# Sequential Sensitivity Analysis of Expensive Black-Box Simulators With Metamodelling

Tom Van Steenkiste*, Joachim van der Herten, Ivo Couckuyt, Tom Dhaene

*Ghent University - imec*
*Technologiepark-Zwijnaarde 15*
*9052 Ghent*
*Belgium*

**Abstract**

Simulators are an essential tool in the design and analysis of complex systems. For many tasks in engineering, highly accurate simulators are used instead of expensive real-life experiments. As accuracy increases, these simulators are becoming computationally more expensive to evaluate. Sensitivity analysis provides information on the importance of the inputs with regard to the outputs. This does not only provide critical information about the workings of the system but also allows to discard inputs with low impact on the outputs. However, sensitivity analysis is an expensive process in terms of number of evaluations. Hence, metamodelling techniques are used to reduce the computational burden. In this paper we present, discuss and evaluate a novel algorithm for sequential variance-based and derivative-based sensitivity analysis of expensive black-box simulators using metamodelling. Two new stopping criteria are proposed on top of the traditional model error based stopping criteria. Extensive testing on benchmark problems and engineering use cases shows the sensitivity indices can be efficiently and accurately computed with a limited number of simulator evaluations.

*Keywords:* Sensitivity analysis, metamodelling, simulations, design of experiments, engineering design

---

*Corresponding author
    Email address: `tomd.vansteenkiste@ugent.be` (Tom Van Steenkiste)

## 1. Introduction

Simulations are a valuable tool in the design and analysis of complex systems. They offer researchers and engineers an analysis method without requiring numerous expensive real-life experiments or prototypes [1]. The improvement of simulation accuracy over the years has significantly increased the evaluation time and computational requirements [2]. This increase in complexity gives rise to the *curse of dimensionality* [3].

Data analysis techniques can be categorized as data-driven techniques in which data generated by a natural process is used or as model-driven techniques for which data has to be generated by evaluating a model or simulator [4]. In data driven analysis, the dataset is generated beforehand. An example of data driven sensitivity analysis is found in [5]. In this work, a model driven approach is followed, meaning that new samples can be sequentially gathered at the required locations.

Sensitivity analysis is a powerful method to reduce the complexity in subsequent analysis steps of the system and to gain critical insights into the system under study. In engineering, it is used, for example, to test the importance of parameters in crashworthiness tests of thin-walled structures of high-strength steel [6]. Several approaches can be distinguished such as variance-based methods [7] and derivative-based methods [8].

Although sensitivity analysis is used to reduce the complexity and hence evaluation time of simulators, it is often expensive in itself in terms of number of required evaluations. An efficient method to reduce the amount of simulator evaluations, is building a metamodel, also known as a surrogate model or response surface model [9], of the simulator [10].

There are various different metamodel types. Here, we explore the Kriging [11, 12], Gaussian Process (GP) [13], and Least-Squares Support Vector Machine (LS-SVM) [14] metamodels to compute both variance-based and derivative-based sensitivity indices. Another metamodel frequently used for sensitivity analysis is the Polynomial Chaos model [15]. The traditional methods to estimate these sensitivity indices are Monte Carlo or quasi-Monte Carlo methods [16, 17]. However, for metamodels of a

tensor product functional form, analytic derivations of the variance-based sensitivity indices can be made [18]. In this work, we extend this approach to other kernels and the GP and LS-SVM metamodel for variance-based indices. Furthermore, we prove that a similar derivation for the derivative-based sensitivity indices can be made. These equations can easily be applied to other metamodels with the same tensor product form such as Radial-Basis-Function Neural Networks [19]. The Kriging and Gaussian Process metamodels also allow for the derivation of the confidence bound on the sensitivity indices as demonstrated in [20, 21]. The focus of this paper however, is on the sequential computation of the sensitivity indices.

A metamodel can be constructed in several different ways. In the simplest form, a set of data points is chosen and evaluated, after which a metamodel is fitted on the generated data. The metamodelling accuracy and efficiency can be significantly improved using sequential sampling algorithms. These algorithms start with a small set of initial points and iteratively extend the dataset [22, 23, 24].

A stopping criterion is used to determine when the results are sufficiently accurate and the sequential sampling process can be stopped. Such criteria typically use the metamodel accuracy, assessed with k-fold Cross Validation (CV) using a specific error measure such as the Root Relative Squared Error ($CV_{RRSE}$) or Bayesian Estimation Error Quotient ($CV_{BEEQ}$) [25]. For sensitivity analysis however, instead of only using the metamodel accuracy, we show it is more efficient and interpretable to also define a stopping criterion based on the sensitivity indices directly. We define two stopping criteria based on the sensitivity indices called the Sensitivity Cross Validation (SCV) maximal variance criterion, SCV-$\max_{var}$, and mean variance criterion, SCV-$\mu_{var}$.

This work presents a complete, novel algorithm for sequential sensitivity analysis of expensive black-box simulators using the Kriging, GP or LS-SVM metamodel. The algorithm consists of analytic derivations of the variance-based or derivative-based sensitivity indices and the FLOLA-Voronoi sequential sampling strategy with the SCV-$\max_{var}$ or SCV-$\mu_{var}$ stopping criteria. For testing, we evaluate the algorithm on three common sensitivity analysis benchmark functions and two engineering applications.

In Section 2, sensitivity analysis is discussed. In Section 3, the process of metamodelling is discussed. In Section 4, the complete algorithm is presented. In Section 5,

the experimental setup is outlined and in Section 6, the results of these experiments are presented and discussed. Finally conclusions and future work are discussed in Section 7.

## 2. Sensitivity Analysis

Global sensitivity analysis is an important tool to determine how the output behavior is related to changes in the inputs [17]. In variance-based global sensitivity analysis, the variance in the input is related to the variance in the output. The simulator to be analyzed, represented as the $d$-dimensional function $f(x)$, is decomposed using a Hoeffding decomposition [26] (also known as an ANOVA decomposition) and the variance $V$ of the function is specified as a combination of the variances of the decomposed parts [16] according to

$$f(x_1,x_2,...,x_d) = f_0 + \sum_{i=1}^{d} \underbrace{f_i(x_i)}_{\text{main effects}} + \sum_{i_1=1}^{d}\sum_{i_2=i_1+1}^{d} \underbrace{f_{i_1i_2}(x_{i_1},x_{i_2})}_{\text{interaction effects}} + \ldots + \underbrace{f_{1..d}(x_1,...,x_d)}_{\text{interaction effects}},$$

$$V = \sum_{i=1}^{d} \underbrace{V_i}_{\text{main effects}} + \sum_{i_1=1}^{d}\sum_{i_2=i_1+1}^{d} \underbrace{V_{i_1i_2}}_{\text{interaction effects}} + \ldots + \underbrace{V_{1..d}}_{\text{interaction effects}},$$

where the inputs $x_i$ are assumed to be independent. As strong annihilating conditions apply, the decomposition functions are assumed to be orthogonal.

The Sobol sensitivity index for a subset $i_s..i_t$ of inputs is then computed as $S_{i_s..i_t} = V_{i_s..i_t}/V$ [16] where $V_{i_s..i_t}$ is the variance attributed to the subset of inputs $i_s..i_t$. Main effect indexes involve only one input $i$ whereas interaction effects involve multiple inputs. A total index $S_i^T$ consists of the main effect index and all interaction effects containing input $i$ [27]. Sobol indices are bound within $[0,1]$ with 0 indicating no influence on the output and 1 indicating all variance in the output originates from this input.

To compute these variance-based global sensitivity indices generally requires a large amount of function evaluations [8]. In this work, metamodelling is used to reduce this amount of function evaluations. Variance-based sensitivity indices can also be efficiently estimated using the FAST method [28]. However, the focus of this work

4

is to analytically derive the sensitivity indices directly from the metamodel predictive equation which is more efficient and can be more accurate, especially in higher dimensions.

Variance-based sensitivity indices rely on the general assumption that variance is sufficient to describe output variability [17]. However, this assumption does not always hold true. We also investigate another interesting index type called the derivative-based global sensitivity measure (DGSM). This index has evolved from the elementary effects method by Morris [29] into several different definitions. We use a recent definition by Sobol and Kucherenko [30],

$$v_i = \mathbb{E}\left[\left(\frac{\partial f}{\partial x_i}(x_1,...,x_d)\right)^2\right],$$ (1)

where the domain of the function to be analyzed is the unit hypercube $\mathscr{H}^d$. If the inputs $x_i$ are independent and uniformly distributed, these derivative based indices can be linked to the total Sobol indices as an upper bound using the definition in [30, 31],

$$S_i^T \leq S_i^{DGSM} = \frac{v_i}{\pi^2 V},$$ (2)

where $V$ is the variance of the function.

Many different definitions for sensitivity analysis indices have been proposed to overcome some of the deficiencies mentioned. For a review on recent sensitivity analysis methods the reader is referred to [32]. We limit ourselves to the variance-based and derivative-based measures discussed above which are commonly used by engineers and design analysts.

## 3. Metamodelling

### 3.1. Tensor Product Functions

To analytically compute the variance-based and derivative-based sensitivity indices using metamodelling, the metamodel predictive function has to be of a tensor product form , i.e., $\hat{f}(x) = \sum_{i=1}^{N} \alpha_i k(x, x_i)$, where $N$ is the number of basis functions and $k$ is a kernel function. When the kernel $k$ is separable, the formula can be rewritten as

$\hat{f}(x) = \sum_{i=1}^{N} \alpha_i \prod_{l=1}^{d} h_{i,l}(x_{i,l})$ where $h_{i,l}(x_{i,l})$ is the part of the kernel for dimension $l$ out of $d$ dimensions. Two popular kernels are the Radial Basis Function (RBF) kernel $k_{RBF}(x, x_i) = \exp(-\theta||x - x_i||^2)$, also known as the Gaussian or squared exponential kernel, and the Matérn $\frac{3}{2}$ kernel $k(x, x_i) = \sigma^2 (1 + \theta\sqrt{3}||x - x_i||) \exp(-\theta\sqrt{3}||x - x_i||)$.

For the metamodel to be of a tensor product form, the kernel has to be separable. This is the case for the RBF kernel by definition, as illustrated by $k_{RBF}(x, x_i) = \prod_{l=1}^{d} \exp(-\theta||x_l - x_{i,l}||^2)$. However, the Matérn $\frac{3}{2}$ kernel is not separable. A separable version of this kernel is defined as the product of $d$ 1-dimensional kernels. Which leads to $k(x, x_i) = \prod_{l=1}^{d} (1 + \sqrt{3}\theta||x_l - x_{i,l}||) \exp(-\sqrt{3}\theta||x_l - x_{i,l}||)$.

### 3.2. *Sequential Design*

In a basic metamodelling setup, selecting the number and location of sample points for simulator evaluations, known as Design of Experiments (DoE), is performed up front. For an extensive review on the state-of-the-art design of experiments methods, the reader is referred to [33]. One-shot designs have the risk of under-fitting (too few data) or over-fitting (too much data) leading to an inaccurate or inefficient analysis.

A sequential sampling approach can be used to improve metamodelling efficiency and accuracy. A small set of initial data points is iteratively extended with additional samples. Doing so, sequential design can exploit the available information of the simulator response (exploitation) in addition to space-filling criteria (exploration) to modify the distribution of the samples to the specific problem or application. Furthermore, the process can be halted when predefined goals (such as metamodel accuracy) have been met, drastically reducing the amount of required simulator evaluations.

A powerful sequential design strategy is (F)LOLA-Voronoi [22, 23] which is a computationally efficient approach for increasing the number of samples in non-linear regions. These non-linear regions can have a large impact on the sensitivity indices. However, such non-linear regions are more difficult to model. (F)LOLA-Voronoi increases the amount of samples in these regions, leading to a more accurate model and improving the accuracy of the sensitivity indices.

The (F)LOLA-Voronoi algorithm balances two parts: (F)LOLA and Voronoi which are respectively the exploitation and exploration steps of the algorithm. The algorithm

6

selects random candidate points in the design space and scores them according to the (F)LOLA and Voronoi criteria. In the (F)LOLA part, new sample locations are ranked according to their non-linearity which is estimated using a LOcal Linear Approximation (LOLA) of each suggested sample location. The best locally linear approximation is the gradient at that sample location. However, this gradient is often not known in advance. Hence, it is estimated using a set of neighboring points. The FLOLA algorithm offers a fuzzy approach for determining the neighbors, increasing the efficiency in higher dimensions over regular LOLA [22].

In the Voronoi part, Voronoi tessellation is used to detect large regions with no samples. If the distance of candidate points to already selected samples is large, these points get a high score to be selected as the next sample. When the FLOLA and the Voronoi scores have been computed for all candidate points, the two parts are aggregated to compute the final scores. The samples are ranked and the sample with the highest score is returned as the next sample to be selected [23].

*3.3. Stopping Criterion*

Stopping criteria are used to halt the sequential sampling process when the desired task is achieved. Standard stopping criteria are based on the generalization error of the model. These methods often use k-fold cross validation (CV) in which the generated data points are split up in $k$ folds followed by the training of the model with each combination of $k-1$ folds and testing the model using the remaining data points. Error measures are then computed based on the difference between the evaluation of the test set and the original simulator data. Examples of popular error measures using cross validation are the Root Relative Squared Error ($CV_{RRSE}$) and the Bayesian Estimation Error Quotient ($CV_{BEEQ}$) [34] shown by

$$CV_{RRSE}(y, \tilde{y}) = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \tilde{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}}, \tag{3a}$$

$$CV_{BEEQ}(y, \tilde{y}) = \left(\prod_{i=1}^{n} \frac{|y_i - \tilde{y}_i|}{|y_i - \bar{y}|}\right)^{\frac{1}{n}}, \tag{3b}$$

where $y$ is the real value, $\tilde{y}$ is the predicted value and $\bar{y}$ is the mean real value.

The $\mathrm{CV_{RRSE}}$ is intuitive but only measures the improvement of the estimate over the mean model. It does not measure the quality of the fit directly and is rather pessimistic for smooth functions. The $\mathrm{CV_{BEEQ}}$ error measure is an improved version which is less sensitive to large errors. An overview and discussion of various error measures for global surrogate modelling is presented in [34].

## 4. Metamodelling based sensitivity analysis

To compute the sensitivity indices, the exact mathematical representation of the simulator has to be known. However, this expression is in general unknown or too complex. In such cases, an estimator of the sensitivity indices, often based on Monte Carlo methods, is used resulting in many evaluations of the simulator. When dealing with expensive-to-evaluate simulators, such Monte Carlo computations quickly become too costly. A common approach is to build a metamodel of the expensive simulator, to be used as a proxy for the Monte Carlo computations.

Examples of Monte Carlo estimation methods are the approaches of Sobol [35] and Saltelli [36] for variance based Sobol indices. However, to achieve a descent accuracy, Monte Carlo methods require numerous samples and even though the metamodel is cheaper to evaluate, this quickly becomes an expensive procedure. When the dimensionality of the problem increases, the amount of required samples for Monte Carlo increases exponentially.

Using quasi-Monte Carlo sequences [4] or Fourier Amplitude Sensitivity Testing (FAST) [28] can help reduce the amount of metamodel evaluations. However, Monte Carlo methods introduce an additional error on top of the metamodel error. For some metamodels, a more efficient and accurate approach is available based on the analytic evaluation of the sensitivity indices using the mathematical expression of the metamodel. In this section, we introduce a complete algorithm for data-efficient sensitivity analysis with metamodelling using this analytic evaluation.

### 4.1. Variance-based Approach

For metamodels in a tensor-product functional form, the Sobol indices can be derived analytically [18]. For such metamodels, the variance of a subset of input dimen-

8

sions, with $p_l(x_l)$ the distribution of the input $l$, $[a_l, b_l]$ the bounds of the distribution $p_l(x_l)$, $d$ the number of input dimensions and $N$ the number of samples, can be computed using

$$V_{i_s \ldots i_t} = \sum_{i_1}^{N} \sum_{i_2}^{N} \left( \alpha_{i_1} \alpha_{i_2} \prod_{l=1}^{d} (C1_{i_1,l} \, C1_{i_2,l}) \left( \prod_{l \in i_s \ldots i_t} \frac{C2_{i_1,i_2,l}}{C1_{i_1,l} \, C1_{i_2,l}} - 1 \right) \right),$$

$$C1_{i,l} = \int_{a_l}^{b_l} h_{i,l}(x_l) p_l(x_l) dx_l,$$

$$C2_{i_1,i_2,l} = \int_{a_l}^{b_l} h_{i_1,l}(x_l) h_{i_2,l}(x_l) p_l(x_l) dx_l.$$

For a derivation of these formulas we refer to the original works of Jin [18]. These integrals can easily be analytically evaluated using symbolic software without the need of any estimators. By using these analytic formulas, we avoid the use of Monte Carlo methods on the metamodel which would introduce additional errors and would still require many samples for high-dimensional problems.

*4.2. Derivative-based Approach*

Similarly, for the derivative-based sensitivity indices, an analytic form of Equation (1) can also be extracted for specific metamodels. This has been applied to polynomial chaos expansion metamodels in [37]. Here, we prove the formulas for general metamodels in a tensor-product form. This results in the following equations

$$\nu_i = \sum_{i_1}^{N} \sum_{i_2}^{N} \alpha_{i_1} \alpha_{i_2} \left( \prod_{\substack{l=1 \\ l \neq i}}^{d} C2_{i_1,i_2,l} \right) C3_{i_1,i_2,i},$$

$$C3_{i_1,i_2,i} = \int_{a_l}^{b_l} \frac{\partial h_{i,i_1}(x_i)}{\partial x_i} \frac{\partial h_{i,i_2}(x_i)}{\partial x_i} p_i(x_i) dx_i.$$

See Appendix A for the mathematical proof.

### 4.3. Stopping Criterion

The metamodel accuracy is assessed using the $CV_{RRSE}$ and the $CV_{BEEQ}$ criterion. However, while the metamodel itself might still not be completely accurate, the general trend of variability of the outputs might already be visible. Furthermore, the error-based stopping criteria are difficult to interpret when evaluating the accuracy of the sensitivity indices. Hence, we introduce two new stopping criteria based on the variability of the sensitivity indices to make the metamodelling process more efficient and interpretable:

$$SCV\text{-}\mu_{var} = \sum_{i=1}^{d} \frac{Var(I_{i,1},..,I_{i,k})}{d},$$

$$SCV\text{-}max_{var} = \max_{i=1..d} \frac{Var(I_{i,1},..,I_{i,k})}{d},$$

where $k$ is the number of folds and $I_{i,j}$, a specific sensitivity index (either a variance-based Sobol index $S_{i,j}$ or a derivative-based index $v_{i,j}$) for variable $i$ in fold $j$. Note that the $SCV\text{-}max_{var}$ definition is a more conservative approach than the $SCV\text{-}\mu_{var}$ definition. The division by $d$ in both equations makes the result more interpretable.

For derivative-based indices, the indices are first normalized across dimensions per evaluation, before computing the criterion to overcome the interpretation problem of derivative-based sensitivity indices.

As the sensitivity indices have values between 0 and 1, and the proposed criteria are based on the variance of these sensitivity indices, an interpretable stopping criterion is achieved even for black-box simulators as this criterion, in contrast to the error-based criteria, is not dependent on the output values of the simulator. The analyst can easily set a desired goal for the variance of the sensitivity indices. This goal will depend on the application requirements.

### 4.4. Complete Algorithm

The complete work flow of the algorithm is shown in Figure 1. First, an initial set of points is chosen using for example a Latin hypercube design after which the metamodelling loop starts. The size of this initial set is problem-depended and determined by the analyst. All selected points are evaluated using the simulator and a metamodel is built.

Next, the sensitivity indices are evaluated analytically based on the metamodel predictive equation. At the end of each iteration, the stopping criteria are computed. This involves a model accuracy based measure ($CV_{RRSE}$ or $CV_{BEEQ}$) and one of the criteria based on k-fold cross validation of the sensitivity indices (SCV-$\mu_{var}$ or SCV-max$_{var}$). If the stopping criterion threshold is reached, the algorithm terminates. Otherwise, a new set of points is selected using FLOLA-Voronoi after which another iteration begins.
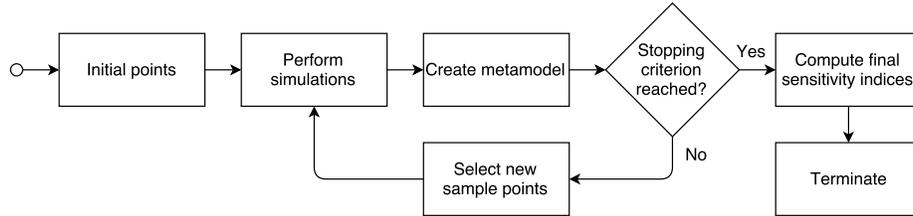


Figure 1: Sequential sensitivity analysis of expensive black-box simulators with metamodelling.

## 5. Experimental Setup

The performance of the proposed sequential global sensitivity analysis algorithm is evaluated on three typical sensitivity analysis benchmark functions and two engineering applications. Table 1 provides an overview of the mathematical formulation and dimensionality of the functions.

Table 1: Sensitivity analysis benchmark functions.

| Name | Equation | Inputs |
|---|---|---|
| **Ishigami** [38] | $\sin(x_1) + 7\sin(x_2)^2 + 0.1x_3^4\sin(x_1)$ | 3 |
| **G-function** [39] | $\displaystyle\prod_{i=1}^{3} \frac{|4x_i - 2| + (i-2)/2}{1 + (i-2)/2}$ | 3 |
| **Loeppky** [40] | $6x_1 + 4x_2 + 5.5x_3 + 3x_1x_2 + 2.2x_1x_3$ $+ 1.4x_2x_3 + x_4 + 0.5x_5 + 0.2x_6 + 0.1x_7$ | 10 |

The first engineering application is the simulation of a satellite braking system [41, 42]. The system has a five-dimensional input which contains, among others, the mass

11

of the braking system.

The second engineering application is a Fluid-Structure Interaction (FSI) simulation of an aneurysm in an artery [43]. In this application, the flow through an artery is modelled by a simulator with a 20D input. Each input dimension represents the spring characteristic of a slice of the artery except for the last input which influences all other spring characteristics.

*5.1. Configuration*

The experiments are performed using the SUMO Toolbox [9] which provides a complete framework for developing and testing metamodelling techniques. The Kriging model [44] and the Gaussian process model [45] are both optimized using maximum likelihood estimation. The Kriging and GP models use a separable Matérn $\frac{3}{2}$ kernel and the LS-SVM model uses a Gaussian RBF kernel, the hyperparameters of this model are optimized with $CV_{RRSE}$. For the low-dimensional problems ($< 20D$), a separate hyperparameter is trained for each dimension of the Matérn kernel. For high-dimensional problems ($\geq 20D$), a single hyperparameter is trained to avoid a complex and lengthy high-dimensional optimization of the hyperparameters. For the Gaussian RBF kernel, a single hyperparameter is trained regardless of dimensionality.

The initial design used to build the metamodels is a Latin hypercube constructed using the Translational Propagation algorithm [46]. During each step of the process, 10 new sampling points, determined by FLOLA-Voronoi are evaluated up to a total of 300 samples. For the artery FSI application, a total of 500 samples are evaluated.

The complete process is repeated 10 times for statistical robustness. The accuracy of the metamodel itself is evaluated using 10-fold cross validation with the $CV_{RRSE}$ and $CV_{BEEQ}$ error functions. Both the variance-based and the derivative-based sensitivity indices are computed. The novel SCV-$\text{max}_{var}$ and SCV-$\mu_{var}$ stopping criteria are computed as well. For variance-based indices, the Monte Carlo estimates are computed as well for comparison.

## 6. Results and Discussion

### 6.1. Benchmark functions

In the following discussion, the Ishigami function will be used to demonstrate the various aspects of the algorithm and Kriging is used as metamodel type, unless otherwise specified. The figures illustrate the mean (full line) and standard deviation (shaded area) across 10 complete runs.

#### 6.1.1. Sensitivity Indices

The most important aspect of the algorithm is the accurate determination of the sensitivity indices of the simulator. For variance-based analysis, both the main effects and total effects (= main effects + interaction effects) have to be accurate. Figure 2 shows the evolution of the various sensitivity indices, as a function of the number of data samples. The cross at the end of the figure indicates the exact target value of the sensitivity indices, analytically computed on the equation of the Ishigami function. The figures clearly demonstrate the MC version struggling in the beginning to get a stable and accurate value. It also shows the large differences across iterations in the MC version while the evolution of the sensitivity indices for the analytic version is much more similar during each iteration of the experiment. In a black-box setting, where the amount of required samples and the sampling budget is unknown, it is desirable to get a stable estimate as fast as possible. By using these analytic computations, this desired stable and accurate evolution can be achieved.

Note the difference in ranking of the sensitivity indices when comparing variance-based with derivative-based analysis. As the sensitivity indices have a different focus, they come to different results. This demonstrates the value in computing multiple types of indices.

The accuracy of the computed indices can be evaluated by comparing the experimental results with the exact theoretical values. For the Ishigami function, this is shown in Table 2. These experimental results were gathered after 300 samples with a final variance-based SCV-max$_{var}$ score of $2.4496 \times 10^{-6}$, total variance-based SCV-max$_{var}$ score of $2.6112 \times 10^{-6}$ and derivative-based SCV-max$_{var}$ score of $3.6477 \times 10^{-5}$.

(a) Variance-based sensitivity.

(b) Variance-based sensitivity MC.

(c) Variance-based total sensitivity.

(d) Variance-based total sensitivity MC.

(e) Derivative-based sensitivity.
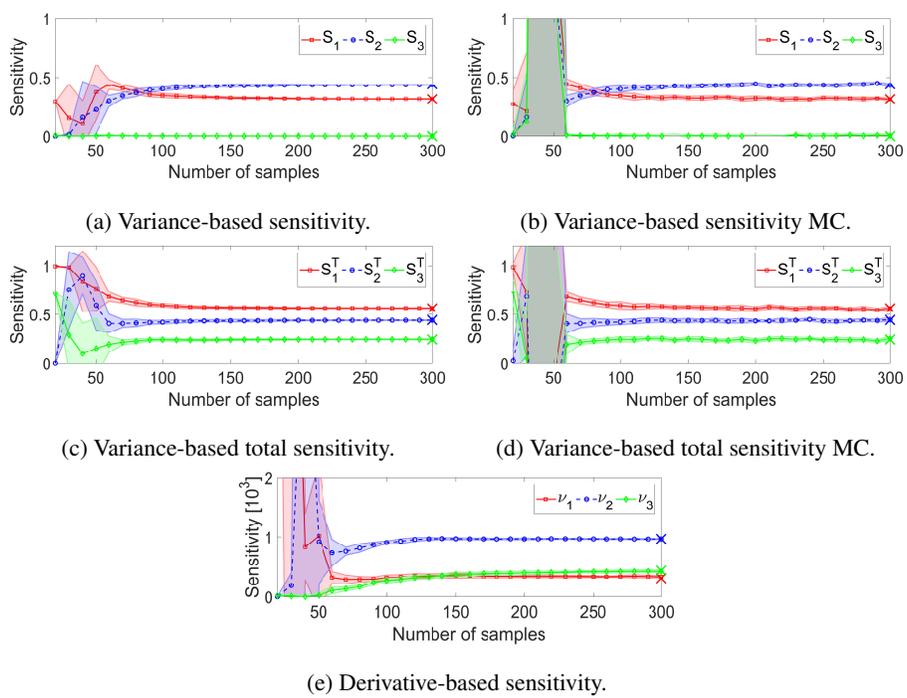
Figure 2: Sensitivity indices for the Ishigami function with Kriging and FLOLA-Voronoi.

14

As detailed in Section 3, the FLOLA-Voronoi sampling criterion guides the algorithm to acquire more samples in the difficult-to-model regions which are the most interesting regions for sequential sensitivity analysis. This helps to reduce the overall required samples and increase the accuracy of the analysis method when compared to recent literature [47].

Table 2: Comparison of metamodel based sensitivity indices with exact analytic values for the Ishigami function with Kriging and FLOLA-Voronoi.

| Index | 1 | 2 | 3 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|
| Type | **Proposed algorithm** | | | **Exact** | | |
| **Variance** | 0.3152 | 0.4393 | 0.0002 | 0.3139 | 0.4424 | 0 |
| **Total-variance** | 0.5603 | 0.4397 | 0.2453 | 0.5576 | 0.4424 | 0.2437 |
| **Derivative** | 329.3 | 965.6 | 420.5 | 304.8 | 967.2 | 433.8 |

### 6.1.2. Stopping Criterion

An important aspect of the algorithm is the stopping criterion to determine when sufficient samples have been gathered. Apart from the two typical error based cross validation criteria, the two newly introduced stopping criteria, SCV-$\max_{\text{var}}$ and SCV-$\mu_{\text{var}}$, are also computed for both the variance and the derivative-based sensitivity indices as shown in Figure 3. All stopping criteria show the same pattern, indicating that when the metamodel for the Ishigami function becomes sufficiently accurate, the sensitivity indices also become stable. Note how the value of the variance across the sensitivity indices, represented by the SCV-$\max_{\text{var}}$ and SCV-$\mu_{\text{var}}$ criteria can easily be interpreted by an analyst while this interpretation is more difficult for the error based stopping criteria when dealing with a black-box function. This is because the sensitivity indices always have values between 0 and 1 while the error based measures are dependent on the output variability of the simulator.

### 6.1.3. Different Models

The algorithm can be used with various different metamodel types. This is useful as the appropriate metamodel type is dependent on the problem being modelled. The

(a) Variance-based SCV.

(b) Total variance-based SCV.

(c) Derivative-based SCV.

(d) Error based CV.
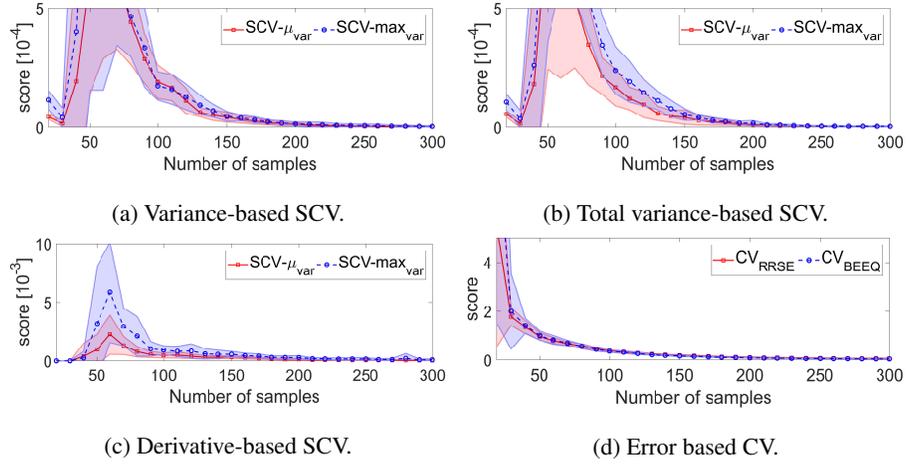
Figure 3: Cross validation measures for the Ishigami function with Kriging and FLOLA-Voronoi.

evolution of the sensitivity indices for the GP and LS-SVM metamodels is shown in Figure 4 and Figure 5 respectively. The results after 300 samples are shown in Table 3 and Table 4 for the GP and LS-SVM metamodels respectively.



(a) Variance-based sensitivity.

(b) Variance-based total sensitivity.

(c) Derivative-based sensitivity.
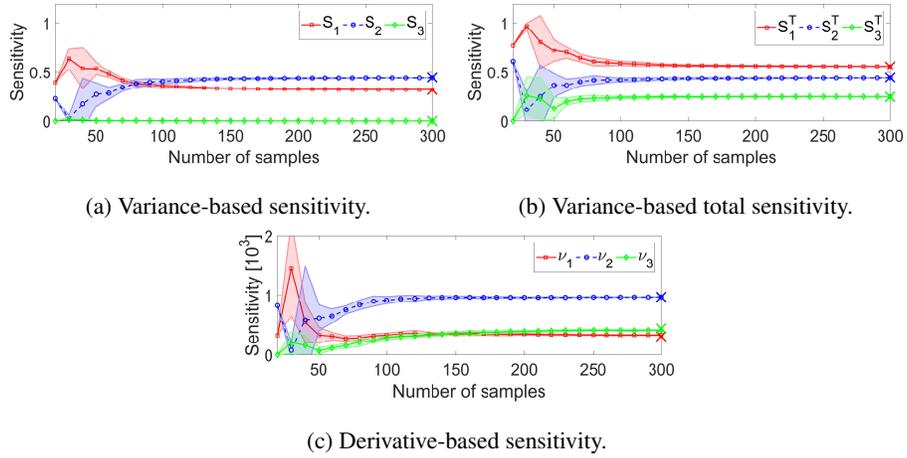
Figure 4: Sensitivity indices for the Ishigami function with GP and FLOLA-Voronoi.

### 6.1.4. G-function and Loeppky Benchmark Functions

The results for the other benchmark functions are similar to those for the Ishigami function. An overview of the results is given in Figure 6 and Figure 7 for the G-

(a) Variance-based sensitivity.

(b) Variance-based total sensitivity.



(c) Derivative-based sensitivity.
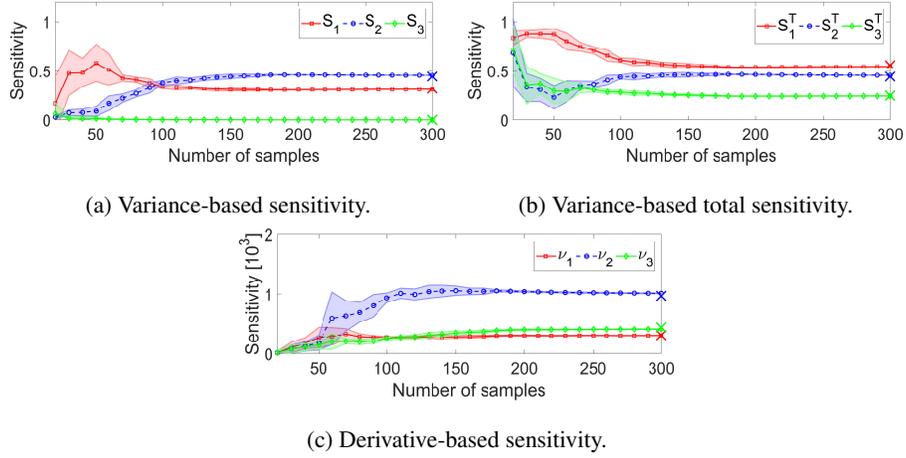
Figure 5: Sensitivity indices for the Ishigami function with LS-SVM and FLOLA-Voronoi.

Table 3: Metamodel based sensitivity indices for the Ishigami function with GP and FLOLA-Voronoi.

| Index | 1 | 2 | 3 | SCV-max$_{var}$ |
|---|---|---|---|---|
| **Variance** | 0.3157 | 0.4393 | 0.0001 | $2.2758 \times 10^{-6}$ |
| **Total-variance** | 0.5605 | 0.4397 | 0.2449 | $2.4638 \times 10^{-6}$ |
| **Derivative** | 326.3 | 965.9 | 405.9 | $3.0642 \times 10^{-5}$ |

Table 4: Metamodel based sensitivity indices for the Ishigami function with LS-SVM and FLOLA-Voronoi.

| Index | 1 | 2 | 3 | SCV-max$_{var}$ |
|---|---|---|---|---|
| **Variance** | 0.3066 | 0.4546 | 0.0000 | $3.0547 \times 10^{-6}$ |
| **Total-variance** | 0.5453 | 0.4552 | 0.2387 | $2.5661 \times 10^{-6}$ |
| **Derivative** | 301.1 | 1010.7 | 406.6 | $1.7899 \times 10^{-5}$ |

17

function and Loeppky function respectively. For the Loeppky function, only the three most significant inputs are provided in the figure legend to keep the figures clear.



(a) Variance-based sensitivity.

(b) Error based CV.

(c) Variance-based total sensitivity.

(d) Variance-based SCV.

(e) Derivative-based sensitivity.

(f) Derivative-based SCV.
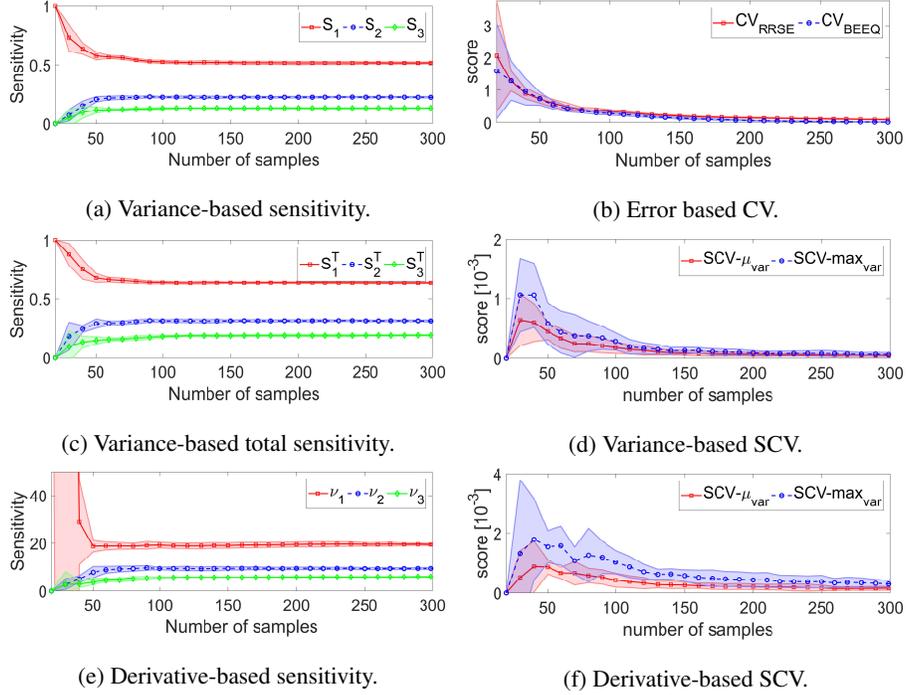
Figure 6: Sensitivity analysis of G-function with Kriging and FLOLA-Voronoi.

## 6.2. Satellite Braking System Application

For the satellite braking system application, the evolution of the corresponding sensitivity indices is shown in Figure 8 where data is shown for an LS-SVM model. The figures show large similarities between the variance-based and total variance-based sensitivity indices. From this, we learn that the simulator has very few interaction effects.

This example clearly shows the advantages of using the SCV-$\mathrm{max_{var}}$ and SCV-$\mu_{\mathrm{var}}$ stopping criteria. While the error-based metrics are still at a relatively high value, and furthermore dependent on the simulator output, the sensitivity-based stopping criteria converge to a small value which represents the variance of the sensitivity indices. The

(a) Variance-based sensitivity.

(b) Error based CV.

(c) Variance-based total sensitivity.

(d) Variance-based SCV.

(e) Derivative-based sensitivity.

(f) Derivative-based SCV.
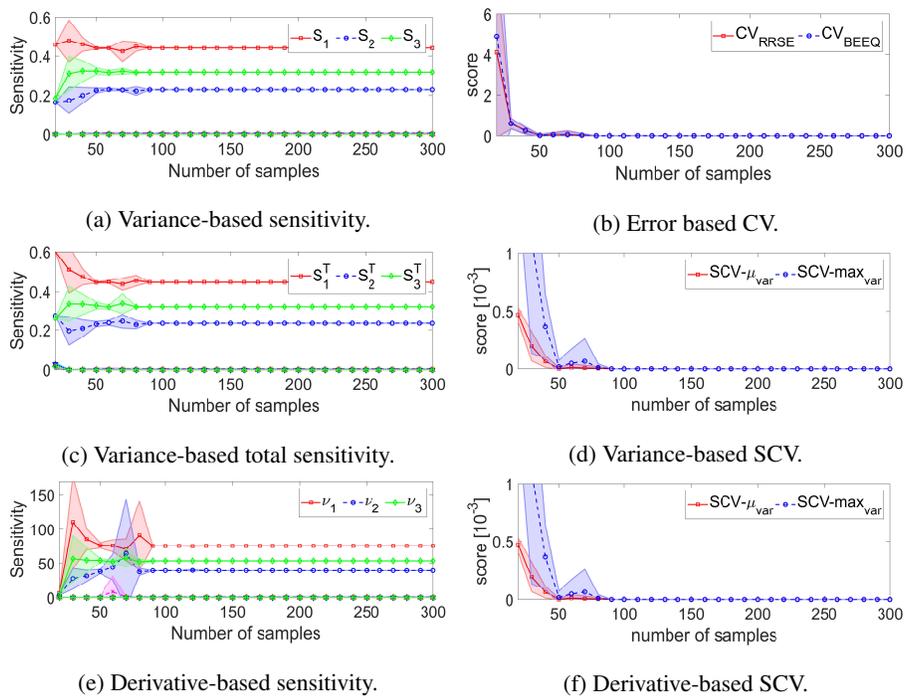
Figure 7: Sensitivity analysis of Loeppky function with Kriging and FLOLA-Voronoi.

analyst can easily interpret this and set a desirable goal for this variance depending on the application requirements.



(a) Variance-based sensitivity

(b) Error based CV

(c) Variance-based total sensitivity

(d) Variance-based SCV

(e) Derivative-based sensitivity

(f) Derivative-based SCV
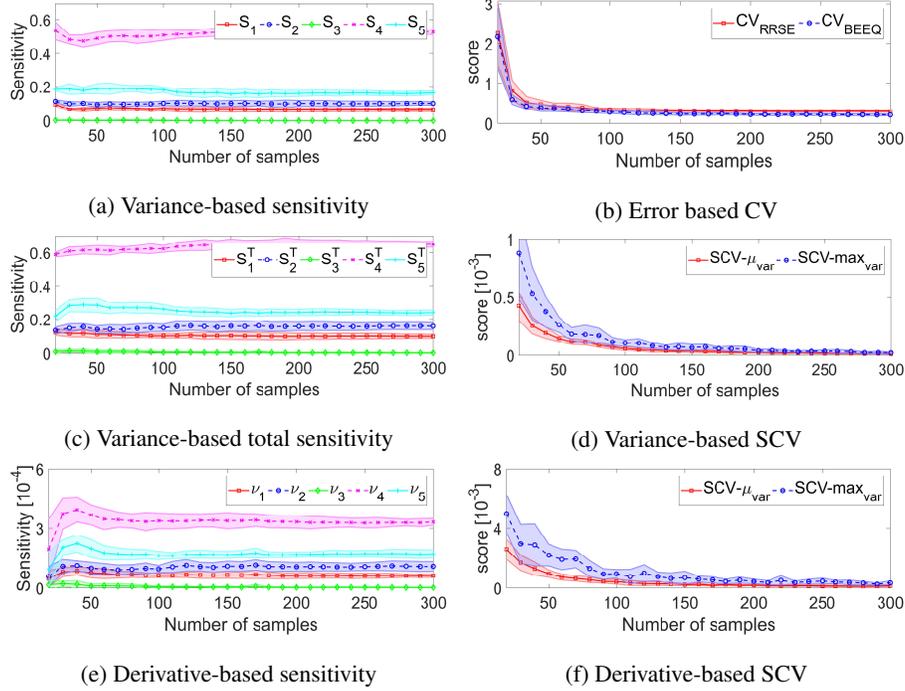
Figure 8: Sensitivity analysis of satellite braking system application with LS-SVM.

The results of the sensitivity analysis teach us that the third input, representing the mass of the satellite braking system, has a negligible impact on the output of the simulator. This is valuable information for the analysis of the system as well as for efficient modelling of the simulator response.

### 6.3. Artery FSI Application

The sensitivity analysis results for the Artery FSI application are shown in Figure 9 where we show the data for a GP model with FLOLA-Voronoi sequential sampling. In these figures, only the last two input dimensions are mentioned in the figure legend as all 18 other inputs have a relative negligible impact. The calculated sensitivity indices confirm the importance of the last input dimension which influences the spring characteristic of all other inputs.

The final engineering example also demonstrates the advantages of the novel stopping criteria. The $CV_{RRSE}$ and $CV_{BEEQ}$ measures, which are based on the error of the model and hence also on the simulator output, show little convergence and remain at a relatively high value. The novel sensitivity based criteria on the other hand, quickly converge to a small and interpretable value as the sensitivity information is already available without the need for small $CV_{RRSE}$ and $CV_{BEEQ}$ scores.

The results of the sensitivity analysis teach us some interesting insights. It shows that the final output indeed has the most influence on the output of the system. If a screening of the output variables was desired, the results have already been achieved. If the goal was to create an accurate model of the simulator, the results provide useful insights into where the most interesting data samples can be gathered.



(a) Variance-based sensitivity.

(b) Error based CV.

(c) Variance-based total sensitivity.

(d) Variance-based SCV.

(e) Derivative-based sensitivity.

(f) Derivative-based SCV.
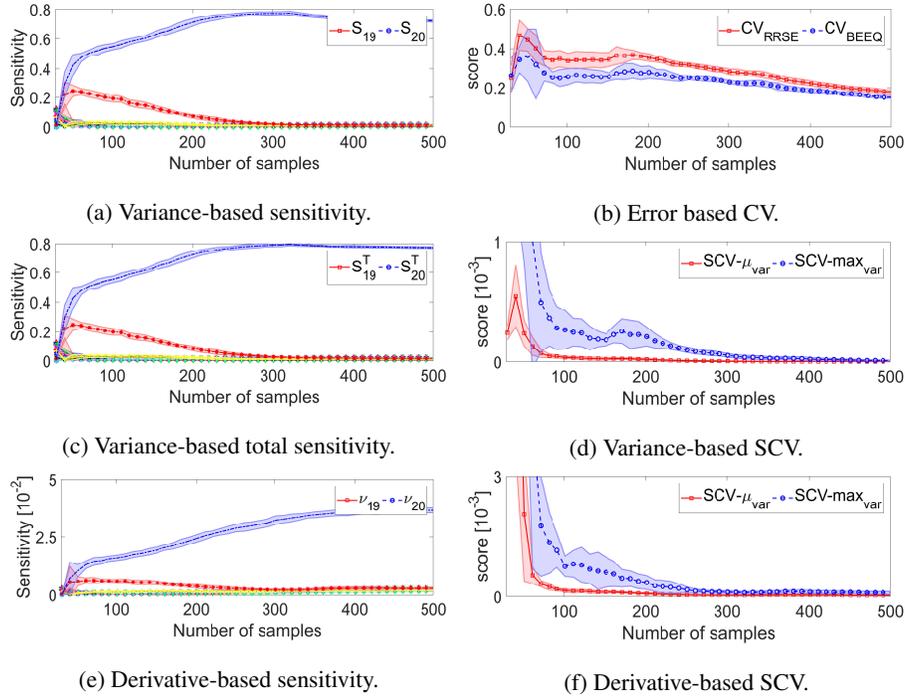
Figure 9: Sensitivity analysis of Artery FSI Application with GP and FLOLA-Voronoi.

## 7. Conclusion

We introduced a complete algorithm for sequential variance-based and derivative-based sensitivity analysis of expensive black-box simulators using a metamodelling approach. The analytic evaluation of the sensitivity indices on the metamodel predictive function, together with the sequential sampling approach and the stopping criterion provides an efficient and accurate estimation process. The tests demonstrate the accurate and efficient computation of the sensitivity indices and the strength of the SCV-max$_{\text{var}}$ and SCV-$\mu_{\text{var}}$ criteria next to the traditional metamodel accuracy measures. These novel criteria represent an intuitive measure of the variance of the sensitivity indices. If the end goal is to have a screening of the input variables, the new stopping criteria allow to stop the metamodelling process without having to create an extensive globally accurate metamodel. If the end goal is to create an accurate metamodel, the new stopping criteria indicate when the operator can remove inputs from the design space to reduce the dimensionality and increase the efficiency of the modelling. The results of this work allow researchers to get better and faster insights into expensive black-box problems and can help to reduce the dimensionality of these complex design problems.

## References

[1] T. Simpson, A. Booker, D. Ghosh, A. Giunta, P. Koch, R.-J. Yang, Approximation methods in multidisciplinary analysis and optimization: a panel discussion, Structural and multidisciplinary optimization 27 (5) (2004) 302–313.

[2] L. Gu, A comparison of polynomial based regression models in vehicle safety analysis, in: ASME Design Engineering Technical Conferences, 2001.

[3] G. Hughes, On the mean accuracy of statistical pattern recognizers, Information Theory, IEEE Transactions on 14 (1) (1968) 55–63.

[4] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, Global sensitivity analysis: the primer, Wiley, 2008.

[5] A. Marrel, B. Iooss, F. Van Dorpe, E. Volkova, An efficient methodology for modeling complex computer codes with gaussian processes, Computational Statistics & Data Analysis 52 (10) (2008) 4731–4744.

[6] X. Song, G. Sun, Q. Li, Sensitivity analysis and reliability based design optimization for high-strength steel tailor welded thin-walled structures under crashworthiness, Thin-Walled Structures 109 (2016) 132–142.

[7] I. M. Sobol, Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates, Mathematics and computers in simulation 55 (1) (2001) 271–280.

[8] S. Kucherenko, M. Rodriguez-Fernandez, C. Pantelides, N. Shah, Monte carlo evaluation of derivative-based global sensitivity measures, Reliability Engineering and System Safety 94 (2009) 1135–1148.

[9] D. Gorissen, K. Crombecq, I. Couckuyt, P. Demeester, T. Dhaene, A surrogate modeling and adaptive sampling toolbox for computer based design, Journal of Machine Learning Research 11 (2010) 2051–2055.

[10] D. Stephens, D. Gorissen, K. Crombecq, T. Dhaene, Surrogate based sensitivity analysis of process equipment, Applied Mathematical Modelling 35 (4) (2011) 1676–1687.

[11] T. Santner, B. Williams, W. Notz, The design and analysis of computer experiments, Springer series in statistics, Springer-Verlag, New York, 2003.

[12] A. Forrester, A. Sobester, A. Keane, Engineering Design Via Surrogate Modelling: A Practical Guide, Wiley, Chichester, 2008.

[13] C. E. Rasmussen, C. K. I. Williams, Gaussian Processes for Machine Learning, MIT Press, 2006.

[14] J. Suykens, T. V. Gestel, J. D. Brabanter, B. D. Moor, J. Vandewalle, Least Squares Support Vector Machines, World Scientific Publishing Co., Pte, Ltd., Singapore, 2002.

[15] J. Hu, S. Zhang, Global sensitivity analysis based on high-dimensional sparse surrogate construction, Applied Mathematics and Mechanics 38 (6) (2017) 797–814.

[16] I. Sobol, Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates, Mathematics and Computers in Simulation 55 (1-3) (2001) 271–280.

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

[18] R. Jin, Enhancements of metamodeling techniques in engineering design, Ph.D. thesis, University of Illinois at Chicago (2004).

[19] D. S. Broomhead, D. Lowe, Radial basis functions, multi-variable functional interpolation and adaptive networks, Tech. rep., DTIC Document (1988).

[20] M. De Lozzo, A. Marrel, Estimation of the derivative-based global sensitivity measures using a gaussian process metamodel, SIAM/ASA Journal on Uncertainty Quantification 4 (1) (2016) 708–738.

[21] L. Le Gratiet, C. Cannamela, B. Iooss, A bayesian approach for global sensitivity analysis of (multifidelity) computer codes, SIAM/ASA Journal on Uncertainty Quantification 2 (1) (2014) 336–363.

[22] K. Crombecq, D. Gorissen, D. Deschrijver, T. Dhaene, A novel hybrid sequential design strategy for global surrogate modelling of computer experiments, SIAM Journal of Scientific Computing 33 (4) (2010) 1948–1974.

[23] J. van der Herten, I. Couckuyt, D. Deschrijver, T. Dhaene, A fuzzy hybrid sequential design strategy for global surrogate modeling of high-dimensional computer experiments, SIAM Journal of Scientific Computing 37 (2) (2015) 1020–1039.

[24] J. Wu, Z. Luo, J. Zheng, C. Jiang, Incremental modeling of a new high-order polynomial surrogate model, Applied Mathematical Modelling 40 (7) (2016) 4681–4699.

[25] X. R. Li, Z. Zhao, Evaluation of estimation algorithms part I: incomprehensive measures of performance, IEEE Transactions on Aerospace and Electronic Systems 42 (4) (2006) 1340–1358.

[26] W. Hoeffding, A class of statistics with asymptotically normal distribution, The annals of mathematical statistics (1948) 293–325.

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

[28] A. Saltelli, S. Tarantola, K.-S. Chan, A quantitative model-independent method for global sensitivity analysis of model output, Technometrics 41 (1) (1999) 39–56.

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

[30] I. Sobol, S. Kucherenko, Derivative based global sensitivity measures and their link with global sensitivity indices, Mathematics and Computers in Simulation 79 (10) (2009) 3009–3017.

[31] M. Lamboni, B. Iooss, A. Popelin, F. Gamboa, Derivative-based global sensitivity measures: general links with sobol' indices and numerical tests, Mathematics and Computers in Simulation 87 (2013) 44–54.

[32] E. Borgonovo, E. Plischke, Sensitivity analysis: a review of recent advances, European Journal of Operational Research 248 (3) (2016) 869–887.

[33] J. P. Kleijnen, Design and Analysis of Simulation Experiments, Vol. 230, Springer, 2015.

[34] D. Gorissen, I. Couckuyt, E. Laermans, T. Dhaene, Multiobjective global surrogate modeling, dealing with the 5-percent problem, Engineering with Computers 26 (1) (2010) 81–98.

[35] I. Sobol, Sensitivity analysis for nonlinear mathematical models, Mathematical Modeling and Computational Experiment 1 (1993) 407–414.

[36] A. Saltelli, Making best use of model valuations to compute sensitivity indices., Computer Physics Communications 145 (2002) 280297.

[37] B. Sudret, C. Mai, Computing derivative-based global sensitivity measures using polynomial chaos expansions, Reliability engineering and system safety 134 (2015) 241–250.

[38] T. Ishigami, T. Homma, An importance quantification technique in uncertainty analysis for computer models, in: Proceedings of First International Symposium on Uncertainty Modeling and Analysis, IEEE, 1990, pp. 398–403.

[39] A. Saltelli, I. Sobol, Sensitivity analysis for non-linear mathematical models: numerical experience, Matematicheskoe Modelirovanie 7 (11) (1994) 16–28.

[40] J. Loeppky, B. Williams, L. Moore, Global sensitivity analysis for mixture experiments, Technometrics 55 (1) (2013) 68–78.

[41] A. Kestil, T. Tikka, P. Peitso, J. Rantanen, A. Nsil, K. Nordling, H. Saari, R. Vainio, P. Janhunen, J. P. et al., Aalto-1 nanosatellite - technical description and mission objectives, Geoscientific Instrumentation, Methods and Data Systems 2 (1) (2013) 121–130.

[42] O. Khurshid, T. Tikka, J. Praks, M. Hallikainen, Accommodating the plasma brake experiment on-board the aalto-1 satellite, in: Proceedings, Vol. 63 of 2S, Estonian Academy of Sciences, 2014, pp. 258–266.

[43] J. Degroote, M. Hojjat, E. Stavropoulou, R. Wüchner, K.-U. Bletzinger, Partitioned solution of an unsteady adjoint for strongly coupled fluid-structure interactions and application to parameter identification of a one-dimensional problem, Structural and Multidisciplinary Optimization 47 (1) (2013) 77–94.

[44] I. Couckuyt, T. Dhaene, P. Demeester, ooDACE toolbox: A flexible object-oriented kriging implementation, Journal of Machine Learning Research 15 (2014) 3183–3186.

[45] C. E. Rasmussen, H. Nickisch, Gaussian processes for machine learning (gpml) toolbox, Journal of Machine Learning Research 11 (2010) 3011–3015.

[46] F. A. Viana, G. Venter, V. Balabanov, An algorithm for fast optimal latin hypercube design of experiments, International journal for numerical methods in engineering 82 (2) (2010) 135–156.

[47] K. Cheng, Z. Lu, Y. Zhou, Y. Shi, Y. Wei, Global sensitivity analysis using support vector regression, Applied Mathematical Modelling.

### Appendix A. Derivative-based Sensitivity Analysis of Tensor Product Functions

In this appendix the general equation for derivative-based global sensitivity indices for tensor product functions is derived. The domain of the function is assumed to be the unit hypercube $\mathcal{H}^d$.

$$
\begin{aligned}
v_i &= \mathbb{E}\left[\left(\frac{\partial f}{\partial x_i}(X)\right)^2\right] \\
&= \mathbb{E}\left[\left(\frac{\partial}{\partial x_i}\left(\alpha_0 + \sum_j^N \alpha_j B_j(X)\right)\right)^2\right] \\
&= \mathbb{E}\left[\left(\left(\sum_j^N \alpha_j \frac{\partial}{\partial x_i} B_j(X)\right)\right)^2\right] \\
&= \mathbb{E}\left[\sum_{j=1}^N \sum_{k=1}^N \alpha_j \alpha_k \frac{\partial}{\partial x_i} B_j(X) \frac{\partial}{\partial x_i} B_k(X)\right] \\
&= \sum_{j=1}^N \sum_{k=1}^N \alpha_j \alpha_k \int_{\mathcal{H}^d} p(x) \frac{\partial}{\partial x_i} B_j(X) \frac{\partial}{\partial x_i} B_k(X) dX \\
&= \sum_{j=1}^N \sum_{k=1}^N \alpha_j \alpha_k \int_{\mathcal{H}^d} \frac{\partial}{\partial x_i} B_j(X) \frac{\partial}{\partial x_i} B_k(X) dX \\
&= \sum_{j=1}^N \sum_{k=1}^N \alpha_j \alpha_k \int_{\mathcal{H}^d} \frac{\partial}{\partial x_i}\left(\prod_{l=1}^M h_{l,j}(X)\right) \frac{\partial}{\partial x_i}\left(\prod_{l=1}^M h_{l,k}(X)\right) dX \\
&= \sum_{j=1}^N \sum_{k=1}^N \alpha_j \alpha_k \int_{\mathcal{H}^d} \left(\prod_{\substack{l=1 \\ l \neq i}}^M h_{l,j}(X)\right)\left(\prod_{\substack{l=1 \\ l \neq i}}^M h_{l,k}(X)\right) \frac{\partial}{\partial x_i} h_{i,j}(X) \frac{\partial}{\partial x_i} h_{i,k}(X) dX \\
&= \sum_{j=1}^N \sum_{k=1}^N \alpha_j \alpha_k \left(\prod_{\substack{l=1 \\ l \neq i}}^M \int_0^1 h_{l,j}(X) h_{l,k}(X) dx_l\right)\left(\int_0^1 \frac{\partial}{\partial x_i} h_{i,j}(X) \frac{\partial}{\partial x_i} h_{i,k}(X) dx_i\right) \\
&= \sum_{j=1}^N \sum_{k=1}^N \alpha_j \alpha_k \left(\prod_{\substack{l=1 \\ l \neq i}}^M C2_{j,k,l}\right) C3_{j,k,i}
\end{aligned}
$$

In this derivation, $C2$ is the same factor as in variance-based global sensitivity analysis. For domains other than the unit hypercube $\mathcal{H}^d$, the equations are easily updated.