison between the results of these two strategies allows us to show that, even if the expert judgment is bad (as in the second strategy), an increase of the degree based on the size of the confidence interval leads to convergence. Nevertheless, if  $p_{est}$  is far from the final value of  $p$ , the convergence becomes longer as illustrated by the case  $p_{est} = 3$ .

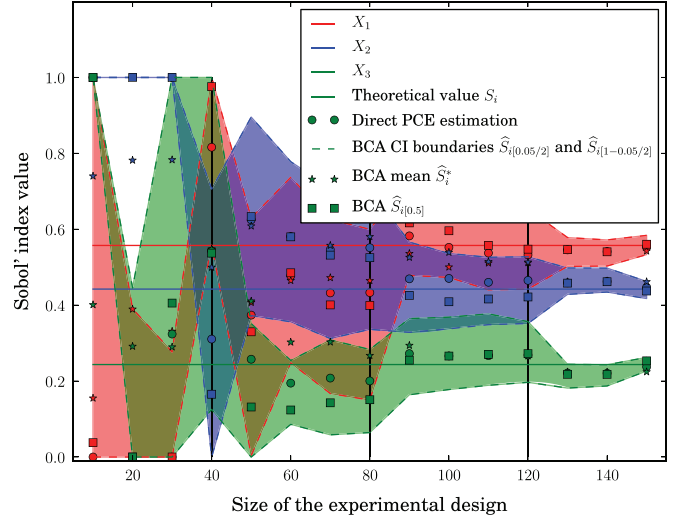
In conclusion, this section shows the importance of complementarity between a correct degree  $p$  for the candidate basis  $\mathcal{B}$  and enough points in the design of experiments. It also shows the capability of the method to converge even if the degree  $p_{est}$  is chosen far from what would be needed.

#### 4.1.5. Type of confidence interval

Section 3.1 introduced three different bootstrap confidence intervals. We recall that the *standard interval* is based on an asymptotic approximation of the bootstrap distribution by a normal one. Despite the fact that [12] shows the asymptotic normality of Sobol' index estimator computed by Monte Carlo or by a convergent meta-model, the *standard interval* is not studied here because the hypothesis of normality of the bootstrap distribution is never verified in practice (rejection of the assumption by the Kolmogorov normality test). In fact, as we proposed an adaptive design of experiment, at first iterations this asymptotic results are not yet verified and this type of CI is not adapted.

So this section discusses the difference between percentile bootstrap and the bias corrected and accelerated bootstrap ( $BC_a$ ). First of all, some remarks have to be done on correction terms in the  $BC_a$  method. These terms are numerically instable for small size of design of experiments (first iterations of the proposed methodology). For example the acceleration terms  $a$  (see Eq. (7)) needs the calculation of statistics  $\hat{S}_{i-j}$  (evaluation of the sensitivity indices removing the  $j$ th sample point from the experimental design) and for small design of experiments, it happens that the best PCE is a constant. Then the sensitivity indices cannot be calculated and  $\hat{S}_{i-j}$  has no sense. The results for small design of experiments are not relevant and must be canceled from the conclusion. The last remark concerns the computational cost of the  $BC_a$ . The calculation of the acceleration term is not negligible because of the terms  $\hat{S}_{i-j}$ .

Total Sobol' indices, PCE order  $p=3$



**Fig. 9.** Evolution of total sensitivity indices—initial value of  $p_{est} = 3 - p$  increases by 1 at each vertical line— $BC_a$  method.

**Table 5**

Direct PCE estimation,  $R^2$  and  $Q^2$  values—total sensitivity indices— $p_{est} = 3$ ,  $BC_a$  method.

ExD size	$p$	$R^2$	$Q^2$
10	3	0.480	0.171
20	3	0.000	-0.108
30	3	0.373	0.195
40	3	0.561	0.309
50	4	0.820	0.699
60	4	0.863	0.784
70	4	0.812	0.760
80	4	0.810	0.764
90	5	0.909	0.871
100	5	0.894	0.854
110	5	0.865	0.839
120	5	0.861	0.837
130	6	0.990	0.984
140	6	0.989	0.984
150	6	0.985	0.981

In order to compare the two types of confidence intervals, the example presented in Fig. 8 is considered. Total sensitivity indices are computed by a sparse PCE, the first degree  $p$  is chosen equal to 3, the experimental design is built by Monte Carlo sampling and the number of bootstrap samples is  $B=700$ . Fig. 9 presents results obtained with  $BC_a$  method. As previously, the  $R^2$  and  $Q^2$  values of the PCE built on the experimental design at each iteration are provided in Table 5. It should be noted that these values are quite similar with the ones in Table 4 as the same Monte Carlo sampling is used. Only the evolution of the degree  $p$  is different.

This figure clearly illustrates the stability problem of the  $BC_a$  method for small design of experiments. Then, compared to Fig. 8 (same condition but CI constructed by percentile method), convergence is reached with a smaller design of experiments (150 points instead of 180) and final confidence intervals are centered on the theoretical values.

Finally, percentile and  $BC_a$  methods give almost the same results at the last iteration but  $BC_a$  converges faster and it is more accurate in this example. Nevertheless, considering the instability problems for small design of experiments and the computational cost of this method, we will use the percentile method in the following example and compare with the  $BC_a$  method only in the industrial application.

#### 4.1.6. Convergence criterion

As defined in Section 3.2, convergence is reached when all the CIs sizes are less than  $x$  percent of the maximum bootstrap mean of the sensitivity indices. This part deals with the choice of parameter  $x$ . As an example, all previous figures were obtained with  $x = 10\%$ .

The choice of value  $x$  must be made according to the goal of the sensitivity analysis. If it only aims at ranking variables with a poor accuracy, a large value of  $x$  can be chosen (like 20 or 30). Conversely, if the sensitivity study has to be accurate (in order to work on variance reduction of some variables for example), value of  $x$  must be small (like 5 or 10).

Anyway, convergence is also depending on the maximum bootstrap mean of the sensitivity indices. It is impossible to know this value *a priori* but, as the sum of all first order sensitivity indices is one if there is no interaction, orders of magnitude given above will lead to correct results in almost every case. Problems will occur when the model has many variables of equal sensitivity indices or when the variables have their major effects in interaction. Then, a solution consists in following the convergence graph at each iteration and in deciding manually when to stop.

#### 4.1.7. Conclusion

The previous numerical experiments on Ishigami function help us to draw some general conclusions on the methodology. First, Section 4.1.2 justifies the arbitrary choice  $B = 700$ , as it shows that, when the experimental design size increases, CIs boundaries become less sensitive to the value of  $B$ . Then, even if LHS leads to a better  $Q^2$ , numerical tests using LHS to build and increase the design of experiments are not enough discriminatory (in terms of CIs size) to conclude about the efficiency of such method in our study. The choice of the basis degree  $p$  is, in an industrial context, linked to previous knowledges in the area. Anyway, Section 4.1.4 presents a simple way to increase this degree and find the correct one within a few iterations. Finally, Table 6 gives some numerical comparison between the theoretical values and the results of the proposed algorithm. It allows one to be confident in the capability of the method. These results are obtained at last iteration of the algorithm with the following parameters:  $B = 700$ ,  $p = 10$ , design of experiments built and increased by Monte Carlo sampling, convergence criterion fixed to 10%.

Considering all points discussed in this section, the methodology is now applied to the g-Sobol' function [21] with 8 parameters then to an industrial case.

#### 4.2. Application to g-Sobol' function

This function reads

$$Y = \prod_{i=1}^8 \frac{|4X_i - 2| + a_i}{1 + a_i},$$

where every  $X_i$  has a uniform distribution between 0 and 1. Parameters  $(a_i)_{1 \leq i \leq 8}$  take their values in  $\{0, 1, 4, 5, 9, 99, 99, 99\}$ . Theoretical sensitivity indices are given in Table 7.

**Table 6** Ishigami function—comparison between theoretical and numerical sensitivity indices.

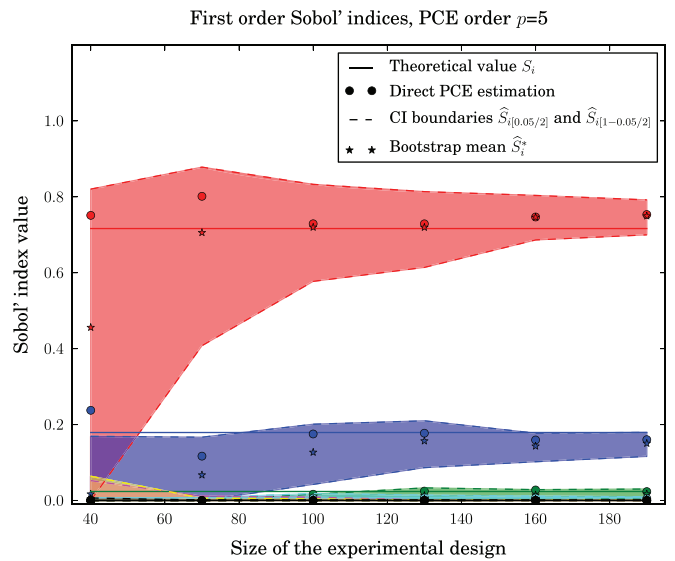
	$S$	$\hat{S}_{[\alpha/2]}$	$\hat{S}^*$	$\hat{S}_{[1-\alpha/2]}$	$S_T$	$\hat{S}_{T[\alpha/2]}$	$\hat{S}_T^*$	$\hat{S}_{T[1-\alpha/2]}$
$X_1$	0.3138	0.3099	0.3143	0.3189	0.5574	0.5432	0.5561	0.5602
$X_2$	0.4424	0.4396	0.4435	0.4565	0.4424	0.4400	0.4441	0.4572
$X_3$	0.0000	0.0000	0.0000	0.0001	0.2436	0.2319	0.2421	0.2451

Our methodology is applied in order to determine the first order sensitivity indices. Parameters of the method, discussed in Section 4, are set up as follows:

- Number of bootstrap samples:  $B = 700$ .
- Experimental design built by Monte Carlo sampling.
- Degree of the polynomial basis  $\beta$ :  $p = 5$ .
- Type of confidence interval: percentile
- Convergence criterion:  $\max(IC \text{ sizes}) < 0.1 \max(E[\hat{S}_i])$ .

**Table 7** g-Sobol function—first order and total sensitivity indices.

Variable	$S_i$	$S_{T_i}$
$X_1$	0.7165	0.7875
$X_2$	0.1791	0.2423
$X_3$	0.0237	0.0343
$X_4$	0.0072	0.0105
$X_{5,6,7,8}$	0.0001	0.0001



**Fig. 10.** Evolution of first order sensitivity indices— $p=5$ .

**Table 8** Direct PCE estimation,  $R^2$  and  $Q^2$  values—total sensitivity indices— $p_{est} = 5$ .

ExD size	$R^2$	$Q^2$
40	0.843	0.788
70	0.940	0.919
100	0.986	0.977
130	0.991	0.981
160	0.978	0.971
190	0.985	0.978

**Table 9** g-Sobol function—comparison between theoretical and numerical first order sensitivity indices.

Variable	$S_i$	Lower CI boundary	$\hat{S}_i^*$	Upper CI boundary
$X_1$	0.7165	0.7000	0.7496	0.7919
$X_2$	0.1791	0.1162	0.1504	0.1794
$X_3$	0.0237	0.0069	0.0172	0.0301
$X_4$	0.0072	0.0000	0.0038	0.0095
$X_{5,6,7,8}$	0.0001	0.0000	0.0000	[0.0001–0.0009]

Fig. 10 shows the evolution of confidence intervals versus the number of iterations. As previously,  $R^2$  and  $Q^2$  values are also given by Table 8. Table 9 gives a numerical comparison between theoretical values and estimations of the first order sensitivity indices obtained at last iteration.

This case shows that, when some variables have a major impact, it is not necessary to have a perfect approximation of the function to obtain accurate results on sensitivity indices. For example, the accuracy, related to results given in Table 9, is enough to select, rank variables and even to have correct relative importance information. Let us also remark that  $R^2 = 0.985$  and  $Q^2 = 0.978$  at the end of iterations. So an adaptive method, guided by values of  $R^2$  and  $Q^2$  (targeting  $R^2 > 0.99$  or  $Q^2 > 0.99$  for example), would lead in this case to a kind of over quality in the meta-model.

## 5. An industrial application: TARANIS

### 5.1. Context of the study

This example takes place in a reliability analysis of a satellite structure called TARANIS developed by CNES (Centre National d'Etudes Spatiales). Fig. 11 shows the finite element model used for calculations (about 380,000 degrees of freedom). The aim of this work was to study the reliability under static load of the satellite using second order surface response (see [19]). Several methods were used to select the most relevant variables for every response and it appeared that 14 variables were sufficient.

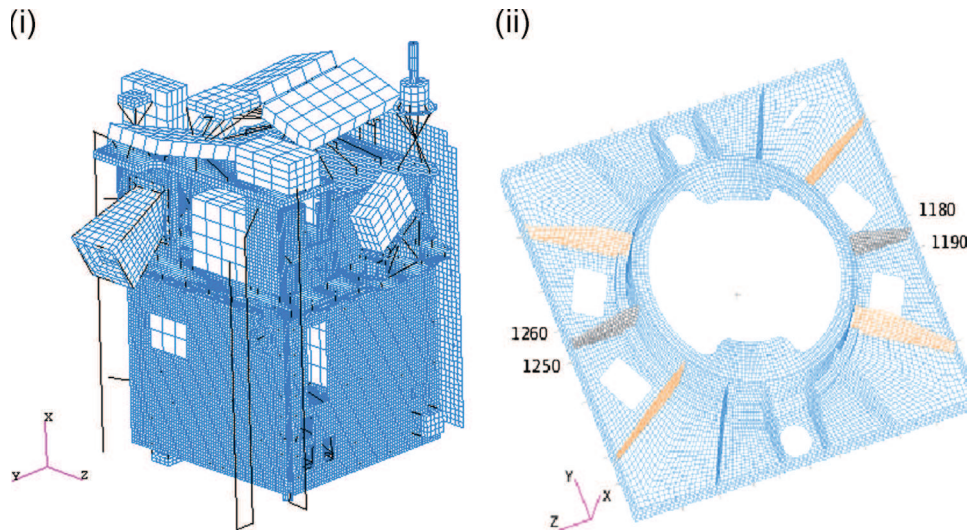


Fig. 11. (i) Finite element model of Taranis structure, (ii) Minus X panel.

More precisely, we focus here on one particular quantity of interest which is the maximum Von Mises stress in a particular panel under the following loading case: 9.75 g following X, -3.68 g following Y and 3.68 g following Z where  $g = 9.81 \text{ m s}^{-2}$  is the gravity intensity.

The selected 14 variables are described in Table 10. All probability densities are supposed to be Gaussian, means and variation coefficients (VCs) being given in Table 10. For this study, we aim at estimating the total sensitivity indices of every variable.

Following [19], a global sensitivity analysis was carried out using the calculation of sensitivity indices by Monte Carlo method on the surface response. This one was built using an experimental design of 186 points. Then, 2,000,000 Monte Carlo simulations of the surface response were necessary to obtain the sensitivity indices. Results are presented in Table 11, Column R-S. It can be noticed that the sum of all indices is close to one, which means there is almost no interaction between variables.

### 5.2. Results obtained with bootstrap re-sampling and polynomial chaos expansion

The algorithm presented in Section 3.2 is applied to the study of the total sensitivity indices of the 14 variables introduced previously. The parameters of the method are

- Number of bootstrap samples:  $B=700$ .
- Experimental design built by Monte Carlo sampling.

Table 10

Gaussian variables for the TARANIS model.

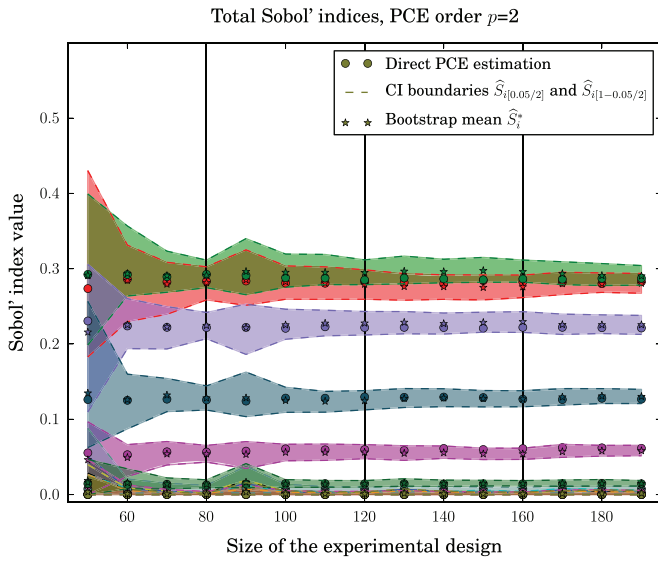
Variable	Short name	Mean	VC (%)
Acceleration following X	grav_X	-9.75 g	4
Acceleration following Y	grav_Y	-3.68 g	4
Acceleration following Z	grav_Z	-3.68 g	4
Thickness of the lower skin, panel -Z	pmz_tck_inf_skin	0.6 mm	6
Thickness of the upper skin, panel -Z	pmz_tck_sup_skin	0.6 mm	6
Thickness of the honeycomb, panel -Z	pmz_tck_honey	17.6 mm	6
Young modulus of the skins, panel -Z	pmz_E_skins	72.e6 MPa	4
Thickness of the lower skin, panel +Z	ppz_tck_inf_skin	0.6 mm	6
Thickness of the upper skin, panel +Z	ppz_tck_sup_skin	0.6 mm	6
Thickness of the honeycomb, panel +Z	ppz_tck_honey	17.6 mm	6
Young modulus of the skins, panel +Z	ppz_E_skins	72.e6 MPa	4
Young modulus of the panel -X panel	pmx_E	72.e6 MPa	4
Thickness of orange stiffeners in Fig. 11(ii), panel -X	pmx_tck_or_stif	18 mm	6
Thickness of black stiffeners in Fig. 11(ii), panel -X	pmx_tck_bl_stif	9.2 mm	6



**Table 11**

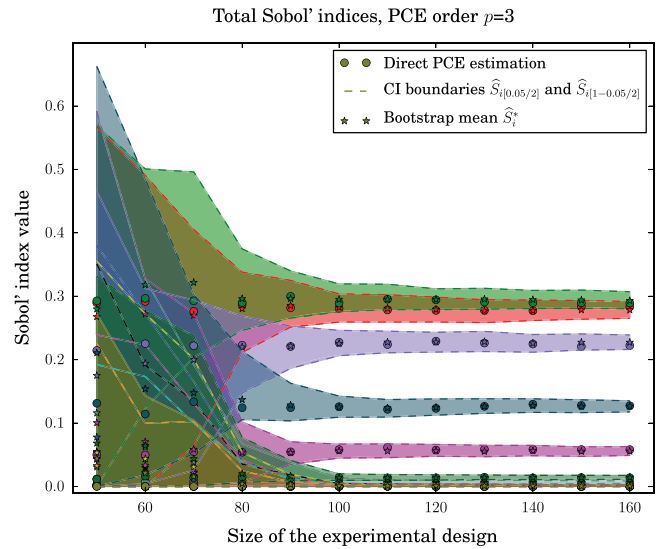
Comparison between surface response method and bootstrap re-sampling on sensitivity indices computed by PCE.

Variable	$p=2$ to $p=5$			
	R-S	B-Inf	Mean	B-Sup
ppz_tck_sup_skin	0.3047	0.2782	0.2903	0.3041
ppz_tck_inf_skin	0.2786	0.2674	0.2806	0.2937
ppz_E_skins	0.0577	0.0516	0.0582	0.0654
grav_Z	0.2174	0.2128	0.2255	0.2379
grav_Y	0.1286	0.1209	0.1302	0.1397
		$p=3$		
		B-Inf	Mean	B-Sup
ppz_tck_sup_skin		0.2803	0.2929	0.3071
ppz_tck_inf_skin		0.2652	0.2783	0.2917
ppz_E_skins		0.0490	0.0555	0.0627
grav_Z		0.2162	0.2272	0.2388
grav_Y		0.1176	0.1264	0.1349
		$p=4$		
		B-Inf	Mean	B-Sup
ppz_tck_sup_skin		0.2802	0.2936	0.3075
ppz_tck_inf_skin		0.2656	0.2792	0.2929
ppz_E_skins		0.0488	0.0557	0.0627
grav_Z		0.2128	0.2247	0.2371
grav_Y		0.1198	0.1295	0.1386
		$BC_a, p=2$		
		B-Inf	Corrected 50% quantile	B-Sup
ppz_tck_sup_skin		0.2822	0.2961	0.3081
ppz_tck_inf_skin		0.2730	0.2846	0.3004
ppz_E_skins		0.0508	0.0540	0.0610
grav_Z		0.2079	0.2162	0.2280
grav_Y		0.1208	0.1285	0.1382

**Fig. 12.** Evolution of total sensitivity indices— $p$  increases by 1 at each vertical line.

- Convergence criterion:  $\max(IC \text{ sizes}) < 0.1 \max(E[\hat{S}_i])$ .
- Type of confidence interval: percentile.
- Degree  $p$  of the polynomial basis  $\mathcal{B}$ : three choices are tested. First,  $p=2$  and it may increase; second,  $p=3$  and it cannot increase; finally,  $p=4$  and it cannot increase.

**Fig. 12** gives the evolution of the sensitivity indices when  $p=2$ , the vertical lines meaning an increase of one in the maximal

**Fig. 13.** Evolution of total sensitivity indices— $p=3$ .

degree of the candidate basis. **Figs. 13** and **14** depict this evolution for  $p=3$  and  $p=4$  respectively.

These figures exhibit an interesting phenomenon. If the degree  $p$  of the full candidate basis  $\mathcal{B}$  is high (**Fig. 14**), at first iterations, 100 points are necessary to reach *correct* CIs sizes. This shows that, if the candidate basis size is too large, the selection algorithm is disturbed and this leads to large CIs. Anyway, as a severe convergence criterion is chosen ( $\chi=10$ ), all choices converge between 160 ( $p=3$ ) and 190 ( $p=2-5$ ). But if the convergence



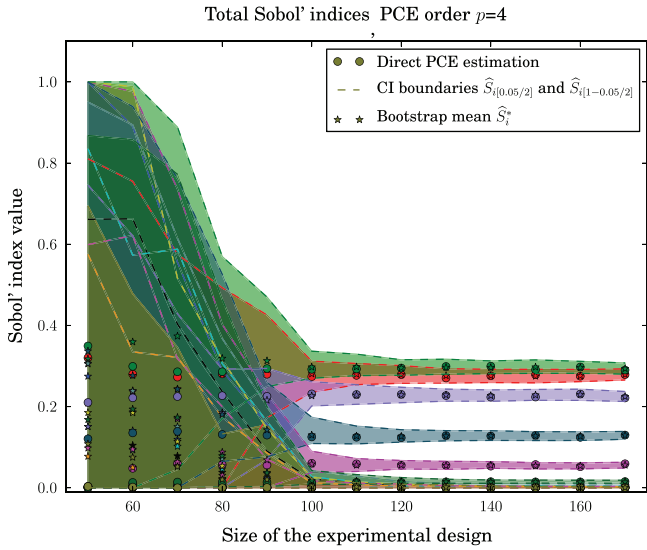


Fig. 14. Evolution of total sensitivity indices— $p=4$ .

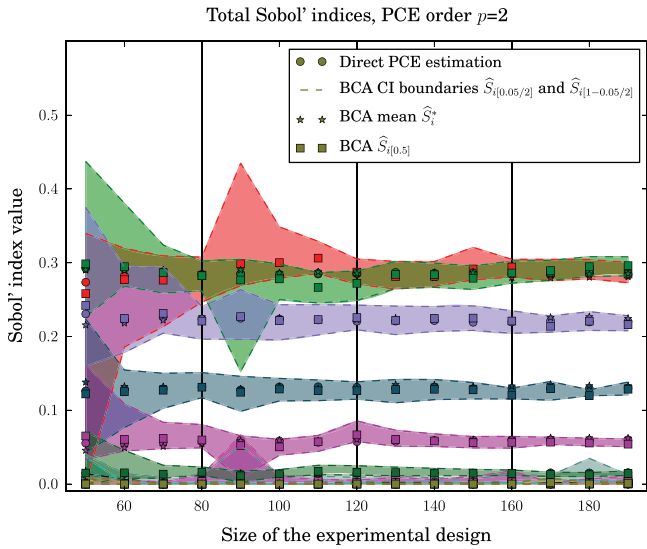


Fig. 15. Evolution of total sensitivity indices— $p=2$ ,  $BC_a$ .

criterion is relaxed at  $x=20$ , the first choice ( $p=2-5$ ) converges faster than the two others.

Finally, the  $BC_a$  method for the construction of sensitivity indices is also tested, with  $p=2$ , all other parameters remaining the same. Fig. 15 presents this case.

The comparison between Fig. 15 (using  $BC_a$ ) and Fig. 12 (using percentiles) reveals almost no difference between the two construction methods for the confidence intervals. This seems to show that for smooth responses, the  $BC_a$  method gives no advantage compared to the percentile one. Table 11 summarizes all results.

A comparison with the results obtained by the response surface method allows one to be confident in the capability of the method for solving industrial problems, as both methods lead to the same conclusion. Moreover confidence intervals, that are built by bootstrap re-sampling, also contain results given by the response surface method, which show the relevance of these intervals.

## 6. Conclusion

The methodology proposed in this paper aims at combining advantages of PCE with bootstrap re-sampling for the determination

of sensitivity indices of industrial models with a given degree of confidence. First of all, an efficient computation of sensitivity indices is obtained thanks to the PCE construction presented in [3]. Then, in order to evaluate confidence intervals for these estimated sensitivity indices, bootstrap re-sampling is applied to the design of experiments that is used to build the PCE. New PCE are constructed on each bootstrap sample, which leads to a collection of sensitivity indices and, finally, to empirical confidence intervals for every sensitivity index. Moreover, the iterative procedure introduced in Section 3.2 guides the construction of an adaptive experimental design to reach an accuracy objective, expressed on the sensitivity indices and not on the quality of the meta-model. Finally, this leads to an optimized design of experiments for the determination of sensitivity indices. Comparisons with classical meta-model error estimators show that, in some cases, confidence intervals on sensitivity indices are accurate enough whereas global error estimators on meta-model are bad (for example, see values of  $Q^2$  and  $R^2$  in the  $g$ -Sobol function compared to the size of confidence intervals). This reveals the interest of a sensitivity-indices-oriented methodology.

Nevertheless, this algorithm requires to set up five parameters. Their influence is discussed along Section 4 and allows to draw some conclusions:

- The number of bootstrap re-sampling  $B$  has to be sufficient to guarantee the convergence of the empirical confidence intervals. A convergence analysis as a function of  $B$  was carried out and the choice  $B=700$  appears to be a good one. This analysis also indicates that it is a conservative way to choose  $B$ .
- To create and increase the design of experiments, we use Monte Carlo sampling and LHS. It seems that LHS do not improve efficiency of the method.
- The degree  $p$  of the full candidate polynomial basis  $\mathcal{B}$  has to be chosen in priority according to previous knowledge. If it is not possible, it seems that an acceptable choice for most elastic stress analysis problems is  $p=3$ . Anyway, a method to increase this degree, linked to the size of confidence intervals, is tested and always leads to correct results in the presented applications. Note that, if it seems comfortable to choose an *a priori* high degree, it may perturb the PCE construction and slow down the convergence, as shown by the industrial application.
- Confidence intervals are mainly constructed using the percentile method. In the case of the Ishigami function the  $BC_a$  method converges faster but this conclusion is not confirmed by the industrial example. To conclude, in our application area, it seems that the  $BC_a$  procedure is not recommended as its numerical cost is important (compared to percentile method) and improvement in the convergence is not guaranteed.
- A convergence criterion, based on the size of the confidence intervals, has to be chosen according to the aim of sensitivity analysis. As seen in Section 4, one must keep in mind that a low convergence criterion can lead to an important number of iterations. A safer way is then to start from a high value, to observe results and then to restart the algorithm from this new starting point with a lower convergence criterion if previous results are not accurate enough.

Finally, let us discuss the numerical cost of the methodology. The number of evaluations of the numerical model used to build the PCE (*i.e.* the size of the final design of experiments) is supposed to be almost optimized for the aim of the PCE. Nevertheless, using a LHS design of experiments generally decreases the number of points needed to build PCE with a given accuracy on a global error criterion. The fact that our methodology is not sensitive to this LHS property seems to be due to bootstrap re-sampling. The computational time devoted to bootstrap re-sampling, is a function of the number  $B$  of samples and the complexity in the construction of the

PCE (number of input variables, size of the candidate polynomial bases), as  $B$  new PCE are built. In comparison with the solutions used in [20] and [13], where the polynomial basis does not change and only coefficients of the expansion are recalculated (least square minimization), our methodology is more expensive ( $B$  basis selections instead of 1) but it allows to take into account variations in the polynomial selection process.

## Acknowledgments

This work was partially supported by Centre National d'Etudes Spatiales and Thales Alenia Space. The authors would like to thank Nicolas Roussouly for fruitful discussions and reviewers for their relevant comments.

## References

- [1] Blatman G. Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis. PhD thesis, Université Blaise Pascal, Clermont-Ferrand, 2009.
- [2] Blatman G, Sudret B. An adaptive algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis. *Probabilistic Engineering Mechanics* 2010;25:183–97.
- [3] Blatman G, Sudret B. Adaptive sparse polynomial chaos expansion based on *least angle regression*. *Journal of Computational Physics* 2010;230:2345–67.
- [4] Blatman G, Sudret B. Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. *Reliability Engineering & System Safety* 2010;95(11):1216–29.
- [5] Castaings W, Borgonovo E, Morris MD, Tarantola S. Sampling strategies in density-based sensitivity analysis. *Environmental Modelling & Software* 2013;38:13–26.
- [6] Cornillon PA, Matzner-Lober E. *Régression théorie et applications*. Springer; 2007.
- [7] DiCiccio TJ, Efron B. Bootstrap confidence intervals. *Statistical Science* 1996;11(3):189–228.
- [8] Efron B. Better bootstrap confidence intervals. *Journal of the American Statistical Association* 1987;82(397):171–85.
- [9] Efron B, Hastie T, Johnstone I, Tibshirani R. Least angle regression. *Annals of Statistics* 2004;32:407–99.
- [10] Efron B, Tibshirani RJ. *An introduction to the bootstrap*. New York: Chapman & Hall; 1993.
- [11] Gayton N, Bourinet JM, Lemaire M. CQ2RS: a new statistical approach to the response surface method for reliability analysis. *Structural Safety* 2003;25(1):99–121.
- [12] Janon A, Klein T, Lagnoux A, Nodet M, Prieur C. Asymptotic normality and efficiency of two Sobol' index estimators. Technical report, INRIA, 2012.
- [13] Janon A, Nodet M, Prieur C. Confidence intervals for sensitivity indices using reduced-basis metamodels. Technical report, INRIA, 2010.
- [14] Lemaire M. *Fiabilité des structures –Couplage mécano-fiabiliste statique*. Hermès; 2005.
- [15] Mallows CL. Some comments on Cp. *Technometrics* 2000;42(1):87–94.
- [16] Notin A, Gayton N, Dulong JL, Lemaire M, Villon P. RPCM: a strategy to perform reliability analysis using polynomial chaos and resampling—application to fatigue design. *European Journal of Computational Mechanics* 2010;19(8):795–830.
- [17] Plischke E, Borgonovo E, Smith CL. Global sensitivity measures from given data. *European Journal of Operational Research* 2013;226:536–50.
- [18] Rahman S. Global sensitivity analysis by polynomial dimensional decomposition. *Reliability Engineering & System Safety* 2011;96(7):825–37.
- [19] Roussouly N. *Approche probabiliste pour la justification par analyse des structures spatiales*. PhD thesis, Université de Toulouse, Toulouse, 2011.
- [20] Roussouly N, Petitjean F, Salaün M. A new adaptive response surface method for reliability analysis. *Probabilistic Engineering Mechanics* 2013;32:103–15.
- [21] Saltelli A, Chan K, Scott EM, editors. *Sensitivity analysis*. John Wiley & Sons; 2000.
- [22] Sobol' IM. Sensitivity estimates for nonlinear mathematical models. *Mathematical Modeling* 1993;1:407–14.
- [23] Sobol' IM. Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematical Models and Computer Simulations* 2001;255:271–80.
- [24] Soize C, Ghanem R. Physical systems with random uncertainties: chaos representations with arbitrary probability measure. *SIAM Journal on Scientific Computing* 2004;26(2):395–410.
- [25] Storlie CB, Swiler LP, Helton JC, Sallaberry CJ. Implementation and evaluation of nonparametric regression procedures for sensitivity analysis of computationally demanding models. *Reliability Engineering & System Safety* 2009;94:1735–63.
- [26] Sudret B. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety* 2008;93(7):964–79.
- [27] Xiu D, Karniadakis GE. The Wiener–Askey polynomial chaos for stochastic differential equations. *SIAM Journal on Scientific Computing* 2002;24(2):619–44.