1

Introducing the common methodological framework

COPYRIGHTED MATERIAL

1.1 Quantitative uncertainty assessment in industrial practice: a wide variety of contexts

Quantitative uncertainty assessment in industrial practice typically involves, as is shown in Figure 1.1 below:

- a pre-existing physical or industrial system or component lying at the heart of the study, represented by a pre-existing model;
- a variety of sources of uncertainty affecting this system;
- industrial stakes and decision-making circumstances motivating the uncertainty assessment. More or less explicitly, these may include: safety and security, environmental control, process improvement, financial and economic optimization, etc. They are generally the rationale for the pre-existing model, the output and input of which help to deal with the various stakes in the decision-making process in a quantitative manner.

As will be illustrated later, mainly in Part II, these three basic features cover a very wide variety of study contexts:

The **pre-existing system** may encompass a great variety of situations, such as: a metrological chain, a mechanical structure, a maintenance process, an industrial or domestic site threatened by a natural risk, etc. In quantitative studies of uncertainties, that system will generally be modelled by a single numerical model or

Uncertainty in Industrial Practice Edited by E. de Rocquigny, N. Devictor and S. Tarantola,
© 2008 John Wiley & Sons, Ltd

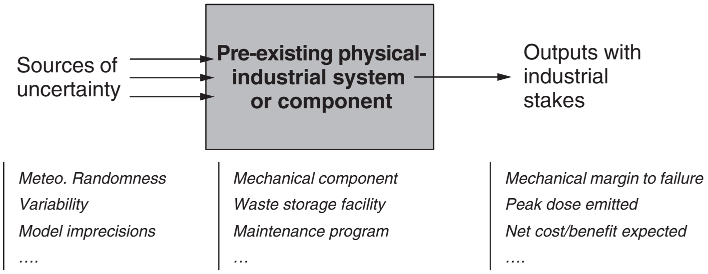

Figure 1.1 Schematic context of quantitative uncertainty assessment in industrial practice.

chain of models. The complexity of the models may vary greatly: from straightforward analytical formulae to physical models based on unsteady partial differential equations, coupled 3-D finite element models, intrinsically probabilistic models, e.g. a Boolean system reliability model predicting probabilities, etc.

The **sources of uncertainty** may include a great variety of uncertain variables of a quantitative nature, including the classical categories of: ‘aleatory phenomena’, ‘lack of data or knowledge’ or ‘epistemic uncertainties’, ‘variability’, ‘measurement errors’, etc. They may affect the pre-existing model in various ways: through uncertain values for model inputs, model errors or even uncertain (or incomplete) structures of the model itself, etc.

The **decision-making process** depends considerably on the **industrial stakes** involved: the study may be designed to answer regulatory requirements attached to licensing or certification of a new process or product; to ensure quality control or more robust designs; to internally optimize a technical-economic performance indicator; to feed into a larger decision model, such as a system reliability model with component uncertainties; or to help understand, in preliminary R&D stages, the importance of the various parts or parameters of the model.

1.2 Key generic features, notation and concepts

Notwithstanding the wide variety of contexts, the case studies of Part II will show that the following key generic features can be derived.

1.2.1 Pre-existing model, variables of interest and uncertain/fixed inputs

Whatever the nature or complexity of the **pre-existing model**, ‘in so far as uncertainties are concerned’ it may, *a priori*, be viewed conceptually as a **numerical**

function linking inputs (uncertain or fixed variables) to outputs (upon which decision criteria are established).

Formally, it is sufficient for the model to link the important **output variables of interest** (denoted $\underline{z}$) to a number of continuous or discrete inputs through a deterministic function $\underline{z} = G(\underline{x}, \underline{d})$, where some **inputs** (denoted $\underline{x}$) are **uncertain** – subject to randomness, lack of knowledge, errors or any other sources of uncertainty – while **other inputs** (denoted $\underline{d}$) are **fixed** – considered to be known – as represented in Equation 1.1 and Figure 1.2 below.

$$\underline{x},\ \underline{d} \Rightarrow \underline{z} = G(\underline{x},\ \underline{d}) \tag{1.1}$$

Note that $G(.)$ may represent any sort of function, including any deterministic physical or economic model (analytical or coupled 3-D finite element), or even an *intrinsically probabilistic* (or stochastic) model viewed from the perspective of a deterministic relation between some pre-existing input and output variables (e.g. failure rates at component or system level in risk analysis, or transition reaction probabilities and fluence expectation in Monte Carlo neutron physics). The computation of the pre-existing model for a given – not uncertain – point value $(\underline{x}, \underline{d})$ may hence require a very variable CPU time: from 10^{-4} s to several days for a single run, depending on the complexity of the simulation code.

Note also that the model output **variables of interest** (v.i.) are all included formally within the vector $\underline{z} = (z^l)_{l=1\ldots r}$. Most of the time $\underline{z}$ is a *scalar* or a *small-size vector* (e.g. $r = 1$ to 5), since the decision-making process involves essentially one or few variables of interest, such as: a physical margin to failure, a net cost, a cumulated environmental dose, a failure rate in risk analysis, etc. But in some cases $\underline{z}$ may be of large dimension (e.g. predicted oil volumes at several potential well sites) or even a *function* (e.g. the mechanical margin as a function of the number of fatigue cycles, the net cost of oil production as a function of the time, etc). Vector notation will be maintained, as it is appropriate in most situations.

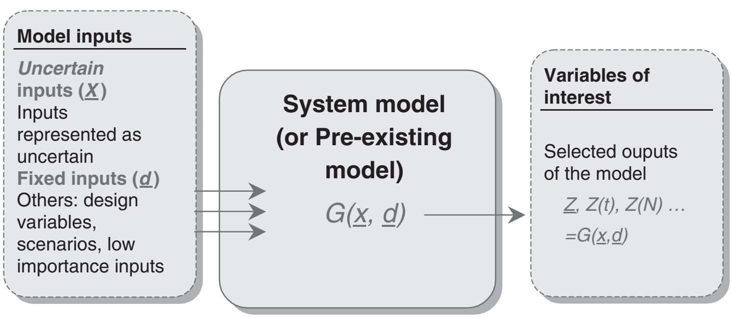

Figure 1.2 The pre-existing model and its inputs/outputs.

Regarding the **uncertain model inputs**, the vector $\underline{x} = (x^i)_{i=1\ldots p}$ could gather formally all sources of uncertainty, whatever their nature or type (parametric, model uncertainties, etc) while in other interpretations or settings it would be restricted to some of them. The dimension of $\underline{x}$ may be very large (for example, in the case studies in Part II, p may range from 3 to several hundreds). Some components of $\underline{x}$ may be continuous, while others could be discrete or branching variables (e.g. a variable indicating the most likely model of a portfolio of uncertain models in competition with each other). It could even formally include situations where there is a spatial *field* of uncertain inputs (such as uncertain subsurface porosities) or even uncertain *functions* (such as scenarios over time); as for $\underline{z}$, the vector notation will be maintained, as it appears appropriate in most situations.

Some **model inputs** may be **fixed** – as their role is different from that of the uncertain inputs, they are given the notation ($\underline{d}$). This is the case for a number of reasons:

- some model inputs represent variables under full control: for example, the major operating conditions of an installation to be certified;
- uncertainties affecting some model inputs are considered to be negligible or of secondary importance with respect to the output variables of interest;
- for some model inputs, the decision process will conventionally fix the values despite uncertainties: for comparative purposes, it will do so by a conventional 'penalization', i.e. the choice of a fixed 'pessimistic' scenario, etc.

In industrial practice, the categorization of model inputs as 'uncertain' or 'fixed' is a matter of choice rather than theory. It can change over the course of the study and the decision-making process. Sensitivity analysis based on importance ranking and model calibration steps play key roles with respect to that choice, as will be discussed later.

Note that, for any given pre-existing system, the choice of the output variables of interest and of the appropriate chain of models to predict them depends on the industrial stakes and decision-making process. Note also that the number of model inputs varies according to the choice of pre-existing model for a given system; if the industrial stakes or regulatory controls change, this will give rise to very different uncertainty assessments even if the pre-existing system is the same.

1.2.2 Main goals of the uncertainty assessment

Industrial practice shows that the goals of any quantitative uncertainty assessment usually fall into the following four categories:

U (*Understand*): To understand the influence or rank importance of uncertainties, thereby to guide any additional measurement, modelling or R&D efforts.

A (*Accredit*): To give credit to a model or a method of measurement, i.e. to reach an acceptable quality level for its use. This may involve calibrating sensors, estimating the parameters of the model inputs, simplifying the system model physics or structure, fixing some model inputs, and finally validating according to a context-dependent level.

S (*Select*): To compare relative performance and optimize the choice of maintenance policy, operation or design of the system.

C (*Comply*): To demonstrate compliance of the system with an explicit criterion or regulatory threshold (e.g. nuclear or environmental licensing, aeronautical certification, etc).

There may be several goals in any given study and they may be combined over the course of a more-or-less elaborate decision-making process. Goals S and C refer to more advanced steps in operational decision-making, while Goals U and A concern more upstream modelling or measurement phases. Importance ranking may serve for model calibration or model simplification at an earlier stage, which becomes, after some years of research, the basis for the selection of the best designs and the final demonstration of compliance with a decision criterion. Compliance demonstration may explicitly require importance ranking as part of the process, etc.

However, as will be discussed later, the proper **identification of the most important goal(s)** of a given uncertainty assessment, as well **as of the quantities of interest** that are attached to them, are **key steps in choosing the most relevant methodologies**: this point often seems to be insufficiently appreciated in theoretical publications or prior comparisons of methodologies in official regulations or standards.

1.2.3 Measures of uncertainty and quantities of interest

The quantitative treatment of the inputs and outputs of the model may vary according to the main goal of the uncertainty assessment. However, they will, more or less explicitly, involve some '**quantities of interest**' (q.i.) in the output variables of interest. For instance, when the goal is to demonstrate compliance or to compare the results of different options, a regulation or a decision-making process involving uncertainty assessment will require the consideration of:

- percentages of error or variability in the variable(s) of interest (i.e. coefficient of variation) in measurement qualification or robust control;

- expected value of the variable of interest, such as a cost or utility in economics;
- confidence intervals of the variable(s) of interest, for instance in quality control; capabilities in robust design (i.e. ratios of a maximal acceptable range divided by 6 standard deviations);
- quantiles of the variable of interest in nuclear safety, mechanical characteristic values or corresponding to the VaR (Value at Risk) in finance;
- probabilities of exceeding a threshold or failure frequency in safety or reliability;
- ranges or simply the maximal value of the variable of interest in process control, etc.

For those quantities, an acceptable maximal value may be explicitly specified in the regulation or decision-making process, hence generating an explicit *decision criterion*. For example, for an installation, a process or a system to be licensed or certified, or for it to respect robust design objectives, the following criteria may have to be considered:

- 'there should be less than 3% uncertainty in the declared value for the output of interest';
- 'the physical margin should remain positive in spite of uncertainty, with a probability of less than 10^{-b} of being negative';
- 'the frequency of failure should be less than 10^{-b} per year, at a 95% confidence level covering the uncertainties';
- 'in spite of uncertainties, scenario A should be better (with respect to a given output variable of interest) than scenario B, to a level of confidence of at least 95%';
- 'the range of the output variable of interest should always be less than 20%' or 'the maximal value of the variable of interest should stay below a given absolute threshold', etc.

There may not be any thresholds or criteria as explicit as these, especially if the uncertainty practice is relatively recent in the given industrial field, as the case studies in Part II will demonstrate. However, there will generally be a certain '*quantity of interest*' in the output variables of interest, or more generally a '*measure of uncertainty*' on which the uncertainty assessment will issue results to be discussed in the decision process. To be more precise, what will be called a *quantity of interest* is a scalar quantity that summarizes mathematically the degree of uncertainty in the variable of interest, while the *measure of uncertainty* is the more complete mathematical distribution function comprehensively representing the uncertainty. As will be detailed in Part III, their mathematical content depends crucially on the paradigm chosen to represent uncertainty, which will hereafter be called the *uncertainty setting*. But the general structure stands, as for instance:

- in a probabilistic framework, the measure of uncertainty will be the probability measure, i.e. generally the cumulative distribution function (cdf) of the variable (or vector) of interest; the quantities of interest may be coefficients of variation, exceedance probabilities, standard deviations, or more generally any quantity derived from the cdf;
- in a non-probabilistic framework, the measure of uncertainty could be a Dempster-Shafer couple of plausibility/belief functions, while the quantity of interest might be the cumulative belief in not having exceeded a given safety threshold;
- formally, even in a deterministic framework, the measure of uncertainty could be considered the maximal range of some outputs, while the quantities of interest might be each bound of that interval.

The specification of quantities of interest and measures of uncertainty is quite natural if the final goal is of type C (*Comply*) or type S (*Select*), but it is also necessary for other types of goals, such as type U (*Understand*), or type A (*Accredit*). As will be discussed in Part III, it appears, for instance, that the importance ranking of the sources of uncertainties (Goal U) depends on the quantity of interest selected: most sensitivity analysis publications refer implicitly to variance as the quantity of interest for importance ranking, but in some cases the probability of threshold exceedance is much more relevant for industrial practice and produces very different results. Similarly, a system model can be satisfactorily calibrated (Goal A) as regards the variance of a given output of interest, but may be less acceptable regarding behaviour in the distribution tail of the output.

1.2.4 Feedback process

According to the final goal(s) motivating the uncertainty assessment, there may also be a more or less explicit *feedback process* after the initial study. Typical functions of this step might be:

- (**Goal C**) to adjust the design or the controlled variables/scenarios; to improve measurements, etc, so that the criteria can be met;
- (**Goal S**) to shift to another scenario that would further enhance performance in spite of uncertainties, e.g. by reducing the uncertainty in a critical output or by reducing costs while maintaining a given safety level, etc;
- (**Goal U/A**) to change the description of uncertainties (e.g. by removing some unimportant sources), to refine the system model to reduce uncertainties, to simplify the system model while maintaining acceptable accuracy despite inevitable uncertainties.

Of course, this feedback process can involve more than just one action, and it may be more or less strictly regulated or well defined according to the maturity of regulation or internal decision-making processes. It is generally considered essential in industrial practice, in which uncertainty assessment is often just one step in a larger or dynamic process.

1.2.5 Uncertainty modelling

Once the sources of uncertainty and corresponding input variables have been identified, there is inevitably a stage of uncertainty modelling (or quantification and characterization of the sources of uncertainty) which depends on the type of measure of uncertainty or quantities of interest chosen:

- In a probabilistic framework the uncertainty model will theoretically be a joint pdf of the vector of uncertain inputs ($\underline{x}$), although it may be specified more simply as a set of simple parametric laws for the components (e.g. Gaussian) with some independence hypotheses or approximate rank correlations.
- In an extended probabilistic framework the uncertainty model would be, for instance, a Dempster-Shafer couple of plausibility/belief functions for $\underline{x}$.
- In a deterministic framework, the maximal range of each component of $\underline{x}$.

Whatever the framework, there is always, however, a need to take into account the largest possible amount of information in order to build a satisfactory 'uncertainty model' (i.e. to choose the measure of uncertainty in the inputs). This information could include:

- direct observations of the uncertain inputs, potentially treated in a statistical way to estimate statistical models;
- expert judgement, in a more or less elaborate elicitation process, and mathematical modelling, from the straightforward choice of intervals to more elaborate Bayesian statistical modelling, expert consensus building, etc;
- physical arguments, e.g. that, however uncertain, the input should remain positive or below a known threshold for physical reasons;
- indirect observations (this is the case when the model is calibrated/validated and may involve some inverse methods under uncertainty).

As will be illustrated in Part II, uncertainty modelling may be a resource-consuming step for data collection; it appears, however, to be a crucial step to which the results of the uncertainty study may prove very sensitive, depending on the final goal and the quantities of interest involved. For instance, the choice of upper bounds or distribution tails becomes very sensitive if the quantity of interest is an exceedance probability.

1.2.6 Propagation and sensitivity analysis processes

Once an uncertainty model has been developed, the computation of the quantity (or quantities) of interest involves the well-known *uncertainty propagation* step (also known as uncertainty analysis). The uncertainty propagation step is needed to transform the measure of uncertainty in the inputs into a measure of uncertainty in the outputs of the pre-existing model. In a probabilistic setting, this implies estimating the pdf of $\underline{z} = G(\underline{x}, \underline{d})$, knowing the pdf of $\underline{x}$ and being given values of $\underline{d}$,

$G(.)$ being a numerical model. According to the quantity of interest and the system model characteristics, it may be a more or less difficult numerical step involving a wide variety of methods, such as Monte Carlo Sampling, accelerated sampling techniques, simple quadratic sum of variances, FORM/SORM or derived reliability approximations, deterministic interval computations, etc. Prior to undertaking one of these propagation methods, it may also be desirable to develop a surrogate model (equally referred to as response surface or meta-model), i.e. to replace the pre-existing system model with another which produces comparable results with respect to the output variables and quantities of interest, but which is much quicker or easier to compute.

The *sensitivity analysis* step (or *importance ranking*) refers to the computation and analysis of so-called sensitivity or importance indices of the components of the uncertain input variables $\underline{x}$ with respect to a given quantity of interest in the output $\underline{z}$. In fact, this involves a propagation step, e.g. with sampling techniques, but also a post-treatment specific to the sensitivity indices considered. This typically involves some statistical treatment of the input/output relations which control quantities of interest involving the measure of uncertainty in both the outputs and inputs (see Part II). The large variety of probabilistic sensitivity indices includes, for instance, graphical methods (scatterplots, cobwebs), screening (Morris, sequential bifurcations), regression-based techniques (Pearson, Spearman, SRC, PRCC, PCC, PRCC, etc.), non-parametric statistics (Mann-Whitney test, Smirnov test, Kruskal-Wallis test), variance-based decomposition (FAST, Sobol', correlation ratios), or local sensitivity indices of exceedance probabilities (FORM).

Note that the expression 'sensitivity analysis' is taken here in its comprehensive meaning, as encountered in the specialized uncertainty and sensitivity literature; in industrial practice the same expression may refer more generally to certain elementary treatments, such as one-at-a-time variations of the inputs of a deterministic model or partial derivatives. These two kinds of indices are usually not suitable for a consistent importance ranking, although they may be a starting point.

Part II and Part III will discuss the practical challenges involved in undertaking these two steps, for which the choice of the most efficient methods has to be carefully made. It will be shown that **it does not depend on the specificities of a physical or industrial context** as such, but rather on the **generic features** identified above: the computational cost and regularity of the system model, the principal final goal, the quantities of interest involved, the dimensions of vectors $\underline{x}$ and $\underline{z}$, etc. This is one of the important messages in this book. Having historically been designed for certain physical applications (e.g. structural reliabilistic methods to compute uncertainties in mechanics and materials), some methods do not reach their full generic potential in industrial applications when adhering to unnecessary cultural norms.

1.3 The common conceptual framework

The consideration of key generic features and concepts requires that the schematic diagram of Figure 1.2 be developed into a full conceptual framework (see

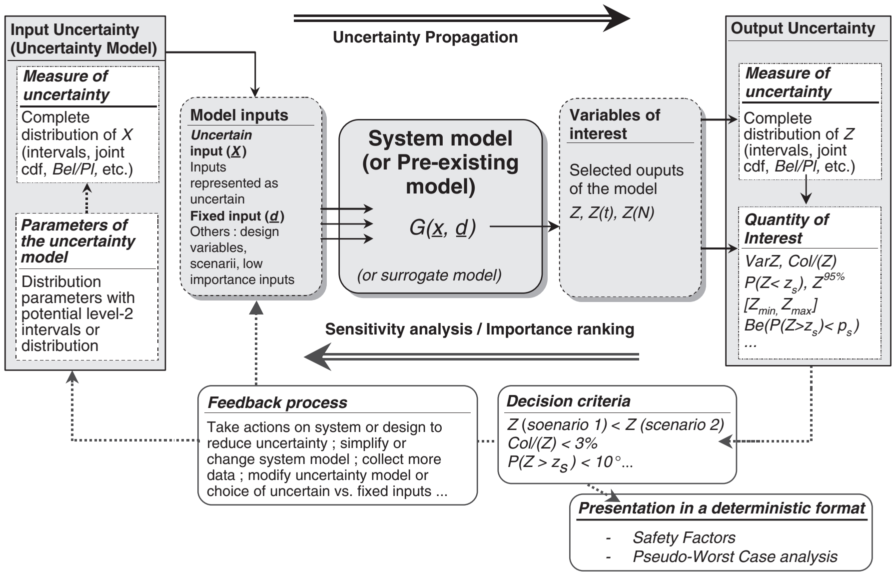

Figure 1.3 Common framework.

Figure 1.3, below) which will be used systematically throughout the case studies in Part II.

Note that the item 'presentation in a deterministic format' has been added. In industrial practice, after a more elaborate uncertainty assessment, there may be a sort of posterior process translating a probabilistic criterion into an easier-to-manipulate deterministic format, depending on the decision-making process and the operational constraints (e.g. set of partial safety factors, deterministic envelope of variation capable of guaranteeing the criteria, etc). This may be due to:

- cultural habits, when the regulation comes with more traditional deterministic margins;
- operational constraints which make it difficult to undertake systematically, on a large industrial scale, certain elaborate assessment processes;
- requirements of acceptability or facilitation of understanding.

The process may also feature when dealing with very low probabilities or highly unlikely uncertain results. Presentation in the form of mixed probabilistic/deterministic results proves, in fact, to be common in such cases, as will be demonstrated by the case studies in Part II and discussed more extensively in Chapter 19.

1.4 Using probabilistic frameworks in uncertainty quantification – preliminary comments

As already mentioned in the introduction, this book takes a practitioner's point of view on the use of mathematical settings in describing uncertainty, assuming firstly that the reader knows the basics of probability calculus, statistics and simulation methods. If necessary, textbooks such as (Bedford and Cooke, 2001; Aven, 2003; Granger Morgan and Henrion, 1990; Melchers, 1999; Rubinstein, 1981; Saltelli *et al*., 2004) can be consulted for a refresh in risk analysis, uncertainty quantification, Monte Carlo and sensitivity analysis. The book will then focus on guiding the overall consistency and practical trade-offs to be made in choosing one of those methods in industrial practice.

Additionally, theoretical approaches of uncertainty involve a long-standing literature and significant on-going controversies, including those associated to the rationale and conditions of the use of deterministic, probabilistic (classical or Bayesian), or extra-probabilistic settings. Indeed, epistemological considerations are necessarily underlying any description of the uncertainty surrounding a system study; and there is a close link to make to decision theory paradigms (and particularly decision-making in risk analysis and management) in choosing a framework or setting to represent uncertainty. Existing regulations, standards or codes of practices do in fact refer more or less explicitly to a wide spectrum of different settings and interpretations dealing with uncertainty over the scope of the book, as will be illustrated in Part II. A very brief panel of interpretations will be introduced hereafter while the reading of useful references might be considered for deeper understanding both on founding theory and applied interpretations, including Savage, 1972; De Finetti, 1974; Kaplan and Garrick, 1981; Helton and Burmaster, 1996, and more recently, for example, Bedford and Cooke, 2001; Aven, 2003; Helton and Oberkampf, 2004.

1.4.1 Standard probabilistic setting and interpretations

Consider for instance what will be referred to as the *standard probabilistic* setting, whereby probability distributions are assigned to the components of the input $\underline{x}$ (more precisely a joint distribution on vector $\underline{x}$). Computations are then made on quantities of interest derived from the resulting distribution of the random vector $\underline{Z} = G(\underline{X}, \underline{d})$ or on sensitivity indices involving both $\underline{X}$ and $\underline{Z}$.

A first interpretation, say *frequentist* or *classical*, of that setting would consider $\underline{x}$ and $\underline{z}$ as observable realisations of uncertain (or variable) events (or properties of the system), which would occur several times independently so that, at least in theory, frequency records of both variables would allow the inference and validation of the probability distribution functions (for both inputs and output). In that context, modelling probabilistic distribution functions on the inputs may be seen as a basis for the inference of some output quantities of interest, such as a probability to exceed a regulatory threshold or an expected cost. Taking such bases for decision-making enjoys straightforward risk control interpretations and has the advantage of a potential validation through long-term observations, as daily practiced in environmental control, natural risk regulations or insurance records.

Other views, involving totally different interpretations, are however possible on the same mathematical setting (say standard probabilistic) to quantify uncertainty. A classical *subjective* interpretation of the same setting might lead to consider the probability distributions as a model of the decision-maker subjective preferences following a 'rational preference' set of axioms (such as in Savage, 1974), or degrees of belief without necessary reference to frequency observations of physical variables. A quantity of interest such as the expected utility of the output random vector $\underline{Z}$ may then be used in a decision-making process, enjoying solid decisional properties. This may not necessarily need validation by long-term observations, which, in fact, may often be impractical in industrial practice: think about such cases as the design choices of an industrial product that does not yet exist.

Alternatively, when considering global sensitivity analysis of complex physical or environmental system model in upstream model development stages, one may rely on a sort of *functional analysis* interpretation. Using probabilistic distributions on the vector of inputs $\underline{x}$ of a system and considering variance as a quantity of interest on the model output $\underline{z}$ enjoys some desirable numerical space-exploration or global averaging properties that allow well-defined sensitivity ranking procedures. However, inputs and outputs of such model may not be observable at all, as being rather abstract model parameters in upstream research processes or not corresponding to reproducible random experiments.

While being quite different, these competing interpretations of a standard probabilistic setting imply rather similar practical implementation features, as will appear later in the book, such as: the need to carefully specify the quantity of interest and select the uncertainty model with all information available; the use of numerical methods for propagation and sensitivity analysis with delicate compromise when addressing complex models G etc. Indeed, frequent uses are made of such a standard probabilistic setting in regulated practice such as metrology (ISO, 2005), pollutant discharge control or nuclear licensing without positively choosing a single theoretical interpretation.

1.4.2 More elaborate level-2 settings and interpretations

More elaborate interpretations and controversies come up when considering the issue of the lack of knowledge regarding the uncertainty description, as generated for instance by small data sets, discrepancies between experts or uncertainty in the system model. This is particularly the case in the field of risk analysis or reliability. The incorporation of observed data to infer or validate the probabilistic modelling of inputs $\underline{x}$, when done through classical statistical estimation, generates statistical fluctuation in the parameter estimates (such as the expectation as estimated from the empiric mean on a finite dataset of temperatures, or the failure rate from limited records of lifetimes); this is even more so when considering the choice of the distribution for an input x^i for which traditional hypothesis-testing techniques give at most only incomplete answers: is the Gaussian model appropriate for the distribution of x^i, as opposed to, for instance, a lognormal or beta model?

This results in a sort of 'uncertainty about the uncertainty (or probabilities)', or sometimes referred to as epistemic uncertainty about the aleatory characteristics, although this formulation is controversial in itself (see Chapter 14). Disagreement between experts (or hesitation of one expert) that are consulted to help building the uncertainty model may also be viewed as generating similar *level-2 uncertainty*, although this all depends on the way the so-called expertise *elicitation* procedure is organised and theoretically formulated (*cf.* Granger Morgan and Henrion, 1990; Cooke, 1991). A *Bayesian* setting further formalises this second probabilistic level by deliberately considering a pdf for the parameters of the pdf of an uncertain model input, representing prior (lack of) knowledge of the uncertainty model, or the 'posterior' situation after incorporation of observed data.

Various settings can be found in the literature to answer that tricky, although often inevitable, issue. Some authors may informally stick to a standard probabilistic (i.e. level 1 setting) using point estimates for the pdf parameters involved in the uncertainty model, assumed to represent the 'best estimates': the majority of the industrial case studies illustrated in Part II will evidence its practicality and pervasiveness, although seldom justified in its precise theoretical foundations. In some cases, a deterministic sensitivity analysis is undertaken on the parameters of the pdf that describes the uncertainty of a model input: let them vary in order to investigate the variation of the output quantity of interest, and possibly retain, for decision-making, its maximal value (e.g. a 'penalised' probability of exceeding a threshold, encountered in some nuclear safety studies), although the real decisional properties of such an approach is rarely discussed. Such a setting will be called *probabilistic with level-2 deterministic* in the rest of the book, as illustrated by one case study in Part II.

Probabilistic modelling of the pdf parameters of the inputs has become popular in the industry since the late 1980s (particularly in the nuclear field) and is often referred to as distinguishing epistemic and aleatory components (see in particular Helton, 1994; Helton and Burmaster, 1996; Apostolakis, 1999). Such setting will be called *double probabilistic* (instead of *probabilistic with level-2 probabilistic*); it generates not only point values for the output quantities of interest but entire distributions (or epistemic distributions of the aleatory pdf of the variable of interest). It may also be encountered in some natural risk regulations, whereby upper confidence interval estimates are preferred to the central estimate of a quantile (such as the 1000-year return intensity) although practices are not homogeneous. Note finally that recent pilot studies have introduced extra-probabilistic settings (such as DST, fuzzy logic, possibility theory etc.) to model either the level 2 on top of a probabilistic setting (i.e. a DST description of the uncertainty of the parameters of the input pdf), as will be illustrated in the frontier case study of Part II, or simply the overall uncertainty model (see Helton and Oberkampf, 2004).

Essential controversies regard the interpretation of those various settings, and even the meaning of the double probabilistic settings alone. Note firstly that, as an equivalent of the *functional analysis* interpretation mentioned earlier, one may practically consider a double probabilistic setting as just another standard probabilistic one; switch simply the definition of the pre-existing model from the one relating input to output values to an intrinsically probabilistic model relating input to output distribution parameters or probabilities. This is often the case with fault trees or event trees in reliability, where the simple deterministic functions resulting from the causal trees hide somehow the aleatory nature of the model, as illustrated by the frontier case study in Part II. Then, level-2 uncertainty analysis may be viewed as a space-exploring numerical approach in order to understand the underlying aleatory model.

In a so-called *probability of frequency approach*, which may be the most frequent approach supporting double probabilistic settings, a more formal interpretation is given. An underlying assumption is that there is a theoretically-unique uncertainty (or aleatory) model, with correct shapes of distributions for the inputs and point values for their parameters as well as true value for the quantities of interest (e.g. a true failure frequency or probability of exceeding a threshold), as could be theoretically evidenced through unlimited long-term frequency observations. However, dataset limitations or lack of expertise generate uncertainty upon those characteristics, so that a complete quantification of uncertainty includes elaborate results, such as a quantity of interest (e.g. failure frequency) and a level of confidence around it. A *classical* interpretation might handle this approach through the use of the distribution of the estimators (e.g. Gaussian distributions underlying maximal likelihood estimation confidence intervals) to model level-2 input uncertainty when data is available or more generally subjective degree of beliefs for other types of level-2 uncertainty. These level-2 distribution on parameters of the level-1 input uncertainty model are later propagated to issue level-2 estimation uncertainty of level-1 quantities of interest. A *Bayesian (or classical Bayesian)* interpretation handles it in a more formalised way by consolidating subjective preferences or expertise (incorporated in the prior uncertainty model) and observed data in a mathematical updating procedure of the level-2 pdf describing the level-1 pdf parameters. This may include mixture of distributions to account to a wider extent for the level-2 uncertainty.

A *fully Bayesian* or *predictive Bayesian* approach (as advocated by Aven, 2003) refers to a quite different interpretation. The focus is placed on observable quantities meaning the observable realisation of some physical (or economic) variables (or states, events) characterising the pre-existing system. As a simplified presentation of that approach in correspondence to the common methodology of the book, this means that inputs $\underline{x}$ and variables of interest $\underline{z}$ should correspond to (at least theoretically) observable variables (or states) of the system. The corresponding q.i, such as a system failure frequency, the expected value of the variable of interest or its probability to exceed a threshold, should be understood as decision-making intermediate tools rather than real properties of the system enjoying one true value. Hence, it is not relevant to distinguish within the results a level-2 uncertainty about the q.i. that embeds the level-1 uncertainty. Uncertainty should be eventually quantified within one single figure for the q.i., which corresponds to the expectation over the level-2 (posterior) distribution of the q.i.; or within the single distribution of the variable of interest, called the predictive distribution, which also corresponds to the level-1 distribution averaged over the level-2 uncertainty of its parameters. For instance, one would not consider a credibility interval for a system failure rate[1], but merely its expected value, consolidating all components of uncertainty into a single probabilistic figure. Note that, beyond significant differences of interpretation, it may be practically linked to the output of a Bayesian *probability of frequency approach* (e.g. taking the expectation over the level-2 distribution generated for the q.i.), provided it was specified along the same information basis on the inputs and definition of observable variables of interest.

1.5 Concluding remarks

The book does not suggest the choice of one interpretation or another recognising that, from a practical point of view, the implementation involves important common features, while of course the interpretation of the results may differ. As will be recalled in Chapter 14, the scientific literature presents a number of classifications regarding the nature of uncertainty (such as aleatory or epistemic uncertainty, variability, imprecision, etc). Some authors link the interpretation of the appropriate mathematical setting to the nature of uncertainty involved in the system studied. On this quite controversial issue the book will not propose a unique preferred interpretation; it will rather illustrate the various settings that may be found in industrial practice, in response to varying regulatory specifications and interpretations, as well as their practical implementation consequences (such as data collection, computing needs, etc). For instance, the previous section has evidenced that while there may be quite different competing interpretations, standard probabilistic and double probabilistic settings share practical implementation features, on which the book will focus. This may result in a certain incompleteness of specification of the precise words employed, such as the use of confidence intervals that would rather deserve the name of credibility intervals or prediction intervals in some cases according to the way the input uncertainty model may be theoretically interpreted and practically elicited. Once again, the reader may refer to the publications cited in the next section discussing theoretical foundations, notably on the use of probabilistic or non-probabilistic settings in representing uncertainty, and also to recent benchmarking or comparison exercises such as that reported in (Helton and Oberkampf, 2004).

[1]except possibly when working on a true population of similar systems running together.

References

Apostolakis, G. (1999) The distinction between aleatory and epistemic uncertainties is important; an example from the inclusion of aging effects in the PSA, in *Proceedings of PSA '99*, Washington DC.

Aven, T. (2003) *Foundations of Risk Analysis*, Chichester: John Wiley & Sons, Ltd.*

Barberà, S., Hammond, P.J. and Seidl, S. (1998) (Eds) *Handbook of Utility Theory, Volume 1 Principles*, Hingham: Kluwer Academic Publisher.

Barberà, S., Hammond, P.J. and Seidl, S. (1998) (Eds) *Handbook of Utility Theory, Volume 2 Extensions*, Hingham: Kluwer Academic Publisher.

Bedford, T. and Cooke, R. (2001) *Probabilistic Risk Analysis – Foundations and Methods*, Cambridge: Cambridge University Press.*

Cooke, R.M. (1991) *Experts in Uncertainty; Opinion and Subjective Probability in Science*, Oxford: Oxford University Press.

De Finetti, B. (1974) *Theory of Probability*, volumes I and II, New York: John Wiley & Sons, Ltd.

De Neufville, R. (2003) Architecting/Designing engineering Systems using Real Options, *MIT report* ESD-WP-2003-01.09 (http://esd.mit.edu).

De Neufville, R. (2004) Uncertainty management for engineering systems planning and design, *MIT Engineering Systems Monograph* (http://esd.mit.edu/symposium/pdfs/monograph/uncertainty.pdf).

Granger Morgan, M. and Henrion, M. (1990) *Uncertainty – A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge: Cambridge University Press.*

Helton, J.C. (1994) Treatment of Uncertainty in Performance Assessments for Complex Systems, *Risk Analysis*, **14**, pp. 483–511.

Helton, J.C. and Burmaster, D.E. (1996) (eds.) Treatment of Aleatory and Epistemic Uncertainty, Special Issue of *Reliability Engineering & System Safety*, **54**(2–3).

Helton, J.C. and Oberkampf, W. (2004) (eds.) Alternative representations of epistemic uncertainty, Special Issue of *Reliability Engineering & System Safety*, **85**(1–3).

Helton, J.C., Cooke, R.M., McKay, M.D. and Saltelli, A. (2006) (Eds) Sensitivity Analysis of Model Output: SAMO 2004, Special Issue of *Reliability Engineering & System Safety*, **91**(10–11).

ISO (1995) *Guide to the expression of uncertainty in measurement (G.U.M.)*, European Prestandard Env. 13005.

Kaplan, S., and Garrick, B.J. (1981) On The Quantitative Definition of Risk, *Risk Analysis*, **1**(1), 11–27.

Knight, F.H. (1921) *Risk, Uncertainty and Profit*, Hart, Schaffner & Marx.

Law, A.M. and Kelton, W.D. (2000) *Simulation Modelling and Analysis* (3rd edition), London: McGraw Hill.

Melchers, R.E (1999) *Structural reliability analysis and prediction* (2nd edition), Chichester: John Wiley & Sons, Ltd.*

Nilsen, T. and Aven, T. (2003) Models and model uncertainty in the context of risk analysis, *Reliability Engineering & System Safety*, **79**(309–317).

*Also suitable for beginners.

Oberkampf, W.L., DeLand, S.M., Rutherford, B.M., Diegert, K.V and Alvin, K.F (2002) Error and uncertainty in modelling and simulation, Special Issue of *Reliability Engineering & System Safety*, **75**(3), 333–357.

Paté-Cornell, M.E. (1996) Uncertainties in risk analysis; Six levels of treatment, *Reliability Engineering & System Safety*, **54**(2–3), 95–111.

Quiggin, J. (1982) A theory of anticipated utility, *Journal of Economic Behavior and Organization*, **3**, pp. 323–343.

Rubinstein, R.Y. (1981) *Simulation and the Monte-Carlo Method*, Chichester: John Wiley & Sons, Ltd.*

Saltelli, A., Tarantola, S., Campolongo, F. and Ratto, M. (2004) *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*, Chichester: John Wiley & Sons, Ltd.*

Savage, L.H. (1954, 1972) *The Foundations of Statistics*, Dover: Dover Publication Inc.

*Also suitable for beginners.