PCE-based Sobol' indices for probability-boxes

R. Schöbi[*] and B. Sudret[*]

[*] *ETH Zürich, Institute of Structural Engineering,*
Chair of Risk, Safety & Uncertainty Quantification
Stefano-Franscini-Platz 5, CH-8093 Zürich

April 22, 2016

Abstract

In modern engineering, complex systems and processes are modelled by computational simulation tools, such as finite element models (FEM), to avoid performing expensive physical experiments. Such simulations map a set of input parameters through a computational model to a quantity of interest (QoI). The input parameters are often not perfectly known, because they are estimated based on noisy measurements and expert judgement or because they have an intrinsic variability. This, in turn, introduces uncertainty in the QoI.

The uncertainty in the input parameters are often represented probabilistically by means of joint probability distributions. The distributions are whenever possible inferred from available data. However, a common situation in practice is to have only scarce or incomplete data to characterize a probabilistic random variable. This introduces epistemic uncertainty (lack of data, lack of knowledge) alongside aleatory uncertainty (natural variability) as a source of uncertainty. In such cases, a more general framework, such as probability-boxes (Ferson and Ginzburg, 1996; Oberguggenberger et al., 2009), is required to describe the input uncertainty appropriately. Probability-boxes (p-boxes) describe the cumulative distribution function (CDF) F_X of a random variable X by lower and upper boundary curves, *i.e.* $\underline{F}_X$ and $\overline{F}_X$, respectively. The true but unknown CDF lies between those boundaries, *i.e.* $\underline{F}_X(x) \leq F_X(x) \leq \overline{F}_X(x),\ \forall x \in X$. A p-box provides a natural framework to distinguish the two sources of uncertainty: its shape captures aleatory uncertainty, whereas its width captures epistemic uncertainty.

In this context, engineers are concerned with the contribution of the uncertainty in each input variable to the uncertainty in the QoI. Global sensitivity analysis provides a powerful set of tools to quantify the relative importance of each input parameter of the computational model with respect to the QoI. In particular, Sobol' indices are a popular variance-based global sensitivity measure, based on decomposing the variance of the output variable in terms of the variances of the input parameters (Sobol', 1993). However, Sobol' indices have been applied mainly to probabilistic variables rather than p-boxes.

In this paper, we extend the usage of Sobol' indices to the context of p-boxes. In particular, we make use of two recent developments. First, it has been shown in the recent literature, that the Sobol' indices can be estimated efficiently using sparse polynomial chaos expansions (PCE) (Sudret, 2008; Blatman and Sudret, 2010). PCE is a meta-modelling technique which approximates the computational model by a sum of weighted multivariate polynomials orthogonal with respect to the distributions of the input parameters (Blatman and Sudret, 2011). Second, Schöbi and Sudret (2015) introduced a generalized form of PCE suitable for p-boxes, based

on the definition of a suitable augmented space. In this contribution, we propose to build Sobol' indices for p-boxes by post-processing the corresponding augmented-space PCE.

The resulting Sobol' indices are interval-valued due to the definition of the input p-boxes. The interval width of each Sobol' index accounts for the epistemic uncertainty of the sensitivity measure. Hence, the interval width is in principle reducible by better characterizing the input variable. The intervals are an additional piece of information compared to the Sobol' indices based on probabilistic random variables. This information can be used for scheduling additional data generation for input variables when a constraint budget for additional measurements is available. Variables with large uncertainty bounds are preferred candidates for data enrichment.

The proposed approach is illustrated on a number of benchmark application examples. They show that an efficient estimation of PCE-based Sobol' indices is feasible in the context of p-boxes. Despite the increased complexity of the analysis, a small number of evaluations of the computational model is sufficient to estimate accurately the Sobol' indices.

References

Blatman, G. and B. Sudret (2010). Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. *Reliab. Eng. Sys. Safety 95*, 1216–1229.

Blatman, G. and B. Sudret (2011). Adaptive sparse polynomial chaos expansion based on Least Angle Regression. *J. Comput. Phys 230*(6), 2345–2367.

Ferson, S. and L. R. Ginzburg (1996). Different methods are needed to propagate ignorance and variability. *Reliab. Eng. Sys. Safety 54*(2-3), 133–144.

Oberguggenberger, M., J. King, and B. Schmelzer (2009). Classical and imprecise probability methods for sensitivity analysis in engineering: A case study. *Int. J. Approx. Reason. 50*(4), 680–693.

Schöbi, R. and B. Sudret (2015). Propagation of uncertainties modelled by parametric p-boxes using sparse polynomial chaos expansions. In *Proc. 12th Int. Conf. on Applications of Stat. and Prob. in Civil Engineering (ICASP12), Vancouver, Canada.*

Sobol', I. (1993). Sensitivity estimates for nonlinear mathematical models. *Math. Modeling & Comp. Exp. 1*, 407–414.

Sudret, B. (2008). Global sensitivity analysis using polynomial chaos expansions. *Reliab. Eng. Sys. Safety 93*, 964–979.