Journal of Computational Physics 361 (2018) 167–198

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

An adaptive least-squares global sensitivity method and application to a plasma-coupled combustion prediction with parametric correlation

Kunkun Tang [a,*], Luca Massa [a,d], Jonathan Wang [a,b], Jonathan B. Freund [a,b,c]

[a] *The Center for Exascale Simulation of Plasma-coupled Combustion (XPACC), Coordinated Science Laboratory (CSL), University of Illinois at Urbana–Champaign, Urbana, IL 61801, USA*
[b] *Mechanical Science & Engineering, University of Illinois at Urbana–Champaign, Urbana, IL 61801, USA*
[c] *Aerospace Engineering, University of Illinois at Urbana–Champaign, Urbana, IL 61801, USA*
[d] *Aerospace and Ocean Engineering, Virginia Tech, Blacksburg, VA 24061, USA*

A R T I C L E   I N F O

Article history:
Received 12 November 2017
Received in revised form 22 January 2018
Accepted 23 January 2018
Available online 21 February 2018

Keywords:
Covariance-based sensitivity analysis
Correlated inputs
Adaptive ANOVA
Dimension reduction
Polynomial dimensional decomposition (PDD)
Plasma-coupled turbulent combustion

A B S T R A C T

We introduce an efficient non-intrusive surrogate-based methodology for global sensitivity analysis and uncertainty quantification. Modified covariance-based sensitivity indices (mCov-SI) are defined for outputs that reflect correlated effects. The overall approach is applied to simulations of a complex plasma-coupled combustion system with disparate uncertain parameters in sub-models for chemical kinetics and a laser-induced breakdown ignition seed. The surrogate is based on an Analysis of Variance (ANOVA) expansion, such as widely used in statistics, with orthogonal polynomials representing the ANOVA subspaces and a polynomial dimensional decomposition (PDD) representing its multi-dimensional components. The coefficients of the PDD expansion are obtained using a least-squares regression, which both avoids the direct computation of high-dimensional integrals and affords an attractive flexibility in choosing sampling points. This facilitates importance sampling using a Bayesian calibrated posterior distribution, which is fast and thus particularly advantageous in common practical cases, such as our large-scale demonstration, for which the asymptotic convergence properties of polynomial expansions cannot be realized due to computation expense. Effort, instead, is focused on efficient finite-resolution sampling. Standard covariance-based sensitivity indices (Cov-SI) are employed to account for correlation of the uncertain parameters. Magnitude of Cov-SI is unfortunately unbounded, which can produce extremely large indices that limit their utility. Alternatively, mCov-SI are then proposed in order to bound this magnitude $\in [0, 1]$. The polynomial expansion is coupled with an adaptive ANOVA strategy to provide an accurate surrogate as the union of several low-dimensional spaces, avoiding the typical computational cost of a high-dimensional expansion. It is also adaptively simplified according to the relative contribution of the different polynomials to the total variance. The approach is demonstrated for a laser-induced turbulent combustion simulation model, which includes parameters with correlated effects.

* Corresponding author.
E-mail address: ktg@illinois.edu (K. Tang).

https://doi.org/10.1016/j.jcp.2018.01.042

1. Introduction

This paper presents an uncertainty quantification (UQ) and global sensitivity analysis (GSA) method designed for relatively high-dimensional problems. It is based on variance-based GSA with the global sensitivity indices (SI) introduced by Sobol' [1] using the Analysis of Variance (ANOVA) expansion for independent inputs. The use of Sobol' indices is now one of the most common GSA techniques. It uses a functional decomposition to incorporate component functions involving single or groups of uncertain parameters, and the computation of the sensitivity measures of each component function is usually done by Monte Carlo (MC) or quasi Monte Carlo (quasi-MC) methods. Theoretical properties of the global SI are established [2–6], and it has proven useful in applications [7]. Similar sensitivity measures were developed independently by Wagner [8] in operations research. However, here at the outset, we should recognize that sensitivity analysis is limited to global methods and to variance-based definitions. Borgonovo and Plischke [9] and Ghanem et al. [10] review local sensitivity measures, derivative-based global measures, moment-independent (density-based) measures, and value-of-information-based measures. Several global measures (including variance-, density-, and value-of-information-based sensitivities) satisfy a common rationale [11], and quantify the discrepancy between unconditional and conditional probability. Borgonovo et al. [11] address the estimation methods for this class of global measures.

It is well-understood that GSA seeks to quantify the overall influence of input parameters on output quantities of interest (see for example Borgonovo et al. [12]). However, a broad evaluation of the parameter space can lead to high cost in, for example, MC or quasi-MC methods. Sudret [13] used generalized Polynomial Chaos expansions (gPC) to build surrogate models for computing Sobol' sensitivity indices based on the surrogate form. Blatman and Sudret [14] further introduced the sparse gPC expansions in order to efficiently estimate the global SI. To avoid the cost of brute-force evaluation in high dimensions, they employ a stepwise approach to select only most significant polynomials by recursively resolving a least-squares regression (LSR) system. The LSR approach is an efficient tool to compute polynomial expansion coefficients, by minimizing the error of the surrogate model representation in the mean-square sense [13–15]. Compared to a projection approach [16–18], where each polynomial coefficient is obtained by computing a multi-dimensional integral, the LSR approach is more flexible (in choosing sampling points), which seems to be generally advantageous for problems involving a potentially large number of uncertain parameters.

In contrast to the Blatman and Sudret [14] approach, Tang et al. [19] used ANOVA expansion and Polynomial Dimensional Decomposition (PDD) [20–24] to represent the expansion's component subspace functions. The main challenge, and motivation for improving PDD methods, is reducing the required number of samples, which corresponds to the cost of the method. Standard ANOVA components increase exponentially with the number of uncertain parameters, with a corresponding polynomial increase of the number of PDD terms for each component function. This limits the LSR approach even for a truncated low-order ANOVA expansion; indeed, for the LSR problem to be well-posed, the number of deterministic model evaluations is necessarily larger than the total polynomial expansion size [14]. Thus, Tang et al. [19] introduced an advantageous combination of the adaptive ANOVA method [25,26] and the stepwise regression technique [14], with which a sparse surrogate model representation can be efficiently constructed. A key to doing this is the use of variance contribution to select the most important polynomials. Tang et al. [19] demonstrated that this can be more efficient. Thus, the present work is based on the use of this approach to take advantage of its efficiency for building sparse surrogates when the superconvergence of PDD/gPC expansions cannot be achieved, as is often the case in practice.

Correlative effects are particularly important in the current combustion application introduced in Sections 9 and 10. Definitions and numerical techniques can be developed for global sensitivity analysis with correlated inputs [27–29]. Kucherenko et al. [29] employed the same definition as Sobol' SI [1] for cases with correlative inputs, by using the conditional expectation for first-order SI, and the conditional variance for total sensitivity indices (TSI). These variance-based SI (Var-SI) [29] are consistent with Sobol' SI in case of independency, and in case of correlation their first-order SI are still strictly positive and bounded by $[0, 1]$. The computation relies on the use of MC methods with the knowledge of conditional probability density function (pdf). Alternatively, Li et al. [27] originally introduced the covariance-based sensitivity indices (Cov-SI) to account for correlative effects among ANOVA component functions due to input correlation. They introduced three Cov-SI (structural, correlative, and total contribution) for each component function, which reduce to a single index for independent inputs. Rahman [30] later defined the same sensitivity indices [31]. Under a boundedness type assumption on the joint pdf of inputs, Chastaing et al. [28] showed the availability of a generalized ANOVA decomposition, based on which they defined generalized sensitivity indices that are very similar to Cov-SI [27]. Li and Rabitz [31] recently established for correlated inputs a relationship between Var-SI and Cov-SI, and showed that Var-SI can be estimated from Cov-SI by using a modest number of model evaluations. An advantage of the Cov-SI [27] compared to Var-SI [29] is that a conditional pdf is not required when computing component functions, which is an attractive property since conditional distribution is rarely known completely, such as in our combustion application. Thus, we use the approach of Li et al. [27]. However, a disadvantage of this approach is that neither the Cov-SI are strictly non-negative nor their magnitude bounded. Consequently, one can obtain extremely large Cov-SI in presence of strong negative correlation between ANOVA components. We thus introduce generalized covariance-based sensitivity indices (mCov-SI), with magnitude $\in [0, 1]$.

The specific formulation we use is developed in Section 2, with emphasis on the new contributions. Overall it includes the ANOVA decomposition, polynomial representation of ANOVA components, and LSR approach for computation of coefficients. The adaptive sparse meta-modeling approach with stepwise LSR is summarized in Section 3. The variance-based SI (Var-SI) are presented in Section 4 and the original Cov-SI for correlated inputs in Section 5 as a foundation. Only key