Chapter 2
Types of Uncertainty

Peter F. Pelz , **Marc E. Pfetsch** , **Sebastian Kersting, Michael Kohler, Alexander Matei, Tobias Melz, Roland Platz, Maximilian Schaeffner** , **and Stefan Ulbrich**

Abstract The goal of this chapter is to define different types of uncertainty in technical systems and to provide a unified terminology for this book. Indeed, uncertainty comes in different disguises. The first distinction is made with respect to the knowledge on the source of uncertainty: stochastic uncertainty, incertitude or ignorance. Then three main occurrences of uncertainty are discussed: data, model and structural uncertainty.

In this book we focus on physical and cyber-physical systems that are designed, manufactured and used. Hence, our context is that of engineering design, production and usage, in combination with applied mathematics providing methods and strategies as well as law providing a social and judicial framework. Uncertainty occurs in every step of system design, production and usage and needs to be anticipated in the design phase. Supporting the analysis, this chapter is concerned with different types of uncertainty and their quantification.

Indeed, before mastering uncertainty, uncertainty has to be identified. In order to do so, it is helpful to define individual uncertainty types. We classify uncertainty using two independent classifiers identifying its appearance and effect. The first classifier captures the effect of the uncertainty on the system at its core. It distinguishes between *stochastic uncertainty*, *incertitude* and *ignorance*. The resulting decision diagram is shown in Fig. 2.1. The second classifier distinguishes data, components and structures. Together they lead to the 3×3 matrix shown in Fig. 2.2.

P. F. Pelz (✉) · T. Melz · R. Platz · M. Schaeffner
Department of Mechanical Engineering, TU Darmstadt, Darmstadt, Germany
e-mail: peter.pelz@fst.tu-darmstadt.de

M. E. Pfetsch · S. Kersting · M. Kohler · A. Matei · S. Ulbrich
Department of Mathematics, TU Darmstadt, Darmstadt, Germany

T. Melz
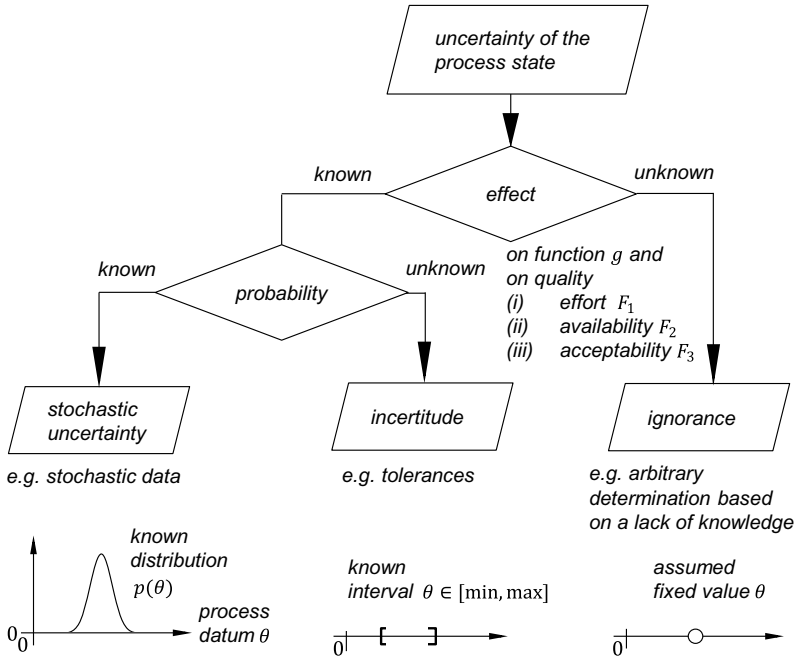Fraunhofer Institute for Structural Durability and System Reliability LBF, Darmstadt, Germany

Fig. 2.1 Classifier using effect and probability to separate into stochastic uncertainty, incertitude and ignorance [8, 18]

This matrix can be applied in all phases of the product life cycle, i.e. design, production and usage. This is even more important, since uncertainty quantification has become a thriving research field in engineering, computer science and mathematics over the last twenty years [20, 39].

Classification by effect and probability

Fig. 2.1 shows the first classifier as a decision diagram. The first decision is whether the *effect* of an uncertain process property on the process or the structure's function is known or unknown. This includes the decision of whether the effect on the system function and quality is known or unknown; recall that the system function is usually represented as a constraint $g(x) \leq 0$ and quality is measured using effort F_1, availability F_2 and acceptability F_3, see Sect. 1.6.

If the effect is unknown, we speak of *ignorance*. If the effect is known, then we speak of *probability*. The second question is whether the probability of the effect is known or unknown. If the probability of the effect is only partially quantified, we speak of *incertitude*. If the probability of the effect is sufficiently quantified, as shown schematically in Fig. 1.4, then we speak of *stochastic uncertainty* [8, 18].

The category *stochastic uncertainty* implies that a probability density function of the process state as sketched in Fig. 1.4 is known. In this case, it is possible to describe, quantify and evaluate uncertainty. The category *incertitude* is very common in mechanical engineering. It is used for many processes in mass production and, for example, in the *Austauschbau* described by Franz Relaux in the year 1899 [47], i.e. a part A fits to a part B even though different people manufacture the two parts on different machines. The incertitude of the *Austauschbau* is mastered by measurement data from systematically drawn samples or by experience manifested in tolerance classes. The two categories stochastic uncertainty and incertitude lead to a non-deterministic system design whereas *ignorance*, i.e. disregarded uncertainty, implies a deterministic system design.

Classification by data, component and structure

The first classifier distinguishes between the effect and quantification of uncertainty. The second classifier is motivated by system design [42]. This classifier can best be understood by keeping in mind a physical system, such as one of the three demonstrators presented in Sect. 3.6, i.e. the lightweight structure MAFDS, the Active Air Spring and the 3D Servo Press, or the hydrostatic transmission as depicted in Fig. 1.7. A process chain, a system and a structure consist of components or individual processes that fulfil the sub-functions of a system. In the following, when mentioning a model and model uncertainty, we may refer to the model as the individual process or component of the system. However, we may also refer to the composed system satisfying one or more specific system functions g_s. As pointed out in Chap. 1, different systems may satisfy the same function. Some possible systems may not be evaluated. We call this nescience *structural uncertainty*.

Applying the second classifier yields (a) data, (b) model and (c) structural uncertainty. These classes form the columns of Fig. 2.2. The first classifier shown in Fig. 2.1 leads to the rows (i) stochastic uncertainty, (ii) incertitude and (iii) ignorance of the 3×3 matrix shown.

From top to bottom, the confidence in data, models and structures is decreasing. This distinction into different types of uncertainty in data, i.e. *data uncertainty*, is then as follows. Data $\theta = \bar{\theta} + \delta\theta$ is subject to stochastic uncertainty if it can be modelled as realisations of a random variable with a distribution $P(\theta)$ or density $p(\theta)$ and expected value/mean $\bar{\theta}$. Incertitude appears if the data is only known to lie within a given fuzzy set or interval. If uncertainty of the data is not considered and thus ignored in the problem analysis, we speak of ignorance. For a more detailed discussion, see Sect. 2.1.

For *model uncertainty* the classification is as follows: a validated and verified model is subject to quantified stochastic uncertainty. There is incertitude, as long as the model is only suspected, i.e. assumed without experimental evidence. If the model is unknown, this is called ignorance, consistent with the scheme shown in Fig. 1.5. The presented classification is supplemented by the required model characteristics given in Sect. 1.3; a model has the following three qualities here: consistency, correctness and conciseness, see Sect. 1.3.
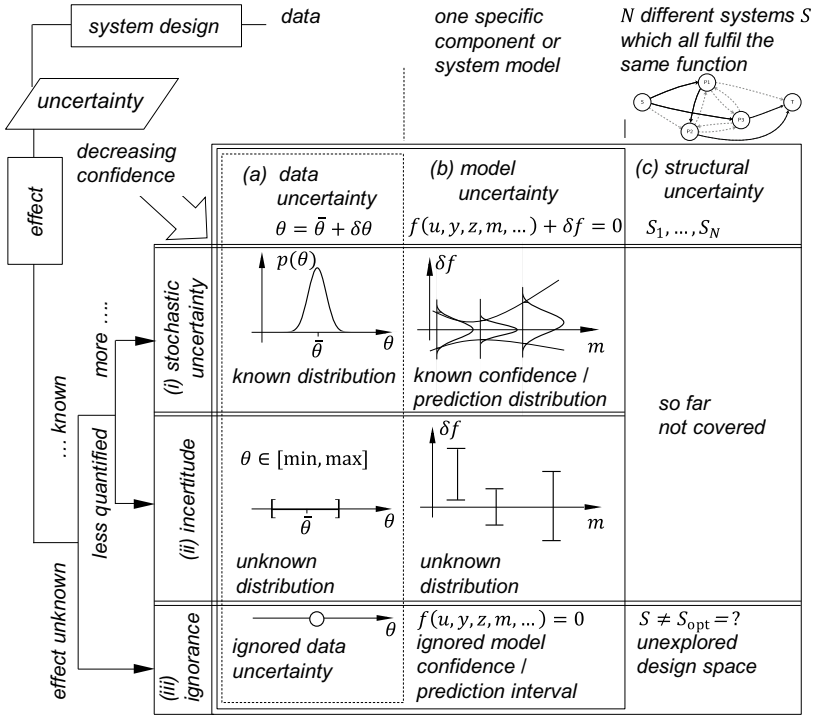
Fig. 2.2 Classification of uncertainty into (i) stochastic uncertainty, (ii) incertitude, (iii) ignorance; and into **a** data, **b** model, **c** structural uncertainty

To represent specific components or systems, a model in *implicit* form

$$f(u, y, z, m, \ldots) = 0, \tag{2.1}$$

is used, where f is the model function, u are inputs, like control or boundary values, y are internal variables, such as states, z is the model's output, i.e. the quantities of interest, and m are the model parameters which need to be calibrated. Sometimes u is split into binary design or other decision variables. In many cases, the above equation can be solved for given u and m such that y and z are uniquely determined. The model is then reduced to an *explicit* form. This often occurs if the model represents an input-output relation.

As mentioned in Chap. 1, in engineering and natural sciences, exact models do not exist. Consequently, to represent reality, one can use a model discrepancy function $\delta f(\ldots)$ that captures the difference between reality and the model given by f. This leads to the "real" model

$$f(u, y, z, m, \ldots) + \delta f(\ldots) = 0. \tag{2.2}$$

Hence, δf is in most cases different from zero and its analytical expression is in general unknown. If δf can be modelled as a random variable, we consider it as stochastic model uncertainty. In this case, the distribution of δf depends on the parameters m and on additional prior assumptions. If a non-probabilistic approach is adopted to model the discrepancy function, then the min-max values of δf have a functional relation to the parameters m. If δf cannot be quantified, non-exactness is assumed and ignorance prevails, as pointed out in Chap. 1. For a more detailed discussion, we refer to Sect. 2.2.

With respect to *structural uncertainty*, there are $N \gg 1$ competing structures S_i, $i = 1, \ldots, N$, all satisfying the same specific system function g_s within an accuracy interval $g_s - g = \pm \delta g$. But the structures may differ in quality F. If the complete design specification is not explored, i.e. $S \neq S_{opt}$ we speak of ignored structural uncertainty. Structural uncertainty is discussed in more detail in Sect. 2.3.

2.1 Data Uncertainty

Sebastian Kersting, Roland Platz, Michael Kohler, and Tobias Melz

In engineering sciences, generating and evaluating data for and from numerical simulations, experimental tests of technical systems with high safety requirements, a representative process or archived data play an important role to adequately predict and evaluate the system's performance, cf. Sect. 1.4. Data uncertainty is present, if the amount, type and distribution of required data, such as model parameters, is incomplete, unknown or insufficient; and in this context data quality as discussed in Sect. 1.4 is an important factor. This section clarifies the expression *data uncertainty* and classifies various approaches to describe different forms thereof. For the latter, a brief overview of probabilistic with further differentiation between frequentist and Bayesian inferences, and non-probabilistic, as well as parametric and non-parametric approaches to analyse data uncertainty is given. Following these approaches, the classification illustrated in Fig. 2.1 into *stochastic data uncertainty* and *incertitude* is discussed in this section.

2.1.1 Introduction

Within this book, we distinguish between two general types of data: model parameters m and state variables u and z that have a quantifiable value. Model parameters describe the technical system's characteristics, such as geometrical and material properties for a mathematical or computer model, e.g. length or width of a beam element and mass, Young's modulus, etc. Aggregating geometrical and material properties of one or multiple components that effect processes in a system or structure may lead to new model parameter expressions like stiffness or damping as important quantities in

the system's computer models for load-bearing systems. The models relate to them in a mathematical way and *model uncertainty* may become relevant, see Sect. 2.2. State variables describe the input and output conditions, such as mechanical loading, stress and strain, displacement, velocity, acceleration etc. They are mathematically related to the model parameters via the models.

On the one hand, model parameters m and state variables u are the input data for numerical simulations with computer models for predicting the system's behaviour, for example the dynamic behaviour of a structure due to vibrational excitation. They are also the input quantities for the experimental tests to validate the prediction. On the other hand, the system's behaviour as the output z of both simulation and test are mostly state variables or other measured quantities that can be inferred from model parameters by mathematical conversions. For example, the measurement of a lateral force and a resulting lateral deflection of a beam element leads to its stiffness.

Data in form of model parameters or state variables may have a single or a distributed value. Both are subject to uncertainty. Single or distributed values may vary in many possible ways depending on the designer's knowledge about the data and other conditions that make them uncertain. For example, a model parameter, such as the length of a beam element, may vary due to production tolerances, which influences the output in computer simulations or experimental tests.

In general, two basic types of data uncertainty occur: *aleatoric* or *epistemic* data uncertainty. *Aleatoric* uncertainty [67] is also known as irreducible uncertainty [62] or variability [67]. It is objective [32], and mostly characterised by a probabilistic distribution function. In [7], it is presumed to be the intrinsic randomness of a phenomenon. *Epistemic* uncertainty is also known as reducible uncertainty [50, 62], ignorance uncertainty [50], or simply uncertainty [67]. It is reducible [62], subjective [32], and occurs due to a lack of knowledge [62], insufficient or incomplete data [7]. Both types of uncertainty may be described via probabilistic and non-probabilistic approaches, cf. [33, 38, 44]. The probabilistic approaches can be further divided into *parametric*, with *frequentist* or *Bayesisan* inference approaches, or *non-parametric*.

Taking into account *aleatoric* and *epistemic* data uncertainty, this book distinguishes between *stochastic data uncertainty* and *incertitude*. They depend on the knowledge and assumptions about the data distribution and are explained in the following.

2.1.2 Stochastic Data Uncertainty

We assume that data is subject to stochastic uncertainty, if it can be modelled as realisations of a random variable Θ with a distribution $P(\theta)$. In this case, a *parametric* or a *non-parametric* approach as well as a *frequentist* or a *Bayesian* inference approach may be used to approximate the distribution for further uncertainty analysis. The approaches highly depend on the knowledge about the data, e.g. if a sample

of measured data exists and, if applicable, its sample size. All approaches to analyse stochastic data uncertainty are conducted under the assumption that an event is possible with a given probability [33].

In the *parametric* approach, one assumes that the underlying distribution depends on finite dimensional parameters that fully describe the distribution $P(\theta)$. In case of measuring errors or production tolerances, a typical choice would be a normal distribution. Here the distribution of a random variable Θ is uniquely identified by the parameters mean $\bar{\theta}$ and standard deviation $\sigma(\theta)$. If the distribution is incomplete or no samples are available, engineering or physical knowledge can be used to approximate the parameters. Otherwise, if an adequate sample is on hand, a maximum likelihood estimator can be applied to estimate the distribution parameters, cf. [21].

In the *non-parametric* approach, the description of distributions is based solely on observations, cf. [57, Chap. 4.8]. The underlying distribution does not need specific finite dimensional parameters. Instead, e.g. a kernel density estimator can be applied to estimate the corresponding probability density function $p(\theta)$ of $P(\theta)$, cf. [43, 49]. In most cases, if only a relatively small sample with less than 50 data points is available, the maximum-likelihood approach yields more adequate results with respect to high convergence than the kernel density estimator—if the distribution assumption is correct. However, the actual sample size needed to achieve sufficient results may vary and depends on the specific application.

If the distribution of input data is known or estimated as described above, several methods may be used for a probabilistic computer simulation for obtaining its probabilistic output prediction, e.g. Monte Carlo Simulation (MCS) methods as described in [5, 10, 15, 39, 41, 51, 54].

MCS methods depend on the selected inference approach, meaning that output data distributions may vary from a *frequentist* or *Bayesian* perspective. They differ in the underlying assumptions made regarding the nature of data distributions [57]. In the *frequentist* view, probabilities are defined as the frequency that an event occurs if an experiment is repeated a large number of times. The *Bayesian* perspective treats probabilities as a distribution of subjective values based on prior knowledge and assumptions. They are constructed or updated as data is observed; algorithms are, e.g. Markov Chain Monte Carlo Techniques, Metropolis-Hastings Algorithms etc., see [57]. *Bayesian* inference-based approaches are mainly used for model parameter calibration and to determine model uncertainty, see Sects. 2.2, 4.1 and 4.3. However, they are computationally demanding because of the need to infer the posterior distributions of each parameter [9].

2.1.3 Incertitude

We say that data is subject to incertitude, if it *can not* be modelled as realisations of Θ with a distribution $P(\theta)$. The distribution of the input data is unknown. Instead, the analysis may be conducted based on *fuzzy set theory* or direct interval analysis that provide information about the possibility that a certain data value lies between

a minimum and a maximum. These *possibilistic* approaches, in contrast to *probabilistic* approaches to assess stochastic data uncertainty, basically analyse whether an event is possible or impossible [33]. They are explained briefly in the following.

Fuzzy data uncertainty

The fuzzy set theory was introduced by Zadeh in 1965 [66]. Since then, numerous subdomains like fuzzy logic, fuzzy modelling, fuzzy arithmetic etc. have emerged [19]. Within the Fuzzy framework, data is expressed as member elements of a set $\mathcal{A}$. They can be defined by using a characteristic or, respectively, membership function $\mu_A \colon \mathcal{A} \to [0, 1]$. A membership of an element x is given if $\mu_A(x) = 1$, non-membership if $\mu_A(x) = 0$, [19]. The membership functions vary in form, they can be triangular, Gaussian, exponential etc. A practical way is to use so called α-cuts that divide the membership function into intervals with the aim to determine possibilities that a value is inside the interval, [33]. An overview of fuzzy methods for data uncertainty analysis in engineering applications is given in [48], applied fuzzy arithmetic can be found in [19] and fuzzy set theory based on fuzzy arithmetic is discussed in [66]. However, due to the absence of a structured and systematic elaboration of the theory, only a few practical approaches have been conducted [19].

Interval based data uncertainty

If neither distribution nor membership functions are available to describe the occurrence of data, an interval based approach may be useful. In this case, it is commonly assumed that the uncertain data lies between a minimum and a maximum value. Using a pessimistic perspective, a worst-case scenario is then the object of further investigations. Each parameter interval consists of a pair of min/max values, the four basic computation rules for adding, subtracting, multiplying, and dividing are valid for each interval parameter. However, the quality of the direct interval arithmetic evaluation depends on how often the interval parameters are present in a governing function of the deterministic computer model, i.e. the intervals become larger when they are propagated. Moreover, building a computer model with interval parameters that include paired min/max values can be demanding and time-consuming. Also, the analysis tends to overestimate uncertainty by using only extreme values that occur only rarely within the interval arithmetic [1].

Intervals can be stochastically motivated, e.g. $\bar{\theta} \pm 3\,\sigma(\theta)^2$-intervals specify minimum and maximum values, if a normal distribution of data is assumed. Eventually, possibilistic methods can be applied as shown in [1, 40].

In higher dimensions and to avoid using extreme values that lead to overestimation, the limiting intervals are typically replaced by ellipsoids for which the worst-case analysis becomes more complicated. For example and as shown in [11, 28–30, 55], sophisticated optimisation techniques introducing uncertainty sets are necessary to master data uncertainty in this pessimistic setting, see also Sect. 6.1.

2.2   Model Uncertainty

Alexander Matei, Roland Platz, Stefan Ulbrich, and Maximilian Schaeffner

In science and technology, mathematical models are frequently employed for the explanation of natural phenomena and for the description, quantification and control of engineering processes. We specifically focus on *mathematical* models and their accuracy, for other types of models we refer to Chap. 1. We do so because mathematical models enable numerical simulations to *predict* the behaviour, outcome or result of real technical products, systems and processes along their life cycle, see Sect. 1.2. However, it is a common observation that the usage of these models is affected by uncertainty, which can be traced back to the system design phase. Several causes can be identified for this uncertainty in early stage product development and, particularly, in the mathematical modelling. Our ignorance about the physical behaviour of a technical system leads to models that are only approximations of reality and may only be valid for a particular range of inputs and parameters, cf. Fig. 1.5.

There are two categories of ignorance which we briefly want to mention. The first is called *lack of knowledge* and stands for objects or processes which are unknown, unfamiliar and nameless to us. Examples for lack of knowledge arise whenever a novel material is exposed to new circumstances. Its reaction to the environment, its behaviour under load or pressure and its wear, all of which needs to be observed, evaluated and generalised. This is a challenge to scientists and engineers alike. The second category comprises effects that are known to us, but they are neglected, ignored and kept out of consideration in the modelling. We call it *disregard of knowledge*. To give an example, one may think of a linear elastic spring under load where the deformation of the spring is proportional to the loading force, only if the latter stays below a certain threshold. For loads above this limit, the spring material shows nonlinear, plastic or hysteresis-type behaviour which is difficult to model and thus it is often neglected.

Another source of model uncertainty arises from the numerical approach used to discretise such equations that are impossible to solve analytically. In engineering applications, the finite element method is a common numerical approximation scheme. In most cases, uncertainty caused by the numerical discretisation with finite elements can not only be quantified but also mastered by standard approaches, e.g. by developing error estimates [46]. In most cases, numerical errors happen on a relatively small scale, whereas the more severe sources of model uncertainty are missing or incomplete physical or empirical relations. In addition, human factors also contribute to model uncertainty. The methods and technologies to detect, quantify and master model uncertainty, which are presented in this book, see Sect. 4.3, are generally applicable, irrespective of the above mentioned causes.

Model uncertainty exists, if the functional relations between input and output, model parameters and other internal variables, as well as the scope and complexity of the model, are unknown, incomplete, inadequate or unreasonable. The dilemma the designer encounters is that in early stage design, before calibration, verification and validation processes start, the extent of uncertainty is difficult to detect. Even in

the usage phase, after the final product has been assembled, a degree of uncertainty remains, since the full-scale product does not exactly match the small-scale prototype, neither does the computer model. Various mathematical approaches using different prior knowledge about functional relations, scope and complexity as well as the required data, which can be uncertain as well, have been developed to deal with model uncertainty.

It is one aim of this book to understand and to evaluate the uncertainty especially in load-bearing mechanical systems with high safety requirements. This section gives a brief overview of different approaches to describe, quantify and master model uncertainty. In this context, the term 'master' means to be aware and to be able to quantify model uncertainty in verification and validation processes, e.g. by data-based training. This leads to an adapted mathematical model that, eventually, adequately describes and predicts the system and process behaviour observed in reality.

2.2.1   Functional Relations, Scope and Complexity of Mathematical Models

In this chapter, we consider *models* as images of reality in the domain of mathematical abstraction. These mappings describe or represent knowledge about a system in the language of functional relations given in implicit form

$$f(u, y, z, m) = 0 \tag{2.3}$$

between input u and output data z, model parameters m, such as material properties, and internal variables y, like states. Accordingly, we mean by *model uncertainty* that these images of reality are imperfect, i.e. the governing *physical* relations f between inputs, outputs, parameters and internal variables are unknown or incomplete, or they are partly reduced to *physically-inconsistent* approximations (Sect. 1.3) of more complex, but expensive, expressions. Thus, in the presence of model uncertainty, Eq. (2.3) does not reflect reality. A common representation of model uncertainty introduces a *model discrepancy function* δf which accounts for lack or disregard of knowledge, numerical errors and human factors in the mathematical modelling as mentioned before. The "real" functional relation is then considered to be implicitly given by

$$f(u, y, z, m) + \delta f(\ldots) = 0. \tag{2.4}$$

In many cases, for given u and m, the implicit equation (2.3) is solved to obtain the model's output z explicitly. If the model function f is differentiable and its derivative is invertible with respect to the internal variables y, then the implicit function theorem yields a *reduced model* η that represents an input-output relation to the quantity of interest z. The model equation can now be written in explicit form

$$\eta(u, m) = z, \tag{2.5}$$

and likewise for the case where the model discrepancy function is added, cf. Eq. (2.4). In the sequel, we follow the main literature on this topic that considers only computer models η, i.e. reduced models, which describe an input-output relation.

Evidently, different assumptions about functional relations like axiomatic or empirical, linear or non-linear, and time-invariant or time-variant alter the output of the model significantly [37]. Moreover, the model's scope and complexity can be varied. Thus, any computer simulation that is based upon uncertain models leads to erroneous predictions of the quantity of interest that affect the verification and validation process necessary to prove the model's consistency and correctness, cf. Sect. 1.3.

The designer has several options to choose from possible modelling assumptions about functional relations, as well as scope and complexity, as mentioned above. When a mathematical model has been built, the functional relations are subject to a verification and calibration process to prove and to update the numerical simulation, the computer code and, if applicable, the model parameters [2]. Eventually, the computer model is validated against experimental tests of a real system, like the Modular Active Spring-Damper System, see Sect. 3.6.1. As for the scope and complexity, the number of degrees of freedom can be high and costly depending on the form and discretisation of the model, e.g. analytical, finite elements or multi-body-models that also influence the results of verification, calibration and validation processes. Furthermore, Occam's razor can be used as a guiding principle to keep models as simple as possible, see Sect. 1.3, because more complex models often tend to be more susceptible to uncertainty.

The functional relations as well as the scope and complexity of models are considered as being independent from how model parameters and input data are present—as a simple value, randomly distributed or as intervals. These are subject to data uncertainty, which is covered in Sect. 2.1.

2.2.2 Approaches to Detect, Quantify, and Master Model Uncertainty

Basically, two different approaches give information about detectable, quantifiable and masterable model uncertainty: a *deterministic* analysis and a *probabilistic frequentist* or *probabilistic Bayesian inference-based* perspective. The latter needs subjective prior data distribution information as discussed in Sect. 2.1. The deterministic and Bayesian inference-based approach allow data-based training, which is an important criterion for the verification and validation processes. However, finding the adequate correction terms or prior information, especially for models with high complexity, remains a challenge [9]. Using a probabilistic frequentist approach, see Sect. 2.1, however, does not take into account prior information other than randomness of data.

In accordance to this book's notation, deterministic approaches do not master stochastic uncertainty in a mathematical model. However, they may be suited to quantify and master incertitude, e.g. when only extreme values like minima and maxima are known or assumed, or when parts of a model are missing to detect ignorance, cf. Fig. 2.2. Probabilistic approaches, however, describe and master all three types of uncertainty. In the following, we give an overview of the two approaches.

Deterministic framework

On the one hand, a *deterministic* analysis can be used, for example, in fault diagnosis [56] to quantify uncertainty in the model equations. To do so, the model is adjusted by a model discrepancy function $\delta\eta$ which is assumed to stay within a bounded uncertainty set $\mathcal{U}_\eta$:

$$\eta(u, m) + \delta\eta(u) = z, \quad \delta\eta \in \mathcal{U}_\eta. \tag{2.6}$$

A standard residual analysis then enables the engineer to distinguish component failures from the effects of model uncertainty. However, this method strongly relies on the assumption that the model discrepancy function $\delta\eta$ stays within a bounded uncertainty set $\mathcal{U}_\eta$.

On the other hand, a deterministic analysis may be used in model verification and validation processes. In [9, 57], the possibility to approximate the residual between the model output and the observed quantity of interest via a polynomial p is mentioned. In this case, the polynomial p takes the role of the discrepancy function:

$$\eta(u, m) + p(u, m, \theta) = z. \tag{2.7}$$

This necessarily leads to an augmented parameter set (m, θ) consisting of the original physical axiomatic and empiric parameters m as well as the non-physical polynomial parameters θ, which do not give an enhanced physics-based understanding of the model's uncertainty or shortcoming in predicting reality [9]. The augmented parameter set needs to be calibrated, which is usually performed by an optimisation scheme.

Probabilistic framework

Within the probabilistic framework, *Bayesian inference-based* approaches are frequently used to assess the prediction quality of a mathematical model under given experimental data [9, 16, 52, 61]. In [34] or [37], Bayesian calibration techniques and a plausibility prior argument are used for model selection. Another important approach introduces a stochastic process δ for the discrepancy function [22]:

$$\eta(u, m) + \delta(u, m, \phi) + \varepsilon = z \,, \tag{2.8}$$

where ε is stochastic noise, usually produced by the measurement error, and ϕ are hyperparameters, which need to be tuned by an optimisation scheme. Based on this setting, [64] introduce a technique for model validation. A method designed for time-dependent systems is proposed in [3]. Combinations of a high- and a low-fidelity model and experimental data are used to construct a multi-fidelity model, which is proposed in [13] as an improved predictor. A scaled Gaussian Process is used in [17] for model calibration and prediction. It is claimed that the method bridges the gap between the least-squares calibration and the Gaussian Process calibration and, thus, is an improvement of the method introduced in [22]. Furthermore, a Bayesian interval hypotheses-based approach [36] and a Bayesian inference-based approach [37] are used to compare different models based on their internal functional relations from axiomatic or empiric assumptions, see also Sect. 4.3.3 for assessing model uncertainty in the Modular Active Spring-Damper System, see Sect. 3.6.1.

In order to apply the Bayesian methods, it is necessary to select prior distributions. Often, this is subjective and it is unclear how to choose them, which may lead to unrealistic assumptions. Furthermore, it is pointed out in [60] that the approach in [22] may lead to inadequate results in case of an imperfect computer model. As a consequence, further verification of the computer codes needs to be conducted.

From the *frequentist's* perspective, a simple approach to quantify model uncertainty is to use validation metrics, such as the area validation metric [35] or the Mahalanobis distance [68]. They give a quantitative measure of disagreement between the model output and the observed quantity of interest based upon observed measurements. In order to select the most adequate model, an arbitrary threshold on the metric is imposed, or classical hypothesis testing is performed. Surrogate models are often used to quantify the uncertainty in a technical system. Usually, these methods use computer simulations and a small sample of experimental data to estimate properties of probability distributions, such as quantiles [25–27] or densities [14, 23, 24]. A detailed description for a method based on an imperfect computer model is given in Sect. 4.3.8. A case where computer models are assumed to fit the reality is shown in Sect. 5.2.6. In [12], another method to detect model uncertainty is proposed which is based on optimum experimental design and hypothesis testing, see also Sect. 4.3.1. In [65], another approach to quantify the model error is developed. Here, the model discrepancy function is estimated on bootstrap samples via a regression estimation, e.g. smoothing splines or artificial neural networks.

However, the application of frequentist methods often demands a larger sample size, which can become a problem, since generating data is an expensive and time-consuming process.

2.3   Structural Uncertainty

Peter F. Pelz and Marc E. Pfetsch

Besides data and model uncertainty introduced in Sects. 2.1 and 2.2, structural uncertainty forms the third important pillar of uncertainty under consideration in this book. Structural uncertainty refers to the fact that only part of the possible solutions are evaluated with respect to uncertainty. In this sense, the model of the system is incomplete. However, the focus is on the system and not on the ignorance of the models of the system components, as in model uncertainty.

Let us start with the viewpoint of classical product design. Given a requested system function, the designer plans the system structure. The systems are assembled by components and modules, which then form the system. In production, this arises when combining single processes to process chains. When a given system function can be generated by a multitude of different function structures, and each function structure can be generated by a multitude of elements, this results in a "combinatorial explosion" of possibilities that, in general, cannot be evaluated by humans anymore. This lack of knowledge on other possibilities is then called *structural uncertainty*.

The consideration of this type of uncertainty seems to be new, but the term "structural uncertainty" is sometimes used in the sense of "model structure uncertainty", i.e. the structure of the model is uncertain. Some examples from different disciplines can be found in [4, 6, 58, 59]. In our book, the latter meaning is captured by the term "model uncertainty", see Sect. 2.2.

In comparison to data and model uncertainty, structural uncertainty has the advantage that its presence has no direct negative effect on product safety. However, economically better solutions might be lost.

Structural uncertainty can be tackled by using (discrete) mathematical optimisation methods that allow to consider all possibilities and select the best system choice with respect to a predefined objective function. These techniques require a combination of domain knowledge in order to set up a physical model and define the allowed elements. This is then integrated into a mathematical optimisation model, which is solved using optimisation software. As usual, one needs to balance the exactness of the model with the effort to obtain optimal solutions; see Sect. 1.3 for a general discussion of this balance. Often tailored solution methods need to be developed in order to achieve practically feasible solution times. The method for the quantification of structural uncertainty is therefore highly context-dependent.

This book contains examples that illustrate this approach, see, for example, Sects. 1.5 and 6.3.5. Many more examples can be found in the literature, e.g. [31, 45, 53, 63]. These examples show the flexibility of mathematical programming to deal with different manifestations of structural uncertainty. Nevertheless, currently, expert knowledge is needed to derive and efficiently solve appropriate models of reality.

Structural uncertainty has two other aspects that we want to briefly mention.

1. When using models in order to compose a system, the model predetermines the possible components that can be chosen, i.e. the model can only be optimised over its model horizon, see Sect. 1.3. Again, possible interesting solutions might be lost due to the choice of the model. One needs to be aware of this restriction, similar to the fact that models are always an approximation of reality.

2. The first aspect arises from the fact that the system is built from smaller elements. However, a quantitative evaluation of uncertainty usually only takes place on the level of the single elements, since an analysis for the complete system would be too complex or taking measurements would be too expensive. This book discusses several methods to deal with this uncertainty. For instance, a common way to handle the corresponding uncertainty is by flexibility, see Sect. 6.2. Another method is to make the system robust or to consider the effect of uncertain parameters already in the mathematical model and perform a robust optimisation, see Sect. 6.1.

References

1. Alefeld G, Mayer G (2000) Interval analysis: Theory and applications. J Comput Appl Math 121(1–2):421–464
2. American Society of Mechanical Engineers (ASME) Standards Committee on Verification and Validation in Computational Solid Mechanics (PTC 60/V&V 10) (2007) Guide for verification and validation in computational solid mechanics. ASME publications
3. Bayarri MJ, Berger JO, Cafeo J, Garcia-Donato G, Liu F, Palomo J, Parthasarathy RJ, Paulo R, Sacks J, Walsh D: Computer model validation with functional output. Annals of Statistics 35(5), 1874–1906 (2007). https://doi.org/10.1214/009053607000000163
4. Bojke L, Claxton K, Sculpher M, Palmer S (2009) Characterizing structural uncertainty in decision analytic models: A review and application of methods. Value in Health 12(5):739–749. https://doi.org/10.1111/j.1524-4733.2008.00502.x
5. Braaten E, Weller G: An improved low-discrepancy sequence for multidimensional quasi-Monte Carlo integration. Journal of Computational Physics 33(2), 249–258 (1979)
6. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD (2012) Model parameter estimation and uncertainty: A report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. Value in Health 15(6):835–842. https://doi.org/10.1016/j.jval.2012.04.014
7. Der Kiureghian A, Ditlevsen O: Aleatory or epistemic? Does it matter? Structural Safety 31(2), 105–112 (2009)
8. Engelhardt RA, Koenen JF, Enss GC, Sichau A, Platz R, Kloberdanz H, Birkhofer H (2010) A model to categorise uncertainty in load-carrying systems. In: 1st MMEP international conference on modelling and management engineering processes, pp 53–64
9. Farajpour I, Atamturktur S: Error and uncertainty analysis of inexact and imprecise computer models. Journal of Computing in Civil Engineering 27(4), 407–418 (2012)
10. Felber T, Kohler M, Krzyżak A (2015) Adaptive density estimation from data with small measurement errors. IEEE Trans. Inform. Theory 61(6):3446–3456. https://doi.org/10.1109/TIT.2015.2421297
11. Gally T, Gehb CM, Kolvenbach P, Kuttich A, Pfetsch ME, Ulbrich S (2015) Robust truss topology design with beam elements via mixed integer nonlinear semidefinite programming. In: Pelz PF, Groche P (eds) Uncertainty in mechanical engineering II, vol 807. Applied mechanics and materials. Trans Tech Publications, pp 229–238

12. Gally T, Groche P, Hoppe F, Kuttich A, Matei A, Pfetsch ME, Rakowitsch M, Ulbrich S (2021) Identification of model uncertainty via optimal design of experiments applied to a mechanical press. Optim Eng, to appear
13. Goh J, Bingham D, Holloway JP, Grosskopf MJ, Kuranz CC, Rutter E (2013) Prediction and computer model calibration using outputs from multifidelity simulators. Technometrics 55(4):501–512. https://doi.org/10.1080/00401706.2013.838910
14. Götz B, Kersting S, Kohler M (2020) Estimation of an improved surrogate model in uncertainty quantification by neural networks. Ann Inst Stat Math. To appear
15. Graham C, Talay D (2013) Stochastic Simulation and Monte Carlo Methods: Mathematical Foundations of Stochastic Simulation, vol 68. Springer
16. Green PL: Bayesian system identification of a nonlinear dynamical system using a novel variant of simulated annealing. Mechanical Systems and Signal Processing 52, 133–146 (2015)
17. Gu M, Wang L (2018) Scaled Gaussian stochastic process for computer model calibration and prediction. SIAM/ASA Journal on Uncertainty Quantification 6(4):1555–1583. https://doi.org/10.1137/17M1159890
18. Hanselka H, Platz R (2010) Ansätze und Maßnahmen zur Beherrschung von Unsicherheit in lasttragenden Systemen des Maschinenbaus. Konstruktion (11/12):55–62
19. Hanss M (2005) Applied Fuzzy Arithmetic: An Introduction With Engineering Applications. Springer
20. Imholz M, Vandepitte D, Moens D (2015) Analysis of the effect of uncertainty clamping stiffness on the dynamical behaviour of structures using interval field methods. In: Pelz PF, Groche P (eds) Uncertainty in mechanical engineering II, vol 807. Applied mechanics and materials. Trans Tech Publications, pp 195–204
21. Kalbfleisch JG (1979) Probability and Statistical Inference. II. Springer, New York
22. Kennedy MC, O'Hagan A (2001) Bayesian calibration of computer models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63(3):425–464. https://doi.org/10.1111/1467-9868.00294
23. Kersting S, Kohler M (2019) Uncertainty quantification based on (imperfect) simulation models with estimated input distributions. Submitted for publication
24. Kohler M, Krzyżak A (2017) Improving a surrogate model in uncertainty quantification by real data. Submitted for publication
25. Kohler M, Krzyżak A (2018) Estimating quantiles in imperfect simulation models using conditional density estimation. Ann Inst Stat Math. https://doi.org/10.1007/s10463-018-0683-8
26. Kohler M, Krzyżak A (2019) Estimation of extreme quantiles in a simulation model. Journal of Nonparametric Statistics 31(2):393–419. https://doi.org/10.1080/10485252.2019.1567727
27. Kohler M, Tent R (2019) Nonparametric quantile estimation using surrogatemodels and importance sampling. Metrika 83:141-169.https://doi.org/10.1007/s00184-019-00736-3
28. Kolvenbach P, Lass O, Ulbrich S (2018) An approach for robust PDE-constrained optimization with application to shape optimization of electrical engines and of dynamic elastic structures under uncertainty. Optim Eng 697–731. https://doi.org/10.1007/s11081-018-9388-3
29. Kuttich A, Ulbrich S (2017) Feedback controller design and topology optimization for truss structures under uncertain dynamic loads. In: von Scheven M, Keip MA, Karajan N (eds) 7th GACM colloquium on computational mechanics for young scientists from academia and industry
30. Lass O, Ulbrich S (2017) Model order reduction techniques with a posteriori error control for nonlinear robust optimization governed by partial differential equations. SIAM J Sci Comput 39(5):S112–S139. https://doi.org/10.1137/16M108269X
31. Leise P, Altherr LC (2018) Optimizing the design and control of decentralized water supply systems – a case-study of a hotel building. In: International conference on engineering optimization. Springer, pp 1241–1252
32. Lemaire M (2014) Mechanics and uncertainty. Wiley Online Library
33. Li S, Platz R (2017) Observations by evaluating the uncertainty of stress distribution in truss structures based on probabilistic and possibilistic methods. Journal of Verification, Validation and Uncertainty Quantification 2(3):031006. https://doi.org/10.1115/1.4038486

34. Lima E, Oden JT, Wohlmuth B, Shahmoradi A, Hormuth II DA, Yankeelov TE, Scarabosio L, Horger T: Selection and validation of predictive models of radiation effects on tumor growth based on noninvasive imaging data. Computer Methods in Applied Mechanics and Engineering 327, 277–305 (2017)
35. Liu Y, Chen W, Arendt P, Huang HZ: Toward a better understanding of model validation metrics. Journal of Mechanical Design 133(7), 071005 (2011)
36. Mallapur S, Platz R (2018) Quantification of uncertainty in the mathematical modelling of a multivariable suspension strut using Bayesian interval hypothesis-based approach. In: Pelz PF, Groche P (eds) Uncertainty in mechanical engineering III, vol 885. Applied mechanics and materials. Trans Tech Publications, pp 3–17
37. Mallapur, S., Platz, R.: Uncertainty quantification in the mathematical modelling of a suspension strut using Bayesian inference. Mechanical Systems and Signal Processing 118, 158–170 (2019)
38. Melzer C, Platz R, Melz T (2015) Comparison of methodical approaches to describe and evaluate uncertainty in the load-bearing capacity of a truss structure. In: Fourth international conference on soft computing technology in civil, structural and environmental engineering. Paper 26
39. Melzer CM, Krech M, Kristl L, Freund T, Kuttich A, Zocholl M, Groche P, Kohler M, Platz R (2015) Methodical approaches to describe and evaluate uncertainty in the transmission behavior of a sensory rod. In: Pelz PF, Groche P (eds) Uncertainty in mechanical engineering II, vol 807. Applied mechanics and materials. Trans Tech Publications, pp 205–217. https://doi.org/10.4028/www.scientific.net/AMM.807.205
40. Moore, R., Lodwick, W.: Interval analysis and fuzzy set theory. Fuzzy Sets and Systems 135(1), 5–9 (2003)
41. Müller-Gronbach T, Novak E, Ritter K (2012) Monte Carlo-Algorithmen. Springer
42. Pahl G, Beitz W, Feldhusen J, Grote KH (2007) Engineering design: a systematic approachD, 3rd edn. Springer, London. https://doi.org/10.1007/978-1-84628-319-2
43. Parzen E (1962) On estimation of a probability density function and mode. Ann Math Stat 33:1065–1076. https://doi.org/10.1214/aoms/1177704472
44. Platz R, Ondoua S, Habermehl K, Bedarff T, Hauer T, Schmitt S, Hanselka H (2010) Approach to validate the influences of uncertainties in manufacturing on using load-carrying structures. In: Proceedings of USD2010 international conference on uncertainty in structural dynamics, pp 20–22
45. Pöttgen P, Ederer T, Altherr L, Lorenz U, Pelz PF (2016) Examination and optimization of a heating circuit for energy-efficient buildings. Energ Technol 4(1):136–144
46. Repin SI, Sauter SA (2020) Accuracy of mathematical models. EMS tracts in mathematics, vol 33. European Mathematical Society Publishing House. https://doi.org/10.4171/206
47. Reuleaux F (1899) Der Konstrukteur: Ein Handbuch zum Gebrauch beim Maschinen-Entwerfen. Vieweg
48. Reuter U, Möller B (2007) Uncertainty Forecasting in Engineering. Springer
49. Rosenblatt M (1956) Remarks on some nonparametric estimates of a density function. Ann Math Stat 27:832–837. https://doi.org/10.1214/aoms/1177728190
50. Roy CJ, Oberkampf WL (2011) A comprehensive framework for verification, validation, and uncertainty quantification in scientific computing. Comput. Methods Appl. Mech. Engrg. 200(25–28):2131–2144. https://doi.org/10.1016/j.cma.2011.03.016
51. Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, Saisana M, Tarantola S (2008) Global Sensitivity Analysis: The Primer. John Wiley & Sons
52. Sankararaman S, Mahadevan S (2011) Model validation under epistemic uncertainty. Reliability Engineering & System Safety 96(9):1232–1241
53. Schänzle C, Altherr LC, Ederer T, Lorenz U, Pelz PF (2015) As good as it can be – ventilation system design by a combined scaling and discrete optimization method. In: Proceedings of the FAN
54. Schueller G, Pradlwarter H: Uncertainty analysis of complex structural systems. International Journal for Numerical Methods in Engineering 80(6–7), 881–913 (2009)

55. Sichau A, Ulbrich S (2012) A second order approximation technique for robust shape optimization. In: Hanselka H, Groche P, Platz R (eds) Uncertainty in mechanical engineering, vol 104. Applied mechanics and materials. Trans Tech Publications, pp 13–22. https://doi.org/10.4028/www.scientific.net/AMM.104.13

56. Simani S, Fantuzzi C, Patton RJ (2003) Model-based fault diagnosis techniques. In: Model-based fault diagnosis in dynamic systems using identification techniques. Springer, pp 19–60

57. Smith RC (2014) Uncertainty Quantification; Theory, Implementation, and Applications, Computational Science & Engineering, vol 12. SIAM, Philadelphia, PA

58. Strong M, Oakley JE (2014) When is a model good enough? Deriving the expected value of model improvement via specifying internal model discrepancies. SIAM/ASA Journal on Uncertainty Quantification 2(1):106–125. https://doi.org/10.1137/120889563

59. Thore P, Shtuka A, Lecour M, Ait-Ettajer T, Cognot R (2002) Structural uncertainties: Determination, management, and applications. Geophysics 67(3):840–852. https://doi.org/10.1190/1.1484528

60. Tuo R, Wu CFJ (2015) Efficient calibration for imperfect computer models. Ann Stat 43(6):2331–2352. https://doi.org/10.1214/15-AOS1314

61. Tuomi M, Pinfield D, Jones HRA (2011) Application of Bayesian model inadequacy criterion for multiple data sets to radial velocity models of exoplanet systems. Astronomy & Astrophysics 532:A116

62. Vandepitte D, Moens D (2011) Quantification of uncertain and variable model parameters in non-deterministic analysis. In: IUTAM symposium on the vibration analysis of structures with uncertainties. Springer, pp 15–28

63. Vergé A, Pöttgen P, Altherr LC, Ederer T, Pelz PF (2016) Lebensdauer als Optimierungsziel – Algorithmische Struktursynthese am Beispiel eines hydrostatischen Getriebes. O+P – Ölhydraulik und Pneumatik 60(1–2)

64. Wang S, Chen W, Tsui KL (2009) Bayesian validation of computer models. Technometrics 51(4):439–451. https://doi.org/10.1198/TECH.2009.07011

65. Wong RKW, Storlie CB, Lee TCM (2017) A frequentist approach to computer model calibration. J R Stat Soc Ser B Stat Methodol 79(2):635–648. https://doi.org/10.1111/rssb.12182

66. Zadeh LA (1965) Fuzzy sets. Inf Control 8(3):338–353

67. Zang TA, Hemsch MJ, Hilburger MW, Kenny SP, Luckring JM, Maghami P, Padula SL, Stroud WJ (2002) Needs and opportunities for uncertainty-based multidisciplinary design methods for aerospace vehicles. Technical report, TM-2002-211462, NASA

68. Zhao L, Lu Z, Yun W, Wang W (2017) Validation metric based on Mahalanobis distance for models with multiple correlated responses. Reliability Engineering & System Safety 159:80–89