Efficient Utilization of Surrogate Models for Uncertainty Quantification

Maria Böttcher¹, Ferenc Leichsenring¹, Alexander Fuchs¹, Wolfgang Graf¹, and Michael Kaliske¹,∗

¹ Institute for Structural Analysis, Technische Universität Dresden, 01062 Dresden, Germany

Numerical simulations for the analysis and the design of structures or systems are often based on deterministic characteristics, whereas the reality is determined by data and information which are characterized by various types of uncertainty (variability, imprecision, inaccuracy, incompleteness). Besides traditional probabilistic approaches, possibilistic uncertainty models are most recently in the focal point of research. Combining the characteristics of aleatoric and epistemic uncertainties, polymorphic uncertainty yields various uncertainty approaches related to the application field of imprecise probability models. Uncertainty analysis schemes, based on pointwise evaluation of a fundamental solution (structural analysis), usually lead to high computational costs, due to repetitive evaluations. Especially for complex uncertainty models, each uncertainty characteristic demands a separate quantification. Hence, surrogate models are indispensable in uncertainty quantification.

Considering complex, high dimensional structural models, even single model evaluations are defined by significant computational cost. Therefore, the transition from traditional space-filling sampling schemes towards adaptive sampling schemes allows for an overall decrease of computational cost while maintaining a necessary accuracy of the prediction. A variance based adaptive sampling strategy is proposed, constituted by the combination of exploration and exploitation of a given input space. The predictions’ variances of multiple surrogate models are hereby utilized to detect regions of interest and define sampling points accordingly.

The contribution briefly introduces the fundamentals of fuzzy-set theory and interval based uncertainty analysis as well as a polymorphic uncertainty model. A generalized computational framework for uncertainty analysis is presented with special respect to corresponding computational effort. Subsequently, a concept for an variance based sampling scheme is explained and further utilized in an uncertainty analysis of a multi story structural example with uncertain material properties.

© 2021 The Authors Proceedings in Applied Mathematics & Mechanics published by Wiley-VCH GmbH

1 Introduction

In the process of structural design optimization, a variety of uncertainties regarding the design parameters, including natural variability of material parameters, production tolerances and load scenarios, are to be considered in order to determine not only a prediction of the structures performance, but also its uncertainty. A short overview of approaches for uncertainty quantification and analysis is given. As the evaluation of a complex simulation model predicting a systems behavior is computationally expensive, making an uncertainty analysis unfeasible in most cases, surrogate models are being utilized to severely decrease the computational effort. In order to obtain a sufficiently accurate model, a comprehensive training data set, representing the input output relationship over the entire input domain, is crucial. To reduce the amount of samples necessary and thereby decreasing the number of needed simulation model evaluations, adaptive sampling strategies, which are selectively described in this contribution, are being developed with the objective of maximizing the information gained by each new sample. By employing both uncertain model parameters and a trained surrogate model, the applicability for an engineering task is shown in an exemplary uncertainty analysis.

2 Framework for Polymorphic Uncertainty Quantification

It is common to distinguish between two general concepts of uncertainty, namely aleatoric and epistemic uncertainties. Where aleatoric uncertainty models, such as randomness, incorporates the variability of data or measurements, epistemic uncertainty considers e.g. incompleteness due to lack of knowledge or a small amount of available data. Interval variables can be utilized in cases of e.g. few to none data samples or if no realistic assessment of certain data points is possible or reasonable. An interval quantity is defined by its bounds

\[ I = [x_1, x_r] = \alpha \land x_1 < x_r . \]  

(1)

Fuzzy-sets allow an assessment of possibility (e.g. due to expert knowledge) by supplementing weights to certain parameter ranges, yielding a corresponding membership function holding the following properties according to [1, 2]

\[ \mu : \mathbb{R} \mapsto [0, 1] , \]  

(2)

\[ \exists x \ | \ \mu(x) = 1 , \]  

(3)

\[ A_\alpha^f = \{ x \in \mathbb{R} \ | \ \mu_A \geq \alpha \} \text{ with } \alpha \in [0, 1] . \]  

(4)
In terms of computational evaluation, fuzzy-variables could be interpreted as stacked interval quantities at each discretized level of $\alpha$. A convex fuzzy variable is therefore composed by multiple discretized $\alpha$-levels $A^\alpha = (A^\alpha_i)_{\alpha \in [0,1]}$. A combination with probabilistic approaches towards polymorphic uncertainty methods is extensively discussed in e.g. [2–4]. The definition of e.g. fuzzy probability based random variables (fp-r) is founded on the assumption, that the probability distribution of a random variable $X$ cannot be described exactly due to lack of information, see e.g. [3, 4]. The fuzzy probability distribution might be represented by a fuzzy cumulative distribution function (CDF) $\tilde{F}_X$, which, again, is defined as a family of $\alpha$-cuts with an arbitrary cumulative distribution function $F(x)$ as well as distribution parameters $\theta_i$ in terms of $F(x, \theta)$. Then, the fuzzy cumulative distribution function $\tilde{F}_X$ can be described by fuzzy distribution parameters $\theta^i = (\theta_{i,1}, \theta_{i,2})_{\alpha \in [0,1]}$. For instance, a two parametric distribution function with parameters $\theta_1$ and $\theta_2$ yields

$$\tilde{F}_X = \{F_{\theta_1 \times \theta_2} \mid \theta_1 \in \theta^i_{1,\alpha}, \theta_2 \in \theta^i_{2,\alpha}\}_{\alpha \in [0,1]}.$$  

This formulation is referred to as bunch parameter representation, since the fuzzy cumulative distribution and the fuzzy probability density function can be considered as assessed bunches of functions which are described by bunch parameters $\theta^i$. The computational framework is based on staggered evaluation loops, which are either defined as fuzzy analysis or stochastic analysis. Uncertainty quantification based on various uncertainty models in the field of imprecise probability, such as, e.g., p-boxes, is determinable by these general concepts of evaluation. The schematic depiction of an exemplary computational order for fp-r variables in Fig. 1 shows, that the total amount of samples is equal to the product of each separate amount of evaluation samples. Hence, surrogate models for the pointwise approximation of the fundamental model input output dependency are almost always necessary in order to keep the computational cost within a feasible range.

### 3 Adaptive Sampling-based Surrogate Model

In order to investigate the uncertainty of a structural design’s performance, i.e. the uncertainties of the computed responses for uncertain design parameters, an uncertainty analysis is performed, which requires a large number of model evaluations. For FEM simulations in standard engineering tasks, this is infeasible due to the large amount of time necessary for a single prediction of the structure’s response. Contrary, a surrogate model $M$ is able to give the response in a fraction of time, allowing the uncertainty quantification of the response parameters. Basis for such a model is a training data set, which is covering the entire input and output domain sufficiently in order to describe the behavior of the function with the desired accuracy. The expensive simulation model is employed to generate the training data, which in itself can be a time-consuming task for a large amount of data samples. With the objective of minimizing the computational effort, the training data set should be designed in such a way as to minimize the number of samples necessary. Therefore, each sample has to be selected where the information gain about the input to output relationship is the highest. However, often little to no knowledge about the function behavior is available in advance, making it difficult to find such sample points.

Different approaches are used in order to choose the training data, which can be classified into non-adaptive, all-at-once sampling strategies, e.g. Random Sampling or Latin Hypercube Sampling, where one Design of Experiments (DoE) is created with the maximum affordable number of sample points $n_{\text{max}}$ in order to most efficiently fill the input domain, and adaptive sampling strategies. The latter are based on a small, initial DoE using non-adaptive sampling strategies, and select iteratively new samples based on information from the existing situation, i.e. the previously selected input parameter combinations with their respective response obtained by the simulation model. The benefits here are for one, that, due to the iterative increase in amount of samples, the process can be stopped as soon as the surrogate model’s prediction reaches a desired accuracy, and the information about the response can be utilized in addition as a sample selection criterion, by prioritizing regions of high non-linearity in form of an exploitation criterion. The space-filling of the input domain is regarded as an exploration criterion, and a trade-off of both is frequently utilized. Most challenging thereby is the definition of “neighboring samples” for multi-dimensional input domains to approximate the response’s non-linearity regarding the exploitation objective.

A sub-classification of adaptive sampling strategies can be made into two groups. Firstly, data-based adaptive sampling strategies consider the existing data samples to formulate a sample selection criterion. Examples are the input space-filling Maximin approach [5, 6] and LOLA-Voronoi [7], where space-filling of the input domain is combined with the non-linearity
of the response as a hybrid criterion. The second group is that of surrogate model-based approaches, where the existing training data is used to train the surrogate model at each step. In addition to the criteria of the data-based approaches, the information about the surrogate model’s prediction is applied for the selection of the next sample. Those methods include, e.g., the KRGING-based sampling strategy [8, 9], where the value of the predictions’ standard deviation of a KRGING model is used as measure of how accurate the prediction is over the input domain, and Leave-One-Out-based approaches [10, 11], which train similar surrogate models multiple times on the training data, but always omitting one (or several) different samples from the set and thereby finding those samples with high information content based on the decrease of the surrogate models’ accuracy.

Another adaptive, surrogate model-based sampling strategy is the Ensemble Learning Sampling Approach (ELSA) presented in [12]. Similar to the idea of Ensemble Learning, where, in order to reduce the risk of local inaccuracies of a model, multiple surrogate models, instead of merely one, are used to make a prediction, e.g. by selecting the mean of all the models’ output. For ELSA, an ensemble of \( n_E \) artificial neural networks, each with randomly initialized weights, is trained, and the variances of the predictions is used as exploitation criterion

\[
V(\tilde{x}_j) = \text{Var}(\tilde{Z}_j), \quad \text{Var}(\tilde{Z}_j) = [\text{Var}(\tilde{\varepsilon}_{1j}) \ldots \text{Var}(\tilde{\varepsilon}_{nj})]^T \forall \tilde{\varepsilon}_{ij} \in \tilde{Z}_j, \quad \tilde{Z}_j = [M_1(\tilde{x}_j) \ldots M_{nj}(\tilde{x}_j)]^T,
\]

as it can be observed, that the predictions vary the most in areas, where samples, which lie within close proximity to another, display a comparatively high variation of the respective responses, i.e. high non-linearity. To additionally give preference to those regions of the input domain, which contain a low number of samples, a Kernel Density Estimation (KDE) using the GAUSSIAN kernel

\[
\hat{f}(x) = \sum_{i=1}^{n_{\text{max}}} k(x, \tilde{x}_i, l) \quad \text{with} \quad k(x, \tilde{x}_i, l) = \exp\left(-\frac{(x - \tilde{x}_i)^2}{2l^2}\right), \tag{7}
\]

with the bandwidth \( l \) is utilized. The resulting hybrid criterion for selecting a new sample is

\[
S(\tilde{x}_j) = w \cdot V(\tilde{x}_j) + (1 - w) \cdot D(\tilde{x}_j) \in [0, 1] \quad \text{with} \quad D(\tilde{x}_j) = 1 - \hat{f}(\tilde{x}_j), \tag{8}
\]

where \( D \) describes the individual criterion normalized onto the interval \([0, 1]\) and \( w \) defining the weight of the variance compared to the density criterion, i.e. the trade-off between exploitation and exploration.

### 4 Numerical Example for Multi-dimensional Uncertainty Quantification

An undamped three story frame structure with concentrated masses, as depicted in Fig. 2, is investigated regarding the eigenvalues. For demonstration purposes, the degrees of freedom are decreased from 18 to 3 by considering the bars as rigid \((EI_B = \infty)\) as well as the axial rigidity of all elements as \(EA = \infty\) and, thereby, eliminating the nodes’ vertical displacement and torsion. The individual masses \(m_i \in [0.5, 5.0] \text{Mg}, i \in \{1, 2, 3\}\) and the stiffness of the columns \(EI_C \in [2.0, 10.0] \times 10^6 \text{kNm}^2\) as design parameter \(\tilde{x}\) constitute the four-dimensional input domain \(\mathcal{X}\). The three eigenvalues \(\omega_{EI}, i \in \{1, 2, 3\}\) are the response parameters \(\tilde{x}\), which can be obtained analytically by computing the linear algebraic eigenvalue problem

\[
(K - \omega_{EI}^2 M) \cdot \tilde{x}_{EI} = 0 \quad \text{with} \quad K = \begin{bmatrix}
k_{1} + k_2 & -k_2 & 0 \\
-k_2 & k_2 + k_3 & -k_3 \\
0 & -k_3 & k_3
\end{bmatrix}, \quad M = \begin{bmatrix} m_1 & 0 & 0 \\
0 & m_2 & 0 \\
0 & 0 & m_3
\end{bmatrix}, \tag{9}
\]

and \(k_i = 12 \cdot (2EI_C)/l_i\). As the initial training data set, \(n_t = 25\) data samples are randomly generated within the input domain \(\mathcal{X}\), and the adaptive sampling procedure ELSA, with \(n_{\text{max}} = 300\), is performed with an ensemble of \(n_E = 5\) neural networks, each consisting of two hidden layers with 50 neurons each and the activation function \(\text{tanh}\). Per iteration step, 4 new sample points are selected and the responses computed. The resulting surrogate model has a mean squared error of \(\varepsilon = 5.25 \cdot 10^{-7}\), which is about 14 times smaller compared to the error \(\varepsilon = 7.53 \cdot 10^{-6}\) of a similar neural network trained with data originating from Random Sampling. The evaluation of the surrogate model provides an approximately tenfold decrease in computation time by comparison to the analytical solution.

The best resulting surrogate model is now utilized for an uncertainty analysis of the structure. Given the assumption, that very few information about the design parameters is available, their uncertainties are quantified by using fuzzy variables \(x_i^f\), specifically the convex fuzzy variables of trapezoid shape \(m_1^f = m_2^f = (1.9, 1.98, 2.02, 2.1) \text{Mg}, m_3^f = (2.375, 2.475, 2.525, 2.625) \text{Mg}\) and \(EI_C^f = (4.275, 4.455, 4.545, 4.725) \times 10^6 \text{kNm}^2\). To compute the corresponding fuzzy responses \(\tilde{x}_i^f\), an optimization
using $\alpha$-level discretization for the three levels $\alpha = \{0.0, 0.5, 1.0\}$ is performed. The responses $z^f$, i.e. the three fuzzy eigenvalues, resulting from the uncertainty analysis, are shown in Fig. 3, where a positive correlation between the responses is clearly visible. Compared to an independent consideration of the eigenvalues $\lambda^f$, the volume of the fuzzy responses $z^f$ at each discretized $\alpha$-level is reduced to less than 6%, as shown in Table 1.

![Fig. 2: Three story frame structure.](image)

![Fig. 3: Result of uncertainty analysis with fuzzy input parameters.](image)

### 5 Conclusion and Outlook

With increasing integration of uncertainty quantification as part of a structural design process, surrogate models are being utilized with the objective of decreasing the computation effort. However, given that a non-negligible amount of samples is necessary in order to train such a model, adaptive sampling strategies have been developed to reduce this amount. The ELSA, combining an exploitation criterion based on the predictions variances of an ensemble of neural networks and an exploration criterion based on a Kernel Density Estimation detecting undersampled regions of the input domain, is presented according to [12]. A neural network as a surrogate model, trained on the thusly selected samples, is utilized to perform an uncertainty analysis of a frame structure with fuzzy design parameters and a multi-dimensional response, showing the benefit of considering dependent fuzzy responses over independent ones.

### Acknowledgements

This work is supported by Continental Reifen Deutschland GmbH as well as the German Research Foundation (DFG) within Research Training Group GRK 2250/1, Project B3 and Priority Program SPP 1886/1 “Polymorphic Uncertainty Modelling for the Numerical Design of Structures”, which is gratefully acknowledged.

Open access funding enabled and organized by Projekt DEAL.

### References

[1] M. Götz, Numerische Entwurfsmethoden unter Berücksichtigung polymorpher Unschärfe, PhD thesis, TU Dresden, 2017.

[2] B. Möller and M. Beer, Fuzzy Randomness – Uncertainty in Civil Engineering and Computational Mechanics (Springer, Berlin, 2004).

[3] S. Pannier, M. Waurick, W. Graf, and M. Kaliske, Mechanical Systems and Signal Processing 37, 105–120 (2013).

[4] M. Götz, W. Graf, and M. Kaliske, International Journal of Reliability and Safety 9, 112–131 (2015).

[5] M. E. Johnson, L. M. Moore, and D. Ylvisaker, Journal of Statistical Planning and Inference 26, 131–148 (1990).

[6] P. Luc, Journal de la Société Française de Statistique 158, 7–36 (2017).

[7] K. Crombez, D. Gorissen, D. Deschrijver, and T. Dhaene, SIAM Journal on Scientific Computing 33, 1948–1974 (2011).

[8] L. Zhao, K. K. Choi, and I. Lee, 13th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference (2010).

[9] H. Liu, J. Cai, and Y. S. Ong, Computers & Chemical Engineering 106, 171–182 (2017).

[10] R. Jin, W. Chen, and A. Sudjianto, Proceedings of the 28th ASME Design Engineering Technical Conference 2, 539–548 (2002).

[11] G. Li, V. Aute, and S. Azarm, Structural and Multidisciplinary Optimization 40, 137–155 (2010).

[12] M. Böttcher, A. Fuchs, F. Leichsenring, W. Graf, and M. Kaliske, Advances in Engineering Software (2020), submitted 23.06.2020.