Advances in Reliability, Safety and Security

ESREL 2024
Collection of Extended Abstracts

Introduction Of Modified Sample Selection Scheme For Adaptive Ensemble Learning-Based Sampling Approach

Maria Böttcher, Wolfgang Graf, Michael Kaliske

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

Keywords: adaptive sampling, metamodeling, design of experiments, space-filling, neural networks

Increasing computational capabilities have long been successfully utilized for design optimisation in structural engineering, investigating structural responses with high-fidelity simulation models of possible design parameter combinations. However, the more detailed these simulation models are, the more expensive a single performance prediction becomes, which can make the determination of the optimal design a time-consuming task. For considerations of robustness, uncertainty analyses need to be performed, entailing a large number model evaluations for each design, which again raises the computational effort significantly. A solution approach to this challenge is to create a fast to evaluate metamodel trained on computationally expensive data samples generated from the numerical simulation, which the metamodel then replaces for the design optimization.

The utilization of metamodels has been successfully applied to perform uncertainty analysis for various design tasks, see, e.g. (Serafinska et al., 2016; Götz et al., 2018; Freitag et al., 2020). However, this raises a new challenge: the generation of a high-quality training data set which allows the metamodel to learn the input-to-output relationship of the original model over the entire input domain. The acquisition of these data samples is an expensive task. Therefore, the objective of the sample selection is to maximize the information gained by each new sample. However, since often little to no prior knowledge exists about the model and which regions are of interest, i.e., show a highly nonlinear response, adaptive sampling strategies have been developed. Starting from a small initial data set, new sample points are iteratively selected based on information resulting from the already existing data samples.

Many different adaptive sampling strategies have been developed, see, e.g. (Crombecq et al., 2011; Liu et al., 2017; Fuhg et al., 2021). This contribution focuses on improvements for the method ELSA, introduced in (Böttcher et al., 2021). An adaptive sample selection scheme is investigated to improve the efficiency of the method.

Adaptive, or sequential, sampling strategies can reduce the amount of training data samples needed in order to train a metamodel with sufficient accuracy. An overview can be found in, e.g. (Fuhg et al., 2021). Starting with a small initial data set created by, e.g. Design of Experiments, the goal is to iteratively create samples in "interesting" regions, so that this new sample yields as much new information as possible. The "interesting" regions can be, on the one hand, regions with a highly non-linear functional response, which will be considered by an exploitation criterion. The other aspect which is to be regarded in the adaptive sampling procedure is the training data density in the input domain. If the density is low, then these areas will be prioritized for selecting a new sample to prevent areas of interest in the response domain to remain undetectable. The sample density of the input domain is considered by an exploration criterion.

By combining both criteria, adaptive sampling strategies aim to select new sample points based on a trade-off of exploration and exploitation. Various strategies exist, which differ in their basis for these criteria, and can be separated into two groups: data-based strategies, which select new samples solely based on the existing samples, and metamodel-based strategies, where the trained metamodel(s) contribute(s) to evaluating possible new sample points.

In ELSA (Böttcher et al., 2021), a novel metamodel-based adaptive sampling strategy is proposed, which uses the kernel density estimate (KDE) as the exploration criterion $D(\underline{x})$ and the prediction variance of an ensemble of neural networks as the exploitation criterion $V(\underline{x})$. A new sample is selected based on a hybrid criterion

$$S(\underline{x}) = w \cdot \overline{V}(\underline{x}) + (1 - w) \cdot \overline{D}(\underline{x}) \in [0, 1], w \in [0, 1], \tag{1}$$

which combines the $[0, 1]$-scaled values of exploration and exploitation criteria $\overline{D}(\underline{x})$ and $\overline{V}(\underline{x})$ with a weight parameter w. The method is illustrated in Figure 1. In the original publication, the weight parameter w is fixed throughout the iteration process.

To further improve the algorithm, a scheme to adaptively modify the weight parameter based on the previous iteration steps is investigated in this contribution. Furthermore, the possibility to adjust the $[0, 1]$-scaling of each iteration is examined. The aim is to include information of the importance of each criterion in the current iteration step, since the original approach does not yet allow for a reduction of one of the criteria's importance in case it assumes almost identical values throughout the input domain.

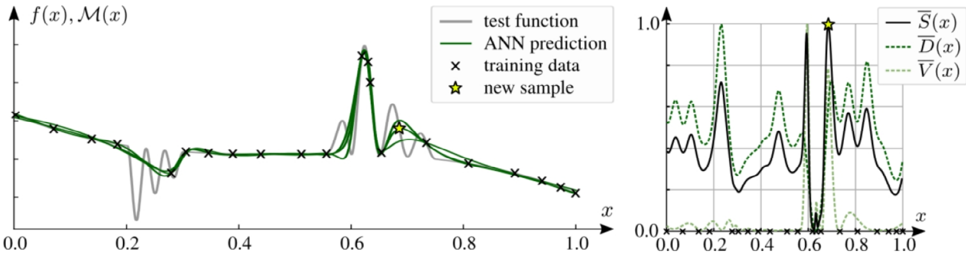

Fig. 1. Selection of new sample point using ELSA. Left: prediction of ensemble of ANNs. Right: scaled values of selection criteria.

Adaptive sampling strategies have shown to be capable to significantly reduce the necessary amount of training data for metamodels to reach acceptable approximation errors. The improvement of ELSA by introducing adaptive weighing and a modified scaling scheme is investigated for various numerical test functions. When the unscaled values of the criteria are considered as well, the algorithm could further reduce the necessary number of samples for a sufficiently accurate metamodel. For example, if the unscaled exploitation criterion decreases first rapidly and then much slower, an exploration needs to be favoured to create the next new samples.

Acknowledgements

This work is funded by Continental Reifen Deutschland GmbH as well as by the Deutsche Forschungs-gemeinschaft (DFG, German Research Foundation) – CRC/TRR 339, Project-ID 453596084, which is gratefully acknowledged.

References

Böttcher, M., Fuchs, A., Leichsenring, F., Graf, W., Kaliske, M. 2021. ELSA: An efficient, adaptive Ensemble Learning-based Sampling Approach. Advances in Engineering Software 154, 102974.

Crombecq, K., Gorissen, D., Deschrijver, D., Dhaene, T., 2011. A Novel Hybrid Sequential Design Strategy for Global Surrogate Modeling of Computer Experiments. Society for Industrial and Applied Mathematics 33, 1948-1974.

Freitag, S., Edler, P., Kremer, K., Meschke, G. 2020. Multilevel surrogate modeling approach for optimization problems with polymorphic uncertain parameters. International Journal of Approximate Reasoning 119, 81-91.

Fuhg, J.N., Fau, A., Nackenhorst, U. 2021. State-of-the-Art and Comparative Review of Adaptive Sampling Methods for Kriging. Archives of Computational Methods in Engineering 28, 2689-2747.

Götz, M., Leichsenring, F., Kropp, T., Müller, P., Falk, T., Graf, W., Kaliske, M., Drossel, W.-G. 2018. Data Mining and Machine Learning Methods Applied to A Numerical Clinching Model. Computer Modeling in Engineering and Sciences 117, 387-423.

Liu, H., Cai, J., Ong, Y.-S. 2017. An Adaptive Sampling Approach for Kriging Metamodeling by Maximizing Expected Prediction Error. Computers and Chemical Engineering 106, 171-182.

Serafinska, A., Hassoun, N., Kaliske, M. 2016. Numerical optimization of wear performance – Utilizing a metamodel based friction law. Computers & Structures 165, 10-23.