AdCo GW ENGINEERING

Advanced Computational Engineering
for Your
Advances via Computational Engineering

Uncertainty Quantification: Why and How

Executive Summary

As the digitization puts an ever greater pressure on companies to shorten their development cycles and reduce costs, the demand on the predictive capability of modern Computer Aided Engineering (CAE) toolchains increases. This implies that simulation models are not only required to correctly represent the underlying physical system, but also that real-world variability and uncertainty must be taken into account for enabling a shifting of the product development cycle more and more from physical experiments to the computer. Furthermore, one has to consistently incorporate available data from physical tests into the modeling pipeline. Uncertainty quantification (UQ) is key in improving the predictive capabilities of state-of-the-art CAE workflows. Furthermore, it speeds up the product design and development processes, particularly in the early phases of those processes, which not only reduces the time required to develop a product, but also the number of potentially expensive physical experiments. Below, we will show that efficient UQ techniques available today will change the way design decisions are made. By incorporating imperfect and incomplete information into the computational product design process, the uncertainty in the results can be quantified. The implications of this paradigm shift are manifold. First and foremost, it means that an even larger portion of the design process can be carried out by using computers, which leads to reduced cost and shorter product development cycles. In the end, the eventual product designs will be less prone to failure, while at the same time being less costly to manufacture. AdCo Engineering^{GW} bundles know-how from statistics, machine learning and engineering to provide innovative algorithms for the quantification of uncertainties in our software QUEENS. Thereby, we enable our clients to harness the full potential that digital product development tools and processes can offer.

1 Introduction

Decreasing costs of high-performance computing platforms as well as recent advances in computational methods along with corresponding software solutions have led to tremendous progress in the field of Computer Aided Engineering (CAE) in recent decades. Simulations have become omnipresent and irreplaceable in product development processes and in the planning of production processes. Computer-based simulation methods have reached a quite high degree of maturity, and a multitude of complex processes, for example, by taking various physical, chemical and biological fields into account, can nowadays be reproduced with the aid of computational models.

However, there is still often a discrepancy between the results of computational models, on the one hand, and the results of physical experiments, on the other hand. The main reasons for this discrepancy may be found, for instance, in the inadequacy of the chosen model to reproduce physical reality, experimental measurement errors, and idealized assumptions for model input parameters, which do not hold in reality. Particularly the latter point has received increased attention in academic research during the last decades and led to the rise of a new sub-discipline of CAE, referred to as Uncertainty Quantification (UQ). Methodological breakthroughs in recent years have made UQ approaches mature enough for industrial application. The incorporation of real-world variability and randomness into CAE models and the rigorous quantification of these uncertainties result in improvements of the digital design process. The consequences are that a more and more increasing portion of the product or process design can be carried out by using computers, which substantially abridges development cycles and reduces overall costs.

UQ enables engineers to move beyond unspecific and empirical safety factors and achieve a considerably more detailed, nuanced, and comprehensive understanding of the system under investigation via computational methods. Furthermore, with UQ, it is possible to incorporate available data and knowledge in the model in a mathematically rigorous way. This is particularly important if - as often - the simulation of a system or process is the only reasonable or even possible method of investigation.

2 What is UQ and how does it help: examples and benefits

UQ means characterizing imperfect and/or unknown information in engineering simulations with the goal of enabling reliable predictions and facilitating decision making. In any real-world engineering setting, there is a plethora of sources of uncertainty due to the fact that there are typically a multitude of parameters related to the model which are based on assumptions, idealizations, and insufficient data. Examples for uncertain parameters are, e.g., geometric uncertainties due to manufacturing imperfections and tolerances, varying properties of the material used for a product, or uncertain load or bearing conditions. In addition to these parametric uncertainties, the model itself is often uncertain as well. Therefore, purely deterministic analyses often fall short of expectations in terms of prediction accuracy and reliability, since critical parameters included in the model are either not determinable with sufficient accuracy or subject to random fluctuations, respectively.

A UQ analysis taking into account the uncertainties in the model input variables significantly improves the predictive capability of computational design methods. Thus, it should be aimed at quantifying these uncertainties in the simulation results as far as possible. In the context of UQ, this uncertainty is typically represented by means of probability distributions or confidence intervals.

In addition to estimating the expected accuracy of the simulation model, the identification of the most influential parameters by means of global sensitivity analyses and the robust design with regard to parameter uncertainties play an important role. In almost all industries and scientific disciplines in which simulation models are used today, e.g., in mechanical and plant engineering, aerospace engineering, mining, in the energy sector and even in financial mathematics, there are numerous lines of evidence supporting the enormous advantages of taking uncertainties into account in simulation models. The following examples from mechanical and biomedical engineering should demonstrate exemplarily these advantages and the potential of UQ.

Example 1:

In some high-precision manufacturing processes, e.g., deep drawing with high-strength steels, the final component geometry is partly influenced by sheet batch-dependent fluctuations in material properties. To compensate for these fluctuations, iterative process and tool changes have to be made in series production, which are particularly cost-intensive. By including such fluctuations as uncertainties in high-resolution simulation models, they can be identified and taken into account at an early stage of the development process. Thus, process steps and component geometries can be adapted to reduce the influence of sheet-batch dependent fluctuations in material properties. This means that process modes can be avoided during series production, resulting in decisive productivity advantages and significantly reduces production costs.

Example 2:

Another application for which predictive simulations using a stochastic approach are particularly valuable are scenarios which neither allow for experiments nor theoretical considerations and simplified analytical models due to their complexity. Various biomedical applications, which will become increasingly relevant in the future, fall into this category. An important biomedical application of numerical simulation models, which is expected to be used more and more routinely in clinics in the foreseeable future, is the prediction of rupture risks of abdominal aortic aneurysms. Other application scenarios include the patient-specific design of stent grafts for endovascular therapy of these aneurysms or the design of osteosynthesis implants for therapy of bone fractures. In the examples mentioned above as well as in many others, the consideration of uncertainties of certain model input parameters, e.g., patient-specific material properties, is of utmost importance, since incorrect predictions usually have serious consequences for the individual patient.

The rigorous quantification of uncertainties in CAE models thus presents a major paradigm shift from idealized parameter settings and modelling assumptions, which typically require safety factors when designing real-world systems, to the incorporation of actual variability and uncertainty in the CAE process. In our view, there are numerous benefits of incorporating UQ in the product development workflow, and two arguments for using UQ will be outlined exemplarily in the following subsections of this white paper.

Benefit 1: Probabilistic design yields more accurate and reliable results

If one is interested in the average response of a model or system, respectively, it is important to note that – given the fact that the model is nonlinear with respect to the input parameters - its response computed with statistically averaged input parameters does not equal its average response. Hence, even if one is "merely" interested in the average model/system response, incorporating uncertainty will yield more accurate results and provide a more realistic prediction on how the system under investigation will behave on average.

Of course, properly conducted UQ analyses can provide much more information than average system responses. Uncertainties in system strength due to, e.g., variations in manufacturing or material properties combined with uncertainties in field conditions, e.g., variations in loading, result in uncertainties in the system performance. Traditionally, such uncertainties were addressed by using rather large safety factors. This implicit way of taking uncertainties into account often leads to "over-designed" systems, which are, in fact, reliable but notably more expensive to produce than actually required. Moreover, safety factors cannot encompass all combinations of uncertainties in the system, which might propagate and amplify through the system in ways that are difficult or impossible to anticipate. This can lead to system or component failures, increasing the overall lifecycle cost. UQ techniques can make uncertainties explicit and quantifiable, thus providing a more reliable way to assess designs or product variants. In the end, the overall costs of product development processes are substantially reduced.

The difference between a deterministic design and decision-making process and a workflow incorporating UQ is shown in Figure 1. It is evident that additional information obtained by UQ techniques might lead to a different product design.

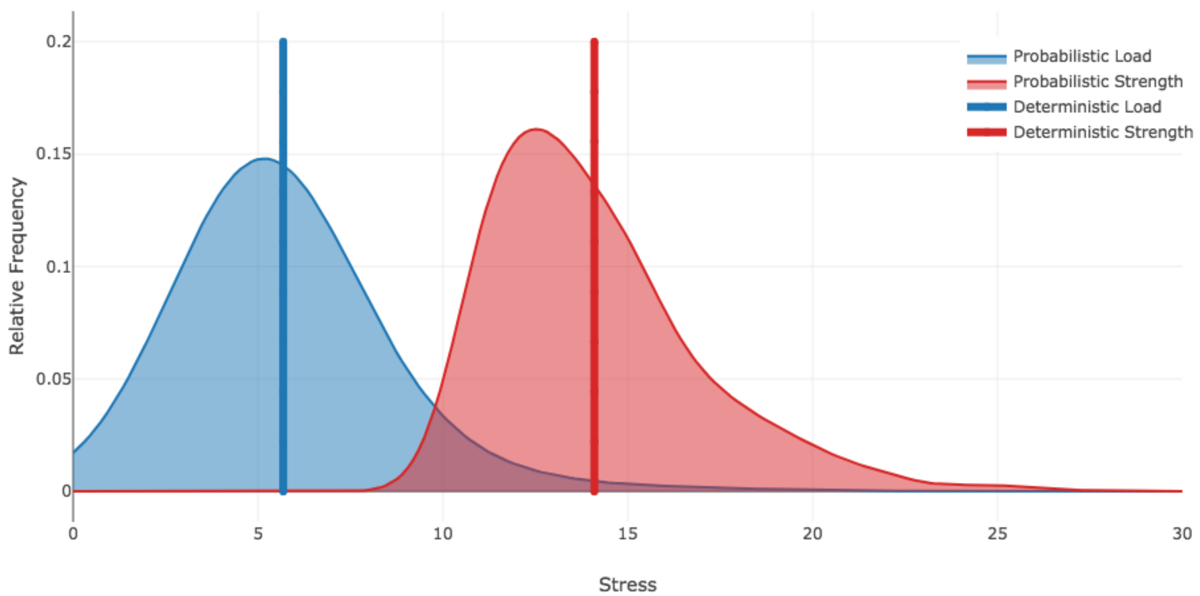

***Figure 1**: Comparison of probabilistic and deterministic system design: here, a deterministic analysis yields a different assessment of system performance compared to a probabilistic approach including uncertainty in both strength and loading conditions. The safety margin between deterministic strength of the system and average loading conditions in this example is rather large. The strength-to-load ratio is approximately 2.4, which is often deemed sufficient. The probabilistic analysis, however, provides an entirely different picture. The relatively large area where the two probability densities intersect indicates clearly that the chance for the load to exceed the strength is quite significant. Thus, the current design based on purely deterministic principles is not safe and needs to be improved.*

Benefit 2: Identification of the most important parameters and efficient resource allocation

Another important benefit of UQ is the identification of the most influential input parameters via global sensitivity analysis. Figure 2 depicts the so-called main effect indices of an exemplary system. The depicted indices indicate which fraction of the output variance results from the isolated variation of each of the individual parameters. The parameters with the largest sensitivity indices should obviously be the focus of subsequent investigations.

By computing global sensitivity indices, one can determine the most influential parameters regarding the output uncertainty. Once these have been identified, it is possible to efficiently allocate resources to reduce the uncertainty in the outcome. This could, e.g., trigger additional experiments with the aim to collect more data on some of the input parameters, to build more accurate, less uncertain probabilistic models for the respective parameters. Another option would be the refinement and improvement of the manufacturing processes to reduce the variability of a specific input parameter.

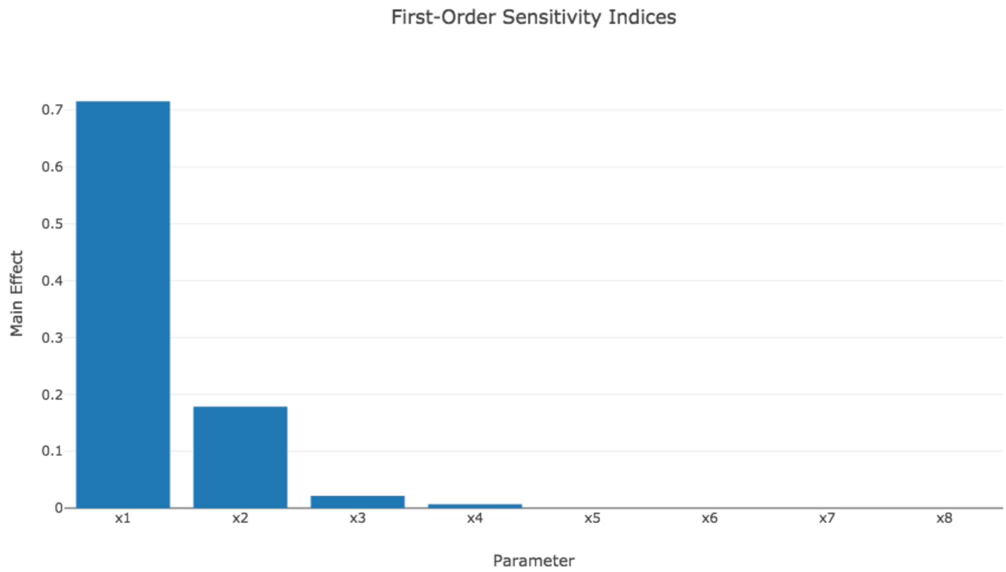

Figure 2: *Main effect indices of an example function. The bar graph shows that the parameter x1 clearly has the largest impact on the result. The parameters x2 and x3 have only a minor influence in comparison and the effect of the remaining parameter is essentially negligible.*

3 How to perform efficient UQ analysis

Performing a UQ analysis is challenging for several reasons, such as potential lack of the required technical know-how and in-depth statistical knowledge to model uncertainties and choose an appropriate, efficient method for the propagation of these uncertainties through the model or system, respectively. The second hurdle used to be the substantially increased computational cost compared to standard deterministic CAE analysis of a system. While the computational cost associated with a stochastic analysis used to be orders of magnitude larger than a deterministic analysis when using traditional UQ approaches, modern UQ methods have reduced this cost to a rather moderate manifold of that of a deterministic analysis. In the following, we will briefly sketch how uncertainties can be described mathematically and present some approaches to the propagation of uncertainties through a simulation model.

Mathematical description of uncertainties

First, an appropriate mathematical description or statistical model for the uncertainties, which influence the system under investigation is required. Typically, random variables and random fields or a combination thereof are used to model uncertain input parameters of simulation models. Random fields are essentially the extension of multivariate random variables to an infinite number of dimensions. Figure 3 depicts two example realizations of two non-Gaussian random fields for illustration. It is a mathematical construct to model a process which dependence on a parameter can be captured by probabilistic laws. This way, spatial or temporal correlation can be considered and incorporated in the model. The notion of a random field allows to model parameters that are subject to spatial and stochastic variation at the same time.

Examples where random fields are used to model uncertainty include spatially varying material and friction parameters, variations in component dimensions due to unavoidable manufacturing tolerances or blurred geometries of biological structures due to insufficient resolution of medical imaging. In the context of UQ, it is not sufficient to determine a probabilistic model for the uncertain input parameters, but one also needs to be able to generate samples from this probabilistic model, to obtain different realizations of the random quantities for which the solution of the model can actually be computed. For random variables, this is rather straightforward. For random fields, though, this task is considerably more complex and requires special techniques and algorithms.

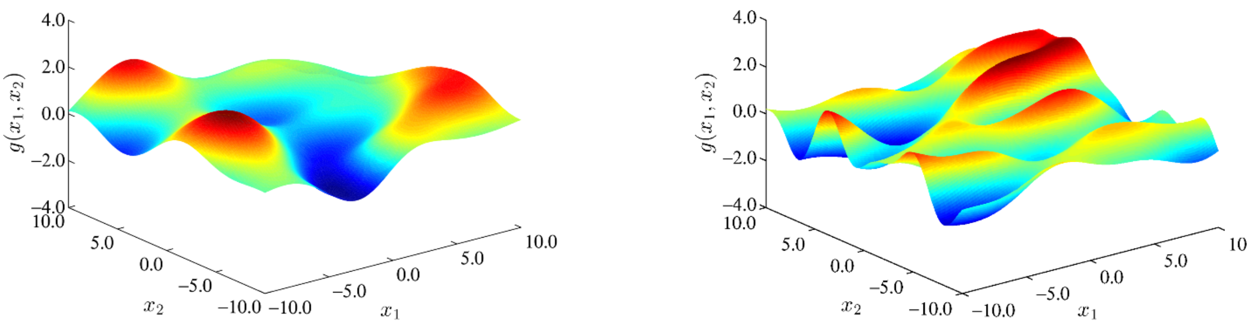

Figure.3: *Realizations of two-dimensional random fields.*

Propagation of uncertainties through the model

As soon as a suitable probabilistic description of the uncertainties in the model has been determined, the next and crucial step in UQ is the so-called propagation of uncertainties through the simulation model. This is typically achieved by evaluating the model for a certain number of realizations of the random inputs. UQ approaches differ in the way these realizations are chosen. The choice varies between generating realizations entirely randomly, on the one hand, and choosing them in a purely deterministic fashion at a number of integration or collocation points, on the other hand. Moreover, UQ methods use the information obtained from individual model evaluations in different ways and impose different assumption about the relationship between model input and model output.

In the following, we will briefly address three classes of methods that are used to propagate uncertainty through computational models. It is impossible to provide blanket advice on which method is the best, since the efficiency of each individual approach is highly dependent on several factors. These factors are, amongst others, the number of random inputs, i.e., the so-called stochastic dimension of the problem, and the characteristics of the problem, that is, for instance, the nonlinearity and the type of nonlinearity in the mapping between inputs and outputs are crucial for the efficiency of a chosen method.

We merely consider UQ approaches that are applicable in combination with industrially applicable simulation models. More precisely, we do not consider so-called intrusive UQ methods, which would require access for being able to modify the source code of the CAE software, but rather methods that can be wrapped around existing CAE software packages without requiring any modification of that software.

Sampling-based approaches

The oldest and conceptually simplest UQ approaches are based on random or quasi-random sampling, i.e., drawing samples from the random input parameters and the subsequent evaluation of the simulation model for these parameters settings. Conclusions about the quantity of interest are then drawn based on the law of large numbers.

The most prominent example of sampling-based methods is the Monte Carlo (MC) method, where the sampling is entirely random. While conceptually rather simple and easy to implement, the computing effort associated with an MC analysis is so demanding that it is usually infeasible for all but the simplest simulation models. The use of more elaborate sampling strategies, such as Latin Hyper Cube, sampling can alleviate this computational burden to some extent. Especially for the computation of rare events, sampling analysis can be significantly expedited by the use of sampling schemes that were designed particularly for the purpose of assessing small failure probabilities in the context of reliability analysis. State-of-the-art results can be obtained with, e.g., subset simulation or sequential importance sampling approaches.

Surrogate-model-based approaches

Generating so-called surrogate models or emulators is another powerful technique to reduce the computational cost associated with UQ. Essentially, the process of surrogate-based UQ consists of two steps. In the first step, a statistical model of the actual simulation model is generated based on a number of sample evaluations. In the second (offline) step, the surrogate model can then be used instead of the simulation model in combination with various sampling techniques to compute the desired quantities of interest. This is a feasible approach due to the fact that evaluating the surrogate model is computationally rather cheap. In fact, one computation often takes only a fraction of a second.

Bayesian surrogate models based on Gaussian processes, often referred to as emulators, have additional capabilities in the sense that they provide estimates of the expected local accuracy of the surrogate model. This information can be readily used to construct sequential experimental design schemes to refine the surrogate model locally where needed.

Multi-fidelity approaches

Particularly for complex nonlinear simulation models, a new approach has emerged from academic research, which improves the efficiency of both sampling- as well as surrogate-based UQ schemes by incorporating information from cheap-to-evaluate computational models with reduced fidelity into the UQ scheme. The computational effort associated with the evaluation of modern simulation models can often be drastically reduced when sacrificing accuracy. Reduced spatial or temporal resolution, loser solver tolerances, reduced complexity of the physical model are only a few aspects that can be adapted to reduce computational complexity. While this approach should obviously not be used in a deterministic setting, it is indeed feasible for UQ, since almost all UQ techniques are based on multiple evaluations of the simulation model. The premise of multi-fidelity UQ schemes is that a large portion - though not all - of these evaluations can be shifted from the high-fidelity model to a cheaper-to-evaluate low-fidelity model without significant loss of accuracy overall.

Depending on the nature of the problem and the desired quantity of interest, various multi-fidelity methods are suitable for substantially decreasing the computational costs associated with an UQ analysis. The methods differ in the way samples are chosen and how the information from the different levels of fidelity is essentially fused to produce an accurate probabilistic estimate of the quantity of interest. Compared to traditional UQ approaches, the computational costs are substantially reduced by employing a multi-fidelity approach, typically by orders of magnitude.

4 Conclusions

Efficient UQ techniques available today will change the way design decisions are made. By incorporating imperfect and incomplete information into the computational product design process, the uncertainty in the results can be quantified. The implications of this paradigm shift are manifold. First and foremost, it means that an even larger portion of the design process can be carried out by using computers, which leads to reduced cost and shorter product development cycles. In the end, the eventual product designs will be less prone to failure, while at the same time being less costly to manufacture. Particularly for complex nonlinear simulation models, multi-fidelity UQ methods represent a novel approach with which the computational costs associated with UQ can be drastically reduced, typically by orders of magnitude.

5 Our Solution for You

AdCo Engineering^{GW} bundles know-how from statistics, machine learning and engineering to provide innovative algorithms for the quantification of uncertainties in our software QUEENS. Since some of these very powerful methods only require information about input and output variables, these approaches can be applied to a variety of problems. Interfaces to commercial CAE software packages, such as ANSYS, are also available.